Comparison of Classifiers in Sentinel-2 and Landsat 9 Datasets for Land Use and Land Cover Classification in GEE

Tejash Anand[1]*, Aadrita Chowdhury[1], Srashti Singh[2], & Anugya Shukla[3]

[1] Center for Geoinformatics, Jamsetji Tata School of Disaster Studies, TISS, Mumbai.
[2] Civil Engineering Department, IIT - Roorkee.
[3] Centre for Emerging Technology for Sustainable Development, IIT – Jodhpur.
*Corresponding Author's email: tejash.85214@gmail.com

Abstract

High-resolution multispectral remote sensing images offer crucial insights into diverse land characteristics, ensuring accurate spatial representations. Classification precision often falls short within intricate urban settings when using the entire original multispectral bands. This study investigates the effectiveness of employing multispectral satellite bands, spectral indices (including NDVI (Normalized Difference Vegetation Index), NDBI (Normalized Difference Built-up Index), and NDWI (Normalized Difference Water Index)), and principal component analysis for feature extraction and selection. These methods are evaluated individually and in combination for classifying various land use and cover (LULC) features. The classification process employs three machine learning (ML) algorithms—Minimum Distance (MD), Random Forest (RF), and Support Vector Machine (SVM) executed on the Google Earth Engine (GEE) platform. GEE offers a wide range of tools and functions for processing and analyzing satellite imagery and geospatial data due to its ability to handle and process large datasets efficiently. An accuracy assessment is carried out to assess each classifier's performance. The quantitative analysis is done by comparing results from these three ML algorithms on Landsat 9 and Sentinel-2 datasets. The study provides insights into the LULC results by each classifier using both datasets. The study helps decision-makers, practitioners, and researchers to select the appropriate feature extraction and classification algorithm.

Keywords: Image classification, SVM, Random Forest, Sentinel-2, Landsat–9, Google Earth Engine

Introduction

Land Use Land Cover (LULC) is used in many studies around the world to study the change detection in the environment, such as biodiversity, water, soil quality, etc. It is primarily used by the decision-maker regarding the development of the city and site selection. There is a need to have good accuracy of the map to avoid wrong decisions and make the process easy. The image classification is categorized into two parts: supervised and unsupervised classification of the image (Forkuor et al., 2018). Google Earth Engine (GEE) was launched in 2010 at the ICC Conference in Mexico. The platform is ready to use data for scientific analysis, transformation, processing, disaster response, and water resource mapping (*Introducing Google Earth Engine*, 2010).

Satellite image classification can be defined based on predefined characteristics and features. These categories include supervised classification, which uses a training sample to

classify the image. It uses Nearest Neighbours, Navies Bayes, Support Vector Machine (SVM), Minimum Distance (MD), Random Forest (RF), Decision Tree, etc as a method to conclude the task. The unsupervised classification is based on the similarity without prior knowledge of the classes using methods like K-means, Gaussian mixture, etc. (Borra et al., 2019; Srivastava, 2020). At the same time, Object-based classification is highly used in high-resolution images to separate object characteristics. With the rise in ML technologies in the past two decades, few advanced methods have grabbed people's attention in the process of LULC classification. RF, a regression and classification method, uses random selection and a decision tree to give the result by averaging it. It is highly accurate for a large number of data sets because of its unbiased estimates, estimation of variables, and effective method of estimating the data and maintaining the accuracy. It can also have a few drawbacks since sometimes it is having over-fitting may give noisy classification (Butt et al., 2020; *RandomForest*, 2023). MD classifier calculates the minimum distance from the mean value of a particular class from the digital number of each pixel (Abinaya & Poonkuntran, 2019). Various mathematical operations can be used to calculate this minimum distance, such as Euclidean, Mahalanobis, etc (we are using Euclidean in this research). SVM determines small subsets as support points and uses them to classify large classes based on the convex optimization and kernel function, allowing efficient learning of generalized functions. It is largely used due to its performance on the large-scale dataset (Mathar et al., 2020).

SVM is a supervised learning algorithm for classification and regression. It finds a hyperplane to separate data into classes, maximizing the margin between them, making it effective for binary classification (Mathar et al., 2020). RF is a robust classifier using an ensemble of decision trees for accurate predictions. MD classification assigns data points to classes based on their proximity to class centroids, which is calculated using the Euclidean distance metric, making it a simple and intuitive classification method (Abinaya & Poonkuntran, 2019).

