COMPARISON OF MODEL ACCURACY IN TREE CANOPY DENSITY ESTIMATION USING SINGLE BAND, VEGETATION INDICES AND FOREST CANOPY DENSITY (FCD) BASED ON LANDSAT-8 IMAGERY (CASE STUDY: PEAT SWAMP FOREST IN RIAU PROVINCE)

Faisal Ashaari[1] , Muhammad Kamal[2] , Dede Dirgahayu[1]*

[1]Remote Sensing Applications Center, LAPAN

[2]Faculty of Geography, Gadjah Mada University

*e-mail: dede.dirgahayu@lapan.go.id

Received: 6 November 2017; Revised: 11 May 2018; Approved: 22 June 2018

Abstract. Identification of a tree canopy density information may use remote sensing data such as Landsat-8 imagery. Remote sensing technology such as digital image processing methods could be used to estimate the tree canopy density. The purpose of this research was to compare the results of accuracy of each method for estimating the tree canopy density and determine the best method for mapping the tree canopy density at the site of research. The methods used in the estimation of the tree canopy density are Single band (green, red, and near-infrared band), vegetation indices (NDVI, SAVI, and MSARVI), and Forest Canopy Density (FCD) model. The test results showed that the accuracy of each method: green 73.66%, red 75.63%, near-infrared 75.26%, NDVI 79.42%, SAVI 82.01%, MSARVI 82.65%, and FCD model 81.27%. Comparison of the accuracy results from the seventh methods indicated that MSARVI is the best method to estimate tree canopy density based on Landsat-8 at the site of research. Estimation tree canopy density with MSARVI method showed that the canopy density at the site of research predominantly 60-70% which spread evenly.

Keywords: *Tree canopy density, single band, vegetation indices, FCD.*

1 INTRODUCTION

Forest is a unity of ecosystem in the form of expanse of land containing biological resources dominated by trees in its natural environment, one with the others cannot be separated (Article 1 of the Law of Republic of Indonesia No. 41 1999). Ayat and Tarigan (2010) state that the forest as a natural resource that is priceless because it can be a source of life for humans. However, there has an increasing pressured on forest resources due to deforestation and degradation (Wibowo and Ginting 2010). Based on data from Forest Watch Indonesia (FWI) (2014), the pressure is indicated by the extent of forest cover decreasing from 87 million hectares to 82 million hectares in the period 2009 - 2013.

Forest canopy covered is the top of the vegetation that provides protection to the underlying environment (Lund 2002). The canopy cover, which refers to the proportion of the whole canopy to the unit area per surface area, is called the density of the canopy. The existence of canopy density information is important to know because the canopy of forest cover can assess the forest status or intervention indicators in forest resource management as indicated by the decrease in standing quality, although

http://dx.doi.org/10.30536/j.ijreses.2018.v15.2845

the forest cover class does not change (Tohir *et al.* 2014).

According to Panta (2003), information about forest canopy density is directly or indirectly related to environmental problems such as erosion, resource degradation, biodiversity loss, and climate change. The tree part of the forest structure that has the highest canopy height is very important to protect the bottom as fauna habitat (APOEC.org 2003). Therefore, the canopy density information, especially tree canopy density is considered significant to support the further studies.

Forest inventories such as tree canopy density information requires a long time and high cost, especially in extensive forest coverage. Remote sensing technology that directly captures the surface vegetation character (Campbell *et al.* 2011) with extensive viewing (synoptic view) coverage (Matthews 2013), can be used as an alternative to perform such inventory activities as Landsat-8 image which have the advantages like spectral, temporal, radiometric resolution and free access.

The basic study of the canopy density through remote sensing can use a variety of approaches. Some approaches such as a single band, vegetation index and developed FCD models have different characteristics, thus potentially result in different relationships and different accuracy rates in the estimation of tree canopy density. The analysis of the relationship and the accuracy of each method approach require density data of tree canopy in the field.

The research site was conducted in Riau peat swamp forest. Locations in the area of Biosphere Reserves Giam Siak Kecil-Bukit Batu (CB GSK-BB) precisely in the core zone of Bukit Batu Wildlife Sanctuary. GSK-BB CB area has a very high biodiversity living in various ecosystems that are still intact (MAB-Indonesia.org 2016a). But nowadays there have been forest encroachment activities that reduced the forest cover in this area. These activities have reached buffer zones and core zones in 2015 (Tempo.co 2015a and 2015b). In fact, these two zones should be protected from the negative impacts of human activities (MAB-Indonesia.org 2016b). The existence of such things, it is a necessary activities such as forest inventory density tree canopy for forest management, especially in the reserve of the biosphere as a protected area of the world.

