# Mapping the spatial distribution and changes of oil palm land cover using an open source cloud-based mapping platform

Nur Shafira Nisa Shaharum[a], Helmi Zulhaidi Mohd Shafri[a,b]*, Wan Azlina Wan Ab Karim Ghani[c], Sheila Samsatli[d], Husni Mobarak Prince[a], Badronnisa Yusuf[a], and Ahmed Mohamed Hamud[a]

[a]*Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia;* [b]*Geospatial Information Science Research Centre (GISRC), Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia;* [c]*Department of Chemical and Environmental Engineering/Sustainable Process Engineering Research Centre (SPERC), Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia;* [d]*Department of Chemical Engineering, University of Bath, Claverton Down, BA2 7AY, United Kingdom.*

*corresponding author: helmi@upm.edu.my

# Mapping the spatial distribution and changes of oil palm land cover using an open source cloud-based mapping platform


## ABSTRACT

Oil palm has become well known for its oil palm yields that can be used to produce food, biodiesel and biogas. The rapid expansion of oil palm plantations over large areas has changed the land use and land cover of surroundings. Changes in land covers can be mapped and later used for further analysis. However, obtaining and classifying large coverages require massive amounts of data and computing resources and the skills and time of analysts. The remote ecosystem monitoring assessment pipeline (REMAP) provides a cloud computing platform that hosts an open-source stacked Landsat data that allows land cover classification to be implemented using a built-in random forest supervised machine learning algorithm. Classifications were performed with the aid of predictor layers to discriminate the following land covers in Peninsular Malaysia: oil palm, built-up, bare soil, water, forest, other vegetation and paddy. The classification performed on Period 1 (1999–2003) and Period 2 (2014–2017) data produced an overall accuracy of 80.34% and 79.53% respectively. The analysis of the changes in oil palm distributions from Period 1 to Period 2 indicated an increment of 17.24%. Further analysis revealed that oil palm expansion in Peninsular Malaysia only minimally affected forested area and is mostly resulted from the conversion of less productive crops to oil palm. Results prove the land cover mapping and change detection capabilities of REMAP as a cloud computing platform for large areas. Despite its limitations, REMAP has the potential to achieve fast-paced mapping over large areas and monitor land changes in oil palm distributions.


## 1. Introduction

Oil palm is one of the important crops that can produce the highest amount of biomass. It is mainly used for frying, and its high contents of good cholesterol or high-density lipoproteins and bad cholesterol or low-density lipoproteins enables it to balance the cholesterol in our bodies (Ong and Goh 2002; Mukherjee and Mitra 2009). In addition, the powerful antioxidants in palm oil provide protection to the body and may prevent cancer (Marcene 2018). Oil palm plantations store higher amounts of carbon than other agriculture land uses, such as rubber, and are thus among the most profitable land uses (Rival 2014). Over the past decades, oil palm has been processed and extracted for many uses. After Indonesia, Malaysia is the largest oil palm exporter; consequently, Malaysia must properly manage its oil palm plantations, especially as the number of planted oil palm areas is expected to grow in the future (Sumathi, Chai, and Mohamed 2008; Chong et al. 2017). Globally, many issues have been reported due to oil palm activities, including uncontrolled deforestation, loss of biodiversity and energy crisis (Chuah et al. 2006; Fitzherbert et al. 2008; Sulaiman et al. 2011; Vijay et al. 2016). During the 73rd United Nations General Assembly, oil palm plantations in Malaysia were identified as a cause of severe damage to wildlife habitat, and the produced palm oil was declared harmful to health. However, Malaysia's Prime Minister Tun Dr Mahathir emphasised the environmental aspect that significantly had been taken into account before planting the oil palm trees, and that approximately 48% of the land cover in Malaysia is still covered by virgin forest (Mohamad 2018). To ensure that oil palm production is well established, the Malaysian Palm Oil Certification Council (MSPO) will become mandatory. The Roundtable on Sustainable Palm Oil (RSPO) was formed to promote the growth and usage of sustainable palm oil products to lessen negative impacts on the environment. This initiative aims to help reduce deforestation and preserve biodiversity. Eight principles must be followed by growers to ensure that they are RSPO certified, including the use of appropriate best practices by growers and millers, environmental responsibility, conservation of natural resources and biodiversity, and commitment to continuous improvement in key areas of activity (RSPO 2013; Tillack 2013). The International Union for Conservation of Nature (IUCN) introduced 17 sustainable development goals (SDGs) that must be achieved by 2030 (Blanc 2015). The impact of energy crisis has forced Malaysia to find alternatives to fossil fuel. Aside from having health benefits and contributing to food production, oil palm is the most suitable element for use as a renewable energy source (Chuah et al. 2006; Shuit et al. 2009). Even though oil palm provides countless benefits, it has negative impacts when not properly managed. Hence, producing oil palm land cover maps is important to monitor and study the pattern that contributes to oil palm changes and to examine the environmental impacts, especially in fulfilling the RSPO's requirements. Indirectly, the study can contribute to some of the SDGs that were presented by the IUCN, namely, life habitat and providing affordable and clean energy. Even though oil palm biomass provides benefits as the major biomass crop to produce a source of renewable energy, it requires proper monitoring especially in large areas (Ibrahim 2014). Therefore, sufficient data are required to ensure efficient and proper planning and management of oil palm products. Oil palm monitoring is time-consuming and requires expensive tools. Furthermore, many countries have insufficient funds and lack resources for regular surveys, making palm monitoring expensive (Chong et al. 2017).

