RESEARCH

Multi-temporal analysis of past and future land cover change in the highly urbanized state of Selangor, Malaysia

Majid Azari1*, Lawal Billa2 and Andy Chan1

Abstract

Background: This study analysed the multi-temporal trend in land cover, and modelled a future scenario of land cover for the year 2030 in the highly urbanized state of Selangor, Malaysia. The study used a Decision Forest-Markov chain model in the land change modeller (LCM) tool of TerrSet software. Land cover maps of 1999, 2006 and 2017 were classified into 5 classes, namely water, natural vegetation, agriculture, built-up land and cleared land. A simulated land cover map of 2017 was validated against the actual land cover map 2017. The Area Under the Curve (AUC) value of 0.84 of Total Operating Characteristics (TOC) and higher percentage of components of agreement (Hits + Correct rejection) compared to components of disagreement (Misses + False alarm + Wrong hits) indicated successful validation of the model.

Results: The results showed between the years 1999 to 2017 there was an increase in built-up land cover of 608.8 km² (7.5%), and agricultural land 285.5 km² (3.5%), whereas natural vegetation decreased by 831.8 km² (10.2%). The simulated land cover map of 2030 showed a continuation of this trend, where built-up area is estimated to increase by 723 km² (8.9%), and agricultural land is estimated to increase by 57.2 km² (0.7%), leading to a decrease of natural vegetation by 663.9 km² (8.1%) for the period 2017 to 2030. The spatial trend of land cover change shows built-up areas mostly located in central Selangor where the highly urbanized and populated cities of Kuala Lumpur and Putrajaya and the Klang valley are located.

Conclusion: The future land cover modelling indicates that built-up expansion mostly takes place at edges of existing urban boundaries. The results of this study can be used by policy makers, urban planners and other stakeholders for future decision making and city planning.

Keywords: Land cover modelling, Land use and land cover, Land change modeller, Markov chain model, Selangor, Malaysia

Introduction

Land use and land cover (LULC) change is the human modification of the earth’s terrestrial surface. Over the past 50 years, the driving force behind LULC change has been the increase in agricultural land, and since 1992 the rapid urbanization has also added to the ever-growing change of land use and land cover (IPBES 2019), and worldwide half of the habitable land is used for agriculture (Ritchie and Roser 2019). The changes from one LULC to another can affect the biogeochemical cycles and biogeophysical processes occurring between the surface and atmosphere (Zhou et al. 2020). LULC change can impact biogeochemical cycles like the carbon cycle by altering carbon sinks and carbon dioxide (CO₂) emissions (Li et al. 2020a, b). LULC change affects biogeophysical processes like surface albedo, roughness, and evapotranspiration (Lejeune et al. 2017; Winckler et al. 2020).
2017; Hirsch et al. 2018), and hence alter the energy budget, water budget, and atmospheric variables like temperature and rainfall.

Many studies have shown that urbanization can result in increasing of surface temperature, which leads to Urban Heat Island (UHI) effect (Chao et al. 2020; Hafoud et al. 2020; Qiu et al. 2020; Son et al. 2020; Sultana and Satyanarayana 2020), this can furthermore alter rainfall over urban areas (Pielke et al. 2002; Liang and Ding 2017; Schmid and Niyogi 2017; Liu and Niyogi 2019; Singh et al. 2020; Yu et al. 2020). Urbanization can also affect the hydrological cycle, where impermeable urban surfaces increase runoff and lead to an increased risk of flooding (Li et al. 2018; Ohana-Levi et al. 2018; Hu et al. 2020).

Besides affecting biogeochemical and biogeophysical processes, LULC change also has a great impact on biodiversity and ecosystem services. The conversion of natural vegetation to agricultural land and urban areas can lead to the degradation of the ecosystem services and loss of biodiversity, particularly in biodiversity-rich hotspots like South East Asia. Forests and protected areas play an essential part in providing a healthy environment that benefits both people and nature, by protecting biodiversity, culture, and livelihoods of indigenous people and local communities within these areas (Shadeed and Almasri 2010; UNEP-WCMC and IUCN 2016), and providing important ecosystem services such as carbon sequestration, landscape value, and regulation of major element cycles (Ronchi and Salata 2017). Deforestation can amplify and increase the severity of floods, hence forests and protected areas are important in flood mitigation (Bhattalachjee and Behera 2018; Hasyim et al. 2020; Tembata et al. 2020).

Furthermore, LULC change modelling can be used to estimate and predict future LULC change scenarios and their effects on the environment, by analysing past and current LULC change trends, and the variables that bring about these changes. Moreover, these LULC change models can be integrated with other models, like climate models (Fattah et al. 2021), hydrological models (Shirmo-hammadi et al. 2020; Tankpa et al. 2020; Galleguillos et al. 2021), and ecosystem models (Krause et al. 2019; Li et al. 2020a, b; Rocca and Milanesi 2020), to better understand the impact of LULC change. Modelling future urban land use change, in particular has been of interest to the scientific community in recent times, due to the growing global population and urbanization. In the past few years many studies have used various LULC models, to model future urban expansion and to study their effects on the environment. For example, studies (Chen et al. 2020; Rimal et al. 2020; Grigorescu et al. 2021; Okwuashi and Ndehedehe 2021) all modelled future urban land use change, and in all the studies it was estimated urban land will increase and other land uses will decrease.

