Jahangirnagar University
Journal of Geospatial Sciences
Volume 1, 2025; DOI: https://doi.org/10.59185/jgs.v1i1.267

Research Article

Adaptability of Machine Learning-based LULC Classification Techniques for Natural Resource Monitoring: A Study on an Oxbow Lake in Manikganj, Bangladesh

Sakib Rahman Siddique Shuvo [1*], Sharmin Sultana Alis [2], Md. Nurul Islam [3]

[1] Postgraduate Student, Institute of Water and Flood Management, Bangladesh University of Engineering and Technology (BUET), Dhaka-1000, Bangladesh
[2] Assistant GIS Specialist, Sheltech Consultants Pvt. Ltd., Panthapath, Dhaka-1205, Bangladesh
[3] Professor, Department of Geography and Environment, Jahangirnagar University, Dhaka-1342, Bangladesh

**Corresponding Author: shuvosrahman@gmail.com*

Abstract
Accurate Land Use Land Cover (LULC) information can help with numerous research efforts relating to floods, droughts, migration, and climate change at several scales. In decision support systems for sustainable development and natural resource management, machine learning-based methods for detecting changes in land use and land cover using Remote Sensing (RS) data have become indispensable tools. However, Smaller areas like the selected study area may have issues regarding the spatial resolution of satellite imagery. Although the development of Machine Learning (ML) algorithms has boosted the accuracy and use of classification systems, the accuracy level of different classifiers in small areas is often not tested. In order to determine the most accurate classifier for a small-area LULC mapping, the study compares three classifiers in ArcGIS and six classifiers in Google Earth Engine (GEE) against manually classified high-resolution Google Earth data. According to the study, the Support Vector Machine (SVM) classifier outperformed the other one in terms of accuracy for both Sentinel and Landsat satellite imagery. But because of the small number of training samples and spectral similarity between the classes, all six classifiers—Classification and Regression Tree (CART), Random Forest (RF), Support Vector Machine (SVM), Minimum Distance (MD), Naive Bayes (NB), and Maximum Likelihood Classifier (MLC)—had trouble correctly differentiating between bare land and built-up areas. The study suggests taking adequate samples for classification, using higher resolution imageries, reinforcement-based classification or indices for specific feature-oriented research. The study also urges a high-resolution earth resource satellite in Bangladesh to monitor the fragile environment and disaster situations better.

Keywords: LULC, Classifiers, Google Earth Engine, Satellite Imageries, Machine Learning, Wetland Monitoring

1 Introduction

Accurate land use land cover (LULC) information can help with numerous research efforts relating to floods, droughts, migration, and climate change at several scales. LULC changes (deforestation and urbanisation) significantly contribute to global climate change. The continuous and exact examination of LULC is essential to any region's sustainability and growth. For various decision support systems for natural resource management and sustainable development, land use and land cover change detection, based on remote sensing (RS) data, has been a vital tool (Basheer et al. 2022).

According to Palmer et al. (2014) and Tockner et al. (2009), riverine landscapes are dynamic interfaces between land and water that are essential for biodiversity, environmental services, and human well-being worldwide. Because of their functions in flood reduction, nutrient cycling, and providing habitat for aquatic species, oxbow lakes-which are naturally generated by rivers meandering through them-are especially important (Naiman et al., 2006). Due to their susceptibility to human activity, the effects of climate change, and hydrological changes, riverine habitats require an understanding of the dynamics of Land Use, Land Cover, or LULC (Kayitesi et al., 2022) The ecosystem of floodplains, sediment dynamics, and water quality can all be impacted by changes in LULC, which can have an impact on human communities as well as natural systems that depend on these resources (Junk et al., 2015; Vörösmarty et al., 2010). For efficient environmental management, conservation planning, and resilience-building against natural hazards, accurate mapping and monitoring of LULC in riverine landscapes are essential (Allan et. al, 2013). This research addresses potential and problems in environmental monitoring within these biologically significant ecosystems by evaluating several classification techniques for mapping land use and land cover (LULC) within oxbow lakes and their surrounding areas.

To monitor LULC change using remote sensing data, different studies used different tools and techniques, and even the study areas and the temporal extent of the articles sometimes varied. There are both short- and long-term analyses. Study areas vary from coastal wetlands to watersheds. Some papers monitored the change of meander also (Al-Husban 2018). All the research reviewed used satellite imageries like LISS, Landsat, Cartosat, SRTM, and KOMPSAT. However, the Landsat series are the most selected satellite for its long-term continuity and global coverage.

Furthermore, the selection of satellites also varied because of the spatial extent – as the Landsat series has low resolution, whereas KOMPSAT – 2 has a relatively higher resolution (Sunwoo, Nguyen, and Choi 2018). Google Earth Imageries mainly produce ground control points or GCP. However, Tyagi and Sahoo (2022) used historical imagery of google earth to study smaller areas. Zhu and Li (2013) also use google earth high-resolution imageries. However, Smaller areas like the selected study area may have issues regarding the spatial resolution of satellite imagery. Such as Landsat is the most famous earth resource satellite in terms of LULC change detection. This satellite has a spatial resolution of 30m, which may give misleading information regarding small-scale research. Medium-resolution imageries provided by the Sentinel-2 satellite may solve the issue with a slightly higher resolution (10m). However, the temporal resolution of that satellite is low. On the other hand, Google Earth Imageries have a high resolution and a more extended temporal resolution than Sentinel imageries.

Furthermore, methods and approaches for collecting correct data about LULC from remote sensing photos for LULC categorisation are incredibly customisable. The accuracy of classification approaches can be influenced by various parameters, including the choice of initial tests, the heterogeneity of the research region, the sensors employed, and the number of classes to be defined. Classifiers are classified according to the approach and technology used, such as supervised and unsupervised classification, border and non-border classification, hard and soft (ambiguous) classification, or classifications based on Pixel and sub. Traditional visual interpretation and statistical statistics are no longer accurate for LULC categorisation. Training samples, classifiers, and different datasets are the three primary components of supervised LULC classification (Atef, Ahmed, and Abdel-maguid 2023). Using remote sensing datasets, distinct LULC categories may be analysed using machine learning algorithms on a geographic information system platform (e.g., ArcGIS Pro) or Google Earth Engine. The geographic information system can help with geographical and temporal analyses linked to LULC categorisation. These analyses can help us comprehend the LULC pattern of change. Google Earth Engine is a platform that integrates remote sensing data (i.e., satellite images from various sources) with high-performance computer services, allowing for faster and easier satellite imagery processing (Basheer et al., 2022).

The development of machine learning (ML) algorithms has boosted the accuracy and use of classification systems. Several ML methods are used for LULC classification, but random forest (RF) and support vector machine (SVM) approaches have emerged as notable and frequently utilised in recent years. Compared to standard approaches, these methods outperform them, particularly in distant sensing applications(Sridhar, Ali, and Sample 2021). These methods, however, do not provide the same results when using different satellite data. It is critical to compare applications to assess these algorithms' performance using the amount of training and test points (Avci et al., 2023). The quality, quantity, and distribution of the training and test data set all significantly impact the success of the chosen method.

The study followed two objectives: to classify the satellite images using different classifiers of ArcGIS and Google Earth Engine and to compare the LULC results with the LULC classification done using high-resolution Google Earth Imageries. Different studies tried to compare these classifiers between the google earth engine (Atef et al., 2023, Avci et al., 2023), and even a cross-platform (GEE and ArcGIS) comparison also has been conducted (Basheer et al., 2022). However, studying areas to a smaller extent may need a better classification technique. In light of this, the study aims to compare the classifiers available in Google Earth Engine and ArcGIS with traditionally classified higher resolution google earth imageries to identify the most accurate classifier for a smaller area.

