PLANNING MALAYSIA:
Journal of the Malaysian Institute of Planners
VOLUME 23 ISSUE 2 (2025), Page 363 – 378

URBAN SPRAWL TRANSITION RULE ALGORITHM CONCEPT IN CELLULAR AUTOMATA FRAMEWORK: CASE STUDY OF MALALAYANG DISTRICT, MANADO CITY, INDONESIA

Octavianus H.A. Rogi[1], Michael M. Rengkung[2], Amanda S. Sembel[3]

[1,2,3] *Urban & Regional Planning Study Program,*
SAM RATULANGI UNIVERSITY, INDONESIA

Abstract

This study aimed to develop an urban sprawl model transition rule algorithm based on land coverage indicators in a cellular automata grid framework. This study attempts to identify various attributes that impact land coverage in a spatial grid cell using statistical and spatial analysis methods. This concept provides a theoretical and methodological basis for developing a more comprehensive urban sprawl simulation model. The results indicated three determinant factors for the condition of land coverage in a cellular automata spatial grid cell: built-up land conditions, slope gradients, and the availability of road network infrastructure. The concept of the algorithm found can be expressed through the following statements: 1) built-up land in a particular spatial grid will be influenced by the condition of built-up land in neighboring grid cells, with a determination of 67%; 2) every one unit increase in the average area of built-up land in neighboring cells will be associated with the same phenomenon in a particular grid cell of 1.08 units; 3) built-up land on each spatial grid is correlated with the slope gradient and the availability of road network infrastructure on the spatial grid; 4) the flatter the slope condition of a spatial grid will be associated with higher built-up land in the spatial grid cell; and 5) the better the road infrastructure availability on a particular spatial grid will be associated with higher built-up land.

Keywords: Urban Sprawl, Transition Rule Algorithm, Cellular Automata, Built-up Land

[1] Corresponding author. Email: ottyrogi@unsrat.ac.id

INTRODUCTION

Urban sprawl, or the expansion of urban areas characterized by low-density development and outward expansion, is often viewed negatively because of its association with various environmental and social problems such as increased land consumption, traffic congestion, and air pollution. Urban planners and policymakers should implement strategies to address the challenges of urban expansion, involving a smart growth approach that emphasizes efficient development management through mixed-use land zoning, the optimization of green open spaces, and community involvement in planning. These strategies aim to improve infrastructure efficiency and reduce open space fragmentation. In addition, urban densification through vertical development or more efficient land use helps reduce the need for expansion to the outskirts, thereby protecting ecosystems and productive land around the city.

Manado City, the capital of North Sulawesi Province in Indonesia, is an urban area that is inseparable from urban sprawl. This development is characterized by low population density, unplanned development patterns, and inconsistent changes in land use, especially on the outskirts. The spatial structure of Manado City is monocentric within certain boundaries of the city center but is starting to shift to a polycentric pattern in the suburbs. In peri-urban zones, a significant challenge is urban sprawl and its associated impacts, including fragmentation of green open spaces, increased dependence on private vehicles, and higher infrastructure investment and maintenance costs (Rogi et al., 2024). This condition often occurs because of leapfrogging development, reducing spatial efficiency. Consequently, green open spaces are converted into settlements or commercial areas, separate from the city center. (Bambulu et al., 2018). Urban sprawl in Manado has negative consequences such as increasing infrastructure management costs, decreasing environmental quality, and increasing transportation burdens. One solution could be the application of the compact city concept, which encourages the efficient land use and focuses on settlement density and diversity of land functions to support sustainability (Tilaar et al., 2017).

Understanding the characteristics of urban sprawl in an area provides insights into its growth patterns and determinants, allowing better prediction. One approach for understanding urban sprawl is the cellular automata method. An advantage of this method is its ability to capture complex spatial dynamics. Each cell in the model represents a land unit that can change its status based on certain transition rules, allowing for the observation of how changes in one location can trigger changes in others. Cellular automata are useful tools for spatial planning and decision-making related to urban growth management.

This study aimed to develop a transition rule algorithm to model urban sprawl based on built-up land or open-space indicators within the framework of a cellular automata grid. Through an analysis of the attributes of an urban cell, this study attempted to identify the determinant factors that influence the conditions of built-up land or open space, which were then used to formulate the algorithm concept. Statistical and spatial analysis methods were used to link these attributes to built-up land or open space conditions. These results provide a theoretical and methodological basis for developing a more comprehensive urban sprawl simulation model.

LITERATURE REVIEW

On Urban Sprawl

Urban sprawl is a phenomenon of uncontrolled urban expansion characterized by high land consumption, low population density, and landscape fragmentation (Sudhira et al., 2012). According to Soetomo (2013), urban sprawl is a model of urban extension that occurs horizontally, leading to the formation of "mega urban" areas. Urban sprawl is an inefficient form of development (Bento et al. 2006). Sprawling negatively impacts the environment, resulting in loss of natural habitat, decreased water quality, and greenhouse gas emissions (Ewing & Rong, 2008). Urban expansion can lead to longer travel times, increased dependence on motorized vehicles, and reduced walking ability (Ewing & Cervero, 2010).