In the 25 years (1990–2015) the largest conversion of forest area to other land uses occurred in the tropics, with South East Asia having the highest rate of deforestation since 1990 (Masum et al. 2017). In Malaysia agriculture especially palm oil plantations, was the main contributor to the economy until 1987, and then manufacturing took over as the main contributor to the economy as the government shifted its development policy to the manufacturing sector, and as a result by the year 2000 other sectors like infrastructure and commercial development started growing, giving a rise to urbanization (Abdullah and Hezri 2008). However, palm oil is still a major part of the Malaysian economy and continues to expand. Between 1990 and 2017 agricultural land increased by 55.7% with 98.2% of this area being plantations (Yan et al. 2020). And as of 2017 palm oil plantations occupy 17.62% of the landmass in Malaysia, which has led to a 20% loss in forest land (Ezechi and Muda 2019).

Urbanization on the other hand has seen a rapid increase, and Malaysia is one of the most urbanized countries in East Asia (Plecher 2020). 50.4% of the population lived in urban areas in 1991, this number reached 65% in 2010 and as of 2020 the number of the population living in urban areas has reached 75%, and with projections by 2040, it will reach 85% (Samat et al. 2020). This has resulted in the expansion of urban areas at the expense of other land covers, urban areas increased from 1793.2 ha in 1992 to 3235.4 ha in 2002, and in 2010 urban areas reached 3987.8 ha, on the other hand, agricultural land decreased from 6171.3 ha (53.8%) in 1992 to 3883 ha (35%) in 2010 (Mohammed et al. 2016). At the state level, Selangor and Penang have historically been the most urbanized states in Malaysia and the rate of urbanization has continued to rise over the years. On the other hand, the federal territories of Kuala Lumpur and Putrajaya located within Selangor have been 100% urbanized since 2010 (Hasan and Nair 2014). Overall the level of urbanization in Malaysia as of 2020, stands at 77.16% (O’Neill 2021).

In Malaysia, there have been few studies that have used LULC models, GIS, and remote sensing to study LULC change, with these studies having a varying spatial scope. For example, Gambo et al. (2018) and Rafaai et al. (2020) used LULC change modelling to study the changes within and around protected areas, whereas Verburg et al. (2002), Memarian et al. (2012), Ibrahim and Ludin (2016), Kamarudin et al. (2018) and Majid et al. (2018) studied LULC change at the basin level, and others like Boori et al. (2015), Almdhun et al. (2018) and Samat et al. (2020) studied land use change in cities/towns. In Selangor, several studies have carried out LULC change
modelling. For example, Boori et al. (2015) and Nourqolipour et al. (2015a, b, 2016) analysed LULC change for certain parts of Selangor. Therefore, there are still gaps in knowledge in the trends and changes in LULC in Malaysia, and only a few handful of studies have modelled future LULC changes.

LULC change models are great tools for researchers and professionals to explore the dynamics and drivers that bring about change in LULC (Agarwal et al. 2002). LULC change models are capable of capturing and reproducing these complex dynamics of LULC change and used to extrapolate future land use scenarios (Soesbergen 2016), which can help to inform policies affecting such change. A broad array of models and modelling methods are available to researchers, and each type has certain advantages and disadvantages depending on the objective of the research. There are statistical and empirical models like logistic regression and Markov chain, dynamic models like Cellular Automata (CA), and integrated models (Al-sharif and Pradhan 2014). The Markov chain and CA are the most commonly used methods in LULC change and many studies use an integration of the CA-Markov method (Hamad et al. 2018; Karimi et al. 2018; Huang et al. 2020; Khawaldah et al. 2020; Mansour et al. 2020).

In recent years several machine learning (ML) methods have been commonly used in LULC change modelling, like artificial neural network (ANN), support vector machine (SVM), and Random Forest (RF). For example, Mirici et al. (2017) used a Multi-Layer Perceptron (MLP), ANN, and Markov chain approach to simulate future land use change, whereas Samardžić-Petrović et al. (2016) used the SVM method to model urban land use change. The RF method, which is an ensemble of Decision Trees (DT), has rarely been used in land use modelling (Al-sharif and Pradhan 2015; Samardžić-Petrović and Dragićević 2015; Karimi et al. 2021).

The RF algorithm has several advantages over other machine learning methods: it is faster and easier to understand and interpret, the algorithm is completed at a fixed number of operations, it can process large volumes of data, a small quantity of parameters is needed to be adjusted during modelling, and has higher accuracy compared to other machine learning algorithms (Kamusoko and Gamba 2015; Legdou et al. 2020; Mao et al. 2020).

