On Markov Process and Simulation of Land Use Change

[1]Kennet G. Cuarteros

Abstract

Land use is a major activity happening in many parts of the world, hence several rules and regulations have been implemented to minimize the cause of destruction and over-usage of the land. Remote Sensing and GIS technologies have been the common technique of absorbing data for land use but these methods are costly and time-consuming. Other studies produce the transition matrices through cross-tabulation. In this study, the transition matrices of the actual land use data gathered from the Cagayan de Oro Socio-Economic Profile were computed through Simulation using the statistical software. The transition probability matrix with the least error was taken as the predictor of the state vector. The future land use of Cagayan de Oro was then projected by an application of stochastic modeling through the use of a Markov process. The results indicate that for the next seven years there will be a substantial agricultural and forest land loss, and an increase in urban land and other land types. The results also indicate a stabilization of land use. Hence, the Markov process is a useful tool in forecasting land use change. To better acquire observed data, it is suggested to use GIS technologies and remote sensing.

Keywords*: Land use, Simulation, Markov process, Prediction, Philippines*

Corresponding Author: kennet.cuarteros@ustp.edu.ph

1.0 Introduction

Land use is a major issue and challenge for the town and country planners to design eco-friendly and sustainable economic growth. The causes in change for land use may be due to socio-economic development or due to changes in the environment or both. Land can be gained by clearing forests or by converting agricultural fields [1].

Urban studies are becoming important tools for planners knowing that in 2015 more than half of the world's population will be living in cities [2]. Urban studies include fields such as sociology, economics, political science, anthropology, geography, and urban planning and design. Architecture, and Urban growth, mostly the movement of residential and commercial land use to rural areas at the edge of cities and towns, has long been considered a sign of regional economic vitality. However, its benefits are increasingly balanced against ecosystems including degradation of environmental quality, especially loss of farmland and forests, and also socioeconomic effects [3]. Many cities and towns in Africa have seen marked increases in urban growth and the associated impacts of environmental degradation [4]. To support sustainable development, local, regional land managers and policymakers need to consider carefully the land changes and its consequences.

The land use change in large city areas is a complicated process; several factors have influenced this process, including both physical aspects and human aspects. On the one hand, accelerated urban expansion is usually associated with and driven by social-economic factors; on the other hand, the process of urbanization has a considerable impact on the economics of the society in that area. For substantial development, municipal authorities need tools to monitor how the land is currently used, assess future demand, and take steps to assure adequacy of future supply; for better planning of future urban development, municipal authorities need to know the situation of urban expansion and in what way it is likely to move in the years to come [5].

The models of land use and land cover change process fall into two groups: regression-based and spatial transition-based models. The majority of research utilizes a regression-based approach, which relates the locations of land use and land cover change to a set of spatially explicit variables, and uses models such as logistic [6,7,8], and hedonic price models [9]. Spatial transition-based models often refer to cellular automation simulation models, which allow for predicting future land development based on probabilistic estimates with Monte Carlo or other methods [10,11]. One crucial limit to the development of the process models is, however, the deficiency of explicit modeling tools for change processes in the current generation of remote sensing and geographical information system (GIS). Equally important is the issue of data availability [12]. Moreover, few studies have attempted to link satellite remote sensing and GIS to stochastic modeling methods in land use and land cover change studies, even though the techniques for such linkages have become mature in recent years due to advances in the technology of GIS and its integration with remote sensing. This missing linkage has hindered modeling and assessing the dynamics of land use and land cover change, and significantly impeded progress towards the understanding of earth-atmosphere interactions, biodiversity loss, and global environmental change.

The city of Cagayan de Oro (CDO) has already taken into account how to use its land with the aid of the 1979 Zoning Ordinances. The Housing and Land Use Regulatory Board (HLURB) is the government institution that enacts the rules and regulations of using land, not to mention also Department of Environment and Natural Resources (DENR) and other land-related agencies. HLURB also does the land use planning for the incoming year. Realistically, these rules and regulations may not be followed. A lot of negotiations will be done, and the final decision would still be made by the Mayor of the city.

CDO is one of the most progressive cities in the Philippines. Economists, national and local officials said that CDO is blooming, blossoming, and booming. The metropolis provides lots of opportunities to attract a huge amount of migrants from all over Northern Mindanao. This kind of increasing population pressure is putting an adverse impact on CDO city rapidly transforming water bodies and vegetation lands to human settlements, workplaces, hotels and malls, and other infrastructures such as roads and bridges. If the population and urban growth trend of CDO continue over the next few years then it is clear how spontaneously the city is growing. If this situation continues, the city will soon become an urban slum with the least livable conditions for the city dwellers. The rapid economic development has brought about fundamental changes in land use and land cover patterns. The study aims to understand the land use and land cover dynamics using the techniques of simulation and stochastic modeling.

