

Land Use and Land Cover Analysis of Koppal District Using ArcGIS and Machine Learning Techniques

Abdul Muqeed L^{1*}, Bindu T¹, Deeksha¹, Abdul Rahaman K¹

¹Department of Civil Engineering, Ballari Institute of Technology & Management, Ballari – 583101, Karnataka, India

*Corresponding Author: dniharika406@gmail.com

Abstract - Land use and land cover (LULC) analysis is an essential tool for understanding the spatial distribution of human activity and natural features across a geographical area, supporting resource monitoring and land management decisions. This study presents a LULC analysis of Koppal district, Karnataka, India, using a satellite image acquired on 27 October 2024, processed through ArcGIS and classified using machine learning algorithms. The district was classified into six categories: agriculture, settlements, forest, water bodies, grazing land, and mining areas. Random Forest and Support Vector Machine (SVM) classifiers were applied to the pre-processed Landsat-8 and Sentinel-2 imagery, and the classification was validated through a confusion matrix and Kappa coefficient. The results indicate that agriculture remains the dominant land-use class, though its extent has declined over the past two decades in favour of settlements and mining. The classification achieved an overall accuracy of approximately 86%, with a Kappa coefficient of about 0.83, confirming the reliability of the approach. The findings highlight encroachment of settlements into agricultural land and the growing footprint of mining activity, and they provide a replicable framework for land-use monitoring in semi-arid districts undergoing rapid transformation.

Key Words: Land use and land cover, remote sensing, ArcGIS, Random Forest, Support Vector Machine, Koppal.

1. INTRODUCTION

Land use and land cover (LULC) analysis plays a pivotal role in understanding the spatial patterns and dynamics of the Earth's surface, reflecting how natural processes and anthropogenic activities shape landscapes over time. LULC studies underpin a wide range of applications, including environmental management, urban planning, agricultural monitoring, resource allocation, and biodiversity conservation, and their importance has grown with increasing concern over climate change, population growth, and sustainable development [7], [19].

This study applies satellite imagery, machine learning techniques, and ArcGIS software to conduct a detailed LULC analysis of Koppal district, Karnataka, focusing on six land categories: agriculture, settlements, forest, grazing land, mining areas, and water bodies. The specific objectives are to classify and quantify the distribution of LULC categories for the year 2024; to analyse the extent of human activities such as urbanisation, mining, and agricultural expansion; to evaluate the effectiveness of machine learning algorithms for

LULC classification; and to derive recommendations for sustainable land-use planning in the district.

Koppal was selected for this study owing to its diverse landscape of fertile plains, forested hills, grazing lands, and mineral-rich zones, and its predominantly rural, agriculture-dependent population. The district is undergoing rapid transformation driven by population growth, urbanisation, and mining, which together create measurable pressure on land resources. Machine learning classifiers such as Random Forest and Support Vector Machines (SVM) were used to classify the satellite imagery based on spectral, spatial, and contextual features, while ArcGIS Pro provided the platform for spatial analysis, data integration, and cartographic representation [15].

2. STUDY AREA

2.1 Geographical Setting

Koppal district lies in the northern part of Karnataka, India, between latitudes 15.35°N and 15.70°N and longitudes 76.05°E and 76.55°E, covering an area of approximately 7,190 km². The region experiences a semi-arid climate with hot summers (30–40°C, March–June), a monsoon season contributing 600–700 mm of annual rainfall (June–September), and mild winters (November–February). The Tungabhadra River is the principal water resource, supporting irrigation and rural livelihoods through the associated canal network.

The district's topography ranges from the fertile alluvial plains of the Tungabhadra basin in the east and southeast, suited to paddy cultivation, to the undulating, mineral-rich terrain of the central and western parts, and the dry deciduous forests of the south. Black cotton soils support cotton, maize, and pulse cultivation, while red lateritic soils in less fertile tracts are used mainly for grazing and afforestation. Nearly 70% of the population depends on agriculture and allied activities, and rising urbanisation around Gangavati and Koppal town, together with expanding mining activity, is progressively converting agricultural land to residential, commercial, and industrial use.

