Examining spectral properties of Landsat 8 OLI for predicting above-ground carbon of Labanan Forest, Berau

A Suhardiman*, B A Tampubolon, M Sumaryono

Faculty of Forestry, University of Mulawarman

*Corresponding Author: suhardiman@fahutan.unmul.ac.id

Abstract. Many studies revealed significant correlation between satellite image properties and forest data attributes such as stand volume, biomass or carbon stock. However, further study is still relevant due to advancement of remote sensing technology as well as improvement on methods of data analysis. In this study, the properties of three vegetation indices derived from Landsat 8 OLI were tested upon above-ground carbon stock data from 50 circular sample plots (30-meter radius) from ground survey in PT. Inhutani I forest concession in Labanan, Berau, East Kalimantan. Correlation analysis using Pearson method exhibited a promising results when the coefficient of correlation (r-value) was higher than 0.5. Further regression analysis was carried out to develop mathematical model describing the correlation between sample plots data and vegetation index image using various mathematical models. Power and exponential model were demonstrated a good result for all vegetation indices. In order to choose the most adequate mathematical model for predicting Above-ground Carbon (AGC), the Bayesian Information Criterion (BIC) was applied. The lowest BIC value (i.e. -376.41) shown by Transformed Vegetation Index (TVI) indicates this formula, $AGC = 9.608*TVI^{21.54}$, is the best predictor of AGC of study area.

1. Introduction

The advancement of satellite-based remote sensing technology brings the opportunity to study forests much better. Various types of satellite imagery are available in many different spatial resolution. Landsat 8 OLI (Operational Land Imager) is the latest generation of NASA (National Atmospheric and Space Administration) earth resource satellite series which was launched in 2013. This new satellite persist its 30-meter spatial resolution to match with the previous generation. The good thing about Landsat is that all image products are completely free and available worldwide through USGS website. For country like Indonesia, Landsat images have been used as main source of land cover mapping since 1990. On the other hand, for remote sensing communities, Landsat spectral properties is always interesting to be explored and studied in various fields of research and applications.

Although there are encouraging results in utilizing remote sensing data to study forest attributes such as basal area, stem volume or aboveground biomass in larger areas [1], correlation between the vegetation cover on the satellite images with forest attributes (e.g. above-ground biomass) can be very poor [2] especially in the complex tropical forest where vegetation are composed from hundreds of species and having different ages. Higher variability in tropical forest structure is major reasons for uncertainties in estimating forest attributes [3], [2]. The variation depends on factors such as soil type, topographic position [4], soil nutrients [5] and levels of human disturbance [6].

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

In order to increase accuracy, some studies suggested utilization of high spatial resolution image or Lidar data [7]. Nonetheless, in case of Indonesia, only a few of government bodies or private companies including forest concession holders willingly spent their money to purchase high spatial resolution image. Depends on the project scale, some high spatial resolution images were bought limited to particular location. The fact is that most of forest concessions are still using moderate spatial resolution images such as Landsat [8] to be used for forest areas mapping and monitoring activities.

Hence, the accurate estimation of structural characteristics of tropical forest vegetation remains as a major obstacle [9]. The promising results are mostly obtained from studies on conifer forest, which consist of pure stands dominated by a single tree species [10], [11], [12]. For optical remote sensing, various image transformation methods such as vegetation indices and texture analysis have been used to estimate forest attributes and so far the accuracy depends on the site, forest type as well as spatial resolution of the image. In this study, the spectral properties of relatively new Landsat 8 OLI were examined in order to predict above-ground carbon (AGC) at large scale based on pixel values and sample plots data.

2. Method

2.1. Study Area

The study area covered approximately 1,380 square kilometers (km^2)of PT. Inhutani I forest concession at Labanan, Berau District, East Kalimantan. Various land cover types were presented in this area but two most dominant are primary lowland Dipterocarp forest (51%) and secondary lowland Dipterocarp forest (23%) which is source for most valuable and commercial timber from this forest. Figure 1 below shows the situation map of study area and the distribution of sample plots.

