Proceedings of Engineering and Technology Innovation, vol. 23, 2023, pp. 23-35

Modeling and Forecasting Urban Sprawl in Sylhet Sadar Using Remote Sensing Data

Md Aminul Islam, Tanzina Ahmed Rickty, Pramit Kumar Das, Md Bashirul Haque*

Department of Civil and Environmental Engineering, Shahjalal University of Science and Technology, Sylhet, Bangladesh

Received 03 March 2022; received in revised form 25 June 2022; accepted 26 June 2022

DOI: https://doi.org/10.46604/peti.2023.9617

Abstract

Forecasting urban sprawl is important for land-use and transport planning. The aim of this study is to model and predict the future urban sprawl in Sylhet Sadar using remote sensing data. The ordinary least square (OLS) regression model and the geographic information system (GIS) are used for modeling urban expansion. The model is calibrated for the years 2014 to 2017 using eight explanatory variables extracted from the regression model. The regression coefficients of the variables are found statistically significant at a 99% confidence level. The cellular automata (CA) model is then used to analyze, model, and simulate the land-use and land-cover (LULC) changes by incorporating the algorithm of logistic regression (LR). The calibrated model is used to predict the 2020 map, and the result shows that the predicted map and the actual map of 2020 are well agreed. By using the calibrated model, the simulated prediction map of 2035 shows an urban cell expansion of 220% between 2020 and 2035.

Keywords: geographic information system, remote sensing, land use and land cover, urban sprawl, ordinary least square regression, cellular automata

1. Introduction

Urbanization is one of the most significant human activities, influencing the quality of urban life and its long-term development all over the world [1]. Before the 1950s, developed countries experienced rapid urbanization, resulting in an increase in urban growth and a major decline in agricultural fields [2]. However, urbanization in developing countries in recent years has outpaced that of developed countries [3]. Urbanization can be the index of measuring the economic progress and development of a country, but urbanization in developing countries is accompanied by unemployment, poverty, crumbling infrastructure, environmental degradation, and the unmanageable spread of informal settlements [4]. In many parts of the world, urban sprawl can have a significant impact on the quality of life. As a result, developments in urbanization and their magnitude should be regularly observed [5]. Then, it can be possible to control the urbanization process and create sustainable developments [6].

Inadequate land-use planning, poor urban governance, and the lack of plans, rules, and implementation are all factors that contribute to urban sprawl [7]. For efficient planning, city governments and municipal corporations require an understanding of the urban sprawl phenomenon and its future movement. One of the most basic operations necessary for studying urban sprawl is mapping the built-up region. Using traditional surveys to assess urban sprawl is complex, time-consuming, and expensive. Furthermore, such data is not available for municipal corporations, particularly in developing nations. Taking into

* Corresponding author. E-mail address: bashir-cee@sust.edu
Tel.: +880-821-713491; Fax: +880-821-715257

account the cost-effectiveness, researchers must therefore rely on geographic information system (GIS) and remote sensing (RS) technologies to estimate the volume and direction of urban expansion [8]. Hence, it is required to justify artificial simulations for modeling the complexity of land-use change dynamics [9].

In Bangladesh, unplanned and haphazard urbanization is a common phenomenon [10]. Following the British departure from the Indian subcontinent, urban areas began to rise steadily. After the independence in 1971, the rate increased significantly. However, it was solely confined to Dhaka city at that time. After 1991, rapid urbanization began in Sylhet as well [11]. Over the last few decades, Sylhet has experienced fast but chaotic urbanization. According to a report published by the United Nations Population Fund (UNFPA) in 2016 [12], the urban area of Sylhet City Corporation grew from 10.49 square kilometers to 26.50 square kilometers in the 20 years between 1991 and 2011. The city is the sixth most urbanized in Bangladesh according to the 2001 population census [13]. This city's population is growing at a rate several times faster than most of the cities in the country. The rapid and unregulated urbanization of Sylhet, as well as the loss of arable lands and water bodies, are a matter of concern for this city.

To control the haphazard expansion of Sylhet city, it is essential to know the expansion patterns and the factors driving it. Although many studies in the literature have been found to explore the causes and patterns of urban expansion in the case of developed countries [14-16], this type of study for the cities in Bangladesh like Sylhet is very limited. The land-use and land-cover (LULC) of Sylhet city and its vicinity have altered due to unplanned urbanization and the capturing of lowland by real estate companies [17]. It was reported that water bodies of Sylhet district were reduced by 57% and 17% by the years 1988-1997 and 1997-2006 respectively, and unplanned urbanization played the key role in this reduction [18]. Although existing studies have captured the phenomena of haphazard urban expansion, none of them have explored the driving factors behind it, let alone modeling and predicting the future scenario of urban sprawl.

Thus, the goal of this study is to understand the urban sprawl of Sylhet Sadar, as well as the driving factors behind it, and to forecast future urban expansion. The following sections present the materials and methods used in this study, the analysis of the result obtained, and finally the discussion conducted regarding the findings.

2. Materials and Methods

This section describes the design adopted in this research, which includes data collection and image pre-processing, LULC map preparation, accuracy assessment regarding classification, selection, and preparation of independent variables, urban sprawl modeling using the ordinary least square (OLS) regression, model calibration and validation using the cellular automata (CA) model, and future LULC prediction.

2.1. Study area

The Sylhet Sadar Upazila (SSU), located in the Sylhet district, is a rapidly developing and resourceful area in Bangladesh's northeast part. Gowainghat, Companiganj, and Jaintiapur Upazila border it on the north, Dakshin Surma Upazila on the south, Golapganj and Kanaighat Upazila on the east, and Chhatak (Sunamganj district) and BishwanathUpazila on the west. Its coordinates are 24.8917°N 91.8833°E (Fig. 1). According to the 2011 census, the city has a population of 531,663 people and a density of 995 people per sq. km. The city's population growth rate is currently 1.73 percent, down from 1.93 percent in 1991. A thematic presentation of the methodology to be adopted under this study is shown in Fig. 2 below.

