III. DEVELOPMENT OF URBAN GROWTH SIMULATION MODEL USING INTEGRATION OF GIS AND CELLULAR AUTOMATA

3.1. Introduction

3.1.1. Background

In this research, Integration of GIS and Cellular automata modelling was used to develop a simulation model that can represent spatial distribution pattern of urban growth in Bandung Area. CA can be considered as analytical engine of GIS, which plays in important role in modelling and simulation of spatial-temporal processes. Another approach to integrate CA and GIS in this research are to develop a standalone CA program that can use data from GIS. Data interchange and compatibility can be achieved through file conversion protocols.

Two scenarios were applied in urban growth cellular automata simulation model to calibrate the model. Meanwhile, Validation are comparing and evaluating of the predicted modelling results with actual data. In the first scenario, driving forces and constrained factors (slope > 15%, protected forest, open green space and water body) as an input parameter in the model simulation. Whereas, the second scenario is the simulation model will be run without slope effect as an input parameter. The best scenario from the result will be applied in model simulation and can be used to predicting spatial distribution of urban growth simulation.

3.1.2. Objective

The objectives of this study are:

1) To develop model capable of representing spatial distribution pattern of urban growth in Bandung area using integration of GIS and Cellular Automata Modelling.

2) To generate alternative scenario was implemented in urban CA simulation model.
3.2. Literature Review

3.2.1. The Concept of Cellular Automata

Cellular Automata (CA) refer one kind of dynamic system, which are discrete in time and space (Zhou Chenghu, Sun Zhanli et al. 1999). CA were first devised by John von Neumann and Stanislaw Ulam in the 1940s as a framework for investigating the logical underpinnings of life.

CA are models in which contiguous or adjacent cells, such as those that might comprise a rectangular grid, change their states their attributes or characteristics - through the repetitive application of simple rules (G.W. Flake's, 2000). According to (Torrens, 2000) and (Itami, 1994; Wolfram, 1984), CA has 5 primary elements:

1) Lattice: the space on which the automaton exists
2) Cell: in which the automaton resides and thus constrained its states
3) Neighborhood: the cells surrounding the automaton
4) Transition rules: the behavior of the automaton
5) Temporal space: discrete time steps in which the automaton evolves

Every cell in the CA has a single value or state out of a finite value range. The simplest range of the states value could be as simple as 0 and 1. Each cell changes its states in the next time step according to the states of its immediate adjacent neighbours (neighbourhood). The function of the neighbourhood states, which is used to change the current cell state, is expressed as transition rules. In the sequence of time (t, t+1, t+2…), each cell in the CA lattice update their states based on the transition rules at the same time.

3.2.2. CA in Urban Modelling

Cellular Automata (CA) refer one kind of dynamic system, which are discrete in time and space (Zhou Chenghu, Sun Zhanli et al. 1999). CA were first devised by John von Neumann and Stanislaw Ulam in the 1940s as a framework for investigating the logical underpinnings of life. CA are models in which contiguous or adjacent cells, such as those that might comprise a rectangular grid, change their states their attributes or characteristics - through the repetitive application of simple rules (G.W. Flake's, 2000).
Urban models are an abstraction, simplified versions of real world objects and phenomena that are used as laboratories for exploring ideas about how things work in cities. Whilst CA models shows high potential usability and suitability in modelling urban dynamics, the basic definition of CA by Wolfram and Conway cannot be used immediately in this field without some extent of modification. Typically a CA model is a close system, which means no information, or energy or mass is exchanging with its outside world while it is not the case in an urban context. Any city, or complex system, is an open system which exchanging energy and/or mass with its environment. The classical CA definition is too strict to accommodate all the components need to represent in a CA urban model. Thus, modifications on classical CA are inevitable.

The modifications have to be applied to almost every element in CA model to make it capable of representing urban dynamics. Conclelis (1985 cited by Itami, 1994) proposed a generalized CA (GCA) model to relax the basic CA limitations to situate CA in to urban simulation realm (Itami, 1994).

To be able to use in urban simulation, classic CA has to be modified, mainly to relax the restrictions on the CA elements.

1) **Cell Space**

In basic CA, cell space is a close system. It does not accept the information from outside. While in urban context, one cannot look the city as a close system because so many exogenous links exist. To incorporate this exogenous influence into CA, urban modellers often use constrains and algorithms applied to transition rules.

Some researchers have divided the cell states into two groups, there are fixed and functional. These applications treat those attributes that are not changed in the urban development process as fixed states, such as watershed and terrain. Meanwhile, sites that are changed or active in the developing processes are considered as functional which can evolve with time, such as population and land use.

