Cellular Automata and Markov Chain Spatial Modeling for Residential Area Carrying Capacity in Samarinda City, East Kalimantan Province

N. I. Fitri¹, A. Damayanti¹, T. L. Indra¹ and M. Dimyati¹

¹ Department Geography, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok City, Indonesia.

Abstract. The increase of population could lead to residential area availability. This could lead to imbalance between population and available housing and could result in higher population pressure on the available area. Spatial modeling prediction is needed as a prevention step to prevent excessive land cover change in the future. This research aims to analyze residential area carrying capacity and spatial modelling of land cover change of Samarinda in 2006, 2014 and 2020 using Cellular Automata Markov Chain (CAMC) and residential area carrying capacity index. The Cellular Automata Markov Chain (CAMC) results show that there is an expansion of residential area land cover which affected by driving factors that consist of distance from the nearby road, distance from the river, distance from the point of interest (health facility and education facility), slope, and elevation. Residential area land carrying capacity affected by population density, standard needed land area, and residential area extent in Samarinda City. Thus, it is needed to analyze the model to see the available area for residential area development sustainably.

1. Introduction

Land cover change happened as a process of dynamic regional development. Samarinda City has 10 sub districts and 53 villages as its administrative area division, with a total area of 718 km², population number of 812,597 person, and population density of 37.8 person/km² [1]. Residential area carrying capacity is a capability of an area in providing a residential lot to accommodate a number of population to settle decently [2].

In the recent decades some studies were trying to find the best method to predict land cover change. Generally, there are 3 kinds of land cover change modelling. Those are empirical/statistic modelling, dynamic modelling, and hybrid modelling. Hybrid modelling known as a better method than dynamic and empirical ones because the first one combining the advantages of the later ones [3]. One of the most common hybrid modelling used in predicting land cover change is Cellular Automata Markov Chain with emphasis of spatial information [3]. Cellular Automata Markov Chain combines Markov Chain as an empirical modelling method and Cellular Automata as the dynamic modelling method which could be applied in GIS platforms [3].

Responding to the troubles of imbalance between higher population pressure and limited carrying capacity of residential area in Samarinda, this research aims to analyze the land cover change of Samarinda in 2006, 2014 and 2020, and then predicting the residential area carrying capacity in 2034 using Cellular Automata Markov Chain. This Research considered important to understand the future
conditions of Samarinda City. This research could also be used as a reference to create the best residential area development management sustainably.

2. Methods
The study area is in Samarinda City, East Kalimantan Province, Indonesia, at 117° 03 ’00’ - 117 ° 18 ’14’ East Longitude and 00° 19 ’02’ - 00° 42 ’34’ South Latitude.6°41’10’S. There are some driving factors which were used in this research such as distance from nearest street, distance from nearest river, elevation, slope, and point of interest or POI, and population number. The data collection was started by downloading United States Geological Survey (USGS) Landsat 5 with the record of 18th May 2006, Landsat 8 with the record of 23rd February 2014, and Sentinel 2A with the record of 8th January 2020.

2.1 Land Cover
Based on No. 4th Constitution of 2011 about Geospatial Information, “Land Coverer” is defined as a line that draws the border of a visible cover of an area above the earth which consists of natural or man-made landscape. Land cover could also mean as biophysical cover of earth surface which could be seen and is a result of human adjustments, activities, and behavior which done to a specific land cover to do productions, changes, or maintenances to the area in context [4]. Based on SNI 7645: 2010, there are 5 classes of basic land cover (Table 1). The aforementioned classes and its descriptions could be seen in the table.

| No. | Classes       | Definitions                                                                 |
|-----|---------------|------------------------------------------------------------------------------|
| 1.  | Residential   | Built up area which characterized by substitution of artificial land cover and oftenly waterproofed. (eg: residential area, industrial area, roads, health facilities, education facilities, etc) |
| 2.  | Agricultural  | Area which is used to cultivate plant based produces (eg: rice fields, dry croplands, plantations, etc) |
| 3.  | Non Agricultural | Area which is not being used to gain plant based produces (eg: dryland forests, wetland forests, bushes, etc) |
| 4.  | Open Lands    | Land without any natural or non natural coverings. (eg: barren fields, sand dunes, mining fields, riverbeds, etc) |
| 5.  | Water Body    | All water coverers (eg: rivers, reservoirs, seas, etc) |

