Available online at www.humg.edu.vn

ESASGD 2016

GIS-IDEAS (2016)

Conference Title: International Conference on GeoInformatics for Spatial-Infrastructure Development in Earth & Allied Sciences (GIS-IDEAS)

Predicting land use change affected by population growth by integrating Logistic regression, Markov chain and Cellular automata models

Nguyen Van Trung[a] , Le Thi Thu Ha[a,]*

[a] *University of Mining and Geology, 18 Vien St, Duc Thang Ward, Bac Tu Liem District, Hanoi City 1000000, Vietnam*

Abstract

Demographic change was thought to be the most major driver of land use change although there were several interacting factors involved, especially in the developing countries. This paper presents an approach to predict the future land use change at the Balat estuary using a hybrid model. A hybrid model consisting of logistic regression model, Markov chain (MC), and cellular automata (CA) was designed to improve the performance of the standard logistic regression model. Demography and socio-economic variables dealing with urban sprawl were operationalised to create a probability surface of spatio-temporal states of built-up land use for the years 2009, 2019, and 2029. The predicted future land use maps for the years 2019 and 2029 show substantial urban development in the area, much of which are located in areas sensitive to source protections. The results of the analysis provide valuable information for local planners and policy makers, assisting their efforts in constructing alternative sustainable urban development schemes and environmental management strategies.

Keywords: Land use change, Population growth, Logistic Regression, Markov Chain, Cellular Automata, Giao Thuy district.

1. Introduction

The intensity of land use change in response to world population growth and its consequences for the environment warrant in-depth studies of these transformations (Wu et al., 2006). Land is an important and finite resource for most human activities such as settlement, agriculture, forestry, animal husbandry, industry, transportation and recreation. It has been tightly coupled with economic growth (Richards, 1990). One of the six possible forces driving land-use and land-cover changes is population increase and its level of affluence, technology, political economy, political structure, and attitudes and values (Meyer and Turner, 1992). Population increase arise a sequence of immediate life sustaining needs such as residence space, food and fiber. However, due to the finite amount of available land, fast economic development and population growth lead to deforestation and loss of arable land and biodiversity, and reduction of environmental services (Lambin et al., 2001).

In recent years, the Land-Use and Cover-Change (LUCC) community has produced a large set of operational models that can be used to predict or explore possible land use change trajectories (Verburg et al., 2006). Models cannot only support the exploration of future land use changes under different scenario conditions, Scenario analysis with land use models can but also support land use planning and policy. Logistic regression (McCullagh and Nelder, 1989) analysis has been one of the most frequently utilized approaches

* Corresponding author. Tel.: +84-043-838-4973;
E-mail address: lethuhatdm1979@yahoo.com.

during the past two decades for predictive land use modeling by means of variation of inductive modeling. Thereby, it is crucial to consider spatial effects, namely spatial autocorrelation and spatial heterogeneity, to challenge regression assumptions (Anselin, 1988; Fotheringham et al., 2000). However, the logistic regression model suffers from the quantification of change and temporal analysis (Hu and Lo, 2007). Thus, empirical estimation and dynamic simulation models have been used to simulate land use change. Various types of rule-based modeling, for instance CA, are the most appropriate with incorporating spatial interaction effects and the treatment of temporal dynamics. Whereas CA models focus on the simulation of spatial patterning rather than on the interpretation of spatiotemporal processes of urban sprawl, there is a deficiency of incorporation among dynamic simulation models and socio-economic and demographic variables (Hu and Lo, 2007). Due to limitations of each individual modeling technique (Poelmans and Rompaey, 2009) proposed a hybrid approach based on logistic regression coupled with CA transition rules, which results in an improved model quality, nevertheless, their model was not able to quantify the amount of land use change.

A need for spatial models of land use change was therefore identified (Mertens and Lambin, 1997). In this paper, an approach to land use change model is to integrate CA, logistic regression, and MC models in order to produce temporal outputs from the logistic regression model. The methodology is applied to the Giao Thuy to analyse recent and near-future developments in land use pattern under conditions of increasing population pressure.

