

Separability Analysis for Wetland Mapping using Segmented Principal Component Analysis in Pulicat Lake, India

Abhi Roop, M.* and Gowri Sankar Reddy, D.

Department of Electronics and Communication Engineering, Sri Venkateswara University College of Engineering, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

E-mail: abhiroop.mule@gmail.com*

*Corresponding Author

DOI: <https://doi.org/10.52939/ijg.v21i12.4645>

Abstract

Lakes and coastal lagoons are ecologically sensitive systems that require continuous monitoring to understand the impacts of environmental change and human activities. Land cover classification plays a critical role in understanding the spatial distribution and temporal dynamics of key landscape elements, thereby aiding in effective management and conservation strategies for these sensitive regions. This study analyzes land cover dynamics in the Pulicat Lake region using multi-temporal Landsat 8 imagery from 2014, 2019, and 2024. Principal Component Analysis (PCA) and Segmented PCA (SPCA) were integrated with Support Vector Machine (SVM) and Random Forest (RF) classifiers to improve land cover mapping. Four major land cover types, namely Water, Vegetation, Bareland, and Wetland, were mapped and analysed. Accuracy assessment confirmed that SPCA-based classifications consistently outperformed PCA, achieving the highest overall accuracy of 95.83%, 94.47% and 96.79% with RF for the years 2014, 2019 and 2024 respectively. The land cover change assessment revealed notable shifts in water, vegetation, bareland, and wetland extents over the study period, indicating ongoing ecological transitions in the lagoon system. The findings highlights the effectiveness of combining SPCA with Random Forest algorithm for accurate and efficient land cover monitoring and supports informed decision-making in the sustainable management of lake ecosystems.

Keywords: Change Detection, Land Cover, PCA, Remote Sensing, RF Algorithm, Segmented PCA, SVM Classifier

1. Introduction

Lakes and their surrounding landscapes are highly dynamic ecosystems that play a crucial role in maintaining regional ecological balance. However, these areas are increasingly subjected to anthropogenic pressures, particularly from rapid urban development and land use transformation. Change detection of land cover in lake regions has become an essential approach for understanding the extent and impact of such changes over time [1].

Urban expansion near lake environments often leads to the conversion of natural landscapes such as wetlands, forests, and agricultural lands into impervious surfaces like roads, buildings, and industrial zones. These changes can disrupt hydrological cycles, degrade water quality, increase sedimentation, and threaten the biodiversity of aquatic and riparian habitats [2]. Monitoring land cover change helps in assessing the ecological

consequences of urbanization and in formulating sustainable land and water management strategies. This study focuses on Pulicat Lake, the second-largest brackish water lagoon in India, located along the southeastern coast bordering Andhra Pradesh and Tamil Nadu. Pulicat Lake is an ecologically vital habitat supporting a diverse array of flora and fauna, including migratory birds, mangroves, and aquatic life, and plays a significant role in fisheries, groundwater recharge, and climate regulation [3].

In recent decades, Pulicat Lake has experienced extensive ecological degradation due to urban expansion, industrialization, and environmental changes, including reduced freshwater inflow, increased salinity, pollution, and encroachment. These factors have led to significant alterations in land cover, such as the changes of water bodies, wetland degradation, and the rise of built-up areas [4].

In this context, land cover classification plays a crucial role in assessing ecosystem health, as it enables the identification and monitoring of spatial and temporal patterns of landscape change. The main objective of this study is to classify the major land cover types, assess the performance of Principal component Analysis (PCA) and Segmented PCA in the classification and perform change detection analysis in and around Pulicat Lake using multi temporal Landsat 8 imagery, thereby providing insights for the sustainable management of this vulnerable ecosystem [5] and [6]. Landsat data was specifically selected for this study owing to its moderate spatial resolution, open-access availability, and inclusion of thermal bands, which are essential for monitoring land surface temperature and supporting land cover classification and change detection analysis [7]. Recent studies have demonstrated that segmented or region-based PCA approaches can enhance spectral separability for land cover mapping, particularly in heterogeneous landscapes (e.g., wetland and mixed vegetation systems). Several remote sensing studies related to Segmented PCA approach reported improved separability of spectrally overlapping classes, especially in comparison to global PCA approaches for land cover mapping in medium resolution Landsat 8 imagery [8].

Despite extensive research on PCA-based feature extraction, limited studies have explored the use of Segmented PCA (SPCA) for LULC classification in complex coastal ecosystems. SPCA applies PCA on specific spectral band combinations, enabling improved representation of localized spectral variability that traditional PCA tends to suppress [9]. While SPCA has shown promising results in hyperspectral and agricultural applications, its utility for wetland and lagoon environments, where subtle spectral differences between water, wetland, vegetation, and bareland are critical which remains underexplored.

This study contributes in analyzing the land cover dynamics and change detection of Pulicat Lake by integration of remote sensing data with machine learning algorithms. Firstly, to enhance feature extraction from Landsat 8 imagery Principal component Analysis (PCA) and Segmented PCA are applied [10][11] and [12]. This approach is particularly useful in regions like Pulicat Lake, where water, wetland, vegetation, and bareland exhibit overlapping spectral signatures. Secondly, machine learning classifiers such as Support Vector Machine (SVM) and Random Forest (RF) algorithms are integrated with the dimensionality reduction techniques to evaluate the land cover mapping and change detection [13][14] and [15]. Finally, this study

produces land cover maps for 2014, 2019 and 2024 with a comparative analysis and performance of PCA and SPCA based classifications facilitating an assessment of long-term spatial and ecological dynamics. These contributions collectively underscore the potential of advanced remote sensing and machine learning techniques to inform sustainable monitoring and management of lake ecosystems.

2. Materials and Methods

2.1 Study Area

The present study focuses on Pulicat Lake, a prominent brackish water lagoon located along the southeastern coast of India, straddling the border of Tamil Nadu and Andhra Pradesh. Pulicat Lake is the second-largest brackish water lagoon in India, covering an approximate area of 759 km², and plays a critical role in supporting regional biodiversity, fisheries, and local livelihoods. The lake exhibits low relief conditions with an elevation mostly less than 10 m, gentle slopes toward the sea, and extensive mudflats that influence tidal flow patterns. The hydrology of Pulicat is mainly governed by monsoonal freshwater inflow from the Arani and Kalangi rivers and tidal exchange with the Bay of Bengal through the Pulicat mouth. Surrounding land use comprises fishing villages, salt pans, aquaculture ponds, mangroves, agriculture, and expanding settlement areas, all of which contribute to heterogeneity in land cover mapping. The lake lies between latitudes 13°25'N to 13°55'N and longitudes 80°03'E to 80°18'E, and is directly influenced by both marine and freshwater inputs, making it highly sensitive to environmental and anthropogenic changes. The inset image of ROI and the Landsat 8 layer stacked image are depicted in Figure 1.

2.2 Methodology

The methodology adopted in this study follows a structured multi-stage workflow designed to assess land cover dynamics, evaluate the performance of PCA and SPCA-based classification, and conduct change detection for the Pulicat Lake region using Landsat 8 imagery. The process begins with the acquisition of multi-temporal Landsat 8 imagery for the years 2014, 2019, 2024 and preprocessing the data. Dimensionality reduction techniques, PCA and SPCA were then applied to enhance feature separability. SPCA was performed using specific band combinations to capture localized spectral variability. The resulting principal components served as inputs for supervised land cover classification using machine learning algorithms. Accuracy assessment was carried out using random sampling method to evaluate the performance of each

classification approach. Finally, land cover maps from the three time periods were compared to perform change detection analysis for the period 2014–2024. The overall workflow is illustrated in Figure 2.

2.2.1 Satellite image acquisition

For this study, multi-temporal Landsat 8 Operational Land Imager (OLI) Level-2 surface reflectance data were employed to analyse land cover changes in the Pulicat Lake region. Images were selected for the years 2014 (02-04-2014), 2019 (15-03-2019), and

2024 (12-03-2024), corresponding to Path 142/Row 51, with priority given to scenes containing less than 10% cloud cover and acquired during comparable seasonal periods to ensure consistency in temporal analysis. The datasets, downloaded from the USGS Earth Explorer platform [16], were provided in *GeoTIFF* format with a spatial resolution of 30 m. Each scene comprises 11 spectral bands, enabling detailed land cover classification, and was projected to the Universal Transverse Mercator (UTM) coordinate system using the WGS 84 datum (EPSG:32644).

