Pertanika J. Sci. & Technol. 34 (3): 1905 - 1932 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Quantitative Evaluation of Machine Learning-based Land Use Classification with Multispectral Satellite Data

Krati Bansal[1]*, Anindita Nath[2], Tanupriya Choudhury[1], Bappaditya Koley[2], and Jitendra Rajpurohit[3]

[1]Department of Computer Science and Engineering, School of Computer Science, University of Petroleum and Energy Studies (UPES), 248007 Dehradun, Uttarakhand, India
[2]Department of Geography, Bankim Sardar College, 743329 South 24 Parganas, West Bengal, India
[3]Department of Computer Science and Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Lavale Campus, 412115 Pune, Maharashtra, India

ABSTRACT

Land use and land cover modification have been observed at an intense level in the Rupnarayan basin, West Bengal, India, due to the rapid expansion of fallow land and built-up area, and the decrease in water bodies and vegetation area. However, monitoring and quantifying the effects of rural catchment area modification is challenging. The heterogeneous landscape of the Rupnarayan catchment area was assessed using multispectral satellite data for 2000, 2010 and 2020 to identify land-cover change. Quantitative Evaluation of various Machine Learning techniques is adopted for the current research work, i.e., Support Vector Machine (SVM), Random Forest (RF) and Maximum Likelihood Classifier (MLC). The novelty of this work lies in benchmarking a parametric (MLC) classifier against two non-parametric classifiers (RF and SVM) within a single Google Earth Engine workflow and validating every annual map with both the Kappa statistic and AUC-ROC analysis for a previously unassessed rural reach of the basin. Five features have been taken for assessing the change detection over the region, i.e., built-up, agricultural land, fallow land, vegetation and water body. After performing each ML algorithm, SVM performed the best result among others with the overall accuracy rate of 97% (2000), 95% (2010) and 95% (2020) and a Kappa value of 0.95, while the other two scored in kappa 0.89 and 0.87 in MLC and RF, respectively. The analysis has been cross-checked by the AUC-ROC curve model. In 2000, the AUC value obtained 0.778 in MLC, RF AUC is 0.689, and SVM AUC is 0.946. In 2010, the MLC AUC is 0.875, the RF AUC value is 0.766, and the SVM value is obtained as 0.969.

ARTICLE INFO

Article history:
Received: 28 November 2025
Accepted: 08 June 2026
Published: 25 June 2026

DOI: https://doi.org/10.47836/pjst.34.3.26

E-mail addresses:
kratibansal1992@gmail.com (Krati Bansal)
aninditan286@gmail.com (Anindita Nath)
tanupriya@ddn.upes.ac.in (Tanupriya Choudhury)
bappadityakoley2012@gmail.com (Bappaditya Koley)
jitendra.rajpurohit@sitpune.edu.in (Jitendra Rajpurohit)
* Corresponding author

e-ISSN: 2231-8526

Finally, the 2020 map of each algorithm has been applied for the AUC curve, where the MLC curve value is 0.898, the RF value is 0.794, and the SVM obtained an AUC of 0.986. From the entire validation model, it has been revealed that SVM acquired the maximum AUC value for each year. This research will be beneficial for future researchers, policymakers, and land planners in enhancing management strategies for sustainable development.

Keywords: AUC-ROC, kappa coefficient, landsat images, LULC, maximum likelihood classifier, random forest, support vector machine

INTRODUCTION

Landscape contains heterogeneous factors with diverse geological, biological, social, ecological, hydrological and environmental characteristics. Land cover with local ecosystems is the primary component of the Earth's surface (Forman, 1995). Land cover is the biophysical cover of the Earth's surface, that is, vegetation and non-vegetation cover. A Land use/land cover (LULC) analysis denotes the contrast of several anthropogenic activities, i.e., built-up, mining, agriculture, etc. Spatial distribution and changes have been assessed by the quantitative LULC analysis (Ramachandra & Bharath, 2012), which also indicates the human interaction with the environmental phenomena.

Natural landscapes have changed due to anthropogenic activities, the growth of the population and climatic change over the globe. Anthropogenic activities have gradually risen over the biosphere, resulting in a large-scale alteration of the land surface of the earth, which impacts the global environment (Lambin et al., 2001). Changing of the land use is significant for effective environmental and watershed management (Twisa & Buchroithner, 2019). In developing countries, rapid land use and land cover changes (Hegazy & Mosbeh, 2015) have caused various vital natural resources to be reduced, such as vegetation, water and soil (Cheruto et al., 2016). Changes like urbanisation, expansion of the agricultural land, deforestation have a directly impacts on infiltration of groundwater, runoff and evapotranspiration (Dires & Temesgen, 2020; Ullah et al., 2024a). Climate, natural disasters and patterns of socio-economic dynamics at the global and local levels negatively affect land use and land cover change (Chakilu & Moges, 2017; Mahata et al., 2024; Ullah et al., 2024b). Ecosystem services are broadly defined as the functions that humans directly and indirectly derive from ecosystems, with terrestrial carbon storage and its consequences being widely studied and one of the most important ecosystem services (IPCC, 2006). In recent decades, land use change resulting from anthropogenic activities, namely timber and food production, has led to changes in the land use ecosystem (Fan et al., 2016).

LULC modify the land surface, fragmenting the contiguous vegetation into small patches that destroy the sustenance of carbon sedimentation, biodiversity and several geo-environmental sequences at the regional to global scale (Liu et al., 2019; Riitters et al., 2012; Ramachandra & Kumar, 2011). The Earth's surface has faced significantly decreasing

the vegetation cover due to a large amount of anthropogenic activities. Consequently, the human activities gradually increased along with river catchment areas, which caused changing are observed in land cover dynamics and the budget of surface temperature, producing changes in global, regional and local climates (Islam & Islam 2013). These dynamic changes have impacts on agriculture, the ecosystem and human society in the river basin. Present research has been carried out on the dynamic changes of the Land use and land cover (LULC) and its relations to the distributions of surface temperature (Ahmed et al. 2013). Several methods have been applied for the analysis and estimation of LULC and Land Surface Temperature (LST) (Rozenstein et al., 2014).

In recent decades, a wide range of Remote sensing and GIS classifiers technologies have been used for changes of Land used and Land covers dynamics and mapping. Multispectral spatial and temporal satellite data provide information about the land cover condition of any specific landscape, which is significant to manage the resources, monitor the modification of land cover and ecosystem, as well as mapping of LULC with change detection (Chowdhury, 2023; Farhan et al., 2024a; Manjunatha & Basavarajappa, 2020). From the recent past, various machine learning (ML) approaches (Khan et al., 2021; Shafi et al., 2023) are employed to monitor and assess the modification of LULC, i.e., Random Forest (RF) (Fu & Zhang, 2022), Support Vector Machine (SVM) (Liu et al., 2022), Maximum Likelihood Classifier (MLC) (Batar et al., 2017), Decision Trees (DT) (Zenouzi et al., 2022), Artificial Neural Network (ANN) (Raczko & Zagajewski, 2017). Among the other ML algorithms, SVM has reflected the most accurate result in limited training samples (Guo et al., 2019). The current study area is a very dynamic and rural developing catchment. In this landscape condition, SVM is a more suitable algorithm for assessing the LULC classification. But the study also applies two other ML algorithms to compare the accuracy of SVM, i.e., RF and MLC. Research in Malaysia mangrove forest has proven that SVM accuracy was 90% in overall accuracy, which is better than the MLC with 70% accuracy in overall accuracy (Deilmai et al., 2014; Halder et al., 2021). The current study aims to analyse and quantify the spatial and temporal modification of LULC in the diverse catchment area of the Rupnarayan catchment for 20 years of time span (2000 - 2020). Such land surface is the best classifier of the SVM algorithm to assess the LULC (Yousefi et al., 2022; Ge et al., 2020). The SVM algorithm has the best potential to classify the features with limited training samples (Bouslihim et al., 2022). Three classifiers were purposely chosen to represent the principal supervised families - a parametric statistical classifier (MLC), an ensemble tree-based classifier (RF) and a margin-based kernel classifier (SVM). CART is not treated separately because RF is itself an ensemble of CART trees and therefore subsumes it, while ANN, gradient boosting and deep-learning classifiers were not adopted because the limited and partly historical (2000 and 2010) training samples available for this rural basin make such data-hungry models prone to overfitting and difficult to tune reliably within the GEE environment (Fe & Zhang, 2022).

The rest of these methods, supervised classifiers, are most significant and widely applied for approaches of LULC dynamics changes (Chowdhury, 2023; Farhan et al., 2024a; Kogo et al., 2019; Langat et al., 2019). Remote-sensing classification models draw on multi-temporal, spatial, spectral and multi-sensor information, but every classified image inevitably carries some misclassification error (Chowdhury, 2026; Farhan et al., 2024b). Accuracy assessment, or verification, is performed as a validation step in remote sensing data processing, as it determines the informational value of the generated maps for end users. Effective utilisation of geodata is feasible only when the quality of the data is understood. The overall accuracy of the classified image evaluates how each pixel is classified compared to the actual land cover conditions derived from corresponding ground truth data. The producer's accuracy assesses errors of omission, indicating how effectively real-world land cover types can be classified. Conversely, commission error in measuring user's accuracy, reflecting the probability that a classified pixel corresponds accurately to the actual land cover type at its real-world location. Error metrics and the kappa coefficient have become ideal for evaluating image classification accuracy. The current accuracy assessment has also adopted the AUC-ROC curve model for cross-validation to validate the generated result by three specific supervised algorithms for each year 2000, 2010 and 2020. In the ROC curve, a statistical graphical representation has been prepared to show the accuracy of algorithms.

