DETERMINATION OF STRATIFICATION BOUNDARY FOR FOREST AND NON FOREST MULTITEMPORAL CLASSIFICATION TO SUPPORT REDD+ IN SUMATERA ISLAND

Tatik Kartika[1*], Inggit Lolita Sari[2], and Bambang Trisakti[1]
[1]Remote Sensing Application Centre, LAPAN, Jakarta
[2]Remote Sensing Technology and Data Center, LAPAN, Jakarta
[*]E-mail: tatikkartika@yahoo.com

Abstract. Multi-temporal classification is a method to determine forest and non-forest by considering a missing data, such as cloud cover using correlations value from the other data. This circumstances is frequently occured in a tropical area such as in Indonesia. To gain an optimum result of forest and non-forest classification, it is needed a stratification zone that describes the difference of vegetation condition due to different of vegetation type, soil type, climate, and land use/cover associations. This stratification zone will be useful to indicate the different biomass volume relating to carbon content for supporting the REDD+ project. The objective of this study was to determine stratification boundary by performing multi temporal classification in Sumatera Island using Landsat imagery in 25 meter resolution and Quick Bird imagery in 0.6 meter. Rough stratification was made by considering land use/cover, DEM and landform, using visual interpretation of moderate spatial resolution of satellite data. High spatial resolution data was also provided in some areas to increase the accuracy level of stratification zone. The stratification boundary was evaluated using forest classification indices, and it was redetermined to obtain the final stratification zone. The indices was generated by Canonical Variate Analysis (CVA) method, which was depend on training samples of forest and non-forest in each previous stratification zone. The amount of indices used in each zone were two or three indices depending on the separability of the forest and non-forest classification. The suitable indices used in each zone described forest as 100, non-forest as 0, and uncertain forest between 50-99. The result showed 20 stratification zones in Sumatera spreading out in coastal, mountain, flat area, and group of small islands. The stratification zone will improve the accuracy of forest and non-forest classification result and their change based on multi temporal classification.

Keywords: *Muti temporal classificatioon, stratification zone, Forest, CVA, Landsat, Quick Bird*

1 INTRODUCTION

People believe that forest is a part of supporting system of human life, especially in terms of climate change due to green house gases (GHG) emissions which has an important role as a carbon sink. Forest is a minimum area of land of 0.05-1.0 hectares with tree crown cover (or equivalent stocking level) of more than 10-30 per cent with trees with the potential to reach a minimum height of 2-5 metres at maturity in situ (Neef, 2006). The Ministry of Forestry such as in Indonesia and Australlia, and Food and Agriculture Organisation (FAO) defined forest as a set of trees with certain height, area, and crown cover (BSN, 2010; http://www.daff.gov.au/brs/forest-veg/nfi/forest-info/).

Stern (2006), stated that the changes of land use (land use, land use change, and forestry/LULUCF) in developing countries have contributed emission from carbon dioxide (CO_2) approximately 20% to global emissions. This emissions will increase in line with the rising of forest conversion into non-forest landuse. However, one of the solutions is to minimize the emissions level toward reducing emission from deforestation and forest degradation (REDD). Therefore, Indonesian National Carbon Accounting System (INCAS) has been established in order to build the capacity, calculate the biomass and carbon stocks, as well as the level of emissions and carbon sequestration occurred in forest area (Wardoyo, 2009). INCAS is a project which is mainly adopted the system of National Carbon Accounting System (NCAS) of Australia that had been a referrences for carbon accounting system acording to Kyoto Protocol and national GHG inventaritation at United Nation Convention about climate change (http://www.climatechange.gov.au/ncas/index.html). In INCAS, there are many of plantations such as oil palm, coconut, ruber plantations which are not considered as forest because of the relation of their product. The integration of several components in INCAS program include land cover change in forest and non-forest area, that analyzed using remote sensing data (Figure 1).

