Comparative Assessment of Classification Algorithms for Land Cover Mapping Using Multispectral and PCA Images of Landsat

Zohaib[1], Nawai Habib[1], Abu Talha Manzoor[1], Sawaid Abbas[1,2*], Samawia Rizwan[1]
[1]Smart Sensing for Climate and Development, Centre for Geographic Information System, University of the Punjab, Lahore, Pakistan.
[2]Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong
***Correspondence:** S.A: sawaid.gis@pu.edu.pk

Citation | Zohaib, Habib. N, Manzoor. A. T, Abbas. S, Rizwan. S, "Comparative Assessment of Classification Algorithms for Land Cover Mapping Using Multispectral and PCA Images of Landsat", IJIST, Special Issue pp 225-239, June 2024
Received | May 31, 2024 **Revised** | June 06, 2024 **Accepted** | June 10, 2024 **Published** | June 14, 2024.

The advancement of remote sensing technologies and the availability of free satellite data have significantly enhanced the precision of land use and land cover (LULC) mapping, facilitating the analysis of landscape transformations and ecosystem changes. However, selecting the most suitable classifier for LULC mapping remains a complex challenge. Therefore, it is essential to evaluate the accuracy of various LULC modeling algorithms to determine their effectiveness in different applications. This study conducted a comprehensive evaluation of both supervised machine learning algorithms and traditional classification methods applied to Landsat 8 imagery with a 30-meter spatial resolution, covering the Shangla and Battagram districts in Khyber Pakhtunkhwa (KPK), Pakistan. The study focused on three classification algorithms: Maximum Likelihood Classification (MLC), Support Vector Machines (SVM), and Random Forest (RF). The performance of these algorithms was assessed on both multispectral images and composite images derived from Principal Component Analysis (PCA) and Band Ratioing and/or Normalized Indices. Additionally, the accuracy of these algorithms, when applied to different datasets, was compared with the recently released World Cover LULC product by the European Space Agency (ESA). The results indicated that the SVM algorithm outperformed the others, achieving an overall accuracy of 90.43% and a kappa coefficient of 0.8792. The MLC and RF algorithms also produced promising results, with overall accuracies of 85.58% and 88.46%, respectively. Furthermore, the study found that the overall accuracy of ESA's World Cover LULC product was 70.67% in the study area, based on similar validation samples. These findings underscore the strengths and limitations of each algorithm, providing valuable insights into their suitability for LULC classification and the applicability of existing global LULC maps.

Keywords: MLC; SVM; RF; PCA, Normalized Indices; Landsat-8, ESA, LULC.

Introduction:

Environmental monitoring techniques are crucial for understanding the Earth's changing climate, including the rise in surface temperatures and the need to track its effects on the environment [1]. Land use and land cover (LULC) data are particularly important in this context. Accurate LULC data support research efforts on floods, droughts, and migration across various scales [2][3]. The demand for and availability of LULC maps have recently increased, driven in part by the expansion of publicly available satellite imagery [4]. Assessing LULC is essential for the sustainability and development of any region, as LULC changes—such as deforestation and urbanization—are major drivers of climate change globally [3]. LULC change is also a critical factor in environmental degradation caused by human activities and natural disasters, making it a significant global concern [5]. Moreover, climate change affects water balance, geomorphology, water quality, groundwater management, resource management, and land monitoring, all of which require detailed LULC maps.

In developing countries like Pakistan, rapid population growth has intensified LULC changes, particularly in urban areas. Urban sprawl often encroaches on agricultural land, raising concerns about food security. LULC changes have far-reaching implications for both natural and human-induced extreme events. However, most previous research has focused primarily on evaluating the accuracy of LULC classification using geographic data, rather than on selecting appropriate LULC change detection strategies [6][7]. To manage agriculture, natural resources, and energy effectively in the face of population growth, it is crucial to choose the proper technique for monitoring LULC changes, which requires accurate LULC maps.

Recently, machine learning (ML) algorithms have gained popularity in remote sensing applications due to their ability to extract vast amounts of information from satellite images. Significant advances have been made in land cover mapping, particularly with the use of Support Vector Machines (SVM) and Random Forest (RF) algorithms. For instance, [6] demonstrated that SVM achieved significantly higher classification accuracies than other algorithms when applied to Landsat 8 images for land cover mapping in various terrains. Similarly, [8] found that RF outperformed other methods in classifying land cover classes using Landsat 8 imagery, thanks to its ability to handle large datasets and reduce overfitting. Other studies have also employed ML models for LULC mapping [2][5][7][9]. These findings suggest that SVM and RF algorithms are promising options for accurate and reliable land cover mapping, often outperforming alternative classification techniques.

To determine the superior model between these two, this research evaluated the performance of the SVM and RF algorithms for land cover mapping. The study compared the effectiveness of these ML-based classification methods using three different datasets: a Landsat 8 image, its Principal Component Analysis (PCA) image, and derived indices from the Landsat 8 image. Specifically, this research assessed the performance of three common supervised classification algorithms: Maximum Likelihood Classification (MLC), RF, and SVM.

In addition to comparing these classification methods, this study investigated the use of Principal Component Analysis (PCA) as a dimensionality reduction tool to enhance the efficiency of information extraction. PCA reduces the dimensionality of multispectral data while retaining most of the variance, thereby improving computational efficiency and classification accuracy [10]. The study also explored the use of spectral indices, such as the Normalized Difference Vegetation Index (NDVI), Normalized Difference Built-up Index (NDBI), and Normalized Difference Snow Index (NDSI), as supplementary features for classification. These indices provide valuable information about specific land cover categories—such as vegetation, water bodies, and snow cover—allowing for better differentiation between distinct land cover classes.

The aim of this study was to analyze and compare the performance of three machine learning methods—MLC, SVM, and RF—for land cover mapping using Landsat 8 imagery.

Their accuracy was evaluated using three datasets: the multispectral Landsat 8 image, the PCA image, and derived indices such as NDVI, NDBI, and NDSI. Furthermore, this research explored the effectiveness of PCA and spectral indices in improving classification outcomes. The most accurate map was validated by comparison with the European Space Agency's World Cover dataset. The goal of this research is to identify the most effective ML algorithm for land cover mapping while providing accurate LULC data for environmental monitoring. This comparative study offers insights into the consistency and reliability of classification results in relation to existing land cover datasets.

