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 development of remote sensing techniques and access to free satellite data has enabled accurate mapping of land use land cover (LULC) to analyze landscape transformations and changes in the ecosystem. However, it remains challenging to choose the best classifier for LULC mapping. Therefore, it is required to assess the accuracy of multiple LULC modeling algorithms for their effective applications. This study performed a comprehensive assessment of supervised machine learning and conventional classification algorithm, applied to a Landsat 8 image with 30 m spatial resolution of the Shangla and Battagram districts, KPK, Pakistan. Three classification algorithms, including Maximum Likelihood Classification (MLC), Support Vector Machines (SVM), and Random Forest (RF), were explored. This research evaluated the performance of these algorithms on multispectral images as well as on composite images obtained from Principal Component Analysis (PCA) and Band Ratioing and/or Normalized Indices. Furthermore, the accuracy of these algorithms with different datasets was compared with the recently released World Cover LULC by the European Space Agency (ESA). The results revealed that the SVM algorithm exhibited superior performance, and achieved an overall accuracy of 90.43% and a kappa coefficient of 0.8792. MLC and RF algorithms also yielded promising results, with overall accuracies of 85.58% and 88.46%, respectively. Additionally, when assessed the accuracy of the World Cover, with similar validation samples, research found that the overall accuracy of the ESA's LULC was 70.67% in the study area. These findings highlight the strengths and limitations of each algorithm, offering insights into their suitability for LULC classification, as well as the applicability of available global LULC maps.

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

Introduction:

Environmental monitoring techniques are necessary to understand the earth's changing climate, including the rise in the planet's surface temperature and the necessity to track its effects on the surrounding environment [1]. Land use and land cover (LULC) is one of the most known variables in that case. Accurate LULC data can help with various research efforts on floods, droughts, and migration at multiple scales [2][3]. Demand and availability for LULC maps have increased recently [4], partly as a result of the expansion of publicly available satellite imagery. Assessment of LULC is an essential component of any region's sustainability and development. LULC changes (such as deforestation and urbanization) are a major driver of climate change [3] around the world. Extreme occurrences of environmental degradation caused by humans and natural disasters are greatly impacted by LULC change, which is an important global issue [5]. Furthermore, climate change substantially impacts water balance, geomorphology, water quality, groundwater management, resource management and its effects on humans and the environment, and land monitoring, all of which require detailed LULC maps.

LULC changes triggered by a fast-rising population are one of the most pressing challenges in developing nations such as Pakistan, particularly in cities. Urban sprawl encroaches on agricultural territory in most developing countries, posing food security concerns. LULC change is a major global concern with far-reaching consequences for both natural and human-caused severe events. Most past research has concentrated solely on evaluating the accuracy of LULC classification using geographic data, rather than selecting appropriate LULC change detection strategies [6][7]. It is important to select the proper technique to monitor changes in LULC to manage agriculture, natural resources, and energy under the population. For this purpose, accurate LULC maps are required.

Machine learning (ML) algorithms have recently gained popularity in remote sensing applications due to their ability to successfully extract enormous amounts of information from satellite images. Previous research in the field of land cover mapping has made substantial advances, particularly in the use of Support Vector Machines (SVM) and Random Forest (RF) approaches. Some studies have demonstrated the effectiveness of SVM and RF algorithms in reliably classifying land cover types using multispectral imagery collected from Landsat 8 satellite data. For example: [6] SVM obtained much greater 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, owing to its capacity to handle large datasets and reduce overfitting. Other than these studies several studies use ML models for LULC [2][5][7][9]. These findings combined revealed that SVM and Random Forest algorithms have emerged as potential options for accurate and reliable land cover mapping, outperforming alternative classification techniques in previous research attempts. To find the best one between these two models, this research was conducted to evaluate the performance of two machine learning algorithms, SVM and RF, for land cover mapping. The study compared the effectiveness of three 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 examined the performance of three common supervised classification algorithms: Maximum Likelihood Classification (MLC), Random Forest (RF), and Support Vector Machine (SVM).

In addition to examining these classification methods, this study investigates the use of Principle Component Analysis (PCA) as a dimensionality reduction tool to improve the effectiveness of quacking extraction of information. PCA has the advantage of lowering the dimensionality of multispectral data while retaining the majority of the variance, which improves computing efficiency and classification accuracy [10]. Also investigate 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

categorization. These indices provide useful information about certain land cover categories, such as vegetation, water bodies, and snow cover, allowing for better discrimination between distinct land cover classes.

This study aims to analyze and compare the performance of three machine learning methods for mapping land cover using Landsat 8 imagery: MLC, SVM, and RF. This study evaluated their accuracy with three datasets: the Multispectral Landsat 8 image, the Principal Component Analysis (PCA) image, and derived indices such as NDVI, NDBI, and NDSI. In addition, this research work explored the efficiency of PCA and spectral indices in classification. The most accurate map is validated by comparing it to the European Space Agency's World Cover dataset. The goal of this research is to find the best ML algorithm for land cover mapping while also providing accurate land use and land cover (LULC) data for environmental monitoring. This comparative research will shed light on the consistency and reliability of classification results to existing land cover datasets.