Remote sensing can collect ground information from a large area in a short time. It captures spatial data without any direct contact and has been applied in various fields, such as urban areas, agricultural land, biomass estimation, oil palm disease detection, hazard prediction, object detection and biodiversity monitoring (Thenkabail et la. 2004; Shafri et al. 2012; Cammalleri et al. 2014; Dudley at al. 2017; Gambo et al. 2018). In response to the growing demand of oil palm plantations, several studies on oil palm mapping have been conducted using remote sensing. Koh et al. (2011) conducted a study using 250 m spatial resolution data to map oil palm distributions over 2 million hectares. In a study that aimed to detect oil palm disease, Shafri and Hamdan (2009) utilised hyperspectral data obtained using an advanced imaging spectrometer for applications to produce a map of disease infection in oil palms. A study on oil palm tree counting application was successfully conducted using high-spatial-resolution data acquired from airborne imagery (Shafri, Hamdan, and Saripan 2011). Li et al. (2015) used 50 m orthorectified mosaic phased array L-band synthetic aperture radar images to map oil palms in Cameroon and completed their study by using machine learning algorithms.  Despite the good quality of data, obtaining data from a large area entails high costs. Therefore, study areas are limited to a small coverage. Few authors have been able to overcome this limitation by using coarse spatial resolution, such as Landsat data. Landsat provides open-source remote sensing data and has a long record of continuous data acquisition that has made it suitable not only for monitoring but also for change detection analysis (Azzari and Lobell 2017; Zhu 2017; Gambo et al. 2018). Landsat offers 30 m spatial resolution images and is thus suitable for use in crop monitoring, land cover mapping and other analyses. Wahid, Nordiana, and Tarmizi (2005), Cheng et al. (2016), Asari, Suratman, and Jaafar (2017) and Miettinen, Gaveau, and Liew (2018) successfully used Landsat data for oil palm mapping. They produced maps using a maximum likelihood classifier (MLC), vegetation indices (VIs) and support vector machine (SVM). Lee et al. (2016) conducted a study using random forest (RF) classification and regression tree (CART) and minimum distance algorithms to map oil palm land cover with the Landsat data obtained from Google Earth Engine (GEE). However, one of the drawbacks of Landsat is the presence of clouds in the data, which can affect image quality. Furthermore, clouds might cover the most crucial area that contains ground information essential for analysis. Even though the previous studies produced acceptable maps using Landsat data, they were implemented in small areas, mostly in tropical countries such as Malaysia, because image processing and classifications for large areas require a considerable amount of time, resources and effort (Franklin et al. 2015; Fahnestock et al. 2016). Additionally, in Malaysia, mapping and detecting changes in oil palm spatial distributions suffer from several setbacks due to limited computing resources, skilled manpower and cloud-free data. To counter these issues, this study utilises an open-source cloud-based analysis platform called remote ecosystem monitoring assessment pipeline (REMAP) to produce oil palm land cover maps from two periods of Landsat data compositions (1999–2003) and (2014–2017) for use in change detection analysis in Peninsular Malaysia. To the best of our knowledge, this study is the first to use and test a cloud computing remote sensing tool for oil palm mapping in Malaysia.

## 2. Data and methods

### 2.1 *Study area*

Malaysia is the second largest oil palm producer in the world, next to Indonesia. As a country with a humid climate and land that covers more than 5 million hectares of oil palm plantations, Malaysia has abundant oil palm biomass crops (Shuit et al. 2009). According to the Malaysian Palm Oil Board (MPOB), the oil palm plantation area of the country has increased over the past few years, resulting in changes in the land cover. To test the efficiency of mapping over a large area using cloud-based REMAP, Peninsular Malaysia was chosen as the study area, as shown in Figure 1. Peninsular Malaysia or West Malaysia has 11 states and a land area of approximately 132,265 km$^2$.

**Figure 1.** [near here]

Peninsular Malaysia has various land covers, including green vegetation, urban areas, water bodies and bare land. Firstly, this study aims to map the land cover of Peninsular Malaysia by using the Landsat data acquired from Period 1 (1999–2003) and Period 2 (2014–2017). Then, the classified maps are assessed in terms of the changes in oil palm distributions throughout the periods. To achieve these objectives, this study implements a cloud computing technique using an open-source cloud-based analysis platform. In addition, the machine learning approach is applied to classify the maps. The flow of this study is shown in Figure 2.

**Figure 2.** [near here]

### 2.2 ***Random forest image classification***

RF is a type of machine learning algorithm. Machine learning is a subset of artificial intelligence and has several categories, including supervised and unsupervised classifications. For image classification, supervised machine learning works by classifying the image from the known data, whereas unsupervised classification classifies the image with no known data (Hasmadi, Pakhriazad, and Shahrin 2009). RF is a powerful machine learning algorithm that performs well in image classification and regression (Svetnik et al. 2003). Besides RF, there are other machine learning algorithms have been used for image classification, such as SVM, artificial neural network (ANN) and decision tree (DT) (Belgiu and Drăgut 2016; Lary et al. 2016; Singh et al. 2016). RF is a supervised machine learning algorithm that works similarly to DT by combining decisions into a tree-like model. However, RF is more powerful and robust than DT because it combines tree-like models and becomes a forest, as shown in Figure 3 (Breiman 2001; Feng, Liu, and Gong 2015). Then, when the system receives new input, it will go through the trees in the forest.

**Figure 3.** [near here]

Rodriguez-Galiano et al. (2012) utilised RF to map land cover from Landsat data. They tested several trees and determined the best classified map by obtaining the highest kappa index. RF can classify data despite the missing values within the trees. The obtained information or value is assigned in each node and allows RF to study and identify the feature. The structure that consists of the combination of many trees that carry large amounts of information makes RF a powerful machine learning algorithm (Gislason, Benediktsson, and

Sveinsson 2006). A more recent study on the implementation of RF for crop classification was applied by Tatsumi et al. (2015) using Landsat 7 ETM+ data. They managed to classify eight types of crops from the medium-resolution Landsat data and obtained 81% of overall accuracy.