Material and Methods:

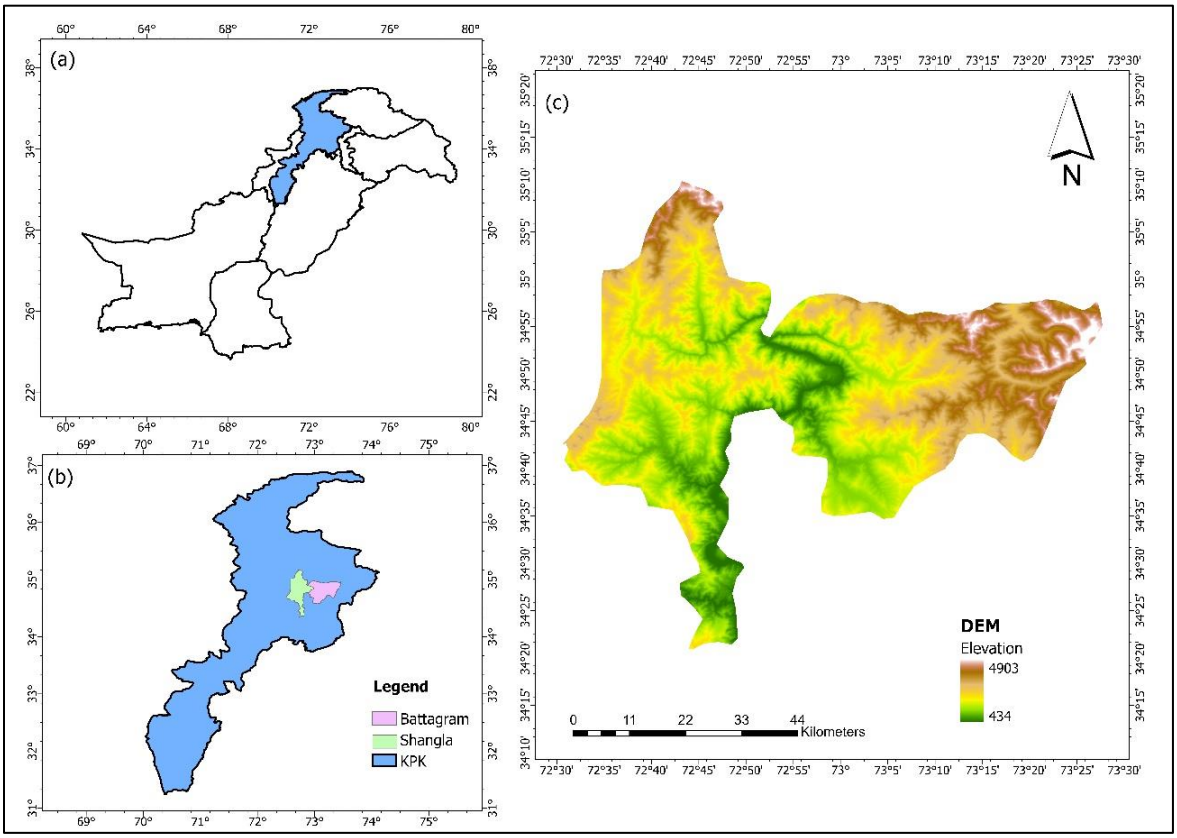


Figure 1: a) International and provincial boundaries of Pakistan **b)** The study area exists in Khyber Pakhtunkhwa, a province of Pakistan **c)** Location of the study area with elevation derived from Shuttle Radar Topography Mission (SRTM) (30 m) data.

Study Area:

The study area encompasses the districts of Shangla and Battagram in Pakistan's Khyber Pakhtunkhwa province, as illustrated in Figure 1. Shangla district is situated between latitudes 34.667°N and 35.033°N and longitudes 72.467°E and 72.917°E, while Battagram district lies between latitudes 34.267°N and 34.767°N and longitudes 72.767°E and 73.117°E. These districts, located in northern Pakistan, are characterized by diverse geographical features, including rocky hills, valleys, and agricultural plains. Shangla district is renowned for its picturesque landscapes, steep hills, terraced fields, and forest cover [11]. Battagram district, on the other hand, features a combination of rugged terrain and fertile valleys intersected by rivers and streams [12]. The study area holds significant environmental and socioeconomic importance, with agriculture serving as the primary source of livelihood for the local population. Additionally, the region is prone to natural disasters such as landslides and floods, making accurate land cover mapping essential for disaster management and land use planning. By focusing on these two districts, this study aims to provide valuable insights into the performance of various classification algorithms for land cover mapping in hilly terrains.

Methodology:

The study evaluated the effectiveness of various image classification techniques applied to Landsat 8 imagery in generating accurate LULC maps. The classification process involved multiple steps, as depicted in Figure 2. These steps included data acquisition, pre-processing, training sample collection, and supervised image classification. The study employed three machine learning algorithms—Support Vector Machine (SVM), Random Forest (RF), and Maximum Likelihood Classifier (MLC)—using multispectral bands, Principal Component Analysis (PCA) bands, and normalized indices such as the Normalized Difference Vegetation Index (NDVI), Normalized Difference Built-up Index (NDBI), and Normalized Difference Snow Index (NDSI). The final step involved assessing the accuracy of the classifications and comparing the results with the European Space Agency (ESA) land cover map.

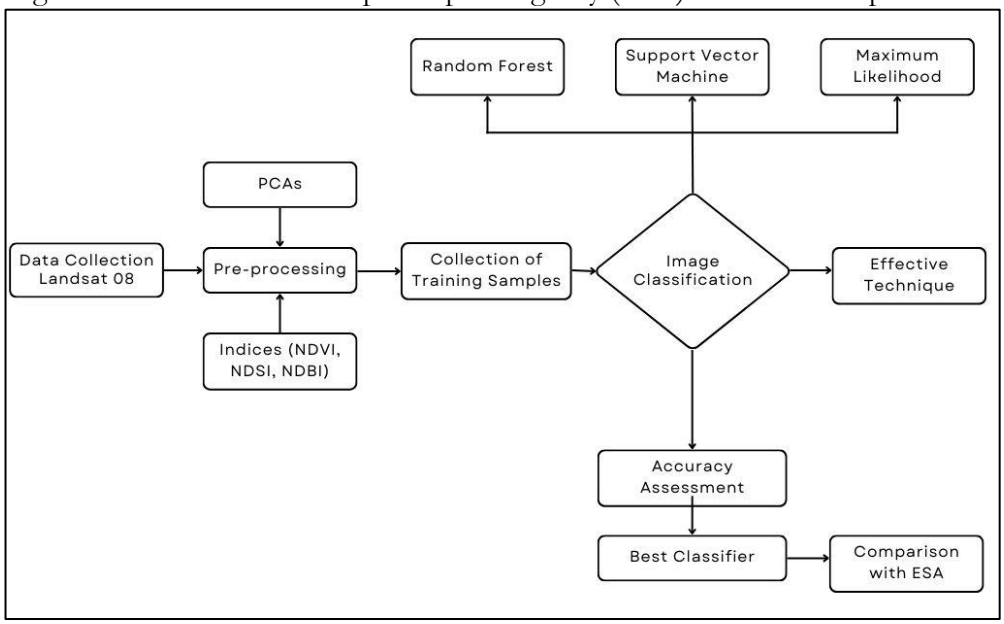


Figure 2: Representation of adopted Methodology.

Data Collection:

In this study, Landsat 8 Operational Land Imager (OLI) satellite imagery served as the primary data source for land cover classification (Table 1). Landsat satellites are widely used for constructing LULC maps due to their favorable spatial and temporal resolution [13][14][15]. Landsat 8 offers multispectral imagery with a 30-meter spatial resolution, making it well-suited for detailed land cover mapping across large areas. The dataset was obtained from the USGS Earth Explorer platform (https://earthexplorer.usgs.gov/). Additionally, the ESA's World Cover LULC map was used as a reference dataset for comparative analysis, accessed from (https://esa-worldcover.org/en). The World Cover LULC map provides a standardized depiction of land cover across different regions, derived from satellite imagery and ground truth data.

Table 1: Description of primary datasets used in the study.