Material and Methods:

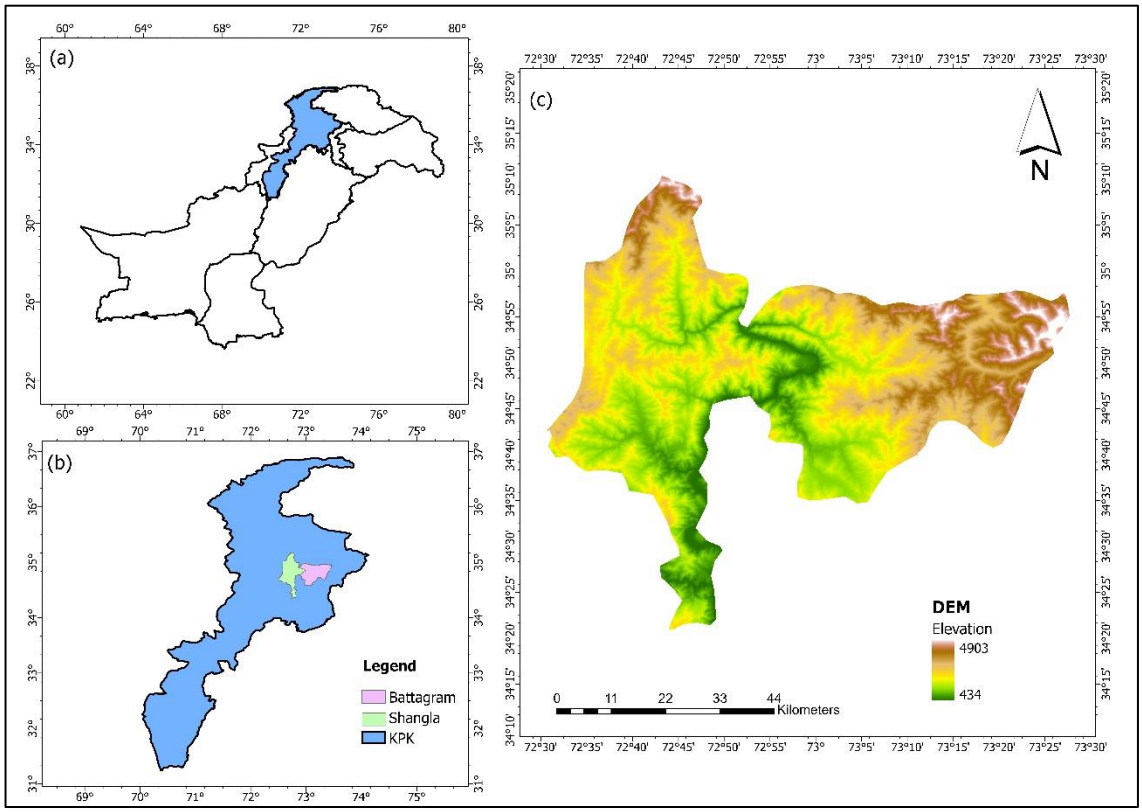


Figure 1: a) International and provincial boundaries of Pakistan **b)** The study area exists in Khyber Pakhtunkhwa, a province of Pakistan **c)** Location of the study area with elevation derived from Shuttle Radar Topography Mission (SRTM) (30 m) data.

Study Area:

The study area includes the districts of Shangla and Battagram in Pakistan's Khyber Pakhtunkhwa province as shown in Figure 1. Shangla district is around 34.667°N to 35.033°N latitude and 72.467°E to 72.917°E longitude, whereas Battagram district is approximately 34.267°N to 34.767°N latitude and 72.767°E to 73.117°E longitude. These districts are located in Pakistan's northern area and are distinguished by various geographical features, including rocky hills, valleys, and agricultural plains. Shangla district is known for its attractive landscapes, which include steep hills, and terraced fields, and its forest cover [11] Battagram district has a mix of rocky terrain and fertile valleys crossed by rivers and streams [12]. The research area has great environmental and socioeconomic value, with agriculture providing the principal source

of income for the local inhabitants. Furthermore, the region is prone to natural disasters like landslides and floods, therefore precise land cover mapping is critical for disaster management and land use planning activities. By focusing on these two areas, this study hopes to gain useful insights into the performance of various classification algorithms for land cover mapping in a hilly area.

Methodology:

The study compared the performance of different image classification techniques on Landsat 8 imagery to assess effectiveness in producing accurate LULC maps. The procedures for these classifications involved several steps as shown in Fig.2, including data acquisition, pre-processing, training sample collection, supervised image classification, extraction, and performing the three ML algorithms i.e. Support Vector Machine (SVM), Random Forest (RF) and Maximum Likelihood Classifier (MLC) using Multispectral bands, Principle Component Analysis (PCA) bands, and normalized indices i.e. Normalized Difference Vegetation Index (NDVI), Normalized Difference Built-up index (NDBI), and Normalized Difference Snow Index (NDSI), accuracy assessment, and comparison with an ESA land cover map.

