Beni-Suef University Engineering Journal
Contents lists available at Science Direct

Journal of Engineering Science and Sustainable Industrial Technology

Journal homepage: www. jessit.journals.ekb.eg

Evaluation And Comparison of Simulating Urban Growth Using The Markov Chains and Logistic Regression Models (The Case of Beni-Suef–Menya–Asuit Cities)

Amira M.El-Gabalawy[1]*, Tark Abdel-Latif Aboul Atta[2]

[1] Assistant Lecturer, Department of Architecture, Beni-Suef University, Egypt.
[2] Professor of Urban and Regional Planning, Department of Architecture, Cairo University, Egypt.

A R T I C L E I N F O

Article history:
Received: 31 Dec. 24
Accepted: 22 March 25
Online:03 May 25

Keywords: Urban Growth, Cellular Automata, Markov Chain, Logistic Regression

A B S T R A C T

As a result of urban growth, especially in developing countries and its negative effects its results are pressure on the natural environment, increased pollution, loss of biodiversity, loss of open spaces, climate changes and many other negative impacts. This research is therefore interested in the identification of modern technological methods for the simulation of urban growth and the comparison between the two models (Cellular Automata-Markov Chain) (CA-MC) and (Cellular Automata-Markov Chain-Logistic Regression) (CA-MC-LR) by application to the cities (Beni-Suef-Menya-Asuit).The research results showed that both models have a high ability to simulate and predict urban growth and therefore government must integrate new research methodologies and modern technologies into the process of urban modeling and simulating urban growth and land use changes. Also, what determines which of the two models is better is the goal of the study. If the goal of the study is predict urban growth only the model (Cellular Automata - Markov Chain) is better to use. However, if the goal of the study is to predict urban sprawl and determine the degree of influence of each of the independent factors in the process of urban growth, the model (Cellular Automata-Markov chain-Logistic Regression) is better to use. The results also showed that the factors influencing the urban growth process vary across the study cities, with each city experiencing different degrees of influence. Therefore, it is essential to analyze the influencing factors for each city individually, considering the variations in their social and economic environments.

1. Introduction

The rapid increase in population growth worldwide, especially in developing countries, is one of the main challenges faced by governments. Currently, 54% of the population lives in urban areas, which is estimated at 3.9 billion people, and this number is expected to reach 6.3 billion people by 2050, or nearly 90% of the future population increase in cities [1]. It is certain that the large increase in population affects the environment, which has

* Corresponding author. Amira El-Gabalawy, Teaching Assistant, Faculty of Engineering, Beni-Suef University, Beni-Suef, Egypt. Mobile no. 01032840843. Email: amiramagdy@eng.bsu.edu.eg.

limits and the capacity to absorb this increase in population, leading to pressure on the natural environment and the depletion of natural resources. This is evident in the transformation of agricultural lands into urban lands in a random manner, as well as the loss of biodiversity due to the removal of trees and forests, loss of water bodies and open spaces, and many climate changes. Rapid population growth has also led to pressure on facilities and unequal distribution of population and facilities.

To deal with this growth and achieve sustainability, decision-makers and urban planners need accurate information about urban growth. Therefore, the analysis of land use change (LUC) has received considerable attention. Moreover, urban growth is one of the most widely discussed topics in urban studies. The process of urban growth and land use change (LUC) is a complex process that occurs as a result of a set of social, economic, and physical factors, leading to the emergence of many theories that discuss urban growth processes, either in terms of their causes or effects resulting from them [2].

Traditional spatial and temporal simulation and analysis methods are limited to extracting temporal data, which makes it difficult to study changes associated with many important geographical phenomena. On the other hand, land use change data are available over multiple time periods through Remote Sensing (RS), such as the Landsat satellite. Land use changes can be studied, and these complex dynamic phenomena can be simulated and predicted.

Monitoring urban growth and land use changes (LUC) using Geographic Information Systems (GIS) and Remote Sensing (RS) has proven to be a very effective tool for monitoring urban growth over time and as a source of primary information for simulating land use changes (LUC).

This study examined studing, analysis, and simulation of urban growth from theoretical and applied perspectives. From a theoretical perspective, the modern methods of modeling urban growth and land use changes applied in this study were addressed. The Markov chain (MC), Cellular Automata (CA), and Logistic Regression (LR) models were presented in terms of the inputs and outputs of each model and the mathematical representation of each model. The strengths and weaknesses of these models are also presented. The integration between the Cellular Automata model and the Markov Chain (MC-CA) and the integration between the Cellular Automata model, Markov Chain and the Logistic Regression (CA-MC-LR) are discussed.

As for the practical aspect of the study, the Remote Sensing technology (RS) was used to obtain historical data for the study areas, as satellite images were obtained from the United States Geological Survey (USGS). The correction, initial processing, and classification of satellite images were carried out using ENVI 5.3. Using Geographic Information Systems (GIS), a geographic database was created for the three study areas after verifying the accuracy of the maps using the Kappa coefficient (KAPPA) when compared to reference maps from Google Earth.

Therefore, the main objective of the study is to model and simulate urban growth and land use changes using modern modeling methods by applying the integrated model (Cellular Automata model -Markov Chain) (MC-CA) and the integrated model (Cellular Automata model-Markov Chain-Logistic Regression) (CA-MC-LR), and compare the accuracy results of both models and conclude which model is more accurate, and compare the outputs of both models, and predict urban growth for years 2030-2040-2050.

2. Aims of the Study

In recent years, several studies have proven the importance of using modern technologies to make planning decisions. Therefore, the aims of this study are as follows:

• Evaluateg andcompareg the results of urban growth simulation using Markov chain (MC) and logistic regression (LR) models andconcludeg which of the two models is more accurate in the process of predicting urban growth.

• Modeling and simulating urban growth and land use changes using modern modeling methods such as applying the integrated model (Cellular Automata model-Markov chain) (MC-CA, and the integrated model (Cellular Automata-Markov chain-Logistic Regression) (CA-MC-LR).

• Knowing and identifying trends in urban growth and land use according to scientific foundations contributes to understanding the future and identifying the changes that affect the environment, thus making the right decisions and developing future plans to preserve and develop the environment.

3. Literature Review

The literature review indicates that previous studies analyzing future urban growth patterns through the Cellular Automata (CA), Markov Chain (MC), and Logistic Regression (LR) models fall into two categories: those that focus on theoretical exploration and those that emphasize practical application.

A review of previous studies revealed the following:

-There is a scarcity of studies that compare different models for the prediction and simulation of future urban growth, which is more accurate.

Previous studies have emphasized the importance of spatiotemporal analysis of urban growth. This is because decision-makers and urban planners need accurate and detailed information about urban growth to evaluate the size of development, its location, characteristics, and the consequences of previous and subsequent urban development.

The studies agreed on the data sources used in the study, as almost all studies used satellite images to obtain maps of previous land uses due to the scarcity and unavailability of sources.

- These studies varied in terms of the tools used to measure urban growth and predict its future status. The programs used to correct and classify the satellite images varied. Some studies used the ENVI program, others used the IDRISI Selva program, and others used the MAP ARC program.

The discrepancy between the results of studies related to the factors affecting urban growth, as the study by Jafari et al. [3] indicated that the most influential factor is proximity to the existing infrastructure and along the main roads. Omar Hamdy and others [4] also proved that proximity to service areas is the most influential factor on urban growth. The study by Asfa Siddiqui [5] also indicated that proximity to the city service center and main roads are the most influential factors on urban growth. It is clear that the results vary from one country to another; therefore, the driving factors affecting urban growth should be studied separately for each case study. Finally, the studies showed that the integration of RS and GIS is an effective method for measuring land-use change.

4. Research tools

The research tools used in this study vary according to the research problem. In terms of the programs used, the ENVI 5.3 program was used, which is one of the programs used to analyze and process geographic spatial images by scientists and researchers in various parts of the world. It was used in this research in the stage of classifying satellite images, with the help of the United States Geological Survey (USGS) website.

The ARC GIS program is an integrated geographic information system issued by (ESRI) company, which performs many functions, including editing, displaying digital data, dealing with layers, and adding elements to maps. This was used in this study to create a database and produce the final maps.

IDRISI Selva is integrated geospatial software for monitoring and modeling the Earth system and is remote sensing software developed by Clark Laboratories at Clark University to analyze and display digital geospatial information. This system provides tools for researchers and scientists involved in the analysis of land-use changes to make effective and responsible decisions for environmental management and sustainable resource development. It also includes tools for importing and exporting all the major file formats and images. It was used in this study to apply the models used in this study (CA-MC) and (CA-MC-LR) and the Land Change Modeler tool (LCM) within the program.

5. Research methods

5.1. Cellular Automata (CA)

It is a separate model with a spatially extended dynamic system the new previous state of the type of land use/land cover (LULC) and the state of the cells adjacent to it. It also has the ability to represent nonlinear and complex processes distributed spatially; thus, it has the ability to predict future changes in land use and land cover. It is one of the techniques that benefits from the integration of Remote Sensing and Geographic Information Systems [6].

A ((CA) consists of four main components (Figure 1) [7]. **Area**: It is a group of cells distributed on a regular grid, for example, 3*3 5,*5, and it can be a multi-dimensional matrix D2, D3. This area lies the geographical map, which consists of cells (pixels) that can be described as 0, 1, or full and empty. In an urban environment, it can be described in terms of use (residential/commercial), density (high/low), and land cover (forests/desert land). **Cell**: This is a single unit of area, and the cells interact with each other according to a set of urban growth rules, where each cell can transform in the future to different states from its current state depending on its geographical location, characteristics, and urban growth rules. **Cell State**: The state of each cell is expressed at a certain time in an urban environment. The cell state can be represented as (urban/non-urban) or (0,1). **Neighborhood**: The neighborhood in Cellular Automata (CA) is defined as the group of cells surrounding the cell under study that may affect its state in the future. There are two commonly used types of neighborhoods: Moore and Von. The neighborhood is called Moore if the cell under study is surrounded by eight cells and Von if the cell is surrounded by four cells. The Von Neighborhood consists of five adjacent cells, and the Moore neighborhood consists of nine adjacent cells.