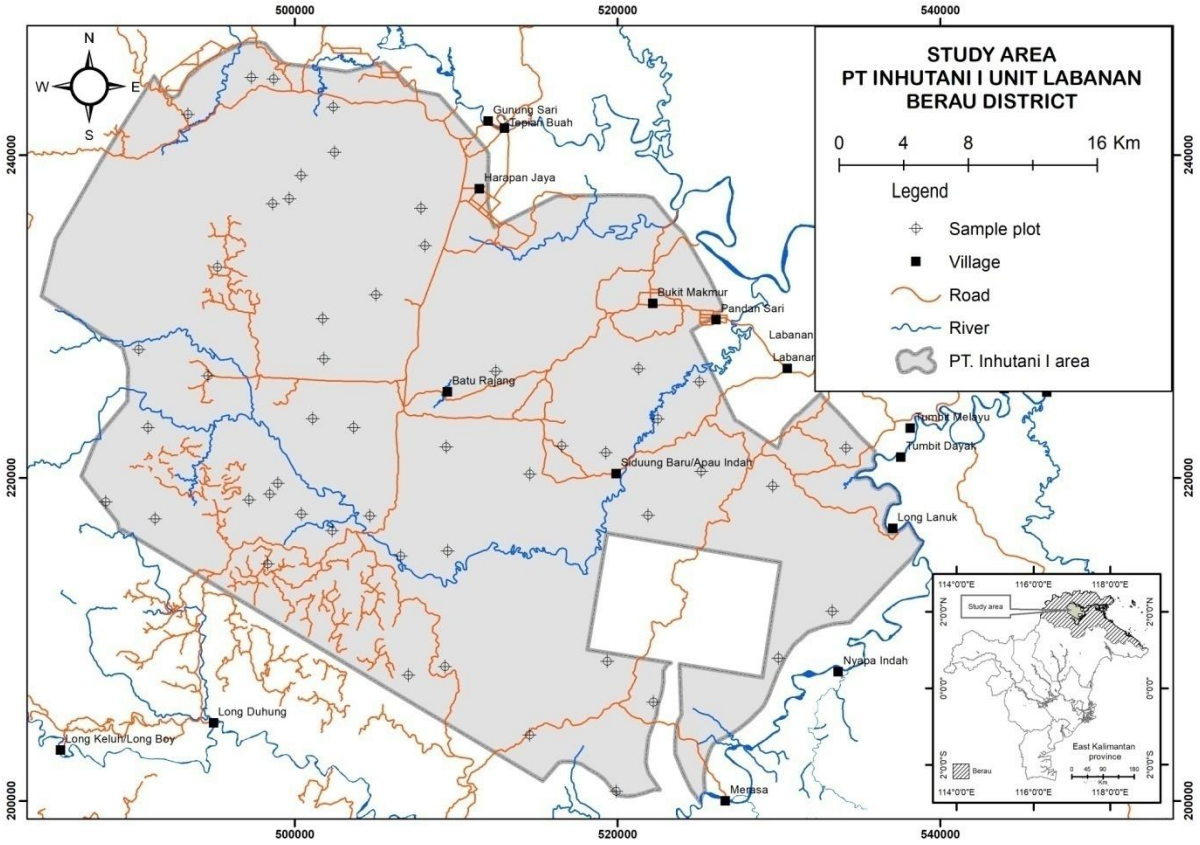


Figure 1. Study area and sample plot distribution

2.2. Materials

There are two set of data needed for this study and therefore will be explained in this section. First data related to the Landsat 8 OLI images which have been transformed into three spectral vegetation indices. Landsat imagery is continuous type of data contains grid pixel values which indicate the magnitude of spectral reflectance over objects on earth. Other dataset was AGC (ton C/ha) from sample plots which was collected in September 2016.

2.2.1. Satellite data. Landsat 8 OLI path/row 117/059 images were downloaded from http://earthexplorer.usgs.gov and two acquisition periods were selected on May and September 2016. In tropics, the persistent of cloud is major issue, especially when dealing with optical remote sensing products such as Landsat. Therefore finding images on Landsat catalog sometimes very challenging. One image scene may not be sufficient due to cloud which was cover up the interest point or location of the study.

