Integrating GIS And CA-MARKOV Model In Evaluating Urban Spatial Growth

NARIMAH SAMAT

ABSTRACT

The physical development is usually undertaken by an expansion of the urban fabric, with agricultural land and green fields on the fringe experiencing industrial and residential development. In Malaysia, for example, although planners know that amount of land needed to cater for urban population expansion, there is no systematic planning tool available to predict to location or the resulting pattern of such development on existing urban landscape. Thus, this study applies of Geographic Information System and CA-Markov model as a planning tool in evaluating urban spatial growth. The result suggests that this model can potentially be used along with existing planning document in order to evaluate the impact of the proposed planning policy on existing urban landscape.

ABSTRAK

Pembangunan fizikal biasanya melibatkan satu perluasan fabrik bandar, dengan tanah pertanian dan kawasan hijau di pinggir bandar dijadikan kawasan perindustrian dan pembangunan perumahan. Di Malaysia, sebagai contoh walaupun para perancang mengetahui tentang jumlah keluasan tanah yang perlu disediakan untuk menampung pertambahan populasi bandar, namun tiada alat perancangan sistematik untuk meramalkan lokasi atau pola perkembangan yang wujud dalam lanskap bandar sedia ada. Oleh Itu, kajian ini mencadangkan pengaplikasian Sistem Maklumat Geografi dan model CA-Markov sebagai alat perancangan dalam menilai pertumbuhan reruang bandar. Hasil kajian mencadangkan bahawa model ini berpotensi digunakan bersama dengan dokumen perancangan sedia ada bertujuan menilai kesan dasar perancangan yang dicadangkan ke atas landskap bandar sedia ada.

INTRODUCTION

Many developing countries are undergoing rapid urbanization mainly due natural population growth and rural urban migration as well as the moving of urban population to urban fringe areas due to changing lifestyle that centered on spacious and environmentally friendly environment (Kivell, 1993; Pacione, 2001). In Malaysia, for example, rigorous industrialization policy has had an impact on many aspects on Malaysian life. Land use has gradually change, with industry, housing estates and new town replacing agricultural activities (Suriati, 1999). Urbanization has increased from 27.6% in 1970 to 65.4% in 2000 and it is projected to achieve 75.0% in 2020 (Ghani Salleh, 2000). In 2001, built-up area was approximately 3.3% or 437,100 hectares of the total area of Peninsular Malaysia. However, this built-up area is expected to increase to 5.8% or 768,600 hectares in order to cater for urban population expansion by 2020 (Mohd Atan, 2005). Furthermore, significant increase of urban population will created demand for residential places and amenities especially within major urban areas and spread towards the fringe areas. These demand, however, are constantly limited due to scarcity of natural resources especially in land. Moreover, experiences have shown that rapid and uncontrolled urban expansion has caused amongst others the deterioration in the quality of urban environment (Petterson et al., 1999; Irwin and Geoghegan, 2001). Thus, various planning and urban management program have been devised to ensure sustainable use of land and sustain the quality of life (Mohd Atan, 2005). Furthermore, such plans should attempt to control against any waste of the natural resources or must be left intact for our posterity (Abdul Samad Hadi et al., 2006).

In order to ensure sustainable use of land, the Town and Country Planning Department of Malaysia launched the National Physical Plan (RFN) in April 2005. This is a long-term strategic plan which will be used to guide strategic development, management and conservation of land in Peninsular Malaysia (Mohd Atan, 2005). The goal of this plan is to establish an efficient, equitable and sustainable national spatial framework to guide the overall development of the country towards achieving developed nation status by 2020 (Town and Country Planning Department, 2008). Although such plan attempts to guide physical urban development in the country, it is a written document prepared for the period of 20 years and reviewed at every five years. This period is considered too long taking into account that socio-economic variables change very fast (Clarke et

al., 1998). This plan, therefore, needs to be supported by more systematic planning tool such as a computer model that can be used to evaluate any planning strategy proposed within RFN.

