View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Universiti Teknikal Malaysia Melaka: UTeM Open Journal System

Classification of Landsat 8 Satellite Data Using NDVI Thresholds

Afirah Taufik[1], Sharifah Sakinah Syed Ahmad[2], Asmala Ahmad[2]
[1]*Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal Melaka, Malaysia.*
[2]*Department of Industrial Computing, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.*
afirahtaufik@gmail.com

***Abstract*—This study aims to classify Landsat 8 satellite data using NDVI thresholds. Initially, visible and near infrared bands of Landsat 8 satellite were used to derive Normalized Different Vegetation Index (NDVI) image. Vegetation, non-vegetation and water areas were then analyzed where thresholds for separating them are carefully determined with the aid of ground truth information of the study area. Density slicing was performed in order to separate the image into different land covers. Eventually, color mapping and class labeling were done to complete the classification process. The accuracy of the classified image is then assessed using a confusion matrix where overall classification accuracy and Kappa coefficient are computed. The result shows that NDVI-based classification is able to classify the Landsat 8 satellite data with a high accuracy.**

***Index Terms*—Accuracy Assessment, Multispectral, NDVI Supervised, Threshold, Vegetation.**

I. Introduction

Vegetation index can be used as an indicator to quantify the greenness of plants within satellite data. There are several vegetation indices, but the most frequently used index is the Normalized Difference Vegetation Index (NDVI) [1]. By analyzing images recorded from visible red and near-infrared (NIR) wavelengths, researchers can determine the coverage of vegetation on the surface of the Earth. NDVI can be expressed as (1):

$$\text{NDVI} = (\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED}) \tag{1}$$

In this study, we used multispectral satellite data that acquired from Landsat 8 for a vegetation classification. A multispectral data that are used for NDVI classification is compared to the ground truth data.

Supervised classification is a process where the user selects representative samples for each land cover class in the digital image. These sample land cover classes are called training data. The classification of land cover is based on the spectral signature defined in the training set. The digital image classification software determines each class on what it is similar to most in the training set. In the supervised classification, the NDVI images are categorized into certain classes according to the NDVI value. Some techniques were applied to execute the process of vegetation classification using NDVI.

II. Related Works

It is important to know if the spectral features of Landsat 8 are of the same standard as previous Landsat data because Landsat 8 data have narrower bands, especially because of the normalized difference vegetation index (NDVI) calculation which is the most popular vegetation index [2]. This study identifies land use changes in the metropolitan region of Klang-Langat Valley focusing on urban sprawl and green space. A technique called Normalized Difference Vegetation Index (NDVI) is used to quantify temporal urban green space dynamics [3]. [4]presented a method for a supervised classification of NDVI time series to identify vegetation type and vegetation coverage, absolute in percentage coverage or relative to a difference NDVI cycle.

[5] presented an improved method for the analysis of satellite image based on NDVI. The method employs the multispectral remote sensing data technique to find spectral signature of different objects such as vegetation index, land cover classification, concrete structure, road, urban areas, rocky areas and remaining areas. The study [6]presented a simple method for rapidly and accurately mapping rubber plantations in the Xishuangbanna region of southwest China using phenology-based vegetation index differencing. Temporal profiles of the NDVI, Enhanced Vegetation Index (EVI), and others indices were constructed using 11 Landsat 8 OLI bands acquired within one year.

The study [7] present recent high-resolution multispectral satellite data were used to produce land use or land cover classification and NDVI mapping for the delta. Many algorithms have been developed for the remote estimation of biophysical characteristics of vegetation, in terms of combinations of spectral bands, derivatives of reflectance spectra, neural networks, inversion of radiative transfer models, and several multi-spectral statistical approaches. However, the most widespread type of algorithm used is the mathematical combination of visible and near-infrared reflectance bands, in the form of spectral vegetation indices. The general objective of this study [8] is to evaluate different vegetation indices for the remote estimation of the green leaf area index (Green LAI) of two crop types (maize and soybean) with contrast canopy architectures and leaf structures.

