Arabian Journal of Geosciences

Spatio-temporal simulation of future urban growth trends using an integrated CA-Markov model

--Manuscript Draft--

Manuscript Number:	AJGS-D-18-02765
Full Title:	Spatio-temporal simulation of future urban growth trends using an integrated CA-Markov model
Article Type:	Original Paper
Corresponding Author:	Maher Milad M Aburas LIBYA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	
Corresponding Author's Secondary Institution:	
First Author:	Maher Milad M Aburas
First Author Secondary Information:	
Order of Authors:	Maher Milad M Aburas
	Biswajeet Pradhan
	HO YUEK MING
	Abdul Hakim Bin Salleh
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	Urban growth, a dynamic and demographic phenomenon, refers to the increased spatial value of urban areas, such as cities and towns, due to social and economic forces. Nowadays, urban lands are rapidly increasing, replacing non-urban lands such as agricultural, forest, water, rural, and open lands. In this study, a CA-Markov model was utilized to predict the growth of urban lands and their spatial trends in Seremban, Malaysia. The performance of the CA-Markov model was also assessed. The Markov chain model was applied to produce the quantitative values of transition probabilities for urban and non-urban lands. Subsequently, the CA model was used to predict the dynamic spatial trends of land changes. The change in urban and non-urban land use from 1984 to 2010 was modeled using the CA-Markov model for calibration purposes and to compute optimal CA transition rules as well as to predict future urban growth. In the accuracy assessment process, the CA-Markov model was validated using a Kappa coefficient. The overall accuracy of the Kappa index statistics was 83%, which indicates the excellent performance of the model proposed in this study. Finally, based on the CA transition rules and the transition area matrix produced from the calibration process using the Markov Chain model, future urban growth in Seremban for 2020 and 2030 was simulated.
Suggested Reviewers:	Hossein Mojaddadi hossein_mras@yahoo.com
	Goma Ahmed joma762001@gmail.com
	Issa Oskoui issaoskoui@gmail.com
	Motasem Azaiza my.azaiza@gmail.com

Spatio-temporal simulation of future urban growth trends using an integrated CA-Markov model

Maher Milad Aburas[1], Yuek Ming Ho[2], Biswajeet Pradhan[3], Abdul Hakim Salleh[1*]

1. School of Civil Engineering,Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia
2. Department of Environmental Management, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor Darul Ehsan, Malaysia
3. Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW 2007, Australia

Spatio-temporal simulation of future urban growth trends using an integrated CA-Markov model

Abstract. Urban growth, a dynamic and demographic phenomenon, refers to the increased spatial value of urban areas, such as cities and towns, due to social and economic forces. Nowadays, urban lands are rapidly increasing, replacing non-urban lands such as agricultural, forest, water, rural, and open lands. In this study, a CA-Markov model was utilized to predict the growth of urban lands and their spatial trends in Seremban, Malaysia. The performance of the CA-Markov model was also assessed. The Markov chain model was applied to produce the quantitative values of transition probabilities for urban and non-urban lands. Subsequently, the CA model was used to predict the dynamic spatial trends of land changes. The change in urban and non-urban land use from 1984 to 2010 was modeled using the CA-Markov model for calibration purposes and to compute optimal CA transition rules as well as to predict future urban growth. In the accuracy assessment process, the CA-Markov model was validated using a Kappa coefficient. The overall accuracy of the Kappa index statistics was 83%, which indicates the excellent performance of the model proposed in this study. Finally, based on the CA transition rules and the transition area matrix produced from the calibration process using the Markov Chain model, future urban growth in Seremban for 2020 and 2030 was simulated.

Keywords: Urban growth; Markov Chain; Cellular Automata; prediction; modelling.

1. Introduction

Urban development has become a global issue, and has resulted in planners and decision makers becoming increasingly concerned over its future impacts on the ecosystem (Bihamta et al., 2014). Simulating and predicting urban sprawl patterns has become essential to ecosystem protection and sustainable development (Barredo et al., 2003). In addition, the complex structure of the urban environment must be understood to simulate urban dynamics correctly (Barredo et al., 2003). In the process of urban growth simulation, the chronology of the issue of sprawl and significant historical information must be considered, so that spatial and temporal relationships can be accurately understood (Sudhira et al., 2004). Hence, the process of obtaining the actual knowledge of growth factors that affect future land uses can be improved using simulation techniques (Pijanowski et al., 2002). Understanding of spatial and temporal changes, as well as all effective elements, can be facilitated using remote sensing (RS) and geographic information system (GIS) techniques (Punia & Singh, 2012).

RS and GIS techniques are commonly used to monitor and control urban growth patterns (Zhang et al., 2011). In recent years, RS and GIS techniques have been considered as effective tools for helping planners and decision-makers formulate sustainable policies. These modern techniques have several advantages, such as their low cost (Yeh & Li, 2001), effective visual interpretation (Epsteln et al., 2002), updatable spatial and temporal databases (Punia & Singh, 2012), monitoring and controlling tools (Doygun, 2009; Tran, 2008), and accurate tools for evaluating, analyzing, and simulating spatial phenomena (Ren et al., 2013). For these reasons, environmental planners and urban designers have relied heavily on RS and GIS techniques to model urban growth patterns and future land-use changes.

Currently, various types of models and methods utilizing the RS and GIS techniques are being employed for the general modeling of urban growth patterns and simulation of land-use changes (Mohammad et al., 2013). There are studies that have used traditional models, which depend on the assessment of the dynamic growth of urban areas, such as the CA models (Aburas et al., 2017). Some of these studies have also relied on quantitative models, such as logistic regression (LR), for simulation and prediction (Alsharif & Pradhan, 2014). Other studies have relied on the integration of different types of models, such as the Markov chain (MC) and the CA models, to achieve accurate and realistic results (Al-sharif & Pradhan, 2013). The modeling of urban growth patterns based on RS and GIS techniques is done in order to understand the spatial process of urban movement within a specific time toward the creation of future policies of sustainable development (Wang & Maduako, 2018).

The use of GIS and RS techniques to model urban growth patterns and future land-use changes can greatly benefit land-use planning and the 'cause-and-effect' analysis of land-use movement. Sites that are facing environmental change and urban sprawl as well as potential critical sites can be identified using several types of models, such as quantitative or spatio-temporal models (Verburg et al., 2002). Spatial modeling is used to simulate land-use patterns that are indispensable towards supporting the development and implementation of urban planning policies (Inouye et al., 2015). In general, planners and policy makers are looking at useful measurements that depend on wide-reaching information, data integration, and qualitative criteria (Celio et al., 2014).

The Cellular Automata (CA) model has an open structure and can be integrated with other models to simulate and predict urban growth patterns (Clarke, 1997). The CA model's flexibility, intuitiveness, and ability to integrate spatial and temporal dimensions of the processes, as well as the capability to model complex dynamic systems, are major reasons for its widespread application in the simulation of urban growth patterns and future land-use changes in recent years (Santé et al., 2010). Tobler (1979) first proposed the application of cellular space models for geographic modeling. Following this, theoretical approaches for simulating urban growth using CA-based models started to emerge in the 1980s (Batty & Xie, 1994; Couclelis, 1985; White & Engelen, 1994).

The conceptual growth of CA studies and the evolution of computing capability contributed to the first operational urban CA model, which first saw use in real-world urban systems in the 1990s. The capability of the urban CA model to simulate and predict land-use changes is based on the assumption that previous urban growth affects future patterns through local and regional interactions among different types of land uses (Santé et al., 2010). Moreover, the urban CA model can be easily integrated with the GIS environment (Wagner, 1997); thus, the CA model has a high spatial resolution and computational efficiency (Santé et al., 2010). The other key fields of urban CA models, which are considered powerful spatial dynamic modeling techniques that represent a major development over previous conventional models, are: (i) spatiality; (ii) the linking of macro to micro approaches; (iii) the integration between GIS and RS techniques; (iv) dynamics: and (v) simplicity and visualization (Batty & Xie, 1994; Clarke, 1997; White & Engelen, 1994, 2000; Wu, 1998).

