ORIGINAL PAPER

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

Maher Milad Aburas[1,2] · Yuek Ming Ho[3] · Biswajeet Pradhan[4] · Abdul Hakim Salleh[5] · Motasem Y. D. Alazaiza[6]

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. For accuracy assessment, the CA-Markov model was validated using a kappa coefficient. An 83% overall accuracy was observed for the kappa index statistics, which indicates the excellent performance of the proposed model. Finally, based on the CA transition rules and the transition area matrix produced from the Markov chain model-based calibration process, the future urban growth in Seremban for 2020 and 2030 was simulated.

Keywords Urban growth · Markov chain · Cellular automata · prediction · Modeling

Introduction

Urban development has become a global issue, resulting in a rising concern among planners and decision-makers over the

✉ Maher Milad Aburas
    boras222@yahoo.com

1   The High Institute for Engineering Professions ALmajurie, Benghazi, Libya

2   Faculty of Engineering, Department of Surveying, Benghazi Private University, Tabalino, Benghazi, Libya

3   Department of Environmental Management, Faculty of Forestry and Environment, Universiti Putra Malaysia, 43400 Serdang, Selangor Darul Ehsan, Malaysia

4   Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW 2007, Australia

5   School of Civil Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia

6   Department of Civil and Environmental Engineering, College of Engineering (COE), A'Sharqiyah University, 400 Ibra, Oman

future impacts of urban development on the ecosystem (Bihamta et al. 2014). The simulation and prediction of urban sprawl patterns have 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 correctly simulate urban dynamics (Barredo et al. 2003). In urban growth simulation, the sprawl chronology and significant historical information must be considered to accurately determine spatial and temporal relationships (Sudhira et al. 2004). In this case, the process of obtaining actual growth factors that affect future land uses can be improved using simulation techniques (Pijanowski et al. 2002). Understanding spatial and temporal changes, as well as all effective elements, can be facilitated using remote sensing (RS) and geographic information system (GIS) techniques (Punia and 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 proven to be effective tools for helping planners and decision-makers formulate sustainable policies. These modern techniques have several advantages. First, these tools come at a low cost (Yeh and Li 2001) and offer effective visual interpretation (Epsteln et al. 2002). The tools also have updatable spatial and temporal

databases (Punia and Singh 2012). Both are effective monitoring and controlling tools (Tran 2008; Doygun 2009) and are accurate tools for evaluating, analyzing, and simulating spatial phenomena (Ren et al. 2013). Therefore, 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 to model general urban growth patterns and simulate land-use changes (Mohammad et al. 2013). Some studies have used traditional models that depend on the dynamic growth assessment of urban areas, such as CA models (Aburas et al. 2017). Others have also relied on quantitative models, such as logistic regression (LR), for simulation and prediction (Alsharif and Pradhan 2014). Still, others 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 (Alsharif and Pradhan 2013). The modeling of urban growth patterns based on RS and GIS techniques is done to understand the spatial process of urban movement within a specific time to facilitate the development of future policies for sustainable development (Wang and Maduako 2018).

The GIS and RS-based modeling of 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 facing environmental change and urban sprawl, as well as potential critical sites, can be identified using several models, such as quantitative or spatio-temporal models (Verburg et al. 2002). Spatial modeling is used to simulate land-use patterns that are indispensable in supporting the development and implementation of urban planning policies (Inouye et al. 2015). In general, planners and policymakers 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 the spatial and temporal dimensions of several processes, as well as its capability to model complex dynamic systems, are major reasons for its widespread application in urban growth pattern simulation 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 (!!! INVALID CITATION !!! (Couclelis 1985; Batty and Longley 1994; White and 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 high computational efficiency (Santé et al. 2010). The other key fields of urban CA models that are considered powerful spatial dynamic modeling techniques representing the 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) dynamic modeling, and (v) simplicity and visualization (!!! INVALID CITATION !!! (Batty and Longley 1994; White and Engelen 1994; Clarke 1997; Wu 1998; White and Engelen 2000)).

The Markov chain is usually used to model and predict the changes, dimensions, and trends in urban growth patterns (Aburas et al. 2017). The Markov chain model can be used to analyze and summarize the changes in urban and non-urban lands based on the number of transition area probabilities from one status to various other statuses over a certain period (Coppedge et al. 2007). The Markov chain model cannot simulate changes in spatial trends. However, it is still a powerful model, with the capability to predict the extent to which land has changed (Yang et al. 2012). The integration between the CA and Markov Chain models is effective for estimating quantities and for modeling spatio-temporal dynamics because both are GIS and RS models that can be proficiently incorporated (Al-sharif and 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, i.e., the difficulty in dynamically or statistically simulating urban issues. In other words, the one complements 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 (CA-Markov) model to simulate the rapid urban growth in Seremban City from 1990 to 2010. Following that, the future land changes were quantitatively and spatially predicted. To the best of the authors' knowledge, this is the first study to be done in this city. Additionally, this study combined the cellular automata (CA) model with the Markov chain model to improve the simulation process.