2 Materials and Methods

2.1 Study Area

An alluvial river Oxbow is a meander loop that has been abandoned. It is the most prevalent variety of fluvial-derived lakes. It emerges from the interaction of channel erosion and accumulation separated in space as a logical result of the lateral shifting and the overall downstream migration of meanders. The selected study area is an oxbow lake spreading between two unions of the Manikganj district of Bangladesh situated between 90° 0' 0" East longitude to 90° 4' 0" East longitude and 23° 46' 0" North Latitude to 23° 52' 0" North Latitude (Fig. 1).

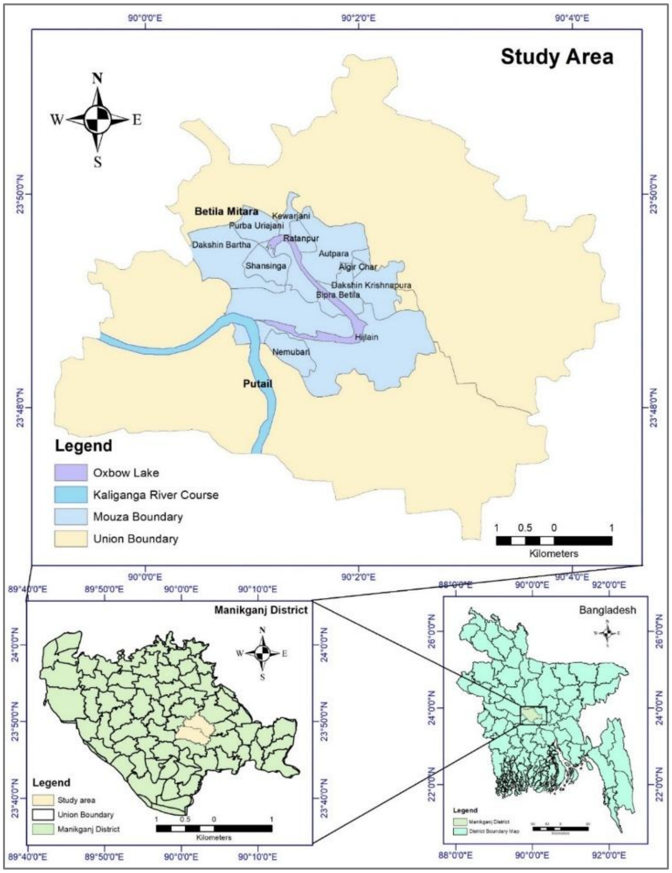

Figure 1: Study Area Map

The present border of Oxbow Lake is 8.50 Kilometres, while the broad area of Oxbow Lake is 0.45 square kilometres or 45.29 Hectares, according to Google Earth Pro satellite imagery. The study area is a changing channel of the Kaliganga River, which runs through the Putail, Nabagram, and Betila - Mitora Unions of Bangladesh's Manikganj district. The region includes twelve Mouzas from these unions: one from Nabagram, seven from Betila-Mitara Union, and four from Putail Union. The land was formed chiefly by sediments transported by the fertile Kaliganga River, and the residents of this area are predominantly involved in agricultural activity. The Kaliganga, a 78-kilometre-long tributary of the Jamuna River with a width of 242 meters, was a meandering river that frequently changed the channel (Razzak 2015). The river took a U-shaped curve in the chosen place, and lately, through erosion, the river followed a new shorter course, and the former channel became primarily a perineal lake, known as an oxbow formation. The study area faced a significant change of river shifting in the recent decades which makes the place a subject of study.

2.2 Satellite Imageries

For a better comparison, the study used low-resolution Landsat 9 (30m) imageries, a medium resolution Sentinel 2 (10m) imageries and a high-resolution image (3 cm) from Google Earth. Where Landsat 9 is the latest sensor of the Landsat program, operating since 1974, the USGS National Land Imaging (NLI) Program has assured data continuity, dependability, and comparability. The spatial resolution of Landsat is 30 m (43 of 1). Sentinel 2 is a multispectral imaging system with a high resolution (10 m) and a large sweep. Sentinel 2 is equipped with thirteen spectral bands (44 of 1). Which helps Copernicus land monitoring research, such as assessing water cover and water bodies, vegetation, soil, and coastal and inland water cover observations.

Google Earth Pro is a geospatial application that offers a variety of tools and capabilities for exploring and studying the Earth's surface. Google Earth Pro's enormous collection of historical satellite images is one of its standout features. These historical images, acquired by several satellite missions, provide information on previous land cover changes, urban growth, and environmental dynamics. The availability of historical satellite images in Google Earth Pro, according to Jaeger (2024), enhances the investigation of long-term landscape changes and aids in evaluating ecosystem dynamics. Furthermore, Goodchild and Janelle (2004) found that Google Earth Pro's historical imagery is helpful for urban planning and disaster management applications. The selected imageries and their date of capture has been described in table 1:

Table 1: Data used in the study

Satellite Name	Spectral Resolution	Date of Capture	Source
Landsat-9 Operational Land Imager 2 (OLI-2)	30m	10th January, 2023	Google Earth Engine Data catalogue and USGS Earth Explorer
Sentinel-2A and 2B – Copernicus Sentinel-2 MSI (Multispectral Instrument)	10m	11th January, 2023	Google Earth Engine Data catalogue and USGS Earth Explorer
Google Earth High-Resolution Imagery	~3cm	10th December, 2022	Google Earth Pro

2.3 Methodological Framework

Table 2 shows the names and description schemes of LULC classes used in the study. By satellite image interpretation and observation, we have selected Waterbodies, Agricultural Land, Built-up Areas, Barelands, and Homestead Vegetation, which are the main classes of the study area. An overall flowchart of the total procedure has been illustrated in Figure 2, modified after Basheer et al., 2022. The study used a confusion matrix as the method for accuracy assessment.

Table 2: Description of LULC classes

LULC Class	*Colour*	*Description of Class*
Waterbodies		Rivers, ponds, lakes
Built-up Areas		Settlements with no or very few trees
Barelands		Point bars, brickfields, lands for developments
Homestead Vegetations		Settlements covered with residential forest patches
Agricultural Lands		Croplands, shrublands, grass

Cloud-free images from January 2023 have been selected for the study. As January is a dry month, the monsoon dynamics will not be applicable here. Maps have been produced from the study results. To avoid sampling bias, the samples for supervised classification remained the same during every analysis in the Google Earth Engine in terms of both images.

2.4 Classification Methods

A collection of training samples specific to the image was produced to perform pixel-based supervised classification. Based on extra data such as Google Street View and orthophoto pictures, each training sample pixel was assigned to a LULC class (Gordon et al. 1984). In GEE, 30 points and ten polygons were utilised as training samples and 40% were used as validation points and selected randomly. The majority of the classifiers encountered default settings. In terms of the analysis on ArcGIS, 125 polygons (60%) have been used as training samples, and 84 polygons (40%) have been used for validation points with help from google earth imageries and ESRI premium base maps. Classifiers used in the study follows:

2.4.1 Classification and Regression Tree (CART)

CART is a decision classification tree that simplifies decision-making and regression analysis (Shetty 2019). CART operates by separating nodes based on a predetermined threshold until it reaches the terminal nodes. This approach separates input data into additional group sets, and trees are constructed except for one of those group sets. The max node and minimum leaf number have been set up as 10 and 5, respectively. The "classifier.smile.Cart" technique from the Google Earth Engine library was employed in this project.

2.4.2 Random Forest (RF) Classifier

Random Forest Classifier, also known as Random Tree Classifier, is a popular method which mixes various CART and constructs an ensemble classifier. RT/RF classifiers generate many decision trees using a randomly nominated training dataset. The tree count is a necessary input criterion for Random Forest classification (Xie et al., 2019, Pal, 2005). In this study, the tree count is set up at 50, which is the ideal value. In ArcGIS, the maximum number of trees is 50, the Max tree depth is 30, and the Max sample per class is 1000 for the random forest classifier.

