Heliyon 10 (2024) e25137

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Application of geospatial techniques and binary logistic regression model for analyzing driving factors of urban growth in Bahir Dar city, Ethiopia

Kenu Getu [a,*], H. Gangadhara Bhat [b]

[a] Department of Geography and Environmental Studies, Debre Tabor University, P.O.Box 272, Ethiopia
[b] Department of Marine Geology, Mangalore University, Mangalagangothri, Karnataka, India

ARTICLE INFO

Keywords:
Bahir dar
Urban growth
Binary logistic regression
Land use land cover
Driving factor

ABSTRACT

Understanding the drivers of urban growth and spatiotemporal land use change is important for rational land use and sustainable urban development. Based on the land use data, GIS data of explanatory variables, experts' knowledge and field observation, the study used a binary logistic regression model (BLRM) to analyze factors that drive rapid urban growth in Bahir Dar city, Ethiopia, using the LOGISTICREG module in IDRISI Selva software. Nine factors were used to reflect the influence of proximity and physical factors on urban growth from 1984 to 2019. This model helped in quantifying and identifying the factors of urban growth, which includes topography (slope, elevation and aspect) and accessibility (Dis. to the main road, Dis. to international airport, Dis. to CBD, Dis. to existing built-up area, Dis. to forest land and Dis. to water body). Furthermore, urban growth probability maps were created based on LRM results, revealing that the biggest urban growth would occur around existing built-up areas along the main roads and near Bahir Dar international airport. The Relative Operating Characteristic (ROC) values of 0.85, 0.90 and 0.93 and PCP values of 96.72 %, 98.46 % and 98.51 % indicate the urban growth probability maps are valid and BLRM had an ideal ability to predict urban growth. So, the study highlighted the relation between urban growth and its drivers in Bahir Dar, giving a decision making framework for better land use management and resource allocation.

1. Introduction

Urbanization is a process of socio-economic changes and a major indicator of regional development [1]. The cityscape has rapidly transformed in recent decades in response to fast population growth and transition from rural to urban areas ([2,3]). Unprecedented growth in the global urban population over the past few decades has rapidly expanded urban areas, particularly in developing Asian and African nations [4,5]. Urban growth has been accelerating because of rapid population growth and high demand of urban land for settlement and different industrial expansions [6,7]. The principal sources of change in natural systems on earth are human-induced ecological and urban landscape transformations, which have an impact on the biosphere's ability to sustain life [8–11]) stated that man depends on the environment for survival and the environment also depends on human beings for sustenance. Land use land cover change in urban areas has significantly increased in the past decades due to rapid urban growth [12,13]. Therefore, different land

* Corresponding author.
E-mail addresses: kenugetu@gmail.com (K. Getu), gangadharrbhat@gmail.com (H. Gangadhara Bhat).

https://doi.org/10.1016/j.heliyon.2024.e25137
Received 13 March 2023; Received in revised form 16 January 2024; Accepted 22 January 2024
Available online 26 January 2024
2405-8440/

cover classes such as forests, farms, mountain ranges and water bodies have changed into urban areas and infrastructures [14,15] (see Table 10).

Spatiotemporal analysis and monitoring of urban growth dynamic as well as its driving factors is very important to understand the process of urbanization and its corresponding environmental effects throughout the world [5,16]. Long term and consistent analysis and mapping of urban growth is also crucial to track, project and understand trends of urban growth and address problems affecting sustainable urban development [16,17]. Human induced land use land cover changes is also another crucial issue to determine and address the challenges of sustainable urban growth [17,18]. So, monitoring and evaluating urban growth and its driving factors is very crucial and area of research for different scholars [18–20].

The rapidly growing cities and towns are diminishing the major natural resources through overutilization, resulting in unplanned and unsustainable development situations [7,17]. Urban growth can be managed by planning the future scenarios for which LU/LC change dynamics is crucial to understand [21,22] So, to ensure sustainable development, cities and towns must have a plan and management which balances the unlimited demand of human beings and the limited resources provided by the natural environment through careful and wise utilization of resources and transfer it to the next generation [23]. To accommodate changes in planning, monitoring, managing and evaluating urban planning, localization of Sustainable Development Goals (SDGs) in towns and cities necessitates up-to-date spatial information [24]. Sustainable development must address the requirements of the current generation without jeopardizing the future [17]. Cities and towns in developing nations struggle to alleviate the problems related to access to up-to-date information, reflecting urban dynamics to improve and protect environmental conditions [25–27] So, Geographic Information Systems (GIS) and Remote Sensing techniques (RS) are useful tools for tracking and studying changes in urban land use and land cover through space and time [28,29].

Unplanned and unmanaged urban growth has been hampering the quality of growth in different regions of the world [20]. This type of growth, which lacks spatial plan, impacts the natural ecosystem, including wildlife, through uncontrolled sprawling of cities and towns [16,30,31]. Rapid urban growth diminishes the amount of agriculture, forest and water bodies [32]. It also affects human health caused by increasing trends of environmental pollution [33,34]. The logistic regression model (LRM) was chosen and applied for this study because the Relative Operating Characteristic (ROC) has an excellent potential to compare the actual and projected maps (Osman & Monir, 2018). ROC of 0 indicates no urban growth and 1 implies the presence of growth [3,19,35,36].

Understanding and simulating urban growth and land use changes are crucial topics for environmental management, planning, and civil engineering [37]. Models may be the most effective tools for comprehending changes in land use and land cover as well as for careful planning to improve the management of cities and towns [32,38]. Rapid developments in modeling have made it possible to simulate and plan unique land use and land cover changes [14,39,40]. To model, predict and simulate urban growth and LULC changes, many scholars have employed various empirical and theoretical modeling techniques [41–44]. The logistic regression model is one of these techniques for empirical simulation. The logistic regression model has been used for studding different research problems such as forest coverage [45–47], urban development [8,43] and land use land cover change [48–50], predicting future trends in land use change [44,51] and urban sprawl [52–54]. Literatures revealed that logistic regression presents a clear picture of the weight of independent variables and their respective functions, as well as a good grasp of urban growth and urban growth processes (Aithal et al., 2017; Ramachandra et al., 2013 [55,56]; LRM allows and considers all socio-economic, physical, and demographic factors involved in the land use transformation and urban growth process, which are not available in many models and techniques [57]. Furthermore, LRM takes into account the spatial effects, autocorrelation and heterogeneity [58]. LRM also demands cautions when dealing with spatial autocorrelations that are common in geographically referred data, as these autocorrelations may cause the LRM hypothesis to be violated [59,60]. The essential features of most urban growth models, including LRM, are the relationships between urban growth changes and its explanatory variables. However, understanding, monitoring and analyzing the interactions between factors of urban growth in LRM is a challenging and time consuming process. As a result, it is vital to address the misunderstandings and lack of knowledge regarding these driving factors. There are certain limitations in using the LRM that should be considered [61, 62].

Literatures indicate an increasing demand for modeling and quantifying the extent and location of urban growth over space and time. Because of its significant growth in recent decades, Bahir Dar was chosen as a case study. Furthermore, its urban growth has never been thoroughly studied. The city is considered multifunctional because it is the political, economic, and social hub of the Amhara regional state and Ethiopia. Despite Ethiopia's government urban plan, Bahir Dar has experienced rapid but chaotic urban planning in recent decades. Corruption, mismanagement, political upheaval, and economic concerns may impact urban planning, resulting in huge and rapid urban growth. Rapid and uncontrolled urban growth and the quick conversion of productive agricultural fields, forests, and environmentally reserved areas have brought socio-economic and environmental difficulties in Bahir Dar. No research has been conducted on factors of urban growth in Ethiopia as well as Bahir Dar city using binary logistic regression model (BLRM) until this study was carried out.

Furthermore, the growth of Bahir Dar city is very fast as compared to other cities with similar socio-economic and political situation in Ethiopia. Previous studies consider few proximity and physical factors such as distance from the main road and slope as the only factors affecting the growth of urban areas. These factors were even studied in other cities and towns but not in the study area. This study also tried to identify many physical and proximity factors affecting urban growth in the study area. The application of relative operative characteristic (ROC) method for validation of the model is also another new concept in the study areas as well as Ethiopia till now. In addition, the use of geospatial data and technologies to determine the driving factors of urban growth is not given attention by researchers in Ethiopia. Survey method of analyzing urban growth related data is common method of research. Furthermore, this study attempted to assess the long term spatiotemporal urban growth in Bahir Dar City and its explanatory variables for the last three and half decades using statistical methods. Most previous studies tried to monitor short-term spatiotemporal urban growth and land use change. So, the study tried to fill these gaps which were not practically alleviated by previous research because of different reasons. Previously, the urban growth pattern and driving factors for long term rapid urban growth of Bahir Dar City has never been analyzed. Application of this method in other study areas can not directly indicate the situations in Bahir Dar. Because the urban growth rate, growth pattern, study period as well as spatiotemporal distribution is not the same. So, BLRM is new for analysis of urban growth and land use land cover study and has good predictive ability. The PCP values were also computed to validate the BLRM in this study. This was also not commonly applied to evaluate the model in most previous studies. But it has good ability to evaluate the quality of the model. The study is, to our knowledge, the only study that uses both PCP and ROC at the same time to evaluate the quality of BLRM. Both of them are excellent methods to compare the actual maps with that of predicted as well as suitability maps.

This paper shows the use of LRM technique in simulating urban growth. This quantitative technique was used to monitor, quantify and analyze urban growth processes in Bahir Dar and its environs. Its goal is to improve our understanding of biophysical and socio-economic elements and their interactions and forecast future urban growth scenarios in Bahir Dar. Specifically, the objectives of this study are to detect and analyze the spatiotemporal changes in urban growth in Bahir Dar between 1984 and 2019, monitor the underlying causes that drive rapid urban growth and last but not least, to predict urban growth changes using Logistic Regression Model. The current study differs from the other studies because it uses nine variables that have a greater impact on urban growth. The study indicates that urban growth is geared towards agricultural lands, forests and wetlands. So this will help policy makers and planners make optimum land-use planning and management decisions.

2. Materials and methods

2.1. Study area

The study was carried out in Bahir Dar city and its environs, which is found in the northwest part of Ethiopia (Fig. 1). The city covers an area extending from 11°30′0″ to 11°41′ N longitude and 37°16′ to 37°30′0″ E latitude. The total area of Bahir Dar is around 213.55 km^2, which is relatively small in the country's total area. It is divided into nine sub-cities and 32 rural kebeles. Bahir Dar is located in the Northwest part of Ethiopia at road distance of 565 km far from the capital city of Ethiopia. The altitude of the city and its surroundings ranges from 1717 to 2008 m above mean sea level. Now it is the administrative seat of Amhara National Regional State.

