RELATIONS BETWEEN LANDSAT ETM+ IMAGERY AND FOREST STRUCTURE PARAMETERS IN TROPICAL RAINFORESTS: A CASE STUDY FROM LORE-LINDU NATIONAL PARK IN SULAWESI, INDONESIA

Pavel Propastin

University of Göttingen, Department of Geography, Göttingen, Germany;
ppropas(at)uni-goettingen.de

ABSTRACT

Forest structure parameters such as diameter at breast height (*DBH*), tree height (*H*), stem density (*STDENS*), basal area (*BA*) and crown extension (*CROWEX*) are major variables of a forest inventory, which can also be investigated using satellite data. Remote sensing is often the only practical means of acquiring information on these variables over large areas, particularly in tropical regions with inaccessible terrain. This paper investigates the suitability of LANDSAT ETM+ imagery for remote estimation of forest structure parameters in a tropical rainforest, namely the Lore-Lindu National Park in Central Sulawesi, Indonesia. All LANDSAT ETM+ bands (except the 6th band) and four multi-spectral vegetation indices derived from these bands were examined for their performance to express the forest structure parameters *DBH*, *H*, *STDENS*, *BA*, and *CROWEX*. The field data used for the correlation and regression analysis were collected from 60 plots throughout the area. The study showed that middle-infrared wavelengths of LANDSAT (bands 5 and 7) were best correlated with the forest structural parameters and could be used for prediction of *DBH*, *H*, *BA* and *CROWEX*. The tree height was predicted most reliably, while predictions of diameter at breast height, basal area and crown extension were less accurate. Stem density could not be predicted with sufficient accuracy from canopy reflectance only. The vegetation indices derived from red and near-infrared wavelengths (bands 3 and 4) such as the normalised difference vegetation index and the simple ratio were weakly correlated with the selected forest parameters. The vegetation indices incorporating band 5 showed much stronger correlations with forest parameters. The results revealed a great potential for the estimation of forest structure parameters with satellite optical data in tropical regions, where an extensive field survey is not feasible.

INTRODUCTION

Forest structure parameters, such as volume of biomass, average tree diameter and height, basal area, and stem density are important data needed to assess forest resources. Periodical updating of forest data is an indispensable condition for monitoring the change of forest resources and assessing the impact of change on sequestering carbon. One of the major problems with updating is that accurate data on forest structure are often difficult to obtain for regional assessments, particularly in the tropical regions. The lack of periodically updated and accurate spatial data for forest resources has been considered one of the most persistent uncertainties in estimating global carbon budget.

Satellite data provide a means by which forest stand structure parameters can be mapped and monitored over a large geographical region in a consistent and replicable manner. During the last three decades great efforts have been done in the field of remote sensing-based estimating and mapping forest structure parameters at different spatial scales (1,2). Multi-spectral reflectance data acquired by a large number of commercial and military satellites have been widely used by researches for forest inventories in different regions all over the world. Images acquired by LANDSAT Thematic Mapper (TM), Enhanced Thematic Mapper (ETM+) and ASTER have been most popular data for applications at sub-regional and regional scales (3,4,5,6,7), while Moderate

Resolution Imaging Spectroradiometer (MODIS), Advanced Very High Resolution Radiometer (AVHRR), and the VEGETATION instrument onboard of Satellite Pour l'Observation de la Terre (SPOT) have been effectively used for applications at regional to global scale (8,9,10,11).

Even though great progress in the field of remote sensing applications in forestry has been achieved and the corresponding study reports climb in recent literature, many problems of such applications wait for their solution. For instance, despite the fact that during the last decade numerous studies have reported about effective estimations of forest structural parameters and other characteristics from satellite data (12,13,14,15), the application of remote sensing in equatorial and tropical forest zones is highly problematic. Many of the approaches that have been widely used were developed for application in other environments and are often inappropriate in these zones. For example, a number of multispectral vegetation indices derived from satellite data have been established as general indicators for conditions of vegetation cover and vegetation activity (16,17). The remotely sensed derived vegetation indices have also been successfully used for estimating the above ground biomass from remotely sensed data in the temperate forest zone (18,19,9). On the contrary, studies by (4,5) showed that vegetation indices demonstrate only a weak or no sensitivity to biophysical properties at high amounts of vegetation and are rather useless for application in tropical forests.

Different parametric and non-parametric methods, such as statistical regressions, neural networks (NN) and k-nearest neighbours (KNN), have often been employed to link field estimations of forest structure parameters to satellite image data (3,4,5,6,8,9,12). The regression analysis, which is the most used simple method in the estimation of forest attributes from remote sensing data, has proved to be as effective as the more complicated NN and KNN techniques (7). However, when implementing a regression analysis, one has to check the linearity of the dataset. Previous studies have shown that non-linear regression models have significantly improved the accuracy of prediction in many cases (6,7,15).

The aim of this paper is to investigate the potential of using satellite imagery to map the structure of tropical rainforest over a large geographical region. Structural forest data on diameter at breast height (*DBH*), total tree height (*H*), stem density (*STDENS*), crown extension (*CROWEX*) and basal area (*BA*) collected across the Lore-Lindu National Park in Sulawesi, Indonesia, were related to canopy reflectance data derived from LANDSAT ETM+ data. The relations are used to investigate which wavebands and vegetation indices may be most reliable for mapping these forest parameters in the tropical moist forest zone.