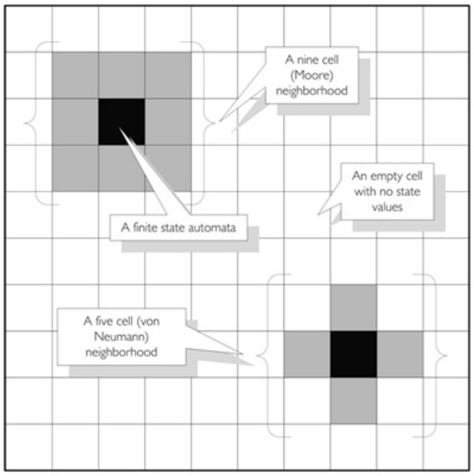

Fig. 1. Components Of Cellular Automata (CA).

Urban growth rules: Urban growth rules are an important element of the Cellular Automata model (CA) is the urban model. It is a set of rules to determine the probability of changing the state of each cell and is represented by the weights associated with the different spatial variables (distance from the city center, slopes, neighboring areas, etc.). It determines the state of the cell in the next time period based on its current state and the state of the neighboring cells.

The Cellular Automata model (CA) can be presented by applying (Equation 1) [6,8,9].

$$\mathrm{S}(\mathrm{t}+1)=\mathrm{f}\left[(St)\times\left(\mathrm{I}_{\mathrm{t}}^{\mathrm{h}}\right)\times(\mathrm{V})\right] \qquad (1)$$

Equation.1 Cellular Automata Model

Where: S(t+1): the state of the cell at time(t +1); S(t): the state of the cell at time(t); ($\mathrm{I}_{\mathrm{t}}^{\mathrm{h}}$): the neighboring cells; (v):the suitability of the cell for urban growth; f(): the rules of urban growth; T: the time periods of the study; H: the size of the neighboring cells; S: are the states of the specific cells

The Cellular Automata (CA) model demonstrates high spatial accuracy, as it focuses on the spatial analysis of the elements under study and examines the relationships between adjacent cells within the neighborhood [10,11]. Moreover, CA can be integrated with Geographic Information Systems (GIS) and Remote Sensing (RS). The model is also characterized by its simplicity; compared to many other urban models, its structure and components are relatively straightforward, making it easier to understand and implement [10].

However, like any model, it has advantages and limitations. The main developers of CA (Turing and von Neumann) described it as having a fixed structure, and this structure works well in simulating physical systems in biology, chemistry, and physics, but it is less suitable as a method for studying social, environmental, and economic studies in cities. It does not consider external factors; the influential factor is more interested in the spatial-temporal study of the elements. The CA model is limited to specific general rules, and in reality, not all urban studies affect all variables with the same effect. The idea of CA is that it is an infinite (two-dimensional) and regular space, and this cannot be applied to cities because cities are not infinite or regular [12].

5.2. Markov Chain Model (MC)

Markov Chain (MC) is a popular model that describes the probability of a land use transitioning from one state (X) to another state (X+1) during a specified time period [6]. Future predictions were made using the Markov Chain model based on the analysis of two land-cover images at different dates [13]. The state of the land cover at a specified time (t) depends on its state in the previous time period (t-1) [14]. The Markov Chain model is known to be more powerful and effective for simulating changes in any region, and is widely applicable for predicting rapid urban growth areas [4]. It has been widely used to study urban growth in cities such as Uttar Pradesh in India [5], the Greater Cairo area [15], Siliguri [16], and China [17].

The model outputs consist of three elements: a transition probability matrix, transition probability area, and urban growth probability images, which represent the probability that each pixel will transform into another specific category in the future [14,6].**Urban growth probability matrix**: The Markov matrix can be defined as a series of random variables containing urban growth probabilities in a square matrix (N*N) called the urban growth matrix (transition Matrix) or Markov matrix and it is characterized by the fact that each of its elements is not negative. It is produced in the form of a text file containing the probability of change, and the sum of each row equals one (Figure2).

	Cl. 1	Cl. 2	Cl. 3	Cl. 4
Class 1 :	0.8980	0.1020	0.0000	0.0000
Class 2 :	0.1241	0.8710	0.0024	0.0025
Class 3 :	0.0886	0.0496	0.8618	0.0000
Class 4 :	0.0003	0.1000	0.0000	0.8997

Fig. 2. Urban Growth Probability Matrix.

Urban growth probability areas: This is a matrix that shows the number of pixels that are expected to change from one state to another during a specific period. This matrix determines the amount of change that can be used to study future changes (Figure 3).

	Cl. 1	Cl. 2	Cl. 3	Cl. 4
Class 1 :	53324	6058	0	0
Class 2 :	78682	552441	1547	1574
Class 3 :	970	543	9432	0
Class 4 :	15	4551	0	40945

Fig. 3. Urban growth probability area

In both matrices, the rows represent old land use, and the columns represent new land use. **A probability image** expresses the probability that each pixel will change from one category to another in the future. The Markov chain (MC) model explains the probability of urban growth over two time steps but does not explain the effect of neighborhood cells on the prediction of future growth [5].

The Markov chain (MC) model for predicting land-use changes can be described mathematically using the probability equation described by (Equation 2,3) [6,8,9,11,18].

$$\mathrm{S(t+1) = Pij \times S(t)} \qquad (2)$$

Equation.2 Markov Chain Model

$$\mathrm{Pij} = \begin{bmatrix} \mathrm{P11} & \cdots & \mathrm{P1n} \\ \vdots & \ddots & \vdots \\ \mathrm{Pn1} & \cdots & \mathrm{Pnn} \end{bmatrix} (0 \le \mathrm{Pij} \le 1 \text{ and } \sum_{j=1}^{n} \mathrm{Pij} = 1, \mathrm{i, j} = 1,2, \dots, \mathrm{n} \qquad (3)$$

Equation.3 Urban Growth Probability Matrix

Where: j: the urban growth probability matrix; i,j: the land uses in the time period t,t+1; N: the number of land uses; s(t),s(t+1) the state of the land cover in time t ,t+1

Markov Chain (MC) model gives the size and quantity of land uses changes and their prediction. It is possible to determine the probability of each use changing from one land use to another [6,10]. However, it does not provide the correct spatial distribution of land cover changes. The Markov Chain model is not a spatial model, so the Markov Chain model is not suitable for determining the location of the change [6,11,10].

5.3. Integration between the Markov Chain Model and the Cellular Automata Model (MC-CA)

The Markov Chain model alone cannot predict future urban growth because it provides quantitative information about the amount of change and not spatial information. To solve this problem, the Markov chain model (MC) was integrated with the CA model. The CA-MARKOV model is considered one of the most effective models for modeling spatial-temporal changes in land cover and land use in addition to Geographic Information Systems (GIS) technology [9]. It is an integrated model that effectively combines the advantages of the Markov chain to predict changes in land cover and land use, and the CA model to simulate future spatiotemporal changes [6].

5.4. Logistic Regression Model (LR)

MC Fadden (1973) developed the Logistic Regression model. This model represents the relationship between the dependent variable (urban growth) and independent variables (driving factors). The dependent variable was binary and the independent variables did not have a linear relationship. These variables can be nominal, ordinal, or a function of the probability of their occurrence or non-occurrence, depending on the relative. The dependent variable in this model takes the value (0,1) [19]:

Logistic Regression was preferred if the variables were a set of categorical and continuous variables. The expected dependent variable in a Logistic Regression model is a function of the probability that it is a factor on a set of in one of the categories; for example, the probability of a change to a particular land use based on a particular influence on land use, such as proximity to a transportation network. [19].

In the IDRISI Selva program, the LOGISTICREG tool applies binary Logistic Regression, in which the dependent variable must be binary, as it can take two values (0,1). The basic hypothesis is that the probability of a dependent variable taking the value 1 (positive response) can be calculated through (Equation 4). To convert the model to linear and remove the probability limits (0,1) for the dependent variable, transformation is applied. After this transformation, it can be assumed that the probability value lies between positive and negative Malanhais (Equation 5). The logit transformation of both sides of (Equation 5) leads to the formulation of a linear regression model, presented in (Equation 6). After the Logit transformation, the dependent variable continues to be in the range of 0 to1 [16]:

$$P\left(Y = \frac{1}{X}\right) = \frac{(exp \sum BX)}{1 + \exp\left(\sum BX\right)} \qquad (4)$$

$$P` = \ln \frac{P}{1-P} \qquad (5)$$

$$\ln \frac{P}{1-P} = b_0 + b_1 X_1 + b_2 X_2 + \cdots + b_n X_n \qquad (6)$$

Equation 4,5,6 : General Equations For The Logistic Regression Model (LR)

Where: P: the probability of the occurrence of the dependent variable; X: the independent variables $X = (X_{0,} X_1, X_2, \dots, X_n)$; B:Regression coefficients $B = b_0, b_1, b_2, \dots, b_n$.

The regression coefficients from b_1 to b_n indicate the effect of each independent variable on the probability value. A positive value indicates a positive effect and a negative value indicates a negative effect.

Logistic Regression Model (LR) can predict the relationship between independent and dependent variables[16], explain the degree of influence of independent variables on dependent variables, find the most likely locations for future urban growth through a map of urban growth potential[3,4,16], and clarify the relative importance of factors and their importance on the dependent factor[20].