2.4.3 Support Vector Machine (SVM) Classifier

SVM is a supervised learning approach to address various regression and classification issues. During the training phase, SVM classifiers construct an ideal hyperplane that divides input datasets into various classes based on the least misclassified pixels. When determining support vectors, kernel functions, cost parameters, and gamma are essential (Talukdar et al. 2020). In this study, the SVM classifier inputs have been kept in the default settings.

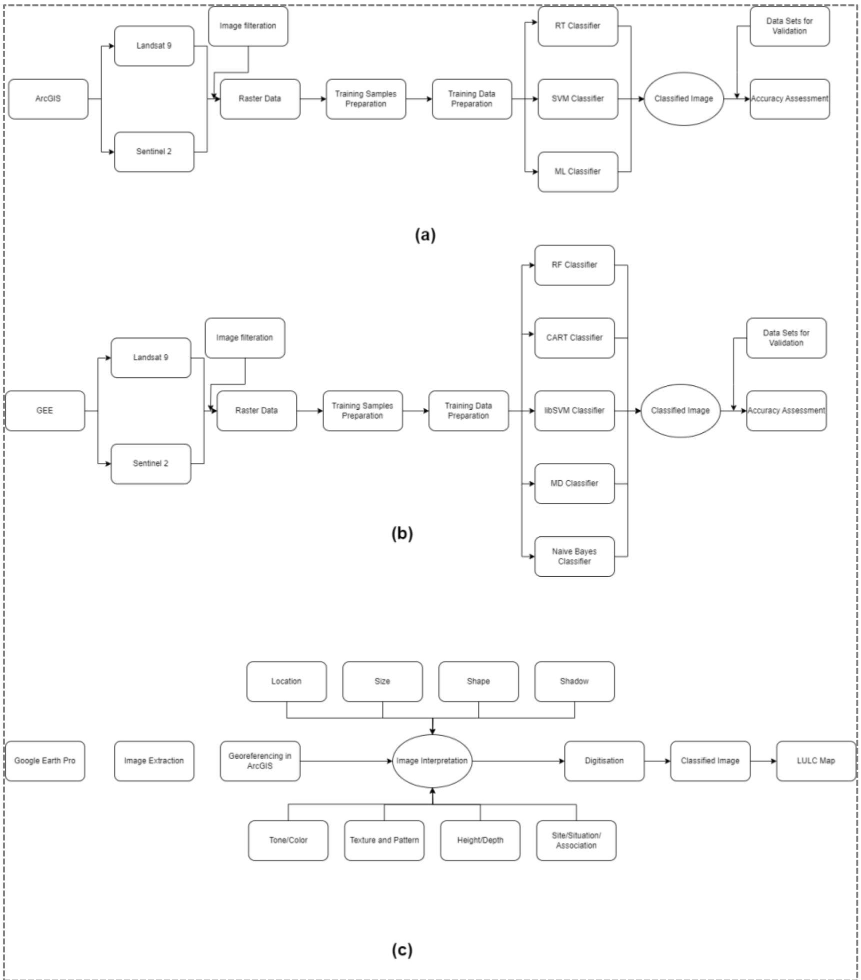

Figure 2: LULC classification using (a) ArcGIS and (b) Google Earth Engine (c) Google Earth Pro Areal Imageries

2.4.4 Minimum Distance (MD) Classifier

Minimum distance classification methods are explored in the context of remote sensing, particularly for tasks like crop species identification. These methods categorize groups of measurement vectors based on their resemblance to known or estimated class distributions. The approach involves comparing distributions using distance measures in distribution function spaces. Literature reviews categorize minimum distance classification problems by assumptions about class distributions. Experimental results indicate that minimum distance classifiers generally outperform maximum likelihood classifiers by 5% to 10% in sample classification accuracy. Parametric and nonparametric versions show minimal performance differences, with nonparametric methods offering slight improvements overshadowed by increased complexity and slower speed (Wacker and Landgrebe, 1972).

2.4.5 Naive Bayes (NB) Classifier

It is a classification strategy based on Bayes' Theorem and the assumption of predictor independence. A Naive Bayes classifier, in essential words, posits that one feature in a class is independent of the presence of any other feature. The main benefit of this classifier is that it just needs a modest amount of training data to generate the decision surface. At the same time, its development is predicated on the assumption that the underlying distributions are Gaussian, which may limit its applicability (Haykin 2009; Shelestov et al. 2017).

2.4.6 Maximum Likelihood Classifier (MLC)

MLC or A maximum likelihood classifier is a supervised classification approach that uses a normal distribution to represent each band. The Byes theorem underpins this supervised classification approach(Norovsuren et al. 2019). In this study, this classification technique is only used in ArcGIS-based analysis.

2.5 Data Processing

The study used Landsat 9 and Sentinel 2 imageries in GEE. When working on ArcGIS, the study also used the same image from the 10^{th} of January (Landsat 9) and the 11^{th} of January (Sentinel 2). In GEE, 200 samples were used, from which 60% were used in classification, whereas 40% were used for validation. In most of the classifications, the settings were in default mode. In ArcGIS, however, a total of 209 samples has been trained where the same ratio (3:2) has been maintained.

2.6 Accuracy Assessment

The confusion matrix (CM) is critical in the map classification validation procedure. A CM contained in GEE is used to assess the accuracy of the LULC categorisation in the research region. The LULC associated with the validation point is compared statistically to the output categories. The kappa coefficient and the values of the confusion matrix's overall user, producer, and accuracy coefficients were compared for each class to evaluate classification performance (Forghani-Zadeh and Rincón-Mora 2007). RF's confusion matrix was used to evaluate picture classification accuracy (Congalton and Green 2019). To assess the accuracy of images in ArcGIS, a set of 84 polygons have been used as validation points, whereas 80 points and polygons selected randomly have been used as validation points to assess Google Earth Engine's image accuracy.

$$Overall\ accuracy = \frac{Number\ of\ total\ correctly\ classified\ pixels}{Number\ of\ total\ validation\ pixels} \times 100 \ldots\ldots\ldots\ldots\ldots\ldots \text{(1)}$$

$$Overall\ accuracy = \frac{Number\ of\ total\ correctly\ classified\ pixels}{Number\ of\ total\ validation\ pixels} \times 100 \ldots\ldots\ldots\ldots\ldots\ldots \text{(2)}$$

Classification accuracy has been calculated in two different categories. Firstly, an overall accuracy has been calculated using equation 1; secondly, another feature wise accuracy has been calculated using equation 2.

2.7 Google Earth Image Interpretation

The Google Earth Image has been downloaded from the Google Earth Pro Software and then georeferenced using the Georeferencing toolbox of ArcGIS, then the image has been digitised using manual draw tool and aerial photo interpretation techniques (Badamfirooz and Mousazadeh 2019). Figure 2(c) explains the methodological procedure of aerial photo interpretation using google earth pro imageries.