This paper presents a method that combines simulation and Markov modeling to analyze and predict land use and land cover changes in Cagayan de Oro City between 1995 and 2011 with a total land area of 48,885.83 hectares. The yearly transitional probability matrices were derived through a random simulation using statistical software. Markovian modeling is then used to examine the land cover change data and to predict the stability of future land in the city.

Table 1 presents the land use classification of the study area.

[1]University of Science and Technology of Southern Philippines

Table 1. Land use classification scheme in the study area

Land Use Type	Description
Agricultural	Crop fields, fallow lands, and vegetable lands
Forest	Open spaces, deciduous forest, mixed forest lands, palm, and others
Urban	Residential, commercial and services, industrial, socio-economic infrastructures, mixed built up and other built-up areas, transportation, roads and airport
Miscellaneous	Government properties, institutional, exposed soils, landfill sites, exempt properties, special class, and areas of active excavation

2.0 Basic Concept and Methodology

2.1 Markov Chains

Markov chain models have several assumptions [13,14,15,16]. One basic assumption is to regard land use and land cover change as a stochastic process, and different categories are the states of a chain. A chain is defined as a stochastic process having the property that the value of the process at time *t*, X_t, depends only on its value at time t-1, X_{t-2}, and not on the sequence of values X_{t-2}, X_{t-3}… X_0, that the process passed through in arriving at. It can be expressed as

$$P\left(X_t = a_j \middle| X_0 = a_0, X_1 = a_1, \dots, X_{t-1} = a_i\right) = P\left(X_t = a_j \middle| X_{t-1} = a_i\right) \quad (1)$$

Moreover, it is convenient to regard the change process as one which is discrete in time (*t*=0,12,...). The $P(X_t=a_i \mid X_{t-1}=a_i)$, known as the one-step transitional probability, gives the probability that the process makes the transition from state a_i to a_j in one time period. When *n* steps are needed to implement this transition, the $P(X_t=a_i \mid X_{t-1}=a_i)$ is then called *n*-step transition probability, $P_{ij}^{(n)}$. If the $P_{ij}^{(n)}$ is independent of times and dependent only upon states a_i , a_j and *t*, then the Markov chain is said to be homogeneous. The treatment of Markov chains in this study will be limited to first-order homogenous Markov chains. In this event

$$P\left(X_t = a_j \middle| X_{t-1} = a_i\right) = P_{ij} \quad (2)$$

where P_{ij} data by tabulating the number of times the observed data went from state *i* to *j*, n_{ij}, and by summing the number of times that state a_i occurred, n_i. Then

$$P_{ij} = \frac{n_{ij}}{n_i} \quad (3)$$

As the Markov chain advances in time, the probability of being in state *j* after a sufficiently large number of steps becomes independent of the initial state of the chain. When this situation occurs, the chain is said to have reached a steady-state. Then the limit probability, P_j, is used to determine the value of $P_{ij}^{(n)}$:

$$\lim_{n} P_{ij}{}^{(n)} = P_j \quad (4)$$

2.2 Methodology

- Acquire land use statistics of Cagayan de Oro from the City Assessors Office or Cagayan de Oro Socio-Economic Profile
- Simulate the transitional probability matrix of the actual land use of Cagayan de Oro through a statistical software program.
- Forecast the land use change of Cagayan de Oro and validate results.

3.0 Highlights of the findings

The statistics of land use in the Cagayan de Oro Socio-Economic Profile is divided into two major land types; agricultural and non-agricultural. Non-agricultural land type is distributed into residential, commercial, industrial, special class, other properties, exempt properties, and open spaces. For non-complexity purposes, the researcher classified the land use into

The yearly transitional probability matrices are calculated through simulation using statistical software. Codes are developed and several iterations were made to come up with less than 0.01 or 1% error.

The yearly derivation of the transitional probability matrices shows that they are not dependent on time or non-homogeneous. Thus, the average of the probability transition matrices is calculated. The average transitional probability matrix then is:

$$P = \begin{array}{c} \\ A \\ F \\ U \\ M \end{array} \begin{array}{c} \begin{array}{cccc} A & F & U & M \end{array} \\ \begin{bmatrix} 0.4572 & 0.4069 & 0.0984 & 0.0375 \\ 0.4318 & 0.4245 & 0.1153 & 0.0284 \\ 0.3485 & 0.3028 & 0.2818 & 0.0669 \\ 0.3554 & 0.1959 & 0.2275 & 0.2212 \end{bmatrix} \end{array} \quad (5)$$

It denotes that 45.72% of the agricultural land will remain. The probability that the agricultural land will be converted into forest land is 32.18%, its probability to be converted into an urban is 40.85%, and its probability to be converted into other land types is 37.54%. The probability of the forest land to be converted into agricultural is 35.69%, 42.45% chance to remain forest land, 40.28% to be converted into urban, and 19.59% into other land types. And so on.