Methodology

Study area

Seremban River Basin is the largest district in the Negeri Sembilan State (Fig. 1). Seremban is also the capital of

Fig. 1 Location of the study area
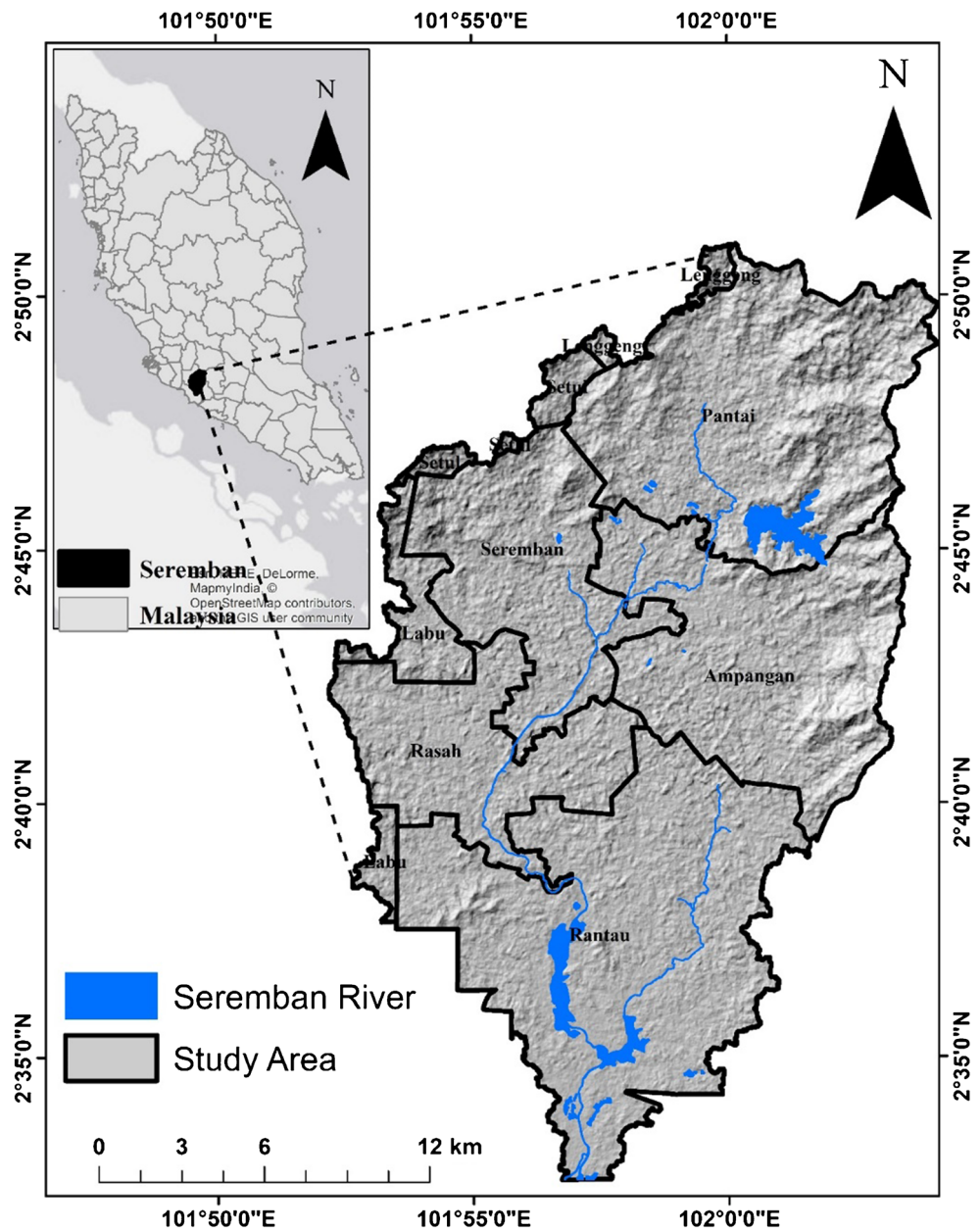
Negeri Sembilan. It occupies a total land area of approximately 935.78 sq. km and includes the districts of Seremban Town, Setul, Labu, Rasah, Ampangan, Rantau, Pantai, and Lenggeng. Seremban is located approximately 20 km from Putrajaya, the national capital of Malaysia, and 67 km from Kuala Lumpur, the economic center of Malaysia. Seremban has a population of more than 500,000 people, with the number expected to increase to 1,000,000 people in 2020 (DOSM 2011). Seremban city was selected as the study area because (i) it is the biggest city in Negeri Sembilan; (ii) it is the economic center of Negeri Sembilan; (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 a future center for urban development.