3 Results and Discussion

In every classification map, the agricultural land is the most significant feature, followed by Homestead vegetation and waterbodies. Barelands and Built-up Areas sometimes switch their places for the least identified features. The maps have been divided into five categories according to their data source and analysis software for better visualisations and understanding. Maps produced using ArcGIS and Google Earth Engines are divided into two categories, whereas the map produced using areal imageries of Google Earth Pro is considered another category for its unique methodology. A brief description for each map has been constructed below:

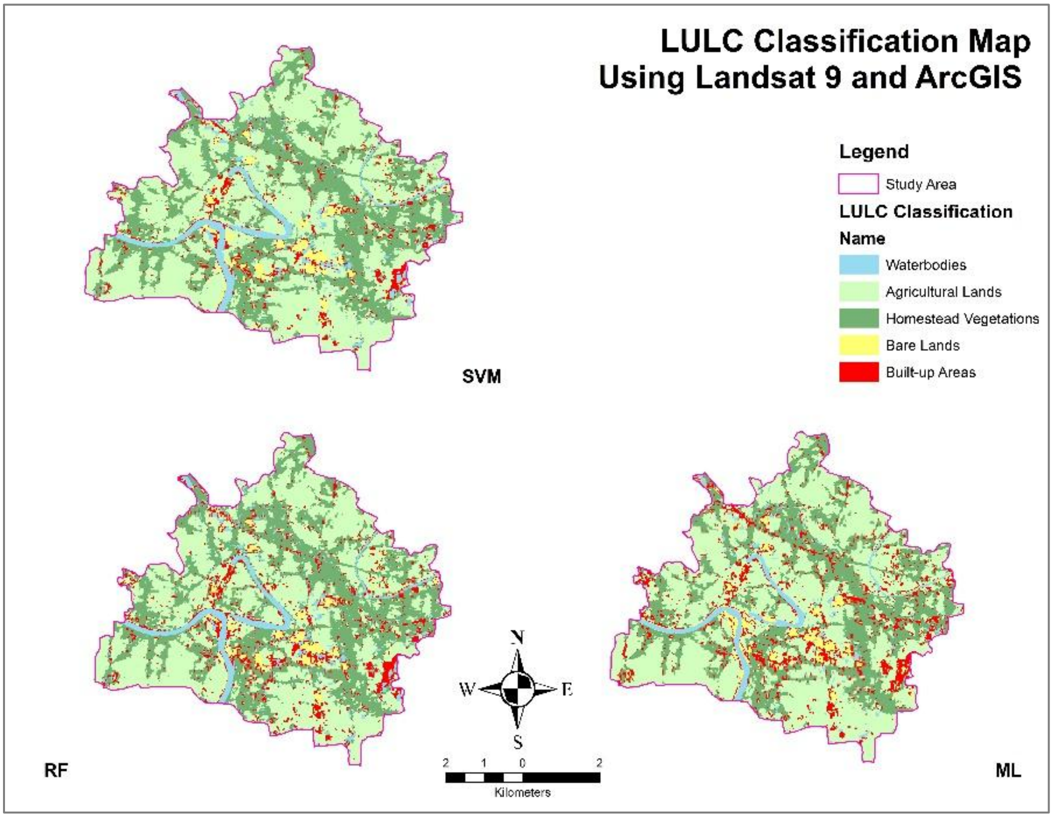

Figure 3: LULC map produced using ArcGIS and Landsat 9 Imageries

ArcGIS's Landsat 9 Map: Figure 3 contains the three maps produced using SVM, RT and MLC classifiers. Agricultural Land (45% - 53%) is the prominent feature, followed by Homestead vegetation (30% - 38%), and Barelands are the lowest prominent features of the maps ranging from 2.5% to 7%. The amount of Agricultural land is significantly low in RT classification, whereas Homestead Vegetation and Built-up Area is also higher than other classifiers. The amount of water bodies is slightly higher in the RT classification, and Agricultural Land is slightly higher in SVM. The SVM classifier identified fewer Built-up areas and bare lands than the other two, whereas the MLC classifier identified the highest amount.

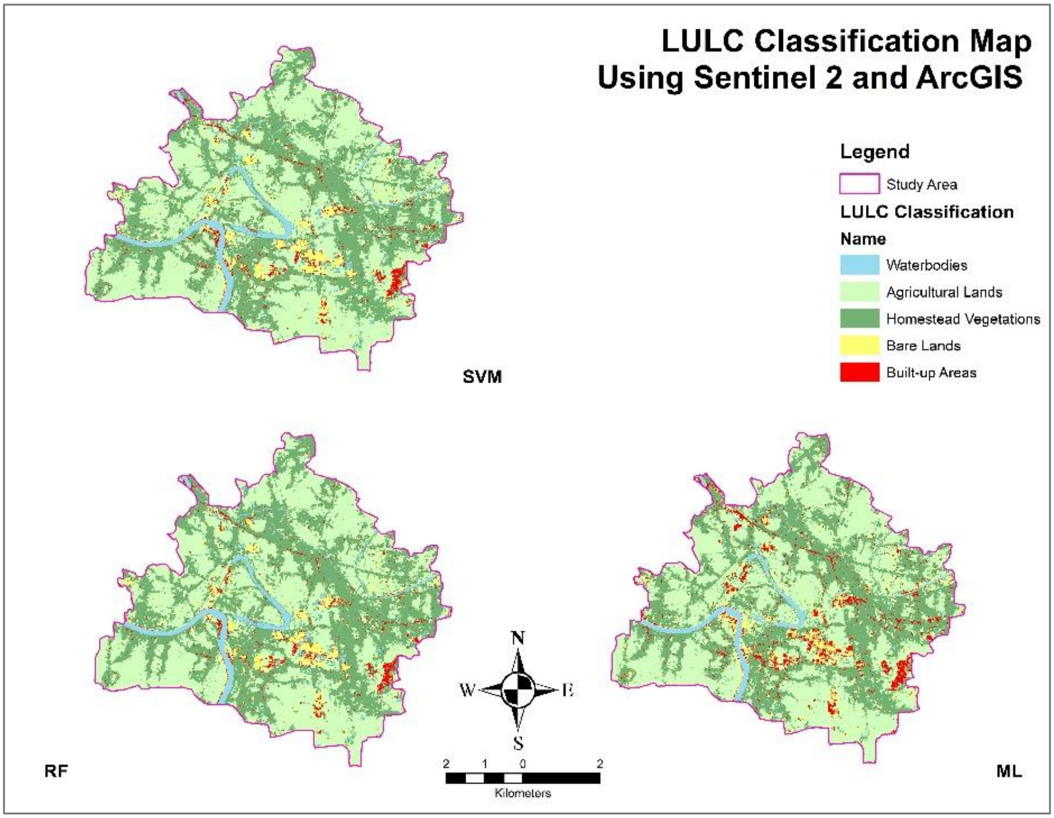

Figure 4: LULC map produced using ArcGIS and Sentinel 2 Imageries

ArcGIS's Sentinel 2 Map: Scenarios are almost identical regarding Sentinel 2 maps produced by ArcGIS. Nevertheless, unlike Landsat maps, the RT classifier identified less agricultural land and more waterbodies than the other classifiers (10 to 20 percent). RT classifier also identified more homestead vegetation, and Built-up areas are significantly low in this classification (more than 20%). However, the MLC classifier identified relatively low waterbodies (10%) and a slightly higher built-up area (5%). SVM classifier identified a higher number of bare lands (5%).

GEE's Landsat 9 Map: In Google Earth Engine, we ran five different algorithms for five different classifiers and got some exciting overviews. With almost 50% of the classes, SVM identified the Agricultural Lands as the most significant feature in the area. Random Forest (RF Method) identified a slightly lower agricultural land. Regarding homestead vegetation, SVM identified 30% of the features, whereas RF and NB classifiers identified almost 39% of the area. Waterbodies vary from 3.4% (RF) to 6.11% (CART), and built-up area varies from 7% (SVM) to almost 10.5% (CART) from classifier to classifier. SVM classifier identified a significant amount of bare land (6.8%), which is higher than any other classifier, which shows 2.20% to 2.74% of the land is bare.