CELLULAR AUTOMATA MODEL AND GEOGRAPHIC INFORMATION SYSTEMS

To date, various models had been developed in order to monitor land use changes, evaluate planning strategy and plan for new development (Yeh, 1999; Brail and Klostermann, 2001; Geertman and Stillwell, 2003). These models have been integrated with Geographic Information Systems (GIS) in order to incorporate spatial and non-spatial elements of urban systems within the modeling framework. Such an approach allows us to experiment on social and environmental systems in a way that would be impossible in the real world (Batty et al., 1997; Samat, 2002). Furthermore, it helps planners and policy makers in predicting from current situations the implications of specific planning actions undertaken on the future urban conditions. Various models ranging from single sector classical models to multiple sectors econometric models were developed and used to represent urban systems. These models, however, lacked the dynamic components and local interaction of various agents within urban systems. Thus, it was unable to capture the behavior of various interacting individuals or top-down policy implement for the area.

Currently, planners and policy makers started to realize the need to capture the behavior of local urban actors in shaping the landscape of urban area. Such factors are significant since the decisions made locally can help generate the global urban pattern. Cellular automata are one of the models that can help to evaluate local actions and its the resulting impact on global pattern (Couclelis, 1989; Batty et al., 1997; Engelen et al., 1999). This model has gained significant interest in the recent literature since it is a powerful modeling technique that can be used to monitor and evaluate complex urban systems. Cellular automata have been used to simulate either synthetic or actual urban form, which are complex non-linear systems (Batty and Xie, 1994; Clake et al., 1997). Furthermore, this model is simple and can be used to imitate the behavior of urban systems on the basis of realistic transition rule.

Cellular automata were first devised by John Von Neumann and Stanislaw Ulam in 1940s as a frame work in investigating the logical underpinnings of life. They were attempting to explore the possibility of

using purely mathematical formulation to reproduce biological automata (Batty and Xie, 1994; Torrens, 2000). This concept, then, was used by Tobler (1979) to model land use changes and named it geographical model. Since then, various approaches had been used to modify the original cellular automata model in order to replicate complex dynamic phenomena such as forest fire, predator and prey and urban systems (Torrens, 2000). In modeling urban systems, this model was integrated with GIS in order to incorporate both spatial and non-spatial elements of urban systems.

GIS technology, which has been used to store, manage, manipulate and display spatial and non-spatial data, is very useful in managing urban information (Geertman and Stillwell, 2003; Yeh, 1999). The integration of GIS and cellular automata model can be utilized to simulate the dynamic spatial pattern of urban systems (Clarke and Gaydos, 1998; Samat, 2002). This approach is better than previously used approach such as the integration of GIS and statistical model such as multiple regression model or spatial interaction model (Lodgson et al., 1996; Clarke and Gaydos, 1998; Scholten et al., 1999). The implementation of GIS and cellular automata allows planners and policy makers to creatively devised unique transition rules to represent specific urban areas. This, therefore, can be easily be used to understand and solve local urban problem.

METHOD

GIS and CA-Markov model have been used to define the relationship between urban development and its driving factors. It calculates the transition probabilities of urban land use categories on the basis of its suitability, transition probabilities from one category to another and numbers of developed neighbors. The model used for this study is developed using 2-dimensional CA. The structure and concept of the model are shown in Figure 1. First, land use changes between two time periods were evaluated in order to evaluate the size, location and types of land underwent transformation. This would also provide information regarding the factor driving such changes. Factors driving land use changes, then, would be inputted into MCE in order to generate land suitability index maps. These maps along with transition probability maps would be entered into CA-Markov model. Assessment was made by comparing simulated land use pattern with historical land use data. Finally, predicted land use pattern could be produced.

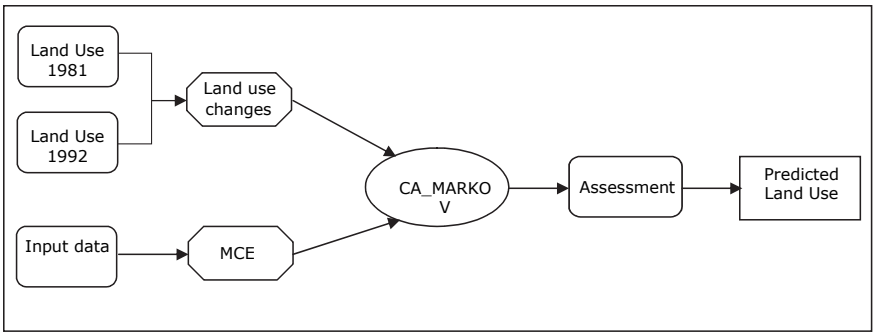


Figure 1: Schematic diagram of the proposed modeling framework.