Data and methods

This study used land-use maps for the years 1984, 1990, 2000, and 2010 obtained from the Department of Agriculture of Malaysia (Fig. 2). These land-use maps were extracted from SPOT 2, 4, and 5 images, with 10-m and 2.5-m spatial resolution SPOT 2.4 and SPOT 5 images, 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
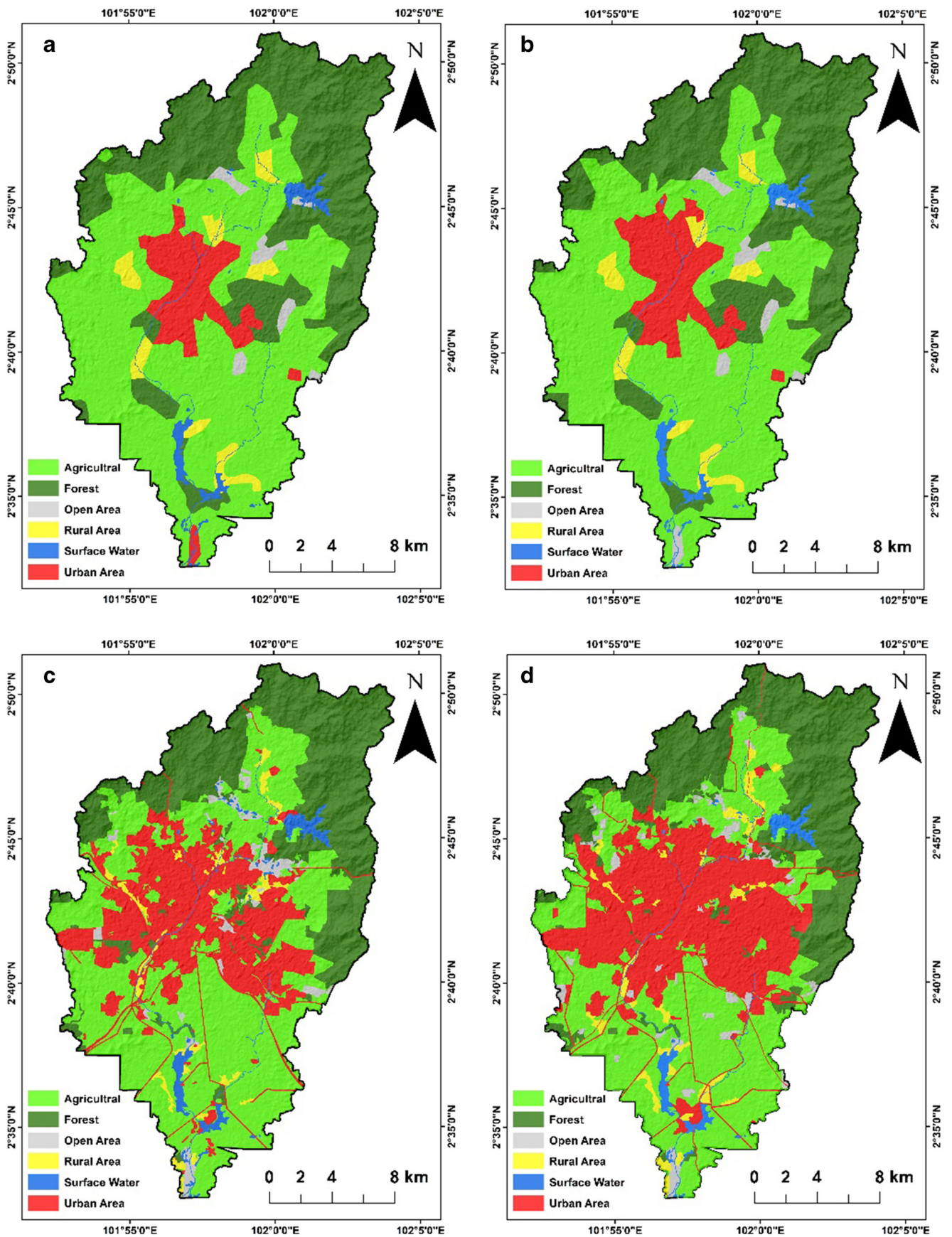
Fig. 2 Land-use maps of Seremban River Basin: **a** 1984, **b** 1990, **c** 2000, and **d** 2010

each type of land use. The accuracy assessment showed acceptable kappa index values, indicating that the image classification is acceptable. Based on the Anderson scheme, an acceptable, accurate kappa index value 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 each district in Seremban (Table 1).

Land-use maps were reclassified into two types of land use, urban areas and non-urban areas, to comply with the general objective of the study. Urban areas were defined as 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. In contrast, non-urban areas were defined as 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 urban growth patterns. The models used to predict urban growth in Seremban are discussed in more detail below:

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 whether or not the model is integrated with other models (Aburas et al. 2016); (ii) the selected 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 the cell state, identifying the neighborhood properties, and identifying the transition rules that will be used (!!! INVALID CITATION !!! (Clarke 1997; White and Engelen 2000; White et al. 2000)); and (iii) validation and calibration of the model using an actual land-use model against the kappa index (Al-sharif and Pradhan 2013; Mohammad et al. 2013). Subsequently, the simulation and prediction of future land use were undertaken (Fig. 3).
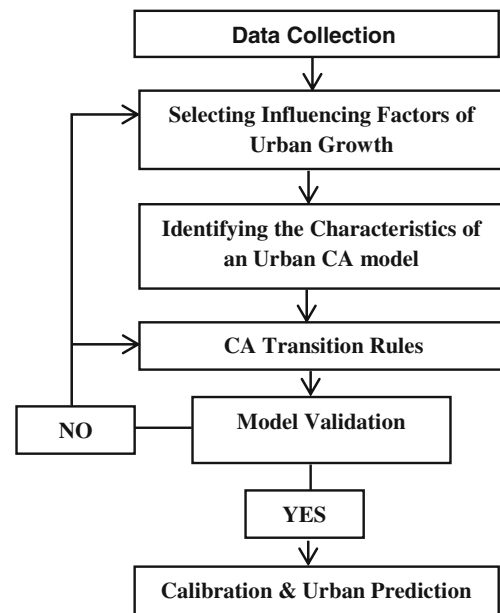
Fig. 3 Flow chart of the urban CA model

CA models use a simple mechanism to identify the future conditions of cells, defined by identifying the actual condition of 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 (!!! INVALID CITATION !!! (White et al. 1997; White and Engelen 2000; Mohammad et al. 2013)); (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) (White and Engelen 2000; Barredo et al. 2003); 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 result in a more

Table 1 Data used in the study

Material	Source	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

dynamic CA simulation model (Wu 1998). The basic 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, N is the cellular field, and f is the transition rule of cellular states in the local space.

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 and 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 (Fig. 3). Then, the transition area probability matrix and the probability map for the specified period 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):

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

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

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

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.

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 to predict dynamic spatial issues, such as urban growth and future land-use change (Wang et al. 2012). In addition, the integration of CA and Markov chain models is considered appropriate for the spatial modeling of urban growth because it capitalizes on the advantages of the Markov chain in predicting urban quantitative change, as well as 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 RS techniques, together with the CA-Markov model, will result in the efficient prediction of spatial and temporal urban growth phenomena (Guan et al. 2011, Wang et al. 2012).