Brueckner (2000) states that the normal expansion of urban spaces is primarily caused by an increasing population and income, and decreasing travel costs. However, incorrect policies can negatively distort this process and justify negative perceptions surrounding the sprawl phenomenon. Burchell & Mukherji (2003) indicate that compared to urban sprawl, intelligently managed physical growth of cities can save human and natural resources extensively.

Hasse & Lathrop (2003) measure sprawl in housing units using five variables as sprawl characteristics, namely building density, leapfrog development patterns, segregated land use, development within the reach of the road network (highway strip), and distance to the city center (node inaccessibility). Burchfield et al. (2005) define the sprawl index as the percentage of undeveloped land around an average urban residence. Angel et al. (2007) state that five attributes should be considered while measuring urban sprawl, one of which is the reduced continuity of built-up areas and the fragmentation of open spaces.

A significant indicator of sprawl is the open-space ratio. This metric is defined as the proportion of unbuilt areas within an area (Tan & Sia, 2019). Open spaces have many benefits such as providing space for recreation, improving air and water quality, and reducing the heat island effect (Benedict & McMahon, 2006). One study showed that sprawl is negatively correlated with the ratio of open spaces: the higher the level of sprawl, the lower the ratio of open spaces

(Sudhira et al., 2012). This was because of the conversion of open land into built-up areas for housing, commerce, and infrastructure.

Cellular Automata and Its Transition Rules Algorithm

The cellular automata method is a computational model that can be used to simulate complex spatial phenomena such as urban sprawl. With its ability to simulate spatial and temporal changes in land use, this method allows us to determine how complex and nonlinear urban growth patterns are. Cellular automata can integrate various factors that influence urban growth, such as government policies, accessibility, and community preferences. This method can also be used to project future scenarios and evaluate the impacts of various policies on urban growth patterns (Batty & Xie, 1994; Batty et al., 1999).

Cellular automata are dynamic models that are discrete in time, space, and state (Balzter et al. 1998). The basic concept of cellular automata is that a spatial grid cell can be conditioned based on the variation in its neighborhood cells, that can be viewed through two approaches: the Moore neighborhood concept, which consists of eight cells surrounding a central cell in two dimensions, and the Von Neumann neighborhood, which consists of four cells orthogonally surrounding a central cell in two dimensions (Basse, 2013).

The primary component of cellular automata is the definition of transition rules that determine changes in the conditions of certain attributes in a geographic process. Various methods have been used to describe the transition rules of cellular automata, such as multi-criteria evaluations, regression models, and artificial neural networks (Cao et al., 2015). Using its transition algorithm, cellular automata can predict the potential sprawl development, evaluate the impact of different policies on sprawl, and identify effective strategies for controlling sprawl.

Several studies have shown that cellular automata can be used to simulate sprawl with high accuracy. These models consider factors such as population density, accessibility, and land-use zoning (Wu & Webster, 1998; Tsai, 2005; Pinto et al., 2021).

RESEARCH METHODOLOGY

In general, this research was conducted with the following steps:

1) Preparation. At this stage, the scope of the study area in the peri-urban area of Manado City was determined and divided into a cellular automata grid framework.
2) Data Collection. The collected data included the spatial data of the study area, describing the variation in cell attributes in the area grid.

3) Data Analysis. The analysis included a) spatial analysis to explore the sprawl pattern in the study area, especially based on open space indicators and built-up land as a complement; and b) statistical analysis of the condition of urban cell attributes according to the grid framework. In this analysis, the intercorrelation patterns between various urban cell attributes and built-up land and open space conditions were examined to identify the factors that contribute to sprawl.
4) Conceptualization of the Transition Algorithm. The concept of the urban cell condition transition algorithm, which indicates the sprawl phenomenon, was developed based on the intercorrelation tendencies between urban cell attributes, particularly built-up land and open space.
5) Drawing Conclusions and Formulating Recommendations.

RESULT AND DISCUSSION

Overview of Study Area and Grid Division of Cellular Automata Framework

This study is located in Malalayang District, one of the peri-urban areas of Manado City. It was chosen because it exhibits the phenomenon of urban sprawl, where built-up land distribution indicates a significant fragmentation of green open spaces in the city image map. For data collection in the cellular automata framework, the study area was derived from several grids measuring 500 × 500 m, which were observation units for collecting attribute data. Each grid had an area of 250,000 m^2 (25 Ha). The delineation of the study area and division of the cellular automata framework grid are shown in Figure 1.