The study area was divided into cell of 90m resolutions, since previous study revealed that this size was probably the best scale to represent urban landscape in this study area such that it produced the best prediction accuracy and maintained the morphology of urban areas (Samat, 2006). Each cell has five different states representing land use activities in the study area. It is given by:

$$ {}^{t}LU_{i,j} = \begin{cases} 1 = \text{Residential} \\ 2 = \text{Commercial and public facilities} \\ 3 = \text{Industrial} \\ 4 = \text{Agricultural} \\ 5 = \text{Others} \end{cases} \tag{1} $$

These five land use categories represent main land use activities in the study area. The evolution of cells from time *t* to *t+1* is determined by a function of its state, its neighborhood space and a set of transition rule. It is given by equation 2 below.

$$ {}^{t+1}LU_{i,j} = f\left(\left({}^{t}LU_{i,j}\right).\left({}^{t}S_{i,j}\right).\left({}^{t}P_{x,y,i,j}\right).\left({}^{t}N_{i,j}\right)\right) \tag{2} $$

Where

${}^{t+1}LU_{i,j}$	=	the potential of cell *i,j* to change at time *t+1*,
${}^{t}LU_{i,j}$:	=	states of cell *i,j* at time *t*,
${}^{t}S_{i,j}$:	=	suitability indexes of cell *i,j* at time *t*,
${}^{t}P_{x,y,i,j}$:	=	probability of cell *i,j* to change from state *x* to state *y* at time *t*, and;
${}^{t}N_{i,j}$:	=	neighborhood index of cell *i,j*.

The transition rule is formulated based on suitability indexes, transition probabilities and neighborhood indexes. The MCE-suitability index maps were produced using weighted linear combination approach shown in equation 3 below. This approach is simple and can easily be implemented.

$$ {}^{t}S_{i,j} = \sum_{m=1}^{M} {}^{t}x_{i,j} \cdot w_m \cdot c_m \tag{3} $$

where

${}^{t}S_{i,j}$:	=	suitability indexes for cell *i,j* at time *t*,
${}^{t}x_{i,j}$:	=	score of criteria m at cell *i,j* at time *t*,
w_m	=	weight for criterion *m*, and;
c_m	=	Boolean value for constraints.

In addition to suitability index, transition rule was calculated on the basis of transition potential of each cell. It was calculated based on land use changes of two different periods. It is given below.

$$ {}^{t}P_{x,y,i,j} = P\{X_t = a_y \mid X_{t-1} = a_x \} \tag{4} $$

${}^{t}P_{x,y,i,j}$ represents the probability of cell i,j to change from activity ax to activity ay,

Finally, the transition rule was calculated on the basis of number of developed neighbors as given below.

$$ {}^{t}N_{i,j} = \sum^{t} N_{i,j} / 24 \tag{5} $$

The value for ${}^{t}N_{i,j}$ was calculated on the basis of *extended moore neighborhood or* 5 x 5 cells. This size of neighborhood was used since it allowed the influence of surrounding cells on central cell to be taken into consideration. This uniform transition rules were repeatedly applied in order to simulate land use changes from 1981 to 1998. It was implemented for the Seberang Perai Region, Penang State Malaysia.

STUDY AREA AND DATA

The Seberang Perai region Penang State, Malaysia was chosen as a study area. Seberang Perai is part of Penang State and is located in the Northwest of Peninsular Malaysia, centered at 5o 20'N latitude and 100o 25'E longitude, with an area about 738.4km2. It population was 432,982 in 1980. However, there was significant increase of Seberang Perai population in 1991, where 545,680 people lived in this area and increased to 699,084 in 2000. It was projected that by the year 2015 and 2020, more than 990,000 and 1.1millions people would be living in this region respectively (SPMC, 1998; JPBD, 2005). Such an increase of its population would indeed require significant amount of land to accommodate the need for housing areas and other facilities. Figure 2 shows the location of the study area and the urban centers evaluated in the analysis. This region is undergoing quite rapid urbanization resulted from the spill-over demand from Penang Island for residential place and other developments. Seberang Perai is relatively flat. Approximately, 92.8 percent of this area has slope of less than 5 percent (with height less than 50m). Only 5.2 percent of the Seberang Perai region has slope of more than 15 percent. Physically, this area is suitable for many development activities.

This area is suitable for testing the proposed model since it has experience significant urbanization mainly resulted from industrialization and resulted residential growth in the last 30 years. This area also has been planned to play a prominent role in attracting economic growth to newly planned Northern Corridor Economic Region (NCER) and as a part of Indonesia-Malaysia-Thailand Growth Triangle (JPBD, 2005; Samat, 2007).