III. Materials And Methods

A. Data

The study was conducted in Klang, Selangor. This area is located between longitudes 101° 16' to 102º 00' E and latitude 3º 05' to 3º 22' N. The satellite data were acquired from USGS Landsat 8 data that were used in this study. The data were selected based on the high quality of data acquired by the Landsat satellite. The data were acquired on24th March 2014. By using spectral subset in ENVI software, the selected regions of interest (ROI) were chosen as training pixels. The DAT file contains 7 bands with the 349 pixels × 329 pixels.

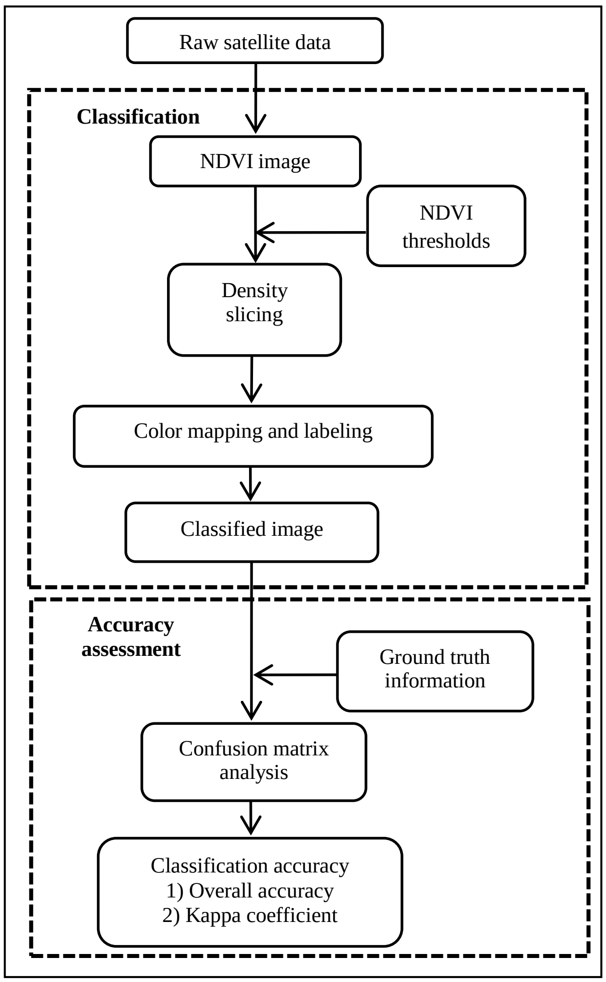

Figure 1: The flow chart of experiment classification.

B. Classification

Before the classification, the process was begun with image enhancement by using a decor relation stretch to enhance the image for more effective visualization. Figure 1 shows the flow chart of classification and accuracy assessment process for this study. Afterwards, the experiment was continued by computing the NDVI. In this study, different combination bands of 5-4-3 (NIR, Red, and Green) are constructed to red, green and blue (RGB color). The RGB image is a standard color for/of infrared (CIR) image. The NDVI was determined using (2):

$$NDVI = \frac{Band\ 5 - Band\ 4}{Band\ 5 + Band\ 4} \quad (2)$$

The experiment continued to locate the vegetation by setting the threshold of NDVI image. Based on the [10], the values of NDVI threshold are shown in Table 1.

Table 1
NDVI values

Value of NDVI	Descriptions
0.1 or less	Very low NDVI
0.2 to 0.5	Moderate NDVI
0.6 to 0.9	High NDVI

The training data are divided into 3 major classes: Water, Non-vegetation and Vegetation area. The results are recorded. Density slicing is a process where the NDVI are set as a modified gray level image. Afterwards, the color mapping and labeling of the satellite image for three classes were performed. The process of classification based on NDVI threshold is explained and compared to the classified image with the ground truth.

C. Accuracy assessment (Confusion matrix)

The information of ground truth data was compared to the classified image in order to check the accuracy. The accuracy of user and producer were carried out to measure the classification accuracy. Producer's accuracy is where the individual class accuracy can be acquired by dividing the sum of correctly classified pixels. The error of misclassified pixels and also misclassified into another class were recorded. Meanwhile, the user's accuracy is a measurement where the individual class acquired from the classified pixels in same group [9]. The overall accuracy (3) was calculated based on the confusion matrix that obtained for user's accuracy and producer's accuracy. The description of overall accuracy is shown below:

$$Overall\ accuracy = \frac{Total\ number\ of\ correct\ classified}{Total\ number\ of\ pixels} \times 100 \quad (3)$$