Figure 1. INCAS components

The information of land cover change is derived from monitoring the changes of forest by interpreting the remote sensing data in a particular area and classified into forest and non-forest. Land cover classification can be done visually by interpreting the image and delineate the boundaries between the classes, or can be done by digitally using supervised or unsupervised method. Visual classification is highly dependable on the ability of interpreter to interpret and delineate the object, while digital classification relies upon the accuracy of training samples. However, due to time considerations the digital classification has more benefit, because the process is faster especially to classify a large amounts of data within a wide areas.

Multi temporal classification for forest and non-forest is a digital supervised method to classify forest and non-forest by considering the missing data caused by cloud cover that frequently occur in tropical area such as Indonesia. The accuracy of the classification result depends on how much the series of the data is provided. The principle of multi-temporal classification for forests is shown in Figure 2.

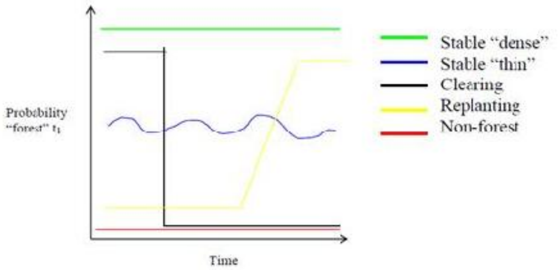


Figure 2. Typical temporal signatures for forest and non-forest cover

Forest and non-forest classification in INCAS program, requiring a forest base probability that will become references for determining the probability of forest in subsequent years. Once, the forest base probability is made, the border of stratification zone can be determined in order to achieve the accuracy of classification between forest-non forest.

This paper discuss how to determind the stratification zone boundaries, as a first step of multitemporal classification process to produce a border of typical area for the forest base probability. In addition, the forest base probability then become a base to calculate the forest probability for other years. The determination of stratification zones is made by using data reference from Landsat TM / ETM + imagery as a medium resolution and Quickbird imagery as high resolution data which has 25 m and 60 cm in spatial resolution, respectively. Therefore, the stratification zone is important as a part of INCAS method to produce an accurate classification of forest in order to account the biomass in terms of supporting the REDD program.

2 MATERIALS AND METHOD

2.1 Time and location

The project was located on Sumatera Island on 95°E - 105°E and 6°N - 6°S, while the primary data of mosaic Landsat 5 TM and Landsat 7 ETM + were acquired in 2008. The data were chosen because of relatively free from clouds and the completeness of Quickbird imageries were used to validate the result. The mosaic of Landsat data in 2006 and 2000 were also provided to fill the gaps caused by the missing data. The mosaic of data and information of acquisition in 2008 is shown in Figure 3.

2.2 Data

The main data used in this study were the mosaic of Landsat 5 TM and Landsat 7 ETM +, in 25 meters resolution acquired in 2008 from LAPAN, USGS, Australia, and Thailand. The data available were already corrected in ortho geometric, radiometric, cloud masking, and mosaicing.

Some QuickBird Images acquired in 2006-2009 were used to validate the training samples. Figure 4 shows area covered by QuickBird imagery (red dot) in Sumatera Island. LAPAN was using high-resolution satellite imagery, provided through the INCAS Program, to support this analysis and subsequently produce annual land cover change maps of all of Indonesia which will form a key input to the pilot system (http://www.climatechange.gov.au/ncas/index.html).

Spatial informations of several forest land cover in Sumatera were obtained from the Ministry of Forestry. List of informations of survei fields was retrieved from the National Agency of Geospatial and Information of Indonesia (BIG).

Knowledge and experiences from local experts to interpret the existing land use were people from Forestry Services of Jambi Province Authority.