Another reason for selecting the Seberang Perai region as the study area is data availability. Datasets used for this project was acquired over prolong period for Seberang Perai. These include digital datasets of roads (at 1:50,000), sub-districts (at 1:75,000), slope (at 1:50,000), and land use (at 1:75,000) of 1992. Other data such as road network and public

facilities were digitized from topographic maps (Department of Survey and Mapping, 1986). Land use 1998 data was obtained from Seberang Perai Municipal Council (SPMC, 1998). Soil data was digitized from soil map obtained from Department of Agriculture, Malaysia (Soo and Selvadurai, 1969). Data from various sources, scale and format, at different temporal scale were converted into IDRISI format.

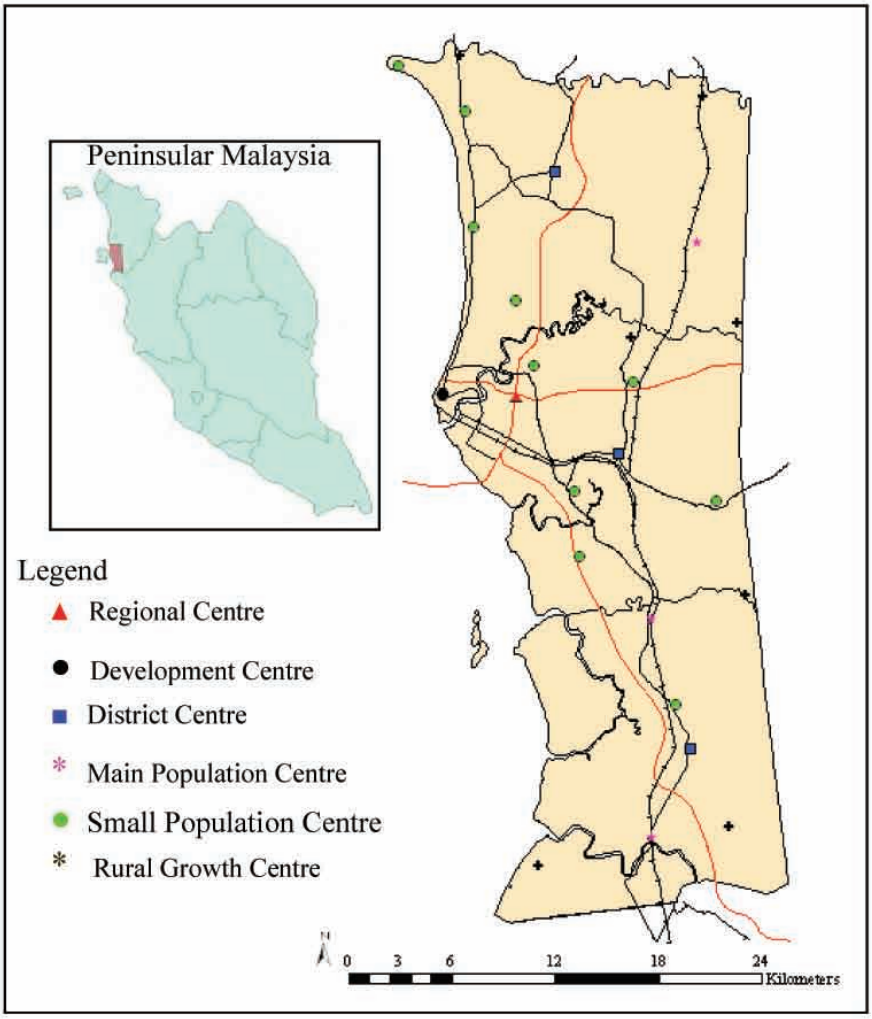
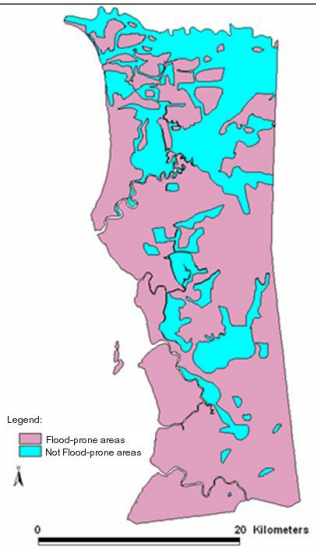


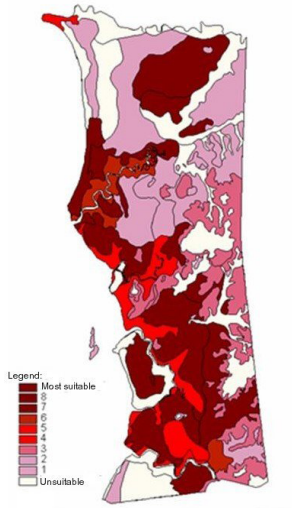
Figure 2: Study area used to test the proposed model.