The population of the city has been growing very fast. Its population was 54, 800 during the first census (1984) of Ethiopia and grew to 96, 140 in the second census (1994). This figure shows there was an increment of the population by 75 % within 10 years of

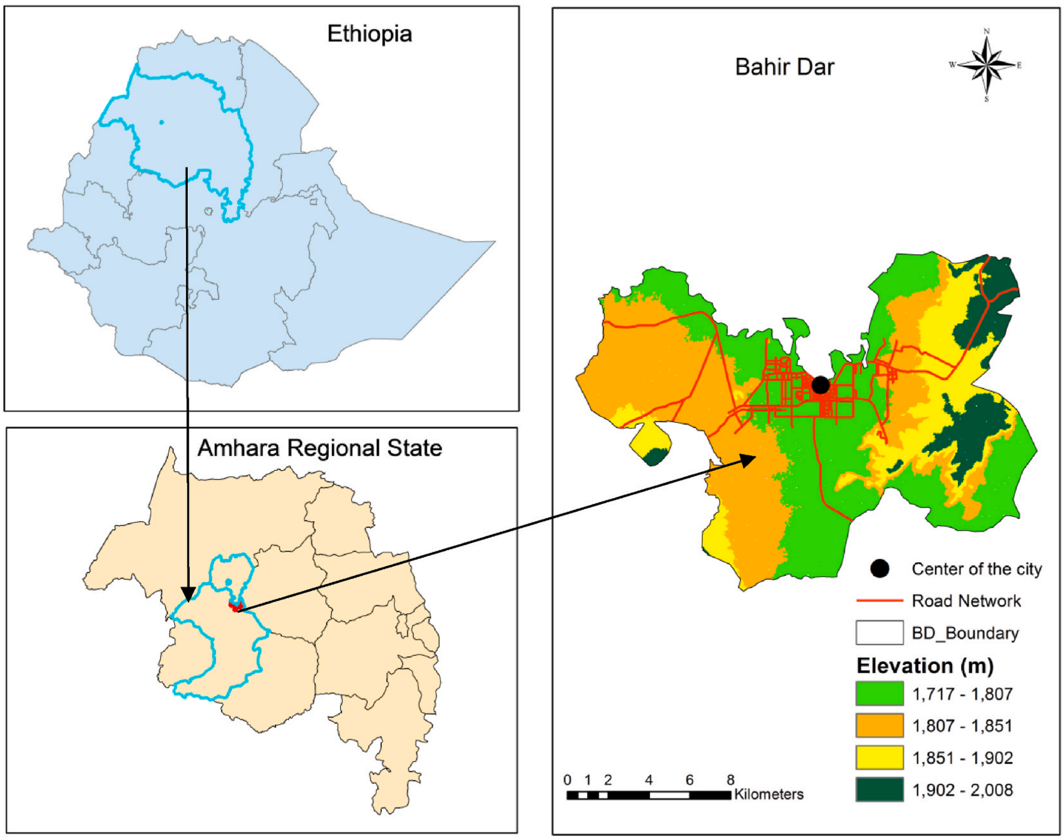

Fig. 1. Location map of the study area.

interval, which is one census period of the country. In the third census (2007), the population of Bahir Dar was increased to 221, 991, with an increase of about 87 % from 1994 census. Census has been conducted in every 10 years of interval in Ethiopia which is carried out by Ethiopian Central Statistics Agency (CSA).

However, due to different socio-economic and political problems, the current population of Bahir Dar is not exactly known, census has not been carried out recently. Based on the sample survey conducted on 2017, the population of Bahir Dar was 348, 429. Based on these figures, the population of Bahir Dar was projected for this study and the projected population for 2019 was 389, 177.

2.2. Data and methods

This paper was conducted based on the data obtained from various sources. Landsat images obtained from different sensors (TM & ETM+) acquired in 1984, 1994, 2009 and 2019 were freely downloaded from the USGS website (Table 1). Furthermore, this study used other geospatial data sources to get auxiliary data. In order to select the appropriate and quality Landsat satellite images, different factors such as type of sensors, resolution of images, costs, quality of images and availability of the required images should be considered. This study selected these dates by considering data availability, data quality and dry seasons. Landsat 7 ETM^+ has not been used for this study because images acquired by the sensor from 2002 to 2008 for this study area were distorted with high strips. So, the researcher has been forced to use Landsat TM 2009 instead of Landsat 7 ETM + for the study because of the lack of adequate data. The images were taken with every 10 years of interval (Table 1). However, the Landsat images from 2002 to 2008 were not good in quality and not convenient for urban growth study. So, the researcher has taken Landsat TM 2009 instead of 2004 due to the poor quality of images of the target year. This affects the fixed interval dates to be considederd between 1994 and 2009 in this research, unlike the other study years taken in a fixed interval of 10 years. So, the selection of Landsat satellite images dates was influenced by the quality of images in this study.

ArcGIS 10.3 has been used for the different steps of image processing, which enables the pre-processing and different spatio-temporal analysis required for this study. Moreover, the Landsat images were classified using Maximum Likelihood supervised classification techniques after pre-processing. These classified thematic maps have six classes: built–up, forest, water body, crop land, Wetland and open spaces. The accuracy of the maps was evaluated and validated using ground control points (GCPs), Google earth maps and high resolution images (Spot 1.5×1.5 m and 5×5 m). Each LULC map used in this study has an accuracy level greater than 80 %, indicating the high image classification accuracy. In addition, the LULc maps and other auxiliary data were geometrically corrected and referenced to WGS 1984, UTM zone. Following this, in ArcMap 10.3, thematic raster maps of the dependent ($Y = 1$) and independent (Xi) variables with a resolution of 30×30 were prepared. Finally, raster data (dependent and independent variables) was converted to ASCII format and imported to IDRISI Selva for further processing, running the LRM, model calibration and validation.

2.3. Logistic regression model

In contrast to knowledge based techniques, LRM is a data driven, empirical statistical model used to identify and determine the drivers of urban growth change. In developing countries like Ethiopia, the availability of well-organized and high quality data connected to the study influences the choice of explanatory variables. Most geospatial data utilized to determine the most significant variables and their correlation with urban growth were obtained from free and open sources.

Previously, numerous researches were carried out by applying spatial logistic regression models like linear regression and logistic regression to determine the relationships between urban growth and its drivers [63–66]. For instance, linear regression was utilized to predict LULC changes of urban fringes, and logistic regression was applied to explain the spatial patterns of land development [67,68]. But the most widely used statistics is BLRM, which enables exploring the association between binary dependent and multiple predictor variables [67,69–71].

A statistical model called BLR was applied in this project to determine the relations between urban expansions (built-up area growth) the predictor variables and to produce probability maps of urban growth. BLRM model is a commonly utilized statistical technique to predict the future changes in urban growth and trends of LULC changes [72] and is purposely used to determine the

Table 1
Data used and their sources.

Data	Data sources
Landsat TM (1984)	USGS earth explorer
Landsat TM (1994)	USGS earth explorer
Landsat TM (2009)	USGS earth explorer
Landsat ETM+ (2019)	USGS earth explorer
SPOT 6	EMA
SPOT 6/7	EMA
Administrative boundary map	EMA
SRTM Dem	USGS earth explorer
Road map	Open Street Map Downloaded
Active economic center (CBD)	Google earth
International Airport	Google earth
Ground control points	GPS

empirical relationships between binary dependent and numerous independent variables [73,74]. The dependent variable is a binary/dichotomous class variable with presence (1) and absence of urban growth (0), and the independent variables are categorical/continuous variables. The study has been conducted based on the hypothesis that the probability of a nonurban cell converting to an urban cell would follow the logistic regression curve. So, the common and widely used formula to compute logistic regression is given below.

$$Y = a + b_1x_1 + b_2x_2 + \ldots + b_mx_m \quad (1)$$

$$Y = log_e\left(P/1 - P\right) = log\ it\ (P) \quad (2)$$

$$P(Z = 1) = e^y/1 + e^y \quad (3)$$

Where $X1$, $X2$ ….Xm are predictor variables and Y is the linear combination of predictor variables indicating linear relationships. The parameters b1, b2 ….bm are the regression coefficients to be estimated in the analysis processes. The other parameter Z denotes the binary response of the variables (0 or 1) and P represents the probability of occurrences of a new unit, i.e., Z = 1.

For this study, urban growth in Bahir Dar was used as a dependent variable, where 1 represents urban growth and 0 for nonurban growth for the study years from 1984 to 1994, 1994–2009 and 2009–2019. So, the driving factors/explanatory variables were selected based on different procedures. First and foremost, the probable explanatory variables were identified based on literature review [11,43,55,61,75,76]. Then, a total of 13 explanatory variables were identified and categorized into three groups: environmental, socio-economic and proximate factors. The explanatory or predictor variables that have been used in this research were Dis. to international airport, Dis. to forestland, Dis. to water bodies, dis. to CBD, Dis. to existing built-up area, Dis. to the main road, slope, aspect and elevation. These variables were chosen depending on literature reviews, field observation and discussion with local experts.

Distance to Bahir Dar international airport is one of the driving factors of urban growth in Bahir Dar city. This means the physical distance that each cell has from the Bahir Dar international airport. Bahir city has an international airport which have been giving international flight services for long period of time.

Distance to forestland is the other factor incorporated in this study as one of the driving factor influencing urban growth in Bahir Dar. Forest is one of the land use land cover class in Bahir Dar city. Distance from forest land in this study indicates the distance that each cell in Bahir Dar has from forestlands.

Distance to water bodies is also another driving factor influencing the growth of Bahir Dar city. There are different water bodies in Bahir Datr city; Abbay River (Blue Nile, longest river in the world) and Tana Lake (shallowest and Wider Lake in Africa). This water bodies are influencing the growth of Bahir Dar city since its establishment till now. So, distance from water bodies in this study refers to the distance each cell in Bahir Dar has from these water bodies.

Distance to central business district (CBD) is also used as factor of urban growth in Bahir Dar city. CBD shows parts of Bahir Dar city which contain the main commercial streets and personal buildings in the center of the city. Distance from CBD in this study indicates the distance each pixel in the raster map has from the center or commercial business district of Bahir Dar city.

Distance to existing built-up areas is the other proximity factor influencing the growth of Bahir Dar city. Built-up areas include settlements, industrial parks, pavements, roads and other artificial surfaces in Bahir Dar city. So, distance from existing built-up areas indicates each pixel's distance from these built-up areas or impervious surfaces.

Distance to main roads is also investigated as one of the driving factors influencing Bahir Dar city's growth. There are major roads that provide transport services for urban dwellers and also pass through the city. The major roads that pass through the city are the Addis Ababa – Gonder and Addis Ababa Dessie road. These roads provide main transport services to the city, regional state and the country. So, distance to main road in this study indicates each pixel's distance from these major roads.

Slope and aspect are the other topographic factors influencing the growth of Bahir Dar city. Slope refers to the rate of elevation for each digital elevation model (DEM) pixel, measured in degrees. It represents the steepness of the surface in Bahir Dar city. Aspect

Table 2
Variables of logistic regression model.

Variables	Description of variables	Nature of data	Shape
Dependent (Y)	Urban growth 0 = no urban growth 1 = urban growth	Dichotomous	Polygon
Independent			
Dis_Airport	Distance to international airport	Continuous	Point
Dis_Forest	Distance to forestland	Continuous	Line
Dis_Water	Distance to water bodies	Continuous	Line
Dis_CBD	Distance to CBD	Continuous	Point
Dis_Urban	Distance to existing built -up	Continuous	Point
Dis_Mird	Distance to main road	Continuous	Line
Slope	Slope	Continuous	Line
Aspect	Aspect	Continuous	Line
Elevation	Elevation	Continuous	Line

indicates the direction of slope. It identifies the down-slope direction of the maximum rate of change in value from each pixel to its neighbors.