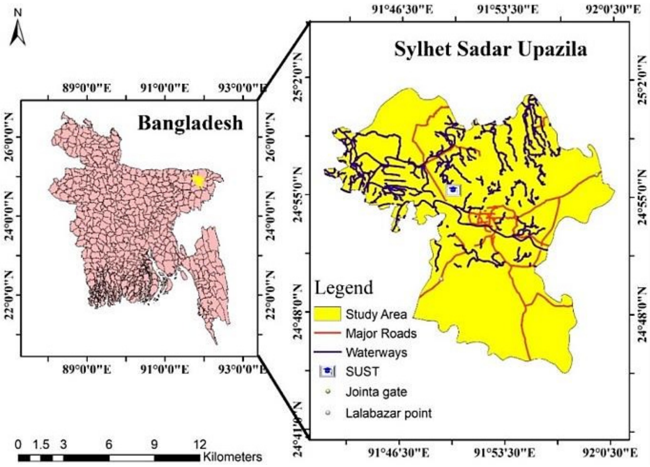


Fig. 1 Study area (Sylhet Sadar Upazila)

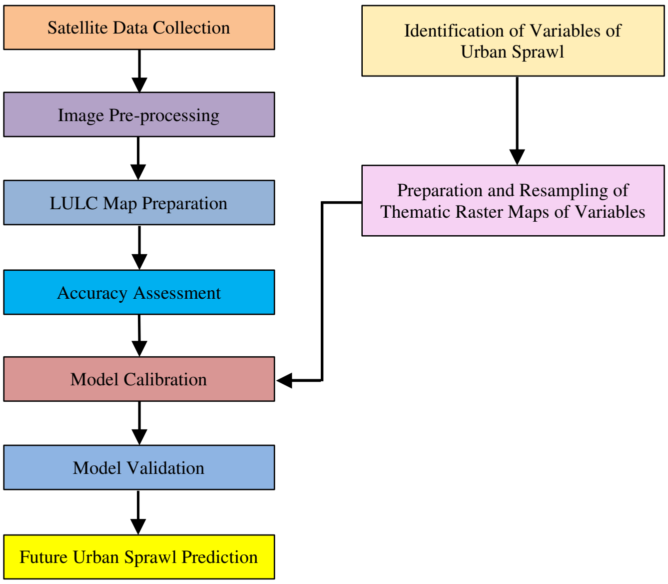


Fig. 2 Methodology of this study

2.2. Data collection

Various raster data and vector data are collected for this study. Raster data includes satellite data which consists of Landsat images of various sensors. Table 1 shows different data types and the related details such as their sources, date of acquisition, pixel size, and cloud cover.

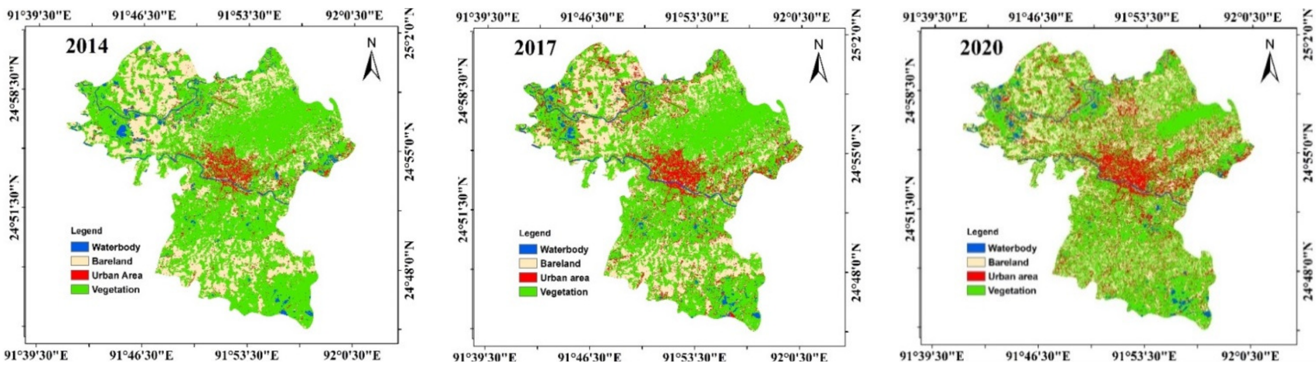
Table 1 Data used and its source

Data	Source	Time	Cloud cover (%)	Pixel size
Landsat image 2014 (Landsat 8 OLI)	United States Geological Survey	19-02-2014	0	30m
Landsat image 2017 (Landsat 8 OLI)	United States Geological Survey	26-01-2017	0	30m
Landsat image 2020 (Landsat 8 OLI)	United States Geological Survey	04-02-2020	0	30m
Shapefiles of variables	Open Street Maps	-	-	-

2.3. Data preparation

2.3.1. Classification of satellite images

The classification of images is one of the most important aspects to consider when performing any kind of LULC analysis [19]. Landsat images of the year 2014, 2017, and 2020 are classified into four specific LULC groups, i.e., (a) urban areas, (b) vegetation, (c) water bodies, and (d) bare land, by using the maximum likelihood classification in ArcMap. The result of the classification is visualized in ArcGIS, as shown in Fig. 3.



(a) 2014 (b) 2017 (c) 2020

Fig. 3 LULC maps of the Landsat images

2.3.2. Accuracy assessment of classified images

For this study, the images are categorized into four specific LULC groups, i.e., (a) urban areas, (b) vegetation, (c) water bodies, and (d) bare land, by using the maximum likelihood classification. For accuracy assessment of the classification, each classified map is evaluated by selecting 300 random points on them and cross-checked with Google Earth Pro images to assure that the right land-cover class is assigned to them. Random points are selected considering that all pixels for each land-cover class for accuracy assessment are not feasible. This is done in ArcGIS.

