Coupling of Markov chains and cellular automata spatial models to predict land cover changes (case study: upper Ci Leungsi catchment area)

K. Marko, F. Zulkarnain, and E. Kusratmoko
Department of Geography, Faculty of Mathematics and Natural Science, Universitas Indonesia, Kampus UI Depok, Depok 16424, Indonesia

Email : kuswantoro@ui.ac.id; kusrif@yahoo.com

Abstract. Land cover changes particular in urban catchment area has been rapidly occur. Land cover changes occur as a result of increasing demand for built-up area. Various kinds of environmental and hydrological problems e.g. floods and urban heat island can happen if the changes are uncontrolled. This study aims to predict land cover changes using coupling of Markov chains and cellular automata. One of the most rapid land cover changes is occurs at upper Ci Leungsi catchment area that located near Bekasi City and Jakarta Metropolitan Area. Markov chains has a good ability to predict the probability of change statistically while cellular automata be believed as a powerful method in reading the spatial patterns of change. Temporal land cover data was obtained by remote sensing satellite imageries. In addition, this study also used multi-criteria analysis to determine which driving factor that could stimulate the changes such as proximity, elevation, and slope. Coupling of these two methods could give better prediction model rather than just using it separately. The prediction model was validated using existing 2015 land cover data and shown a satisfactory kappa coefficient. The most significant increasing land cover is built-up area from 24% to 53%.

1. Introduction
Land cover changes is a complicated, dynamic, and spatially explicit phenomena. Land cover changes affected by many factors related spatial-temporal phenomena, for example: proximity, physical conditions, and also socio-economics factors [1,2,3]. In order to detect land cover changes, geographic information system (GIS) and remote sensing are powerful tools to generate land cover changes effectively and efficiently [4,5,6]. Furthermore, advanced GIS technique can even predict land cover changes [7].

Over last decades, some studies trying to find out the best method for predicting land cover changes. In general, there are three types of land cover changes modelling i.e. empirical/statistical model, dynamic model, and hybrid model. A hybrid model known as a better model than empirical/statistical model and dynamic model because it combined two advantages both empirical/statistical model and dynamic model [8]. One of hybrid model that commonly used to predict land cover changes is geographic information system (GIS) based Markov chains-cellular automata model. This research is coupling two different methods: Markov chains, which is empirical/statistical model, and cellular automata, which is dynamic model that incorporate using GIS platform.
The purpose of this research is to predict future land cover changes in upper Ci Leungsi catchment area based on Markov chains-cellular automata spatial model, and to review the effectiveness of MC-CA spatial model for predicting land cover changes in catchment area scale. This research is important to recognize the future condition of upper Ci Leungsi catchment area with assumption by business as usual approach (BAU). In addition, the result of this research can be used as a reference to formularize best management practices of integrated water resources management in the future.

2. Area of Study and Data
Upper Ci Leungsi catchment area is located in Bogor Regency, about 15 kilometers from Bekasi City and 25 kilometers from Jakarta, the capital city of Indonesia. This catchment area is a subsystem of Bekasi watershed. The total catchment area is approximately 231 km² that cover five districts namely Gunung Putri, Citeureup, Klapanunggal, Babakan Madang, and Sukamamakmur.

![Figure 1. Location of study area](image)

This research used multi-temporal land cover data to process the Markov chains - cellular automata spatial models. Multi-temporal land cover obtained by digital image processing from Landsat 5 and Landsat 8 that downloaded from [http://earthexplorer.usgs.gov](http://earthexplorer.usgs.gov). This research also used physical characteristics maps such as: slope, elevation, and infrastructure proximity maps such as road network, central business district locations, and highway gate locations. Physical characteristics and infrastructure maps is using as driving factors to develop cellular automata while multi-temporal land cover is using to produce Markov transitional probabilistic map. Here is a list of data that was used in this research:

| Datasets                      | Source                      | Date or Year | Resolution |
|-------------------------------|-----------------------------|--------------|------------|
| Landsat imageries             | United States Geological Survey | Acquisitions : | Raster (30 meters) |
| • LT51220642005183BKT01       |                             | 02/07/2005   |            |
| • LT51220652005183BKT01       |                             | 02/07/2005   |            |
| • LT51220642010213BKT01       |                             | 01/08/2010   |            |
| • LT51220652010213BKT01       |                             | 01/08/2010   |            |
| • LC81220642014256LGN00       |                             | 13/09/2014   |            |
| • LC81220652014256LGN00       |                             | 13/09/2014   |            |
| Digital elevation model       | ASTER Imagery               | 2009         | Raster (30 meters) |
| Road network map              | Badan Informasi Geospasial  | 2009         | Vector     |
3. Methodology

Markov chains is a dynamic processes based on Markovian random process that calculate the probability of changes from particular object (for example: forest) into another objects (for example: built-up area). Markov chains describes as transitional probability matrix [1]:

\[
p = (P_{ij}) = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\
p_{21} & p_{22} & \cdots & p_{2n} \\
p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}
\]

The probability of changes from particular object (ith) into another object (jth) is described as a transformation probability (\(P_{ij}\)); n is the land cover types of the catchment area. \(P_{ij}\) must meet the following condition below [1]:

\[
0 \leq P_{ij} \leq 1 (i, j = 1, 2, 3, ..., n) \\
\sum_{i=1}^{n} P_{ij} = 1 (i, j = 1, 2, 3, ..., n)
\]

Based on non-after effect of Markovian random processes and probability of Bayes condition, Markov chains model is obtained by [1]:

\[
P_{(n)} = P_{(n-1)} p_{ij}
\]

Where:

- \(P_{(n)}\) is state probability of any times
- \(P_{(n-1)}\) is preliminary state probability

Cellular automata developed by Von Neumann for self-reproducible systems [8]. Cellular automata characteristics are spatially explicit and dynamics that can simulation the evolution of two-dimensional objects [9, 10]. GIS raster data is known have resemblance with cellular automata concept hence many studies generate future land cover using cellular automata [9]. In order to generate future land cover map, there are many variables that drive the change of objects, called driving force [11]. Driving force consists of many factor for example: proximity, elevation, slope, etc. [1, 2, 3, 11].

Markov chains is just calculating the probability of changes but cannot represent the changes in a spatial-explicit term [12, 13]. Combining Markov chains and cellular automata is effort to minimalize the weakness of Markov chains [13]. Cellular automata represent the spatial-explicit term and location of changes while Markov chains predict changes quantitatively. Combining two methods would give better result than perform a model separately [14]. The procedure of Markov chains-cellular automata model is presented in figure below:

**Figure 2.** The procedure of Markov chains – cellular automata model
Accuracy test was conducted in order to know how the reliability, capability, and limitation of Markov chains is – cellular automata spatial models to predict land cover changes. Accuracy test of land cover model is an important part of the model that become the standard evaluation model whether the result of the model can be used for policy maker or not [21, 22]. The accuracy of the predicted land cover model was assessed using the kappa statistics. Using the same methods and steps, Markov chains – cellular automata spatial model was applied to create the projection of 2014 land cover based on 2005 land cover and 2010 land cover then the 2014 projected model is being evaluated with kappa statistics based on 400 ground truth derived from high resolution satellite imagery, Quickbird images. The kappa statistics calculated by equation below:

\[
K = \frac{N \sum_{i=1}^{r} X_{ii} - \sum_{i=1}^{r} (X_{ij} \cdot X_{ji})}{N^2 - \sum_{i=1}^{r} (X_{ij} \cdot X_{ji})}
\]

