International Journal of Applied Mathematics

Volume 38 No. 5 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

COMPARATIVE EVALUATION OF MACHINE LEARNING ALGORITHMS FOR LULC CLASSIFICATION USING LANDSAT 9 AND SENTINEL-2 DATA ON GEE: A CASE STUDY OF COASTAL KARNATAKA CITIES OF MANGALURU AND UDUPI, INDIA

Sumangala Nooji[1, 2*] Shashidhar Kini[3]

[1]Research Scholar, Institute of Computer Science and Information Science, Srinivas University, Mangaluru, India.

[2]Assistant Professor, Department of Computer Applications, St Joseph Engineering College, Mangaluru, India.

[3]Professor, Srinivas Institute of Technology, Mangaluru, India

E-Mail : skinimca@sitmng.ac.in

[*]**Corresponding E-mail:** n.sumangala@gmail.com

Abstract

Accurate mapping of Land Use and Land Cover (LULC) is the fundamental requirement for urban policy, environmental science, and sustainable development goals, as it ensures informed decisions that promote organized urban growth, protect ecosystems, and enable sustainable land use. Large-scale satellite data analysis has become computationally efficient with the emergence of platforms like Google Earth Engine (GEE). This study compares five machine learning methods, including Random Forest (RF), Gradient Boosting Machine (GBM), Classification and Regression Trees (CART), Support Vector Machine (SVM), and k Nearest Neighbor (kNN), to classify LULC in the region covering Mangaluru and Udupi. It also includes the transport corridor buffer between the two cities and uses data from both Sentinel-2 and Landsat 9. Seasonal composites for the wet and dry periods were created and improved using several remote sensing indices, including NDVI, NDWI, MNDWI, NDBI, SAVI, and IBI. Digital Elevation Model (DEM) data were also added to help separate the land-cover classes more effectively. The key contribution of this study lies in incorporating spectral indices, DEM data, and seasonal composites to evaluate their impact on classification performance. RF and GBM algorithms have given competitive results, achieving high accuracies of 98.48% and 98.99%, respectively, with Sentinel-2 data outperforming Landsat-9 by generating more accurate LULC maps. The study highlights the capacity of the modern classifiers on GEE to deliver high-accuracy, reproducible LULC products.

Keywords: Machine Learning, Land Use and Land Cover (LULC), GEE, Remote Sensing

1. Introduction

Changes in Land Use Land Cover(LULC) impact the ecosystem, biodiversity, and, in turn, affect human well-being. Such studies or applications developed for urban growth analysis, agricultural pattern analysis, and analysis of variations in climate, pollution, and forest degradation consider the Classification of Land Use Land Cover(LULC) as the crucial first step [1]. In the field of LULC analysis, traditional manual interpretation of maps is often time-consuming and challenging. Traditional manual interpretation of maps is challenging because it is difficult to analyze large geographic areas across multiple maps. To overcome these challenges, today GIS techniques are used, which make use of remote sensing data for efficient monitoring of LULC, and it is also easy to incorporate these data derived from such tools in various related applications. GEE is an effective cloud platform that allows the processing of remote sensing data using various Machine Learning(ML) algorithms across large study areas. Its strength lies in combining an open-source framework with massive storage capacity for geospatial datasets, including the entire Landsat and

Received: August 04, 2025

International Journal of Applied Mathematics

Volume 38 No. 5 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Sentinel archives, and substantial processing power with the help of Google's computational network. Crucially, GEE integrates various ML classification algorithms, allowing researchers to perform rapid, automated, and accurate LULC mapping and change analysis without the need for local data storage or powerful personal computers[2].

Ensemble-based models are recognized for their superior performance in LULC classification. The most popular approaches among them are Support Vector Machine(SVM), Gradient Boosting Machines (GBM), and Random Forest (RF)[3]. Individual studies have successfully applied these algorithms and compared data from earth observation sensors such as Sentinel and Landsat[4][5]. However, few studies have systematically compared these algorithms across Sentinel-2, Landsat 9 datasets within a reproducible GEE workflow. This study systematically addresses the existing research gap in LULC classification by carrying out a comprehensive study on the GEE platform. The methodology involves generating wet and dry seasonal composites for both Sentinel-2 and Landsat 9 imagery, and further enriching the dataset by computing six spectral indices to enhance class distinction. A direct, quantitative comparison is performed across five key ML algorithms, RF, GBM, a distance-based k-Nearest Neighbor (kNN), SVM, and a decision tree approach called Classification and Regression Trees (CART), for classifying five distinct LULC classes.

2. Literature Survey

Remote sensing has come to the forefront as a key approach for mapping LULC because it allows for repeated observations over large areas and across different time periods. Landsat, Sentinel, and MODIS are the widely used satellite platforms that are commonly used for studies involving LULC analysis. These satellites are popular due to their spectral range, high frequency revisit, and spatial resolution[6]. High-precision LULC maps can be generated using ML techniques as they are capable of processing large amounts of data and better handling complex landscapes. Possible misclassifications when a single pixel of an image covers more than one type of land cover class can be handled in a better way, thereby by errors can be reduced with ML methods when compared to the conventional pixel-based classification methods [7].

The LULC classification made using ML techniques is more reliable and accurate for various applications, including environmental monitoring, urban growth analysis, study of ecological changes, and resource management. Various such studies are found in the literature demonstrating the application of ML algorithms for this purpose, highlighting their capability in capturing complex spatial patterns and improving thematic accuracy. A study focusing on urban analysis using time series data from Landsat 5, Landsat 7, and Landsat 8 has used RF within GEE to produce high-accuracy maps [8]. This study used auxiliary data such as Surface thermal data and a Digital Elevation Model (DEM) along with spectral bands. In addition to this, the authors have pansharpened the image to reduce the spatial resolution of Landsat images to 15m. The generated LULC maps with 94.44% accuracy were targeted for long-term environmental monitoring applications.