The selection of the site is located in part of CG GSK-BB peat swamp forest aims to know the best method of digital image processing, especially in peat swamp forest because there is currently no standard method that related to the characteristics of the region. Lillesand *et al.* (2014) said that the existence of spatial effects in different geographic locations will affect the spectral reflection even though the object is similar. In addition, each of methodical approaches has the different characteristics, thus it potentially has different relationships and different levels of accuracy. From both of these things, it is necessary to study the comparative results of several methods in obtaining spatial information density of tree canopy properly and correctly. The objective of this research is to find the best parameter to estimate Tree Canopy Density using Landsat 8 data.

2 MATERIALS AND METHODOLOGY

2.1 Tools and Materials

Tools:

1. Personal Computer
2. Software for data processing and analysis: ENVI 4.8, FCD Mapper

V.2, CANEYE V6.3.13, Ms. Office 2016 and SPSS 17.

3. Field equipment:
   -Global Navigation Satellite System (GNSS).
   -Nikon D5100 DSLR Camera + Fisheye Lens and Niveau.
   -Tape measure, Compass and Checklist.

Data:

1. Landsat-8 image path / row 126/59 level 1T recording on July 10, 2015.
2. Indonesian Topographical Map Bukit Batu Sheet 0817-34 scale 1: 50.000.
3. Map of Forest Utilization Agreement (TGHK) Riau Province 2014.
4. Vector Data of Land Use of Riau Province SK. 878 / Menhut-II / 2014.
5. Vector Data of Biosphere Reserves of Giam Siak Kecil - Bukit Batu.

2.2 Data Processing

a. Image correction

- Geometric correction

This research used the level-1T of Landsat-8 image, the level-1T indicates that the image has been corrected geometrically and radiometrically globally. However, the correction is required locally, geometric correction was done by image-to-map correction with reference data used is the RBI map. The interpolation method used is nearest neighbor.

- Radiometric correction

The radiometric correction was applied to correct Top of Atmosphere (ToA) effect of the OLI sensor for band 2 – 7, while the TIRS sensor was corrected to Brightness Value.

- Atmospheric correction

The ToA Reflectance result still has an error due to atmospheric disturbance where there is a scattering of energy in the atmosphere so that the atmospheric correction is needed. The method used for atmospheric correction is Histogram Adjustment that correct the ToA reflectance to the surface reflectance (at surface reflectance).

b. Image masking

Image masking is the process of selecting the area of the certain information. Masking is conducted to separate the forest from the others because the research focused on tree canopy in the forest. The TGHK 2014 map as the boundary of the study area and land use vector data SK 878 / Menhut-II / 2014 were used as masking data.

c. Single band usage

The single band used consists of green, red and infrared bands. The Landsat-8 image is in band 3 (green), 4 (red), and 5 (near infrared) that has been masked.

d. Transformation of Vegetation Index

The vegetation indices used in this research were NDVI, SAVI, and MSARVI which were described in the following equation.

$$NDVI = \frac{\rho IMD - \rho M}{\rho IMD + \rho M} \quad (2\text{-}1)$$

$$SAVI = \frac{\rho IMD - \rho M}{\rho IMD + \rho M + L} * (1+L) \quad (2\text{-}2)$$

$$MSARVI = \frac{2\rho IMD + 1 - \sqrt{[(2\rho IMD+1)^2 - \gamma(\rho IMD - \rho RB)]}}{2} \quad (2\text{-}3)$$

RB = NIR - ɣ* (Blue-Red)
NIR = Near Infrared Band,
M = Reflection value of Red Band,
B = Reflection value of Blue Band,
L = Enlightenment correction of ground background

e. FCD model

The processing FCD model was done using FCD Mapper software which can only accept the 8-bit data specification. The corrected pixel value in Landsat-8 band was converted back to 0-255 (8-bit). Furthermore, the FCD model stage consists of the following:

- Masking Noise (Clouds, Cloud Shadows and Water Body).
- Calculating the Vegetation Density (VD) and Scaled Shadow Index (SSI) that were integrated together to obtain the density of the canopy from the FCD.

f. Field Survey

Measurement of tree canopy density using hemispherical photography approach was taken by looking up from under the tree canopy (upward). Measurement through this approach also pays attention to some technical issues related to the use of tools such as camera positions. The photograph with the same height in all the sample quadrate is 2 meters and it is to avoid understory vegetation disturbance.

g. Calculation of Tree Canopy Density

The value of a tree canopy density in this study was a canopy density which was representing 30 meters that related to the size of spatial resolution of Landsat-8 image. The value of a tree canopy density can be obtained from hemispherical photographs using CAN-EYE software. Processing consists of two stages. First, setting the processing characteristics related to image size, calibration parameters, Center of Interest (COI) and angular resolution. Second, doing a binary photo classification to separate vegetation and sky.

h. Statistical analysis

The study used regression statistical analysis. This analysis was to know the relation between independent and dependent variable. The independent variable was called the predictor variable which was a manipulated variable to observe its effect on the dependent variable. Thus, in this study was using the independent variable in the form of image data and dependent variable of field data. Regression analysis aims to find out the correlation and relation from one value to other values. In this study, the analysis was used to construct tree canopy density data from the single band method and vegetation index by using canopy density data from field survey.

Tree canopy density information that was obtained from single band, vegetation index and FCD model were tested base on field measurement data (accuracy test samples) through a hemispherical photography approach that has been processed in CAN-EYE software. Test accuracy aims to determine the level of error of each model. Standard Error of Estimate (SE) was used to assess the accuracy. The smaller SE the smaller errors of estimating the density of the tree canopy will become smaller.

$$SE = \sqrt{\frac{\sum_i^n (y' - y)^2}{n-2}} \qquad 2\text{-}4$$

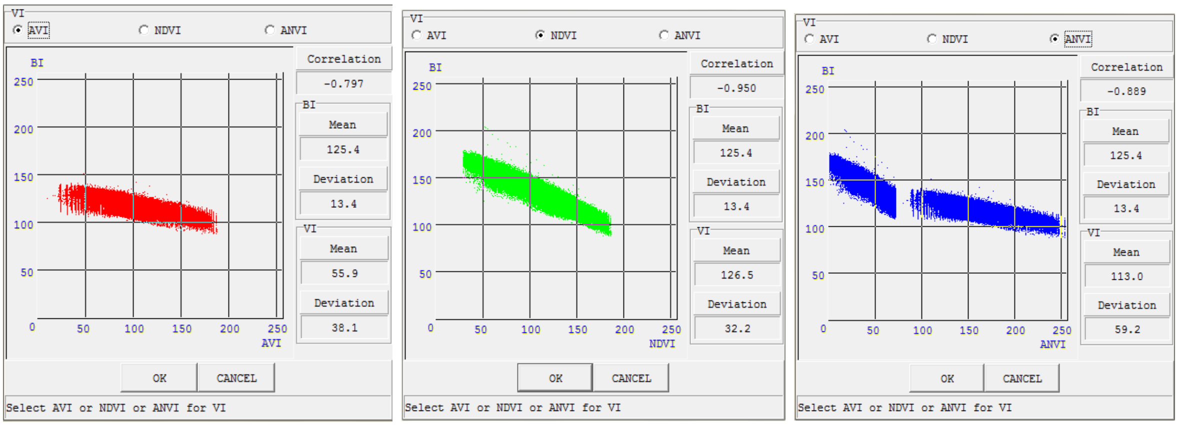
y '= modeled data (estimated) data,
y = field measurement data,
n = number of samples

