

PAPER • OPEN ACCESS

The utilisation of cloud computing and remote sensing approach to assess environmental sustainability in Malaysia

To cite this article: N S N Shaharum *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **230** 012109

View the [article online](#) for updates and enhancements.

You may also like

- [Land use change impacts on European heat and drought: remote land-atmosphere feedbacks mitigated locally by shallow groundwater](#)
Samuel C Zipper, Jessica Keune and Stefan J Kollet
- [Observed rainfall changes in the past century \(1901–2019\) over the wettest place on Earth](#)
J Kuttippurath, S Murasingh, P A Stott et al.
- [Projecting the land cover change and its environmental impacts in the Cedar River Basin in the Midwestern United States](#)
Yiping Wu, Shuguang Liu, Terry L Sohl et al.



Free the Science Week 2023 April 2–9

Accelerating discovery through
open access!

 www.ecsdl.org [Discover more!](#)

The banner features a dark blue background with a futuristic, glowing blue interface. A hand is shown interacting with a circular element containing a padlock icon, symbolizing open access. The text is in white and light blue, with the ECS logo and website URL in white.

The utilisation of cloud computing and remote sensing approach to assess environmental sustainability in Malaysia

N S N Shaharum¹, H Z M Shafri^{1,2}, W A W A K Ghani³, S Samsatli⁴ and B Yusuf¹

¹ Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia, Selangor, Malaysia.

² Geospatial Information Science Research Centre (GISRC), Faculty of Engineering, Universiti Putra Malaysia Selangor, Malaysia

³ Department of Chemical and Environmental Engineering/Sustainable Process Engineering Research Centre (SPERC), Faculty of Engineering, Universiti Putra Malaysia, Selangor, Malaysia

⁴ Department of Chemical Engineering, University of Bath, Claverton Down, United Kingdom

E-mail: fieranisa94@gmail.com

Abstract. Monitoring of the environment over a large area will require huge amount of data. The implementation of conventional methods will be time consuming and very costly. Furthermore, one way to assess the environmental sustainability is by analysing the changes of land cover over two different periods. Therefore, this study implemented a cloud computing approach utilising an open source Remote Ecosystem Monitoring Assessment Pipeline (REMAP) to map the changes of Land use land cover (LULC) over Peninsular Malaysia utilising Landsat data obtained from two different periods (2003 and 2017). This approach has utilised a powerful inbuilt machine learning algorithm, Random Forest (RF) to test the performance of cloud computing using REMAP to produce LULC maps over Peninsular Malaysia. The results showed an acceptable LULC maps and the changes between two periods were analysed. Therefore, the utilisation of machine learning algorithm with the integration of cloud computing using REMAP can reduce the cost, lessen the processing time, produce LULC maps, and perform change analysis over large area.

1. Introduction

Sustainability is the development or a concept that is carried out to meet the needs of present without endangering the future generation in many aspects including economic health, environment and ecological. It usually uses alternative approaches to replace current method or material in order to reduce the negative impacts especially towards human, animals and environment. To support this idea, United Nations (UN) has introduced the 17 Sustainable Development Goals (SDGs) where each goal has its target to be achieved by 2030 [1]. The goals include to protect the planet, counter climate change and its impacts, produce clean energy and ensuring health lives. To contribute to the 17 SDGs,



this paper will be focusing on the environmental sustainability where two sets of remote sensing data taken from different periods will be used to produce the Land Use Land Cover (LULC) maps which later can be used to assess the environmental sustainability.

Machine learning is a subset of Artificial Intelligence (AI) that works well in prediction and classification. It is an application or algorithm that learns and improves from experience without being explicitly programmed [2-4]. Previous studies had used supervised machine learning algorithms to perform image classification [5,6]. For example, Support Vector Machine (SVM) is able to handle the complexity in the dimensionality and it classifies the data by separating the support vectors using a hyperplane. Random Forest (RF) is one of the powerful machine learning algorithms that uses thresholds to distinguish the features based on the defined parameters. It combines several numbers of trees and finally forms a forest-structure [7]. Machine learning gives good results for mapping LULC and it has been used in many applications including urban and agriculture. However, the study areas involved were small and do not cover large areas due to the higher cost and longer time to process the area [8].