LULC change is increasing and surpassing climate change as having the most significant impact on environmental change dynamics. Selangor has been experiencing rapid LULC change and urbanization, with the federal territories of Kuala Lumpur and Putrajaya being 100% urbanized. This has therefore had a significant impact on the environmental dynamics in Selangor. The LULC change and urbanization have increased the risk of flooding and air pollution, increased urban temperatures, and resulted in the degradation of the natural ecosystem within Selangor. With Kuala Lumpur and Putrajaya being 100% urbanised, the continuous expansion of these urban centres could affect the surrounding areas and result in urban sprawl.

Therefore, for local decision-makers and urban planners to mitigate the future impacts of the urban expansions and LULC changes, and to implement better land use policies, it is important to know what the future scenarios of LULC change are, determine the spatial changes, and quantify these changes. A few studies have focused on LULC change modelling in Selangor; however, there has not been any study that has analysed the trend and changes of LULC for Kuala Lumpur, Putrajaya, and Selangor as a whole, and what the future scenario could be. Therefore, the objective of this study is to determine what future land-use change scenarios in Selangor are and how the expansion of the urban centres could affect the surrounding areas.

The Decision Forest algorithm, which is a model rarely used in LULC change modelling, has shown to have high accuracy in past studies, and has several advantages over other models. An integrated Decision Forest algorithm and Markov Chain model in the LCM tool of the TerrSet2020 software developed by Clark Labs (Clark Labs 2021) was used in this study. The DF algorithm is a modification of the original RF algorithm by Leo Breiman and Adele Cutler (Breiman 2001; Cutler et al. 2012).

**Methodology**

**Study area**

The state of Selangor, with an area of 8200 km², is located in the western part of Peninsular Malaysia and lies near to the equator (Fig. 1). Selangor is one of the 13 states of Malaysia. It is located at latitudes 2°35′–3°60′ N and longitudes 100°45′–102°00′ E on the west coast of Peninsular Malaysia. It is bordered by Perak to the north, Pahang to the east, Negeri Sembilan to the south, and the Strait of Malacca to the west. The state capital is Shah Alam, while Klang serves as the royal capital. It is the most populated state in Malaysia, with a population of about 6.5 million (Department of Statistics Malaysia 2021) and it’s highly urbanized.

**Data**

The study used Landsat 5 TM (1999 and 2006) and Landsat 8 OLI (2017) satellite images obtained from United State Geological Services (USGS) website (http://earthexplorer.usgs.gov/). The SRTM digital elevation map was also obtained from the USGS website. The ancillary data like road networks and rivers were obtained from...
the open street website at (https://www.openstreetmap.org) which were used for land use change modelling. The slope map was created from the DEM and using the slope tool in ArcGIS 10.2.2 software developed by ESRI (2014).

**Land cover classification and accuracy assessment**

For the land cover classification ERDAS IMAGINE 2020 software developed by Hexagon Geospatial (2020) was used. First, image mosaic was carried out to merge
images which give the full extent of the study area, this was then followed by a haze reduction process to remove or reduce any haze in all the images. For the 2017 image, the panchromatic band of Landsat 8 was used to pan sharpen the image and improve its spatial resolution from 30 m to 15 m for better interpretation and classification of land cover. The maximum likelihood algorithm under supervised classification which is a commonly used method for classification was performed to classify Landsat images of 1999, 2006, and 2017. The following 5 land cover classes were generated: water, natural vegetation, agriculture, built-up land, and cleared land.

For accuracy assessment, the accuracy assessment tool in ERDAS software was used. The accuracy assessment tool allows the comparison of random sample points in the classified map with reference pixels with known class labels. For each map 150 sample points were generated using stratified random sampling, and an error matrix was constructed for each land cover map. The reference data were obtained using Google Earth and its historical imagery were unable to obtain onsite ground truth data. This was then followed by the calculation of the producer accuracy, user accuracy, and the overall accuracy.

**Land cover change modelling**

To simulate a future land cover map, a Decision Forest- Markov chain model is used. The Land Change Modeler (LCM) in TerrSet2020 software developed by Clack Labs (Clark Labs 2021) is used for the modelling. The LCM is based on historical land cover data, transition potential maps, and Markov matrices, to simulate future LC change. The LCM consists of 3 main steps, change analysis, transition potential modelling, and change prediction.

**Change analysis**

The change analysis step calculates the nature and extent of land cover change between time 1 and time 2 and between 2 land cover maps. The changes that are identified are transitions from one land cover state to another. The change analysis evaluates gains and losses, detects net gains, and creates change maps.

**Transition potential modelling**

In this step, the potential of land to transition is identified, and transition potential maps for each transition are created. The transition potential maps that have the same underlying driver variables are grouped within an empirically evaluated transition sub-model. A transition sub-model can consist of a single land cover transition or a group of transitions that are thought to have the same underlying driver variables. These driver variables are used to model the historical change process.

The driver variables used in this study are: distance to rivers, distance to roads, distance to urban area, DEM, and slope (Figs. 2, 3). The driver variables were selected based on the literature review (Camara et al. 2020; Rafaai et al. 2020) and the author’s knowledge of the study area. The transition potential maps are created using the Decision Forest algorithm, which is an implementation of the Random Forest method.