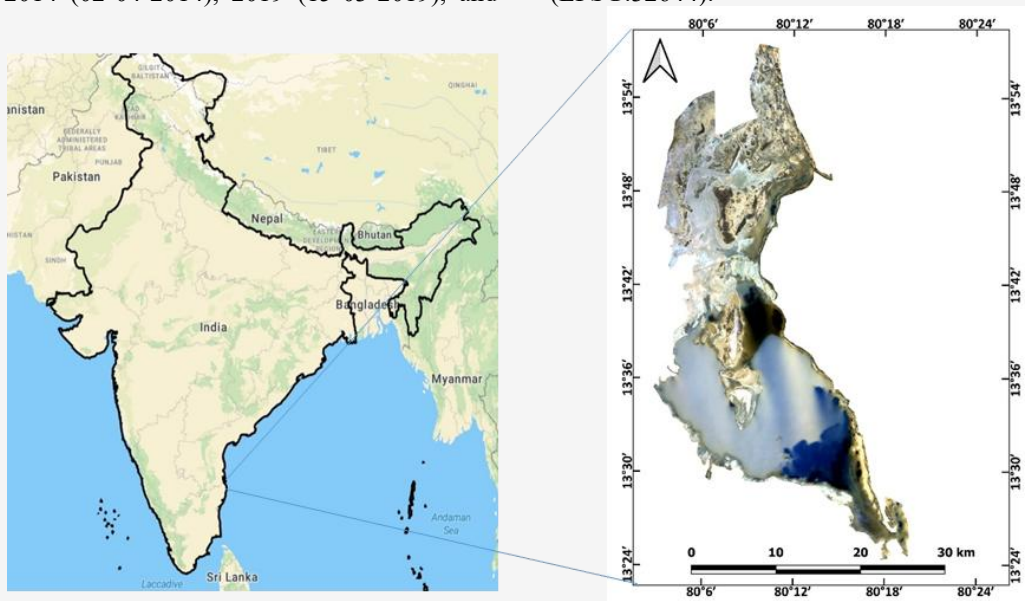


Figure 1: Pulicat lake, India

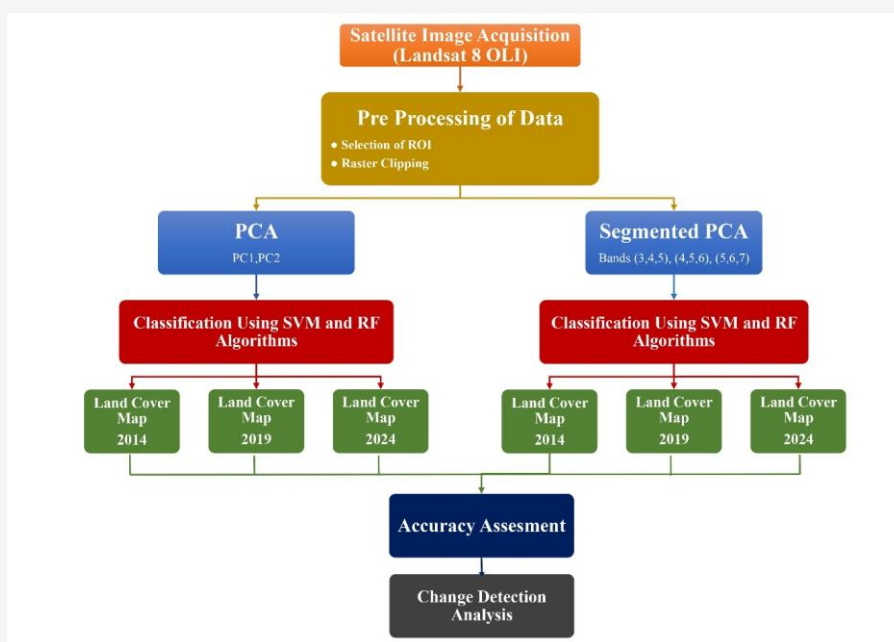


Figure 2: Workflow for land cover classification and change detection analysis

2.2.2 Preprocessing of data

The Landsat -8 Level - 2 Surface Reflectance products were used in the study. These products are already corrected for atmospheric effects including aerosol scattering, Rayleigh scattering, and water-vapor absorption. They also incorporate solar angle correction during the atmospheric correction process, providing surface reflectance values that are directly suitable for further analysis.

Selection of Region of Interest (ROI): The Pulicat Lake boundary and surrounding zones were delineated using shapefiles and Raster Clipping was performed for all Landsat data and the images were clipped to the ROI to eliminate extraneous data and improve processing efficiency.

2.2.3 Application of PCA and Segmented PCA

To improve classification performance, two feature extraction approaches were implemented:

PCA was applied to the full Landsat 8 multispectral stack to generate orthogonal components capturing global variance in the dataset. The first two components (PC1 and PC2) were evaluated based on explained variance and separability. Although PC1 retained the highest variance, it tends to often suppress local spectral differences between wetland and shallow water areas. To overcome this limitation of global PCA, SPCA was implemented by applying PCA separately to three spectral band combinations: Bands 3-4-5, 4-5-6, and 5-6-7. These combinations were selected based on evidence that the visible-NIR-SWIR transitions enhance discrimination of water, vegetation, and wetland features. For each segment:

- PC1 from 3, 4, 5 (Green, Red, NIR) captured dominant variance in vegetation and shallow water
- PC1 from 4-5-6 (Red, NIR, SWIR-1) highlighted moisture and soil-vegetation transitions
- PC2 from 5-6-7 (NIR, SWIR-1, SWIR-2) retained subtle variance associated with wetland - bareland boundaries

These PCs were selected based on their higher localized explained variance and superior separability scores between the classes compared to other components. These selected principal components were stacked to form a composite PCA image representing key localized spectral information. This optimized feature space was then used as input to supervised classification. This approach enhances the classification of spectrally complex regions such as wetlands and vegetation

mosaics by retaining subtle spectral differences that are often lost in global PCA transformations.

2.2.4 Land cover classification

Land cover classification was performed on both the PCA-derived and Segmented PCA-derived feature datasets for the years 2014, 2019, and 2024. The classification focused on four major land cover classes that are ecologically and hydrologically significant in the Pulicat Lake region: Water, Vegetation, Bareland and Wetland. These classes were selected based on their environmental importance, spectral distinctiveness, and the historical land use changes observed in the study area. For each year and each PCA method (traditional and segmented), supervised classification was performed using machine learning algorithms:

- Support Vector Machine (SVM)
- Random Forest (RF)

The classifications resulted in twelve land cover maps in total, six from standard PCA and six from segmented PCA for the years 2014, 2019, and 2024.

2.2.5 Accuracy assessment

To evaluate the performance of the land cover classification results, an accuracy assessment was conducted using the random stratified sampling method. This approach ensures a balanced and representative selection of validation points across all classified land cover classes. For each classified image derived from both PCA and segmented PCA approaches for the years 2014, 2019, and 2024, a set of 50 random samples were generated for each class in proportion to its spatial extent, ensuring adequate sampling even in minority classes. The generated sample points were then visually interpreted and validated using high-resolution Google Satellite Imagery, serving as an up-to-date and spatially accurate reference. Reference raster datasets and auxiliary land use maps were also used to support validation, particularly in ambiguous areas such as wetland-vegetation transitions or seasonal water bodies.

The classification accuracy was quantified using standard statistical metrics derived from the confusion matrix, including Overall Accuracy (OA), User's Accuracy (UA), Producer's Accuracy (PA), and Kappa Coefficient.

2.2.6 Change detection analysis

Post-classification comparison was carried out using the classified maps to assess temporal land cover changes across the three selected years. The comparative analysis between PCA and Segmented PCA-based classifications provided insight into the effectiveness of each method in capturing landscape

transitions and distinguishing between complex class boundaries.

3. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a statistical technique widely applied in remote sensing for dimensionality reduction, feature extraction, and data decorrelation. It transforms a set of correlated spectral bands into a new set of uncorrelated variables known as principal components (PCs), which are linear combinations of the original variables arranged in descending order of variance contribution. The first few components typically capture the majority of the spectral variance within the dataset, enabling the reduction of input dimensionality with minimal information loss.

Figure 3 illustrates the distribution of data points in a two-dimensional space illustrating Principal Component Analysis (PCA).

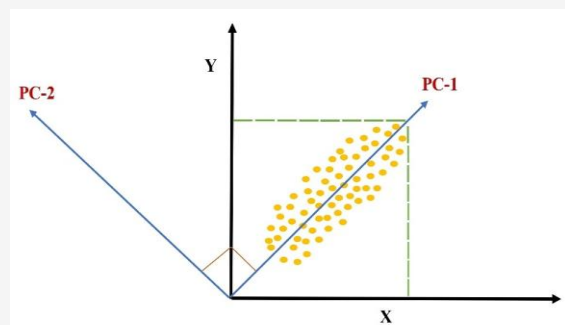


Figure 3: Principal Component Analysis (PCA) for a two dimensional data set

The primary principal component (PC-1) captures the direction of maximum variance in the dataset. The secondary component (PC-2) is orthogonal to PC-1 and represents the next highest variance. PCA aids in transforming correlated variables into a set of uncorrelated components through a sequence of optimized linear combinations, thereby facilitating dimensionality reduction and enhanced data interpretation. In the context of satellite image analysis, particularly with data acquired from the Landsat series, PCA serves as an effective preprocessing step prior to classification. The application of PCA reduces this redundancy and enhances the spectral distinction between land cover types. The benefits of using PCA in classification workflows are Dimensionality Reduction, Noise Suppression and Enhanced Feature Separability. Figure 4 illustrates the concept of spectral class separability using Principal Component Analysis (PCA). This distance serves as a quantitative measure of spectral separability between the two classes.

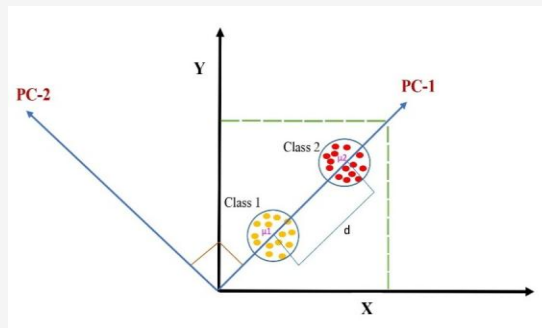


Figure 4: Spectral separability of land cover classes in PCA space using Euclidean distance

Two distinct land cover classes, Class 1 (yellow) and Class 2 (red), are projected into a new coordinate system defined by the first (PC-1) and second (PC-2) principal components, which capture the maximum variance in the dataset along orthogonal axes.

Each class is represented by a cluster of data points in this PC space, and the centroids of these clusters are denoted by $\mu_1(x_1, y_1)$ and $\mu_2(x_2, y_2)$ respectively. The Euclidean distance (d) between these centroids is calculated using the formula represented in Equation (1):

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Equation 1

A larger Euclidean distance indicates better discrimination and reduced overlap, thereby enhancing the effectiveness of the Random Forest classifier. The strategic selection of principal components that maximize this distance improves the overall accuracy of land cover classification.

3.1 Segmented Principal Component Analysis (SPCA)

Segmented Principal Component Analysis (SPCA) is an extension of the conventional PCA technique, developed to improve classification performance in heterogeneous landscapes. Traditional PCA reduces spectral redundancy and enhances feature separability by transforming multispectral data into a new set of orthogonal components that maximize global variance. However, because PCA applies a single global transformation across all spectral bands, it tends to emphasize dominant spectral patterns while suppressing subtle local variations. This limitation reduces its effectiveness in complex environments such as lakes, where classes like water, vegetation, bareland, and wetlands often display overlapping spectral responses. As a result, global PCA may fail to adequately capture localized spectral differences, leading to reduced class separability for spectrally similar features such as

wetlands, shallow waters, and mixed land covers in coastal zones. SPCA addresses these shortcomings by dividing the dataset into specific band combinations before applying PCA. This localized approach captures fine-scale spectral variability and enhances the discriminability of spectrally similar classes, which are often masked in a global PCA dominated by larger homogeneous areas [8]. By focusing on carefully selected spectral ranges, SPCA improves the representation of localized spectral differences and increases classification accuracy in heterogeneous ecosystems.

Table 1: Eigen vector matrices
SPCA of bands 3,4,5

	PC 1	PC 2	PC 3
Band 3	0.110	0.608	0.785
Band 4	0.277	0.740	-0.612
Band 5	0.954	-0.285	0.087

SPCA of bands 4,5,6

	PC 1	PC 2	PC 3
Band 4	-0.165	-0.801	-0.574
Band 5	-0.561	0.556	-0.613
Band 6	-0.811	-0.220	0.541

SPCA of bands 5,6,7

	PC 1	PC 2	PC 3
Band 5	-0.467	-0.840	0.273
Band 6	-0.692	0.155	-0.704
Band 7	-0.549	0.518	0.654

In this study, SPCA was applied to three Landsat 8 band combinations: 3-4-5, 4-5-6, and 5-6-7. From these, the most informative components, PC1 from 3-4-5, PC1 from 4-5-6, and PC2 from 5-6-7, were extracted and used as input features for supervised classification. PC2 from the 5-6-7 combination was selected because it exhibited the maximum Euclidean distance between Wetland and Bareland classes, indicating its capability to capture subtle moisture-soil transitions that PC1 failed to represent (Table 1). These components provided enhanced spectral variance tailored to the internal heterogeneity of the Pulicat Lake region, thereby improving the separability of water, vegetation, bareland, and wetland classes.

4. Land Cover Classification Using Machine Learning Algorithms

4.1 Support Vector Machine (SVM) Classifier

The Support Vector Machine (SVM) is a supervised machine learning algorithm widely used for classification tasks in remote sensing. It operates by identifying the optimal hyperplane that maximally separates data points of different classes in a high-

dimensional feature space. To handle non-linear relationships between features, SVM employs kernel functions, such as the Radial Basis Function (RBF), polynomial, or sigmoid kernels, which enable the transformation of input data into a higher-dimensional space where linear separation becomes feasible [17]. In this study, SVM was utilized to classify land cover types in the Pulicat Lake region based on two distinct feature extraction methods: Principal Component Analysis (PCA) and Segmented Principal Component Analysis (SPCA). These techniques were employed to enhance spectral separability and reduce data dimensionality prior to classification. The SVM model was configured using a regularization parameter (C) set to 1.0000, which provided an effective balance between model complexity and classification accuracy for the heterogeneous land cover conditions in the study area.

4.2 Random Forest (RF) Algorithm

Random Forest (RF) is an ensemble-based supervised classification algorithm that constructs multiple decision trees during the training phase. Each tree is built using a randomly selected subset of the input features and training samples, introducing diversity into the model. The final class prediction is determined through majority voting among all the individual trees, which helps reduce the risk of overfitting and enhances the model's generalization capability [18]. In the present study, the RF classifier was applied to both PCA-derived and Segmented PCA (SPCA)-derived feature sets for land cover classification of the Pulicat Lake region. The classification focused on key land cover types, including Water, Vegetation, Bareland, and Wetland. The number of trees was set to 10, which was found to be optimal for achieving a balance between computational efficiency and classification accuracy in the context of the study area. The supervised classification was carried out using QGIS software with the Semi – Automatic Classification (SCP) plugin [19] and [20].

Using this approach, land cover maps for the years 2014, 2019, and 2024 were generated based on the classification outputs. These maps served as the foundation for subsequent change detection analysis, enabling the identification and assessment of temporal transitions and land cover dynamics across the Pulicat Lake ecosystem over the past decade.