Despite the many advantages of the model, it suffers from some difficulties in modeling variables with multiple timescales. If both land use and common variables change values over time, Logistic Regression must use the value at one time or the average over time, which leads to a loss of information [16]. The actual change in the area over time cannot be determined [3].

An integrated approach combining technology based on a previous theoretical background. In the current study, the Remote Sensing and Geographic Information Systems approach is used to simulate and predict future changes in the study areas based on the (CA-MC) and (CA-MC-LR) models. Thus, the future spatial and temporal patterns of the study areas can be predicted. It is expected that the results of this study will benefit planners and decision makers to understand the future impacts of changes in land cover, which will enable them to develop appropriate policies to achieve social and economic development in the study area.

6. Study methodology:

The methodology followed for modeling and simulating urban growth and land cover change (LULC) consists of several stages, as shown (Figure 4).

The methodology is discussed in the following section.

6.1. Data source

A variety of data were collected from multispectral satellite images. According to the main objective of the study, which was to evaluate and compare the results of urban growth simulation using the Markov chain model and logistic regression, it was necessary to obtain satellite images at different time periods. In this study, images from the Landsat satellite for the sensors (ETM+ and OLI) were used for the past 20 years according to the coordinate system (N36 Zone84 WGS-UTM). (Table 1) shows the data and their sources.

6.2. Data processing

This stage is concerned with the initial processing of satellite images and an increase in the accuracy of satellite images. To do this step, it is necessary to use some programs that are concerned with sensing operations. In this study, the ENVI 5.3 program was used due to the accuracy of the satellite image analysis process and its proven effectiveness in previous studies concerned with this field.

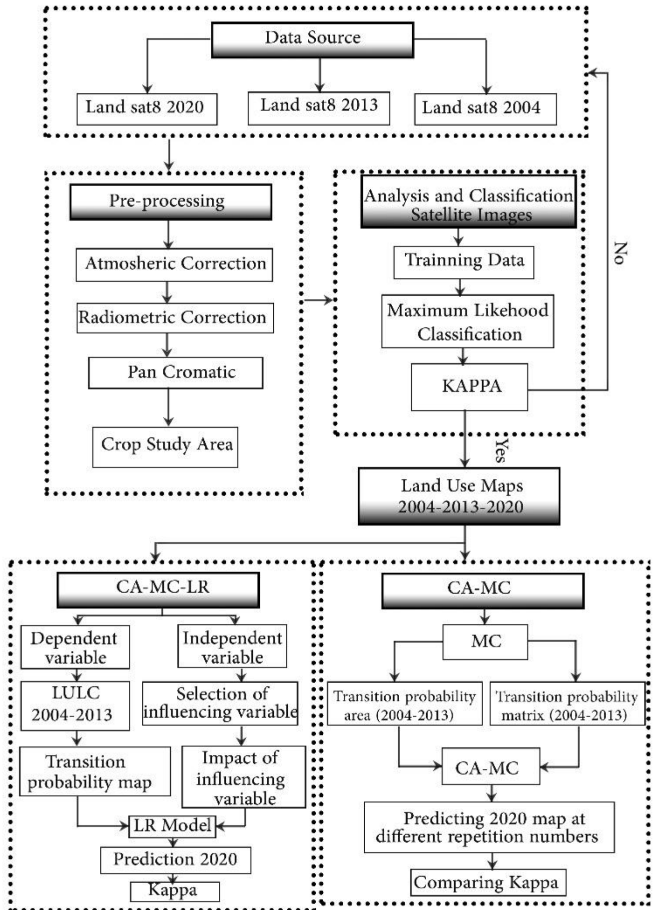

Fig. 4. Study Methodology.

Table 1
Data Source.

Data	Scale	Source
Landsat (OLI,+ETM)	30m	USGS [21].
Roads ,Rivers , Canals	1:10000	Urban Planning Authority.
Boarders	-	https://www.diva-gis.org [22].
DEM ,Slope	30m	STRM[21].
Reference Maps	2.4m	Google Earth.

The radiometric correction was done in the ENVI 5.3 program. And also increasing the accuracy of the satellite images by using the Ban chromatic image to have an accuracy of 15 meters, then cutting the study area from the satellite image because the satellite image covers a large area, according to the UTM Coordinate System.

6.3. Data classification

Then comes the process of classifying satellite images, which are transformed from satellite images into digital maps that show the land cover. The controlled classification method was used as three types of land cover were chosen: urban areas, non-urban areas, and water. The maximum likehood classification method was used to classify satellite images by ENVI 5.3 program.

6.4. Map accuracy evaluation

After obtaining land use maps from satellite images, it was necessary to compare the actual maps (Google Earth) with the maps extracted from satellite images. After the map classification process in ENVI 5.3 program, the maps were exported to the ARC MAP program, and the map accuracy evaluation stage was presented. The Matrix Error method and kappa coefficient were used to evaluate the accuracy of the maps by comparing the land cover maps and reference maps. The Google Earth program was used with an accuracy of 2.4 meters [23,24], and with the help of the ARC MAP program, 150 random points were selected based on previous studies.

The accuracy of the maps of Beni-Suef city has reached 90.51, 91.55, 91.64 for years 2004 , 2013 , 2020. The accuracy of the maps of Menya city has reached 94, 93, 95 for years 2004 , 2013 , 2021. The accuracy of the maps of Asuit city has reached 93, 91.5, 94.68 for years2004 , 2013 , 2021.

After deducing land cover maps from satellite images, the model must be verified before applying it to future prediction processes. The 2004 and 2013 maps will be entered for predicting with the 2020 map and the simulation map (prediction map) will be compared with the reference map (which was deduced from Landsat satellite images) and its accuracy will be measured using the KAPPA coefficient

6.5. Cellular Automata -Markov Chain Model (CA-MC)

Using the IDRISI Selva program, the CA-MARKOV model was applied; the maps were entered and defined within the program. To apply the CA-MARKOV model, several steps are followed. First: The Markov chain model is applied, by entering two land cover maps during two time periods, which are land use maps for the years 2004 and 2013, to predict the 2020 map. Then, the number of years between these two maps is determined, which is 9 years, as well as the number of years between the prediction map and the last time map, which is 7 years. This results in an urban growth probability matrix, urban growth probability area, and probability images. Second, the CA-MARKOV model was applied using the urban growth matrix and the urban growth probability image for the regions. These data are entered and different repetition numbers are entered (5-10-20-30-40-80-100-160-200-250-300). Then, the process of comparing the accuracy of the simulation maps with the reference maps at these different repetition numbers is carried out. The repetition number at which the map reaches the greatest accuracy is chosen.

The outlined methodology will be implemented in a series of case study cities:

- **Beni-suef :**

The Markov chain model (MC) was applied by entering two land cover maps from 2004 and 2013 to predict the 2020 map. From previous step we have urban growth probability matrix, urban growth probability areas, and probability images.

The CA-MARKOV model was applied using the urban growth probability matrix and the urban growth probability image and then entering different repetition numbers. As a result, multiple prediction maps were generated corresponding to different numbers of iterations, as shown(Figure 5).

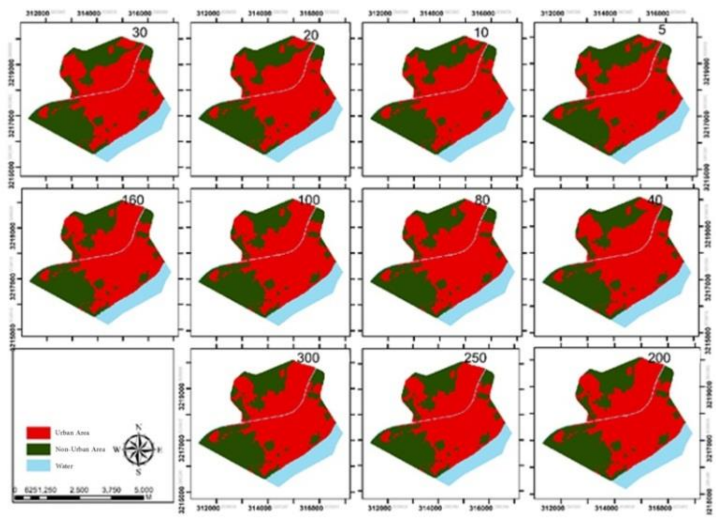

Fig. 5. Simulation maps for 2020 At Different Frequency Numbers (Beni-Suef).

The accuracy of all simulated maps was assessed using the Kappa. Subsequently, the simulation maps were compared with the reference maps across the various iteration counts. The highest accuracy was recorded at an iteration count of 10, with a Kappa value of 0.95 (Figure 6).

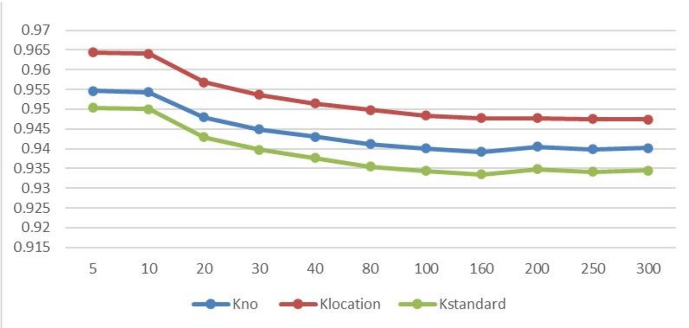

Fig. 6. Comparison Of Kappa Coefficient At Frequency Numbers (Beni-Suef).

- Menya

The same methodology was applied to the city of Beni-Suef, resulting in the generation of multiple prediction maps based on varying numbers of iterations, as illustrated in (Figure 7). The accuracy of the simulation maps was then compared with the reference maps across these different iteration counts. The highest accuracy was achieved at 250 iterations, with a Kappa value of 0.966 (Figure 8).