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

1. The urban and non-urban maps were 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 were 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 to 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 used to predict future urban growth in 2020. Additionally, land maps of 2010 and 2020 were used to predict future urban growth in 2030. The transition probability matrices provided the transformation rules and the change probability of different land-use layers into other layers. In contrast, the quantity of land change (i.e., urban or non-urban lands) into another land layer in the predicted future was reflected by the transition area matrices.

3. The AC filter was determined; standard 7×7, 5×5, and 3×3 contiguity kernels were designated as the neighborhoods in this study to identify the appropriate contiguity filters for predicting urban growth. In the end, the 5×5 contiguity filter was selected, meaning 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 the starting point for the CA were determined. The CA-Markov model was applied, utilizing various iteration numbers from 1 iteration to 200 iterations, to identify the appropriate iteration number. This study found that the iteration numbers all produced different performances, meaning that this study could use certain iteration numbers to yield reliable future predictions.

In this study, the years 1990 and 2000 were taken as starting points in the calibration and validation process, which was done using the Kappa index. Meanwhile, the years 2000 and 2010 were used as starting points to predict future urban

Fig. 4 The stepwise approach of
the CA-Markov model
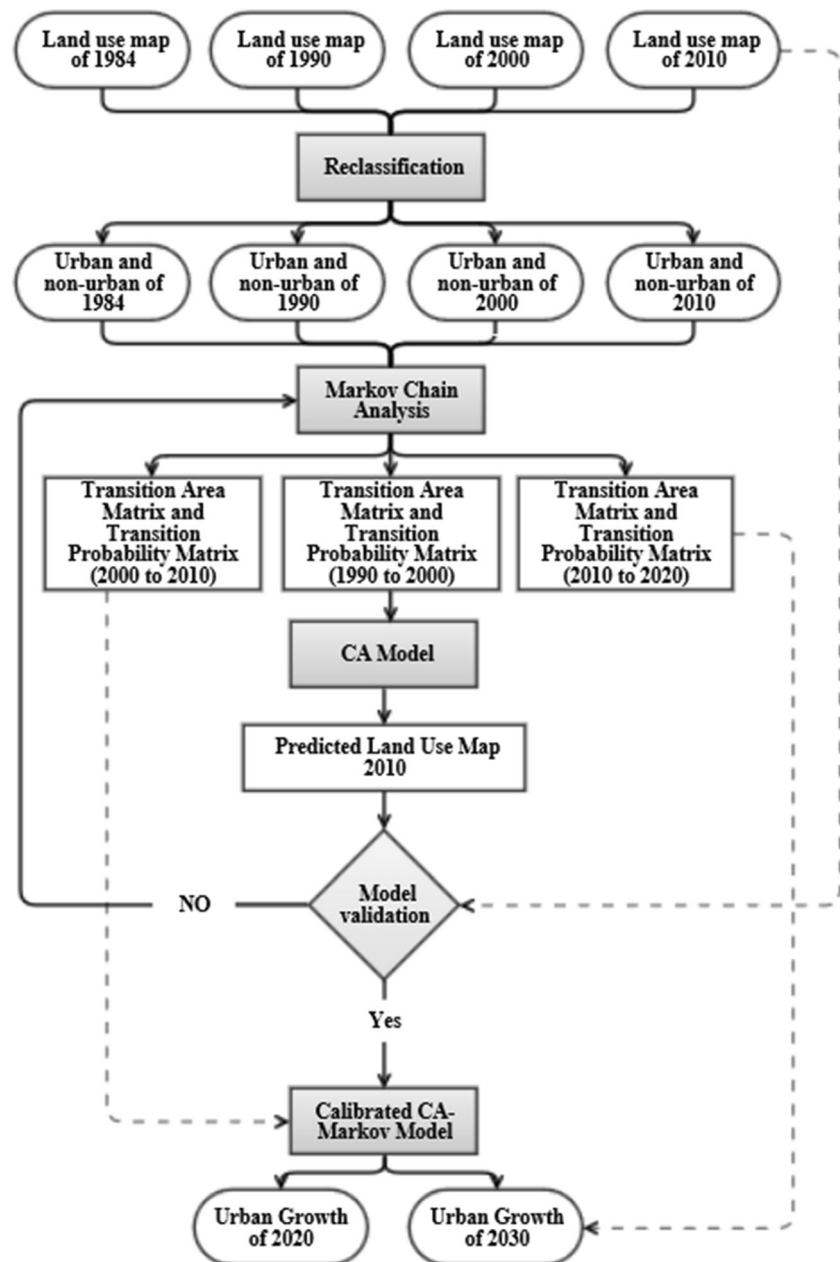
growth in 2020. Additionally, the years 2010 and 2020 were used as starting points to predict future urban growth in 2030.

Results and discussion

The change in urban and non-urban areas

The findings on the change in urban and non-urban areas in the study area are presented in Table 2 and Fig. 5. The results show the changes in urban and non-urban areas between 1984 and 2010. From the analysis of the results, the behavior, patterns, and speed of land-use changes can be better understood.

The significance of these findings is as follows: (i) these results would be very useful as a scientific basis for planners and decision-makers when creating future urban policies and (ii) these results will also be effective for achieving urban growth sustainability. The results confirm that a major increase in urban growth has occurred in the period between 1990 and 2000, which equates to 58 km^2 of urban area, due to population and economic growth. In contrast, the total amount of non-urban areas decreased from 1984 to 2010 by 92 km^2, considered to be a significant change in such a short period. 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

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

	Urban area	Non-urban area
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

and non-urban areas has led to many questions regarding the effectiveness of urban policies, environmental policies, and sustainability policies implemented in the study area.