Multi-source data or multi-indices, if used appropriately, can increase the accuracy of the ML supervised classification. There are several indices based on the band statistics such as NDVI is used to determine the type of vegetation in a particular area, NDWI is used to determine the presence of water and whether it is flood water or permanent source of water, NDBI is the to get the built-up of the region and is also very helpful the determine the barren or fallow land available in the region (Serrano et al., 2019).

Materials and methods

Study Area: Bhopal is a city located in Madhya Pradesh in Central India. The coordinates of Bhopal, Madhya Pradesh, are approximately 23.2599° N latitude and 77.4126° E longitude (Duhon, 2015). Bhopal city, the 16th largest city in India and the 131st largest city in the world is known as the City of Lakes because of its natural and artificial lakes (Singh & Jain, 2022). This city is at the upper limit of the Vindhya Mountain ranges, comprising uneven elevation and small hills within the boundaries of its region. The average elevation of this city is 500m (1401 ft) (Ghosh, 2019). The forested area represents a mixed deciduous forest, making it suitable for cultivating various crops. Bhopal City is situated in a semi-arid climate. Studies have been done in the past examining the usefulness of supervised classification techniques for detecting land use/land cover changes in semi-arid regions using GIS and

remote sensing (Ghurah et al., 2018). In recent years, the city has faced rapid expansion in the urban agglomeration and growth in population, too (Singh & Jain, 2023).

Data and Tools Used: Landsat 9 and Sentinel 2 data are used for the last week of December. Landsat 9 provides data with a spatial resolution of 30m (panchromatic at 15m) and a temporal resolution of 16 days. The spatial resolution of the Sentinel 2 is 10m, with a temporal resolution of 5 days. It provides the data in 13 spectral bands.

GEE is a cloud-based platform that uses JavaScript specially designed for performing the spatial analysis, development and customization of huge datasets. The current repository has information about various satellites, including Landsat-9 and Sentinel-2B, and Geographic Information System (GIS), such as elevation models, climate data, demographics, etc. (Mutanga & Kumar, 2019).