Cloud computing is the delivery of computing services such as storage, servers and databases. It can store huge amount of data and process big data in the cloud [9]. By using conventional method, few hard drives with large storage will be essential to store the data and on top of that, having high computational power is crucial to process big data. However, to fulfil these requirements will be very costly and time consuming. Therefore, to overcome these issues, this study had utilised a cloud-based platform with the integration of machine learning and big data to perform change analysis in Peninsular Malaysia.

2. Material and Methods

2.1. Study area and satellite data

For this study, Peninsular Malaysia that is located at 3.1351° N, 101.7400° E with an approximate of 132 156 km² land cover area was chosen as the study area. Using high resolution data to cover huge area will be expensive and therefore, open source Landsat data with 30m spatial resolution was used in this study.

2.2. Methods

This study utilised cloud computing approach where a cloud-based platform namely Remote Ecosystem Monitoring Assessment Pipeline (REMAP) was used to map the LULC of Peninsular Malaysia. Stacked Landsat data obtained from 1999-2003 (noted as 2003) and 2014-2017 (noted as 2017) were analysed and classified in REMAP using an inbuilt machine learning algorithm. The best combination of predictors was used to classify the data and the best outputs were exported to produce the LULC maps. To assess the changes of the land covers in Peninsular Malaysia, few classes were classified: paddy, other agriculture, forest, water bodies and built-up. The procedure of the work was shown in Figure 1.

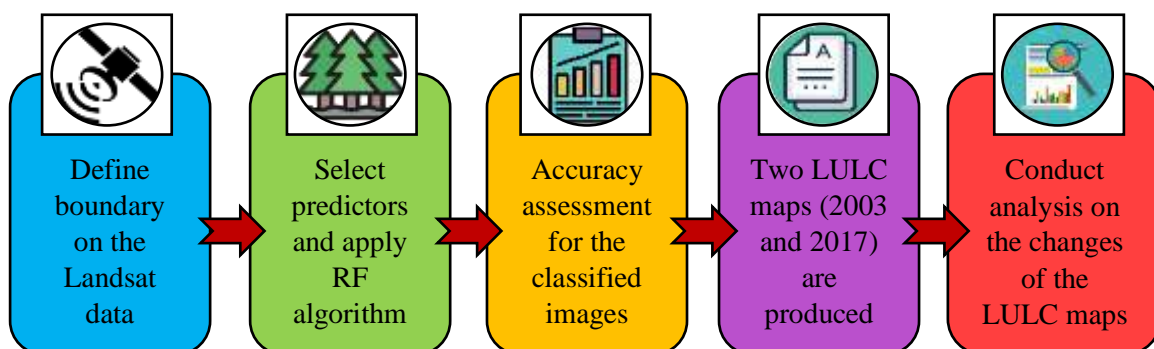


Figure 1. Work flow of the study.

2.3. Image classification

Both 2003 and 2017 data were classified using RF algorithm with the aid of 13 available predictors in REMAP [10]. The training samples were produced in REMAP and the produced samples can be exported as keyhole mark-up language (KML) and comma-separated values (CSV) file formats. Then, the accuracy of the classified images was assessed, and the best outputs were exported as GeoTiff which later be imported to other software for further analysis and map making. Changes occurred on the area of the LULC maps produced in 2003 and 2017 were calculated where the differences will then be tabulated in the next sub-section.

2.4 Change detection analysis

Change detection analysis is conducted to assess and measure the changes of the land cover between two or more periods. Change detection is a post-classification technique that is conducted after that images have been classified and it usually involves comparing two or more satellite imagery of the same area taken at different periods. Previous studies have conducted change detection analysis on several applications including biodiversity, forest, and urban [11-13]. Change detection methods can be grouped into several categories which include classification, algebra, visual analysis, and transformation [14]. This study conducted a post-classification comparison technique where the classified thematic maps were assessed by comparing the pixel from different thematic map. To assess the changes of the land cover, the total number of pixels for each class obtained from classified LULC map in 2003 was compared with the total number of pixels obtained from classified LULC map in 2017.

3. Results and Discussion

Two Land use land cover (LULC) maps of Peninsular Malaysia (2003 and 2017) and the changes of the land covers between 2003 to 2017 can be seen in Figure 2 and Figure 3. Therefore, the changes for each class can be analysed and the results were tabulated in Table 1.