Another study in the fast-growing city of Lahore demonstrated the capability of three ML classifiers, namely CART, SVM, and RF, for accurate LULC classification. The study processed Landsat 7, Landsat 8, and Landsat 9 imagery for the years 2008, 2015, and 2022 on GEE, giving a prominent result of 95.2% Overall accuracy with a Kappa value of 0.87. The study highlighted the usefulness of ML algorithms and recommended RF as a reliable algorithm for similar studies focusing on monitoring LULC changes in urbanized environments [9].

SVM, Random Trees, and CART were implemented and compared for mapping LULC classes by integrating Sentinel-1 and Sentinel-2 satellite data. The evaluation showed that the Radial Basis Function SVM significantly performed better than others, achieving the top Overall Accuracy of 89.6% and F1 of 89.5%. This research concludes that integrating ML algorithms like SVM is highly effective for generating accurate LULC maps for spatiotemporal analysis of the landscape [10].

A study on aquatic and tree varieties investigated the impact of disparities in radiometric resolution across Landsat 8 and Landsat 9 on classification accuracy. The study did not find a significant impact of such disparity in accuracy and also found that Landsat 9 is better for such studies, as it enhances accuracy compared to Landsat 8 [11]. Another recent

International Journal of Applied Mathematics

Volume 38 No. 5 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

study incorporated multiple sensor data effectively by considering the vertical structure of trees from GEDI LiDAR, spectral information from Sentinel-2, and Sentinel- 1 radar data for mapping tree crops [12] Facing a scarcity of ground reference data, the researchers first combined and evaluated the distinct features from these three sensors at the single GEDI footprint scale using advanced models like Dynamic Time Warping, incorporating time weights and RF. This high-quality, footprint-level classification was then leveraged as a robust training dataset to generate a map that provides complete, continuous geographic coverage of tree crops across the eastern Mediterranean and southern France.

3. Data And Methods

3.1 Study Area

Mangaluru, the administrative center of Dakshina Kannada District in southern Karnataka, is the largest coastal city in the region. As a rapidly urbanizing Tier II city, it is part of India's Smart Cities Mission launched in 2015. Udupi district, located in Karnataka's north Malabar coastal region, has its administrative headquarters in the city of Udupi. The study area covers these two growing cities of Karnataka and the connecting National Highway, including a one-kilometer buffer zone on both sides of the highway. The study area was chosen with the intention of capturing heterogeneity, as it includes wide-ranging classes such as built environments, forests, water-filled areas, agricultural land, and more, rather than being dominated by a single land type. The location map, presented in Fig. 1, depicts the geographic background of the study region.

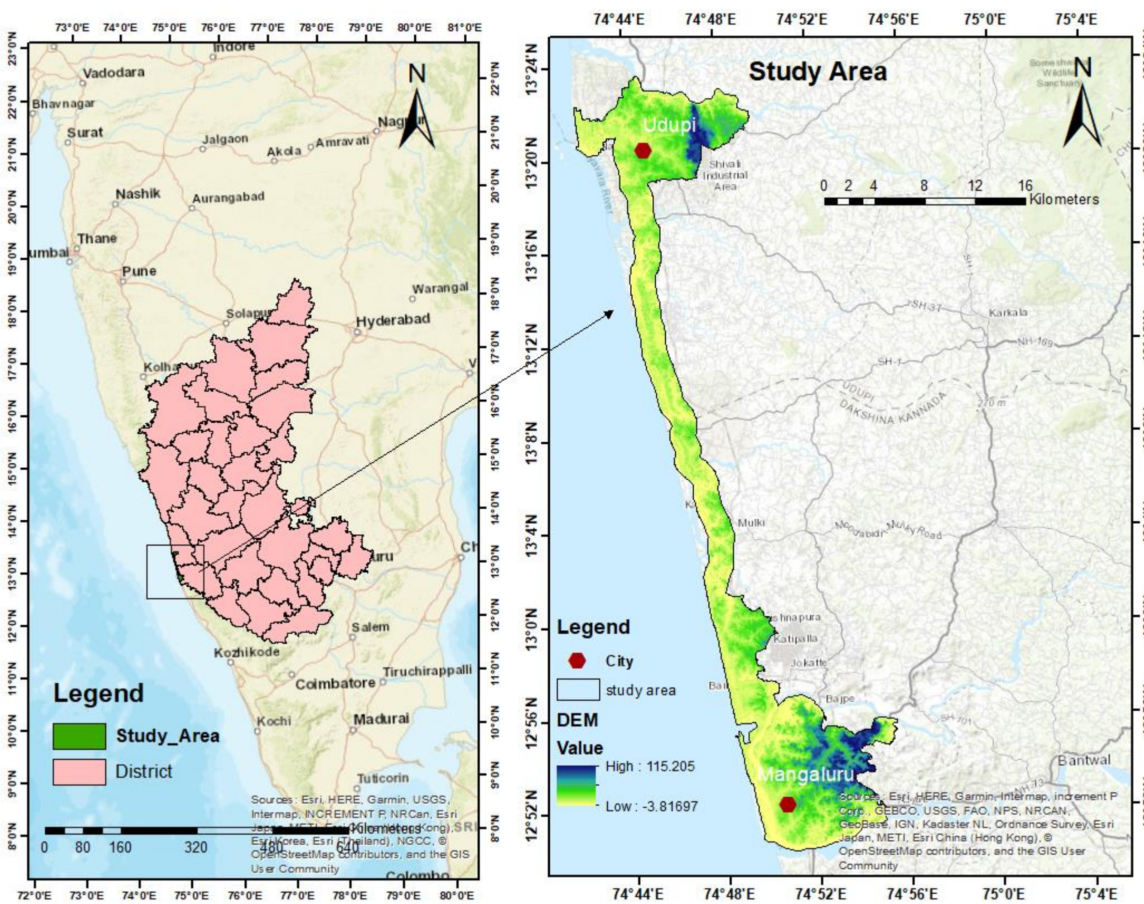

Fig 1. Map depicting the study location

Received: August 04, 2025

Volume 38 No. 5 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.2 Sources of Data

This investigation made use of Landsat-9 and Sentinel-2 satellite data sets retrieved from the Earth Engine collection for LULC classification. Satellite images collected over a full one-year period, from June 2024 to May 2025, were used in this study.