2.3 Data Processing and Analyses

There were five stages for creating the stratification zone boundaries in Sumatra Island (Furby, 2011) i.e.,

1) Make a combination of channels 453 from Landsat imagery
2) Perform a rough delineation of the stratification zone boundaries by visually seeing the difference of land cover, forest type, and Landform. This process can guide in selecting the training sample, so that the entire region in Sumatra Island is already divisible.
3) The training sample is chosen by visually interpreting the Landsat imagery through the combination of Chanels 453 and the information from the high resolution data. Moreover, there are several requirement to

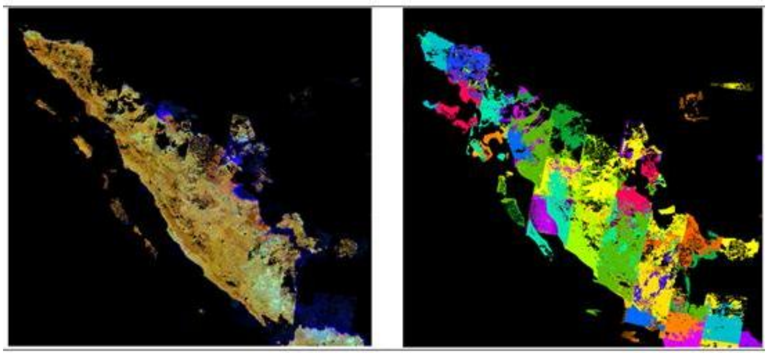


Figure 3. Mosaic of Sumatera Island using Landsat imagery in 2008 (left) and mosaic of raster date image that shows information of path-row and date (right)

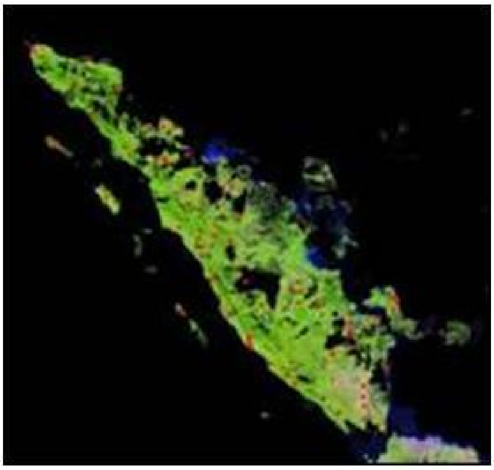


Figure 4. QuickBird data distribution for Sumatra Island (red dot)

make a training sample, such as; a). Homogeneous areas with a known ground cover types, which was assisted by the use of high spatial resolution data from QuickBird Images. This training sample was devided into two classes, namely forest and non-forest, b). the number of pixels for each training sample were vary between 10-100 pixels, c). Representing a variety of colors and tones, such as setlements, palm oil plantations, primary forest, secondary forest, mangrove, and peat forest, d). the information of landcover had also been verified by local and expert people, and e). perform naming standards of each training samples.

4) the determination of the index and the threshold in each stratification zones were obtained by using CVA method. The aim of this method was for analyzing the structure of multivariate groups of the data that mathematically equivalent to a one-way multivariate analysis and reknown as the canonical discriminated analysis. Furthermore, CVA analysis can be use to separate between forest and non-forest using its spectral values, determined the stratification zone, identify the best index over multiple image dates, and specify the required number of indices.

The basic methodology of CVA was described detail in Campbell (1981). He said that the CVA mainly consisted of two stages orthogonal rotation. The first stage was using a principal component analysis (PCA) of the original variable. The second stage was for applying PCA for the group means of the ortho normal variables from the first stage eigen analysis. Therefore the index was derived by a two group, namely the within group and the between group. The geomatrically to determine the within group and the between group are shown in the equation below:
The within group;

$$c^T\{\sum_{k=1}^{g}\sum_{m=1}^{n_k}(x_{km}-\bar{x}_k)(x_{km}-\bar{x}_k)^T\}c \quad (1)$$

In those equation, the term in {...} reflects the squared deviations and cross deviations of each trianing site observation from the mean of its corresponding group. On the other hand, the formula to determine the between group mentioned below:
The between group;

$$c^T\{\sum_{k=1}^{g} n_k(\bar{x}_k-\bar{x}_T)(\bar{x}_k-\bar{x}_T)^T\}c \quad (2)$$

The term in {...} mentioned above is describes the squared deviations and cross deviations of each mean from the mean of the means. It is noted that C is the canonical vector which represents the direction of the calculated canonical vector and X_{km} represents all the observation.