**Change prediction**

In the final step, the historical change of rates calculated in the change analysis step and the transition potential maps, are used to predict a future scenario for a specified future date. The Markov Chain determines the amount of change using the earlier and later land cover maps along with the date specified. The procedure determines exactly how much land would be expected to transition from the later date to the prediction date based on a projection of the transition potentials into the future and creates a file of transition probabilities. The file of transition probabilities is a matrix that records the probability that each land cover category will change to every other category. In this study, the 1999 and 2006 land cover maps are first used for the calibration and validation stages of the model, and a land cover map of 2017 is simulated and validated with a reference land cover map of 2017. Then in the next step, a future land cover map of 2030 is simulated using the 2006 and 2017 maps. The gains and losses and net gain for the year 1999–2006 and 2006–2017 from the change analysis step is shown in Fig. 4.

**Validation**

For the validation of the model the area under the curve (AUC) of Total Operating Characteristic (TOC) and the 3 maps comparison cross tabulation method which measures components of agreement and disagreement are used. TOC method indicates how well the model is predicting change, while the 3 maps cross-tabulation matrix provides detailed information on the accuracy of predicted change and persistence of each land cover class. The 3 maps cross tabulation method uses a reference map of time 1, a reference map of time 2, and a simulated map of time 2 to create the cross-tabulation matrix. In this study reference map 2006 (t1), reference map 2017 (t2), and simulated map 2017 (t2) are used. There are 2 components of agreement called Hits and Correct Rejection and 3 components of disagreement called Misses, False alarm, and Wrong hits. These metrics of agreement and disagreement are recommended by Pontius et al. (2011) and Pontius and Millones (2011) as an alternative to Kappa statistics since Kappa indices attempt to compare accuracy to a baseline of randomness, but randomness is
Fig. 2 Raster maps of distance to roads, distance to rivers, and distance to urban area, used as driver variables for the transition potential modelling step.
Results

Land cover classification and accuracy assessment

The spatio-temporal land cover classification maps of 1999, 2006, and 2017 are shown in Fig. 5. In the period 1999 to 2017 built-up land increased by 608.8 km² (7.5%), agricultural land increased by 285.5 km² (3.5%), and water bodies increased by 21.1 km² (0.3%), whereas natural vegetation decreased by 831.8 km² (10.2%) and cleared land decreased by 83.7 km² (1%) (Table 1). The majority of the urban expansion took place in central Selangor, where the federal territories of Kuala Lumpur and Putrajaya are located. The agricultural expansion from 1999 to 2006 took place mostly in the North of Selangor.

The overall accuracy for the land cover map 1999, 2006, and 2017 are 84%, 92.74%, and 88.67% respectively, and the producer accuracy and user accuracy are shown in Table 2.

For land cover maps 1999 and 2017, the water has low producer accuracy and user accuracy of 50%. In the accuracy assessment of both land cover maps, there were only 2 random samples for the water, and in both cases, the model wrongfully classified one sample point as another class, hence the 50% producer and user accuracy, as shown in the error matrix in Table 3.

not a reasonable alternative for map construction, hence Kappa statistics can give an illusion of high accuracy.
Land cover change modelling and validation

The simulated land cover map of 2030 is shown in Fig. 6. The future simulation of land cover change shows an increase of 723 km² (8.9%) in built-up land and an increase of 57.2 km² (0.7%) in agricultural land from the year 2017 to 2030, on the other hand, there is a decrease
of 663.9 km² (8.1%) in natural vegetation, a decrease of 118.7 km² (1.5%) in cleared land and a very small decrease in water bodies (Table 4).

Model validation
The area under the curve (AUC) of the Total Operating Characteristic (TOC) is shown in Fig. 7. The validation had an AUC of 0.84.

The hits, misses, false alarms, wrong hits, and correct rejections of the cross-tabulation matrix are shown in Fig. 8. The total components of agreement were 71.1% which is the sum of Hits (3.1%) and Correct rejections (68%), and the total components of disagreement were 28.9% which is the sum of Misses (14.9%), False alarm (9.4%) and Wrong hits (4.7%).

Discussion
This study developed a DT-MC model to analyse the trend in land cover and to estimated future change in Selangor. The trend in land cover change from the period 1999 to 2017, shows an increase of 608.8 km² (6.8%) in built-up area. The majority of the built-up areas are located in central Selangor, where the highly populated federal territories of Kuala Lumpur and Putrajaya are located. This increase in built-up areas in Selangor could be mostly attributed to the expansions of Kuala Lumpur, Putrajaya, and the surrounding Klang Valley, as previous studies have shown these cities have expanded and resulted in urban sprawl (Rosni et al. 2016; Almdhun et al. 2018). The AUC value of 0.84 and the higher percentage of total components of agreement (71.1%) compared to total components of disagreement (28.9%), show that the model is adequately calibrated and validated, and it’s suitable for simulating future land use maps. This is comparable with the study by Samardžić-Petrović et al. (2015), that had similar AUC value and accuracy, which indicates that the DF-MC is a suitable model in land cover change modelling.