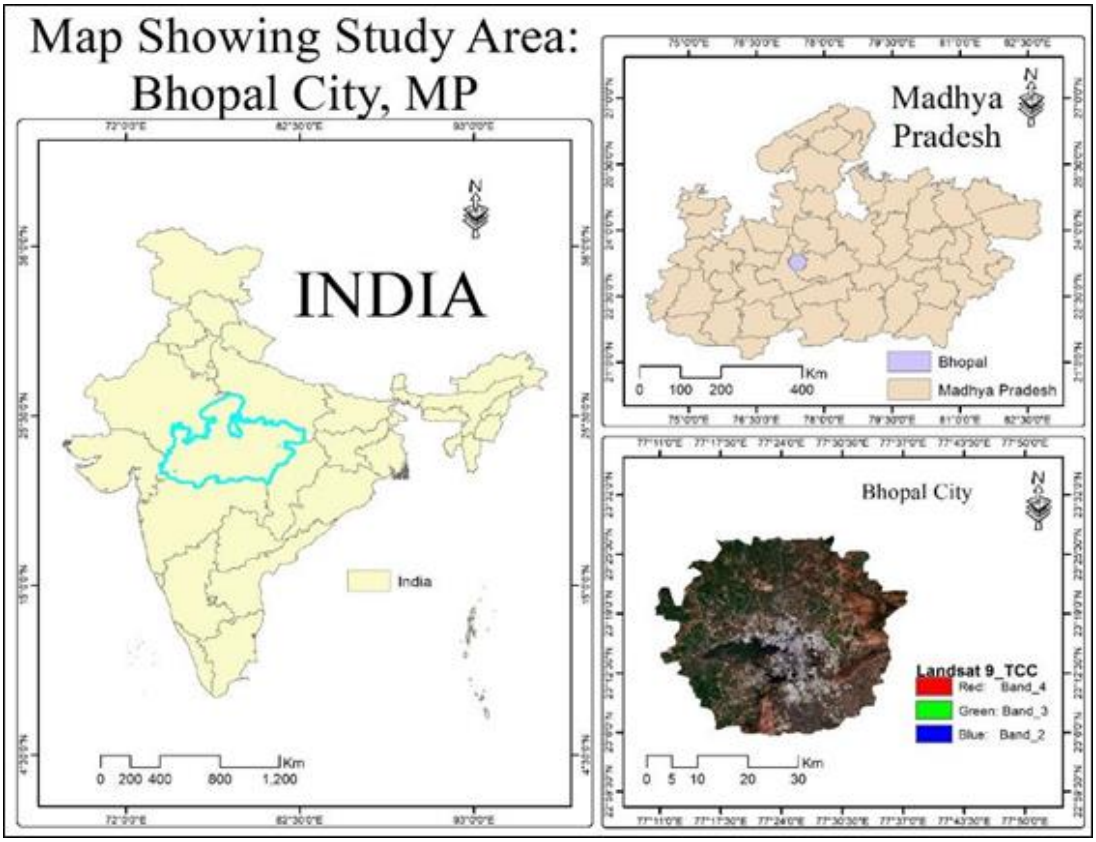

Fig. 1 Study Area of Bhopal, Madhya Pradesh, India (TCC – True Color Composition).

Methodology: LULC maps can be helpful in planning and decision-making. The high-resolution multispectral datasets make the development of the regional and global LULC possible. However, still, there is a question on the accuracy of the LULC datasets available. The accuracy is a concern in the open source dataset and when mapping LULC using various ML classifiers (Forkuor et al., 2018). The pre-processing of the data is done after being taken out from the GEE data catalogue. This pre-processing involves clipping the region of interest and filtering the particular tile comprised of Bhopal. We calculated NDVI, NDBI and NDWI, as shown in Figure 2 and Figure 3, using the equations that follow.

NDVI = (NIR – Red)/ (NIR + Red) Eq. (1)

NDVI is calculated by the difference between the NIR (Near Infrared) band and the red band divided by the sum of both. It ranges from -1 to 1, where all the negative values represent water bodies, and the positive values from 0 to 0.05 represent the built-up and the barren or fallow land. As the value increases, the density of the vegetation also increases from sparse to moderate to dense vegetation (Atun et al., 2020).

NDWI = (Green – NIR)/(Green + NIR) Eq. (2)

NDWI is calculated by dividing the difference between Green and NIR bands by their sum. It ranges from -1 to 1 where all value from -1 to -0.3 represents the drought and non-aqueous surface, from -0.3 to 0 represents moderate drought or the non–aqueous surfaces, 0 to 0.2 represents the flooding and non-aqueous surfaces and the value from 0.2 to 1 represents the water surface in the region (*NDWI*, 2021).

NDBI = (SWIR – NIR)/ (SWIR + NIR) Eq. (3)

NDBI is calculated by the difference between the SWIR (Shortwave Infrared) band and NIR (Near-Infrared) band divided by their sum. It ranges from -1 to 1, where all the values from -1 to 0.01 represent water, -0.01 to 0.49 represent soil, vegetation and minor built-up, and -0.49 to 1 represent the significant built-up region (Watik & Jaelani, 2019).