### 2.3 *Cloud-based remote ecosystem monitoring assessment pipeline*

Cloud computing is the delivery of computing services, such as servers, storage, networking, databases and analytics (Hashem et al. 2015). It can be used to perform data analysis, create new apps and host websites, as shown in Figure 4. The development of cloud computing has had a considerable impact on information technology, and cloud computing is widely used in large companies, including Google, Amazon and Microsoft (Armbrust et al. 2010). It was developed to reduce the time and cost required to perform related works. Fortunately, cloud computing is not limited to data management but also offers an effective way of executing remote sensing computing (Wang et al. 2013).

**Figure 4.** [near here]

REMAP is a cloud-based platform that enables users to perform image classification. REMAP was introduced by Murray et al. (2018) and it can be used to perform land cover mapping and change detection analysis. In addition, REMAP uses geospatial data and the storage capacity of GEE, thus allowing REMAP to process and develop classified maps in the cloud in just few minutes without the need of high computational computers. REMAP is an open-source cloud-based analysis platform that provides fast land cover classification using a built-in RF machine learning algorithm. On the other hand, the utilisation of machine learning algorithms for image classification via software such as ENVI 5.3 (Exelis Visual

Information Solutions, Boulder, CO, USA) and Erdas Imagine (ERDAS Inc., Atlanta, GA, USA) software will entail huge amount of time to complete the processing (Shaharum et al. 2018). Furthermore, it requires knowledge of remote sensing to perform image pre-processing, including handling the tools provided in the software. On top of that, massive effort, cost and time are needed to process vast amount of remote sensing data covering Peninsular Malaysia. Nevertheless, REMAP provides user-friendly platform that allows a user to access massive satellite data archives directly and handle the technical details of remote sensing that focus on training, classifying and improving the generated classified maps because it uses GEE to perform the workflow shown in Figure 5 (Murray et al. 2018). In addition, REMAP allows a beginner in remote sensing to perform image classification and monitor land use change over time (Shih 2018).

### 2.4 *Satellite data*

This study utilised the optical remote sensing data acquired from Landsat Enhanced Thematic Mapper and Operational Land Imager. Landsat data provide a moderate scale with 30 m spatial resolution images. However, optical sensors are sensitive to clouds, thereby reducing the quality of images in the presence of clouds. Gambo et al. (2018) and Shaharum et al. (2018) performed image stacking via Smart GeoFill to produce cloud-free image of a protected area, Krau Wildlife Reserve. Although they have successfully produced cloud-free image using several number of Landsat data, certain amount of time was required to process and stack all the data. To address this issue, REMAP provides ready-stacked images by compiling numerous images from two different periods, 1999–2003 (historical) and 2014–2017 (current). In this paper, 1999–2003 is Period 1 and 2014–2017 is Period 2. The implementation of image stacking using the FMASK algorithm in GEE led to the production of quality Landsat images with fewer clouds, thereby allowing users to perform remote sensing application analysis by supporting environmental conservation, including biodiversity, land monitoring, hotspot identification and ecosystem mapping through the workflow shown in Figure 5 (Murray et al. 2018).

**Figure 5.** [near here]

**2.5** ***Sample collection and predictor selection***

The training and testing samples were divided into seven classes: oil palm, forest, paddy, water, bare soil and other vegetation. The samples were selected based on the high-resolution image provided in REMAP, as shown in Figure 6.

**Figure 6.** [near here]

The samples were chosen using points that can be downloaded as JSon and CSV formats. REMAP allows users to upload their own training samples to perform classification. In this study, samples were created separately for each state, and 70% were used for training and 30% for testing. The training samples were used to perform image classification, and the testing samples were used to validate the classified maps. REMAP provided several predictor layers as shown in Table 1, and these layers are additional information that complements the primary information for RF image classification in REMAP. The user can easily choose and test different combinations of predictor layers for image classification to obtain the best output, given that RF is a powerful machine learning algorithm that classifies data based on assigned parameters (Svetnik et al. 2003).

**Table 1.** [near here]

Layers, such as NDVI and slope, are useful for extracting oil palm features. Shafri and Hamdan (2009) conducted a study using NDVI to detect and map disease in oil palm and accurately determined the condition of oil palm trees. A study on biomass estimation of forest and oil palm was conducted using several indices, including NDVI and NIR (Morel, Fisher, and Malhi 2012). Therefore, before any image classification can be performed in REMAP, different combinations of predictor layers were tested in the Selangor area, and the results were evaluated. After several evaluations of the classified Selangor maps, the combination of all layers listed in the spectral and topographic predictor layers was found to be the best and thus used to perform image classification in other states.

### 2.6 *Accuracy assessment*

Accuracy assessment is important in remote sensing applications. It compares the classified image with the ground truth data to measure the consistency between the classified map and the actual data. It also calculates the accuracy of the classified map, including classification errors. In addition, this method has also been used to measure and compare the ability of remote sensing algorithms in classifying images (Shaharum et al. 2018). Many approaches can be adopted to measure the accuracy of a classified map, but a common technique via the confusion matrix that was explained by Foody (2002) was chosen in this study to assess the accuracy of the classified maps. With the help of the data provided by Department of Agriculture, this study created training and testing samples. Of the total samples, 30% were used as testing samples to validate the accuracy of the classified maps. To avoid sample overlapping, 30% of the testing samples were generated in REMAP and exported to ArcMap version 10.4.1 to perform the assessment via the confusion matrix.