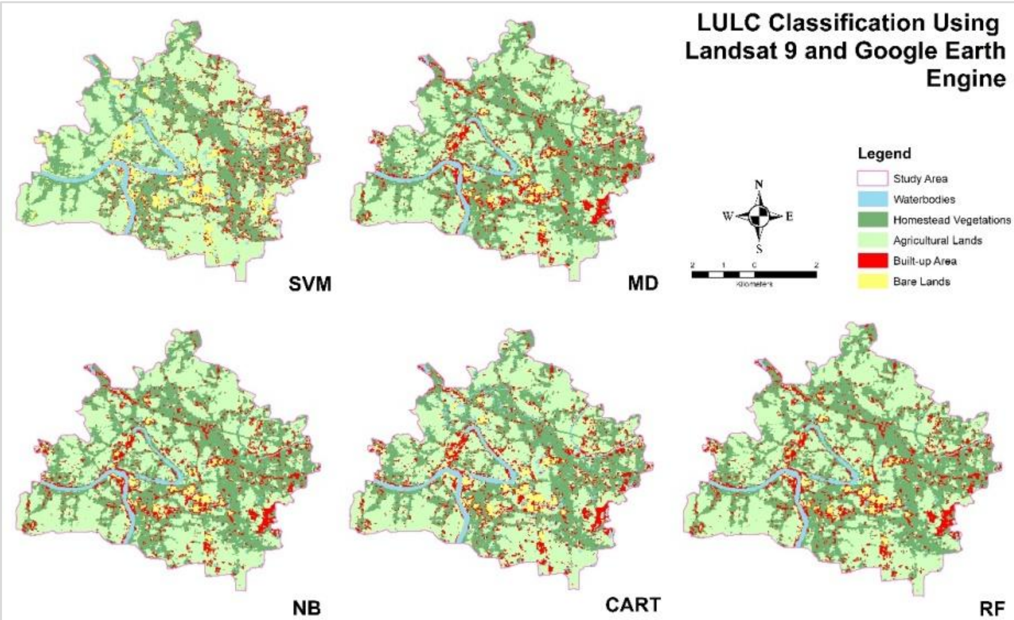

Figure 5: LULC map produced using GEE and Landsat 2 Imageries

GEE's Sentinel 2 Map (Figure 5): 55.7% of the land is identified as agricultural land in the SVM classifier, which is higher than any other classifier, with a lowest of 47.8% (MD and NB). Waterbodies are higher according to the CART classifier (38.5%) and lower in the SVM classifier's classification (34%). MD classifier identified the lowest amount of Bare lands (<2%), and the NB classifier identified the most (>3.5%). Built-up areas are higher in the MD classification (8.75%) and NB classification (8%) than in the other three classifiers' results.

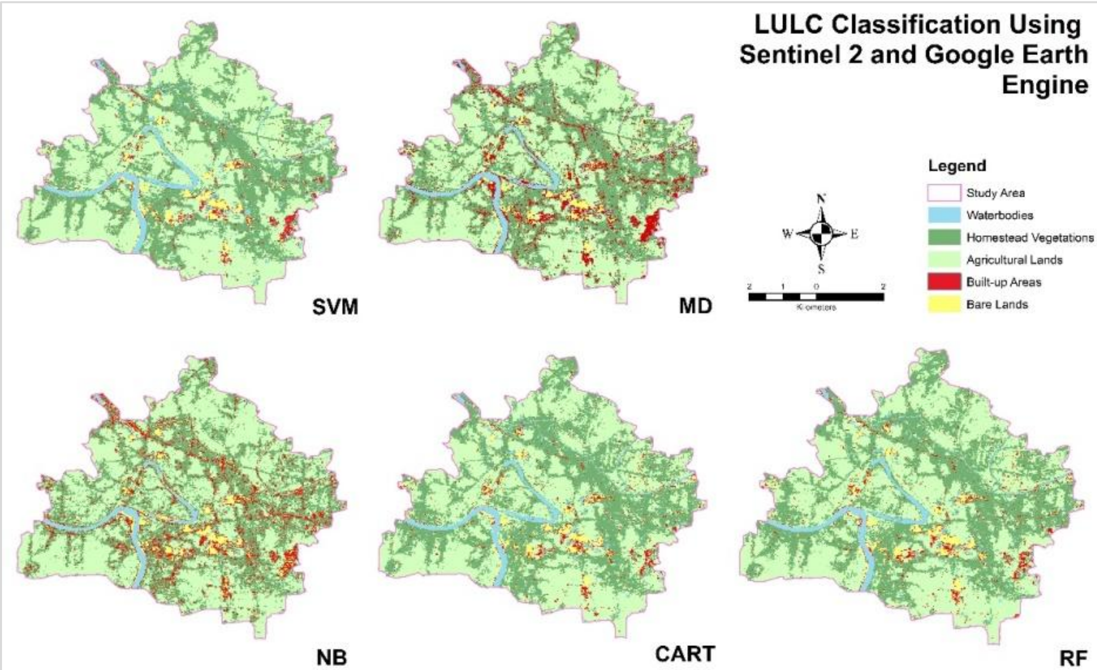

Figure 6: LULC map produced using GEE and Sentinel 2 Imageries

Accuracy Assessment: In this study, we employed a validation dataset that was different from the training datasets to test the accuracy of each classifier. We distributed these locations over the research region using stratified random sampling to ensure all LULC classes were correctly and uniformly represented. Following LULC classification using all classifiers, total accuracy was evaluated using a confusion matrix to assess the correctness of these LULC categorised maps in ArcGIS Pro and Google Earth Engine.

Overall Accuracy: Figure 7 shows the overall accuracy of GEE-based analysis, and Figure 8 shows the overall accuracy of Landsat imageries. Among the five classifiers, SVM showed the highest accuracy in both imageries. However, the RF classifier stands slightly higher for Landsat imageries than the RF classifier and 2% lower for Sentinel Imageries. NB is significantly lower in both cases, and the CART classifier is more than 90% accurate for both imageries.

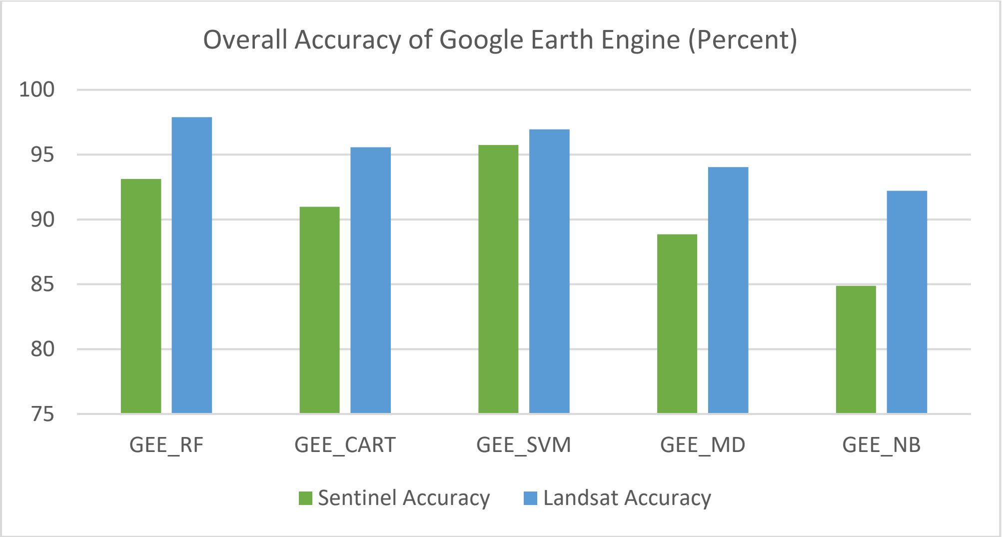

Figure 7: Overall accuracy of Google Earth Engine classifiers in percent.

Among the three classifiers, the result of SVM is more than 90% for both imageries. Although the accuracy result for MLC and SVM is similar for Landsat images, the inaccuracy of MLC is significant in classifying Sentinel images.