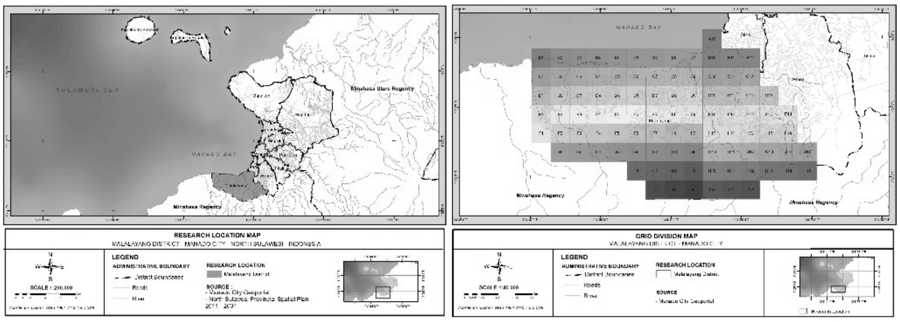
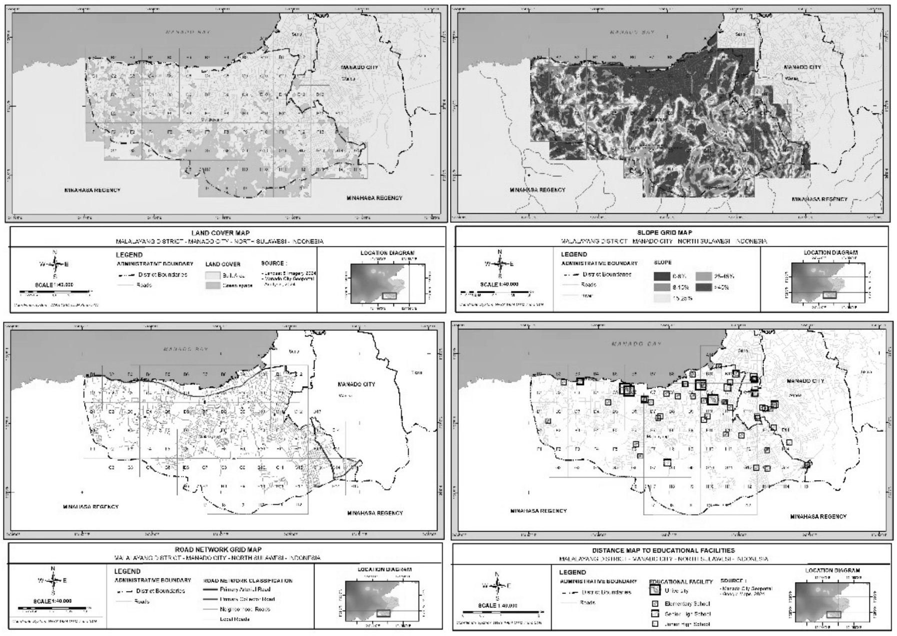


Figure 1. Delineation & Cellular Automata Grid Framework of Study Area
Source: Geoportal Site Data, Field Observation & Interpretation of Research Team

Basic Data According to Cellular Automata Grid

Based on grid division, some grid attributes were recorded, which were considered as variables in the development of urban sprawl transition rules in the cellular automata framework. In general, two groups of grid data attributes exist. The first is an attribute that has the potential to be a dependent variable in the transition rule algorithm as well as the main indicator of the urban sprawl phenomenon, namely, the area of built-up land that is complementary to the area of open space in each grid. This attribute is also considered an independent variable that determines the sprawl conditions in the neighboring grid cells. Second, other attributes have the potential to be independent variables (determinants) in the transition rule algorithm, including slope conditions, road infrastructure availability conditions, and accessibility distance conditions for certain public facilities such as education, health, trade and services, and government. Built-up land attributes were recorded in units of hectare (Ha) per grid. Other attributes were recorded in score units, each with a certain range, with low scores representing conditions that were perceived as bad and high scores representing conditions that were perceived as good in the context of encouraging physical growth in the area in question and its neighboring grids. These data were obtained from several secondary data sources, including image maps and shape files accessed at the geoportal site, along with field observations. Figure 2 shows the spatial data for each attribute presented above for each grid of the study area.



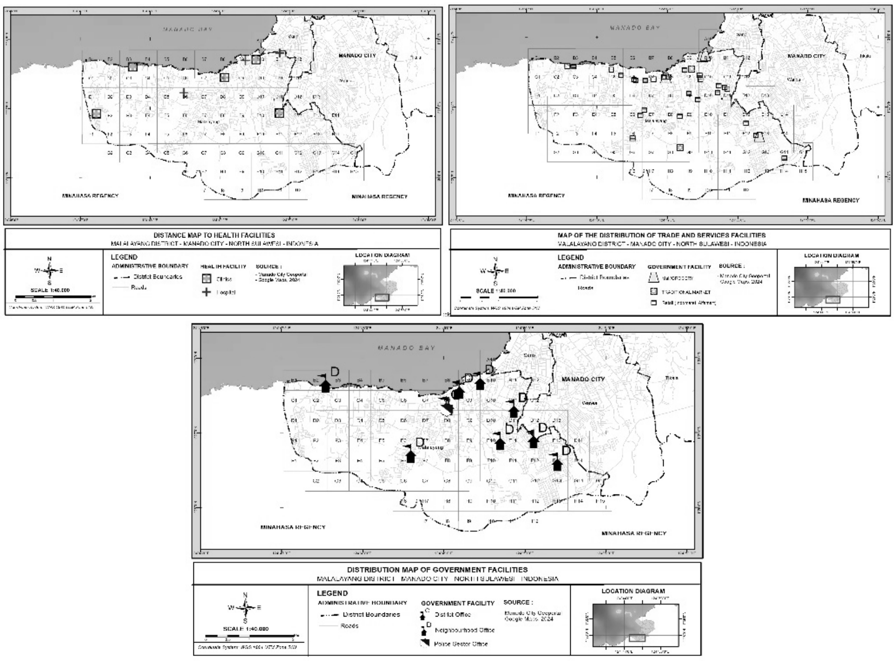


Figure 2. Thematic Maps of Grid Attribute Data of Study Area
Source: Geoportal Site Data, Field Observation & Research Team Interpretation

Referring to the thematic map above as well as the quantification of each attribute, basic data were obtained from seven attributes observed in a total of 96 (ninety-six) grid cells in the study area.