Landsat 8 OLI brought more sensors than its predecessor. The sensing product of sensor is called band. The pixel value of band represents bit-package combination of surface, atmosphere, and sensor condition that can affect the overall usefulness of a given pixel [13]. Furthermore [14]mentioned that vegetation indices may be sensitive to the atmospheric condition. Therefore in this study, radiometric correction was applied to pixel band so that atmospheric factor can be excluded using Top of Atmosphere (ToA) formula as provide in [15].

Among nine Landsat 8 OLI spectral bands, near-infrared (NIR) and red (R) band are essential to study biomass content of vegetation. NIR and R band were used in this study to transform pixel values to vegetation index values. There are numerous vegetation index which has been developed and some of them were exhibited a positive correlation withstand or other forest properties.In this study, three vegetation indices that is Normalized Difference Vegetation Index (NDVI), Single Ratio Vegetation Index (SRVI) and Transformed Vegetation Index (TVI) were examined. Landsat pixel value was transformed to vegetation index value using these following formula:

$$NDVI = \frac{NIR - R}{NIR + R} \quad \ldots\ldots (1)$$

$$SRVI = \frac{NIR}{R} \quad \ldots\ldots (2)$$

$$TVI = \sqrt{NDVI + 0.5} \quad \ldots\ldots (3)$$

Where NIR is near infrared band and R is red band of Landsat 8 OLI. The transformation resulted a continuous layer of NDVI, SRVI, and TVI with similar format to Landsat 8 OLI unless the pixel values were altered. The new pixel value were critical in this study since it role as predictor of value from ground sample plot data.

2.2.2. Sample plot data. As many as 50 sample plots were established during field surveying September 2016. Simple random sampling were chose to pinpoint the sample plots location so that bias would be minimized. Four team of surveyor were set up to do forest inventory in the plots. GPS receiver helped teams to navigate plots. Once the survey team arrived at targeted location, a 30-meter radius of circular sample plots were then established. Sample plots were also nested to capture different strata of vegetation. Smaller radius of plots captured smaller trees. A 30-meter radius sample plot is set to collect data for big trees (dbh $\geq$ 50 cm) while the smallest subplot (radius equals to 1.5 m) were collected seedling data (juvenile trees where dbh < 2 cm and height < 1.5 meter). Figure 2 shows layout of the sample plot established on the ground.