In the Rupnarayan catchment, very little research work has been conducted to assess the LULC using several ML algorithms. As the study is based on the Quantitative Evaluation of ML algorithms, the chief objectives of the current research have aimed i) to identify the LULC modification over the basin area; ii) to understand the change detection of the study area; iii) finally, to evaluate the performance of classified algorithms. "To the best of knowledge, no prior study of the Rupnarayan catchment has benchmarked a parametric classifier (MLC) against two non-parametric machine-learning classifiers (RF and SVM) within a single cloud-based (Google Earth Engine) workflow and then cross-validated every annual map with the AUC-ROC technique. The scientific contribution of this work is therefore three-fold: (i) a reproducible, GEE-based multi-classifier benchmark for a data-scarce rural tidal-fluvial reach; (ii) dual validation combining the Kappa confusion-matrix statistic with AUC-ROC analysis, which is rarely reported together in regional Indian LULC literature; and (iii) a 20-year (2000-2020) change-detection record quantifying the conversion of water body and vegetation into fallow and built-up land in the lower basin."

STUDY AREA CHARACTERISTICS

West Bengal is surrounded by a drainage system with many interlinking channels, rivulets, streams, and torrents, thus forming a very complex topography. In southern West Bengal, rivers such as the Damodar, Dwarakeswar, Silabati, Rupnarayan, Kangsabati and Haldi

are seen to be rapidly getting choked up. Based on the drainage networks and water availability for agriculture, water retention, and an excess resulting in floods, which is a major problem in these river basins (Mukhopadhyay & Dasgupta, 2010). The Rupnarayan River basin suffers from the highest depositional threat at its lower reaches. The Silabati and Dwarakeswar flows from opposite directions in Bandar (Ghatal), and their swollen mass is known as the Pangar, which joins the Hoogly River at Geonkhali after passing through 788 km with a watershed area of 10,797 sq. km. The Rupnarayan River is again subdivided into three reaches depending on the channel pattern of the river and associated riverine characteristics. The upper reach is a 28 km stretch beginning from Bandar to Jasar. The 10 km elongation of the middle reach is between Jasar and Kolaghat. The climate and meteorological characteristics of this basin are mainly humid subtropical in nature. However, the maximum rainfall occurs only during the rainy season, and its amount varies between 1804 mm/y to 1077 mm/y (IMD). As a result, this basin has gradually become a drought-prone area due to high summer temperatures, lack of water retention capacity in the lateral layer, and high evaporation and transpiration rates, respectively (Mittal et al., 2016). The Rupnarayan River basin reach, which is the main focus of this investigation, covers a distance of 40 km from Kolaghat (Dainan) to Geonkhali with the approximate latitudes of 22°01'20" N to 22°02'60" N and longitudes of 87°05'30" E to 88°00'30" E as shown in Figure 1. This lower reach is further divided into five sub-reaches: Geonkhali to Dhanipur, Dhanipur to Pyratungi, Pyratungi to Anantapur, Anantapur to Soyadighi, and Soyadighi to Dainan. This is done based on the needs of the study.

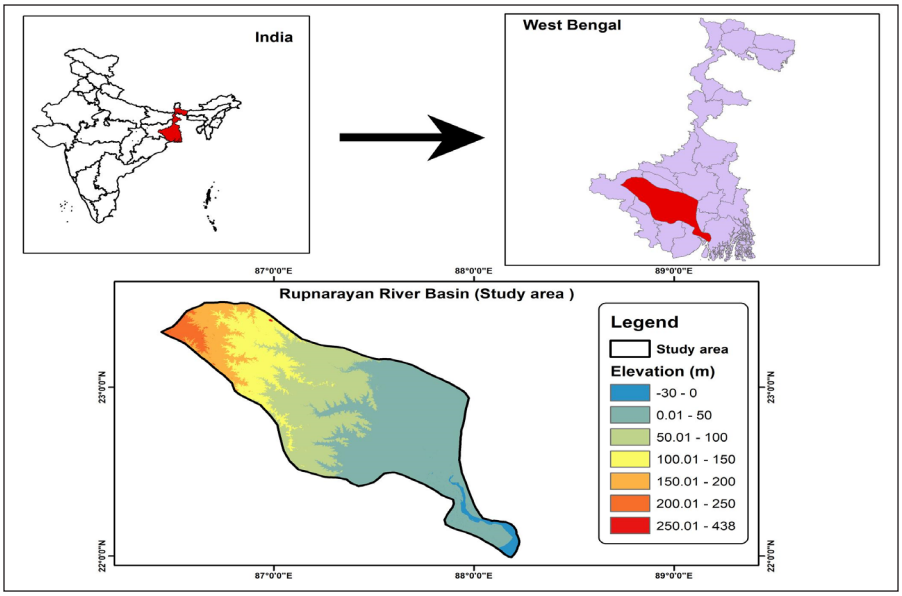


Figure 1. Location of the Rupnarayan river basin

METHODS AND MATERIALS

Data Used and Methodology

To monitor changes in land use and land cover dynamics, it is necessary to have data comparisons from at least two time periods. The LULC analysis of the present area has been performed based on a couple of steps as expressed in Figure 2. Spatial data has been acquired from a space-borne satellite sensor of the public domain and pre-processed the data through radiometric and geometric rectification to analyse the LULC of the corresponding study area (Phiri et al., 2018). LULC analyses have been implemented through machine learning supervised algorithms applying the Google Earth Engine system as a cloud-based open-source platform. The analyses of spatial data of LULC have involved i) finding out the heterogeneous areas through FCC (False Colour Composite) acquired from the wavelength of NIR, Red and Green. ii) Polygons in the virtual data were separated into polygons corresponding to heterogeneous patches in the FCC, and polygons evenly distributed throughout the region were separated from the FCC and complementary. The polygons have been covered at least with a spatial extent of 15% throughout the region, as uniformly distributed. iii) LULC analyses have been executed by the parametric and non-parametric supervised classifiers, i.e., maximum likelihood classifier (MLC), random forest (RF) and Support Vector Machine (SVM). This all-classifier method has been executed on the Google Earth Engine platform. After all the analysis, each algorithm has been validated through the Kappa coefficient statistics and the AUC-ROC curve model. LULC classifications were generated by supervised classification and the maximum likelihood method in the years 2000, 2010, and 2020 in the study area, as shown in Table 1. Scenes were selected primarily from the post-monsoon dry season, November to December, to ensure cloud-free conditions and comparable low-moisture, low-flow surface states across years; March and May scenes were used only where a cloud-free post-monsoon scene was unavailable for a given path/row. Because the basin is strongly monsoon-driven, residual seasonal differences in agricultural, fallow-land and water-body extent cannot be fully eliminated, and this is acknowledged as a limitation.

Table 1
List and sources of data collected

Data	Date Acquired	Path/Row	Resolution	Source
Landsat 5 TM	2000.11.09 2000.03.21 2000.11.23			
Landsat 5 TM	2010.12.23 2010.12.14 2010.02.20	138/045 139/044 140/044	30 m (Cloud Free)	USGS: https://earthexplorer.usgs.gov/
Landsat 8/9 (OLI) /(TIRS)	2020.12.18 2020.11.23 2020.05.22			
Base Map	Topographical Map	Study area	1:50000	Survey of India

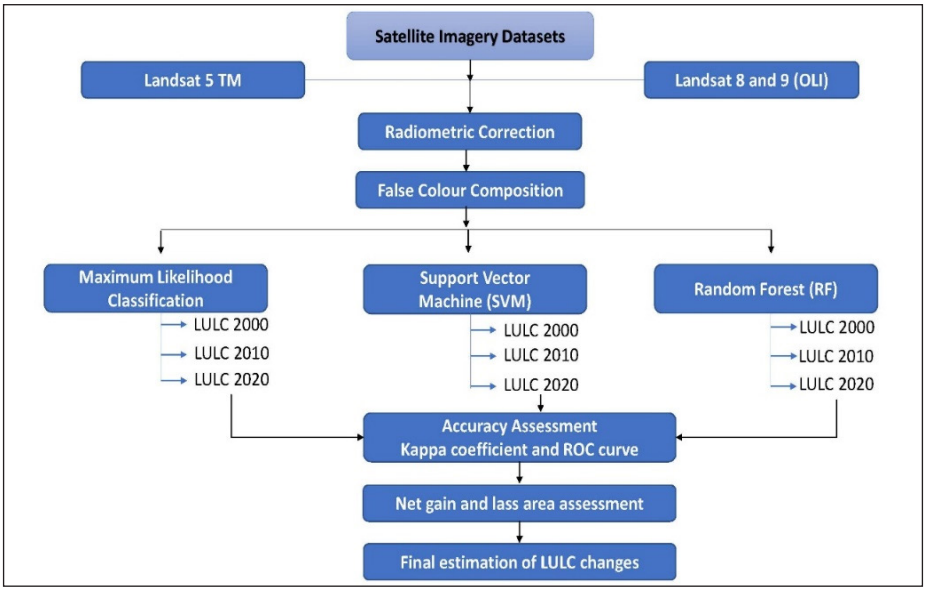


Figure 2. Flowchart of the adopted methodology

Maximum Likelihood Classification and Accuracy Assessment

Maximum likelihood classification (MLC) is a technique used to identify a known class of distributions by maximising a specific statistic. The origins of MLC can be traced back to electrical engineering. This method assumes that the training samples follow a normal distribution (Nilsson, 1956). The algorithm constructs probability density functions for each category. During the classification process, unclassified pixels are assigned to categories based on the relative likelihood (or probability) of each pixel fitting within the probability density function of each category (Hagner & Reese, 2007).

Each class spectral reflectance has been evaluated based on the qualitative method of variance and covariance as per Equation 1.

$$L = \frac{1}{(2\lambda)^{\frac{n}{2}}(\sum_k \frac{1}{2})} \exp \frac{1}{2} (X - \mu_k) \sum_k^{-1} (X - \mu_k) t \quad [1]$$

Here, L represent the likelihood of X that belongs to the K class, n denotes the spectral resolution, μ_k reflects the K class mean vector, $\sum K$ represent the matrix of the K class of variance and covariance.