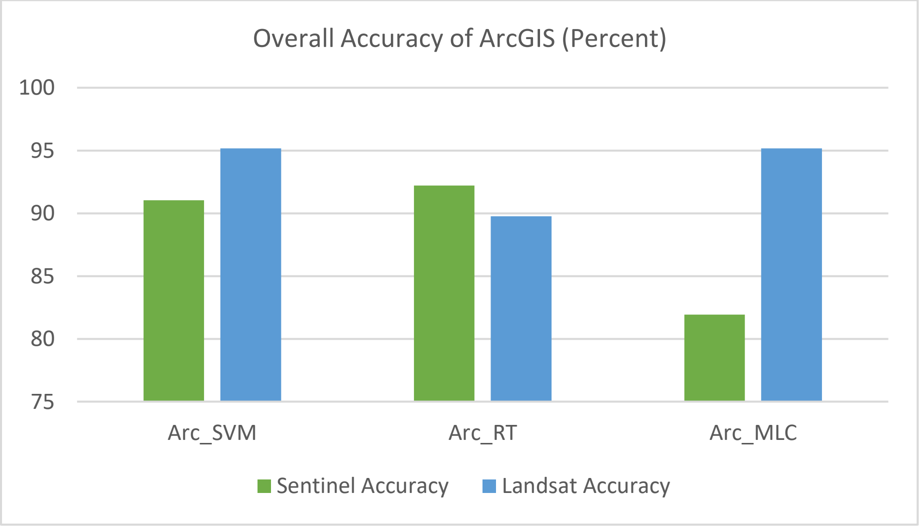

Figure 8: Overall accuracy of ArcGIS classifiers in percent.

Feature-wise accuracy: The best advantage of a confusion matrix as an accuracy assessment method is that it can identify feature-wise accuracy. Figure 9 shows the feature-wise accuracy of classifiers in terms of Sentinel imageries, and Figure 10 shows the accuracy of Landsat imageries.

All classifiers are almost 100% correct in identifying Waterbodies apart from the GEE's NB classifier, which identified almost 1.5% fewer waterbodies than the other classifiers. NB does a better job identifying Barelands along with ArcGIS's RT classifier, which is 100% correct according to the confusion matrix assessment. Only GEE's RF classifier can classify agricultural land with a 98% accuracy level. GEE's SVM classifier is almost as accurate as RF (97%). NB and MD identified agricultural land with an accuracy level lower than 90%. GEE's RF, CART and SVM all showed an accuracy level of more than 95%, the MLC classifier is almost 94.5% accurate, and the rest are less than 90%, with NB in a 67.73% accuracy level. Every classifier struggled in terms of identifying built-up areas. Google Earth Engine's RF, CART, and SVM classifier is less than 35% accurate in identifying built-up areas. ArcGIS's SVM classifier showed 71%, and GEE's MD classifier showed a 65% accuracy in identifying built-up areas. Google Earth Engine's SVM classifier drops accuracy significantly in identifying bare lands, whereas ArcGIS's SVM classifier remains 93.85% accurate in identifying bare lands from Sentinel 2 imageries.

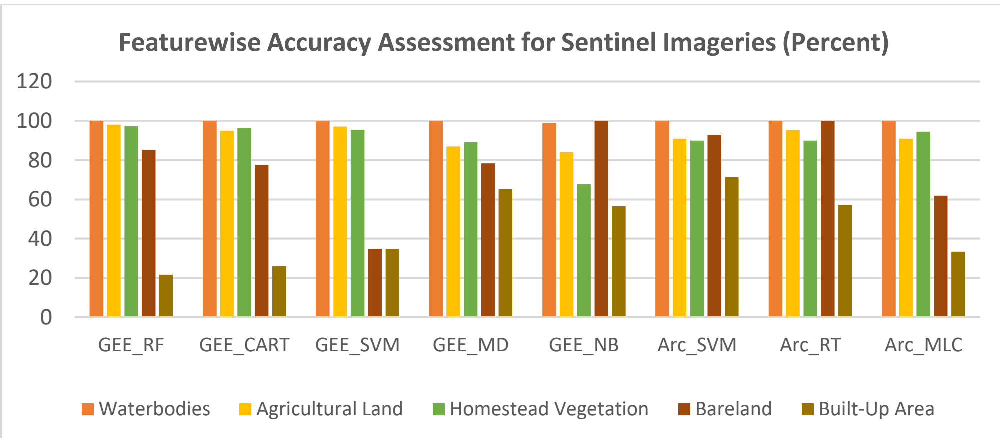

Figure 9: Feature-wise accuracy of classifiers for sentinel imageries.

Regarding identifying waterbodies from Landsat imageries, the scenario is the same as the Sentinel imageries, with GEE's NB and MD dropping 2% accuracy than the other classifiers. Every classifier showed improvements in identifying agricultural lands, with a 2% increase for NB's classification. However, ArcGIS's SVM dropped a significant accuracy (74%) in identifying agricultural land. ArcGIS's RT and MLC classifier identified homestead vegetation, and GEE's SVM almost showed perfect accuracy. However, ArcGIS's SVM classifier drops almost 5% more accuracy than GEE's SVM classifier identifying homestead vegetation. Henceforth, every classifier has more than 95% accuracy in identifying homestead vegetations. Unlike the cases of Sentinel imageries, the classifier has a higher accuracy in terms of lower-resolution Landsat imageries.

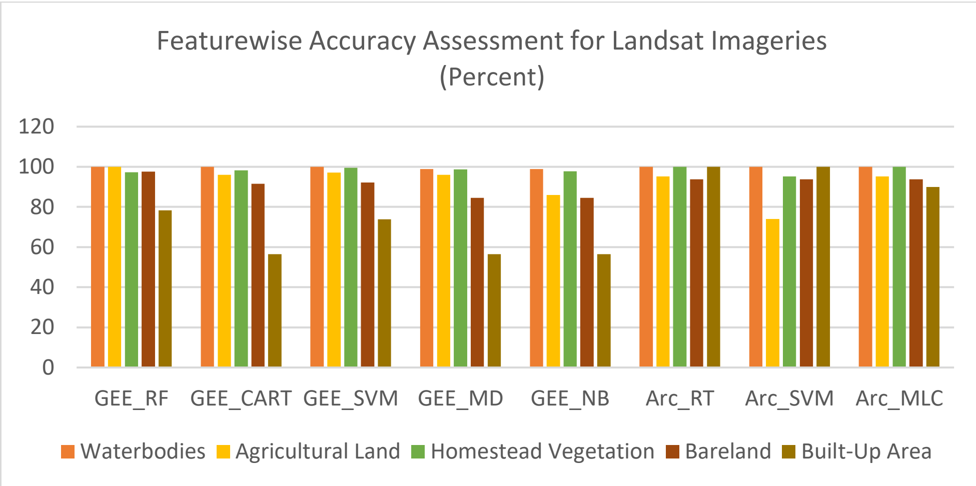

Figure 10: Feature-wise accuracy of classifiers for Landsat Imageries.

Although it varies from feature to feature, the SVM classifier is better positioned to identify waterbodies, agricultural lands, and homestead vegetation. Struggles in identifying bare lands and build-up areas can be caused by inadequate samples and similarities between DN values, as they are both unique reflectors. Henceforth, a higher resolution of the Sentinel image may create higher inaccuracy as it has a higher pixel count.

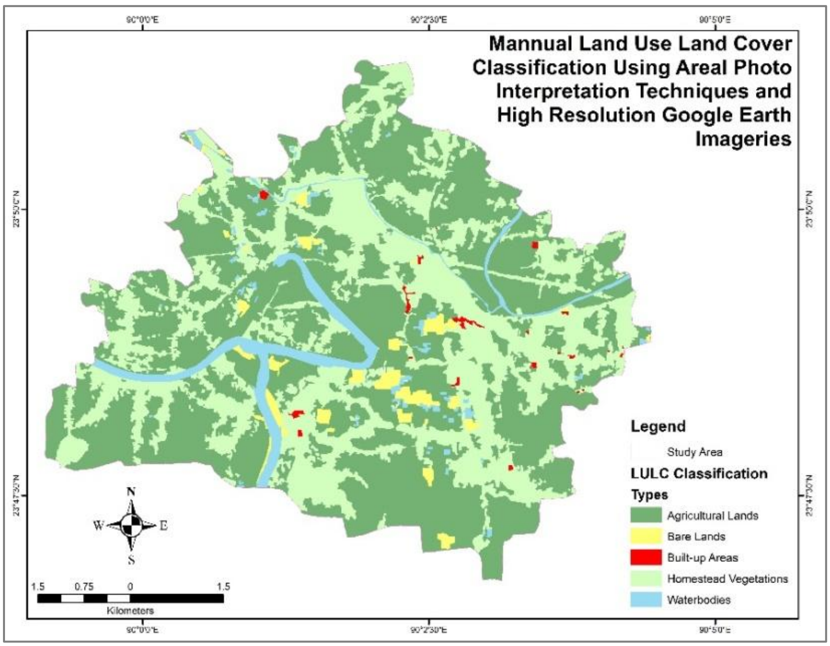

Figure 11: Manual LULC Classification Using High Resolution Images.