3.2.1 Sentinel- 2

The Earth Engine asset ee.ImageCollection ('COPERNICUS/ S2_SR_HARMONIZED') was imported, and the data set refers to the Harmonized Sentinel-2 Surface Reflectance collection. This dataset provides high-resolution, multi-spectral imagery for land monitoring, which has been processed to create a consistent time series from both Sentinel-2A and 2B satellites. Bands B2 (Blue), B3 (Green), B4 (Red), and B8 (NIR) at 10-m resolution, along with two 20-m resolution shortwave infrared spectral channels, namely B11(SWIR 1) and B12(SWIR 2), were considered in this study.

3.2.2 Landsat 9

The dataset USGS Landsat 9 Level 2, Collection 2, Tier 1 was used on Earth Engine, which offers surface reflectance and land surface temperature corrected for atmospheric effects. This is generated from the data produced by the Landsat 9 OLI/TIRS instrument. 6 spectral bands, namely SR_B2 (Blue), SR_B3 (Green), SR_B4 (Red), SR_B5 (NIR), and SR_B6 (SWIR 1), each with a spatial resolution of 30 m, were utilized.

3.3 Methodology

A comprehensive feature-engineering workflow was implemented to prepare inputs for LULC classification. The raw spectral bands were first acquired from the satellite data and complemented by a collection of spectral indices that highlight key biophysical properties as specified in Table 1. In addition, topographic variables derived from the DEM—such as slope and aspect were incorporated to account for terrain-driven variations. All spectral, index-based, and topographic layers were then stacked to create a multiband feature composite, ensuring each pixel carried integrated information on reflectance, moisture, vegetation, and terrain. As the next step, season-specific wet and dry composites were generated through direct band-ratio processing, enabling enhanced characterization of seasonal surface conditions. Finally, the season-specific wet and dry composites, along with all associated spectral, index, and topographic layers from both seasons, were layer stacked to produce a unified feature space derived from multi-temporal data for the LULC categorization model. The present method is useful as the study area experiences very significant monsoon and dry season fluctuations, as it has a distinct tropical monsoon climate, and land cover appears differently across seasons. Following the feature preparation and composite generation steps, various ML algorithms were applied to perform the LULC classification. High-accuracy classified maps were generated for both Landsat and Sentinel datasets. The methodology is described in Fig. 2. Feature engineering and LULC classification using machine-learning algorithms were performed on GEE through the JavaScript API. High-resolution maps were produced in ArcGIS, and the other visualizations, such as charts and heat maps, were generated using the Python programming language on the Google Colab platform.

International Journal of Applied Mathematics
Volume 38 No. 5 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

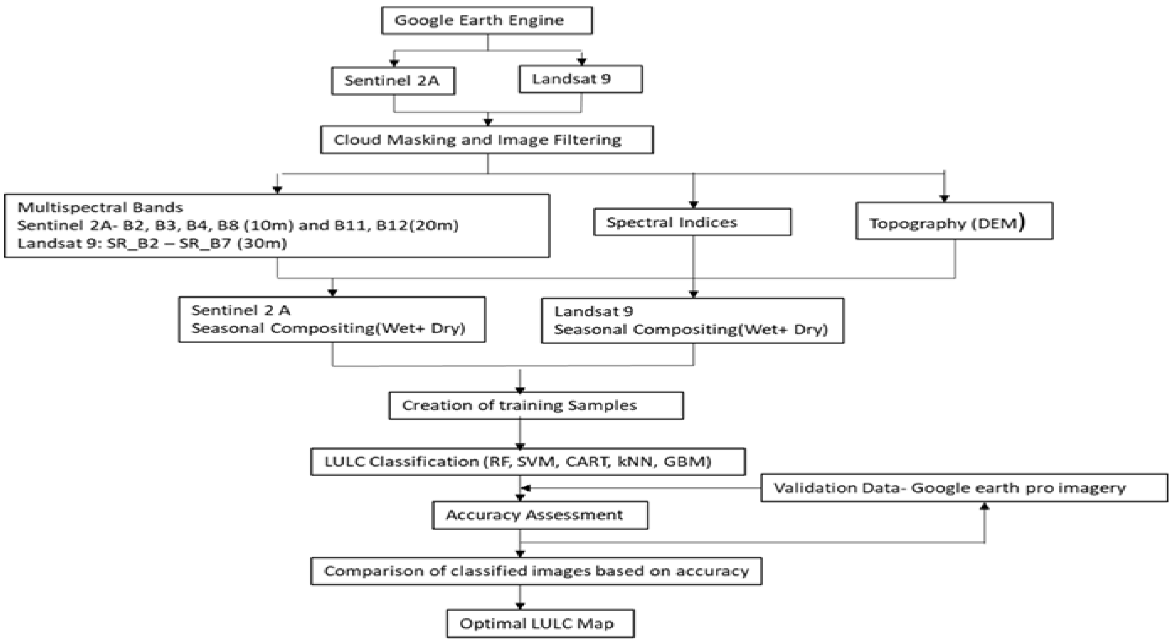

Fig. 2 Flow diagram of the proposed methodology

3.3.1 Spectral Derived Metrics

Table 1 presents the spectral reflectance indices used in the classification process to enhance the spectral distinction between specific land cover classes.