2. Study area

Giao Thuy is a rural district (belong to Balat estuary) located in Nam Dinh Province in the Red River Delta in Vietnam. In 2003 population of the district was 207,273. The district covers an area of 166 km² and has a central town named Ngo Dong. Besides, this district included 20 communes and a small town. Giao Thuy district has the Xuan Thuy Natural Wetland Reserve, which is the only Ramsar site in Vietnam (Ramsar Convention Bureau, 1997). In 1988, 120 km^2 of mangroves were designated for inclusion for a reserve (Fig. 1)

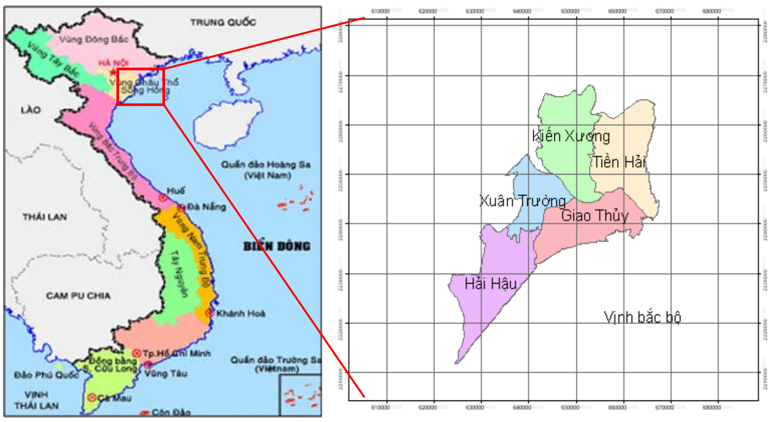


Fig.1. Red River Delta and the Giao Thuy study area

3. Material, data sources and analysis method

3.1 Materials and data sources

Table 1. List of variables included in the logistic regression-Markov-CA model

Variable	Meaning	Nature of variable
Dependent		
Y1 (1989-1999)		
Y2 (1999-2009)	0 – no change; 1 – change	Dichotomous
Y3 (1989-2009)		
Independent		
X1	Population density (person/km2)	Continuous
X1	Households density (Number of households /km2)	Continuous
X1	The proportion of people in working age (% person/km2)	Continuous
X1	Distance to active economy centers (km)	Continuous
X1	Distance to the nearest major road (km)	Continuous
X1	Distance to the sea dike (km)	Continuous

3.2 Methods

This section discusses the essential characteristics of the utilized models, which are integrated in this approach. An overview is given in Fig. 2. First, land use maps of 1989, 1999, and 2009 were produced by the processing of Landsat TM images of the aforementioned years; additionally, temporal land change mapping was implemented. Second, the main driving forces determining land use change, using logistic regression, were investigated. The resulting probability surface of future land change was used in the third step to estimate the quantity of change based on the MC model. Fourthly, whereas the MC model is not able to allocate the estimated amount of change and has to be integrated with other geospatial models, a customized CA model was designed in order to achieve the desired objective. In order to verify the results, the land use map of 2009 was estimated and compared against actual land use maps in the fifth step. Finally, the model was used to simulate future land use maps of 2019 and 2029.