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 use in the periods of 1990-2000, 2000–2010, and 2010–2020 can be ascertained using transition probability matrices. Moreover, the results in Table 3 showed a 25% future

Fig. 5 Urban growth in Seremban River Basin between 1984 and 2010
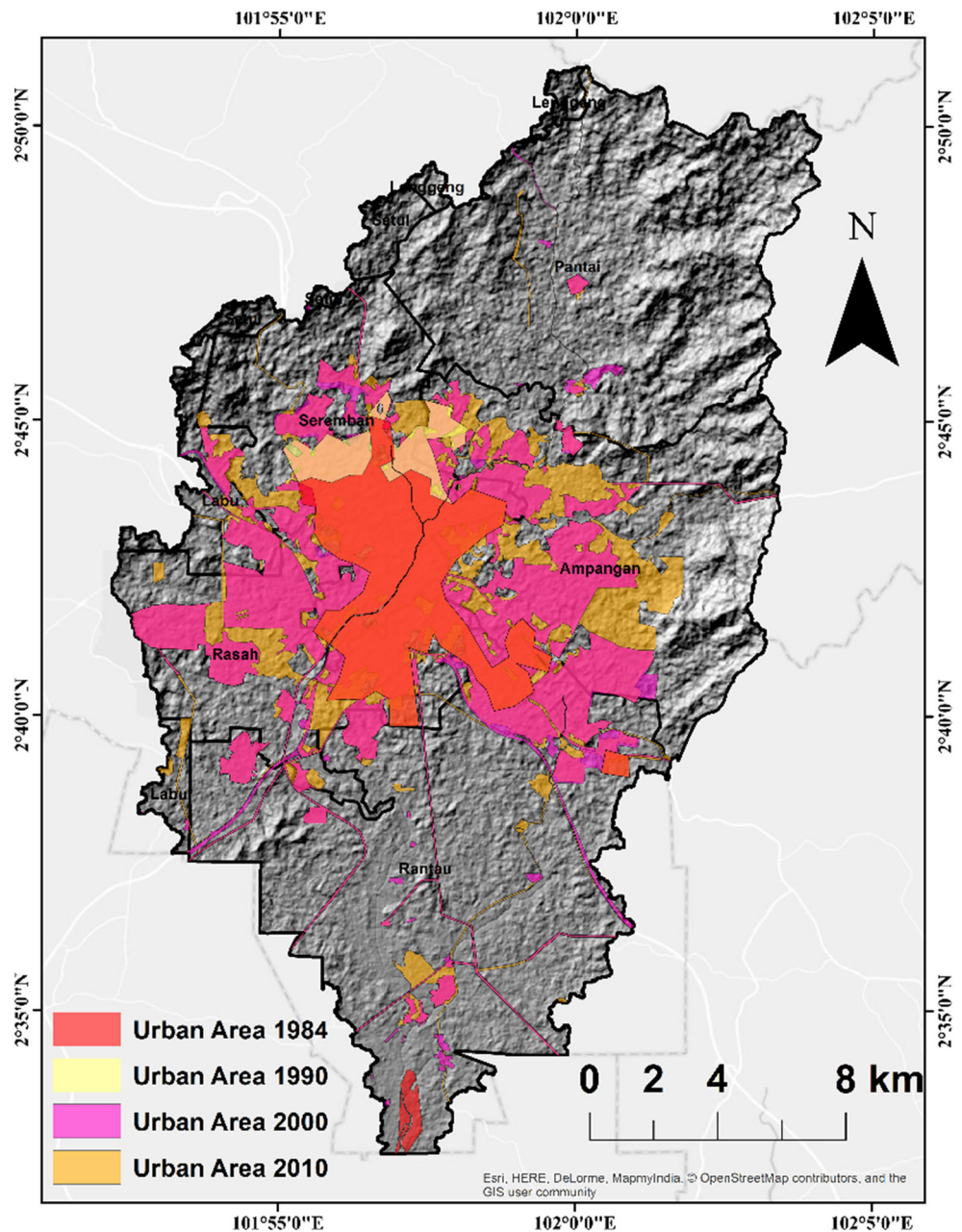
Table 3 Transition probability matrices for the periods 1990–2000, 2000–2010, and 2010–2020

		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

Fig. 6 A comparison of urban growth between the observed and simulated maps of 2010.

transition probability of non-urban to urban areas from 1990 to 2000, while 2000 to 2010 saw a decrease in the same transition probability to 21%. One explanation for this decline is that the urban process in Seremban, had decreased between 2000 and 2010 in comparison with 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 transition value 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. Therefore, more action is needed to analyze and simulate the urban growth patterns in this city.

Model validation and prediction of future urban growth

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 (Fig. 6). This study used different iteration numbers (i.e., the appropriate iteration numbers) to achieve the best performance for the CA-Markov model.

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 measures quantity and location validity (Al-sharif and Pradhan 2013; Zhang et al. 2011). Figure 7 illustrates the variation in the kappa coefficient with various iteration numbers from 1 to 200. From Fig. 7, it can be observed that, when predicting the urban and non-urban areas of 2010, the CA-Markov model performed best at 40 and 60 iterations. A high kappa coefficient value was also achieved with these iterations, namely, (i)
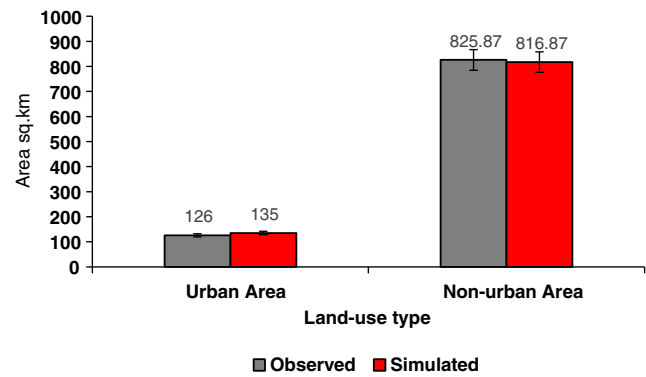
a kappa standard index of 0.83, (ii) a kappa location index of 0.86, and (iii) and a kappa index no. of 0.83.

