Land Use Land Cover (LULC) Classification of Boalkhali Upazila using Supervised and Unsupervised Classification Techniques in Google Earth Engine

1. Introduction

An LULC (Land Use/Land Cover) map is a type of map that shows how the land surface of an area is being used and what is on it. It is a critical tool for understanding, monitoring, and managing an area's natural resources and development by categorizing land into types like forest, water, urban, and agricultural areas. These maps are created using satellite imagery and Geographic Information Systems (GIS), and can be used to track changes over time.

Land Use/Land Cover (LULC) maps classify land into broad categories such as water, crops, or built area. They are useful for visualising the dominant land uses in a given area. The total area or proportion of different classes can also be calculated for a specified area. Land Use Land Cover (LULC) mapping plays a vital role in monitoring environmental changes, planning land management, and assessing ecosystem health.

In this study, Sentinel-2 imagery of the year 2023 was used to prepare two LULC maps of Boalkhali Upazila, Chattogram District, Bangladesh—one using unsupervised classification (K-Means clustering) and another using supervised classification (smileRandomForest). The goal was to compare the outcomes of both techniques and identify which method produces a more accurate land cover map.

2. Methodology

2.1 Data Collection and Pre-processing

a) Sentinel-2 Level-2A imagery for the year 2023 was obtained from the *COPERNICUS/S2_SR_HARMONIZED* dataset in GEE Data Catalogue.
b) Cloudy scenes were filtered using the condition `CLOUDY_PIXEL_PERCENTAGE < 10`, and a median composite was taken using .median().
c) The boundary of Boalkhali Upazila was extracted from an administrative boundary shapefile.
d) The imagery was clipped to the desired boundary.

2.2 Unsupervised Classification

a) The K-Means clustering algorithm (`ee.Clusterer.wekaKMeans`) was applied using 5 clusters (K=5) to automatically classify pixels based on the similarity of the land use type.
b) The obtained TIFF file was imported to Qgis and masked within the boundary of Boalkhali Upazila to eliminate the pixels outside the region.
c) The clusters were visualized using a separate color palette and interpreted manually as vegetation, waterbody, settlement, and barren land based on visual comparison.

2.3 Supervised Classification

a) Training points were manually digitized for Vegetation, Waterbody, Settlement, and Barren Land and Agricultural Land using high-resolution imagery from Google Earth.
b) A Random Forest (RF) classifier with 50 trees was trained using these samples.
c) The trained model was applied to the Sentinel-2 composite to produce a supervised LULC map.

3. Results and Discussion

The K-Means clustering and the Random Forest classifier both successfully differentiated general land cover patterns. Vegetation areas were distinctly mapped in Dark green, settlements appeared in red, and waterbodies in blue, agricultural land use in light green and barren land in brown colour.

Even though it was expected for the unsupervised classification to lack precision in separating classes successfully in absence of user-defined training points, observations show that there is almost no difference between the supervised and unsupervised classification methods.