Table 1. List of spectral indices used in the current research

Spectral Index	Formula	Key Role in Classification	Reference
NDVI	$NDVI = \frac{(NIR - RED)}{RED + NIR}$	Delineating and assessing the health/density of green vegetation.	[13]
NDWI	$NDWI = \frac{(GREEN - NIR)}{(GREEN + NIR)}$	Detecting and mapping open water features	[14][15]
NDBI	$NDBI = \frac{(SWIR - NIR)}{SWIR + NIR}$	Identifying urban and settlement areas.	[16]
SAVI	$SAVI = (\frac{(NIR - RED)}{(NIR + RED + L)}) \times (1 + L)$ Where L is the soil brightness correction factor	Estimating vegetation cover while minimizing the influence of background soil brightness.	[17]
MNDWI	$MNDWI = \frac{(GREEN - SWIR)}{(GREEN + SWIR)}$	More accurately extracting water, particularly in urban areas, by suppressing noise from built-up areas and soil.	[15]
IBI	$IBI = \frac{(NDBI - (MNDWI + SAVI)/2)}{(NDBI + (MNDWI + SAVI)/2)}$	Optimizing built-up area extraction by maximizing its separability from bare soil and vegetation simultaneously.	[18]

International Journal of Applied Mathematics
Volume 38 No. 5 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.3.2 LULC Classification Algorithms

Five ML algorithms were implemented on the GEE platform to classify LULC types identified for the study area, namely vegetation, agriculture, openland, water bodies, and built-up area during the period from June 2024 to May 2025. The LULC types identified for this study were based on the existing research on land use analysis of Mangaluru agglomeration[19].

RF is a machine learning approach that creates a collection of decision trees by randomly sampling both the training data and the input features for each tree. Introducing this randomness helps make each tree more unique, which in turn boosts the model's overall accuracy and reliability. As more trees are added, the model's error rate generally begins to level off. This behavior depends mainly on how accurate each tree is and how similar the trees are to one another. When more trees are included, the model's error rate tends to stabilize. This is largely determined by the strength of each tree and how much the trees are correlated with one another[20].

SVMs work by finding the best boundary that separates different classes of data. This boundary is chosen to be as wide as possible, which helps the model make accurate predictions on new data. SVM uses kernel functions to change the dimensionality of the input space to a higher level, where the class boundaries are found to be distinct and make them linearly separable. Choosing a suitable kernel function and setting up the other connected parameters are important to achieve improved performance of SVM[21].

Classification of complex environmental data sets can be achieved with the help of the CART approach, which is a flexible decision tree method. The approach is also suitable for regression tasks. It uses predictor variables to split the dataset into uniform subsets, which are based on predictor variables, and this step is repeated to produce a clear and instinctive tree structure. The model is capable of generating easily interpretable data modelling nonlinear patterns, variable interactions, and incomplete data. This approach is able to manage different response variables, maintain robustness, and provide transparent decision rules. So it can be considered as an alternative to more traditional statistical techniques[22].

The kNN method groups the closest neighbors and assigns the labels to new samples based on the dominant class in the group. Commonly, Euclidean distance is used to identify the neighbors in this distance-based approach. The computational cost of kNN increases with the increase in the size of the data set, and also selecting the right value of k is an important factor to be considered while incorporating this approach[23].

GBM is considered an effective technique to generate accurate predictive models. It constructs decision trees by adding weak learners. Residual errors left by the preceding models are minimised while training each new tree. This method gradually improves the performance of the ensemble. Important features are selected at the time of training, and it maintains reliability even when there are missing values or outliers in the data, and it is able to work with diverse data types[24].

3.3.3 Classification Accuracy Assessment and Validation

Using the reference data sets, the results of the LULC calculation were evaluated. The confusion matrices were generated in GEE comparing the land cover data with the predicted classes. The standard error matrices generated organize the actual classes in rows and predicted classes in columns. The diagonal elements of the matrix represent the correctly classified samples of each category. Using this matrix, the accuracy metrics such as User's Accuracy, Producer Accuracy, Overall Accuracy, and Kappa coefficients were computed. The User's accuracy was computed using Equation 1, which indicates the stability of the classified categories for end-users. The Producer's Accuracy was calculated using Equation 2, which reflects how accurately each LULC type on the ground was mapped. The proportion of correctly classified samples is provided by the Overall Accuracy, which was computed using Equation 3, and the Kappa coefficient was computed using Equation 4, which measures the degree of agreement between the reference and predicted categories beyond random chance[25][26].

International Journal of Applied Mathematics
Volume 38 No. 5 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$User's\ Accuracy = \frac{Number\ of\ correctly\ classified\ pixels\ in\ each\ category}{Total\ number\ of\ classified\ pixels\ in\ that\ category(Row\ Total)} \times 100 \tag{1}$$

$$Producer's\ Accuracy = \frac{Number\ of\ correctly\ classified\ pixels\ in\ each\ category}{Total\ number\ of\ reference\ pixels\ in\ that\ category(Column\ Total)} \times 100 \tag{2}$$

$$Overall\ Accuracy = \frac{Total\ number\ of\ correctly\ classified\ pixels(diagonal)}{Total\ number\ of\ refrence\ pixels} \times 100 \tag{3}$$

$$Kappa\ Coefficient = \frac{Observed\ Accuracy - chance\ agreement}{1 - chance\ agreement} \tag{4}$$

ArcGIS was used for creating validation points and which were labelled carefully using Earth Pro imagery, ensuring reliable ground truth information. This validation framework minimized bias, strengthened the assessment of the machine-learning models, and confirmed that the final classified maps met the required accuracy standards.

4. Results And Discussion

The study area was grouped into five LULC classes using Landsat 9 and Sentinel-2 imagery during the study period. The feature engineering was performed by layer stacking spectral indices along with spectral band data and seasonal composites. The effectiveness of ML algorithms in this classification process using these two types of satellite images was compared. The study used five ML algorithms, namely SVM, kNN, CART, GBM, and RF, for classification. As mentioned in the Methodology section, User Accuracy, Producer Accuracy, Overall accuracy, and the Kappa coefficient were computed from the error matrices of respective classification results on GEE. The paper includes heat maps, as shown in Fig. 3 and Fig. 4, showing the performance of these algorithms across the two satellite data sets.