STUDY AREA

The study area is located in Central Sulawesi, Indonesia (Latitude 0°55'- 01°54' South, Longitude 119°40'- 120°29' East) and comprises the region of the Lore-Lindu National Park together with the bordering areas (Figure 1). The area has a very complicated relief with elevations from zero in the north to more than 2600 m in the middle part and is separated by four river valleys: Palolo in the north, Napu in the east, Bada in the south and Kulawi in the west. The highest peaks are Mt. Nokilalaki (2355 m) and Mt. Rorekatimbu (2610 m). The extent of the study area is about 100 km from north to south and about 60 km from east to west.

Large parts of the study area, especially major parts of the mountain ranges, are covered by undisturbed tropical forest, whose natural vegetation is generally classified into two major vegetation types based on altitudinal distribution with lowland rainforest below 1000 m and mountain rainforest above 1000 m. These forests are dominated by wood species such as *Aglaia argentea, Pimelodendron amboinicum, Bischofia javanica, Cananga odorata, Meliosma sumatrana.* The natural forest has been subject to remarkable forest conversion activities along the park boundaries. Near the national park boundaries, many previously forested areas were transformed into perennial agro-forestry areas with cocoa and/or coffee cultivation. Some of these areas were abandoned after short-term cultivation and reverted to secondary forest. The secondary forests are dominated by *Acalypha caturus, Grewia*

glabra, Homalanthus populneus, Macaranga hispida, Mallotus mollissimus, Pipturus argenteus. Most areas of the river valleys are completely deforested and used for production of paddy rice.

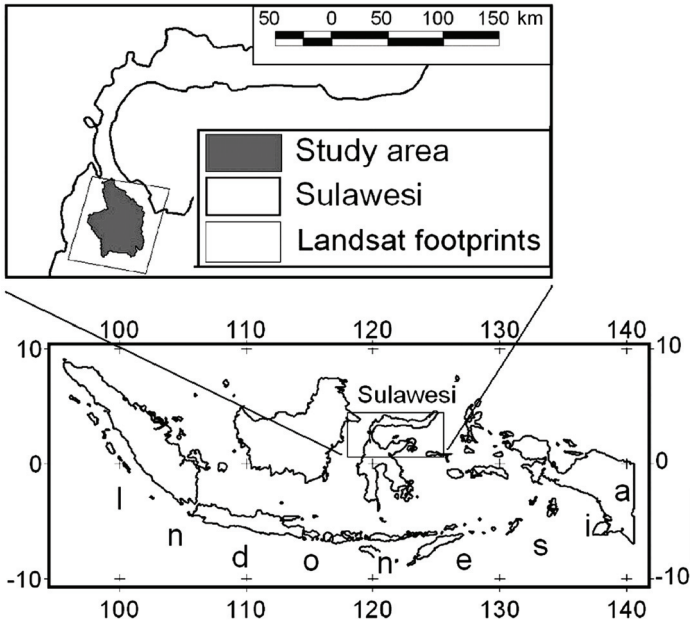

Figure 1: The study area is located in Sulawesi, Indonesia. In the inset (top) the Lore-Lindu-National Park and the footprint of LANDSAT ETM+ image employed in the study.

DATA AND METHODS

Preparation of the field data

A ground survey of 60 forest plots in natural undisturbed forest and secondary forest was carried out during a number of field survey campaigns between 2003 and 2007. Every effort was made to select study plots throughout the study region, but in practice, because of a very difficult terrain, the selection of survey sites was influenced by accessibility. Originally, the field survey plots were selected based on stratified random sampling using a LANDSAT classified image (24) and a digital terrain model in combination. The stratification was based on the natural vegetation distribution in the region. Two strata were determined: the stratum of lowland rainforest and the stratum of mountain rainforest. In cases where the targeted plots were not accessible (due to impassable vegetation and difficult terrain), the specific plots were replaced by new easily accessible plots.

A nested sampling strategy organised in plots and subplots was employed in order to inventory the sample sites (Figure 2). The dimensions of the plots were 60 m $\times$ 40 m and each of them was divided into 24 subplots with dimensions of 10 m $\times$ 10 m (Figure 2, right). Once a plot was selected and demarcated, geographic coordinates of the corners as well as the elevation above sea level were measured by a GPS device. Many times the terrain and the crown closure of the forest hindered the reception of the satellite signal by the GPS. Therefore, according to the device specifications and standards the error of the GPS measurements in the study area must be considered between 5 and 8 m. For every plot, 6 out of 24 nested subplots were randomly selected in order to conduct the field measurements. These measurements included *DBH* (cm), *H* (m), and *CROWEX* (m) of all individual trees diameter with a breast height higher than 10 cm. Diameter at breast height was measured at 1.3 m height using the common girth tape. Tree heights were measured to the tallest live portion of the crown using a Vertex height finder. The crown extension of an individual tree was calculated as ½ of a mean from four measurements of crown radius. The measurements of crown radius were carried out using the perpendicular projection of the crown edge of a tree on the ground. The basal area (*BA*, m²/ha)

was calculated from the measured *DBH* and *STDENS*. The results of the sub-plot measurements for *DBH*, *H* and *CROWEX* were re-calculated to the corresponding plot. The re-calculation of these three structure parameters to the plot level was conducted using the following equations in (18):

$$DBH_{plot} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} DBH_{tree}}{m} \quad (1)$$

$$H_{plot} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} H_{tree}}{m} \quad (2)$$

$$CROWEX_{plot} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} CROWEX_{tree}}{m} \quad (3)$$

The stem density and the basal area were re-calculated from the level of ground measurements to the hectare level (ha^{-1} and m^2/ha) using the equations:

$$STDENS_{ha} = 10000 \cdot \frac{\sum_{i=1}^{m} n}{\sum_{i=1}^{m} A_{sub-plot}} \quad (4)$$

$$BA_{m^2/ha} = STDENS_{ha} \cdot \frac{0.0314\ DBH_{plot}^2}{2} \quad (5)$$

In these equations, *n* is the total tree number in a sub-plot, *m* is the number of sub-plots in a plot, DBH_{tree} is *DBH* of the individual tree, H_{tree} is *H* of the individual tree, $CROWEX_{tree}$ is *CROWEX* of the individual tree, $A_{sub\text{-}plot}$ is the sub-plot area (m^2) within the selected plot.

