Journal of Urban Management 14 (2025) 451–467

Contents lists available at ScienceDirect

Journal of Urban Management

journal homepage: www.elsevier.com/locate/jum

Research Article

Assessment of urban expansion susceptibility in major urban units of Bangladesh leveraging machine learning and geostatistical approach

Mafrid Haydar *, Sakib Hosan, Al Hossain Rafi

Department of Urban and Regional Planning, Khulna University of Engineering & Technology (KUET), Khulna, 9203, Bangladesh

ARTICLE INFO

Keywords:
Urban expansion
Machine learning
Geostatistical modeling
Data-driven modeling
Remote sensing

ABSTRACT

This study examines the governing factors and susceptibility zones for urban expansion in Bangladesh's major urban areas, including Dhaka, Barisal, Chittagong, Comilla, Narayanganj, Gazipur, Khulna, Sylhet, Rajshahi, Rangpur, and Mymensingh. The main goals of this research are to determine the impact of governing factors and to identify susceptibility zones for urban expansion in major urban units using a data-driven approach. By using governing factors (DEM, Slope, LST, NDVI, Population, distance to (industry, growth center, settlement, facilities, waterbody, road), and machine learning (Random Forest) and geostatistical approach (Binary Logistic Regression), the research identifies the most important factors influencing urban expansion, including NDVI, LST, waterbodies, roads, and settlements. The RF model's ROC-AOC values showed the highest accuracy (1.00) in Comilla and Mymensingh, moderate accuracy (0.99) in Barisal, Chittagong, Narayanganj, Gazipur, Khulna, and Rajshahi, and lower accuracy in Dhaka (0.98), Sylhet (0.89), and Rangpur (0.85). For the Binary Logistic Regression model, Comilla, Narayanganj, Gazipur, and Mymensingh had the best fit (Nagelkerke $R^2 = 1.00$), while Sylhet had the lowest significance (0.482). Furthermore, Khulna, a major urban unit, is the highest urban expansion susceptibility zone which is 35.72%. Rajshahi and Barisal are the moderate and low urban expansion susceptibility where 83.17% and 0.88% respectively. This unplanned and rapid urban expansion zone has also confronted policymakers and planners with an insurmountable challenge and stressed local governments' ability to manage and use their scarce land-based resources with geospatial data. Thus, this study's machine learning and geostatistical findings will help explain land cover change and urban expansion in Bangladesh's eleven metropolitan areas. This study will improve urban development understanding in Bangladesh. Findings will help planners, stakeholders, and policymakers understand urban expansion patterns, enabling better environmental planning.

1. Introduction

Urban expansion (UE), indicating the conversion of agricultural or natural terrains into developed areas, is a major environmental issue stemming from urban growth. Furthermore, land use change is perceived as the outcome of the interaction between human and natural systems, and the proliferation of urban areas has emerged as a global phenomenon (Gao et al., 2014). The conversion of natural habitat to urban land uses, caused by this rapid urbanization, can substantially degrade biodiversity (Anderson et al., 2013; Foley et al.,

* Corresponding author.
E-mail addresses: mafrid@urp.kuet.ac.bd (M. Haydar), sakibkuet19@gmail.com (S. Hosan), alhossainrafi@gmail.com (A.H. Rafi).

https://doi.org/10.1016/j.jum.2024.11.011
Received 6 May 2024; Received in revised form 10 October 2024; Accepted 25 November 2024
Available online 6 December 2024

2005; Hochstrasser, 2007). Urban development, traditionally seen as an indicator of regional economic success, has resulted in various environmental problems due to human activities and changes in land use. Deforestation, urban development, and fragmented landscapes resulting from urbanization can exacerbate soil erosion and disrupt climate control (M. Rahman et al., 2024; Rimi et al., 2024, pp. 1–14). These activities can also contribute to reduced food availability, carbon sequestration, soil water storage, and biodiversity (Huilei et al., 2017; Pullanikkatil et al., 2016). Many American cities have grown a lot since the middle of the 20th century and it's mostly due to more people moving into cities and out to the neighborhoods (Gao et al., 2014). China's rapid urbanization and dramatic urban landscape changes have affected the economy and environment (Yue et al., 2014). Rapid population, industrial, and commercial growth resulted in the urbanization of 41 Indonesian cities (Olivia et al., 2018). In the last few decades, both the size and number of people living in cities in developing countries, especially in South Asia, have grown dramatically. Urbanization in developing countries like India is linked to high population increase, which in turn leads to several socioeconomic and environmental issues (Capps et al., 2016). Furthermore, urban land has expanded while agricultural, grassland, and barren lands have been displaced in Pune, India (Kantakumar et al., 2016). Bangladesh is witnessing fast urbanization, with the capital city Dhaka being the 11th largest megacity, seeing urban growth of 81.54% (Morshed et al., 2017). In response to this growth, urban areas are spreading both vertically and horizontally, exerting strain on natural resources by displacing marshes, flora, and agricultural land (Roy et al., 2020).

There are various model-based methodologies available to evaluate urban expansion. Scholars (Sun et al., 2018) used a land use simulation and Markov-logistic-CA model for urban expansion in the USA. Time-series spatial data analysis has also been applied to research examining the evolution of particular land uses or land-use systems (Fuchs et al., 2015; Gerard et al., 2010; Hatna & Bakker, 2011 & Levers et al., 2018). A limited number of scientists offer illustrations of a multi-disciplinary methodology to investigate the factors influencing land-use conversion on a national or regional scale in Europe as well as worldwide (Verburg et al., 2004; Chakir & Parent, 2009; Jiang et al., 2012 and Mazzocchi et al., 2013). Researchers have used remote sensing and geographic information system (GIS) tools to get information on how cities are growing and changing land use and also, and they have also looked at how this affects green space, urban heat islands (UHIs), and the quality of the air and water in Chinese towns (Hua et al., 2008; Yunhao et al., 2003 & Zhang et al., 2010). Besides, the assessment of urban expansion can be done using several models in India like logistic regression (Sarkar & Chouhan, 2020), remote sensing, and GIS (Gibson et al., 2015; Kantakumar et al., 2016), etc. Urban expansion has been complicated by environmental, economic, and infrastructural issues. Here governing factors were chosen DEM, slope, LST, NDVI, population, and distance to key urban elements like industry, growth centers, settlements, facilities, waterbodies, and roads for this study. These variables directly or indirectly affect land use changes, making them crucial to understanding urban expansion in Bangladesh. The formation of DEM and Slope is crucial for understanding the topographical limitations that influence urban development. Steeper gradients or elevated terrains generally prevent construction, rendering flatter regions more amenable to development (Ma & Xu, 2010). Land Surface Temperature (LST) indicates heat retention in urbanized regions, correlating with developed areas and serving as a significant metric of urbanization intensity (Wan et al., 2002). NDVI improves the evaluation of vegetation cover, with diminishing vegetation frequently indicating escalating urban sprawl, so underscoring the environmental deterioration resulting from growth. Urban expansion depends on population density, which increases housing, infrastructure, and service demand. The research shows that urban regions with higher population densities have higher land resource pressures and urbanization rates (Seto et al., 2012). Industrial, growth, and facility proximity shows urban areas' economic pull, where jobs and services drive urban migration and development (Pradhan et al., 2017). The distance from to roads and water bodies affects accessibility and defines the natural limits of urban expansion (Habib et al., 2020). Distance to roadways enhances the likelihood of development owing to increased connectivity, however water bodies frequently act as natural barriers that influence urban expansion patterns. In Bangladesh, several studies have been conducted with just the one goal of investigating the impacts of urban expansion such as land surface temperature in Chattogram Metropolitan Area (Roy et al., 2020), LULC with driving factors for sustainable development in Gazipur (Arifeen et al., 2021), measuring urban expansion pattern using spatial matrices in Khulna city (Alam et al., 2023) and using support vector machine in Chattogram (Bajracharya & Sultana, 2022). Moreover, changing patterns of land use and land cover in Bangladesh as a result of urban expansion is linked to a variety of factors, including agriculture, build-up, water, vegetation, fellow land, low land, etc (Arifeen et al., 2021). When evaluating urban expansion, factors such as land use, land conversion, digital elevation models (DEMs), climate data, soil qualities, and crop production are considered, as suggested by (Sun et al., 2018). Furthermore (Ma & Xu, 2010), created a model to assess the susceptibility of urban expansion by considering factors such as build-up area, water body, vegetation, and dry land. Previous research (Bajracharya & Sultana, 2022) used population density, slope, road, commercial, rail, pond, and river distances, distance to forest, neighborhood urban cells, vegetation cells, agriculture cells, water cells, and barren cells. to create a support vector machine for future urban expansion to predict susceptibility.

There is no research that has been conducted in Bangladesh that takes into consideration the expansion of urban areas in major urban units. The focus of this research is on determining the degree to which districts that contain city corporations (large urban entities) are susceptible to urban expansion. Numerous studies have been conducted on urban expansion utilizing different metrics to understand and assess the processes and functions of cities. Nevertheless, previous researches have neither examined urban expansion through the prism of governing factors, nor has it employed data-driven models to evaluate susceptibility to urban expansion. In addition, there has never been any study on urban expansion in Bangladesh that has been undertaken using a data-driven machine learning method, although this research has been done on data driven approach. Also, no research found on governing factors (DEM, slope, Land surface temperature, NDVI, Population, distance to industry, (distance to growth center, settlement, facilities, waterbody, road) on urban expansion susceptibility in Bangladesh. Subsequently, there are significant gaps in the research, but this study can adequately fill each of those gaps to predict the susceptibility of urban expansion in Bangladesh. Thus, the main aim of this research is to use machine learning (ML) methods to identify the urban expansion susceptibility zones in Bangladesh's key metropolitan units. To address the research gap, the objectives can be evaluated; (i) to examine the influence of governing factors in major urban units in Bangladesh; (ii) to assess the significance of governing factors in urban expansion; (iii) to use a data-driven approach to identify urban expansion susceptibility zones in Bangladesh. Urban planners must prioritize the major urban units of Bangladesh and this study's findings and scientific comprehension will provide useful insights into the aspects contributing to urban expansion susceptibility in Bangladesh. Furthermore, it will offer solutions to mitigate the potential incidence of future urban expansion.