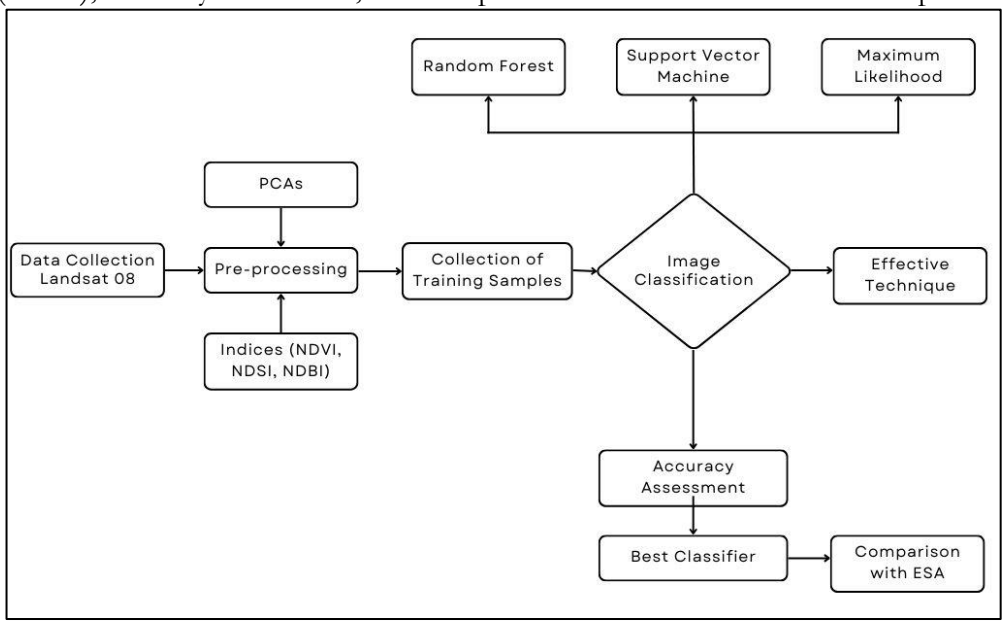


Figure 2: Representation of adopted Methodology.

Data Collection:

For this study, Landsat 8 Operational Land Imager (OLI) satellite image was used as a primary data source (Table 1) for land cover classification. Landsat satellite is a very common dataset for building LULC maps due to decent spatial and temporal resolution [13][14][15]. Landsat 8 provides multispectral imagery at a spatial resolution of 30 meters, making it ideal for comprehensive land cover mapping of large areas. The dataset was accessed from USGS Earth Explorer (https://earthexplorer.usgs.gov/). Also, the ESA's World Cover LULC map was utilized as a reference dataset for comparative analysis and accessed from (https://esa-worldcover.org/en). The World Cover LULC map provides a standardized representation of land cover across various 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 used as a dimensionality reduction technique to improve the accuracy of land cover mapping with Landsat 8 OLI satellite images. PCA

transforms multispectral data into a set of uncorrelated variables known as principal components, which account for the bulk of the variance in the original dataset. By maintaining only, the most informative components and removing redundant information, PCA decreases the dataset's dimensionality, boosting computational efficiency and having a great impact on the spatial and spectral resolution [16][17][18]. In this study, PCA was derived from Landsat 8 OLI images to identify the most important spectral features for LULC classification. Following that, supervised classification algorithms like Maximum Likelihood Classifier (MLC), Support Vector Machines (SVM), and Random Forest (RF) used the principal components as input variables.

Band Indices:

In remote sensing, band indices are mathematical combinations of spectral bands from remote sensing data, such as satellite or aerial imagery. These indices can help to highlight specific features of interest, such as vegetation, water bodies, and soil types [17]. Therefore, in this study, NDVI, NDSI, and NDBI band indices are used for the extraction of 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. Features that are extracted by NDVI are shown 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).

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

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 the presence of snow and ice cover in satellite imagery. Features that are extracted by NDSI are shown in Fig.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 in Equation 2.

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

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 detect built-up areas and urban infrastructure in satellite imagery. Features that are extracted by NDBI are shown in Fig.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 in Equation 3.

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

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

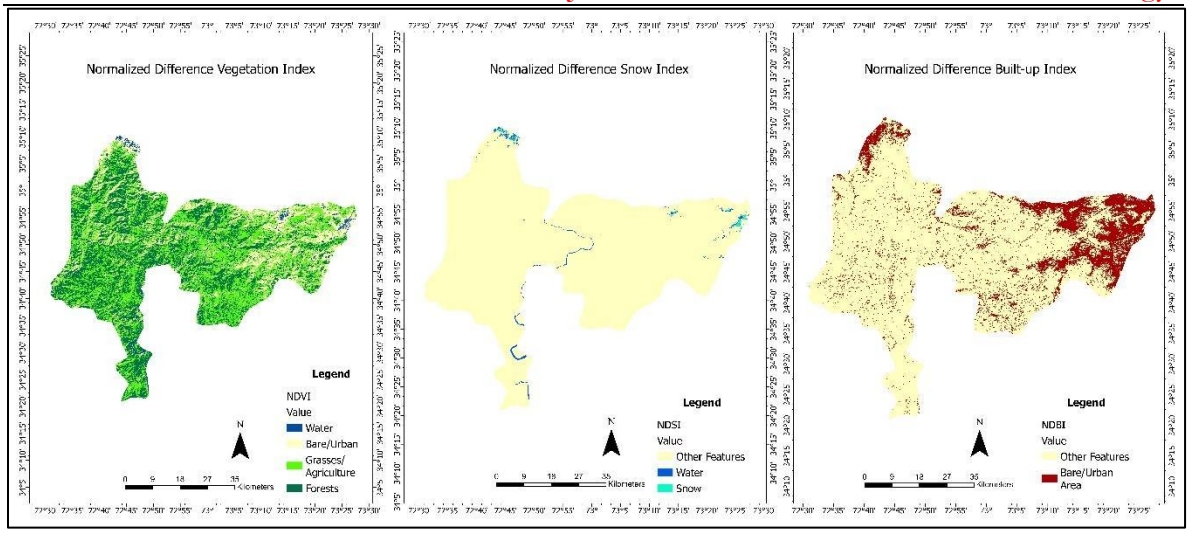


Figure 3: Layers of normalized indices that were used for classification **(a)** NDVI, **(b)** NDSI, and **(c)** NDBI