According to the conducted measurements, *DBH* ranged from 10.18 to 34.03 cm with a mean value of 20.66 cm. Total tree heights varied from 5.45 to 33.27 m with a mean value of 14.30, and the stem density ranged from 209 to 972 ha^{-1} with a mean of 601.51. The crown extension ranged from 2.2 to 11.6 m. The mean value of CROWEX was 6.03 m and the standard deviation showed a value of 2.3 m. A summary of forest structure statistics for the 60 plots surveyed are given in Table 1.

Table 1: Summary statistics of the field measurements.

Parameters	Minimum	Maximum	Mean	Standard dev.
DBH, cm	10.18	34.03	20.66	7.17
H, m	5.45	33.27	14.30	7.33
STDENS, ha-1	209.87	972.13	601.51	221.97
CROWEX, m	2.25	11.61	6.03	2.31
BA, m²/ha	4.49	84.39	33.94	15.50

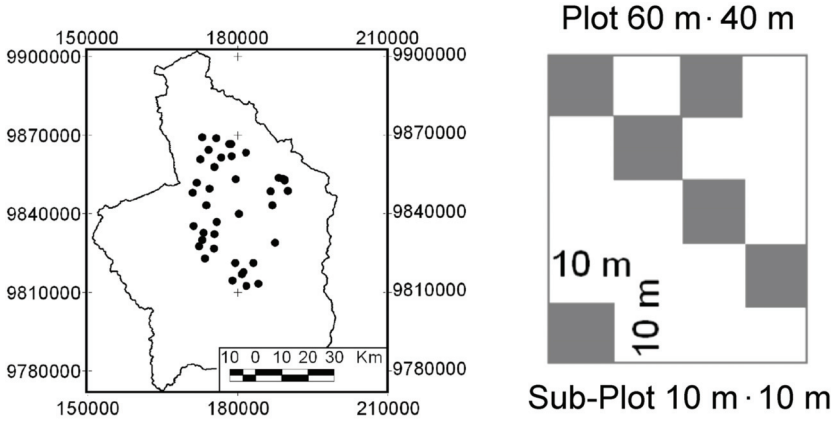

Figure 2: Design of the field data sampling: distribution of the sample plots inside the study area (left side), and subplot design (random selection) inside sample plots (right side).

Preparation of the spectral data

A LANDSAT ETM+ image (WRS-2, Path 114 / Row 61) covering the entire study area was acquired on 28 September 2002. Image pre-processing included all common steps of the satellite image treatment such as geometric, radiometric, and atmospheric corrections, ortho-rectification and the correction of terrain illumination effects ? . The image was rectified into UTM projection (WGS84, Zone 51) using 65 control points taken from topographic maps on a 1:50,000 scale. A nearest-neighbour resampling technique was used and a root mean square error of less than 0.5 pixels was obtained for the rectified image. The considerable effect of terrain illumination in the research area has been addressed using a topographic normalization methodology based on the non-lambertian Minnaert model (20,21). The empirical constants for this correction were separately computed for different sub-regions of the dataset according to image based structural surface characteristics. A detailed description of the pre-processing using LANDSAT ETM+ scene is given by (22). The software package ENVI 4.3 was used for image processing (23).

In order to examine which of the most used vegetation indices respond to forest structure parameters in the study area, four multi-spectral vegetation indices were incorporated in the analysis (Table 2). The vegetation indices used in the present work cover two different groups. Simple Ratio (*SR*) and Normalized Difference Vegetation Index (*NDVI*) are the most common vegetation indices that employ two bands in the red (R) and near-infrared (NIR) domain. Three wavelengths (R, NIR, and MIR) are used to form the reduced simple ratio (RSR) and the middle infrared corrected normalised difference vegetation index (*NDVIc*).

Table 2: Multispectral vegetation indices employed in the analysis.

Vegetation index	Equation	Reference
NDVI	$\frac{NIR - R}{NIR + R}$	(4,5,16,18)
NDVIc	$NDVI \cdot \left(1 - \frac{MIR - MIR_{\min}}{MIR_{\max} - MIR_{\min}}\right)$	(18,19,27,28)
SR	$\frac{NIR}{R}$	(16,25,26)
RSR	$SR \cdot \left(1 - \frac{MIR - MIR_{\min}}{MIR_{\max} - MIR_{\min}}\right)$	(27,28)

Statistical analysis

In general, the geometric precision of the satellite data and the positional accuracy of ground plots are related to the spatial fit of spectral and field data. In the present study, given the fact that the positional RMS error of the field data was less than 8 m (determined by the GPS device specifications), while the geometric RMS error of the image was 0.5 pixels (i.e., 15 m), the total spatial fit RMS error between the two was less than 23 m. However, for a better spatial fit of the current image and field data, the image was spatially degraded by two times using a nearest neighbour resampling technique. This resampling resulted in scaling up spectral data approx. to the field plot size (60 m $\times$ 60 m vs. 60 m $\times$ 40 m), thus guaranteeing a better spatial fit of the image and field datasets.