Using this average transition matrix, the future land use of Cagayan de Oro can be constructed. Based on the observed state vectors, the error of each transitional probability matrix was computed with the expected state vectors, that is:

$$Error = \frac{|Observed - Expected|}{Observed} \quad (6)$$

Using the statistical software, the prediction of the land use statistics or the state vectors were obtained. The mean error for the year 1995 is zero because it is the first observed state which means that the Markov process is yet to be conducted for that year. The transitional probability matrix for the year 2006-2007 is an identity matrix since no conversion took place on the actual data. 2011 is also predicted for validation.

To forecast the land use/cover, we will choose any initial state vector. Suppose we begin with $X^{(0)}=[0\ 1\ 0\ 0]^T$ which means that in the initial year, the land use of Cagayan de Oro is 100% forested.

For next year:	$X^{(1)} = PX^{(0)}[0.4069 \quad 0.4245 \quad 0.3028 \quad 0.1959]^T$
For the next two years:	$X^{(2)} = PX^{(1)}[0.3959 \quad 0.3964 \quad 0.3688 \quad 0.3400]^T$
For the next three years:	$X^{(3)} = PX^{(2)}[0.3913 \quad 0.3914 \quad 0.3847 \quad 0.3775]^T$
For the next four years:	$X^{(4)} = PX^{(3)}[0.3902 \quad 0.3902 \quad 0.3885 \quad 0.3868]^T$
For the next five years:	$X^{(5)} = PX^{(4)}[0.3899 \quad 0.3899 \quad 0.3895 \quad 0.3891]^T$
For the next six years:	$X^{(6)} = PX^{(5)}[0.3898 \quad 0.3898 \quad 0.3897 \quad 0.3896]^T$
For the next seven years:	$X^{(7)} = PX^{(6)}[0.3898 \quad 0.3898 \quad 0.3898 \quad 0.3898]^T$

The results above may vary depending on the assumption of the initial state vector. $X^{(1)}= PX^{(0)}[0.4069\ 0.4245\ 0.3028\ 0.1959]^T$ will serve as the validation period. Thus, in comparison with the actual land use data of 2011, the mean error of the forecast is 19.06 as shown in Table 2.

Table 1. Validation of the Prediction

	2011 Land Use		
	Actual	**Predicted**	**Error**
Agricultural	19740.6741	21108.9014	0.0693
Forest	19563.8374	20747.1463	0.0605
Urban	7195.1795	5636.5362	0.2166
Miscellaneous	2386.1387	1393.2462	0.4161
		Average Error	0.1906

It is expected to have a large error for the prediction of the 2011 land use. This scenario is caused by using the average transition matrix for the prediction. But as the years go on until the forecast reaches 2017, the error tends to lower down to 5%.

Table 3 shows the actual land use from 1995 to 2011 and the forecasted land use from 2012 to 2017. From the table, it is clear that there has been a considerable change. In 2012, Agricultural land and forest land decreased in area by 6.73% and 5.05%, respectively. An increase of urban and other land types by 30.66% and 4.63%, respectively. In 2013 before 2011, agricultural and forest land decreased in area by 7.80% and 6.68%, respectively. Urban and other land types increased in area by 21.66% and 41.61%, respectively.

