Spatiotemporal Land Use and Land Cover Change Analysis (2005–2025) and Future Projections to 2040 in Prayagraj, India, Using Machine Learning and Google Earth Engine

Thammaboribal, P.,[1*] Agrawal, N.,[2] Srivastava, A.,[2] Maurya, M.[2] and Tripathi, N. K.[1]
[1]Remote Sensing and Geographic Information Systems FoS., Asian Institute of Technology, Pathumthani, Thailand, E-mail: prapasgnss@gmail.com
[2]Department of Computer Science, United University, Prayagraj, Uttar Pradesh, India
**Corresponding Author*

DOI: https://doi.org/10.52939/ijg.v22i10.5242

Abstract
Rapid urbanization has significantly transformed land use and land cover (LULC) patterns in Indian cities, creating a need for reliable spatial information to support urban planning and sustainable land management. This study investigates LULC dynamics in Prayagraj, Uttar Pradesh, India, from 2005 to 2025 and projects potential changes for 2030, 2035, and 2040 using multi-temporal Landsat imagery and Google Earth Engine (GEE). Four LULC classes: water bodies, vegetation, built-up land, and barren land were mapped using three supervised machine learning algorithms: Random Forest (RF), Support Vector Machine (SVM), and Classification and Regression Trees (CART). Classification performance was evaluated using overall accuracy (OA) and the Kappa coefficient. Among the three classifiers, SVM demonstrated the most consistent performance, achieving a mean OA of approximately 0.93 across the five reference years. The historical analysis revealed substantial landscape transformation, particularly the expansion of built-up areas and the overall decline in vegetation cover. The RF-based prediction model was evaluated by comparing a simulated LULC map for 2025 with the corresponding classified map derived from satellite imagery. Although the model reproduced several broad spatial patterns, it overestimated built-up and barren land in some areas, indicating uncertainty in the projected results. The future projections suggest increasing built-up land and declining vegetation between 2030 and 2040, while water bodies and barren land exhibit more variable patterns. These findings demonstrate the potential of integrating satellite remote sensing, machine learning, and cloud-based geospatial processing for long-term LULC assessment. However, the projected changes should be interpreted cautiously because of classification uncertainty, spatial-resolution constraints, and limitations in representing the complex processes governing urban expansion. Further validation and comparison with alternative spatial prediction models are necessary to improve the reliability of future LULC scenarios.

Keywords: CART, Google Earth Engine, LULC prediction, Prayagraj, RF, SVM, Urban expansion

1. Introduction

Rapid urbanization, population growth, and changes in land use have significantly altered urban landscapes, making land use and land cover (LULC) monitoring an important component of urban and regional planning [1] and [2] Conventional land surveys rely heavily on field measurements and manual data collection, which can be time-consuming and difficult to maintain consistently over large areas and extended periods. In contrast, satellite remote sensing provides repeated observations of the Earth's surface, allowing changes in land use and land cover to be mapped and compared across different time periods [3]. The integration of satellite imagery, cloud-based geospatial platforms, and machine learning algorithms has further expanded the capacity to monitor these changes and assess potential future development patterns [4].

Prayagraj, located in Uttar Pradesh, India, provides an important setting for investigating long-term land use and land cover change. Formerly known as Allahabad, the city is one of India's historically significant urban centres, with a strong religious, cultural, and administrative identity. Its location at the Sangam, where the Ganga and Yamuna rivers meet, has made it a major pilgrimage destination for centuries [5]. The city also hosts the

Kumbh Mela, one of the world's largest religious gatherings, which attracts millions of pilgrims and places substantial demands on local infrastructure, transportation, accommodation, and public services [6] . In addition to its religious importance, Prayagraj serves as a centre for administration, judicial institutions, and education, contributing to continued urban development [7] .

The combination of cultural significance, riverine geography, and expanding urban activities makes Prayagraj a relevant case for examining the spatial and temporal dynamics of LULC change [8] . Urban expansion, infrastructure development, and the increasing demand for residential and commercial land can alter the distribution of built-up areas, vegetation, and other land cover types. The periodic organisation of the Kumbh Mela also creates additional demands for temporary facilities and supporting infrastructure, although the extent to which these activities contribute to long-term land cover change requires systematic assessment. Meanwhile, the conversion of vegetated and undeveloped land into built-up areas may affect green space availability and the conservation of natural resources, particularly along the city's expanding urban fringe and river corridors [9] . Identifying where these changes occur, how rapidly they progress, and how land cover patterns may develop in the future can provide useful evidence for urban planning and sustainable land management.

This study examines LULC changes in Prayagraj district over the 20-year period from 2005 to 2025 using multi-temporal Landsat imagery and Google Earth Engine. Three supervised machine learning classifiers; Random Forest (RF), Support Vector Machine (SVM), and Classification and Regression Trees (CART) are compared to assess their performance in LULC classification. The resulting maps are then used to analyse changes in land cover over the study period and project future LULC patterns to 2040. By combining historical change analysis with future projections, the study aims to improve understanding of the spatial dynamics of land use change in Prayagraj and provide information that may support future urban and regional planning.

2. Literature Review

Machine learning algorithms, notably Random Forest (RF), Support Vector Machine (SVM), and Classification and Regression Trees (CART) have become foundational to satellite-based land use/land cover (LULC) classification [30] [31] and [32]. While SVM resolves complex, high-dimensional boundaries and CART provides interpretable rule-based partitions, ensemble approaches like RF mitigate overfitting and improve classification stability. However, classifier performance remains heavily contingent upon training data quality, feature selection, and local spectral separability. Spectral overlap among built-up surface classes can reduce classification accuracy, while the RF classifier helps mitigate overfitting and improve classification performance [10] , bare soil, and sparse vegetation frequently introduces errors in heterogeneous urban settings, making localized benchmarking and rigorous accuracy assessment via confusion matrices, user's/producer's accuracies, and cross-temporal validation essential to prevent propagating spurious land transitions.

The emergence of cloud-based platforms, particularly Google Earth Engine (GEE), has facilitated large-scale, multi-temporal LULC analyses by providing direct access to petabyte-scale archives and standardized preprocessing pipelines, including cloud masking and image compositing [11]. These tools enforce uniform processing workflows across temporal intervals, substantially reducing sensor- and scene-level discrepancies [12] and [13]. Beyond historical mapping, LULC research increasingly integrates predictive modeling such as Cellular Automata-Markov (CA-Markov) [14] , Artificial Neural Network-coupled CA (CA-ANN) [15], CA-Logistic Regression, CA-SVM, CA-MLP, CA-RF [16] and time-series approaches like ARIMA to simulate future spatial trajectories [17]. These models combine probabilistic transition matrices with spatial drivers, including transport corridors and expressway infrastructure, which heavily steer urban expansion into agricultural and undeveloped land. Because projections inherently reflect the parameterization and historical assumptions of their parent models, robust calibration and validation remain imperative when employing simulated LULC scenarios for spatial planning and policy formulation.

Prayagraj, formerly known as Allahabad, provides a relevant setting for examining urban expansion and changes in land cover in a riverine environment. Located at the confluence of the Ganga and Yamuna rivers, the city has considerable religious, cultural, and administrative importance. Its urban development is associated with demand for residential areas, commercial activities, transport infrastructure, and public services. Previous studies have examined the conversion of agricultural and peri-urban land into built-up areas, indicating changes in the spatial distribution of land cover associated with urban growth [8] [9] and [18] . Other studies have investigated changes in impervious surfaces and urban land cover, highlighting the importance of monitoring development patterns to support land-use planning [19] [20] and [21].

While satellite remote sensing and machine learning have advanced urban land use/land cover (LULC) monitoring, consistently evaluating long-term dynamics in Prayagraj requires addressing key methodological challenges. First, classifier performance varies with local landscape characteristics and class spectral overlap, meaning results from Random Forest (RF), Support Vector Machines (SVM), and Classification and Regression Trees (CART) cannot be transferred without localized benchmarking [22] . Second, multi-temporal change detection remains vulnerable to false transitions caused by sensor, acquisition, or training data discrepancies, necessitating a standardized processing pipeline to isolate [23]. This study utilizes Google Earth Engine to evaluate multi-temporal Landsat imagery across Prayagraj district from 2005 to 2025. We systematically compare RF, SVM, and CART classifiers under a unified framework to quantify historical LULC trajectories and project future spatial patterns through 2040. These findings establish a reliable, standardized baseline of urban expansion to support evidence-based urban and regional planning.

3. Methodology

3.1 Study Area

Prayagraj (formerly Allahabad) is situated in the southern part of Uttar Pradesh, India, between latitudes 24°50′N to 25°45′N and longitudes 81°40′E to 82°20′E (Figure 1), at an average elevation of 98 meters above mean sea level [24].

