Comparison of Supervised and Unsupervised Classification for LULC Mapping of Hathazari Upazilla

1. Introduction

Land use-land cover (LULC) classification is essential for understanding the spatial distribution of natural and human-induced features on the Earth's surface. In this study, both unsupervised and supervised classification methods were applied to satellite imagery of Hathazari Upazilla, located in Chattogram, Bangladesh. The purpose of this comparison is to evaluate which approach provides a more accurate and meaningful representation of land cover patterns such as vegetation, settlements, barren land, and water bodies.

2. Methodology

2.1 Study Area

Hathazari Upazilla is in the Chattogram District of southeastern Bangladesh, between approximately 22°24'–22°36' N latitude and 91°41'–91°53' E longitude, covering an area of about 251 km². The region contains agricultural land, settlements, vegetation-covered areas, and waterbodies.

2.2 Data Source

Satellite imagery of Hathazari Upazilla was used as the primary dataset. This study used Sentinel-2A MSI Level-2A multispectral satellite imagery for generating the land use-land cover maps. The image was accessed and processed using the Google Earth Engine (GEE) cloud platform. A cloud free Sentinel-2 scene covering Hathazari Upazilla was selected and clipped to the study area boundary. The spectral bands with 10 m spatial resolution (B2, B3, B4, and B8) were used for classification.

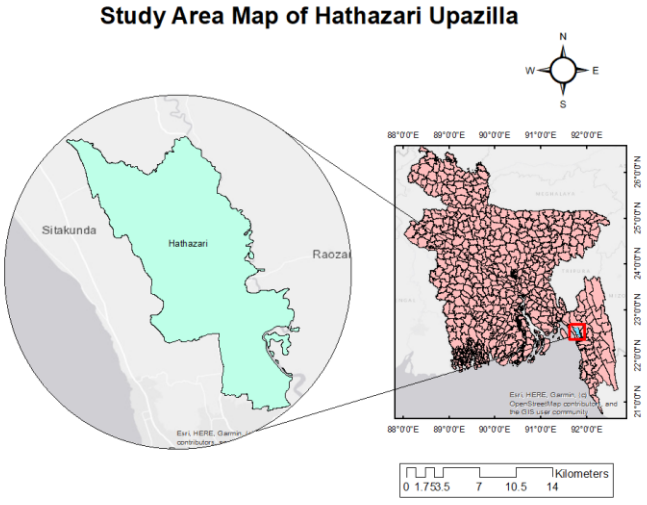

Figure 1 Study area map

2.3 Unsupervised Classification

The unsupervised classification was performed using the K-means clustering algorithm. The number of clusters (K) was set equal to the number of land cover classes expected in the study area. In this analysis, four major land cover classes were focused on: Vegetation, Waterbody, Settlement, and Barren Land. The clustering algorithm grouped pixels based on their spectral reflectance similarity. After clustering, each cluster was interpreted and assigned to the respective land cover class through manual visual comparison and reference information.