2. Methods and materials

2.1. Description of the study area

The study area included eleven major districts in the country of Bangladesh, which is located between 20°34′ and 26°38′ North Latitude and between 88°01′ and 92°41′ East Longitude (Bangladesh Ministry of Foreign Affairs, 2017). The area of the districts is respectively 1464 km^2, 2785 km^2, 5283 km^2, 3146 km^2, 684 km^2, 1806 km^2, 4394 km^2, 3452 km^2, 2425 km^2, 2401 km^2, and 4395 km^2 and the density of the population according to the census 2022 is respectively 10067/km^2, 923/km^2, 1736/km^2, 1974/km^2, 5712/km^2, 2914/km^2, 595/km^2, 1117/km^2, 1202/km^2, 1320/km^2, and 1342/km^2 for Dhaka, Barisal, Chittagong, Comilla, Narayanganj, Gazipur, Khulna, Sylhet, Rajshahi, Rangpur, and Mymensingh (Bangladesh Bureau of Statistics, 2022). In Bangladesh, city corporations have a significant influence on district area development that drives economic growth via investment, industry development, and employment

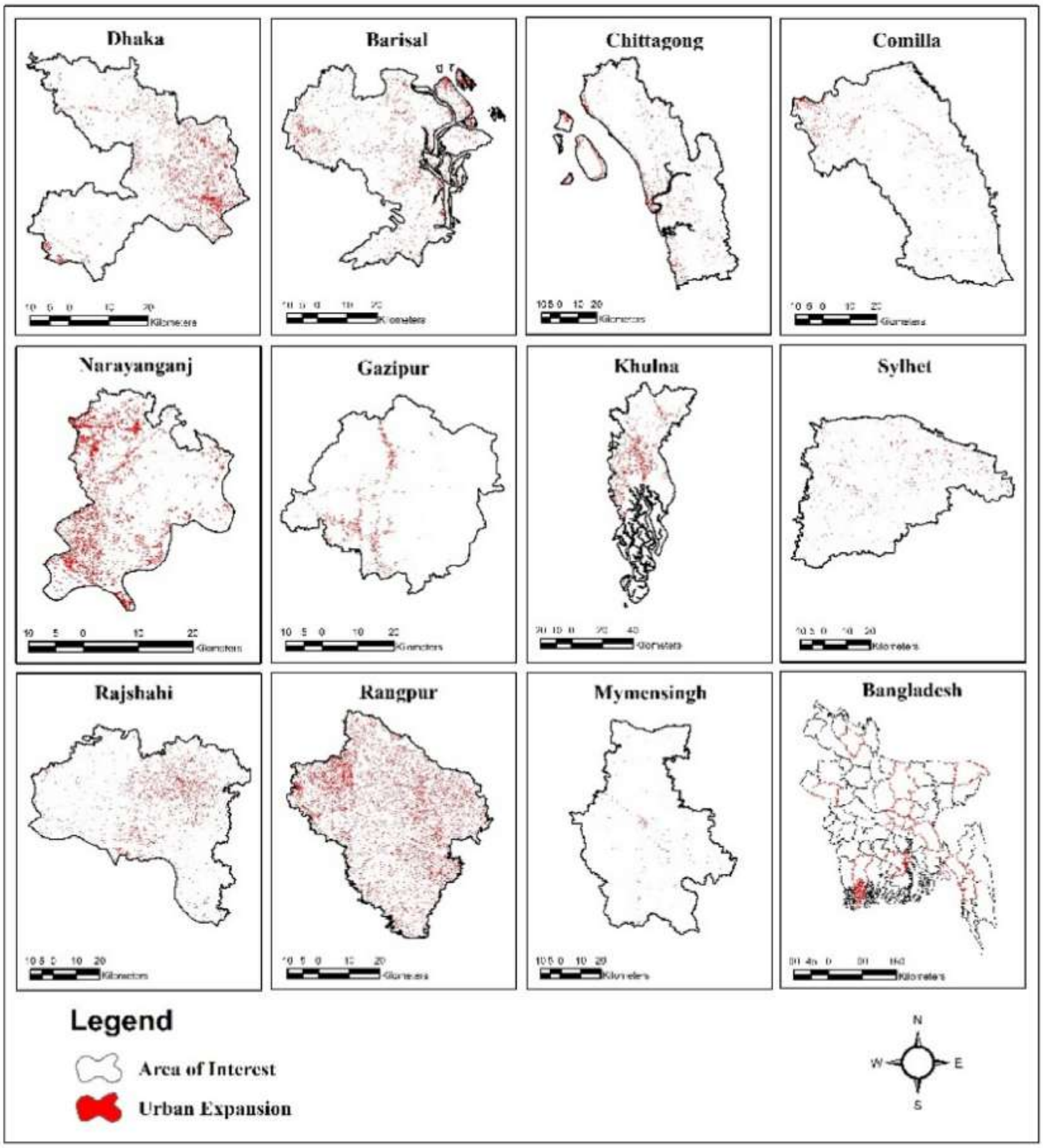

Fig. 1. Research location and urban expansion inventory.

generation and are also responsible for managing essential services like waste management, transportation, sanitation, and water supply, all of which affect residents quality of life (Zafarullah & Ferdous, 2024). City corporations also provide health, education, and social welfare services that have an impact on residents in nearby communities in addition to city dwellers (Zafarullah & Ferdous, 2024). They serve as the national hub for overseeing transportation, education, and other infrastructure in the country's largest cities. People may be able to meet their needs for a higher standard of living as cities expand. These districts significantly influence the growth and development of Bangladesh. The districts show the urban expansion from 2018 to 2023 (Fig. 1). In Narayanganj and Rangpur, there is rapid urban growth due to population influxes. The excellent roads and services in these locations support the expansion where the expansion of cities like Chittagong, Comilla, Sylhet, and Mymensingh is slower due to several factors, including greater concern for environmental preservation, fewer employment possibilities, and inadequate infrastructure (Fig. 2).

2.2. Description of data

There were thirteen parameters used to find the urban expansion susceptibility (Table 1), including the Digital Elevation Model (DEM), Land Surface Temperature (LST), Land Use and Land Cover (LULC), Normalized Difference Vegetation Index (NDVI), Regional Facilities, Growth Centre, Industry, Population, Precipitation, Road, Settlement, Slope, and Waterbody. A quantitative representation of the Earth's surface, the Digital Elevation Model (DEM) offers fundamental details about the relief of the terrain that influences urban expansion by providing detailed elevation data, including trees, buildings, and any other surface objects (USGS, 2022). We extracted DEM data from SRTM Datasets. An essential component of biology and climate, land surface temperature (LST) influences animals and ecosystems on a local to global scale due to increased absorption and retention of heat by built surfaces, reduced vegetation cover, heat emissions from human activities, and altered wind patterns that affect urban expansion (Liu et al., 2022). LST data extracted from LANDSAT – 8 image datasets. The classification of natural features and human activities on the landscape during a certain period using accepted scientific and statistical methods of analysis of relevant source materials is known as land use/land cover, or LULC which impacts urban expansion because if natural resources decrease, it means urban expansion increases (Singh, 2013). For LULC data, firstly, obtained satellite imagery and ground truth data and then, corrected image distortions. Next, extracted features and apply a classification algorithm to assign land cover classes. After classification, assessed accuracy using validation data and extracted LULC. By using the Google Earth Engine. The Normalized Difference Vegetation Index (NDVI) quantifies vegetation by measuring the difference

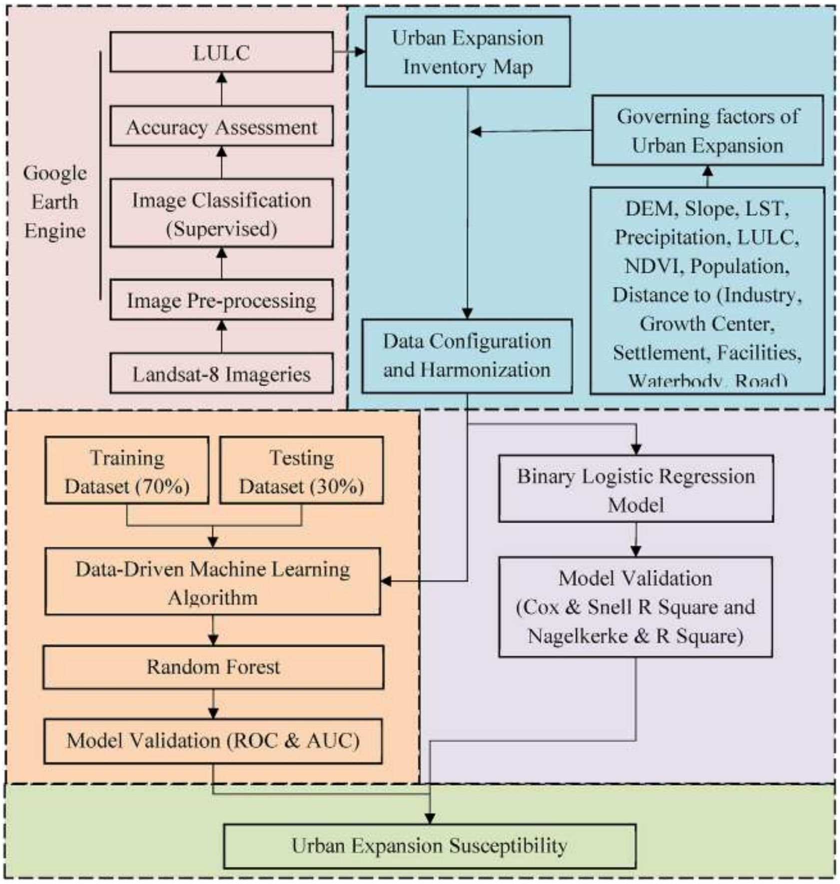

Fig. 2. Methodological framework.

Table 1
Data sources of urban expansion parameters for urban expansion susceptibility.

Description of parameters	Sources	Resolution	Data Type	Year
DEM	SRTM Datasets	30 × 30 m	Raster	2023
LST	LANDSAT – 8	30 × 30 m	Raster	2023
LULC	LANDSAT – 8	30 × 30 m	Raster	2018 & 2023
NDVI	LANDSAT – 8	30 × 30 m	Raster	2023
Precipitation	CHIRPS Datasets	30 × 30 m	Raster	2023
Slope	Extracted from DEM	30 × 30 m	Raster	2023
Regional Facilities	LGED Datasets		Vector	2023
Growth Center	LGED Datasets		Vector	2023
Industry	OpenStreetMap		Vector	2023
Population	BBS (2022)		Vector	2022
Waterbody	Humanitarian Datasets		Vector	2023
Road	LGED Datasets		Vector	2023
Settlement	Humanitarian Datasets		Vector	2023
Urban Expansion Inventory	Extracted from LULC data	30 × 30 m	Raster	

between near-infrared (which vegetation strongly reflects) and red light (which vegetation absorbs) (GISGeography, 2021). Degradation of vegetation caused by extensive urban expansion increases the decline of many natural habitats (K. Yang et al., 2021). NDVI data was also extracted from LANDSAT – 8 image datasets. Regional facilities include the supply of basic utilities and services necessary for an area to function and develop. In the study, the use of educational facilities and medical facilities as regional facilities, are extracted from LGED datasets. The growth center acts as the hub for social, cultural, and economic activity (Pradhan et al., 2017). It impacts urban expansion by transforming natural areas into places used for residential, commercial, and industrial purposes (Mondal & Das, 2010). Growth center data was also extracted from LGED datasets. Industry is a collection of profitable businesses or organizations that manufacture or provide goods, services, or sources of income and drive urban expansion by creating employment opportunities, stimulating economic growth, and necessitating the development of supporting infrastructure and land use changes (Gorton, 2023). We extracted industry data from OpenStreetMap. Population refers to the total number of individuals or residents in the area that impact urban expansion by increasing the demand for housing, infrastructure, and services (Momoh, 2022). Precipitation refers to any liquid or frozen water that condenses in the atmosphere and descends to Earth, which is essential for maintaining human existence, promoting industrial development, and ensuring ecological security, which impacts urban expansion (Trenberth, 2011). Rapid urban expansion is an important period for developing countries due to the growing economy as well as the road network expansion, where a road is a travel path used by humans, animals, and vehicles (G. Shi et al., 2019). We extracted road data from LGED datasets. Settlement refers to a community or group of people living together in a particular area that occurs through the transformation of physical landscapes and ecological systems into developed land, which impacts urban expansion (Leyk et al., 2020). We extracted settlement data from humanitarian datasets. The measurement of a line's steepness and direction is called its slope, which determines whether the lines are parallel, perpendicular, or none at all, which helps to find urban expansion susceptibility (K. Shi et al., 2023). We extracted slope data from DEM. A waterbody is any significant accumulation of water, generally on a planet's surface, that impacts urban expansion by influencing land use patterns, serving as natural barriers, shaping infrastructure development, and offering economic opportunities (Habib et al., 2020). We extracted waterbody data also from humanitarian datasets. There several processes are required to use Land Use/Land Cover (LULC) data to assess urban expansion between 2018 and 2023. Urban expansion indicates the transformation of non-urban areas, including vegetation and water bodies, into urbanized regions driven by causes such as industry, population growth, infrastructure, settlements, growth centers, and regional facilities. This research calculated urban expansion using raster datasets of land use and land cover (LULC) from 2018 to 2023. Initially, we identified the raster cells corresponding to urban areas for both 2018 and 2023. Subsequently, the raster datasets were evaluated using a raster calculator to subtract the 2018 Land Use and Land Cover (LULC) from the 2023 LULC, thereby pinpointing locations where non-urban land transitioned to urban use. This technique facilitated the identification and quantification of urban expansion by establishing an inventory of land cover alterations. Supplementary geographical and geostatistical analyses were conducted to assess the total area and patterns of urban expansion, providing insights into urban sprawl and aiding sustainable urban planning initiatives.