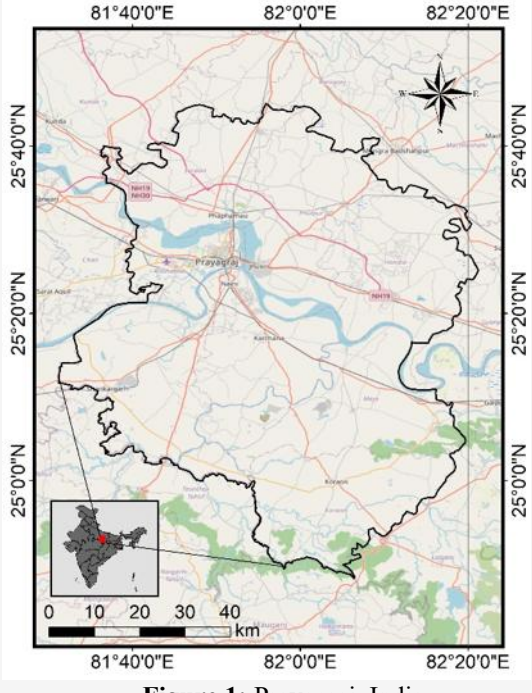


Figure 1: Prayagraj, India

Geographically positioned within the fertile alluvial plains of the Indo-Gangetic Basin, the city lies at the sacred confluence of the Ganga, Yamuna, and the mythical Saraswati rivers. The region experiences a humid subtropical climate, characterized by hot summers, a monsoon rainy season delivering an average annual rainfall of approximately 950-1,020 mm [25] , and mild winters [26]. Over the past three decades, Prayagraj has evolved from a historic administrative, educational, and pilgrimage hub into a rapidly expanding urban center under the Smart Cities Mission. This transition has catalyzed swift shifts in LULC. The geographic constraints posed by the Ganga and Yamuna River channels have historically directed spatial growth along distinct linear and peri-urban corridors, resulting in the widespread conversion of prime alluvial farmland, riparian vegetation, and wetlands into impervious built-up terrain [5]. Concurrently, periodic socio-religious events most notably the Kumbh and Maha Kumbh congregations exert massive, recurring infrastructural pressure on riverbank floodplains, driving short- and long-term land modifications.

3.2 Methods

The methodology in Figure 2 outlines a multi-stage remote sensing and spatial modeling workflow for analyzing and projecting Land Use and Land Cover (LULC) changes in Prayagraj. The workflow begins with data acquisition and pre-processing, where Landsat 5, 8, and 9 satellite imagery covering the Prayagraj boundary from 2005 to 2025 was obtained. These satellite scenes are pre-processed within the Google Earth Engine (GEE) Code Editor using cloud masking, mosaicking, and spatial clipping. Following this, training samples are manually delineated, collecting 50 sample points across 4 distinct classes for a total of 200 training points.

The subsequent phase involves image classification and validation. Supervised machine learning classifications are implemented using Random Forest (RF), Support Vector Machine (SVM), and Classification and Regression Tree (CART) algorithms. Classification performance is evaluated through accuracy assessment metrics including confusion matrices, Overall Accuracy (OA), and the Kappa Coefficient. The optimal classifier for each year is retained based on threshold benchmarks of OA ≥ 0.85 and Kappa ≥ 0.80 [27]. In the final phase, spatiotemporal change detection is conducted across the multi-decadal reference timeframe from 2005 to 2025. Using a Random Forest model, future LULC dynamics are simulated and predicted for the target years 2030, 2035, and 2040. Lastly, the resulting data are exported as PNG,

CSV, and GeoTIFF formats to facilitate cartographic visualization and statistical reporting.

3.3 Data Used and Processing

Multi-temporal satellite imagery was acquired from the Landsat series of Earth observation satellites, selected based on data availability, spatial resolution, and temporal continuity throughout the study period (2005–2025 with 5 years interval). The satellite platforms and spectral bands used varied across the study years, depending on data availability and image quality. Landsat 5 Thematic Mapper (TM) imagery was used for 2005 and 2010, Landsat 8 Operational Land Imager (OLI) imagery was used for 2015, and Landsat 9 imagery was used for 2025. Landsat is uniquely suited for spatiotemporal analysis primarily because of its continuous, multi-decadal historical archive and standardized sensor continuity spanning several satellite generations, including Landsat 5, 8, and 9 [28] and [29]. This uninterrupted temporal timeline allows researchers to capture long-term environmental trends and LULC trajectories across consistent multi-year or decadal intervals. Furthermore, its balanced 30-meter spatial resolution and consistent multi-spectral bands provide sufficient detail to monitor regional-scale landscape transformations such as urban expansion, vegetation fragmentation, and hydrological shifts without the extreme computational overhead required by very high-resolution datasets. When integrated into cloud computing platforms like Google Earth Engine, Landsat's standardized, pre-processed collections enable highly scalable, reproducible, and uniform change detection across vast spatial and temporal scales.

All image preprocessing, image classification, accuracy assessment, and change detection procedures were performed using Google Earth Engine (GEE), a cloud-based platform with a browser-accessible integrated development environment that enables large-scale satellite image processing and geospatial analysis through JavaScript programming. The use of GEE eliminated the need to store large raster datasets locally and facilitated efficient, reproducible, and scalable geospatial processing.

3.4 LULC Classification

Four broad land use/land cover (LULC) classes were defined for Prayagraj to represent the dominant land cover types within the study area. The classification scheme, summarized in Table 1, was applied consistently across all five reference years. Supervised classification requires a labelled reference dataset comprising known LULC locations, commonly referred to as training samples. In this study, training samples were generated through manual digitization in the Google Earth Engine Code Editor. Reference points were manually placed on satellite imagery for each of the four LULC classes, based on their spectral characteristics and prior knowledge of the study area.

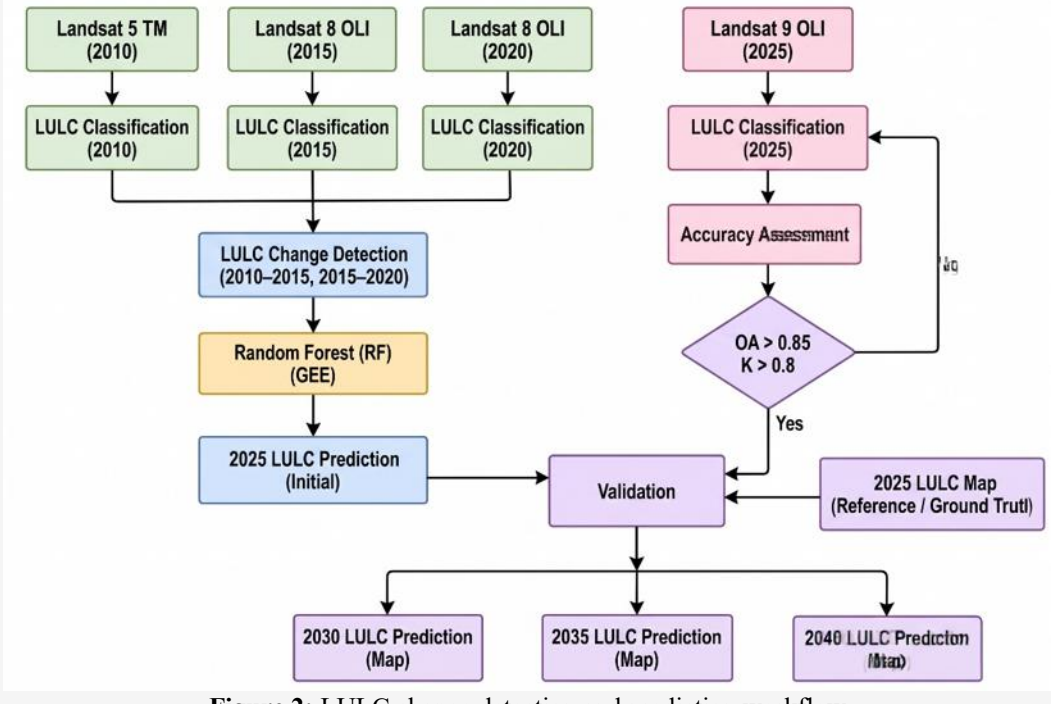


Figure 2: LULC change detection and prediction workflow

Table 1: LULC classification scheme adopted for the study

Class No.	LULC Class	Description
1	Water Bodies	Rivers, ponds, reservoirs and other perennial or seasonal water surfaces.
2	Vegetation	Forest cover, dense and sparse vegetation identified using vegetation indices such as NDVI.
3	Build-up	Residential, commercial, industrial and other constructed/impervious surfaces.
4	Barren	Fallow land, cropland and exposed soil falling under the above categories

Approximately 50 reference points were collected for each LULC class, yielding a total training dataset of approximately 200 points. Each point was assigned the corresponding class label and associated with the spectral band values of the corresponding pixel to construct the labelled feature collection used to train the classifier. The complete dataset of 200 labelled points was randomly divided into two subsets: a training subset used to train the classifier and an independent testing subset withheld from the training process and used exclusively for accuracy assessment. This partitioning approach follows a common practice in supervised image classification.