The number of indeces depends on the relatives sizes of the canonical root, which is the corresponding ratio of the canonical vector. Therefore, it is hope that the number of indices can be an appropriate configuration between two or three dimentional subspace defined by the first, two or three canonical vector. It means that if there is a significant contribution from the second and subsequent canonical root (more than 10%), it will be necessary to consider a second index to differentiate between forest and non-forest cover for each images. On the other hand, a threshold is needed to determine the probability between certain forest or the within group and non-forest. This is done in order to avoid the commission error in the between group.

This stage produce an index for each zone as well as the threshold. Afterwards, the image of the forest was presented by the percentage of probability for non-forest (0) into the woods (100).

5) To revise the rough stratification boundary by checking the results of forest base probability that was applied in every area in Sumatera Island, so that there would be a separation or merging into a new zone in order to gives a good between forest and non-forest.
6) Mosaicing of all the forest base probability using fixed zones stratification, this can be used as a reference to make a forest probability for the other years.

3 RESULTS AND DISCUSSION

There were three UTM zones used in Sumatra Island, which are 47 S, 47 N and 48 N. Each zones then divided into some quadrants, such as Zone 48 was divided into 6 quadrants and they were namely as CE (Central East), CW (Central West), NE (North East), NW (North West), SE (South East), and SW (South West), whereas others were divided into four quadrants. The detail of each UTM zone is shown in Figure 5.

The process was started by dividing Sumatera Island into 5 zones as showed at Figure 6. First zone was containing Jambi, Bengkulu, West Sumatera, and a part of West Sumatera provinces which will be divided into 3 zones: coastal, lowlands, and highland zones. Two other stratification zones were in the southern part of of Sumatera Island, located in Palembang and Lampung region. Lastly, in the northern part were starting from Riau, North Sumatera, and Aceh Provinces.

The next stage was choosing the training site as an input of CVA calculation for determining the index and threshold. Figure 7 shows an examples of training site in zone 3 which was a mountain area. In each zone, the numbers of sample were taken about 60 of training site that contain of 30 samples of forest and 30 samples of non- forest. The distribution of training samples should be representative of existing conditions of land cover. After that, the sample of forest and non-forest were labeled and put additional

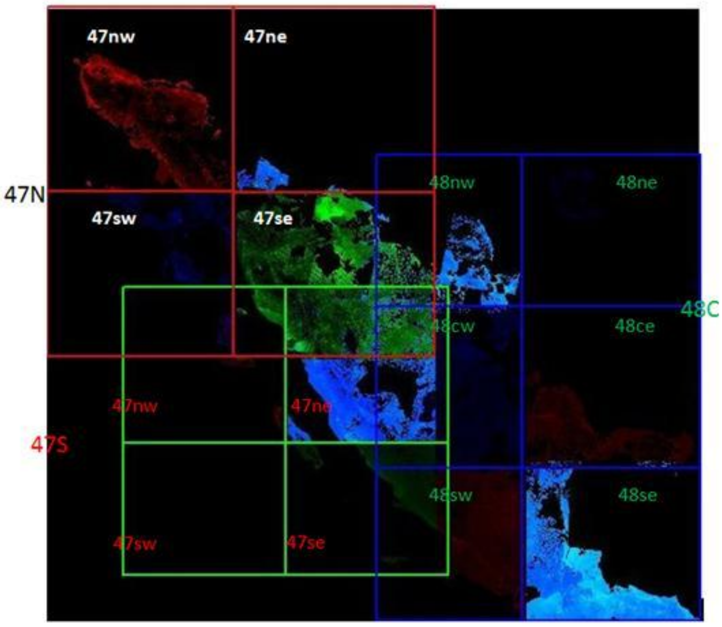


Figure 5. The detail of UTM sub-zone in Sumatera Island.