Accordingly, distance from Wetland, distance from satellite town, soil type and population density were excluded from the list because of lack of well–organized data and insignificant roles of the variables for urban growth. The probability of land becoming urban is generally taken as a function of the determined predictor variables. So, nine driving factors were taken for analysis and prediction purposes for the proposed model (Table 2). All the data sets procedures and analysis methods used in this research has been indicated in the figure bellow (Fig. 2).

2.4. Variables used in the model

2.4.1. Dependent variable

This study used four binary land cover maps of 1984, 1994, 2009 and 2019 as dependent variables to run LRM. The LULC maps were classified into two land cover classes 0 and 1. In these reclassified maps, 0 represents urban class and 1 denotes nonurban land use classes. The conversion rule assumes that land cover only changes from non –urban to urban land use class. This is because the probability of land cover converting from urban to non – urban is very unseal and rare in developing nations like Ethiopia. Urban growth in this research has been detected using binary logistic regression method, which uses the net urban growth from 1984 to 1994, 1994–2009 and 2009–2019 as dependent variables and explanatory variables as predictor variables. The dependent variable was produced by using image calculator tool in IDRISI Selva software.

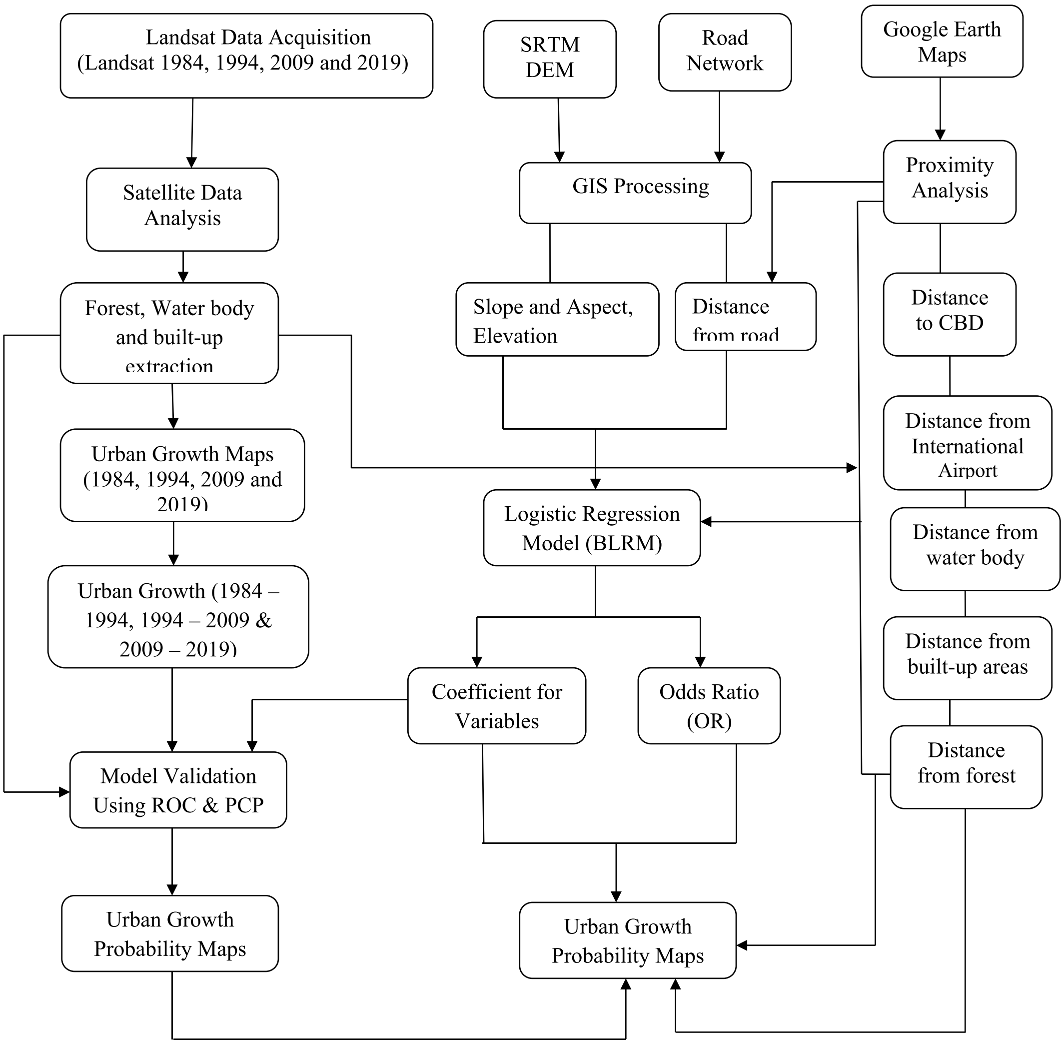

Fig. 2. Methodological flowchart of the study.

2.4.2. Explanatory variables of urban growth

Finally, based on the literatures reviewed, field observation and discussion with local experts, nine predictor variables were selected to run the LRM. It is obvious that most LU/LC models are data driven. The variables included in the model are Dis. to international airport, Dis. to forestland, Dis. to water bodies, Dis. to CBD, Dis. to existing built-up area, Dis. to main road, slope, aspect and aspect elevation. These variables were produced in ArcMap 10.3 and had the same grid cell size of 30 × 30 m. Slope, elevation and aspect maps were generated from STRM DEM of the study area. Distance maps were computed in ArcMap using the Euclidean distance function tool. The shortest Dis. between the center of the source cell and the center of each surrounding cell was calculated using this Dis. function tool of ArcMap 10.3. The values of all these explanatory variables were standardized into the range from 0 to 1.

The aim of LRM is to better understand and quantify the interplay between urban growth and its factors. During the modeling processes, urban growth is treated as binary/dichotomous dependent variable that must be explained in terms of the nine independent variables involved in the model. The cells in independent variables have only two options, conversion to urban growth cell (1) or stay as no urban growth (0). Based on the assumptions, the model allowed only the conversions from nonurban to urban (0–1). But the explanatory variables involved in LRM are continuous.

The LRM is undertaken for the three time periods: 1984–1994, 1994–2009 and 2009–2019. In each period, the number of predictor variables used is the same. Table 2 contains the list of all variables included in the logistic regression and raster maps of all predictor variables (Fig. 3).

2.5. Multicollinearity analysis of explanatory variables

Multicollinearity refers to a scenario in which two or more variables are more strongly linked than dependent variables [77–79]. Multicollinearity in the model might result in a very high standard error and low t-statistics, unexpected changes in the coefficient magnitudes/signs and insignificant coefficients despite a high R^2 values. When the output is interpreted incorrectly, such issues can lead to erroneous conclusions about the relationships between the explanatory and dependent variables. So, computing multicollinearity analysis is a pre-condition for conducting logistic regression [80].

In logistic regression, the presence or absence of multicollinearity is not the major concern of researchers; rather the degree of presence is important. The more multicollinearity there is, the greater will be the likelihood that multicollinearity cause problems.

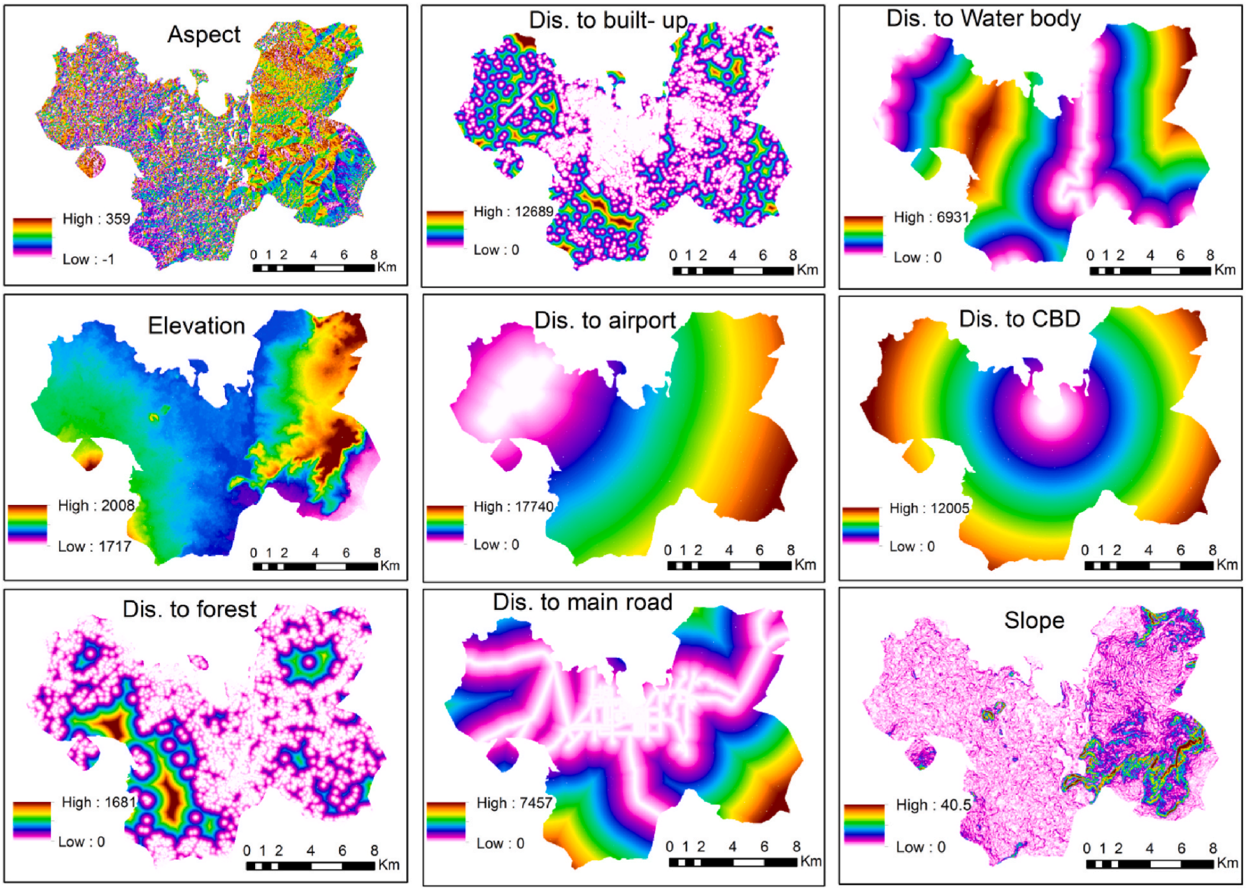

Fig. 3. Raster layers of independent variables of logistic regression.

Different methods of evaluating the presence and degree of presence of multicollinearity in spatial logistic models are available in literatures. In statistical analysis, VIF/variance inflation factor/is the most extensively used test for determining multicollinearity. This statistical test measures how far the variance of the estimated regression coefficient are inflated. VIF for K_{th} predictor can be computed:

$$VIF_{k=}(1-R_k^2)^{-1} \ldots \tag{4}$$

Where R^2 is the R^2 value obtained by regressing the K_{th} predictor on the remaining predictors. According to most previous research and reference books/statistical manuals, VIF>10 indicates the presence of severe or serious multicollinearity among independent variables (Marcoulides & Raykov, 2019; O'Brien, 2007). Statisticians recommended several approaches to alleviate the multicollinearity problem. The simplest and usual method is to remove some of the variables that are highly associated with each other in the model because the variables that are highly correlated provide the same information about the study.