FACTORS INFLUENCING URBAN DEVELOPMENT

Factors influencing urban land use development can be divided into three categories namely physical, socio-economic and environmental criteria. Five physical factors describing spatial relationship between sites and others facilities located within the study areas were used. Socio-economic factor was limited to the availability of land value data which was interpolated from property evaluation report (Valuation and Property Department, 1999). Finally, environmental factors considered in this model were flood-prone areas and valuable paddy field agriculture, which were utilized as constraints towards urban development. Figure 3 shows maps used as constraints on urban development in this study area. These maps illustrate flood-prone areas (on the left) and highly suitable agriculture land (on the right) as hindrance for any types of development.



a) Flood-prone area was considered as hindrance towards any development.



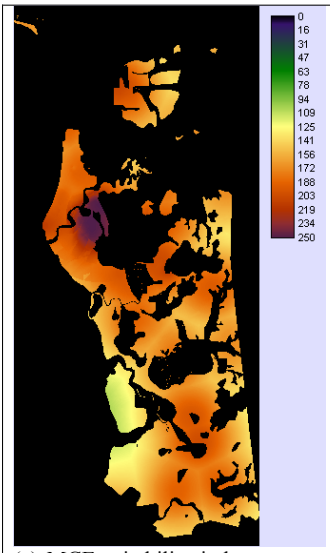
b) Highly suitable agriculture land was assumed as unsuitable for any type of development.

Figure 3: Maps used as constraints for urban development.

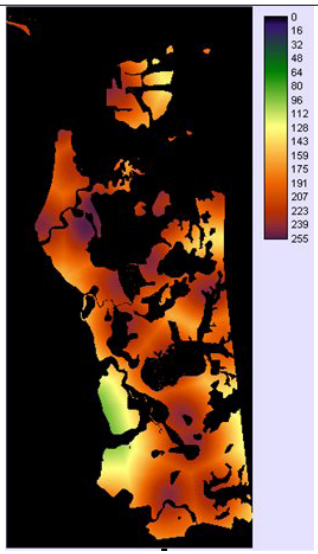
Weights were determined based on the interviewed conducted with five local planners, who were directly involved in land use planning in the study area (Samat, 2002). Different sets of weights were derived for different land use activities namely residential, commercial and public facilities, and industry. Table 1 shown the weights used for the three major land use activities evaluated in this study. The weights used, however, were held constant throughout the simulation, changes in those weights might have significant influence on urban spatial growth (Samat, 2006). These weights were used to produce suitability index maps as shown in Figure 4. These maps shown that potential sites for urban development were clustered around existing urban development corridors. Land use 1981 was used as initial state of the model. The output produced was validated using actual land use data of 1992 and 1998. Then, the model was used to simulate urban spatial growth until 2010.

Table 1: Weights used for three major land use activities modeled in the study.

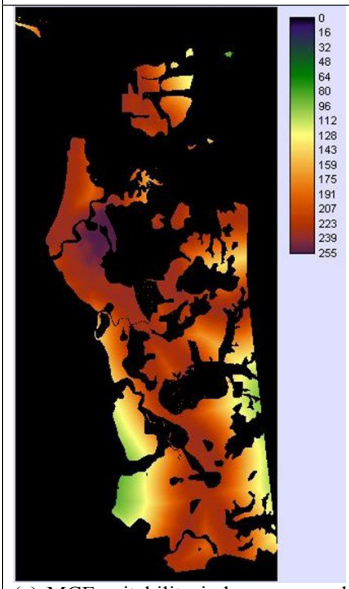
Criteria Used	**Residential**	**Commercial and Public Facilities**	**Industrial**
Proximity to employment centers	0.14	-	-
Proximity to major roads	0.08	0.26	0.73
Proximity to public facility	0.05	-	-
Proximity to population centers	0.28	0.64	0.08
Land Value	0.44	0.11	0.19
Consistency Ratio	*0.08*	*0.03*	*0.06*



(a) MCE-suitability index map produced for residential activity.