Data	Satellite	Resolution (m)	Date/Year	Location
Satellite Imagery	Landsat 08 (OLI)	30	2021/10/06	Battagram, Shangla
ESA World Cover	Sentinel-1/2	10	2021	Battagram, Shangla

Principal Component Analysis:

Principal Component Analysis (PCA) was employed as a dimensionality reduction technique to enhance the accuracy of land cover mapping using Landsat 8 OLI satellite images. PCA transforms multispectral data into a set of uncorrelated variables, known as principal components, which capture the majority of the variance in the original dataset. By retaining only, the most informative components and eliminating redundant information, PCA reduces the

dataset's dimensionality, thereby improving computational efficiency and significantly impacting spatial and spectral resolution [16][17][18]. In this study, PCA was applied to Landsat 8 OLI images to identify the most significant spectral features for LULC classification. These principal components were then used as input variables for supervised classification algorithms, including Maximum Likelihood Classifier (MLC), Support Vector Machines (SVM), and Random Forest (RF).

Band Indices:

In remote sensing, band indices are mathematical combinations of spectral bands derived from remote sensing data, such as satellite or aerial imagery. These indices are instrumental in highlighting specific features of interest, including vegetation, water bodies, and soil types [17]. In this study, NDVI, NDSI, and NDBI band indices were utilized to extract seven feature classes (Figure 3).

Normalized Difference Vegetation Index (NDVI):

The Normalized Difference Vegetation Index (NDVI) is a widely used spectral index in remote sensing for assessing vegetation health and abundance. The features extracted by NDVI are illustrated in Figure 3a. NDVI is calculated as the normalized difference between the near-infrared (NIR) and red bands of satellite imagery, with higher NDVI values indicating denser and healthier vegetation cover (Equation 1).

$$NDVI = \frac{NIR-RED}{NIR+RED}, \quad (1)$$

Where NIR and RED represent the corresponding bands of the image.

Normalized Difference Snow Index (NDSI):

The Normalized Difference Snow Index (NDSI) is a spectral index used to detect snow and ice cover in satellite imagery. The features extracted by NDSI are shown in Figure 3b. NDSI is calculated as the normalized difference between the green and shortwave infrared (SWIR) bands, with higher NDSI values indicating the presence of snow or ice, as shown in Equation 2.

$$NDSI = \frac{Green-SWIR}{Green+SWIR} \quad (2)$$

Where Green and SWIR represent the corresponding bands of the image.

Normalized Difference Built-up Index (NDBI):

The Normalized Difference Built-up Index (NDBI) is a spectral index used to identify built-up areas and urban infrastructure in satellite imagery. The features extracted by NDBI are illustrated in Figure 3c. NDBI is calculated as the normalized difference between the shortwave infrared (SWIR) and near-infrared (NIR) bands, with higher NDBI values indicating the presence of built-up structures, as shown in Equation 3.

$$NDBI = \frac{SWIR-NIR}{SWIR+NIR} \quad (3)$$

Where NIR and SWIR represent the corresponding bands of the image.

Methods for LULC Classification:

Assigning predetermined classes to individual pixels or image segments based on their spectral characteristics is a crucial step in the classification process [19]. In the Shangla and Battagram districts of Khyber Pakhtunkhwa province, Pakistan, Landsat 8 multispectral satellite imagery, along with its PCA image and stacked band ratios derived from the original image, was classified using supervised classification algorithms. These classifications resulted in seven LULC classes: grasses, urban areas, forests, agriculture, water, snow, and bare areas, as detailed in Table 2. A total of 530 sample points were used in this study, with 350 allocated for training the algorithms and the remaining 180 samples reserved for validation and accuracy assessment.

Support Vector Machine:

Support Vector Machine (SVM) was originally developed to address binary classification problems and is based on the Structural Risk Minimization (SRM) principle. It utilizes a

hyperplane to classify data points [20], aiming to maximize the margin between different classes through the support vectors that define this margin. SVM is capable of handling both continuous and categorical variables and can classify linear and non-linear samples with varying class memberships. Although radial basis function (RBF) and polynomial kernels are commonly used in remote sensing, RBF is preferred for LULC classification due to its superior accuracy compared to other conventional techniques [5].

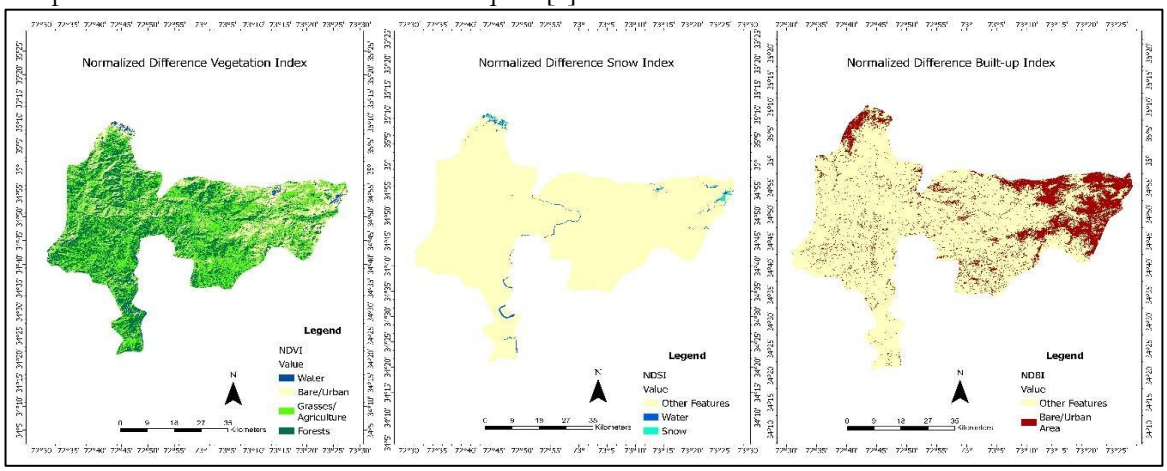


Figure 3: Layers of normalized indices that were used for classification **(a)** NDVI, **(b)** NDSI, and **(c)** NDBI

Table 2: Explains the description of all LULC classes

Classes	Description
Grasses	Areas dominated by natural or cultivated grasslands
Urban Area	Developed areas like buildings, infrastructure, and human settlements
Forest	Dense areas covered by trees
Agriculture	Cultivated areas such as farming, crop cultivation, and grazing.
Water	Bodies of water including rivers, lakes, reservoirs, and ponds.
Snow	Areas covered by snow and ice
Bare Area	Open areas like barren lands and bare soils

The SVM approach begins with a dataset and seeks to identify the hyperplane that best separates the data into distinct classes. To achieve this, SVM relies on an appropriate kernel function to minimize classification errors and create effective hyperplanes. The choice of kernel is crucial, as it influences the smoothness of the separating surface. The performance of SVM, particularly when applied to hyperspectral satellite data, is optimized with a well-chosen kernel and may benefit from genetic optimization methods. The primary objective of SVM is to determine the optimal boundary that maximizes the distance between all support vectors [15].

Random Forest:

Random Forest (RF) is a widely used classifier for LULC classification, often referred to as an ensemble classifier because it relies on a multitude of decision trees for classification [7]. RF generates numerous random decision trees by using predictions from previously established trees [21]. While a subset of the input data is used for training the RF model, the remaining data is reserved for objective validation. The final classification result is derived by averaging the predictions of each tree to determine the likelihood of each class. RF is known for its robustness and effectiveness in managing issues like noise and overfitting [5]. Additionally, RF is well-suited for handling high-dimensional data. Various factors can influence the final classification outcome of the RF model [1].