2.6. Validation and calibration of LRM

The Relative Operative Characteristic (ROC) was used in this instigation to test the validation of LRM and evacuate probability maps that resulted from it. The ROC has been used to assess and validate the proposed model in urban growth research and is a reliable tool [3,67,81].

The ROC tool evaluates/tests the relationships between the actual and the predicted changes. It computes the percentages of false positives, true-positives for a range of thresholds and relates them to each other [82]. The area under the curve (AUC), which can range from 0.5 to 1, is calculated using the ROC technique. The value of 0.5 shows that the probabilities are assigned at random, implying that the expected agreement is due to change. On the other hand, a value of 1 implies the presence of a perfect probability assignment, which means that the actual urban growth and the predicted urban probability maps have a perfect spatial agreement [82].

LRM validation was carried out by comparing the probability maps of the future urban growth generated by the logistic regression

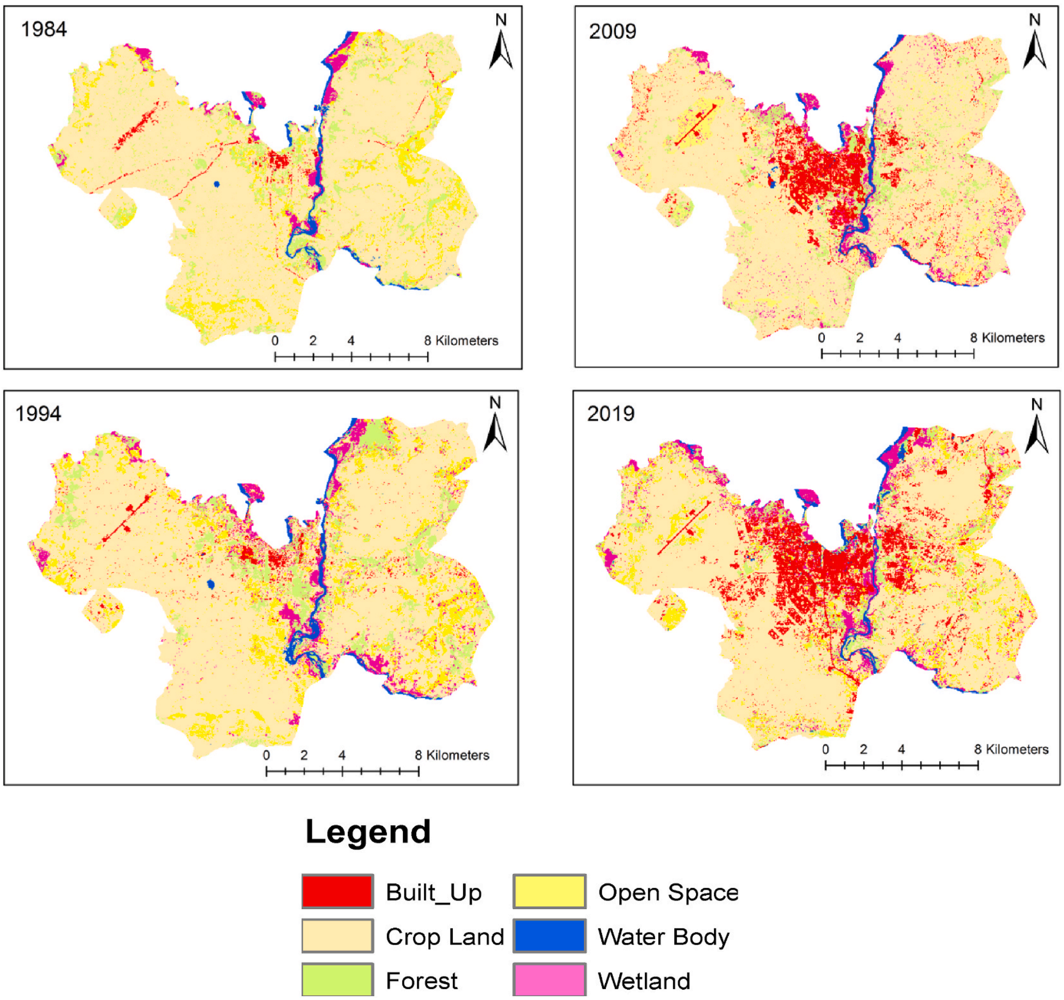

Fig. 4. LULC change maps of Bahir Dar in 1984, 1994, 2009 and 2019.

model with the actual urban growth from 1984 -1994, 1994–2009 and 2009–2019.

3. Result analysis

3.1. Analysis of LULC changes

The LULC maps and changes in the area are depicted in Fig. 4 and Table 3, respectively. Table 3 shows that crop land was the most prominent land cover type in 1984, with an area of 16568.10 ha representing 77.66 % of the total area. The built-up area class represents only 243.94 ha (1.1 %) in the same period. However, between 1994 and 2009, the urban land cover (Built-up) expanded dramatically, reaching 2529.98 ha of total area. Furthermore, forest cover decreased from 2163.95 ha in 1984 to 1589.39 ha 1994 to 2019 due to rapid land use change and urban growth. Fig. 3 depicts the evolving pattern of LULC changes in the research area from 1984 to 2019.

3.2. Multicollinearity analysis

Multicollinearity diagnosis was performed for each set of variables utilized in different time periods to remove redundant variables from the model and ensure coefficient stability. This analysis was carried out in SPSS 20 statistical software by iteratively regressing one of the independent variables against the remaining eight variables. So, variables that have variance inflation factor greater than 10 (VIF>10) were excluded from the model. The analysis indicated that the variables used in this model have a score of VIF<10. So, the proposed independent variables have been measured the different aspects of the dependent variable being modeled. These analyses' detailed descriptions have been elaborated and shown in the table below (Table 4).

3.3. Logistic regression analysis

Before applying the BLRM model, the urban growth maps for 3 periods (1984–1994, 1994–2009 & 2009–2019) have been produced (Fig. 5). This enables to properly detect the changes that have been occurred in the dependent variable for each period. As indicated in the figure (Fig. 5), there was an increased urban growth throughout the study period (1984–2019) in Bahir Dar city.

3.3.1. Urban growth model 1984–1994

The result explained the linkage between urban growth and its determinants in Bahir Dar and its environs. It can be interpreted that pixels with a higher coefficient near any independent variable have a larger development likelihood for change. As indicated below (Tables 5–7), the signs given for each coefficient indicate the existing positive/negative relation to the reaction of the dependent variables. For example, the negative correlation coefficient (−) refers to if Dis. increased, urban growth decreased and vice versa.

Among the explanatory variables, distance from water body, slope and distance from existing built-up area are positively correlated with urban growth. The other six variables are negatively correlated with urban growth (Table 5). The significance of explanatory variables can be measured using odds ratio. So, odds ratio (OR) is a crucial metric for interpreting logistic regression model. Explanatory factors with a higher odds ratio (>1) in the LRM are considered as strongly impacted variables. In this study, all the independent variables have OR>1 except elevation, which reveals that these predictor variables play a significant role in urban growth processes of the specified period.

3.3.2. Urban growth model 1994–2009

As indicated from the table, Dis. to main road, aspect and Dis. to exiting built-up land was negatively correlated with the dependent variable. The remaining variables such as Dis. to international airport, Dis. to forest land, Dis. to water body, Dis. to city center (CBD), slope and elevation are observed positively correlated with urban growth in this period of urban expansion (Table 6). Variables that have odds ratio>1 are significantly fitted in the model and considered influential in urban growth. All the variables, except distance to water body, have odds ratio (OR >1). This indicates the presence of these variables involved in this logistic regression model in high probability of urban growth.

Table 3
Land use land cover change over the study period.

Land Use Types	1984		1994		2009		2019	
	area (ha)	%	area (ha)	%	area (ha)	%	area (ha)	%
Built - up	243.94	1.14	357.56	1.68	1768.09	8.29	2529.98	11.86
Crop Land	16568.10	77.66	15319.84	71.82	14947.79	70.07	14183.18	66.49
Forest	2163.95	10.14	1993.85	9.35	1998.68	9.37	1589.39	7.45
Open Space	1478.74	6.93	2196.13	10.29	1371.67	6.43	1428.98	6.70
Water Body	356.98	1.67	392.66	1.84	332.22	1.56	361.05	1.69
Wetland	523.13	2.45	1071.93	5.02	913.68	4.28	1239.12	5.81

Table 4
Results of multicollinearity analysis.

Variables	Description of variables	VIF		
		(1984–1994)	(1994–2009)	(2009–2019)
Dis_Airport	Distance to international airport	5.456	5.172	4.985
Dis_Forest	Distance to forestland	6.543	6.983	6.012
Dis_Water	Distance to water bodies	4.373	4.015	3.987
Dis_CBD	Distance to CBD	3.398	3.837	3.509
Dis_Urban	Distance to existing built -up	2.357	2.059	2.067
Dis_Mird	Distance to main road	3.362	3.879	3.994
Slope	Slope (%)	4.587	4.901	4.342
Aspect	Aspect	4.593	4.893	4.572
Elevation	Elevation (m)	2.758	2.391	2.039

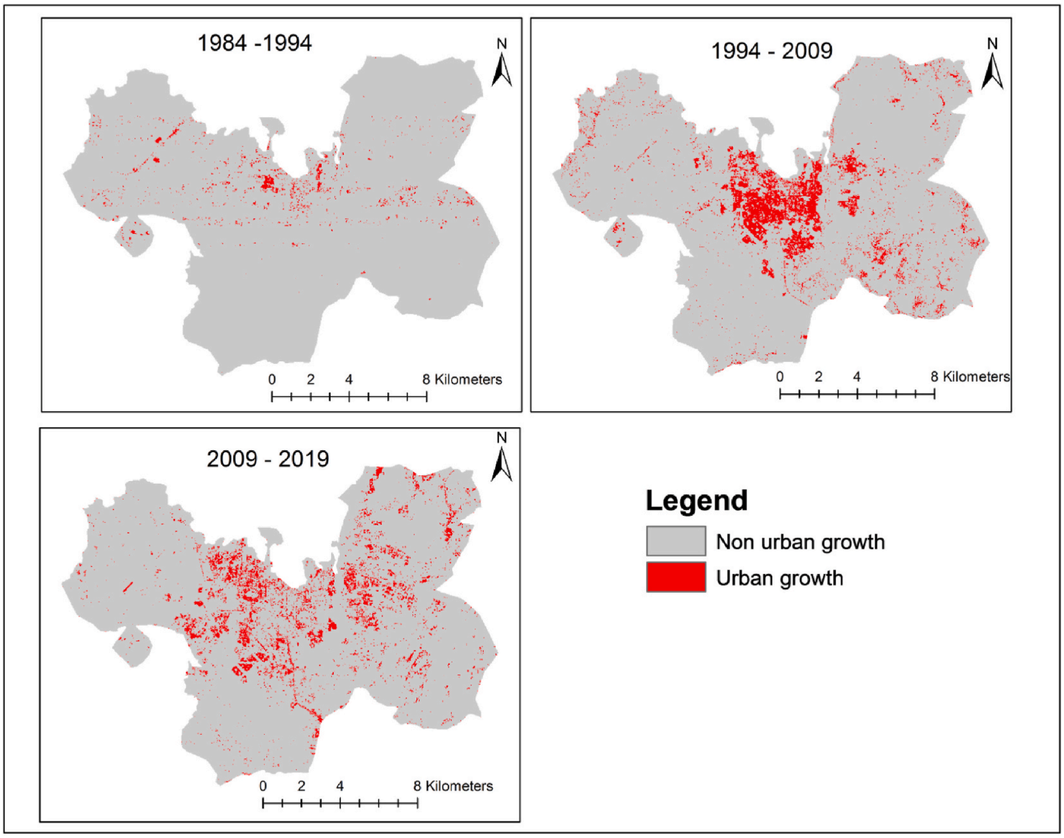

Fig. 5. Dependent variable: urban growth from 1984 to 2019.

