

PAPER • OPEN ACCESS

## An integrated framework for assessing land-use/land-cover of Kelantan, Malaysia: Supervised and unsupervised classifications

To cite this article: L B Yeo *et al* 2022 *IOP Conf. Ser.: Earth Environ. Sci.* **1053** 012026

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
- [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.*
- [Biophysical and socioeconomic drivers of oil palm expansion in Indonesia](#)  
Yu Xin, Laixiang Sun and Matthew C Hansen



## ECS Membership = Connection

**ECS membership connects you to the electrochemical community:**

- Facilitate your research and discovery through ECS meetings which convene scientists from around the world;
- Access professional support through your lifetime career;
- Open up mentorship opportunities across the stages of your career;
- Build relationships that nurture partnership, teamwork—and success!

**Join ECS!**

**Visit [electrochem.org/join](https://electrochem.org/join)**



# An integrated framework for assessing land-use/land-cover of Kelantan, Malaysia: Supervised and unsupervised classifications

L B Yeo<sup>1\*</sup>, W S N Wan Mohamad<sup>1</sup>, R Hasan<sup>1</sup>, N I Othmani<sup>1</sup>, N H Abdul Hamid<sup>1</sup>, N Ramlee<sup>1</sup>, O T S Yeo<sup>2</sup>

<sup>1</sup> Faculty of Architecture and Ekistics, Universiti Malaysia Kelantan, Bachok 16300, Malaysia

<sup>2</sup> Department of Crop Science, Faculty of Agriculture, Universiti Putra Malaysia, 97008 Bintulu, Sarawak, Malaysia

\* Corresponding author: yeoleebak@umk.edu.my

**Abstract.** The change of land-use/land-cover (LULC) is inexorable to cater to the needs of the people, from a natural land use dominance to the built-up areas' dominance. Many geospatial studies have examined the LULC, yet the data used were not up-to-date. Hence, this paper proposes an integrated framework to assess the LULC by using the latest complimentary data of Landsat 8 from Earth Explorer-USGS, and subsequently processed in Geographic Information System application (ArcGIS 10.8) by executing two classification techniques: i) interactive supervised classification (ISC) and ii) Iso cluster unsupervised classification (ICUC). The study area is Kelantan, a state located in Peninsular Malaysia. Three hundred sampling points were randomly generated within the boundary of Kelantan to compare the classified map of 2021 with ground data using Google Earth Pro 7.3 and Google Street View. Results show that the ISC method presents an overall accuracy of 87.7% ( $K=0.79$ ,  $p \leq 0.001$ ) compared to the ICUC method, with an accuracy rate of 67.7% ( $K=0.47$ ,  $p \leq 0.001$ ). ISC method with the proposed integrated framework can be viable techniques to assess LULC and can be used to study the change of LULC periodically. Thus, it helps stakeholders make better decisions in resource planning and distribution, especially in Kelantan, where the natural resources are rich.

**Keywords:** LULC; image classifications; GIS; Kelantan.

## 1. Introduction

The increasing population has contributed to the demand for more living space where forested areas were tardily converted to agricultural lands and built-up areas, particularly housing, commercial lots, and industrial zones [1]. The growth rate of Melaka, Pulau Pinang, Selangor and Wilayah Persekutuan have exceeded the national urban population rate of 70%. Meanwhile, Kedah, Melaka, and Perak increased significantly from around 20% to 70%, whereas Kelantan increased gradually from approximately 15% to 40% [2]. It means Kelantan may still retain substantial natural resources compared to other urbanized states in Peninsular Malaysia.

The studies of LULC allow researchers to understand, analyze and explore geospatial information of a town or city. One of them is related to Urban Heat Island (UHI) and human thermal comfort. For instance, Yeo *et al.* [3] examined various spatial settings of LULC. They suggested that people living in urban areas tend to experience heat stress due to asphalt streets, car parks, concrete areas, and vacant land. Meanwhile, suburban and rural areas with tree-covered areas are less likely to experience heat stress due to shades and evaporative cooling effects. Assessing LULC allows researchers to monitor the change of land use and predict the dynamic landscape pattern [1,4]. Other than that, the studies on ecosystem services provision, biodiversity, habitat conservation, and environmental sustainability



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

utilize the LULC data set as well [5]. Despite the fact that LULC studies are widely explored in the field of city planning, one common problem facing is that outmoded data set from open public sources were often utilized [4]. The outdated geospatial mapping can hardly be used in the decision-making process, especially concerning sustainable development and planning of a city or town, unless the nature of the study itself is to scrutinize urban morphology that focuses on dynamic landscape simulation. Thus, the objectives of this paper are: i) to examine LULC classification with post-classification techniques using Landsat 8 imagery via GIS, and ii) to compare the accuracy of both ISC and ICUC classification approaches.