3 RESULTS AND DISCUSSION

Landsat-8 image is used for analysis was an image that has been corrected to surface reflectance. The Landsat-8 image consists of several single bands to be used. Single band is band of recording one sensor with certain range. Each band has a different range of reflectance values because the response of each wave to an object is also different. Single band used is green, red and infrared (Near IM) band. Some single bands can be used to produce a vegetation index is aimed at obtaining vegetation information. Many vegetation transformation indices have been constructed with certain characteristics in obtaining the vegetation information such as tree vegetation density. In this study is used the several index vegetation of each index representative (generic, pressing the soil and pressing the atmospheric disturbances). The generic vegetation index is represented by NDVI, the vegetation index suppresses the effect of soil represented by SAVI and the vegetation index suppresses atmospheric disturbance i.e. MSARVI. Table 3-1 shows the pixel value on a single band and a vegetation index based on Landsat-8 imagery on the study area. In addition to the single band and vegetation index, this study uses the FCD model. The FCD model is built on the Vegetation Density (VD) and Scaled Shadow Index (SSI). Vegetation Density is calculated from the several vegetation indices (vegetation set) and several bare-soil indices (bare-soil set). The vegetation set is compiled based on the selected index of the three vegetation indices, namely AVI, NDVI, and ANVI. The selected vegetation index is an index that has a high correlation value and a low deviation. Figure 3-1 shows the correlation and deviation values of the three vegetation indices. It processes by using the Principal Component Analysis (PCA) technique between vegetation index and bare land index (BI). The value of PCA results near 0% indicates the area of bare land and otherwise, for 100% indicates the vegetated area. Based on the results of Landsat-8 image processing, the index is used to design the Vegetation set is the NDVI index because it has a correlation of -0.950 and a deviation of 13.8 (Figure 3-1b). Then the threshold of the Vegetation and Bare-soil (non vegetation) area is obtained to obtain Vegetation Density.

Table 3-1: Pixel values on single band and vegetation index.

Reflectance / Index	**Minimum**	**Maximum**	**Mean**	**Std. Dev**
Green	0.013112	0.042608	0.021087	0.002165
Red	0.007126	0.044644	0.012961	0.001425
NIR	0.189117	0.394593	0.280255	0.019087
NDVI	0.526524	0.740005	0.711457	0.009261
SAVI	0.408479	0.625464	0.542792	0.018294
MSARVI	0.032157	0.053153	0.043186	0.002002



(a) (b) (c)

Figure 3-1: PCA results between vegetation index and bare land index (BI). (a) AVI, (b) NDVI, and (c) ANVI