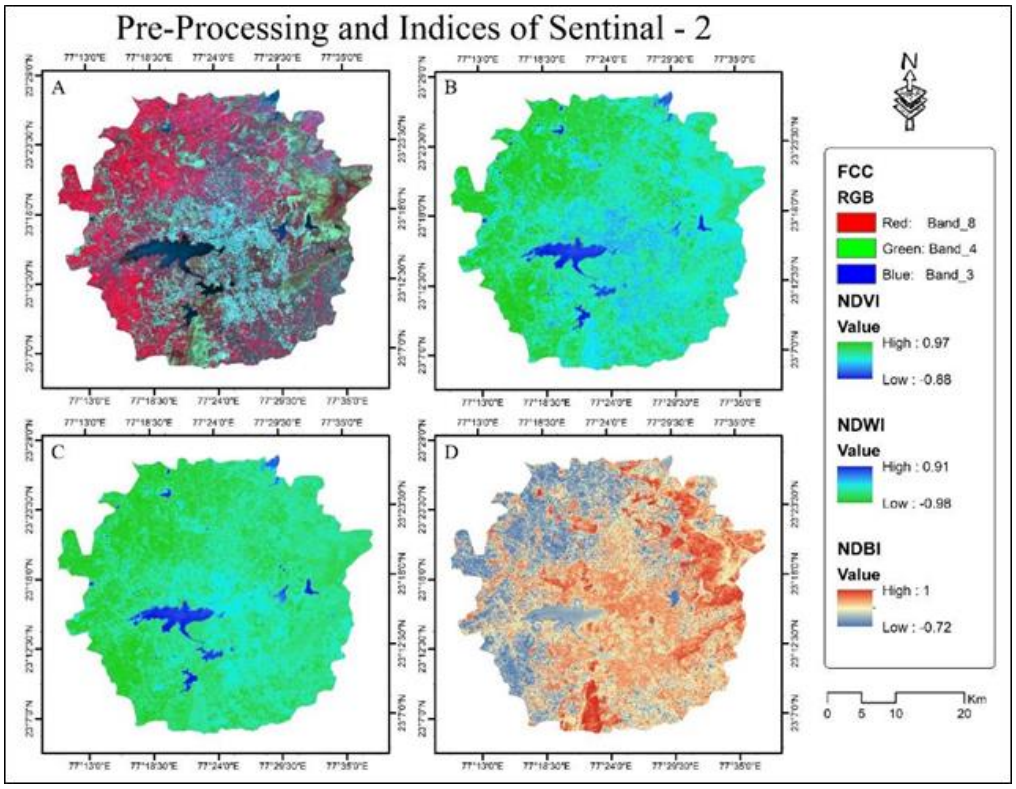

Fig. 2 Preprocessing of Sentinel 2 image for the training points collection; (A) False Colour Composite (FCC), (B) NDVI, (C) NDWI, (D)NDB.I

With the help of the calculated spectral indices, i.e., NDVI, NDBI and NDWI, training points were taken to train the data for the LULC classification. Classification was done for six classes, namely: water, fallow land, hilly terrain, agriculture, dense vegetation and built-up area. The same training data is now used to conduct the LULC classification using three ML classifiers: SVM, MD and SMILE (Statistical Machine Intelligence and Learning Engine) Random Forest. Then the common testing samples are taken to check the accuracy of the image classification using a confusion matrix. We calculated the Overall Accuracy (OA) and Kappa Coefficient of the image using the following formulae.

OA = (Total No. of correct pixels in all categories)/ (Total No. of pixels) Eq. (4)

Kappa Coefficient (K_{hat}) = (OA – Expected Accuracy)/ (1 – Expected Accuracy) Eq. (5)

After making the LULC by all three classifiers, the accuracy of the classified image obtained by each classifier was checked separately using the same testing sample.