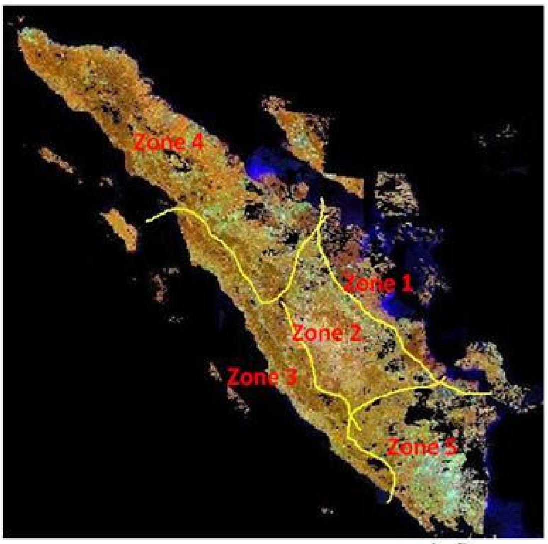
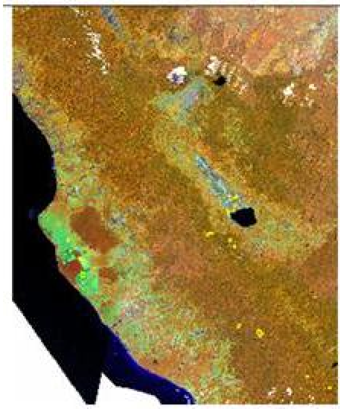


Figure 6. The Rough stratification zones in Sumatera Island.



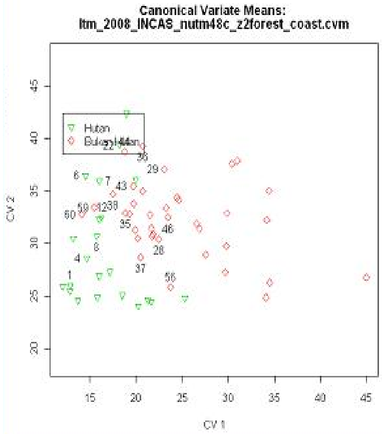
18	VECTOR	NH_455_sawit
19	VECTOR	NH_451_sawit
20	VECTOR	NH_450_sawit
21	VECTOR	HT_445_dense
22	VECTOR	HT_444_dense
23	VECTOR	HT_443_dense
24	VECTOR	HT_441_dense
25	VECTOR	HT_440_dense
26	VECTOR	HT_431_dense
27	VECTOR	HT_430_dense
28	VECTOR	HT_429_dense
29	VECTOR	NH_423_crop

Figure 7. The distribution of training sample vector and high resolution data also their list of training samples labeled.

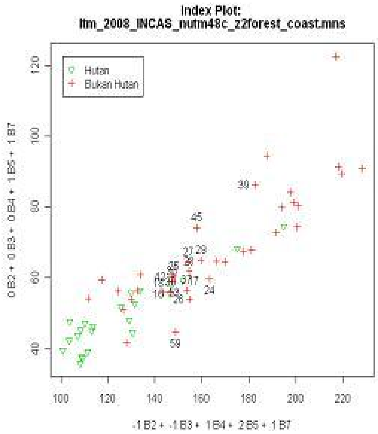
information based on local knowledge such as the types of forest or plantation to make it easier to evaluate and revise the training sample.

Next, the result of training samples were calculated using CVA method to obtained the information of distribution and separation between forest and non-forest as well as the index and the threshold. Figure 8 shows the canonical variate means for the indices 1 and 2, where the forest and non-forest were well separated unless a few of mixed. The mixed training simple was needed to be check and revise, whether it had labeled correctly or no, for instance some plantation were classified as forest. On the other hand, adding some training site was better to increase the accuracy, as it shown in Figure 9.