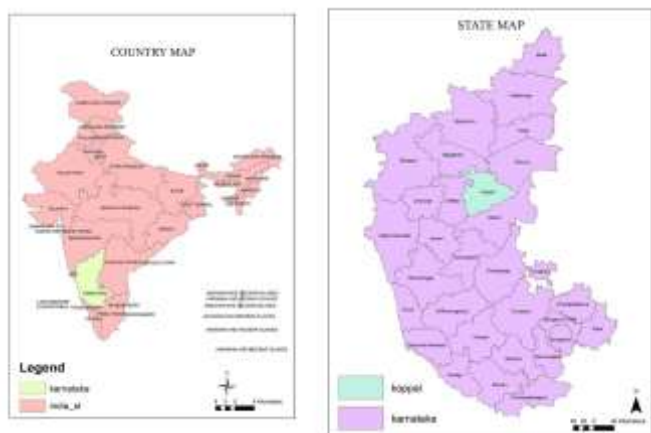


Fig -1: Location of the study area — (a) India, showing the position of Karnataka state; (b) Karnataka state, showing the location of Koppal district

3. LITERATURE REVIEW

Gopalakrishnan et al. [1] analysed LULC change in an urban expansion zone using Landsat imagery classified with Maximum Likelihood Classification, and reported that rapid urbanisation converted agricultural land and wetlands into built-up area, recommending green infrastructure to mitigate the impact.

Patil et al. [2] conducted a land suitability analysis for agriculture combining GIS with Multi-Criteria Decision Analysis, weighting soil, topographic, and climatic factors through the Analytical Hierarchy Process; only about 65% of the study area was found highly suitable for farming.

Sharma et al. [3] used MODIS and Sentinel-2 imagery to study deforestation trends over a decade, finding that forest loss reduced groundwater recharge and increased surface runoff and soil erosion, and recommended afforestation and stricter land-use regulation.

Kumar et al. [4] applied Random Forest and SVM classifiers to map urban sprawl, reporting a 40% expansion of built-up area over two decades that reduced vegetation cover, fragmented habitats, and intensified urban heat islands.

Mehta et al. [5] integrated rainfall data, digital elevation models, and drainage patterns in a GIS-based hydrological model to identify flood-prone zones, showing that encroachment on natural drainage channels was a major contributor to flooding.

Reddy et al. [6] used the Normalized Difference Vegetation Index (NDVI) to assess land degradation, linking poor land management, deforestation, and overgrazing to loss of soil fertility, and recommended afforestation and controlled grazing as mitigation measures.

Collectively, these studies confirm that remote sensing and GIS-based classification, increasingly supported by machine learning, are effective tools for tracking land-use transitions, and they motivate the integrated ArcGIS and machine-learning workflow adopted for Koppal district in the present study.

4. DATA AND METHODOLOGY

4.1 Satellite Data

LULC classification was based on Landsat-8 and Sentinel-2 imagery acquired on 27 October 2024, at Level-2 (surface reflectance) processing. Landsat-8 imagery, with its 30 m spatial resolution and long-term consistency, and Sentinel-2 imagery, offering higher spatial resolution, were used in combination to improve classification accuracy [12], [13]. Key specifications of the two platforms are summarised in Table 1.

Table -1: Specifications of satellite data used for LULC classification

Parameter	Landsat-8	Sentinel-2
Sensor	OLI & TIRS	MSI
Spatial resolution	30 m	10/20/60 m
Temporal resolution	16-day revisit	5-day revisit
Processing level	Level-2 SR	Level-2 SR
Date of acquisition	27 Oct 2024	27 Oct 2024



Fig -2: Landsat/Sentinel image of Koppal district used for LULC analysis

4.2 Pre-processing and Feature Extraction

Raw imagery was radiometrically corrected to convert digital numbers to surface reflectance, and atmospheric correction was applied using Dark Object Subtraction and Sentinel-2 Level-2A surface-reflectance products. Geometric correction using ground control points aligned the imagery to real-world coordinates, and cloud masking (Fmask) removed cloud and shadow-affected pixels. Multi-scene composites were mosaicked and clipped to the district boundary, followed by histogram equalisation and contrast stretching to enhance visual separability of land-cover classes.

Feature extraction combined spectral indices — the Normalized Difference Vegetation Index (NDVI) for vegetation, the Normalized Difference Water Index (NDWI) for water bodies, and the Normalized Difference Built-up Index (NDBI) for built-up areas — with Principal Component Analysis for dimensionality reduction and Gray-Level Co-occurrence Matrix (GLCM) texture measures (contrast, entropy, homogeneity) to aid discrimination of built-up, barren, and vegetated surfaces. Object-Based Image Analysis (OBIA) was used to segment the imagery into spectrally and

spatially homogeneous objects, reducing pixel-level misclassification.

