Taraqqiyot spektri
ILMIY JURNALI

Vol 2 № 5 2026

Comparative Analysis of Supervised and Unsupervised Classification of Sentinel-2 Imagery for Land Use and Land Cover Mapping: A Case Study of Gurlan District, Uzbekistan

Masharipov Shakhzod

Master student, “Tashkent Institute of Irrigation and Agricultural Mechanization Engineers” National Research University, Tashkent, Uzbekistan

Romanov Jamshid

Master student, “Tashkent Institute of Irrigation and Agricultural Mechanization Engineers” National Research University, Tashkent, Uzbekistan

Abstract: Accurate land use and land cover (LULC) mapping is critical for sustainable land management, particularly in agricultural regions like Gurlan District, Uzbekistan. This study evaluates the performance of supervised (Random Forest, RF) and unsupervised (K-Means) classification methods using Sentinel-2 imagery for 2016, 2020, and 2024. Focusing on agricultural lands, built-up areas, and water bodies, the study assesses classification accuracy and detail in a semi-arid region. Results show that supervised classification outperforms unsupervised methods, with Sentinel-2 achieving higher overall accuracy (93% in 2024) due to its superior spatial resolution and spectral bands. Temporal analysis revealed a decline in agricultural lands and an increase in built-up areas. These findings underscore the efficacy of Sentinel-2 and supervised classification for precise LULC mapping, offering insights for regional planning and resource management.

Keywords: Gurlan District, Land Use and Land Cover, Sentinel-2, Supervised Classification, Unsupervised Classification, Remote Sensing

1. Introduction

Land use and land cover (LULC) mapping is essential for monitoring environmental changes, urban expansion, and agricultural productivity.

Remote sensing technologies, such as Landsat and Sentinel-2, provide cost-effective tools for large-scale LULC analysis. Gurlan District, located in Uzbekistan's Khorezm Region, is an agriculturally significant area reliant on irrigation systems, making it an ideal case study for evaluating remote sensing applications.

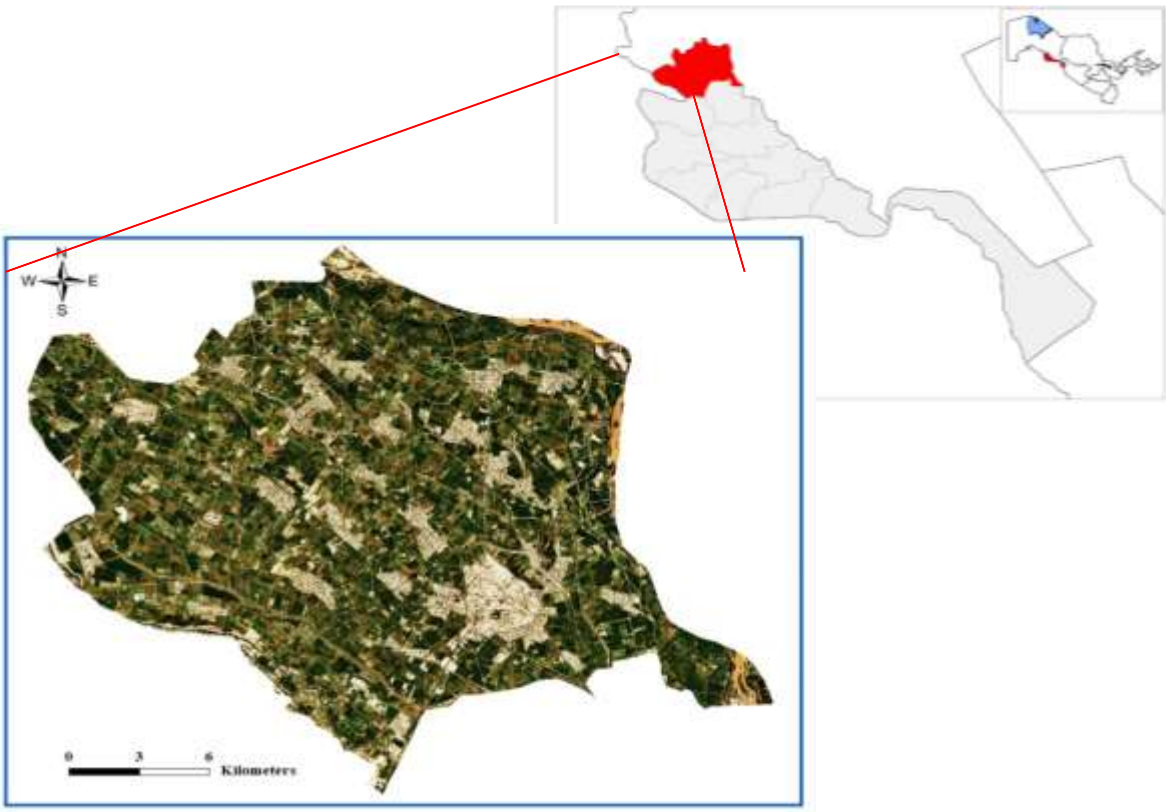
Land use maps are fundamental data sources for land planning and management [1,2]. Accurate and up-to-date land use/land cover (LULC) mapping has always been of interest to geoscience and remote sensing societies [3–5], mainly because it is a provider of valuable information to understand human–environment relationships [6,7]. The starting point for LULC mapping was that of using mono-temporal and mono-source satellite images [8]. Sentinel-2, with 10-20 m resolution and red-edge bands, provides enhanced detail for vegetation and water monitoring. Classification methods, such as supervised (e.g., Random Forest) and unsupervised (e.g., K-Means), differ in their reliance on training data and computational complexity, impacting their suitability for specific applications.