Data Processing for Analysis Input

Basic data obtained is further processed and prepared as input for analysis to identify potential sprawl growth transition rules within the cellular automata grid framework. A correlation analysis is conducted between the attributes of the sprawl indicator variables in a given grid cell, particularly, the condition of built-up land, and various attributes in directly adjacent grids cells. These neighboring grid cells are considered determinant variables influencing the sprawl indicator in the observed grid cell. Figure 3 shows a map of the grid positions in the study area as a cellular automata framework. The lighter-colored grid cells are the observed grid cells because each has directly adjacent grid cells on the top, bottom, left, and right sides of the cell. For example, grid cell C2, as one of the observed grid cells, has 4 adjacent grid cells, each of which is B2, C1, C3, and D2.

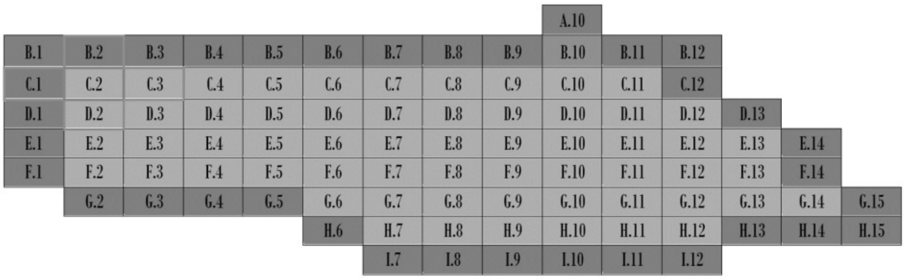


Figure 3. Cellular Automata Grid Framework of Study Area
Source: Processed Data Results of Research Team

In the correlation analysis, there were two groups of data: 1) internal attribute or observed grids, which were theoretically positioned as dependent variables, and 2) adjacent grid attributes, which were viewed as independent variables or those estimated to influence the condition of the internal attribute or observed grids. The first group includes the built-up land area (Y1), which is complementary to the open space area (Y2) and the open space fragmentation index (Y2/Y1). The value of the first data attribute in each observed grid cell is the same as that of the basic data obtained. The second dataset includes the average value of built-up land area (X1), slope gradient score (X2), road infrastructure availability score (X3), and the score of the distance of accessibility to educational facilities (X4), health (X5), trade and services (X6), and government (X7) in the four grid cells that are directly adjacent to the observation cell. Referring to the description of data processing above, Table 1 shows a recapitulation of the distribution of the values of each variable in each observation grid cell, which includes the values of the variables in the 61 observed grid cell units. These processed data are then used as input at the analysis stage, especially for correlation and regression analyses, to identify the possibility of a sprawl growth transition rule algorithm in the cellular automata grid framework.

Table 1. Data Processing Results for Analysis Input

Data Number	Grid Name	Y1 (Ha)	Y2 (Ha)	(Y2/Y1)	X1 (Ha)	X2 (Score)	X3 (Score)	X4 (Score)	X5 (Score)	X6 (Score)	X7 (Score)
1	B.10	23,49	0,00	0,00	18,54	4,78	0,86	2,15	1,62	2,01	1,17
2	C.2	19,41	5,59	0,29	13,72	4,02	1,45	1,95	1,23	1,67	1,09
3	C.3	14,05	10,95	0,78	13,15	4,08	1,49	2,06	1,26	1,68	1,11
4	C.4	16,25	8,75	0,54	16,27	4,26	1,59	2,30	1,38	1,71	1,19
5	C.5	23,44	1,56	0,07	17,30	4,61	1,41	2,48	1,56	1,72	1,27
6	C.6	23,87	1,13	0,05	17,52	3,47	0,98	2,42	1,63	1,79	1,33
7	C.7	24,83	0,17	0,01	18,38	3,65	1,04	2,25	1,44	1,87	1,34
8	C.8	24,83	0,17	0,01	19,40	4,80	1,28	2,13	1,43	1,85	1,27
9	C.9	25,00	0,00	0,00	22,59	4,83	1,25	2,19	1,43	1,79	1,24
10	C.10	23,97	1,03	0,04	23,25	4,51	1,19	2,28	1,41	1,84	1,15