The Landsat-based heat maps shown in Fig. 3 show that the algorithms do not perform equally across the different LULC classes. It can be observed that RF and GBM came out as the most reliable models, producing steady and high user and producer accuracies for every category. Compared to the other ensemble approaches, the results of CART are found to be a little lower, but it has also performed quite well in terms of accuracy values. Mixed outcomes are observed with kNN, which fairly identified water and vegetation categories, but the performance is found to be low for agriculture and open land. Most variation in accuracy results is observed with SVM, achieving good accuracy for water and vegetation but dropping noticeably for built-up and open land classes. Overall, the performance of ensemble methods with both Landsat and Sentinel imagery is found to outperform other approaches, and the heat maps clearly highlight the class-specific limitations seen in kNN and SVM.

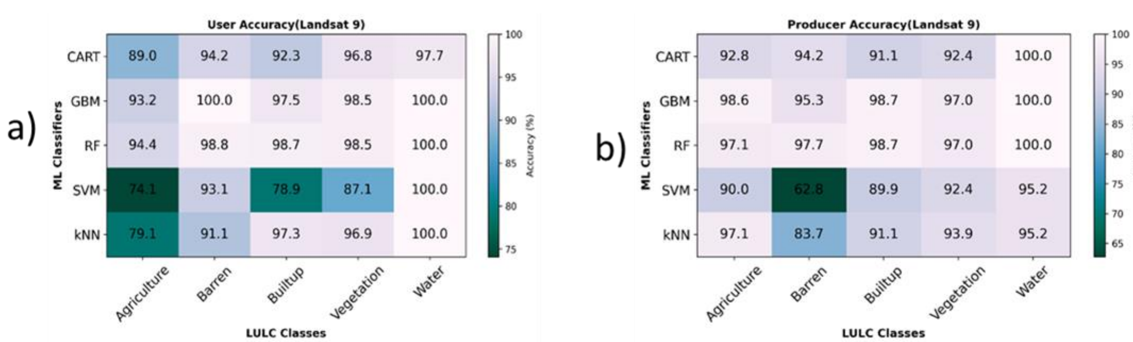

Fig. 3 Heat maps showing (a) User Accuracy and (b) Producer Accuracy for the different LULC classes derived from Landsat-9 imagery.

International Journal of Applied Mathematics
Volume 38 No. 5 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The Sentinel-based results shown in Fig. 4 show that all five classifiers perform quite well, though the differences between them become clearer when looking at the heatmaps. RF and GBM again come out on top, offering the most stable and highest accuracies across every land-cover category. kNN and SVM also produce strong results, particularly for built-up areas, vegetation, and water, while their accuracy drops slightly for agriculture and barren land. CART performs reliably overall but shows a small decline in accuracy for the built-up and barren classes when compared with the ensemble methods. Taken together, the Sentinel-2 outcomes suggest that each algorithm is capable of handling the dataset effectively, with RF and GBM providing the most consistent and well-balanced performance.

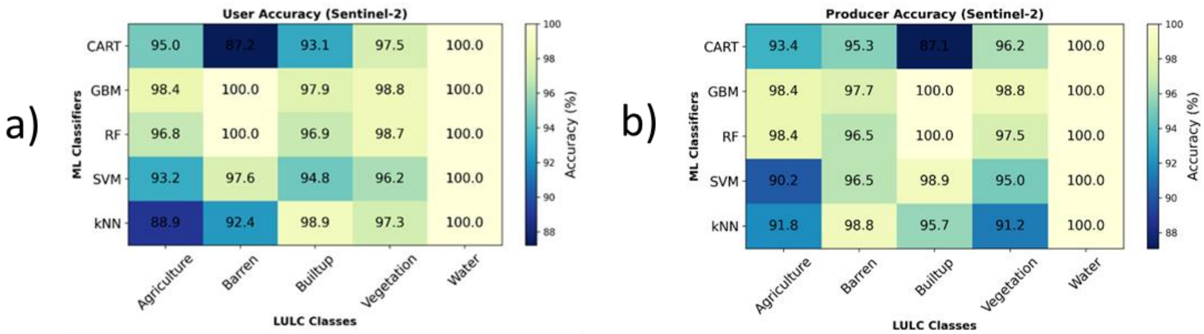

Fig. 4 Heat maps presenting (a) User Accuracy and (b) Producer Accuracy for all LULC classes based on the Sentinel-2 classification results.

The overall accuracy and kappa comparison of the five ML algorithms for both the Landsat and Sentinel datasets is shown in Fig. 5. The comparison of overall accuracy and kappa values between Landsat and Sentinel datasets shows a consistent improvement in classification performance when using Sentinel imagery. All algorithms achieve higher accuracies with Sentinel, reflecting the benefit of its finer spatial and spectral resolution. RF and GBM deliver the strongest results for both sensors, with Sentinel accuracies exceeding 98 percent and kappa values of 0.98. kNN and CART also show noticeable improvement with Sentinel, indicating better class separability in the higher-resolution data. Among the ML approaches, SVM shows a significant increase in overall accuracy of 96.47% from 84.25% and an increase in the kappa coefficient of 0.95 from 0.80 with Sentinel from Landsat. Overall, the comparative analysis highlights that all classifiers benefit from the enhanced quality of Sentinel imagery, leading to more reliable land cover mapping. In addition to this, we can say that the ensemble algorithms are the top-performing approaches.