2.3. Analytical methods

2.3.1. Machine learning algorithm: random forest

Machine learning, a facet of artificial intelligence, employs algorithms trained on datasets to automatically generate models capable of predicting outcomes and classifying data without direct human intervention (Coursera Staff, 2024). Supervised, semi supervised, unsupervised, and reinforcement machine learning are among the several varieties available (Sarker, 2021). To make predictions based on known instances, we used Random Forest, a supervised machine-learning algorithm that learns from labeled data (Sarker, 2021). It constructs an effective ensemble framework by combining many decision trees (Sruthi E R, 2024). It specifies the classification task with two class labels, such as "yes and no" or "true and false," where a class may represent the normal condition and a different class the abnormal one (Ha et al., 2011). For instance, the normal state of the task is "Area has expanded," while the abnormal state is "Area has not expanded". With a special focus on the random forest approach, we used the Python programming language to explore the field of machine learning. Python is an effective tool for developing machine-learning algorithms because of its large library and simple syntax.

Table 2
The scale for AUC values (Metz, 1978).

AUC Values	Test Quality
0.91–1.00	Excellent
0.81–0.90	Very Good
0.71–0.80	Good
0.61–0.70	Satisfactory
0.51–0.60	Unsatisfactory

On the other hand, Jupyter Notebook offers an interactive environment that is suitable for data exploration, code execution, and result visualization. For data exploration, the data has been divided into two categories: testing and training sets, with 70% for testing and 30% for training. Considering the training data, the model generates a random forest classifier. Next, generate predictions according to the test data, and then utilize the ROC (Receiver Operating Characteristics) curve and AUC (Area Under the Curve) score to determine the classifier's accuracy (Table 2). A higher AUC indicates greater performance and an ROC curve touching the top-left corner characterizes a perfect classifier (Evidently AI, 2020). The Random Forest approach was employed to forecast the susceptibility to urban expansion. The Random Forest machine learning method employs 100 estimators. The dataset is partitioned into two subsets: 70% allocated for training and the remaining 30% reserved for testing. A state coded 42 was selected at random for this procedure.

This work employed the Random Forest classifier with Python's scikit-learn module (Weiss, 2012), which offers a comprehensive array of machine learning techniques. Hyperparameter improving was conducted by grid search with cross-validation to optimize the model. The parameters modified to optimize performance were the number of trees (n_estimators), the maximum depth of trees (max_depth), and the minimum samples necessary to divide an internal node (min_samples_split). Cross-validation was utilized to avert overfitting and guarantee the model's effective generalization to novel data (Bickel et al., 2009). We employed the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) score for model performance assessment, ROC and AUC are calculated from training and testing dataset. The ROC curve illustrates the true positive rate in relation to the false positive rate, while the AUC signifies the probability that the model prioritizes a randomly selected positive instance above a randomly selected negative one (Richardson et al., 2024). An elevated AUC value (close to 1) indicates superior performance. Furthermore, accuracy, precision, recall, and the F1-score were computed to deliver a thorough assessment of the model's performance across many dimensions (Yacouby, 2020).

If the AUC is near to 1, the model performs excellently, clearly distinguishing between positive and negative cases, if the AUC is around 0.5, the model isn't performing effectively and is almost like making random guesses. if the AUC is close to 0, the model is doing poorly, often confusing positive cases with negative one (Prasanna, 2024).

2.3.2. *Binary logistic regression model*

Binary Logistic Regression predicts urban development or non-expansion based on a set of independent variables such as DEM, LST, NDVI, among others. The binary logistic regression model assumes independent observations, linearity, and no perfect multicollinearity. Descriptive statistics, assumptions checked and met, model significance, and model fit are part of binary logistic regression model reporting. Also, the model employs a logit link function to represent the log-odds of the dependent variable as a linear combination of the independent variables (Harris, 2021, pp. 1–7). To make meaningful inferences, logistic regression model assumptions must be addressed. First, logistic regression requires linearity between dependent and independent variable log-odds. We checked interaction and higher-order terms for non-linearity to test this assumption. Second, predictor multicollinearity can skew model coefficients and statistical significance. Multicollinearity was checked using variance inflation factor (VIF) to ensure no predictor variables were substantially linked (Statistics, 2024). The residuals confirmed that binary logistic regression implies error independence. A binary logistic regression model describes how a collection of independent variables and a dependent variable relate to each other (UCLA, 2020). In this instance, urban expansion serves as the dependent variable (response), while the other variables (DEM, LST, LULC, NDVI, etc.) are our independent variables (predictors). We used the Statistical Package for the Social Sciences (SPSS) software to analyze the relationship between urban expansion and several independent factors using a binary logistic regression model. We build a logistic regression model to predict the chance of urban expansion based on the values of the independent variables. SPSS simplifies this process by providing an intuitive interface for variable selection, model fitting, a nd parameter evaluation. After establishing the model, we evaluated its performance using statistical metrics, specifically focusing on the Cox & Snell R Square and Nagelkerke R Square, which offer a summary of the model's fit to the data. An improved variant of the Cox & Snell R Square, Nagelkerke's R Square, shows a better fit to the data with a higher value (Patel, 2021). The model range of values is 0%–100%, with values below 25% indicating the level of explanation is zero, which means there are no relationships between the independent and dependent variables. 25%–50% indicate a weak explanation, 50%–75% indicate a moderate explanation, 75%–99% indicate a strong explanation, and 100% indicate a perfect explanation, which means a strong fit (UCLA, 2020). We acquired Cox & Snell R Square and Nagelkerke R Square values as measures of the model's goodness-of-fit after doing logistic regression analysis in SPSS. After calculation, the Cox & Snell R Square value, statistic shows the percentage of the dependent variable's volatility that the model can account for. Additionally, the Nagelkerke R Square value offers an adjusted measure of explained variance in the dependent variable.

The comprehensive calibration of Random Forest and the compliance with logistic regression assumptions augment the validity of the findings. The integration of machine learning and statistical models enables a comprehensive analysis of the determinants of urban growth, while reducing biases and providing dependable predictions.

3. Results

3.1. Description of parameters

Urban hubs pulsate with life in Dhaka, the bustling capital, and neighboring Narayanganj, where commerce and culture intertwine in a vibrant tapestry of urban activity and the districts like Barisal, Chittagong, Rajshahi, and Khulna, verdant vegetation landscapes dominate, providing pockets of green amidst urban sprawl and serving as essential ecological havens (Fig. 3(a)). Dhaka, Chittagong, and Khulna have emerged as vital economic arteries, drew investments, and nurtured urban expansion, while districts like Mymensingh, Rangpur, and Sylhet exhibit fewer growth centers, reflecting nuanced economic landscapes (Fig. 3(b)). In Dhaka, Gazipur, and Narayanganj, population densities paint a vivid picture of urban dynamics and demographic pressures, with bustling city centers pulsating with energy and activity, while districts like Barisal, Rajshahi, and Khulna accommodate smaller populations, fostering a more tranquil pace of life (Fig. 3(c)). The lush greenery of Barisal, Rajshahi, and Khulna contrasts with the concrete jungle of Dhaka and Narayanganj, offering a visual testament to the diverse ecological tapestry of Bangladesh's districts (Fig. 3(d)). Urban settlements cluster densely in Barisal, Khulna, and Chittagong, echoing the heartbeat of economic and social life, while Comilla, Sylhet, and Rangpur embrace a more scattered settlement pattern, reflecting a decentralized population distribution and a quieter rural lifestyle (Fig. 3(e)). The dichotomy of precipitation patterns paints a vivid picture of Bangladesh's hydrological landscape, with Chittagong and Dhaka basking in higher rainfall, while districts like Rajshahi, Rangpur, Mymensingh, Narayanganj, and Comilla face water scarcity challenges (Fig. 3(f)). Dhaka, Barisal, Sylhet, Rajshahi, and Khulna boast a rich array of regional facilities, including educational institutions and healthcare facilities,

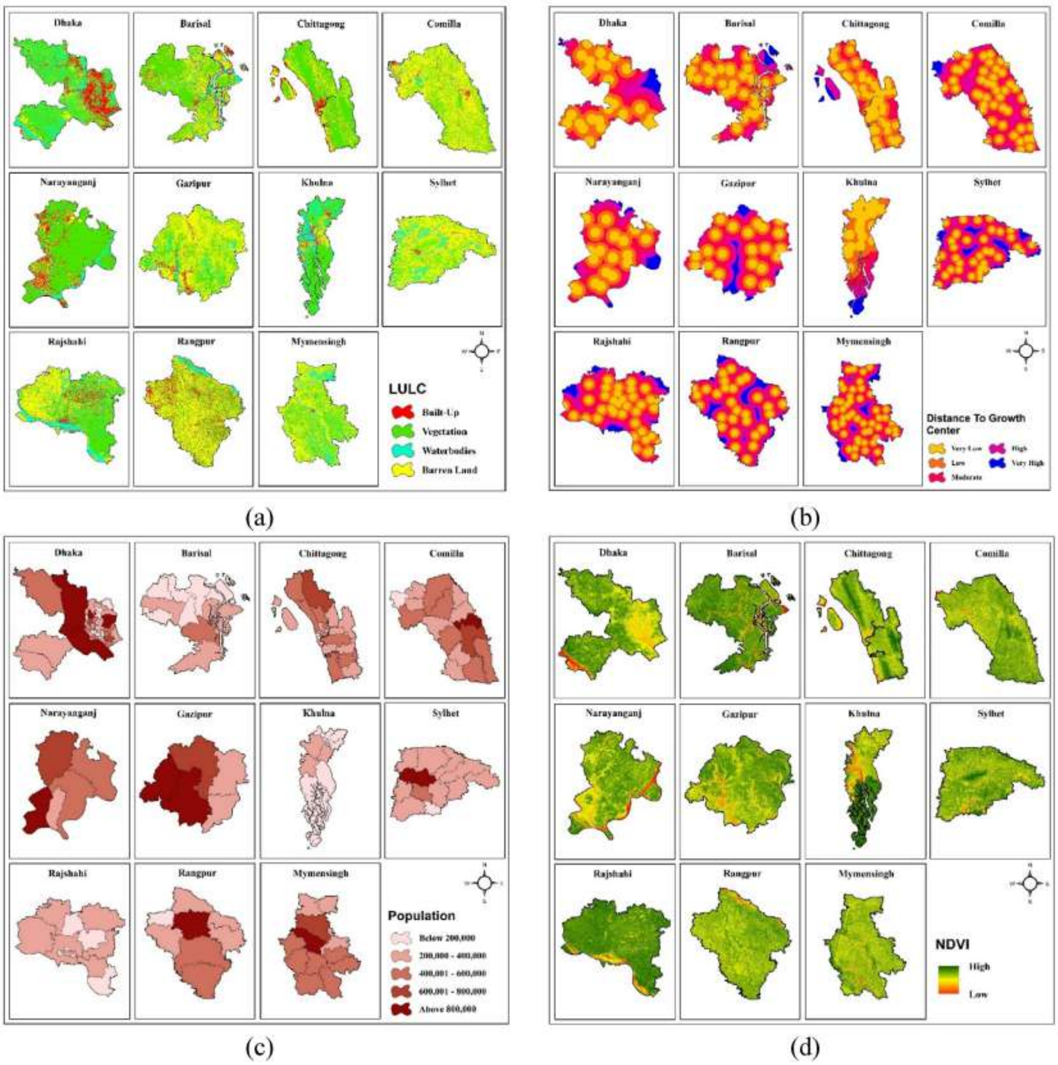

Fig. 3. Urban Expansion Governing Parameters (a) LULC, (b) Distance to growth center, (c) Population, (d) NDVI, (e) Distance to settlement, (f) Annual Average Precipitation, (g) Distance to facilities, (h) DEM, (i) Distance to industry, (j) Slope, (k) Distance to the road, (l) LST, (m) Distance to waterbody.