The user's accuracy is the accuracy from the point of view of a map user, not the map maker. The user's accuracy essentially tells how often the class on the map will be present on the ground. The producer's accuracy is the map accuracy from the point of view of the map maker (the producer). This refers to how often the real features on the ground are correctly shown on the classified map or the probability that a certain land cover of an area on the ground is classified as such. A summary of the classification accuracies is shown in Table 2, indicating a substantial agreement between the user's accuracy and the producer's accuracy. The kappa value of the accuracy assessment of all the three maps is greater than 0.8, which indicates a substantial agreement between the classified map and the reference data.

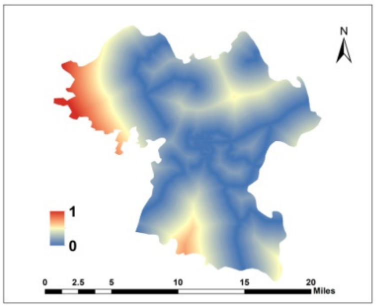
Table 2 Summary of classification accuracies for LULC maps

Respective year	User's accuracy (%)	Producer's accuracy (%)	Kappa
2014	89.34	81.66	0.903
2017	88	76	0.82
2020	89.09	76.6	0.81

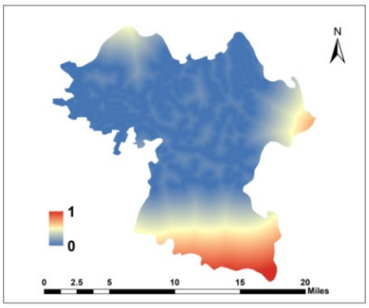
2.3.3. Thematic raster maps of independent variables

In regression models, defining and analyzing the relationship between the driving forces of land-use/cover change is a complex and challenging task. Since the drivers of land-use change are the key factors that can support understanding the processes of land-use shifting from non-urban to urban, selecting the driving forces of land use is an important part of urban growth modeling.

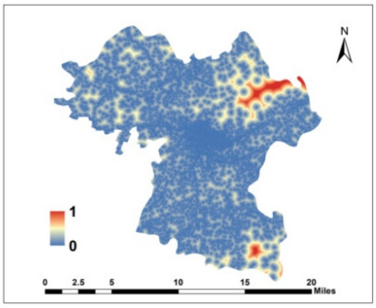
Eight independent variables are projected into the same spatial resolution and geographical coordinate system as that of the LULC maps. To input these factors into the OLS model, thematic raster maps are prepared for all the variables using the Euclidean distance function built-in ArcGIS (Fig. 4). The measured distance from each cell to the nearest source is stored in the Euclidean distance output raster. The values are stretched along the color ramp in this case. The Euclidean distance output raster is categorized into two classes: zero and one. One represents cells further away from the sources (independent variables), and zero represents cells closer to the sources. For example, for the variable "distance to Shahjalal University of Science and Technology (SUST)" (Fig. 4(f)), the blue zone denotes cells that are closer to SUST, while the red zone denotes cells that are further away.



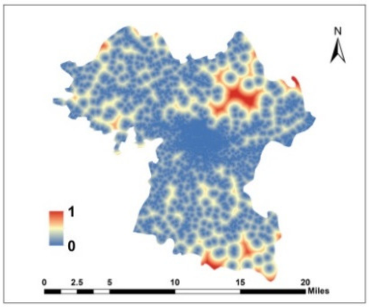
(a) Distance to major roads



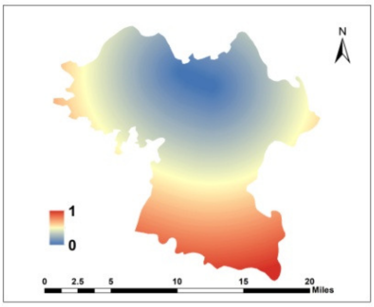
(b) Distance to waterways



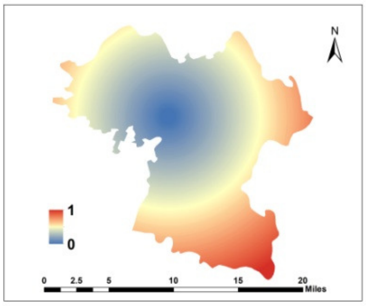
(c) Distance to dense urban areas



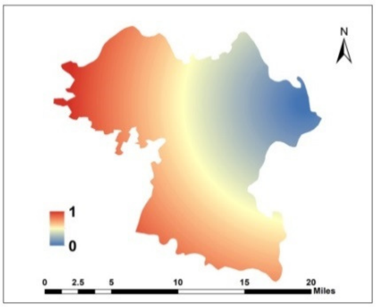
(d) Distance to less dense urban areas



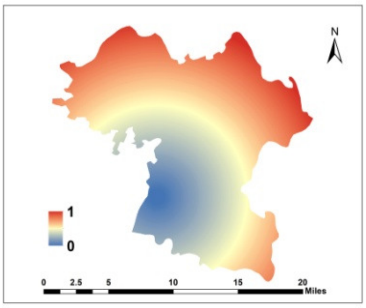
(e) Distance to Osmani international airport



(f) Distance to SUST



(g) Distance to Jointa-gate



(h) Distance to Lalabazar point

Fig. 4 Thematic raster maps of variables

For choosing the drivers of land-use changes, no fixed or global formula exists. Therefore, the reivers of land-use changes are selected on a case-to-case basis. The search for drivers of urban sprawl in this study starts with a trial-and-error procedure consisting of many variables (e.g., distance to forests, distance to religious monuments, distance to tourist spots, slope, etc.) but ends up selecting the eight factors found to be significant in the regression model.