Random Forest (RF)

Google Earth Engine (GEE) is the most significant tool that provides various machine learning algorithm applications through a cloud computing online platform, including random forest (RF), support vector machine (SVM), decision trees (DT), maximum likelihood classifier (MLC), etc. (Shafi et al., 2023) to analyse the land use/ land cover change. In the current analysis, the random forest classifier method of 2000, 2010 and 2020 was also adopted for a specific region.

Random Forest is a combination of non-parametric methods in machine learning that is excellent for solving the pre-direction problem as well as classification. An accumulation of decision trees, the algorithm has been generated for each of the different subsets of input data. Each subset has been generated by random choices of instances from the provided dataset with replacement and using the bootstrap sampling method; the decision trees are generated (Mateen et al., 2023). To enhance the generalising model, a random subset of data has been considered for splitting at each node of the decision trees. To decrease the correlation between trees and obtain the maximum accuracy in the outcome, a random subset of the available input dataset to make the splits during the training stage. The decision trees can be used to classify the new dataset by applying the class probability of each tree after the trees are trained well and then averaging the results over all the trees. To reduce the noise, the diverse tree facility and random features combined selection has been mitigating the overfitting of subsets. For accurate outcomes, each feature has been extracted to assess its contribution. The cloud GEE open platform is one of the most accurate and reliable algorithms applying through large scale remote sensing data (Ge et al., 2020).

Training samples have been collected through GPS-based ground truth field investigation during 2020 for LULC analysis. Maximum field data were collected according to the accessibility and portion of the selected LULC region, randomly. Previous years' field investigation was done by high-resolution images of Google Earth Proando visual interpretation system. Then uploaded the training samples in GEE as assets from ArcGIS Pro, where the field data was stored. Near bout 3000 reference points have been generated to assess the LULC in the GEE platform. After that, 80% of the dataset was randomly set for the training sample, and 20% considered as testing or validation to verify the accuracy of the RF classifier, as Feng et al. (2022) proposed. Reference data for 2020 were obtained from a GPS-based ground-truth survey, whereas reference polygons for 2000 and 2010 were generated by visual interpretation of the False-Colour-Composite of the corresponding Landsat-5 TM scenes, supported by the nearest-date archival imagery available in Google Earth Pro and by Survey of India 1: 50,000 topographic sheets. High-resolution Google Earth coverage contemporaneous with the 2000 acquisition was sparse; for such cases, interpretation relied on the Landsat FCC, and the temporally closest archival image, and ambiguous polygons were cross-checked against the topographic sheets.

Approximately 3,000 reference points were distributed across the basin and randomly partitioned so that the 20% testing subset was spatially held out and never used during training. A single random 80/20 hold-out partition was applied to all three classifiers for each year; the study did not use k-fold cross-validation and repeated random subsampling.

Support Vector Machine

In machine learning, SVM is a non-parametric binary supervised classification process. From structural risk minimisation (SRM), the SVM method was formed by dividing and maximising the data closest to the hyperplanes of spectral angle mapper (SAM). The method involves the division of the feature point into several land cover classes. The vector width of the margin line has been maximised while processing the SVM. To acquire a more precise result from SVM, a kernel function was required while performing the SVM algorithm. To enhance the distance in the vector boundary of support vectors is the actual tendency of the SVM algorithm. SVM is applied for LULC classification as well as regression assessment (Vapnik, 1999). Some binary formula, Equation 2, has been adopted for the classification:

$$(x) = \mathrm{Sin}(\sum_{i=1}^{n} \propto i\, \gamma i K(x, xi) + b) \qquad [2]$$

Here, f(x) denotes the decision function which predicts the input classes, x, αi is the multiplier of language that was acquired from the training sample, yi is the training sample of xi class label, and K (x, xi) is the kernel function. The bias term is represented by b.

Accuracy Assessment

Accuracy assessment is crucial for supervised classification within a GIS environment. The Kappa coefficient and overall accuracy are key techniques for analysing accuracy and indicating the quality of any classified image algorithm. The Kappa coefficient is derived from a confusion matrix and is applied to current land cover mapping (Chowdhury & Hafsa, 2022; McHugh, 2012). In this study, various Kappa indices were calculated, including overall accuracy, user accuracy, producer accuracy, and the Kappa coefficient itself. Overall accuracy reflects the true values of the entire Land Use and Land Cover (LULC) map, while user accuracy estimates the false values for a specific class. Producer accuracy assesses the true values of a particular class. The Kappa coefficient and its associated metrics are determined using a specific applied Equation 3 (Das et al., 2021).

$$K = \frac{[n \sum_{i=1}^{r} x_{ij} - \sum_{i=1}^{r} x_i x_j]}{[n^2 - \sum_{i=1}^{r} (x_i x_j)]} \qquad [3]$$

Where K indicates Kappa Coefficient; n indicates the total number of samples; x_{ij} is the total corrected sample; x_i is the producer total, and x_j is the user total

Overall accuracy is estimated using Equation 4:

$$\text{OA} = \frac{\sum CP_{ii}}{N} \quad [4]$$

OA indicates the Overall accuracy. ΣCP_{ii} indicates the total observed samples classified correctly for each feature. N indicates the total number of samples observed.

Model Validation by AUC-ROC

The present analysis of LULC by various methods, i.e., SSVM, RF, and MLC, has been validated by a true statistical method, AUC (area under the curve), and ROC (receiver operating characteristics), which is one of the most suitable and mathematical approaches for model validation, also applied by several researchers in the past. In the model, the actual position rate has been assessed by the false position rate of the hypothesis. The following formula, Equation 5, has been applied here to calculate the AUC-ROC:

$$\text{AUC} = \sum_{i=1}^{n=100} \frac{(x_1 + x_2)}{2(y_2 + y_1)} \quad [5]$$

RESULT AND DISCUSSION

Comparative Assessment of LULC Change Dynamics by MLC, RF and SVM

Temporal changes in land use/land cover analysis between the years 2000 and 2020 have been performed through supervised classification algorithms of maximum likelihood classifier (MLC), random forest (RF) and support vector machine (SVM) systems. The analysis of each algorithm has been calculated for the years of 2000, 2010 and 2020 to depict the land use and land cover condition of the Rupnarayan river basin area. First, the MLC has been generated over the area for 2000, 2010 and 2020, where five distinct features have been selected, i.e., built up. Vegetation, fallow land, agricultural land and water body. From the LULC map through MLC, it has been revealed that 164.2 sq.km area was found under built-up area in the year 2000, where it has been expanded in 2010 to 3207.06 sq.km and 1976.46 sq.km area was covered with built-up in 2020.

From the analysis of MLC agricultural land area, in 2000 it covered only 308.8 sq.km, but in 2010 it increased to 2167.89 sq.km and decreased to 693.2 sq.km in 2020. Fallow land in the study area reveals the major changes where a 5001.1 sq.km area fell under fallow land in 2000 and 1893.54 sq.km in 2010; however, increases in 2020 to 7867.55 sq.km, that denote a very positive expansion of fallow land. Vegetation area by MLC in 2000, showing the 1365.2 sq.km and expanded to 3237.89 sq.km in 2010, but extremely decreased in the year 2020 to 45.46 sq.km. In the LULC assessment, major modification has been observed in the water body where a 3957.7 sq.km area in the year 2000 reduced to 290.66 sq.km and 214.37 sq.km in the years 2010 and 2020, respectively, as shown in Figure 12. The map of land use/land cover has been generated to assess the change detection of each year through three different supervised algorithms and spatial data derived from the prepared map to identify the loss and gain areas within the period, as shown in Figures 3 to 11.

From the analysis of the random forest, it has been observed that some spatial changes were made in each feature area with the temporal condition. In case of built-up, 1931.00 sq.km area in 2000 has been enhanced in 2010 to 5839.88 sq.km, which again reduced in 2020 to 3569.12 sq.km. A very positive change has been found in agricultural land, where a 338.11 sq.km area in 2000 expanded to 1366.81 sq.km in 2020. In fallow land, a 3703.42 sq.km area in 2000 shifted to a 330.59 sq.km area in 2010, which was modified to 53.1539 sq.km in 2020. A drastic change has been observed in the vegetation area, where a 1286.27 sq.km area was found in 2000, which increased to 4164.59 sq.km in 2010, but was extremely reduced in 2020 to 71.197 sq.km by RF. Water bodies represent a negative shift over 20 years, where 3538.21 sq.km in 2000 shifted to 305.46 sq.km and 474.51 sq.km in 2010 and 2020, respectively, as shown in Figure 12.

Finally, the study generates a map of support vector machine (SVM) for each year to show the land use/land cover condition. In the year 2000, a map of the SVM built-up area observed in 469.24 sq.km, which expanded to 6227.37 sq.km in 2010 and again reduced in 2020 to 2121.09 sq.km. Agricultural area has been shifted from 471.377 sq.km in 2000 to 2392.68 sq.km and 1529.66 sq.km area in 2010 and 2020, respectively. After the analysis of SVM, the fallow land has been denoted as the major modification where 5609.54 sq.km area was modified to 197.62 sq.km and 6170.21 sq.km in 2010 and 2020, respectively. Similarly, the vegetation area in 2000 was revealed to be 2160.74 sq.km, which shifted to 1614.40 sq.km and 658.43 sq.km in 2010 and 2020, respectively. The water body has been gradually decreasing through SVM analysis, where 2086.13 sq.km area in 2000 regressed to 364.97 sq.km and 317.64 sq.km in 2010 and 2020, respectively, as shown in Figure 12.