This study aims to compare the accuracy and detail of supervised and unsupervised classification methods using Sentinel-2 imagery for LULC mapping in Gurlan District over 2016, 2020, and 2024. The research questions are: (11) Which classification method provides higher accuracy for LULC mapping? (10) How do Sentinel-2 compare in terms of detail and precision? (9) What temporal LULC changes are observed in Gurlan District?

The rest of the paper is organized as follows. In Section 2, the study area, Data Sources, and methodologies of LULC classification and accuracy assessment are described. The results and analyses are presented in Section 3. A discussion is presented in relation to other studies in Section 4, and finally, the paper concludes in Section 5.

2. Materials and Methods

2.1 Study Area



Gurlan District (41.8°N, 60.4°E) in Khorezm Region, Uzbekistan, spans approximately 438.3 km² (170 sq mi) along the Amu Darya River. The region's economy is predominantly agricultural, with cotton, rice, and orchards as key land uses, supported by extensive irrigation networks. The capital lies at the town Gurlan. The relief consists of a slightly undulating plain, sloping toward the northwest. The district's climate is continental. The average temperature in July is 26–28 °C, and in January - 6 °C. The average annual precipitation is 100 mm. The vegetation period lasts 180–200 days. The Amu Darya River flows through the northern and northeastern parts of the district.

Figure 1. The geographic location of the Gurlan District. Blue box: the Sentinel-2 (S-2) true color composite of Gurlan District (RGB,256).

2.2 Data Sources

In this study, the S-2 Operational Land Imager (OLI) time series data were used for mapping the LULC classes. The Sentinel-2 Multispectral imagery with 10-20 m resolution, acquired for 2016, 2020, and 2024 from (https://sentinel.esa.int/web/sentinel/missions/sentinel-2) Copernicus Open Access Hub.

Ancillary Data: Local LULC maps, agricultural statistics, and Google Earth imagery for validation.

Images were selected from the vegetation season (June-July) with cloud cover <10%.