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

Methods for LULC Classification: Assigning predetermined classes to individual pixels or image segments based on their spectral characteristics is a key task in the classification procedure [19]. In the Shangla and Battagram districts of Khyber Pakhtunkhwa province, Pakistan, Landsat 8 multispectral satellite imagery, its PCA image, and stacked band ratios derived from the original image were classified using supervised classification algorithms into seven LULC classes including grasses, urban area, forest, agriculture, water, snow, and bare area as shown in Table 2. This research study used a total of 530 sample points of which 350 were used for training the algorithms and the remaining 180 samples were used for validation or accuracy assessment.

Support Vector Machine:

SVM was initially created to solve binary classification issues. It is built upon the SRM principle. It makes use of a hyper-spectral plane to classify data points [20]. The margin width is maintained at its maximum during this process by the vectors. SVM can handle many continuous and categorical variables, as well as linear and non-linear samples with different class membership levels. The surrounding samples that define the margin or hyperplane of the SVM are called support vectors. While radial basis function (RBF) and polynomial kernels have been employed more often in remote sensing, RBF is the methodology of choice for LULC classification due to its superior accuracy over other conventional techniques [5]

The original SVM approach starts with a collection of data and seeks to find the hyperplane that can divide the datasets into many classes since the goal of SVM is to locate the best-separating hyperplane among the available hyperplanes, the SVM technique also requires a good kernel function to reduce classification mistakes and effectively create hyperplanes. It matters what kind of kernel the SVM technique uses. While the higher kernel density determines the resemblance of a smooth surface, the SVM's functionality is primarily governed by the kernel size. Hyperspectral satellite data, both synthetic and real-world, yield the greatest results for the

genetically optimized SVM. The main task of SVM is to draw the optimal border that maximizes the distance between the entire set of support vectors [15].

Random Forest:

The most popular classifier for LULC classification is RF, also referred to as a collective classifier since it uses a large number of autonomous trees to carry out the classification [7]. RF uses the predictions of previously specified trees to create numerous random decision trees [21]. The remaining input data is used for objective validation, while a subset of the inputs supplied to RF is used for training. The final classification result is determined by averaging the likelihood of each class for each unique tree. Regression analysis (RA) is regarded as an extremely stable classifier that can effectively handle problems such as noise and overfitting [5]. Moreover, RF can handle high-dimensional data. Numerous factors may have an impact on the ultimate RF classification [1].

Maximum Likelihood:

The Maximum Likelihood (ML) classifier is a popular supervised classification technique in remote sensing for land cover mapping. In this study, the ML classifier was used on Landsat 8 OLI satellite images to assign each pixel to the most likely land cover class based on its spectral signature. The ML classifier assumes that pixel values within each class have a multivariate normal distribution in the feature space [22]. The ML classifier assesses 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 training data [23]. The ML classifier uses spectral information from various bands of satellite imagery to distinguish between distinct land cover types, making it especially useful for complicated landscapes with a variety of land cover classifications.

In this study, the ML classifier was used alongside other classification algorithms, such as Support Vector Machines (SVM) and Random Forest (RF), to compare their performance in accurately mapping land cover in the Shangla and Battagram districts of Khyber Pakhtunkhwa province, Pakistan. The ML classifier's capacity to describe the statistical distribution of spectral data and make probabilistic predictions makes it an important tool for land cover categorization and environmental monitoring [18].

Accuracy Assessment:

The accuracy of a classifier depends on input data, study region, and satellite sensors. Various researchers have demonstrated varying LULC accuracy for different classifiers in different study areas [14][24]. The accuracy assessments of both techniques were made through a confusion or error matrix. A confusion matrix contains information about actual and predicted classifications done by a classification system. The pixel that has been categorized from the image was compared to the same site in the field. The result of an accuracy assessment typically provides the users with an overall accuracy (OA) of the map and the accuracy for each class in the map. The percentage of overall accuracy was calculated using Equation 4.

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

Besides the overall accuracy, the classification accuracy of individual classes was calculated similarly. The two approaches are the user's accuracy (UA) and the producer's accuracy (PA). The producer's accuracy is derived by dividing the number of correct pixels in one class by the total number of pixels as derived from reference data revealed in Equation 5. In this study, the producer's accuracy measures how well a certain area has been classified.

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

Meanwhile, the user's accuracy is computed by dividing the number of correctly classified pixels in each category by the total number of pixels that were classified in that category. User's accuracy refers to the likelihood of a pixel falling into the forecasted class as shown in Equation 6 [25].

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

Kappa Coefficient:

Kappa coefficient (Kc) is another measurement used in many studies [24][26][27]. It is calculated by multiplying the total number of pixels in all the ground verification classes (N) by the sum of the confusion matrix diagonals (Xkk), subtracting the sum of the ground verification pixels in a class time the sum of the classified pixels in that class summed over all classes (ΣXkΣ YkΣ), where XkΣ is row total and YkΣ is column total, and dividing by the total number of pixels squared minus the sum of the ground verification pixels in that class times the sum of the classified pixels in that class summed over all classes. The value of Kappa lies between 0 and 1, where 0 represents agreement due to chance only. 1 represents complete agreement between the two data sets. Negative values can occur but they are spurious. It is usually expressed as a percentage (%). Kappa statistic can be a more sophisticated measurement of classifier agreement [28] and thus gives better interclass discrimination than overall accuracy. The kappa coefficient is calculated by using Equation 7.