The spectral values corresponding to each plot were extracted after overlaying the field plots on the resampled image. No spectral values were extracted from the thermal band, due to the fact that this band had a coarser spatial resolution (120 m) and was influenced by many other factors not relevant to canopy temperature. The field data were modelled with LANDSAT spectral data using linear and non-linear regression formulas (7,8,19). The non-linear formula employed was the following:

$$y = a x_1^{b_1} x_n^{b_n} \tag{6}$$

where y is one of the forest structure parameters under study, x_1 and x_2 is the reflectance in LANDSAT ETM+ bands, and a, b_1 and b_n are regression constants.

A number of performance (goodness-of-fit) and accuracy statistics of the model were used for the evaluation of individual regression models by the cross-validation technique, in which the observed values were assessed against the predicted values. *F*-value and the value of the determination coefficient R^2 served as guides for assessing the model suitability, whereas the root mean squared error (*RMSE*) and bias were indicators of the model predictive ability. The accuracy statistics were derived using *RMS* error and *bias* equations:

1. $$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{y}_i\right)^2} \tag{7}$$

2. $$Bias = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right) \tag{8}$$

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 samples. The dataset was tested for normality and homogeneity using the formulas supplied by (34) and passed both tests. A statistical analysis was conducted using the Microsoft EXCEL software package.

RESULTS

Relations between forest structure and spectral parameters

A visual inspection of scatter diagrams between ETM+ data and forest structure parameters was done prior to the correlation analysis. All the relationships observed in the scatter diagrams demonstrated a clear non-linear nature. For this reason, the original data were transformed to logarithms to linearise the relationship between these variables and improve the correlation.

A correlation analysis between the transformed data on spectral reflectance in each LANDSAT ETM+ band and the forest parameters was performed. The resulted correlation coefficients, confidence intervals, and the significance level of correlation coefficients are summarised in Table 3.

The individual LANDSAT ETM+ bands showed negative response to the *DBH*, *H*, *CROWEX* and *BA*. On the contrary, *STDENS* correlated positively with the reflectance in the individual bands.

The examined forest structure parameters showed statistically significant response to the reflectance signal in the LANDSAT ETM+ individual bands. In most cases the relationship between the structural parameters and the LANDSAT ETM+ bands was significant at the confidence level $p<0.01$, even though the values of the correlation coefficient were not particularly high. Thus, *DBH*, *H*, *CROWEX* demonstrated statistically significant correlations with all individual LANDSAT ETM+ bands, whereas *BA* showed no significant correlation with band 4. Among structural parameters, *H* showed the strongest correlations with all the individual LANDSAT ETM+ bands. The *H*-correlations ranged from -0.70 to -0.85, with the slightest correlation with band 4 and the strongest correlation with bands 5 and 7. The response of *STDENS* to the reflectance in individual LANDSAT ETM+ bands was very weak with values of the correlation coefficient ranging from 0.21 to 0.34. *STDENS* was slightly correlated only with bands 1, 2 and 7 ($p<0.05$).

Table 3: Correlation coefficient between DBH, H, STDENS, CROWEX, BA and LANDSAT ETM+ spectral response. The original data sets were transformed to logarithms.

	DBH	H	STDENS	CROWEX	BA
B1	-0.52	-0.73	0.34	-0.61	-0.61
B2	-0.63	-0.80	0.34	-0.69	-0.69
B3	-0.64	-0.74	0.32*	-0.65	-0.70
B4	-0.42	-0.62	0.06*	-0.59	-0.30*
B5	-0.70	-0.85	0.21*	-0.77	-0.71
B7	-0.74	-0.85	0.35	-0.72	-0.76
NDVI	0.42	0.38	-0.30*	0.31*	0.49
NDVIc	0.73	0.84	-0.37	0.76	0.76
SR	0.40	0.36	-0.26*	0.27*	0.46
RSR	0.62	0.69	-0.29*	0.54	0.67

* Correlation coefficient is not significant at $p<0.05$.

Among the visible wavelengths, band 3 was the band most strongly correlated with DBH, H and BA. In comparison with other wavelengths, the near-infrared wavelength (band 4) appeared to have a relatively weak response to the examined forest parameters. The middle-infrared wavelengths correlated strongly with DBH, H, BA and CROWEX. Band 7 had the all strongest correlation with DBH, H and BA, whereas band 5 was the one that showed the best correlation with CROWEX.

The results summarised in Table 3 show that not all forest structure parameters are significantly related to the employed vegetation indices. Also vegetation indices are not definitely better related to the forest structural parameters than ETM+ spectral signatures. Overall, the "two-bands" vegetation indices, *NDVI* and simple ratio, provided weaker response to the structural parameters than the individual LANDSAT ETM+ bands (except band 4). The relationship between the "two-band" *VI*s and *H* and *CROWEX* was not significant, whereas with *DBH* and *BA* it showed statistical significance only at the 0.05 level. *STDENS* responded to the vegetation indices negatively and showed a statistically significant correlation only with the middle-infrared corrected *NDVI*. However, the correlation coefficient of the *STDENS*-*NDVIc* relationship was very weak ($r = -0.37$, $p<0.05$). In contrast to the "two-band" vegetation indices, *RSR* and *NDVIc*, which incorporated the middle-infrared wavelength as the third band, provided a much stronger correlation with *DBH*, *H*, *STDENS*, *CROWEX* and *BA* than *NDVI* and *SR*. Thus, *NDVIc* showed values of the correlation coefficient similar to the best correlations provided by the individual ETM+ bands. The correlations provided by *RSR* were somewhat lower than by *NDVIc*. These results suggest that different vegetation indices have different potentials for estimation of relevant forest parameters.

Models for predicting the forest structural parameters from remote sensing data

Simple and multiple regression approaches were employed for establishing models between the structural forest parameters and the spectral data. When selecting bands to be incorporated into models (Table 4), only bands with statistically significant different variance were taken as explanatory variables. For the evaluation of the models the cross-validation procedure was employed.