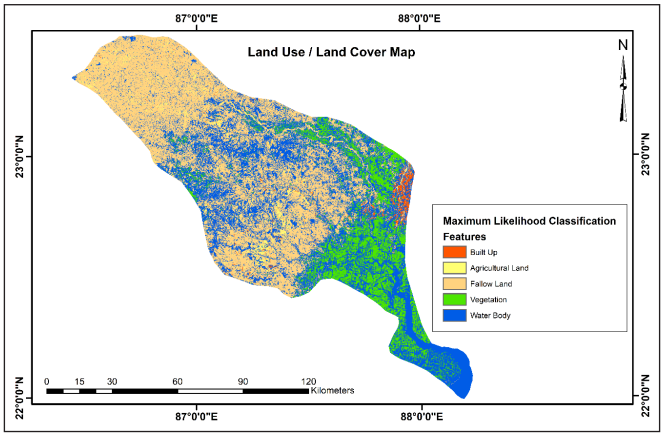


Figure 3. LULC map by maximum Likelihood Classifier (MLC), year 2000

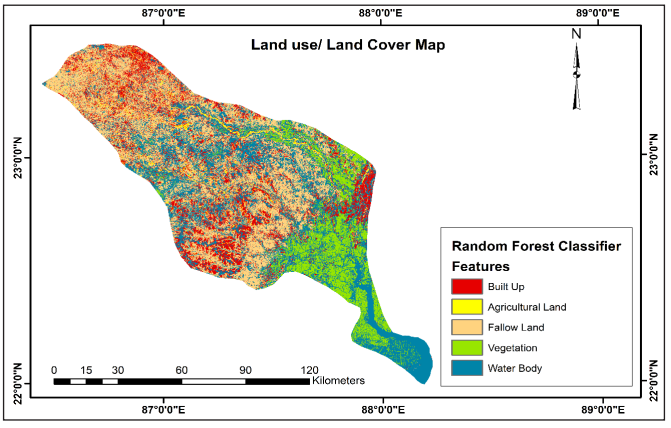


Figure 4. LULC map by Random Forest (RF), year 2000

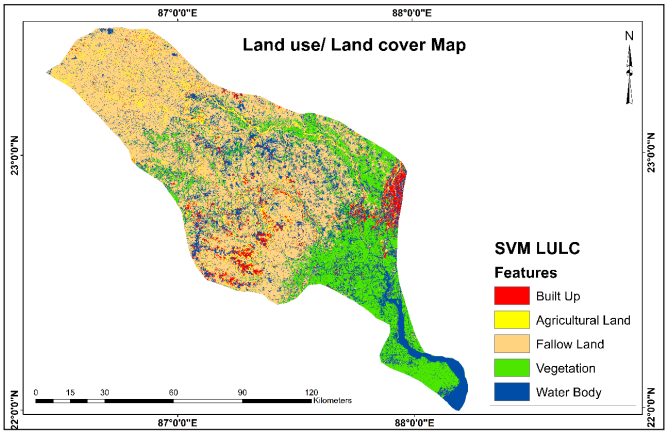


Figure 5. LULC map by Support Vector Machine (SVM), year 2000

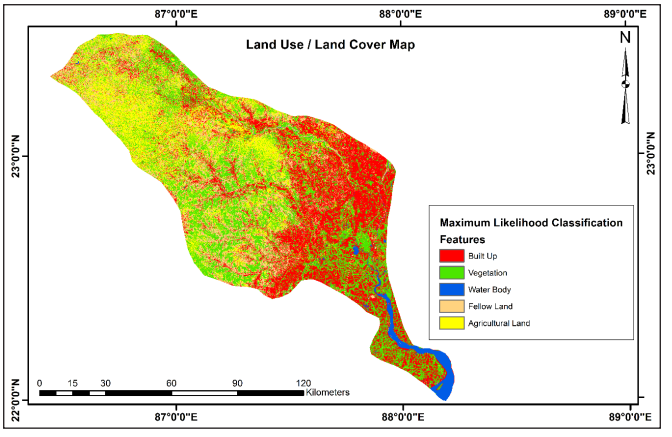


Figure 6. LULC map by Maximum Likelihood Classifier (MLC), year 2010

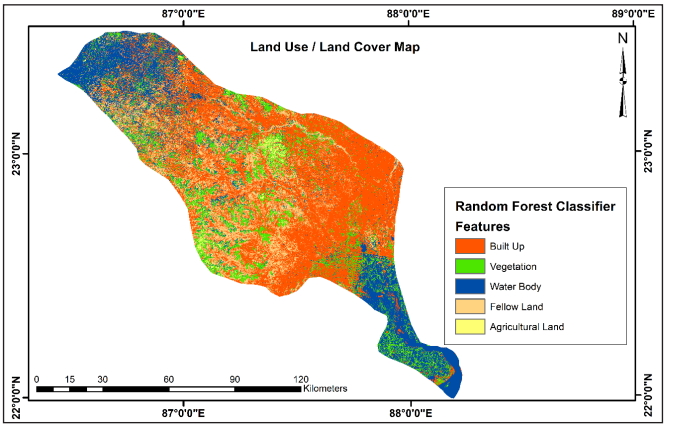


Figure 7. LULC map by Random Forest (RF), year 2010

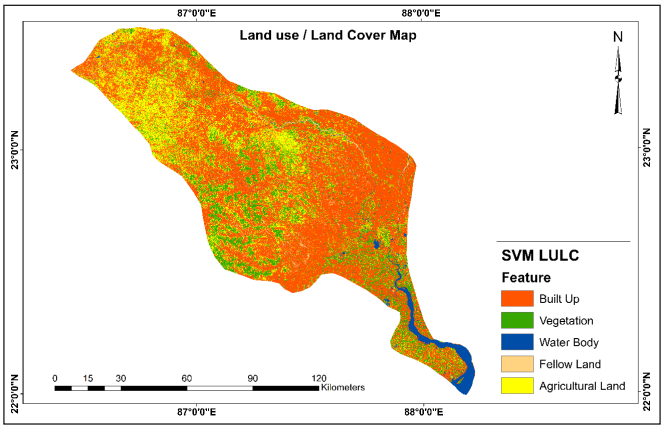


Figure 8. LULC map by Support Vector Machine (SVM), year 2010

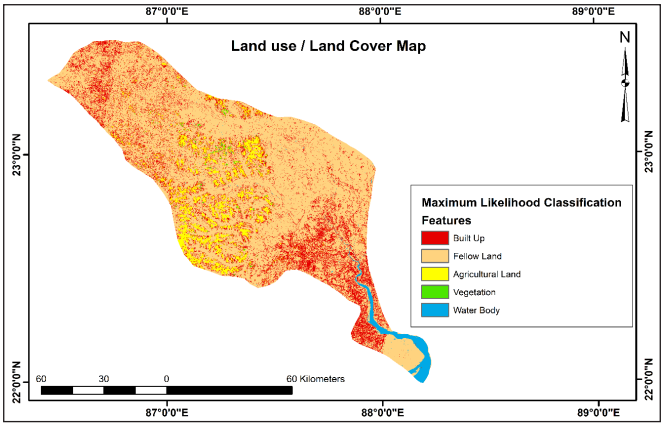


Figure 9. LULC map by Maximum Likelihood Classifier (MLC), year 2020

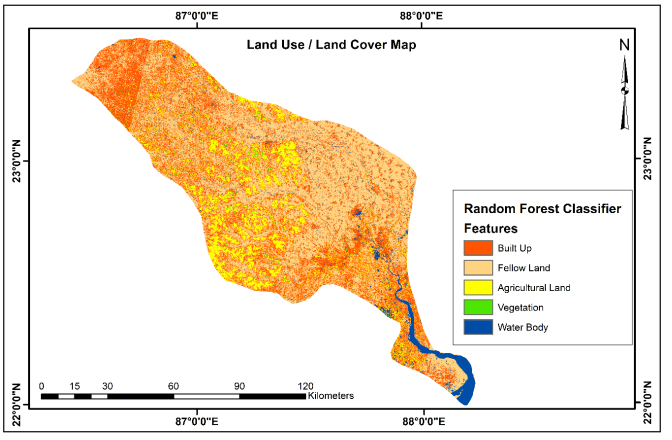


Figure 10. LULC map by Random Forest (RF), year 2020

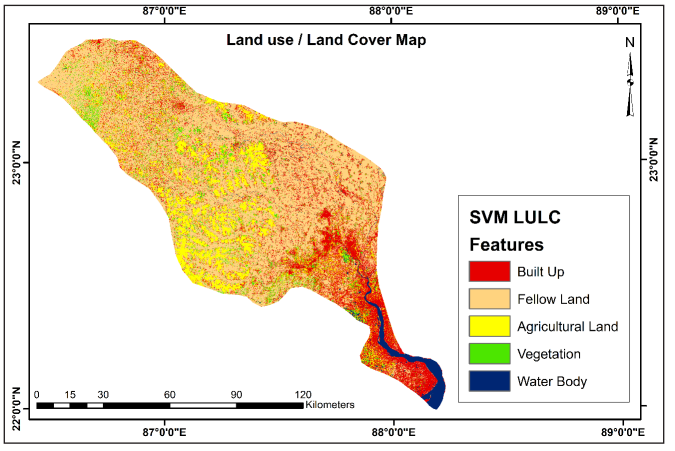
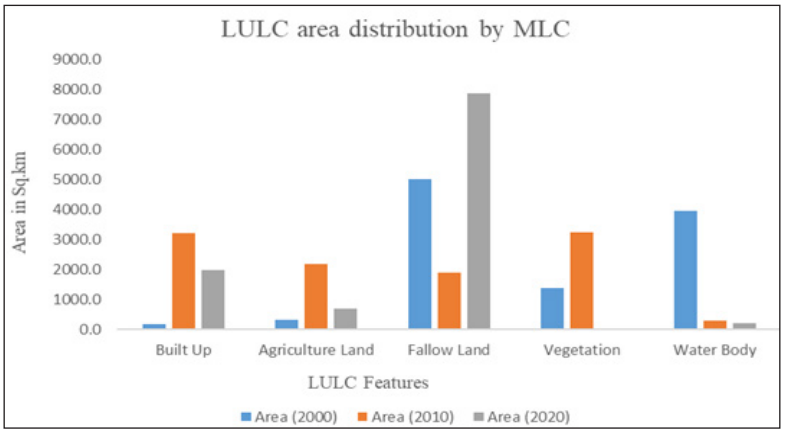
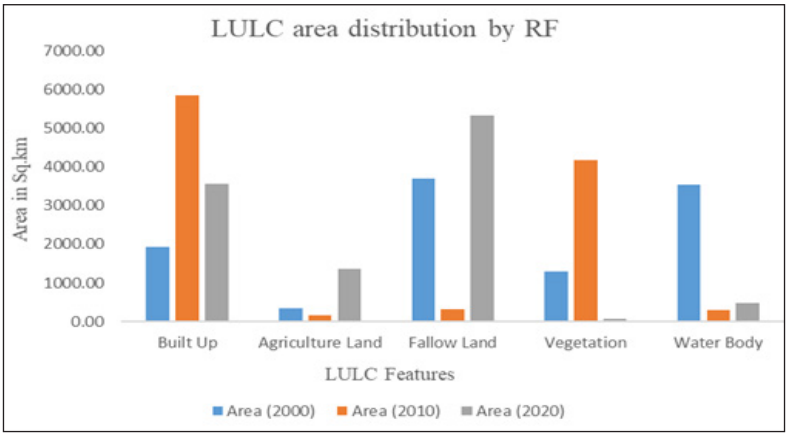