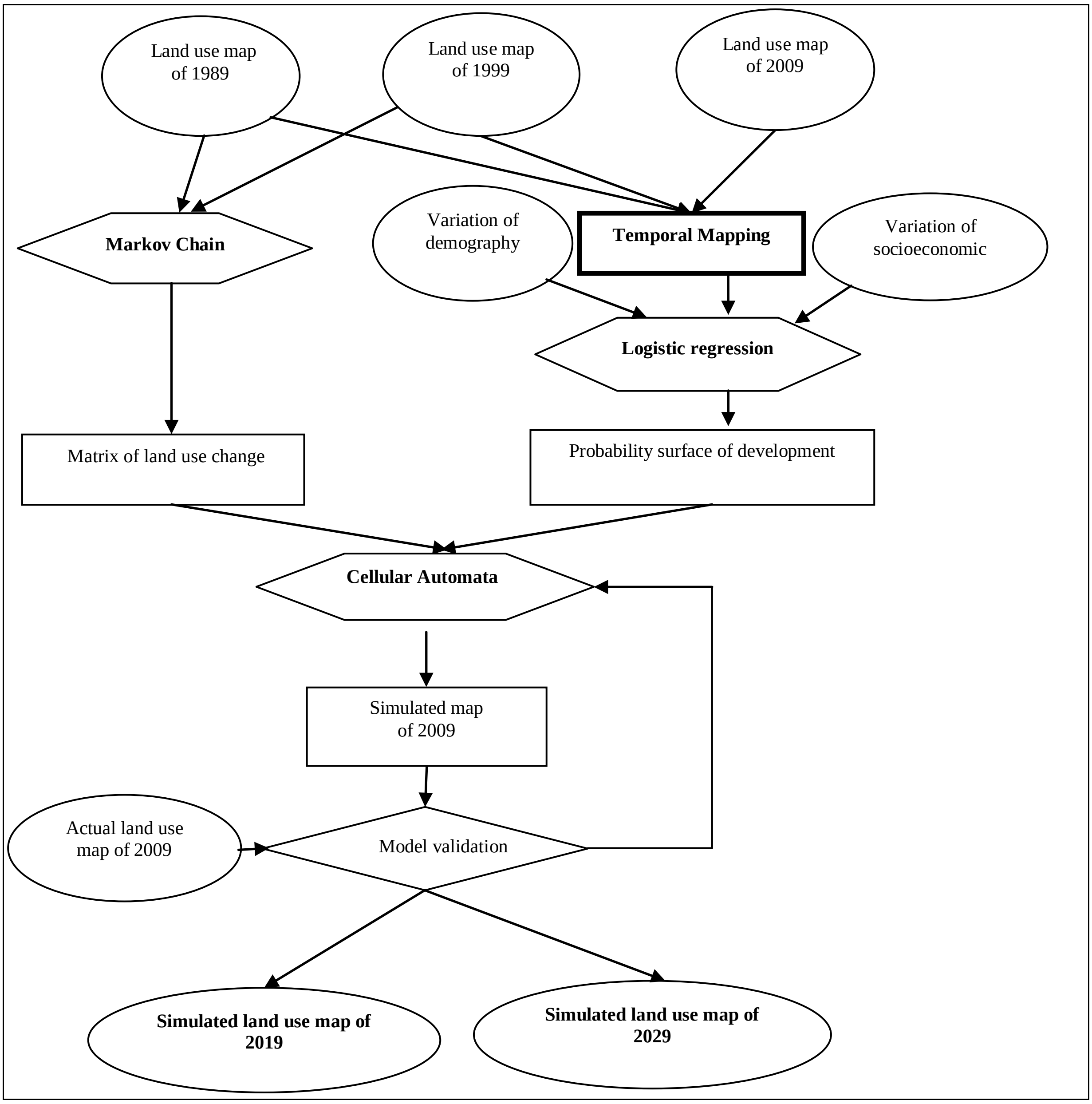


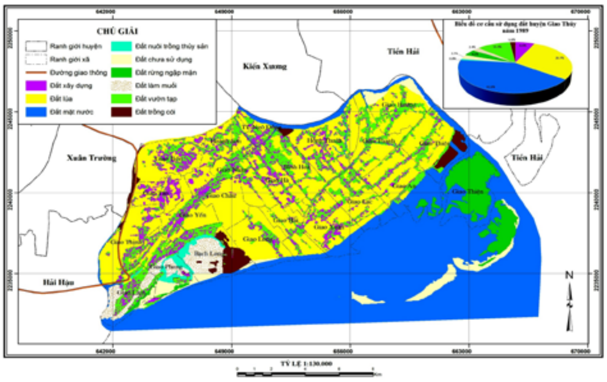
Fig. 2. Flowchart of the Logistic–Markov–CA approach.

4. Results

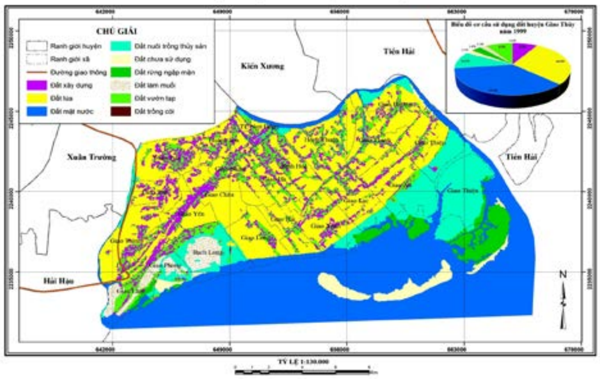
4.1. Temporal land use mapping

For temporal land use mapping of the study area a set of Landsat TM and ETM+ images for the period 1989–2009 was chosen to extract land use maps. Images of 1989, 1999, and 2009, constituting regular 10-year cycles were chosen to be synchronized with the environmental and socioeconomic data. Accuracy assessment by means of cross-tabulation analysis of the map classifications was conducted in order to ensure the accuracy of the maps at 71,88%, 78,54% and 80%, respectively. Eight land use categories were retrieved, i.e., agricultural lands, water, aquacultural lands, mangrove, sedge-land, open land and built up areas. Fig. 3 illustrates the produced land use maps.