The hypothesis regarding the impact of the driving factors of urban sprawl is drawn based on the finding of similar studies in the literature. For example, “distance from roads” is an important factor for urban sprawl as it has a significant impact on human activities, the environment, and socioeconomic development. “Distance to existing roads” controls human choices on housing, workplace, and schools as it has an impact on transportation costs, time, and comfort. Similarly, “proximity to urban cores” is a significant factor as it is related to good accessibility of settlements, affordable housing, job opportunities, transportation facilities, etc.

2.4. Modeling

Models are the tools used by researchers for studying the behavior of urban sprawl and examining the driving factors which influence this expansion. The structure simulation model (SEM) has been commonly used to model urban sprawl. It is a type of multiple regression analysis used to analyze structural relationships between dependent and independent variables. However, SEM is not appropriate to use when dependent variables have spatial characteristics. Rule-based simulation models (RSMs), such as the CA model, can be used with this kind of variable. Although the CA model can simulate the spatial pattern, it cannot interpret the driving factors of urban sprawl. Since regression is a widely used statistical tool for examining and modeling the relationships between variables, a combination of OLS regression and CA is used in this study.

2.4.1. Ordinary least square regression model

OLS is used for estimating unknown parameters in a linear regression model. OLS estimates the parameters of a linear function by minimizing the sum of squares of the differences between the observed and predicted values of the dependent variable. The general equation of the OLS model can be written as:

$$y_i = \beta_0 + \sum_{j=1}^{p} X_{ij}\beta_j + \varepsilon_i \tag{1}$$

where β_0 is the intercept coefficient, β_j is the slope coefficient for the j^{th} independent variable X_j, ε_i is the random error term with $N(0,\sigma^2 I)$, and I is the $n \times n$ identity matrix.

The dependent variable, i.e., urban expansion, is presented in a raster map. The role of independent explanatory factors is reflected in the regression coefficient. The dependent variable indicates urban growth from 2014 to 2017. Data from 2014 to 2017 are used in the modeling process for model calibration. Validation is done with the actual 2020 growth map and the predicted one, while future pattern projection is done with the data from 2020.

2.4.2. Cellular automata (CA) model

The CA model is used to simulate future LULC changes using the MOLUSCE plugin in QGIS 2.18.24. The plugin covers several procedures such as area change analysis, modeling techniques, simulation, and validation. The CA model accounts for both static and dynamic aspects of LULC changes, resulting in higher prediction accuracy [20]. To map realistic urban expansion patterns and LULC change dynamics, the CA model uses a bottom-up approach [21-22]. The expression of the CA model is shown in Eq. (2) [23].

$$S(t, t+1) = f(S(t), N) \tag{2}$$

where $S(t + 1)$ is the system status at the time of $(t, t + 1)$ that is functioned by the state probability of any time (N).

2.4.3. Model calibration

Researchers have used hybrid models such as a combination of CA and LR to model urban growth patterns and their driving factors [24]. Although the logistic regression (LR) learning model is well-known for its adaptability to both classification and regression problems [25], its shortcomings can be addressed by using a combination of CA and LR [26]. The model is first calibrated by simulating the period 2014 to 2017 in ArcMap using OLS. Next, the model is trained by using the LR under CA. For calibration from 2014 to 2017 (by overlaying the 2014 and 2017 maps), the urbanized areas are extracted for creating an overlay map. From the map overlaid, the following classes of land use are found: (i) undeveloped region, (ii) developed region, and (iii) change region. Overall, 28,500 points are randomly selected from the undeveloped cells and developed cells for training, whereas the selected point from the change region is null as this is not common in both the 2014 and 2017 maps. The dependent variable includes zeros (0) and ones (+1) in compliance with the regression model where zero and one stand for non-urban and urban areas respectively.

2.4.4. Model validation

For model training in QGIS, the explanatory variables (used for model calibration) and the LULC maps (2014 and 2017) are imported into QGIS. Then, the model is trained using LR. A prediction of 2020 is made and validated using the actual LULC map of 2020. For evaluating the performance matrix, Kappa statistic is used. Kappa statistic can be presented mathematically as shown in Eqs. (3)-(5) [27].

$$K = \frac{P_0 - P_c}{1 - P_c} \tag{3}$$

$$P_0 = \sum_{i=1}^{m} P_{ii} = \frac{1}{N}\sum_{i=1}^{m} n_{ii} \tag{4}$$

$$P_c = \sum_{i=1}^{m} P_{i+}P_{+i} = \frac{1}{N^2}\sum_{i=1}^{m} n_{i+}n_{+i} \tag{5}$$

where P_0 is the observed proportion; P_c is the expected proportion; n_{ii} is the total number of correctly classified points by the class along the diagonal of the error matrix; N is the total number of points checked (sampled); P_{ii} is the proportion of correctly classified sample points by the class at the diagonal of the error matrix (i.e., n_{ii}/N); P_{i+} is the marginal distribution of the sample data (n_{i+}/N where n_{i+} is the row sum by class); P_{+i} is the marginal distribution of the reference data (n_{+i}/N where n_{+i} is the column sum of class); m is the total number of classes

The Kappa statistic is more reliable than other validation techniques because it can evaluate the actual agreement and chance agreement [28]. Kappa statistic is calculated from a performance matrix obtained from the evaluation of the reference data with the predicted values.

3. Analysis of Result

3.1. Visualization of land-use change from 2014 to 2017

To visualize the changes in built-up areas over three years, a change map is created by overlaying the LULC maps of 2014 and 2017 (Fig. 3). The green color (change zone) represents cells that change over three years (2014-2017), whilst the blue and red colors represent zones that remain non-urban and urban during that time, respectively (Fig. 5).