The Markov chain is usually utilized to model and predict changes, dimensions, and trends of urban growth patterns (Aburas et al., 2017). Changes in urban and non-urban lands can be analyzed and summarized by the number of transition area probabilities from one status to various other statuses during a certain period of time using the Markov chain model (Coppedge et al., 2007). The Markov chain model does not have the ability to simulate changes in spatial trends. However, it is a powerful model, which has the capability to predict the quantity of land change (Yang et al., 2012). The integration between the CA and Markov Chain models is an effective technique to estimate quantities and to model spatio-temporal dynamics because this type of GIS and RS model and data can be proficiently incorporated (Al-sharif & Pradhan, 2013). The integration of dynamic simulation models (such as the CA model) with that of statistical and empirical models (such as the Markov chain) has overcome the shortcoming inherent in each of them, i.e., the difficulty in dynamically or statistically simulating urban issues, and one will therefore complement the other (Guan et al., 2011).

In this research, the city of Seremban, Malaysia, was chosen as a case study. Seremban has faced rapid urban growth over the last two decades. This growth has led to the continuous, rapid change of non-urban lands into urban lands. This study used an integrated Markov chain and Cellular Automata model (CA-Markov) to simulate rapid urban growth in Seremban City from 1990 to 2010, and then to predict future land changes quantitatively and spatially. To the authors' best knowledge, no study of this kind has ever been done in this city before.

2. Methodology

2.1. Study area

Seremban River Basin is the largest district in the Negeri Sembilan State (Figure 1). Seremban is also the capital of Negeri Sembilan State. It occupies a total land area of approximately 951.87 sq. km and includes the districts of Seremban town, Setul, Labu, Rasah, Ampangan, Rantau, Pantai, and Lenggeng. Seremban is located approximately 20 km from Putrajaya, the national capital of Malaysia, and 67 km from Kuala Lumpur, the economic center of Malaysia. The population of Seremban is more than 500,000 and is expected to increase to 1,000,000 in 2020 (DOSM, 2011). Seremban city was selected as the study area because: (i) it is the biggest city in the Negeri Sembilan State; (ii) it is the economic center of the Negeri Sembilan State: (iii) it is located near the main developed areas in Malaysia, such as Kuala Lumpur, Putrajaya, and Selangor; (iv) it is an extension of the urban mass of Kuala Lumpur; and (v) it is the future center for urban development.

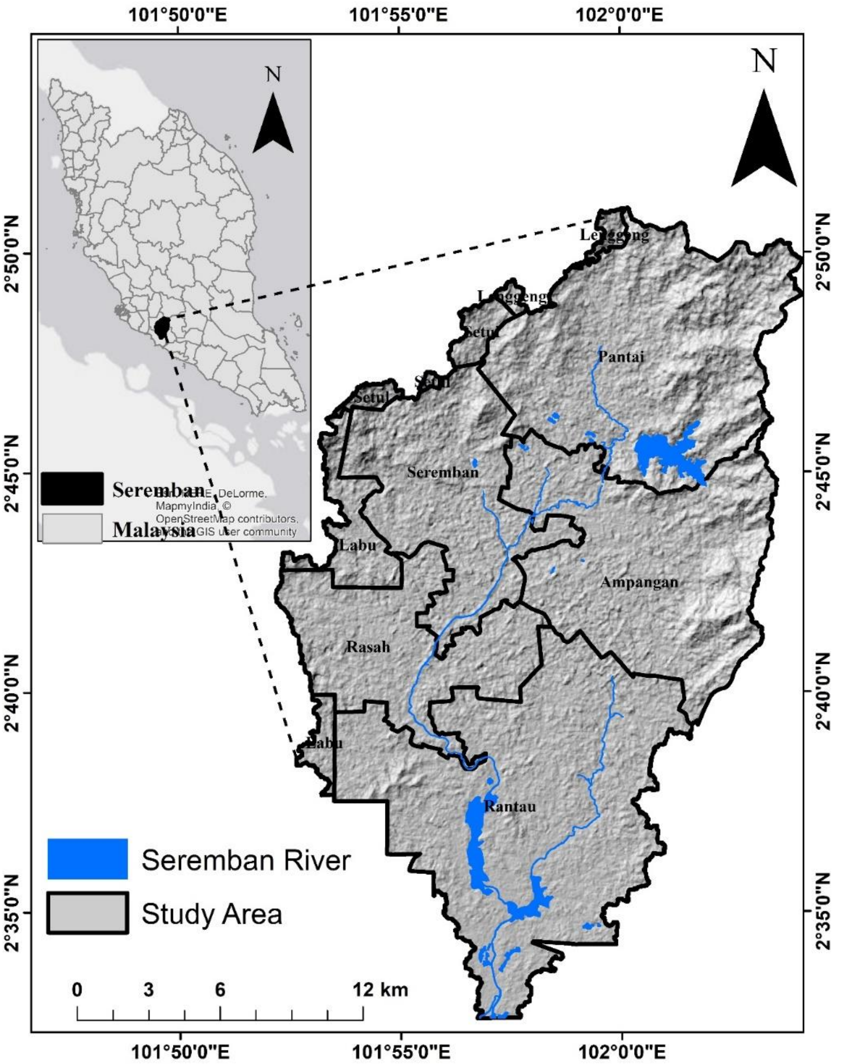

Figure.1. Location of the study area.

2.2. Data and Methods

This study utilized land-use maps of 1984, 1990, 2000, and 2010 from the Department of Agriculture of Malaysia (Figure 2). These land-use maps were extracted from SPOT 2, 4, and 5 images, with a 10-m and 2.5 m spatial resolutions of SPOT 2.4 and SPOT 5, respectively. All SPOT images were registered and geo-corrected with ground control points using a Global Positioning System (GPS) and were classified using image-enhancement techniques. A supervised classification method was used to group and extract all clipped images into land-use categories. Field data were collected using GPS to assess the accuracy of classification by comparing the classified images with GPS points from the field for each type of land use. The accuracy assessment values reached acceptable Kappa index values, indicating that the image classification is acceptable. Based on the Anderson scheme, an acceptable Kappa index value to yield an accurate assessment should be higher than 0.85 (Anderson, 1976). The total accuracies of the land-use maps were 92%, while the Kappa coefficient values were 0.90. Thus, the classification of land-use maps by the Malaysian Department of Agriculture meets the present study's requirements. The topographic map of 2012 was collected and used to identify the administrative boundary of the whole Seremban area and that of each district (Table 1).