Land use map of 1989



Land use map of 1999



Land use map of 2009

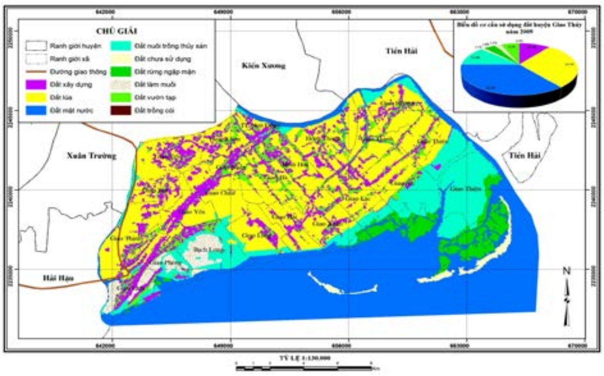


Fig. 3: Extracted land use maps of 1989 (upper left), 1999 (upper right), and 2009 (lower center)

Table 2. Quantity of land use change over time in terms of hectare and percent of each category

Year	1989		1999		2009		1989-1999	1999-2009	1989-2009
Category	(ha)	%	(ha)	%	(ha)	%	(ha)	(ha)	(ha)
Built-up	4525	16,2	4792	17,1	5239	18,7	+ 267	+447	+714
Agricultural	9415	33,7	9200	32,9	8578	30,6	-215	-622	-837
Water	10256	36,7	8188	29,3	8491	30,4	-2048	+303	-1745
Mangrove	1304	4,7	1039	3,7	1174	4,2	-265	+135	-130
Aquacultural	248	0,9	3208	11,6	3528	12,6	+2960	+320	+3280
Salt-land	728	2,6	682	2,4	648	2,3	-46	-34	-80
Open land	884	3,2	764	2,7	314	1,2	-120	-450	-570
Sedge-land	600	2,1	87	0,3	8	0,0	-513	+79	-434

4.2. Quantification of future changes

The MC model was run to quantify changes, with a pair of land use images as input and a transition probability matrix, a matrix of transition areas, as well as a set of predicted transition probability surface maps as output of logistic regression model (Fig. 4). The results (transition probability surface maps, matrices) were used for further change analysis and determine the estimated quantity of change that is assumed to be an input for the CA model. MC is not a spatially explicit model; therefore, it is not an appropriate model to estimate the location of change, which needs to be integrated with other spatial models and in this investigation, logistic regression and CA models were chosen to spatialize the estimated change quantity. Finally, the simulated land use for 2019 and 2029 has been demonstrated in Table 3 and Fig. 5.