Table 5
Estimated coefficient and odds ratio for logistic regression model of 1984–1994.

Variables	Description of variables	Coefficient (β)	Odds ratio	Standard Error
*X*1	Distance to international airport	−6.5362	3.1367	0.090787
*X*2	Distance to forestland	−4.8786	14.3798	0.090663
*X*3	Distance to water bodies	2.530	3.5404	0.090604
*X*4	Distance to CBD	−8.6190	14.3798	0.090784
*X*5	Distance to existing built -up	2.5687	1.0914	0.090920
*X*6	Distance to main road	−1.451	1.4348	0.090733
*X*7	Slope	9.1673	1.3818	0.090626
*X*8	Aspect	−6.650	4.0262	0.090629
*X*9	Elevation	−5.8980	0.7822	0.090627

Table 6
Estimated coefficient and odds ratio for logistic regression model of 1994–2009.

Variables	Description of variables	Coefficient (β)	Odds ratio	Standard Error
*X*1	Distance to international airport	2.871	25.7326	0.205305
*X*2	Distance to forestland	6.765	515.3326	0.204675
*X*3	Distance to water bodies	1.901	0.7571	0.205690
*X*4	Distance to CBD	4.988	109.2833	0.206850
*X*5	Distance to existing built -up	−2.024	1.0805	0.08378
*X*6	Distance to main road	2.463	1.1893	0.205690
*X*7	Slope	3.162	1.7377	0.204523
*X*8	Aspect	−2.884	1.3415,	0.204541
*X*9	Elevation	2.625	1.2801	0.204552

Table 7
Estimated coefficient and odds ratio for logistic regression model of 2009–2019.

Variables	Description of variables	Coefficient (β)	Odds ratio	Standard Error
*X*1	Distance to international airport	5.547	30.9772	0.191265
*X*2	Distance to forestland	4.715	1.2343	0.192058
*X*3	Distance to water bodies	2.599	3.2949	0.191742
*X*4	Distance to CBD	9.879	320.8942	0.192025
*X*5	Distance to existing built -up	−1.896	1.2571	0.191528
*X*6	Distance to main road	1.547	1.3377	0.195113
*X*7	Slope	−1.586	1.8277	0.194970
*X*8	Aspect	−2.089	2.2061	0.194119
*X*9	Elevation	−1.201	1.3360	0.195101

3.3.3. Urban growth model 2009–2019

From 2009 to 2019, the results of LRM revealed that all the predictor variables included in the model were able to explain the dependent variables at varying levels of significance. Here, slope, aspect, elevation and distance to existing built-up area are negatively correlated with urban growth (Table 7). However, the remaining variables: distance to airport, distance to forest, distance to water body, distance to CBD, and distance to main road are positively correlated with urban growth in the specified period. The odds ratio of all variables of this period are greater than 1 (OR>1), showing that these predictor factors had a very high influence on the probability of urban growth during this period.

3.4. Comparative analysis of the driving factors of urban growth

The table below (Table 8) indicates that all the variables except elevation incorporated in this investigation had high odd ratio (OR>1) from 1984 to 1994. This figure implies that all the variables in this study had great influence on the probability of urban growth in Bahir Bad city. However elevation had odd ratio below 1 (OR<1) which indicates its impact on urban growth not that much significant. Even though, it's value (OR) is less than 1, it is nearly approaching to 1 i. e 0.78. But scientific studies indicates that predictable variables that have OR>1 has no significant influence on the probability of urban growth. Based on the OR values, distance to forestland (OR = 14.3818), distance to CBD (OR = 14.3798), aspect (OR = 4.0262), distance to water bodies (OR = 3.5404) and distance to Bahir Dar international airport (3.1367) had OR>3. This indicates that the above listed variables had very high impact on the probability of urban growth from 1984 to 1994 in Bahir Dar city. But this does not mean the other explanatory variables had less impact on the urban growth in the same time period. All the 8 variables except elevation had significant impact on urban growth from 1984 to 1994 in the study area (see Table 9).

From 1994 to 2009, all the explanatory variables incorporated in this model had significant impact on the growth of Bahir Dar city

Table 8
Comparative analysis on temporal changes of factors of urban growth in Bahir Dar from 1984 to 1994, 1994–2009 and 2009–2019.

Variables	Description of variables	Odds ratio		
		1984–1994	1994–2009	2009–2019
*X*1	Distance to international airport	3.1367	25.7326	30.9772
*X*2	Distance to forestland	14.3798	515.3326	1.2343
*X*3	Distance to water bodies	3.5404	0.7571	3.2949
*X*4	Distance to CBD	14.3798	109.2833	320.8942
*X*5	Distance to existing built -up	1.0914	1.0805	1.2571
*X*6	Distance to main road	1.4348	1.1893	1.3377
*X*7	Slope	1.3818	1.7377	1.8277
*X*8	Aspect	4.0262	1.3415,	2.2061
*X*9	Elevation	0.7822	1.2801	1.3360

Table 9
ROC and PCP value for BLRM.

Binary Variable	Period	ROC	Accuracy	PCP
Urban growth (1) & Non urban growth (0)	1984–1994	0.8558	85.58 %	96.72 %
	1994–2009	0.9082	90.82 %	98.46 %
	2009–2019	0.9391	93.91 %	98.51 %

Table 10
2 × 2 contingency table showing the number of gride cells in actual vs. predicted map.

		Actual Map		Total
		Urban growth (1)	No urban growth (0)	
Predicted probability map	Urban expansion (1)	A	B	A + B
	No urban expansion (0)	C	D	C + D
Total		A + C = 3, 501	B + D = 416, 477	A + B + C + D = 419, 978

except distance to water body. All factors listed below in the table had high odds ratio (OR>1) except distance to water bodies. This implies that all 8 variables incorporated in this model had impacts on the probability of urban growth in Bahir Dar city. However, distance to water body had odds ratio less that 1 (OR>1), indicating that this variable had no significant influence on urban growth from 1994 to 2009 in Bahir Dar city. It had odds ratio (OR = 0.7571) i. e below the expected threshold (OR<1). As per odds ratio values obtained from logistic regression analysis, distance to forestland (OR = 515.3326), distance to CBD (OR = 109.2833), distance to international airport (OR = 25.7326) had OR>3. This values implies that had very high impact on the probability of urban growth in Bahir Dar from 1994 to 2009. The other remaining explanatory variables had also OR>1 except distance to water bodies. So, these factors had also significant impact on the probability of urban growth in Bahir Dar in the same study area.

In the third study period (2009–2029), all the nine explanatory variables had significant impact on the probability of urban growth in Bahir Dar. The variables which had high odds ratio were distance to CBD (OR = 320.8942), distance to international airport (OR = 30.97772), distance to water body (3.2949) and aspect (OR = 2.2061). The other 5 explanatory variables had also OR>1, implies that urban growth in Bahir was influenced by all the variables incorporated in the logistic regression model of this study from 2009 to 2019. Even though, the 4 explanatory variables (distance to CBD, distance to water body, aspect and distance to Bahir Dar international airport) had very high impact the probability of urban growth in Bahir Dar from 2009 to 2019, the remaining 5 variables (distance to forestland, distance to existing built-up area, slope, distance to main road and elevation) had also significant impact on the growth of urban growth.

There was changes that has been taken place on the driving factors of urban growth in Bahir Dar over time. This would be related with the frequent change in the land use land cover of the city, land use policy of the country, socio – economic status of the city and urban dwellers and difference in spatiotemporal distribution of environmental factors.

3.5. Probability of urban growth

The coefficient LRM encompassing the nine predictor variables was employed to predict urban growth probability of Bahir Dar in the three specified period and the following general formula was applied:

$$P = \left(Y = 1\ X_{1,}X_2, \ldots\ldots..X_k\right) = \frac{1}{1 + e^{-\left(\propto + \sum_{i=1}^{K} \beta i X i\right)}} \tag{5}$$

Where P(Y = 1 X_1, X_2, … ….X_k) is the probability of dependent variable Y being 1 given (X_1, X_2, ….X_k) is the probability of a cell being urbanized; is an independent variable representing a driving forces of urban growth; and βi the coefficient for variable Xi.

The urban growth probability maps of Bahir Dar indicated in Fig. 6 (A-C) are 0–0.167 (1984–1994), 0–0.555 (1994–2009) and 0–0.982 (2009–2019) classified colors of urban growth probability values. These maps identified and showed the future patterns of urban growth, which is likely to occur around the existing built-up area, near roadways and international airport, particularly in the central, west and east directions.

3.6. Statistical validation of logistic regression model

The ROC technique is excellent for validating the LRM that predicts the occurrence of a class by comparing the actual image with the predicted probability maps. It computes the percentage of false positive and true positive for a set of thresholds. The ROC technique calculates the area under the curve, which fluctuates between 0.5 and 1. A value of 0.5 shows that the probabilities are assigned at random, that predicting agreement is due to chance, and that a value of 1 implies presence of perfect probability assignment of perfect fit. The LRM validation of this study has been carried out through comparing simulated urban growth and actual urban growth maps from 1984 to 1994, 1994–2009 and 2009–2029 using stratified sampling of 5000 cells in the two maps. The ROC values of the BLRM were 0.85, 0.90 and 0.93 for the urban growth periods of 1984–194, 1994–2009 and 2009–2019, respectively. Based on the results,