3.5 LULC Classification

The images were obtained from the manually mapped points and was used to train and compare three widely used supervised machine learning techniques within the Google Earth Engine: Random Forest (RF), Support Vector Machine (SVM), and Regression Tree (CART). Each algorithm was applied independently to the same set of training and validation samples for every reference year, allowing their classification performance to be directly compares and the most suitable classifier selected for final mapping

3.5.1 Random forest

Random Forest (RF) is an ensemble learning algorithm that constructs multiple decision trees using randomly selected subsets of the training data and predictor variables. Each tree generates a prediction, and the final classification is determined by majority voting among the individual trees [30]. By combining multiple decision trees, RF generally reduces the risk of overfitting compared with a single decision tree and performs effectively on large, complex datasets [31]. However, its ensemble structure makes the model less interpretable than a single decision tree [32].

3.5.2 Support vector machine

Support Vector Machine (SVM) is a margin-based classification algorithm that identifies an optimal hyperplane to separate different classes while maximizing the margin between them. The data points closest to the decision boundary, known as support vectors, play a critical role in defining the optimal hyperplane [33] . SVM can also classify nonlinearly separable data using kernel functions, which enable the modelling of nonlinear relationships in the feature space [31]. This algorithm is particularly effective for small- to medium-sized datasets with high-dimensional features. However, its performance depends on appropriate parameter selection, and its computational requirements may increase with dataset size [34].

3.5.3 Classification and regression tree

Classification and Regression Tree (CART) is a decision tree algorithm that recursively partitions a dataset into smaller subsets based on predictor variables, producing a tree-like structure. For classification tasks, each terminal node assigns a class label based on the majority class of the observations within that node, whereas regression tasks estimate continuous values based on the distribution of the target variable [35] . CART is relatively simple to implement and interpret and can accommodate both categorical and numerical predictor variables. However, individual decision trees are prone to overfitting and may exhibit lower predictive performance than ensemble methods, such as Random Forest, particularly when the data are complex [36].

3.6 Accuracy Assessment

The reliability of the classified LULC maps generated by the three classification models: RF, SVM, and CART was quantitatively evaluated. Two standard accuracy metrics were used to evaluate classification performance. Overall accuracy (OA) represents the proportion of correctly classified validation samples relative to the total number of validation samples. It provides an overall measure of classification performance across all LULC classes. A minimum OA of 0.85 was established as the acceptance threshold [37]. The Kappa coefficient (κ)

is a statistical measure of agreement between the classification results and the reference data that accounts for agreement expected by chance. It complements overall accuracy by incorporating both observed agreement and the agreement expected under random classification. A minimum κ of 0.80 was established as the acceptance threshold [15].

3.7 Change Detection

Change detection was performed by comparing LULC maps from different reference years to identify the nature, direction, and magnitude of land cover transformations within the study area. Pixel-based comparisons and class-area statistics were used to assess changes between successive and non-successive reference years throughout the study period (2005–2025), thereby quantifying the gains and losses associated with each LULC category. This analysis facilitated the identification of long-term land cover trends, including potential decreases in forest and vegetation cover and corresponding increases in built-up and urban areas over the study period.

3.8 LULC Prediction

Multi-temporal LULC maps for 2010, 2015, 2020, and 2025, representing five-year intervals, were used as the baseline dataset for predictive modelling. The 2010–2025 period was used to examine long-term historical LULC changes, while the 2015–2020 interval served as the initial training period for model development and validation. The 2020–2025 interval was subsequently used to calibrate the model for future LULC projections.

Future LULC patterns were projected using the Random Forest (RF) machine learning algorithm. Rather than classifying individual satellite images independently, the model was designed to learn LULC transitions between two time steps and use the resulting relationships to simulate future land cover patterns [17] and [38].

The model was initially trained using the 2015–2020 transition period and a stratified random sample of 3,000 points, of which 80% were allocated to training and 20% to testing. The trained model was then applied to the 2020 LULC map to simulate the 2025 LULC state. The simulated map was evaluated against the actual satellite-derived 2025 LULC map using 5,000 independent spatial validation points. Model performance was assessed using an overall accuracy (OA), and the Kappa coefficient. Following validation, the model was recalibrated using the more recent 2020–2025 transition period and another stratified random sample of 3,000 points. The recalibrated model was then applied iteratively to project future LULC states at five-year intervals. Specifically, the actual 2025 LULC map served as the baseline for the 2030 projection, the predicted 2030 map served as the baseline for the 2035 projection, and the predicted 2035 map served as the baseline for the 2040 projection.

4. Results and Discussions

4.1 Accuracy Assessment

The multi-temporal performance of the three supervised machine learning algorithms was quantitatively assessed across five discrete observation intervals (2005, 2010, 2015, 2020, and 2025) using OA and the Kappa coefficient (κ) are presented in Table 2. All three classifiers satisfied the accepted threshold for LULC mapping, with OA consistently exceeding 0.85 and κ remaining above 0.80 throughout the 20-year observation period. However, substantial variations in classification fidelity, inter-annual stability, and predictive consistency were observed among the algorithms.

Table 2: LULC classification accuracy assessment

Year	Metric	CART	RF	SVM
2005	OA	0.95	0.93	0.93
	Kappa	0.90	0.90	0.93
2010	OA	0.89	0.87	0.91
	Kappa	0.88	0.82	0.85
2015	OA	0.87	0.85	0.93
	Kappa	0.80	0.91	0.83
2020	OA	0.87	0.86	0.90
	Kappa	0.82	0.87	0.93
2025	OA	0.95	0.96	0.96
	Kappa	0.94	0.94	0.93

Support Vector Machine (SVM) demonstrated the highest overall performance and longitudinal robustness across the monitoring period. SVM attained the highest Overall Accuracy in four out of five evaluation years (2010, 2015, 2020, and tied for highest in 2025), achieving a multi-year mean OA of 0.93 and a mean κ of 0.89. The superior capability of SVM in handling complex, high-dimensional spectral feature spaces is evident in its stability; its OA never fell below 0.90 across any epoch. This resilience is attributable to the structural risk minimization principle inherent to SVM, which optimizes the separating hyperplane between classes rather than relying on orthogonal recursive splitting, thereby mitigating misclassification along spectral transition zones and mixed pixels.

In contrast, the ensemble-based Random Forest (RF) classifier exhibited greater variability across intermediate observation windows. Although RF achieved its maximum classification accuracy in 2025 (OA = 0.96, and κ = 0.94) and demonstrated high statistical agreement in 2015 ($\kappa = 0.91$), its

performance declined in 2010 (OA = 0.87) and 2015 (OA = 0.85). Consequently, RF yielded a multi-year mean OA of 0.89 and a mean Kappa coefficient of 0.89. While RF remains widely acknowledged for its capacity to reduce overfitting relative to single-tree architectures, its comparatively lower accuracy in intermediate years suggests sensitivity to class imbalance or localized spectral overlap within intermediate-stage Landsat imagery.

The single-tree CART algorithm produced competitive results in initial and terminal baseline years, achieving an OA of 0.95 in both 2005 and 2025. Nonetheless, CART recorded the lowest multi-year average agreement (mean $\kappa = 0.87$; mean OA = 0.91), exhibiting pronounced susceptibility to inter-class spectral confusion during transitional years (e.g., $\kappa = 0.80$ in 2015). This performance trajectory corroborates established empirical findings: individual decision trees are inherently prone to higher variance and sub-optimal boundary delineation in landscapes characterized by high spatial heterogeneity and spectral fragmentation. The comparative evaluation identifies SVM as the most statistically rigorous and longitudinally consistent classifier for long-term LULC characterization in this study. While RF performed comparably to SVM in the final monitoring period (2025), the persistent multi-year accuracy and higher average agreement of SVM establish it as the most reliable algorithm for tracking multi-decadal landscape dynamics.

4.2 Spatiotemporal Dynamics of LULC (2005–2025)

The classified LULC maps in Figure 3 provide a spatial view of how land cover has changed across Prayagraj over the study period. The maps were produced using the SVM classifier, which showed the highest classification performance among the three methods evaluated. Comparison of the five years allows the main patterns of expansion and land-cover conversion to be examined, particularly the growth of built-up areas around the urban core and the associated changes in vegetation and barren land. The multi-temporal spatial distribution of LULC across Prayagraj over the 20-year observation period (2005–2025) illustrats four distinct thematic classes were delineated: Water, Built-up, Vegetation, and Bareland. The classified series reveals pronounced landscape transitions characterized by sustained urban intensification, spatial reorganization of agrarian and barren zones, and hydrological continuity along the primary drainage corridors.

4.2.1 Urban expansion and built-up intensification

In the baseline year of 2005, built-up surfaces were predominantly concentrated within the compact historical core near the confluence (*Sangam*) of the Ganga and Yamuna river systems, with secondary clusters limited to prominent transport arteries and localized southern enclaves. Over the subsequent observation windows (2010–2020), built-up areas expanded radially outward from the core. By 2025 (Figure 3(e)), substantial peri-urban consolidation and ribbon development are evident, forming a continuous, dense urban footprint that encroaches heavily into the adjacent vegetative and fallow lands to the north, west, and southeast. Secondary built-up clusters in the south and southwest also demonstrate progressive spatial coalescence.