## 2. Methods

### 2.1 Study area

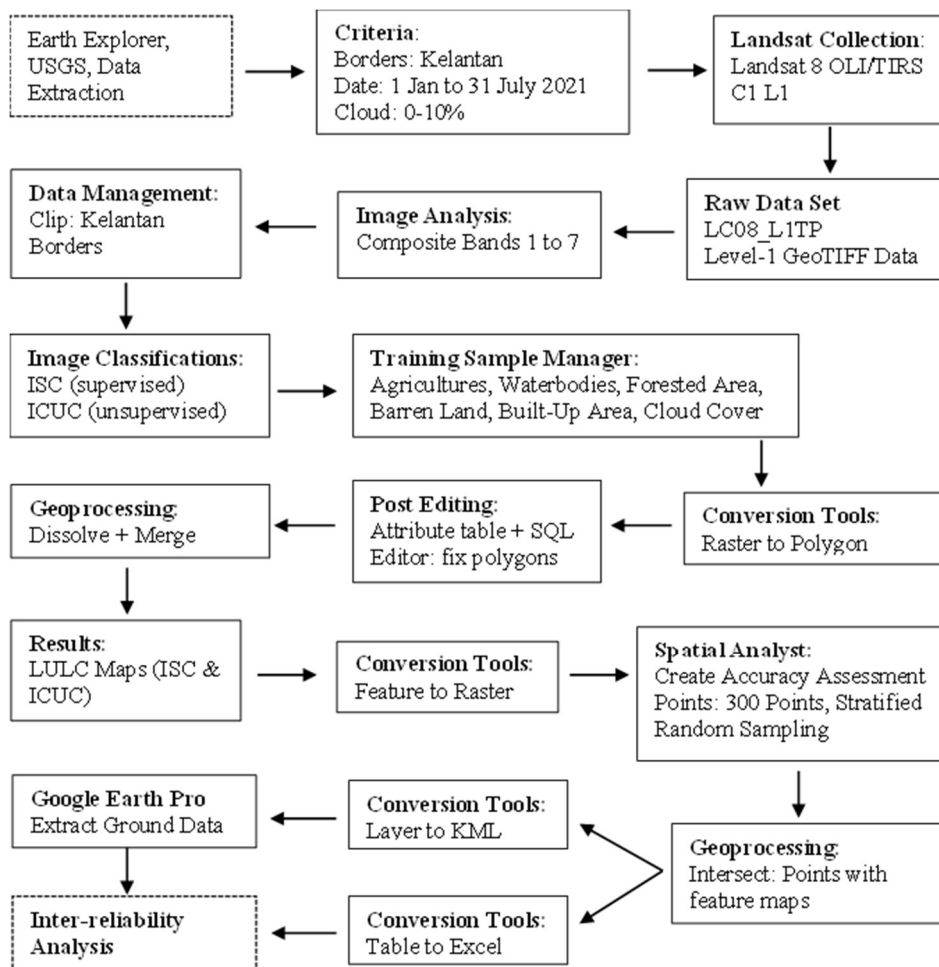
Kelantan is in the northeast part of peninsular Malaysia under the Universal Transverse Mercator (UTM) coordinate system of Zone 47N to 48N, with approximately 15,040 km<sup>2</sup>. The northern part of Kelantan faces Thailand's border, with Golok River defining the border. The eastern and southern sides face Terengganu and Pahang, respectively, with hilly terrains and forestry. Lastly, the western side faces Perak.