The projected 2030 land cover map shows this urban expansion to continue, and the model estimated that urban land covers to increase by 723 km² (8.9%) in the period 2017 to 2030 (Fig. 9). The model shows the urban expansion taking place mostly at the edges of existing urban boundaries, and road networks, an indication of the distance to roads and distance to exiting urban areas, are the most impactful variables on urban expansion.

The increase of transportation networks has played a major role in the expansion of urban areas in Malaysia. For example, in George Town, Northern Malaysia, it is estimated by 2030 urban land to expand from 925.77 to 1253.95 km², with the North–South highway and the

| Area               | 1999 | 2006 | 2017 | Change 1999–2017 |
|--------------------|------|------|------|------------------|
| Water              | 74.5 | 92.4 | 95.5 | 21.1             |
| Natural vegetation | 4439.0| 3830.7| 3607.1| -831.8          |
| Agriculture        | 2336.9| 2691.8| 2622.4| 285.5            |
| Built-up land      | 989.8| 1187.5| 1598.6| 608.8            |
| Cleared land       | 316.5| 354.2| 232.8| -83.7            |

| Table 1 | Land cover area of each cover type for the years 1999, 2006, and 2017, and the changes for each cover type from 1999 to 2017
|---------|--------|--------|--------|-------------------|
|         | 1999 (km²) | 2006 (km²) | 2017 (km²) | Change (km²) |
| Water   | 74.5      | 92.4      | 95.5      | 21.1           |
| Natural vegetation | 4439.0 | 3830.7 | 3607.1 | -831.8 |
| Agriculture | 2336.9 | 2691.8 | 2622.4 | 285.5 |
| Built-up land | 989.8 | 1187.5 | 1598.6 | 608.8 |
| Cleared land | 316.5 | 354.2 | 232.8 | -83.7 |

| Table 2 | The producer’s accuracy and user’s accuracy for the land cover classification maps of 1999, 2006 and 2017, showing the accuracy of the land cover maps
|---------|--------|--------|--------|--------|--------|
|         | 1999 (Producers accuracy (%)) | Users accuracy (%) | 2006 (Producers accuracy (%)) | Users accuracy (%) | 2017 (Producers accuracy (%)) | Users accuracy (%) |
| Water   | 50.00 | 100.00 | 100.00 | 100.00 | 50.00 | 50.00 |
| Natural vegetation | 97.14 | 80.95 | 98.18 | 91.53 | 96.49 | 83.33 |
| Agriculture | 70.00 | 89.74 | 88.64 | 92.86 | 86.27 | 91.67 |
| Built-up land | 82.61 | 90.48 | 85.00 | 100.00 | 84.85 | 96.55 |
| Cleared land | 60.00 | 60.00 | 100.00 | 80.00 | 71.43 | 100.00 |
second bridge between Georg Town and Penang island playing a major role in this expansion (Samat et al. 2020). Similarly, a study in Northern part of Selangor shows an estimated urban expansion of about 33% from 2015 to 2033, with distance to road and distance to build-up areas, being the major variables (Camara et al. 2020).

Similar to build-up areas, agricultural land has seen an increase in area of 285.5 km² (3.5%) from 1999 to 2017, and the model estimated that agricultural land will increase by 57.2 km² (0.7%) from 2017 to 2030. The increase of built-up area and agricultural land on the other hand has resulted in a decrease in natural vegetation by 831.8 km² (10.2%) from 1999 to 2017, where natural vegetation is turned into build-up and agricultural land cover. The hotspots of these agricultural land use change, are peat swamps in the north east of Selangor and the coastal regions, where distance to transportation infrastructure plays a major role, as distance to transport increases, agricultural land decrease (Olaniyi et al., 2015). The model estimates that there will be a loss of 663.9 km² (8.1%) in natural vegetation from 2017 to 2030. Water and cleared land both had very minimal changes, therefore it shows that the majority of the land for urban development and agriculture comes from natural vegetation, as the population and the economy of Selangor continue to increase.

Table 3 Error matrix for the land cover maps 1999, 2006, and 2017, showing the accuracy of the land cover maps

| Classified data         | Reference data       | Classified data | Reference data       |
|-------------------------|----------------------|-----------------|----------------------|
| Water                   | Natural vegetation  | Water           | Natural vegetation  |
| Water                   | 1                    | 1               | 1                    |
| Natural vegetation      | 68                   | 14              | 4                    |
| Agriculture             | 2                    | 35              | 39                   |
| Built-up land           | 0                    | 0               | 19                   |
| Cleared land            | 0                    | 0               | 2                    |
| Column total            | 2                    | 70              | 50                   |