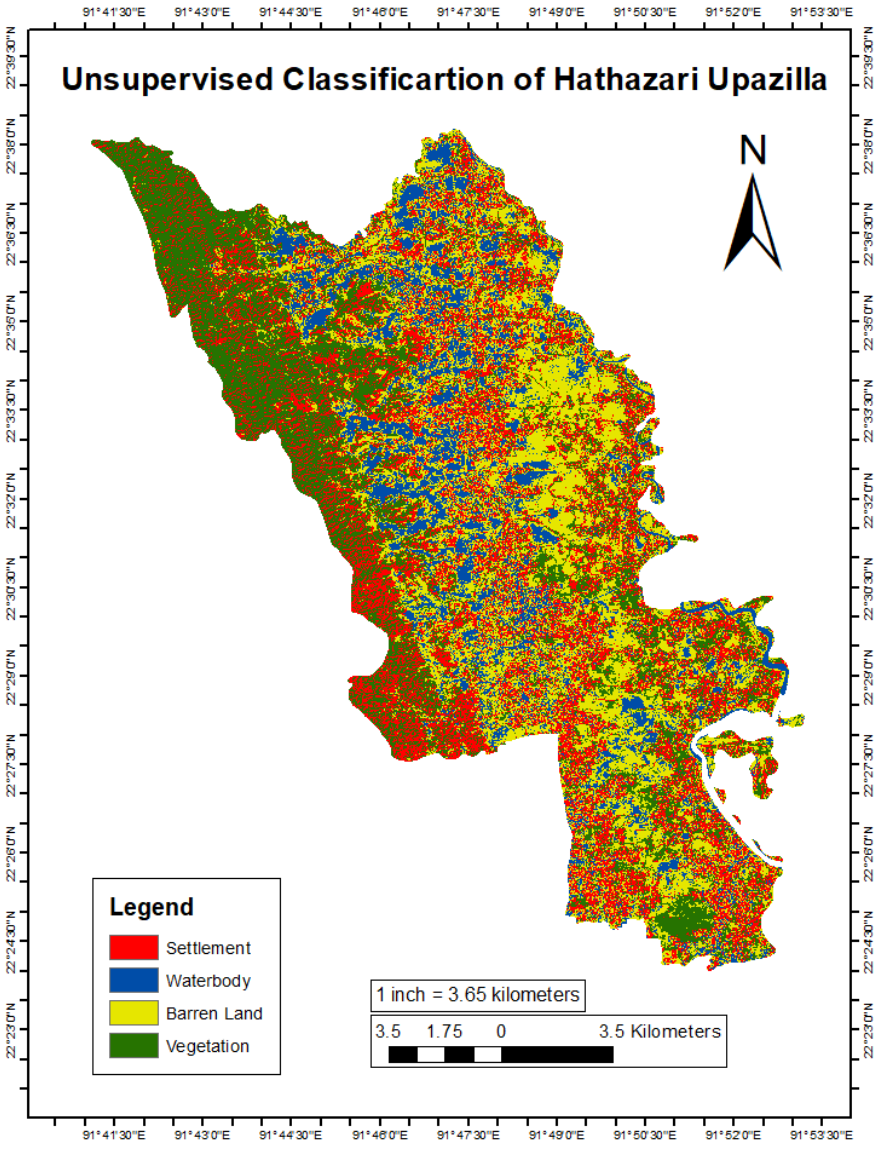

Figure 2 Unsupervised classification

2.4 Supervised Classification

The supervised classification approach required prior knowledge of the land cover categories. Training samples for each class (Vegetation, Waterbody, Settlement, and Barren Land) were collected using visual interpretation and known ground reference points. A Maximum Likelihood Classification (MLC) algorithm was applied to categorize all pixels. This method utilizes statistical probability distributions of the spectral signatures from the training data to classify unknown pixels.