Land cover change in an area affected by some factors. Those factors could be used as a rule in land cover change prediction analysis [5]. Driving factors and development progress is important to project the future. Thus, driving factor is absolutely needed to model the land cover change. Driving factors used to determine the locations which could change from each type of land cover pixels into other type [3].
2.2 Cellular Automata Markov Chain

Basically, model is a simplification of a reality or of a real world phenomenon. Cellular Automata Markov Chain is a common hybrid model to predict GIS based land cover. Cellular Automata Markov Chain combines 2 different methods which consists of Markov Chain which is an empirical/statistic model. The other one is Cellular Automata which is a dynamic model that could be inputted to a GIS platform [3].

Cellular Automata (CA) is a dynamic model which integrates space and time dimensions. Cellular Automata concepts had been developed since 1940s in Computer Sciences related fields by John Von Neumann and Stanistaw Ulam [6]. Cellular Automata model is a discrete model in a shape of cell/grid where each cell could change its status and properties based on time function, which controlled by some determined rules/scenario, and its surrounding other cells. Cellular Automata contains cells which had a certain values. Each cells could change its values following a transition rule

Markov Chain modelling is a stochastic progress which defines transition probability from one state to another using a simple sequential procedure [7]. Markov Chain is a concept which introduced for the first time by Russian Mathematician, Rusia Andrey Markov. Based on Markov’s theory, probability of an occurring event, determined by the preceding events before, and could be used to predict the next event. That probability is defined as transition probability and considered as a stationary process [6]. Markov Chain could be drawn as a transition probability matrix as follows:

\[
p = (P_{ij}) = P_{11} \quad P_{12} \quad \ldots \quad P_{1n} \\
P_{21} \quad P_{22} \quad \ldots \quad P_{2n} \\
P_{n1} \quad P_{n2} \quad \ldots \quad P_{nn}
\]

\(P_{ij}\) is a transformation probability of a land cover into another. \(P_{ij}\) should fulfill these rules:

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

Based on transition probability matrix, and Bayes probability theory, Markov Chain could be defined as follows [8]:

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

Note:

\(P_{(n)}\) = Condition’s probability in a time frame

\(P_{(n-1)}\) = Probability in the preceding conditions

2.3 Residential Area Carrying Capacity

Residential area carrying capacity is an ability of an area to provide residential lots to accommodate a number of population to settle decently [9]. The area of land requirement refers to the Indonesian National Standard (SNI) Number 03-1733-2004 on the procedures for planning a residential environment, the standard area of space / capita needs (m² / capita) is 26 m² [10]. Population number is the first indicator used to determine the suitable area for ideal residential area carrying capacity. Determination of the calculation method of population projection in 2034 can be done with 2 calculation methods, namely the arithmetic method and the geometric method. The results of calculating population projections using these 2 methods are then calculated the standard deviations and their correlations [11]. The model chosen is the model that has the smallest standard deviation and the correlation coefficient between the actual population data and the projected population close to 1[11]. Residential area carrying capacity could be determined using residential area carrying capacity formula [12] as follows:
\[
DDP_m = \frac{(L_{Pm} \cdot J_P)}{\alpha}
\]

Note:
\( DD_{Pm} \) = Residential Area Carrying Capacity
\( L_{Pm} \) = Residential field area
\( J_P \) = Population number
\( \alpha \) = Needed area coefficient

Referring from [12] the range of residential area carrying capacity are as follows:
1. If \( DD_{Pm} > 1 \) m²/capita, residential area carrying capacity is high, and could accommodate population to settle in the area.
2. If \( DD_{Pm} = 1 \) m²/capita, carrying capacity is optimal, there is a balance between settling population and the available area.
3. If \( DD_{Pm} < 1 \) m²/capita, residential area carrying capacity is too low and therefore not capable to accommodate population to settle in the area.