From 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 the 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 Fig. 8.

From the accuracy results, it can be concluded that the CA-Markov model used in this study obtained acceptable and reliable values compared with other past models used in this field. For example, ANN is one of the most used models to predict urban growth. The main advantage of this model is that it can be integrated with dynamic models, such as CA models, which have achieved acceptable simulation accuracy with a kappa index of 0.72 (Maithani 2009). Ballestores Jr and Qiu (2012) concluded that the decision tree (DT) is one of the most effective machine learning models, achieving a 0.84 simulation accuracy. Meanwhile, Pijanowski et al. (2002) used the LTM model and recorded a more than 0.65 simulation accuracy. Alsharif and Pradhan (2014) obtained 0.86 simulation accuracy by using the logistic regression (LR) model. Al-sharif and Pradhan (2015) used evidential belief functions (EBF) and frequency ratio (FR) models. The validation results indicated 83% simulation accuracy for the EBF model and 84% for the FR model. Abdullahi et al. (2015) confirmed that the weights-of-evidence (WoE) model was successfully used to predict future urban growth. The validation results indicated 75% simulation accuracy for the WoE model using the ROC technique for validation. Abdullahi and Pradhan (2016) clearly concluded that the AUC validation

**Fig. 7** Kappa index value vs. the number of iterations
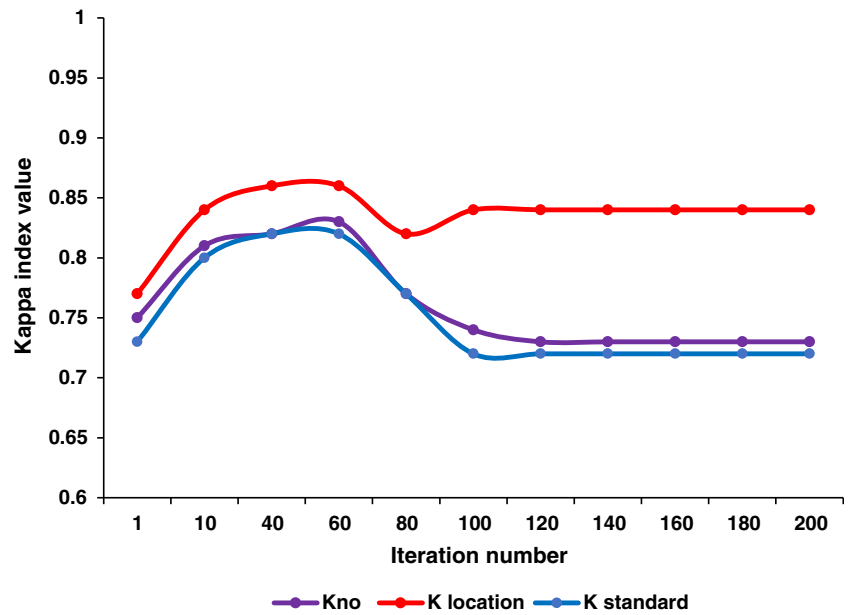
technique that was used to confirm the simulation process, which was conducted using WoE model, indicated 77.4, 78, and 67%, for residential, commercial, and industrial land-use, respectively. The present results confirm that the CA-Markov has a high simulation ability, so it has a high potential to predict future urban growth trends.