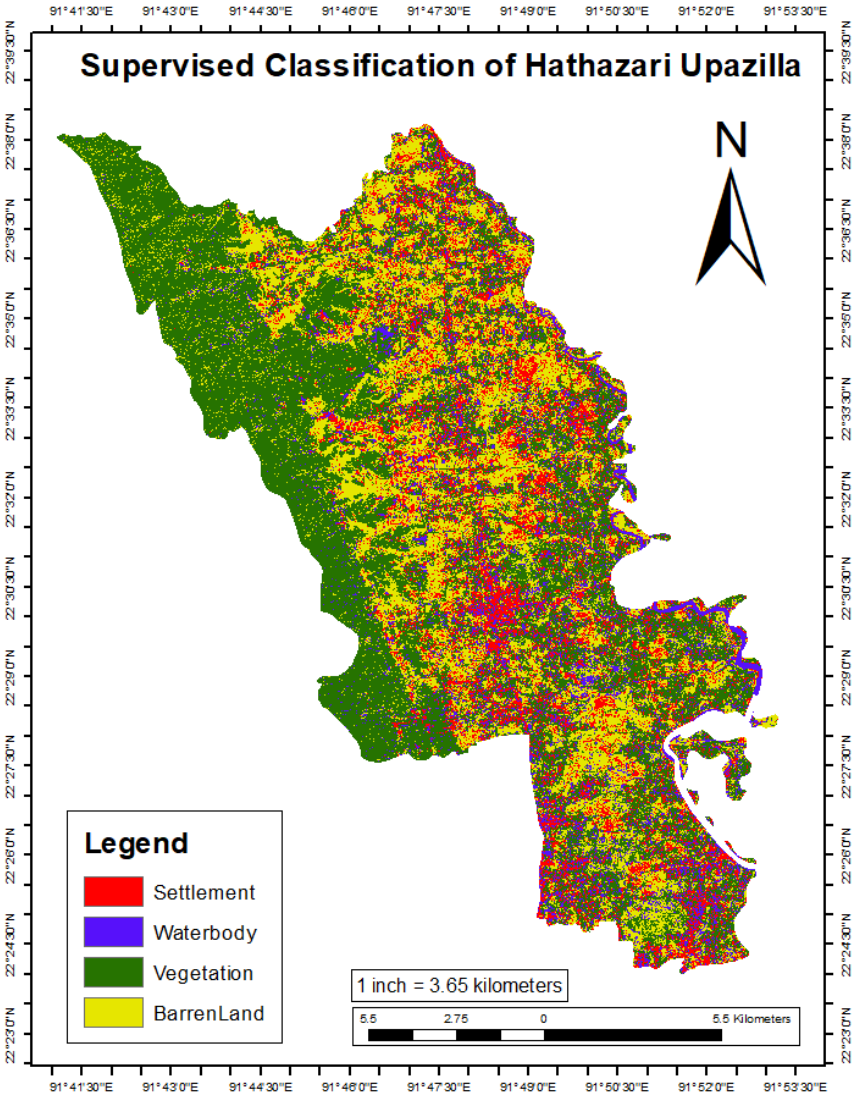

Figure 3 Supervised classification

2.5 Map Production

Both classified outputs were exported from Google Earth Engine and further processed in ArcGIS 10.8. The maps were clipped to the Hathazari Upazilla boundary, and necessary cartographic elements such as the legend, north arrow, scale bar, and coordinate grid were added to ensure proper map visualization and interpretation.

3. Results and Discussion

The land use-land cover classification of Hathazari Upazilla produced two outputs using unsupervised (K-means) and supervised (Maximum Likelihood Classification) methods. The unsupervised classification map presents a general distribution of land cover patterns; however, it shows notable class confusion. Vegetation and agricultural lands appear mixed due to their similar spectral reflectance signatures, particularly during the active growing season. Likewise, settlement and barren land areas tend to overlap, especially where settlement patterns are dispersed or semi-rural in nature. Although the unsupervised classification is simple, fast, and does not require prior training data, the results demonstrate limited thematic accuracy for heterogeneous landscapes such as Hathazari.

In comparison, the supervised classification produced a more spatially coherent and thematically accurate land cover map. The use of manually selected training samples enabled the classifier to capture both spectral and contextual differences among classes. Vegetation areas were clearly delineated in the northern elevated and forested zones, while settlement areas were more accurately defined, displaying continuous spatial patterns in developed regions. Water bodies including the Halda River and small ponds were distinctly mapped with minimal misclassification. Barren land areas were correctly identified where agricultural fields were temporarily unused, indicating the classifier's effectiveness in discriminating seasonal agricultural variation.

The overall results suggest that supervised classification provides a higher level of reliability in representing land cover in Hathazari Upazilla. Its strength lies in guided training data, which improves class separability and reduces misclassification. In contrast, unsupervised classification is more suitable for preliminary analysis, rapid mapping, or scenarios where ground reference or training data are not available.

Table 1. Comparison of Unsupervised and Supervised Classification Methods

Criteria	Unsupervised Classification (K-means)	Supervised Classification (Maximum Likelihood)
Prior Knowledge Requirement	Not required	Training samples required
Class Separability	Moderate; spectral confusion common	High; clearer separation between classes
Performance in Mixed Land Areas	Limited accuracy	High accuracy and consistency
Representation of Vegetation	Often mixed with cropland	Clearly delineated and continuous
Mapping of Settlement Areas	Overlaps with barren land	Well-differentiated and spatially coherent
Mapping of Water Bodies	May contain minor noise	Distinct, minimally misclassified
Overall Suitability	Useful for exploration/initial classification	Recommended for detailed LULC mapping and planning

4. Conclusion

Both classification methods generated LULC maps for Hathazari Upazilla; however, the supervised classification produced a significantly more accurate and interpretable outcome. The ability to incorporate training data allowed the supervised method to better represent the spatial and thematic characteristics of the land cover classes. The unsupervised classification, while useful for preliminary pattern recognition, exhibited notable class confusion. Thus, supervised classification is recommended for future land-use analysis and planning efforts in the region.