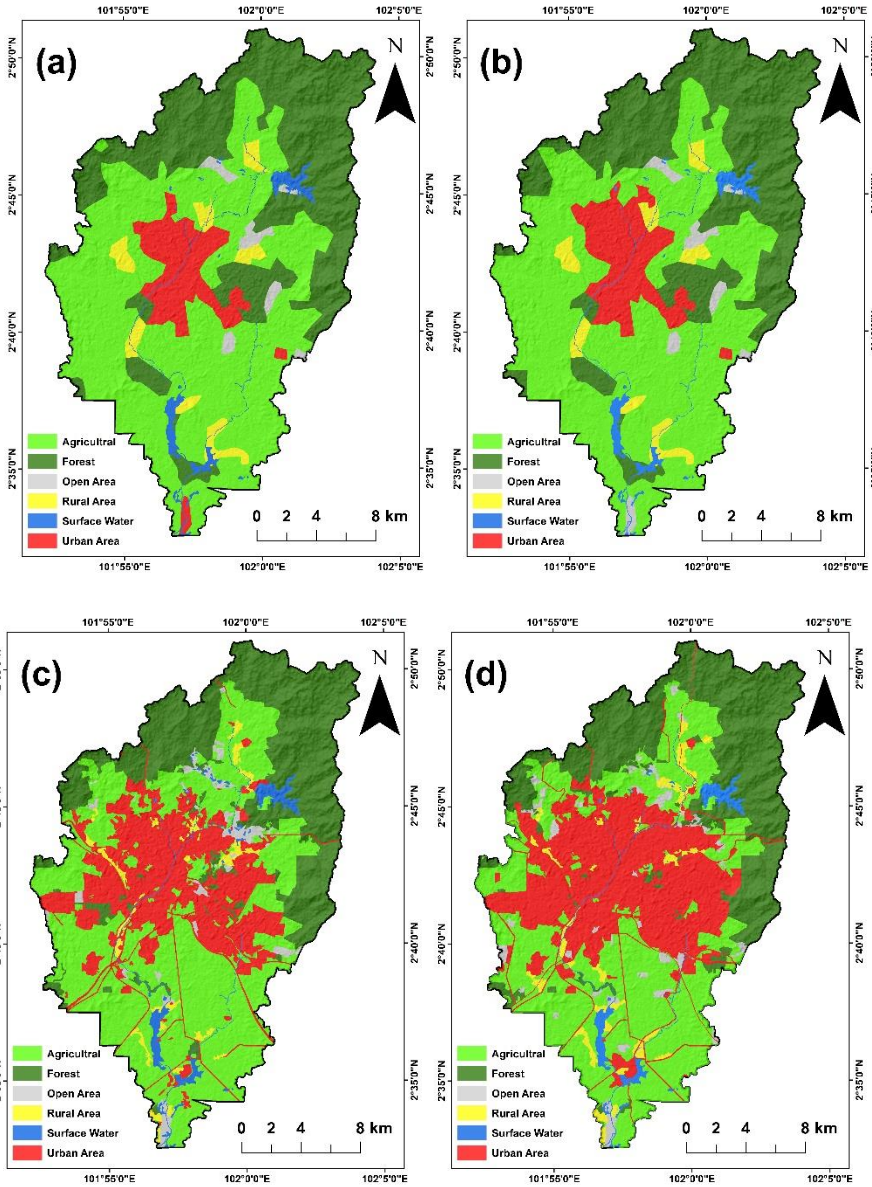

Figure.2. Land-use maps of Seremban River Basin: (a) 1984; (b) 1990; (c) 2000; and (d) 2010.

Table.1. Data used in the study.

Materials	Sours	Type of data	Scale
Land-use maps of 1984, 1990, 2000, 2010	Department of Agriculture, Malaysia	Grid	10*10
Topographic map of 2012	Department of Surveying and Mapping, Malaysia (JUPEM)	Map	1:25,000
DEM	USGS	Grid	30*30

Land-use maps were reclassified into two types of land use, urban area and non-urban area, to comply with the general objective of the study. Urban areas include residential, commercial and services, industrial, transportation, communications, and utility areas, as well as mixed urban or built-up lands and other urban or built-up lands, while non-urban areas include other types of land use, such as water bodies, agricultural lands, forests, and open areas. Land-use maps were classified into urban and non-urban area classes mainly because spatial simulation was applied in this study to predict the urban growth patterns. The models used to predict urban growth in Seremban are discussed in more detail below:

2.2.1. Urban CA model

An urban CA model can be designed based on multiple phases, namely: (i) the data collection phase, which requires different types of data according to the type of model, data availability, and the existence of a type of integration with other models (Aburas et al., 2016); (ii) selection of factors influencing urban growth patterns (Aburas et al., 2017); (iii) identification of the characteristics of CA that are used for simulation, such as defining the lattice, determining cell state, identifying the neighborhood properties, and identifying the transition rules that will be used (Clarke, 1997; White & Engelen, 2000; White et al., 2000); and (iv) validation and calibration of the model using an actual land-use model with the Kappa index (Al-sharif & Pradhan, 2013; Mohammad et al., 2013). Subsequently, simulation and prediction of future land use are undertaken (Figure 3).

CA models use a simple mechanism to identify future conditions for cells, where their future condition is defined by identifying the actual condition for each cell and by determining the real condition of neighboring cells (Couclelis., 1997). The CA model, considered the simplest type of dynamic spatial model, essentially consists of: (i) the cell lattice (i.e., the urban CA model consists of a grid containing square cells or other geometrical shapes, such as hexagonal shapes), where all cells in the CA grid should be of equal size; (ii) the state of each cell in the CA grid, which is usually represented either by land use or land cover, but can sometimes be used to show spatial distributions of variables to model spatial movement (Mohammad et al., 2013; White & Engelen, 2000; White et al., 1997); (iii) the CA neighborhood space (i.e., the neighborhood effect in urban CA is calculated for each state using the positive and negative effects of each cell in terms of the conversion or non-conversion of the cell to another state via the surrounding cells) (Barredo et al., 2003; White & Engelen, 2000); and (iv) the CA transition rules, wherein the behaviors that occur in the actual world can be understood through the transition rules in the CA models (Mohammad et al., 2013). The state of each cell can be converted to another state using the CA transition rules that can make a CA model more dynamic for simulation (Wu, 1998). The basic expression of a CA model is expressed by Equation (1):

$$S(t, t+1) f(S(t), N) \tag{1}$$

Where, S represents the states of discrete cells, t is the time instant, t +1 is the coming future time instant respectively, N is the cellular field, and f is the transition rule of cellular states in local space.

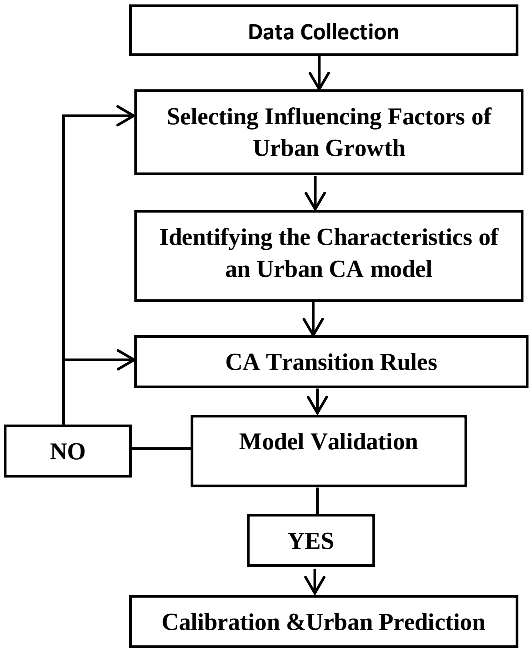

Figure. 3. Flow Chart of the Urban CA Model.

2.2.2. Markov chain model

The Markov chain model is used to predict the status of a cell that is converted to different statuses according to the progression of the formation of Markov stochastic process systems (Muller & Middleton, 1994). This model is commonly used to simulate urban growth because it does not need rich data (Sun et al., 2007). This model is also used to compute the probabilities of transition areas from one land-use status to another (Coppedge et al., 2007). In this study, the urban and non-urban classes were used as input data for the model (Figure 3). Then, the transition area probabilities matrix and the probability map for the specified period time were generated using this model. The prediction of urban growth can be computed according to the conditional probability formula outlined in Equations (2), (3), and (4):

$$S\,(\,t+1\,) = P_{ij}\ \times S(t) \tag{2}$$

$$P_{ij} = \begin{vmatrix} P_{11} & P_{12} & P_{n1} \\ P_{21} & P_{22} & P_{n2} \\ P_{n1} & P_{n2} & P_{nn} \end{vmatrix} \tag{3}$$

$$\left(0 \le P_{ij} < 1\ and\ \sum_{i=1}^{N} P_{ij} = 1, (i, j = 1,2, \ldots\ldots n)\right) \tag{4}$$

Where, $S(t)$ is the state of the system at time, t, $S(t+1)$ is the state of the system at time, $(t+1)$, and P_{ij} is the matrix of transition probability in a state.