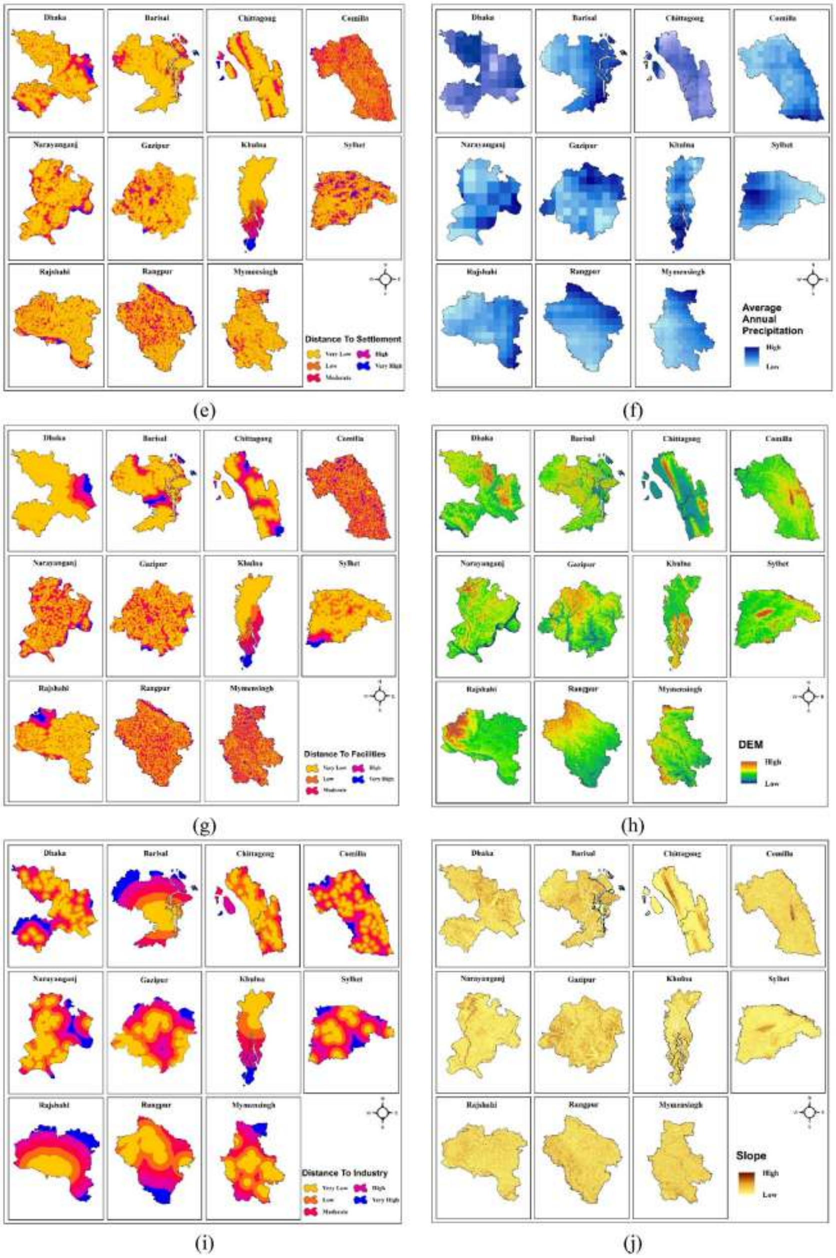

Fig. 3. . (*continued*).

anchoring their status as socio-economic hubs, while districts like Comilla, Rangpur, and Mymensingh grapple with fewer amenities, highlighting regional disparities in infrastructure (Fig. 3(g)). The rugged terrain of Chittagong and Sylhet offers panoramic vistas and biodiversity hotspots, contrasting with the flat plains of Dhaka, Narayanganj, and Gazipur, which facilitate urban expansion and infrastructure development (Fig. 3(h)). Chittagong, Narayanganj, and Gazipur emerge as industrial juggernauts, fueling Bangladesh's economic engine with manufacturing prowess, while Barisal, Rajshahi, and Khulna embrace a more diversified economic landscape, blending agriculture, commerce, and industry (Fig. 3(i)). The topographical nuances of districts like Dhaka, Narayanganj, and Gazipur, with their flat terrain, stand in stark contrast to the undulating slopes of Chittagong and Sylhet, shaping accessibility and urban development patterns (Fig. 3(j)). Dhaka, Gazipur, and Comilla boast extensive road networks, facilitating trade, commerce, and mobility, while districts like Rajshahi and Khulna grapple with less developed transportation infrastructure, posing challenges for connectivity and accessibility (Fig. 3(k)). The thermal tapestry of Bangladesh unfolds with Rajshahi higher temperatures, reflecting urban heat island phenomena, juxtaposed with the cooler climes of Gazipur and Rangpur, offering respite from the sweltering heat (Fig. 3(l)). Barisal, Comilla, and Khulna serve as aquatic sanctuaries, with extensive water bodies supporting biodiversity and fisheries,

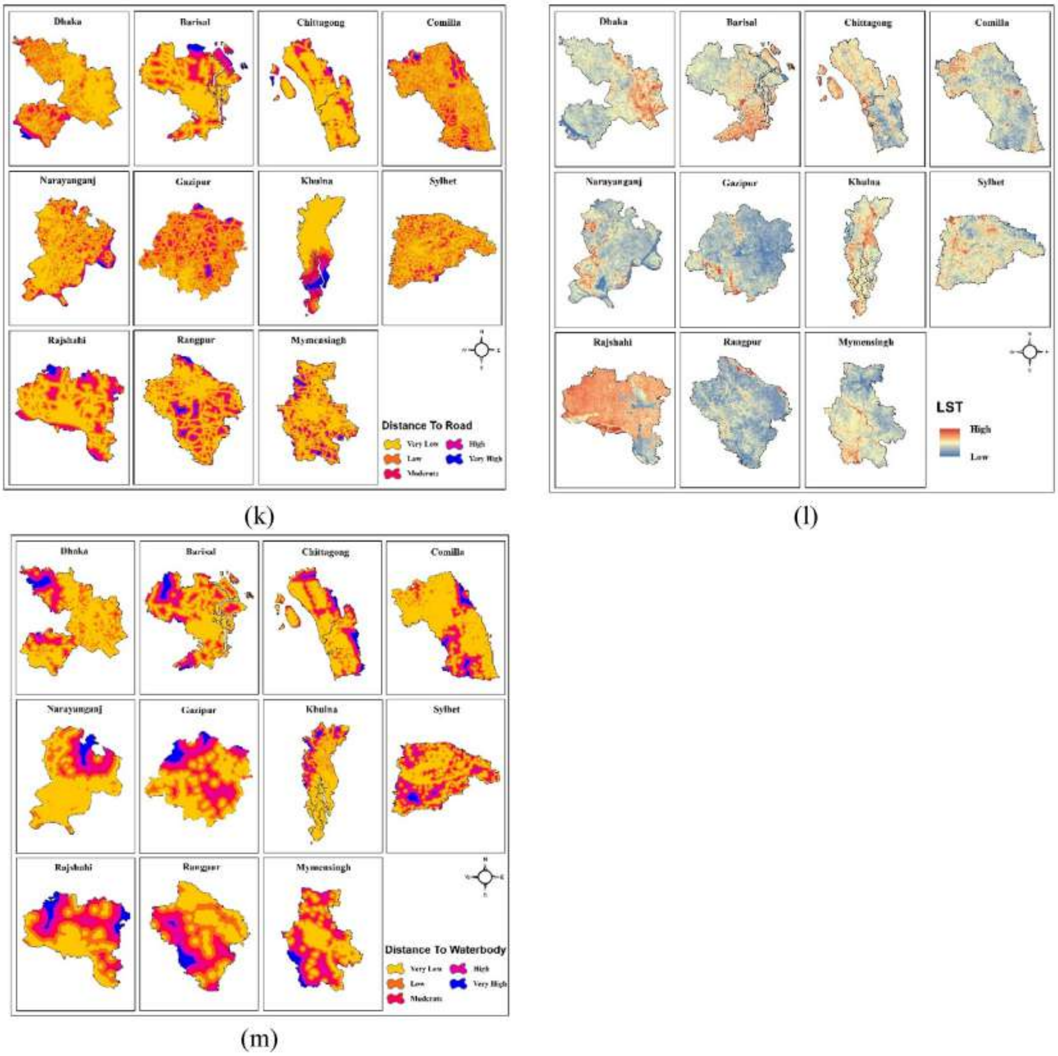

Fig. 3. . (*continued*).

while Gazipur, Rajshahi, and Rangpur rely on fewer water resources, highlighting the importance of water management and conservation efforts (Fig. 3(m)).

3.2. *Description of binary logistic regression model*

Table 3 depicts, the digital elevation model (DEM) significantly affects the vulnerability of urban expansion in districts like as Barisal, Chittagong, Khulna, Sylhet, Rajshahi, and Rangpur, with a statistical significance ($P < 0.05$). The population affected in Barisal ($P = 0.049$), Comilla ($P = 0.046$), Narayanganj ($P = 0.048$), and Gazipur ($P = 0.001$). Furthermore, the existence of amenities, growth centers, and industrial zones has been identified as important factors that have a notable impact on districts like as Dhaka, Comilla, and Narayanganj. This impact is statistically significant, with a significance level of less than 0.05. Precipitation had a significant impact on the growth of urban areas in Dhaka ($P = 0.035$), Chittagong ($P = 0.001$), Gazipur ($P = 0.048$), Khulna ($P = 0.008$), Rajshahi ($P = 0.001$), Rangpur ($P = 0.042$), and Comilla ($P = 0.040$), with all P values being less than 0.05. The industry shown substantial impacts in Narayanganj, Barisal, Comilla, Khulna, Rangpur, Mymensingh, and Gazipur. Land surface temperature (LST) and normalized difference vegetation index (NDVI) have complex connections with urbanization patterns, affecting areas such as Barisal, Chittagong, and Rajshahi. The slope has a substantial impact on the urban expansion in Barisal ($P = 0.038$), Chittagong ($P = 0.013$), Narayanganj ($P = 0.029$), and Mymensingh ($P = 0.001$). The score in Table 3 denotes a metric for assessing the model's fit to the data, with higher values signifying a better fit. Furthermore, almost all variables demonstrate an excellent fit with the model across all cities. The existence of infrastructure elements such as roads and settlements demonstrated significant correlations with the vulnerability to urban expansion in many districts, including Dhaka, Barisal, Comilla, Rangpur, Khulna, and Sylhet. Water bodies significantly contributed to the urban development of Dhaka, Narayanganj, Khulna, Rajshahi, and Rangpur, as evidenced by statistical study.

Table 3
Summary of scores and significances for various variables across districts in Bangladesh.