2) **Lattices**

The lattice of CA comprises the space in which the CA exists and evolves over time. In an elementary CA, this lattice is one-dimensional, while CA designed for geographic purposes such as urban modelling are generally defined in two-
dimensions. To basic CA, the lattices are often defined in a regular fashion, as grid squares or other combinations of regular shapes (Torrens, 2000c).

3) Neighbourhoods

A neighbourhood is any set of one or more locations that bear a specified distance and/or directional relationship to a particular location, the neighbourhood focus (Tomlin, 1991, p. 96). The operations that are used to calculate neighbourhood characteristics are called convolution, spatial filtering, or focal functions (Bonham-Carter, 1994; Burrough & McDonell, 1998). Various statistics can be used to characterise the neighbourhood of a location. In raster-based geographic analysis, neighbourhood operations are used to compute a new value for every location as a function of its neighbourhood. The neighbourhood of a cell in the CA formalism consists of an individual cell itself as well as a set of adjacent cells. There are two types of neighbourhood in strict two-dimensional CA, Von Neumann neighbourhood and Moore neighbourhood.

4) Time

In classic CA, time is discrete and the cells are updated synchronously between time steps as transition rules are applied simultaneously at every location. Recently, researchers tried to asynchronous update cell-state via the actions of agents in a CA space.
5) **Transition rules**

The essence driving force behind the CA model is the transition rules. It translates the behaviour of the real world into CA models. In classic CA, transition rules are deterministic and will not change during the evolution of CA. While in urban simulation, transition rules have to be opened to exogenous effects and they are modified into a probabilistic expression. This introduced an element of randomness, or called noise, into the model. In addition, it can be made dependent on other rules within the model to reflect the idea that urban systems operated as a tangled web of co-dependent subsystems and phenomena (Torrens, 2000). Other researchers have used a probability function to control the model action, like the work of Batty (Batty et al., 1999), the action of transition rules is a contingent action upon a certain probability.

3.2.3. **Integrating GIS and CA for Urban Dynamics Modelling**

There is a real advantage to integrate the potentialities of CA (dynamic, neighbourhood coherence and flexibility) with the data manipulation ability of GIS. Urban researchers agree on the principle, the usefulness and the necessity to link CA with GIS to achieve more realistic and informed urban dynamics model. Three main approaches are often used (White and Engeelen, 1995).

One approach argues for building a CA modelling application using the programming language within a GIS language protocol. This option requires familiarity with the programming language embedded in the GIS package in use. Such an option is potentially problematic given that the flexibility of the language is not always guaranteed, and the scope for applying the skills learn in the process is limited.

Another approach to integrating CA and GIS is to develop a stand alone CA program that can use data from GIS. Data interchange and compatibility can be achieved using file conversion protocols (Baatky et al, 1994; Meaile et al, 1991; Yeh et al, 2001). The problem with this approach, however is the compatibility of the data. If the program cannot access and modify the data from and to GIS environment, then the process of reformatting the input and output is not only more likely to mislead the representation, it could also be time consuming and error prone.
A further possibility, and what has been argued to be an efficient approach, is the coupling of GIS environment with modelling and simulation tools to handle spatio-temporal processes (Bivand et al., 2000). In this case, CA and GIS are used together and each technology supports the function it best performs. This approach, known as coupling, is generally subdivided into loose and tight-coupling, depending on the level of linkage between the CA and GIS technologies. Loose or tight coupling provides a great level of flexibility and has been suggested to produce better integration model (Almeida et al., 2002). This integration has been further enhanced by the possibility to amend and relax the original cellular automata framework (for instance the number of neighbouring cells on a CA lattice) to generate real world application.

3.2.4. Model Calibration and Validation

Models are simplified representations of real systems. As a result of this simplicity, a model can never completely reproduce the structure or process of the modelled system (Norlén 1975). The model leaves part of the reality behind. This is both its power and its weakness. The power lies in the clarity of the essentials and the manipulative nature of the symbols; the weakness lies in a small but necessary degree of invalidity (Kilbridge, O’Block, and Teplitz 1970). Therefore, testing or verifying how well a model matches its specifications, and minimising or controlling the model’s degree of invalidity, become two critical tasks in model construction (Naylor and Finger 1967).