4.2.2 Vegetation dynamics and landscape fragmentation

Vegetation represents the primary land cover of the study area across all five time steps, maintaining extensive spatial coverage across both northern and southern tracts. However, the vegetation exhibits progressive localized fragmentation, particularly along the urban fringes and transport corridors where agricultural parcels have been converted to impervious surfaces. The 2010 interval (Figure 3(b)) shows a temporary, widespread dispersion of bare/fallow pixels across the vegetation background, likely reflecting seasonal cropping cycles, post-harvest conditions, or inter-annual precipitation variability during satellite data acquisition. By 2020 and 2025, the agricultural matrix displays increased homogeneity, though with clear localized retreats around expanding development nodes.

4.2.3 Bareland fluctuations and peripheral transitions

Bareland dynamics exhibit distinct temporal fluctuations throughout the study period. During 2005–2015, bare surfaces were largely distributed as diffuse, fine-scale patches across agricultural areas, representing active fallow fields and localized excavation. While during 2020–2025, the bareland shifts into contiguous, prominent zones, most notably concentrated along the southeastern periphery and in proximity to the river meanders. This prominent expansion of bareland in 2025 frequently adjacent to newly developed built-up sectors, suggests extensive land preparation, infrastructural earthworks, and speculative clearing preceding further urban sprawl.

4.2.4. Hydrological Stability

The river network, defined by the meandering courses of the Ganga and Yamuna rivers, maintains consistent geometric alignment throughout the 2005–2025 timeline.

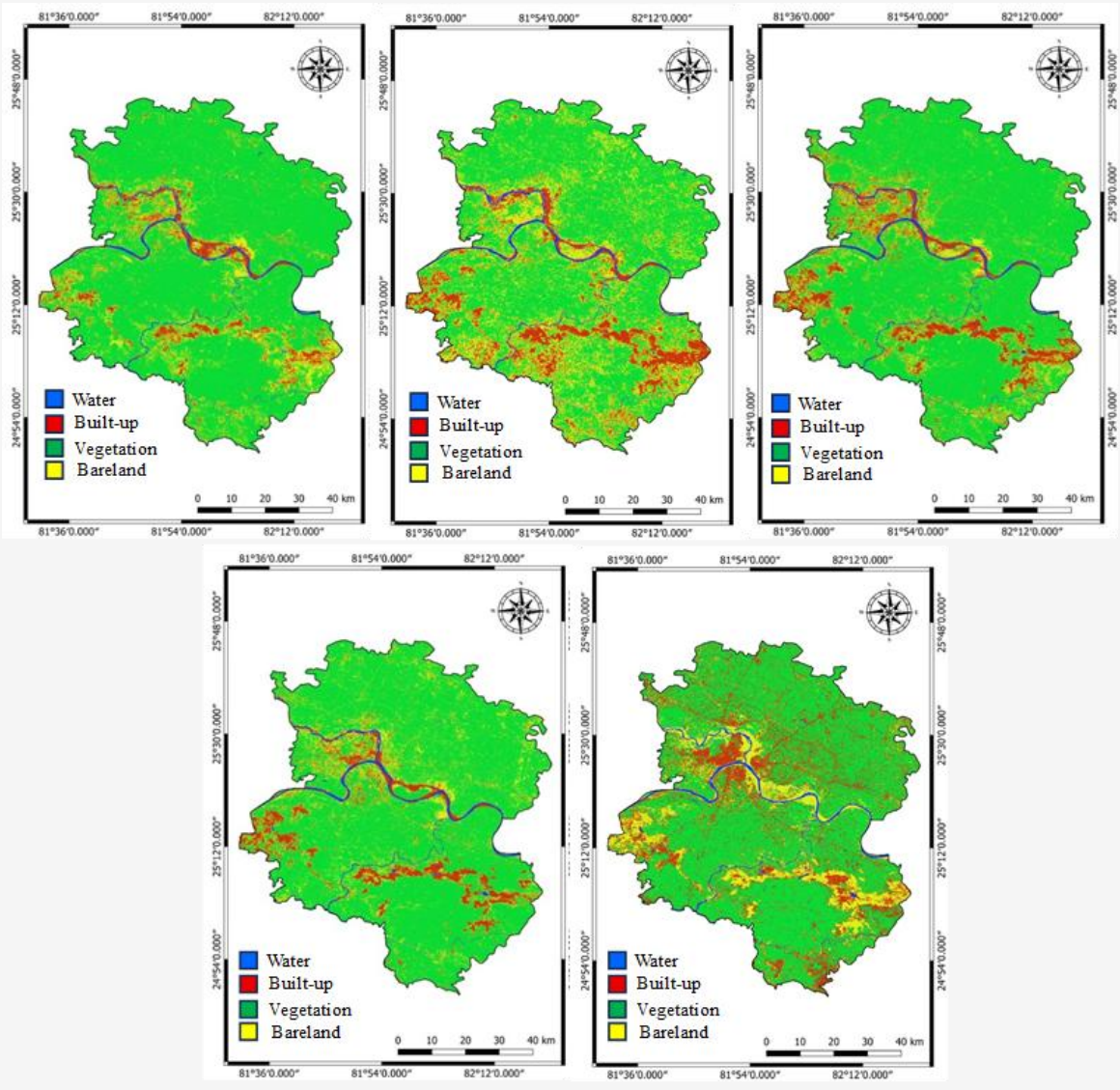


Figure 3: Spatial distribution of LULC: (a) 2005, (b) 2010, (c) 2015, (d) 2020, and (e) 2025

Variations within the water class are primarily confined to lateral sandbar emergence, channel braiding, and fluctuating active-channel widths. These subtle shifts reflect natural riverbed sedimentation and hydrologic seasonality at the dates of image capture rather than long-term geomorphic diversion.

4.3 LULC Variation During 2005 to 2025

The spatiotemporal distribution of LULC across the study area reveals substantial landscape transformation between 2005 and 2025. Across a constant total area of 5,463.58 km^2 the landscape is primarily dominated by vegetation, though human-induced transformations, predominantly urbanization exert increasing pressure on natural covers over the 20-year period. The LULC change is illustrated in Figure 4. According to Figure 4, vegetation constituted the largest LULC category, accounting for 78.37% (4,281.69 km^2) of the total landscape in 2005 and remaining above 56% across all epochs. Conversely, waterbodies represented the smallest and most stable class, initially comprising 1.61% (87.94). The remaining surface was partitioned between bareland (765.57 km^2; 14.01%) and built-up land (328.38 km^2; 6.01%). Over the 20-year span, the landscape experienced two distinct transformation phases: an initial phase of widespread surface disturbance and land clearance leading up to 2010, followed by a secondary phase of rapid, consolidated urban expansion by 2025. Built-up area exhibited the most pronounced net growth among all classes.

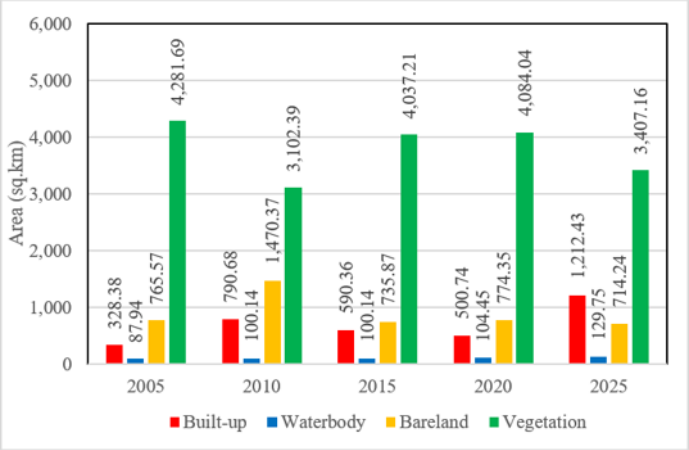


Figure 4: LULC changes from 2005 to 2025

Table 3: Summary of decadal LULC change

LULC	Area (km^2)		Area change
	2005	2025	(km^2 [%])
Built-up	328.38	1,212.43	+ 884.05 [+269.22]
Waterbody	87.94	129.75	+41.81 [+47.54]
Bareland	765.57	714.24	-51.33 [-6.70]
Vegetation	4,218.69	3,407.16	-874.53 [-19.24]

Between 2005 and 2010, built-up surfaces expanded by 140.78% (from 328.38 to 790.68). Although a temporary contraction or reclassification occurred between 2010 and 2020 decreasing to 590.36 km^2 in 2015 and 500.74 km^2 in 2020, the class underwent an unprecedented surge between 2020 and 2025, reaching a study-period peak of 1,212.43 km^2. Overall, built-up land expanded by 269.22% (+884.05 km^2) relative to the 2005 baseline, elevating its landscape share from 6.01% to 22.19%. This dramatic inflection in 2025 indicates accelerated infrastructural development, peri-urban encroachment, or residential sprawl.