The confusion matrix is normally expressed in table form and contains correctly and incorrectly classified pixel values. The classification example produced using three classes is shown in Table 3.

**Table 2.** [near here]

| | |
|---|---|
| True positives (TP) | : Model detects the condition when the condition is present |
| True negatives (TN) | : Model does not detect the condition when the condition is absent |
| False positives (FP) | : Model detects the condition when the condition is absent |
| False negatives (FN) | : Model does not detect the condition when the condition is present |

**Table 3.** [near here]

The summation of TP and TN defines the accuracy of the model, while FP and FN are the errors of the classification model, which are known as Type 1 and Type 2 errors, respectively. Several types of information can be extracted from the table, such as the producer's accuracy (PA), the user's accuracy (UA), the overall accuracy (OA) and the kappa coefficient. Since this study focuses on the changes of oil palm from Periods 1 to 2, the calculation of PA and UA will be done only for oil palm classification. PA is the probability of the pixel to be classified correctly to the ground truth, which indirectly indicates how many pixels show the same feature as the reality. UA is the probability of the pixels predicted in several classes that belong to that class. OA is calculated based on the number of correctly classified pixels divided by the total number of pixel values. The kappa coefficient measures the agreement between the classified and true values (Foody 2002). The confusion matrix shown in Table 3 is expressed in Table 4.

**Table 4.** [near here]

## 2.7 *Change detection of land use and land cover*

Aside from producing maps, one of the fundamentals in remote sensing image analysis is detecting changes in land use and land cover. Several methods can be used to perform change detection analysis, such as image differencing, image rationing and post-classification comparison (Afify 2011). Ideally, change detection is a post-classification technique that identifies changes or differences by using remotely sensed data. It provides a worldwide monitoring program by integrating the use of spatial, historical and spectral datasets. These datasets can be utilised to measure changes in vegetation, soil content, land cover and urban expansion (Mouat, Mahin, and Lancaster 2008; Wu et al. 2017).

Several studies implementing change detection analysis were conducted using Landsat data. One of these studies was conducted by Zhu and Woodcock (2014) to detect changes in forest land cover. Sets of images obtained from 2001 to 2002 were used to monitor forest changes (Zhu, Woodcock, and Olofsson 2012). El-Kawy et al. (2011) performed change detection analysis to observe the conditions of the western Nile delta. They utilised Landsat data obtained from four different years. Change detection was applied using the post-classification approach by classifying the image via supervised classification to detect changes in agriculture and barren lands. The changes from one feature to another can be seen and observed. Thus, the results can be analysed for further action. The present study utilised the same post-classification technique in the change detection analysis of oil palm land cover over a large area. This technique was applied on two sets of stacked Landsat data from Periods 1 and 2, which were classified using RF.

Utilising the post-classification technique in change detection requires at least two different images from two different periods. It is one of the easiest methods that result in a direct from-to results from the classified images. The changes for each class were compared and analysed by using different classified images (Almutairi and Warner 2010). The changes can be calculated using the following formula:

$$A_1 - A_2 = A \qquad \text{(Eq. 1)}$$

$A_1$ = Area of a class in Period 1
$A_2$ = Area of a class in Period 2
A = Difference between the areas of the two periods

## 3. Results and discussion

### 3.1 *Land cover assessments*

Figures 7(a) and 7(b) show the land cover maps of Peninsular Malaysia produced for Periods 1 and 2. The OA, PA and UA of oil palm for all states are listed in Table 5.

**Figure 7.** [near here]

The classified maps were evaluated by using the testing samples (30% of the total samples), the assessment was conducted by using a tool provided in ArcMap version 10.4.1, and the calculation was performed using the confusion matrix as explained in Section 2.6. A forest consists of dense trees, mangroves and forest, while other vegetation feature plantations, less dense forests and agriculture crops other than oil palm and paddy. Bare soil consists of bare land, that is, open space, sand and areas where trees have been cleared. Lastly, built-up areas consist of buildings and roads.

**Table 5.** [near here]

The OA of the classified maps produced for Periods 1 and 2 are 80.34% and 79.53% respectively. For Period 1, Selangor obtained the highest OA of 87.39% with a kappa value of 84%, while Kelantan obtained the least OA of 72.99% with a kappa value of 67.17%. For Perlis, the oil palm area was small (MPOB 2018). Although the samples for oil palm in Perlis were created, the samples were limited, and the area was misclassified as forest. Perak produced the highest OA of 85.43% for Period 2, while Selangor had the least OA of 74.24%. Table 6 shows that most of the obtained PAs and UAs have an accuracy of more than 75%. However, the UAs and PAs produced for Periods 1 and 2 varied and were further investigated based on the oil palm area tabulated in Table 5.

**Table 6.** [near here]

Table 6 shows the oil palm area obtained from REMAP and the MPOB inventory for each state in Peninsular Malaysia for Periods 1 and 2. By comparing the results obtained from REMAP with the inventory from MPOB, the errors produced do not exceed 15%. Terengganu produced the nearest oil palm area to the MPOB statistics for Period 1, while Melaka produced the nearest oil palm area for Period 2. Although the UA and PA produced for Melaka in Period 1 exceeded 70%, the mapped oil palm area was sparser than indicated by the MPOB statistics. Kelantan produced an overestimated oil palm area by having more than twice the oil palm area provided by MPOB. Misclassifications of oil palm as other classes, such as forest and other vegetation, probably occurred due to the similarity of the spectral signal. Moreover, the effect of image stacking, which probably contains different illuminations, could contribute errors to image classification (Cheng, Han, and Lu 2017; Gambo et al. 2018; Shaharum et al. 2018).