3. Results and Discussion
3.1 Development of Changes Land Cover Samarinda City in 2006-2020

In figure 1, it is shown that Samarinda City’s residential area and open lands land cover change were increasing, while agricultural and non agricultural area were decreasing. Then, the land cover
class which did not experiencing increase or decrease is just water body. The development of residential land cover is most directed to the north and south. The land cover that has changed the most into residential land cover is agricultural land cover with an area of 16.11 km², land cover for non-agricultural areas with an area of 4.4 km², and open land with an area of 0.18 km². In figure 2, residential land coverage composition in 2006-2020 period always shown an increase 118 km² (17%) in 2006, 128 km² (18%) in 2014 and 146 km² (22%) in 2020. Open lands coverage also shown increase in each period 20 km² (3%) in 2006, 63 km² (9%) in 2014 and 102 km² (14%) in 2020. Agricultural land coverage is reduced in each period 221 km² (30%) in 2006, 192 km² (27%) in 2014, 136 km² (19%) in 2020. Non agricultural land cover also experiencing decrease. 332 km² (46%) in 2006, 308 km² (42%) in 2014, and 307 km² (41%) in 2020. The water body class had the same area from 2006-2020, around 27 km² (4%) of total Samarinda City area.

![Figure 2. Land Cover Change Composition Year 2006-2020](image-url)
Figure 3. Samarinda Residential Area Composition Change per Sub District Year 2006-2020

Residential area composition change graphs for last 14 years from are shown in figure 3. Residential area which experiencing increase, is predicted as a residential center in the future. The sub district which experiencing aforementioned conditions were Sungai Pinang, Sungai Kunjang, Sambutan, Samarinda Utara, Samarinda Ilir, Loa Janan Ilir, Samarinda Seberang, and Palaran. The area which experiencing decrease in residential area composition was Samarinda Ulu Sub District and the sub district which had no difference is Samarinda City.
3.2 Prediction of Residential Area Carrying Capacity in 2034

Figure 4. Driving Factors of Samarinda City

Table 2. Residential Area Suitability of Samarinda City

| Residential Area Suitability | Area (Km²) | Percentage |
|-----------------------------|------------|------------|
| Suitable                    | 380        | 53%        |
| Not Suitable                | 338        | 47%        |

Table 3. Residential Area Suitable of Samarinda City

| Residential Area Suitable | Area (Km²) | Percentage |
|---------------------------|------------|------------|
| Already Used              | 160.67     | 42%        |
| Not yet Used              | 219.33     | 58%        |

Figure 4 shows overlays from driving factors. The one on the right shows the fuzzy overlay result and the one on the left shows the residential area suitability. Fuzzy overlay generated from the combination of 5 parameters which had been explained in physical conditions and infrastructures before. Fuzzy values shown in black gradient from black to white where values are continuous. Thus, the brighter the fuzzy value is, it is considered as more suitable, and the darker the fuzzy value is, it is considered more unsuitable. Residential area suitability generated from overlay of 5 parameters of physical conditions and infrastructures scoring of Samarinda City. From table 2, residential area suitability could be seen that the suitable area composed of 380 km² (53%) from total area. While the
unsuitable area composed of 338 km² (47%) from total area. From table 3, residential area suitable could be seen that the already used area composed 160.67 km² (42%) from suitable area. While the not yet used area composed 219.33 km² (58%) from suitable area.

In this research, the prediction of year 2020 land cover change is used to test the accuracy of the whole prediction by comparing the prediction and the actual land cover change with the aid of Kappa Index. Validation with Kappa index could show and determine the accepted value between land cover model (prediction of land cover year 2020) and existing land cover (google earth verification). This is done to understand the model’s level of accuracy. Kappa value between the number of row and column is maxed at 1. In figure 5 about prediction model validation with Kappa test, it is shown a K-Stadard value, this value shows 0.8949 or 89% which means this accuracy considered as very good - more than 75%. The accuracy test which resulted in very good accuracy makes this model passed to be used to predict land cover change in 2034.