11	C.11	24,34	0,66	0,03	23,73	4,52	1,31	2,16	1,52	1,81	1,13
12	D.2	14,94	10,06	0,67	15,59	4,13	1,13	2,02	1,23	1,72	1,13
13	D.3	13,14	11,86	0,90	15,45	4,06	1,19	2,10	1,31	1,74	1,15
14	D.4	23,35	1,65	0,07	17,24	4,13	1,10	2,29	1,48	1,77	1,21
15	D.5	23,86	1,14	0,05	21,12	4,37	1,24	2,37	1,67	1,76	1,26
16	D.6	21,79	3,21	0,15	23,43	4,49	1,23	2,27	1,68	1,76	1,28
17	D.7	24,83	0,17	0,01	23,65	4,60	1,19	2,13	1,57	1,82	1,31
18	D.8	24,65	0,35	0,01	24,49	4,54	1,24	2,04	1,38	1,75	1,33
19	D.9	23,99	1,01	0,04	23,03	4,48	1,20	2,14	1,40	1,71	1,24
20	D.10	20,18	4,82	0,24	23,68	4,27	1,19	2,27	1,35	1,64	1,15
21	D.11	24,88	0,12	0,00	22,76	3,97	1,06	2,15	1,33	1,76	1,13
22	D.12	22,55	2,45	0,11	23,08	4,01	1,02	1,87	1,37	1,80	1,08
23	E.2	11,07	13,93	1,26	12,35	3,93	1,10	2,09	1,22	1,78	1,18
24	E.3	9,48	15,52	1,64	13,76	4,01	1,14	2,23	1,35	1,79	1,19
25	E.4	15,70	9,30	0,59	13,01	3,95	1,22	2,29	1,50	1,78	1,20
26	E.5	15,88	9,12	0,57	16,63	4,01	1,05	2,24	1,65	1,73	1,20
27	E.6	21,15	3,85	0,18	20,50	4,26	1,15	2,16	1,70	1,69	1,20
28	E.7	23,34	1,66	0,07	20,94	4,07	1,17	2,01	1,58	1,72	1,24
29	E.8	24,33	0,67	0,03	21,54	4,43	1,11	2,11	1,46	1,72	1,27
30	E.9	22,28	2,72	0,12	21,80	3,95	1,27	2,17	1,38	1,61	1,24
31	E.10	21,89	3,11	0,14	20,87	3,93	1,19	2,14	1,28	1,73	1,17
32	E.11	23,97	1,03	0,04	20,38	3,56	1,13	2,06	1,24	1,75	1,07
33	E.12	21,09	3,91	0,19	23,16	3,89	1,17	1,94	1,24	1,90	1,09
34	E.13	23,20	1,80	0,08	23,47	3,86	0,61	1,72	1,34	1,95	1,08
35	F.2	9,63	15,37	1,60	12,53	4,24	0,85	2,17	1,25	1,81	1,21
36	F.3	15,13	9,87	0,65	9,86	3,88	0,82	2,27	1,37	1,80	1,21
37	F.4	3,33	21,67	6,51	13,22	3,92	1,06	2,28	1,50	1,76	1,20
38	F.5	5,82	19,18	3,30	14,28	3,84	1,14	2,20	1,59	1,67	1,15
39	F.6	20,98	4,02	0,19	11,86	3,64	1,19	2,08	1,62	1,62	1,14
40	F.7	13,47	11,53	0,86	15,55	3,95	0,83	2,05	1,57	1,59	1,17
41	F.8	15,92	9,08	0,57	16,60	4,03	1,32	2,15	1,47	1,65	1,22
42	F.9	16,99	8,01	0,47	16,10	4,03	1,41	2,25	1,39	1,70	1,22
43	F.10	17,04	7,96	0,47	16,60	3,64	1,17	2,20	1,29	1,78	1,15
44	F.11	13,63	11,37	0,83	17,70	3,78	1,12	2,10	1,21	1,91	1,11
45	F.12	22,92	2,08	0,09	18,89	3,63	1,13	1,92	1,26	1,94	1,07
46	F.13	24,25	0,75	0,03	20,04	3,84	1,19	1,81	1,33	1,99	1,07
47	G.6	6,98	18,02	2,58	10,85	3,45	0,55	2,08	1,58	1,59	1,13
48	G.7	1,98	23,02	11,66	8,15	3,63	1,02	2,10	1,55	1,64	1,16
49	G.8	11,61	13,39	1,15	8,14	3,83	0,76	2,23	1,48	1,66	1,22
50	G.9	9,17	15,83	1,73	10,87	3,78	0,95	2,31	1,40	1,77	1,22
51	G.10	13,87	11,13	0,80	10,31	3,46	1,35	2,26	1,33	1,88	1,18
52	G.11	6,87	18,13	2,64	15,16	3,66	1,02	2,13	1,29	1,96	1,15
53	G.12	16,61	8,39	0,51	15,14	3,81	1,00	2,00	1,30	1,99	1,11
54	G.13	22,36	2,64	0,12	18,18	3,77	1,30	1,86	1,34	1,94	1,08
55	G.14	15,00	10,00	0,67	17,53	3,70	1,59	1,71	1,38	1,93	1,09
56	H.7	0,52	24,48	46,88	2,80	3,41	0,00	2,17	1,53	1,68	1,19
57	H.8	5,50	19,50	3,55	3,66	3,63	0,36	2,27	1,48	1,76	1,23
58	H.9	1,02	23,98	23,55	6,10	3,41	0,75	2,34	1,42	1,85	1,24
59	H.10	8,18	16,82	2,06	9,20	3,08	0,50	2,29	1,37	1,95	1,22
60	H.11	16,54	8,46	0,51	6,50	3,04	0,75	2,19	1,34	1,99	1,18
61	H.12	8,42	16,58	1,97	15,14	3,59	0,75	2,05	1,34	1,97	1,14