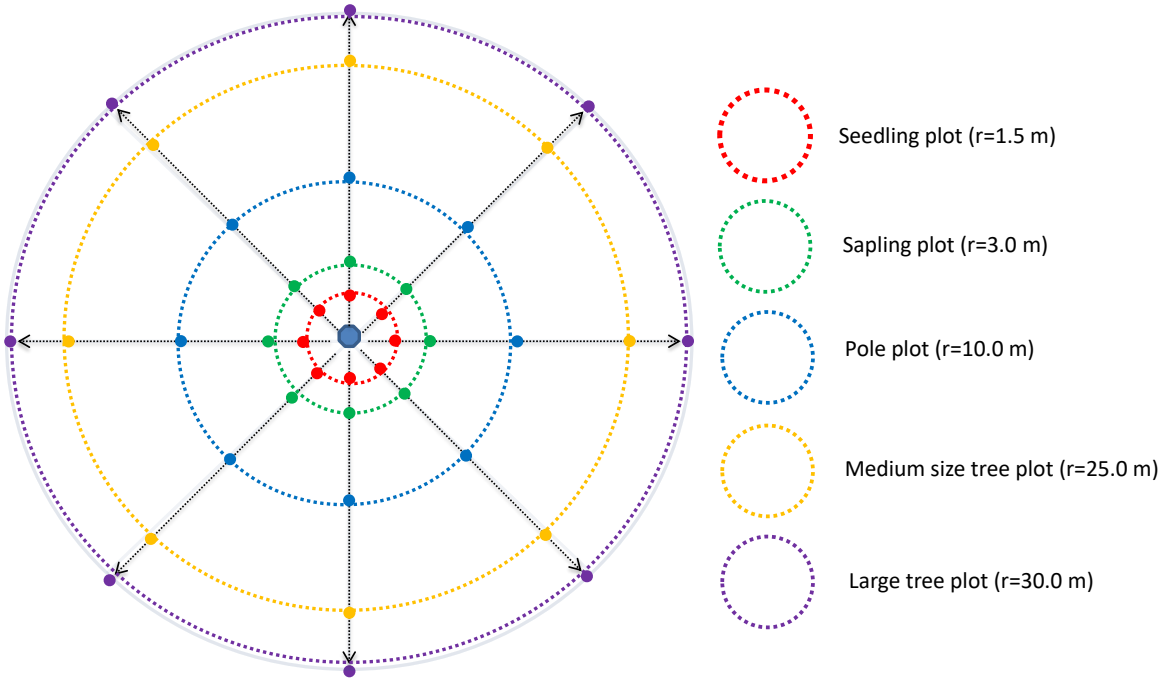


Figure 2. Layout of sample plots on the ground

Diameter at breast height (dbh) of the trees, pole, and sapling were measured in each sample plots and subplots while species/genus/family name and land cover types were carefully identified. The center of the sample plot were marked using GPS receiver in order to update the plot position on maps.Eventually, allometric equation of [16] were used to estimate above-ground biomass (AGB) content for each trees.

$$Ln(AGB) = -2.289 + \left(2.649 \cdot Ln(Dbh)\right) - (0.021 \cdot Ln(Dbh)^2). \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Where AGB above-ground biomass (kg) and Dbh is diameter at breast height (cm). Conversion of AGB to AGC simply by multiplying AGB with a constant value of 0.47 [17]. The estimation of AGC (ton C ha^{-1}) for each sample plot was calculated by summing up all carbon trees and multiplied by area conversion factor.

2.2.3. Statistical data analysis. The position of 50 sample plots were overlaid with three-layer data of vegetation indices in geographic information system (GIS) environment. Position of sample plots were drew as circle polygon with radius 30 meter rather than points. Pixel values that correspond to the polygon of sample plot position were averaged. At the end, 50 carbon stocks from sample plots have their corresponding pixel value.

Furthermore, statistical analysis was carried out to denote potential relationship between carbon stock and pixel value. Different mathematical function were tested i.e. linear and non-linear (logarithmic, quadratic, power and exponential) to develop befittingrelationship models. The coefficient of Pearson correlation (r) were used to measure how well two sets of data are related to each other.Theoretically r-value closed to 1 exhibits strong positive correlation.

In this study, models with r-value greater than 0.5 were selected to be further analyzed using Bayesian Information Criterion (BIC). BIC is a tool to estimate the quality of mathematical or statistical model of a given dataset. The model with lowest BIC is preferred [18].For multi datasets, BIC is calculated using following formula.

$$BIC = \sum_{i=1}^{N} \left(n_i \cdot ln \frac{RSS}{n_i} \right) \cdot + k \cdot N \cdot \ln(n_t) \quad (5)$$

Where BIC is Bayesian Information Criterion, RSS is residual sum of square of regression, k is number of parameters used in the model, N is number of datasets, n_t= total number of data point used in the set (i.e. in all N datasets), i is index of dataset 1, 2,..., N, and n_i is the number of points in dataset 1, 2,..., N.

3. Results and discussion

3.1. Above-carbon stock of sample plots.

The characteristics of Labanan forest of PT.Inhutani I could be recognized from sample plots data. Tropical forest of Kalimantan has complex stand structure and diverse species composition. That is the reason why variation among forest properties (e.g. stand volume, AGB or AGC) is relatively high in tropics especially where forest has been disturbed by logging or any other human interference. Area of PT Inhutani I is an example of lowland dipterocarp forest that has been logged for years so that different forest cover type exist as an impact of logging activities including encroachment and et cetera.