Table 4 shows the parameters of the non-linear regression models and the performance statistics. The explanatory power (R^2) of the LANDSAT ETM+ reflectance was highest for the tree height parameter. The best regression model for *H* explained 80% of spatial variance in tree height. Models for estimating *DBH*, *CROWEX* and *BA* are characterised by significantly lower explanatory power (R^2 = 0.60-0.61). The poorest R^2 resulted when modelling the stem density parameter (R^2 = 0.29), indicating that it is rather difficult to predict stand density from this satellite image. In spite of the fact that the R^2-values of the models were not particularly high except for the R^2 of the *H*-model, the predictions of the models were relatively accurate: the *RMSE* values varied between 15-27% of the corresponding mean values calculated in the ground data sets (compare with Table 1). For instance, the best-fitted *H*-model produces a value of *RMSE* = 3.50 m which accounts to approx. 17% from the mean value of *H* in the ground-based data set (14.30 m). The highest relative *RMSE* value was produced by the *STDENS* model: 164.57 against 601.28 (about 27%). All models are characterised by slight positive biases.

The values predicted by the models (except *STDENS*) are shown together with the 95% confidence intervals of the individual predictions in Figure 3. Note that the confidence intervals are for individual predictions and so, are wider than those for the expected values. The models seem to explain reasonably well the variation of the dependent variables. The residuals of the models against the predicted values are shown in Figure 4. A statistical test of regressing the models squared residuals on the corresponding predictors (LANDSAT ETM+ bands) revealed no significant relationships between these variables. The results of the tests proved that no heteroscedacity was presented in the models residuals. This means good suitability of the produced regression models to reproduce the ground-based values of the structural parameters.

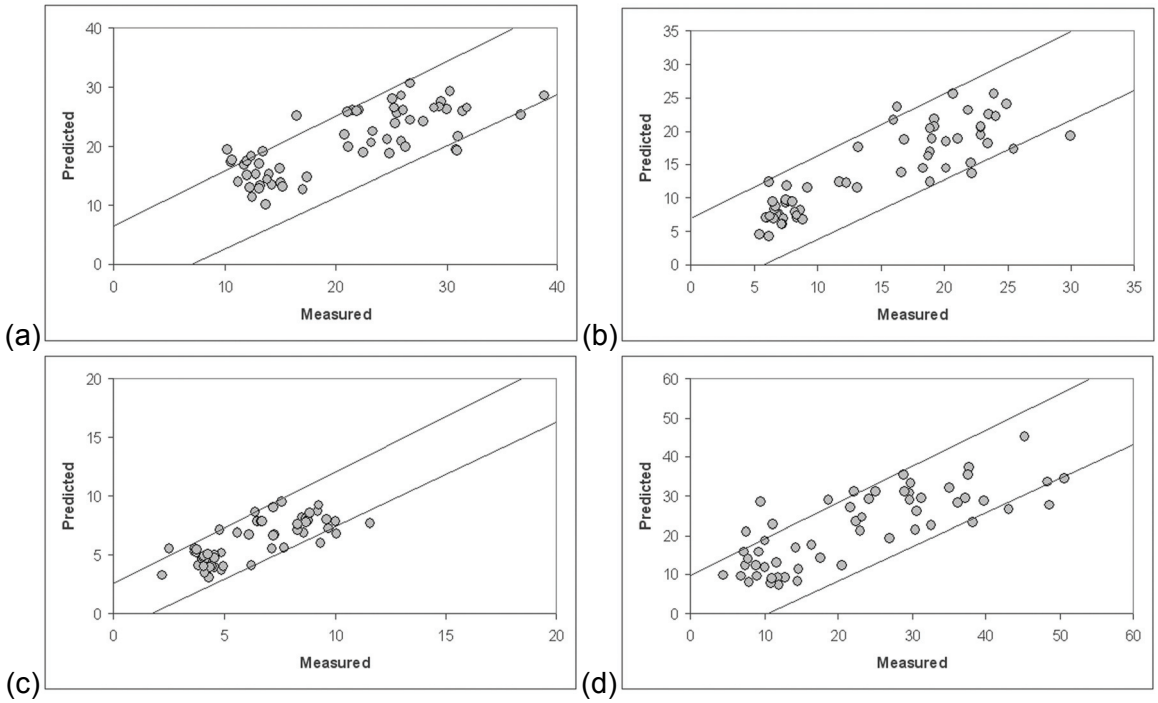

Figure 3: LANDSAT ETM+ estimates of the forest structural parameters against the ground measured values together with 95% confidence intervals of individual predictions for (a) diameter at breast height (DBH), (b) tree height (H), (c) crown extension (CROWEX), and (d) basal area (BA).

Table 4: Non-linear regression models for estimating DBH, H, CROWEX, and BA from image spectral response.