2.3 Data Preprocessing

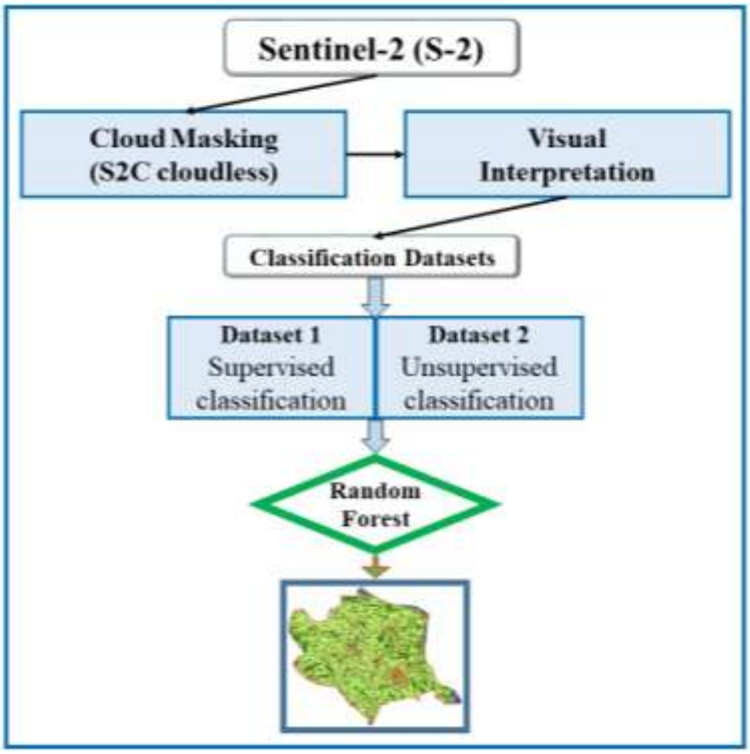
Atmospheric Correction: Sentinel-2 imagery was corrected using the Sen2Cor.

Geometric Alignment: Images were clipped to Gurlan District boundaries using ArcGIS 10.8.

2.4 Classification Methods

Supervised Classification: Random Forest (RF) was applied with 300 training samples per class (agricultural lands, Ricelands, Settlement (built-up areas) lands, waterbody lands, Roads and empty lands, Degradated lands and others), derived from Google Earth and local LULC maps. RF was implemented in ArcGIS 10.8.

Unsupervised Classification: K-Means clustering was used with eighteen clusters, later merged into the target six classes. The algorithm was executed in Google Earth Engine (GEE).



Class Definitions:

1. Agricultural Lands: Croplands and orchards.
2. Ricelands.
3. Settlement (built-up areas) lands: Residential and industrial zones.
4. Waterbody lands: Rivers, canals, and small lakes.
5. Roads and empty lands.
6. Degradated lands: Bare soil and sparse vegetation.

Figure 2. Flow chart of the methodology used in this study for land use/land cover (LULC) mapping.

2.5 Accuracy Assessment

Classification accuracy was evaluated using an error matrix, calculating Overall Accuracy (OA) and Kappa coefficients [13]. A stratified random sample of 100 validation points was compared against ground truth data from local sources and Google Earth.

$$K = \frac{P_0 - P_e}{1 - P_e}$$

Where:

K – The Kappa coefficient

P_0 – Observed accuracy (how often the classifier was correct)

P_e – Expected accuracy by chance.

Figure 4. The Kappa coefficient Interpretation table [13].

№	Kappa value	Agreement Level
1.	< 0	Poor
2.	0.01 - 0.20	Slight agreement
3.	0.21 – 0.40	Fair agreement
4.	0.41 – 0.60	Moderate agreement
5.	0.61 – 0.80	Substantial agreement
6.	0.81 – 1.00	Almost perfect agreement

Variable importance stands for the variables' contribution to distinguish between LULC classes, which helps by improving the classification accuracy while reducing data redundancy and processing workload. In this study, variable importance was derived from the RF model to estimate the contribution of variables (i.e., spectral bands and indices) to the obtained accuracy of the model [12].

3. Results

3.1 Supervised Classification

Figure 3. The Kappa coefficient **(**also known as Cohen's Kappa**)** [13].

Sentinel-2: OA was 90% (2016), 92% (2020), and 93% (2024), with Kappa coefficients of 0.87, 0.89, and 0.91, respectively. Sentinel-2's higher resolution improved the detection of small agricultural plots and irrigation canals.

3.2 Unsupervised Classification

Sentinel-2: OA was 82% (2016), 84% (2020), and 85% (2024), with Kappa coefficients of 0.78, 0.80, and 0.81, respectively.

Unsupervised classification struggled with distinguishing built-up areas due to spectral overlap with bare soil.