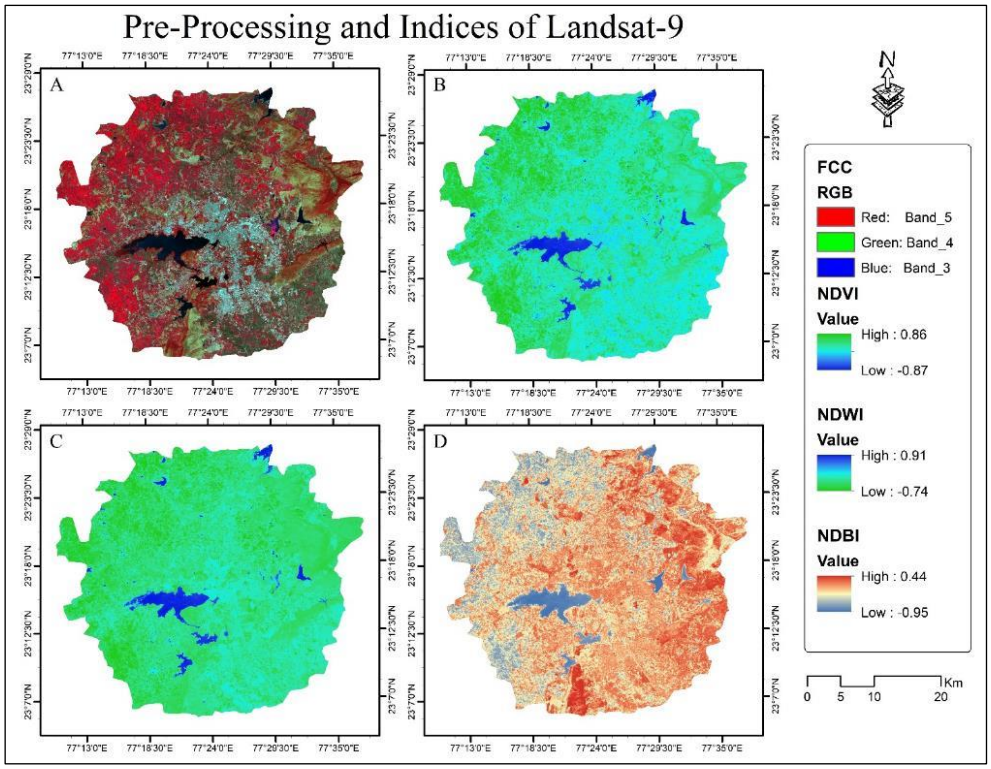

Figure 3. Preprocessing of Landsat 9 image for the training points collection; (A) False Color Composite (FCC), (B) NDVI, (C) NDWI, (D) NDBI.

Results and Discussion

The training sample for classifying the imageries was obtained using the calculated indices and FCC visualization. Using the three ML classifiers, LULC maps were obtained for both the satellite images, which were then tested for accuracy. Kappa coefficient and OA are used as the two indicators of accuracy.

Table 6. Accuracy Assessment of LULC images by different ML classifiers.

	Landsat 9		Sentinel 2	
	Kappa Coefficient	**OA**	**Kappa Coefficient**	**OA**
SVM	64.65%	71.42%	85.82%	88.52%
MD	71.90%	75.22%	72.27%	77.04%
RF	90.42%	92.20%	87.71%	90.16%

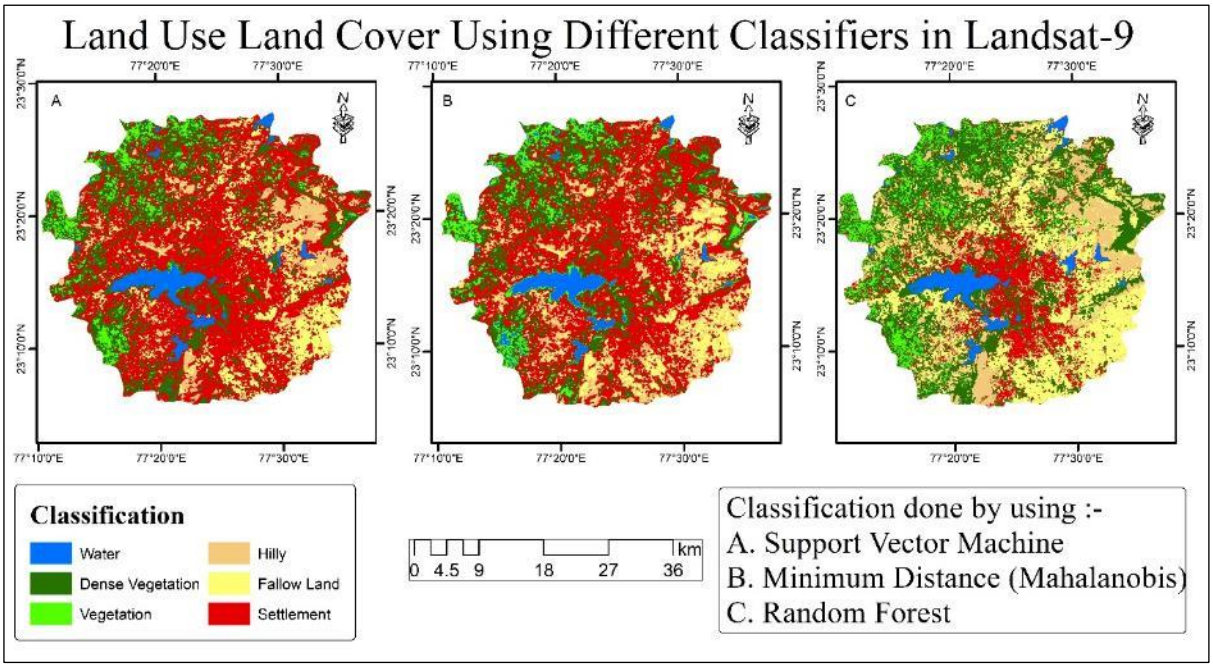

Fig. 4 LULC classification using Landsat 9.