Table 4 The area of land cover for the simulated year 2030, and changes between 2017 and 2030

|          | 2030     | Change from 2017 to 2030 |
|----------|----------|--------------------------|
|          | Area (km²) | Area (%)     | Area (km²) | Area (%)     |
| Water    | 93.9      | 1.2                  | −1.6       | 0.0          |
| Natural vegetation | 2943.3    | 36.1                | −663.9     | −8.1         |
| Agriculture | 2679.6    | 32.9                | 57.2       | 0.7          |
| Built-up land | 2321.6    | 28.5                | 723.0      | 8.9          |
| Cleared land | 114.0     | 1.4                 | −118.7     | −1.5         |
in 2014 to construct an expressway, in 2016 around 30 ha of Sungai Puteh North and South Forest Reserves were de-gazetted for a highway, and 28.3 ha of Bukit Lagong Forest Reserve was de-gazetted to make room for a housing project, recently the government has proposed to de-gazette 985 ha (97%) of the Kuala Langat North Forest Reserve for urban development (Ravindran and Rajendra 2020).

As the demand for urban land and agricultural land increases, more of the protected areas and urban green spaces will be used to meet these demands. The model shows the urban and agriculture expansion by 2030 will result in more loss of natural vegetation and urban green spaces, and the expansion of the major urban centres will affect their surrounding areas. Therefore, the results of this study can help local policymakers and urban planners to visualise and compare the impacts of future land cover change, which can help to formulate and implement better land use policies and structural plans, which can limit or prevent urban sprawl and uncontrolled urban expansion into surrounding areas. The future land cover change projection can also help local government to mitigate the effects of flooding, and develop a more resilient structural plan.
Conclusion

This study assessed the trend in land cover change in the highly urbanized and populated state of Selangor, and modelled a future land use scenario using a Decision Forest-Markov model in the TerrSet software. Land use maps of 1999, 2006, and 2017 were classified into 5 classes, namely water, natural vegetation, agriculture, built-up land and cleared land. To validate the model, the simulated land use map of 2017 was compared with the reference map of 2017, and AUC of TOC and disagreement parameters were used. The AUC of 0.84 and the higher percentage of total components of agreement indicate the model is sufficiently validated. After the successful validation of the model, the land use maps of 2006 and 2017 are used to simulate a future land use map of 2030.

The results of this study show that, at the current rate of development and urbanization, large parts of natural vegetation will be lost by 2030. Therefore, the findings of this study can help Selangor state government, policymakers, urban planners, and other stakeholders to better understand and manage future development in Selangor and to achieve the objectives of the Selangor State Structural Plan 2035, where one of the objectives is to maintain and preserve 32% of forest areas. Furthermore, the results of this study can be incorporated in other studies, like climate and hydrological studies, where effects of land cover change and increasing urbanization on temperature and flood risk can be studied.

There are some limitations in using the DF-MC model. The Markov chain is essentially a projection model and not policy sensitive, therefore making it difficult to include policy variables like socioeconomic and population variables into the modelling (Iacono et al. 2015). To overcome this limitation would need to integrate the MC model with other models that can incorporate socioeconomic variables (Hamad et al. 2018). Moreover, the Decision Forest is a complex model that requires high computational power and is more time-consuming. The complexity of the model can be reduced by reducing the number of decision trees and lowering the number of variables used, and therefore it is advisable to carry out variable selection processes during modelling (Samardžić-Petrović et al. 2015). Overall, the model is capable of estimating future land cover change and can be used in future studies.

Abbreviations

LULC: Land use and land cover; LC: Land cover; AUC: Area under the curve; TOC: Total operating characteristics; CO₂: Carbon dioxide; UHI: Urban heat island; PA: Protected areas; CA: Cellular automata; ML: Machine learning; ANN: Artificial neural network; SVM: Support vector machine; RF: Random forest; MLP: Multi-layer perceptron; DT: Decision trees; DEM: Digital elevation model; LCM: Land change modeller.
Acknowledgements
Not applicable.

Authors’ contributions
MA and LB conceptualized and designed the study and MA carried them out. MA developed the model and performed the simulations. MA prepared the manuscript with contributions from LB and AC. AC and LB provided supervision and reviewed the final draft of the manuscript. All authors read and approved the final manuscript.

Funding
Not applicable.

Availability of data and materials
All Landsat satellite images and DEM data are freely available from US Geological Survey at https://earthexplorer.usgs.gov/. The ancillary data like roads network and rivers are freely available at www.openstreetmap.org.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no conflict of interest.

Author details
1Department of Civil Engineering, University of Nottingham Malaysia, Jalan Broga, 43500 Semenyih, Selangor, Malaysia. 2School of Environmental and Geographical Sciences, University of Nottingham Malaysia, Jalan Broga, Selangor 43500 Semenyih, Malaysia.

