Otokhonov O. D., master's student,
https://orcid.org/0009-0000-4972-8025
Ashirov Y. R., associate professor,
https://orcid.org/0009-0001-1358-7790
Nurullayev M. J., master's student,
https://orcid.org/0009-0003-0639-3626
Tashkent State Agrarian University

REMOTE SENSING AND MACHINE LEARNING FOR MAPPING IRRIGATED AGRICULTURAL LANDS (A CASE STUDY OF URGENCH DISTRICT)

Abstract. This study presents an operational framework for monitoring irrigated croplands and identifying agricultural species using Remote Sensing and Machine Learning. To process multi-temporal Sentinel-2 data for the heterogeneous Urgench district (Khorezm region), Google Earth Engine (GEE) was leveraged. An optimized Random Forest classifier was trained by integrating spectral indices (NDVI, NDWI, EVI, SAVI) with Red Edge channels. Post-classification, a focal mode filter was applied to reduce "salt-and-pepper" noise. Validation using an independent testing subset and field-collected control points demonstrated high accuracy and spatial precision in categorizing complex agricultural landscapes.

Keywords: Google Earth Engine (GEE), agricultural monitoring, Sentinel system, spectral vegetation indices, Machine Learning (Random Forest), multi-temporal mapping, irrigated croplands.

Introduction. Satellite data and geospatial platforms are essential for modern agricultural management. Open-source satellite programs like Sentinel, integrated with cloud architectures like Google Earth Engine (GEE), have modernized the continuous observation of large agrolandscapes [1, 2]. High temporal and multi-band sensing provide advanced capabilities for tracking plant phenology. This is critical in fragmented, densely irrigated agricultural zones like the Urgench district, where isolating specific land-use is challenging due to soil salinity, structural crop similarities, and very small field geometries that hinder standard classification [3]. Research indicates that relying solely on unitemporal optical indices often fails to separate structurally analogous vegetation [4]. Consequently, modern mapping is shifting toward multi-temporal optical chronologies processed by machine learning models [5]. Recent studies in Uzbekistan have validated these methodologies, with researchers identifying

Random Forest and SVM as highly effective classifiers for irrigated landscapes in the Khorezm and Tashkent provinces [4, 7].

Furthermore, the evolution toward adaptive feature-fusion networks has proven effective in large-scale cotton mapping across diverse Uzbek agro-climatic zones, achieving high precision even under limited sample conditions [6]. Moreover, recent studies emphasize that the accurate spatial monitoring of primary crops like cotton and winter wheat serves as a fundamental requirement for optimizing irrigation systems and enhancing water use efficiency, particularly under the increasingly water-scarce conditions characteristic of Uzbekistan [9]. While previous assessments reported high accuracies, the fragmentation of parcels in districts like Urgench continues to necessitate the 10-meter spatial precision provided by Sentinel-2.

Motivated by these shifts, this study develops a robust mapping protocol for the Urgench district. By fusing multi-temporal Sentinel-2 surface reflectance data with soil-compensating algorithms (SAVI, EVI) and dynamically tuning a Random Forest classifier followed by spatial focal filtering, the objective is to provide an empirical foundation for regional agricultural resource planning.

Methodology. Conducted within the GEE environment, this study utilized the Sentinel-2 Surface Reflectance (SR) Harmonized archive covering a complete agricultural cycle (March 1-September 30, 2025). Spectral integrity was maintained via systematic cloud and shadow masking using the Scene Classification Layer (SCL). To address heterogeneous agricultural challenges, a suite of spectral indices was computed. The Soil Adjusted Vegetation Index (SAVI) was prioritized to suppress background interference during early crop emergence, supported by NDVI, EVI, NDBI, and NDWI to concurrently map biomass, hydration, and soil exposure thresholds.

Month	Number of Scenes	Tiles Covered	Date Range
March	7	T40TGL	March 03 - March 28
April	7	T40TGL, T40TGM	April 02 - April 27
May	8	T40TGL	May 02 - May 27
June	7	T40TGL	June 01 - June 26
July	9	T40TGL	July 01 - July 31

August	6	T40TGL	August 05 - August 30
September	7	T40TGL, T41TKF	Sept 01 - Sept 29
Total	51	-	March - September

Table 1. Temporal distribution of Sentinel-2 satellite imagery used in the study (March - September 2025).

Based on this temporal density, seasonal median composites (April-May, June-July, August, September) were generated to track fine-scale phenology. During the August peak, Sentinel-2's Red Edge bands (B5, B6, B7) were included to capture variations in canopy chlorophyll content, thus differentiating structurally similar crop types.