Maximum Likelihood:

The Maximum Likelihood (ML) classifier is a widely used supervised classification technique in remote sensing for land cover mapping. In this study, the ML classifier was applied

to Landsat 8 OLI satellite images to assign each pixel to the most probable land cover class based on its spectral signature. The ML classifier operates under the assumption that pixel values within each class follow a multivariate normal distribution in the feature space [22]. It calculates the probability of a pixel belonging to each class and assigns it to the class with the highest likelihood, based on the mean vector and covariance matrix estimated from the training data [23]. By utilizing spectral information from various bands of satellite imagery, the ML classifier effectively differentiates between diverse land cover types, making it particularly useful for complex landscapes with multiple land cover classifications.

In this study, the ML classifier was evaluated alongside other classification algorithms, such as Support Vector Machines (SVM) and Random Forest (RF), to compare their effectiveness in accurately mapping land cover in the Shangla and Battagram districts of Khyber Pakhtunkhwa province, Pakistan. The ML classifier's ability to model the statistical distribution of spectral data and make probabilistic predictions underscores its importance for land cover categorization and environmental monitoring [18].

Accuracy Assessment:

The accuracy of a classifier is influenced by factors such as input data, the study region, and satellite sensors. Different researchers have reported varying LULC accuracy for different classifiers across diverse study areas [14][24]. Accuracy assessments for these techniques were conducted using a confusion matrix, which compares the actual and predicted classifications made by the classification system. This matrix evaluates how well the categorized pixels from the image match field observations at the same locations. The results of an accuracy assessment typically provide the overall accuracy (OA) of the map as well as the accuracy for each individual class. The overall accuracy percentage was calculated using Equation 4.

$$\textbf{Overall Accuracy} = \frac{\textbf{Corrected samples} * \textbf{100}}{\textbf{total samples}} \quad (4)$$

In addition to overall accuracy, the classification accuracy for individual classes was assessed using two key metrics: user's accuracy (UA) and producer's accuracy (PA). Producer's accuracy is calculated by dividing the number of correctly classified pixels in a given class by the total number of pixels in that class, as indicated by reference data, and is detailed in Equation 5. In this study, producer's accuracy measures the effectiveness of classifying specific areas.

$$\textbf{Producer's Accuracy} = \frac{\textbf{properly classified pixels of class} * \textbf{100}}{\textbf{total classified pixels of class}} (\textbf{the column total}) \quad (5)$$

Meanwhile, user's accuracy is calculated by dividing the number of correctly classified pixels in each category by the total number of pixels that were classified into that category. User's accuracy reflects the probability that a pixel classified into a particular category actually belongs to that category, as shown in Equation 6 [25].

$$\textbf{User's Accuracy} = \frac{\textbf{properly classified pixels of class} * \textbf{100}}{\textbf{total classified pixels of class}} (\textbf{the row total}) \quad (6)$$

Kappa Coefficient:

The Kappa coefficient (Kc) is another important measure used in many studies [24][26][27]. It is calculated by multiplying the total number of pixels across all ground truth classes (N) by the sum of the confusion matrix diagonals (Xkk), then subtracting the sum of the products of ground truth pixels in each class and classified pixels in that class, summed over all classes ($\Sigma X k\Sigma\ Yk\Sigma$), where $Xk\Sigma$ represents the row totals and $Yk\Sigma$ represents the column totals. This result is divided by the total number of pixels squared minus the sum of the products of ground truth pixels and classified pixels for each class. The Kappa value ranges from 0 to 1, with 0 indicating agreement due only to chance and 1 indicating perfect agreement between the datasets. Negative values, though rare, can occur but are considered spurious. The Kappa statistic is often expressed as a percentage (%) and provides a more nuanced measure of classifier agreement compared to overall accuracy [28], offering better interclass discrimination. The Kappa coefficient is calculated using Equation 7.

$$KC = \frac{N\sum cr\ i=1\ xjk\ \sum cr\ i=1\ xj+\ x+j\ )}{N2\ \sum r\ i=1\ xj+\times x+j\ )} \quad (7)$$

Where, xij = number of counts in the ijth cell of the confusion matrix

- N = total number of counts in the confusion matrix
- xi+ = marginal total of row i
- x+i = marginal total of column i

Result and Discussion:

LULC Classifications:

This study analyzed seven different types of land use and land cover (LULC) in the Battagram and Shangla districts for the year 2021 using Landsat 8 imagery and its derivatives (PCA and band ratios). The analysis employed three commonly used supervised learning algorithms—Maximum Likelihood Classification (MLC), Random Forest (RF), and Support Vector Machines (SVM)—each yielding slightly different results. Figure 4 illustrates the spatial distribution of all LULC classes as determined by these three machine learning algorithms using Landsat 8 multispectral imagery, PCA, and stacked indices.

Classification Using Multispectral Bands of Landsat 8 Image:

The results reveal that forest cover is the most dominant land cover type in Shangla and Battagram, as shown in Figure 5a. Using the Landsat 8 image, the SVM classifier identified 2,145 ha as water bodies, 12,563 ha as urban areas, 109,329 ha as forest, 1,828 ha as snow, 48,676 ha as bare areas, 90,906 ha as grasslands, and 48,021 ha as agricultural land.