### 3.2 *Analysis of the changes in the oil palm area*

The changes were observed and analysed using a post-classification technique. The increment in oil palm area that occurred from Periods 1 to 2 for each class is illustrated in Figure 8.

**Figure 8.** [near here]

Figure 8 shows the estimated percentage of each class in Period 1 that was converted to oil palm as classified in Period 2. The rate of conversion from forest to oil palm was 12.96%. Despite the misclassifications between oil palm and forest, this rate alone cannot explain the loss of forest to oil palm land cover (Vijay et al. 2016). Gatti et al. (2019) stated that oil palm plantation has caused deforestation and increased in number of tree loss. They conducted a study using data obtained from Global Forest Watch (GFW), and concluded that certified productions of oil palm are not sustainable. However, Hegarty and d'Enghien (2018) found that data obtained from GFW has several limitations: cannot distinguish between oil palm and natural forest, failed to distinguish between tree cover loss and forest cover loss and plantation regrows is considered as tree loss. Moreover, zero tree removal was detected in both certified and uncertified oil palm productions. It is because replanting was taking place and although oil palm area continued to grow from Periods 1 to 2, the deforestation rate started to slow down from 1980s (Miyamoto et al. 2014). These findings revealed that oil palm plantation in Peninsular Malaysia is sustainable and it is not the main proximate cause of deforestation. Figure 8 clearly shows that other vegetation is the biggest contributor to the 17.24% of oil palm expansion measured from Periods 1 to 2 (14 years). A total of 70.98% of other vegetation classified in Period 1 was converted to oil palm partly due to the reduction of rubber trees within the 14 years (Miyamoto et al. 2014). Furthermore, Wicke et al. (2011) found that permanent crops, mainly export crops natural rubber and coconut, decreased significantly. These findings explain the high conversion rate of other vegetation in oil palm expansion. The increment of the oil palm area was also due to the replanting activities that converted bare soil and plantations into oil palm crops (Agus et al. 2013). Meanwhile, the degraded land remains stagnant over time (Wicke et al. 2011), and oil palm was reported to grow at a slow rate after 2010 (MPOB 2018).

## 4. Conclusion

With a powerful built-in RF machine learning algorithm and the availability of the cloud service, REMAP has become a powerful tool that can be used to generate land cover maps over a large area in a short time. Through the cloud computing service, REMAP provides quality images with fewer clouds worldwide. For the first time, the combination of cloud computing and machine learning via REMAP was utilised and proved successful in conducting oil palm mapping and change detection in Peninsular Malaysia. REMAP produced maps of oil palm distributions for the state of Peninsular Malaysia, with accuracies of 80.34% and 79.53% respectively for Periods 1 and 2. Furthermore, via the change detection approach, the rate and patterns of change could be observed and analysed (Wulder et al. 2018). The results produced from REMAP indicated that oil palm plantation in Peninsular Malaysia is sustainable and does not result in adverse effects on the forest environment, thus achieving the sustainable development objectives of the country. Although REMAP provides fast image processing and map making, it is limited only to

Landsat data and the RF classifier. Furthermore, the data are fixed and cannot be improved or adjusted. We have several recommendations to improve this study, such as the utilisation of other sensors, such as Sentinel 2 and Sentinel 1. Furthermore, optical and radar data can be integrated to improve image quality. Then, the utilisation of other machine learning algorithms, such as SVM, ANN and the deep learning approach, in the REMAP environment can be tested.

**Acknowledgements**

The authors would like to thank UPM for their facilities and for funding this research. Our warm gratitude also goes to the Engineering and Physical Sciences Research Council for their financial support through the BEFEW project under Grant EP/P018165/1. The comments from anonymous reviewers in improving this article are highly appreciated.

## References

Afify, H. A. 2011. "Evaluation of change detection techniques for monitoring land-cover changes: A case study in new Burg El-Arab area." *Alexandria Engineering Journal* 50 (2): 187-195. doi:10.1016/j.aej.2011.06.001.

Agus, F., P. Gunarso, B. H. Sahardjo, N. Harris, M. van Noordwijk, and T. J. Killeen. 2013. "Historical CO2 emissions from land use and land use change from the oil palm industry in Indonesia, Malaysia and Papua New Guinea." *Reports from the technical panels of the second greenhouse gas working group of the roundtable for sustainable palm oil (RSPO)* 65-88. http://www.rspo.org/file/GHGWG2/5_historical_CO2_emissions_Agus_et_al.pdf.

Alberg, A. J., J. W. Park, B. W. Hager, M. V. Brock, and M. Diener- West. 2004. "The Use of "Overall Accuracy" to Evaluate the Validity of Screening or Diagnostic Tests." *Journal of General Internal* Medicine 19 (5): 460-465. doi:10.1111/j.1525-1497.2004.30091.x.

Almutairi, A., and T. A. Warner. 2010. "Change Detection Accuracy and Image Properties: A Study Using Simulated Data." *Remote Sensing* 2 (6): 1508-1529. doi:10.3390/rs2061508.

Armbrust, M., A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, and M. Zaharia. 2010. “A view of cloud computing.” *Communications of the ACM* 53 (4): 50-58. doi:10.1145/1721654.1721672.

Asari, N., M. N. Suratman, and J. Jaafar. 2017. “Modelling and mapping of above ground biomass (AGB) of oil palm plantations in Malaysia using remotely-sensed data.” *International Journal of Remote Sensing* 38 (16): 4741-4764. doi:10.1080/01431161.2017.1325533.

Azzari, G., and D. B. Lobell. 2017. “Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring.” *Remote Sensing of Environment* 202: 64-74. doi:10.1016/j.rse.2017.05.025.