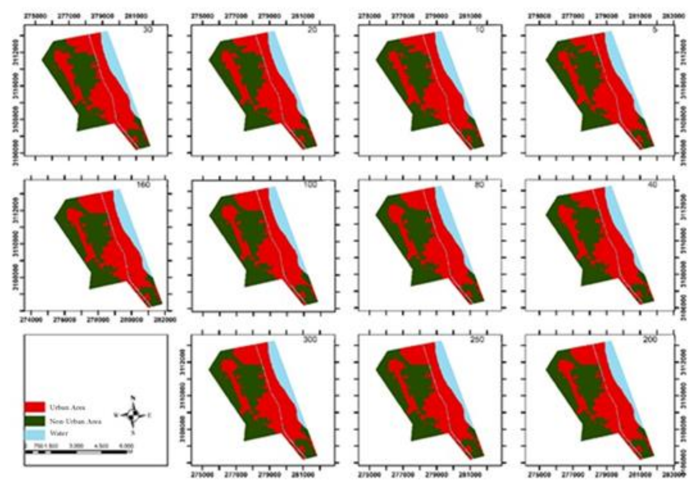

Fig. 7. Simulation maps for 2021 at different frequencies (Menya).

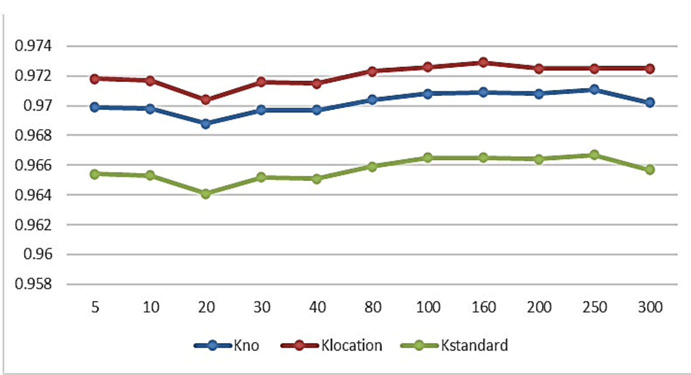

Fig. 8. Comparison of kappa coefficients at frequency numbers (Menya).

- Asuit

The same steps were carried out in the cities of Beni-Suef and Menya will be followed. We produced different prediction maps for different numbers of repetitions, as shown in (Figure 9). The accuracy of the simulation maps was compared with the reference maps across different iteration counts. The highest accuracy was observed at 10 iterations, with a Kappa value of 0.94 (Figure 10).

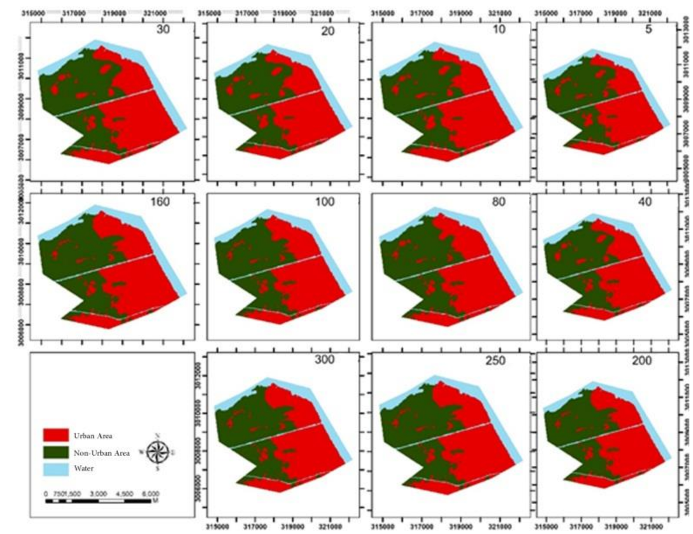

Fig. 9. Simulation maps for 2021 At Different Frequency Numbers (Asuit).

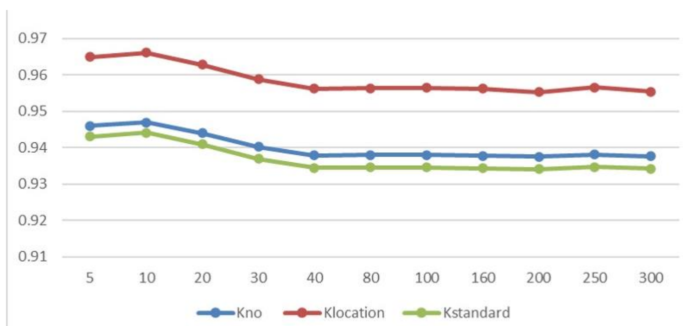

Fig. 10. Comparison Of Kappa Coefficient At Frequency Numbers (Asuit).

6.6.Cellular Automata-Markov Chain- Logistic Regression Model (CA-MC-LR)

In this part of the study, urban growth was simulated using the Logistic Regression model (LR) for each of the study cities (Beni-Suef-Menya-Asuit), the dependent and independent variables were determined, and the effect of the independent variables on the dependent variable was studied. The accuracy of the model and the results were verified. To apply the model, several steps are followed:

First, determining the dependent variable: The dependent variable in this study is the amount of change from a non-urban area to an urban area during two time periods, represented in a binary map that takes values (0,1) where the value 1 indicates the occurrence of a change (the land uses that were converted from non-urban to urban) and the value 0 indicates the absence of change. To obtain the map of the dependent variable in this study, the IDRISI Selva program was used. IDRISI using the Land Change Modeler tool (figure 11). Second, determining Independent variables: The process of selecting variables and factors affecting the growth process urban growth is a difficult process because the urban growth process itself is complex and it is affected by a large group of factors and these factors are also dynamic and vary according to place and time. By studying the study area, previous studies [5,10,15,16], and the available data, 10 variables were identified for the study (Table 2).

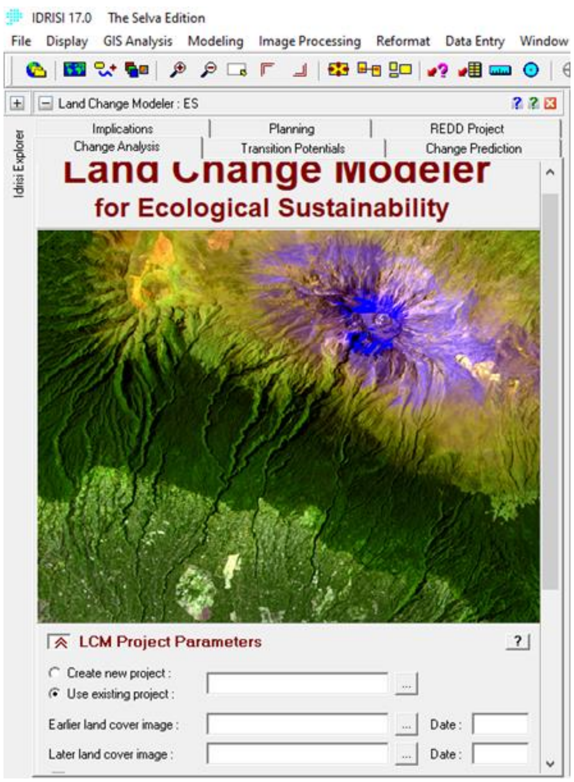

Fig. 11. Land Change Modeler Tool In IDRISI Selva

Table 2
Variables For Logistic Regression Model

Variable		Type
V0	LULC (0,1)	In-Dependent variable
V1	Slope	Dependent variable
V2	DEM	Dependent variable
V3	Distance from health services	Dependent variable
V4	Distance from educational services	Dependent variable
V5	Distance from services	Dependent variable
V6	Distance from roads	Dependent variable
V7	Distance from railways	Dependent variable
V8	Distance from river	Dependent variable
V9	Distance from canals	Dependent variable
V10	Distance from urban area	Dependent variable

7. Calibration of independent variables and logistic regression (LR) model

First, calculating the value of the Odds Ratio: It necessary to calculate the value of the Odds Ratio for all variables to study the independent variables affecting the dependent variable using the LOGISTICREG tool within the IDRISI program. The LOGISTICREG tool was applied ten times for each variable separately to obtain the Odds Ratio.

Second, calculating Cramer's V: This test was conducted using the LCM tool (inside the IDRISI Selva program). The Cramer's V test was performed for all variables. If the value of Cramer's V was greater than 0.15. This means that the variable has a strong relationship with urban growth.

Third, Applying the (MC) model: After verifying the calibration process of the independent factors and ensuring their impact on the urban growth process, the urban growth matrix for the period 2004-2013 was calculated to predict the 2020 map by applying the Markov chain model and specifying the prediction year using the LCM tool in the IDRISI Selva program.

To verify the accuracy of the Logistic Regression model, the Kappa coefficient was calculated by comparing the actual map for 2020 and the predicted map from the (CA-MC-LR) model by applying the VALIDATE tool.

- **Beni-suef**

The CA-MC-LR model was applied to the city of Beni-Suef and calibrated and validated to ensure its accuracy. The odds ratios were calculated for all variables (Table 3) (Figure 12). The results showed that the ten factors affected urban growth to varying degrees, as indicated by the Odds Ratio. The Odds Ratio for distance from educational services is 0.1630, which means that the expected urban growth in an area close to educational services is estimated at 6.13 times more than the expected urban growth in an area one kilometer away from educational services.

When Cramer's V was calculated for the variables, all values were found to be greater than 0.15, indicating a strong relationship with the urban growth process. Therefore, no variable needed to be excluded (Table 3).