Google Earth Pro's LULC Map: Manual areal photo interpretation is a time-consuming Land Use Land Cover classification method. It is also prone to geometrical errors and other human errors. However, this can be an excellent tool for studying small areas like this study. Figure 11 shows a LULC map of the area, illustrated using the Areal Photo Interpretation technique.

In this method, the waterbodies under 30m resolution can be identified. Furthermore, the human eye can detect bodies covered with water hyacinths. Though the overall ratio of various features is almost the same as the ML-based classifications, there are some anomalies in total areas feature-wise.

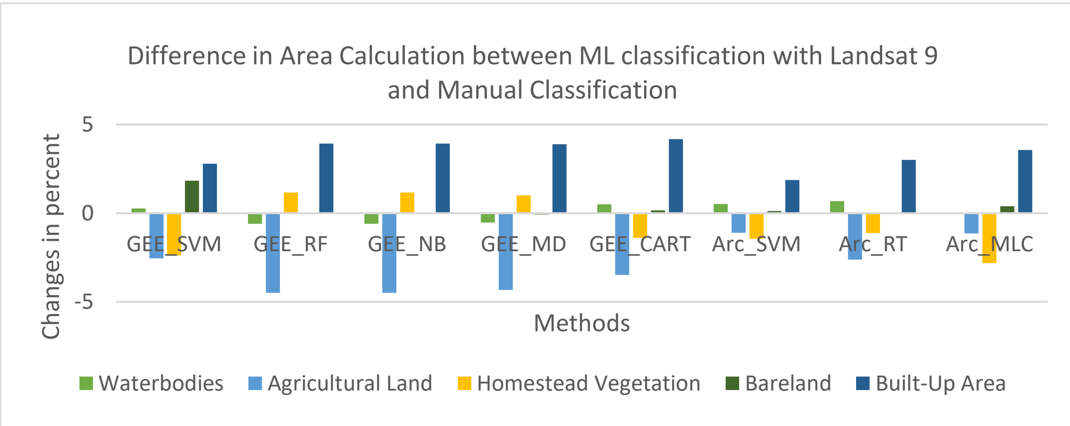

Figure 12: Areal difference between Manual Classification and ML-based Classification of Landsat Imagery.

Figure 12 shows the difference between manual classification and ML-based classifications with Landsat imageries. Three classifiers identified more homestead vegetation than the manual classification, whereas differences in the identification of waterbodies are minimal for every classifier. It is evident from the graph that, ML techniques identified more built-up area, which may cause more inaccuracy, and less agricultural land, which indicated ML classifiers are prone to error differentiating agricultural land and built-up areas.

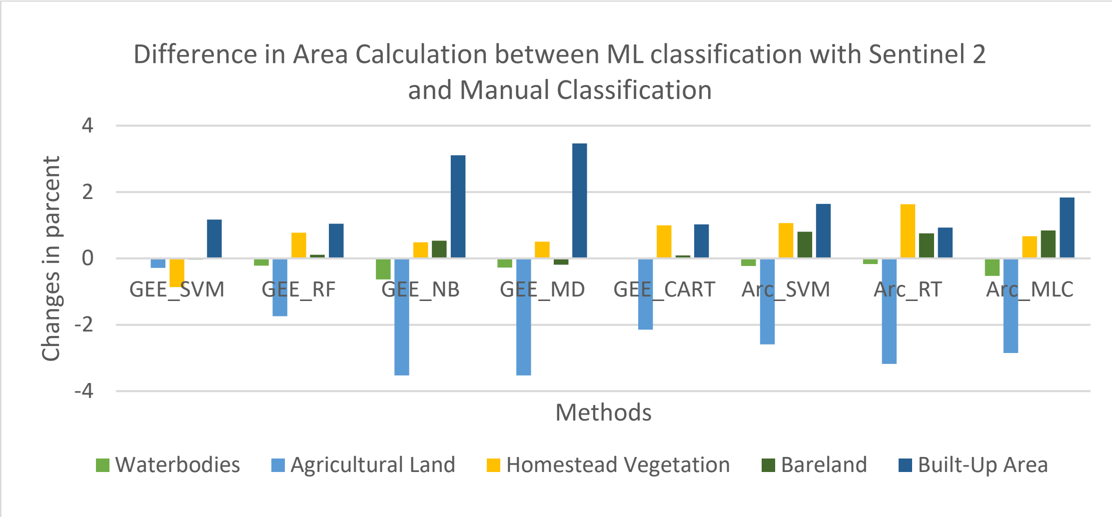

Figure 13: Areal difference between Manual Classification and ML-based Classification of Landsat Imageries.

Although the accuracy of Sentinel imageries is lower than the Landsat imageries, the difference between manual LULC and ML-based classification is lower here as Sentinel image has a higher resolution. Still, the ML classifiers struggled to identify/differentiate the agricultural lands and built-up areas, whereas the Google Earth Engine SVM method provided almost the same results as the human-operated classifications. Despite less identifying significant agricultural land, Google Earth Engine's CART and ArcGIS's SVM also provided a decent result. Nevertheless, a field verification may add significant value in the study, however, due to budget and time constraint, we could not do any field verification.

4 Conclusion

The Landsat satellite series has collected imagery for over fifty years, whereas Sentinel created ten years of continuous monitoring data in a higher resolution. On the other hand, human activities' changing land use patterns triggered climate change since the pre-industrial area (IPCC, 2021; Islam 2019). Free and paid areal photography and the aforementioned machine learning techinques also triggered earth scientist to do more and more research on LULC change detection (Basheer et al., 2022). Henceforth, studying smaller-scale LULC classification is becoming more and more relevant due to the complex environmental scenario. Moreover, the accuracy of classification techniques should always be kept in mind to avoid wrong interpretations which may influence the later discussion. Although Basheer et al. (2022) conducted a study comparing classifiers, the purpose of this study was to focus on a small-scale study area and conduct a comparative study for Bangladesh, which has unique geography (Sanyal and Shuvo 2020; Islam and Neelim 2010) and machine-based classification may give a wrong result if the samples are not carefully trained (because of the similarities between spectral resolution of homestead vegetations, forest area and even water hyacinths over waterbodies).

A comparison of classifiers revealed different aspects of LULC classification. In ArcGIS, SVM and MLC, both classifiers showed higher accuracy in terms of classifying Landsat Imageries. However, MLC struggled to classify Sentinel imageries, whereas RT showed a better result for classifying Sentinel Imageries. Random Tree (RT) classifier also showed a better accuracy feature-wise for Landsat imageries, whereas SVM struggled to identify Agricultural Land, and MLC slightly struggled to classify Built-up Areas. Overall, classifiers struggle to identify bare lands and built-up areas. A low sample count may trigger this, and most of the built-up in the study area has some amount of vegetation with them and, finally, due to similar reflectivity of bare lands and built-up area. So, an indices-based analysis may be better if someone wants to analyse the area with more feature-oriented objectives.

The scenario is slightly different with sentinel imageries. The difference between these two satellite's resolutions may cause the difference. As Sentinel has 3x more resolution, the accuracy rate can be reduced because of inadequate samples, so it is better to train more samples for Sentinel Imageries to get higher accuracy. The calculation of Geometry also varies because of the spatial resolution of satellites. Google Earth imageries can significantly help in this regard; however, the irregularity of coverage may create a problem regarding time series analysis. Finally, a new method based on reinforcement classification is also suggested to reduce the errors of supervised learning; reinforcement-based learning can be a great tool to classify remotely sensed images.