Model Calibration is the process of modifying the input parameters in an attempt to match field conditions within some acceptable criteria until the output from the model matches an observed set of data. It is often used to estimate a model’s parameters that provide the best fit to an observed set of data (Bell, Dean, and Blake 2000: 574). The validation of a model is the determination of the model is an accurate representation of the system. Validation is usually achieved through the calibration of the model, an iterative process of comparing the model to actual system behaviour and using the discrepancies of the two, and the insights gained to improve the model. The validity of a simulation model is determined by the accuracy of its predictions.
Simulation accuracy assessments were applied in model validation process. It can be defined as the task of comparing two maps: one generated by the model (data to be assessed), and the other based on the ground truth (the reference data). The reference data were assumed accurate and forms the standard for comparison. The error matrix approach and kappa analysis are methods in accuracy assessment that applied in this study.

The error matrix approach is a common means of expressing the accuracy of landcover classification for remotely sensed data (Jensen 1996; Lillesand and Kiefer 1994; Story and Congalton 1986). This matrix summaries data from two different sources (one from reference data representing the actual data, and the other from classified image data or simulation model), and compares the relationship between the two on a site-by-site basis. An error matrix is a square array, the rows and columns of which represent the number of categories whose classification accuracies are being assessed. Typically, the columns of an error matrix represent the reference data, whereas rows indicate the classified image data. The terms used in an error matrix and their calculations are defined in Table 3.1.

Table 3.1 Definitions of Terms Used in an Error Matrix and Their Computations

<table>
<thead>
<tr>
<th>Terms</th>
<th>Definitions</th>
<th>Computations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Producer’s accuracy</td>
<td>Probability of reference cells being correctly categorized in the classification data. This measures the omission error.</td>
<td>Number of cells on the major diagonal divided by the column total of each category</td>
</tr>
<tr>
<td>User’s accuracy</td>
<td>Probability that cells in the classification data actually belong to the same category as in the reference data. This measures the commission error.</td>
<td>Number of cells on the major diagonal divided by the row total of each category</td>
</tr>
<tr>
<td>Omission error</td>
<td>Cells that are excluded (omitted) from the categories that they belong to in the reference data</td>
<td>Total of the off-diagonal column cells divided by the column total of each category</td>
</tr>
</tbody>
</table>
Commission error: Cells that are included (committed) in the categories that they do not belong to in the reference data.

Overall accuracy: A measurement of the overall proportion of correctly categorized cells in relation to the total number of cells under assessment.

Total of the off-diagonal row cells divided by the row total of each category.

Total number of cells along the major diagonal of the matrix divided by the total number of all the cells.

Chrisman (1980); Congalton and Mead (1983); and Congalton, Oderwald, and Mead (1983) proposed an application of the Kappa analysis, as described by Bishop, Fienber, and Holland (1975) and Cohen (1960), as a means of improving the interpretation of the error matrix.

3.3. Methodology

3.3.1. Database and Model Development

Single cell many table or grid-like shape format method has been developed as a Spatial Database Management System (SDBMS) in spatial data model. This method was compiled using vector spatial data format with polygon or point data type. Each point or polygon spatial data was compiled in unique ID like cell characterization in raster data format (Risdiyanto et al., 2003). Figure 3.2 was shown the schematic database.

![Database schematic](image)
Each cell in the database has unique coordinate system which describe in Cartesian coordinate. The same column in the lattice has the same of \( y \)-coordinates, the other way the same row has the same of \( x \)-coordinate. The attributes cell contains information of driving forces (Slope, Neighbourhood, distance from road and Urban Hierarchy Index), landcover priority and cell state in year \( -t \) (Urban existing, Urban allocation and constraint area).

Hybrid approach was implemented to integrate between tabular and spatial data as a Spatial Database Management System (SDBMS) in development of Urban Growth Spatial Distribution Model (UGSDM) in Bandung Area. In this study, UGSDM was developed to modelling and simulating spatial distribution of urban growth in Bandung area using integration of GIS and cellular automata methods. Figure 3.3 was showing diagram of hybrid system to develop UGSDM application Software.

![Figure 3.3 Hybrid approach to build UGSDM Application Software](image_url)

Various software were used to build UGSDM Application Software. The following application software can be seen in Table 3.2.

<table>
<thead>
<tr>
<th>No</th>
<th>Software</th>
<th>Type</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Microsoft Visual Basic 6.0</td>
<td>Programming Software</td>
<td>Developing User Interface and Building Simulation Model</td>
</tr>
</tbody>
</table>
2. ESRI Map Object  
2.4 Integrating Spatial data into Programming Software.