### 2.2 Data processing and analysis

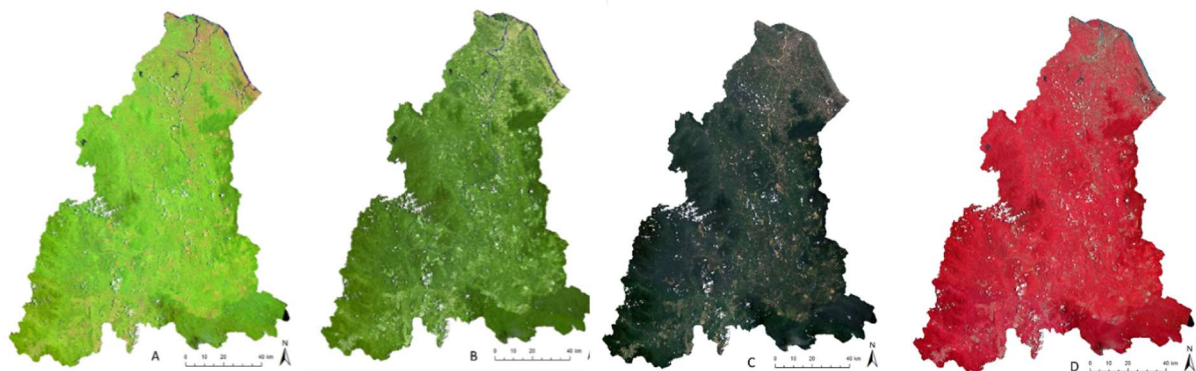
There are various models or frameworks to assess LULC. For instance, the manual processing method uses geo-processing and editor tools in ArcGIS to update the data by overlaying the latest remote sensing data [1]. However, this method is only suitable for small-scale editing based on the data provided by local authorities. Another method is through supervised or unsupervised classification methods, suitable for macro scale processing by letting the software identify the statistical patterns or using predefined decision rules to train the sample pixels without ground data [6]. Little to no literature explicitly explains how the open-source data can produce the LULC maps. Thus, this study aims to propose an integrated framework to produce high-accuracy LULC maps (refer to Figure 1).

Firstly, we downloaded the Landsat 8 data from Earth Explorer, USGS [7], and set the predefined criteria like the site's boundary, date of the map, and cloud cover percentage. Two Landsat 8 images were downloaded, one covered the northern part, and another covered the southern part of Kelantan, Malaysia. The projected coordinate system is WGS 1984, UTM Zone: 47N. There are a total of 11 bands images for each set of data. We only used bands 1 to 7 because bands 8 to 11 were irrelevant and needed. The composited images are as shown in Figure 2, where the combinations of the bands are: Map A (6,5,2), Map B (7,6,4), Map C (4,3,2), and Map D (5,4,3) based on RGB colour model.

It is important to note that the composite of bands may influence the identification of different LULC. For instance, Map D's infrared colour is suitable for identifying vegetation, where darkest red indicates forested area and the lightest red indicates agricultural land. Similarly, Map A can also identify agricultural land and forested areas where lighter green indicates agricultural land and darker green indicates forested areas. Nevertheless, both Map A and Map D are rather hard to distinguish the barren land and built-up areas because both appear almost the same colours. Thus, we suggest using Map B to differentiate between barren land and build-up area. Indeed, we can always refer to Map C, the natural colour. Despite the different band combinations helping us identify the type of LULC, there are always limitations. For instance, it is hard to differentiate agricultural land and forested areas, especially when the agricultural lands are fruit tree plantations, rubber estates, or oil palm estates. The ISC or ICUC classification might miscode between the two.