Structural parameter	Regression parameters ($y = ax_1^{b_1} x_n^{b_n}$)					Model performance statistics		
	a	$x_1^{b_1}$	$x_2^{b_2}$	$x_3^{b_3}$	$x_4^{b_4}$	R^2	*RMSE*	*Bias*
DBH	0.3560	$B3^{-0.9623}$	$B4^{0.5328}$	$B7^{-0.5615}$		0.66	4.56	0.51
	0.3875	$B3^{-1.1307}$	$B4^{0.6608}$	$B5^{-0.5325}$		0.65	4.67	0.60
	0.1635	$B3^{-0.8728}$	$B7^{-0.7281}$			0.60	4.85	0.53
	1.0259	$B4^{0.3923}$	$B7^{-1.3541}$			0.59	4.91	0.54
	0.7633	$B3^{-0.5453}$	$B5^{-0.8503}$			0.57	4.92	0.38
	0.9881	$B7^{-1.2021}$				0.58	4.91	0.42
	50.3770	$NDVIc^{1.9256}$				0.55	5.01	0.56
H	0.0101	$B3^{-1.2251}$	$B7^{-1.1950}$			0.80	3.50	0.39
	0.0174	$B2^{-0.9438}$	$B7^{-1.3977}$			0.78	3.55	0.41
	0.0368	$B4^{-0.6849}$	$B7^{-2.0080}$			0.78	3.58	0.37
	0.0393	$B7^{-2.2733}$				0.76	3.61	0.38
	67.9498	$NDVIc^{3.6848}$				0.71	3.82	0.41
CROWEX	2.8058	$B1^{-0.8681}$	$B2^{1.0599}$	$B5^{-3.1923}$	$B7^{1.6219}$	0.62	1.62	0.20
	5.9981	$B2^{0.5534}$	$B5^{-3.3357}$	$B7^{1.5231}$		0.61	1.62	0.23
	2.1242	$B5^{-3.0591}$	$B7^{1.6303}$			0.60	1.64	0.21
	0.2111	$B5^{-1.9126}$				0.60	1.64	0.21
STDENS	7886.07	$B2^{0.8735}$	$B4^{-1.3536}$	$B7^{0.4536}$		0.29	164.57	20.25
BA	0.0032	$B1^{-0.6437}$	$B3^{-0.5777}$	$B7^{-1.8023}$		0.62	10.59	2.02
	0.0038	$B3^{-2.1593}$	$B5^{-0.7413}$			0.60	11.05	2.11
	0.0079	$B3^{-0.8535}$	$B7^{-1.9197}$			0.60	10.95	2.08
	145.8396	$NDVIc^{4.4964}$				0.58	11.92	2.21

DISCUSSION

Relationships between forest structural parameters and LANDSAT ETM+ bands

DBH, *H*, *STDENS*, *BA* and *CROWEX* were found to be related to reflectance in individual LANDSAT ETM+ spectral bands and the selected vegetation indices. The relationships between *DBH*, *H*, *BA* and *CROWEX* to individual LANDSAT ETM+ bands were negative, which is typical for both coniferous and deciduous forest stands in boreal and tropical zones (6,7,18,29). The reason for the negative correlations between these forest attributes and the reflectance in LANDSAT bands is the multiple layer absorption in the crown space which decreases the reflectance as the vertical dimension of the canopy increases. Therefore, the relationship is negative even in the near infrared band, although the NIR absorption of a single leaf or needle is low (29). The reflectance

saturates when the incoming radiation is almost completely absorbed by the canopy and the observed reflectance originates only from the uppermost layers of the canopy (6,30). Shadows are also likely to decrease the reflectance in all spectral bands as biomass increases. On the contrary, stem density correlated positively with LANDSAT ETM+ data. *STDENS* is a variable that generally depends on DBH and H. Stem density is lower in the forest stands with higher and thicker trees. Therefore, the relationships between *STDENS* and other forest structure parameters are strongly negative. Decreasing *STDENS* causes an increase of the vertical and horizontal dimensions of the canopy. That leads, as a consequence, to a decrease of reflectance in LANDSAT ETM+ bands due to the multiplayer absorption in the crown as it was suggested by (29).

The strength of the response of remote sensing data to forest structure parameters varied significantly between the individual LANDSAT ETM+ bands and the vegetation indices used in the analyses. Thus, visible bands, particularly bands 1 and 3 were statistically significant related to diameter at breast height, tree height, basal area, crown extension and stem density. However, the bands 5 and 7 showed higher response to all the examined forest structural parameters than the visible bands. With respect to *DBH*, our results agree with the results presented by (18), which found the best correlation of *DBH*, *H* and above ground biomass with the reflectance in bands 5 and 7. In this study, we did not investigate relationships of aboveground biomass to LANDSAT ETM+ bands. With regard to the value of the correlation coefficient, the total tree height was the forest parameter with the strongest observed correlations to individual LANDSAT ETM+ bands. The values of the correlation coefficient ranged from -0.70 to -0.86, demonstrating a good response of canopy spectral reflectance to the height of forest trees. Relationships of stem density and LANDSAT data from tropical rainforest regions have not been found in recent literature, while studies from the temperate and boreal zones have presented contradictory results. Some authors have found statistically significant relations between tree density and canopy reflectance (31,32). Others have found no significant relations between stem density and remote sensing data (6).

Generally, NIR band is considered to have a low response to forest structure parameters. A number of studies reported about much lower correlations of this band with forest structure parameters in comparison to other bands (6,31). This is one of the reasons for the low suitability of *NDVI* and *SR* for the prediction of forest structure parameters. (6) gave two explanations for the poor NIR-structure relation: (a) increase in NIR reflectance associated with tree growth and canopy closure is counteracted by decrease in reflectance associated with reductions in the amounts of background vegetation, which is visible to the sensor; (b) stronger spectral variability of background vegetation in the NIR than at other wavelengths, which is caused by variations in species composition and the age of background vegetation, confuses relationships between NIR reflectance and forest structure. However, the poor response of NIR to forest parameters is not expected in every case. The value of the correlation coefficient between NIR band and forest structure parameters is site-dependent and can vary in a wide range as it was shown by (18); they found strong correlations between NIR and forest parameters in three regions from four studies. Their results correspond with the findings in the present study. It was found that NIR band had a weak response to two of the four forest parameters under study. The values of the correlation coefficient between *DBH* and NIR were very low in comparison to other individual bands, and NIR showed no significant correlation with *STDENS*. Nonetheless, the correlations of band 4 with *H*, *BA* and *CROWEX* were only at some extent weaker than those of other individual bands.