Vegetation cover exhibited a dynamic, non-linear pattern characterized by strong reciprocal interactions with bareland and built-up features. During 2005–2010, vegetation contracted sharply from 4,281.69 km^2 to 3,102.39 km^2 (a net loss of 1,179.30 km^2; -27.54%). Concurrently, bareland nearly doubled, surging from 765.57 km^2 to a maximum of 1,470.37 km^2 (+92.06%). This synchronicity strongly suggests widespread land clearing, timber harvesting, or preparatory site grading for urban/agricultural projects. Following the 2010 disruption, vegetation demonstrated substantial recovery, expanding back to 4,037.21 km^2 in 2015 and 4,084.04 km^2 in 2020. This regeneration corresponded with a return of bareland to near-baseline levels (~735–774 km^2), reflecting either ecological succession, afforestation efforts, or the maturation of agricultural canopies. Between 2020 and 2025, vegetation contracted again to 3,407.16 km^2 (-676.88 km^2), directly mirroring the exponential growth of built-up areas. By 2025, vegetation experienced a cumulative net loss of 874.53 km^2 (-20.42%) compared to 2005, highlighting permanent canopy replacement by impervious surfaces.

Waterbodies maintained a persistent, monotonic upward trend across the two decades, increasing from 87.94 km^2 in 2005 to 100.14 (stable across 2010 and 2015), 104.45 km^2 in 2020, and culminating at 129.75 km^2 in 2025. This represents an overall net increase of 47.54% (+41.81 km^2). In the context of regional environmental studies, such steady increases in surface water are frequently attributed to anthropogenic hydrological interventions (e.g., retention pond construction, aquaculture pond excavation, reservoir development) or localized drainage changes induced by urban runoff. The spatiotemporal dynamics observed between 2005 and 2025 in Table 3 demonstrate a landscape undergoing pronounced anthropogenically driven restructuring, characteristic of rapidly growing urban centers across the Indo-Gangetic Plain. Situated at the confluence of the Ganga and Yamuna rivers, Prayagraj historically functioned within an agro-ecological matrix dominated by fertile alluvial plains and expansive agricultural vegetation. Over the two-decade observation period, this baseline underwent marked transformation, primarily governed by the aggressive expansion of built-up surfaces. This urban

growth trajectory aligns with the city's strategic elevation as an administrative, commercial, and religious hub. In particular, peri-urban infrastructural initiatives such as the expansion of the Naini industrial and residential corridor across the Yamuna, arterial transport corridor upgrades, and large-scale permanent civil works executed to accommodate massive pilgrim influxes during the Ardh Kumbh and Mahakumbh events have functioned as primary catalysts driving physical urban expansion [39].

The steady outward progression of built-up land has occurred overwhelmingly at the expense of regional green cover. The continuous net contraction of vegetation corroborates empirical observations from regional remote-sensing literature, which document widespread fragmentation and permanent conversion of prime agricultural tracts, riparian grasslands, and orchards into impervious surfaces [40]. Concurrently, the behavior of bareland characterized by a pronounced mid-period surge followed by consolidation toward near-baseline levels exemplifies the typical land-transition lifecycle observed in urbanizing Indian landscapes [41]. In this trajectory, vegetated and agricultural parcels undergo initial clearance, remaining temporarily fallow or barren during speculative acquisition and phased site preparation, before being permanently occupied by dense infrastructure.

Conversely, surface waterbodies exhibited an upward trend over the study period. Within the local context of Allahabad, this trajectory reflects deliberate hydrological management and river restoration efforts undertaken across the region. Initiatives under the Namami Gange mission, urban drainage channel improvements, and the revival and construction of community retention ponds and artificial wetlands under state groundwater recharge schemes have contributed to expanded surface-water retention [42]. Furthermore, the complex hydromorphology of the braided Ganga–Yamuna channel system, coupled with seasonal shifts in post-monsoon discharge and sandbar inundation during satellite image acquisition periods, likely contributed to the observed expansion in detectable surface water [43].

4.4 LULC Prediction Model Validation

To assess the reliability of the prediction model before projecting future LULC scenarios, a map for 2025 was generated using the historical classified LULC maps from 2005 to 2020 and compared with the actual classified map for 2025 as presented in Figure 5. The predicted map generally captured the overall spatial patterns of land-cover change observed in the actual 2025 map, particularly the continued expansion of built-up areas around the central urban core and the increased presence of barren land in the western and southern parts of the district. However, the predicted map overestimated the extent of both built-up and barren land relative to the actual 2025 classification, suggesting that the model projects a more rapid expansion of these land-cover classes than occurred during the observed period. Consequently, the projected LULC patterns for 2030–2040 should be interpreted with caution, as they may represent an accelerated-growth scenario rather than precise predictions of future land-cover conditions.

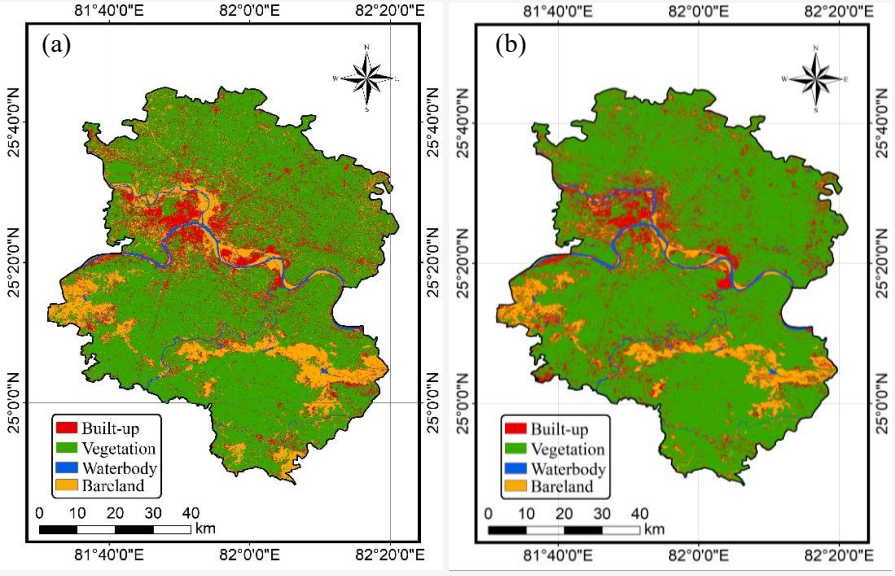


Figure 5: LULC in 2025 (a) actual LULC, and (b) predicted LULC

The spatial comparison between the actual LULC map for 2025 (Figure 5(a)) and the predicted LULC map (Figure 5(b)) illustrates the predictive capability and spatial fidelity typical of land-change simulation framework using RF algorithm. Overall, the model demonstrates high spatial consistency in capturing the macro-scale geographic configuration of dominant land categories. Vegetation remains the prevailing matrix across the landscape, while the primary hydrological features notably the main river channels and their confluences are preserved with high geometric fidelity. The model accurately captures the core urban agglomeration and major peripheral built-up clusters around the central river bend. The concentric expansion of the primary urban core, along with linear development along infrastructural corridors, is realistically reproduced. In both figures, the core built-up zone exhibits substantial contiguity, reflecting persistent anthropogenic pressure and conversion of surrounding agricultural or vegetated land into urban surfaces.

Notable spatial discrepancies emerge between the actual and simulated states, predominantly regarding the distribution of bareland and peri-urban built-up patches. In the predicted map (Figure 5(b)), the model tends to smooth out finer-scale landscape fragmentation. The actual LULC map (Figure 5(a)) exhibits larger, contiguous patches of bareland across the south-central/southeastern corridor and a more pronounced cluster along the western border. In contrast, the simulated map underestimates or fractures these continuous bareland expanses, partially misallocating them as surrounding vegetation matrix or over-predicting small, scattered built-up/bareland pixels. These variations underscore the intrinsic challenges of modeling dynamic land-cover classes subject to both natural seasonality and localized socio-economic interventions. While the simulation successfully validates the macro-scale growth vectors and long-term urban trajectories, localized discrepancies highlight the necessity of integrating finer-scale socio-economic drivers, zoning regulations, and multi-temporal remote sensing datasets to capture stochastic land-conversion processes more effectively.

4.5 LULC Prediction

Figure 6 illustrates the predicted LULC patterns in Prayagraj, India, for 2030, 2035, and 2040. The predictions indicate the continued predominance of vegetation across the study area, accompanied by the persistence of built-up development concentrated primarily in the central urban area and along the river corridor. Bareland remains particularly evident in the southwestern and southern portions of the study area, while the waterbody network is characterized by the main river and its associated channels. Although the overall spatial configuration of the four LULC classes remains broadly consistent across the three predicted years, local variations in their distribution suggest continued land-cover transitions over the projection period. In 2030 (Figure 6(a)), vegetation constitutes the visually dominant land-cover class, covering extensive areas in the northern, eastern, and central portions of the study area, as well as substantial parts of the south. Built-up areas are concentrated around the urban core, particularly near the river corridor, where development forms a relatively complex and fragmented pattern. Bareland is prominent in the southwestern sector and occurs in several patches across the southern portion of the study area.