3. Microsoft Access  
2010 Database Application  
Developing attribute data, Storing database as tables.

### 3.3.2. Development of CA Model

A cellular automata is a discrete dynamic system in which space is divided into regular spatial units called cells and time progresses in discrete steps. Each cell in the system has one of a finite number of states. The state of each cell is updated according to local rules, i.e., the state of a cell at a given time depends on its own state and the states of its nearby neighbours at the previous time step (Wolfram, 1984). Cellular automata can be represented as follows:

\[
S_{t+1} = f(S_t, \Omega_t, T)
\]  

(2.1)

Here \(S_{t+1}\) means the exam cell’s state in Time \(t+1\), and \(S_t\) is the cell’s state in Time \(t\), \(\Omega_t\) is the cell’s neighbourhood situation in Time \(t\). \(T\) donate the transition rules of the CA. Using this formula, we can get the situation of whole cells after each CA iteration. This formula is the core part of CA, it articulate the CA’s evolution process clearly.

The essence driving force behind the CA model is the transition rules. It translates the behaviour of the real world into CA models. There are two transition rules, which control the cell states transition: calculate potential rule and update rules. In each time step \(t\) to \(t+1\), the cell in location \((x, y)\) changes its state according to the states of its neighbourhood and considering the other factors or constraints. To determine which cell should be converted into built-up area, an appraisal to each cell has to be carried out and a potential of transition to built-up area was calculated according to the natural condition and its neighbourhood situation. This is done by applying transition rule Calculate Potential Value. In Calculate Potential value, a multi criteria evaluation approach has been used to appraise the potential of conversion. Following the equation 3.2 was used in Calculate Potential value:
\[ P_v = \left( \frac{W_1 \times N + W_2 \times R + W_3 \times UH + W_4 \times S}{4} \right) \times L_p \] (2.2)

\[ W_1 + W_2 + W_3 + W_4 = 100\% \] (2.3)

Where; \( P_v \) denotes Potential Value; \( L_p \) denotes Landcover priority, \( N \) denotes Neighbourhood effect; \( R \) denotes distance from road effect; \( UH \) denotes Urban Hierarchy Index; \( S \) denotes Slope effect; \( W_1 \) to \( W_4 \) separately denotes the weight system applied to Neighbourhood effect, distance from road effect, Urban Hierarchy Index and Slope effect. The weights system \( W_i \) to \( W_4 \) also range from 0 to 100\% and the sum of all weight equals to 100\%.

Figure 3.4 Schematic diagram of the proposed

At each time step in urban cellular automata modelling, after all the cells already have been done calculate potential rule to define potential value, the update rule was applied to convert those qualified cells into built-up area cells. Cell will be update on land demand of built-up area and ranking of potential value. These simple rules were relatively easy to understand especially among planners, who potentially would utilize this model (Batty et al., 1997). Figure 3.4 was explaining Schematic diagram of the proposed model modified from Chapin et al., 1962.
3.3.3. Population Projection

Population projection is the prediction of future populations based on the present population, and with the present urban growth rates. Population projection can be estimated using Geometric method, the equation used is:

\[ P_t = P_0 (1 + r)^n \]  

Where; Estimated population; \( P_0 \) denotes Base year population, \( r \) denotes Population Growth Ratio (%); \( N \) difference year between \( P_t \) and \( P_0 \).

3.3.4. Relation Between Amount of Population and Built-up Area

Regression and correlation analysis was applied to find relation between urban land use and population projection. The result from this step is relation between increases of population and built-up area. Linier regression equation used to predict land demand of urban built up area in the feature year’s base on increase of population.

Figure 3.5 Regression and correlation analysis to get demand for prediction model
3.3.5. Configuration of Model, Calibration and Validation

Model calibration is the process of modifying the input parameters in an attempt to match field conditions within some acceptable criteria until the output from the model matches an observed set of data. In this study, two scenarios were applied to calibrate the model. In the first scenario, driving forces and constrained factors (slope > 15 %, protected forest, open green space and water body) as an input parameter in the model simulation. Whereas, the second scenario is the simulation model will be run without slope as an input parameter.

Meanwhile, Validation is the evaluation of the predicted modelling results. In this research, Model simulation was running from 1991 – 2000 and 2000 – 2007 to calibrate and validate the model result.

Simulation accuracy assessments base on the error matrix and kappa coefficient were applied in model validation process. Kappa analysis yields a $K_{hat}$ coefficient that measures the difference between the observed agreement of the two maps and the agreement that might be attained by chance matching (Campbell 1996). The $K_{hat}$ coefficient was computed as follows:
Where:

\[ K_{hat} = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{ir} \times x_{ri})}{N^2 - \sum_{i=1}^{r} (x_{ir} \times x_{ri})} \]  

(2.5)

\( K_{hat} \): The number of categories in the error matrix

\( x_{ii} \): The number of cells in row \( i \) and column \( i \)

\( x_{ir} \): The total number of cells in row \( i \) (shown as the row total in the matrix)

\( x_{ri} \): The total number of cells in column \( i \) (shown as the column total in the matrix)

\( N \): The total number of observations included in the matrix

Landis and Koch (1977) characterize agreement as follows: values > 0.75 are very good to excellent, values between 0.4 and 0.75 are fair to good and values of 0.4 or less indicate poor agreement.