The study did all of the classifications in default settings. However, the best settings for accuracy in a small-scale study area or Bangladesh is still unexplored and may be a subject of further study. Finally, the study also urges arranging necessary policy lobbying for a high-resolution Earth resource satellite in Bangladesh to monitor the changing dynamics of land use, land cover, and other environmental phenomena.

5 References

Al-Husban, A. K. (2018). Meandering and Land Use /Cover Change Detection in the Lower Jordan River (LJR), 1984-2016, using GIS and RS. *Environmental Research Engineering and Management*, *74*(2). https://doi.org/10.5755/j01.erem.74.2.20917

Allan, J. D., McIntyre, P. B., Smith, S. D. P., Halpern, B. S., Boyer, G. L., Buchsbaum, A., Burton, G. A., Campbell, L. M., Chadderton, W. L., Ciborowski, J. J. H., Doran, P. J., Eder, T., Infante, D. M., Johnson, L. B., Joseph, C. A., Marino, A. L., Prusevich, A., Read, J. G., Rose, J. B., . . . Steinman, A. D. (2012). Joint analysis of stressors and ecosystem services to enhance restoration effectiveness. *Proceedings of the National Academy of Sciences*, *110*(1), 372–377. https://doi.org/10.1073/pnas.1213841110

Atef, I., Ahmed, W., & Abdel-Maguid, R. H. (2023). Modelling of land use land cover changes using machine learning and GIS techniques: a case study in El-Fayoum Governorate, Egypt. *Environmental Monitoring and Assessment*, *195*(6). https://doi.org/10.1007/s10661-023-11224-7

Avci, C., Budak, M., Yağmur, N., & Balçik, F. (2021). Comparison between random forest and support vector machine algorithms for LULC classification. *International Journal of Engineering and Geosciences*, *8*(1), 1–10. https://doi.org/10.26833/ijeg.987605

Badamfirooz, J., & Mousazadeh, R. (2019). Quantitative assessment of land use/land cover changes on the value of ecosystem services in the coastal landscape of Anzali International Wetland. *Environmental Monitoring and Assessment*, *191*(11). https://doi.org/10.1007/s10661-019-7802-8

Basheer, S., Wang, X., Farooque, A. A., Nawaz, R. A., Liu, K., Adekanmbi, T., & Liu, S. (2022). Comparison of land use land cover classifiers using different satellite imagery and machine learning techniques. *Remote Sensing*, *14*(19), 4978. https://doi.org/10.3390/rs14194978

Congalton, R. (1998). *Assessing the accuracy of Remotely sensed Data: Principles and Practices, second edition*. https://openlibrary.org/books/OL9451454M/Assessing_the_Accuracy_of_Remotely_Sensed_Data

Forghani-Zadeh, H. P., & Rincon-Mora, G. A. (2007). An accurate, continuous, and lossless Self-Learning CMOS Current-Sensing scheme for Inductor-Based DC-DC converters. *IEEE Journal of Solid-State Circuits*, *42*(3), 665–679. https://doi.org/10.1109/jssc.2006.891721

Haykin, S. S. (2010). *Neural networks and learning machines*. http://ci.nii.ac.jp/ncid/BB00465945

Intergovernmental Panel on Climate Change [IPCC]. (2021). Climate Change 2021—The Physical Science Basis. In *IPCC* (Vol. 43, Issue 4, pp. 22–23). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. https://doi.org/10.1515/ci-2021-0407

Islam, S. T. (2019). Climate Change Crisis in Bangladesh. In *Geography in Bangladesh* (1st ed., pp. 146–163). Routledge India.

Islam, T., & Neelim, A. (2010). Climate change in Bangladesh : a closer look into temperature and rainfall data. In *University Press eBooks*. http://ci.nii.ac.jp/ncid/BB05957448

Jaeger, A. J. (2024). Google Earth as a tool for supporting geospatial thinking. *Land*, *13*(12), 2218. https://doi.org/10.3390/land13122218

Junk, W. J., Wittmann, F., Schöngart, J., & Piedade, M. T. F. (2015). A classification of the major habitats of Amazonian black-water river floodplains and a comparison with their white-water counterparts. *Wetlands Ecology and Management*, *23*(4), 677–693. https://doi.org/10.1007/s11273-015-9412-8

Kayitesi, N. M., Guzha, A. C., & Mariethoz, G. (2022). Impacts of land use land cover change and climate change on river hydro-morphology- a review of research studies in tropical regions. *Journal of Hydrology*, *615*, 128702. https://doi.org/10.1016/j.jhydrol.2022.128702

Naiman, R. J., Décamps, H., & McClain, M. E. (2006). Riparia: ecology, conservation, and management of streamside communities. *Choice Reviews Online*, *43*(10), 43–5880. https://doi.org/10.5860/choice.43-5880

Norovsuren, B., Tseveen, B., Batomunkuev, V., Renchin, T., Natsagdorj, E., Yangiv, A., & Mart, Z. (2019). Land cover classification using maximum likelihood method (2000 and 2019) at Khandgait valley in Mongolia. *IOP Conference Series Earth and Environmental Science*, *381*(1), 012054. https://doi.org/10.1088/1755-1315/381/1/012054

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, *26*(1), 217–222. https://doi.org/10.1080/01431160412331269698

Palmer, M. A., Hondula, K. L., & Koch, B. J. (2014). Ecological restoration of streams and rivers: shifting strategies and shifting goals. *Annual Review of Ecology Evolution and Systematics*, *45*(1), 247–269. https://doi.org/10.1146/annurev-ecolsys-120213-091935

Razzak, M. M. (2015). *Bangladesher nod nodi: Bortoman gatiprokriti.* (1st ed.). Kothaprokash.

Sanyal, N., & Shuvo, S. R. S. (2020, October 29). *Coastal hazards in Asian countries in the context of climate change*. Sanyal | International Energy Journal. http://rericjournal.ait.ac.th/index.php/reric/article/view/2531

Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., & Skakun, S. (2017). Exploring Google Earth Engine platform for big data processing: Classification of Multi-Temporal Satellite Imagery for crop mapping. *Frontiers in Earth Science*, *5*. https://doi.org/10.3389/feart.2017.00017

Shetty, S. (2019). *Analysis of machine learning classifiers for LULC classification on Google Earth Engine*. https://essay.utwente.nl/83543/

Sridhar, V., Ali, S. A., & Sample, D. J. (2021). Systems analysis of coupled natural and human processes in the Mekong River Basin. *Hydrology*, *8*(3), 140. https://doi.org/10.3390/hydrology8030140

Sunwoo, W., Nguyen, H. H., & Choi, M. (2018). Coastal wetland change detection using high spatial resolution KOMPSAT-2 imagery. *Terrestrial Atmospheric and Oceanic Sciences*, *29*(5), 509–521. https://doi.org/10.3319/tao.2018.05.18.01

Talukdar, S., Singha, P., Mahato, S., Shahfahad, N., Pal, S., Liou, Y., & Rahman, A. (2020). Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A review. *Remote Sensing*, *12*(7), 1135. https://doi.org/10.3390/rs12071135

Tockner, K., Uehlinger, U., & Robinson, C. T. (2009). *Rivers of Europe*. http://ci.nii.ac.jp/ncid/BA9001633X

Tyagi, N., & Sahoo, S. (2022). Assessing the status of changing regimes of water bodies in Gorakhpur District, Uttar Pradesh, India. *Environmental Monitoring and Assessment*, *194*(2). https://doi.org/10.1007/s10661-021-09630-w

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R., & Davies, P. M. (2010). Global threats to human water security and river biodiversity. *Nature*, *467*(7315), 555–561. https://doi.org/10.1038/nature09440