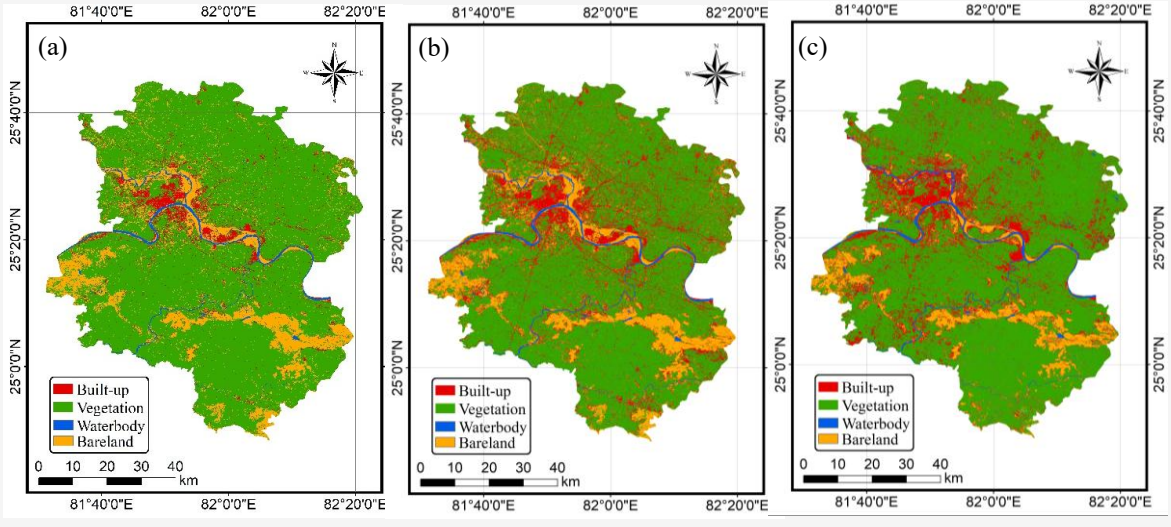


Figure 6: LULC prediction: (a) 2030, (b) 2035, and (c) 2040

The main river remains a distinct linear feature separating parts of the urbanized and vegetated landscape. This configuration suggests that future land-cover dynamics will involve both the continued consolidation of existing urban areas and localized transitions among vegetation, bareland, and built-up land.

By 2035 (Figure 6(b)), the overall LULC pattern remains similar to that predicted for 2030. Vegetation continues to dominate the landscape, although localized differences in the distribution of built-up areas and bareland are visible. The urban core and its surrounding areas retain a concentration of built-up land, while scattered patches of development extend into adjacent areas. Bareland remains concentrated in the southwestern and southern sectors, with additional fragmented patches distributed across the landscape. These patterns suggest that urban development may continue through the expansion and densification of existing built-up areas, together with localized land-cover conversion at the urban fringe. However, because the maps provide categorical spatial distributions rather than numerical transition statistics, the magnitude and rate of these changes cannot be established from visual interpretation alone.

The predicted LULC pattern for 2040 (Figure 6(c)) retains the broad spatial characteristics observed in 2030 and 2035, with vegetation remaining the dominant class and built-up areas concentrated around the central urban zone. Localized differences in the distribution of built-up land and bareland are apparent, particularly around the urban core and in portions of the southwestern and southern landscape. The persistence of extensive vegetated areas indicates that vegetation is expected to remain a major component of the regional landscape throughout the prediction period. Nevertheless, the continued presence of fragmented built-up patches within or adjacent to vegetated areas highlights the potential for localized land-cover conversion. The main river remains a prominent landscape feature in all three projections, indicating that the broad spatial configuration of the river corridor is maintained in the predicted scenarios.

The predicted spatial patterns are broadly consistent with previous LULC modelling studies conducted in Prayagraj [44], using a multilayer perceptron–Markov chain (MLP–MC) approach, reported substantial growth in built-up land alongside reductions in agricultural/open land and forest cover. Their projections also indicated continued built-up expansion toward the northern, southwestern, and southern parts of the study area. Similarly, a cellular automata–artificial neural network (CA–ANN) model was utilized to predict further urban expansion in Prayagraj, with agricultural and forest land facing increasing pressure from urban development. The findings provide a regional context for interpreting the persistence of built-up clusters and their surrounding land-cover mosaics in the present predictions [45]. Further evidence is provided by the SLEUTH-based urban growth modelling study of Prayagraj, which reported an increase in built-up area from 40.22 km² in 1990 to 85.89 km² in 2020 and projected further expansion by 2040. The study identified road accessibility and terrain-related factors as important influences on the spatial distribution of urban growth [46]. In addition, the changes in land cover in Prayagraj between 2000 and 2020 were associated with increased built-up and cropland areas, declining tree cover, and a rise in land surface temperature[41]. Taken together, these studies highlight the importance of examining the potential environmental implications of future urban expansion, particularly its interactions with vegetation and other open-land classes.

In the present study, the continued dominance of vegetation in the 2030–2040 projections contrasts with the documented expansion of built-up land in previous studies of Prayagraj. This difference should not necessarily be interpreted as a contradiction, because the studies may differ in their geographical boundaries, classification schemes, historical training periods, spatial resolutions, and modelling methods. Moreover, the four classes used in the present study aggregate land-cover types that other studies distinguish separately. Consequently, direct comparisons of class proportions or rates of change require consistent class definitions and study-area boundaries. From a land-management perspective, the predicted patterns emphasize the importance of monitoring built-up expansion and its potential encroachment into vegetated and bareland areas. Maintaining vegetation within and around expanding urban areas is particularly relevant given the relationship between land-cover composition and land surface temperature documented by [41]. The predicted maps can therefore provide a spatial basis for identifying areas that warrant closer monitoring and for informing future urban and environmental planning. Figure 7 presents the predicted changes in the areas of four (LULC) classes in Prayagraj, India, for 2030, 2035, and 2040. The results indicate contrasting trends among the classes, particularly a substantial expansion of urban land and a continuous decline in vegetation. Urban land is predicted to increase from 302.54 km² in 2030 to 914.05 km² in 2035 and 945.53 km² in 2040. This represents a net increase of 642.99 km² (212.5%) over the ten-year projection period.

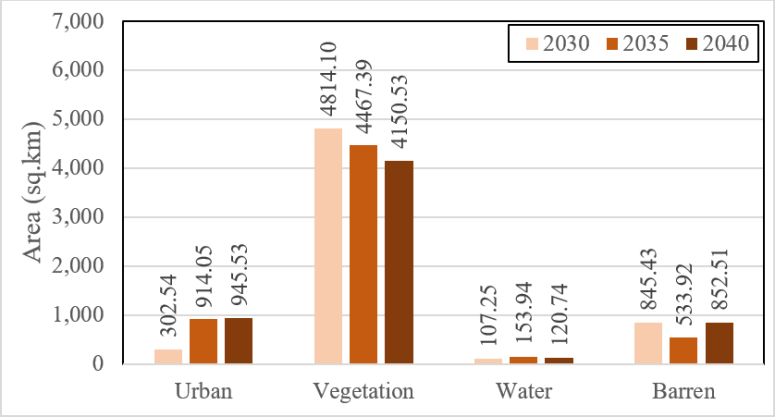


Figure 7: The LULC area changes derived from predicted LULC

The most pronounced increase occurs between 2030 and 2035, followed by a considerably smaller increase between 2035 and 2040. This pattern suggests that the predicted urban expansion is concentrated in the earlier projection interval, with a slower rate of growth during the subsequent five years. In contrast, vegetation is projected to decline consistently from 4,814.10 km² in 2030 to 4,467.39 km² in 2035 and 4,150.53 km² in 2040. The total reduction of 663.57 km² represents a decrease of approximately 13.8% over the projection period. Despite this decline, vegetation remains the largest LULC class throughout all three prediction years. The simultaneous increase in urban land and decrease in vegetation suggests a potential association between urban expansion and the conversion of vegetated areas. However, the area statistics alone do not establish that vegetation is the direct source of the additional urban land; a class-to-class transition matrix would be required to quantify these conversions. Nevertheless, the projected trends highlight the potential pressure of urban development on vegetation and the importance of incorporating green-space conservation into future land-use planning.