4.3 Classification and Accuracy Assessment

Random Forest and SVM classifiers were trained on manually selected, representative samples for each LULC class, split into training and validation subsets. Random Forest, an ensemble of decision trees, was selected for its robustness to high-dimensional data and resistance to overfitting, while SVM with a Radial Basis Function kernel was used to separate classes in high-dimensional feature space. Both classifiers were implemented in ArcGIS Pro using the Spatial Analyst and Image Analyst extensions, with ArcGIS ModelBuilder used to automate the processing workflow [15]. Hyperparameters, including tree depth for Random Forest and kernel parameters for SVM, were tuned to maximise classification accuracy.

Validation followed a stratified random sampling design, with independent samples compared against classified output to compute Overall Accuracy, Producer's Accuracy, User's Accuracy, and the Kappa coefficient [8], [9]. Post-classification refinement, including majority filtering and contextual editing, was applied to reduce salt-and-pepper noise and improve boundary delineation. The overall analytical workflow is summarised in Figs. 3 and 4.

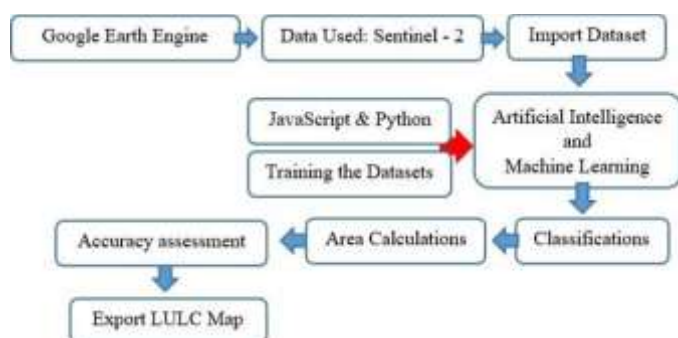


Fig -3: Flowchart of the methodology adopted for LULC classification

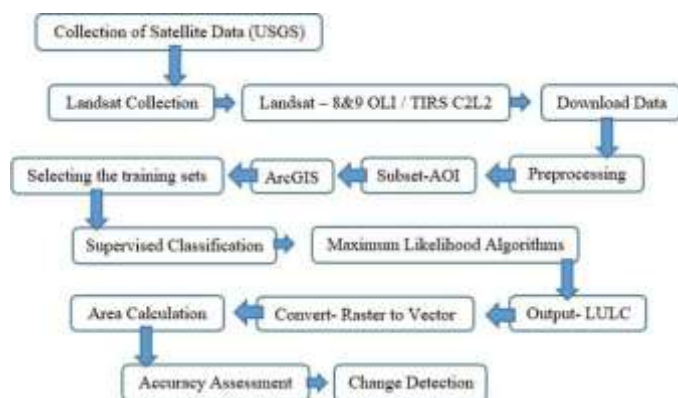


Fig -4: Workflow for LULC analysis of the Koppal study area

5. RESULTS AND DISCUSSION

5.1 LULC Classification Results

The 2024 LULC map of Koppal district (Fig. 5) classifies the region into six categories: agriculture, settlements, forest, water bodies, grazing land, and mining areas. Agriculture is the dominant class, covering roughly half of the district area, followed by mining areas and grazing land, with forest, settlements, and water bodies occupying smaller shares. The

estimated area and percentage share of each class are summarised in Table 2.