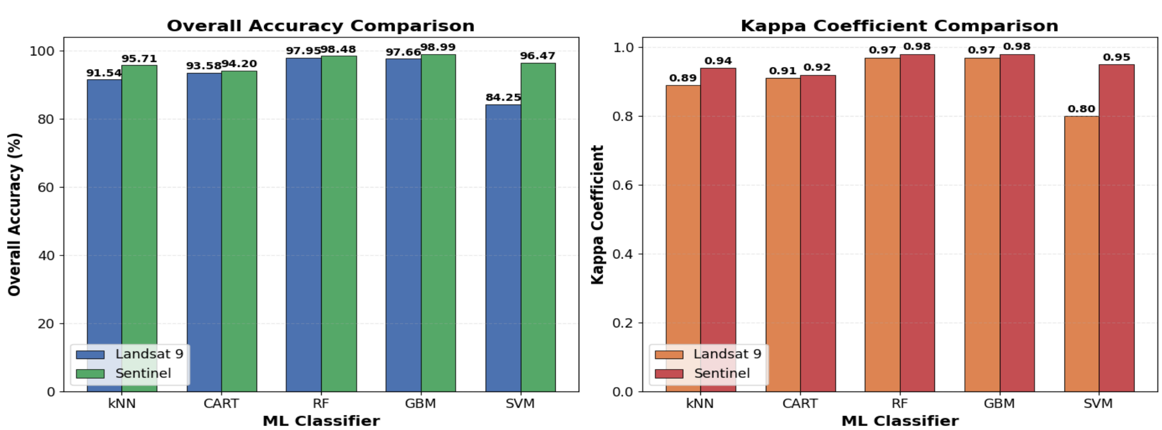

Fig. 5 Comparison of ML classifiers in terms of (a) Overall Accuracy and (b) Kappa Coefficient

Finally, RF was chosen for generating the final LULC maps because it delivered consistently high accuracy and kappa values for both Landsat and Sentinel datasets. Although GBM also performed exceptionally well and emerged as a

International Journal of Applied Mathematics

Volume 38 No. 5 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

strong competitor, RF offered greater stability, robustness to noise, and reduced risk of overfitting, making it more suitable for reliable map production. Its balanced performance across all land cover classes and computational efficiency further supported its selection over GBM as the preferred classifier for this study. The LULC maps generated from the Landsat 9 and Sentinel-2 datasets are presented in Figure 6.

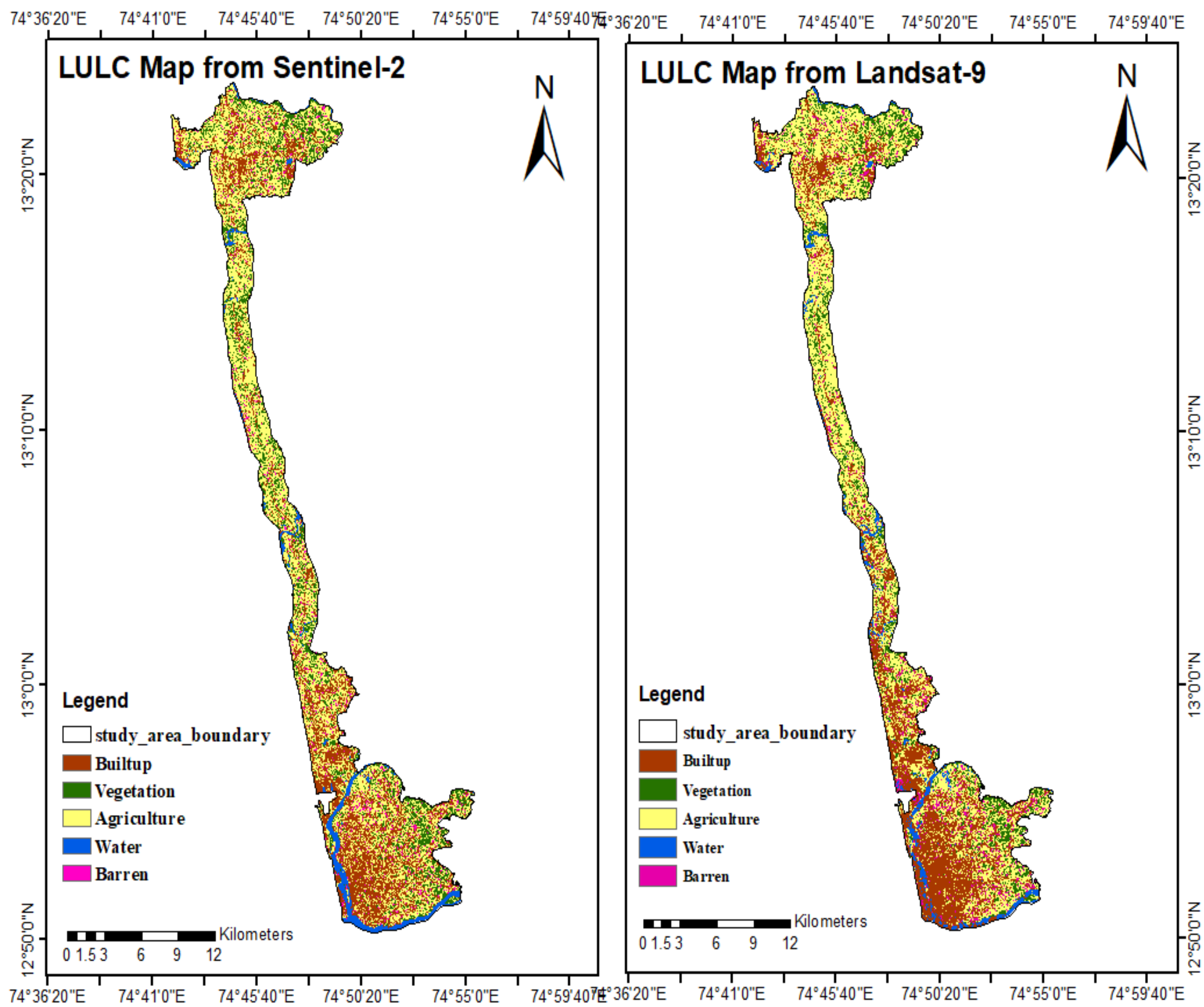

Fig 6. LULC classification maps generated using Random Forest for (a) Sentinel-2 and (b) Landsat 9 imagery.