Variables	Dhaka		Barisal		Chittagong		Comilla		Narayanganj		Gazipur		Khulna
	Score	Sig.	Score	Sig.	Score	Sig.	Score	Sig.	Score	Sig.	Score	Sig.	Score
DEM	0.000	0.994	16.239	0.001	7.064	0.008	1.776	0.183	0.425	0.514	0.196	0.658	28.938
Facilities	29.964	0.001	0.120	0.729	0.005	0.944	5.526	0.019	0.490	0.484	2.097	0.148	8.195
Growth Center	10.742	0.001	7.466	0.006	0.740	0.390	13.219	0.001	0.772	0.380	0.003	0.957	9.219
Industry	0.732	0.392	4.548	0.033	1.130	0.288	2.966	0.045	6.941	0.008	6.007	0.014	3.150
LST	27.900	0.001	5.862	0.015	6.088	0.014	21.580	0.001	24.595	0.001	34.050	0.001	24.576
LULC	63.482	0.001	92.350	0.001	53.311	0.001	66.068	0.001	52.263	0.001	30.717	0.001	161.751
NDVI	22.486	0.001	5.332	0.021	40.491	0.001	1.947	0.163	5.116	0.024	14.310	0.001	14.874
Population	0.009	0.926	3.512	0.049	1.236	0.266	3.848	0.046	3.545	0.048	13.779	0.001	0.041
Precipitation	4.433	0.035	0.751	0.386	13.399	0.001	4.234	0.040	0.002	0.963	2.918	0.048	7.071
Road	0.058	0.810	3.808	0.041	0.206	0.650	6.580	0.010	1.337	0.248	2.604	0.107	13.370
Settlement	23.476	0.001	11.397	0.001	2.268	0.132	13.624	0.001	0.912	0.340	0.002	0.963	8.428
Slope	0.957	0.328	4.297	0.038	6.182	0.013	0.886	0.347	4.742	0.029	0.740	0.390	2.299
Waterbody	5.346	0.021	0.003	0.956	1.503	0.220	0.068	0.794	3.755	0.049	1.779	0.182	2.919

*df for all is 1.

The "Model Range" in Table 4 indicates the model's predictive capacity, which ranges from 0% to 100%. A range that is closer to 100% suggests a more robust fit, which implies that the model offers a more comprehensive explanation of the data. Model ranges can be interpreted as high fit (80%–100%), moderate fit (50%–79%), and low fit (<50%) in order to categorize the results in Table 4. According to Tables 4 and in Dhaka, the model elucidates between 28.7% and 73.5% of the variability in urban expansion susceptibility. Meanwhile, Barisal exhibits a broader model range, spanning from 33.6% to 94.2%. Chittagong, on the other hand, presents an extensive model range, ranging from 16.2% to 99.9%. Comilla's model range extends from 12.1% to 100%. In Narayanganj and Gazipur, the model demonstrates a model range from 48.4% to 100% and 13.2%–100%. Khulna's model range spans from 30.8% to 92.4%. However, Sylhet's model range ranges from 30.8% to 48.2%. Rajshahi model range, from 68.3% to 94.2%, while Rangpur's range, from 40.4% to 59.8%, denotes an extent of variability accounted for. Finally, Mymensingh showcases the model range of 44.4%–100%. Narayanganj or Gazipur indicates that the model is a robust predictor, indicating a high fit and accounting for a significant percentage of the variation in the dependent variable. Additionally, Rangpur (40.4%–59.8%) demonstrates a moderate level of clarity, wherein the model accounts for some variance yet allows for enhancement. Finally, districts that have lower model ranges, such as Sylhet (30.8%–48.2%), indicate that the model accounts for a lesser proportion of the variation, implying a diminished fit in those regions.

3.3. Urban expansion susceptibility

Fig. 4 and Table 5 illustrate the susceptibility of major urban units in Bangladesh to urban expansion, categorized into three tiers: low, moderate, and high. The categories were determined based on the raster class produced from the outputs of the Random Forest model, integrating the governing parameters. The low susceptibility area in Dhaka covers 12.88 square kilometres (0.88%), the intermediate susceptibility zone comprises 943.55 square kilometres (64.45%), and the high susceptibility area contains 507.57 square kilometres (34.67%). In Barisal, the area of low susceptibility spans 24.51 square kilometres (0.88%), the moderate susceptibility zone comprises 1839.49 square kilometres (66.05%), and the high susceptibility region stretches over 921.00 square kilometres (33.07%). The region of Chittagong may be categorized into three zones according to vulnerability. A low susceptibility area encompasses 183.32 square kilometres (3.47%), a moderate susceptibility region spans 3759.38 square kilometres (71.16%), while a high susceptibility zone covers 1340.30 square kilometres (25.37%). Comilla exhibits a low vulnerability, encompassing an area of 118.29 square kilometres (3.76%). It also displays a moderate susceptibility across 2335.28 square kilometers (74.23%) and a high susceptibility across 692.12 square kilometers (22.00%). The area with low susceptibility in Narayanganj includes 37.62 square kilometers, which accounts for 5.50% of the whole area. The moderate susceptibility zone spans 462.38 square kilometers, making up 67.60% of the total area. Lastly, the high susceptibility area stretches over 184.00 square kilometers, representing 26.90% of the total area. Gazipur exhibits a low susceptibility area spanning 121.72 square kilometers (6.74%), a moderate susceptibility region covering 1469.54 square kilometers (81.37%), and a high susceptibility zone spanning 214.73 square kilometers (11.89%). The low susceptibility region in Khulna spans an area of 855.95 square kilometers, which accounts for 19.48% of the whole area. The moderate susceptibility zone includes 1968.51 square kilometers, making up 44.80% of the total area. The high susceptibility area stretches over 1569.54 square kilometers, representing 35.72% of the total area. Sylhet exhibits a small susceptibility area spanning 111.15 square kilometers (3.22%), a moderate susceptibility area covering 2510.98 square kilometers (72.74%), and a large susceptibility area including 830.21 square kilometers (24.05%). Rajshahi demonstrates a small susceptibility area of 67.42 square kilometers (2.78%), a moderate susceptibility area of 2016.87 square kilometers (83.17%), and a large susceptibility area of 340.71 square kilometers (14.05%). Rangpur has a small susceptibility area of 209.37 square kilometers (8.72%), a moderate susceptibility area of 1636.04 square kilometers (68.14%), and a large susceptibility area of 555.59 square kilometers (23.14%). In Mymensingh, the area with low susceptibility spans 308.97 square kilometers (7.03%), the area with moderate susceptibility contains 3440.41 square kilometers (78.28%), and the area with high susceptibility covers 645.63 square kilometers (14.69%).

Khulna	Sylhet		Rajshahi		Rangpur		Mymensingh	
Sig.	Score	Sig.	Score	Sig.	Score	Sig.	Score	Sig.
0.001	9.436	0.002	52.926	0.001	3.557	0.049	2.472	0.116
0.004	2.520	0.112	56.733	0.001	0.401	0.527	0.427	0.513
0.002	0.304	0.581	11.816	0.001	1.974	0.160	0.480	0.489
0.046	0.257	0.612	2.394	0.122	2.939	0.046	3.572	0.049
0.001	0.105	0.746	21.751	0.001	1.040	0.308	38.521	0.001
0.001	197.341	0.001	311.321	0.001	177.293	0.001	21.688	0.001
0.001	26.508	0.001	40.231	0.001	1.012	0.314	9.913	0.002
0.839	0.840	0.359	0.103	0.748	0.012	0.913	0.819	0.365
0.008	1.774	0.183	44.410	0.001	4.138	0.042	0.375	0.540
0.001	4.604	0.032	7.733	0.005	4.340	0.037	1.610	0.205
0.004	4.860	0.027	5.082	0.024	0.516	0.473	1.813	0.178
0.129	0.968	0.325	0.522	0.470	0.651	0.420	11.024	0.001
0.048	0.746	0.388	24.627	0.001	3.525	0.049	0.535	0.465

3.3.1. Accuracy assessment

In Dhaka, Barisal, Chittagong, Comilla, Narayanganj, Gazipur, Khulna, Rajshahi, and Mymensingh, the AUC values ranged from 0.98 to 1.00, signifying excellent predictive accuracy. Conversely, Sylhet and Rangpur exhibited slightly lower AUC values of 0.89 and 0.85, respectively, categorizing their predictive performance as “very good” (Fig. 5).

4. Discussion

This research evaluated the susceptibility to urban expansion in eleven major cities in Bangladesh, determining Khulna as the most susceptible and Dhaka and Barisal as the least susceptible. The research effectively identified significant parameters of urban expansion, including NDVI, LST, and proximity to industry, through the application of machine learning and geostatistical models. It is essential to critically assess these findings, particularly in relation to previous literature and recognizing possible anomalies. Numerous studies have demonstrated that urban expansion poses a significant risk to the world's biodiversity (Di Giulio et al., 2009; Seto et al., 2012; Vimal et al., 2012). In the context of the globe as a whole, the most significant factor for urban expansion is the presence of green infrastructure, urban parks, and recreational facilities (Larondelle & Lauf, 2016). Throughout the metropolis, leap-frog and discontinuous development occurred as a result of anthropogenic activities such as urbanization, which continued to indicate that landscapes were becoming increasingly complex in shape and structure (Yang and Zhou, 2014). Similarly, vegetation is also considered the most influencing factor for urban expansion in Ethiopia (Barow et al., 2019). Moreover, machine learning models have been used widely to predict the susceptibility environmental climate change (Hassan & Nazem, 2016; Haydar et al., 2024; Khosravi et al., 2023; Sadia et al., 2023) Urban growth can enhance the chance potential of people's activities by providing a more favorable natural environment (Li et al., 2016). Urban build-up areas, agricultural areas, bushland, forests, residential, industrial, and waterbody are taken as major factors for the urban expansion where urban and built-up areas, agricultural areas, and water bodies have experienced an increase, while the residential, industrial areas, bushland, and forests have seen a decrease (El Garouani et al., 2017; Mundia & Aniya, 2005). Over a 28-year period, the urban area grew at a pace of 2.25% in the Chittagong region, becoming the most prominent land use type as a result of the conversion of vegetation and settlement (Roy et al., 2020). Similarly, Chittagong has shown high urban expansion susceptibility in this study. The land use in the urban core consisted of a combination of major and minor industries, commercial structures, office spaces, and residential units in Bangladesh that there is similarity in Dhaka and Chittagong for urban expansion (Hassan & Nazem, 2016). The

Table 4
Model fitting information's for the Distracts.

Model Summary				
	−2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square	Model Range
Dhaka	62.576	0.287	0.735	28.7%–73.5%
Barisal	17.422	0.336	0.942	33.6%–94.2%
Chittagong	0.250	0.162	0.999	16.2%–99.9%
Comilla	0.000	0.121	1.000	12.1%–100%
Narayanganj	0.000	0.484	1.000	48.4%–100%
Gazipur	0.000	0.132	1.000	13.2%–100%
Khulna	32.452	0.308	0.924	30.8%–92.4%
Sylhet	611.863	0.308	0.482	30.8%–48.2%
Rajshahi	80.954	0.683	0.942	68.3%–94.2%
Rangpur	388.933	0.404	0.598	40.4%–59.8%
Mymensingh	0.000	0.444	1.000	44.4%–100%