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

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 examined the seven different types of LULC (water bodies, forest, urban area, agriculture, snow, grasslands, and bare area) of Battagram and Shangla districts for the year 2021 using Landsat-8 imagery and its derivatives (PCA and Band ratios). The study utilized three commonly used supervised learning algorithms (MLC, RF, SVM) and every algorithm has produced results slightly different from others. Fig. 4 shows the spatial distribution of all LULC classes done by three machine learning algorithms using Landsat-8 multispectral, PCA, and Stacked Indices imagery.

Classification Using Multispectral Bands of Landsat 8 Image:

Results indicated that forest cover is the most dominant LC of Shangla and Battagram City as shown in Figure 5a. By using the Landsat 8 image out of 305,464 ha the SVM classifier classified 2,145 ha as Water bodies, 12,563 ha as Urban Areas, 109,329 ha as Forest, 1828 ha as Snow, 48,676 ha as Bare area, 90,906 ha as Grasslands and 48,021 ha as Agricultural land. 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 area, 87,197 ha as Grasslands and 46,675 ha as Agricultural land and the Maximum Likelihood classifier classified 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 area, 90,906 ha as Grasslands and 48,021 ha as Agricultural land.

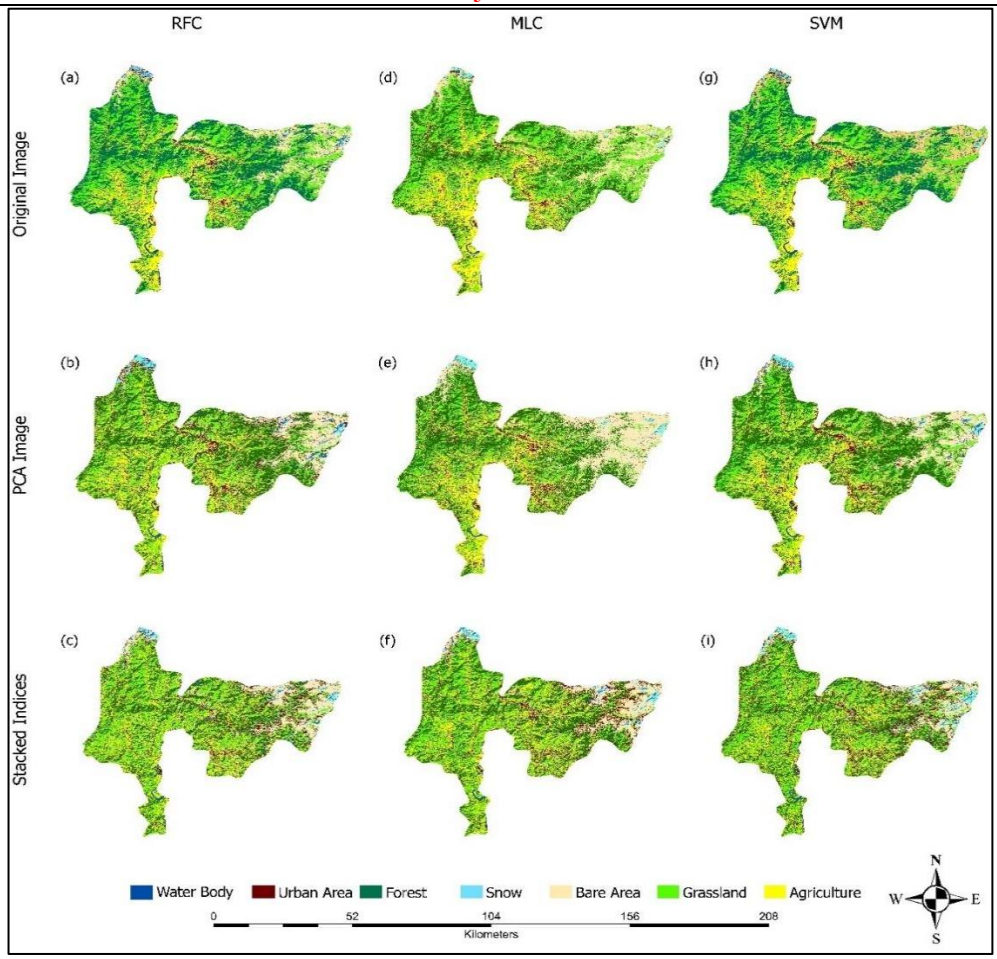
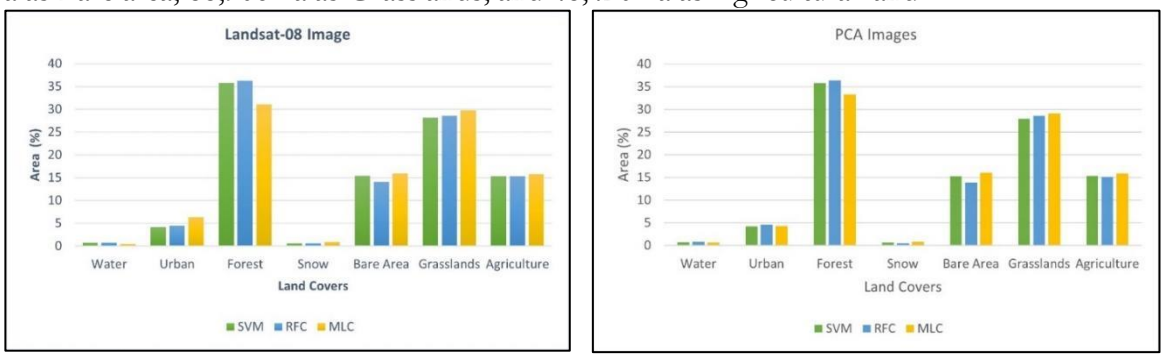