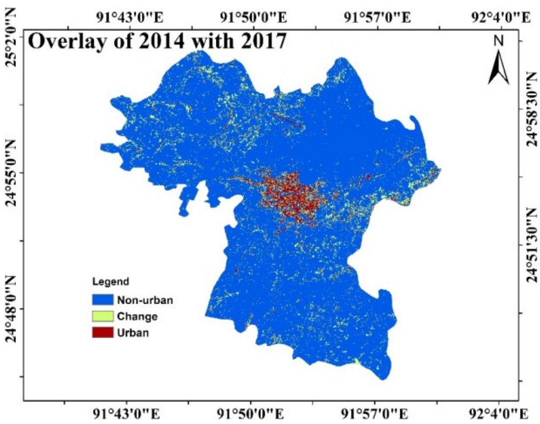


Fig. 5 Land-use change map from 2014 to 2017

3.2. Regression model results

The findings of the simulation of urban sprawl behavior are shown in Table 3. The correlation of independent variables of the regression model is examined by checking multicollinearity. The variance inflation factor (VIF) is used for checking multicollinearity. The VIF should not exceed ten as a general rule of thumb [29]. Although the VIF of the factors such as "distance to the airport" and "distance to SUST" is found higher than the threshold, these variables are kept in the model for a strong hypothesis. However, the limitation of the model, due to the potential risk of multicollinearity of the variables, is acknowledged. The multicollinearity issue is not found critical for other variables.

Table 3 OLS results of the simulation for 2014-2017

Variable	Coefficient	Standard error	t-statistics	Probability	VIF
Intercept	0.819644	0.018762	43.69	0.000000*	-
Distance to dense urban areas	-0.000280	0.000014	-19.31	0.000000*	1.84
Distance to less dense urban areas	-0.000635	0.000009	-70.58	0.000000*	1.97
Distance to Osmani international airport	0.000036	0.000001	25.76	0.000000*	12.11
Distance to major roads	-0.000009	0.000002	-5.15	0.000001*	2.20
Distance to waterways	0.000026	0.000002	13.88	0.000000*	3.88
Distance to SUST	-0.000073	0.000002	-46.80	0.000000*	11.49
Distance to Jointa-gate	-0.000018	0.000001	-36.79	0.000000*	1.80
Distance to Lalabazar point	0.000022	0.000001	19.50	0.000000*	6.31

The statistically significant factors ($p < 0.01$) are shown using an asterisk sign (*) in Table 3. The table shows that dense urban areas, which mostly include industrial and commercial centers such as Zindabazar, Ambarkhana, and Nayasharak, have an impact on urban expansion. This result reflects the study area's polycentric nature. The most striking feature is that the coefficients for the densely urbanized areas and the less densely urbanized area variables are both negative. Urban sprawl is more likely to occur in the areas adjacent to the urbanized clusters, i.e., the estimate for the variable distance to the nearest highly urbanized area is found at 0.000280, implying that increasing the distance from the urban area decreases the probability of urbanization.

For large cities, the expansion of cities tends to occur in areas that are further away from waterways. In a developing country like Bangladesh, however, the situation is the opposite due to a lack of water transportation and infrastructure development near waterways. The model shows that the variable of "distance to highways" has a substantial impact on urban development, which is clear. "Distance to existing roadways" influences human decisions on housing, employment, and education.

The model demonstrates that the variable "distance to SUST (educational area)" has a significant effect on urban development. Each year, thousands of students from all over the country come to study at SUST; most of them are non-residents of the city. As a result, shops, markets, hospital facilities, etc. have developed near SUST.

The factors "distance to Lalabazar point" and "distance to the airport" are located at the southern and northern sides of the study area respectively. They have coefficients with positive signs indicating that urban growth does not tend to occur close to those areas, whereas it is found that urban expansion tends to take place near Jointa-gate on the eastern side. This indicates an unbalanced urban expansion pattern in different directions of the city.

3.3. Cellular automata (CA) model calibration

To calibrate the CA model, the variables found to be significant in the regression model are used as the independent variables, and the dependent variables comprise urban and non-urban areas. To simulate the LULC map of 2020, the LULC map of 2017 is used as the base map, and the matrix conversion probabilities and conditional probability images from 2014 to 2017 are used as inputs.

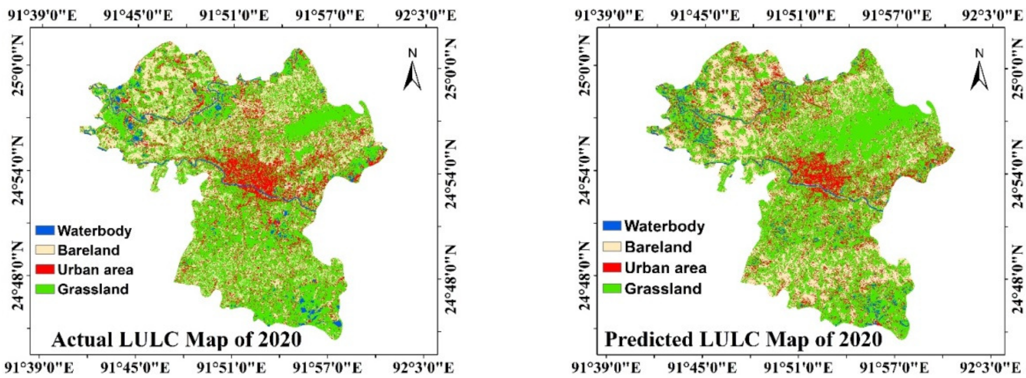
3.4. Validation

Kappa statistic is used to analyze the performance of the CA model and the consequent predicted maps used in this study. The performance matrix obtained from the comparison of the reference data with the anticipated values is used to calculate the Kappa statistic. The calculated Kappa statistic value is 0.69633 which indicates a strong agreement between the actual map and the simulated map for the year 2020 [30].