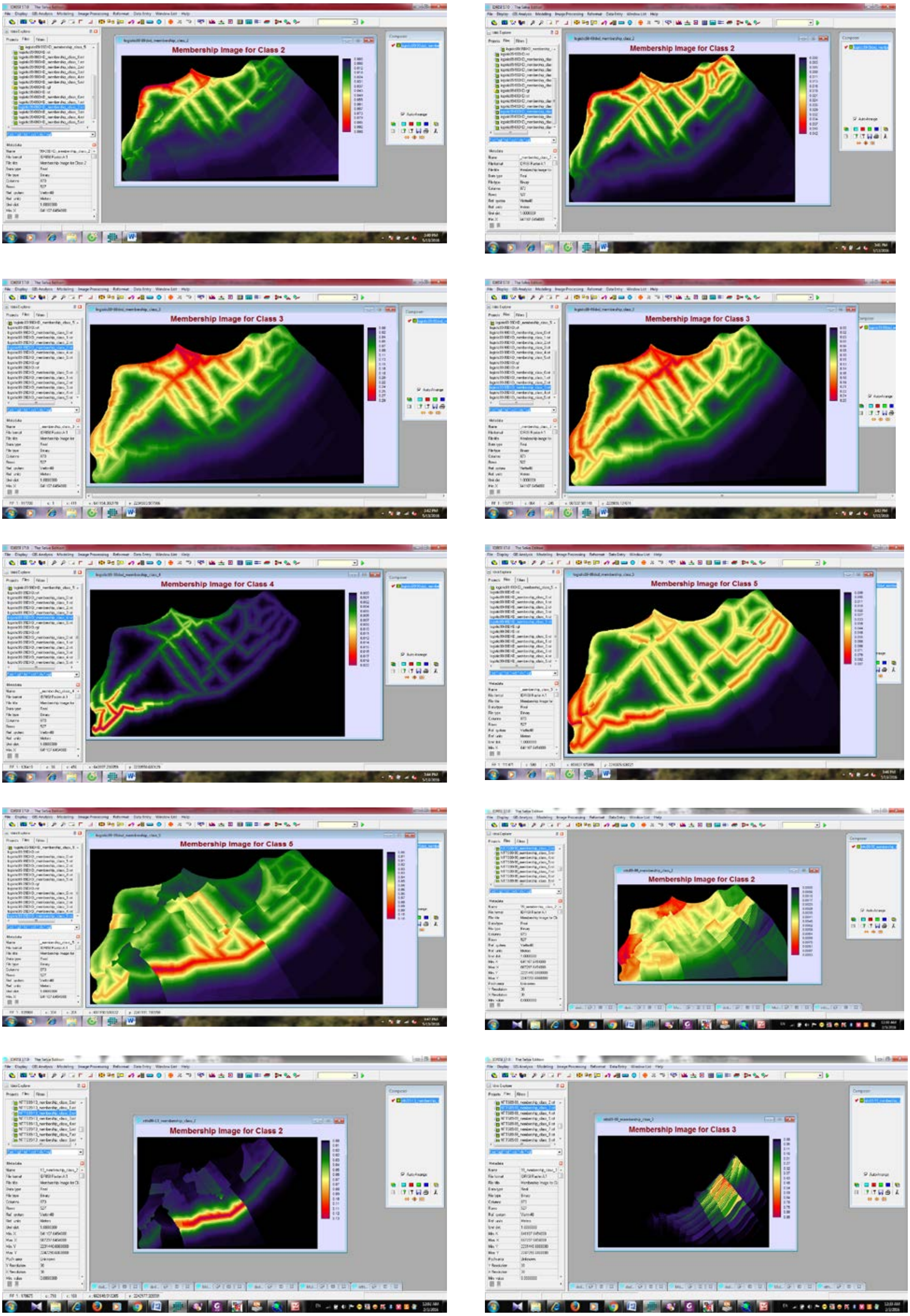


Fig. 4. Predicted transition probability surface map of Logistic regression

Table 3. Quantity of land use through the Markov chain model for 2019 and 2029 in hectare

Category Year	Built-up (ha)	Agricultural (ha)	Water (ha)	Mangrove (ha)	Aquacultural (ha)	Salt-land (ha)	Open land (ha)	Sedge-land (ha)
2019	6105	7731	7930	1207	4508	302	177	0
2029	6834	7146	6118	1301	5843	211	107	0

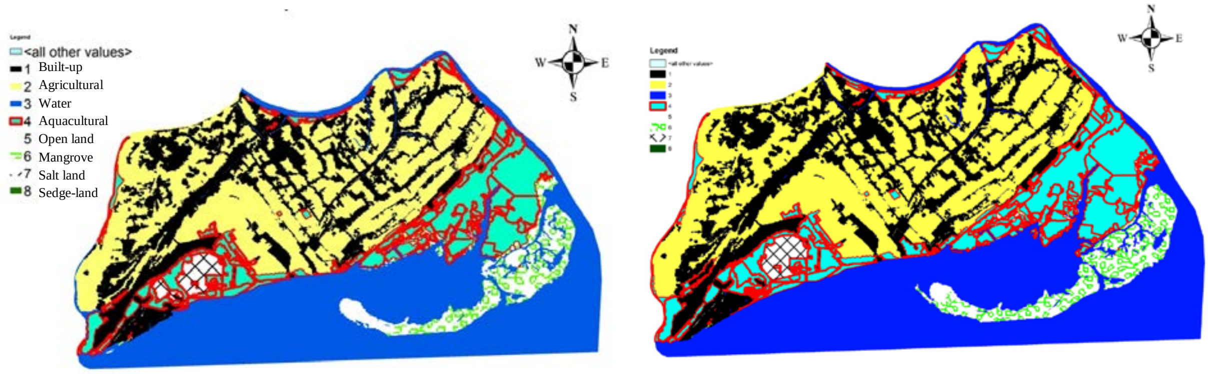


Fig. 5: Simulated land use maps of 2019 (left) and 2029 (right) through the Logistic-Markov-CA approach