5. Results

This section presents the classification outputs, accuracy assessment, and land cover change analysis for the Pulicat Lake region using PCA and SPCA in combination with SVM and RF classifiers. The emphasis is placed on comparing the performance of

PCA- and SPCA-derived feature sets, identifying dominant land cover transitions, and interpreting ecological and anthropogenic drivers shaping the observed changes. Figures 5 to 7 present the land cover classification maps derived from PCA-transformed Landsat data using Support Vector Machine (SVM) and Random Forest (RF) classifiers for the years 2014, 2019, and 2024. The corresponding land cover area distribution charts are illustrated in Figure 8. Both classification strategies categorize the Pulicat Lake region into four dominant

land cover types: Water, Vegetation, Bareland, and Wetland. The land cover classification maps of the Pulicat Lake region for the years 2014, 2019, and 2024, generated using the Segmented Principal Component Analysis (SPCA) approach integrated with Support Vector Machine (SVM) and Random Forest (RF) classifiers is shown in the Figures 9 to 11 respectively. Figure 12 illustrates the areas of land cover types determined from the classification models.

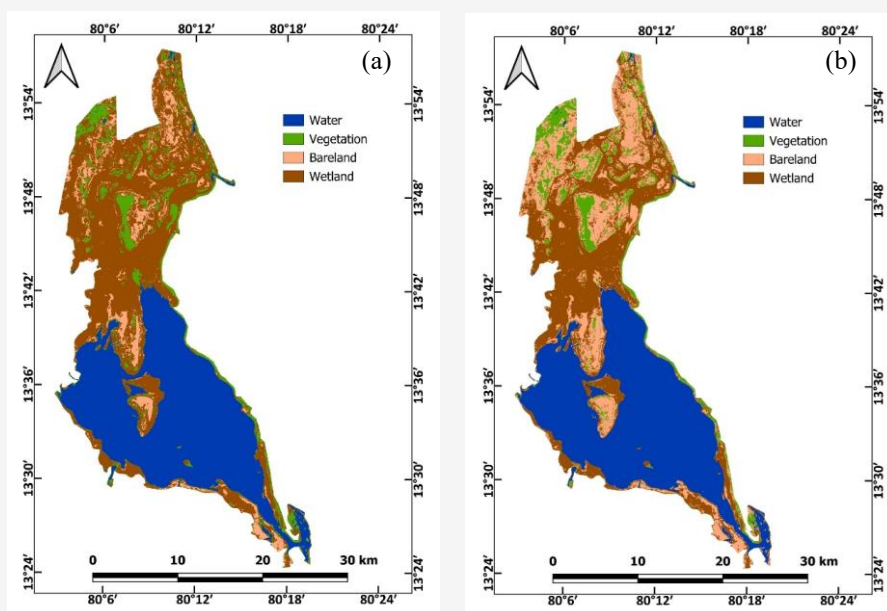


Figure 5: Land cover classification of 2014: (a) PCA/SVM, and (b) PCA/RF

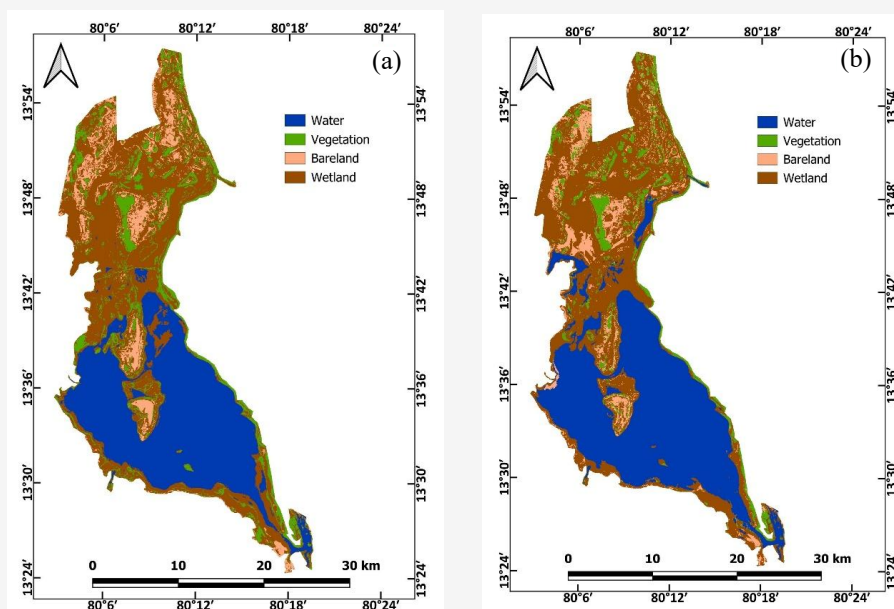


Figure 6: Land cover classification of 2019 (a) PCA/SVM, and (b) PCA/RF

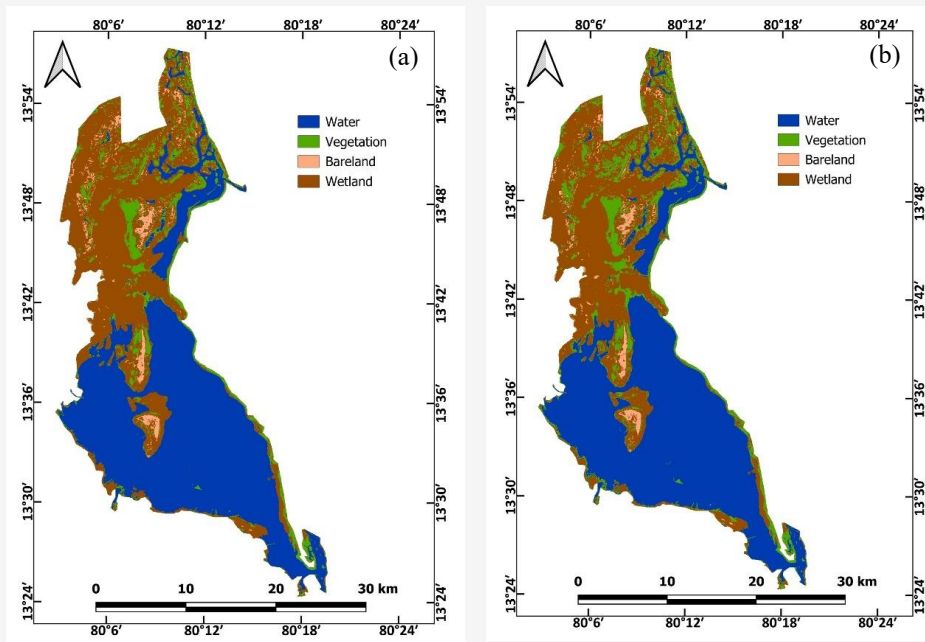


Figure 7: Land cover classification of 2024 (a) PCA/SVM, and (b) PCA/RF

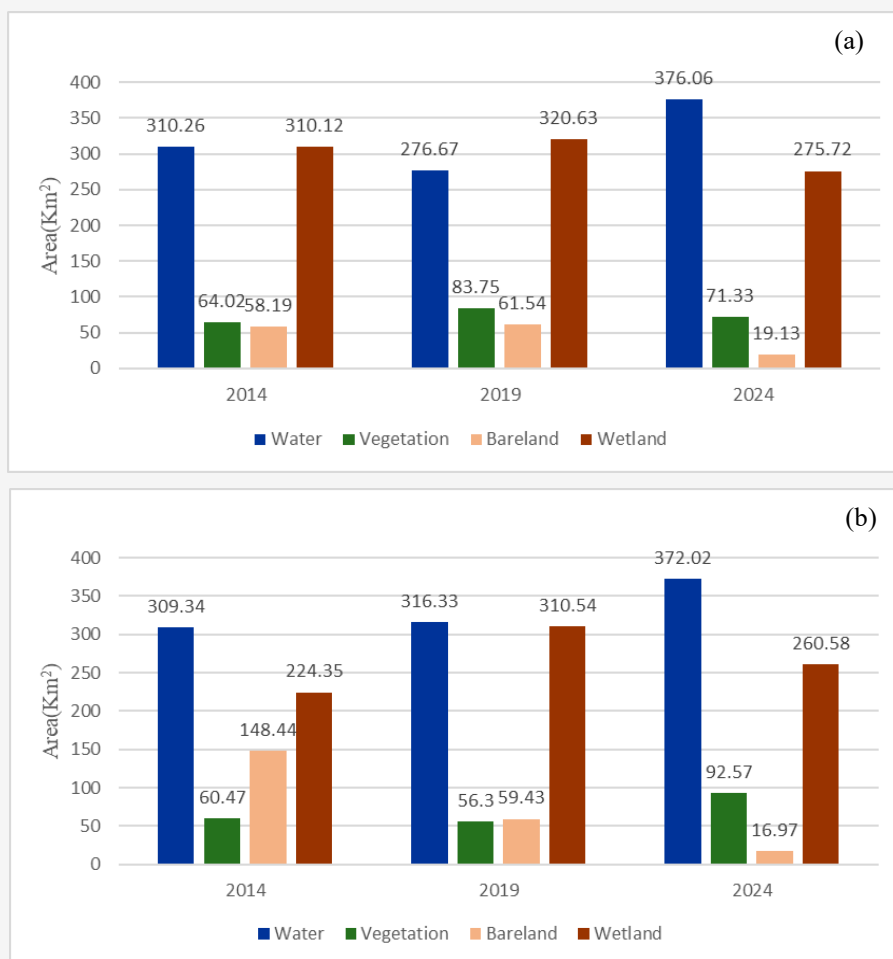


Figure 8: Land cover area classified using different classifier techniques: (a) PCA/SVM and (b) PCA/RF