According to Figure 10, the indices and the threshold in the eastern coastal areas of Sumatra were quite good grouping between forest, non-forest, and uncertain, but there was still found that some swamp area were classified as forest. Therefore, the addition of index 3 was needed to differentiate the swamp areas. The using of three commbination of indices resulted in good

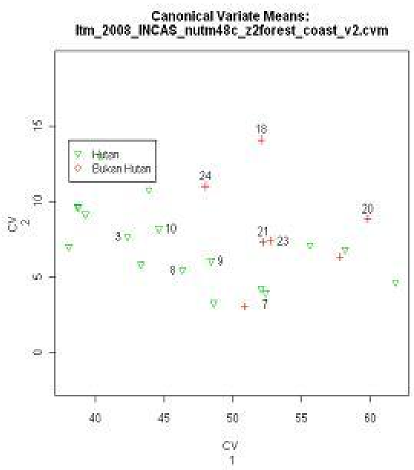


A picture of canonical variate means which shows a separation between a sample of forest and non-forest

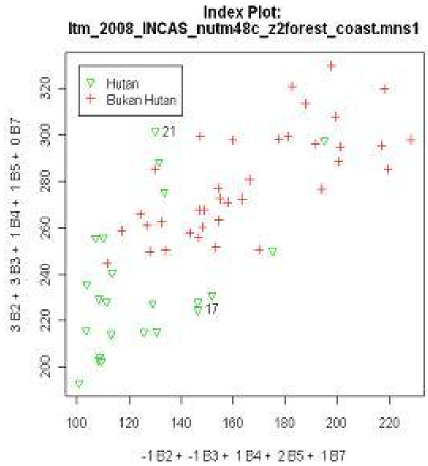


A picture of index 1 and 2

Figure 8. A picture of CVA results for determining the indices (before correction).



A picture of canonical variate means which shows a separation between a sample of forest and non-forest



A picture fixed of index 1 and 2

Figure 9. A picture of CVA results for determining the indices (after correction).

separation between the forests along the coastal which mostly consist of mangrove and non-forest such as rice fields and coconut plantations. The result of classification on mangrove then identified as high probability or 100% of probability.

After the appropriate indices and the threshold was already derived, then tested to the other coastal areas especially in the coastal areas along zone NUTM 48 in central east. This process was to check whether the index and the threshold were suitable and could be applied in another coastal areas as well as in the lowlands and highlands. In this case, it was found that in Lampung province which mostly contained of cinnamon forest, needs a different index and threshold, therefore a new separate zone was needed for this area.

It was obviously showed that the accuracy of training sample, CVA calculation for creating the indices and the threshold, also several trials testing of the indices and the threshold to be applied in other area was important to perform a best stratification zone. Moreover, this could be done by editing the rough stratification zones using merging or dividing into some zone into detail zones. As a final result, there were 20 stratification zones can be made in Sumatra Island in order to achieved he best separation of forest and non-forest, as shown in Figure 11. Figure 12 showed the forest probability of Sumatera Island as a forest base where the indices at stratification zone in related can be used for the other years with threshold adjusment. The indices and the threshold used for all of zones stratificatins in Sumatera Island is presented in Table 1.