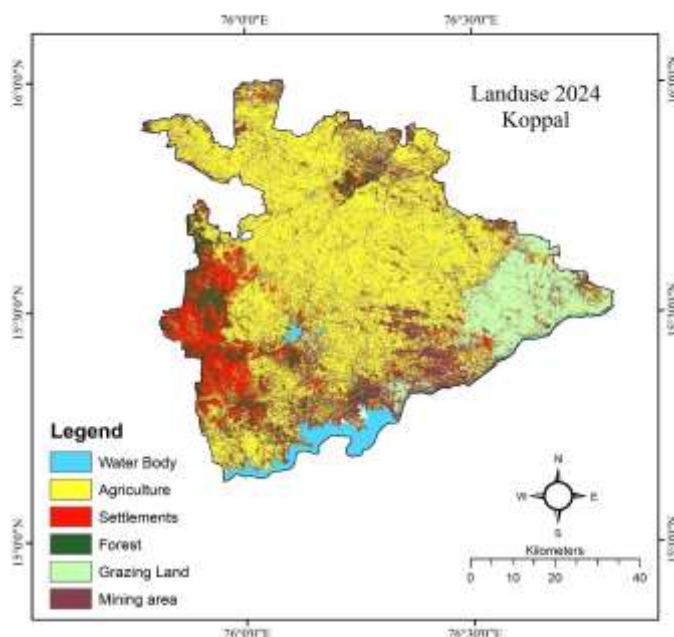


Fig -5: LULC classification map of Koppal district, 2024

Table -2: Estimated area and percentage share of LULC classes

LULC category	Colour code	Area (sq. km)	% of total area
Agriculture	Yellow	2780	50.24%
Settlements	Red	500	9.00%
Forest	Dark green	332	6.00%
Water bodies	Blue	193	3.48%
Grazing land	Light green	548	9.90%
Mining areas	Brown	1180	21.31%

Agriculture remains the most extensive land-use category, reflecting the district's continued dependence on farming, but its share has fallen markedly from earlier decades (Table 3), largely due to encroachment by settlements and mining. Mining areas occupy a substantial 21.3% of the district, concentrated in the eastern and southwestern zones, indicating an active and expanding extractive economy. Forest cover, at 6%, and grazing land, at 9.9%, together indicate ongoing pressure on the district's natural vegetation and pastoral resources, while water bodies (3.48%) remain the smallest category, consistent with the semi-arid setting.

5.2 Comparison with Historical LULC Trends

Table -3: Comparison of LULC class share across 2000, 2010, 2020, and 2024

LULC category	2000 (%)	2010 (%)	2020 (%)	2024 (%)
Agriculture	70	67	63	50.24
Settlements	7	9	12	9
Forest	12	11	10	6
Water bodies	5	5	5	3.48
Grazing land	6	6	8	9.9
Mining areas	0	2	2.5	21.31

The temporal comparison shows a consistent decline in agricultural share, from approximately 70% in 2000 to about 50% in 2024, alongside a sharp rise in mining area, from negligible levels in 2000 to over 21% by 2024. Settlement area increased through 2020 before the 2024 estimate returned closer to earlier levels, while forest cover has declined steadily across the period. These trends point to a structural shift in the district's land economy, from an agriculture-dominated pattern toward one increasingly shaped by mineral extraction, with corresponding implications for soil stability, vegetation cover, and water resources.

5.3 Accuracy Assessment

Classification accuracy was assessed using a confusion matrix comparing classified output against reference (ground-truth) samples for each LULC class, summarised in Table 4. Water bodies (95%) and settlements (88%) achieved the highest user's accuracy owing to their distinct spectral signatures, while forest and grazing land (80% each) showed comparatively lower accuracy due to spectral overlap between vegetated classes. The overall classification accuracy was approximately 86%, with an estimated Kappa coefficient of 0.83, indicating strong agreement between the classified map and reference data and confirming the reliability of the Random Forest and SVM classification approach.

Table -4: Confusion matrix and accuracy assessment for the 2024 LULC classification

Reference/Class	Agri.	Settl.	Forest	Water	Grazing	Mining	User's acc.
Agriculture	85	5	2	1	6	1	85%
Settlements	4	88	1	0	2	5	88%
Forest	3	2	80	2	8	5	80%
Water bodies	0	1	2	95	1	1	95%
Grazing land	6	3	7	2	80	2	80%
Mining areas	1	3	3	1	2	90	90%
Producer's acc.	85%	88%	80%	95%	80%	90%	Overall=86%

5.4 Discussion

The results indicate a landscape in transition: agriculture continues to anchor the district's economy but is under sustained pressure from both settlement expansion and, more markedly, from mining activity. The clustering of settlement growth around major towns and transport corridors, combined with encroachment into agricultural and forested land, points to the need for planned, vertically oriented urban development rather than unregulated sprawl. The scale of mining expansion warrants stronger environmental impact assessment, land reclamation, and monitoring, given its implications for soil erosion, vegetation loss, and water contamination. Declining forest cover and the moderate accuracy achieved for forest and grazing classes together suggest a need for targeted afforestation and more detailed ground-truthing in vegetated zones to improve both ecological outcomes and future classification accuracy. Stable but limited water body coverage, in a semi-arid setting with rising demand, further underscores the importance of water conservation and groundwater recharge measures.