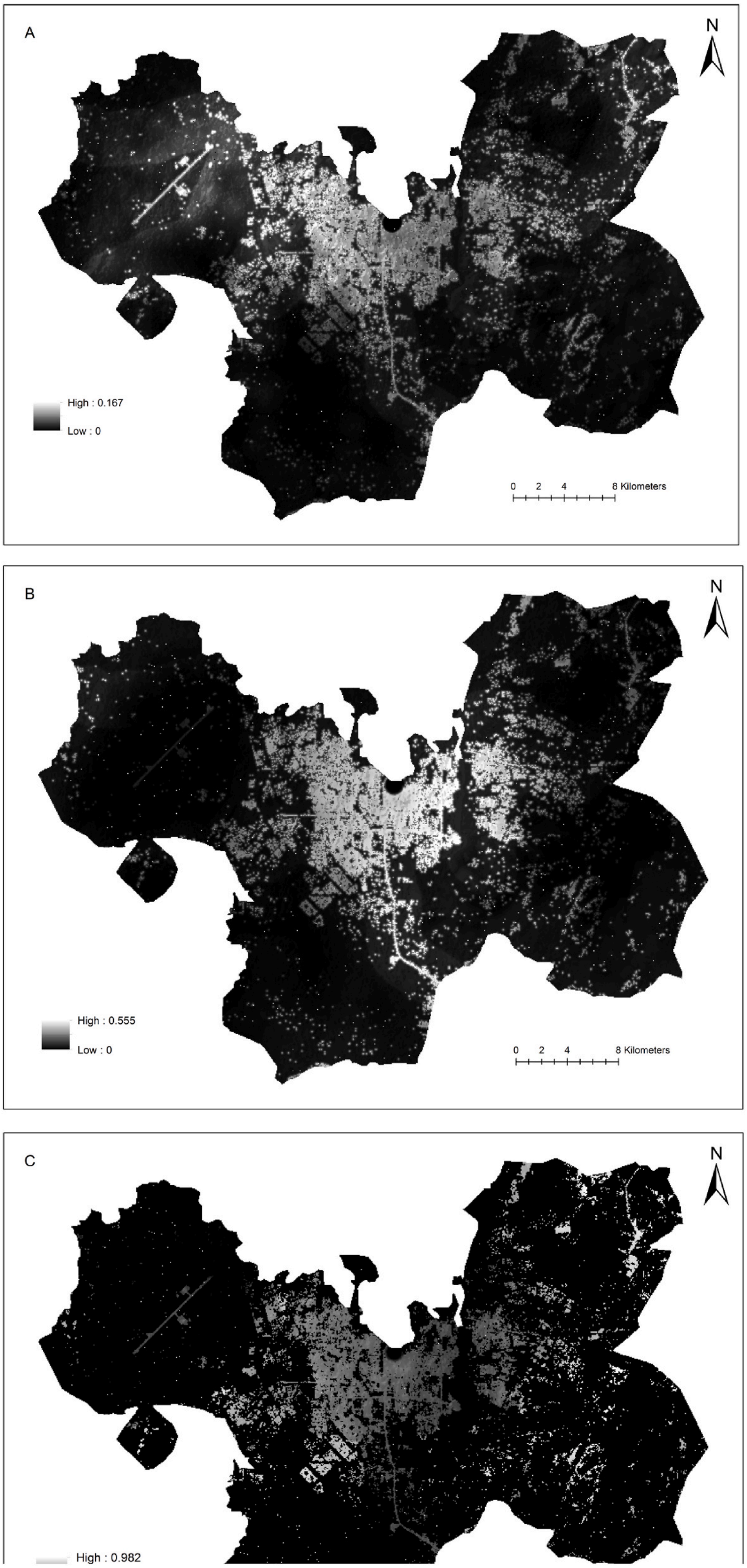

Fig. 6. Urban growth probability maps for Bahir Dar city and its surroundings in A (1984–1994), B (1994–2009) and C (2009–2019), predicted by logistic regression model.

BLRM was an effective technique for the prediction of urban growth probability and analysis. The PCP values were also computed to validate the proposed model. These values have some variations in different urban growth periods. It was 96 % in the growth periods between 1984 and 1994, 98 % from 1994 to 2009 and 95 % from 2009 to 2019. This PCP value indicates that the model had a good predictive ability for urban growth in the specified period. Furthermore, the ROC and PCP values indicate the presence of high agreement between the predicted the actual maps.

4. Discussion

The study of urban growth, land use change and its influencing variables as well as projection of these landscape changes are fundamental tools for developing efficient management and allocation of resources and effective future change scenarios. This study illustrated how to estimate urban growth changes while using remote sensing and a geographic information system simultaneously. In order to collect information on the phenomenon of land cover changes in the examined area, remote sensing data were used. The acquisition of the necessary input data for the models was aided significantly using GIS. Bahir Dar's expected urban growth, like that of cities of other developing nations, would be chaotic, resulting in the establishment of slum areas in crop lands, forests and other reserved green areas. The finding of this study explained the drivers of urban growth that affect urban growth processes to varying degrees, as detailed below.

The results of this study indicated that urban built-up areas have increased in the period from 1984 to 1994 by 0.54 %. In other words, built –up area have increased by 113.62 ha (11.3 ha/year). However, cropland and forest have been declined by 5.8 % or 1248.26 ha (124.83 ha/year) and 0.8 % or 170.1 ha (17 ha/year) in the same study period. In the second study period (1994–2009), built-up surfaces have increased by 6.61 %. In other words, built-up area have increased by 1410.53 ha (94.3 ha/year). In contrast to this, crop land and forest have been declined 1.8 % (372.05) and 0.8 % (4.83 ha) in the same study period. The figures in this study period indicates that crop land and forest have been decreased by 24.48 ha/year and 0.32 ha/year. In the period 2009–2019, built-up surfaces of Bahir Dar city have increased by 3.6 % (761.89 ha). Whereas, crop land and forest have been declined by 3.6 % (372.05 ha) and 1.9 % (409.29 ha) in the same study period. In this study period, built-up areas have increased by 76.2 ha/year, while crop and forest have been decreased by 76.46 ha/year and 40.93 ha/year, respectively. In the past 35 years (1984–2019), built-up areas have been increased by 2286.04 (11 %), but crop land forest areas have been declined by 2384.92 (11.17 %) and 574.56 (2.69 %). In other words, built-up have been increased at the expanse other land use classes by 65.32 ha/year. However, crop land and forest areas have been decreased by 68.14 ha/year and 16.42 ha/year in the past 35 years.

Studies on urban growth and land is change have been carried out by various researchers in different cities around the world. For instance Ref. [82], have conducted a study on urban sprawl and its impact on land use change in central Ethiopia, Dukem town. This study reported that the built up surface of Dukem occupied only 648.7 ha of total area. the other study carried out by Ref. [83] on dynamics of urban expansion and land cover change in Northern Ethiopia, Mekele city and indicated that built-up areas have been increased annually by an average of 8 % in past three decades (1984–2014) [82].also conducted research on urban growth and land use change in the capital city of Ethiopia, Addis Ababa. Based on the results of this study built-up areas of Addis Ababa has been increased by 2.27 km^2/year from 1990 to 2020. Agriculture and forest has been converted to built-up area at fastest rate during this period. However, Bahir Dar city has been growing in fastest rate as compared to other cities of Ethiopia like Mekele, Jigjiga, Hawassa, Jimma and even Addis Ababa, Capital city of the country. Studies were also conducted in cities of other developing countries. For example [82],conducted study on urban land use change and peri – urban land patterns in Takoradi and Bolgatanga, Ghana. The study reveals that urban areas have been increased by 7.1 % in Takoradi and 1.1 % in Bol-gatanga between 2007 and 2013. The study demonstrated this growth rate of two districts of Ghana, West Africa as fastest urban growth. However, this growth rate is not faster than the growth rates observed in Bhir Dar city and other Ethiopian towns and cities. The other study carried out in Africa was conducted by Ref. [82] on land use land cover change in Blatyre city, Malawi and reported that built-up areas have been increased by 1.43 km^2 in the past 20 years from 1999 to 2019. Study conducted in Dar es Salaam city, Tazania by Ref. [82]on analysis of current and future land cover change. The findings of this study reveal that built-up areas have been increased by 8.1 % in the past two decades from 1999 to 2019 and forest has been declined by 4.8 %, due to rapid urban growth. The other researcher Rahimi, (2016) carried out study on urban growth changes modeling in Tabrize, Iran. The results of this study indicated that built up areas have been increased in fastest rate i. e 8.5 % (4860.217 ha) from 1989 to 2021. Though urban growth against other land use classes such as forest and crop land is common phenomenon in developing countries, the growth rate of Bahir Dar city has been increased at fastest rate.

In the period 1984–1994, Dis. to forest land, Dis. to CBD, aspect, dis. to water body and Dis. to international airport has odds ratio of 14.37, 14.37, 4, 3.54 and 3.13, which indicates that these variables had a higher influence on urban growth processes than others. The remaining variables include Dis. to existing built-up land, Dis. to main road and slope were significant enough to affect urban growth processes with odds ratio of 1.09, 1.43 and 1.38. However, elevation had an odds ratio of 0.78, which is not significant enough in affecting urban growth processes during this period. Study conducted by Ref. [84] indicated that Dis. to main road has a lesser effect on urban growth in Greater Cairo Region, Egypt, which is not similar with the results of this study and other similar studies such as [67,85,86]. Based on study conducted by Ref. [87] demonstrated that the main factors affecting urban growth using logistic regression in New Castle County are population density, slope, proximity to roads, and surrounding land use with high odds ratio greater than 1 [87]. also conducted study on spatio-temporal urban growth modeling using binary logistic regression and GIS in Colombo urban fringe, Sri Lanka. This study identified three main factors influencing urban growth like land value, proximity to green areas, and population density in the study area.

From 1994 to 2009, all the factors incorporated in the model, except Dis. to water body, had odds ratio>1. This indicates the factors had a great influence on the processes of urban growth during the specified period. Distance to forest land, Dis. to CBD, and Dis. to

Bahir Dar international airport had high influence than others with odds ratio 515, 109 and 25.73. the other factors such as Dis. to main road, slope, aspect, elevation and Dis. to exiting built-up land also had a significant influence on urban growth processes with odds ratio of 1.18, 1.73, 1.34, 1.28 and 1.08. These results are similar with the previous studies conducted in similar cases in different study areas [87–89] conducted study on spatio-temporal dynamics and driving factors analysis of urban growth in Zhuhai city, China. The findings of this study indicated that elevation, slope, distance from built-up land and growth rate of built-up land are important factors influencing the spatio-teporal variation and transformation of land use land cover from 1991 to 2018 in the study area [87]. also tried to investigate the urban growth patterns of Ahvaz, Iran 1991–2000 using GIS and logistic regression. The factors influencing urban growth in Ahvaz are distance to the bridge, rural Areas, planned town and industry activities (all with odds ratios<1_or coefficient <0); and number of urban cells within a 5·5 cell window, distance to the hospitals, main road, high road, rail Line, river, CBD and secondary centers, agriculture areas and vacant areas. The present study from the period 1994–2009 indicates Dis. to river is not influencing urban growth of Bahir Dar, but the study conducted by Ghalehnoee in Iran reported that Dis. to water body was influencing urban growth.

The last but not the least urban growth period (2009–2019), Dis. to CBD, Dis. to international airport and Dis. to water body were more influential factors that affect urban growth processes with odds ratio 320.89, 30.97 and 3.29. The other factors like Dis. to forest land, Dis. to existing built-up land, Dis. to main road, slope, aspect and elevation had odds ratio of 1.23, 1.25, 1.33, 1.82, 2.20 and 1.33. All the explanatory variables incorporated in the model had odds ratio>1. This reveals that all the factors used in the model significantly influenced urban growth processes during this period. According to Ref. [87], Dis. to water body and Dis. to existing urban area are less influential factors for urban growth. In contrary to this, Dis. to main road was the main driving forces of urban growth as stated by the same study. The other study conducted on driving factors of urban growth indicated that Dis. to main road and Dis. to international airport were the main driving factors of urban growth [41]. Study conducted on urban growth modeling by Ref. [87] in Irbid city, North Jordan identified factors influencing urban growth. Findings in this study indicated that the main factors influencing the spatiotemporal urban distribution in the study area are slope, distance to road, distance to urban centers, distance to commercial, density, elevation, and land fertility with odds ratio greater than 1. This study better describes and identifies the factors of urban growth compared to other previous studies that gave emphasis only for few proximity and physical factors. Gharaibeh et al. (2020) also carriedout study on application of binary logistic regression model on analysis of urban growth factors in Limbe, Cameroon from 1986 to 2019 and the study reported that built-up areas has been increased from 3.6 % in 1986 to 17.6 % in 2019. The factors influencing urban growth in this city during the study area are Dis. to urban areas, Dis. to main road, Dis. to agriculture and Dis. to river. The other study conducted on spatiotemporal analysis of urban land use change in West Bengal, India also reported that Dis. to the nearby city and population density as main factors of urban growth in the study area.