Blanc, D. L. 2015. “Towards Integration at Last? The Sustainable Development Goals as a Network of Targets.” *Sustainable Development* 23 (3): 176-187.

Belgiu, M., and L. Drăguţ. 2016. “Random forest in remote sensing: A review of applications and future directions.” *ISPRS Journal of Photogrammetry and Remote Sensing* 114: 24-31. doi:10.1016/j.isprsjprs.2016.01.011.

Breiman, L. 2001. “Random forests.” *Machine learning* 45 (1): 5-32. doi:10.1023/A:101093340.

Cammalleri, C., M. C. Anderson, F. Gao, C. R. Hain, and W. P. Kustas. 2014. “Mapping daily evapotranspiration at field scales over rainfed and irrigated agricultural areas using remote sensing data fusion.” *Agricultural and forest meteorology* 186: 1-11. doi:10.1016/j.agrformet.2013.11.001.

Chuah, T. G., W. A. W. A. K. Ghani, R. Yunus, & R. Omar. 2006. “Biomass as the Renewable Energy Sources in Malaysia: An Overview.” *International Journal of Green Energy* 3 (3): 323-346. doi:10.1080/01971520600704779.

Cheng, Y., L. Yu, A. P. Cracknell, and P. Gong. 2016. “Oil palm mapping using Landsat and PALSAR: a case study in Malaysia.” *International Journal of Remote Sensing* 37 (22): 5431-5442. doi:10.1080/01431161.2016.1241448.

Cheng, G., J. Han, and X. Lu. 2017. “Remote Sensing Image Scene Classification: Benchmark and State of the Art.” *Proceedings of the IEEE* 105 (10): 1865-1883. doi:10.1109/JPROC.2017.2675998.

Chong, K. L., K. D. Kanniah, C. Pohl, and K. P. Tan. 2017. “A review of remote sensing applications for oil palm studies.” *Geo-spatial Information Science* 20 (2): 184-200. doi:10.1080/10095020.2017.1337317.

Dudley, N., N. Ali, M. Ketiunen, and K. MacKinnon. 2017. “Editorial essay: Protected areas and the sustainable development goals.” *Parks* 23 (2): 10-12.

El-Kawy, O. A., J. K. Rød, H. A. Ismail, and A. S. Suliman. 2011. “Land use and land cover change detection in the western Nile delta of Egypt using remote sensing data.” *Applied Geography* 31 (2): 483-494. doi:10.1016/j.apgeog.2010.10.012.

ERDAS Inc. 2001. *Erdas Field Guide.* Atlanta: Georgia.

Fahnestock, M., T. Scambos, T. Moon, A. Gardner, T. Haran, and M. Klinger. 2016. “Rapid large-area mapping of ice flow using Landsat 8.” *Remote Sensing of Environment* 185: 84-94. doi:10.1016/j.rse.2015.11.023.

Fitzherbert, E. B., M. J. Struebig, A. Morel, F. Danielsen, C. A. Brühl, P. F. Donald, and B. Phalan. 2008. “How will oil palm expansion affect biodiversity?” *Trends in Ecology & Evolution* 23 (10): 538-545. doi:10.1016/j.tree.2008.06.012.

Franklin, S. E., O. S. Ahmed, M. A. Wulder, J. C. White, T. Hermosilla, and N. C. Coops. 2015. “Large Area Mapping of Annual Land Cover Dynamics Using Multitemporal Change Detection and Classification of Landsat Time Series Data.” *Canadian Journal of Remote Sensing* 41 (4): 293-314. doi:10.1080/07038992.2015.1089401.

Feng, Q., J. Liu, and J. Gong. 2015. “UAV Remote Sensing for Urban Vegetation Mapping Using Random Forest and Texture Analysis.” *Remote Sensing* 7 (1): 1074-1094. doi:10.3390/rs70101074.

Foody, G. M. 2002. “Status of land cover classification accuracy assessment.” *Remote sensing of environment* 80 (1): 185-201. doi:10.1016/S0034-4257(01)00295-4.

Gambo, J., H. Z. M. Shafri, N. S. N. Shaharum, F. A. Z. Abidin, and M. T. A. Rahman. 2018. “Monitoring and predicting land use-land cover (LULC) changes within and around krau wildlife reserve (KWR) protected area in Malaysia using multi-temporal Landsat data.” *Geoplanning: Journal of Geomatics and Planning* 5 (1): 23-44. doi:10.14710/geoplanning.5.1.17-34.

Gatti, R. C., J. Liang, A. Velichevskaya, and M. Zhou. 2019. “Sustainable palm oil may not be so sustainable.” *Science of The Total Environment* 652: 48-51. Advance online publication. doi:10.1016/j.scitotenv.2018.10.222.

Gislason, P. O., J. A. Benediktsson, and J. R. Sveinsson. 2006. “Random Forests for land cover classification.” *Pattern Recognition Letters* 27 (4): 294-300. doi:10.1016/j.patrec.2005.08.011.

Hashem, I. A. T., I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, and S. U. Khan. 2015. “The rise of “big data” on cloud computing: Review and open research issues.” *Information Systems* 47: 98-115. doi:10.1016/j.is.2014.07.006.

Hasmadi, M., H. Z. Pakhriazad. and M. F. Shahrin. 2009. “Evaluating supervised and unsupervised techniques for land cover mapping using remote sensing data.” *Malaysian Journal of Society and Space* 5 (1): 1-10.

Hegarty K., and P. B. d’Enghian. 2018. “Purdue’s Sub-par Palm Research.” Accessed June, 2018. http://palmoilmonitor.org/2018/12/12/purdues-sub-par-palm-research.