Table 3
Oddes Ratio, Odds Ratio and Coefficient For Dependent Variable (Beni-suef)

Variable	Odds Ratio	Cramer's V	Coefficient
V1	2.1980	0.2354	0.01020514
V2	0.3756	0.3356	0.18087514
V3	0.8382	0.3892	0.00075965
V4	0.1630	0.4657	- 0.00186803
V5	1.1986	0.4225	- 0.00137068
V6	0	0.3465	- 0.00673695
V7	0.2743	0.5541	0.00166505
V8	1.6910	0.6154	- 0.00095545
V9	0.3192	0.5725	- 0.00196483
V10	0.2977	0.4671	- 0.00013403

After verifying the calibration process of the independent factors and ensuring their impact on the urban growth process, the transition potential from non-urban to urban areas was produced (Figure13) and The urban growth matrix for the period 2004-2013 was calculated to predict the 2020 map by applying the Markov chain model and specifying the prediction year within the LCM tool in the IDRISI SELVA program. The Kappa coefficient was calculated by comparing the actual map for 2020 and the predicted map from the (CA-MC-LR) model. By applying the VALIDATE tool, the Kappa coefficient was found to be 0.9510, indicating the validity of the model and its applicability.

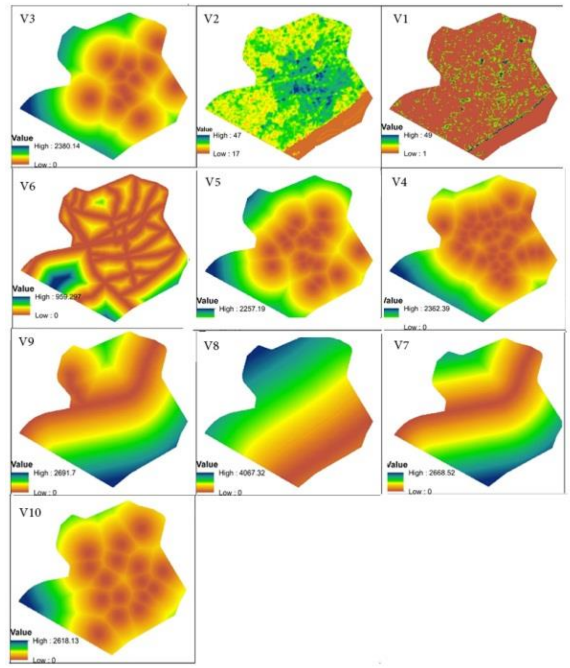

Fig. 12. Dependent Variable For Beni-suef City.

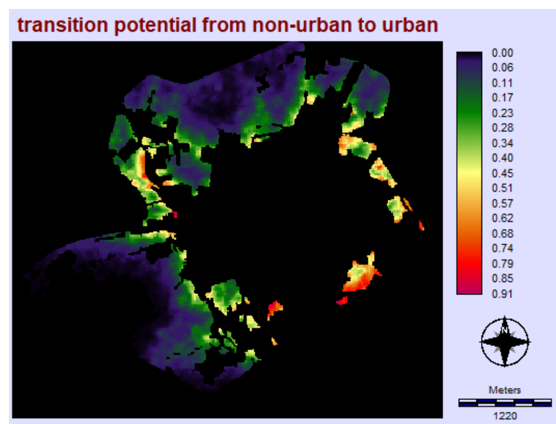

Fig. 13. Transition Potential From Non-Urban To Urban (Beni-suef).

- **Menya**

The CA-MC-LR model was applied to the city of Menya, and the model was calibrated and validated to ensure its accuracy. The odds ratios were calculated for all variables (Table 4) (Figure 14). The results showed that the ten factors affected urban growth to varying degrees, as shown in the Odds Ratio. The Odds Ratio for distance from rivers is 0.0929, which means that the expected urban growth in an area close to rivers is estimated at 10.76 times more than the expected urban growth in an area one kilometer away from educational services.

When Cramer's V was calculated for the variables, all values were found to be greater than 0.15, indicating a strong relationship with the urban growth process. Therefore, no variable needed to be excluded (Table 4).

Table 4
Oddes Ratio, Odds Ratio and Coefficient For Dependent Variable (Menya)

Variable	Odds Ratio	Cramer's V	Coefficient
V1	2.1355	0.5366	0.02841157
V2	0.1363	0.3364	0.03997126
V3	0.1708	0.4575	-0.00275405
V4	0.5578	0.4103	-0.00060755
V5	0.2430	0.4782	0.00698663
V6	0.3553	0.3379	-0.00085765
V7	0.5217	0.5074	-0.00111955
V8	0.0929	0.7173	-0.00189184
V9	0.2881	0.5091	-0.00146336
V10	0.5463	0.5221	-0.00322520

After verifying the calibration process of the independent factors and ensuring their impact on the urban growth process, the transition potential from non-urban to urban areas was produced (Figure 15). The urban growth matrix for the period 2004-2013 was calculated to predict the 2020 map by applying the Markov chain model and specifying the prediction year within the LCM tool in the IDRISI SELVA program. The Kappa coefficient was calculated by comparing the actual map for 2020 and the predicted map from the (CA-MC-LR) model. By applying the VALIDATE tool, the Kappa coefficient was found to be 0.9559, indicating the validity of the model and its applicability.

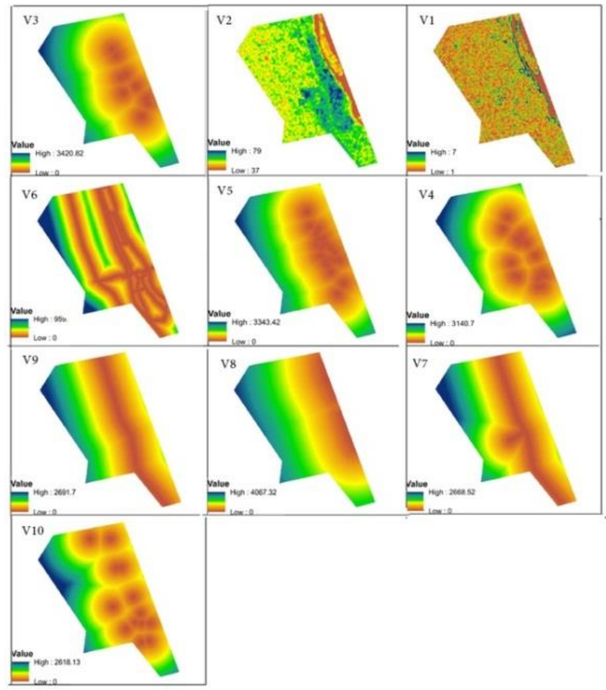

Fig. 14. Dependent Variable For Menya City.

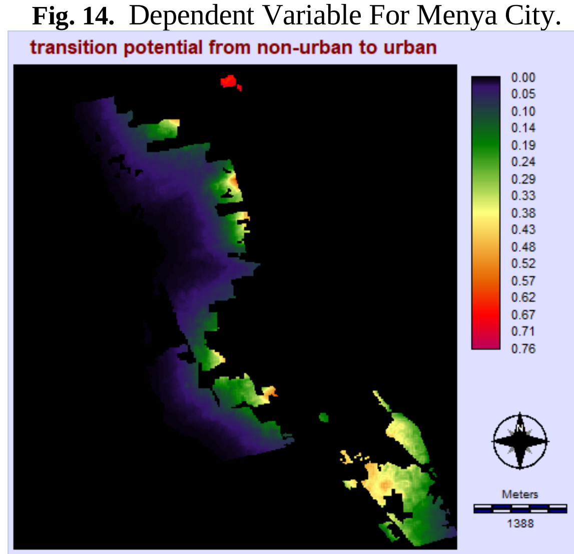

Fig. 15. Transition Potential From Non-Urban To Urban (Menya).

- Asuit

The CA-MC-LR model was applied to the city of Minya , and the model was calibrated and validated to ensure its accuracy. The odds ratios were calculated for all variables (Table 5) (Figure 16). The results showed that the ten factors affected urban growth to varying degrees, as shown in the Odds Ratio. The Odds Ratio for distance from health services is 0.1409, which means that the expected urban growth in an area close to health services is estimated at 7.01 times more than the expected urban growth in an area one kilometer away from educational services.

When Cramer's V was calculated for the variables, all values were found to be greater than 0.15, indicating a strong relationship with the urban growth process, except for slope which have 0.148 which can be excluded from the influencing factors (Table 5).

Table 5
Oddes Ratio, Odds Ratio and Coefficient For Dependent Variable (Asuit).

Variable	Odds Ratio	Cramer's V	Coefficient
V1	0.5610	0.1408	-
V2	0	0.3566	0.19583656
V3	0.1409	0.4429	-0.00195252
V4	0.0776	0.4555	-0.00034158
V5	0.2434	0.4571	-0.00024248
V6	0	0.3513	0.00164211
V7	0.4328	0.3225	0.00156671
V8	2.6234	0.5753	-0.00064962
V9	0	0.3728	-0.00148669
V10	0.1487	0.5256	0.00125805

After verifying the calibration process of the independent factors and ensuring their impact on the urban growth process, the transition potential from non-urban to urban area is produced (Figure 17) and the urban growth matrix for the period 2004-2013 is calculated to predict the 2020 map by applying the Markov chain model and specifying the prediction year within the LCM tool in the IDRISI SELVA program. The Kappa coefficient is calculated by comparing the actual map for 2020 and the predicted map from the (CA-MC-LR) model. By applying the VALIDATE tool, the Kappa coefficient was found to be 0.9434, indicating the validity of the model and its applicability.