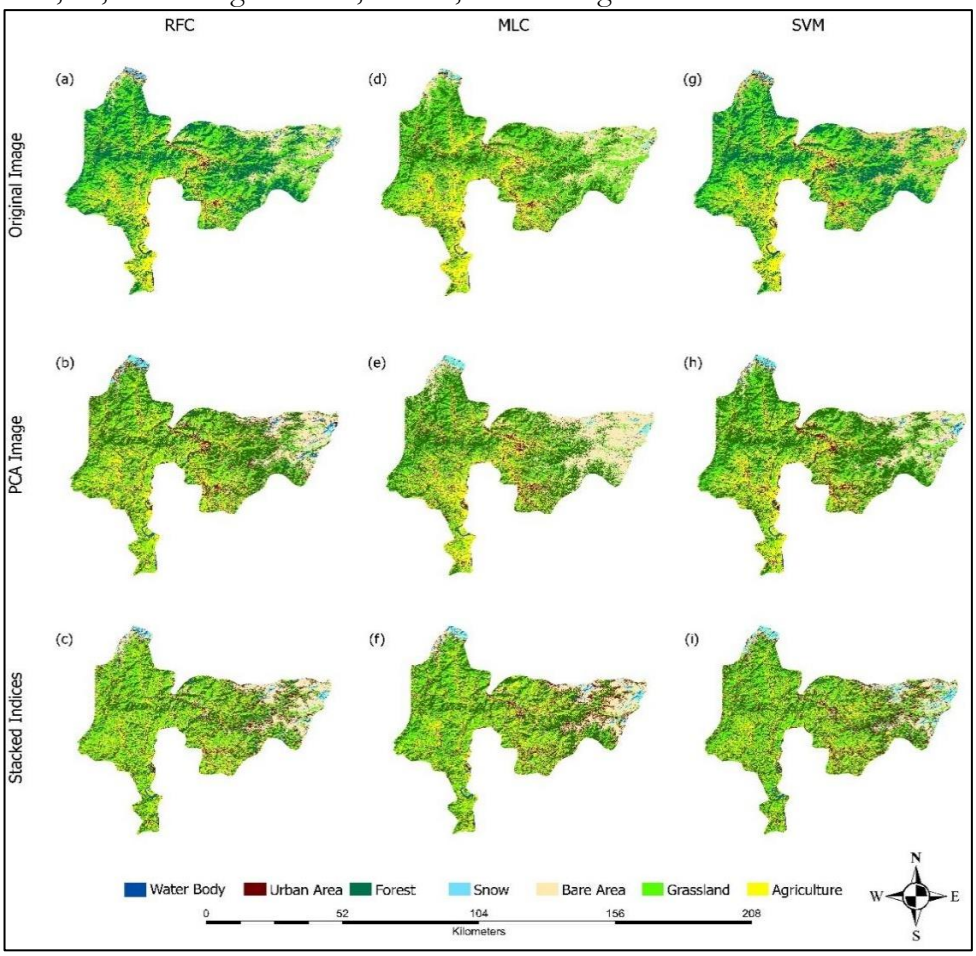


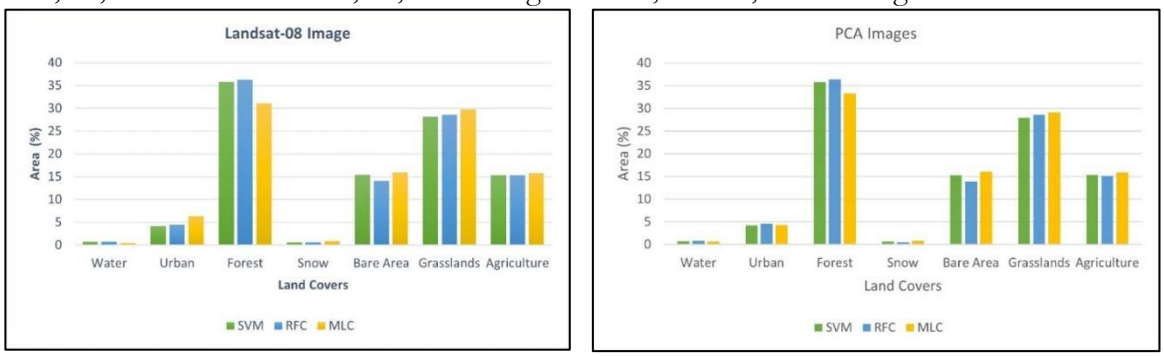
Figure 4: Spatial Distribution of LULC classification by RF, SVM, and MLC by using Landsat 08, PCA, Stacked Indices

The Random Forest classifier classified 2,338 ha as water bodies, 13,556 ha as urban areas, 110,916 ha as forest, 1,802 ha as snow, 42,980 ha as bare areas, 87,197 ha as grasslands,

and 46,675 ha as agricultural land. The Maximum Likelihood classifier identified 1,171 ha as water bodies, 19,404 ha as urban areas, 94,701 ha as forest, 2,585 ha as snow, 48,676 ha as bare areas, 90,906 ha as grasslands, and 48,021 ha as agricultural land.

Classification Using PCA Image:

Results using the PCA images are depicted in Figure 4, with Figure 5b illustrating the area comparison of LULC derived from the PCA image. For the PCA image classification, covering 305,464 ha, the SVM classifier identified 2,243 ha as water bodies, 12,994 ha as urban areas, 109,329 ha as forest, 1,928 ha as snow, 46,696 ha as bare areas, 85,298 ha as grasslands, and 46,976 ha as agricultural land. The Random Forest classifier classified 2,468 ha as water bodies, 14,156 ha as urban areas, 109,141 ha as forest, 1,602 ha as snow, 42,380 ha as bare areas, 87,297 ha as grasslands, and 48,420 ha as agricultural land. The Maximum Likelihood classifier identified 1,971 ha as water bodies, 13,104 ha as urban areas, 101,702 ha as forest, 2,385 ha as snow, 48,976 ha as bare areas, 88,906 ha as grasslands, and 48,420 ha as agricultural land.



(a) (b)

Figure 5: Area comparison of land cover classes **(a)** using Landsat-08 **(b)** using PCA image

Classification Using Stacked Indices:

Results using the stacked indices are illustrated in Figure 4. From the classification of the image, covering 305,464 ha, the SVM classifier identified 2,343 ha as water bodies, 20,994 ha as urban areas, 102,329 ha as forest, 2,328 ha as snow, 43,743 ha as bare areas, 91,258 ha as grasslands, and 42,479 ha as agricultural land. The Random Forest classifier classified 2,284 ha as water bodies, 22,140 ha as urban areas, 101,241 ha as forest, 2,200 ha as snow, 45,560 ha as bare areas, 92,262 ha as grasslands, and 39,780 ha as agricultural land. The Maximum Likelihood classifier identified 2,476 ha as water bodies, 20,099 ha as urban areas, 98,717 ha as forest, 2,480 ha as snow, 50,676 ha as bare areas, 89,306 ha as grasslands, and 41,720 ha as agricultural land, as shown in Figure 6.

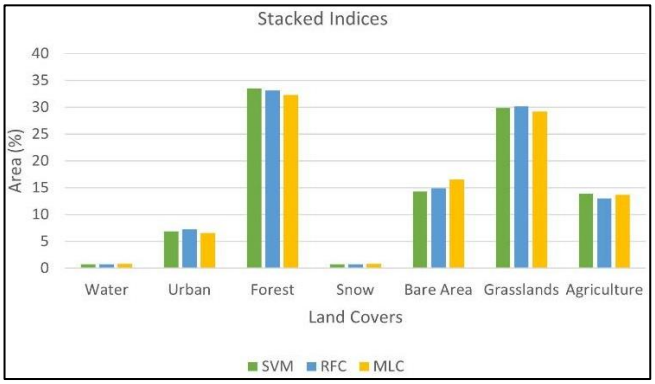


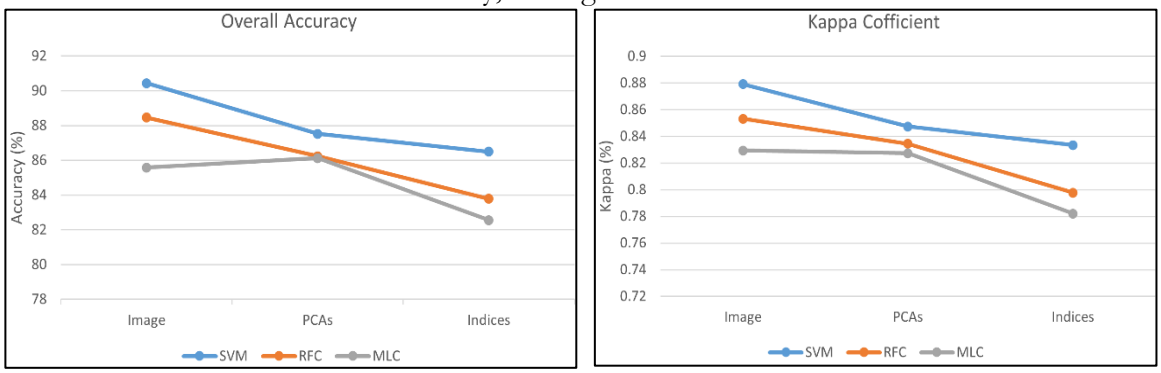
Figure 6: Comparison of the area of LC classes using Stacked Indices Image

This study assesses the performance of classification algorithms using multispectral images, Principal Component Analyses (PCA), and normalized indices. Consequently, the area statistics may not fully align with national estimates for some land cover and land use classes.