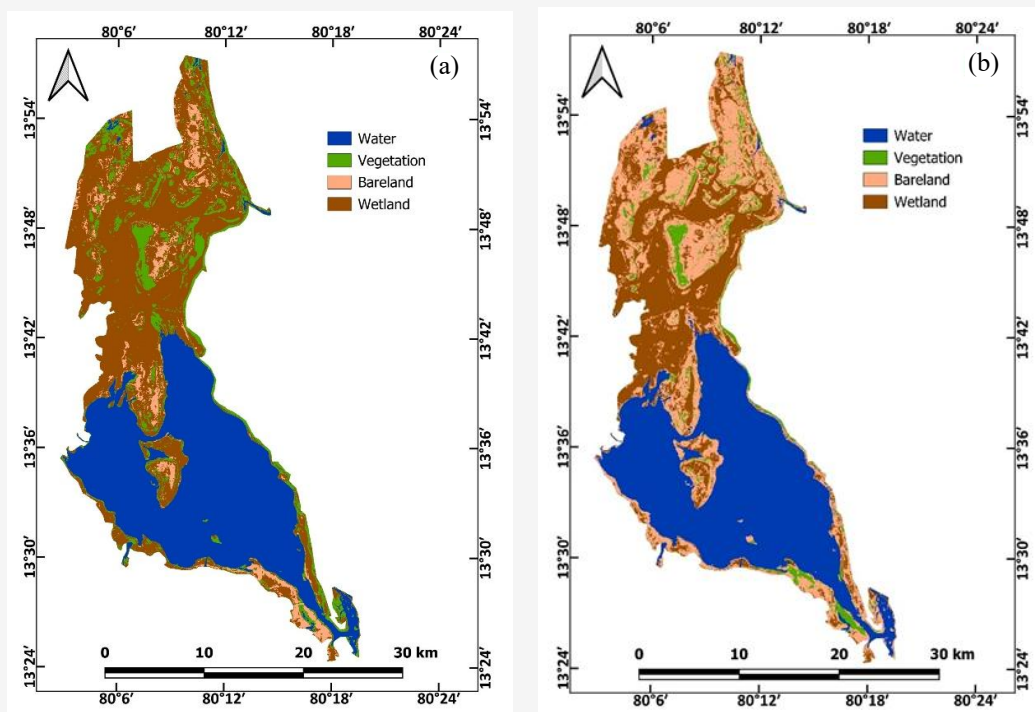


Figure 9: Land cover classification of 2014 (a) SPCA/SVM, and (b) SPCA/RF

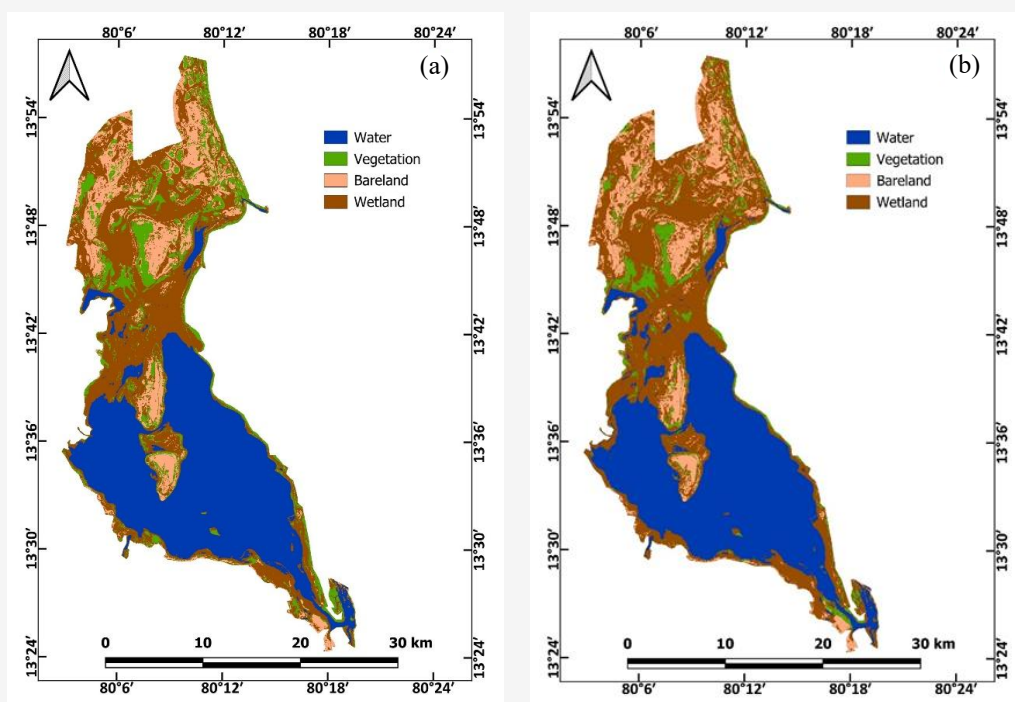


Figure 10: Land cover classification of 2019 (a) SPCA/SVM, and (b) SPCA/RF

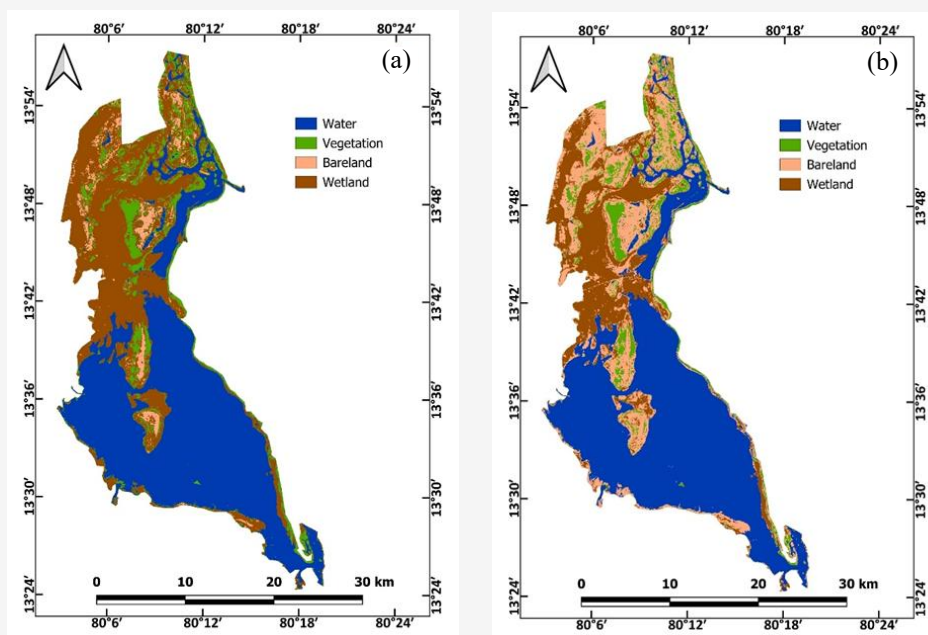


Figure 11: Land cover classification of 2024 (a) SPCA/SVM ,and (b) SPCA/RF

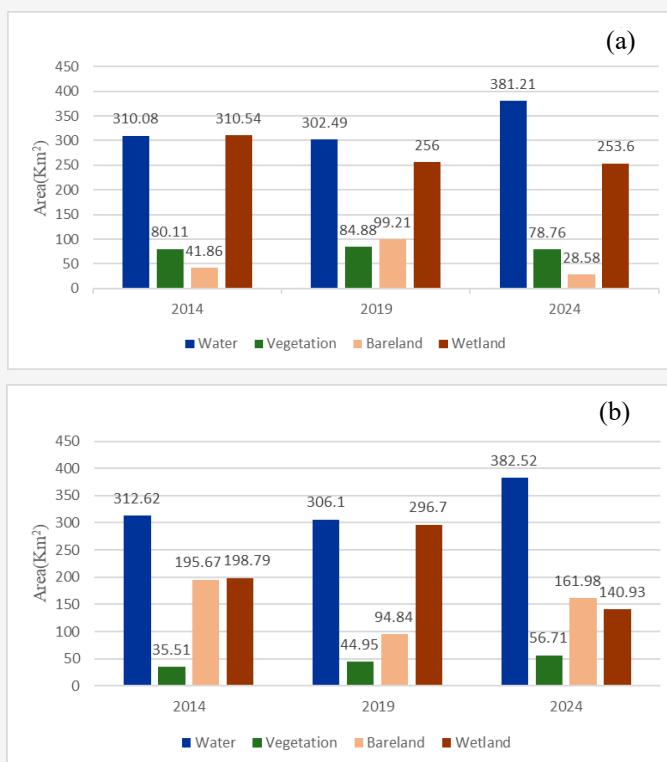


Figure 12: Land cover area classified using different classifier techniques: (a) SPCA/SVM and (b) SPCA/RF

Across all three study years, SPCA-based classifications demonstrated consistently higher accuracy than PCA, particularly when combined with the RF classifier. The enhanced separability of spectrally similar classes especially Wetland and Bareland, under SPCA contributed to this

improvement. The classification results reveal the substantial variation in class-specific area estimates among different classifier–feature combinations. This discrepancy reflects differences in how classifiers handle mixed pixels and class boundary transitions:

SVM tends to under-estimate scattered or spectrally mixed Bareland, misclassifying them into Vegetation or Wetland due to its sensitivity to overlapping class boundaries. SVM specifically works well in binary classifications.

RF, by contrast, captures the spectral heterogeneity more effectively, producing a more detailed representation of fragmented Bareland patches. This outcome is an important methodological insight for future land cover studies in complex lagoon environments where mixed pixels dominate.

5.1 Accuracy Assessment

The accuracy assessment was conducted to evaluate the performance of the land cover classifications derived from both PCA and Segmented PCA (SPCA) methods using Support Vector Machine (SVM) and Random Forest (RF) classifiers. A stratified random sampling approach was employed to ensure representative validation across all land cover classes [21]. In this process, each land cover class was designated with 50 random sample points for accuracy validation, and the samples were selected separately for the years 2014, 2019, and 2024 to account for temporal variation in the study area. The classified outputs were compared with high-resolution reference data obtained from Google

Satellite imagery and existing land cover maps. The metrics considered for the assessment are Overall Accuracy, User's Accuracy, Producer's Accuracy and Kappa Coefficient.

From the results depicted in Tables 2 and 3 below, it is evident that the SPCA-based classifications consistently outperformed PCA-based approaches in terms of overall accuracy. For the PCA method, the SVM classifier yielded accuracies of 88.27%, 86.12%, and 90.87% for the years 2014, 2019, and 2024 respectively, while the RF classifier demonstrated higher accuracies of 92.41%, 90.68%, and 95.79% for the same years. In contrast, the SPCA method delivered enhanced performance across all years. The SVM classifier achieved overall accuracies of 92.36%, 91.09%, and 92.97%, whereas the RF classifier surpassed these results with accuracies of 95.83%, 94.47%, and 96.79% respectively. The detailed accuracy assessment comparison for the PCA and SPCA methods are illustrated in Tables 2 and 3 respectively. These results clearly indicate that integrating SPCA with the RF classifier yields the highest classification accuracies. This reinforces the effectiveness of SPCA in preserving localized spectral variance and improving feature separability, ultimately resulting in more reliable land cover maps for the Pulicat Lake region.