Figure 4: Spatial Distribution of LULC classification by RF, SVM, and MLC by using Landsat 08, PCA, Stacked Indices

Classification Using PCA Image:

Results using the PCA images are shown in Figure 4. Figure 5b shows the area comparison of LULC of the PCA image. From the classification of PCA image out of 305,464 ha, the SVM classifier classified 2,243 ha as Water bodies, 12,994 ha as Urban Areas, 109,329 ha as Forest, 1928 ha as Snow, 46,696 ha as Bare area, 85,298 ha as Grasslands, 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 area, 87,297 ha as Grasslands, 48,420 ha as Agricultural land, and Maximum Likelihood classifier classified 1,971 ha as Water bodies, 13,104 ha as Urban Areas, 101,702 ha as Forest, 2,385 ha as Snow, 48976 ha as Bare area, 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 shown in Figure 4. From the classification of the image out of 305,464 ha, the SVM classifier classified 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 area, 91,258 ha as Grasslands, 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 area, 92,262 ha as Grasslands and 39,780 ha as Agricultural land and the Maximum Likelihood classifier classified 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 area, 89,306 ha as Grasslands and 41,720 ha as Agricultural land as shown in Figure 6.

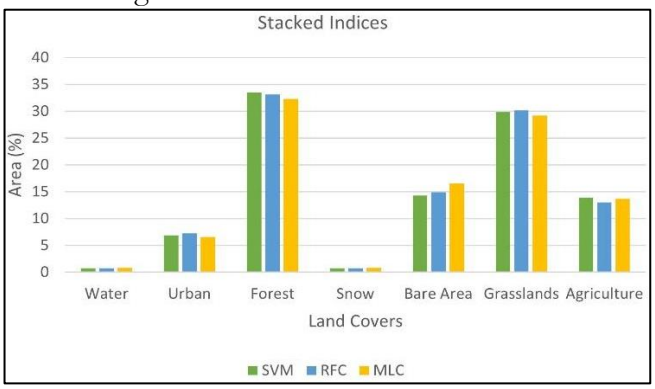


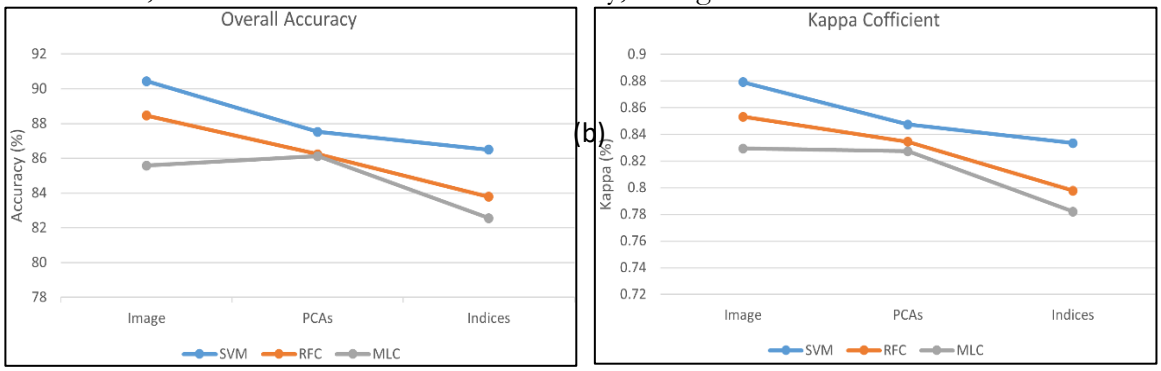
Figure 6: Comparison of the area of LC classes using Stacked Indices Image

This research study evaluates the performance of classification algorithms using multispectral images, PCAs, and normalized Indices, therefore, the area statistics might not align with national estimates for some land cover land use classes.

Comparison of Classifiers:

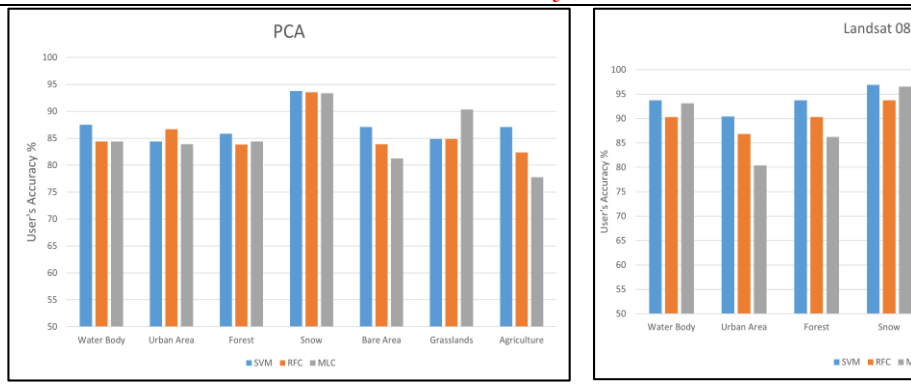
To compare their performance, three different classifiers were applied to three different images (Landsat 8, PCA, and stacking Indices including NDVI, NDSI, and NDBI). Overall accuracy and the Kappa coefficient are the most often used measures for determining the accuracy of any classifier for classification. Figure 7a shows the overall accuracy, whereas Figure 7b shows the Kappa coefficients of each classifier for Landsat-8, PCA, and stacked indices.