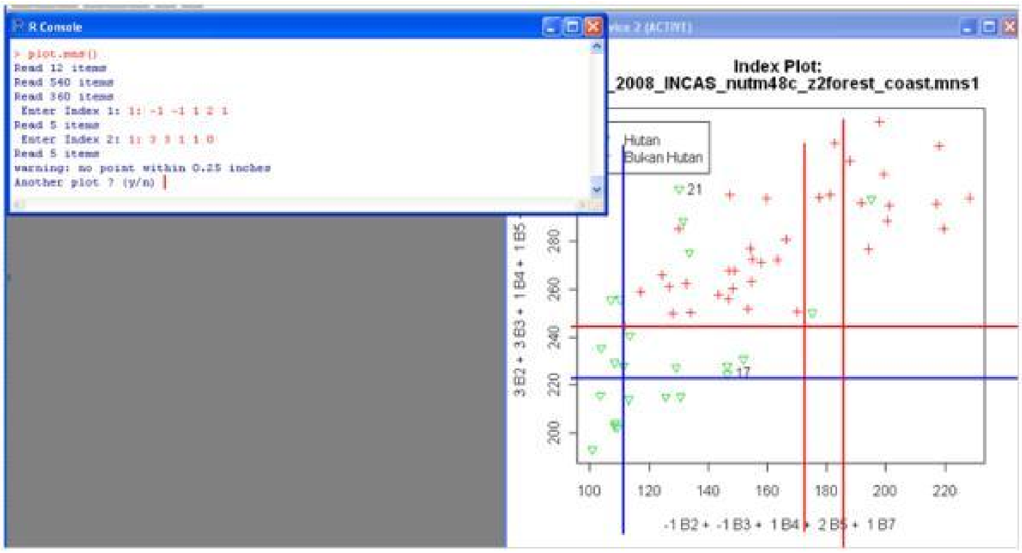


Figure 10. The indices and the threshold Indeks that separated the object.

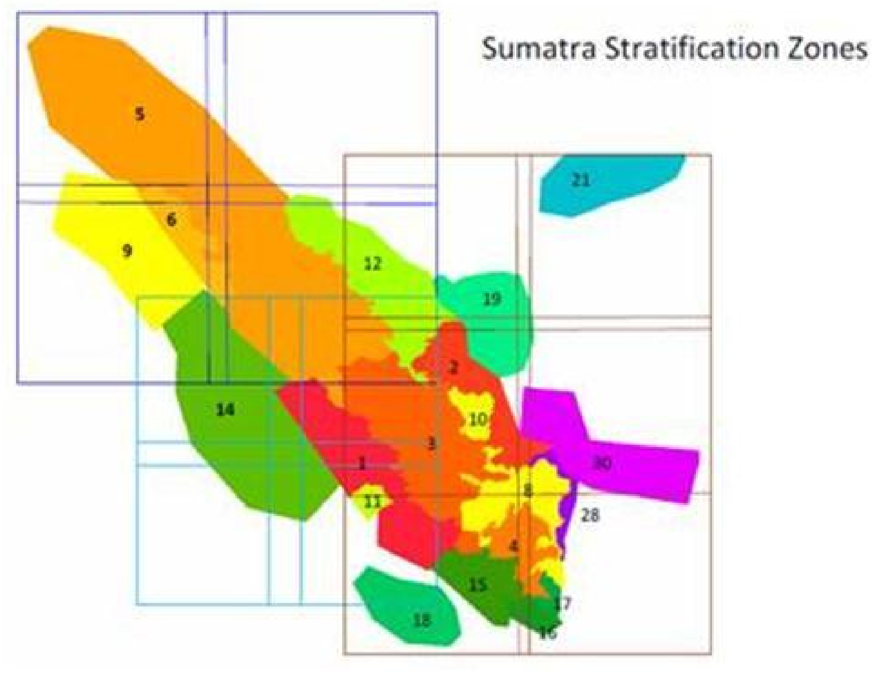


Figure 11. Final stratification zones in Sumatera Island.

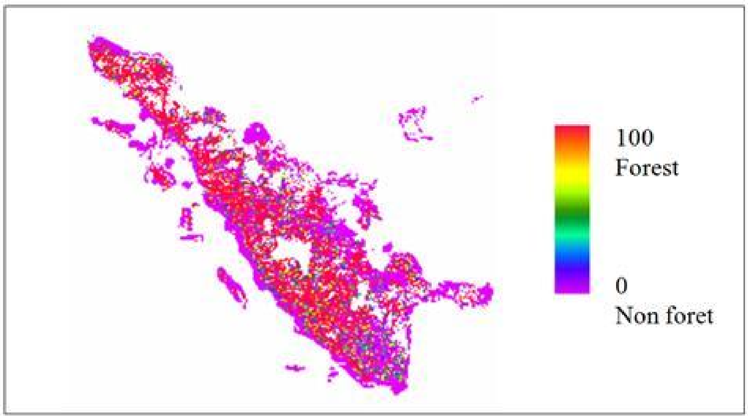


Figure 12. The forest base probability in Sumatera Island.

Table 1. The indices and threshold for all of zones in Sumatera Island.