In general, Dis. to the main road, Dis. to CBD and Dis. to international airport were the main factors of urban growth in Bahir Dar in all study periods. Despite BLRM's ability to include a large number of variables in the analysis, some critical aspects such as government policy issues, economic factors such as GDP and demographic factors like population density cannot be included in the model. This was due to absence of well-organized demographic, socio-economic and policy related data in developing countries especially in the study area, Bahir Dar city, Ethiopia. This study tried to identify various physical and proximity factors influencing the spatio-temporal distribution and growth in the study areas. Furthermore, the study has tried to assess the land use land cover change and urban growth from 1984 to 2019 in Bahir Dar. This makes the study different from other previous researches which gave only for short time period and few factors influencing the study. The results of this study can help the policy makers, researchers and decision makers to give precise and efficient as well as effective policies and decisions. Because it provides necessary information related to driving factors of urban growth and trends of urban growth in Bahir Dar. Hence land is the most important resource in Africa as well as Ethiopia, it needs effective policy and planning for sustainable management of this resources. This study also gives detail and critical analysis of factors of urban growth using quantitative ways, binary logistic regression, which is different from m previous studies conducted in the study area Bahir Dar city as well as in Ethiopia.

5. Conclusion

This study identified nine significant factors affecting spatial distribution of urban growth in Bahir Dar and its surroundings and analyzed their influence on urban growth processes using BLRM. Accordingly, Dis. to CBD, Dis. to international airport, aspect and Dis. to forest land were the most significant driving factors than others based on their odds ratio during the last ten years (1984–1994). In the same way, Dis. to CBD, Dis. to forest land and Dis. to international airport were the main significant driving factors in the period 1994–2009. In the period 2009–2019, Dis. to CBD, Dis. to water body and Dis. to internal airport were more significant explanatory variables which had high influence on urban growth processes in Bahir Dar city. The ROC value was used to examine the accuracy of the model for the probability of urban growth, which indicated the model is valid.

The findings of this study can be applicable not only to Bahir Dar city but also other cities of developing nations including Ethiopia. The study examined and identified the primary factors that influence urban growth processes that would be regulated via the development of appropriate laws and regulations. This study recommends the integration of BLRM with Cellular automata to overcome the limitations of BLRM in studying spatiotemporal dynamics of urban growth. In this study socio –economic factors such as GDP and population density was initially identified as driving factors of urban growth. However, adequate, quality and time series data for these socio – economic factors was not available from the concerned office. So, these factors were rejected from further analysis and not incorporated in the present study. This is the other limitation of this study. The study also used Landsat satellite images with medium resolution (30 m) instead of high resolution images. High resolution images are preferable to obtain more quality results. However, due to lack of access to these high resolution images, this study was done using medium resolution satellite images. So, the study recommends to future potential researchers to use high resolution images to further improve the quality of the results for the same study. Further research should be addressing the role of government policy, GDP and population density in the urban growth processes of Bahir Dar.

Data availability statement

Data will be made available upon reasonable request from the corresponding author.

CRediT authorship contribution statement

Kenu Getu: Formal analysis, Conceptualization. **H. Gangadhara Bhat:** Methodology, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acronyms and Abbreviation

BLRM – Binary logistic regression model
CBD – Central business district
DEM – Digital elevation model
EMA – Ethiopian mapping agency
GIS – Geographic information system
GCPs – Ground control points
LRM – Logistic regression model
LULC – Land use land cover
MLC – Maximum likelihood classification
PCP - Percentage of correct prediction
ROC – Relative operating characteristic
VIF – Variance inflation factor

References

[1] M. Arfanuzzaman, B. Dahiya, Sustainable urbanization in Southeast Asia and beyond: challenges of population growth, land use change, and environmental health, Growth Change 50 (2) (2019) 725–744.
[2] W. Sun, et al., Geospatial analysis of urban expansion using remote sensing methods and data: a case study of yangtze river delta, China, Complexity 2020 (2020).
[3] A. Traore, T. Watanabe, Modeling determinants of urban growth in conakry , Guinea : a spatial logistic approach, Urban Sci. 1 (2017) 12.
[4] UN, World Urbanization Prospects, vol. 12, 2018.
[5] M. Zhao, et al., Mapping urban dynamics (1992–2018) in Southeast Asia using consistent nighttime light data from DMSP and VIIRS, Remote Sens. Environ. 248 (July) (2020) 111980.
[6] N.S. Sumari, P.B. Cobbinah, F. Ujoh, G. Xu, On the absurdity of rapid urbanization: spatio-temporal analysis of land-use changes in Morogoro, Tanzania, Cities 107 (September 2019) (2020) 102876.
[7] B. Bagheri, A. Soltani, The spatio-temporal dynamics of urban growth and population in metropolitan regions of Iran, Habitat Int. 136 (2023) 102797.
[8] Y. Cao, X. Zhang, Y. Fu, Z. Lu, X. Shen, Urban spatial growth modeling using logistic regression and cellular automata: a case study of Hangzhou, Ecol. Indicat. 113 (December 2019) (2020) 106200.
[9] A. khaje borj sefidi, M. Ghalehnoee, Analysis of urban growth pattern using logistic regression modeling, spatial autocorrelation and fractal analysis Case study: Ahvaz city TT -, Int. J. Archit. Eng. Urban Plan 26 (2) (2016) 183–194.
[10] A.R. Pardakhti, M. Mejari, Analysis of urban growth pattern using logistic regression modeling, spatial autocorrelation and fractal analysis Case study: Ahvaz city, Int. J. Archit. Eng. Urban Plan. 26 (2) (2016) 183–194.
[11] Y. Nong, Q. Du, Urban growth pattern modeling using logistic regression, Geo-Spatial Inf. Sci. 14 (1) (2011) 62–67.
[12] G. Xu, et al., Urban expansion and form changes across African cities with a global outlook: spatiotemporal analysis of urban land densities, J. Clean. Prod. 224 (2019) 802–810.
[13] A. Bakshi, M. Esraz-Ul-Zannat, Application of urban growth boundary delineation based on a neural network approach and landscape metrics for Khulna City, Bangladesh, Heliyon 9 (6) (2023) e16272.
[14] S. Vafaei, M.M. Karim, S. Soltanian, S. Rasooli, Simulating the expansion of built-up areas using the models of logistic regression, artificial neural network, and geo-mod in marivan city, Iran, J. Indian Soc. Remote Sens. 49 (5) (2021) 1081–1090.
[15] N. Li, S. Miao, Y. Wang, The future urban growth under policies and its ecological effect in the Jing-Jin-Ji area, China, Heliyon 7 (4) (2021) e06786.
[16] Y. Chen, Z. Chen, D. Guo, Z. Zhao, T. Lin, ScienceDirect Underground space use of urban built-up areas in the central city of Nanjing : insight based on a dynamic population distribution, Undergr. Space 7 (5) (2022) 748–766.
[17] B. Singh, V. Venkatramanan, B. Deshmukh, Monitoring of land use land cover dynamics and prediction of urban growth using Land Change Modeler in Delhi and its environs , India, Environ. Sci. Pollut. Res. (2022) 71534–71554.
[18] K. Kourtit, M. Mazurencu, M. Pele, P. Nijkamp, D. Traian, Safe cities in the new urban world : a comparative cluster dynamics analysis through machine learning, Sustain. Cities Soc. 66 (February 2020) (2021) 102665.