2.2.3. CA-Markov chain model

The reliability of urban growth modeling techniques can be improved and developed by combining two or more prediction techniques to integrate the advantages of these models (Yang et al., 2012). It could be argued that the CA-Markov model has been used recently in order to predict dynamic spatial issues such as urban growth and future land-use change (S. Wang et al., 2012). In addition, the integration of CA and Markov chain models is considered appropriate for spatial modeling of urban growth because it capitalizes on the advantages of the Markov chain in predicting urban quantitative change, and the dynamic explicit spatial simulation strength of the CA model (Yang et al., 2012). Thus, the integration of GIS environment and urban growth maps derived from satellite images and remote sensing techniques together with the CA-Markov model will result in the efficient prediction of spatial and temporal urban growth phenomena (Guan et al., 2011; S. Wang et al., 2012).

The CA-Markov model has been applied to simulate and predict future urban growth in Seremban, as shown in the stepwise approach of the CA-Markov model presented in Figure 4. Four main steps have been applied in the CA-Markov chain modeling using ArcGIS 10.3 and IDRISI Selva software. These are outlined below:

1. The urban and non-urban maps are prepared and loaded into the ArcGIS 10.3 software. Land-use maps of 1984, 1990, 2000, and 2010 were reclassified to suit the objective of predicting urban growth in Seremban. All land-use maps were converted from vector to raster format. After that, the raster maps were converted to ASCII file format using conversion tools in the ArcGIS environment. Then, the ASCII files were reclassified and converted to raster format in the IDRISI Selva environment, so they can be used to predict future urban growth.

2. Urban and non-urban land use transition probability matrix and transition rules utilizing the Markov chain model are identified. Based on the previous land class state, the future urban growth change was modeled, i.e., the transition probabilities among urban and non-urban maps from 1990–2000 were applied to predict the changes in 2010 and to calibrate and validate the model. Meanwhile, urban and non-urban maps of 2000 and 2010 were utilized to predict future urban growth in 2020. Additionally, land maps of 2010 and 2020 were used to predict future urban growth in 2030. The transformation rules and the change probability of different land-use layers into other layers are provided by the transition probability matrices while the

quantity of land change (i.e., urban or non-urban lands) into another land layer in the predicted future is reflected by the transition area matrices.

3. The AC filter is determined; the standard 7×7, 5×5, and 3×3 contiguity kernels were designated as the neighborhoods in this study, so as to identify appropriate contiguity filters to predict urban growth. In the end, the contiguity filter 5×5 was selected; this means that each cell center is surrounded by a matrix space of 5×5 cellular kernels to significantly reflect the cellular changes.

4. The number of iterations and starting point of time for the CA are determined. The CA-Markov model was applied, utilizing various iteration numbers starting from 1 to 200 iterations, in order to identify the appropriate iteration number. This study found that the iteration numbers all showed different performances; which means that this study can use certain iteration numbers to perform future predictions.

In this study, the years 1990 and 2000 were taken as starting points to carry out the calibration and validation process using the Kappa index, while the years 2000 and 2010 were used as starting points to predict future urban growth in 2020. Additionally, the years 2010 and 2020 were used as starting points to predict future urban growth in 2030.

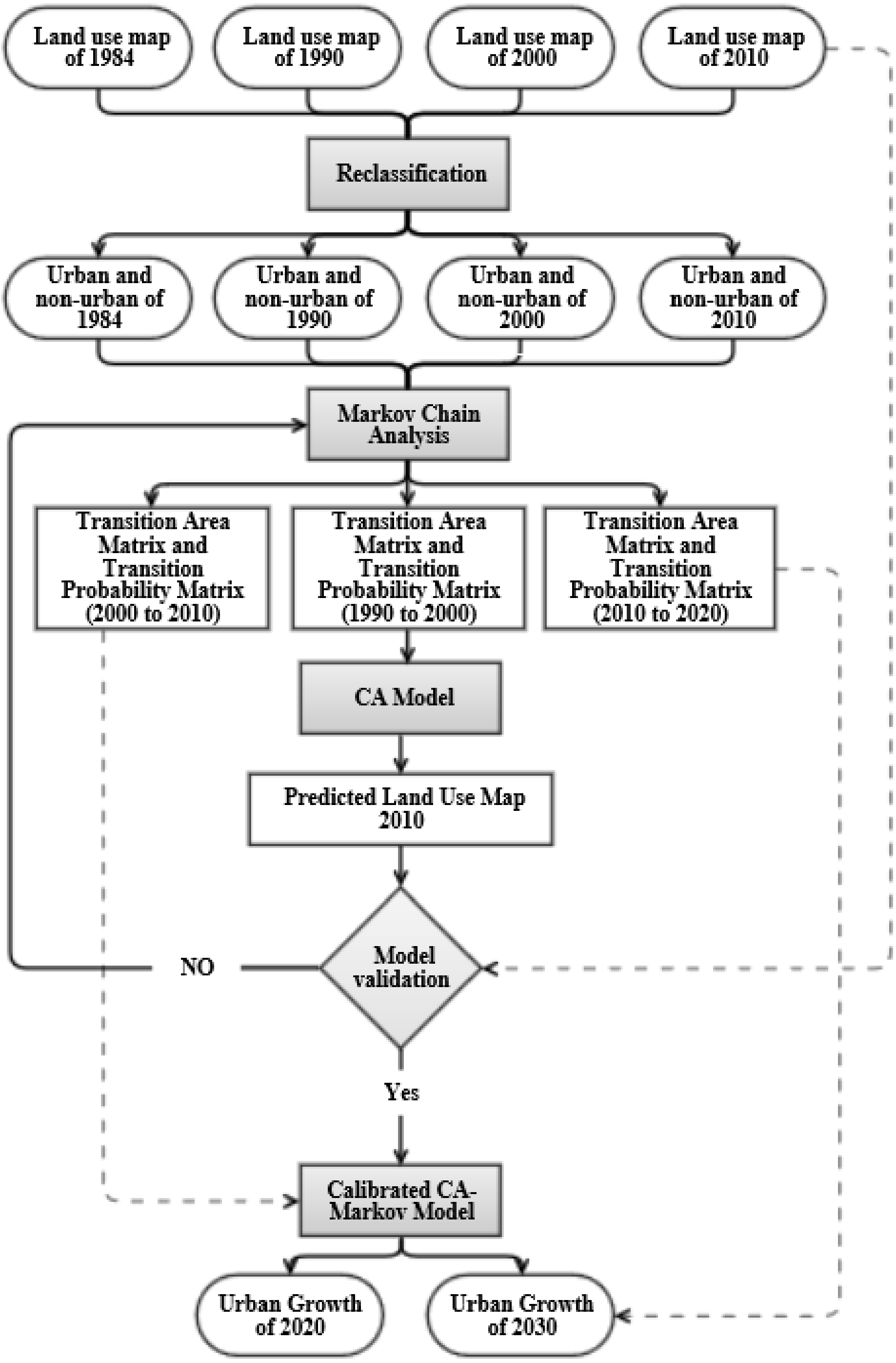

Figure. 4 The stepwise approach of the CA-Markov model.

3. Result and discussion

3.1. The change in urban and non-urban areas

The findings of change in the urban and non-urban areas under study are presented in Table 2 and Figure 5, where the changes of urban and non-urban areas between 1984 and 2010 are shown. From the analysis of these results, the behavior, patterns, and speed of land-use changes can be better understood. The significance of these findings are as follows: (i) these results would be very useful as a scientific basis for planners and decision makers when creating future urban policies; (ii) and will be effective in achieving urban growth sustainability. The results confirm that a major increase in urban growth has occurred in the time period between 1990 and 2000, which equates to 58 km^2 of urban area, due to population and economic growth (Figure 6). In contrast, the total amount of non-urban areas has decreased from 1984 to 2010 by 92 km^2, which is considered to be a significant change in a short period of time. Unfortunately, non-urban areas such as agricultural and forest areas have decreased the most as a result of the urban growth in Seremban. However, this remarkable change in both urban and non-urban areas has led to many question marks about the effectiveness of urban policies, environmental policies, and policies of sustainability implemented in the study area.