Relationships between forest structural parameters and vegetation indices

NDVI and other vegetation indices derived from red and NIR bands have been widely used as general indicators for vegetation activity such as fractional vegetation cover, leaf area index, amount of photosynthetically active radiation (13,32). The use of vegetation indices for estimating forest parameters has brought contradictory results. Some studies have reported about very low performance of *NDVI* and other vegetation indices for satellite-based mapping forest structure parameters (4,5). Other studies have successfully used vegetation indices to predict forest aboveground biomass (9,19) and other forest parameters (18). The results presented here support the

second group. The present study indicated that the remotely sensing-derived multi-spectral vegetation indices can be used for prediction of forest structure parameters, even though the relationship between the variables were not especially strong with the exception of the tree height parameter that demonstrated a value of the correlation coefficient -0.86. The relations of *NDVIc* to *DBH* and *H* were as strong as those of the individual LANDSAT ETM+ bands.

In summary, the results of the correlation analysis in the present study agree with the results of recent studies on remote sensing-based estimations of forest parameters in tropical rainforest regions. In terms of correlation coefficient, the strength of the relationship observed in this study was similar to that reported by (18) for the Brazilian Amazon Basin and for a region in Borneo, Indonesia by (12). Some studies from the boreal regions have reported a somewhat stronger response of satellite data to the forest structural parameters than it was observed in the present study, see e.g. (19). Nevertheless, it is acknowledged that tropical rainforests are characterised by a generally lower response to the reflectance in satellite image than boreal forest. In terms of the correlation direction, the canopy of all forest types usually demonstrates negative correlation with visible, near-infrared and middle-infrared wavelengths, a fact that was revealed in the present study, too.

Predicting forest structural parameters

The present study used multivariate non-linear regression models for predicting forest structural parameters. All the established models were statistically significant at the level of $p<0.05$. Nevertheless, the performance and prediction ability of the models differed between the forest parameters. The study results disclosed that the predicting of *STDENS* is rather impossible with sufficient accuracy. The *STDENS* model had a very weak explanation power ($R^2 = 0.29$) and relatively high values of *RMSE* and bias. The other four models explained much more variance of the dependent variables. The best prediction power and accuracy was achieved by predicting the total tree height. This model explained about 80% of variance in the ground-based *H* data set and was sufficiently accurate with a value of *RMSE* = 3.5 m. The models for *DBH*, *CROWEX* and *BA* explained about 60% of variability of the dependent variables.

Regarding the suitability of the regression models for mapping forest structural parameters, the present results gave low R^2 (except tree height) (0.60-0.61 for *DBH*, *BA* and *CROWEX*). A number of studies on forest structural parameters with remote sensing models gave similar or even lower R^2. A regression model of diameter at breast height presented by (15) explained 65% of the variance in the ground-based data; a regression model used for mapping above ground biomass by (7) gave a value of $R^2 = 0.55$. A neural network based model for estimating aboveground biomass developed by (4) explained about 64% of the variance.

CONCLUSIONS

The present study investigated the suitability of LANDSAT ETM+ imagery for estimating forest structure parameters in a tropical rainforest in Indonesia. Field data from 60 plots including diameter at breast height, total tree height, stem density, basal area and crown diameter were compared with spectral data in individual LANDSAT ETM+ bands and multi-spectral vegetation indices using correlation and non-linear regression analysis. The results show that LANDSAT ETM+ imagery is generally suitable for predicting forest structure parameters over tropical rainforest.

More specifically bands 3, 5 and 7, and the multi-spectral vegetation indices can be used to derive maps of diameter at breast height, total tree height, basal area and crown extension, even though with different certainty. Among the forest structural parameters, the total tree height is the most reliably predicted parameter in this study. Tree density cannot be mapped reliably from the LANDSAT ETM+ imagery. The relationships between the LANDSAT ETM+ data and the other three parameters were not particularly strong. Nevertheless, these parameters could be modelled

with sufficient accuracy. The best models of those developed provide new insights into remote sensing applications for forest structure parameters in tropical regions.

ACKNOWLEDGEMENTS

This study was performed as a part of the research project *Stability of Rainforest Margins in Indonesia* (STORMA) funded by the Deutsche Forschungsgemeinschaft. The author would like to thank BSc C. Zsyschka for her great help during the field survey in November-December 2007. Acknowledgments are also extended to PhD M. Kessler and PhD J. Clough who friendly provided forest inventory data from their surveys for the study.

REFERENCES

1 Brown S, 1997. Estimating biomass and biomass change of tropical forests. FAO Forestry Paper, Vol 134, (Rome, FAO) http://www.fao.org/docrep/W4095E/W4095E00.htm (last date accessed: 20.08.2009)

2 Rosenqvist A, A Milne, R Lucas, M Imhoff & C Dobson, 2003. A review of remote sensing technology in support of the Kyoto protocol. Environmental Science and Policy, 6: 441-455

3 Steiniger 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

4 Foody G M, M E Cutler, J Mcmorrow, D Pelz, H Tangki, D S Boyd & I Douglas, 2001. Mapping the biomass of Bornean tropical rain forest from remotely sensed data. Global Ecology & Biogeography, 10: 379-387

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

6 Puhr C B & D N Donoghue, 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

7 Muukkonen P & J Heiskanen, 2005. Estimating biomass for boreal forests using ASTER satellite data combined with stand-wise forest inventory data. Remote Sensing of Environment, 99: 434-447

8 Muukkonen P & J Heiskanen, 2007. Biomass estimation over a large area based on standwise forest inventory data and ASTER and MODIS satellite data: a possibility to verify carbon inventories. Remote Sensing of Environment, 107: 617-624