Received: 18 April 2021   Accepted: 2 December 2021
Published online: 24 January 2022

References
Abdullah SA, Hezri AA (2008) From forest landscape to agricultural landscape in the developing tropical country of Malaysia: pattern, process, and their significance on policy. Environ Manag 42(5):907–917. https://doi.org/10.1007/s00267-008-9178-3
Agarwal C, Green GM, Grove JM, Evans TP, Schweik CM (2002) A review and modelling of urban sprawl through remote sensing and GIS in Kuala Lumpur, Malaysia. Ecol Process 4:15. https://doi.org/10.1186/s13717-015-0040-2
Breiman L (2001) Random forests. Mach Learn 45(1):5–32. https://doi.org/10.1023/A:1010933404324
Camara M, Jamil NR, Abdullah AFB, Hashim RB (2020) Integrating cellular automata Markov model to simulate future land use change of a tropical basin. Glob J Environ Sci Manag 6(3):403–414. https://doi.org/10.22034/gjesm.2020.03.09
Chao Z, Wang L, Che M, Hou S (2020) Effects of different urbanization levels on land surface temperature change: taking Tokyo and Shanghai for example. Rem Sens 12(12). https://doi.org/10.3390/rs12122202
Chen G, Li X, Liu X, Chen Y, Liang X, Leng J, Xu X, Liao W, Qiu Y, Wu Q, Huang K (2020) Global projections of future urban land expansion under shared socioeconomic pathways. Nat Commun 11:537. https://doi.org/10.1038/s41467-020-14386-x
Clark Labs (2021) TerraSet 2020 geospatial monitoring and modeling software. Clark University 950 Main St, Worcester MA 01610 USA. https://clarklabs.teraset/. Accessed 5 Aug 2021
Cutler A, Cutler DR, Stevens JR (2012) Random forests. In: Ensemble machine learning. Boston, MA: Springer US, pp 157–175. doi: https://doi.org/10.1007/978-1-4419-9326-7_5
Department of Statistics Malaysia (2021) Department of Statistics Malaysia Official Portal. Department of Statistics Malaysia. https://www.dosm.gov.my/v1/index.php?option=com_&_view=menueu_id=weGlY7m9rcEvZ1myYW45dmpnZH4d0z0#. Accessed 30 Jul 2021
ESRI (2014) ArcGIS 10.2.2. Esri Geographic information system company. Esri. https://www.esri.com/arcgisidar. Accessed 3 Aug 2021
Ezechi EH, Muda K (2019) Overview of trends in crude palm oil production and economic impact in Malaysia. Srijiviya J Environ 4(1):19–26. https://doi.org/10.22135/vje.2019.4.119
Fattah MA, Morshed SR, Morshed SY (2021) Multi-layer Perceptron-Markov chain-based artificial neural network for modelling future land-specific carbon emission pattern and its influences on surface temperature. SN Appl Sci 3:359. https://doi.org/10.1007/s42452-021-04351-8
Galleguillos M, Gimeno F, Puelma C, Zambrano-Bigiarini M, Lara A, Rojas M (2021) Disentangling the effect of future land use strategies and climate change on streamflow in a Mediterranean catchment dominated by tree plantations. J Hydrol 595:126047. https://doi.org/10.1016/j.jhydrol.2021.126047
Gambo J, Mohn Shafi HZ, Shaharum NSN, Abidin FAZ, Rahman MT (2018) Monitoring and predicting land use-land cover (LULC) changes within and around Krau wildlife reserve (KWR) protected area in Malaysia using multi-temporal land sat data. Geoplan J Geomat Plan 5(1):17–34. https://doi.org/10.14710/geoplanning.5.1.17-34
Grigorescu I, Kucsicsa G, Popovici E-A, Mitrică B, Mocanu I, Dumitriuč M (2021) Modelling land use/cover change to assess future urban sprawl in Romania. Geocarto Int 36(7):721–739. https://doi.org/10.1080/10704971.2019.1624991
Hafoud S, Boutouil K, Oussama A, Mahjoubi F, Kzaiber F (2020) Urbanization and its impact on land surface temperature changes using landsat image in Dakhila City, Morocco. Int J Adv Res Eng Technol 11(6):143–155
Hamad R, Balzer H, Kolo K (2018) Predicting land use/land cover changes using a CA-Markov model under two different scenarios. Sustainability 10(10):3421. https://doi.org/10.3390/su10103421
Hasan AR, Nair PL (2014) Urbanisation and growth of metropolitan centres in Malaysia. Malays J Econ Stud 51(1):87–101
Hasym S, Abdullah R, Ibrahim H (2020) Forest damage and preservation through forest resources management in Indonesia. GeoJournal 86:2183–2189. https://doi.org/10.1007/s10708-020-10177-5
Hapood S, Boustiel K, Oussama A, Mahjoubi F, Kzaiber F (2020) Urbanization and its impact on land surface temperature changes using landsat image in Dakhila City, Morocco. Int J Adv Res Eng Technol 11(6):143–155
Hirsch AL, Guillod BP, Spenereit SJ, Beyerle U, Boysen LR, Brovkin V, Davin EL, Doelman JC, Kim H, Mitchell DM, Nitta T, Shiozama H, Sparrow S, Stehfest E, van Vuuren DP, Wilson S (2018) Biophysical impacts of land-use change on climate extremes in low-emission scenarios: results from HAPPI-Land. Earth’s Future 6(3):396–409. https://doi.org/10.1002/2017EF000744
Hu S, Fan Y, Zhang T (2020) Assessing the effect of land use change on surface runoff in a rapidly urbanized city: a case study of the central area of Beijing. Land 9:17
Rafai NH, Abdullah SA, Hasan Reza MI (2020) Identifying factors and predicting the future land-use change of protected area in the agricultural landscape of Malaysian peninsula for conservation planning. Rem Sens Appl Soc Environ 18:100298. https://doi.org/10.1016/j.rsa.2020.100298