Comparison of Classifiers:

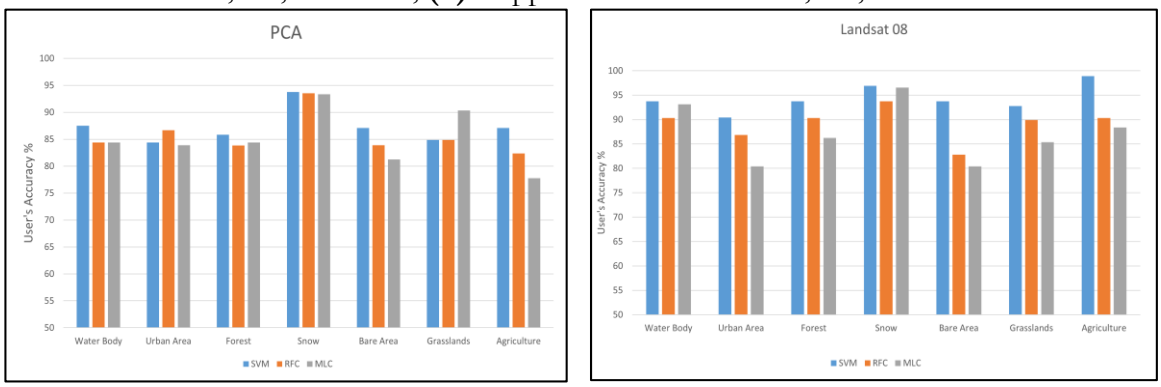
To evaluate classifier performance, three different classifiers were applied to three types of images: Landsat 8, PCA, and stacked indices (including NDVI, NDSI, and NDBI). Overall accuracy and the Kappa coefficient are commonly used metrics for determining classifier performance. Figure 7a presents the overall accuracy, while Figure 7b displays the Kappa coefficients for each classifier across Landsat 8, PCA, and stacked indices.

The results indicate that the Support Vector Machine (SVM) classifier outperformed both the Random Forest Classifier (RFC) and the Maximum Likelihood Classifier (MLC) across all three image types. The SVM classifier achieved a maximum overall accuracy of 90.43% for the Landsat 8 image, compared to 87.52% for RFC and 84.5% for MLC. Among the classifiers, MLC showed the lowest overall accuracy, trailing behind both SVM and RFC.

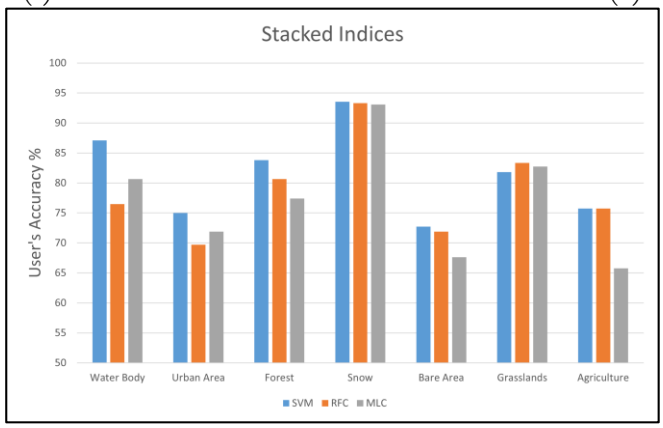


(a) (b)

Figure 7: Performance comparison of Machine Learning algorithms. **(a)** Overall accuracy for MLC, RF, and SVM, **(b)** Kappa coefficient of MLC, RF, and SVM



(a) (b)

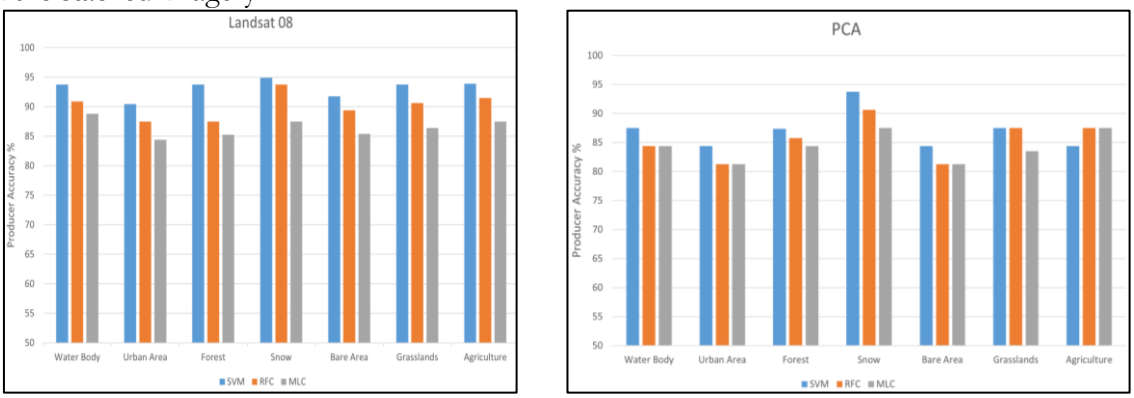


(c)

Figure 8: User's Accuracy of all land cover classes using SVM, RFC, and MLC classifiers on **(a)** PCA image **(b)** Landsat 08 image, and **(c)** Stacked indices image

Among the two derived datasets from Landsat satellite imagery, the Principal Component Analysis (PCA) from Landsat 8 performed best across all three classification methods. PCA achieved overall accuracies of 86.12% for the Maximum Likelihood Classifier (MLC), 86.23% for the Random Forest Classifier (RFC), and 87.52% for the Support Vector Machine (SVM) classifier. In contrast, the Landsat-8-derived stacked indices (NDVI, NDSI, and NDBI) performed the least effectively among the three datasets, with overall accuracies of 82.56% for MLC, 83.78% for RFC, and 86.5% for SVM.

The study assessed the performance of SVM, MLC, and RFC models on three data products for each LULC class, using User Accuracy (UA) and Producer Accuracy (PA) metrics. Results for each class were reported separately due to varying model performance. Figure 8 illustrates the user's accuracy for each classifier using Landsat 8, its PCA, and derived indices imagery. Snow cover generally had the highest user accuracy, while Bare Area had the lowest. The Landsat 8 image consistently outperformed the PCA and stacked indices images across most classes. The SVM classifier demonstrated superior performance compared to the other classifiers. Similar trends were observed in the producer accuracy results. Snow cover achieved the highest producer accuracy in all images, as shown in Figures 9 and 10. Both user and producer accuracies were notably lower for the stacked indices image, as depicted in Figures 8c and 10. This reduction in accuracy is attributed to misclassifications, particularly of Bare and Urban areas, due to the NDBI index used for urban area extraction, which impacted the accuracy of the stacked imagery.