The calibration model is then validated to assess the prediction power of the calibration model (from 2014 to 2017) in forecasting future urban expansion. Therefore, the known urbanized area of 2020 is first forecasted (Fig. 6) and then validated using the reference data for 2020.

True positive (TP), false positive (FP), true negative (TN), and false negative (FN) matches of the simulated 2020 map and actual 2020 map are considered for validation. Fig. 7 presents the overlay map of the actual and simulated 2020 maps. The model prediction accuracy in validation is presented in Table 4. Here, TP denotes urban cells in both the simulated and actual 2020 maps, whereas TN denotes non-urban cells in both maps. FP represents cells that are non-urban in the actual map but urban in the simulated map, whereas FN denotes cells that are urban in the actual map but non-urban in the simulated map. Table 4 shows that 90.17 percent of the total predicted cells are correct, whereas 9.83 percent are incorrectly predicted.

The data from the actual and predicted maps of 2020 (areas in sq km) are plotted as a bar diagram (Fig. 8), and it is found that the values of areas of land-use classes in the two maps are very close. The difference between the two maps is calculated using root mean square error, and it is found to be around 3, indicating a major similarity between the actual and predicted maps.



(a) Actual LULC map of 2020 (b) Predicted LULC map of 2020

Fig. 6 Predicted LULC map of 2020 and actual LULC map of 2020

Table 4 Confusion matrix of the actual and prediction result for 2020

-		Reference data of 2020	
		Developed	Undeveloped
Simulated data of 2020	Developed	49,010 (TP)	30,382 (FP)
	Undeveloped	25,889 (FN)	467,068 (TN)

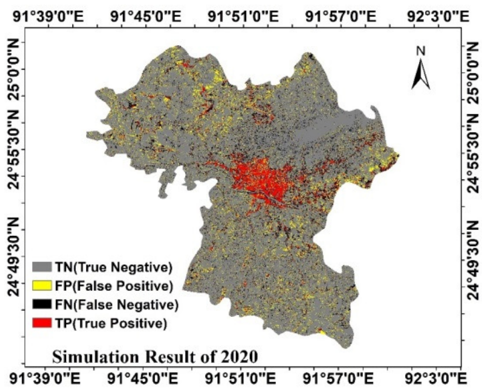


Fig. 7 LULC simulation result of 2020

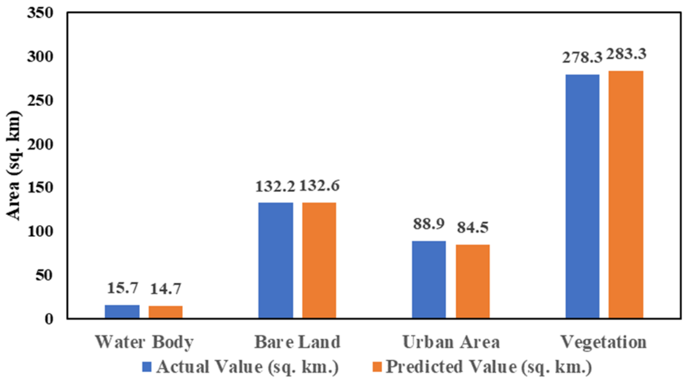


Fig. 8 Comparison between the actual value and predicted value (in terms of area)

3.5. Prediction of future urban sprawl

Calibrating the model by simulating from 2014 to 2017, urban growth prediction for the future is done using the variables mentioned in Table 3. It is assumed that the variables will remain relatively unchanged during the prediction period. The final prediction of the LULC map for 2035 is shown in Fig. 9.

Based on the actual maps, it is found that Sylhet Sadar expands approximately by 52.57% (in terms of cells) from 2014 to 2020. The analysis of the predicted map of 2035 and the actual map of 2020 in ArcGIS shows an urban expansion of 220% in terms of cells, where the cell denotes the smallest unit of information in raster data that represents a portion of the earth (Table 5).

The transition rate between distinct LULC categories is an important factor to be considered when determining the factors that have a substantial impact on LULC alterations. Fig. 10 depicts the change in LULC classes over the years. A negative value indicates a reduction in LULC classes over the period, a positive value indicates the increase in LULC classes, and zero means no change in LULC classes (e.g., urban-urban, vegetation-vegetation, etc.). It is observed that water bodies, bare land, and vegetation decrease in terms of area (square kilometers) whereas urban areas are found to increase over the years. From 2014 to 2035, the urban area expansion is 36.496%, and the vegetation, bare land, and water bodies are found to decrease by 22.3%, 13.192%, and 1.006% in terms of area respectively. The gradual alteration of different LULC classes is seen in nearby metropolitan regions, resulting in intensive change dynamics. Waterbodies, barren terrain, and vegetation next to metropolitan areas are all particularly vulnerable to infrastructure construction.

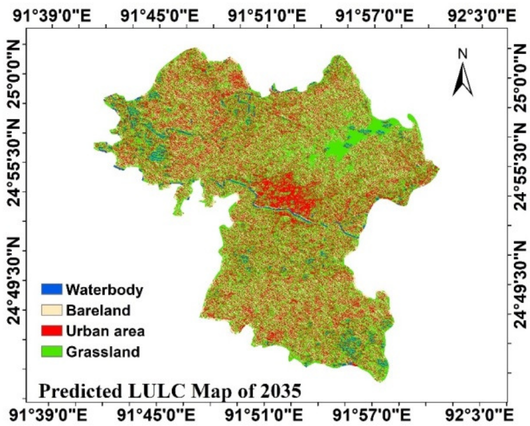


Fig. 9 Predicted LULC map of 2035

Table 5 Urban growth from 2014 to 2035 (predicted)

Year	No. of cells (urban)	Cell expansion (urban)
2014	36,487	-
2020	55,670	19,183 (52.57%) (2014-2020)
2035	178,353	122,683 (220%) (2020-2035)