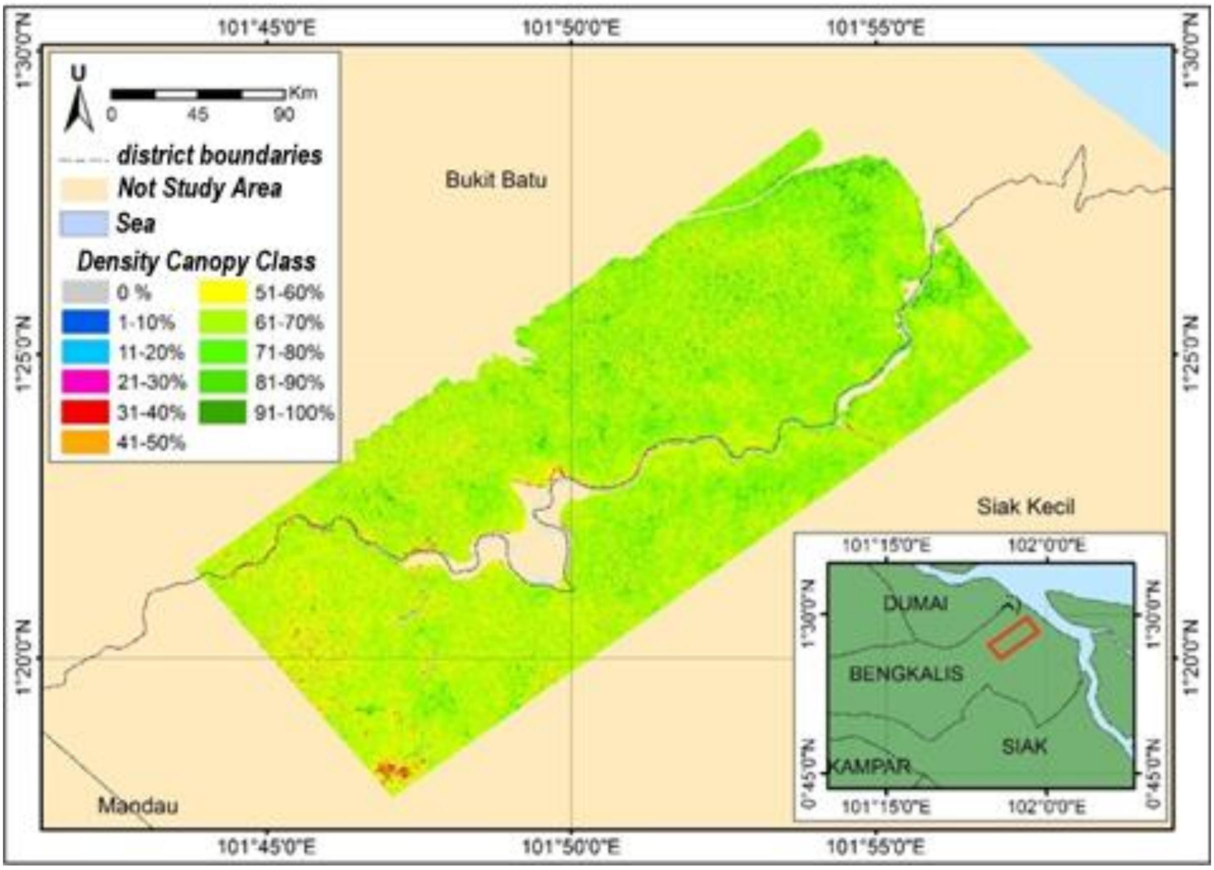


Figure 3-2: Map of canopy density based on FCD classification.

.

Furthermore, SSI design uses of vegetation index, thermal index and shadow index. The design is conducted to cluster forest cover and to distinguish forest and non-forest vegetation. Then, the VI-BI-SI results were thresholded to obtain SSI. FCD classification is conducted after designing VD and SSI. Classification result is the 11 classes of density intervals according to the FCD's classification system. Here the final result of the FCD model can be seen in Figure 3-2. Processing the single band model and vegetation index use the field canopy density data. The canopy density field data ranged from 13.84592 - 91.93871% with the total sample 42, mean 68.94125%, and standard deviation 20.84033%. The density of tree canopy visibility in the field can be seen in Table 3-2. Before processing to build the model using regression, the normal distribution test is required as parametric statistic and regression analysis. Normal distribution test is to see the data of tree canopy density is normal distributed or not. This research used Kolmogorov-Smirnov normal distribution test. The condition of normal data can be seen from the value of significance, based on the test results showed the significance value Kolmogorov-Smirnov for dependent of 0.172. This value indicates that the dependent data in this study is normally distributed because of the significance value> 0,05.

After the normality test, the next step is the correlation test. Correlation shows the strength of the relationship between two variables. The correlation can be shown based on the significance value, if significance value < 0.05 shows real and unreal correlation between two variables. Afterwards, the value of the Pearson correlation indicates a strong or weak relationship. The seven methods have the values of significance below 0.05 that indicates there is a correlation. Hereafter based on Pearson correlation, the seven methods have a strong positive relationship (> 0.5).

Single band and vegetation index that have correlation can be used for further analysis such as regression. The regression used is linear and non-linear (quadratic and exponential). Both types of regression is to obtain an empirical equation in building a good model. Terms of regression equation used can be shown based on the significance value of F-test and T-test (< 0.05). Statistical analysis between pixel value of the single band method and vegetation index (independent variable) with the field survey or density of tree canopy (dependent variable) obtained 8 empirical equations which originally have 18 equations. However, for exponentials do not meet the requirements of F test and T test. The empirical equations of each method can be shown in Table 3-3.

Model results from single band, vegetation index and FCD were tested to know how accurate the model for canopy density estimation based on the model of a regression equation. The model will be tested with the same field data type, but with different sample site from the model sample. The sample used is 12 samples. The accuracy test method used is SE by using confidence level 95%. The smaller of the SE value indicates better accuracy and the bigger of the SE value indicates the worse accuracy.

Table 3-2: The appearance of tree canopy density in the field.

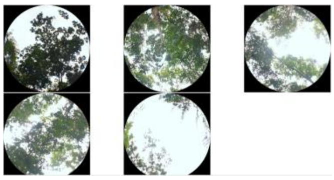
Photo	Value of tree canopy density
	20.46 %
	56.33 %
	86.38 %

Table 3-3: Results of statistical analysis of each method.