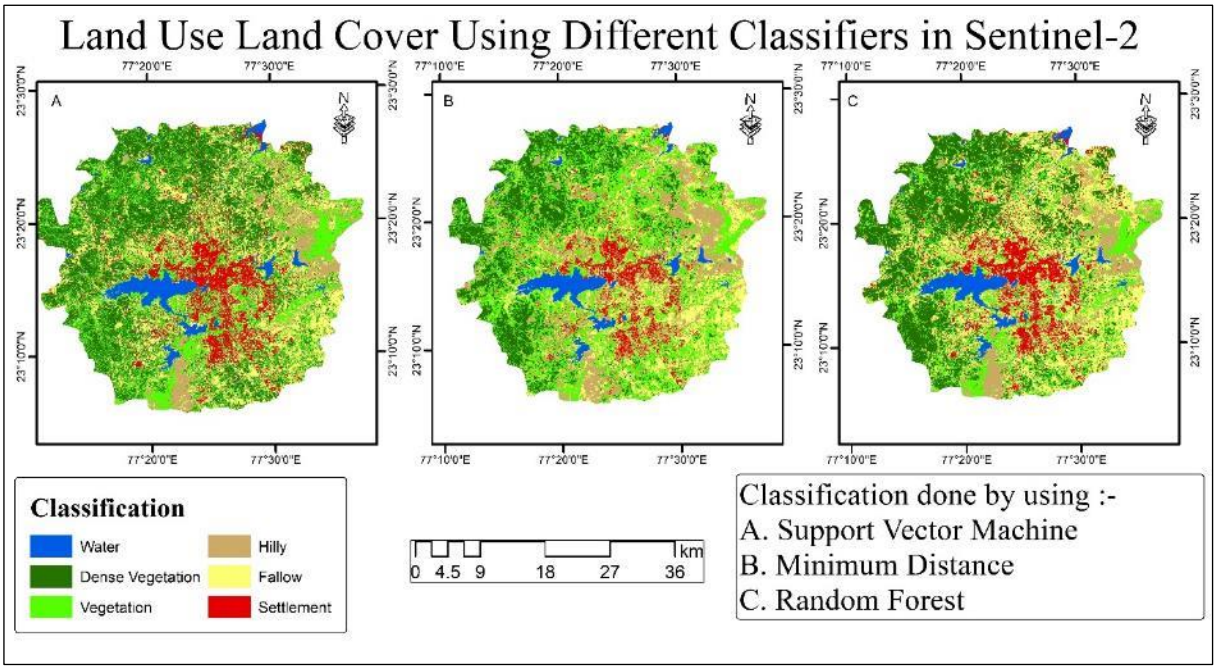

Fig. 5 LULC classification using Sentinel 2.

Figure 4 and Figure 5 represent the output classified LULC images by the three classifiers for Landsat 9 and Sentinel 2, respectively. Table 1 represents the accuracy

assessment results by the three classifiers for both Landsat 9 and Sentinel 2. It is observed in the study that RF and SVM resulted in higher overall accuracies for Sentinel 2 LULC images, and RF and MD classifiers resulted in high overall accuracies in Landsat 9 LULC images. RF proved to be the best classifier in both cases, with OA of 90.16% for Sentinel-2 and 92.20% for Landsat 9. It was also observed that the misclassification of water and fallow land with urbanization was most prominent in Landsat-9 images, particularly in the MD and SVM classifier, as shown in Figure 6. Table 2 represents the area covered by each of the six classes in LULC obtained by each classifier for both Landsat 9 and Sentinel 2.

Further, we compared the number of pixels in the water class from the classified image with the NDWI image. It was revealed that RF (in Landsat 9 and Sentinel 2) and SVM (in Sentinel 2) quantified the closest match in the number of water pixels classified with actual water pixels. Figure 5 for Sentinel 2 shows that RF classified the pixels more accurately than other classifiers. Figure 6 depicts the water body in all classifiers for both datasets. It can be seen from the figure that the SVM classifier extracted water pixels more accurately than MD. It is also observed in the accuracy assessment that RF gives the best results in LULC classification in the Landsat 9 dataset.