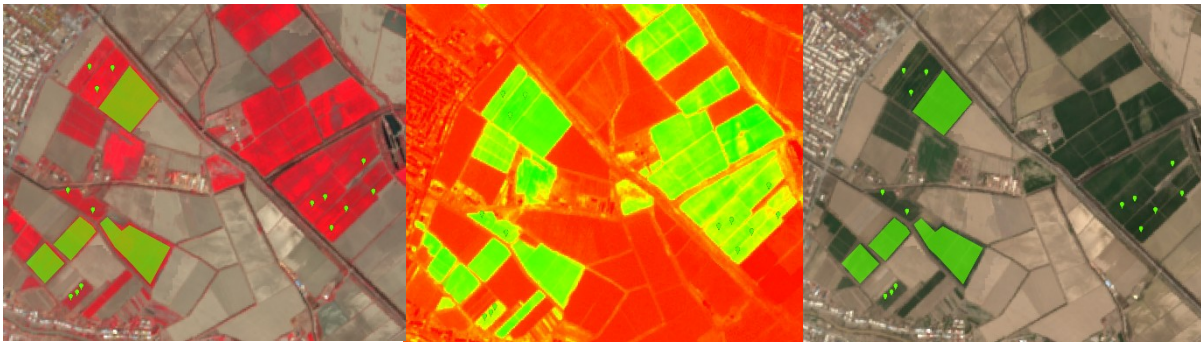

Figure 1: Spectral Comparison of False Color, NDVI, and Natural Color Imagery (April)

To calibrate the classifier, a geographically diverse ground-truth library of 50-60 reference points and polygons per class was established through grid-based visual photo-interpretation and field validation. Training data extraction relied on cross-examining Sentinel-2 False Color composites (B8-B4-B3) with corresponding NDVI profiles. For instance, flourishing winter wheat appeared as bright-red parcels in False Color and was verified against high-NDVI pixels. Conversely, water bodies were isolated via strongly negative NDVI and dark-blue False Color anomalies. This inferential logic, adapted from Otokhonov and Ashirov (2025) [8], was similarly applied to riparian forest networks. These polygons were converted to training pixels via the GEE "sampleRegions" protocol. The Random Forest (RF) algorithm was employed, chosen for its capacity to process high-dimensional temporal data. Utilizing a randomized 70/30 spatial split, 30% of the data was reserved for independent validation. Hyperparameters were tuned to maximize regional generalization: 300 decision trees were used to reduce variance, and a "minLeafPopulation" of 3 was set to

prevent model overfitting and promote the recognition of generalized agronomic signatures.

Results. To reduce localized 'salt-and-pepper' noise and enhance spatial consistency, a focal mode spatial filter with a 1.5-pixel radius was applied to the classification result. This step smoothed isolated misclassifications while preserving the boundaries of agricultural parcels. Statistical validation using a 30% independent testing subset yielded an Overall Accuracy of 86.0% and a Kappa Coefficient of 0.80. To complement these metrics, the results were visually and spatially validated against independent field control points (Figure 3). As demonstrated in the zoomed-in inset of Figure 3, the control points for cotton, wheat, and other crops show precise spatial agreement with the classified parcels, confirming the model's high thematic consistency. This performance is attributed to the multi-temporal Sentinel-2 features, though it is interpreted as internal statistical consistency rather than an exhaustive ground-truth survey. These results confirm the effectiveness of non-parametric Random Forest classifiers for the complex irrigated landscapes of Uzbekistan, as previously established by Basukala et al. (2017) [4]. While earlier studies (Basukala et al., 2017) relied on 30m Landsat-8 data, this study leverages the 10m spatial resolution of Sentinel-2, which substantially reduces spectral mixing challenges and further improves mapping precision across the region.

To validate that specific agricultural species exhibit distinct phenological patterns, this study analyzed the temporal progression of primary classifications using NDVI and NDWI metrics to characterize their seasonal growth curves.