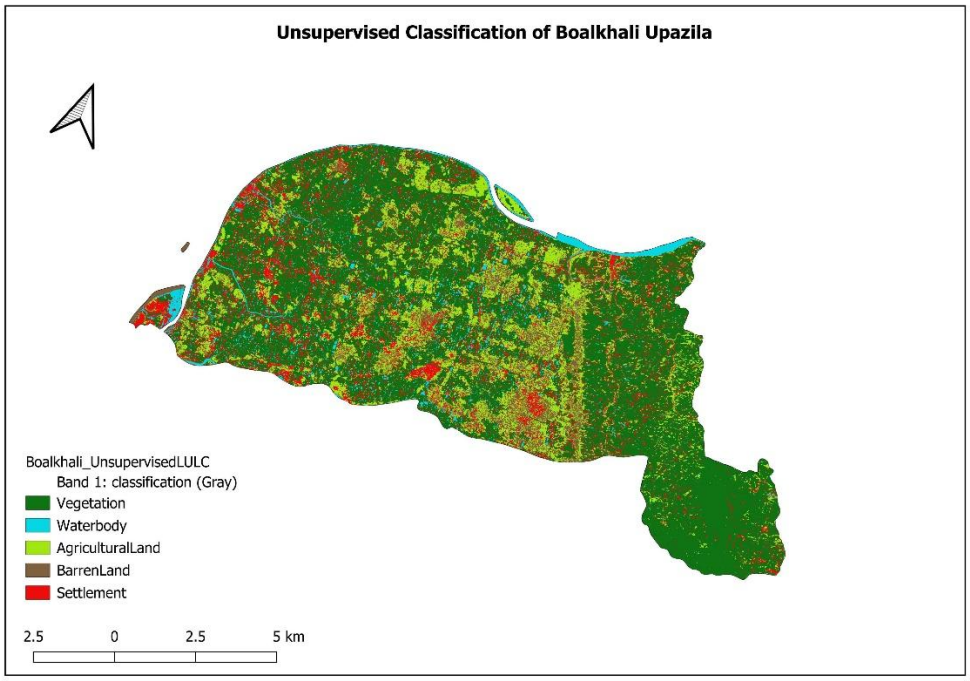
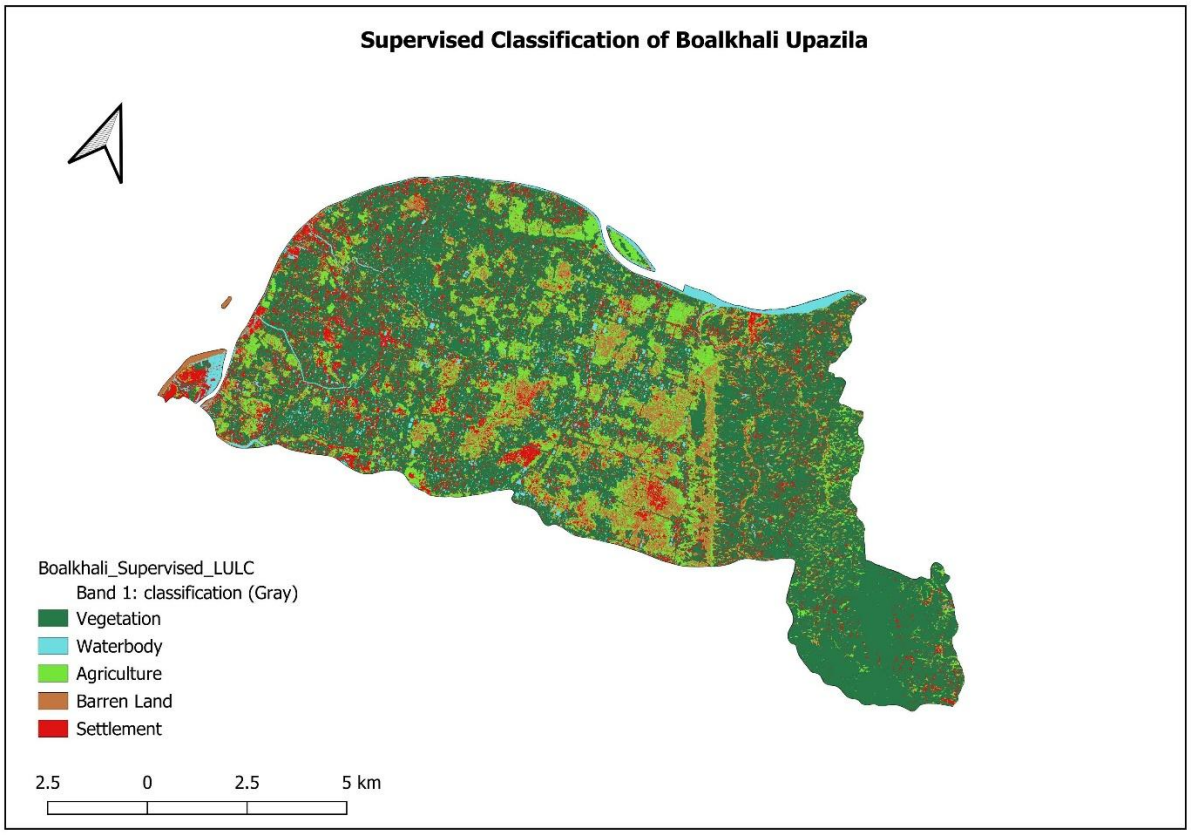


Figure 1 Unsupervised Classification



3. Conclusion

Both classification techniques provide valuable insights into land cover patterns. However, supervised classification proved to be more reliable and accurate for LULC mapping in Boalkhali Upazila.

It is convenient and useful to use the unsupervised approach in classifying land use when ground truth data are unavailable, but for detailed land use mapping and policy-related applications, it is strongly recommended to go with the supervised classification approach.