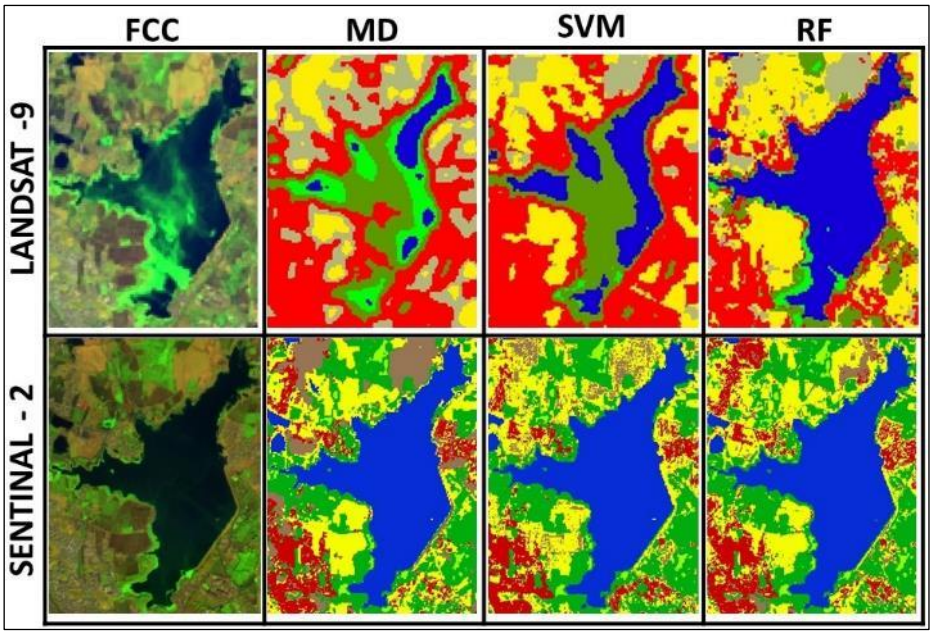

Fig. 6 Surface water body classification result.

Table 7 Areas covered by each class (in km^2).

	Landsat 9			Sentinel 2		
	MD	**SVM**	**RF**	**MD**	**SVM**	**RF**
Water	55.43	51.62	57.09	58.28	62.05	64.80
Dense Vegetation	274.56	273.18	429.49	240.89	431.49	342.51
Vegetation	91.09	64.59	95.08	441.67	284.95	269.88
Hilly	207.91	124.76	191.33	200.11	78.30	98.13
Fallow Land	102.54	140.54	333.33	287.92	343.43	409.16
Settlement	592.07	668.92	217.28	94.72	123.37	139.11
Total	1324	1324	1324	1324	1324	1324

The study provides valuable insights into the effectiveness of combining various inputs from surface reflectance of the satellite data for ML-based land cover classification. It can be concluded that the accuracy of LULC depends upon the spatial resolution of the data used and the classifier selection. The study's specific choice of classifiers and input data limited the analysis. A potential enhancement to this research would involve expanding the sample size used for training and testing, which may yield more resilient and broadly applicable results. Furthermore, using imagery from two satellites, Landsat 9 and Sentinel-2, reveals essential information about sensor-based variations in the methodology and classifier performance.

Conclusion

In this research, remote sensing data was employed to assess the efficacy of three ML classifiers in land cover classification. The study incorporated index-based combinations like NDWI, NDVI, and NDBI to train the classifier. The primary goal was to identify the most proficient classifiers for precise land cover classification. The research also investigated the accuracy of these classifiers specifically for the targeted feature class. The results unveiled substantial variations in accuracy depending on the dataset's spatial resolutions and the choice of classifiers when distinguishing between land cover categories. The study concludes by stating that the RF classifier is best suited among all the classifiers tested, irrespective of the class.

References

Abinaya, V., & Poonkuntran, D. S. (2019). *Classification of Satellite Image using Minimum Distance Classification Algorithm*.

Atun, R., Kalkan, K., & Gürsoy, Ö. (2020). *Determining The Forest Fire Risk with Sentinel 2 Images*.