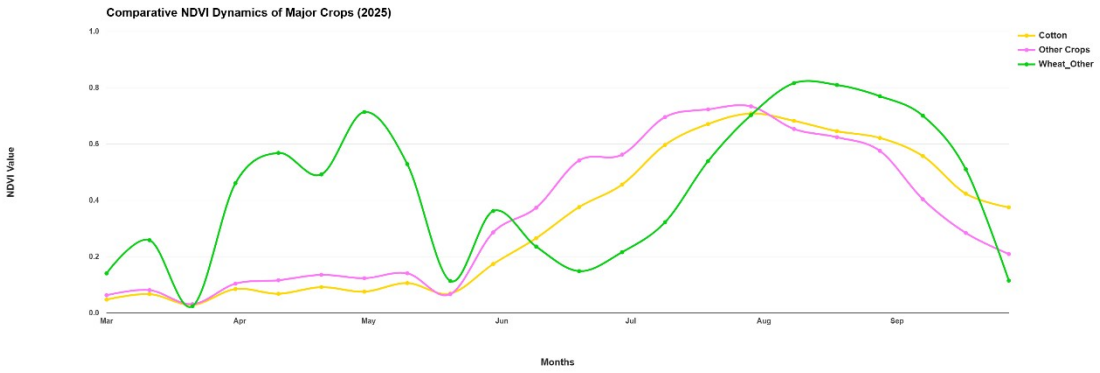

Figure 2: Temporal profiles of NDVI for 'Cotton', 'Wheat_Other', and 'Other Crops' across the 2025 agricultural cycle

Analysis of these phenological profiles (Figure 2) reveals distinct phenological trajectories. Commercial cotton shows a monophasic arc: NDVI scales upward from April, crests past 0.7 in August, and declines during late-September defoliation, while NDWI remains persistently negative. Conversely, intensive double-cropping parcels (winter wheat/catch crops) demonstrate a pronounced diphasic (double-cresting) waveform. A primary vegetative peak in May (NDVI ≈ 0.7) signifies wheat maturity, followed by a sharp June harvest decrease and an immediate secondary NDVI acceleration (>0.8) in August, accurately reflecting regional secondary planting regimens. "Other crops" display a delayed vegetative ascent, plateauing through late summer. Distinguishing these subtle delays supports the use of the multi-temporal framework.

Applying this feature base, the RF model classified the Urgench district's physical landscape. The raw output was subjected to the 1.5-pixel focal filter, removing spatial noise to generate a refined map (Figure 3).