### 3.4. Result and Discussion

#### 3.4.1. Configuration of model structure and elements

The very first step to model the urban growth process in CA model is build cells in discrete space represented by lattice. The space (lattice) of the study area is Bandung Area around of 329.510 Ha, while the city form is a strap shape distributed in the central of Bandung area to contain the whole city and the space for further development. A lattice of 750 x 850 cells was chosen to represent the study area and its surrounding. All of the data are contained in a grid-like shape based vector polygon with the spatial resolution of the cell in 100 x100 m. it’s means, each cell on the model represent an area of 10.000 m² or 1 Ha on the ground. Several reasons to select in this cell size to input into the model development. The first one is the spatial data used in this research have different of spatial resolution and scale. The lowest spatial resolution of spatial data is SRTM data which have spatial resolution in 90 x 90 m. Based on standardization of spatial data, all of the spatial data should be converted into the lowest of spatial resolution, so in this study all of the spatial data have been converted and standardized into spatial resolution for each cell 100 x 100 m. The second one is limitation of the computer specification to run and process the...
spatial database in the model simulation. Bandung as a study area, around of area is for about 329.510 Ha. This cell size was chosen to simplify in computing of the spatial database in the model processing. Lattice configuration of Bandung as study area was illustrated in Figure 3.7.

All of the data are contained in a grid-like shape based vector polygon with the spatial resolution of the cell in 100 x100 m. it’s means, each cell on the model represent an area of 10.000 m² or 1 Ha on the ground. Several reasons to select in this cell size to input into the model development. The first one is the spatial data were used in this research have different of spatial resolution and scale. The lowest spatial resolution of spatial data is SRTM data which have spatial resolution in 90 x 90 m. Based on standardization of spatial data, all of the spatial data should be converted into the lowest of spatial resolution, so in this study all of the spatial data have been converted and standardized into spatial resolution for each cell 100 x 100 m. The second one is limitation of the computer specification to run and process the spatial
database in the model simulation. Bandung as a study area, around of area is for about 329,510 Ha. This cell size was chosen to simplify in computing of the spatial database in the model processing.

The relationship of the Lattice, Cells and transition rules was described in Figure 3.8. Each cell in the lattice has unique coordinate system which describe in Cartesian coordinate. The same column in the lattice has the same of $y$-coordinates, the other way the same row has the same of $x$-coordinate. The attributes cell contains information for driving forces (Slope, Neighbourhood, distance from road and Urban Hierarchy Index), landcover priority and cell state in year $-t$. There are three cell state was defined in this model, Urban, Non-urban and constraint. The weight system is the weights for different driving forces and landcover priority was applied to derive potential ad update cell in transition rule calculation.

Transition rules are implemented in cells and be repeatedly applied to each cells during the model running. The CA transitions are defined for the allocation of new built-up area. The allocation was determined by the demand of built-up area as an effect of population increase. In one-iteration of the CA model, the simulation time is set to one year and so the model would allocate cells into built-up area for the amount determined by the prediction model for demand also established for one-year simulation time.
3.4.2. Population Projection and Land demand for Built-up Area

Population projection and land demand for built-up area are the first steps in model simulation that must be calculate. Actual condition in year 2000 used as a baseline in this research, for simulation and prediction population projection in Bandung Area. Based on Bandung statistical book in year 2000, amount of Population in Bandung area is 6,294,961 people and average of Growth rate in this area is 2.4% / year.

Geometric method used in this research to predict population projection, which assume growth of population in Bandung area is constant for each year. Equation 3.1 was applied to predict population projection in study area which amount of population in 2000 as base year’s population and average of growth rate is 2.4%.

Table 3.3 Data input to Predict Land Demand for Settlement Area.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Year’s</th>
<th>Source of Data Input</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2000</td>
<td>Topographical Map (RBI) in Bandung Area from Bakosurtanal.</td>
</tr>
<tr>
<td></td>
<td>2007</td>
<td>Landuse Map in Bandung Area from BPN and classified from spot 5 imagery in 2007.</td>
</tr>
</tbody>
</table>

Actually, the increase of population will be followed by the increase of land demand for built up area. Linier Regression and correlation methods used in this research to get relation between the increases of population growth and land demand for built up area. In this methods, built up area used as dependent variable – x and population data used as independent variable – y. Table 3.3 showing the data input to predict land demand for built up area using linier regression and correlation methods.
Base on the calculation result in Table 3.3, trend of Population density in Bandung area is increase for each year. These phenomena shown that the increase of population will be followed by the increase of built up area, in the other hand requirement for 1 Ha built-up area will be occupied by over populated. This condition will be affected overcrowding population for each 1 Ha in the future years.