Table.2. Amount of urban growth changes observed in sq. km.

	Urban Areas	Non-Urban Areas
1984	34.00	917.87
1990	39.00	912.87
2000	97.00	854.87
2010	126.00	825.87
Annual growth rate (1984–1990)	2.3%	-0.09 %
Annual growth rate (1990–2000)	9.54%	-0.65 %
Annual growth rate (2000–2010)	2.65%	-0.34 %
Total Change sq. km	+ 92	-92

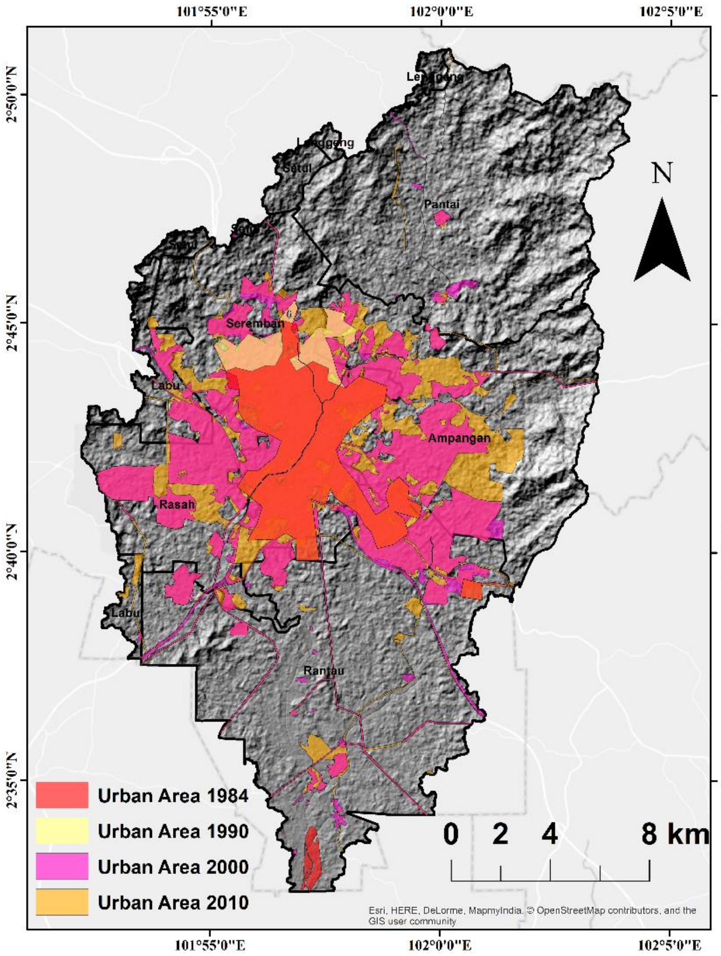

Figure. 5. Urban growth in Seremban River Basin between 1984 and 2010.

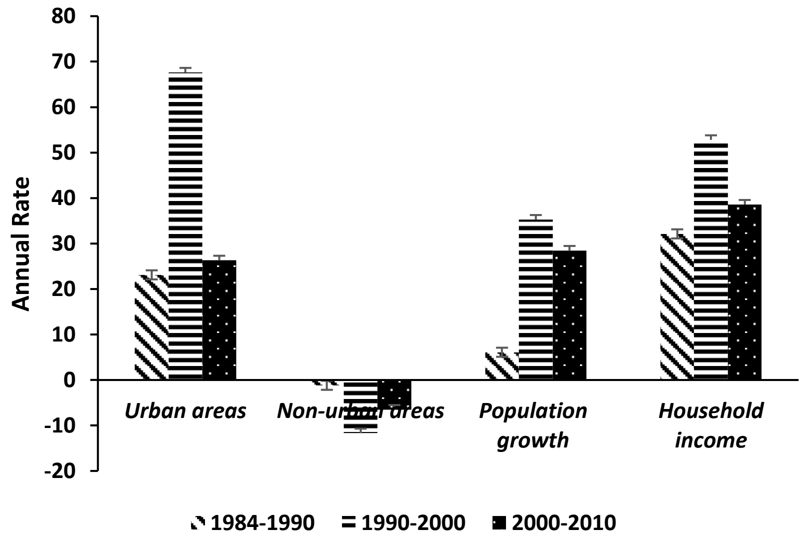

Figure 6. Annual growth rate of urban and non-urban areas, population, and household income in Seremban between 1984 and 2010.

3.2. The transition probability matrices

The Markov chain model was used to calculate the transition probability matrices, as presented in Table 3. In addition, the future potential percentages of change in urban and non-urban land uses in the time periods of 1990-2000, 2000–2010, and 2010–2020 can be ascertained using transition probabilities matrices. Moreover, from further analysis of the results in Table 3, it can be noted that the probability of future transition of non-urban to urban areas from 1990 to 2000 is 25%, while the same probability of transition decreased to 21% from 2000 to 2010. The explanation for this decline is that the urban process in Seremban had decreased between 2000 and 2010 in comparison to 1990 and 2000, which saw a lot of urban development operations, particularly in Seremban and in Malaysia generally (Economic Planning Unit, 2013). However, the probability of the future transition of non-urban to urban areas from 2010 to 2020 is expected to increase to 29%. This high value of transition from non-urban to urban land uses can be seen from the alarming decrease in non-urban areas such as agricultural lands in Seremban. By pondering the findings of the analysis and the classified maps, it can be concluded that Seremban city is facing rapid urban growth, which calls for more action in analyzing and simulating its urban growth patterns.

Table.3. Transition probability matrices for the periods: 1990–2000, 2000–2010, and 2010–2020.

		Urban	Non-urban
1990–2000	Urban	0.6530	0.3470
	Non-urban	0.2553	0.7447
2000–2010	Urban	0.7699	0.2301
	Non-urban	0.2164	0.7836
2010–2020	Urban	0.6653	0.3347
	Non-urban	0.2912	0.7088

3.3. Model validation and prediction of future urban growth

In order to confirm the accuracy of future urban and non-urban land-use predictions in 2010, the CA-Markov model was used. The 1990 and 2000 maps were used to predict land-use state in 2010. After that, the actual 2010 land-use map was compared with the predicted 2010 land-use map to ensure model reliability (Figure 7 and 8). This study used different iteration numbers (i.e., the appropriate iteration numbers) in order to achieve the best performance for the CA-Markov model.