These findings are consistent with the trends reported for other Indian districts undergoing similar transitions [1], [4], [6], and reinforce the value of combining moderate-resolution satellite imagery with machine learning classifiers for cost-effective, repeatable LULC monitoring in resource-constrained regions. The main limitation of the present analysis is its reliance on a single-date classification and estimated historical comparisons rather than a fully validated multi-temporal dataset; future work should extend the

approach using multi-temporal imagery and field-validated training data to strengthen change-detection accuracy.

6. CONCLUSIONS

This study presents a LULC classification of Koppal district for 2024 using Landsat-8 and Sentinel-2 imagery, Random Forest and SVM classifiers, and ArcGIS-based spatial analysis. Agriculture remains the dominant land-use category at approximately 50% of the district area, but its share has declined steadily since 2000 as settlement and, more significantly, mining activity have expanded, the latter now accounting for over one-fifth of the district. The classification achieved an overall accuracy of about 86% and a Kappa coefficient of approximately 0.83, confirming the reliability of the adopted machine-learning workflow.

The results highlight encroachment of settlements into agricultural land, continuing loss of forest and grazing land, and the growing environmental footprint of mining, all of which call for integrated, GIS-supported land-use planning that balances agricultural sustainability, controlled urban growth, and mineral resource development. The methodology developed here — combining multi-sensor satellite imagery, spectral and textural feature extraction, and machine-learning classification within ArcGIS — offers a replicable and scalable framework for LULC monitoring in other semi-arid, agriculture-dependent districts facing similar developmental pressures.

REFERENCES

- [1] S. Gopalakrishnan et al., "Land use and land cover change analysis in an urban expansion zone using remote sensing and GIS," 2015.
- [2] R. Patil et al., "Land suitability analysis for agriculture using GIS and multi-criteria decision analysis," 2017.
- [3] V. Sharma et al., "Assessment of deforestation trends and their impact on water resources using remote sensing," 2018.
- [4] A. Kumar et al., "Mapping urban sprawl and its environmental impact using machine learning and GIS," 2019.
- [5] P. Mehta et al., "GIS-based hydrological modeling for identification of flood-prone areas," 2020.
- [6] S. Reddy et al., "Land degradation assessment using remote sensing and GIS techniques," 2021.
- [7] J. R. Anderson, E. E. Hardy, J. T. Roach, and R. E. Witmer, "A Land Use and Land Cover Classification System for Use with Remote Sensor Data," U.S. Geological Survey Professional Paper 964, 1976.
- [8] R. G. Congalton and K. Green, Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, 3rd ed. Boca Raton, FL, USA: CRC Press, 2019.
- [9] G. M. Foody, "Status of land cover classification accuracy assessment," Remote Sensing of Environment, vol. 80, no. 1, pp. 185–201, 2002.
- [10] J. R. Jensen, Introductory Digital Image Processing: A Remote Sensing Perspective, 4th ed. Boston, MA, USA: Pearson Education, 2016.
- [11] T. Lillesand, R. W. Kiefer, and J. Chipman, Remote Sensing and Image Interpretation, 7th ed. Hoboken, NJ, USA: John Wiley & Sons, 2015.
- [12] National Aeronautics and Space Administration (NASA), "Landsat 9 Data Products User Guide," 2024.

- [13] European Space Agency (ESA), “Sentinel-2 User Handbook,” 2024.
- [14] QGIS Development Team, “QGIS Geographic Information System — Open Source GIS Software,” 2024.
- [15] Esri, “ArcGIS Pro Documentation,” 2024.
- [16] Indian Space Research Organisation (ISRO), “Bhuvan Geoportal for India's Remote Sensing Data,” 2024.
- [17] United Nations Environment Programme (UNEP), Global Environment Outlook 6: Land and Soil Systems, 2023.
- [18] World Bank, “Land Use Planning and Sustainable Development: Policy Guidelines,” 2023.
- [19] C. Giri, Remote Sensing of Land Use and Land Cover: Principles and Applications. Boca Raton, FL, USA: CRC Press, 2012.
- [20] Google, “Google Earth Engine: Cloud-Based Geospatial Analysis Platform,” 2024.
- [21] OpenStreetMap Contributors, “OpenStreetMap: Global Open GIS Data,” 2024.