<table>
<thead>
<tr>
<th>Criteria</th>
<th>Year's</th>
</tr>
</thead>
<tbody>
<tr>
<td>Built up Area (Ha)</td>
<td>33.382</td>
</tr>
<tr>
<td>Population (People)</td>
<td>4,888,439</td>
</tr>
<tr>
<td>Density(People/Ha)</td>
<td>146</td>
</tr>
</tbody>
</table>

Relation between the increase of population and built up area based on data input in Table 3.4 was shown in Figure 3.9. This relation used to predict demand of settlement area for the future years based on population projection simulation.

![Figure 3.9 Relation between The Increase of Population and Built-up Area](image)

\[ y = 0.0019x + 24903 \]

\[ R^2 = 0.946 \]
The calculation result by using linear regression method was get equation between increase of population and built up area, where prediction of demand for settlement area for the future years \(y\) and population projection \(x\) was described by formula \(y = 0.0019x + 24903\). Determinant coefficient \(R^2\) from this formula is 0.946 and correlation is 0.97. This relation is significant enough to describe relation between the increase of population and built up area and can be used to predict and simulate demand for built-up area as an effect from the increase of population.

### 3.4.3. Logical Framework and Program flow

The model development is called UGSDM (Urban Growth spatial Distribution Model) which can be divided into 4 sub-models. Figure 3.10 was shown logical framework of the model development.

<table>
<thead>
<tr>
<th>UGSDM (URBAN GROWTH SPATIAL DISTRIBUTION MODEL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Population projection and land demand for built up area</td>
</tr>
<tr>
<td>(2) Calculate the Potential Value</td>
</tr>
<tr>
<td>(3) Sorting and ranking the potential value</td>
</tr>
<tr>
<td>(4) Update Cell and transit into built-up area</td>
</tr>
</tbody>
</table>

Figure 3.10 Logical Framework of The Model Development.

Data input as actual condition in year 2000 was used as a baseline in year-\(t\) in the simulation model program. In this study, two scenarios rules were implemented to predict and simulate of urban growth spatial distribution. In the
first scenario, driving forces and constrained factors (slope > 15%, protected forest, open green space and water body) as an input in the model simulation. Whereas, the second scenario is the simulation model will be run without slope effect in model input. Weighted for each driving forces parameter in the first and second scenario was defined automatically by model simulation program, based on the result of weighted calculation using Multicriteria evaluation.

After scenario rules has been choose, the model has already to be run to simulate urban growth spatial distribution in the year \(- t+n\) simulation. Data initialization was done to sets all the parameters input in the model such as amount of population, cell state and demand for built-up area in year \(- t\) as a baseline in simulation model.

Population projection and land demand for built-up area calculation is the next process after model environment was set up. All of the cells in the lattice start iteration and the transition rules to calculate potential values will be applied to each cell. In the next steps, the model will be ranking and sorting all of the potential values from the highest to the lowest values. The model will be indicated that the number of the cells who have higher potential values will be translated into built-up area development base on amount of land demand requirement for built-up area in year \(- t+n\). For instance, if there are 50.000 cells available in one year for built-up area development and amount of land demand requirement for built-up area is for about 1.500 cells, after ranking these 50.000 cells by their developing potential values into descend order, the top most 1.500 cell are chosen to be developed into built-up area cells. The developing potential value of the 1.500\(^{th}\) cell will be used as an update cells whereas the other cells were no changes.

The model time step is set to one year, which means one model step equals one year in real world. After all the cells applied this update rule, the model starts iteration for each year, which means a new time step. The all cells calculate their developing potential values again and those qualified cells transit. Figure 3.11 describe a flowchart showing the simulation of the model.
Start Model from year $t_1$, $t_1 = 2000$

Configure scenario Rule

Running model to year $t_1+n$

Data Initialization
- Population in year $t_1$
- Built-up area in year $t_1$
- Cells state in year $t_1$

Population projection and Land demand for built-up area in year $t_1+1$

Calculate Potential Value

Sorting and ranking all the Potential Value

Update cells base on land demand requirement in year $t_1+1$

Transit into built-up area

Built-up area in year $t_1+1$

Simulation for Built-up area in year $t_1+n$

The result of Built-up area in year $t_1+n$

End

Figure 3.11 A flowchart showing The Simulation of The Model
3.4.4. Calibration and Validation Model using Two Scenarios