Table 4. Transitional Probability Matrix year 2034

|     | I   | II  | III | IV  | V   |
|-----|-----|-----|-----|-----|-----|
| I   | 0.8596 | 0.0114 | 0.1066 | 0   | 0.0224 |
| II  | 0.1026 | 0.6057 | 0.2205 | 0   | 0.0712 |
| III | 0.0182 | 0.0037 | 0.9690 | 0   | 0.0091 |
| IV  | 0    | 0    | 0    | 1   | 0   |
| V   | 0.0253 | 0.0002 | 0.0002 | 0   | 0.9743 |

Note:
I = Non Agriculture
II = Agriculture
III = Open Lands
IV = Water Bodies
V = Residential
Table 4 shows transitional probability matrix of 2034. Romanization I shows non-agriculture area, romanization II shows agriculture area, Romanization III shows open lands area, Romanization IV shows water bodies and Romanization V shows residential area. Maximum probability value between columns and rows is 1. Higher the probability of a land cover is, higher the probability of its land cover class to experience change. Value 1 in land cover transitional probability matrix means that the land cover would never change while value 0 means there are no land cover change from a single class to another.

Form the 5 of land cover class aforementioned, it is shown that agricultural area had the smallest TPM value (0.6057) which means it has a huge probability to change into other class of land cover, which is 0.1026 into non-agricultural, 0.2205 into open lands, and 0.0712 into residential area. Then, non-agricultural area would change into agricultural area with the probability of 0.0114, open lands with probability of 0.1066 and residential area with probability of 0.0224. Open lands would change into non-agricultural area with probability of 0.0182, into agricultural area with probability of 0.0037, and into residential area with probability of 0.0091. Then, residential area would change into non-agricultural area with probability of 0.0253, into agricultural area with probability of 0.0002 and open lands with probability of 0.0002.

Figure 6. Samarinda City Land Cover Prediction Year 2034
Figure 7. Land Cover Gains and Losses year 2020 – 2034

Figure 8. Land Cover Change Composition Year 2006-2034

Figure 6 shows model prediction of 2034 land cover in Samarinda City. The land cover model year 2034 is used as a reference in determining residential area land availability to inputted into residential area carrying capacity formula. If we see figure 7 and figure 8, it shows land cover’s gains and losses in the period of last 28 years (2006-2034). Samarinda land cover change between 2020-2034 shown a significant increase of residential area land cover with 14.67 km² of change, open lands with a change of 62.62 km², and non agricultural area with a change of 11.11 km². Agricultural area had a significant decrease with a loss of 51.84 km², and non agricultural area losses up to 35.50 km².
Table 5. Calculation of Prediction of Population in 2034

| No. | Year | Total Population | Arithmetic Method Population Growth Rate | Geometric Method Population Growth Rate |
|-----|------|------------------|------------------------------------------|-----------------------------------------|
| 1   | 2006 | 588135           | 4%                                       | 3%                                      |
| 2   | 2007 | 593827           | 609469                                   | 605643                                  |
| 3   | 2008 | 602117           | 630802                                   | 623671                                  |
| 4   | 2009 | 607675           | 652136                                   | 642237                                  |
| 5   | 2010 | 732161           | 673470                                   | 661355                                  |
| 6   | 2011 | 748102           | 694803                                   | 681042                                  |
| 7   | 2012 | 764908           | 716137                                   | 701315                                  |
| 8   | 2013 | 781015           | 737471                                   | 722192                                  |
| 9   | 2014 | 797006           | 758804                                   | 743690                                  |
| 10  | 2015 | 812597           | 780138                                   | 765829                                  |
| 11  | 2016 | 828303           | 801471                                   | 786266                                  |
| 12  | 2017 | 843446           | 822805                                   | 812102                                  |
| 13  | 2018 | 858080           | 844139                                   | 836276                                  |
| 14  | 2019 | 872768           | 865472                                   | 861171                                  |
| 15  | 2020 | 886806           | 886806                                   | 886806                                  |

Standard Deviation: 95407
Correlation Coefficient: 0.943

Prediction of Population Growth in 2034: 1337150

\[ DDPm_{2034} = \frac{(160.670.000 \text{ m}^2)}{26 \text{ m}^2} \]

\[ DDPm_{2034} = 4,6 \text{ m}^2/\text{capita} \]

4. Conclusions

In last 14 years, land cover changes in Samarinda in 2006, 2014 and 2020 experienced an increase in residential and open lands class while the classes which experienced decrease were agricultural area, and non agricultural area. The land cover class which did not experienced an increase or decrease is just the water body. The development of residential land cover mostly leads to the north and south. Residential area composition which experiences increase is predicted to be the center of residential area because of its available access to more decent facilities. The sub districts which experienced increase were Sungai Pinang, Sungai Kunjang, Sambutan, Samarinda Utara, Samarinda Ilir, Loa Janan Ilir, Samarinda Seberang, and Palaran.