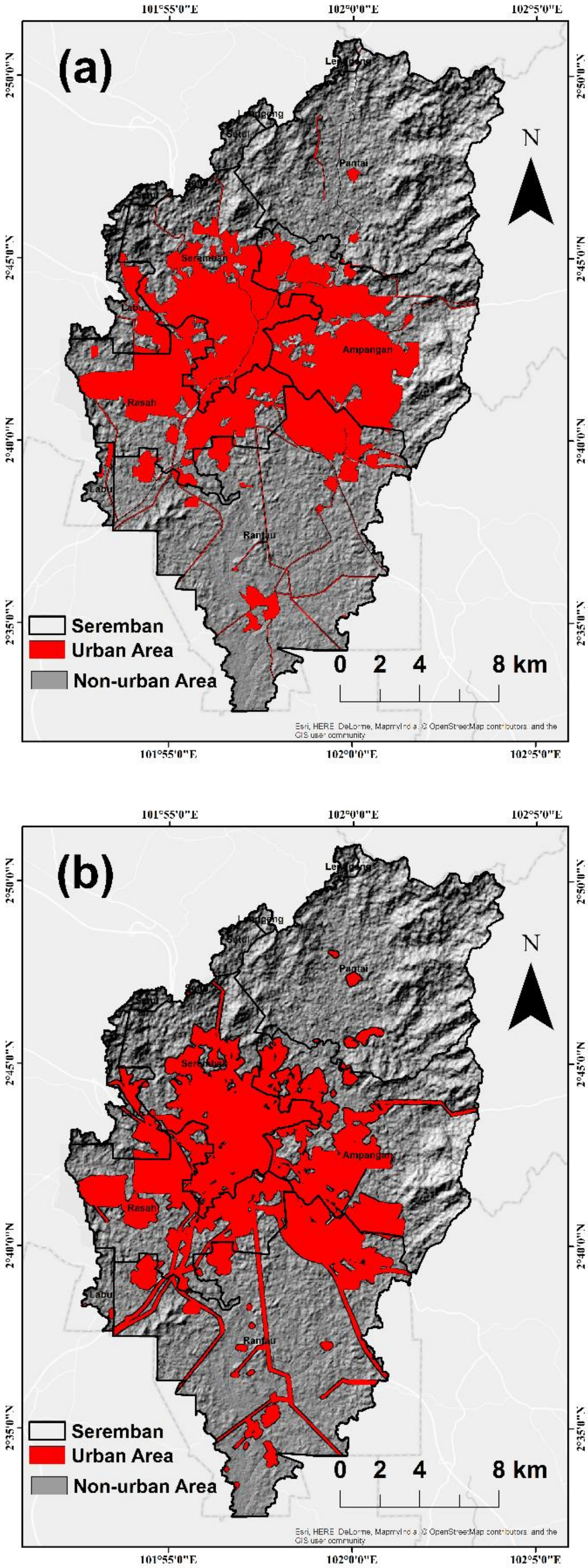

Figure 7. Observed and simulated urban growth in 2010: (a) Observed 2010; and (b) Simulated 2010.

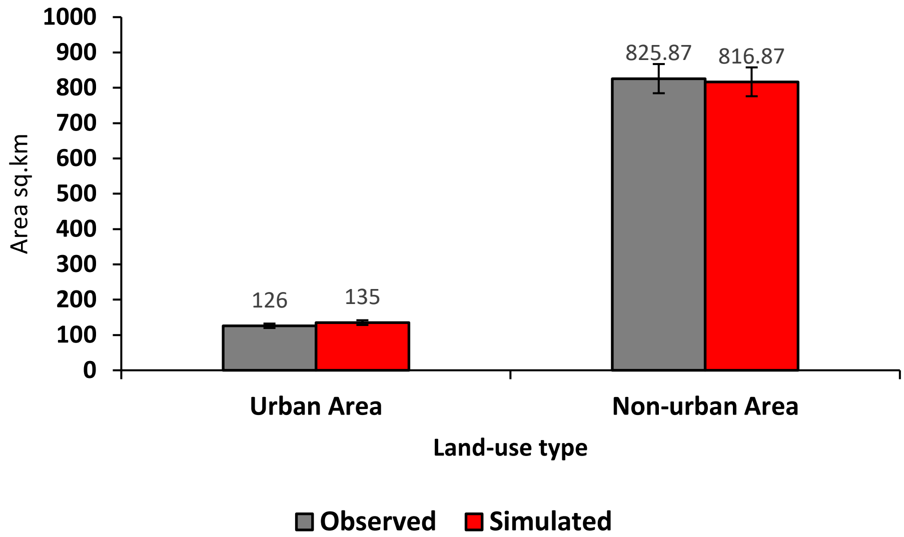

Figure 8. A comparison of urban growth between the Observed and Simulated maps of 2010.

To assess the accuracy of the model, the projected urban and non-urban maps of 2010 were compared with the actual 2010 map using the Kappa index statistic, which will measure its validity in terms of quantity and location (Al-sharif & Pradhan, 2013; Zhang et al., 2011). Figure 9 illustrates the variation of the Kappa coefficient with various iteration numbers from 1 to 200. From Figure 9, it can be observed that, when predicting urban and non-urban areas of 2010, the CA-Markov model performed best at 40 and 60 iterations. High values of the Kappa coefficient were also achieved; (i) Kappa standard index of 0.83; (ii) Kappa location index of 0.86; (iii) and Kappa index no. of 0.83.

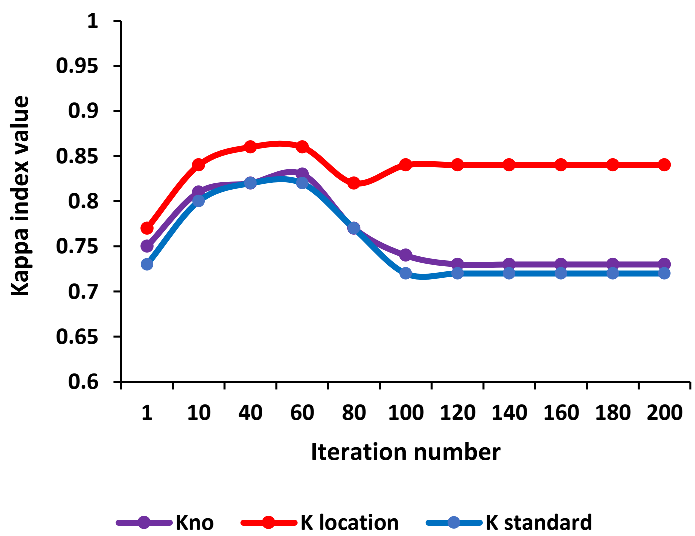

Figure. 9. Kappa index value vs. number of iterations.

From the result of the model's accuracy assessment, a strong agreement between the actual and projected urban and non-urban land-use maps can be observed. From the validation phase, the optimal transition rules for the model were computed using the appropriate iteration numbers (i.e., 40 and 60). After that, these iteration numbers were used to predict land use in 2020 and 2030. According to the successful model validation, the future urban and non-urban land-use maps of 2020 and 2030 were generated using the actual map of 2010 and projected map of 2020, respectively. By using the 2010 and 2020 urban and non-urban land uses as base maps, potential transition maps and transition area matrices of 2002–2010 and 2010-2020 as well as the future of urban growth patterns can be predicted, as presented in Figure. 10.

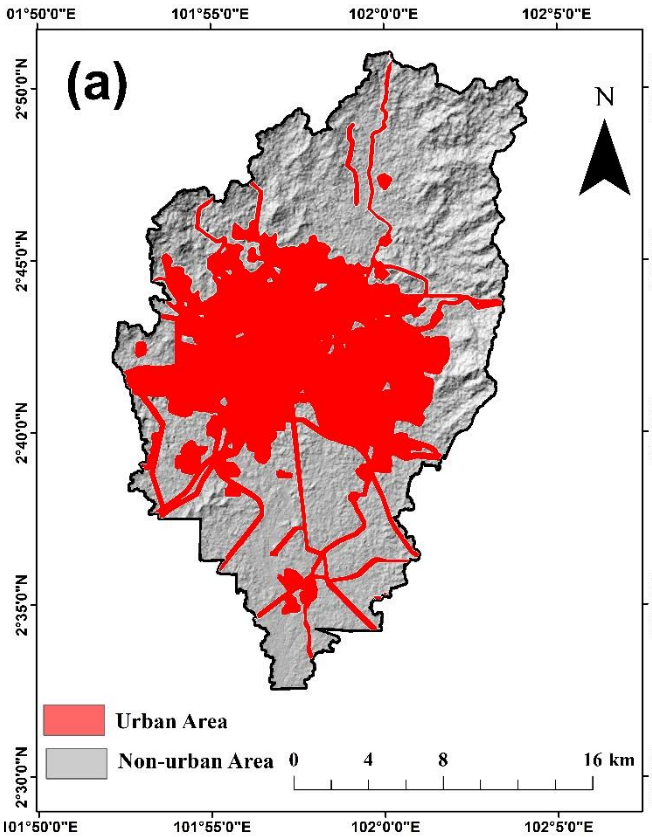

Figure 10. Predicted Maps of Urban Growth in Seremban River Basin: (a) 2020; (b) 2030.