Ibrahim, A. 2014. “Sustainability of Malaysian Palm Oil Industry 2013/2014.” Accessed 17 July 2018. http://www.mpoc.org.my/upload/IPOSC-2014-Sustainability-of-Malaysian-Palm-Oil-Industry_Dr-Ahmad-Ibrahim.pdf.

Koh, L. P., J. Miettinen, S. C. Liew, and J. Ghazoul. 2011. “Remotely sensed evidence of tropical peatland conversion to oil palm.” *Proceedings of the National Academy of Sciences* 108 (12): 5127–5132. doi:10.1073/pnas.1018776108.

Lary, D. J., A. H. Alavi, A. H. Gandomi, and A. L. Walker. 2016. “Machine learning in geosciences and remote sensing.” *Geoscience Frontiers* 7 (1): 3-10. doi:10.1016/j.gsf.2015.07.003.

Lee, J. S. H., S. Wich, A. Widayati, and L. P. Koh. 2016. “Detecting industrial oil palm plantations on Landsat images with Google Earth Engine.” *Remote Sensing Applications: Society and Environment* 4: 219-224. doi:10.1016/j.rsase.2016.11.003.

Li, L., J. Dong, T. S. Njeudeng, and X. Xiao. 2015. “Mapping oil palm plantations in Cameroon Using PALSAR 50-m Orthorectified Mosaic Images.” *Remote Sensing* 7 (2): 1206-1224. doi:10.3390/rs70201206.

Marcene, B. 2018. “11 Amazing Health Benefits of Palm Oil.” Accessed June, 2018. https://www.naturalfoodseries.com/11-benefits-palm-oil/.

Miettinen, J., D. L. A. Gaveau, and S. C. Liew. 2018. “Comparison of visual and automated oil palm mapping in Borneo.” *International Journal of Remote Sensing*, 1-12. doi:10.1080/01431161.2018.1479799.

Miyamoto, M., M. M. Parid, Z. N. Aini, and T. Michinaka. 2014. “Proximate and underlying causes of forest cover change in Peninsular Malaysia.” *Forest Policy and Economics* 44: 18-25. doi:10.1016/j.forpol.2014.05.007.

Mohamad, M. 2018. “Speech text of Dr Mahathir at 73rd UN General Assembly.” *New Straits Times*, September 29. https://www.nst.com.my/news/nation/2018/09/415941/speech-text-dr-mahathir-73rd-un-general-assembly.

Morel, A. C., J. B. Fisher, and Y. Malhi. 2012. “Evaluating the potential to monitor aboveground biomass in forest and oil palm in Sabah, Malaysia, for 2000–2008 with Landsat ETM+ and ALOS-PALSAR.” *International Journal of Remote Sensing* 33 (11): 3614-3639. doi:10.1080/01431161.2011.631949.

Mouat, D. A., G. G. Mahin, and J. Lancaster. 2008. “Remote sensing techniques in the analysis of change detection.” *Geocarto International* 8 (2): 39-50. doi:10.1080/10106049309354407.

MPOB. 2017. Summary of Industry Performance: 2017. Accessed December, 2018. http://econ.mpob.gov.my/economy/EID_web.htm.

Mukherjee, S., and A. Mitra. 2009. “Health Effects of Palm Oil.” *Journal of Human Ecology* 26 (3): 197-203.

Murray, N. J., D. A. Keith, D. Simpson, J. H. Wilshire, and R. M. Lucas. 2018. “REMAP: An online remote sensing application for land cover classification and monitoring.” *Methods in Ecology and Evolution* 9 (9): 2019-2027. doi:10.1111/2041-210X.13043.

Ong, A. S., and S. H. Goh. 2002. “Palm oil: a healthful and cost-effective dietary component.” *Food and nutrition bulletin* 23 (1): 11-22. doi:10.1177/156482650202300102.

Rival, A. 2014. “The oil palm (Elaeis guineensis) : Research challenges beyond controversies." Accessed July, 2018. European Network of Palm Scientists: http://agritrop.cirad.fr/573384/1/document_573384.

Rodriguez-Galiano, V. F., B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez. 2012. “An assessment of the effectiveness of a random forest classifier for land-cover classification.” *ISPRS Journal of Photogrammetry and Remote Sensing* 67: 93-104. doi:10.1016/j.isprsjprs.2011.11.002.

RSPO, “Adoption of Principles and Criteria for the Production of Sustainable Palm Oil, 2013”, 2013. Accessed June, 2018. http://www.rspo.org/file/revisedPandC2013.pdf.

Shafri, H. Z. M., and N. Hamdan. 2009. "Hyperspectral Imagery for Mapping Disease Infection in Oil Palm Plantation Using Vegetation Indices and Red Edge Techniques." *American Journal of Applied Sciences* 6 (6): 1031-1035.

Shafri, H. Z. M., N. Hamdan, and M. I. Saripan. 2011. "Semi-automatic detection and counting of oil palm trees from high spatial resolution airborne imagery." *International Journal of Remote Sensing* 32 (8): 2095-2115. doi:10.1080/01431161003662928.

Shafri, H. Z. M., E. Taherzadeh, S. Mansor, and R. Ashurov. 2012. "Hyperspectral Remote Sensing of Urban Areas: An Overview of Techniques and Applications." *Research Journal of Applied Sciences, Engineering and Technology* 4 (11): 1557-1565. doi: 10.1111/gec3.12017.