Method	Minimum Value	r	R^2	Equation	SE	Maximum Accuracy (%)
Green band	Linear	0.587	0.345	Y = 4660.733X – 42.62	22.06477	74
Red Band	Linear	0.598	0.358	Y = 10552.99X – 80.351	20.41212	76
IM Near Band	Linear	0.681	0.464	Y = 364.673X – 44.951	20.72349	75
NDVI	Linear	0.738	0.545	Y = 906.96X – 572.467	17.23816	79
SAVI	Linear	0.777	0.604	Y = 521.284X – 222.8999	15.10359	82
	Quadratic	0.817	0.667	$Y = -3837.982X^2 + 4759.878X - 1389.495$	15.06552	82
MSARVI	Linear	0.783	0.613	Y = 4529.523X – 136.284	14.80993	82
	Quadratic	0.828	0.686	$Y = -302024.421X^2 + 31523.284X - 735.648$	14.53656	83
FCD					15.68869	81

Based on the results of the accuracy test in Table 3-3, the lowest value of SE is MSARVI with the model of the quadratic regression equation. Overall, we can see the empirical equations of two types of regression (linear and non-linear) that have a low SE that is quadratic. Quadratic regression or 2nd order polynomial is one type of non-linear regression thus the study in this site is in accordance with Myeong *et al.* (2006) where the relationship of vegetation index to vegetation density values formed non-linear curve. Based on the quadratic model constructed from the field data, It shows the SE value ranges from 14.53656. The low value of the SE shows the model of a good quadratic equation used to estimate tree density at the location of the study when compared to the linear model. However, the value of SE is greatly influenced by the number of samples of accuracy and the more samples of the accuracy the result will be become the error value of a mapping model is suitable with field conditions.

The results of the density canopy estimation model using single band tree especially green band, red and near infrared have a SE equal to 22.06477 and 20.41212 and 20.72349. One cause of SE is high on single band because this method is simply using the only one band so it is less suitable for the estimation of tree canopy density. The single band is also influenced by the many external factors that influence the spectral characteristics of the density of the canopy so that the lack of vegetation can be highlighted. Meanwhile the transformation of vegetation index and FCD model using some bands are transformed to suppress the external factor that highlight vegetation aspect or reduce external factor.

The transformation of SAVI and MSARVI vegetation indexes has a low SE among other models of 15.06552 and 14.53656. On the contrary, the NDVI has SE of 17.23816. The magnitude of NDVI errors can be caused by poor attention to external factors such as soil and atmospheric disturbances (Jensen 2007). The SAVI and MSARVI index have a low SE because of their ability to reduce soil disturbance, especially MSARVI reduces soil and atmospheric disturbances. This condition was explained by Huete and Lie (1994, in Jensen 2007), both indexes (SAVI and MSARVI) are better than NDVI in any atmospheric conditions of remote sensing data. Subsequently, SAVI has a high error compared to MSARVI because there is no atmospheric disturbance factor. The results of the tree canopy density model from SAVI and MSARVI can also influence two indices of soil factors in general, while in the peat land have a wet characteristic of the soil. Furthermore, the FCD model has an SE of 15.68869. One of the error factors in the FCD because this model is not built empirically from field data yet the direct processing through FCD Mapper software by limiting the pixel value (threshold) of each process such as VD and SSI so that required the accuracy in determining the threshold value.

Statistical analysis showed a high correlation coefficient (r) and determination (R2) did not close the possibility of having small SE. Estimated canopy density in addition to view r and R2 also consider the small value of SE. MSARVI is statistically qualified with r = 0.828, R2 = 0.686, and the smallest SE of 14.53656. The use of SE as reference model selection is because SE uses field data from accuracy samples in which field data has the highest accuracy. Therefore, the smaller the SE shows the closer to field conditions moreover, it is good to use to estimate the density of the tree canopy.

