Utilizing Landsat 8 OLI for land cover classification in plantations area

S N Anua and W V C Wong*

Faculty of Tropical Forestry, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia

* Corresponding author: w.wilson@ums.edu.my

Abstract. Identifying the land cover in plantations is crucial to assist the management of an area. Today, land cover classification can be achieved using free satellite data. The objective of this study is to perform a supervised classification using LANDSAT 8 OLI to differentiate the land cover in Brumas Camp which consists of non-vegetation, oil palm, forest, and forest plantations. The overall accuracy and Kappa’s coefficients were 71.64% and 0.62, respectively. We found out that the accuracy of classification for non-vegetation is relatively higher compared to vegetation land cover types. The non-vegetation land cover has distinct spectral reflectance which is useful to differentiate between non-vegetation and vegetation land covers.

Keywords: Natural forest; plantations; supervised classification; Landsat 8 OLI.

1. Introduction

Monitoring the land cover and land use is very important to analyze its impacts on the environment. Most of the changes caused by anthropogenic activities such as forest logging, forest fire, and agriculture were hard to see from the ground. Therefore, compared to the conventional method, which is time-consuming and labor extensive, the application of remote sensing can be a reliable tool for monitoring the land-use change.

Many studies had been using remote sensing multispectral data for land cover classification such as Landsat satellite [1], Sentinel [2] and SPOT [3], and many more. Some of these satellite data were cost-free and able to cover a large area of interest. However, the con of these data is when cloud or the cloud shadow that might cover the area of interest which can be affecting the monitoring activity.

Earth’s surfaces have their unique spectral reflectance and emittance properties that are useful for classification [4]. In remote sensing, image classification techniques use to group the pixels of the equivalent properties into the same class [3]. In this study, supervised classification is used for the classification. The supervised Classification concept is to segment the spectral domain into regions that correspond with the land cover class of interest [5]. Supervised classification required the analyst to create training samples that will be trained into the algorithm to apply to the image.

However, the abilities for the remote sensing data to distinguish between natural forest and plantation would be impressive and useful to accurately monitor the natural forest loss and plantation expansion [2]. Therefore, the objective of this study is to perform a supervised classification to differentiate the land cover in the Brumas area using Landsat 8 OLI.
2. Methodology

2.1 Study area
The study site was conducted in Brumas Camp of Tawau district, in Sabah, Malaysia (4°43′31.61″N - 4°30′39.00″N, 117°39′51.45″E - 117°48′1.93″E) (Figure 1). This 41,505-hectare area consists of oil palm plantation, forested area, tree plantations (Eucalyptus pellita (EP), Eucalyptus hybrid (EH), and Albizia falcataria (AF)), and settlement area. The topography condition is mainly low mountains and hills (300 m elevations).

2.2 Image analysis
The data acquired from Landsat 8 OLI (path 117, rows 57) was used in this study with 30 m resolution and 11 spectral bands. The least cloud-covered images were taken on 16 March 2019 and obtained from United States Geological (USGS) Earth Explorer (https://earthexplorer.usgs.gov/).

Figure 2 shows the steps were satellite image acquisition, pre-processing, supervised classification, ground verification, accuracy assessment, and output for the classification map. Pre-processing is to remove image distortion caused by the topography issue, noise from the sensor irregularity, and the atmospheric conditions to the satellite images for further analysis. The topographic and atmospheric correction was done using image processing in Tntmips 2017.

Figure 1. Location of the study area at Brumas Camp, Tawau.

Figure 2. The flow chart of the study.
Image enhancement and band combination was performed using ERDAS IMAGINE 2015 to improve the image visually for better interpretation. Band combination is useful for image interpretation and information on land cover. This study used a ‘natural colour’ combination (4,3,2) for analysis.

The Maximum Likelihood Classification (MLC) method was used for supervised classification using ERDAS IMAGINE 2015 (Table 1). For this method samples were taken for the training set. The training set is then used as the reference for the classifier to predict the class of unseen pixels.

**Table 1.** The land covers classification class.

| No. | Land Cover          | Description                                      |
|-----|---------------------|--------------------------------------------------|
| 1   | Non-vegetation (NV) | Land covered by buildings or exposed soil.        |
| 2   | Oil Palm (OP)       | Lands covered by oil palm.                       |
| 3   | Forest (F)          | Lands covered by naturally grown trees.          |
| 4   | Tree Plantation (TP)| Lands covered with fast-growing tree species (EP, EH, and AF). |

The planting year map was obtained from the company as the secondary data for this study. The Google Earth Pro and Garmin GPSMAP 64s Handheld GPS were used to navigate the location for the ground truth samples. A total of 200 random sample points were collected with their land cover.

An error matrix or contingency matrix used showed the information between the actual and prediction that had been classified through the classification for accuracy assessment. Overall accuracy shows the accuracy of the map. The individual classes’ accuracy is calculated using the user’s accuracy and producer’s accuracy. For the user’s accuracy, the number of correctly classified pixels was divided by the total number of pixels classified in that class. It measures the probability of the correct classified pixel with the actual class on the ground. While for producer’s accuracy is calculated by the number of correctly classified pixels divide with the total number of pixels as the reference data (ground truth point) in the class. It indicates the accuracy of the classifier.