Eighty-four percent of the total sample plots were randomly located at lowland dipterocarp forest as major land cover type over the study area. The average AGC approximately 135.18 ton C/ha with standard deviation is 57.16 ton C/ha. Hence the variation in this area is about 42.28%. The minimum AGC is 11.12 ton C/ha while the maximum one is 287.32 ton C/ha. This average carbon stock value is slightly lower than the average value of Indonesia secondary lowland forest carbon stock which is 169 ton C/ha.

3.2. Correlation of AGC and vegetation index.

Table 1 below summarizes the relationship between AGC and pixel value of three different vegetation indices. The degree of relationship denoted by r-value exhibit a promising result. All three indices showed r-value greater than 0.5. According to [19], R-values between ±0.3 and ±0.7 indicate a moderate positive or negative linear relationship. The relatively low r-value as result of this study was admittedly predicted.[2] mentioned that correlation between the vegetation cover on the satellite images with forest attributes (e.g. aboveground biomass) can be very poor especially in the complex tropical forest where vegetation are composed from hundreds of species and having different ages.Furthermore, [20] stressed and argued that low coefficient of correlation is often showed by most of optical remote sensing data analysis when it used to estimate forest attributes.

Table 1. Pearson correlation coefficient of each vegetation index

Vegetation index	Coefficient of Pearson correlation (r)	Significant (2-tailed)
NDVI (*Normalized Difference Vegetation Index*)	0.523	0.000
SRVI (*Single Ratio Vegetation Index*)	0.528	0.000
TVI (*Transformed Vegetation Index*)	0.517	0.000

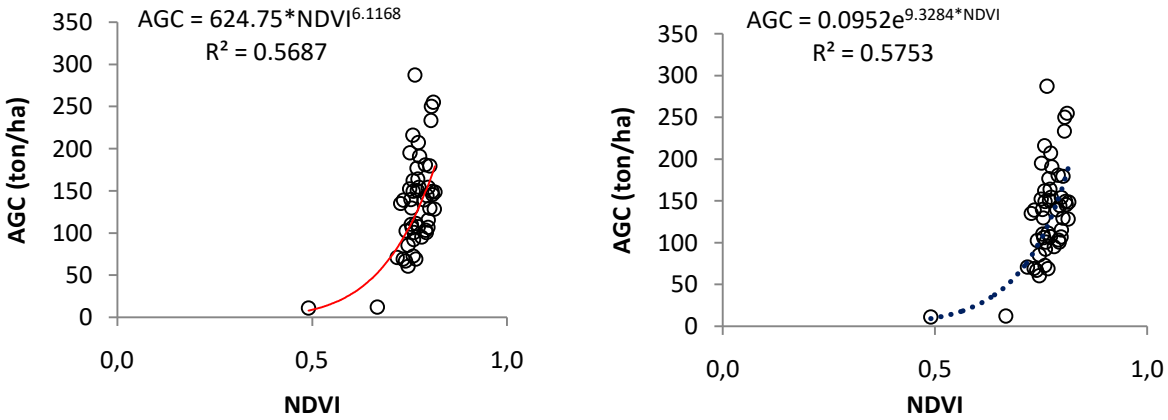
Table 2 showed mathematical models of vegetation index as predictor. Different behavior were demonstrated by each model. Linear model as well as logarithmic and quadratic exhibited lower coefficient of determination (R^2) compare to power and exponential model. This finding suggested

that correlation between AGC and vegetation index is adequate to be modeled by power and exponential function for all vegetation indices.