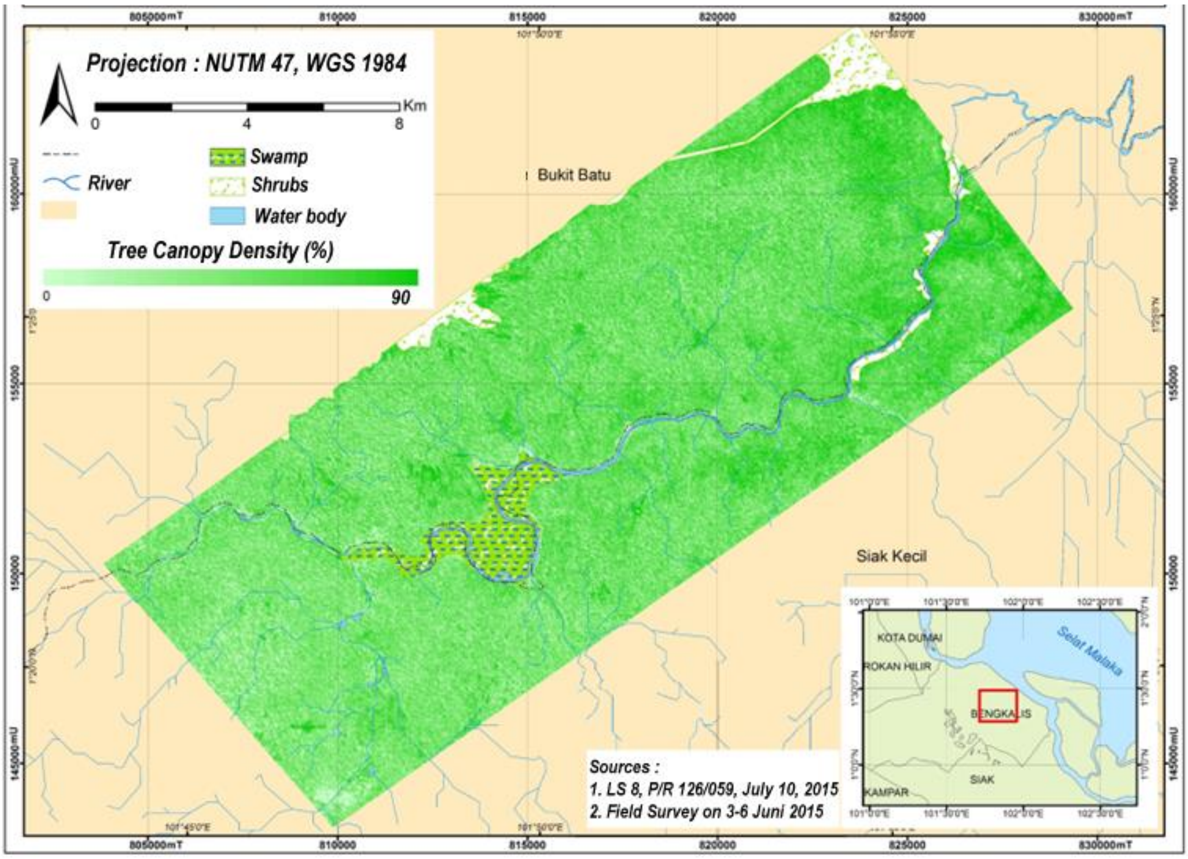


Figure 3-3: Map of Estimated Density of Tree Canopy with the best method (MSARVI).

The MSARVI results show the site of the study has an evenly distributed 60-70% density level (Figure 3-3). MSARVI can be used for mapping tree density on this site because it utilizes three bands at once they are blue, red and IM close so that this method can involve many external factors. This index according to Huete and Liu (1994, in Jensen 2007), is best used when the atmospheric or atmospheric corrected image data is thorough or partial. This study uses the whole atmospheric corrected data. Afterward the use of blue band can have an effect on the results because if it is monitored on aerosol and blue bands, the site of the study there is atmospheric noise in the form of thin clouds. The existence of the blue band in MSARVI according to Kristian (2014) can be a factor of high accuracy in coastal area. The location of the study near the Bengkalis and Malacca Straits allows for high evaporation and evapotranspiration. This condition can cause the high moisture content so that it can disrupt the reflected energy of the object especially the vegetation object. Then, the research sites are located in peatlands that have the wet environmental conditions so it has the evaporation in the study site. Thus, the use of the blue band in MSARVI becomes one of the lowest error factors.

The selection of MSARVI as the best method cannot be a benchmark for the future tree canopy density estimation as it is possible that the method cannot be used for measuring the other objects with different field conditions and methodologies. Later on, the magnitude of the error generated based on the MSARVI model cannot yet be known which spatial location is really accurate or inaccurate. Particularly in the central part of the study location that is not yet known because the sample is not reachable to that site. In addition, the influence of the accuracy of the estimation model of tree canopy density can also be due to environmental and external factors. Some of these factors, first of all, are mixed spectra primarily between trees and understory vegetation such as shrubs and bushes. It can cause of low tree canopy density in field conditions, in which the image can potentially produce high tree density canopy. Afterward, there are influences in some sites that have dry and wet peatlands that will give different ground spectral responses. In addition, the sample site factors are still less representative due to the low accessibility of accommodation and transportation.