Table 2: Comparison of accuracy assessment for PCA based land cover classification

Algorithm	Overall Accuracy (%)	User's Accuracy (%)	Producer's Accuracy (%)	Kappa Coefficient
Accuracy Assessment (2014)				
PCA/SVM	88.27	82.95	80.33	0.83
PCA/RF	92.41	89.98	89.72	0.89
Accuracy Assessment (2019)				
PCA/SVM	86.12	81.71	78.66	0.80
PCA/RF	90.68	89.58	83.44	0.86
Accuracy Assessment (2024)				
PCA/SVM	90.87	90.49	76.73	0.86
PCA/RF	95.79	96.15	84.25	0.93

Table 3: Comparison of accuracy assessment for SPCA based land cover classification

Algorithm	Overall Accuracy (%)	User's Accuracy (%)	Producer's Accuracy (%)	Kappa Coefficient
Accuracy Assessment (2014)				
SPCA/SVM	92.36	82.75	85.98	0.88
SPCA/RF	95.83	94.00	89.96	0.94
Accuracy Assessment (2019)				
SPCA/SVM	91.09	87.47	88.61	0.88
SPCA/RF	94.47	90.00	88.67	0.92
Accuracy Assessment (2024)				
SPCA/SVM	92.97	93.00	81.76	0.89
SPCA/RF	96.79	96.00	94.00	0.95

5.2 Land Cover Change Assessment of Pulicat Region Over the Period of 2014-2024

The percentage change in land cover types Water, Vegetation, Bareland, and Wetland, across three temporal intervals: from 2014 to 2019, from 2019 to 2024, and the overall period of 2014 to 2024 is illustrated in Figure 13. The most notable observation is the significant fluctuation in Wetland and Bareland classes [22][23] and [24]. Wetlands exhibited a marked decrease of approximately 20.99 % between 2019 and 2024, followed by a cumulative decline of 7.79% over the entire 10-year period, highlighting considerable degradation of wetland ecosystems in the Pulicat Lake region. Conversely, Bareland showed a decrease of 13.57% from 2014 to 2019, followed by a sharp increase of 9.05% from 2019 to 2024. The land cover change assessment for the period is illustrated in Table 4. The SPCA/RF results, which demonstrated the highest classification reliability, were used to interpret land cover changes. The following trends highlight landscape processes:

- Water bodies expanded substantially, primarily

due to increased tidal ingress, channel widening, and monsoon-driven inundation, also reported in earlier hydrological assessments [25].

- Vegetation increased, largely along the northern and western margins, suggesting natural regeneration and possible community-based conservation interventions.
- Wetland decline between 2019 and 2024 is consistent with increased sediment deposition, salinity intrusion, and encroachment patterns frequently reported for Pulicat Lake's hydromorphology [26].
- Bareland fluctuations reflect seasonal drying/wetting cycles, shoreline erosion, and exposure of lakebeds during drier periods [27] and [28].

These observed variations emphasize the dynamic nature of land cover in the Pulicat region, with clear evidence of wetland degradation and the encroachment of bare lands during the latter half of the study period.