Source: Recapitulation of Processed Research Team Data Results

Analysis Result & Interpretation

Referring to the processed data, the next stage is a correlation analysis, which aims to identify the strength and direction of the relationships between the observed variables. Correlation analysis was conducted using the Pearson product-moment quantitative data correlation analysis method, which was calculated using Microsoft Excel. The results of the correlation analysis that have been carried out are presented in full in Table 2.

Table 2. Correlation Analysis Results

	Y1	*Y2*	*Y2/Y1*	*X1*	*X2*	*X3*	*X4*	*X5*	*X6*	*X7*
Y1	1									
Y2	-0,99964386	1								
Y2/Y1	-0,56654298	0,565609834	1							
X1	**0,816675855**	**-0,81511646**	**-0,51896945**	1						
X2	**0,537032223**	**-0,54226524**	**-0,31517723**	0,631004578	1					
X3	**0,49634328**	**-0,49193766**	**-0,55767808**	0,534653664	0,473747423	1				
X4	-0,16783185	0,167029184	0,116166895	-0,26224755	0,088211191	-0,07988472	1			
X5	0,049455995	-0,0545889	0,122437174	0,008455046	0,227808592	-0,12635705	0,391161238	1		
X6	0,116865116	-0,12358805	-0,12146288	0,071747833	-0,17052893	-0,11031287	-0,27529187	-0,33602426	1	
X7	0,090471912	-0,08934744	0,04825072	0,004537171	0,295109291	-0,04548726	0,652199358	0,49659291	-0,23201812	1

Source: Results of Research Team Analysis

The results indicate that between the first data group (Y1, Y2, and Y1/Y2) and the second data group (X1 to X7), significant correlation coefficient figures only occur between the three variables of the first data group with three variables from the second data group, namely variables X1, X2, and X3, with correlation coefficient values in the range of (+/-) 0.5 to 0.8, indicating a moderate to high level of correlation, both in the same direction (between Y1 and X1, X2, and X3) and in the opposite direction (Y2 and Y2/Y1 with X1, X2, X3). The correlation between the variables X4 and X7 can be ignored because their values approach 0. This indicates that the observed grid's land cover conditions, both built-up land and open space areas (Y1 and Y2), were significantly correlated with the conditions of built-up land area (X1), slope gradient (X2), and availability of road infrastructure (X3) in adjacent grid cells. The positive correlation between the internal grid variable (Y1) and the three adjacent grid variables (X1, X2, and X3) indicates that high built-up land area, slope gradient score, and road infrastructure availability score in cells directly adjacent to an observed grid cell are also associated with high built-up land area inside the observed grid cell, and vice versa. The negative correlation between the unbuilt land area (Y2) and open space fragmentation index (Y2/Y1) with the other three variables can be understood logically in connection with the fact that the Y2 value is complementary to the Y1 value. This is also confirmed by the correlation coefficient value between the variables of the first data group, namely between variables Y1 and Y2, which have values close to 1, and between Y2/Y1 and Y1 and Y2, which have relatively the same values but in opposite directions. In the scope of the second group of data, it can be observed that between variables X1, X2, and X3, each has a fairly significant (moderate) level of intercorrelation, with a value of approximately 0.5, and is positive in the same direction. A significant and positive correlation was also observed between variables X4 (access to educational facilities) and X5 (access to health facilities), and variable X7 (access to government facilities).

Hence, to identify the urban sprawl transition rule algorithm within the cellular automata framework, the internal grid built-up land area variable (Y1) or the open space area variable (Y2) can be positioned as an indicator of the sprawl phenomenon of a spatial unit grid is influenced by at least three variables in the grid cells directly adjacent to the grid, which in this case include the built-up land area variable (X1), slope gradient score (X2), and road infrastructure availability score (X3).

Furthermore, the transition rule algorithm in question was identified using multivariate regression analysis involving variable Y1 as the dependent variable and variables X1, X2, and X3 as the independent variables. This multivariate regression analysis was conducted through Microsoft Excel. The results are presented in Table 3.