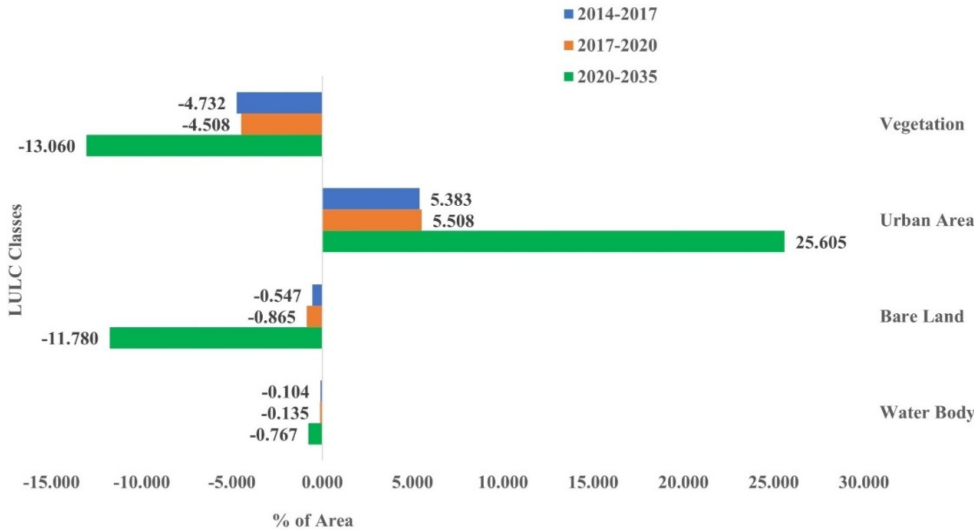


Fig. 10 Change of LULC classes (in terms of area)

4. Discussion

Based on the developed model, the urbanization process is found to be influenced by land-use variables. An econometric variable such as "distance to an educational zone" is added to the model to get a better understanding of the urbanization process. The results indicate that urbanization is happening in all directions of the city, but the pattern is unbalanced.

The urbanization forecasting for 2035 provides a useful insight for planners and policymakers. The growth of Sylhet city by 2035 is a consequence of urban sprawl as urbanization is supposed to expand adjacent to the neighboring urban cluster. Therefore, urban planning and development measures can be considered to control the urban expansion of Sylhet city. Considering the predictive models in the planning stage will ensure effective planning for sustainable urban development. The findings of this study may help urban planners to visualize future urban expansion of the city and to take necessary measures accordingly.

The prediction map also provides insights into the LULC transformation process. From 2020 to 2035, vegetation is predicted to decrease by 13.06%, bare land by 11.78%, and water bodies by 0.76%. However, the urban area is predicted to increase by 25.605%. It is clearly found that waterbody, bare land, and vegetation experience negative growth; hence, positive growth is seen in urban classes. The intense change dynamics are noticed in neighboring urban areas with the steady transformation of different LULC classes. The waterbody, bare land, and vegetation, adjacent to the urban areas, are highly susceptible to transformation by the construction of infrastructure.

Urban expansion is correlated with many dynamic phenomena, and therefore cannot be predicted with 100% accuracy. The limitation of this study is the capacity to identify whether urbanization takes place toward the city or away from the city. Also, the lack of availability of satellite images of proper quality makes it compulsory to use a small interval of input data (satellite data from 2014 to 2017), which places a constraint on the rationality of this study.

The images used in this study are Landsat 8 images. These data files are twice as large as those from previous Landsats because it collects more precise data and includes observations from two new spectral bands. This results in crisper images with colors that do not saturate. However, Landsat 8 images of the study area before 2014 are not available. Therefore, in order to have a more precise classification of images, the input data starting from 2014 are used here. This study only focuses on the pixel change in determining the urban sprawl. The dispersion or concentration of pixels is not considered here and can be addressed in the future study. Also, the accuracy assessment per land-cover class and sensitivity is not considered in this study.

5. Conclusions

A model was developed in this study for predicting the future urban sprawl of Sylhet city. The independent variables used in the model were found statistically significant at a 99% confidence interval, which implies that all the variables estimated have a significant contribution to the urban sprawl of Sylhet city. The land-use change predicted for the year 2035 gives important insight to policymakers and planners. The predicted growth of Sylhet city indicates urban sprawl as urban areas are expanding adjacent to the neighboring urban cluster.