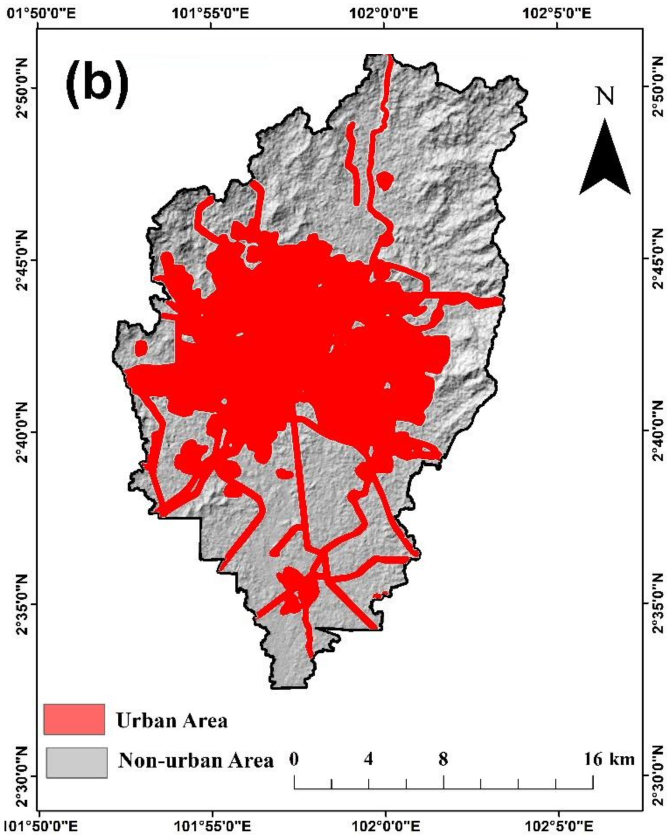

Figure.10. (Continued) Predicted Maps of Urban Growth in Seremban River Basin: (a) 2020; (b) 2030.

The CA-Markov chain model predicted that urban areas in Seremban would increase to 177 km^2 and 195.5 km^2 in 2020 and 2030, respectively (Figure 11). On the other hand, non-urban areas such as agricultural, forest, open, and rural lands, as well as surface water will decrease by 774.87 km^2 and 756.37 km^2 in 2020 and 2030, respectively. Unfortunately, this change will affect the ecosystem and land-use sustainability in Seremban, and cause uncontrolled urban growth.

Generally, it is important to note that the CA-Markov model applied in this study is capable of predicting future urban growth trends using only land-use maps (i.e., it can be used with limited data and still give impactful findings). However, several driving forces also affect urban growth. These forces include physical forces (i.e., slope, elevation, etc.), environmental forces (i.e., land use and cover), socio-economic forces (i.e., population growth, household income, etc.), and infrastructural issues (i.e., road and railway networks, etc.). Accordingly, both the driving forces and their factors can be used for predicting future urban growth rather than relying on land-use maps only. Therefore, incorporating these driving forces within the CA-Markov environment will enhance the simulation and prediction capability of the model.

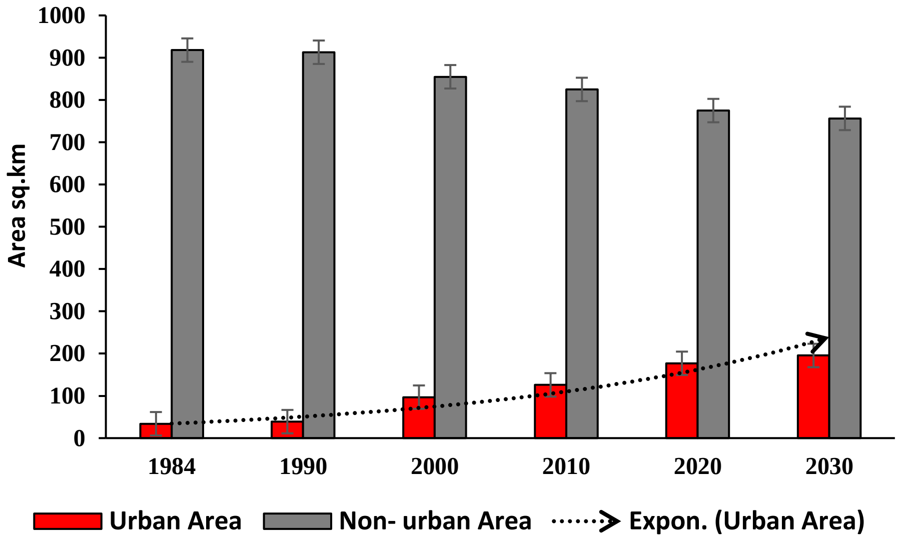

Figure. 11. Quantity of previous and predicted urban and non-urban areas in sq. km.

4. Conclusion

By using multiple classified and unclassified land-use maps, together with the integrated CA-Markov chain model (a combination of the CA and Markov chain models) the urban growth patterns in Seremban, Malaysia, was simulated and predicted excellently. The model achieved 83% accuracy in simulating projected urban and non-urban land-use maps, which is a reflection of the model's success in predicting urban growth patterns. One of the significant advantages of using the CA-Markov chain model is that the prediction of urban growth patterns can be done using limited data (i.e., it requires at least two land-use maps in different time periods). However, it can also be said that there are some limitations when it comes to using the integrated model such as its inability to apply urban growth driving forces such as physical and socio-economic forces in the prediction process. These forces are highly significant to the monitoring and controlling of current processes of urban growth and the preparation of wise policies and plans for future requirements.

The urban and non-urban land-use change analysis has shown that there is a high, continuous decline in non-urban lands in Seremban. This continuous reduction has affected its agricultural, forest, rural, and open lands. On the other hand, the prediction analysis of 2020 and 2030 using the CA-Markov chain model demonstrated that urban areas will continue to increase, which will threaten the arable lands in Seremban in the long term. Moreover, according to simulation findings, the urban sprawl in Seremban will be in disaggregation mode. Thus, the urban growth scenario will become worse in the future. Subsequently, it is important to save and protect the non-urban areas in order to achieve urban sustainability.

Finally, this study shows the significance of using the integrated CA-Markov chain model, which plays an important role in modeling urban growth, especially in developing countries, which have different urban features. However, it is important to assert that the urban growth driving forces should be applied in the prediction process of the CA-Markov chain model in order to obtain a better understanding of the change in urban growth patterns. For this purpose, the CA-Markov chain model should be integrated with other models such as the Analytic Hierarchy Process (AHP), Frequency ratio (FR), and logistic regression (LR) models in order to further improve its capability.

Acknowledgements The study presented here is part of a research project funded by Universiti Sains Malaysia (USM).

5. References

Aburas, M. M., Ho, Y. M., Ramli, M. F., & Ash'aari, Z. H. (2016). The simulation and prediction of spatio-temporal urban growth trends using cellular automata models: A review. *International Journal of Applied Earth Observation and Geoinformation, 52*, 380-389.

Aburas, M. M., Ho, Y. M., Ramli, M. F., & Ash'aari, Z. H. (2017). Improving the capability of an integrated CA-Markov model to simulate spatio-temporal urban growth trends using an analytical hierarchy process and frequency ratio. *International Journal of Applied Earth Observation and Geoinformation, 59*, 65-78.

Al-sharif, A. A., & Pradhan, B. (2013). Monitoring and predicting land use change in Tripoli Metropolitan City using an integrated Markov chain and cellular automata models in GIS. *Arabian Journal of Geosciences*, 1-11.

Alsharif, A. A., & Pradhan, B. (2014). Urban sprawl analysis of Tripoli Metropolitan city (Libya) using remote sensing data and multivariate logistic regression model. *Journal of the Indian Society of Remote Sensing, 42*(1), 149-163.

Anderson, J. R. (1976). *A land use and land cover classification system for use with remote sensor data* (Vol. 964): US Government Printing Office.