Water and barren land exhibit more variable temporal patterns. The predicted water area increases from 107.25 km² in 2030 to 153.94 km² in 2035, before declining to 120.74 km² in 2040. Despite this fluctuation, the water area in 2040 remains approximately 12.6% higher than in 2030. Barren land decreases from 845.43 km² in 2030 to 533.92 km² in 2035, followed by a substantial recovery to 852.51 km² in 2040. Consequently, its net area changes only slightly over the entire period, increasing by 7.08 km² (approximately 0.8%). These fluctuations may reflect the predicted redistribution of land-cover classes, although their underlying causes cannot be established from the area statistics alone. Further examination of the transition matrix and spatial prediction maps is necessary to identify the specific land-cover conversions responsible for these changes. Overall, the predicted LULC patterns suggest that urban expansion and vegetation loss will be the principal changes in Prayagraj between 2030 and 2040. These trends have implications for urban planning, vegetation conservation, and the management of land resources. However, the particularly large increase in urban area between 2030 and 2035 warrants careful examination against the historical change rates, model validation results, and class-transition probabilities. Because the urban increase is unusually pronounced relative to the subsequent interval, the classification, projected class areas, and model assumptions should be checked before interpreting the magnitude of this change as a plausible future trajectory.

5. Conclusion

This study evaluated the long-term spatiotemporal evolution of land use and land cover across Prayagraj District, India, over two decades (2005–2025) and projected spatial scenarios through 2040 using cloud-based geospatial workflows. By benchmarking three machine learning algorithms, SVM demonstrated the greatest robustness and spatial fidelity for long-term multi-temporal classification. The historical analysis captured extensive landscape restructuring, headlined by an expansive 269.22% growth in impervious built-up terrain that progressively encroached upon prime agricultural lands and riparian vegetation. While predictive modeling with Random Forest successfully mapped macro-scale growth vectors and outward peri-urban trajectories, it also revealed persistent developmental pressure on the region's ecological buffers and riverine corridors.

Several methodological limitations should be addressed in interpreting these outcomes. First, relying on the 30-meter moderate spatial resolution of Landsat imagery inevitably introduces spectral

mixing along complex urban-rural transition zones and river sandbars, potentially masking fine-grained infrastructural developments and localized vegetation canopies. Second, the 4-class thematic scheme groups distinct agricultural regimes, dense forests, and open orchards into a single aggregate vegetation category, obscuring internal agro-ecological transitions. Furthermore, the transition-based Random Forest predictive model tended to smooth localized landscape fragmentation and slightly overestimated the pace of built-up expansion during validation. Future research should incorporate higher-resolution satellite observations (such as Sentinel-2 or high-resolution commercial optical constellations) to mitigate mixed-pixel anomalies in dense urban fabrics. To improve predictive simulation fidelity, integrating spatial dynamic models, most notably Cellular Automata coupled with Artificial Neural Networks (CA-ANN) and Cellular Automata-Markov chain (CA-Markov) models is strongly recommended. Coupling these models with spatially explicit driving factors, such as transport accessibility, digital elevation models, flood-hazard zones, and socio-economic master plans, will provide more nuanced, scenario-driven projections to guide resilient infrastructure and land conservation policies.

References

[1] MohanRajan, S. N., Loganathan, A. and Manoharan, P., (2020). Survey on Land Use/Land Cover (LU/LC) Change Analysis in Remote Sensing and GIS environment: Techniques and Challenges, *Environmental Science and Pollution Research*. Vol. 27(24); 29900–29926. https://doi.org/10.1007/s11356-020-09091-7.

[2] Hidayati, I. and Balqis, H., (2026). Urban Growth Dynamics of Special Region of Yogyakarta, Indonesia (1994–2024): Population Trends and Expansion Prediction. *International Journal of Geoinformatics*, Vol. 22(9); 71–86. https://doi.org/10.52939/ijg.v22i9.5192.

[3] Rahul, K. R., (2024). Peri-urban Delineation and Urban Expansion Quantification from 2001 to 2021 of Hisar City, India, using Geospatial Techniques. *International Journal of Geoinformatics*, Vol. 20(4); 86–99. https://doi.org/10.52939/ijg.v20i4.3155.

[4] Thammaboribal, P., Tripathi, N., Lipiloet, S. and Mandadi, R., (2025). Flood Mapping and Damage Assessment using UN-SPIDER Recommended Practices in Google Earth Engine: A Case Study of the 2024 Chiang Rai Flood, Thailand. *International Journal of Geoinformatics*, Vol. 21(3); 165–179. https://doi.org/10.52939/ijg.v21i3.4039.

[5] Vishvkarma, M. and Siddiqui, A., (2026). Trend Analysis of Land Use and Land Cover (LULC) along the Ganga River Corridor (2 Km Buffer) of Prayagraj City during 1985- 2025. *International Journal of Science and Social Science Research*. Vol. 4; 155–166. https://doi.org/10.63671/ijsssr.v4i1.602.

[6] Sarif, Md. O. and Gupta, R. D., (2022). Evaluation of seasonal Ecological Vulnerability using LULC and Thermal State Dynamics using Landsat and MODIS Data: A Case Study of Prayagraj City, India (1987–2018), *Environmental Science and Pollution Research*. Vol. 29(51); 77502–77535. https://doi.org/10.1007/s11356-022-21225-7.

[7] Vishvkarma, M. and Siddiqui, A. R., (2026). Trend Analysis of Land Use and Land Cover (LULC) along the Ganga River Corridor (2 Km Buffer) of Prayagraj City during 1985- 2025. *International Journal of Science and Social Science Research.* Vol. 4(1). https://doi.org/10.63671/ijsssr.v4i1.602.

[8] Tiwari, A. and Siddiqui, A., (2025). A Spatio-Temporal Analysis of Land Use/Land Cover Changes in Prayagraj City, India, *International Journal of Educational Innovations and Research*, Vol. 1; 132–141. https://doi.org/10.63671/ijeir.v1i4.94.

[9] Kumar, V. and Agrawal, S., (2019). Agricultural Land Use Change Analysis using Remote Sensing and GIS: A Case Study of Allahabad, India. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 397–402. https://doi.org/10.5194/isprs-archives-XLII-3-W6-397-2019, 2019.

[10] Sheppert, A., (2026). Techniques for Mitigating Overfitting in Machine Learning: A Comprehensive Review, Taxonomy, and Practical Guide. *Frontiers in Artificial Intelligence*, Vol. 9. https://doi.org/10.3389/frai.2026.1794271.

[11] Herndon, K. E., Griffin, R., Schroder, W., Murtha, T., Golden, C., Contreras, D. A., Cherrington, E., Wang, L., Bazarsky, A., Kollias, G. and Alcover Firpi, O., (2023). Google Earth Engine for Archaeologists: An Updated Look at the Progress and Promise of Remotely Sensed Big Data. *Journal of Archaeological Science: Reports*, Vol. 50. https://doi.org/10.1016/j.jasrep.2023.104094.

[12] Chuengpituk-u-dom, P., Tripathi, N. and Thammaboribal, P., (2025). A Comparative Study of Flood Monitoring Techniques Using the UN-SPIDER Recommendations and a Generative AI-Based Model (SATGPT): A Case Study in Ayutthaya, Thailand. *International Journal of Geoinformatics*, Vol. 21(10); 116–128. https://doi.org/10.52939/ijg.v21i10.4535.

[13] Jagtap, A., Shedge, D. and Mane, P., (2024). Exploring the Effects of Land Use/Land Cover (LULC) Modifications and Land Surface Temperature (LST) in Pune, Maharashtra with Anticipated LULC for 2030. *International Journal of Geoinformatics,* Vol. 20(2); 42–63. https://doi.org/10.52939/ijg.v20i2.3065.

[14] Aldileemi, H., Zhran, M. and El-Mewafi, M., (2023). Geospatial Monitoring and Prediction of land Use/Land Cover (LULC) Dynamics Based on the CA-Markov Simulation Model in Ajdabiya, Libya. *International Journal of Geoinformatics*, Vol. 19(2); 15–29. https://doi.org/10.52939/ijg.v19i12.2973.

[15] Thammaboribal, P. and Tripathi, N., (2024). Predicting Land Use and Land Cover Changes in Pathumthani, Thailand: A Comprehensive Analysis from 2013 to 2023 Using Landsat Satellite Imagery and CA-ANN Algorithm, with Projections for 2028 and 2038, *International Journal of Geoinformatics*, Vol. 20(5); 13–27. https://doi.org/10.52939/ijg.v20i5.3225.

[16] Balqis, H., Hidayati, I. and Susilo, B., (2026). Comparison of Built-Up Expansion Prediction Models for Magelang City and Its Peri Urban Areas: CA-Logistic Regression, CA-SVM, CA-MLP, and CA-RF. *International Journal of Geoinformatics.* Vol. 22(5); 89–112. https://doi.org/10.52939/ijg.v22i5.4982.

[17] Qazi, M. R., Lone, F. A., Ara, S., Khan, I., Ashraf, I., Yatu, K., Ahmed, R., Malik, I. H., Bhat, I. F. and Majid, M., (2026). Assessing and Forecasting LULC Changes in Sindh Forest Range, Kashmir Himalayas, using Remote Sensing, ARIMA, and Random Forest Models. *World Development Sustainability*. Vol. 9. https://doi.org/10.1016/j.wds.2026.100323.