Kappa coefficient (4) is another measurement to measure the training pixels with the ground truth data. The Kappa values are +1.0 to -1.0 to measure where the positive value shows high accuracy. A value of zero shows no correlation in classification.

$$K = \frac{n\sum_{i=1}^{p} x_{ii} - \sum_{i=1}^{p} x_{io}x_{oi}}{n^2 - \sum_{i=1}^{p} x_{io}x_{oi}} \quad (4)$$

where:

n = total number of pixels

p = number of classes

$\sum x_{ii}$ = total number elements of confusion matrix

$\sum x_{i0}$ = sum of row i
$\sum x_{oi}$ = sum of column i

The results of classification based on the accuracy assessment were obtained and recorded in the next section.

IV. Results And Discussion

A. *Accuracy assessment results*

The results of any classification process to satellite data must be quantitatively determine their accuracy of classification. The pixel that has been categorized from the NDVI image was compared with the ground truth information. In order to evaluate the supervised NDVI threshold classification, an area (349 pixels × 329 pixels) of image is classified based on the NDVI ratio. The grayscale image of NDVI is shown in Figure 2.

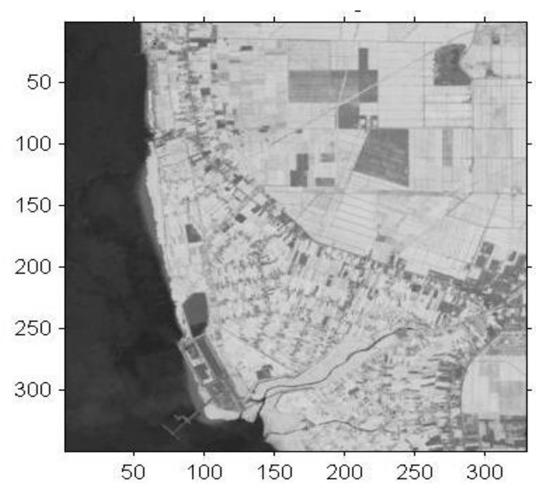

Figure 2: Grayscale image of NDVI

The classification process showed the color mapping and labeling of 3 classes. As we can observe in Figure 3, the images are grouped into 3 categories based on NDVI value. Water class shows the low coverage for NDVI. The other class for non-vegetation (contains urban, bare land, etc.) show the medium value for NDVI. The high value for vegetation contains the oil palm and the other green plant.

In this supervised classification, the NDVI image based on threshold values are calculated and the pixels of image are classified. Thus, those pixels which have NDVI ratio more than 0.35 (contain moderate and high NDVI values) are assigned to the vegetation class.

Table 2 shows the related confusion matrices assessment and classification results (overall accuracy and kappa coefficient). Based on the results, the total pixels are correctly classified is very high rather than misclassified pixels for every category class.

The misclassified pixels for the water class are only 66 pixels into non-vegetation category. The misclassified pixels for the vegetation class are in the other two classes; water and non-vegetation, 448 and 929, respectively. The misclassified pixels for non-vegetation class are also into the other two classes; water and vegetation, 2882 and 780, respectively.

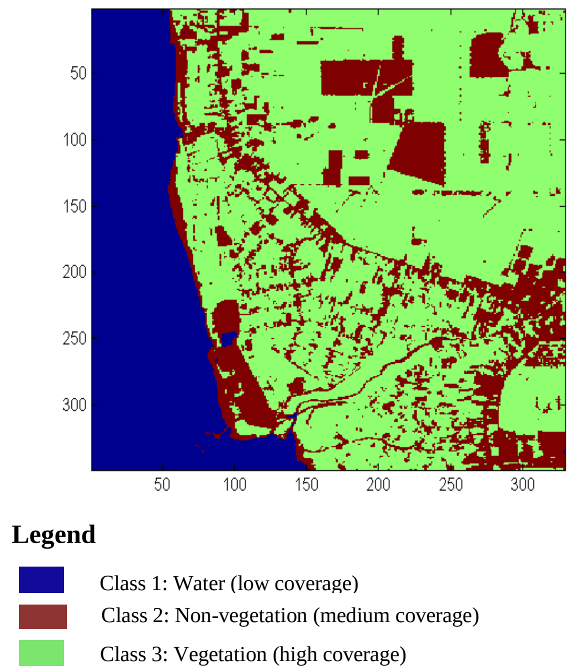

Figure 3: Color mapping and labelling.