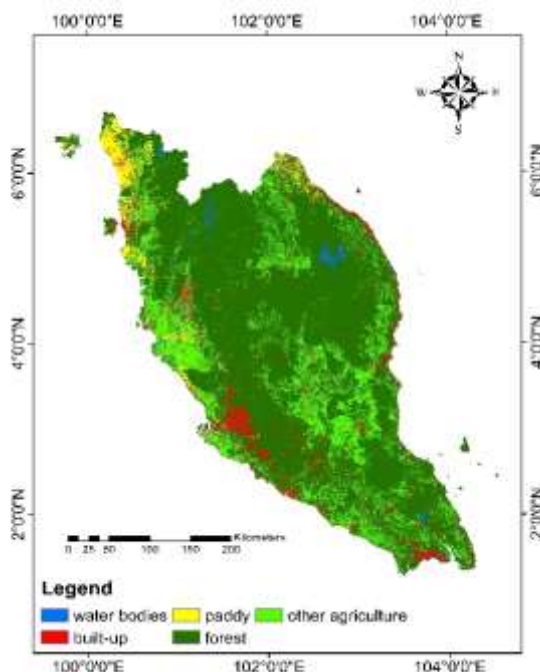


Figure 2. 2003 LULC of Peninsular Malaysia

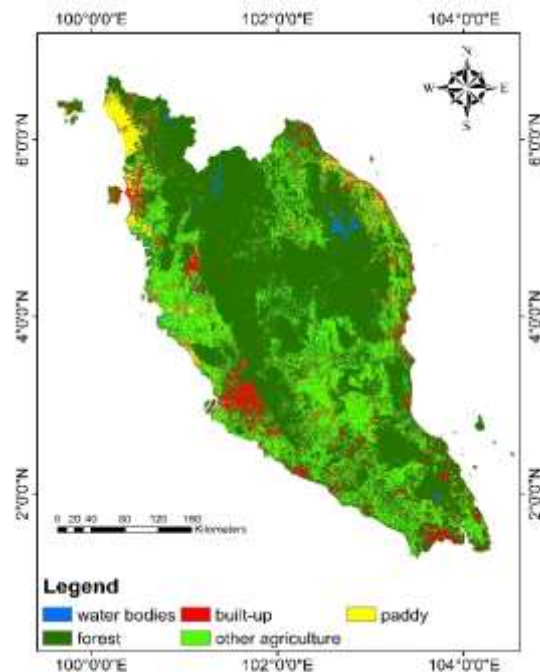


Figure 3. 2017 LULC of Peninsular Malaysia

Table 1. Changes of land cover between 2003 and 2017.

Class	Area (%)
Water bodies	-11.86
Paddy	-5.71
Other agriculture	+59.40
Built-up	+38.91
Forest	-9.70

Utilising optical Landsat data by means of conventional method to produce results shown in Figure 2 and Figure 3 will require massive efforts. Table 1 shows the changes of land cover area for each class occurred from 2003 to 2017. The negative and positive sign indicates the loss and the gain of the land cover area respectively throughout 14 years of period. Traditionally, minimum number of 15 scenes obtained from Landsat data are needed to cover the whole area of Peninsular Malaysia which will require the user to manually download the satellite images, implementing image patching (remove clouds) and process the data from scratch. These steps will require long time, and high computational power is needed to be able to finish the image processing especially on the process of image patching and image classification. Furthermore, to store and compile the processed satellite data, it is vital to have huge storage of hard drives before any LULC map can be produced. Therefore, the utilisation of REMAP has been very handy and has successfully tackled the aforementioned issues. On top of that, the results produced on the changes of the LULC can later be improved and the statistical values can later be used as an aid to assess the sustainability for Peninsular Malaysia.

REMAP uses cloud computing approach which stores the data in the cloud by using cloud-based platform and it makes LULC classification over large area possible. Furthermore, the advantage of using REMAP can be seen clearly when the application of image classification was performed in the cloud using an inbuilt machine learning RF algorithm. On the other hand, utilising conventional method will require additional effort in order the user to set the threshold and parameters in RF algorithm. Although optical sensor like Landsat often having problem with the cloud cover, REMAP provides Landsat data that are nearly to absolute cloud-free. Furthermore, REMAP provides data that covers the whole world which makes it very useful especially for tropical countries like Malaysia.