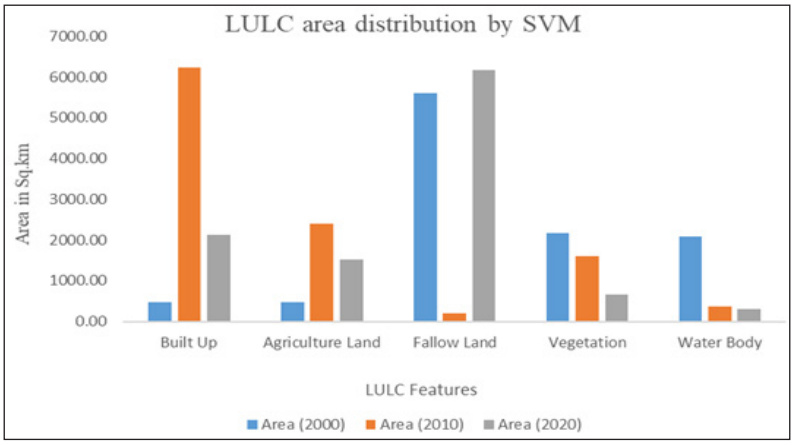
Figure 11. LULC map by Support Vector Machine (SVM), year 2020



(a)



(b)



(c)

Figure 12. LULC area distribution by (a) MLC; (b) RF; (c) SVM

Change Detection of LULC from 2000 to 2010 and 2010 to 2020 by MLC RF SVM

After the analysis, the loss and gain areas have been calculated for each algorithm from 2000 to 2010 and from 2010 to 2020. Table 2 and Figure 13 (a) show the modification of each feature through the MLC method, where the built-up area positively changed from 2000 to 2010 to 3042.85 sq.km, but built up reduce in the next 10 years (2010 to 2020) to -1230.61 sq.km. On the other hand, in the RF method, as shown in Table 3 and Figure 13 (b), it has been depicted that the built-up area increased between the first 10 years, 3908.88 sq.km, but reduced in the next 10 years to -2270.76 sq.km. But maximum built-up reduction has been observed by the SVM method -4106.29 sq.km as shown in Table 4 and Figure 13 (c). In agricultural land -1474.68 sq.km area disappeared from the MLC method from 2010 to 2020, but 1210.31 sq.km area was enhanced in the RF model, where -863.01 sq.km area was again lost in the SVM method from 2010 to 2020. In the case of fallow land, the 2000 to 2010 period depicts the negative change where -3107.55 sq.km area lost by MLC, -3372.83 sq km and -5411.93 sq km by RF and SVM accordingly within that time. But the next 10 years showed the positive change where 5974.01, 4984.79 and 5972.59 sq.km by MLC, RF and SVM, respectively. In the case of vegetation, MLC and RF models remarked the positive change in the 2000 to 2010 period, but the SVM model depicts the negative change in both time spans (-546.34 sq.km in 2000 – 2010 and -955.98 sq.km in 2010 to 2020). A proper negative shifting has been observed in the water body by the MLC and SVM model, where -3232.75 sq.km and -1721.16 sq.km were lost in 2000 to 2010 and -76.27 sq.km and -47.32 sq.km in 2010 to 2020, as shown in Tables 2 and 4 and in Figure 3.

Table 2
LULC change detection by MLC (From 2000 to 2010 and 2010 to 2020)

Feature	Area (Year 2000)	Area (Year 2010)	Area (Year 2020)	Change Detection (Year 2000 to 2010)	Change Detection (Year 2010 to 2020)
Built Up	164.2	3207.06	1976.45	3042.85	-1230.61
Agriculture Land	308.8	2167.88	693.19	1859.11	-1474.68
Fallow Land	5001.1	1893.54	7867.55	-3107.55	5974.01
Vegetation	1365.2	3237.89	45.45	1872.6426	-3192.44
Water Body	3957.7	290.66	214.38	-3667.06	-76.27
			10797.04		

Table 3
LULC change detection by RF (From 2000 to 2010 and 2010 to 2020)

Feature	Area (Year 2000)	Area (Year 2010)	Area (Year 2020)	Change Detection (Year 2000 to 2010)	Change Detection (Year 2010 to 2020)
Built Up	1931.01	5839.88	3569.12	3908.88	-2270.76
Agriculture Land	338.12	156.50	1366.81	-181.62	1210.31

Table 3 (*continued*)

Feature	Area (Year 2000)	Area (Year 2010)	Area (Year 2020)	Change Detection (Year 2000 to 2010)	Change Detection (Year 2010 to 2020)
Fallow Land	3703.43	330.59	5315.39	-3372.83	4984.79
Vegetation	1286.27	4164.59	71.18	2878.32	-4093.39
Water Body	3538.22	305.47	474.51	-3232.75	169.05
			10797.04		

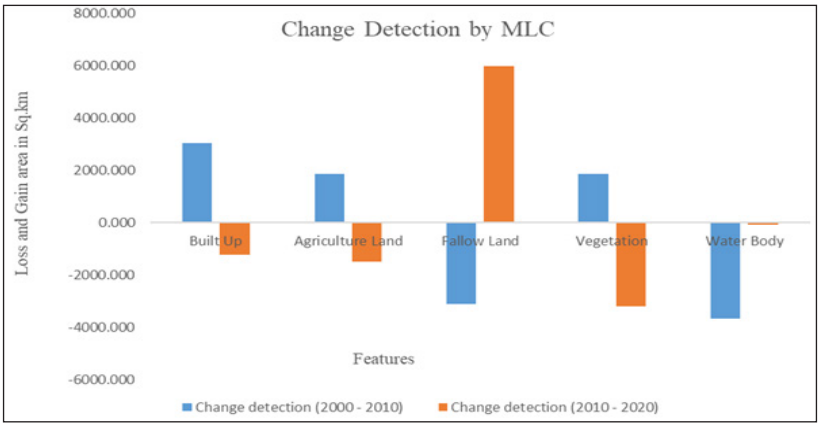
Table 4
LULC change detection by SVM (From 2000 to 2010 and 2010 to 2020)

Feature	Area (Year 2000)	Area (Year 2010)	Area (Year 2020)	Change Detection (Year 2000 to 2010)	Change Detection (Year 2010 to 2020)
Built Up	469.24	6227.36	2121.07	5758.12	-4106.29
Agriculture Land	471.38	2392.68	1529.66	1921.34	-863.01
Fallow Land	5609.54	197.62	6170.21	-5411.93	5972.59
Vegetation	2160.74	1614.40	658.42	-546.34	-955.98
Water Body	2086.13	364.97	317.64	-1721.16	-47.320
			10797.04		

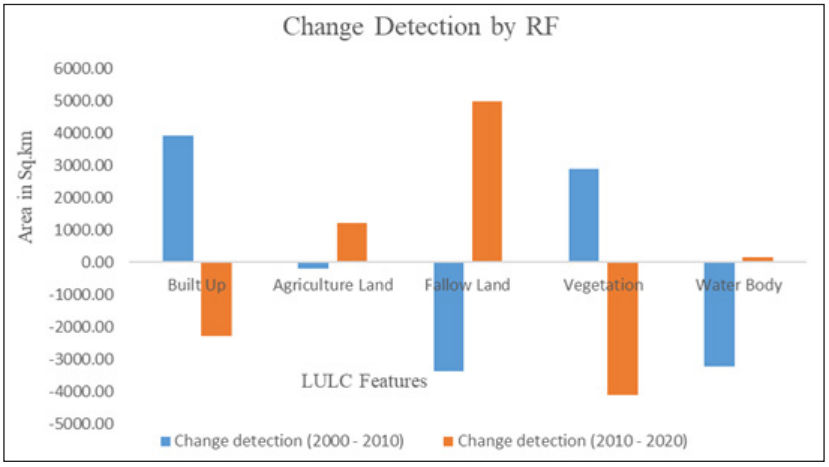
Comparative Accuracy Assessment of MLC RF SVM

The per-class areas estimated by the three classifiers differ in magnitude, which is expected because each algorithm applies a different decision rule to the same imagery; these differences should be read as a comparative evaluation of classifier behaviour rather than as data contradictions. Importantly, all three classifiers agree on the direction of change - a decline of water bodies and vegetation and an expansion of fallow and built-up land over 2000 to 2020, and they differ mainly in how sharply they resolve the spectrally mixed classes. The following paragraphs summarise each classifier in turn, with full figures reported in Tables 2 to 4 and Figures 12 to 13.