Zone	Index 1	Index 2	Index 3
1	3B3+B4-B5+2B7	B3-B4+B5+2B7	-B3+B4
2	-B2-B3+B4+2B5+B7	3B2+3B3+B4+B5	-B3+B4
3	B2+2B3-B4	B2+2B5+B7	
4	6B2+3B3+B4-2B5	3B3-B4	-B3+B4
5	4B2+4B3-B4+2B5	B2+B4-B7	
6	B2+B3-B4+2B5+2B7	B2+2B3+B4-B5-2B7	
7	B2-B4+B5+B7	B2+B5	-B3+B4
8	3B2+B3+B4	B2+B4-4B7	-B3+B4
9	4B2+B5	-4B2-4B3+B4	
10	-B3+B4+2B5	-2B2+2B4-2B5+B7	
11	2B2+B4	2B+2B3-B4+B5+2B7	-B3+B4
12	-B3+B4+2B5	B2+B3-B4+B5	
14	4B2+4B3-B4+2B5	B2+B4-B7	
15	3B3+B4-B5+2B7	B3-B4+B5+2B7	-B3+B4
16	4B2+4B3-B4-B5	2B5+B7	
17	B5+B7	2B2+2B3+B4-B5	-B3+B4
18	4B2+4B3-B4+2B5	B2+B4-B7	
19	-B2-B3+B4+2B5+B7	3B2+3B3+B4+B5	-B3+B4
21	-B2-B3+B4+2B5+B7	3B2+3B3+B4+B5	-B3+B4
22	-B2-B3+B4+2B5+B7	3B2+3B3+B4+B5	-B3+B4
28	-B2-B3+B4+2B5+B7	3B2+3B3+B4+B5	
30	6B2-3B3-B4+2B5+2B7	-B3+B4	

The accuracy assessment was done by taking three point location on forest probability image using a random sample method for each stratification zones, which was verified using Landsat imagery and Quickbird and Ikonos as the two high resolution data provided by INCAS project. Sixty (60) points of location wes derived to be assessed from 20 stratification zones. The result showed that the accuracy of classification was 92%.

4 CONCLUSION

The stratification zones was created in order to increase the accuracy of forest base probability at multi temporal classification. The 20 stratification zones spreading out in coastal, mountain, flat area, and group of small islands were created in Sumatra.

The indices of each stratification zones can be used for another years of images which is located in the same stratification zone by adjusting the threshold.

From 60 points of location to assessed for the quality of classification, it showed that the accuracy was 92%.

ACKNOWLEDGMENT

We would like to thank the people of the following companies for the help, discussion, and the partnership: Nur Hidayat, Orbita, Agus Hidayat, Arum Tjahjaningsih, Kustiyo (LAPAN-Indonesia), Suzanne Furby, Jeremy Wallace, Tony (CSIRO-Australia), Virni Arifanti (The Indonesia Ministry of Forestry), Darmawan (BIG-Indonesia), and Hasvia, Erizal (Forestry Services of Jambi Province, Indonesia).

REFERENCES

BSN, 2010. Land Cover Classification.

Furby, S. And W. Jeremy, 2011, Guidelines for annual forest extent and change mapping, Version 2.2, CSIRO Mathematics, Informatics and Statistics.

Neef, L.S., 2006, Forest and climate change working paper 4, choosing forest definition for the clean development mechanism, FAO.

Wardoyo, 2009, Development of national carbon accounting system in forestry sector. Proceeding of The National Seminar on Indonesia Agriculture Informatics Association (in Indonesian). ISBN 978-979-95366-0-7.

Stern, 2006, The economic effects to climate change. *Population and Dvelopment Review Vol 32, Issue 4*. Available online: http://onlinelibrary. wiley.com/ doi/10.1111/j.1728-4457.2006.00153.x /pdf. [accessed on 9 January 2012].

Campbell, N.A. and W.R. Atchley, 1981, The geometry of canonical variate analysis. *Systematic Zoology*, 30(3): 268–280.