4.3 Model validation

In order to validate the proposed approach, the probability map of change for 2009 was utilized to allocate the attained quantity of change through the customized cellular automata function. It was aimed to simulate the land use map of 2009 and compare it with the actual map of 2009. Being aware of (Pontius and Millones's, 2011) critics, kappa statistic for map classification comparison (actual map vs. simulated maps) was applied and revealed a Kappa index of 0.78.

5. Discussion and conclusions

As shown in Table 2 and 3, after 2009, the land use change trend is agricultural, salt-land, open land, sedge-land and water land area decreased significantly. The maximum rate of reduction is still sedge-land. Mangrove, aquaculture area, and built-up land are increasing year by year. The growth of built-up land and aquaculture area are the biggest in all land use classes. This change will persist for long time, until it reaches a relatively steady state.

These three techniques were combined for the following purposes: firstly, the logistic regression model was utilised to create a probability surface and to determine the most probable sites for development; secondly, the MC model was used to retrieve the quantity of change. Because land development policy has been inconsistent in recent years population growth and land development rates are impossible to synchronise. Thirdly, the CA model is a significant tool to allocate probable changes under predefined conditional rules. This CA model allocated the amount of change, beginning with the cells of highest probability. Therefore, the approach is capable of predicting the most probable sites for development, estimating the likely amount of change as well as allocating the estimated quantity within the study area.

References

Anselin, L., 1988. Spatial Econometrics. Methods and Models. Kluwer Academic Publishers, Dordrecht.

Fotheringham, A.S., Brunsdon, C., Charlton, M., 2000. Quantitative Geography: Perspectives on Spatial Data Analysis. Sage Publications, London; Thousand Oaks, California.

Hu, Z., Lo, C., 2007. Modeling urban growth in Atlanta using logistic regression. Computers, Environment and Urban Systems 31 (6), p.667–688.

Kamusoko, C., Aniya, M., Adi, B., Manjoro, M., 2009. Rural sustainability under threat in Zimbabwe – simulation of future land use/cover changes in the Bindura district based on the Markov-cellular automata model. Applied Geography 29 (3), p. 435–447.

McCullagh, P., Nelder, J., 1989. Generalized Linear Models. CRC Press, Boca Raton.

Mertens, B., Lambin, E.F., 1997. Spatial modelling of deforestation in southern Cameroon. Spatial disaggregation of diverse deforestation processes. Applied Geography 17, p. 143-162.

Meyer W B, Turner B L, 1992. Human population growth and global land-use/cover change. Annual Review of Ecology and Systematic, 23: 39–61.

Lambin E F, Turner B L, Geist H J et al., 2001. The causes of land-use and land-cover change: Moving beyond the myths. Global Environmental Change, 11(4), p. 261–269.

Poelmans, L., Van Rompaey, A., 2009. Complexity and performance of urban expansion models. Computers, Environment and Urban Systems 34 (1), p. 17–27.

Pontius Jr., R.G., 2000. Quantification error versus location error in the comparison of categorical maps. Photogrammetric Engineering and Remote Sensing 66 (8), p. 1011–1016.

Pontius Jr., R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. International Journal of Remote Sensing 32 (15), p. 4407–4429.

Qiong Wu, Hong-qing Li , Ru-song Wang , Juergen Paulussen,Yong He , Min Wang , Bi-hui Wang , Zhen Wang, 2006. Monitoring and predicting land use change in Beijing using remote sensing and GIS, Landscape and Urban Planning 78, p. 322–333.

Richards, J.F., 1990. Land transformation. In:Turner II, B.L., et al. (Eds.), The Earth as Transformed by Human Action. Cambridge University Press, New York, p. 163–178.

Verburg, P.H., Overmars, K.P., Huigen, M.G.A., de Groot, W.T., Veldkamp, A., 2006. Analysis of the effects of land use change on protected areas in the Philippines.Appl. Geogr. 26, 153–173.