(a) (b)

Figure 9: Producer's Accuracy of all land cover classes using SVM, RFC, and MLC classifiers on **(a)** Landsat 08 image **(b)** PCA image

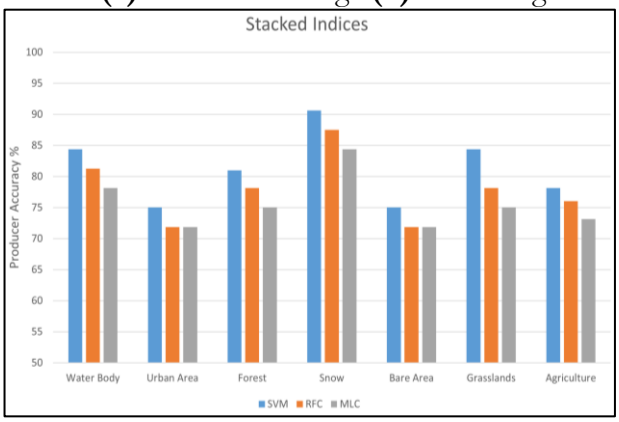


Figure 10: Producer's Accuracy of land cover classes using SVM, RFC, and MLC classifiers on Stacked Indices.

ESA Landcover Validation:

The ESA land cover map, validated with the same points, achieved an overall accuracy of 70.16%. This moderate accuracy can be attributed to several factors. Firstly, the temporal

aspect of the datasets contributed to reduced accuracy. The ESA land cover map, updated annually, was validated using data from October 6, 2021. Consequently, the map may not capture the dynamic changes in land cover throughout the year, leading to class mixing and temporal mismatches. Secondly, the resolution discrepancy between datasets affected the accuracy; our validation points were based on 30-meter resolution, while the ESA map has a finer resolution of 10 meters. These factors highlight the challenges in achieving improved validation accuracy with datasets of varying resolutions and temporal spans.

Discussion:

This study assessed Landsat 8 multispectral images and their derivatives (PCA and indices) to evaluate the accuracy of machine learning classifiers (MLC, SVM, and RF) for land cover classification. Accuracy was assessed using user accuracy (UA), producer accuracy (PA), overall accuracy (OA), and the Kappa coefficient (Kc). The results showed that the SVM classifier outperformed both RF and MLC across all datasets. SVM achieved the highest overall accuracy of 90.43% with the Landsat 8 image, compared to 87.52% for RF and 84.5% for MLC. These findings align with previous research [4][6][19], indicating that SVM provides robust predictions and captures complex linkages effectively. The SVM model exhibited higher OA and Kappa indices than RF, as illustrated in Figure 7. According to [15], RF's stability is influenced by the number of trees, bagging, and random impressions, which can affect efficiency and accuracy. MLC demonstrated lower accuracy compared to RF and SVM, as shown in Figure 7. Research suggests that RF and SVM classifiers perform better due to their resilience to noise [29] and their ability to handle random and systematic noise in training data [9][30], with performance also dependent on data dimensionality [5].

The performance of the classifiers was influenced by both the machine learning models and the data used. PCA images improved classification accuracy for certain classes but introduced discrepancies. For MLC, PCA enhanced accuracy for some classes (Figure 7) but led to misclassifications in others. SVM is effective and efficient in high-dimensional spaces with clear class separations but requires extensive training for large datasets.

In classifying LULC in the Battagram and Shangla districts, this study found that RF and MLC had the lowest accuracy in bare and urban areas, especially with the stacked indices dataset due to NDBI's influence. Unplanned urbanization and limited high-rise constructions in these districts caused minor misclassifications, affecting accuracy. Despite these challenges, all three models demonstrated reliable accuracy in detecting distinct LULC classes, with snow cover being the most accurately classified. Previous studies have shown that different classifiers can yield varying results for LULC mapping depending on the data and study area [31][32].

Conclusion:

LULC mapping is crucial for sustainable development and economic resilience. Machine learning algorithms generally outperform other methods for LULC classification in terms of reliability and accuracy. Understanding the factors affecting algorithm accuracy can enhance classification results. This study evaluated SVM, RF, and MLC algorithms using Landsat 8 multispectral images, PCA, and stacked normalized indices. The findings indicate that machine learning algorithms effectively analyze Landsat data and its derivatives for LULC analysis. SVM emerged as the top-performing algorithm across all datasets, achieving 90.43% accuracy for Landsat 8, 87.52% for PCA, and 86.5% for band indices imagery. RF followed with 88.46% for Landsat 8, 86.23% for PCA, and 83.78% for band indices imagery, while MLC had the lowest accuracy at 85.58% for Landsat 8, 86.12% for PCA, and 82.56% for indices imagery. The Landsat 8 image achieved the highest overall accuracy, and the ESA land cover map achieved 70.16% accuracy using the same validation points.

LULC classification results vary based on satellite data, research region, and algorithms. Different classifiers perform differently across climate zones and geographical areas. For more accurate LULC classification, researchers should use high-resolution imaging, specific class

definitions, field data, and advanced machine learning methods. This study may guide policymakers and planners in mountainous cities. Changes in study area, duration, algorithms, or datasets can significantly impact classification accuracy.

Acknowledgments: We extend our gratitude to the organizations providing satellite-derived data for this study: the Landsat 8 product managed by NASA and the U.S. Geological Survey, and the World Cover map by the European Space Agency. Their efforts in making these resources publicly available have greatly benefited this research.

Author's Contribution: All the authors had different contributions to this research work and are mentioned here accordingly. Conceptualization and supervision (S.A, Z), formal analysis (Z), methodology (Z, N.H, and S.A), writing—original draft preparation (Z, S.A, and A.T.M), writing—review and editing (S.A, Z, N.H, S.R), visualization (Z) All authors have read and agreed to the published version of the manuscript.

Conflict of Interest: All the authors declare that they have no conflict of interest in publishing this manuscript in IJIST.

References:

[1] A. Jamali, "Evaluation and comparison of eight machine learning models in land use/land cover mapping using Landsat 8 OLI: a case study of the northern region of Iran," SN Appl Sci, vol. 1, no. 11, Nov. 2019, doi: 10.1007/s42452-019-1527-8.

[2] S. Aldiansyah and R. A. Saputra, "COMPARISON OF MACHINE LEARNING ALGORITHMS FOR LAND USE AND LAND COVER ANALYSIS USING GOOGLE EARTH ENGINE (CASE STUDY: WANGGU WATERSHED)."

[3] S. Bedi, A. Samal, C. Ray, and D. Snow, "Comparative evaluation of machine learning models for groundwater quality assessment," Environ Monit Assess, vol. 192, no. 12, Dec. 2020, doi: 10.1007/s10661-020-08695-3.

[4] L. Ghayour et al., "Performance evaluation of sentinel-2 and landsat 8 OLI data for land cover/use classification using a comparison between machine learning algorithms," Remote Sens (Basel), vol. 13, no. 7, Apr. 2021, doi: 10.3390/rs13071349.

[5] Z. Zhao et al., "Comparison of Three Machine Learning Algorithms Using Google Earth Engine for Land Use Land Cover Classification," Rangel Ecol Manag, vol. 92, pp. 129–137, Jan. 2024, doi: 10.1016/j.rama.2023.10.007.