[18] Sarif, Md. O.and Gupta, R. D., (2022). Spatiotemporal Mapping of Land Use/Land Cover Dynamics using Remote Sensing and GIS Approach: A Case Study of Prayagraj City, India (1988–2018). *Environment, Development and Sustainability*. Vol. 24(1); 888–920. https://doi.org/10.1007/s10668-021-01475-0.

[19] Srivastava, P. K., Pant, T. and Singh, C. K., (2024). Decadal Urban Growth Analysis of Prayagraj City During 2015-2024 Using Landsat-8 Multispectral Data, *2024 IEEE 11th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*; 1–5. https://doi.org/10.1109/UPCON62832.2024.10983762.

[20] Baba, K., Lal, D. and Bello, A., (2019). Application of Remote Sensing and GIS Techniques in Urban Planning, Development and Management. *International Journal of Scientific and Engineering Research*, Vol. 10. 1127-1134.

[21] Kumar, P. and Siddiqui, A. R., (2025). Detection of Land Use/ Landcover in Prayagraj City. *International Journal of Current Research*, Vol. 17(5); 32980–32986. https://doi.org/10.24941/ijcr.48941.05.2025.

[22] Awwal, Y., Isma'il, M., Ya'u, S., Lemuel Musa, M. and Ahmad, H., (2026). Comparative Assessment of Machine Learning Classifiers for Land Use/Land Cover Mapping in a Tropical Savanna Landscape, *Kaduna Journal of Geography*. Vol. 8; 565–578. https://doi.org/10.47514/kjg.2026.08.01.055.

[23] Sui, H., Feng, W., Li, W., Sun, K. and Xu, C., (2018). Review of Change Detection Methods for Multi-temporal Remote Sensing Imagery, *Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics and Information Science of Wuhan University*. Vol. 43; 1885–1898. https://doi.org/10.13203/j.whugis20180251.

[24] Aliyu Kasim, A., (2015). Derivation of Surface Soil Water Content Using a Simplified Geometric Method in Allahabad District, Uttar Pradesh India, *International Journal of Scientific and Engineering Research*, Vol. 6; 1631–1638.

[25] Swarnim, Tripathi, J. N., Sonker, I. and Singh, R., (2026). Combining MCDM and Geospatial Techniques to Identify Groundwater Potential Zones and Trend Analysis of Rainfall and Well Water Level Data: An Investigation in the Prayagraj and Kaushambi Districts, *Geosystems and Geoenvironment*, Vol. 5(1). https://doi.org/10.1016/j.geogeo.2025.100454.

[26] Sarif, Md. O., Gupta, R. D. and Murayama, Y., (2023). Assessing Local Climate Change by Spatiotemporal Seasonal LST and Six Land Indices, and Their Interrelationships with SUHI and Hot–Spot Dynamics: A Case Study of Prayagraj City, India (1987–2018), *Remote Sensing*. Vol. 15(1). https://doi.org/10.3390/rs15010179.

[27] Seeboonruang, U., Mandadi, R., Thammaboribal, P., Gonzales, A. L. and Ganni, S. V. S. A. B., (2026). Land Use Classification, Prediction, and the Relationship Between Land Use and Sediment Loss in the Lam Phra Phlong Watershed, Thailand, *Agriculture*, Vol. 16(4); https://doi.org/10.3390/agriculture16040448.

[28] Thammaboribal, P., (2024). Investigating Land Surface Temperature Variation and Land Use Land Cover Changes in Pathumthani, Thailand (1997-2023) using Landsat Satellite Imagery: A Comprehensive Analysis of LST and Urban Hot Spots (UHS). *International Journal of Geoinformatics*. Vol. 20(2); 27–41. https://doi.org/10.52939/ijg.v20i2.3063.

[29] Thammaboribal, P., Triapthti, N., and Lipiloet, S., (2025). Comparative Analysis of Land Surface Temperature (LST) Retrieved from Landsat Level 1 and Level 2 Data: A Case Study in Pathumthani Province, Thailand, *International Journal of Geoinformatics*. Vol. 21(4); 160–177. https://doi.org/10.52939/ijg.v21i4.4089.

[30] Breiman, L., (2001). Random Forests, *Machine Learning*, Vol. 45, 5–32.

[31] Salman, H., Kalakech, A. and Steiti, A., (2024). Random Forest Algorithm Overview. *Babylonian Journal of Machine Learning*. Vol. 2024; 69–79. https://doi.org/10.58496/BJML/2024/007.

[32] Halabaku, E. and Bytyçi, E., (2024). Overfitting in Machine Learning: A Comparative Analysis of Decision Trees and Random Forests, *Intelligent Automation & Soft Computing*, Vol. 39; 987–1006. https://doi.org/10.32604/iasc.2024.059429.

[33] Shmilovici, A., (2023). Support Vector Machines, In: Maimon, O., Rokach, L. (eds) Data Mining and Knowledge Discovery Handbook. Springer, Boston. https://doi.org/10.1007/0-387-25465-X_1293–110.

[34] Ghaddar, B., and Naoum-Sawaya, J., (2017). High Dimensional Data Classification and Feature Selection using Support Vector Machines, *European Journal of Operational Research*, Vol. 265. https://doi.org/10.1016/j.ejor.2017.08.040.

[35] Lewis, R., (2000). An Introduction to Classification and Regression Tree (CART) Analysis. Department ofEmergency Medicine. Harbor-UCLA Mewdical Center Torrance, California.

[36] Mankar, A., Bhoite, S., Kharade, K. and Raskar, K., (2024). Metaanalysis of Overfitting of Decision Trees. *Journal of Nonlinear Analysis and Optimization*. Vol. 15(1); 30-36.

[37] Aji, A., Husna, V. and Purnama, S., (2024). Multi-Temporal Data for Land Use Change Analysis Using a Machine Learning Approach (Google Earth Engine). *International Journal of Geoinformatics*. Vol. 20(4); 19–28. https://doi.org/10.52939/ijg.v20i4.3145.

[38] Pan, X., Hou, J., Guo, Q., Li, B., Zhou, Q., Li, D., Li, X., Chen, G., Wang, Y., Xu, L. and Wang, T., (2027). Real-Time Urban Flood Forecasting using ML-Based Surrogate Models: Enhancing Smart and Sustainable City Decision-Making, *Expert Systems with Applications*, Vol. 332. https://doi.org/10.1016/j.eswa.2026.133461.

[39] Gupta, G., Singh, R. and Dimri, A. P., (2026). Society and Culture Development in India Tradition Meets Transformation: Economic and Social Impacts of Maha Kumbh 2025 in Prayagraj, *Society and Culture Development in India*, Vol. 5; 225–241. https://doi.org/10.47509/SCDI.2025.v05i02.02.

[40] Mohan, M., Pathan, S., Narendrareddy, K., Kandya, A. and Pandey, S., (2011). Dynamics of Urbanization and Its Impact on Land-Use/Land-Cover: A Case Study of Megacity Delhi, *Journal of Environmental Protection*, Vol. 02; 1274–1283. https://doi.org/10.4236/jep.2011.29147.

[41] Srivastava, A., Shukla, S., Singh, P. and Jha, P. K., (2023). Spatio-Temporal Dynamics of Land Use/Cover and Land Surface Temperature in Prayagraj City, India. *Indoor and Built Environment*. Vol. 32; 1250–1268. https://doi.org/10.1177/1420326X231159633.

[42] Simon, M. and Joshi, H., (2022). Story of the Ganga River: Its Pollution and Rejuvenation. *Riverine System*. 21–55. https://doi.org/10.1007/978-3-030-87067-6_2.

[43] Punja, V., Pathak, V., Shukla, S. P., Tripathi, R. P. and Dheeraj, V. P., (2026). Analyzing the Spatiotemporal Impact of Land Use/Land Cover Change on Surface Water Quality of Middle Ganga Region, India. *Environment, Development and Sustainability*. https://doi.org/10.1007/s10668-026-08194-4.

[44] Kumar, V. and Agrawal, S., (2023). A Multi-Layer Perceptron–Markov Chain Based LULC Change Analysis and Prediction using Remote Sensing Data in Prayagraj District, India. *Environmental Monitoring and Assessment*. Vol. 195(5). https://doi.org/10.1007/s10661-023-11205-w.

[45] Sarif, Md. O., Gupta, R. D. and Sharifi, A., (2024). Predicting Prayagraj's Urbanization Trajectory using CA-ANN Modelling: Population Pressures and Land Use Dynamics. *Journal of Environmental Management*, Vol. 370. https://doi.org/10.1016/j.jenvman.2024.122469.

[46] Kumar, V. and Agrawal, S., (2023). Urban Modelling and Forecasting of Landuse using SLEUTH Model. *International Journal of Environmental Science and Technology (Tehran)*, Vol. 20(6); 6499–6518. https://doi.org/10.1007/s13762-022-04331-4.