4. Conclusion

Using conventional method to map large area will be very costly and time consuming. Furthermore, the cost to obtain high-resolution data such as LiDAR (Light Detection and Ranging) data will be expensive even for small coverage area. In addition, using open source Landsat data will give cloudy data for tropical countries. Furthermore, to produce almost cloud-free images, additional time and effort are required to remove the clouds. Therefore, the utilisation of cloud computing using REMAP is very effective because it provides efficient way to map large area using medium-resolution Landsat data with least amount of cloud cover. The powerful cloud-based REMAP has provided a fast-paced image classification and with the integration with a powerful RF machine learning algorithm, REMAP has proven to be very efficient. As a conclusion, the integration of cloud computing and machine learning has managed to produce the LULC maps and the changes of the land cover can be analysed for large area covering the Peninsular Malaysia.

Acknowledgments

The authors wish to acknowledge Universitas Brawijaya for the accommodation, daily transportation, meals, excursion and UC SEARCA for the expenses on the airfares. The authors also would like to thank UNSW Centre of Ecosystem Science for assisting on the technical support and UPM for their facilities and funding of this research. Apart from that, our appreciation also goes to Engineering and

Physical Sciences Research Council for their financial support through the BEFEW project (Grant No. EP/P018165/1).

References

- [1] Griggs D, Stafford-Smith M, Gaffney O, Rockström, J, Öhman M C, Shyamsundar P, Steffen W, Glaser G, Kanie N, Noble I 2013 Policy: sustainable development goals for people and planet *Int. J. Sci. Nat.* **495** 305-307
- [2] Ali I, Greifeneder F, Stamenkovic J, Neumann M, Notarnicola C 2015 Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data *Remote Sens.* **7** 12 16398-16421.
- [3] Han J, Zhang D, Cheng G, Guo L, Ren J 2015 Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning *IEEE Trans. on Geosci. Remote Sens.* **53** 6 3325-3337.
- [4] Shafri H Z M 2017 Machine learning in hyperspectral and multispectral remote sensing data analysis *Artificial Intel. Sci. Technol.* 3-9.
- [5] Nooni I K, Duker A A, Van Duren I, Addae-Wireko L, Osei Jnr, E M 2014 Support vector machine to map oil palm in a heterogeneous environment *Int. J. of Remote Sens.* **35** 4778-4794.
- [6] Dibs H, Idrees M O, Alsalhin G B A 2017 Hierarchical classification approach for mapping rubber tree growth using per-pixel and object-oriented classifiers with SPOT-5 imagery *The Egypt. J. Remote Sens. Space Sci.* **20** 21-30.
- [7] Millard K, Richardson M 2015 On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping *Remote Sens.* **7** 8489-8515.
- [8] Belgiu M, Drăguț L 2016 Random forest in remote sensing: A review of applications and future directions *ISPRS J. Photogram Remote Sens.* **114** 24-31.
- [9] Hashem I A T, Yaqoob I, Anuar N B, Mokhtar S, Gani A, Khan S U 2015 The rise of “big data” on cloud computing: Review and open research issues *Inf. Syst.* **47** 98-115.
- [10] Murray N J, Keith D A, Simpson D, Wilshire J H, Lucas R M 2017 REMAP: An online remote sensing application for land cover classification and monitoring *BioRxiv* 1-33.
- [11] Gambo J, Shafri H Z M, Shaharum N S N, Abidin F A Z, Rahman M T A 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 *J. Geom. Plan.* **5** 17-34.
- [12] Dadras M, Shafri H Z M, Ahmad N, Pradhan B, Safarpour S 2014 Land use/cover change detection urban sprawl analysis in Bandar Abbas City, Iran *Sci. World J.* **2014** 1-12.
- [13] Leimgruber P, Kelly D S, Steininger M K, Brunner J 2005 Forest cover change patterns in Myanmar (Burma) 1990-2000 *Environ. Conserv.* **32** 356-364.
- [14] Lu D, Mausel P, Brondizio E, Moran E 2014 Change detection techniques *Int. J. Remote Sens.* **25** 2365-2401.