Based in CA-MC prediction model year 2034, residential area land coverage became 160.67 km², with a total population of 1.337.150. Therefore, the calculation of residential area carrying capacity of 2034 in Samarinda resulted in index value of 4.6 m²/capita, which means if DDPm >1 m²/capita, residential area carrying capacity is still high enough to accommodate population to built housing in the mentioned region. The meaning of 4.6 m²/capita is that the carrying capacity of residential area is still able to accommodate residents to live 4 times the total population in 2034 because it has more area to accommodate the existing population.

5. References

[1] Badan Pusat Statistik Provinsi Kalimantan Timur. (2016). Kota Samarinda dalam angka 2016. Kota Samarinda
[2] Damayanti, A. P., Hardiana, A., & Rahayu, P. (2019). Faktor-faktor yang Mempengaruhi Perkembangan Perumahan di Wilayah Pesisir Kabupaten Purworejo. Jurnal Pembangunan Wilayah dan Perencanaan Partisipatif, Volume 14, No. 2 UNS.

[3] Marko, K., Zulkarnain, F., & Kusratmoko, E. (2016). Coupling of Markov Chains And Cellular Automata Spatial Models To Predict Land Cover Changes (Case Study: Upper Ci Leungsri Catchment Area). IOP Conference Series: Earth and Environmental Science, 47, 1–10.

[4] Badan Standarasisasi Nasional. (2010). SNI 7645-2010 tentang klasifikasi penutup lahan. Jakarta.

[5] Setiady, D. (2016). Prediksi Perubahan Lahan Pertanian Sawah Sebagian Kabupaten Klaten dan Sekitarnya Menggunakan Cellular Automata dan Data Penginderaan Jauh. Jurnal Bumi Indonesia, 1-10.

[6] Susilo, B. (2011). Modeling Spatial Integration Probabilistic Markov Chain and Cellular Automata for The Study of Land Use Changes Regional Scale in DIY Yogyakarta. Gea, Vol. 11 No. 2, Program Studi Kartografi dan Penginderaan Jauh, Fakultas Geografi UGM, Pages 163-178.

[7] Ozturk, D. (2017). Modelling spatial changes in coastal areas of Samsun (Turkey) using a cellular automata-markov chain method. Tehnički vjesnik, Volume 24 (1), pages 99-107.

[8] 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.

[9] Muta'ali, L. (2015). Teknik Analisis Regional. Yogyakarta: Fakultas Geografi (BPFG) UGM.

[10] Badan Standarasisasi Nasional. (2004). SNI 03-1733-2004 tentang tata cara perencanaan lingkungan perumahan di perkotaan. Jakarta.

[11] Widayani, P. (2016). Pemanfaatan Citra Worldview-2 Untuk Analisis Kepadatan Penduduk dan Proyeksi Kebutuhan Permukiman di Kota Magelang Tahun 2012-2022. Jurnal Iptek Pertanahan, Vol. 6, 39-54.

[12] Hasmita, L., Sekarrini, C. E., & Septiana, K. N. (2020). Study of Environmental Carrying Capacity for Settlement Development in Ranah Batahan District, Pasaman Barat Regency. IOP Conf. Series: Earth and Environmental Science, 412, 1-8. doi:10.1088/1755-1315/412/1/012002

[13] Pratami, M. (2018). Pemodelan Cellular Automata Untuk Daya Dukung Lahan Permukiman di Kota Bengkulu. Tesis Magister Ilmu Geografi, Program Magister Ilmu Geografi Universitas Indonesia. Jakarta

Acknowledgments
The author would like to thank the research grant PUTI Proceedings was financially supported by Universitas Indonesia with contract number NKB-966/UN2.RST/HKP.05.00/2020, and the National Institute of Aviation and Space (LAPAN) for the Landsat 5, Landsat 8 and Sentinel-2 satellite imagery that has been provided and is the most crucial requirement in this study.