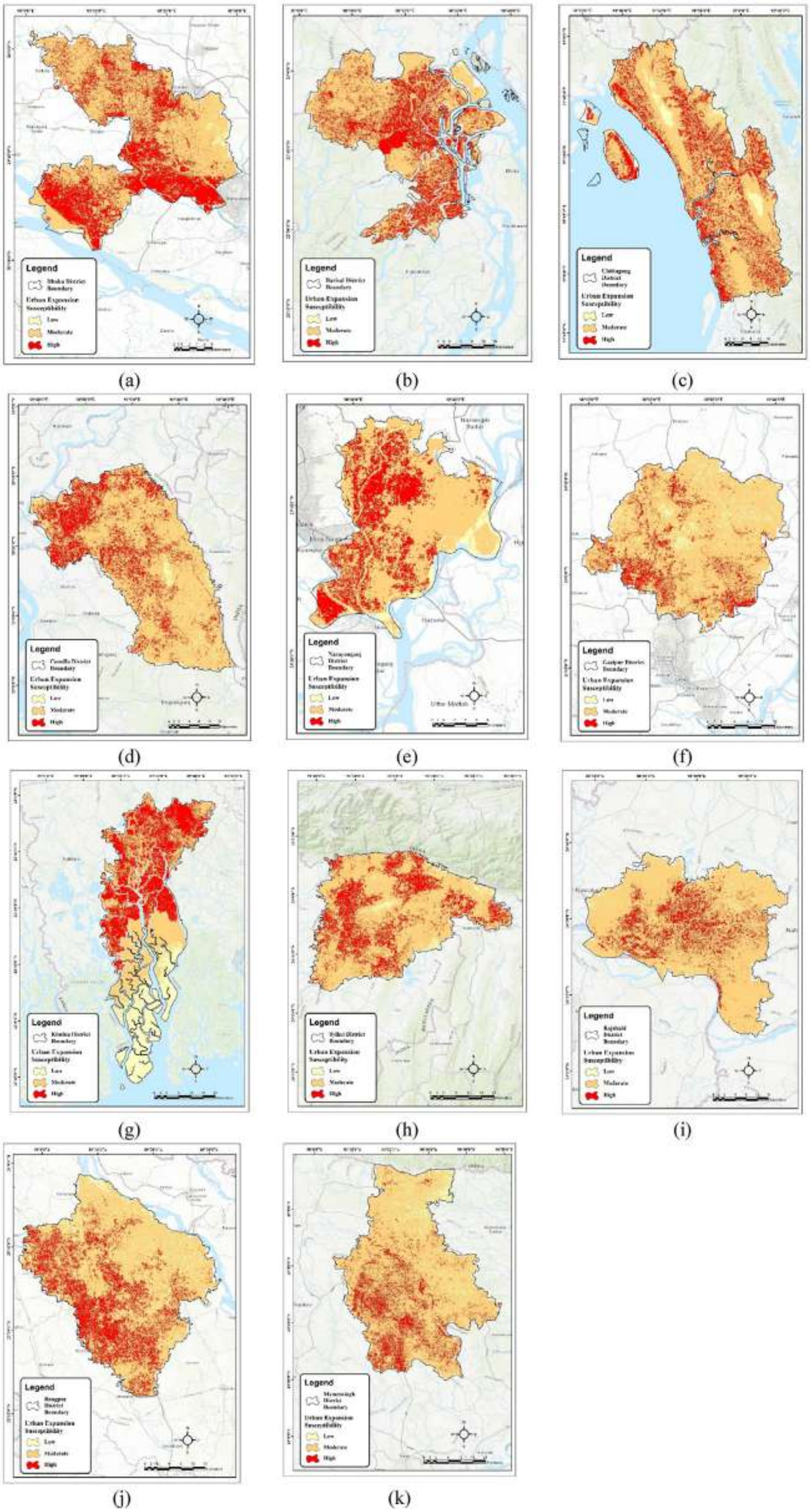

Fig. 4. Urban expansion susceptibility in (a) Dhaka, (b) Barisal, (c) Chittagong, (d) Comilla, (e) Narayanganj, (f) Gazipur, (g) Khulna, (h) Sylhet, (i) Rajshahi, (j) Rangpur, (k) Mymensingh.

Table 5
Distribution of urban expansion susceptibility across different districts in Bangladesh.

Districts Name	Low		Moderate		High	
	Area (Sq.km)	%	Area (Sq.km)	%	Area (Sq.km)	%
Dhaka	12.8832	0.88%	943.55	64.45%	507.57	34.67%
Barisal	24.508	0.88%	1839.49	66.05%	921.00	33.07%
Chittagong	183.3201	3.47%	3759.38	71.16%	1340.30	25.37%
Comilla	118.2896	3.76%	2335.28	74.23%	692.12	22.00%
Narayanganj	37.62	5.50%	462.38	67.60%	184.00	26.90%
Gazipur	121.7244	6.74%	1469.54	81.37%	214.73	11.89%
Khulna	855.9512	19.48%	1968.51	44.80%	1569.54	35.72%
Sylhet	111.1544	3.22%	2510.98	72.74%	830.21	24.05%
Rajshahi	67.415	2.78%	2016.87	83.17%	340.71	14.05%
Rangpur	209.3672	8.72%	1636.04	68.14%	555.59	23.14%
Mymensingh	308.9685	7.03%	3440.41	78.28%	645.63	14.69%

steady shifts in landscape measures across the five cities (Rajshahi, Rangpur, Sylhet, Khulna, and Barisal) show that disturbances and human activity are always transforming landscapes from one dominant land cover to another (Yang and Zhou, 2014). In this study, the similar results have been found on urban expansion in Rajshahi, Rangpur, Sylhet, Khulna, and Barisal. In Rangpur, scholars showed that 5670 hectors of land in all directions with settlement, industrial, and commercial infrastructure, are rapidly increasing, indicating the loss of agricultural land and bare soil with rising temperatures and increased urban expansion due to parameters industry, growth center and settlements (N. Rahman, 2019). Khulna's polycentric urban expansion is due to strong road system development, port formation, and city authority initiatives. The 2015 Mongla Port, 2008 Khulna-Jessore bypass road, and 2016 Khulna-Mongla Rail link are enabling exports and imports to India. This bypass is connected by three auxiliary routes across Khan Jahan Ali, Dumuria, and Batiaghata (Moniruzzaman et al., 2012). The transportation network development led to polycentric urban growth in Khulna (Alam et al., 2023).Furthermore, the primary cause is the recovery of industry and the formation of new industries that impact the urban expansion of both core and periphery regions in Bangladesh. The jute industry in Khulna faced difficulties under several political administrations; nevertheless, a 2010 strategy promoting public-private partnerships facilitated the regeneration of the industrial sector and encouraged urban development. Following to 2017, the revitalization of the industrial sector and the establishment of new firms in urban peripheries generated employment prospects, drawing migrants impacted by catastrophes and poverty to reside or utilize the region as a dormitory (Noorhosseini et al., 2017; Su et al., 2013). Some scholars identified the most important factors; density gradients (Landis, 2017); land-use mix, road network, and activity centering (Ewing & Hamidi, 2017). In the regional context, the urban expansion rate is 2.5% as increasing of the transformation of rural settlements and agricultural land during 28 years in the Chittagong Metropolitan Area

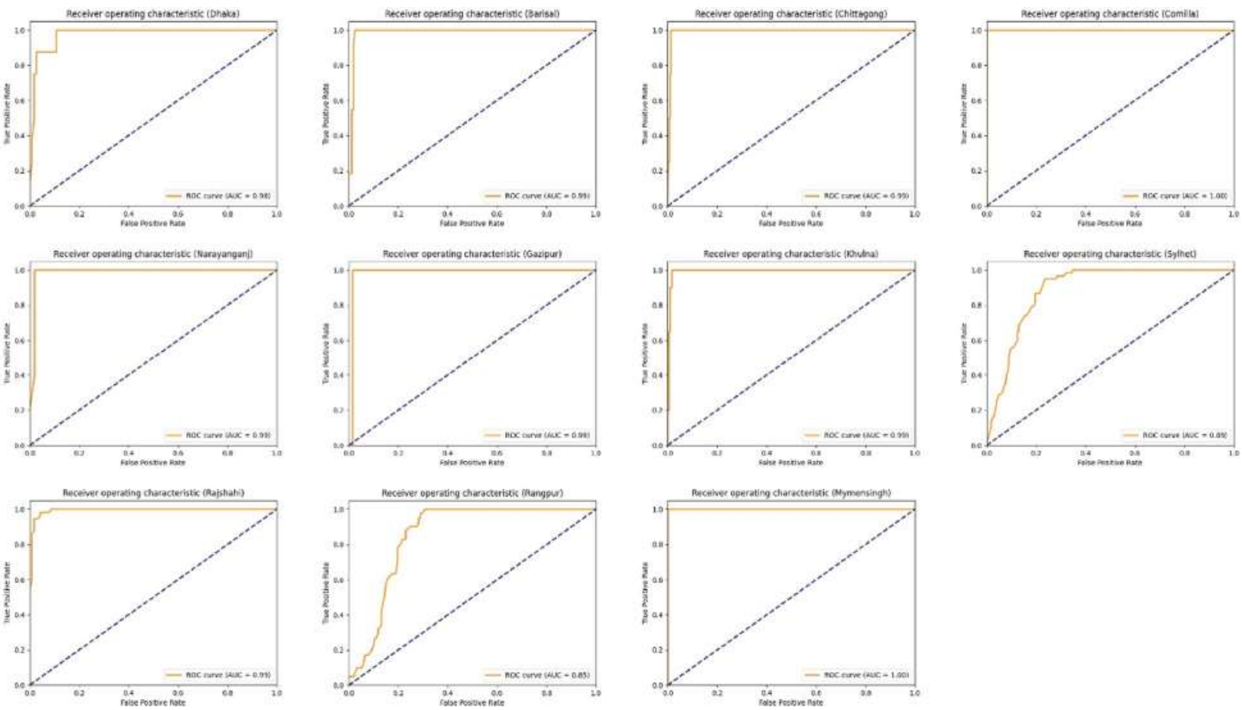

Fig. 5. ROC curve for RF model of the districts.

(CMA) (Roy et al., 2020). Another research (Hassan & Nazem, 2016) stated that Chittagong City has grown by 61.8%, at a yearly growth rate of 17.5%. Nowadays, urban expansion is happening regular year in Bangladesh such as in Dhaka where the built-up area increased from 46% to 58% of the entire study area (Ahmed & Ahmed, 2012), urban growth occurred in Rajshahi, Sylhet, and Rangpur between 1973 and 1989 at rates of 12.66%, 11.02%, and 13%, respectively (Hassan, 2017). The results of this study provide a scientific understanding of the factors that lead to urban expansion and how susceptible they are. It is obvious, based on an analysis of data related to the major urban units, that a significant portion of Khulna (specifically, 35.72%) is situated in an area with the greatest urban expansion. Conversely, 83.17 percent of the main urban unit area in Rajshahi is characterized by urban expansion at a moderate level. On the contrary, approximately 0.88% of the land area in Dhaka and Barisal falls within the lowest urban expansion zone. The differences in economic growth and activities, also geographical size and location influence the rate of urban expansion (Hassan, 2017; M. S. Rahman et al., 2019; Yousaf Raza et al., 2023).

A particular subject of concern is the decreased AUC values observed for Sylhet and Rangpur in comparison to other districts. This may arise from geographical disparities in data accessibility, changes in topography, or inconsistencies in socio-economic activity. Sylhet, characterized by substantial vegetation and high precipitation, may display growth patterns distinct from those of more industrialized regions, hence resulting in problems to model forecasting. Comparable variations have been identified in other areas with distinct environmental characteristics, as documented in research on urbanization patterns in Ethiopia (Barow et al., 2018). Although these constraints, our results concerning NDVI and LST as key elements in urban expansion are consistent with global research on urban sprawl, especially in regions where accelerated vegetation loss is closely associated with urban expansion (Foley et al., 2005). Nonetheless, several of our findings deviate from the established literature. While most studies recognize population density as a principal factor in urban expansion, our findings underscore the significance of industry and regional infrastructure as more pivotal in areas like as Dhaka and Khulna. This discrepancy may stem from Bangladesh's distinct socio-economic environment, where industrial development frequently occurs prior to population expansion, in contrast to other areas where urbanization is a response to demographic pressures (Roy et al., 2020).

The LULC in different regions exhibits different rates and patterns of land conversion from non-urban to urban. Factors such as current land utilization (vegetation, water bodies), accessibility to infrastructure (roads, regional facilities, growth centers, industry, settlements), and the proximity to urban centers substantially affect the probability of an area developing into an urban environment. These findings offer essential guidance for policymakers, indicating that infrastructure development, especially in industrial hubs, requires careful oversight to prevent unsustainable urban expansion. This study advances Sustainable Development Goals (SDGs), specifically Goal 11 ("Sustainable Cities and Communities"), by discerning regions susceptible to urban expansion to enhance land utilization. It further advances Goal 15 ("Life on Land") by safeguarding ecosystems, Goal 9 ("Industry, Innovation, and Infrastructure") through enhanced infrastructure planning, and Goal 13 ("Climate Action") by regulating urban expansion in climate-sensitive regions Moreover, although our models demonstrate robust prediction capability (AUC values ranging from 0.89 to 1.00), subsequent research should incorporate economic factors like GDP and employment to provide a more refined comprehension of urban growth dynamics. Furthermore, employing high-resolution data in forthcoming studies could reduce certain anomalies and improve model accuracy.