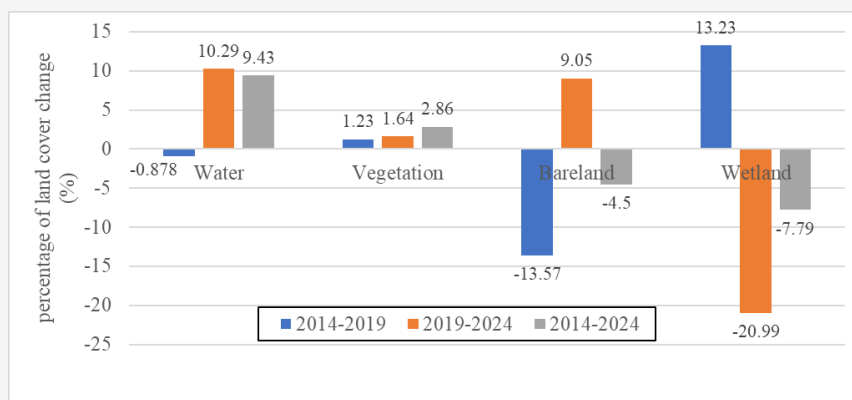


Figure 13: Percentage of land cover change

Table 4: Land cover change assessment of Pulicat region over the period of 2014-2024

Land cover type	Area (Km ²)	% Change area
Land Cover Change From 2014 to 2019		
Water	-6.52	-0.87
Vegetation	+9.12	+1.23
Bareland	-100.83	-13.57
Wetland	+98.23	+13.23
Land Cover Change From 2019 to 2024		
Water	+76.43	+10.29
Vegetation	+12.17	+1.64
Bareland	+67.18	+9.05
Wetland	-155.78	-20.99
Land Cover Change From 2014 to 2024		
Water	+69.97	+9.43
Vegetation	+21.26	+2.86
Bareland	-33.39	-4.50
Wetland	-57.84	-7.79

The results reinforce the importance of continuous monitoring using advanced techniques like Segmented PCA and machine learning classifiers to inform conservation and land management strategies.

6. Discussion

The primary objective of the present study was to analyze the land cover dynamics of the Pulicat Lake region over a ten-year period (2014–2024) using multi-temporal Landsat imagery, and performance assessment of Principal Component Analysis (PCA) and Segmented Principal Component Analysis (SPCA) integrated with machine learning classifiers such as Support Vector Machine (SVM) and Random Forest (RF) classifiers for the land cover classification. The change detection analysis was conducted for the Pulicat Lake due to its ecological significance, biodiversity and increasing exposure to developmental pressures. The lake has become increasingly vulnerable to human-induced activities such as urban expansion, industrial development, and reduced freshwater inflow, which have led to salinity intrusion, wetland degradation, and habitat loss as reported by [2][5] and [6]. The change detection analysis was undertaken to assess the spatial and temporal alterations in land cover associated with these developmental pressures, which have contributed to the gradual degradation of the Pulicat Lake ecosystem.

The results of this study demonstrate that the Segmented Principal Component Analysis (SPCA) method significantly improves land cover classification accuracy compared to conventional PCA. The enhanced performance of SPCA can be attributed to its ability to preserve localized spectral variability by processing specific band combinations independently, thereby improving class separability in heterogeneous landscapes such as Pulicat Lake. Traditional PCA, while effective in reducing redundancy, tends to emphasize dominant spectral features and suppress subtle variations, resulting in spectral mixing among similar land cover types. The present findings align with research of [10], who reported improved classification accuracy using weighted PCA-ICA for hyperspectral image analysis, and [11], who emphasized that the interpretation of principal components varies across spectral ranges, affecting class discrimination. However, by incorporating segmentation, SPCA achieves more refined class boundaries and higher accuracy levels, in accordance with the observations of [12], who highlighted that optimized band combinations enhance separability in land cover classification.

However, SPCA is not universally superior. Its performance depends on choosing appropriate segment combinations. If segments are not

ecologically meaningful or if the landscape is homogeneous, SPCA may offer little advantage over PCA. These limitations highlight the importance of tailoring SPCA band selection to ecological characteristics rather than applying it uniformly.

6.1 Ecological Interpretation of Land Cover Changes

The expansion of water bodies, particularly between 2019 and 2024, aligns with strengthened tidal influence and increased inundation across the central and southern lagoon channels. Similar hydrological fluctuations have been documented in other coastal lagoons responding to monsoon variability and channel deepening [20] and [22]. Wetlands displayed a contrasting pattern: after peaking in 2019, they declined sharply by 2024. This decline is consistent with reported salinity intrusion, sediment accumulation, and reduced freshwater inflow that have long threatened the wetland ecosystems of Pulicat [5] and [6]. The transition of large wetland areas to either water or bareland indicates both prolonged flooding in some sectors and drying/encroachment in others. Bareland dynamics also reflect these ecological processes. The decline in bareland up to 2019, whereas the subsequent increase by 2024 may indicate drying conditions driven by changing monsoon patterns or anthropogenic modification. Vegetation growth along the lake margins, although moderate, points to natural regeneration and possible community-level planting efforts, consistent with vegetative recovery patterns documented in similar coastal systems [8] and [21]. Collectively, these changes underscore the fragility of Pulicat Lake and the urgent need for sustained hydrological and ecological monitoring.

The results also reveal meaningful differences between SVM and RF. RF performed consistently better due to its ability to model class heterogeneity and capture nonlinear spectral relationships. SVM exhibited higher misclassification in mixed pixels, especially between wetland and vegetation, where boundary transitions are gradual. RF occasionally overestimated bareland in fragmented zones, while SVM tended to underestimate it due to tighter decision boundaries. These errors emphasize the need for classifier selection based on landscape complexity and the spectral mixing characteristics of wetlands.

6.2 Broader Applicability of SPCA for Lagoon and Coastal Wetland Monitoring

The integration of SPCA with Random Forest in this study produced higher classification accuracies than those reported in similar PCA-based approaches, confirming its robustness and applicability for large-

scale, heterogeneous ecosystems. The demonstrated effectiveness of SPCA in Pulicat Lake suggests strong potential for its use in other coastal lagoons, estuaries, and wetland complexes where spectral mixing, moisture gradients, and soil–vegetation overlap complicate classification. SPCA can be particularly beneficial for long-term monitoring programs where capturing subtle ecological changes is critical.

7. Conclusion

This study provides an integrated assessment of land cover changes in the Pulicat Lake region over a decade using PCA and SPCA-enhanced feature extraction and machine learning classification. The classification focused on four key land cover types: Water, Vegetation, Bareland, and Wetland. The SPCA–RF combination offered the most reliable and ecologically consistent land cover maps, enabling a clearer understanding of the landscape transformations occurring within this vulnerable lagoon ecosystem. The SPCA method, which utilizes spectral segmented band combinations to enhance spectral feature representation, outperformed standard PCA in classification accuracy across all years. Among the classifiers, Random Forest (RF) achieved superior performance compared to Support Vector Machine (SVM), yielding overall accuracies of 95.83%, 94.47%, and 96.79% for 2014, 2019, and 2024, respectively, using SPCA-derived features. The results indicate major ecological transitions, including long-term wetland degradation, expansion of water bodies, and fluctuations in bareland and vegetation. The findings carry important implications for ecosystem management. The documented wetland losses emphasize the need for enhanced freshwater inflow management, control of shoreline encroachment, and active restoration of marsh areas that buffer the lake’s ecological functioning. The demonstrated effectiveness of integrating SPCA with RF suggests a framework for similar coastal ecosystems and wetland monitoring. Its ability to enhance class separability in heterogeneous environments makes it suitable for long-term change detection in lagoons, estuaries, and ecologically sensitive regions. However, the approach still requires careful selection of band combinations and may benefit from complementary high-resolution datasets to reduce uncertainties related to mixed pixels. Future research can be explored with multi-seasonal imagery to capture intra-annual ecological variability, integrate additional physical parameters such as hydrodynamics or salinity gradients, and apply predictive modeling to anticipate future land cover trajectories to similar ecosystems.