4 CONCLUSION

Landsat-8 image is capable for application in tree canopy density estimation at SM Bukit Batu based on the seven methods shows that MSARVI as the best method with an accuracy of 82.65%. As for the other methods, green band 73.66%, 75.63% red band, near infrared band 75.26%, NDVI 79.42%, SAVI 82.01%, and FCD model 81.27%. The result of the comparison of accuracy test is based on 12 points of samples, where the number of samples greatly affects the results of accuracy.

The Landsat-8 image-based MSARVI model is the best model for tree density estimation at the study site. The estimation result of tree canopy density with this model shows the density of tree canopy of the study area is dominated by a density level of 60-70%. Density levels are distributed evenly throughout the region.

ACKNOWLEDGEMENT

The authors would like to thank to the BKSDA Riau Province for assisting writers in providing data and Mrs. Sachiko Kato - ITTO assisting in providing FCD Mapper V.2 software and Head of Remote Sensing Application Center, LAPAN in Satellite data provision and field information.

REFERENCES

APOEC.org, (2003), Forest Structure, downloaded April 14, 2016 from: http://www.apoec.org.nz/docs/

Ayat S., Tarigan J., (2010), Forest village of Lubuk Beringin: Scenario Conservation of Bungo Regency. Kiprah Agroforestri 6, 3:2, page. 3-5

Campbell JB, Wynne RH, (2011), Introduction to Remote Sensing. New York: Guilford Press.

Forest Watch Indonesia, (2014), Profile of Indonesia's Forests in 2009-2013 Period. FWI.

Jensen JR, (2007), Remote Sensing of the Environment: An Earth Resource Perspective. 2nd Edition, Pearson Prentice Hall, Upper Saddle River.

Lillesand TM, Kiefer RW, Chipman J., (2014), Remote Sensing and Image Interpretation. 7th Edition. USA: John Wiley & Sons, Inc.

Lund HG, (2002), Definitions of old growth, pristine, climax, ancient forests, and similar terms. [Online publication], Manassas, VA: Forest Information Services. Misc. Pagination, downloaded January 20, 2016 from http://home.att.net/~gklund/pristine.html

MAB Indonesia, (2016a), Cagar Biosfer Giam Siak, diunduh 14 April 2016 dari http://www.mab-indonesia.org/cagar.php?i=giam

MAB Indonesia, (2016b), Peta Area Cagar Biosfer, downloaded April 14 , 2016 from http://www.mab-indonesia.org/cagar.php

Matthews JA, (2013), Encyclopedia of Environmental Change. United Kingdom: Sage Publications.

Myeong S., Nowak DJ, Duggin MJ, (2006), A temporal analysis of urban forest carbon storage using remote sensing Remote Sens. Environ. 101 277–82

Panta M., (2003), Analysis of Forest Canopy Density and Factors Affecting It Using RS and GIS Techniques. Unpublished M.sc. Thesis, ITC, Netherland.

Republic of Indonesia, (1999), Law of the Republic of Indonesia No.41 of 1999 on Forestry.

Tempo.co, (2015a), Police Shoot Squadron Giam Siak Kecil Biosphere Reserve Bukit Batu, downloaded April 13, 2016 from: https://m.tempo.co/read/news/2015/06/14/058674941/polisi-tembak-perambah-cagar-biosfer-giam-siak-kecil

Tempo.co, (2015b), Forest Fires in Riau Threatens Biosphere Reserve of Giam Siak Kecil, downloaded April 13 , 2016 from: https://m.tempo.co/read/news/2015/10/22/206711908/kebakaran-hutan-di-riau-ancam-cagar-biosfer-giam-siak

Tohir NR, Prasetyo LB, Kartono AP, (2014), Mapping of Canopy Density Changes in the People's Forest of Kuningan Regency of West Java. National Seminar on Remote Sensing: Detection of Geobiophysical Parameters and Remote Sensing Dissemination.

Wibowo A., Gintings A., (2010), Degradation and Forest Conservation Efforts. Reversing the Degradation Trends of Land and Water Resources. P. 67-87.