Table 2. Regression model of vegetation index as predictor using five mathematical functions

Indices	Regression / Mathematical model	R^2	R^2 Adjusted	Model
NDVI	-330.1 + 608.6*NDVI	0.272	0.257	Linear
SRVI	-50.84 + 23.97*SRVI	0.278	0.263	Linear
TVI	-1319 + 1294*TVI	0.267	0.252	Linear
NDVI	240.4 + 388.3*ln NDVI	0.255	0.239	Logarithmic
SRVI	-191.8 + 160.8*ln SRVI	0.287	0.272	Logarithmic
TVI	-26.34 + 1382*ln TVI	0.262	0.247	Logarithmic
NDVI	$300.7 – 1300*NDVI + 1411*NDVI^2$	0.295	0.265	Quadratic
SRVI	$-161 + 55.7*SRVI - 2.202*SRVI^2$	0.291	0.261	Quadratic
TVI	$7275 - 1.465e+04*TVI + 7379*TVI^2$	0.296	0.266	Quadratic
NDVI	$624.8*NDVI^{6.117}$	0.569	0.560	Power
SRVI	$1.026*SRVI^{2.338}$	0.546	0.536	Power
TVI	$9.608*TVI^{21.54}$	0.573	0.564	Power
NDVI	0.0952*Ln(9.3284*NDVI)	0.575	0.567	Exponential
SRVI	$9.507*1.385^{SRVI}$	0.462	0.451	Exponential
TVI	$(2.035*10^{-8})*(4.875*Ln(20.005*TVI))$	0.574	0.565	Exponential

Figure 3 below showed the scatter plot diagram as well as the trendline of dataset using power and exponential function which explore the relationship between vegetation index and AGC from sample plots. All vegetation indices showed a promising result where R^2 value slightly higher than 0,5 except SRVI model developed using polynomial function (Figure 3d). Either NDVI and TVI, both indices demonstrated narrow value range which indicate variation among pixels are a bit lower. In contrast, the variability of AGC from sample plots showed wider range that indicate higher variation of AGC value between sample plots.