5. Conclusion

This study aimed to determine the susceptibility of urban expansion using the binary logistic regression model and data-driven machine learning algorithm. The study indicates that Khulna has the highest susceptibility to urban expansion, covering 35.72% of its total area, succeeded by Chittagong at 25.37% and Dhaka at 34.67%. Rajshahi demonstrates the greatest moderate susceptibility at 83.17%, while Comilla and Barisal also possess significant proportions in the moderate category, at 74.23% and 66.05%, respectively. Regions with low susceptibilty encompass Dhaka (0.88%) and Barisal (0.88%), signifying negligible danger for imminent urban growth. Gazipur has intermediate susceptibility at 81.37%, with a mere 11.89% indicating high susceptibility. The ROC-AUC of random forest for Dhaka, Barisal, Chittagong, Comilla, Narayanganj, Gazipur, Khulna, Rajshahi, and Mymensingh ranged from 0.98 to 1.00, indicating exceptional predictive accuracy. On the other hand, Sylhet and Rangpur exhibited somewhat lower values of 0.89 and 0.85, respectively. The findings indicate that urban expansion susceptibility is mostly affected by factors including DEM, Industry, LST, LULC, NDVI, precipitation, road, and settlement, also found significant in Binary Logistic Regression model. This investigation is subject to specific limits. The majority of datasets employed in this study were updated to the greatest extent feasible. This research is constrained by various limits, including high-resolution images, spatio-temporal analysis employing machine learning methods, and economic considerations such as GDP and employment. Moreover, alterations in regulating parameters may provide varying outcomes based on the local setting. This strategy will improve our capacity to identify and quantify the elements affecting the variability in urban expansion susceptibility, providing a more comprehensive knowledge of regional urbanization changes. This study strengthens SDG 11 by identifying urban expansion-prone zones to improve land use. It advances Goal 15 by protecting ecosystems, Goal 9 by improving infrastructure planning, and Goal 13 by limiting urban expansion in climate-sensitive zones. Mitigating the effects of growing urbanization on rural regions necessitates proactive development plans, with forthcoming research emphasizing machine learning for enhanced evaluations of urban expansion.

CRediT authorship contribution statement

Mafrid Haydar: Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization. **Sakib Hosan:** Writing – original draft, Visualization, Methodology, Formal analysis, Data curation. **Al Hossain Rafi:** Writing – original draft, Formal analysis, Data curation.

Ethical approval

The authors declare No ethical approval is required. Ethical approval for this type of study is not required by our institute.

Funding

None.

Conflicts of 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.

References

Ahmed, B., & Ahmed, R. (2012). Modeling urban land cover growth dynamics using multioral satellite images: A case study of Dhaka, Bangladesh. *ISPRS International Journal of Geo-Information, 1*(1), 3–31. https://doi.org/10.3390/ijgi1010003

Alam, I., Nahar, K., & Morshed, M. M. (2023). Measuring urban expansion pattern using spatial matrices in Khulna City, Bangladesh. *Heliyon, 9*(2), Article e13193. https://doi.org/10.1016/j.heliyon.2023.e13193

Anderson, P. M. L., Okereke, C., Rudd, A., & Parnell, S. (2013). Regional assessment of africa. In *Urbanization, biodiversity and ecosystem services: Challenges and opportunities: A global assessment*. https://doi.org/10.1007/978-94-007-7088-1_23

Arifeen, H. M., Phoungthong, K., Mostafaeipour, A., Yuangyai, N., Yuangyai, C., Techato, K., & Jutidamrongphan, W. (2021). Determine the land-use land-cover changes, urban expansion and their driving factors for sustainable development in gazipur Bangladesh. *Atmosphere, 12*(10). https://doi.org/10.3390/atmos12101353

Bajracharya, P., & Sultana, S. (2022). Examining the use of urban growth boundary for future urban expansion of Chattogram, Bangladesh. *Sustainability, 14*(9). https://doi.org/10.3390/su14095546

Bangladesh Bureau of Statistics. (2022). *Population & housing cencus 2022, preliminary report* (Vol. 11). Bangladesh Bureau of Statistics.

Bangladesh Ministry of Foreign Affairs. (2017). *Bangladesh–An-Introduction - Ministry of Foreign Affairs-*. Bangladesh Ministry of Foreign Affairs.

Barow, I., Megenta, M., & Megento, T. (2018). Spatiotemporal analysis of urban expansion using GIS and remote sensing in Spatiotemporal analysis of urban expansion using GIS and remote sensing in Jigjiga town of Ethiopia. https://doi.org/10.1007/s12518-018-0245-z.

Barow, I., Megenta, M., & Megento, T. (2019). Spatiotemporal analysis of urban expansion using GIS and remote sensing in Jigjiga town of Ethiopia. *Applied Geomatics, 11*(2), 121–127. https://doi.org/10.1007/s12518-018-0245-z

Bickel, P., Diggle, P., Fienberg, S., Gather, U., Olkin, I., & Zeger, S. (2009). *Springer series in statistics*.

Capps, K. A., Bentsen, C. N., & Ramírez, A. (2016). Poverty, urbanization, and environmental degradation: Urban streams in the developing world. *Freshwater Science, 35*(1), 429–435. https://doi.org/10.1086/684945

Chakir, R., & Parent, O. (2009). The economic model the econometric model Data and Estimation results Spatial Econometric Workshop. *TSE . April 2009 conclusion outline introduction*. April.

Coursera Staff. (2024). *What is machine learning? Definition, types, and examples*. Coursesa.

Di Giulio, M., Holderegger, R., & Tobias, S. (2009). Effects of habitat and landscape fragmentation on humans and biodiversity in densely populated landscapes. *Journal of Environmental Management, 90*(10), 2959–2968. https://doi.org/10.1016/j.jenvman.2009.05.002

El Garouani, A., Mulla, D. J., El Garouani, S., & Knight, J. (2017). Analysis of urban growth and sprawl from remote sensing data: Case of Fez, Morocco. *International Journal of Sustainable Built Environment, 6*(1), 160–169. https://doi.org/10.1016/j.ijsbe.2017.02.003

Evidently, A. I. (2020). *How to explain the ROC curve and ROC AUC score? Classification metrics guide*.

Ewing, R., & Hamidi, S. (2017). Costs of sprawl. *Costs of Sprawl*. https://doi.org/10.4324/9781315628103

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N., & Snyder, P. K. (2005). Global consequences of land use. *Science, 309*(5734), 570–574. https://doi.org/10.1126/science.1111772

Fuchs, R., Herold, M., Verburg, P. H., Clevers, J. G. P. W., & Eberle, J. (2015). Gross changes in reconstructions of historic land cover/use for Europe between 1900 and 2010. *Global Change Biology, 21*(1), 299–313. https://doi.org/10.1111/gcb.12714

Gao, J., Wei, Y. D., Chen, W., & Chen, J. (2014). Economic transition and urban land expansion in Provincial China. *Habitat International, 44*, 461–473. https://doi.org/10.1016/j.habitatint.2014.09.002

Gerard, F., Petit, S., Smith, G., Thomson, A., Brown, N., Manchester, S., Wadsworth, R., Bugar, G., Halada, L., Bezák, P., Boltiziar, M., de Badts, E., Halabuk, A., Mojses, M., Petrovic, F., Gregor, M., Hazeu, G., Mücher, C. A., Wachowicz, M., … Feranec, J. (2010). Land cover change in Europe between 1950 and 2000 determined employing aerial photography. *Progress in Physical Geography, 34*(2), 183–205. https://doi.org/10.1177/0309133309360141

Gibson, J., Boe-Gibson, G., & Stichbury, G. (2015). Urban land expansion in India 1992-2012. *Food Policy, 56*, 100–113. https://doi.org/10.1016/j.foodpol.2015.08.002

GISGeography. (2021). *GIS geography*.

Gorton, D. (2023). *Industry definition in business and investing*. Investopedia.

Ha, J., Kambe, M., & Pe, J. (2011). Data mining: Concepts and techniques. In *Data mining: Concepts and techniques*. https://doi.org/10.1016/C2009-0-61819-5

Habib, M. A., Kabir, F., & Sultana, M. (2020). *Impact of urban expansion on surface water bodies in Dhaka metropolitan area of Bangladesh: A remote sensing and GIS based analysis*. BUFT Journal of Business & Economics (BJBE).

Harris, J. K. (2021). *Primer on binary logistic regression*. https://doi.org/10.1136/fmch-2021-001290

Hassan, M. M. (2017). Monitoring land use/land cover change, urban growth dynamics and landscape pattern analysis in five fastest urbanized cities in Bangladesh. *Remote Sensing Applications: Society and Environment, 7*, 69–83. https://doi.org/10.1016/j.rsase.2017.07.001

Hassan, M. M., & Nazem, M. N. I. (2016). Examination of land use/land cover changes, urban growth dynamics, and environmental sustainability in Chittagong city, Bangladesh. *Environment, Development and Sustainability, 18*(3), 697–716. https://doi.org/10.1007/s10668-015-9672-8

Hatna, E., & Bakker, M. M. (2011). Abandonment and expansion of arable land in Europe. *Ecosystems, 14*(5), 720–731. https://doi.org/10.1007/s10021-011-9441-y

Haydar, M., Rafi, A. H., Sadia, H., & Hossain, T. (2024). Data driven forest fire susceptibility mapping in Bangladesh. *Ecological Indicators*, (April), Article 112264. https://doi.org/10.1016/j.ecolind.2024.112264

Hochstrasser, R. M. (2007). Two-dimensional spectroscopy at infrared and optical frequencies. *Proceedings of the National Academy of Sciences of the United States of America, 104*(36), 14190–14196. https://doi.org/10.1073/pnas.0704079104

Hua, L. J., Ma, Z. G., & Guo, W. D. (2008). The impact of urbanization on air temperature across China. *Theoretical and Applied Climatology, 93*(3–4), 179–194. https://doi.org/10.1007/s00704-007-0339-8

Huilei, L., Jian, P., Yanxu, L., & Yi'na, H. (2017). Urbanization impact on landscape patterns in Beijing city, China: A spatial heterogeneity perspective. *Ecological Indicators, 82*(May), 50–60. https://doi.org/10.1016/j.ecolind.2017.06.032

Jiang, L., Deng, X., & Seto, K. C. (2012). Multi-level modeling of urban expansion and cultivated land conversion for urban hotspot counties in China. *Landscape and Urban Planning, 108*(2–4), 131–139. https://doi.org/10.1016/j.landurbplan.2012.08.008

Kantakumar, L. N., Kumar, S., & Schneider, K. (2016). Spatiotemporal urban expansion in Pune metropolis, India using remote sensing. *Habitat International, 51*, 11–22. https://doi.org/10.1016/j.habitatint.2015.10.007

Khosravi, K., Rezaie, F., Cooper, J. R., Kalantari, Z., Abolfathi, S., & Hatamiafkoueieh, J. (2023). Soil water erosion susceptibility assessment using deep learning algorithms. *Journal of Hydrology, 618*(January), Article 129229. https://doi.org/10.1016/j.jhydrol.2023.129229

Landis, J. D. (2017). The end of sprawl? Not so fast. *Housing Policy Debate, 27*(5), 659–697. https://doi.org/10.1080/10511482.2017.1296014

Larondelle, N., & Lauf, S. (2016). Balancing demand and supply of multiple urban ecosystem services on different spatial scales. *Ecosystem Services, 22*, 18–31. https://doi.org/10.1016/j.ecoser.2016.09.008

Levers, C., Müller, D., Erb, K., Haberl, H., Jepsen, M. R., Metzger, M. J., Meyfroidt, P., Plieninger, T., Plutzar, C., Stürck, J., Verburg, P. H., Verkerk, P. J., & Kuemmerle, T. (2018). Archetypical patterns and trajectories of land systems in Europe. *Regional Environmental Change, 18*(3), 715–732. https://doi.org/10.1007/s10113-015-0907-x

Leyk, S., Uhl, J. H., Connor, D. S., Braswell, A. E., Mietkiewicz, N., Balch, J. K., & Gutmann, M. (2020). *Two centuries of settlement and urban development in the United States*.