Kappa coefficient measures the classifier performance derived from the error matrix without any bias from the chance agreement between the reference data and the classifier output [5]. The range of kappa statistics is between 1 to 0 whereas the higher the value, the better the performance of the classifier. The value of Kappa below 0.59 show (weak), 0.60-0.79 (moderate) and >0.80 show (strong) level of agreement [6]. The calculation of Kappa statistic is as follows;

\[
K = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} x_{i+} x_{+i}}{N^2 - \sum_{i=1}^{r} x_{i+} x_{+i}}
\]

where;

\[
\begin{align*}
    r &= \text{number of rows and columns in error matrix}, \\
    N &= \text{total number of pixels}, \\
    X_{ii} &= \text{observation in row i and column i}, \\
    X_{i+} &= \text{marginal total of row i}, \\
    X_{+i} &= \text{marginal total of column i}
\end{align*}
\]

**3. Results and discussion**

Band combinations 4,3,2 (Figure 3) gave a better interpretation of the land cover classes especially for vegetation and non-vegetation class. Bare land or non-vegetation pixels reflected more solar radiation compared to vegetation. Therefore, we could see darker green in the area with high vegetation especially forest area and mature tree plantation area. While for the area that had been harvested and less vegetation can be seen clearly without the green colour.
3.1 Supervised classification image
After the image enhancement and band combination, the signature information of each class was extracted. Then, the image was trained with the Maximum Likelihood Classifier (MLC) algorithm. Four land cover classes such as non-vegetation, oil palm, forest, and tree plantation classes were classified (Figure 4).

3.2 Accuracy assessment
Accuracy assessment is used to quantify the effectiveness of the pixels being classified at their correct land cover. A total of 281 points were generated in the study area. The overall accuracy of the classification was 81.14%. User’s accuracy range between 58.57% to 100.00 and producer’s accuracy is obtained 58.57% to 100.00%. The result of user’s accuracy and producer’s accuracy for non-vegetation was 69.47% and 89.29% respectively. The spectral response between vegetation cover and bare land is very distinct make it easier to differentiate [1].

However, tree plantation classes show a weak level of agreement for producer’s accuracy based on Table 2 with 58.57%. The classifier misclassified the ground point of tree plantation into the other land cover classes especially, forest. In each class have a specific spectral signature which is unique and
unidentical when interacted with visible and near-infrared [1]. However, forest and tree plantation have almost identical spectral responses and can be the reason for a high number of misclassifications of tree plantation into the forest.

Table 2. Confusion matrix of the supervised classification.

| Classified Data          | NV  | OP  | F   | TP  | Total | User’s Accuracy (%) |
|--------------------------|-----|-----|-----|-----|-------|--------------------|
| Non-Vegetation (NV)      | 71  | 3   | 1   | 8   | 83    | 85.54              |
| Oil Palm (OP)            | 0   | 50  | 0   | 6   | 56    | 89.29              |
| Forest (F)               | 0   | 14  | 66  | 15  | 95    | 69.47              |
| Tree Plantation (TP)     | 0   | 3   | 3   | 41  | 47    | 87.23              |
| Total                    | 71  | 70  | 70  | 70  | 281   | Overall Accuracy   |
| Producer’s Accuracy (%)  | 100.00 | 71.43 | 94.29 | 58.57 | 81.14%            |

Forest showed a moderate level of agreement for user’s accuracy with 69.47% compared to the other classes. Based on Table 2, a total of 29 samples from oil palm and tree plantations were misclassified as forest. User’s accuracy showed if the classification is reliable to the user [7]. The overall Kappa coefficient obtained for this study was 0.75 which is rated as a moderate level of agreement between the classified map and reference data.

The classification results can be improved by using combination between different classifications method [8]. Supervised classification methods have some problem with complex landscapes therefore hybrid approaches can use to cover the weakness of this approach. Moreover, application of Object-based classification (OBIA) can improve the classification since it use geographic objects as basic unit for the classification [9]. This approach able to decrease the salt-and-pepper effects that commonly found in pixel-based classification.

4. Conclusion

The classification of the Brumas area using Landsat 8 concludes with a moderate level of agreement with overall accuracy and Kappa coefficient, 81.14% and 0.75 respectively. MLC can classify the non-vegetation accordingly but have an error in classifying the vegetation land cover, especially forest and tree plantation area. The spectral reflectance that emits by this landcover almost the same, therefore further studies are needed to understand how to differentiate between this landcover.

5. References

[1] Anuar M B, Jusoh A N, Suhaili, I and Munir A 2017 Mapping vegetation cover of Acacia mangium plantation by age. Trans. Sci. Technol. 4(4) 482–488
[2] Spracklen B and Spracklen D V 2021 Synergistic use of Sentinel-1 and Sentinel-2 to map natural forest and acacia plantation and stand ages in North-Central Vietnam. Remote Sens. 13(2) 1–19
[3] Ismail M H, Pakhriazad H Z and Shahrin M F 2009 Evaluating supervised and unsupervised techniques for land cover mapping using remote sensing data. Malaysian Journal of Society and Space 5(1) 1-10
[4] Al-doski J, Mansor S B, Zulhaidi H and Shafri M 2013 Image classification in remote sensing. Journal of Environment and Earth Science 3(10) 141–148
[5] Richards J A 2013 Remote Sensing Digital Image Analysis, 5th ed. Berlin, Heidelberg: Springer Berlin Heidelberg
[6] McHugh M L 2012 Lessons in biostatistics interrater reliability: The kappa statistic. Biochem. Medica 3(22) 276–282
[7] Rwanga S S and Ndambuki J M 2017 Accuracy Assessment of land use/land cover classification using remote sensing and GIS. Int. J. Geosci. 8(4) 611–622
[8] Lu D and Weng Q 2007 A survey of image classification methods and techniques for improving classification performance. Int. J. Remote Sens. 28(1) 823–870
[9] Peña J M, Gutiérrez P A, Hervás-Martínez C, Six J, Plant R E and López-Granados F 2014 Object-based image classification of summer crops with machine learning methods. Remote Sens. 6(6) 5019–5041

Acknowledgments
The authors would like to thank Sabah Softwoods Berhad for allowing their plantation area as the study sites and Universiti Malaysia Sabah G.I.S. Lab members.