The CA-Markov chain model predicted that the urban areas in Seremban would increase to 177 km² and 195.5 km² in 2020 and 2030, respectively (Fig. 9). 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 km² and 756.37 km² in 2020 and 2030, respectively. Unfortunately, this change will
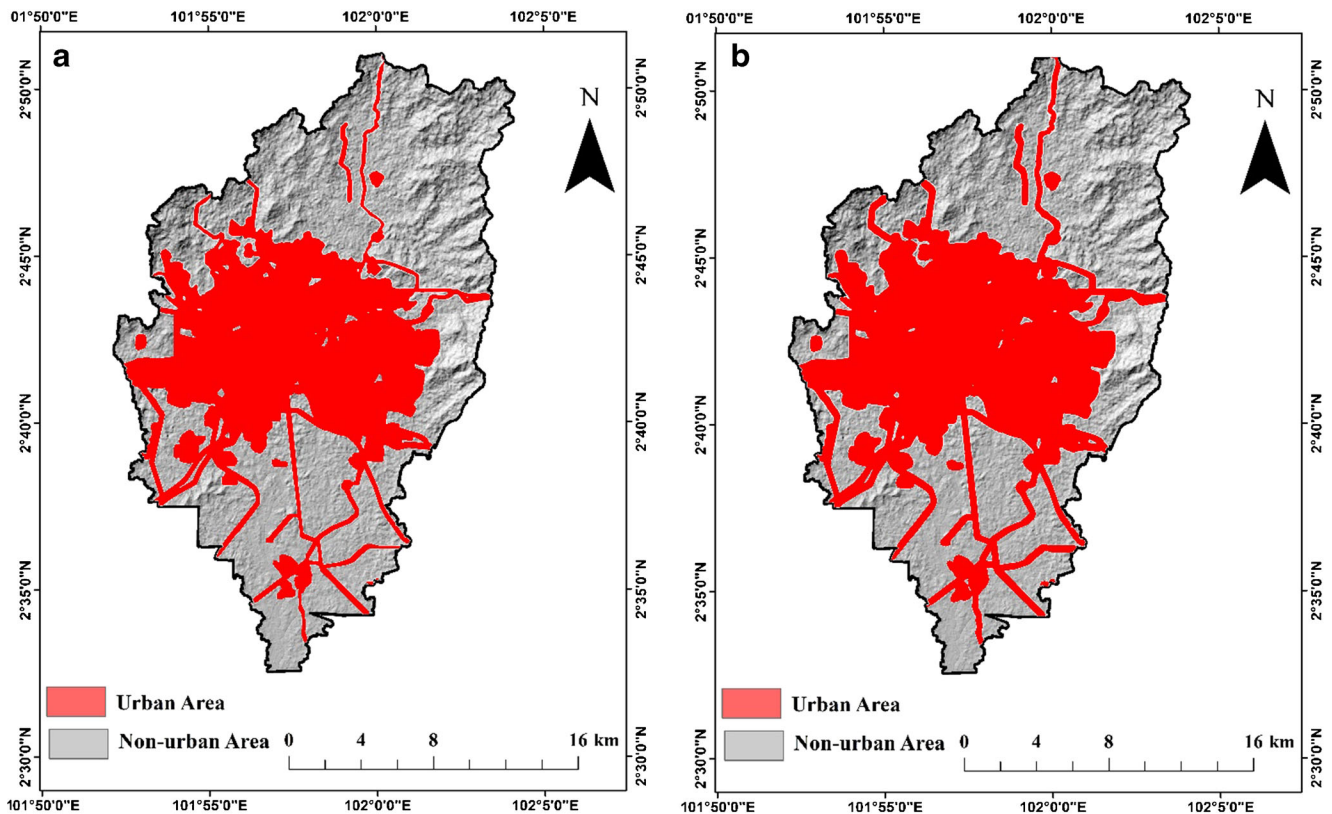
Fig. 8 Predicted maps of urban growth in Seremban River Basin: a 2020 and b 2030

Fig. 9　Quantity of previous and predicted urban and non-urban areas in sq. km
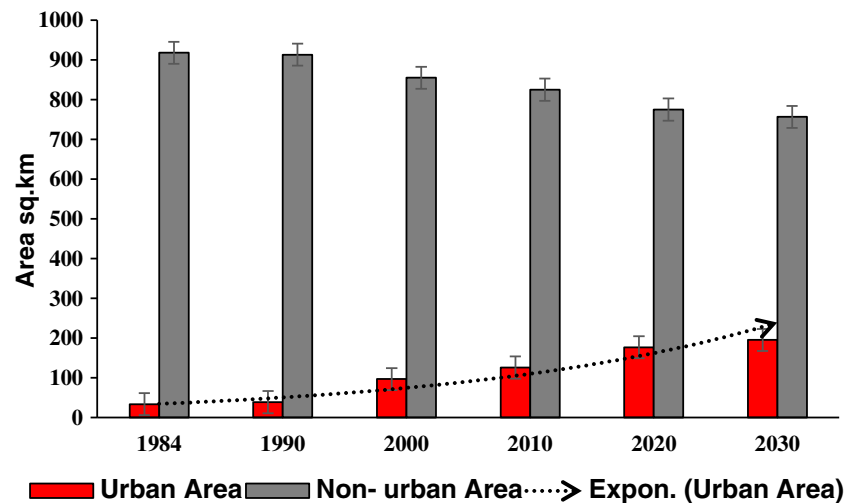
affect the ecosystem and land-use sustainability in Seremban and cause uncontrolled urban growth.

It is important to note that the CA-Markov model applied in this study can predict 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 to predict 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.

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, were simulated and predicted excellently. The model achieved 83% accuracy in simulating projected urban and non-urban land-use maps, indicating 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 periods). However, the integrated model also has some limitations, such as its inability to apply urban growth driving forces, for example, physical and socio-economic forces in the prediction process. These forces are highly significant for monitoring and controlling current urban growth processes and to prepare robust policies and plans for future sustainability.

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 the city's 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 and will threaten the arable lands in Seremban in the long term. Moreover, according to the simulation findings, the urban sprawl in Seremban will follow a disaggregation mode. That is, the urban growth scenario will become worse in the future. Therefore, it is important to save and protect the non-urban areas in Seremban to achieve urban sustainability.

Finally, this study shows the significance of using the integrated CA-Markov chain model for modeling urban growth, especially in developing countries, which have different urban features. However, it is important to assert that the driving forces of urban growth should also be applied when using the CA-Markov chain model in the prediction process to obtain a better understanding of the change in urban growth patterns. Hence, 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 to further improve its prediction capability.

References

Abdullahi S, Pradhan B (2016) Sustainable brownfields land use change modeling using GIS-Based weights-of-evidence approach. Appl Spat Anal Policy 9:21–38

Abdullahi S, Pradhan B, Mansor S, Shariff ARM (2015) GIS-based modeling for the spatial measurement and evaluation of mixed land use development for a compact city. GISci Remote Sens 52:18–39

Aburas MM, Ho YM, Ramli MF, Ash'aari ZH (2016) The simulation and prediction of spatio-temporal urban growth trends using cellular automata models: a review. Inter J Appl Earth Obs 52:380–389

Aburas MM, Ho YM, Ramli MF, Ash'aari ZH (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. Inter J Appl Earth Obs 59:65–78

Al-sharif AA, Pradhan B (2013) Monitoring and predicting land use change in Tripoli Metropolitan City using an integrated Markov chain and cellular automata models in GIS. Arab J Geosci 7(10): 4291–4301