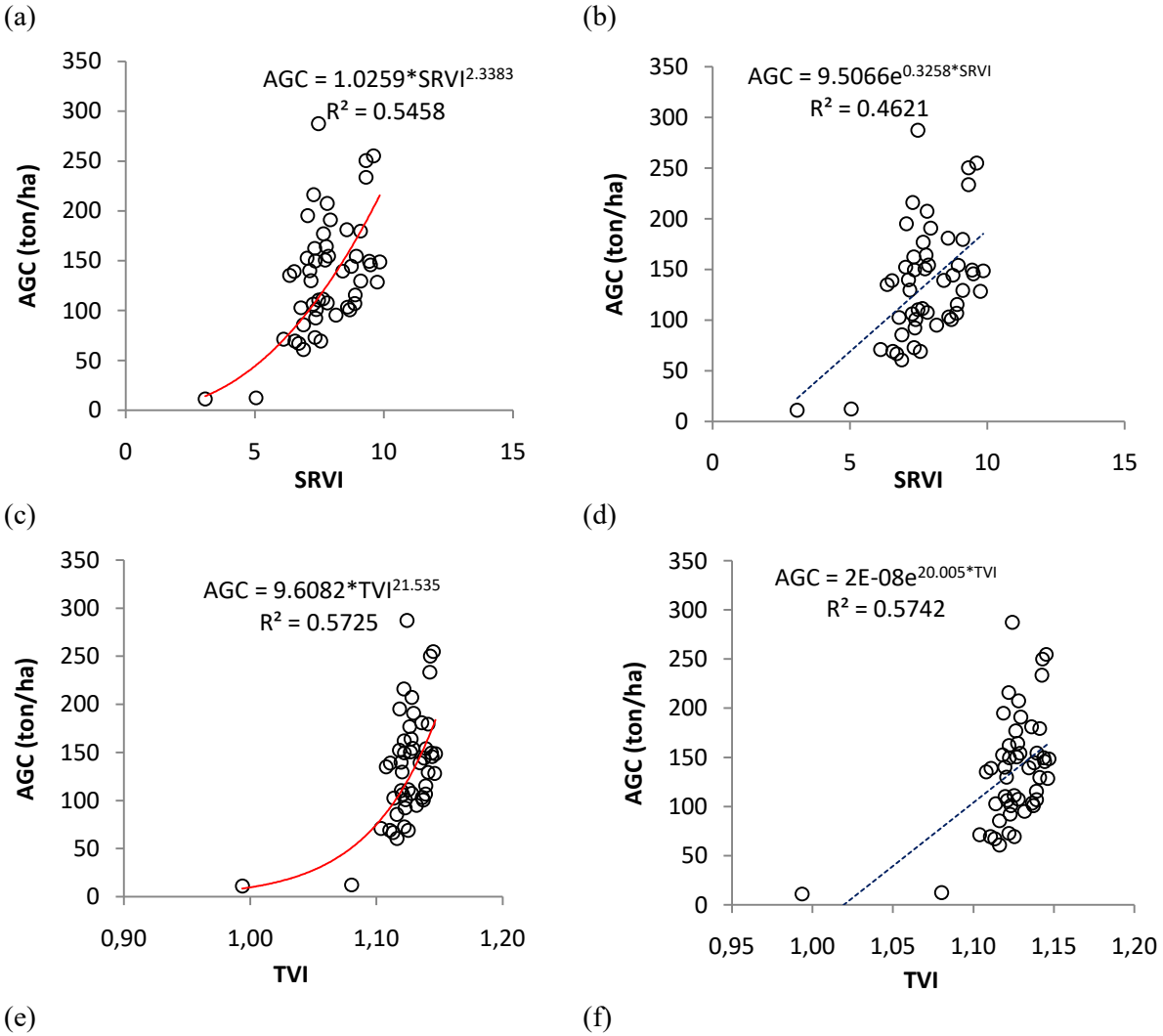


Figure 3. Graphical relationship between AGC and vegetation index using power function (a, c and e) and exponential function (b, d, and f).

Table 3 below showed the calculated BIC values only for power and exponential models. Lowest BIC value is -376.41 and it belongs to the mathematical model where TVI is the predictor. Then the selected mathematical model for predicting AGC (ton/ha) in PT. Inhutani I Unit Labanan in Berau district is AGC = $9.608*TVI^{21.54}$

Table 3. The calculated Bayesian Information Criterion (BIC) for selecting best mathematical model

	Mathematical model	N	n_i	n_t	k	RSS	BIC
Power	$624.8*NDVI^{6.117}$	3	50	150	3	0.051	-299.55
Power	$1.026*SRVI^{2.338}$	3	50	150	3	35.196	27.54
Power	$9.608*TVI^{21.54}$	3	50	150	3	0.011	-376.41
Exponential	0.0952*Ln(9.3284*NDVI)	3	50	150	3	0.050	-300.25
Exponential	$9.507*1.385^{SRVI}$	3	50	150	3	41.709	36.03
Exponential	$(2.035*10^{-8})*(4.875*Ln(20.005*TVI))$	3	50	150	4	0.011	-361.49

4. Conclusion

Landsat products are the most widely used satellite imagery in Indonesia due to its extensive coverage and available at no-cost. Therefore exploration of the capability of Landsat product especially Landsat

8 OLI to predict land-based properties is always interesting subject. This study revealed once again that spectral properties of Landsat image as reflectance of energy from the sun exhibited a moderate correlation with forest properties which is above-ground carbon stock.The complexity of forest structure is one factor preventing coefficient of correlation to meet high value i.e. 0.7 or 0.8. Landsat pixel value captures only the emergent trees canopy while co-dominant and suppressed trees are unable to be recognize. Another factor is related to the pixel size of Landsat that seems too large (30 × 30 m) to capture underlying variability of certain objects on the ground such as AGC. Therefore further study may be needed to discriminate AGC based on the forest stratum before correlation with pixel values is undertaken.

Acknowledgments
This study supported from GIZ Forclime Forests and Climate Change Programme and Forestry office of Berau District.