Table 3. Results of Stage 1 Regression Analysis

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,819838597
R Square	0,672135326
Adjusted R Square	0,65487929
Standard Error	4,201141052
Observations	61

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	2062,393241	687,4644137	38,95073847	7,9451E-14
Residual	57	1006,02641	17,64958614		
Total	60	3068,419651			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-2,985731754	5,877446808	-0,50799809	0,613414999	-14,75511598	8,783652469	-14,75511598	8,783652469
X1	1,003714072	0,137106074	7,320711908	9,29198E-10	0,729163893	1,27826425	0,729163893	1,27826425
X2	0,320625279	1,790675333	0,179052714	0,858531252	-3,265140224	3,906390783	-3,265140224	3,906390783
X3	2,017523626	2,308592199	0,873919451	0,385829309	-2,605352511	6,640399763	-2,605352511	6,640399763

Source: Results of Research Team Analysis

The results indicate that some interpretations can be proposed. First, with an 'R Square' value of 0.67 (level of correlation between the predictive and factual values of the Y1 variable), it can be interpreted that approximately 67% of the variation in the Y1 value is contributed by variables X1, X2, and X3, and 33% in the Y1 value is contributed by several other variables. Second, the results of the ANOVA test indicate that the Significance F value (probability value) was <0.5, demonstrating that the regression analysis results were statistically significant. Third, the regression analysis results indicate that the coefficient and intercept values of the multivariate regression equation were obtained, mathematically describing the influence of variables X1, X2, and X3 on Y1. The regression equation can be written as follows:

Y1 = X1 + 0,32X2 + 2,02X3 – 2,99, *where:*
Y1 = Observed Grid Built-up Land Area
X1 = Average Adjacent Grids Built-up Land Area
X2 = Adjacent Grids Slope Score
X3 = Adjacent Grids Road Infrastructure Availability Score

However, since the previous correlation analysis indicated a relatively high intercorrelation among the three independent variables, the prediction of the dependent variable Y (internal grid built-up land) should be based only on the independent variable with the highest determination level. In this case, the variable is X1 (average area of built-up land in adjacent grids). By eliminating variables X2 and X3, the following are the results of a simple linear regression analysis indicating the determination of variable X1 on Y1.

Table 3. Results of Stage 2 Regression Analysis

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,816675855
R Square	0,666959451
Adjusted R Square	0,661314696
Standard Error	4,161787759
Observations	61

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	2046,511487	2046,511487	118,1556054	1,02992E-15
Residual	59	1021,908164	17,32047735		
Total	60	3068,419651			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,713311509	1,717000229	-0,415440544	0,679326478	-4,149021031	2,722398012	-4,149021031	2,722398012
X1	1,075011329	0,098897628	10,86994045	1,02992E-15	0,877117632	1,272905026	0,877117632	1,272905026

Source: Results of Research Team Analysis

Some interpretations can be made from the results of the regression analysis. First, as in the previous regression, the 'R Square' value of 0.67 indicates that approximately 67% of the variation in the Y1 value is contributed by the independent variable X1, and the remaining 33% is contributed by other variables. Second, the results of the ANOVA test, with a Significance F value of <0.5, indicate the statistical significance of the regression. Third, the coefficient and intercept values for the simple regression equation are also obtained, which can be expressed as follows:

Y1 = 1,08X1 – 0,71, *where:*
Y1 = Built-up Land Area Observed Grid
X1 = Average Built-up Land Area Adjacent Grids

This equation can be observed as the urban sprawl transition rule algorithm, using a land cover condition indicator within the cellular automata grid. It suggests that every one-unit (Ha) increase in the average built-up land area of the four adjacent grids will contribute to an increase of 1.08 units (Ha) of built-up land area of the observed grid, without considering other driving factors.

CONCLUSION

After formulating the problem and research objectives, the results of this study can be summarized as follows:

1) In the context of urban sprawl and the cellular automata grid framework, certain neighboring grid cell attributes can be considered as driving factors for variations in built-up land conditions in a particular spatial grid cell. Out of the seven neighboring grid cell attributes considered as driving factors, only three (built-up land conditions, slope gradient, and availability of road infrastructure) were significantly correlated with the built-up land area variable in the observed cell. However, the correlation analysis also indicates a relatively strong intercorrelation between the three variables. As observed in the regression relationship, the three variables cannot be used simultaneously as predictors because they are not statistically independent. Referring to the correlation coefficient value, the most significant predictor variable for variations in built-up land area within a spatial grid cell in the cellular automata framework is the average built-up land area of neighboring grid cells. The other two variables can be regarded as secondary predictors, influencing the outcomes indirectly.

2) Conceptually, the urban sprawl transition rule algorithm developed in this study can be expressed as follows.

   - Built-up land or open space in a particular spatial grid was influenced by the condition of built-up land and open space in neighboring grid cells, with a determination level of 67%. The rest were influenced by other driving factors.
   - Every one-unit (Ha) increase in the average area of built-up land or a decrease in an open space area in a group of neighboring cells is followed by the same phenomenon in a particular grid cell, with a quantity of 1.08 units (Ha).
   - The condition of built-up land and open space in each spatial grid cell correlated with the slope conditions and availability of road network infrastructure in the spatial grid cell. The flatter the slope conditions and

the higher the availability of road infrastructure in a spatial grid, the higher the condition of built-up land cover in that cell.

ACKNOWLEDGMENT

The authors would like to thank the Dean of the Faculty of Engineering of Sam Ratulangi University, who facilitated this research, mainly through the provision of space and equipment at the Regional & City Planning Laboratory. The authors would also like to thank Editage (www.editage.com) for English language editing.

DISCLOSURE STATEMENT

This research was supported by the Institute for Research & Community Service, Sam Ratulangi University, Indonesia, with funding from the Public Service Agency Budget Implementation List (SP DIPA - 023.17.2.677519/2024).