Borra, S., Thanki, R., & Dey, N. (2019). Satellite Image Classification. In S. Borra, R. Thanki, & N. Dey (Eds.), *Satellite Image Analysis: Clustering and Classification* (pp. 53–81). Springer. https://doi.org/10.1007/978-981-13-6424-2_4

Butt, A. M., Bhatti, Y. K., & Hussain, F. (2020). Emotional Speech Recognition Using SMILE Features and Random Forest Tree. In Y. Bi, R. Bhatia, & S. Kapoor (Eds.), *Intelligent Systems and Applications* (pp. 10–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-29516-5_2

Duhon, H. (2015). Bhopal: A Root Cause Analysis of the Deadliest Industrial Accident in History. *Oil and Gas Facilities*, *3*(03), 24–28. https://doi.org/10.2118/0614-0024-OGF

Forkuor, G., Dimobe, K., Serme, I., & Tondoh, J. E. (2018). Landsat-8 vs. Sentinel-2: Examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso. *GIScience & Remote Sensing*, *55*(3), 331–354. https://doi.org/10.1080/15481603.2017.1370169

Ghosh, S. (2019). A city growth and land-use/land-cover change: A case study of Bhopal, India. *Modeling Earth Systems and Environment*, *5*(4), 1569–1578. https://doi.org/10.1007/s40808-019-00605-y

Ghurah, M. A., Kamarudin, M. K. A., Wahab, N. A., Umar, R., Wan, N. A. F. N., Juahir, H., Gasim, M. B., Hassan, A. R., Lananan, F., Yusra, A. F. I., Sunardi, S., & Hidayat, Y. (2018). Temporal change detection of land use/land cover using GIS and remote sensing techniques in South Ghor Regions, Al-Karak, Jordan. *Journal of Fundamental and Applied Sciences*, *10*(1S), Article 1S. https://doi.org/10.4314/jfas.v10i1S

Introducing Google Earth Engine. (2010, December 2). Google. https://blog.google/outreach-initiatives/sustainability/introducing-google-earth-engine/

Mathar, R., Alirezaei, G., Balda, E., & Behboodi, A. (2020). Support Vector Machines. In R. Mathar, G. Alirezaei, E. Balda, & A. Behboodi (Eds.), *Fundamentals of Data Analytics: With a View to Machine Learning* (pp. 83–105). Springer International Publishing. https://doi.org/10.1007/978-3-030-56831-3_6

Mutanga, O., & Kumar, L. (2019). Google Earth Engine Applications. *Remote Sensing*, *11*(5), 591. https://doi.org/10.3390/rs11050591

NDWI: Index Formula, Value Range, And Uses In Agriculture. (2021, September 29). https://eos.com/make-an-analysis/ndwi/

Palanisamy, P. A., Jain, K., & Bonafoni, S. (2023). Machine Learning Classifier Evaluation for Different Input Combinations: A Case Study with Landsat 9 and Sentinel-2 Data. *Remote Sensing*, *15*(13), 3241. https://doi.org/10.3390/rs15133241

RandomForest. (2023, October 15). https://haifengl.github.io/api/java/smile/regression/RandomForest.html

Singh, S., & Jain, K. (2022). GEOSPATIAL APPROACH FOR URBAN ENVIRONMENTAL QUALITY ASSESSMENT. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIII-B3-2022*, 705–711. https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-705-2022

Singh, S., & Jain, K. (2023). The Changing Face of Urbanization, Population Expansion, and its Driving Forces: A Case Study of an Indian City. *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, 3013–3016. https://doi.org/10.1109/IGARSS52108.2023.10283086

Srivastava, S. (2020). *A survey on satellite image classification approaches*.

Watik, N., & Jaelani, L. (2019). Flood Evacuation Routes Mapping Based on Derived- Flood Impact Analysis From Landsat 8 Imagery Using Network Analyst Method. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-3/W8*, 455–460. https://doi.org/10.5194/isprs-archives-XLII-3-W8-455-2019

Citation

Anand, T., Chowdhury, A., Singh, S., Shukla, A. (2024). Comparison of Classifiers in Sentinel-2 and Landsat 9 Datasets for Land Use and Land Cover Classification in GEE. In: Dandabathula, G., Bera, A.K., Rao, S.S., Srivastav, S.K. (Eds.), Proceedings of the 43rd INCA International Conference, Jodhpur, 06–08 November 2023, pp. 313–321, ISBN 978-93-341-2277-0.