Ravindran S, Rajendia E (2020) Green lung under threat. In: Thesrat. https://www.thesrat.com.my/metro/metro-news/2020/02/18/green-lung-under-threat. Accessed 6 Mar 2021

Rimal B, Sloan S, Keshkar H, Sharma R, Rijal S, Shrestha UB (2020) Patterns of historical and future urban expansion in Nepal. Rem Sens 12(4):6268. https://doi.org/10.3390/RS1204628

Ritchie H, Roser M (2019) Land Use. In: OurWorldInData.org. https://ourworldindata.org/land-use. Accessed 6 Mar 2021

Rocca FD, Milanesi P (2020) Combining climate, land use change and dispersal to predict the distribution of endangered species with limited vagility. J Biogeogr 47(7):1427–1438. https://doi.org/10.1111/JBI.13804

Ronchi S, Salaeta S (2017) Assessment of the ecosystem services capacity in natural protected areas for biodiversity conservation. IOP Conf Ser Mater Sci Eng 245:072031. https://doi.org/10.1088/1757-899X/245/7/072031

Rosni NA, Noor NM, Abdullah A (2016) Managing urbanisation and urban sprawl in Malaysia by using remote sensing and GIS applications. Plan Int Mal My 4(4):17–30. https://doi.org/10.21837/pmjnewsletter.v14i4.145

Samardžić-Petrović M, Dragićević S, Bajat B, Kovačević M (2015) Exploring the decision tree method for modelling urban land use change. Geomatica 69(3):313–325

Samardžić-Petrović M, Dragićević S, Kovačević M, Bajat B (2016) Modelling urban land use changes using support vector machines. Trans GIS 20(5):718–734. https://doi.org/10.1111/TGIS.12174

Samat N, Mahamud MA, Tan ML, Tilaki MJM, Tew YL (2020) Modelling land cover changes in peri-urban areas: a case study of George Town corurbation, Malaysia. Land 9(10):373. https://doi.org/10.3390/land9100373

Schmid PE, Niyogi D (2017) Modelling urban precipitation modification by spatially heterogeneous aerosols. J Appl Meteorol Climatol 56(8):2141–2153. https://doi.org/10.1175/JAMC-D-16-0320.1

Shadeed S, Almasri M (2010) Application of GIS-based SCS-CN method in West Bank catchments, Palestine. Water Sci Eng 3(1):1–13. https://doi.org/10.3882/jisesn.1674-2370.2010.01.001

Shirmohammadi B, Malekian A, Salajegheh A, Taheri B, Azarnivand H, Malek Z, Verburg PH (2020) Scenario analysis for integrated water resources management under future land use change in the Urmia Lake region, Iran. Land Use Policy 90:104299. https://doi.org/10.1016/J.LANDUSEPOL.2019.104299

Singh J, Karnakar S, Painamzumder D, Ghosh S, Niyogi D (2020) Urbanization alters rainfall extremes over the contiguous United States. Environ Res Lett 15(7):074033. https://doi.org/10.1088/1748-9360/ab9880

Soesbergen AV (2016) A review of land-use change models. In: UNEP WCMC. https://www.unep-wcmc.org/system/comfy/cms/files/files/000/000/802/original/Land_Use_Chan ge_Models_2016_WEB.pdf. Accessed Mar 2021

Son NT, Chen CF, Chen CR (2020) Urban expansion and its impacts on local temperature in San Salvador, El Salvador. Urban Clim 32:100617. https://doi.org/10.1016/j.uclim.2020.100617

Sultana S, Satyanarayana AHV (2020) Assessment of urbanisation and urban heat island intensities using Landsat imageries during 2000–2018 over a sub-tropical Indian City. Sustain Cities Soc 52:101846. https://doi.org/10.1016/j.scs.2019.101846

Tembata K, Matsumoto K, Yamamoto M, Yamamoto Y (2020) Forest and floods mitigation: evidence from China. SSRN Electron J. https://doi.org/10.2139/ssrn.3515698

UNE-P-WCMC and IUCN (2016) Protected Planet Report 2016 how protected areas contribute to achieving global targets for biodiversity. Cambridge UK and Gland, Switzerland

Verburg PH, Soepboer W, Veldkamp A, Limpiada R, Espaldon V, Mastura SSA (2002) Modelling the spatial dynamics of regional land use: the CLUE-S model. Environ Manag 30(3):391–405. https://doi.org/10.1007/s00267-002-2630-x

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com