References

- [1] Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y. A. and Rahman, A., (2020). Land-Use Land-cover Classification by Machine Learning Classifiers for Satellite Observations-A Review. *Remote Sensing*, Vol. 12. <https://doi.org/10.3390/rs12071135>.
- [2] Kannan, V., (2022). Vulnerable Ecosystem: The Pulicat Lake Needs Government's Attention. *Earthy Worthy*, Vol. 1(1).
- [3] Illangovan, R., (2007). Restoration of Polluted Lakes – A New Approach. In: *Proceedings of the 12th World Lake Conference (Taal 2007)*, 1321–1328.
- [4] Sanjeeva Raj, P. J., (2006). Macro Fauna of Pulicat Lake. Chennai: National Biodiversity Authority, 1–67.
- [5] Thirunavukkarasu, N., Gokulakrishnan, S., Premjothi, P. V. R. and Inbaraj, R. M., (2011). Need of Coastal Resource Management in Pulicat Lake – Challenges Ahead. *Indian Journal of Science and Technology*, Vol. 4(3), 322–326. <https://doi.org/10.17485/ijst/2011/v4i3/29990>.
- [6] Syamala, R. and Hemavathy, E., (2018). Physico-Chemical Parameters and Land Use Patterns of Pulicat Lake, Tamil Nadu, India. *International Journal of Advanced Scientific and Technical Research*, Vol. 6(8), 10–37. <https://dx.doi.org/10.26808/rs.st.i8v6.02>.
- [7] Chander, G., Markham, B. L. and Helder, D. L., (2009). Summary of Current Radiometric Calibration Coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI Sensors. *Remote Sensing of Environment*, Vol. 113(5); 893–903. <https://doi.org/10.1016/j.rse.2009.01.007>.
- [8] Makar, R. S., Shahin, S. A., El-Nazer, M. and Wheida, A., (2022). Development of a PCA-based Land Use/ Land Cover Classification Utilizing Sentinel-2 Time Series. *Middle East Journal of Agricultural Research*, Vol. 11(2), 630–637. <https://doi.org/10.36632/mejar/2022.11.2.42>.
- [9] Uddin, M. P., Mamun, M. A. and Hossain, M. A., (2020). PCA-based Feature Reduction for Hyperspectral Remote Sensing Image Classification. *IETE Technical Review*, Vol. 38(5), 1–21. <https://doi.org/10.1080/02564602.2020.1740615>.
- [10] Liu, L., Li, C. F. and Lei, Y. M., (2017). Feature Extraction for Hyperspectral Remote Sensing Image Using Weighted PCA-ICA. *Arabian Journal of Geosciences*, Vol. 10. <https://doi.org/10.1007/s12517-017-3090-1>.

- [11] Gupta, R., Tiwari, R., Saini, V. and Srivastava, N., (2013). A Simplified Approach for Interpreting Principal Component Images. *Advances in Remote Sensing*, Vol. 2, 111–119. <https://doi.org/10.4236/ars.2013.22015>.
- [12] Kulkarni, K. and Vijaya, V., (2021). Separability Analysis of the Band Combinations for Land Cover Classification of Satellite Images. *International Journal of Engineering Trends and Technology*, Vol. 69, 138–144. <https://doi.org/10.14445/22315381/IJETT-V69I8P217>.
- [13] Jiang, W., He, G., Long, T. and Ni, Y., (2018). Detecting Water Bodies in Landsat 8 OLI Image Using Deep Learning. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLII-3, 669–672. <https://doi.org/10.5194/isprs-archives-XLII-3-669-2018>.
- [14] Fahrezi, A., Danoedoro, P., and Kamal, M. (2025). Comparison of Detailed Land Cover Mapping Based on SNI Classification Scheme using Conventional and Machine Learning Multispectral Classification. *International Journal of Geoinformatics*, Vol. 21(2), 87–105. <https://doi.org/10.52939/ijg.v21i2.3941>.
- [15] Akar, O. and Gungor, O., (2012). Classification of Multispectral Images Using Random Forest Algorithm. *Journal of Geodesy and Geoinformation*, Vol. 1(2), 105–112. <https://doi.org/10.9733/jgg.241212.1>.
- [16] United States Geological Survey (USGS). (2024). Earth Explorer. [Online] Available: <https://earthexplorer.usgs.gov>. [Accessed: Apr. 4, 2024].
- [17] Shi, D. and Yang, X., (2015). Support Vector Machines for Land Cover Mapping from Remote Sensor Imagery In: *Monitoring and Modeling of Global Changes: A Geomatics Perspective*. Springer, 265–279. <https://doi.org/10.1007/978-3-319-16787-917>.
- [18] Vasavi, S., Chintalapudi, V. K. and Vuppuluri, A. S. R., (2024). Classification of Water Bodies Using Ensemble of U-Net and Random Forest Algorithm. *Journal of Image and Graphics*, Vol. 12(1), 76–89. <https://www.joig.net/2024/JOIG-V12N1-76.pdf>.
- [19] QGIS.org. (2024). Quantum Geographic Information System (QGIS). [Online] Available: <https://qgis.org>. [Accessed: Apr. 4, 2024]
- [20] Congedo, L., (2021). Semi-Automatic Classification Plugin: A Python Tool for the Download and Processing of Remote Sensing Images in QGIS. <https://doi.org/10.13140/R>
G.2.2.29474.02242. [Accessed: Apr. 4, 2024].
- [21] Llano, X., (2025). AcATaMa – QGIS Plugin for Accuracy Assessment of Thematic Maps, Version [24.12C]. SMBByC–IDEAM. [Online] Available: <https://github.com/SMBByC/AcATaMa>. [Accessed: Sep. 3, 2024].
- [22] Aji, A., Husna, V., and Purnama, S. (2024). Multi-Temporal Data for Land Use Change Analysis Using a Machine Learning Approach (Google Earth Engine). *International Journal of Geoinformatics*, Vol. 20(4), 19–28. <https://doi.org/10.52939/ijg.v20i4.3145..>
- [23] Bill Donatien, L. M., Clobite, B. B. and Meris Midel, M. L., (2024). Land Use Land Cover Change Detection Using Multi-Temporal Landsat Imagery in the North of Congo Republic: A Case Study in Sangha Region. *Geocarto International*, Vol. 39(1). <https://doi.org/10.1080/10106049.2024.2425184>.
- [24] Rokni, K., Ahmad, A., Selamat, A. and Hazini, S., (2014). Water Feature Extraction and Change Detection Using Multitemporal Landsat Imagery. *Remote Sensing*, Vol. 6, 4173–4189. <https://doi.org/10.3390/rs6054173>.
- [25] Singh, R., Saritha, V. and Pande, C. B., (2023). Monitoring of Wetland Turbidity Using Multi-Temporal Landsat-8 and Landsat-9 Satellite Imagery in the Bisalpur Wetland, Rajasthan, India. *Environmental Research*, Vol. 241. <https://doi.org/10.1016/j.envres.2023.117638>.
- [26] Sarp, G. and Ozcelik, M., (2016). Water Body Extraction and Change Detection Using Time Series: A Case Study of Lake Burdur, Turkey. *Journal of Taibah University for Science*, Vol. 11, 381–391. <https://doi.org/10.1016/j.jtusci.2016.07.004>.
- [27] Pal, T., Paul, A., Maiti, C. and Podder, T., (2024). Spatiotemporal Analysis of Land Use and Land Cover Dynamics Using Multi-Temporal Remote Sensed Imagery and Quantifying Landscape Fragmentation: A Case Study of Chandannagar Municipal Corporation (CMC), West Bengal, India. *World Journal of Advanced Research and Reviews*, Vol. 23(3), 966–995. <https://doi.org/10.30574/wjarr.2024.23.3.2664>.
- [28] Mohsen, A., Elshemy, M. and Zeidan, B., (2021). Water Quality Monitoring of Lake Burullus (Egypt) Using Landsat Satellite Imageries. *Environmental Science and Pollution Research*, Vol. 28, 15687–15700. <https://doi.org/10.1007/s11356-020-11852-w>.