In this study, calibration and validation the model simulation rules were used to defining the CA rules to minimize the simulation error. Two scenarios was implemented in this simulation model to calibrate the model. In the first scenario, driving forces and constrained factors (slope > 15%, protected forest, open green space and water body) as an input in the model simulation. Whereas, the second scenario is the simulation model which running without slope effect in model input. Weighted for each driving forces parameter in the first and second scenario was defined automatically by simulation model, based on the result of weighted calculation using Multicriteria evaluation. One important aspect of model calibration is to verify the model’s outcomes and evaluate the goodness of fit of those outcomes with the real world system it modelled. Map simulation result for first scenario describe in Figure 3.12 and 3.13 whereas Figure 3.14 and 3.15 for second scenario.

Figure 3.12 Urban Development in 1991 – 2000 using 1st Scenario
Figure 3.13 Urban Development in 2000 – 2007 using 1\textsuperscript{st} Scenario

Figure 3.14 Urban Development in 1991 – 2000 using 2\textsuperscript{nd} Scenario
Figure 3.15 Urban Development in 2000 – 2007 using 2nd Scenario

Figure 3.12 – 3.15 shows that for two time series model, large patches of urban built-up area development had actually been developed in the center parts of Bandung Area. The other was distributing in Cimahi, Bandung municipality and West Bandung region. Simulation accuracy assessments based on the error matrix and kappa coefficient were applied in model validation process. Simulation accuracies from first and second scenario describe in Table 3.5 and 3.6

Table 3.5 First Scenario Simulation Result

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Producer’s accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Built-up area</td>
<td>86,95</td>
<td>90,61</td>
</tr>
<tr>
<td>Non Urban</td>
<td>99,98</td>
<td>99,75</td>
</tr>
<tr>
<td>User’s accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Built-up area</td>
<td>69,16</td>
<td>66,05</td>
</tr>
<tr>
<td>Non Urban</td>
<td>99,63</td>
<td>99,95</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>99,51</td>
<td>99,69</td>
</tr>
<tr>
<td>K_hat Coefficient</td>
<td>76,80</td>
<td>76,26</td>
</tr>
</tbody>
</table>
Table 3.6 Second Scenario Simulation Result

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Producers accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Built-up area</td>
<td>95.81</td>
<td>94.16</td>
</tr>
<tr>
<td>Non Urban</td>
<td>99.96</td>
<td>99.97</td>
</tr>
<tr>
<td>User’s accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Built-up area</td>
<td>72.95</td>
<td>68.64</td>
</tr>
<tr>
<td>Non Urban</td>
<td>99.68</td>
<td>99.97</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>99.64</td>
<td>99.73</td>
</tr>
<tr>
<td>(K_{hat}) Coefficient</td>
<td>82.66</td>
<td>79.27</td>
</tr>
</tbody>
</table>

The results generated by the model with the second scenario, show that producer’s and user’s accuracies for the built-up area category in 1991 - 2000 were only 95.81 % and 72.95 %, respectively. This means that 4.19 % of the actual urban built-up areas were omitted from being selected for development, and 27.05 % of the simulated built-up area were committed to the category by the model incorrectly. On the other hand, for the same built-up area category simulation accuracies produced by the model have been decreasing in 2000 – 2007 simulation model, were only 94.16 % and 68.64 %, respectively. This means that 5.84 % of the actual urban built-up areas were omitted from being selected for development, and 31.36 % of the simulated built-up area were committed to the category by the model incorrectly.

Validation results generated by the model with the configuration without slope > 15 % as a constraint model show that model \(K_{hat}\) coefficient in second scenario is highest than first scenario. In reality, some of urban development in Bandung area was located in Slope area > 15 %. Table 3.7 was shown urban development area in year 2007 base on suitability of slope range.

Table 3.7 Urban Development Base on Suitability of Slope Range

<table>
<thead>
<tr>
<th>Slope %</th>
<th>Urban Area (Ha)</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 15</td>
<td>34.226</td>
<td>89.04</td>
</tr>
<tr>
<td>16 - 30</td>
<td>3.580</td>
<td>9.31</td>
</tr>
<tr>
<td>&gt; 30</td>
<td>631</td>
<td>1.64</td>
</tr>
<tr>
<td>Total</td>
<td>38.437</td>
<td>100.00</td>
</tr>
</tbody>
</table>
\( K_{\text{hat}} \) coefficient was increasing by applying second scenario. The simulation result from two scenarios describe in Table 3.8

Table 3.8 \( K_{\text{hat}} \) Coefficient for Two Scenarios

<table>
<thead>
<tr>
<th>( K_{\text{hat}} ) Coefficient</th>
<th>Simulation from 1991 to 2000</th>
<th>Simulation from 2000 to 2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>76,80</td>
<td>76,26</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>82,66</td>
<td>79,27</td>
</tr>
</tbody>
</table>

Base on \( K_{\text{hat}} \) Coefficient in the Table 3.8, second scenario is the best to be applying in model simulation. This scenario will be reduces the impact of the mismatched cells on both the producer’s and user’s accuracies for the built-up area development and making the simulation accuracies more realistic to reflect the actual performance of the model.