According to the total accuracy and kappa coefficient statistics, the SVM classifier outperformed both classifiers in all three images. The SVM classifier had a maximum overall accuracy of 90.43% for the Landsat-8 image, while the RFC classifier achieved 87.52% overall accuracy and 84.5% accuracy for MLC from Landsat-08, placing it ahead of both classifiers. Furthermore, MLC had the lowest overall accuracy, falling below both classifiers.



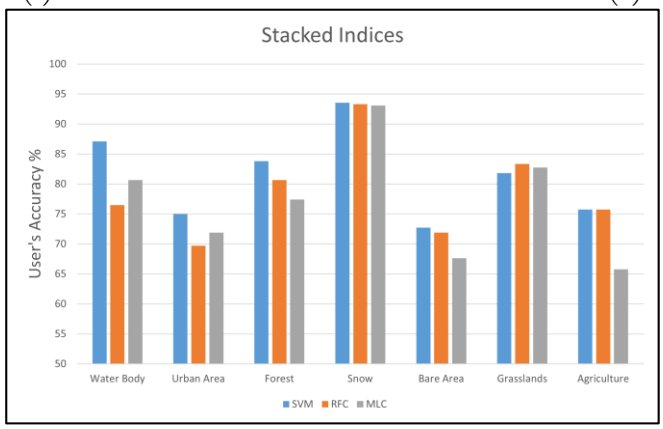
(a) (b)

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)

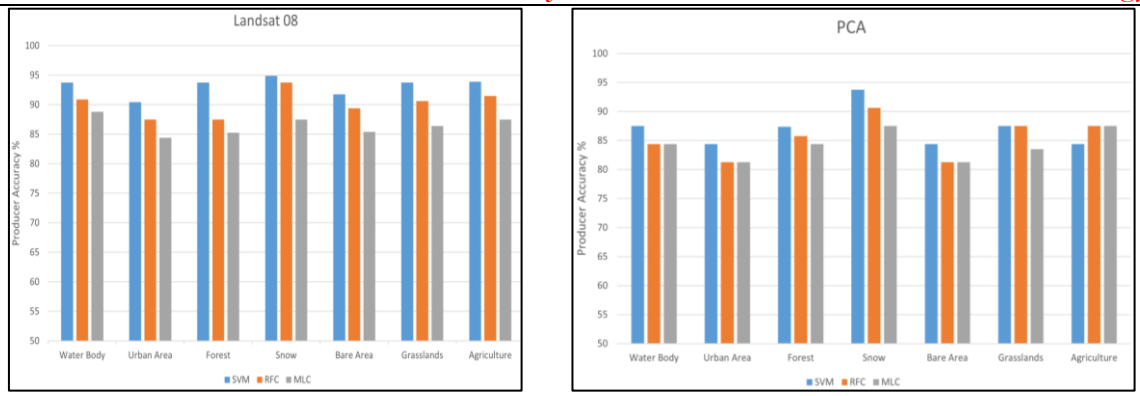


(c)

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

In the case of the Landsat satellite's two derived datasets, the PCA from Landsat-8 performed best for all three classification methods. PCA yielded an overall accuracy of 86.12% for the MLC classifier, 86.23% for the RFC, and 87.52% for the SVM classifier. Likewise, Landsat-8-derived stacked indices (NDVI, NDSI, and NDBI) performed the least well among the three datasets. Overall accuracy was 82.56% for the MLC classifier, 83.78% for the RFC, and 86.5% for the SVM classifier.

This study evaluated the performance of SVM, MLC, and RF models on three data products for each LULC class using UA and PA metrics. Results for each class were described separately due to differences in model performance across different classes. Figure 8 displays the user's accuracy for each classifier using Landsat-8, its PCA, and Derived indices imagery. In terms of class accuracy, other than the Stacked Indices image the users were most accurate with Snow cover and least accurate with Bare Area. Compared to other sensors, the Landsat-8 image outperformed the PCA and Stacked Indices image. The SVM classifier outperformed other classifiers. Furthermore, similar findings were obtained for producer accuracy. SVM outperformed other classifiers and snow cover had the highest producer accuracy in all the images. The results for producer accuracy are displayed in Fig. 9, 10. For the stacked indices image both user and producer accuracies are lower than other images also shown in Fig.8c and Fig.10 because of some misclassification (Bare and urban area) due to the NDBI index that is used for extracting urban area than others so this is the main reason for the least accuracy in 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

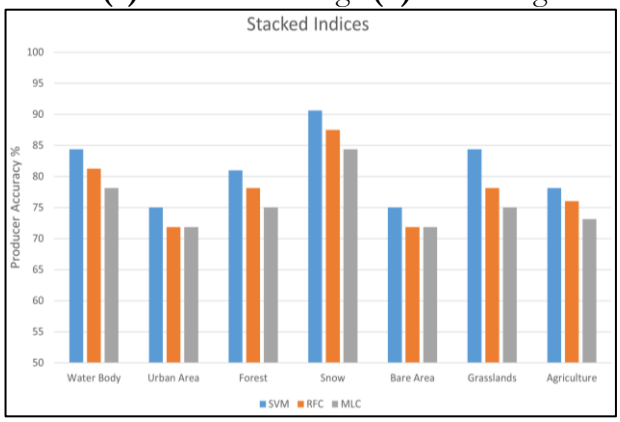


Figure 10: Producer's Accuracy of land cover classes using SVM, RFC, and MLC classifiers on Stacked Indices.