[19] G. Li, S. Sun, C. Fang, Landscape and Urban Planning the varying driving forces of urban expansion in China : insights from a spatial-temporal analysis, Landsc. Urban Plann. 174 (November 2016) (2018) 63–77.
[20] D. K, D.P. Angadi, Urban expansion quantification from remote sensing data for sustainable land-use planning in Mangaluru, India, Remote Sens. Appl. Soc. Environ. 23 (August) (2021) 100602.
[21] K.T. Deribew, Spatiotemporal analysis of urban growth on forest and agricultural land using geospatial techniques and Shannon entropy method in the satellite town of Ethiopia, the western fringe of Addis Ababa city, Ecol. Process. 9 (1) (2020).
[22] A. Saxena, M.K. Jat, M. Choudhary, Analysis of urban growth using geospatial techniques, Int. J. Earth Sci. Eng. 9 (6) (2016) 2855–2861.
[23] I.R. Hegazy, M.R. Kaloop, Monitoring urban growth and land use change detection with GIS and remote sensing techniques in Daqahlia governorate Egypt, Int. J. Sustain. Built Environ. 4 (1) (2015) 117–124.
[24] UN-Habitat, Monitoring Human Settlements Indicators SUSTAINABLE DEVELOPMENT GOALS, 2020.
[25] B. Addae, N. Oppelt, Land-use/land-cover change analysis and urban growth modelling in the greater accra Metropolitan area (GAMA), Ghana, Urban Sci. 3 (1) (2019) 26.
[26] O.A. Frank, A.A. Kwaku, O. Emmanuel, N.O. Samuel, Trends in land use/land cover changes in the Densu River basin and its impact on the Weija reservoirs and the Densu Delta (Sakumo I lagoon) in Ghana, J. Geogr. Reg. Plann. 12 (4) (2019) 76–89.
[27] J. Kleemann, J.N. Inkoom, M. Thiel, S. Shankar, S. Lautenbach, C. Fürst, Peri-urban land use pattern and its relation to land use planning in Ghana, West Africa, Landsc. Urban Plann. 165 (2017) 280–294.
[28] T. Dessu, D. Korecha, D. Hunde, A. Worku, Long-Term Land Use Land Cover Change in Urban Centers of Southwest Ethiopia from a Climate Change Perspective, vol. 2, 2020, pp. 1–23, no. December.
[29] P.S. Pradeeep, Land use and land cover analysis of nellore using GIS and remote sensing techniques, Int. J. Eng. Res. Technol. 8 (6) (2019) 390–394.
[30] A. Heymans, J. Breadsell, G.M. Morrison, J.J. Byrne, C. Eon, Ecological urban planning and design: a systematic literature review, Sustain. Times 11 (13) (2019).
[31] B. Frank, Urban systems: a socio-ecological system perspective, Soc. Int. J. 1 (1) (2017) 1–8.
[32] I. Dutta, A. Das, Application of geo-spatial indices for detection of growth dynamics and forms of expansion in English Bazar Urban Agglomeration , West Bengal, J. Urban Manag. 8 (2) (2019) 288–302.
[33] C. M, C.B. Sarah West, Miranda Loh, Urban air pollution and children in low- and middle-income countries, Glob. Alliance – Cities 4 Child. (2021).
[34] T.N.T. N, Q.H.B.T.D. Pham, V.H. Pham, Q.T. Luu, X.T. Ngo, Analyzing the impacts of urban expansion on air pollution in Vietnam using the SEAP platform Analyzing the impacts of urban expansion on air pollution in Vietnam using the SEAP platform, Earth Environ. Sci. Pap. 266 (2019) 012008.
[35] N. Shrestha, Assessing Discriminatory Performance of a Binary Logistic Regression Model, 2020 no. July 2019.
[36] H.Z.M. S, Abubakr, A.A. Alsharif, Biswajeet Pradhan, Shattri Mansor, Urban expansion assessment by using remotely sensed data and the relative, Theor. Empir. Res. Urban Manag. 10 (1) (2015).
[37] M.M. Aburas, M.S.S. Ahamad, N.Q. Omar, Spatio-temporal simulation and prediction of land-use change using conventional and machine learning models: a review, Environ. Monit. Assess. 191 (4) (2019).
[38] Y. Ren, Y. Lü, A. Comber, B. Fu, P. Harris, L. Wu, Spatially explicit simulation of land use/land cover changes: current coverage and future prospects, Earth Sci. Rev. 190 (2019) 398–415.
[39] S. Buya, P. Tongkumchum, K. Rittiboon, S. Chaimontree, Logistic regression model of built-up land based on grid-digitized data structure: a case study of krabi, Thailand, J. Indian Soc. Remote Sens. 50 (5) (2022) 909–922.
[40] M.W. Naikoo, et al., Analysis of peri-urban land use/land cover change and its drivers using geospatial techniques and geographically weighted regression, Environ. Sci. Pollut. Res. (2022) 0123456789.
[41] K. Mom, S. Ongsomwang, Urban growth modeling of phnom penh, Cambodia using satellite imageries and a logistic regression model, Suranaree J. Sci. Technol. 23 (4) (2016) 481–500.
[42] M. Wang, L. Cai, H. Xu, S. Zhao, Predicting land use changes in northern China using logistic regression, cellular automata, and a Markov model, Arabian J. Geosci. 12 (24) (2019).
[43] P. Myagmartseren, D. Ganpurev, I. Myagmarjav, G. Byambakhuu, G. Dabuxile, Remote sensing and multivariate logistic regression model for the estimation of urban expansion (case of Darkhan City, Mongolia), Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 43 (B3) (2020) 721–726.
[44] X. Chang, F. Zhang, K. Cong, X. Liu, Scenario simulation of land use and land cover change in mining area, Sci. Rep. 11 (1) (2021) 1–12.
[45] M.P. Bavaghar, Deforestation modelling using logistic regression and GIS, J. For. Sci. 61 (5) (2015) 193–199.
[46] C. Chivuraise, T. Chamboko, G. Chagwiza, An assessment of factors influencing forest harvesting in smallholder tobacco production in hurungwe district, Zimbabwe: an application of binary logistic regression model, Adv. Agric. 2016 (2016).
[47] E. Pujiono, R. Sadono, Hartono, M.A. Imron, P.Y.A.P. Wirabuana, Factors contributing to forest degradation in the mountainous tropical forest: a case study of the mutis-timau forest complex, Indonesia, J. Sustain. For. (2022) 1–20, vol. 00, no. 00.
[48] S. Gaur, A. Mittal, A. Bandyopadhyay, I. Holman, R. Singh, Spatio-temporal analysis of land use and land cover change: a systematic model inter-comparison driven by integrated modelling techniques, Int. J. Rem. Sens. 41 (23) (2020) 9229–9255.
[49] J. Cui, M. Zhu, Y. Liang, G. Qin, J. Li, Y. Liu, Land use/land cover change and their driving factors in the yellow river basin of shandong province based on Google earth engine from 2000 to 2020, ISPRS Int. J. Geo-Inf. 11 (3) (2022).
[50] C.N. Huu, C.N. Van, T.N.N. My, Modeling land-use changes using logistic regression in Western Highlands of Vietnam: a case study of Lam Dong province, Agric. Nat. Resour. 56 (5) (2022) 935–944.
[51] E.A. Alshari, B.W. Gawali, Modeling land use change in sana'a city of Yemen with MOLUSCE, J. Sens. 2022 (2022).
[52] O. Hamdy, S. Zhao, M.A. Salheen, Y.Y. Eid, Analyses the driving forces for urban growth by using IDRISI®Selva models abouelreesh - aswan as a case study, Int. J. Eng. Technol. 9 (3) (2017) 226–232.
[53] T. Luo, R. Tan, X. Kong, J. Zhou, Analysis of the driving forces of urban expansion based on a modified logistic regression model: a case study of Wuhan City, Central China, Sustain. Times 11 (8) (2019).
[54] I. Grigorescu, G. Kucsicsa, B. Mitrică, I. Mocanu, M. Dumitrașcu, Driving factors of urban sprawl in the Romanian plain. Regional and temporal modelling using logistic regression, Geocarto Int. 37 (24) (2022) 7220–7246.
[55] A. Eyoh, D.N. Olayinka, P. Nwilo, O. Okwuashi, M. Isong, D. Udoudo, Modelling and predicting future urban expansion of Lagos, Nigeria from remote sensing data using logistic regression and GIS, Int. J. Appl. 2 (5) (2012) 1–9.
[56] B. Bhatta, Analysis of Urban Growth and Sparwl from Remote Sensing Data, Springer-Verlag Berlin Heidelberg, 2010.
[57] A. Wubalem, M. Meten, Landslide susceptibility mapping using information value and logistic regression models in Goncha Siso Eneses area, northwestern Ethiopia, SN Appl. Sci. 2 (5) (2020) 1–19.
[58] O. Hamdy, S. Zhao, T. Osman, M.A. Salheen, 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. 6 (4) (2016).
[59] O.J.G. Somsen, Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land-use change modeling: a case study, J. Sagepub.Com 47 (3) (2015) 220–243.
[60] Z. Mei, H. Wu, S. Li, Simulating land-use changes by incorporating spatial autocorrelation and self-organization in CLUE-S modeling: a case study in Zengcheng District, Guangzhou, China, Front. Earth Sci. 12 (2) (2018) 299–310.
[61] B. Sun, D.T. Robinson, Comparison of statistical approaches for modelling land-use change, Land 7 (4) (2018).
[62] Y.P. Lin, H.J. Chu, C.F. Wu, P.H. Verburg, Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land-use change modeling - a case study, Int. J. Geogr. Inf. Sci. 25 (1) (2011) 65–87.
[63] C. Li, J. Li, J. Wu, Quantifying the speed, growth modes, and landscape pattern changes of urbanization: a hierarchical patch dynamics approach, Landsc. Ecol. 28 (10) (2013) 1875–1888.

[64] Z. Zhang, Y.J. Tu, X. Li, Quantifying the Spatiotemporal patterns of urbanization along urban-rural gradient with a roadscape transect approach: a case study in Shanghai, China, Sustain. Times 8 (9) (2016).
[65] J. Gong, Z. Hu, W. Chen, Y. Liu, J. Wang, Urban expansion dynamics and modes in metropolitan Guangzhou, China, Land Use Pol. 72 (December 2017) (2018) 100–109.
[66] J. Ou, X. Liu, X. Li, Y. Chen, J. Li, Quantifying spatiotemporal dynamics of urban growth modes in metropolitan cities of China: Beijing, Shanghai, tianjin, and guangzhou, J. Urban Plann. Dev. 143 (1) (2017) 04016023.
[67] Z. Hu, C.P. Lo, Modeling urban growth in Atlanta using logistic regression, Comput. Environ. Urban Syst. 31 (6) (2007) 667–688.
[68] P. Christiansen, T. Loftsgarden, Drivers behind Urban Sprawl in Europe, 2011.
[69] R. Xiao, Y. Liu, X. Huang, R. Shi, W. Yu, T. Zhang, Exploring the driving forces of farmland loss under rapidurbanization using binary logistic regression and spatial regression: a case study of Shanghai and Hangzhou Bay, Ecol. Indicat. 95 (May) (2018) 455–467.
[70] A. Achmad, S. Hasyim, B. Dahlan, D.N. Aulia, Modeling of urban growth in tsunami-prone city using logistic regression: analysis of Banda Aceh, Indonesia, Appl. Geogr. 62 (2015) (2015) 237–246.
[71] A. Sarkar, P. Chouhan, Modeling spatial determinants of urban expansion of Siliguri a metropolitan city of India using logistic regression, Model. Earth Syst. Environ. 6 (4) (2020) 2317–2331.
[72] P.H. Verburg, et al., Land use change modelling : current practice and research priorities Linked references are available on JSTOR for this article : land use change modelling : current practice and research priorities, Source Geoj. GeoJoumal 61 (4) (2016) 309–324.
[73] A.F. Alqurashi, L. Kumar, K.A. Al-ghamdi, Spatiotemporal Modeling of Urban Growth Predictions Based on Driving Force Factors in Five Saudi Arabian Cities, 2016.
[74] Lu Jeffery Allen and Kang, Modeling and prediction of future urban growth in the charleston region of South Carolina : a GIS-based integrated approach, Conserv. Ecol. 8 (2) (2003).
[75] A.A.A. Alsharif, B. Pradhan, Urban sprawl analysis of tripoli metropolitan city (Libya) using remote sensing data and multivariate logistic regression model, J. Indian Soc. Remote Sens. 42 (1) (2014) 149–163.
[76] S.K. Shah, Binary logistic model to identify the factors associated with households with bank accounts in Nepal, Quest J. Manag. Soc. Sci. 2 (2) (2020) 323–336.
[77] N. Shrestha, Detecting multicollinearity in regression analysis, Am. J. Appl. Math. Stat. 8 (2) (2020) 39–42.
[78] J.H. Kim, Multicollinearity and misleading statistical results, Korean J. Anesthesiol. 72 (6) (2019) 558–569.
[79] X. Yinghui, Z. Qingming, A review of remote sensing applications in urban planning and management in China. 2009 Jt. Urban Remote Sens. Event, 2009 no. March 2014.
[80] D. Minetos, S. Polyzos, Multivariate statistical methodologies for testing hypothesis of land-use change at the regional level. A review and evaluation, J. Environ. Prot. Ecol. 10 (3) (2009) 834–866.
[81] O. Hamdy, S. Zhao, M.A. Salheen, Y.Y. Eid, Analyses the Driving Forces for Urban Growth by Using IDRISI ® Selva Models Abouelreesh - Aswan as a Case Study, vol. 9, 2017 no. 3.
[82] H. Faramarzi, S.M. Hosseini, H.R. Pourghasemi, M. Farnaghi, Forest fire spatial modelling using ordered weighted averaging multi-criteria evaluation, J. For. Sci. 67 (2) (2021) 87–100.
[83] A.A. Fenta, et al., The dynamics of urban expansion and land use/land cover changes using remote sensing and spatial metrics: the case of Mekelle city of northern Ethiopia, Int. J. Rem. Sens. 38 (14) (2017) 4107–4129.
[84] M. Salem, N. Tsurusaki, 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, Infrastructure 4 (1) (2019).
[85] M. Salem, et al., Urban expansion simulation based on various driving factors using a logistic regression model: Delhi as a case study, Sustain. Times 13 (19) (2021) 1–17.
[86] M. Salem, N. Tsurusaki, P. Divigalpitiya, T. Osman, Driving factors of urban expansion in peri-urban areas of greater Cairo region, in: *23rd* Int. Conf. Urban Planning, Reg. Dev. Inf., 2018, pp. 191–196, vol. 0, no. April.
[87] B. Huang, L. Zhang, B. Wu, Spatio-temporal analysis of rural-urban land conversion, Int. J. Geogr. Inf. Sci. 23 (10) (2009) 1–27.
[88] G. Nduwayezu, R. Sliuzas, M. Kuffer, Modeling Urban Growth in Kigali City Rwanda, vol. 1, 2016 no. Ii.
[89] J. Cheng, I. Masser, Urban growth pattern modeling: a case study of Wuhan City, PR China, Landsc. Urban Plann. 62 (4) (2003) 199–217.