3.4.5. Projection of Urban Growth Spatial Distribution Model

The Urban growth spatial distribution model was implemented to simulate and projection land demand of built up area in Bandung area 2015 – 2030 was shown in Table 3.9.

Table 3.9 Prediction of Land Demand for Built-up Area in 2015 and 2030

<table>
<thead>
<tr>
<th>Years</th>
<th>Population Projection (People)</th>
<th>Land demand for built-up area (Ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>7,439.622</td>
<td>38.437</td>
</tr>
<tr>
<td>2015</td>
<td>9,342.096</td>
<td>41.596</td>
</tr>
<tr>
<td>2030</td>
<td>14.260.742</td>
<td>51.433</td>
</tr>
</tbody>
</table>

The spatial distribution of urban growth development base on Table 3.9 can be seen in Figure 3.17 and 3.18. Comparing the simulated scenarios of urban development in 2015 and 2030 with the actual urban extent in 2007, it is clear that urban development in Bandung Area will continue in two ways: one is through the in-filling of areas within the currently existing urban in Cimahi and Bandung City, and another is through expanding toward the southeast of the areas.
Figure 3.16 Urban Development Simulation in Bandung Area, 2015

Figure 3.17 Urban Development Simulation in Bandung Area, 2030
Through the in-filling process, some of the vacant land within the existing urban areas will be developed into urban areas, and some areas that are already developed to a certain extent will continue to be further developed into a fully urban state. The result from urban development simulation in 2015 and 2030, shows that urban development in Bandung Area will continue in two ways: one is through the in-filling of areas within the currently existing urban in Cimahi and Bandung City, and another is through expanding toward the southeast of the areas.

3.4.6. Urban Land Carrying Capacity in Bandung Area

Analyze of Urban Land carrying capacity was done to know how many amount of people can be limited urban land support in Bandung Area. In the UGSDM application, urban land to support population people in Bandung area is 144.752 Ha or for about 43.93 % from the total of Bandung Area. Amount of people calculation that can be allocated in Urban Land support in Bandung area, can be calculated using formulation bellow:

\[
y = 0.019x + 24.903 \Rightarrow 144.752 = 0.019x + 24.903
\]

\[
x = \frac{(144.752 - 24.903)}{0.019} = 64.728.947
\]

Furthermore, modification from population projection formula can be used to derive how many years will be need to calculate maximum urban land carrying capacity in Bandung Area. The result from the calculation of maximum urban land carrying capacity in Bandung Area can be seen in Figure 3.18.

\[
P_t = P_0(1+r)^n \iff n = \frac{Ln\left(\frac{P_t}{P_0}\right)}{Ln(1+r)} \iff n = \frac{(1+r)}{Ln(1+r)} \log\left(\frac{P_t}{P_0}\right)
\]

\[
n = (1.0286) \log\left(\frac{64.728.947}{6.294.961}\right) \approx 82
\]
3.5. Conclusion

Integration of GIS and Cellular Automata was proved very efficient in simulating the urban growth simulation over time. Two scenarios have been implemented in this simulation model. In the first scenario, driving forces and constrained factors (slope > 15%, protected forest, open green space and water body) as an input in the model simulation. Whereas, the second scenario is the simulation model was running without slope effect in model input. $K_{hit}$ Coefficient for first and Second scenario in 1991 – 2000 simulation model is for about 76.80% and 82.66% respectively, whereas in 2000 – 2007 is for about 76.26% and 79.27% respectively. Based on the result, second scenario is the best to be applying in model simulation and can be used to predicting urban growth simulation. This scenario will be reduces the impact of the mismatched cells on both the producer’s and user’s accuracies for the built-up area development and making the simulation accuracies more realistic to reflect the actual performance of the model.

The result from urban development simulation in 2015 and 2030, shows that urban development in Bandung Area will continue in two ways: one is through the in-filling of areas within the currently existing urban in Cimahi and Bandung City, and another is through expanding toward the southeast of the areas.