The user's accuracy and producer's accuracy were obtained from the submission every row and column from the confusion matrix table. The correctly classified pixels of every row and column are divided with the sum of total pixels for every row and column. The percentage for overall accuracy that obtained from (3) is 95.55%, where the result is highly accurate. Kappa coefficient is 0.915 where in the range of positive value. This result is classified into high accuracy for kappa coefficient.

B. *Discussion*

The confusion matrix analysis of any classification process for satellite data must be quantitatively done to determine their accuracy of classification. In this supervised classification**,** the NDVI images are categorized into certain classes according to the NDVI value. Hence, the training sample for classification of NDVI image is selected based on the knowledge of the user (referring NDVI value on Table 1).

Table 2
Confusion Matrices

	Water	Vegetation	Non-vegetation	Sum	User's Accuracy
Water	25,600	448	2,882	28,930	0.88
Vegetation	0	63,090	780	63,870	0.99
Non-vegetation	66	929	21,026	22,021	0.95
Sum	25,666	64,467	24,688	114,821	
Producer's accuracy	1.00	0.98	0.85		

V. CONCLUSION

In this study, classification of Landsat 8 satellite data was/is carried out based on NDVI thresholds. The three classes of interest are vegetation, non-vegetation and water. Accuracy of the classification has been assessed based on the overall classification accuracy and Kappa coefficient. The result shows that NDVI - based method is able to classify Landsat 8 satellite data with a promising accuracy.

ACKNOWLEDGEMENT

This research was supported by University Technical Malaysia Melaka and Ministry of Education Malaysia under grants no PJP/2015/FTMK(4B)/S01433.

REFERENCES

[1] J. W. Rouse, R. H. Haas, J. A. Schell, and D. W. Deering, "Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation," 1973.

[2] D. Xu, "Compare NDVI Extracted from Landsat 8 Imagery with that from Landsat 7 Imagery," *Am. J. Remote Sens.*, vol. 2, no. 2, p. 10, 2014.

[3] A. Ahmad, B. M. Aboobaider, M. S. M. Isa, N. M. Hashim, M. Rosul, S. Muhamad, and S. Man, "Temporal changes in urban green space based on normalized difference vegetation index," *Appl. Math. Sci.*, vol. 8, no. 55, pp. 2743–2751, 2014.

[4] R. Geerken, B. Zaitchik, and J. P. Evans, "Classifying rangeland vegetation type and coverage from NDVI time series using Fourier Filtered Cycle Similarity," *Int. J. Remote Sens.*, vol. 26, no. 24, pp. 5535–5554, 2005.

[5] a. K. Bhandari, a. Kumar, and G. K. Singh, "Feature Extraction using Normalized Difference Vegetation Index (NDVI): A Case Study of Jabalpur City," *Procedia Technol.*, vol. 6, pp. 612–621, 2012.

[6] H. Fan, X. Fu, Z. Zhang, and Q. Wu, "Phenology-Based Vegetation Index Differencing for Mapping of Rubber Plantations Using Landsat OLI Data," *Remote Sens.*, vol. 7, no. 5, pp. 6041–6058, 2015.

[7] B. Satyanarayana, K. A. Mohamad, I. F. Idris, M.-L. Husain, and F. Dahdouh-Guebas, "Assessment of mangrove vegetation based on remote sensing and ground-truth measurements at Tumpat, Kelantan Delta, East Coast of Peninsular Malaysia," *Int. J. Remote Sens.*, vol. 32, no. 6, pp. 1635–1650, 2011.

[8] A. Viña, A. a. Gitelson, A. L. Nguy-Robertson, and Y. Peng, "Comparison of different vegetation indices for the remote assessment of green leaf area index of crops," *Remote Sens. Environ.*, vol. 115, no. 12, pp. 3468–3478, 2011.

[9] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.*, vol. 37, no. 1, pp. 35–46, 1991.

[10] NDVI, the Foundation for Remote Sensing Phenology.