(b) MCE-suitability index map produced for commercial and public facilities.



(c) MCE-suitability index map produced for industrial activity.

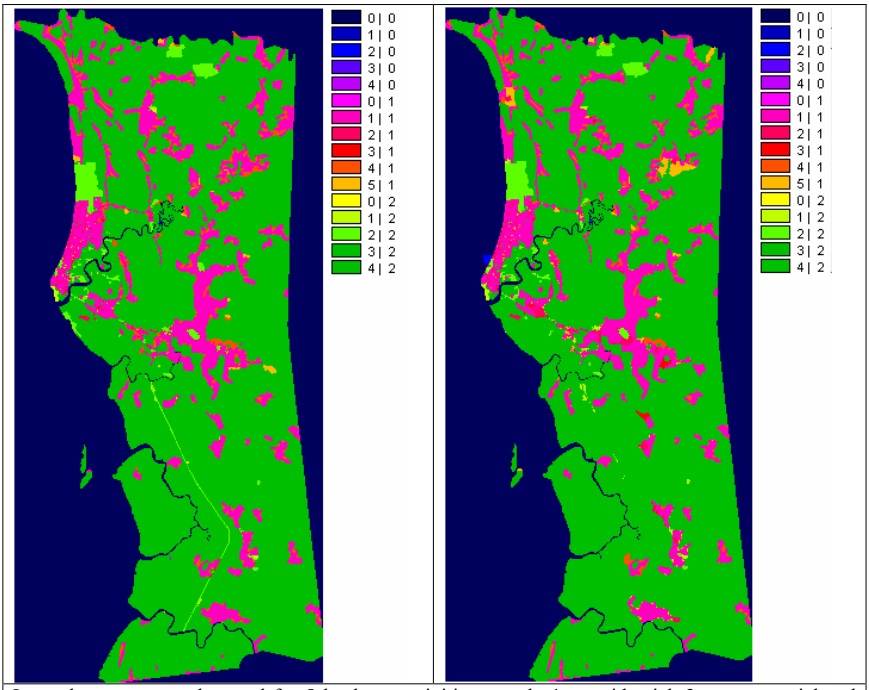
Figure 4: MCE-suitability index maps produced for three major land use activities modeled in this study.

RESULT AND DISCUSSION

The model performed quite well. Figure 5 shows the actual and simulated urban spatial pattern from 1992 to 1998. Table 2 shows the validation results for 1992 and 1998. The table illustrates that overall accuracy is quite high that is 92.3 and 81.0 for 1992 and 1998 respectively. The accuracy for each land use activity is quite high for 1992, but the accuracy for 1998 is a little bit low. This is probably due to the magnitude of error scales in term of the iterative process of the CA model. The accuracy for the three main land use categories namely residential (coded 1|1 in Figure 5), commercial and public facilities (coded 2|2) and industrial (coded 3|3) for 1992 is 83.7, 80.1 and 84.7 respectively. This result is considered as good prediction accuracy (Moonsrud and Leemans, 1992). The accuracy for these land use categories decreases to 70.9% for residential, 40.5% for commercial and public facilities, and 51.9% for industrial activity in 1998. The accuracy for residential land use probably is quite high since the model developed mainly based on physical factors can actually replicate residential development in this study area (Chapin and Kaiser, 1979). Furthermore, residential developments usually are undertaken by private developers who usually used land value and physical factors such as proximity to main roads and proximity to employment centers to determine sites selection. For the development of commercial and public facilities and industrial land use activities, site selection is usually determined by planners or policy makers. The development is usually at a large scale undertaken to promote economic growth within the region (SPMC, 1998). The model, which is developed mainly based physical factors, is unable to simulate these activities very well.

Table 2: The validation of the proposed model for the simulation of urban development between 1981 and 1998.

Land Use Categories	**Kappa Index (x 100) 1992**	**Kappa Index (x 100) 1998**
Residential	83.7	70.9
Commercial and public facilities	80.7	40.5
Industrial	84.1	51.9
Agricultural	86.1	85.3
Others	83.0	44.8
Overall accuracy	90.29	81.04



Legend represents codes used for 5 land use activities namely 1 – residential, 2 – commercial and public facilities, 3 – industrial, 4 – agriculture, and 5 – others. Code on the left represents actual land use categories and code on the right represents simulated land use categories.