The LULC area statistics derived from Landsat and Sentinel (Fig. 7) show broadly similar patterns, with both datasets identifying agriculture as the dominant land cover type, followed by built-up areas and vegetation. Sentinel reports slightly higher vegetation, agriculture, and water coverage, exceeding Landsat estimates by 3.18 ha, 1.95 ha, and 0.41 ha, respectively, whereas Landsat indicates a larger extent of built-up and open land by 4.42 ha and 1.3 ha. The higher resolution of Sentinel images would have contributed to the finer separation of green areas and water bodies, and the coarser resolution of Landsat may have generalized mixed pixels into built-up or open land, giving higher estimates for these classes. Overall, consistent land cover trends could be observed with the usage of Sentinel and Landsat datasets. The minor differences reflect the difference in sensor resolution and spectral detail that affect the LULC mapping accuracy.

Received: August 04, 2025

International Journal of Applied Mathematics
Volume 38 No. 5 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

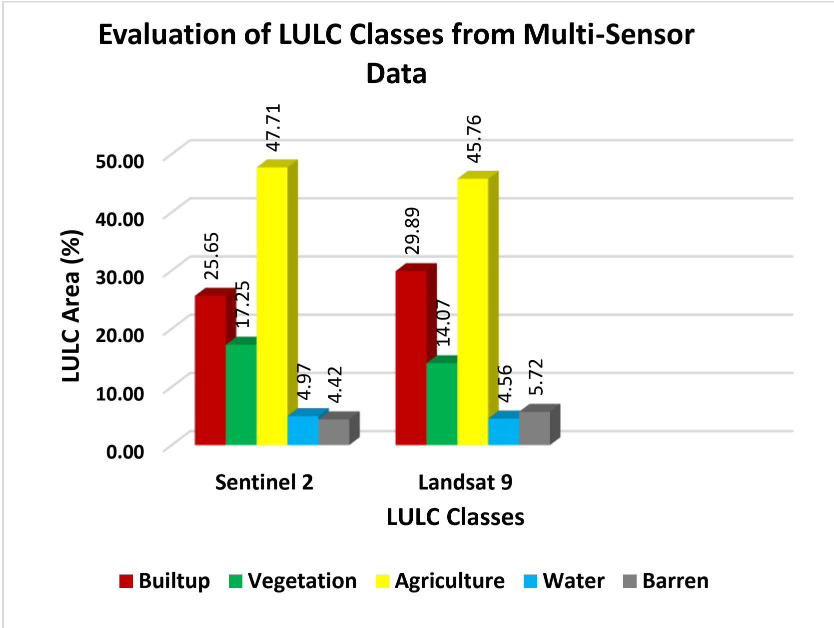

Fig. 7. Evaluation of LULC class areas derived from Sentinel-2-2 and Landsat 9 datasets for the study area.

5. Conclusion

This approach has used the original multispectral bands, spectral indices, and the Digital Elevation Model as input features to improve classification accuracies of ML algorithms. By stacking the dry and wet season composites, the model gets information from both seasons. As the single image does not give the full range of conditions of the land cover throughout a period, this study has used seasonal composites where the wet and dry seasonal images helped in capturing the seasonal differences, including greenness of vegetation, soil wetness, and water spread. Such multiple seasons' data helped in separating the LULC classes, effectively providing better features for the ML models to improve their classification accuracy.

The study has shown an effective way of processing large-scale Sentinel-2 and Landsat 9 data for LULC classification using the advanced ML algorithms on GEE and GIS tools such as ArcGIS for mapping. In comparison with SVM, CART, and kNN, the RF approach has shown extremely good performance in classifying all LULC categories accurately. RF has come out as a competitor for GBM with high accuracy and with the additional advantage of less risk of overfitting and ease of hyperparameter tuning. As a result, the classification became more accurate. The practical usefulness of the study is enhanced by including the transport-corridor buffer zone along with the coastal cities of Karnataka, and it is particularly relevant for policymakers involved in urban planning and infrastructure monitoring. The comparative analysis of ML algorithms with Sentinel and Landsat imagery provides a clear picture of the accuracy and scalability of various ML approaches for LULC mapping. Overall, this comprehensive evaluation is useful for selecting an appropriate combination of satellite imagery and ML algorithms for regional as well as large-scale LULC analysis on the GEE platform.

References

[1] T. Anthony, A. A. A. Shohan, A. Oludare, S. Alsulamy, A. Al Kafy, and K. M. Khedher, "Spatial analysis of land cover changes for detecting environmental degradation and promoting sustainability," *Kuwait Journal of Science*, vol. 51, no. 2, p. 100197, 2024, doi: 10.1016/j.kjs.2024.100197.

[2] W. Tesfaye, E. Elias, B. Warkineh, M. Tekalign, and G. Abebe, "Modeling of land use and land cover changes using google earth engine and machine learning approach: implications for landscape

Volume 38 No. 5 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

management," *Environmental Systems Research*, vol. 13, no. 1, 2024, doi: 10.1186/s40068-024-00366-3.

[3] H. Du, M. Li, Y. Xu, and C. Zhou, "An Ensemble Learning Approach for Land Use/Land Cover Classification of Arid Regions for Climate Simulation: A Case Study of Xinjiang, Northwest China," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 2413–2426, 2023, doi: 10.1109/JSTARS.2023.3247624.

[4] Z. Deng, X. Zhu, Q. He, and L. Tang, "Land use/land cover classification using time series Landsat 8 images in a heavily urbanized area," *Advances in Space Research*, vol. 63, no. 7, pp. 2144–2154, 2019, doi: 10.1016/j.asr.2018.12.005.

[5] L. M. Bill Donatien, B. Biona Clobite, and M. Lemvo Meris Midel, "Comparing Sentinel-2 and Landsat 9 for land use and land cover mapping assessment in the north of Congo Republic: a case study in Sangha region," *International Journal of Remote Sensing*, vol. 45, no. 22, pp. 8015–8036, 2024, doi: 10.1080/01431161.2024.2394238.