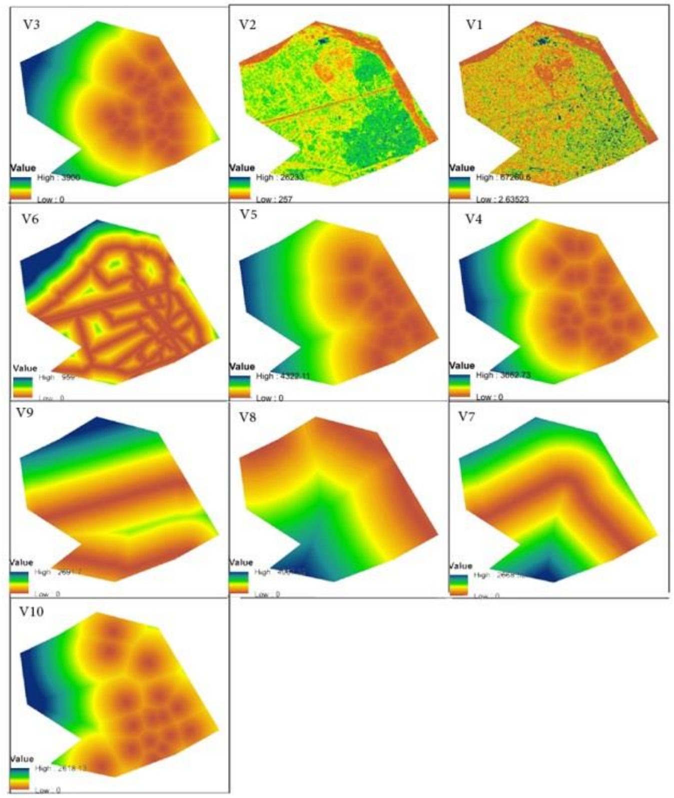

Fig. 16. Dependent Variable For Asuit City.

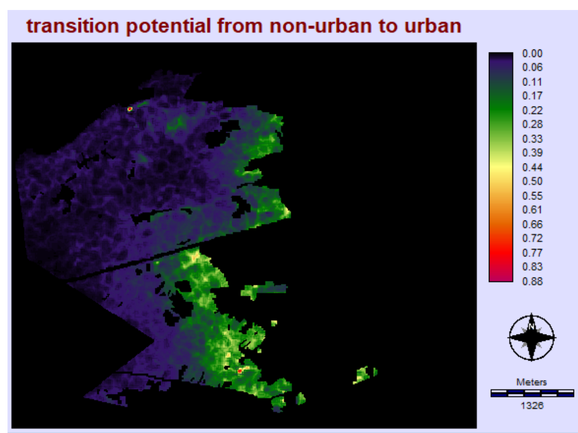

Fig. 17. Transition Potential From Non-Urban To Urban (Asuit).

7.1.Comparing results of two models

This part of the study aimed to compare the results obtained by applying the integrated model between the Markov chain model and the Cellular Automata model (CA-MC), and the integrated model between the Cellular Automata model, Markov chain model, and Logistic Regression (CA-MC-LR).

This study compared the accuracy of the two models for the three cities, as presented in (Table 6), and found that both models performed similarly, showing high accuracy in predicting and simulating urban growth. However, when comparing the inputs and outputs of each of the two models, they differ, and therefore what determines which growth model is preferable to use is the available data and the outputs that decision makers want to obtain. Therefore, if the aim of the study is to predict the future land cover only and the study does not take into account the factors affecting the urban growth process, and the land cover maps have more than two uses, it is better to apply the CA-MC model. However, if one of the study objectives is to know the factors affecting the urban growth process and the extent of the impact of each variable, it is better to use the CA-MC-LR model. In the Logistic Regression model (LR), up to 20 independent factors affecting the urban growth process can be studied.

Table 6
Comparison Of Accuracy Results For The (CA-MC) And (CA-MC-LR) Models.

Variable	Beni-suef	Menya	Asuit
CA-MC	0.9500	0.9667	0.9441
CA-MC-LR	0.9510	0.9559	0.9434

This study found that the maps generated by the CA-MC model exhibited circular borders inconsistent with real urban patterns, whereas those produced by the CA-MC-LR model were more realistic (Figure 18).

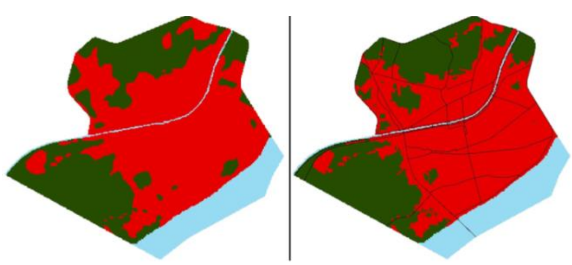

Fig. 18. Difference Between Land use Maps For The (CA-MC) and (CA-MC-LR) models.

7.2. Prsediction of urban growth

This study simulates and predicts urban growth for the years 2030, 2040, and 2050 for the selected cities. Based on previous results, the CA-MC-LR model was chosen for this purpose.

- **Beni-suef**

Preparing Data Required for the Logistic Regression Model (LR). First, Dependent Variable: Using the Land Change Modeler (LCM) tool within the IDRISI software, the land use maps for the years 2013 and 2020, extracted from satellite images, are input to analyze land cover changes and produce the dependent variable map (Figure 19).

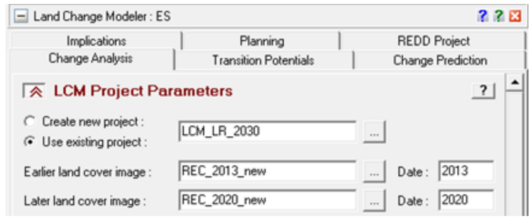

Fig. 19. Preparing Data Inside Land Change Modeler Tool (Beni-Suef).

By analyzing the land cover changes between 2013 and 2020, it was found that approximately 70 hectares of non-urban area were converted into urban area (Figure 20). Second, determining the independent variables; the same variables previously selected were input into the model.

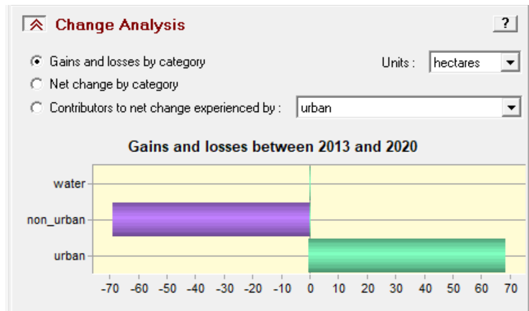

Fig. 20. Land cover changes Between (2013-2020).

Applying the Logistic Regression Model

By inputting the independent and dependent variables into the Logistic Regression model, the following is obtained:

1. Urban growth probability maps

In this phase, the probabilities of urban growth from non-urban to urban are determined by applying the "Potential Transition" tool within the Land Change Modeler (LCM) (Figure 21).

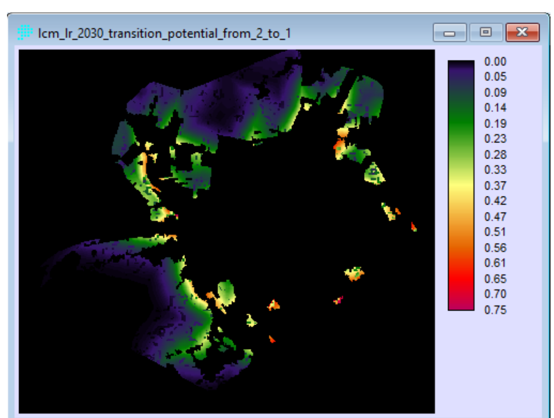

Fig. 21. Urban Growth Probability Map For Beni-Suef City (2013-2020).

Application of the Markov Chain Model

By applying the Markov Chain model to the 2013 and 2020 maps and inputting the prediction year, the urban growth probability matrix was derived for each year. Finally producing prediction maps (Figure 22).

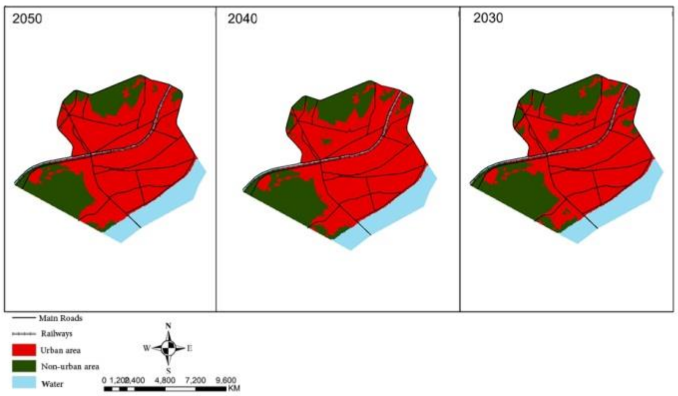

Fig. 22. Urban Growth Simulation For Beni-Suef For Years (2030-2040-2050).

The results of the urban growth simulation study for Beni-Suef City up to 2050 indicated the expansion of the urban area on agricultural land within the boundary from 849.21 hectares in 2020 to 934.90–1069.16 hectares in 2030 and 2050, respectively. the decrease in agricultural land within the boundary from 565.84 hectares in 2020 to 480.42–346.16 hectares in 2030 and 2050, respectively. This urban expansion on agricultural land will lead to negative social, economic, and environmental impacts as well as increased pressure on basic infrastructure (Figure 23).

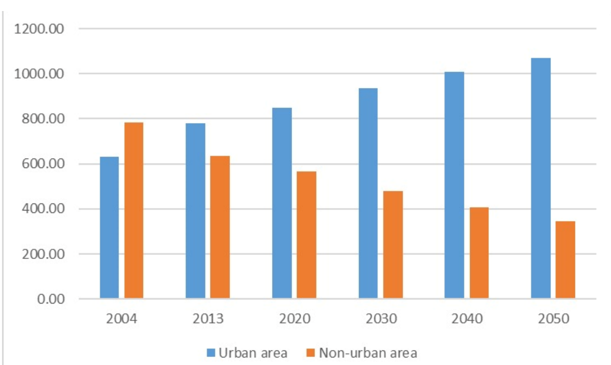

Fig. 23. Land Cover Change In Beni-Suef City (2004:2050).