Barredo, J. I., Kasanko, M., McCormick, N., & Lavalle, C. (2003). Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. *Landscape and Urban Planning, 64*(3), 145-160.

Batty, M., & Xie, Y. (1994). Research Article. Modelling inside GIS: Part 1. Model structures, exploratory spatial data analysis and aggregation. *International Journal of Geographical Information Systems, 8*(3), 291-307.

Bihamta, N., Soffianian, A., Fakheran, S., & Gholamalifard, M. (2014). Using the SLEUTH Urban Growth Model to Simulate Future Urban Expansion of the Isfahan Metropolitan Area, Iran. *Journal of the Indian Society of Remote Sensing*, 1-8.

Celio, E., Koellner, T., & Grêt-Regamey, A. (2014). Modeling land use decisions with Bayesian networks: Spatially explicit analysis of driving forces on land use change. *Environmental Modelling & Software, 52*, 222-233.

Clarke. (1997). A self-modifying cellular automaton model of historical. *Environ Plan B, 24*, 247-261.

Coppedge, B. R., Engle, D. M., & Fuhlendorf, S. D. (2007). Markov models of land cover dynamics in a southern Great Plains grassland region. *Landscape Ecology, 22*(9), 1383-1393.

Couclelis, H. (1985). Cellular worlds: a framework for modeling micro-macro dynamics. *Environment and planning A, 17*(5), 585-596.

Couclelis. (1997). From cellular automata to urban models: new principles for model development and implementation. *Environment and Planning B: Planning and Design, 24*(2), 165-174.

DOSM. (2011). Statistics Yearbook. *Department of Statistics Malaysia*, 367.

Doygun, H. (2009). Effects of urban sprawl on agricultural land: a case study of Kahramanmaraş, Turkey. *Environmental Monitoring and Assessment, 158*(1-4), 471-478.

Economic Planning Unit, M. (2013). Economic History of Malaysia. *Economic Planning Unit, Malaysia*, 4,5.

Epsteln, J., Payne, K., & Kramer, E. (2002). Techniques for mapping suburban sprawl. *Photogrammetric engineering & remote sensing, 63*(9), 913-918.

Guan, D., Li, H., Inohae, T., Su, W., Nagaie, T., & Hokao, K. (2011). Modeling urban land use change by the integration of cellular automaton and Markov model. *Ecological Modelling, 222*(20), 3761-3772.

Inouye, C. E. N., de Sousa, W. C., de Freitas, D. M., & Simões, E. (2015). Modelling the spatial dynamics of urban growth and land use changes in the north coast of São Paulo, Brazil. *Ocean & Coastal Management*.

Mohammad, M., Sahebgharani, A., & Malekipour, E. (2013). Urban Growth Simulation Through Cellular Automata (CA), Anaiytic Hierarchy Process (AHP) and GIS; Case Study of 8th and 12th Municipal Districts of Isfahan.

Muller, M. R., & Middleton, J. (1994). A Markov model of land-use change dynamics in the Niagara Region, Ontario, Canada. *Landscape Ecology, 9*(2), 151-157.

Pijanowski, B. C., Brown, D. G., Shellito, B. A., & Manik, G. A. (2002). Using neural networks and GIS to forecast land use changes: a land transformation model. *Computers, Environment and Urban Systems, 26*(6), 553-575.

Punia, M., & Singh, L. (2012). Entropy approach for assessment of urban growth: a case study of Jaipur, India. *Journal of the Indian Society of Remote Sensing, 40*(2), 231-244.

Ren, P., Gan, S., Yuan, X., Zong, H., & Xie, X. (2013). Spatial Expansion and Sprawl Quantitative Analysis of Mountain City Built-Up Area *Geo-Informatics in Resource Management and Sustainable Ecosystem* (pp. 166-176): Springer.

Santé, I., García, A. M., Miranda, D., & Crecente, R. (2010). Cellular automata models for the simulation of real-world urban processes: a review and analysis. *Landscape and urban planning, 96*(2), 108-122.

Sudhira, H., Ramachandra, T., & Jagadish, K. (2004). Urban sprawl: metrics, dynamics and modelling using GIS. *International Journal of Applied Earth Observation and Geoinformation, 5*(1), 29-39.

Sun, H., Forsythe, W., & Waters, N. (2007). Modeling urban land use change and urban sprawl: Calgary, Alberta, Canada. *Networks and spatial economics, 7*(4), 353-376.

Tran, T. V. (2008). Research on the effect of urban expansion on agricultural land in Ho Chi Minh City by using remote sensing method.

Verburg, P. H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., & Mastura, S. S. (2002). Modeling the spatial dynamics of regional land use: the CLUE-S model. *Environmental Management, 30*(3), 391-405.

Wagner, D. F. (1997). Cellular automata and geographic information systems. *Environment and planning B, 24*, 219-234.

Wang, J., & Maduako, I. N. (2018). Spatio-temporal urban growth dynamics of Lagos Metropolitan Region of Nigeria based on Hybrid methods for LULC modeling and prediction. *European Journal of Remote Sensing, 51*(1), 251-265.

Wang, S., Zheng, X., & Zang, X. (2012). Accuracy assessments of land use change simulation based on Markov-cellular automata model. *Procedia Environmental Sciences, 13*, 1238-1245.

White, R., & Engelen, G. (1994). Cellular dynamics and GIS: modelling spatial complexity. *Geographical systems, 1*(3), 237-253.

White, R., & Engelen, G. (2000). High-resolution integrated modelling of the spatial dynamics of urban and regional systems. *Computers, Environment and Urban Systems, 24*(5), 383-400.

White, R., Engelen, G., & Uljee, I. (1997). The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. *Environment and planning B, 24*, 323-344.

White, R., Engelen, G., Uljee, I., Lavalle, C., & Ehrlich, D. (2000). *Developing an urban land use simulator for European cities.* Paper presented at the Proceedings of the Fifth EC GIS Workshop: GIS of Tomorrow. European Commission Joint Research Centre: S.

Wu, F. (1998). SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. *International Journal of Geographical Information Science, 12*(1), 63-82.

Yang, X., Zheng, X.-Q., & Lv, L.-N. (2012). A spatiotemporal model of land use change based on ant colony optimization, Markov chain and cellular automata. *Ecological Modelling, 233*, 11-19.

Yeh, A. G.-O., & Li, X. (2001). Measurement and monitoring of urban sprawl in a rapidly growing region using entropy. *Photogrammetric engineering and remote sensing, 67*(1).

Zhang, Q., Ban, Y., Liu, J., & Hu, Y. (2011). Simulation and analysis of urban growth scenarios for the Greater Shanghai Area, China. *Computers, Environment and Urban Systems, 35*(2), 126-139.

Cover letter

[MAHER MILAD MOHAMMED ABURAS]
[Universiti Sains Malaysia]
[Penang, Malaysia]

[25-11-2018]

Dear Prof.Abdullah M. Al-Amri,

We wish to submit an original research article entitled "[Spatio-temporal simulation of future urban growth trends using an integrated CA-Markov model]" for consideration by Arabian Journal of Geosciences.
We confirm that this work is original and has not been published elsewhere, nor is it currently under consideration for publication elsewhere.

In this paper, we show that the integration between CA and MC models obtained high value of simulation accuracy which means that CA-MC model is very suitable for simulating and predicting future urban growth.
We believe that this manuscript is appropriate for publication by of Geosciences because it is very appropriate for aims of the journal.

This study used an integrated Markov chain and Cellular Automata model (CA-Markov) to simulate rapid urban growth in Seremban City from 1990 to 2010, and then to predict future land changes quantitatively and spatially. To the authors' best knowledge, no study of this kind has ever been done in this city before.

We have no conflicts of interest to disclose.
Thank you for your consideration of this manuscript.

Sincerely,

[Dr. MAHER MILAD MOHAMMED ABURAS]

Universiti Sains Malaysia (USM)