Table 3. Actual and Forecasted Land Use of Cagayan de Oro City

Year	Agricultural	Forest	Urban	Miscellaneous	Total
1995	21845.4453	18774.8150	5514.3132	2751.2565	48885.83
1996	21903.9611	18714.4794	5516.1332	2751.2563	48885.83
1997	21845.4453	18774.8152	5514.3132	2751.2563	48885.83
1998	20713.7880	20639.6984	6314.6318	1217.7118	48885.83
1999	20742.0100	19439.8800	6172.0700	2531.8700	48885.83
2000	21133.8303	20208.7935	6318.5164	1224.6898	48885.83
2001	21131.4799	20217.3702	6316.1644	1220.8154	48885.83
2002	21048.3335	19790.1421	6676.7438	1370.6106	48885.83
2003	19589.5108	21496.2340	6020.9951	1779.0901	48885.83
2004	20508.1529	17362.0465	7380.7847	3634.8459	48885.83
2005	21374.7473	18133.2343	7783.8779	1593.9705	48885.83
2006	20337.8088	18957.1650	7992.5666	1598.2896	48885.83
2007	20337.8088	18957.1650	7992.5666	1598.2896	48885.83
2008	20337.6920	18960.8335	7984.0018	1603.3027	48885.83
2009	19335.2742	18966.8399	6702.4460	3881.2699	48885.83
2010	20404.7570	16877.9964	7225.7723	4377.3043	48885.83
2011	19740.6741	19563.8374	7195.1795	2386.1387	48885.83
2012	18412.1267	18575.8636	9401.22153	2496.61692	48885.83
2013	18312.7012	18338.0926	9687.95879	2547.07355	48885.83
2014	18283.4009	18250.2072	9766.43126	2585.78907	48885.83
2015	18279.7442	18229.0443	9780.10426	2596.93382	48885.83
2016	18276.0883	18220.1485	9786.95033	2602.64707	48885.83
2017	18276.0883	18220.1485	9786.95033	2602.64707	48885.83

The results showed the same stabilization in the seventh year. Different initial state vectors come up with the same steady-state vector. As the state vector approaches the steady-state vector, the mean error becomes smaller.

4.0 Conclusion and Recommendations

This study has illustrated an application of Monte Carlo methods in computing for the transitional probability matrices of the actual land use data. Mathematical modeling has also been used. Markov chains in particular have been demonstrated in predicting land use of Cagayan de Oro.

Simulation using the statistical software of the transitional probability matrices of the land use is highly efficient with a prediction error of less than 1%. The yearly prediction error lowers as it approaches the steady-state vector. Different initial states lead to the same steady-state vector.

The prediction of forest land to other land classifications may be caused by several factors. The results showed a continuing trend of increase in urban and other miscellaneous covers and decline in agriculture and forest land.

The acquisition of the actual data transitional matrices of land use is not always available. If these transitional matrices are readily available, the results would be more efficient. There has been a popular method of finding the transitional matrices of land use, that is, the use of GIS technologies and remote sensing. Maps, specifically Landsat Thematic Mapper data were commonly used and are analyzed. More time should be invested though and skills are needed but this is the surest way to have observed data.

5.0 References

K. Singh, (Netherlands, 2003) Modelling land use land cover change using cellular automata in a geo-spatial environment. *International Institute for Geo-Information Science and Earth Observation Enschede*.

UNECE, (2003) *Trends in Europe and North America-The statistical yearbook of the economic commission for Europe 2003,* United Nations Economic Committee for Europe. Retrieved from http://www.unece.org/stats/trends/

G. D. Squire, (2002) Urban sprawl and the uneven development of metropolitan America. In Gregory D. Squires (Ed.), *Urban Sprawl: Causes, consequences, and policy responses* (pp. 1-22). Washington, D.C.: Urban Institute Press.

C. Kamusoko and M. Aniya, (2007)Land use/cover change and landscape fragmentation analysis in the Bindura District, Zimbabwe. *Land Degradation and Development 18(*2):221 – 233.

W. L. Huang, H. P. Liu, Q. Z. Luan, Q. Jiang, J. Liu and H. Liu, (2008) Detection and prediction of land use change in Beijing based on remote sensing and GIS. *ISPRS 37*(6b):75–82.

J. D. Landis, (1994) The California urban future model: a new-generation of metropolitan simulation-models. *Environ Plan B. 21*:399–420.

M.G.Turner,(1987)Spatialsimulationoflandscapechangesin Georgia: a comparison of three transition models. *Landscape Ecol. 1*:29–36.

D. N. Wear, M. G. Turner and R. J. Naiman, (1998) Land cover along an urban-rural gradient: implications for water quality. *Ecol Appl. 8*:619–630.

J. Geoghegan, L. A. Wainger and N. E. Bockstael, (1997) Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS, *Ecological Economics 23*: 251-264.

K. C. Clarke, S. Hoppen and L. Gaydos, (1997) A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environ Plan B: Plan Des. 24:*247–261.

K.C.ClarkeandL.J.Gaydos,(1998)Loose-couplingacellularautomaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science, 12*:699–714.

W. L. Baker, (1989)A review of models of landscape change *Landscape Ecol. 2,* 111–133.

E. Parzen, (1962) Stochastic processes. San Francisco: Holden-Day.

C. T. Haan, (1977) *Statistical methods in hydrology*. Ames, Iowa: The Iowa State University Press.

X. Wang, (1986) *Multivariate statistical analysis of geological data*. Beijing, China: Science Press.

W. J. Stewart, (1994) *Introduction to the numerical solution of Markov chains.* Princeton, NJ: Princeton University Press.