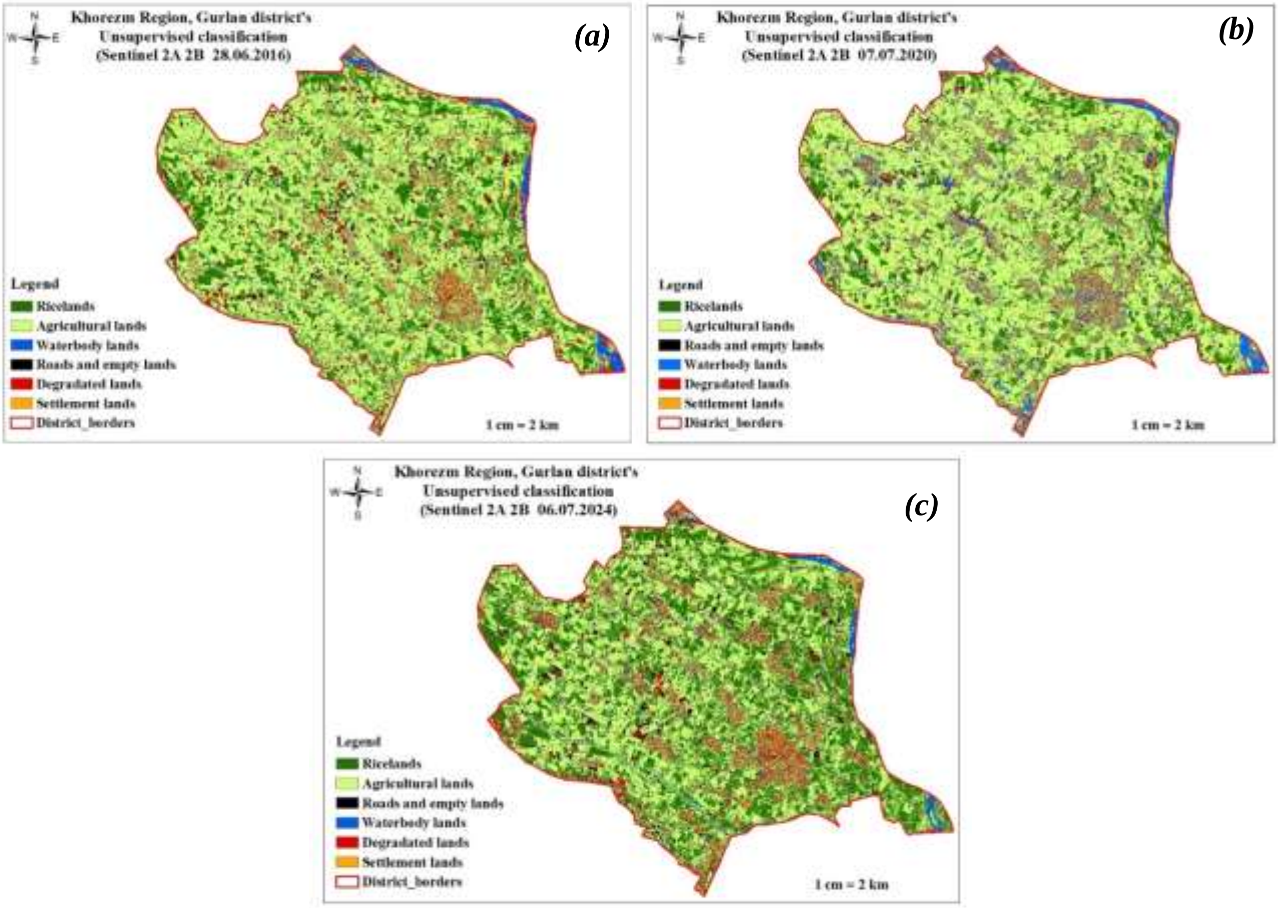


Figure 5, 6, 7. Land use/land cover (LULC) maps of the study area resulting from: **(a)** 2016 S-2 seasonal composites, **(b)** 2020 S-2 seasonal composites, **(c)** 2024 S-2 seasonal composites.

4. Discussion

The superior performance of supervised classification aligns with previous studies (e.g., Belgiu & Drăguţ, 2016), as RF leverages training data to reduce misclassification.