[6] V. C. Radeloff *et al.*, "Need and vision for global medium-resolution Landsat and Sentinel-2 data products," *Remote Sensing of Environment*, vol. 300, no. October 2023, p. 113918, 2024, doi: 10.1016/j.rse.2023.113918.

[7] S. Talukdar *et al.*, "Land-use land-cover classification by machine learning classifiers for satellite observations-A review," *Remote Sensing*, vol. 12, no. 7, 2020, doi: 10.3390/rs12071135.

[8] S. Amini, M. Saber, H. Rabiei-Dastjerdi, and S. Homayouni, "Urban Land Use and Land Cover Change Analysis Using Random Forest Classification of Landsat Time Series," *Remote Sensing*, vol. 14, no. 11, pp. 1–23, 2022, doi: 10.3390/rs14112654.

[9] Z. Zafar, M. Zubair, Y. Zha, S. Fahd, and A. Ahmad Nadeem, "Performance assessment of machine learning algorithms for mapping of land use/land cover using remote sensing data," *Egyptian Journal of Remote Sensing and Space Science*, vol. 27, no. 2, pp. 216–226, 2024, doi: 10.1016/j.ejrs.2024.03.003.

[10] G. Bayable, G. Gebrie, T. Melese, and A. Melaku, "Land use/cover classification using machine learning algorithms and their impacts on land surface temperature and soil moisture in the Alawuha Watershed, Ethiopia," *Environmental and Sustainability Indicators*, vol. 27, no. April, p. 100797, 2025, doi: 10.1016/j.indic.2025.100797.

[11] H. You, X. Tang, W. Deng, H. Song, Y. Wang, and J. Chen, "A Study on the Difference of LULC Classification Results Based on Landsat 8 and Landsat 9 Data," *Sustainability (Switzerland)*, vol. 14, no. 21, 2022, doi: 10.3390/su142113730.

[12] E. Adrah, J. P. Wong, and H. Yin, "Integrating GEDI, Sentinel-2, and Sentinel-1 imagery for tree crops mapping," *Remote Sensing of Environment*, vol. 319, no. December 2024, p. 114644, 2025, doi: 10.1016/j.rse.2025.114644.

[13] K. Mehmood *et al.*, "Exploring spatiotemporal dynamics of NDVI and climate-driven responses in ecosystems: Insights for sustainable management and climate resilience," *Ecological Informatics*, vol. 80, no. February, p. 102532, 2024, doi: 10.1016/j.ecoinf.2024.102532.

[14] S. Liu, J. Qiu, and F. Li, "A Remote Sensing Water Information Extraction Method Based on Unsupervised Form Using Probability Function to Describe the Frequency Histogram of NDWI: A Case Study of Qinghai Lake in China," *Water (Switzerland)*, vol. 16, no. 12, 2024, doi: 10.3390/w16121755.

[15] F. Hossain, M. A. Kamal, and T. Afrin, "Fluvio-geomorphic change of the Padma-Meghna river course using the NDWI and MNDWI techniques," *Water Science*, vol. 38, no. 1, pp. 293–310, 2024, doi: 10.1080/23570008.2024.2344752.

[16] A. K. Bhomi, R. Poudyal, S. K. Tolange, and S. Chaudhary, "Assessing the Impact of Urban Expansion on Forest Cover using LULC Maps, NDVI, and NDBI: A Case Study of Kathmandu District," *Journal on*

Volume 38 No. 5 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Geoinformatics, Nepal, pp. 1–7, 2024, doi: 10.3126/njg.v23i1.66045.

[17] K. Kocur-Bera and A. Małek, "Assessing the Feasibility of Using Remote Sensing Data and Vegetation Indices in the Estimation of Land Subject to Consolidation," *Sensors*, vol. 24, no. 23, pp. 1–18, 2024, doi: 10.3390/s24237736.

[18] S. Singh, R. Avtar, A. K. Jain, W. S. Alkhuraiji, and M. Zhran, "Influence of Land Use/Land Cover Dynamics on Urban Surface Metrics in Semi-Arid Heritage Cities," *Land*, vol. 14, no. 9, pp. 1–20, 2025, doi: 10.3390/land14091834.

[19] K. Dhanaraj and D. P. Angadi, "Land use land cover mapping and monitoring urban growth using remote sensing and GIS techniques in Mangaluru, India," *GeoJournal*, vol. 87, no. 2, pp. 1133–1159, 2022, doi: 10.1007/s10708-020-10302-4.

[20] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[21] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, 2020, doi: 10.1016/j.neucom.2019.10.118.

[22] G. De Ath and K. E. Fabricius, "Death Fabricius_2000 _CART," *Ecology*, vol. 81, no. 11. pp. 3178–3192, 2000.

[23] A. Pandey and A. Jain, "Comparative Analysis of KNN Algorithm using Various Normalization Techniques," *International Journal of Computer Network and Information Security*, vol. 9, no. 11, pp. 36–42, 2017, doi: 10.5815/ijcnis.2017.11.04.

[24] Y. Liu, S. Wu, Z. Wu, and S. Zhou, "Application of gradient boosting machine in satellite-derived bathymetry using Sentinel-2 data for accurate water depth estimation in coastal environments," *Journal of Sea Research*, vol. 201, no. July, 2024, doi: 10.1016/j.seares.2024.102538.

[25] P. Dash, S. L. Sanders, P. Parajuli, and Y. Ouyang, "Improving the Accuracy of Land Use and Land Cover Classification of Landsat Data in an Agricultural Watershed," *Remote Sensing*, vol. 15, no. 16, pp. 1–24, 2023, doi: 10.3390/rs15164020.

[26] C. B. Pande *et al.*, "Characterizing land use/land cover change dynamics by an enhanced random forest machine learning model: a Google Earth Engine implementation," *Environmental Sciences Europe*, vol. 36, no. 1, 2024, doi: 10.1186/s12302-024-00901-0.

Received: August 04, 2025