Al-sharif AA, Pradhan B (2015) Spatio-temporal prediction of urban expansion using bivariate statistical models: assessment of the efficacy of evidential belief functions and frequency ratio models. Appl Spat Anal Policy 9(2):213–231

Alsharif AA, Pradhan B (2014) Urban sprawl analysis of Tripoli Metropolitan city (Libya) using remote sensing data and multivariate logistic regression model. J Indian Soc Remote Sens 42:149–163

Anderson JR (1976) A land use and land cover classification system for use with remote sensor data. US Government Printing Office

Ballestores F Jr, Qiu Z (2012) An integrated parcel-based land use change model using cellular automata and decision tree. Proceedings Intern Acad Ecol Environ Sci 2:53–69

Barredo JI, Kasanko M, McCormick N, Lavalle C (2003) Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. Landsc Urban Plann 64:145–160

Batty M, Longley PA (1994) Fractal cities: a geometry of form and function. Academic Press, London

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. J Indian Soc Remote Sens 43: 407–414

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. Environ Model Softw 52:222–233

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

Coppedge BR, Engle DM, Fuhlendorf SD (2007) Markov models of land cover dynamics in a southern Great Plains grassland region. Landsc Ecol 22:1383–1393

Couclelis H (1985) Cellular worlds: a framework for modeling micro—macro dynamics. Environ Plan A 17(5):585–596

Couclelis (1997) From cellular automata to urban models: new principles for model development and implementation. Environ Plann: Plann Des 24:165–174

DOSM (2011) Statistics Yearbook. Department of Statistics Malaysia 3: 367–381

Doygun H (2009) Effects of urban sprawl on agricultural land: a case study of Kahramanmaraş, Turkey. Environ Monit Assess 158:471–478

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

Epsteln J, Payne K, Kramer E (2002) Techniques for mapping suburban sprawl. Photogramm Eng Remote Sens 63:913–918

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. Ecol Model 222:3761–3772

Inouye CEN, de Sousa WC, de Freitas DM, Simões E (2015) Modelling the spatial dynamics of urban growth and land use changes in the north coast of São Paulo, Brazil. Ocean Coas Manag 108:147–157

Maithani S (2009) A neural network based urban growth model of an Indian city. J Indian Soc Remote Sens 37:363–376

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 8:57-70

Muller MR, Middleton J (1994) A Markov model of land-use change dynamics in the Niagara Region, Ontario, Canada. Landsc Ecol 9: 151–157

Pijanowski BC, Brown DG, Shellito BA, Manik GA (2002) Using neural networks and GIS to forecast land use changes: a land transformation model. Comput Environ Urban Syst 26:553–575

Punia M, Singh L (2012) Entropy approach for assessment of urban growth: a case study of Jaipur, India. J Indian Soc Remote Sens 40:231–244

Ren P, Gan S, Yuan X, Zong H, Xie X (2013) Spatial expansion and sprawl quantitative analysis of Mountain City built-up area. In Geo-Informatics in Resource Management and Sustainable Ecosystem. Springer:166–176

Santé I, García AM, Miranda D, Crecente R (2010) Cellular automata models for the simulation of real-world urban processes: a review and analysis. Landsc Urban Plann 96:108–122

Sudhira H, Ramachandra T, Jagadish K (2004) Urban sprawl: metrics, dynamics and modelling using GIS. Inter J Appl Earth Obs 5:29–39

Sun H, Forsythe W, Waters N (2007) Modeling urban land use change and urban sprawl: Calgary, Alberta, Canada. Netw Spat Econ 7: 353–376

Tobler WR (1979) Cellular geography. Philosophy in geography. Springer, Dordrecht, pp 379–386

Tran TV (2008) Research on the effect of urban expansion on agricultural land in Ho Chi Minh City by using remote sensing method 24:104-111

Verburg PH, Soepboer W, Veldkamp A, Limpiada R, Espaldon V, Mastura SS (2002) Modeling the spatial dynamics of regional land use: the CLUE-S model. Environ Manag 30:391–405

Wagner DF (1997) Cellular automata and geographic information systems. Environ Plann 24:219–234

Wang J, Maduako IN (2018) Spatio-temporal urban growth dynamics of Lagos Metropolitan Region of Nigeria based on Hybrid methods for LULC modeling and prediction. Eur J Remote Sen 51:251–265

Wang S, Zheng X, Zang X (2012) Accuracy assessments of land use change simulation based on Markov-cellular automata model. Procedia Environ Sci 13:1238–1245

White R, Engelen G (1994) Cellular dynamics and GIS: modelling spatial complexity. Geogr Syst 1(3):237–253

White R, Engelen G (2000) High-resolution integrated modelling of the spatial dynamics of urban and regional systems. Comput Environ Urban Sys 24:383–400

White R, Engelen G, Uljee I (1997) The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. Environ Plann B Plann Des 24(3):323–343

White R, Engelen G, Uljee I, Lavalle C, Ehrlich D (2000) Developing an urban land use simulator for European cities. In: Proceedings of the Fifth EC GIS Workshop: GIS of Tomorrow. European Commission Joint Research Centre, Stresa, pp 179–190

Wu F (1998) SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. Inter J Geogr Inf Sci 12:63–82

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. Ecol Model 233:11–19

Yeh AG-O, Li X (2001) Measurement and monitoring of urban sprawl in a rapidly growing region using entropy. Photogrammetric Eng Remote Sens 67:83–90

Zhang Q, Ban Y, Liu J, Hu Y (2011) Simulation and analysis of urban growth scenarios for the Greater Shanghai Area, China. Comput Environ Urban Sys 35:126–139