4. Result and Discussions

4.1. Land Cover Changes Analysis

4.1.1. Spatial Pattern of land cover changes

The classification of land cover divided into 7 types i.e. built-up area (BA), forest (FO), shrubs (SH), bare land/mining (BM), cultivated area (CA), paddy field (PF), and water bodies (WB). Over 10 years, built-up area is growing rapidly in Upper Ci Leungsi catchment area. Generally, the spreading of built-up areas concentrated in the central and northern part of the catchment. The spatial pattern of built-up area also indicated the linearity following the road network. Land cover changes can be seen in figure below:

![Figure 3. Land cover changes between 2005 to 2010 and 2010 to 2014](image-url)
The figure above shows that built-up area spreads significantly from 14% to 23% in period time 2005 to 2010 and 23% to 28% in period time 2010 to 2014. Over a decade, built-up area grew double. On the other hand, other types of land cover such as shrubs and forest is declining distinctly. Forest was declining from 18% to 9% while shrubs was also declining from 58% to 52% for over ten years. It is an evidence that there were much area of shrubs and forest that converted into built-up area between 2005 to 2014.

4.1.2. Transitional probability matrix of land cover changes
Transitional probability matrix is matrix that calculate the probability changing of each types of pixel of land cover into another types of land cover based on Markovian random process [15, 16]. The probability describes in a range of value between 0-1 which zero value means that the event of changing is impossible happen and 1 means is certainly happening. The transitional probability matrix of is shown below:

| From | Period | Land cover | To      | FO | SH | BM | CA | BA | PF | WB |
|------|--------|------------|---------|----|----|----|----|----|----|----|
| 05-10| FO     |            | FO      | 0.54| 0.29| 0.00| 0.16| 0.01| 0.01| 0.00|
| 10-14| FO     |            | SH      | 0.01| 0.83| 0.00| 0.03| 0.10| 0.03| 0.00|
| 05-10| SH     |            | BM      | 0.00| 0.75| 0.01| 0.05| 0.12| 0.07| 0.00|
| 10-14| SH     |            | BM      | 0.00| 0.03| 0.69| 0.00| 0.29| 0.00| 0.00|
| 05-10| BM     |            | CA      | 0.00| 0.68| 0.00| 0.19| 0.04| 0.01| 0.00|
| 10-14| BM     |            | CA      | 0.01| 0.62| 0.01| 0.30| 0.04| 0.02| 0.00|
| 05-10| CA     |            | BA      | 0.00| 0.00| 0.00| 0.00| 1.00| 0.00| 0.00|
| 10-14| CA     |            | BA      | 0.00| 0.00| 0.00| 0.00| 1.00| 0.00| 0.00|
| 05-10| BA     |            | PF      | 0.00| 0.00| 0.00| 0.00| 0.24| 0.12| 0.00|
| 10-14| BA     |            | PF      | 0.08| 0.00| 0.00| 0.06| 0.27| 0.07| 0.00|
| 05-10| PF     |            | WB      | 0.00| 0.00| 0.00| 0.00| 0.00| 0.00| 1.00|
| 10-14| PF     |            | WB      | 0.00| 0.01| 0.00| 0.00| 0.00| 0.00| 1.00|

Based on transitional probability matrix of land cover above, five types of changing have higher probability (higher than 0.2) than the other i.e. forest becomes shrubs, bare land/mining becomes built-up area, paddy field becomes built-up area. The other transitional probability is relatively low therefore the probability of the changes is not significant.

4.2. Land cover Prediction

4.2.1. Driving Factor
Driving factors are used to determine the location that may changes from each pixel of land cover into another type [17]. Driving factors divided into three types, slope, elevation, and infrastructure’s proximity. Infrastructure’s proximity then itemized into specific types that could stimulates the changes such as proximity from road, proximity from highway gates, and proximity from central business district.

Slope is one of important factor that driving the land cover changes [13]. A steep slope could become limiting factor for built-up area. Most of built-up area is found in relatively flat area because related to the easiness of structure and building construction. Furthermore, steep slope is vulnerable.
for mass movement that could stimulate landslide so it become another reason why slope is a limiting factor for built-up area [18].