ESA Landcover Validation:

The ESA landcover map is validated with the same validation points and it gets 70.16% overall accuracy. Several factors contribute to this moderate accuracy. Firstly, the temporal character of the datasets contributed to their poorer accuracy. The ESA land cover map is an annual product and is validated by using data, dated 2021/10/06. As a result, the intrinsic unpredictability and dynamic changes in land cover across the year are not reflected, potentially leading to class mixing and temporal mismatches. Second, there is a resolution difference between the datasets; our validation points were made based on 30-meter resolution, but the ESA land cover map has a finer resolution of 10 meters These characteristics illustrate the difficulties in achieving improved validation accuracy when dealing with datasets of varying resolutions and temporal spans.

Discussion:

The study examined Landsat 08 multispectral and its derivatives (PCA, Indices) datasets to evaluate the accuracy of machine learning classifiers (MLC, SVM, and RF) for LULC classification. As discussed earlier, the current study computed the UA, PA, OA, and Kc were used to evaluate the accuracy of LULC mapping using the implemented algorithm. The results revealed that the SVM classifier outperformed RF and MLC algorithms in three datasets for all land cover classifications. These findings are consistent with previous research [4][6][19]. The SVM model accurately classified land cover captured complicated linkages, and made robust predictions. The SVM produced higher OA and Kappa indices in all the datasets than the RF as shown in Figure 7. [15] found that the stability of the RF classifier is influenced by the number of trees, bagging, and random impressions, which affect its efficiency and accuracy. The study found that MLC had lower accuracy than RF and SVM which is also shown in Figure 7. Research

suggests that RF and SVM classifiers perform better because of their ability to tolerate noise [29] and are more resilient to random and systematic noise in training data [9][30] and their performance also depends on the dimensionality of the data [5].

The performance of the approach is influenced by both the machine learning model and the data used. Implementing PCA images enhanced classification accuracy for some classes while introducing some discrepancies. PCA transforms improved classification accuracy for some classes in the case of MLC as shown in Figure 7 but misclassifies others. SVM is computationally effective and efficient in high-dimensional areas with clear class separations. However, it is challenging to utilize, understand, and fine-tune, and requires lengthy training for large datasets.

This study used machine learning techniques to classify LULC in the Battagram and Shangla districts, determining the best algorithm for these regions. The user and producer accuracy results revealed RF and MLC produced the lowest accuracy values in the bare and urban areas especially in the Stacked indices dataset because of NDBI. Both districts have unplanned urbanization, with few high-rise constructions in the urban centers. This resulted in minor misclassifications between these classes, affecting their accuracy. The study found that all three models performed well with reliable closed-related accuracy in detecting distinct classes of LULC during classification and Snow is the most accurate class for LULC classification. Several studies show that different classifiers perform differently when mapping LULC classes for the same data and study year [31][32] observed a shift in the outcomes of machine learning models used for LULC mapping in the study area.

Conclusion:

LULC mapping plays an important role in sustainable development and economic resilience. Machine learning algorithms better than other strategies for LULC classification in terms of reliability and accuracy. Understanding the parameters that affect the algorithm's accuracy can enhance classification results. This study aims to evaluate the best machine learning algorithm among SVM, RF, and MLC for Landsat-08 multispectral image, its PCA image, and Stacked normalized Indices image. The results demonstrate that machine learning algorithms can effectively analyze Landsat data and its derivative PCA and band indices for LULC analysis. The study's major findings and conclusion are stated below. SVM is the top-performing machine learning algorithm from the others in all the images, with 90.43% accuracy for Landsat-8 image, 87.52% for PCA, and 86.5% for band indices imagery. In comparison, RF is second with 88.46% accuracy for Landsat-8 image, 86.23% for PCA, and 83.78% for band indices imagery. MLC has less accuracy than others which is 85.58% for Landsat 08, 86.12% for PCA, and 82.56% for Indices imagery. The Landsat-8 original image had the highest overall accuracy (90.43% for SVM, 88.46% for RF, and 85.58% for MLC). The ESA Landcover produces 70.16% accuracy by using the same validation points as other LULC maps.

LULC classification results vary depending on the satellite, research region, and algorithms. Studies indicate that classifiers work differently in different climate zones and geographical areas. To better understand LULC, researchers should use high-resolution imaging, specific classes, field data, and advanced machine learning methods for classification. This study may guide policymakers, service planning, management practices, and LULC classification in mountainous cities. Changes in study area or duration can impact the accuracy of machine learning algorithms, while changes in adopted algorithms or datasets could end up in different results.

Acknowledgments:

We are thankful for satellite-derived data products used in our study: Landsat 08 product managed by NASA and the U.S. Geological Survey and WorldCover by the European Space Agency—their efforts to make these beneficial resources public have substantially impacted this study.

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.