Figure 5: Validation of simulated urban pattern for 1992 (on the left) and 1998 (on the right).

Another aspect that is attributable to the poor performance of the model is that urban growth does not occur uniformly across space and time (Pacione, 2001). For example, development centers (refer to Figure 1) experienced more rapid development when compared to district centers and small population centers. The model, however, considered urban growth to be uniformed across the study area (Samat, 2002). Therefore, the model tends to overestimate urban expansion in areas experiencing less rapid urban growth. On the other hand, the model underestimates urban growth in areas experiencing rapid urban development.

Finally, the factor contributing to the poor performance of the model is the inability of the model to recognize new development. In this model, vacant cells are urbanized based on their suitability value, the states of cells within 5 by 5 neighborhood and transition probabilities. This rule results in cells that are highly contiguous to existing urban areas.

This model, however, is not able to capture urban growth that has occurred away from existing urban development (Samat, 2002).

Although the performance of the model was not as high as expected, this model is useful in deriving the spatial pattern of urban growth. This type of information is significant in early stage of the planning process by which planners can test different planning scenarios within computer environment before actually choosing a specific plan. This can prevent mistake in selecting urban planning program, thus reduce its negative impacts (Brail and Klostermann, 2001). This model has been used to simulate urban spatial growth 2010. Figure 6 below represents land use pattern in 2010. This figure illustrates that urban development will continue to be developed around existing urban areas. It seems that the restriction imposed on the development of valuable agriculture land manages to control and direct urban development into available land in the south of the study area.

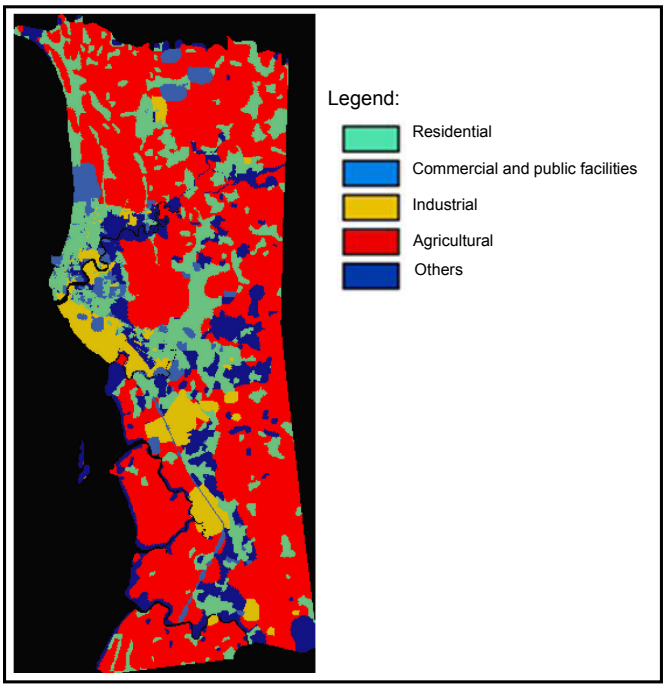


Figure 6: Urban spatial pattern 2010 produced using GIS and CA_Markov model.

CONCLUSION

Current planning projection and related planning policies suggest that Malaysia will experience rapid urbanization at least until 2020. At present, however, land use allocation is still being conducted in rather ad-hoc manner, often on the basis of knowledge of a few decision makers and local planners. Thus, there is no clear policy on land use allocation or resulted urban spatial growth in this region. Although planners are aware that non-urban land especially agriculture land at the urban fringe or linear to major transport network is likely to be converted to urban, there is no planning tool to locate, map or forecast areas that may experience urban pressure. The integration of GIS and CA-Markov model developed here is suitable to be used as a complement to existing strategic plan utilized by planning department in Malaysia.

REFERENCES

Abdul Samad Hadi, Shaharudin Idrus, Ahmad Fariz Mohamed dan Abdul Hadi Harman Shah, 2006. *Perubahan Persekitaran Bandar dan Kemudahanterancaman Lembangan Langat*, Bangi: Penerbit Universiti Kebangsaan Malaysia.

Batty, M. and Xie, Y. 1994. From cells to cities. *Environment and Planning B: Planning and Design*, 21, pp.S31-S48.

Batty, M., Couclelis, H. and Eichen, M. 1997 Urban systems as cellular automata, Environment and Planning B: Planning and Design, 24, 159-164.