Proximity also becomes important factor particular for built-up area invasion. Built-up area will occupy the land that close to vital infrastructure i.e. road networks, central business districts, and highway gates. Determining proximity of highway gates that become one of driving factor is unique because highway is the most important infrastructure for residents to access the capital city of Indonesia, Jakarta where the majority of Ci Leungsi catchment’s residents work there. Ci Leungsi catchment located in peri-urban of Jakarta that many residents of this catchment depends on Jakarta and use highway as a major road to access Jakarta [19, 20]. All of driving factor then will be used for cellular automata based on multi-criteria analysis.

4.2.2. Land cover Change Prediction
Land cover change prediction in 2020, 2025, 2030 is shown in figure 7. The prediction shows expansion of built-up area in the central of catchment area. The expansion might be happening because the central of catchment area is suitable for the development of built-up area since it is
physical characteristics (low to moderate elevation and flat slope) and its proximity to the infrastructure. It also shows the disappearing of paddy field and shrubs that might be converted into built-up area in the future. The forest predicted still exist because pixels of forest located in limiting factor, which has characteristics: steep slope, high elevation, far from road network, central business districts, and highway gate. This limiting factor makes forest might still exist and not decreasing significantly in the future. Based on figure 8 below, built-up area is predicted increasing from 28% to 39%, 46%, and 53% in 2020-2030. Almost all of types except built-up area and water bodies is predicted decreasing. The most decreasing area is shrubs, disappears about 24% from 2014 to 2030.

Figure 5. Land cover prediction (left: 2020, center: 2025, right: 2030)
4.2.3. Accuracy Test of Prediction Model

The overall kappa coefficient test table is shown that the prediction model is 82.58%, tested from 400 ground truth data that acquired from Quickbird image. The best user’s accuracy is built-up area, forest, and water bodies which above 90%. Other types of land cover i.e. shrubs and bare land/mining has adequately accuracy which above 80% but cultivated area and paddy field has lower accuracy than the other land cover, just about below 80%.

Table 3. Kappa coefficient test table

| land cover | ground truth | | ground Truth | user's accuracy (%) |
|------------|--------------|--------|--------------|---------------------|
| FO         | ref. 1 (FO)  | ref.2 (SH) | ref.3 (BM)  | ref. 4 (CA) | ref. 5 (BA) | ref. 6 (PF) | ref. 7 (WB) |          |          |
| SH         | -            | 107     | -            | 5         | 8         | -            | -            | 120       | 89.17    |
| BM         | -            | 5       | 24           | 1         | -         | -            | -            | 30        | 80.00    |
| CA         | -            | 12      | -            | 18        | -         | -            | -            | 30        | 60.00    |
| BA         | -            | 9       | -            | -         | 111       | -            | -            | 120       | 92.50    |
| PF         | -            | -       | -            | 4         | 7         | 19           | -            | 30        | 63.33    |
| WB         | -            | -       | -            | -         | -         | -            | 30           | 30        | 100.00   |
| Ground Truth | 37      | 136     | 24           | 28        | 126       | 19           | 30           | 400       | 82.50    |

Kappa Coefficient 82.58

5. Conclusion and Recommendation

A business as usual land cover projection model in Ci Leungsi catchment was successfully executed using coupling of Markov chains and cellular automata spatial model with adequately kappa coefficient. The result of projection models showed that built-up area will increase almost doubled from 24 to 53%. Almost of land cover types predicted accurately except cultivated area and paddy field. For next research, developing some land cover change scenarios based on land use policy and
best land use management probably can improve the accuracy of projection model and also finding the most influential driving factor that changing the land cover. There are some opportunities to apply the projection model for improving best management practices in Ci Leungsi Catchment to support integrated water resources management program. This research is also very useful as a reference for land use planning in watershed scale for government and stakeholders.

Acknowledgments
The authors would like to thank the Department of Geography, Faculty of Mathematics and Natural Science, University of Indonesia which has facilitated the research activities through excellent research grants.