Li, B., Chen, D., Wu, S., Zhou, S., Wang, T., & Chen, H. (2016). Spatio-temporal assessment of urbanization impacts on ecosystem services: Case study of Nanjing City, China. *Ecological Indicators, 71*, 416–427. https://doi.org/10.1016/j.ecolind.2016.07.017

Liu, Y., Yu, Y., Wang, H., & Yu, P. (2022). Land surface temperature validation. *Field Measurements for Passive Environmental Remote Sensing: Instrumentation, Intensive Campaigns, and Satellite Applications*, 375–389. https://doi.org/10.1016/B978-0-12-823953-7.00016-2

Ma, Y., & Xu, R. (2010). Remote sensing monitoring and driving force analysis of urban expansion in Guangzhou City, China. *Habitat International, 34*(2), 228–235. https://doi.org/10.1016/j.habitatint.2009.09.007

Mazzocchi, C., Sali, G., & Corsi, S. (2013). Land use conversion in metropolitan areas and the permanence of agriculture: Sensitivity Index of Agricultural Land (SIAL), a tool for territorial analysis. *Land Use Policy, 35*, 155–162. https://doi.org/10.1016/j.landusepol.2013.05.019

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine, 8*(4), 283–298. https://doi.org/10.1016/S0001-2998(78)80014-2

Momoh, O. (2022). Population definition in statistics and how to measure it. *Investopedia*.

Mondal, B. K., & Das, K. (2010). Role of growth center: A rural development perspective. *Journal of Bangladesh Institute of Planners, 3*.

Moniruzzaman, M., Alam, M. S., Rashid, J. A., Begum, S. N., & Islam, M. M. (2012). MARKER-ASSISTED backcrossing for identification of salt tolerant rice lines. *2*(2), 1–8.

Morshed, N., Yorke, C., & Zhang, Q. (2017). Urban expansion pattern and land use dynamics in Dhaka, 1989–2014. *The Professional Geographer, 69*(3), 396–411. https://doi.org/10.1080/00330124.2016.1268058

Mundia, C. N., & Aniya, M. (2005). Analysis of land use/cover changes and urban expansion of Nairobi city using remote sensing and GIS. *International Journal of Remote Sensing, 26*(13), 2831–2849. https://doi.org/10.1080/01431160500117865

Noorhosseini, S. A., Allahyari, M. S., Damalas, C. A., & Moghaddam, S. S. (2017). Public environmental awareness of water pollution from urban growth: The case of Zarjub and Goharrud rivers in Rasht, Iran. *Science of the Total Environment, 599–600*(2017), 2019–2025. https://doi.org/10.1016/j.scitotenv.2017.05.128

Olivia, S., Boe-Gibson, G., Stitchbury, G., Brabyn, L., & Gibson, J. (2018). Urban land expansion in Indonesia 1992–2012: Evidence from satellite-detected luminosity. *The Australian Journal of Agricultural and Resource Economics, 62*(3), 438–456. https://doi.org/10.1111/1467-8489.12258

Patel, K. (2021). What is binary logistic regression and how is it used in analysis? *Dataversity*.

Pradhan, B., Abdullahi, S., & Al-sharif, A. A. A. (2017). Introduction to urban growth and expansion. In *Spatial modeling and assessment of urban form: Analysis of urban growth: From sprawl to compact using geospatial data* (pp. 3–15). Springer International Publishing. https://doi.org/10.1007/978-3-319-54217-1_1.

Prasanna, C. (2024). *Understanding ROC and AUC in model evaluation | by chanaka prasanna | medium*.

Pullanikkatil, D., Palamuleni, L. G., & Ruhiiga, T. M. (2016). Land use/land cover change and implications for ecosystems services in the Likangala River Catchment, Malawi. *Physics and Chemistry of the Earth, 93*, 96–103. https://doi.org/10.1016/j.pce.2016.03.002

Rahman, N. (2019). *Urban expansion analysis and land use changes in Rangpur city corporation area , Bangladesh , using remote sensing (RS) and geographic information system (GIS) techniques* (Vol. 2, pp. 22–30). https://doi.org/10.5281/zenodo.3526360, 10.

Rahman, M., Haque, T., Jahan Rakhi, T., Hossain Rafi, A., & Islam Oishi, S. (2024). *Carbon consequences of deforestation in khagrachhari of Chittagong hill tracts: A gis and remote sensing approach*. February.

Rahman, M. S., Mohiuddin, H., Kafy, A. Al, Sheel, P. K., & Di, L. (2019). Classification of cities in Bangladesh based on remote sensing derived spatial characteristics. *Journal of Urban Management, 8*(2), 206–224. https://doi.org/10.1016/J.JUM.2018.12.001

Richardson, E., Trevizani, R., Greenbaum, J. A., & Carter, H. (2024). Article the receiver operating characteristic curve accurately assesses imbalanced datasets the receiver operating characteristic curve accurately assesses imbalanced datasets. *Patterns, 5*(6), Article 100994. https://doi.org/10.1016/j.patter.2024.100994

Rimi, R. A., Rafi, A. H., & Sultana, J. (2024). *Factors affecting sustainable resource management in*. February.

Roy, S., Pandit, S., Eva, E. A., Bagmar, M. S. H., Papia, M., Banik, L., Dube, T., Rahman, F., & Razi, M. A. (2020). Examining the nexus between land surface temperature and urban growth in Chattogram Metropolitan Area of Bangladesh using long term Landsat series data. *Urban Climate, 32*(August 2019), Article 100593. https://doi.org/10.1016/j.uclim.2020.100593

Sadia, H., Kumar, S., & Haydar, M. (2023). Soil erosion susceptibility mapping in Bangladesh. *Ecological Indicators, 156*(October), Article 111182. https://doi.org/10.1016/j.ecolind.2023.111182

Sarkar, A., & Chouhan, P. (2020). Modeling spatial determinants of urban expansion of Siliguri a metropolitan city of India using logistic regression. *Modeling Earth Systems and Environment, 6*(4), 2317–2331. https://doi.org/10.1007/s40808-020-00815-9

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science, 2*(3), 1–21. https://doi.org/10.1007/s42979-021-00592-x

Seto, K. C., Güneralp, B., & Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences of the United States of America, 109*(40), 16083–16088. https://doi.org/10.1073/pnas.1211658109

Shi, K., Cui, Y., Liu, S., & Wu, Y. (2023). Global urban land expansion tends to Be slope climbing: A remotely sensed nighttime light approach. *Earth's Future, 11*(4). https://doi.org/10.1029/2022EF003384

Shi, G., Shan, J., Ding, L., Ye, P., Li, Y., & Jiang, N. (2019). Urban road network expansion and its driving variables: A case study of nanjing city. *International Journal of Environmental Research and Public Health, 16*(13). https://doi.org/10.3390/ijerph16132318

Singh, Y. (2013). SATPALDA. *Geospatial insight*.

Sruthi, E. R. (2024). Building a random forest model: A step-by-step guide. *Analytics Vidhya*.

Statistics, D. (2024). *Andy field*.

Su, S., Xiao, R., Mi, X., Xu, X., Zhang, Z., & Wu, J. (2013). Spatial determinants of hazardous chemicals in surface water of Qiantang River, China. *Ecological Indicators, 24*, 375–381. https://doi.org/10.1016/j.ecolind.2012.07.015

Sun, X., Crittenden, J. C., Li, F., Lu, Z., & Dou, X. (2018). Urban expansion simulation and the spatio-temporal changes of ecosystem services, a case study in Atlanta Metropolitan area, USA. *Science of the Total Environment, 622–623*, 974–987. https://doi.org/10.1016/j.scitotenv.2017.12.062

Trenberth, K. E. (2011). Changes in precipitation with climate change. *Climate Research, 47*(1–2), 123–138. https://doi.org/10.3354/cr00953

UCLA. (2020). *Logistic regression: SPSS annotated output*. Institute for Digital Research & Education.

USGS. (2022). *U.S. Geological survey*. US Geological Survey.

Verburg, P. H., Ritsema van Eck, J. R., de Nijs, T. C. M., Dijst, M. J., & Schot, P. (2004). Determinants of land-use change patterns in The Netherlands. *Environment and Planning B: Planning and Design, 31*(1), 125–150. https://doi.org/10.1068/b307

Vimal, R., Geniaux, G., Pluvinet, P., Napoleone, C., & Lepart, J. (2012). Detecting threatened biodiversity by urbanization at regional and local scales using an urban sprawl simulation approach: Application on the French Mediterranean region. *Landscape and Urban Planning, 104*(3–4), 343–355. https://doi.org/10.1016/j.landurbplan.2011.11.003

Wan, Z., Zhang, Y., Zhang, Q., & Li, Z. liang (2002). Validation of the land-surface temperature products retrieved from terra moderate resolution imaging spectroradiometer data. *Remote Sensing of Environment, 83*(1–2), 163–180. https://doi.org/10.1016/S0034-4257(02)00093-7

Weiss, R. (2012). *Scikit-learn : Machine learning in Python scikit-learn : Machine learning in Python*. January.

Yacouby, R. (2020). *Probabilistic extension of precision , recall , and F1 score for more thorough evaluation of classification models* (pp. 79–91).

Yang, K., Sun, W., Luo, Y., & Zhao, L. (2021). Impact of urban expansion on vegetation: The case of China (2000–2018). *Journal of Environmental Management, 291*, Article 112598. https://doi.org/10.1016/J.JENVMAN.2021.112598

Yang, Y., & Zhou, Q. (2014). *Gradient analysis of landscape spatial and temporal pattern changes in Beijing metropolitan area*. https://doi.org/10.1007/s11431-010-3206-2. January.

Yousaf Raza, M., Hasan, M. M., & Chen, Y. (2023). Role of economic growth, urbanization and energy consumption on climate change in Bangladesh. *Energy Strategy Reviews, 47*, Article 101088. https://doi.org/10.1016/J.ESR.2023.101088

Yue, W., Fan, P., Wei, Y. D., & Qi, J. (2014). Economic development, urban expansion, and sustainable development in Shanghai. *Stochastic Environmental Research and Risk Assessment, 28*(4), 783–799. https://doi.org/10.1007/s00477-012-0623-8

Yunhao, C., Xiaobing, L., Jing, L., & Peijun, S. (2003). Research on urban spatial thermal environment based on remote sensing and GIS. *International Geoscience and Remote Sensing Symposium (IGARSS), 3*(C), 1981–1983. https://doi.org/10.1109/igarss.2003.1294313

Zafarullah, H., & Ferdous, J. (2024). Metropolitan governance and interorganizational coordination: Study of a city corporation in Bangladesh. *Public Organization Review, 24*(1), 97–117. https://doi.org/10.1007/S11115-023-00749-7/METRICS

Zhang, K., Wang, R., Shen, C., & Da, L. (2010). Temporal and spatial characteristics of the urban heat island during rapid urbanization in Shanghai, China. *Environmental Monitoring and Assessment, 169*(1–4), 101–112. https://doi.org/10.1007/s10661-009-1154-8