- **Minya**

Preparing the Data Required for the Logistic Regression Model (LR). As the steps were applied to Beni-Suef City, the same methodology was followed for Menya City. First, for the dependent variable: the Land Change Modeler (LCM) tool was used along with the land use maps of 2013 and 2021 to analyze land cover changes and generate the dependent variable map (Figure 24).

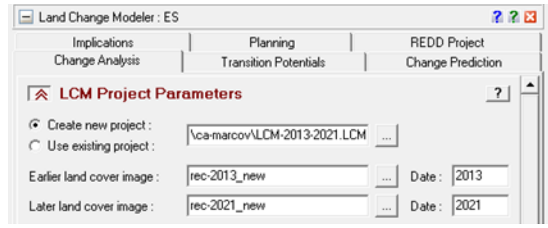

Fig. 24. Preparing data inside the land change model tool (Menya).

By analyzing the land cover changes between 2013 and 2021, it was found that approximately 70 ha of non-urban areas were converted into urban areas (Figure 25). Second, to determine the independent variables, the previously selected variables were input into the model.

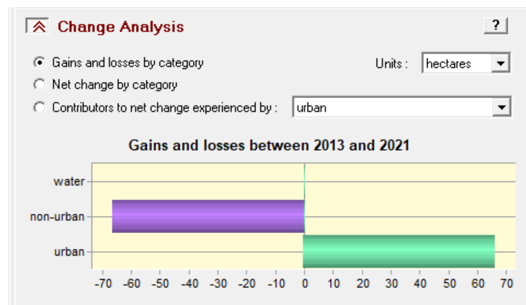

Fig. 25. Land Cover Change Between (2013-2021).

Applying the Logistic Regression Model:

By inputting the independent and dependent variables into the logistic regression model, the following is obtained:

1. Urban Growth Probability Map:

In this phase, the probabilities of urban growth from non-urban to urban are determined by applying the "Potential Transition" tool within the Land Change Modeler (LCM) (Figure 26).

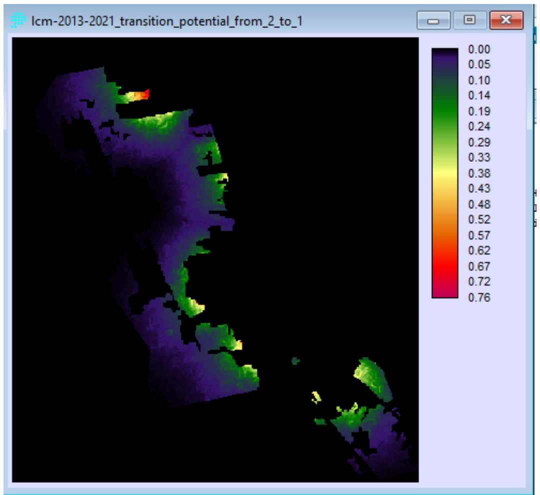

Fig. 26. Urban Growth Probability Map For Menya City (2013-2021).

Applying the Markov Chain Model:

By applying the Markov Chain model to the 2013 and 2021 maps and inputting the prediction year, the urban growth probability matrix was derived for each year. Finally, the prediction maps were produced (Figure 27).

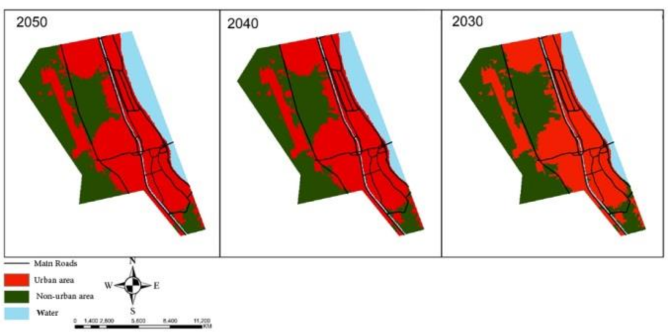

Fig. 27. Urban Growth Simulation For Menya For Years (2030-2040-2050).

The results of the urban growth simulation study for Menya City up to 2050 (Figure 28) indicate the expansion of the urban area on agricultural land within the boundary from 1075.57 hectares in 2021 to 1143.11–1283.74 hectares in 2030 and 2050, respectively. the decrease in agricultural land within the boundary from 977.18 hectares in 2021 to 909.68–769.05 hectares in 2030 and 2050, respectively. This urban expansion on agricultural land will lead to negative social, economic, and environmental impacts as well as increased pressure on basic infrastructure (Figure 28).

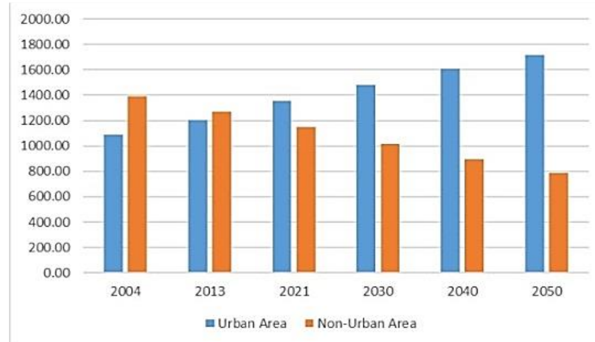

Fig. 28. Land Cover Change In Menya City (2004:2050).

- **Asuit**

Preparing the Data Required for the Logistic Regression Model (LR). As the steps were applied in Beni-Suef and Menya cities, the same methodology was followed in Assiut City. First, for the dependent variable: the Land Change Modeler (LCM) tool was used with land use maps from 2013 and 2021 to analyze land cover changes and generate the dependent variable map (Figure 29).

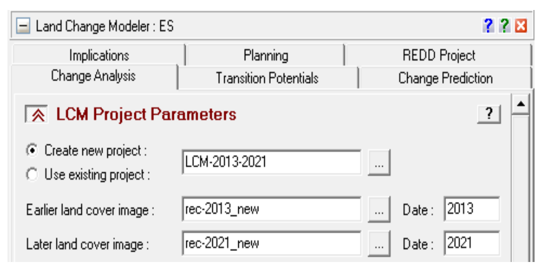

Fig. 29. Preparing data inside the land change model tool (Asuit).

By analyzing the land cover changes between 2013 and 2021, it was found that approximately 140 ha of non-urban areas were converted into urban areas (Figure 30). Second, determining the independent variables: The same variables previously selected were input into the model.

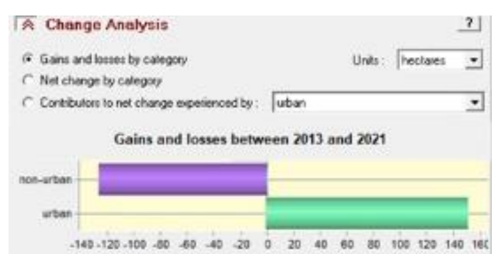

Fig. 30. Land Cover Change Between (2013-2021).

Applying the Logistic Regression Model:

By inputting the independent variables and the dependent variable into the logistic regression model, the following is produced:

Urban Growth Probability Map:

In this phase, the probabilities of urban growth from non-urban to urban are determined by applying the "Potential Transition" tool within the Land Change Modeler (LCM) (Figure 31).

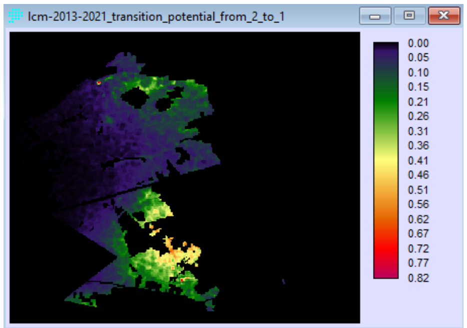

Fig. 31. Urban growth probability map for Asuit City (2013-2021).

Applying the Markov Chain Model:

By applying the Markov Chain model to the 2013 and 2021 maps and inputting the prediction year, the urban growth probability matrix is derived for each year. Finally, the prediction maps were produced (Figure 32).

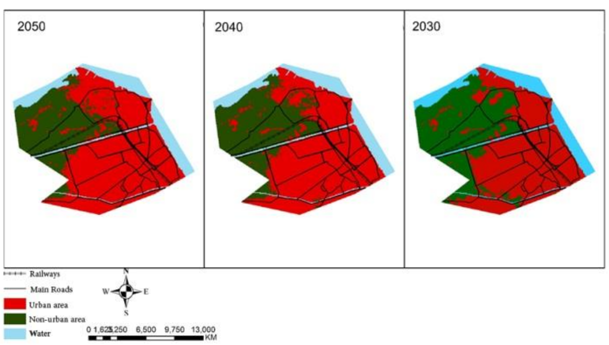

Fig. 32. Urban Growth Simulation For Asuit For Years (2030-2040-2050).

The results of the urban growth simulation study for Asuit City up to 2050 indicate: the expansion of the urban area on agricultural land within the boundary from 1356.23 hectares in 2021 to 1484.51–1718.55 hectares in 2030 and 2050, respectively. The decrease in agricultural land within the boundary from 1146.44 hectares in 2021 to 1017.95–783.90 hectares in 2030 and 2050, respectively. This urban expansion on agricultural land will lead to negative social, economic, and environmental impacts as well as increased pressure on basic infrastructure (Figure 33).

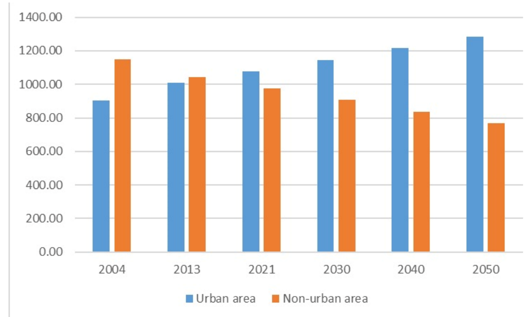

Fig. 33. Land Cover Change in Asuit City (2004:2050).