REFERENCES

Angel, S., Parent, J., & Civco, D. (2007). Urban Sprawl Metrics: An Analysis of Global Urban Expansion Using GIS. *ASPRS 2007 Annual Conference Tampa, Florida*.

Balzter, H., Braun, P. W., & Kohler, W. (1998). Cellular Automata Models for Vegetation Dynamics. *Ecological Modelling*, *107*(2-3), 113–125.

Bambulu, V. J., Tondobala, L., & Takumansang, E. D. (2018). Analisis Karakteristik Urban Sprawl di Kota Manado. *Jurnal Spasial : Perencanaan Wilayah & Kota*, *5*(3), 367-376.

Basse, R. M. (2013). A Constrained Cellular Automata Model to Simulate the Potential Effects of High-Speed Train Stations on Land-use Dynamics in Trans-Border Regions, *Journal of Transport Geography*, *32*, 23-37.

Batty, M., & Xie, Y. (1994). From Cells to Cities. *Environment and Planning B: Planning & Design*, *21*(7), S31-S48. https://doi.org/10.1068/b21S031

Batty, M., Xie, Y., & Sun, Z. (1999). Modeling Urban Dynamics through GIS-based Cellular Automata, *Computers, Environment & Urban Systems*, *23*(3), 205-233.

Benedict, M. A., & McMahon, E. T. (2006). *Green Infrastructure: Linking Landscapes & Communities*. Island Press.

Bento, A. M., Franco, S. F., & Kaffine, D. (2006). The Efficiency and Distributional Impacts of Alternative Anti-Sprawl Policies. *Journal of Urban Economics*, *59*(1), 121-141.

Brueckner, J. K. (2000). Urban Sprawl: Diagnosis and Remedies. *International Regional Science Review*, *23*(2), 160-171. https://doi.org/10.1177/016001700761012710

Burchell, R. W., & Mukherji, S. (2003). Conventional Development Versus Managed Growth: The Costs of Sprawl. *American Journal of Public Health*, *93*(9), 1534-1540. https://doi.org/10.2105/AJPH.93.9.1534

Burchfield, M., Overman, H., Puga, D., & Turner, M. (2005). Causes of Sprawl: A Portrait From Space. *Quarterly Journal of Economics*, *121*(2), 587-633.

Cao, M., Tang, G., Shen, Q., & Wang, Y. (2015). A New Discovery of Transition Rules for Cellular Automata by Using Cuckoo Search Algorithm. *International Journal* *of Geographical Information Science*, *29*(5), 806-824. https://doi.org/10.1080/13658816.2014.999245

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment: A Meta-Analysis. *Journal of the American Planning Association*, *76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Ewing, R., & Rong, F. (2008). The Impact of Urban Form on U.S. Residential Energy Use, Housing Policy Debate, Taylor & Francis Journals, *19*(1), 1-30.

Hasse, J., & Lathrop, R. G. (2003). A Housing Unit Level Approach to Characterizing Residential Sprawl. *Photogrammetric Engineering & Remote Sensing*, *69*(9), 1021-1030.

Pinto, N., Antunes, A. P., & Roca, J. (2021). A Cellular Automata Model for Integrated Simulation of Land Use and Transport Interactions. *ISPRS International Journal Geo-Information*, *10*(3), 149.

Rogi, O. H. A., Waani, J. O., & Supardjo, S. (2024). Understanding The Consequences of a City's Spatial Structure through its Density Profile as Indicator. A Case Study in Manado Municipal, North Sulawesi, Indonesia. *Devotion: Journal of Research and Community Service*, *5*(2), 244-257.

Seipalla, B., Latupapua, L., & Lelloltery, H. (2020). Kajian Potensi Ekowisata di Desa Liliboy Kecamatan Leihitu Barat Kabupaten Maluku Tengah. *Jurnal Hutan Tropis*, *8*(3), 280.

Soetomo, Soegiono. (2013). *Urbanisasi dan Morfologi. Proses Perkembangan Peradaban dan Wadah Ruang. Menuju Ruang yang Manusiawi*. Edisi 2. Graha Ilmu.

Sudhira, H. S., Ramachandra, T. V., & Jagadish, K. S. (2012). Urban Sprawl and Its Impact on Land Use and Water Bodies in Bangalore. *Applied Geography*, *34*, 597-606.

Tan, P. L., & Sia, A. (2019). Urban Sprawl and Its Impact on Open Space in Singapore. *Sustainability*, *11*(12), 3280. https://www.mdpi.com/2071-1050/11/12/3280

Tilaar, S., Sela, R. L., & Tondobala, L. (2017). Analisis Urban Compactness Kota Manado. *Jurnal Spasial: Perencanaan Wilayah & Kota*, *4*(2), 116-122.

Tsai, Y. (2005). Quantifying Urban Form: Compactness Versus Sprawl. *Urban Studies*, *42*(1), 141-161.

Wu, F., & Webster, C. J. (1998). Simulation of Land Development through the Integration of Cellular Automata and Multicriteria Evaluation. *Environment and Planning B-Planning & Design*, *25*(1), 103-126.

Received: 4th January 2025. Accepted: 10th March 2025