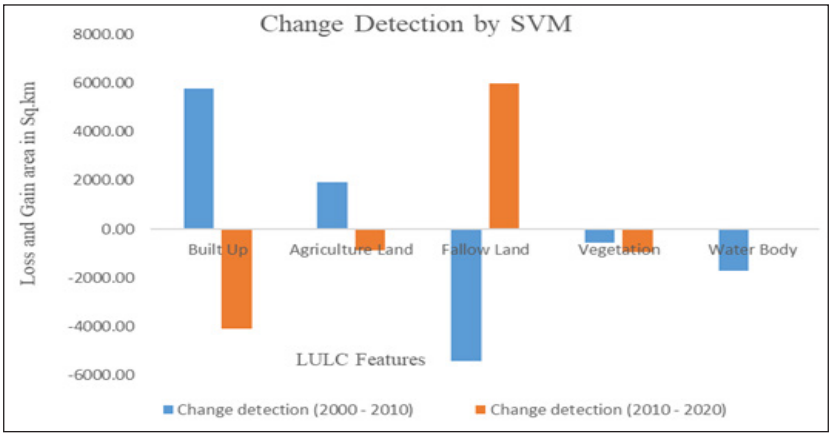
The LULC classification has been evaluated based on the accuracy assessment as shown in Tables 5, 6 and 7. The comparative supervised algorithm reveals that the SVM machine learning algorithm has been more accurate and precise for the current study. The kappa value has been obtained in 0.97 (2000), 0.95 (2010) and 0.95 (2020) in SVM. On the other hand, the second relevant method has been depicted in the MLC classifier, where an accuracy of 0.85 was obtained in the year 2000, 0.88 in 2010 and 0.89 in 2020. The RF model revealed the least accuracy in the current research work, where a 0.85 value was acquired in 2000, 0.85 in 2010 and 0.87 in 2020. From the comparative evaluation, it has been seen that built-up area, agriculture and water body misplaced in the omission error while performing the model.



(a)



(b)



(c)

Figure 13. Change statistics of LULC features by (a) MLC; (b) RF; (c) SVM

Similarly, in the MLC model, built-up and vegetation areas have been misplaced while operating the algorithm. But the SVM algorithm portrayed the maximum accuracy in omission and commission error in the current LULC assessment. Tables 5 to 7 provide the category-wise accuracy assessment of each algorithm through the Kappa coefficient.

Table 5
Accuracy assessment of MLC by the kappa coefficient

LULC Classes	User's Accuracy			Producer's Accuracy			Overall Accuracy			Kappa Coefficient (K)		
	2000	2010	2020	2000	2010	2020	2000	2010	2020	2000	2010	2020
Vegetation	85	90	95	87	88	87	85%	88%	90%	0.85	0.88	0.89
Water Body	87	90	92	92	90	85						
Build up	85	87	90	82	85	85						
Agricultural Land	92	85	88	85	88	87						
Fallow Land	85	88	90	90	85	85						

Table 6
Accuracy assessment of RF by the kappa coefficient

LULC Classes	User's Accuracy			Producer's Accuracy			Overall Accuracy			Kappa Coefficient (K)		
	2000	2010	2020	2000	2010	2020	2000	2010	2020	2000	2010	2020
Vegetation	88	90	87	85	88	92	85%	85%	87%	0.85	0.85	0.87
Water Body	89	85	85	84	89	89						
Build up	88	85	82	87	87	85						
Agricultural Land	89	92	87	90	85	92						
Fallow Land	90	85	92	85	83	88						

Table 7
Accuracy assessment of SVM by the kappa coefficient

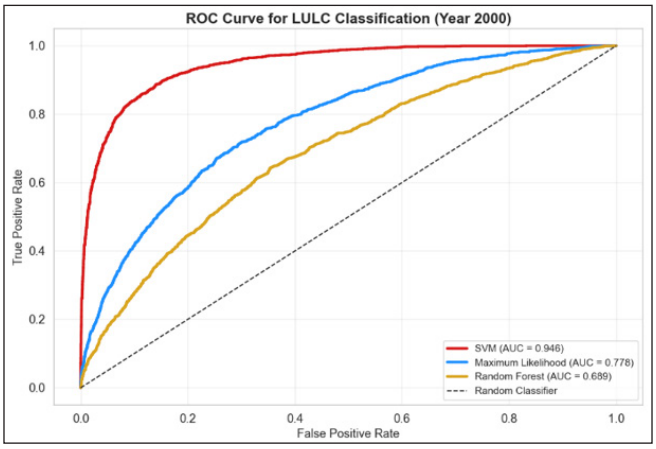
LULC Classes	User's Accuracy			Producer's Accuracy			Overall Accuracy			Kappa Coefficient (K)		
	2000	2010	2020	2000	2010	2020	2000	2010	2020	2000	2010	2020
Vegetation	92	100	100	95	92	100	97%	95%	95%	0.97	0.95	0.95
Water Body	90	92	95	100	100	100						
Build up	92	100	100	90	100	100						
Agricultural Land	90	90	90	95	100	100						
Fallow Land	92	95	100	95	95	100						

Model Validation by AUC-ROC

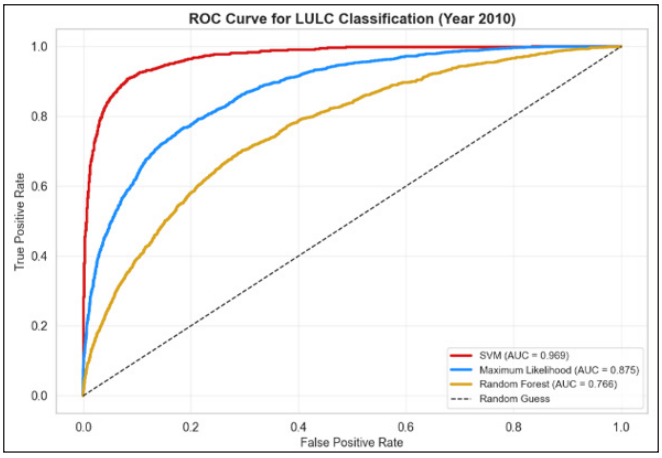
It is significant to ensure the model validation with the real ground conditions of LULC. The spatial position of actual land cover features is required to match the generated land false colour map by SVM, RF and MLC of each year (Ogato et al., 2020; Tadesse et al., 2022). The applied model of each algorithm of each year has been cross checked the accuracy through the AUC-ROC model. All the land use spatial conditions have been cross-verified by the ground position through Global Positioning System (GPS) and high-resolution Google Earth imagery to validate the model. From the ArcSDM extension tool of ArcGIS, the curve has been drawn. In the analysis of 2000, the AUC value obtained 0.778 in MLC, RF AUC is 0.689, and SVM AUC is 0.946. In 2010, the MLC AUC is 0.875, the RF AUC value is 0.766, and the SVM value is obtained as 0.969. Finally, the 2020 map of each algorithm has been applied for the AUC curve, where the MLC curve value is 0.898, the RF value is 0.794, and the SVM obtained an AUC of 0.986. From the entire validation model, it has been revealed that SVM acquired the maximum AUC value for each year, while the other two (MLC and RF) algorithms obtained the lowest range of AUC value, as shown in Figure 14. Moreover, a 0.85 value in the model has been considered to validate the algorithm.

DISCUSSION

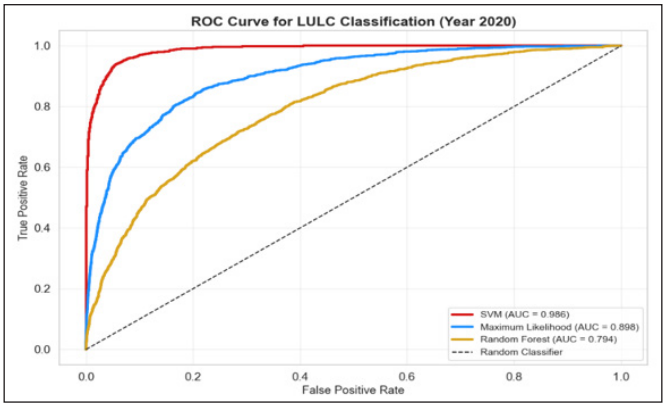
Three different algorithms have been applied for assessing the LULC in the Rupnarayan catchment area, i.e., MLC, RF and SVM. Three model was depicting the spatial extension of the built-up area during the years 2000 to 2010. The spatial extension of fallow land increased, and 5974.01 sq.km area gained during 2010 to 2020 in MLC, 4984.79 sq.km increased by RF, and 5972.59 sq.km area enhanced through the SVM method due to the loss of agricultural land, which reduced -863.01 sq.km during the last ten years. After generating the LULC map, it has been observed that extreme modification depicted in the water body has been reduced within two specific time spans due to the fallow land expansion over the catchment area. In the SVM method, the water body area reduced by -1721.16 sq.km during the years 2000 to 2010 and again -47.32 sq.km was lost between 2010 and 2020. The evaluation application of random forest, maximum likelihood classifier and support vector machine classifier for assessing the land cover change in Rupnarayan River catchment area, it has been revealed that the most accurate outcomes were acquired through the SVM classifier method with a high accuracy of Kappa coefficient value and AUC-ROC model validation of each model. SVM consistently outperformed MLC and RF (Figure 14) because its margin-based decision boundary separates spectrally overlapping classes (fallow vs. agricultural land, built-up vs. bare soil) more effectively than the Gaussian-distribution assumption of MLC or the variance of individual trees in RF, and because it is comparatively robust when training samples are limited - a key advantage given the scarce historical reference data for 2000 and 2010.



(a)



(b)



(c)

Figure 14. Classifier algorithm validation by specificity and sensitivity, AUC-ROC curve of (a) Year 2000; (b) Year 2010; (c) Year 2020

Practically, the consistent loss of water bodies and vegetation against expanding fallow and built-up land signals reduced water-retention capacity and rising land-surface stress in the lower basin, with direct implications for flood management and sustainable land planning.