Shaharum, N. S. N., H. Z. M. Shafri, J. Gambo, and F. A. Z. Abidin. 2018. "Mapping of Krau Wildlife Reserve (KWR) protected area using Landsat 8 and supervised classification algorithms." *Remote Sensing Applications: Society and Environment* 10: 24-35. doi:10.1016/j.rsase.2018.01.002.

Shih, I. 2018. "App helps ecologists map vulnerable ecosystems within minutes." Accessed June, 2018. https://newsroom.unsw.edu.au/news/science-tech/app-helps-ecologists-map-vulnerable-ecosystems-within-minutes.

Shuit, S. H., K. T. Tan, K. T. Lee, and A. H. Kamaruddin. 2009. "Oil palm biomass as a sustainable energy source: A Malaysian case study." *Energy* 34 (9): 1225-1235. doi:10.1016/j.energy.2009.05.008.

Singh, A., B. Ganapathysubramanian, A. K. Singh, and S. Sarkar. 2016. "Machine Learning for High-Throughput Stress Phenotyping in Plants." *Trends in Plant Science* 21 (2): 110-124. doi:10.1016/j.tplants.2015.10.015.

Sulaiman, F., N. Abdullah, H. Gerhauser, and A. Shariff. 2011. "An outlook of Malaysian energy, oil palm industry and its utilization of wastes as useful resources." *Biomass and Bioenergy* 35 (9): 3775-3786. doi:10.1016/j.biombioe.2011.06.018.

Sumathi, S., S. P. Chai, and A. R. Mohamed. 2008. “Utilization of oil palm as a source of renewable energy in Malaysia.” *Renewable and Sustainable Energy Reviews* 12 (9): 2404-2421. doi:10.1016/j.rser.2007.06.006.

Svetnik, V., A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan, and B. P. Feuston. 2003. “Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modelling.” *Journal of chemical information and computer sciences* 43 (6): 1947-1958. doi:10.1021/ci034160g.

Tatsumi, K., Y. Yamashiki, M. A. C. Torres, and C. L. R. Taipe. 2015. “Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data.” *Computers and Electronics in Agriculture* 115: 171-179. doi:10.1016/j.compag.2015.05.001.

Thenkabail, P. S., N. Stucky, B. W. Griscom, M. S. Ashton, J. Diels, B. Van Der Meer, and E. Enclona. 2004. “Biomass estimations and carbon stock calculations in the oil palm plantations of African derived savannas using IKONOS data.” *International Journal of Remote Sensing* 25 (23): 5447-5472. doi:10.1080/01431160412331291279.

Tillack, G. 2013. “Why 'Roundtable on Sustainable Palm Oil (RSPO)' palm oil is neither responsible nor sustainable.” Accessed July, 2018. https://www.ran.org/why_rspo_sustainable_palm_oil_is_not_responsible.

Vijay, V., S. L. Pimm, C. N. Jenkins, and S. J. Smith. 2016. “The Impacts of Oil Palm on Recent Deforestation and Biodiversity Loss.” *Public Library of Science One* 11 (7): 1-13. doi:10.1371/journal.pone.0159668.

Wahid, B. O., A. A. Nordiana, and A. M. Tarmizi. 2005. “Satellite Mapping of Oil Palm Land Use.” *Malaysian Palm Oil Board Information Series* 255: 1511-7871.

Wang, P., J. Wang, Y. Chen, and G. Ni. 2013. “Rapid processing of remote sensing images based on cloud computing.” *Future Generation Computer Systems* 29 (8): 1963-1968. doi:10.1016/j.future.2013.05.002.

Wicke, B., R. Sikkema, V. Dornburg, and A. Faaij. 2011. “Exploring land use changes and the role of palm oil production in Indonesia and Malaysia.” *Land Use Policy* 28 (1): 193-206. doi:10.1016/j.landusepol.2010.06.001.

Wu, C., B. Du, X. Cui, and L. Zhang. 2017. “A post-classification change detection method based on iterative slow feature analysis and Bayesian soft fusion.” *Remote Sensing of Environment* 199:  241-255. doi:10.1016/j.rse.2017.07.009.

Wulder, M. A., N. C. Coops, D. P. Roy, J. C. White, and T. Hermosilla. 2018. “Land cover 2.0.” *International Journal of Remote Sensing* 39 (12): 4254-4284. doi:10.1080/01431161.2018.1452075.

Zhu, Z., C. E. Woodcock, and P. Olofsson. 2012. “Continuous monitoring of forest disturbance using all available Landsat imagery.” *Remote Sensing of Environment* 122: 75-91. doi:10.1016/j.rse.2011.10.030.

Zhu, Z., and C. E. Woodcock. 2014. “Continuous change detection and classification of land cover using all available Landsat data.” *Remote sensing of Environment* 144: 152-171. doi:10.1016/j.rse.2014.01.011.

Zhu, Z. 2017. “Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications.” *ISPRS Journal of Photogrammetry and Remote Sensing* 130: 370-384. doi:10.1016/j.isprsjprs.2017.06.013.

Table 1. Available predictor layers.

Table 2. Confusion matrix table.

Table 3. Confusion matrix using three classes.

Table 4. Calculations of accuracies.

Table 5. Accuracies of each state in Peninsular Malaysia for Periods 1 and 2.

Table 6. Oil palm area of Peninsular Malaysia.

Figure 1. The study area: Peninsular Malaysia.

Figure 2. General workflow of this study.

Figure 3. Structure of the RF algorithm.

Figure 4. Uses of cloud computing.

Figure 5. Workflow using REMAP.

Figure 6. (a) Landsat image (b) High-resolution image.

Figure 7. Classified map of Peninsular Malaysia (a) Land cover period 1 (b) Land cover period 2.

Figure 8. Conversion rate of each class converted to oil palm from Periods 1 to 2.