References
[1] Guan D, Li H, Inohae T, Su W, Nagaie T, Hokao K 2011 Modeling urban land use change by the integration of cellular automaton and markov model Ecological Modelling 222 3761-3772
[2] Al-Sharif A A A and Pradhan B 2014 Monitoring and predicting land use change in Tripoli Metropolitan City using an integrated Markov chain and cellular automata model in GIS Arabian Journal of Geosciences 7 4291-4301
[3] Fuglsang M, Münier B, Sten Hansen H 2013 Modelling land-use effects of future urbanization using cellular automata: An Eastern Danish case Environmental Modelling & Software 50 1-11
[4] Leao S, Bishop I, Evans D 2004 Simulation urban growth in a developing nation’s region using a cellular automata-based model Journal of Urban Planning and Development 130 145-158
[5] Batty M and Xie Y 1997 Possible urban automata Environmental and Planning B: Planning & Design 24 175-192
[6] Dadras M, Shafri H Z M, Ahmad N, Pradhan B, Safarpour S 2015 Spatio-temporal analysis of urban growth from remote sensing data in Bandar Abbas city, Iran The Egyptian Journal of Remote Sensing and Space Sciences 18 35-52
[7] Batty M, Xie Y, Sun Z 1999 Modeling urban dynamics through GIS-based cellular automata Computers, Environment and Urban Systems 23 205-233
[8] Von Neumann J and Burks A W 1966 Theory of Self-reproducing Automata (Urbana: University of Illinois Press)
[9] Ye B and Bai Z 2008 Simulating land use/cover changes of Nenjiang county based on CA-Markov model Computer and Computing Technologies in Agriculture 1 321-329
[10] Long Y, Jin X, Yang X, Zhou Y 2014 Reconstruction of historical arable land use patterns using constrained cellular automata: A case study of Jiangsu, China Applied Geography 52 64-77
[11] Lauf S, Haase D, Hostert P, Lakes T, Kleinschmit B 2012 Uncovering land-use dynamics driven by human decision-making – A combined model approach using cellular automata and system dynamics Environmental Modelling & Software 27-28 71-82
[12] Lopez E, Bocco G, Mendoza M, Duhau E 2001 Predicting land-cover and land-use change in the urban fringe A case in Morelia city, Mexico Landscape and Urban Planning 55 271-285
[13] Shafizadeh Moghadam H and Helbich M 2013 Spatiotemporal urbanization processes in the megacity of Mumbai, India: a Markov chains-cellular automata urban growth model Applied Geography 40 140-149
[14] Arsanjani J J, Helbich M, Kainz W, Boloorani A D 2011 Integration of logistic regression, Markov chains and cellular automata models to simulate urban expansion International Journal of Applied Earth Observation and Geoinformation 21 264-272
[15] Eastman J R 2009 IDRISI 16: The Taiga Edition (Worcester: Clark University)
[16] Brown D G, Pijanowski B C, Duh J D 2000 Modeling the relationship between land use and land cover on private lands in the Upper Midwest, USA Journal of Environmental
[17] Halmy W W A, Gessler P E, Hicke J A, Salem B H 2015 Land use/land cover change detection and prediction in the north-western coastal desert of Egypt using Markov-CA Applied Geography 63 101-112

[18] Vanacker V, Vanderschaeghe M, Govers G, Willems E, Poesen J, Deckers J, De Bievre D 2003 Linking hydrological, infinite slope stability and land-use change models through GIS for assessing the impact of deforestation on slope stability in high Andean watersheds Geomorphology 52 299-315

[19] Winarso H, Hudalah D, Firman T 2015 Peri-urban transformation in the Jakarta metropolitan area Habitat International 49 221-229

[20] Firman T 2009 The continuity and change in mega-urbanization in Indonesia: A survey of Jakarta-Bandung Region (JBR) development Habitat International 33 327-339