CONCLUSION

The current study indicates that the Rupnarayan River Basin has undergone significant dynamic changes in LULC from 2000 to 2020. While land use is increasing, this study highlights a decline in certain areas. Across all three years, the SVM classifier returned the highest accuracy (Kappa 0.97 / 0.95 / 0.95 and AUC 0.946 / 0.969 / 0.986 for 2000 / 2010 / 2020), consistently outperforming MLC and RF, which confirms that a margin-based classifier is best suited to this heterogeneous, data-scarce catchment. The dominant change detected was a marked decline of water bodies and vegetation and a corresponding expansion of fallow and built-up land between 2000 and 2020. The current research work three specific supervised algorithms, i.e., MLC, RF and SVM. After generating the algorithm in high-resolution multispectral satellite data, it has been remarked that the SVM algorithm provides the maximum accuracy on ground truth data and Google Earth imagery. A ROC curve was also generated based on the true and false data to validate the model of each year, which also reflects the highest SVM AUC value in each year. This finding is consistent with earlier comparative studies in which SVM outperformed parametric classifiers (Darabi et al., 2019; Fatah et al., 2025; Javidan et al., 2019; Jodhani et al., 2025; Ruidas et al., 2023; Tripathi et al., 2023), and it extends that evidence to a data-scarce tidal-fluvial Indian catchment validated jointly by Kappa and AUC-ROC.

From the analysis, it has been observed that the most notable decline was observed in the water body. Additionally, vegetation cover decreased in both time periods between 2000 and 2020 through SVM. Conversely, the built-up area increased significantly, primarily at the expense of Fallow land. It is essential to develop innovative land systems tailored to the local context, framed within the global socio-ecological system. These systems should explicitly consider land governance as a key driver of land change and agricultural production.

This study has several limitations. Scenes were drawn from slightly different months, so residual seasonal variation in fallow, agricultural and water classes cannot be fully excluded. Accuracy was assessed with a single random 80/20 hold-out rather than k-fold cross-validation or repeated sub-sampling, so the reported accuracies should be read as indicative. Reference data for 2000 and 2010 relied partly on visual interpretation of archival imagery, which introduces interpretation uncertainty, and the 30 m Landsat resolution produces mixed pixels along class boundaries. Future work should (i) extend the comparison to CART, gradient boosting and deep-learning (CNN) classifiers; (ii) fuse finer-resolution multi-sensor data (e.g. Sentinel-2) to reduce mixed pixels;

(iii) adopt k-fold cross-validation with repeated random sampling to quantify accuracy stability; and (iv) link the observed LULC trajectories to their socio-hydrological drivers.

The significance of this study is that it delivers a validated, reproducible GEE-based workflow and a 20-year change record for the lower Rupnarayan basin, providing policymakers, land planners and watershed managers with a reliable evidence base for sustainable land- and water-resource management in similar rural river catchments

AUTHOR CONTRIBUTIONS

KB, AN, TC and BK contributed to conceptualising the study and writing the introduction, literature review, data analysis, modelling, and interpretation. BK formally analyses the data and provides the figure representation.TC, BK and JR handled formatting, language editing, review the manuscript, and AN and TC supervised this research. Manuscript confirming that all authors have approved the final version of the manuscript.

DECLARATIONS

Ethical Approval and Consent to Participate

No human subjects are involved in this research.

Consent for Publication

Not applicable.

Competing Interests

The authors declare that they do not have any competing interests.

Data Availability

The datasets used in this study were derived from publicly available satellite imagery. All outcomes and generated datasets were cross-verified through ground-truth field surveys to ensure accuracy and reliability. Spatial data processing, analysis, and mapping were conducted using the ArcGIS platform and Google Earth Engine (GEE). The processed data and related materials are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to all individuals and organisations that contributed to the successful completion of this research. The authors also thank the anonymous reviewers and the editor for their constructive comments and suggestions, which helped improve the quality of the manuscript. Additionally, we gratefully acknowledge the

Here Technologies Lab, University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand, India, for providing the necessary research facilities.

REFERENCES

Ahmed, B., Kamruzzaman, M., Zhu, X., Rahman, M. S., & Choi, K. (2013). Simulating land cover changes and their impacts on land surface temperature in Dhaka, Bangladesh. *Remote Sensing, 5*(11), 5969-5998. https://doi.org/10.3390/rs5115969

Batar, A. K., Watanabe, T., & Kumar, A. (2017). Assessment of land-use/land-cover change and forest fragmentation in the Garhwal Himalayan region of India. *Environments, 4*(2), Article 34. https://doi.org/10.3390/environments4020034

Bouslihim, Y., Kharrou, M. H., Miftah, A., Attou, T., Bouchaou, L., & Chehbouni, A. (2022). Comparing pan-sharpened Landsat-9 and Sentinel-2 for land-use classification using machine learning classifiers. *Journal of Geovisualization and Spatial Analysis, 6*, Article 35. https://doi.org/10.1007/s41651-022-00130-0

Chakilu, G., & Moges, M. (2017). Assessing the land use/cover dynamics and their impact on the low flow of the Gumara Watershed, Upper Blue Nile Basin, Ethiopia. *Hydrology: Current Research, 7*(2), Article 268. https://doi.org/10.4172/2157-7587.1000268

Cheruto, M. C., Kauti, M. K., Kisangau, P. D., & Kariuki, P. (2016). Assessment of land use and land cover change using GIS and remote sensing techniques: A case study of Makueni County, Kenya. *Journal of Remote Sensing & GIS, 5*, Article 175. https://doi.org/10.4172/2469-4134.1000175

Chowdhury, M. S. (2023). GIS-based method for mapping actual LULC by combining seasonal LULCs. *MethodsX, 11*, Article 102472. https://doi.org/10.1016/j.mex.2023.102472

Chowdhury, M. S. (2026). Incorporation of seasonal land cover to develop a seasonally integrated land cover map for change detection in Dhaka Mega City. *Scientific Reports, 16*(1), Article 468. https://doi.org/10.1038/s41598-025-17478-0

Chowdhury, M. S., & Hafsa, B. (2022). Multi-decadal land cover change analysis over Sundarbans Mangrove Forest of Bangladesh: A GIS and remote sensing-based approach. *Global Ecology and Conservation, 37*, Article e02151. https://doi.org/10.1016/j.gecco.2022.e02151

Darabi, H., Choubin, B., Rahmati, O., Haghighi, A. T., Pradhan, B., & Kløve, B. (2019). Urban flood risk mapping using the GARP and QUEST models: A comparative study of machine learning techniques. *Journal of Hydrology, 569*, 142-154. https://doi.org/10.1016/j.jhydrol.2018.12.002

Das, N., Mondal, P., Sutradhar, S., & Ghosh, R. (2021). Assessment of variation of land use/land cover and its impact on land surface temperature of Asansol subdivision. *The Egyptian Journal of Remote Sensing and Space Sciences, 24*(1), 131-149. https://doi.org/10.1016/j.ejrs.2020.05.001

Deilmai, B. R., Ahmad, B. B., & Zabihi, H. (2014). Comparison of two classification methods (MLC and SVM) to extract land use and land cover in Johor, Malaysia. *IOP Conference Series: Earth and Environmental Science, 20*(1), Article 012052. https://doi.org/10.1088/1755-1315/20/1/012052

Fan, S., Guan, F., Xu, X., Forrester, D. I., Ma, W., & Tang, X. (2016). Ecosystem carbon stock loss after land use change in subtropical forests in China. *Forests, 7*(7), Article 142. https://doi.org/10.3390/f7070142

Farhan, M., Wu, T., Amin, M., Tariq, A., Guluzade, R., & Alzahrani, H. (2024a). Monitoring and prediction of the LULC change dynamics using time series remote sensing data with Google Earth Engine. *Physics and Chemistry of the Earth, Parts A/B/C, 136*, Article 103689. https://doi.org/10.1016/j.pce.2024.103689

Farhan, M., Wu, T., Anwar, S., Yang, J., Naqvi, S. A. A., Soufan, W., & Tariq, A. (2024b). Predicting land use land cover dynamics and land surface temperature changes using CA-Markov chain models in Islamabad, Pakistan (1992-2042). *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17*, 16255-16271. https://doi.org/10.1109/JSTARS.2024.3441241

Fatah, K. K., Mustafa, Y. T., & Hassan, I. O. (2025). Assessing the influence of climate and land use/land cover changes on groundwater levels in arid and semi-arid regions: A geoinformatics-based machine learning approach in the Iraqi Kurdistan Region. *GeoJournal, 90*(3), Article 125. https://doi.org/10.1007/s10708-025-11374-w

Feng, S., Li, W., Xu, J., Liang, T., Ma, X., Wang, W., & Yu, H. (2022). Land use/land cover mapping based on GEE for the monitoring of changes in ecosystem types in the Upper Yellow River Basin over the Tibetan Plateau. *Remote Sensing, 14*(21), Article 5361. https://doi.org/10.3390/rs14215361

Forman, R. T. T. (1995). Some general principles of landscape and regional ecology. *Landscape Ecology, 10*(3), 133-142. https://doi.org/10.1007/BF00133027

Fu, Y., & Zhang, Y. (2022). Research on temporal and spatial evolution of land use and landscape pattern in Anshan City based on GEE. *Frontiers in Environmental Science, 10*, Article 988346. https://doi.org/10.3389/fenvs.2022.988346

Ge, G., Shi, Z., Zhu, Y., Yang, X., & Hao, Y. (2020). Land use/cover classification in an arid desert-oasis mosaic landscape of China using remotely sensed imagery: Performance assessment of four machine learning algorithms. *Global Ecology and Conservation, 22*, Article e00971. https://doi.org/10.1016/j.gecco.2020.e00971

Guo, Y., Yin, X., Zhao, X., Yang, D., & Bai, Y. (2019). Hyperspectral image classification with SVM and guided filter. *EURASIP Journal on Wireless Communications and Networking, 2019*, Article 56. https://doi.org/10.1186/s13638-019-1346-z