**Figure 1.** Integrated framework for LULC assessment.

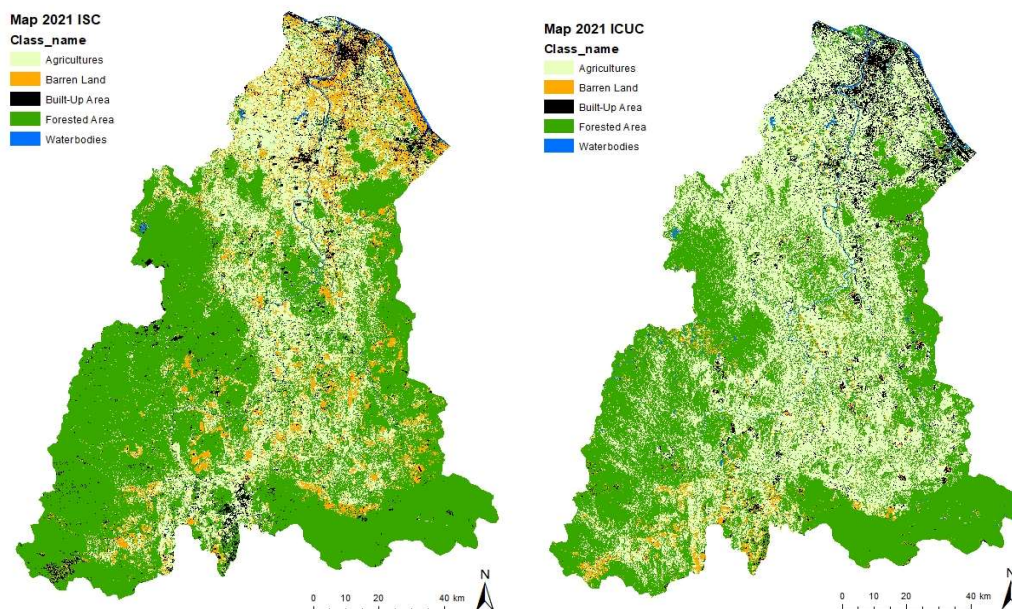


**Figure 2.** Different bands combination for image classification.

For the ISC method, we trained the sample on the map manually by selecting the pixels on the map. We trained the samples according to agricultures, waterbodies, forested area, barren land, built-up area,

and cloud cover. Cloud cover was trained separately to simplify the post-editing process because it tends to be miscoded as built-up areas. After classifying the image, both ISC and ICUC maps were converted to polygon to fix the cloud cover with the editor tool. It is almost impossible to fix every detail in this stage. Hence, we only fixed all the cloud cover in the attribute table and its shadows. After editing the ISC and ICUC maps, the attributes table in the map was *dissolved* and *merged* using geo-processing tools to combine similar types of LULC and both northern and southern parts. It is important to note that it should not be merged at the beginning of the stage because the colour of the pixels is non-identical. Otherwise, it may miscode during the classification process.

The ISC and ICUC generated LULC maps (see Figure 3) were converted to raster using the conversion tool - *feature to raster*. It is not mandatory to convert to raster, but we suggested doing so. Otherwise, computing error might occur due to a large sampling size. Next, executing the spatial analyst tool - *create accuracy assessment* to generate 300 stratified randomized sampling points within the maps across different types of LULC. Each sampling point was *intercepted* with the feature map to obtain the attribute information.



**Figure 3.** LULC maps based on supervised and unsupervised classifications. Map A is by ISC method, while Map B is by ICUC method.

Lastly, the sampling point layer was exported to an excel file using a conversion tool and converted to a *kml* file to open with Google Earth Pro. Next, ground data were collected from the *kml* file open in Google Earth Pro by examining the exact location of the points in the satellite image and Google Street View. Using these methods can help reduce the petrol consumption and time to travel to each destination where the sampling points are. Nevertheless, if the information on the sampling point data is unavailable, an actual ground visit should be carried out.

### 3. Results and discussion

Tables 1 and 2 show the comparison of ground data with the projected result. The map produced using the ISC method has an 87.7% accuracy rate, with  $K = 0.79$  and  $p < 0.001$  (Table 1). Based on the Cohen's Kappa coefficient ( $K$ ),  $K = 0.79$  represents a substantial strength of agreement, whereas the  $p < 0.001$  indicates the result is statistically significant. Meanwhile, the map produced by the ICUC method has

an accuracy of 67.7 %, where  $K = 0.47$  and  $p < 0.001$ . It means the strength of agreement is moderate. Based on the previous literature [6], the supervised classification method is expected to provide a more accurate map. However, the method they applied was maximum likelihood classifier, while we were applying interactive supervised classification (ISC), which does not require a training signature file. The sample points we trained from the map are more than 50 per group, and we found that the signature file could not be generated.

**Table 1.** Accuracy assessment for interactive supervised classification (ISC).

		Ground data					Total	U. Acc., %
		AG	BL	BA	FA	WB		
ISC	AG	57	0	0	21	0	78	73.1
	BL	3	25	2	0	0	30	83.3
	BA	1	2	6	3	0	12	50.0
	FA	5	0	0	165	0	170	97.1
	WB	0	0	0	0	10	10	100.0
Total		66	27	8	189	10	300	
P. Acc, %		86.4	92.6	75.0	87.3	100		

a. Overall Accuracy:  $263/300 * 100 = 87.7\%$

b. Measurement of Agreement,  $K = 0.79$

c.  $p < 0.001$

**Table 2.** Accuracy assessment for Iso cluster unsupervised classification (ICUC).

		Ground data					Total	U. Acc, %
		AG	BL	BA	FA	WB		
ICUC	AG	54	16	0	53	0	123	43.9
	BL	0	3	0	5	1	9	33.3
	BA	1	8	7	0	0	16	43.8
	FA	11	0	1	130	0	142	91.6
	WB	0	0	0	1	9	10	90.0
Total		66	27	8	189	10	300	
P. Acc., %		81.8	11.1	87.5	68.8	90.0		