[6] M. Z. Hasan, R. S. Leya, and K. S. Islam, "COMPARATIVE ASSESSMENT OF MACHINE LEARNING ALGORITHMS FOR LAND USE AND LAND COVER CLASSIFICATION USING MULTISPECTRAL REMOTE SENSING IMAGE," Khulna University Studies, pp. 33–46, Oct. 2022, doi: 10.53808/kus.2022.icstem4ir.0124-se.

[7] V. Sheth, U. Tripathi, and A. Sharma, "A Comparative Analysis of Machine Learning Algorithms for Classification Purpose," in Procedia Computer Science, Elsevier B.V., 2022, pp. 422–431. doi: 10.1016/j.procs.2022.12.044.

[8] 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, Jun. 2024, doi: 10.1016/j.ejrs.2024.03.003.

[9] F. F. Camargo, E. E. Sano, C. M. Almeida, J. C. Mura, and T. Almeida, "A comparative assessment of machine-learning techniques for land use and land cover classification of the Brazilian tropical savanna using ALOS-2/PALSAR-2 polarimetric images," Remote Sens (Basel), vol. 11, no. 13, Jul. 2019, doi: 10.3390/rs11131600.

[10] R. Eastman, "Guide to GIS and Image Processing Volume 2." [Online]. Available: https://www.researchgate.net/publication/242377547

[11] F. M. Qamer, S. Abbas, R. Saleem, K. Shehzad, H. Ali, and H. Gilani, "Forest cover change assessment in conflict-affected areas of northwest Pakistan: The case of Swat

and Shangla districts," J Mt Sci, vol. 9, no. 3, pp. 297–306, Jun. 2012, doi: 10.1007/s11629-009-2319-1.

[12] F. Haq, M. Irfan, and C. R. Fraser-Jenkins, "Multivariate statistical analysis of the Pteridophytic diversity of District Battagram, Khyber Pakhtunkhwa, Pakistan," Acta Ecologica Sinica, vol. 42, no. 4, pp. 322–331, Aug. 2022, doi: 10.1016/j.chnaes.2022.01.003.

[13] M. Mohajane et al., "Land use/land cover (LULC) using landsat data series (MSS, TM, ETM+ and OLI) in azrou forest, in the central middle atlas of Morocco," Environments - MDPI, vol. 5, no. 12, pp. 1–16, Dec. 2018, doi: 10.3390/environments5120131.

[14] B. Rimal, S. Rijal, and R. Kunwar, "Comparing Support Vector Machines and Maximum Likelihood Classifiers for Mapping of Urbanization," Journal of the Indian Society of Remote Sensing, vol. 48, no. 1, pp. 71–79, Jan. 2020, doi: 10.1007/s12524-019-01056-9.

[15] S. Talukdar, P. Singha, Shahfahad, S. Mahato, B. Praveen, and A. Rahman, "Dynamics of ecosystem services (ESs) in response to land use land cover (LU/LC) changes in the lower Gangetic plain of India," Ecol Indic, vol. 112, May 2020, doi: 10.1016/j.ecolind.2020.106121.

[16] M. Dharani and G. Sreenivasulu, "Land use and land cover change detection by using principal component analysis and morphological operations in remote sensing applications," International Journal of Computers and Applications, vol. 43, no. 5, pp. 462–471, 2021, doi: 10.1080/1206212X.2019.1578068.

[17] V. K. A. Somayajula, D. Ghai, and S. Kumar, "Land Use/Land Cover Change Analysis using NDVI, PCA," in Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021, Institute of Electrical and Electronics Engineers Inc., Apr. 2021, pp. 849–855. doi: 10.1109/ICCMC51019.2021.9418025.

[18] T. Li, S. Li, C. Liang, R. T. Bush, L. Xiong, and Y. Jiang, "A comparative assessment of Australia's Lower Lakes water quality under extreme drought and post-drought conditions using multivariate statistical techniques," J Clean Prod, vol. 190, pp. 1–11, Jul. 2018, doi: 10.1016/j.jclepro.2018.04.121.

[19] S. Basheer et al., "Comparison of Land Use Land Cover Classifiers Using Different Satellite Imagery and Machine Learning Techniques," Remote Sens (Basel), vol. 14, no. 19, Oct. 2022, doi: 10.3390/rs14194978.

[20] M. Pal and P. M. Mather, "Support vector machines for classification in remote sensing," Int J Remote Sens, vol. 26, no. 5, pp. 1007–1011, Mar. 2005, doi: 10.1080/01431160512331314083.

[21] L. Breiman, "RANDOM FORESTS-RANDOM FEATURES," 1999.

[22] B. R. Shivakumar and S. V. Rajashekararadhya, "Investigation on land cover mapping capability of maximum likelihood classifier: A case study on North Canara, India," in Procedia Computer Science, Elsevier B.V., 2018, pp. 579–586. doi: 10.1016/j.procs.2018.10.434.

[23] P. V Bolstad and T. M. Lillesand, "Rapid Maximum Likelihood Classification," 1991.

[24] F. Ghasempour, A. Sekertekin, and S. H. Kutoglu, "HOW LANDSAT 9 IS SUPERIOR TO LANDSAT 8: COMPARATIVE ASSESSMENT OF LAND USE LAND COVER CLASSIFICATION AND LAND SURFACE TEMPERATURE," in ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Copernicus Publications, Jan. 2023, pp. 221–227. doi: 10.5194/isprs-annals-X-4-W1-2022-221-2023.

[25] Y. O. Ouma, A. Keitsile, B. Nkwae, P. Odirile, D. Moalafhi, and J. Qi, "Urban land-use classification using machine learning classifiers: comparative evaluation and post-

classification multi-feature fusion approach," Eur J Remote Sens, vol. 56, no. 1, 2023, doi: 10.1080/22797254.2023.2173659.

[26] L. Khaldi, A. Elabed, and A. El Khanchoufi, "Performance evaluation of Machine Learning algorithms for LULC classification: A case study of Fez-Meknes region," E3S Web of Conferences, vol. 527, p. 02012, May 2024, doi: 10.1051/e3sconf/202452702012.

[27] M. Ganjirad and H. Bagheri, "Google Earth Engine-based mapping of land use and land cover for weather forecast models using Landsat 8 imagery," Ecol Inform, vol. 80, May 2024, doi: 10.1016/j.ecoinf.2024.102498.

[28] M. Safabakhshpachehkenari and H. Tonooka, "Assessing and Enhancing Predictive Efficacy of Machine Learning Models in Urban Land Dynamics: A Comparative Study Using Multi-Resolution Satellite Data," Remote Sens (Basel), vol. 15, no. 18, Sep. 2023, doi: 10.3390/rs15184495.

[29] L. Breiman, "Random Forests," 2001.

[30] C. Pelletier, S. Valero, J. Inglada, N. Champion, C. M. Sicre, and G. Dedieu, "Effect of training class label noise on classification performances for land cover mapping with satellite image time series," Remote Sens (Basel), vol. 9, no. 2, 2017, doi: 10.3390/rs9020173.

[31] A. M. Abdi, "Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data," GIsci Remote Sens, vol. 57, no. 1, pp. 1–20, Jan. 2020, doi: 10.1080/15481603.2019.1650447.

[32] A. Tariq et al., "Modelling, mapping and monitoring of forest cover changes, using support vector machine, kernel logistic regression and naive bayes tree models with optical remote sensing data," Heliyon, vol. 9, no. 2, Feb. 2023, doi: 10.1016/j.heliyon.2023.e13212.