Sentinel-2's higher spatial resolution and red-edge bands enhance its ability to distinguish vegetation and water features, corroborating findings by Zhang et al. (2023).

Unsupervised classification, while faster, produced lower accuracy, particularly in heterogeneous landscapes like Gurlan District. The observed LULC trends highlight the need for integrated land-use planning to balance urbanization and agricultural sustainability.

Limitations include the reliance on cloud-free imagery and the need for extensive training data for supervised methods. Future research could explore hybrid classification approaches or integrate Sentinel-2 dataset for improved temporal resolution.

5. Conclusion

This study demonstrates that supervised classification using Random Forest with Sentinel-2 imagery provides the highest accuracy and detail for LULC mapping in Gurlan District, achieving up to 93% overall accuracy in 2024. The observed decline in agricultural lands and rise in built-up areas underscore the need for sustainable land management. Future studies should investigate the integration Sentinel-2 of data and machine learning advancements to enhance Land Use and Land Cover (LULC) monitoring in semi-arid regions.

References:

1. Esfandeh, S.; Danehkar, A.; Salmanmahiny, A.; Sadeghi, S.M.M.; Marcu, M.V. Climate Change Risk of Urban Growth and Land Use/Land Cover Conversion: An In-Depth Review of the Recent Research in Iran. Sustainability 2021, 14, 338. [CrossRef]

2. Yao, Y.; Yan, X.; Luo, P.; Liang, Y.; Ren, S.; Hu, Y.; Han, J.; Guan, Q. Classifying Land-Use Patterns by Integrating Time-Series Electricity Data and High-Spatial Resolution Remote Sensing Imagery. Int. J. Appl. Earth Observ. Geoinf. 2022, 106, 102664. [CrossRef]

3. Qian, X.; Zhang, L. An Integration Method to Improve the Quality of Global Land Cover. Adv. Space Res. 2022, 69, 1427–1438.[CrossRef]

4. Schulz, D.; Yin, H.; Tischbein, B.; Verleysdonk, S.; Adamou, R.; Kumar, N. Land Use Mapping Using Sentinel-1 and Sentinel-2 Time Series in a Heterogeneous Landscape in Niger, Sahel. ISPRS J. Photogramm. Remote Sens. 2021, 178, 97–111. [CrossRef]

5. Viana, C.M.; Girão, I.; Rocha, J. Long-Term Satellite Image Time-series for land use/land cover change detection using refined open source data in a rural region. Remote Sens. 2019, 11, 1104. [CrossRef]

6. Praticò, S.; Solano, F.; Di Fazio, S.; Modica, G. Machine Learning Classification of Mediterranean Forest Habitats in Google Earth Engine Based on Seasonal Sentinel-2 Time-Series and Input Image Composition Optimisation. Remote Sens. 2021, 13, 586. [CrossRef]

7. Sobhani, P.; Esmaeilzadeh, H.; Barghjelveh, S.; Sadeghi, S.M.M.; Marcu, M.V. Habitat Integrity in Protected Areas Threatened by LULC Changes and Fragmentation: A Case Study in Tehran Province, Iran. Land 2021, 11, 6. [CrossRef]

8. Steinhausen, M.J.; Wagner, P.D.; Narasimhan, B.; Waske, B. Combining Sentinel-1 and Sentinel-2 Data for Improved Land Use and Land Cover Mapping of Monsoon Regions. Int. J. Appl. Earth Observ. Geoinf. 2018, 73, 595–604. [CrossRef]

9. Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing, 114*, 24-31.[CrossRef]

10. Claverie, M., Ju, J., Masek, J. G., et al. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment, 219*, 145-161. [CrossRef]

11. ESA. (2024). Sentinel-2 user guide. Retrieved from [CrossRef]

12. Liu, C.; Li, W.; Zhu, G.; Zhou, H.; Yan, H.; Xue, P. Land Use/Land Cover Changes and Their Driving Factors in the Northeastern Tibetan Plateau Based on Geographical Detectors and Google Earth Engine: A Case Study in Gannan Prefecture. Remote Sens. 2020, 12, 3139. [CrossRef]

13. Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. [CrossRef]