a. Overall Accuracy:  $203/300 * 100 = 67.7\%$

b. Measurement of Agreement,  $K = 0.47$

c.  $p < 0.001$

To identify the shortcoming of the proposed integrated framework, we also compared the projected LULC with ground data on each category (i.e., user's accuracy). For instance, waterbodies (WB) show 100% accuracy for the ISC method, while ICUC only shows a 90% accuracy rate. Meanwhile, the built-up area (BA) has the lowest accuracy rate for the ISC method. Initially, this was somewhat unexpected. We thought the forested area should exhibit the highest error due to difficulty differentiating the agriculture (AG) and forested area (FA) while training the samples. After re-assessment, we believe this occurred because the top layer of the built-up areas retained too many pixel values, indirectly causing the problem for the training sample manager to automate the classification algorithm to associate with the pixel values. For ICUC, the lowest accuracy rate goes to barren land (BL) with only 33.3%. It is rather critical because most of the pixels were miscoded as FA. It is the same for AG and BA (see Table

2). In sum, the ICUC method does not produce encouraging results unless post-editing is executed after running the inter-reliability test. Similarly, extra precautions should focus on BA and AG for the ISC method to further improve the accuracy of the LULC map.

#### 4. Conclusion remarks

The proposed integrated framework could effectively produce LULC maps with a higher accuracy rate. The accuracy rate can be improved when thoroughly examined in the post-classification stage. In our case, we only fixed the cloud cover and its shadows. Based on the findings, we found that built-up areas (user's accuracy of 50%) and agricultural lands (user's accuracy of 73.1%) for the ISC method have the lowest user's accuracy rate. Nevertheless, the overall accuracy rate is still more than the usual standard of 85% precision; hence the result is reliable. If both LULC were rectified during the post-editing process using a bottom-up approach, the accuracy might see an overall improvement. We would also like to highlight that the shortcoming of using computational algorithm-based classification is that each LULC's sub-attributes are impossible to identify at this stage. Therefore, future research can improve in this aspect. Practically, this study may benefit stakeholders, local authorities, or practitioners to adapt the framework in their respective fields. An up-to-date map can support sustainable planning of a town or city, including but not limited to monitoring the change of LULC, identifying potential eco-tourism spots, assessing and valuing ecosystem services providers, and perhaps preventing illegal logging. In sum, this study would bring a new contribution to the practical production of a LULC map without depending on the outdated data sources acquired from the open public sources data or local authorities.

#### 5. References

- [1] Yeo L B, Said I, Saito K and Mohd Fauzi A 2017 Mapping land use/ cover changes and urbanization at sub-districts of Muar, Malaysia. *Chemical Engineering Transactions* **56** 289–294
- [2] DOSM 2020 Evolution of Migration for Urban and Rural. ([https://www.dosm.gov.my/v1/uploads/files/6\\_Newsletter/Newsletter%202020/DOSM\\_BPPD\\_4-2020\\_Series-62.pdf](https://www.dosm.gov.my/v1/uploads/files/6_Newsletter/Newsletter%202020/DOSM_BPPD_4-2020_Series-62.pdf): Department of Statistics Malaysia)
- [3] Yeo L B, Ling G H T, Tan M L and Leng P C 2021 Interrelationships between Land Use Land Cover (LULC) and Human Thermal Comfort (HTC): A Comparative Analysis of Different Spatial Setting. *Sustainability* **13** 382
- [4] Mahamud M A, Samat N, Tan M L, Chan N W and Tew Y L 2019 Prediction of future land use land cover changes of Kelantan, Malaysia. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XLII-4/W16** 379-384
- [5] Yeo L B and Said I 2018 Mapping Recreational Ecosystem Service at Sub-Districts of Muar. In: *Asia International Multidisciplinary Conference (AIMC 2017): The European Proceedings of Social & Behavioural Sciences* pp 843-857
- [6] Thapa R B and Murayama Y 2009 Urban mapping, accuracy, & image classification: A comparison of multiple approaches in Tsukuba City, Japan. *Applied Geography* **29** 135-144
- [7] USGS 2021 Earth Explorer. (<https://earthexplorer.usgs.gov/>: USGS)

#### Acknowledgments

We would like to thank the grant provider- Universiti Kelantan Malaysia under scheme UMK- FUND (R/FUND/A1200/01903A/001/2022/01016) for the financial support to this research project.