8. Results

This study applied both the CA-MC and CA-MC-LR models to simulate urban growth and land cover change across the study areas. The comparison of results revealed that both models demonstrated high accuracy in prediction and simulation. This was validated using IDRISI Selva through the CA-Markov and Land Change Modeler (LCM) tools across multiple study locations.

This study found that both models demonstrated high accuracy in predicting and simulating urban growth across the three cities. Accordingly, the selection of the preferable model depends on the data available and the outputs required by decision-makers.

When analyzing the factors influencing urban growth in each city, it was observed that the most influential factor in Beni-Suef City was the distance from railways; in Menya City, the distance from roads; and in Assiut City, the distance from services. However, when comparing these influential factors to the actual urban context of the study areas, some results appear to be inconsistent with reality. Therefore, caution is advised when interpreting the logistic regression outputs regarding influential factors. For instance, in Beni-Suef City, although the model identified distance from railways as the most significant factor, actual urban growth tends to occur near existing urban areas, suggesting that railways may not have a strong influence on urban expansion.

The results of the simulation of the urban growth forecast for Beni-Suef City until 2050 showed a decrease in the area of agricultural land from 565.84 hectares in 2020 to reach 346.16 hectares in 2050. Approximately 39% of the agricultural area has been converted into urban areas. In Menya city, the area of agricultural area decreased from 977.18 hectares in 2020 to reach 769.05 hectares in 2050. Approximately 21% of the agricultural area has been converted into urban areas. In Asuit City, the area of agricultural land has decreased from 1146.44 hectares in 2020 to reach 783.90 hectares in 2050. Approximately 32% of the agricultural area has been converted to urban areas.

9. Conclusion

The process of simulating and predicting urban growth is essential for planners in developing sustainable development strategies. In addition, it helps them predict the areas most likely to grow to provide adequate infrastructure for the population. Therefore, Strict policies must be established to limit urban growth on agricultural lands.

Governmental agencies must integrate new research methodologies and modern technologies into the process of urban modeling and simulation of urban growth and land cover changes, such as Cellular Automata models (CA), Geographic Information Systems (GIS), and Remote Sensing Data (RS) in state institutions such as the General Authority for Urban Planning. This study aims to further integrate these tools in urban planning research, understanding potential negative impacts, and the ability to avoid them.

Due to the lack of data and resources necessary for the process of simulating urban growth and land cover changes, integrating geographic information systems (GIS) and remote sensing systems (RS) is an indispensable mechanism in developing countries that do not have the necessary data in the planning process to achieve sustainability. Its importance may appear during the collection, classification, and analysis of satellite images with the help of some programs. In this study, the Envi 5.3 program was used in the processing and classification of satellite images, which were then converted to the ArcMap program to create a geographic database.

The Markov Chain (MC) model is not a spatial model for identifying changes; however, it is an effective model for predicting the amount of change that can be combined with other modeling methods to spatially determine changes. For example, in this study, the (MC) model was used to predict the expected quantity of change and was integrated with (CA) model to identify the changes spatially.

The (CA-MC-LR) model has proven its ability to study the factors influencing urban growth and their future impacts, allowing the analysis of up to 20 influencing factors. Understanding the effects of these factors on urban growth helps planners and decision makers comprehend the growth process and develop appropriate plans to achieve sustainable planning while avoiding anticipated negative impacts.

10. References

[1] U. N. N. York, "World Urbanization Prospects The 2014 Revision." 2015.

[2] E. F. Lambin *et al.*, "The causes of land-use and land-cover change: moving beyond the myths. GLOBAL ENVIRONMENTAL CHANGE-HUMAN AND POLICY DIMENSIONS," *Glob. Environ. Chang.*, vol. 11, no. 4, pp. 261–269, 2001.

[3] M. Jafari, H. Majedi, S. M. Monavari, A. A. Alesheikh, and M. K. Zarkesh, "Dynamic simulation of urban expansion based on cellular automata and logistic regression model: Case study of the Hyrcanian region of Iran," *Sustain.*, vol. 8, no. 8, 2016, doi: 10.3390/su8080810.

[4] O. Hamdy, S. Zhao, T. Osman, M. A. Salheen, and Y. Y. Eid, "Applying a hybrid model of markov chain and logistic regression to identify future urban sprawl in abouelreesh, aswan: A case study," *Geosci.*, vol. 6, no. 4, 2016, doi: 10.3390/geosciences6040043.

[5] A. Siddiqui, A. Siddiqui, S. Maithani, A. K. Jha, P. Kumar, and S. K. Srivastav, "Urban growth dynamics of an Indian metropolitan using CA Markov and Logistic Regression," *Egypt. J. Remote Sens. Sp. Sci.*, vol. 21, no. 3, pp. 229–236, 2018, doi: 10.1016/j.ejrs.2017.11.006.

[6] M. G. Munthali, S. Mustak, A. Adeola, J. Botai, S. K. Singh, and N. Davis, "Modelling land use and land cover dynamics of Dedza district of Malawi using hybrid Cellular Automata and Markov model," *Remote Sens. Appl. Soc. Environ.*, vol. 17, 2020, doi: 10.1016/j.rsase.2019.100276.

[7] A. Rienow, "The Future of Central European Cities – Optimization of a Cellular Automaton for the Spatially Explicit Prediction of Urban Sprawl," *Optim. Algorithms - Methods Appl.*, Sep. 2016, doi: 10.5772/62424.

[8] S. K. Singh, S. Mustak, P. K. Srivastava, S. Szabó, and T. Islam, "Predicting Spatial and Decadal LULC Changes Through Cellular Automata Markov Chain Models Using Earth Observation Datasets and Geo-information," *Environ. Process.*, vol. 2, no. 1, pp. 61–78, 2015, doi: 10.1007/s40710-015-0062-x.

[9] P. Subedi, K. Subedi, and B. Thapa, "Application of a Hybrid Cellular Automaton – Markov (CA-Markov) Model in Land-Use Change Prediction: A Case Study of Saddle Creek Drainage Basin, Florida," *Applied Ecology and Environmental Sciences*, vol. 1, no. 6. pp. 126–132, 2013, doi: 10.12691/aees-1-6-5.

[10] M. M. Aburas, Y. M. Ho, M. F. Ramli, and Z. H. Ash'aari, "Improving the capability of an integrated CA-Markov model to simulate spatio-temporal urban growth trends using an Analytical Hierarchy Process and Frequency Ratio," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 59, pp. 65–78, 2017, doi: 10.1016/j.jag.2017.03.006.

[11] U. N. N. York, "Department of Economic and Social Affairs, Population Division." 2015.

[12] I. M. B. Mahmoud, "Modelling and Simulation Of Informal Housing Growth," Cairo University, 2010.

[13] H. Sha and M. Helbich, "Spatiotemporal urbanization processes in the megacity of Mumbai , India : A Markov chains-

cellular automata urban growth model," vol. 40, pp. 140–149, 2013, doi: 10.1016/j.apgeog.2013.01.009.

[14] P. Ghosh *et al.*, "Application of Cellular automata and Markov-chain model in geospatial environmental modeling- A review," *Remote Sensing Applications: Society and Environment*, vol. 5. pp. 64–77, 2017, doi: 10.1016/j.rsase.2017.01.005.

[15] M. Salem, N. Tsurusaki, and P. Divigalpitiya, "Analyzing the driving factors causing urban expansion in the peri-urban areas using logistic regression: A case study of the greater Cairo region," *Infrastructures*, vol. 4, no. 1, 2019, doi: 10.3390/infrastructures4010004.

[16] A. S. · P. Chouhan, "Modeling spatial determinants of urban expansion of Siliguri a metropolitan city of India using logistic regression," 2020.

[17] X. Zhang, J. Zhou, and W. Song, "Simulating urban sprawl in china based on the artificial neural network-cellular Automata-Markov model," *Sustain.*. 12, no. 11, 2020, doi: 10.3390/su12114341.

[18] M. S. Mondal, N. Sharma, P. K. Garg, and M. Kappas, "Statistical independence test and validation of CA Markov land use land cover (LULC) prediction results," *Egypt. J. Remote Sens. Sp. Sci.*, vol. 19, no. 2, pp. 259–272, 2016, doi: 10.1016/j.ejrs.2016.08.001.

[19] C. C. and M. C. Chenglin Xie, Bo Huang, "Spatial Logistic Regression and GIS to Model Rural-Urban Land Conversion Chenglin," 2005.

[20] A. Sarkar and P. Chouhan, "Modeling spatial determinants of urban expansion of Siliguri a metropolitan city of India using logistic regression," *Model. Earth Syst. Environ.*, vol. 6, no. 4, pp. 2317–2331, 2020, doi: 10.1007/s40808-020-00815-9.

[21] URL: https://earthexplorer.usgs.gov/

[22] URL: https://www.diva-gis.org

[23] B. Rimal, L. Zhang, H. Keshtkar, B. N. Haack, S. Rijal, and P. Zhang, "Land use/land cover dynamics and modeling of urban land expansion by the integration of cellular automata and markov chain," *ISPRS Int. J. Geo-Information*, vol. 7, no. 4, 2018, doi: 10.3390/ijgi7040154.

[24] D. Abudu, R. A. Echima, and G. Andogah, "Spatial assessment of urban sprawl in Arua Municipality, Uganda," *Egypt. J. Remote Sens. Sp. Sci.*, vol. 22, no. 3, pp. 315–322, 2018, doi: 10.1016/j.ejrs.2018.01.008.