References

[1] Odzemir I, Karnieli A. 2011. Predicting Forest Structural Parameters Using the Image Texture Derived from WorldView-2 Multispectral Imagery in a Dryland Forest, Israel. International Journal of Applied Earth Observation and Geoinformation 13: 701-710.
[2] Saatchi S, Houghton R, Dos Santos Alvala R, Soares J, Yu Y. 2007. Distribution of Aboveground Live Biomass in the Amazon Basin. Global Change Biology 13: 816-837.
[3] Intergovernmental Panel on Climate Change. 2001. Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Editor Houghton J, Ding Y, Griggs D, Noguer M, van Linden P, Dai X. Cambridge: Cambridge University Press
[4] Clark D, Clark D. 2000. Landscape-scale Variation in Forest Structure and Biomass in a Tropical Rain Forest. Forest Ecology and Management 137: 185-198.
[5] Laurence W, Fearnside P, Laurance S, Delamonica P, Lovejoy T, Rankin de Merona J. 1999. Relationship Between Soils and Amazon Forest Biomass: a Landscape-scale Study. Forest Ecology and Management 118: 127-138.
[6] Urquiza-Haas T, Dolman P, Peres C. 2007. Regional Scale Variation in Forest Structure and Biomass in the Yucatan Peninsula, Mexico: Effects of Forest Disturbance. Forest Ecology and Management 247: 80-90.
[7] Powel S, Cohen W, Healey S, Kennedy R, Moisen G, Pierce K. 2010. Quantification of Life Aboveground Forest Biomass Dynamics with Landsat Time-series and Field Inventory Data: A Comparison of Empirical Modeling Approaches. Remote Sensing of Environment 114: 1053-1068.
[8] Lu D. 2005. Aboveground Biomass Estimation Using Landsat TM Data in the Brazilian Amazon. International Journal of Remote Sensing 26 (12): 2509-2525.
[9] Drake JB, Dubayah RO, Clark DB, Knox RG, Blair JB, Hofton MA. 2002. Estimation of Tropical Forest Structure Characteristics Using Large-footprint Lidar. Remote Sensing of Environment 79: 305-319.
[10] Kayitakire F, Hamel C, Defourny P. 2006. Retrieving Forest Structure Variables Based on Image Texture Analysis and IKONOS-2 Imagery. Remote Sensing of Environment 102 (3-4): 390-401.
[11] Sprintsin M, Karnieli A, Berliner P, Rotenberg E, Yakir D, Cohen S. 2007. The Effect of Spatial Resolution on the Accuracy of Leaf Area Index Estimation for a Forest Planted in the Desert Transition Zone. Remote Sensing of Environment 109: 416-428.
[12] Wolter P, Townsend P, Sturtevant B. 2009. Estimation of Forest Structural Parameters Using 5 and 10 meter SPOT-5 Satellite Data. Remote Sensing of Environment 113: 2019-2036.
[13] Anonymous. 2017. Landsat 8 Pre-Collection Quality Assessment Band. Available online at https://landsat.usgs.gov/qualityband. Accessed on November 29, 2017

[14] Freitas S, Mello M, Cruz C. 2005. Relationship Between Forest Structure and Vegetation Indices in Atlantic Rainforest. Forest Ecology and Management 218: 353-362.
[15] Anonymous. 2017. Using the USGS Landsat 8 Product. Available online at http://landsat.usgs.gov/using-usgs-landsat-8-product. Accessed on November 29, 2017.
[16] Brown S. 1997. Estimating Biomass and Biomass Change of Tropical Forests: A primer. Rome: FAO Forestry Paper.
[17] Standard Nasional Indonesia. 2011. Pengukuran dan Perhitungan Cadangan Karbon – Pengukuran Lapangan untuk Penaksiran Cadangan Karbon Hutan (Ground based forest carbon accounting). Badan Standardisasi Nasional. Jakarta.
[18] Navakatikyan MA. 2007. A Model for Residence Time in Concurrent Variable Internal Performance. Journal of the Experimental Analysis of Behavior 87: 121-141.
[19] Ratner B. 2003. Statistical Modeling and Analysis for Database Marketing. Effective Techniques for Mining Big Data. Florida: Chapman & Hall/CRC.
[20] Gibbs H, Brown S, O'Niles J, Foley J. 2007. Monitoring and Estimating Tropical Forest Carbon Stock: Making REDD a Reality. Environ. Res. Lett. 2: 1-13.