Sylhet, one of the largest cities in Bangladesh, receives significant attention from the Bangladesh government. Therefore, integrating the predictive model into city planning will convert the present typical planning into state-of-the-art planning ensuring sustainable development. The findings of this study will help the urban planners and policymakers in foreseeing future urban expansion to take the appropriate planning and policy steps. These study findings will also give a signal to the planners and policymakers in Sylhet and other major cities in Bangladesh which requires immediate steps to control the persistent urban sprawl. Otherwise, the uncontrolled urban sprawl will put intense pressure on transport, infrastructure, housing, and the environment.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Z. Shao, et al., "Urban Sprawl and Its Impact on Sustainable Urban Development: A Combination of Remote Sensing and Social Media Data," Geo-Spatial Information Science, vol. 24, no. 2, pp. 241-255, July 2021.
[2] T. Firman, "Land Conversion and Urban Development in the Northern Region of West Java, Indonesia," Urban Studies, vol. 34, no. 7, pp. 1027-1046, June 1997.
[3] A. M. Youssef, et al., "Integrated Evaluation of Urban Development Suitability Based on Remote Sensing and GIS Techniques: Contribution from the Analytic Hierarchy Process," Arabian Journal of Geosciences, vol. 4, no. 3, pp. 463-473, February 2010.
[4] T. Angotti, Metropolis 2000: Planning, Poverty, and Politics, 1st ed., England: Routledge, 1993.
[5] B. Batsuuri, et al., "Estimating the Impact of Urban Planning Concepts on Reducing the Urban Sprawl of Ulaanbaatar City Using Certain Spatial Indicators," Land, vol. 9, no. 12, Article no. 495, December 2020.
[6] B. Jiang, et al., "Geospatial Analysis and Modeling of Urban Structure and Dynamics: An Overview," GeoJournal Library, vol. 99, pp. 3-11, 2010.
[7] M. H. Bazai, et al., "Assessment of Urban Sprawl and Land Use Change Dynamics through GIS and Remote Sensing in Quetta, Balochistan, Pakistan," Journal of Geography and Social Sciences, vol. 2, no. 1, pp. 31-50, June 2020.
[8] X. Yang, et al., "Modelling Urban Growth and Landscape Changes in the Atlanta Metropolitan Area," International Journal of Geographical Information Science, vol. 17, no. 5, pp. 463-488, July 2003.
[9] P. H. Verburg, et al., "Dynamic Simulation of Land-Use Change Trajectories with the CLUE-s Model," GeoJournal Library, vol. 90, pp. 321-335, 2007.
[10] M. J. B. Alam, et al., "Unplanned Urbanization: Assessment through Calculation of Environmental Degradation Index," International Journal of Environmental Science and Technology, vol. 3, no. 2, pp. 119-130, April 2006.
[11] S. Gupta, et al., "Analysis of Impervious Land-Cover Expansion Using Remote Sensing and GIS: A Case Study of Sylhet Sadar Upazila," Applied Geography, vol. 98, pp. 156-165, June 2018.
[12] United Nations Population Fund, "Independent Country Programme Evaluation Annexes Bangladesh," https://www.unfpa.org/sites/default/files/admin-resource/Bangladesh_CPE_Final_Report_Annexes.pdf, 2016.
[13] Bangladesh Bureau of Statistics (BBS), "Bangladesh Bureau of Statistics (BBS): Area, Population and Literacy Rate by Paurashava—2001," http://203.112.218.65:8008/WebTestApplication/userfiles/Image/LatestReports/YB2011.pdf, September 2012.
[14] S. Leyk, et al., "Two Centuries of Settlement and Urban Development in the United States," Science Advances, vol. 6, no. 23, pp. 1-12, June 2020.

[15] D. Aldogom, et al., "Multi Temporal Satellite Images for Growth Detection and Urban Sprawl Analysis; Dubai City, UAE," Remote Sensing Technologies and Applications in Urban Environments IV, vol. 11157, pp. 71-81, October 2019.

[16] Q. He, et al., "Comparison of Urban Growth Patterns and Changes between Three Urban Agglomerations in China and Three Metropolises in the USA from 1995 to 2015," Sustainable Cities and Society, vol. 50, Article no. 101649, October 2019.

[17] M. H. Rahman, "Unplanned Urbanization and Hill Cutting: A Study on Environmental Change in Sylhet," Brac University Journal, vol. 8, no. 1-2, pp. 13-21, 2011.

[18] A. Haque, et al., "Study on Land Use Pattern Change and Its Causes," International Journal of Environmental Research, vol. 2, no. 2, pp. 199-204, April 2008.

[19] R. Hamad, "A Remote Sensing and GIS-Based Analysis of Urban Sprawl in Soran District, Iraqi Kurdistan," SN Applied Sciences, vol. 2, no. 1, pp. 1-9, December 2019.

[20] I. Santé, et al., "Cellular Automata Models for the Simulation of Real-World Urban Processes: A Review and Analysis," Landscape and Urban Planning, vol. 96, no. 2, pp. 108-122, May 2010.

[21] A. A. Al-Sharif, et al., "Monitoring and Predicting Land Use Change in Tripoli Metropolitan City Using an Integrated Markov Chain and Cellular Automata Models in GIS," Arabian Journal of Geosciences, vol. 7, no. 10, pp. 4291-4301, October 2014.

[22] C. Losiri, et al., "Modeling Urban Expansion in Bangkok Metropolitan Region Using Demographic-Economic Data through Cellular Automata-Markov Chain and Multi-Layer Perceptron-Markov Chain Models," Sustainability, vol. 8, no. 7, Article no. 686, July 2016.

[23] L. Sang, et al., "Simulation of Land Use Spatial Pattern of Towns and Villages Based on CA-Markov Model," Mathematical and Computer Modelling, vol. 54, no. 3-4, pp. 938-943, August 2011.

[24] A. Khan, et al., "Machine Learning-Based Monitoring and Modeling for Spatio-Temporal Urban Growth of Islamabad," The Egyptian Journal of Remote Sensing and Space Science, vol. 25, no. 2, pp. 541-550, August 2022.

[25] S. C. K. Tekouabou, et al., "Reviewing the Application of Machine Learning Methods to Model Urban Form Indicators in Planning Decision Support Systems: Potential, Issues and Challenges," Journal of King Saud University—Computer and Information Sciences, in press.

[26] M. Bosch, et al., "Addressing Urban Sprawl from the Complexity Sciences," Urban Science, vol. 3, no. 2, Article no. 60, June 2019.

[27] Z. Ma, et al., "Tau Coefficients for Accuracy Assessment of Classification of Remote Sensing Data," Photogrammetric Engineering and Remote Sensing, vol. 61, no. 4, pp. 435-439, 1995.

[28] T. Fung, et al., "The Determination of Optimal Threshold Levels for Change Detection Using Various Accuracy Indices," Photogrammetric Engineering and Remote Sensing, vol. 54, no. 10, pp. 1449-1454, 1988.

[29] D. C. Montgomery, et al., Introduction to Linear Regression Analysis, 3rd ed., New York: John Wiley & Sons, 2001.

[30] J. R. Landis, et al., "The Measurement of Observer Agreement for Categorical Data," Biometrics, vol. 33, no. 1, pp. 159-174, March 1977.