9 Gonzalez-Alonso F, S Merido-De-Miguel, A Roldan-Zamarron, S Garsia-Gigorro & J M Cuevas, 2006. Forest biomass estimation through NDVI composites. The role of remotely sensed data to assess Spanish forests as carbon sinks. International Journal of Remote Sensing, 27: 5409-5415

10 Dong J, R K Kaufmann, R B Myneni, C J Tucker, P E Kauppi, J Liski, W Buermann, V Alexeyev & M K Hughes, 2003. Remote sensing estimates of boreal and temperate forest woody biomass: carbon pools, sources, and sinks. Remote Sensing of Environment, 84: 393-410

11 Sivanpillai R, R Srinivasan, C T Smith, M G Messina & X Ben Wu, 2007. Estimating regional forest cover in East Texas using Advanced Very High Resolution Radiometer (AVHRR) data. International Journal of Applied Earth Observation and Geoinformation, 9: 41-49

12 Tangki H & N A Chappel, 2008. Biomass variation across selectively logged forest within a 225-km2 region of Borneo and its prediction by Landsat TM. Forest Ecology and Management, 256: 1960-1970

13 Potapov P, M C Hansen, S V Stehman, T Loveland & K Pittman, 2008. Combining MODIS and Landsat imagery to estimate and map boreal forest cover loss. Remote Sensing of Environment, 112: 3708-3719

14 Hansen M C, D P Roy, E Lindquist, B Adusei, O Justice & A Altstatt, 2008. A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin. Remote Sensing of Environment, 112: 2495-2513

15 Feldpausch T, J McDonald, C A M Passos, J Lehmann & J Riha, 2008. Biomass, harvestable area, and forest structure estimated from commercial timber inventories and remotely sensed imagery in southern Amazonia. Forest Ecology and Management, 233: 121-132

16 Tucker C J, C L Vanpra, M J Sharman & G Van Ittersum, 1985. Satellite remote sensing of total herbaceous biomass production in the Senegalese Sahel: 1980-1984. Remote Sensing of Environment, 17: 233-249

17 Fensholt R, Sandholt I & M S Rasmussen, 2004. Validation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements. Remote Sensing of Environment, 91: 490-507

18 Lu D, P Mausel, E Brondizio & E Moran, 2004. Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. Forest Ecology and Management, 198: 149-167

19 Zheng D, J Rademacher, J Chen, T Crow, M Bresee, J Le Moine & S R Ryu, 2004. Estimating aboveground biomass using Landsat ETM+ data across a managed landscape in northern Winconsin, USA. Remote Sensing of Environment, 93: 402-411

20 Teillet P M, B Guindon & D G Goodenough, 1982. On the slope-aspect correction of multispectral scanner data. Canadian Journal of Remote Sensing, 8: 84-106

21 Smith J A, T L Lin & K J Ranson, 1980. The Lambertian assumption and Landsat data. Photogrammetric Engineering & Remote Sensing, 46: 1183-1189

22 Twele A & S Erasmi, 2005. Evaluating topographic correction algorithms for improved land cover discrimination in mountainous areas of Central Sulawesi. In: Remote Sensing & GIS for Environmental Studies: Applications in Geography, edited by S Erasmi, B Cyffka & M Kappas (Verlag Erich Goltze, Göttingen) 287-295

23 ENVI Software website;
http://www.ittvis.com/ProductServices/ENVI.aspx (last date accessed: 20.09.2009)

24 Erasmi S, C Knieper, A Twele & M Kappas, 2007. Monitoring inter-annual land cover dynamics at the rainforest margin in Central Sulawesi, Indonesia. In: New Developments and Challenges in Remote Sensing, edited by S Bochenek, (Millpress, Rotterdam) 297-308

25 Chen M, G Pavlic, L Brown, J Cuhlar, S G Leblanc, H P White, R J Hall, D R Peddle, D J King, J A Trofimow, E Swift, J Van der Sanden & P K E Pellikka, 2002. Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurements. Remote Sensing of Environment, 80: 165-184

26 Tucker C J & P J Sellers, 1986. Satellite remote sensing of primary vegetation. International Journal of Remote Sensing, 7: 1395-1416

27 Brown L J, J M Chen, S G Leblanc & J Cihlar, 2000. Shortwave infrared correction to the simple ratio: an image and model analysis. Remote Sensing of Environment, 77: 16-25

28 Eklundh L, L Harrie & A Kuusk, 2001. Investigating relationships between Landsat ETM+ sensor data and LAI in a boreal conifer forest. Remote Sensing of Environment, 78: 239-251

29 Häme T, A Salli, K Andersson & A Lohi, 1997. A new methodology for estimation of biomass of conifer-dominated boreal forest using NOAA AVHRR data. International Journal of Remote Sensing, 18: 3211-3243

30 Nilson T & U Peterson, 1994. Age dependence of forest reflectance: analysis of main driving factors. Remote Sensing of Environment, 48: 319-331

31 Danson F M & P J Curran, 1993. Factors affecting the remotely sensed response of coniferous forest plantations. Remote Sensing of Environment, 43: 55-65

32 Gemmel F M, 1995. Effects of forest cover, terrain, and scale on timber volume estimation with thematic mapper data in a Rocky Mountain site. Remote Sensing of Environment, 51: 291-305

33 Asrar G M, M M Fuchs, E T Kanemasu & J L Hatfield, 1984. Estimating absorbed photosynthetically active radiation and leaf area index from spectral reflectance in wheat. Agronomy Journal, 87: 300-306

34 Gujarati D N, 2004. Basic Econometrics. 4th edition (McGraw-Hill, New York) 1032 pp.