Chapin, F.S. and Kaiser, E.J. 1979. *Urban Land Use Planning*. Urbana, University of Illinios Press.

Clarke, K.C., Hoppen, S. and Gaydos, L. 1997. A self-modifying cellular automaton model of historical urbanisation in the San Francisco Bay Area. Environment and Planning B: Planning and Design, 24, 247-261.

Clarke, G., Langley, R., and Cardwell, W. 1998. Empirical applications of dynamic spatial interaction models. *Computers, Environment and Urban Systems*, 22(2), pp 157-184.

Clarke, K. and Gaydos, L. 1998. Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science*, 12(7), pp.699-714.

Couclelis, H. 1985. Cellular worlds: a framework for modelling micro-macro dynamics. *Environment and Planning A*, 1, pp.585-596.

Choguill, C. L. 1994. Crisis, chaos, crunch? planning for urban growth in the developing world. *Urban Studies*, 31 (6), pp.935-945.

Department of Survey and Mapping, Malaysia. 1986. *Topographic Map*. Kuala Lumpur, Department of Survey and Mapping, Malaysia.

Engelen, G., Geertman, S., Smits, P. and Wessels, C. 1999. Dynamic GIS and strategic physical planning support: a practical application. In: Stillwell, J., Geertman, S. and Openshaw, S. eds. *Geographical Information and Planning*, London, Springer, pp.87-111.

Ghani, S. 2000. *Urbanisation & Regional Development in Malaysia*, Kuala Lumpur, Utusan Publications & Distributors Sdn Bhd.

Harris, B. and Batty, M. 2001. Locational Models, Geographic Information and Planning Support Systems, in Brail, R.K and Klosterman, R.E (eds) *Planning Support Systems*, ESRI Press, California pp. 25-57.

Ibrahim Wahab and Abdul Hadi Harman Shah 1992. *Kaedah Analisis Perancangan*. Kuala Lumpur, Dewan Bahasa dan Pustaka.

Irwin, E.G., and Geoghegan, J. 2001. Theory, data, methods: developing spatially explicit economic models of land use change, *Agriculture, Ecosystems and Environment*, 85, 7-23.

Kivell, P. 1993. *Land and the City: Pattern and Processes of Urban Change*, London, Routledge.

Mohd Atan, M. J. 2005. Rancangan Fizikal Negara: Pelaksanaan dan Pemantauan, Senior Planners Meeting, Department of Urban and Country Planning, Malaysia, 25 – 26 July, Penang, Malaysia.

Monserud, R.A. and Leemans, R. 1992. Comparing global vegetation maps with the kappa statistic. *Ecology Modelling*, 62, pp.275-293.

Peterson, P.J, Sham, Sani and M. Nordin 1999. *Indicatiors of Sustainable Development in Industrialising Countries*. Volume III, Key Indicators for Tropical Cities, Universiti.Kebangsaan Malaysia: Institute for Environment and Development.

Samat, N. 2002. A geographic information system and cellular automata spatial model of urban growth for Penang State, Malaysia. Ph.D. Thesis, School of Geography, University of Leeds, Leeds, United Kingdom.

Samat, N. 2006. Characterising scale sensitivity of cellular automata model of urban growth: A case study of Seberang Perai, Penang, Malaysia, *Computers, Environment and Urban Systems*, Elsevier Ltd, Vol 30, 901-920.

Scholten, H.J., van de Velde., Rietveld, P. and Hilferink, M. 1999. Spatial information infrastructure for scenario planning: the development of a land use planner for Holland. In: Stillwell, J, Geertman, S. and Openshaw, S. eds. *Geographical Information and Planning*. London, Springer, pp. 112-134.

SPMC (1998), *Structure Plan* (Reviewed). Seberang Perai Municipal Council, Town and Country Planning Department, Malaysia.

Soo Swee Weng and Selvadurai, K. 1969. Reconnaissance Soil Survey of Penang and Province Wellesley. *Malayan Soil Survey Report No 1/1969*, Ministry of Agriculture & Co-Operatives, Malaysia.

Yeh, A.G-O. 1999. Urban planning and GIS. In: Longley, P.A, Goodchild, M.F, Maguire, D.W, and Rhind, D.W. eds. *Geographic Information Systems: Management issues and applications*, 2nd edition, New York, John Wiley & Sons, pp. 877-888.

School of Humanities
Universiti Sains Malaysia
11800 Penang, MALAYSIA
e-mail: narimah@usm.my