Hagner, O., & Reese, H. (2007). A method for calibrated maximum likelihood classification of forest types. *Remote Sensing of Environment, 110*(4), 438-444. https://doi.org/10.1016/j.rse.2006.08.017

Halder, B., Bandyopadhyay, J., & Banik, P. (2021). Monitoring the effect of urban development on urban heat island based on remote sensing and geospatial approach in Kolkata and adjacent areas, India. *Sustainable Cities and Society, 74*, Article 103186. https://doi.org/10.1016/j.scs.2021.103186

Hegazy, I. R., & Mosbeh, R. K. (2015). Monitoring urban growth and land use change detection with GIS and remote sensing techniques in Daqahlia Governorate, Egypt. *International Journal of Sustainable Built Environment, 4*(1), 117-124. https://doi.org/10.1016/j.ijsbe.2015.02.005

Intergovernmental Panel on Climate Change. (2006). *2006 IPCC guidelines for national greenhouse gas inventories*.https://www.ipcc-nggip.iges.or.jp/public/2006gl/vol4.html

Islam, M. S., & Islam, K. S. (2013). Application of thermal infrared remote sensing to explore the relationship between land use-land cover changes and urban heat island effect: A case study of Khulna City. *Journal of Bangladesh Institute of Planners, 6*, 49-60. https://doi.org/10.3329/jbip.v6i1.76962

Javidan, N., Kavian, A., Pourghasemi, H. R., Conoscenti, C., & Jafarian, Z. (2019). Gully erosion susceptibility mapping using multivariate adaptive regression splines: Replications and sample size scenarios. *Water, 11*(11), Article 2319. https://doi.org/10.3390/w11112319

Jodhani, K. H., Patel, D., Madhavan, N., Gupta, N., Singh, S. K., & Pandey, M. (2025). ML-based land use and land cover classification: Assessing performance and predicting future changes. *Journal of Hydrologic Engineering, 30*(4), Article 05025011. https://doi.org/10.1061/JHYEFF.HEENG-6416

Khan, A. R., Khan, A., Masud, S., & Rahman, R. M. (2021). Analysing the land cover change and degradation in Sundarbans Mangrove Forest using machine learning and remote sensing techniques. In I. Rojas, G. Joya, & A. Català (Eds.), *Advances in computational intelligence* (Lecture Notes in Computer Science, Vol. 12862, pp. 429-438). Springer. https://doi.org/10.1007/978-3-030-85099-9_35

Kogo, B. K., Kumar, L., & Koech, R. (2019). Analysis of spatiotemporal dynamics of land use and land cover changes in Western Kenya. *Geocarto International*, 1-16. https://doi.org/10.1080/10106049.2019.1608594

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. M., Angelsen, A., Bruce, J. W., Coomes, O. T., Dirzo, R., Fischer, G., & Folke, C. (2001). The causes of land-use and land-cover change: Moving beyond the myths. *Global Environmental Change, 11*(4), 261-269. https://doi.org/10.1016/S0959-3780(01)00007-3

Langat, P. K., Kumar, L., Koech, R., & Ghosh, M. K. (2019). Monitoring of land use/land-cover dynamics using remote sensing: A case of Tana River Basin, Kenya. *Geocarto International, 36*(13), 1470-1488. https://doi.org/10.1080/10106049.2019.1655798

Liu, J., Coomes, D. A., Gibson, L., Hu, G., Liu, J., Luo, Y., Wu, C., & Yu, M. (2019). Forest fragmentation in China and its effect on biodiversity. *Biological Reviews, 94*(5), 1636-1657. https://doi.org/10.1111/brv.12519

Liu, J., Xu, Q., Yi, J., & Huang, X. (2022). Analysis of the heterogeneity of urban expansion landscape patterns and driving factors based on a combined multi-order adjacency index and Geodetector model. *Ecological Indicators, 136*, Article 108655. https://doi.org/10.1016/j.ecolind.2022.108655

Mahata, B., Sahu, S. S., Sardar, A., Laxmikanta, R., & Maity, M. (2024). Spatiotemporal dynamics of land use/land cover (LULC) changes and their impact on land surface temperature: A case study in New Town Kolkata, eastern India. *Regional Sustainability, 5*(2), Article 100138. https://doi.org/10.1016/j.regsus.2024.100138

Manjunatha, M. C., & Basavarajappa, H. T. (2020). Mapping of land units and its change detection analysis in Chitradurga Taluk of Karnataka State, India, using geospatial technology. *International Advanced Research Journal in Science, Engineering and Technology, 7*(7), 61-68. https://doi.org/10.17148/IARJSET.2020.7711

Mateen, S., Nuthammachot, N., Techato, K., & Ullah, N. (2023). Billion Tree Tsunami forests classification using image fusion technique and random forest classifier applied to Sentinel-2 and Landsat-8 images: A case study of Garhi Chandan, Pakistan. *ISPRS International Journal of Geo-Information, 12*(1), Article 9.

McHugh, M. L. (2012). Interrater reliability: The kappa statistics. *Biochemia Medica, 22*(3), 276-282.https://doi.org/10.11613/BM.2012.031

Mukhopadhyay, S. C., & Dasgupta, A. (2010). *River dynamics of West Bengal: Physical aspects* (Vol. 1).

Nilsson, N. J. (1965). *Learning machines: Foundations of trainable pattern-classifying systems*. McGraw-Hill Book Company.

Ogato, G. S., Bantider, A., Abebe, K., & Geneletti, D. (2020). Geographic information system (GIS)-based multicriteria analysis of flooding hazard and risk in Ambo Town and its watershed, West Shoa Zone, Oromia Regional State, Ethiopia. *Journal of Hydrology: Regional Studies, 27*, Article 100659. https://doi.org/10.1016/j.ejrh.2019.100659

Phiri, D., Morgenroth, J., Xu, C., & Hermosilla, T. (2018). Effects of preprocessing methods on Landsat OLI-8 land cover classification using OBIA and random forests classifier. *International Journal of Applied Earth Observation and Geoinformation, 73*, 170-178. https://doi.org/10.1016/j.jag.2018.06.008

Raczko, E., & Zagajewski, B. (2017). Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *European Journal of Remote Sensing, 50*(1), 144-154. https://doi.org/10.1080/22797254.2017.1299557

Ramachandra, T. V., & Bharath, H. A. (2012). Spatiotemporal pattern of landscape dynamics in Shimoga, Tier II City, Karnataka State, India. *International Journal of Emerging Technology and Advanced Engineering, 2*(9), 563-576.

Ramachandra, T. V., & Kumar, U. (2011). Characterisation of landscape with forest fragmentation dynamics. *Journal of Geographic Information Systems, 3*(3), 242-252. https://doi.org/10.4236/jgis.2011.33021

Riitters, K. H., Coulston, J. W., & Wickham, J. D. (2012). Fragmentation of forest communities in the eastern United States. *Forest Ecology and Management, 263*, 85-93. https://doi.org/10.1016/j.foreco.2011.09.022

Rozenstein, O., Qin, Z., Derimian, Y., & Karnieli, A. (2014). Derivation of land surface temperature for Landsat-8 TIRS using a split-window algorithm. *Sensors, 14*(4), 5768-5780. https://doi.org/10.3390/s140405768

Ruidas, D., Pal, S. C., Towfiqul Islam, A. R. M., & Saha, A. (2023). Hydrogeochemical evaluation of groundwater aquifers and associated health hazard risk mapping using ensemble data-driven model in a water-scarce plateau region of eastern India. *Exposure and Health, 15*(1), 113-131. https://doi.org/10.1007/s12403-022-00480-6

Shafi, A., Chen, S., Waleed, M., & Sajjad, M. (2023). Leveraging machine learning and remote sensing to monitor long-term spatiotemporal wetland changes: Towards a national Ramsar inventory in Pakistan. *Applied Geography, 151*, Article 102868. https://doi.org/10.1016/j.apgeog.2022.102868

Tadesse, D., Suryabhagavan, K. V., Nedaw, D., & Hailu, B. T. (2022). A model-based flood hazard mapping in Itang District of the Gambella Region, Ethiopia. *Geology, Ecology, and Landscapes*. https://doi.org/10.1080/24749508.2021.2022833

Tripathi, V., Mohanty, M. P., & Singh, H. (2023). Fidelity of machine learning models in capturing flood inundation through geomorphic descriptors over Ganga Sub-basin [Conference paper]. *12th World Congress on Water Resources and Environment (EWRA 2023)*, Thessaloniki, Greece.

Twisa, S., & Buchroithner, M. F. (2019). Land-use and land-cover (LULC) change detection in Wami River Basin, Tanzania. *Land, 8*(9), Article 136. https://doi.org/10.3390/land8090136

Ullah, S., Khan, M., & Qiao, X. (2024a). Examining the impact of land use and land cover changes on land surface temperature in Herat City using machine learning algorithms. *GeoJournal, 89*(5), Article 225. https://doi.org/10.1007/s10708-024-11217-0

Ullah, S., Qiao, X., & Abbas, M. (2024b). Addressing the impact of land use land cover changes on land surface temperature using machine learning algorithms. *Scientific Reports, 14*(1), Article 18746. https://doi.org/10.1038/s41598-024-68492-7

Vapnik, V. (1999). *The nature of statistical learning theory*. Springer. https://doi.org/10.1007/978-1-4757-3264-1

Yousefi, S., Mirzaee, S., Almohamad, H., Al Dughairi, A. A., Gomez, C., Siamian, N., Alrasheedi, M., & Abdo, H. G. (2022). Image classification and land cover mapping using Sentinel-2 imagery: Optimisation of SVM parameters. *Land, 11*(7), Article 993. https://doi.org/10.3390/land11070993

Zenouzi, S. A., Yenneti, K., Teimouri, R., Abbasiyan, F., & Palme, M. (2022). Analysis of changes in vegetation index during the rapid urban spatial development period (1990-2020) in Tehran Metropolis, Iran. *Atmosphere, 13*(12), Article 2010. https://doi.org/10.3390/atmos13122010