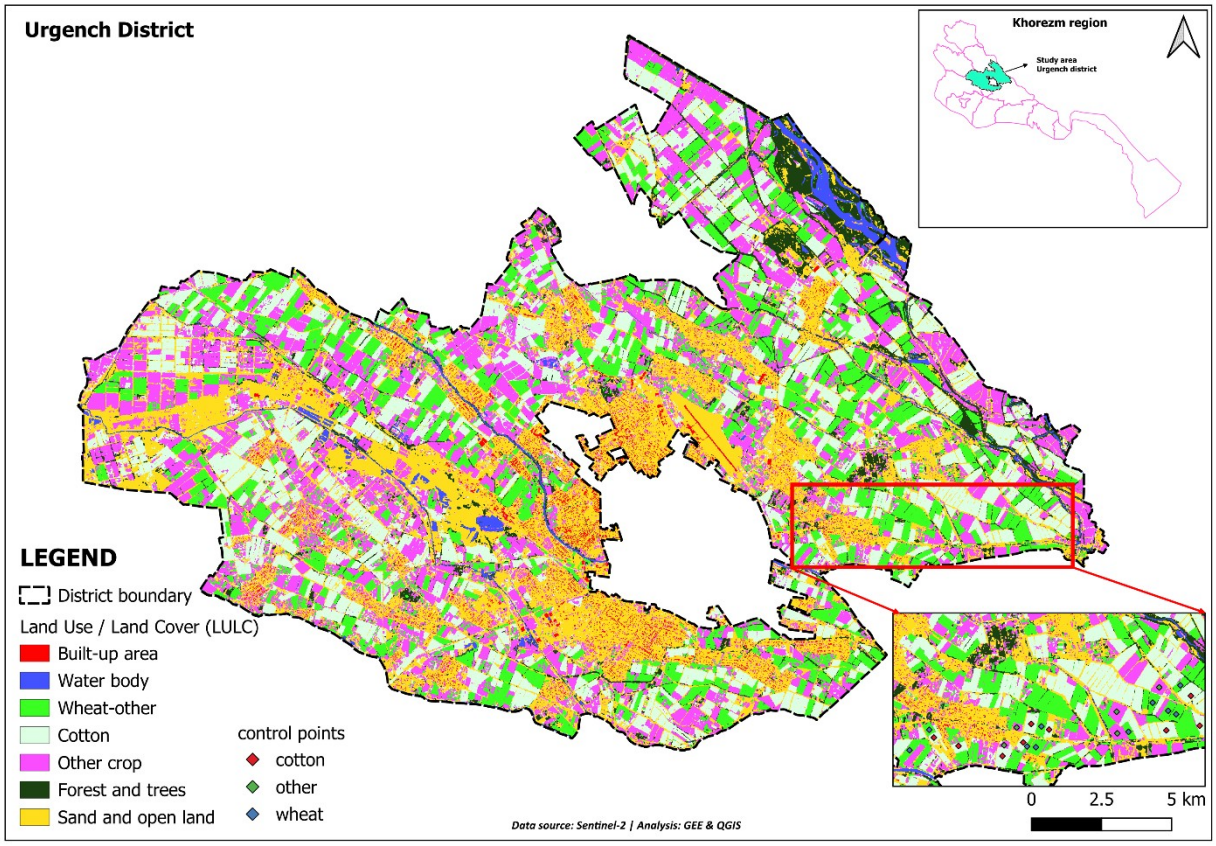

Figure 3: Regional Land Use Classification Map of Urgench District

The finalized classification results were exported from the Earth Engine environment and processed within QGIS for cartographic visualization. Cartographic assessment in QGIS confirms the structural precision of the model, segmenting the extensive agricultural footprint into defined parcels for cotton expanses, complex wheat rotations, and diversified gardens. The map visually validates the effectiveness of the focal mode filter; localized noise is replaced by homogeneous areas that isolate precise ecological and parcel boundaries, resulting in a high-quality cartographic product.

Surface area estimations, calculated in the UTM Zone 41N projection for geometric precision, total 43,881.25 hectares across seven typologies, with active croplands occupying 62.0% of the territory. Among the specific agricultural typologies, 'Other Crops' (Class 5) represent the largest share at 10,556.79 hectares (24.06%), underscoring a significant trend of secondary agricultural diversification in the region. Cotton (Class 4) remains a strategic monoculture, covering 9,624.30 hectares (21.93%) of the landscape.

Furthermore, Winter Wheat and Second Crops (Class 3) encompass 7,010.57 hectares (15.98%); the multi-temporal mapping protocol successfully captured the complex rotational geometries and diphasic growth curves characteristic of these intensive agricultural systems. Regarding natural environments and physical infrastructure (38.0% of the landscape), Sand, Saline, and Open Land (Class 7) is the largest single class at 12,167.18 hectares (27.73%). Notably, this captures both extensive desert areas and narrow bare-earth perimeter tracks separating fields, demonstrating the focal filter’s high spatial precision.

Forest and Trees (Class 6) occupy 2,766.23 hectares (6.30%), isolating mature orchards and protective vegetative shelterbelts. Finally, Built-up area (Class 1) accounts for 788.90 hectares (1.80%), and Water bodies (Class 2) accounts for 967.28 hectares (2.20%).

No.	Types	Area (hectares)
1	Built-up area	788.90
2	Water body	967.28
3	Wheat and other	7010.57
4	Cotton	9624.30

5	Other crop	10556.79
6	Forest and trees	2766.23
7	Sand and open land	12167.18
	Total:	**43881.25**

Table 2. Areal Statistics of Land Cover Categories

Conclusion. This research demonstrates that the integrated use of multi-temporal Sentinel-2 data and machine learning provides a reliable framework for mapping complex irrigated agrolandscapes in the Urgench district. The inclusion of Red Edge parameters and seasonal composites effectively captures the phenological diversity of regional crops, reducing classification errors. To further enhance this methodology, future work should focus on expanding the ground-truth database across more extensive regional scales and integrating land parcel (cadastral) data to achieve more precise spatial refinement of field boundaries. Applying this protocol in practice will provide an objective tool for the continuous monitoring of land resources and agricultural crops, supporting the sustainable management and effective administration of regional soil and water resources.

References

1. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18-27. https://doi.org/10.1016/j.rse.2017.06.031

2. Mutanga, O., & Kumar, L. (2019). Google Earth Engine Applications. Remote Sensing, 11(6), 591. https://doi.org/10.3390/rs11060591

3. Remelgado, R., Zaitov, S., Kenjabaev, S., Stulina, G., Sultanov, M., Ibrakhimov, M., Akhmedov, M., Dukhovny, V., & Conrad, C. (2020). A crop type dataset for consistent land cover classification in Central Asia. Scientific Data, 7(1), 250. https://doi.org/10.1038/s41597-020-00591-2

4. Basukala, A. K., Oldenburg, C., Schellberg, J., Sultanov, M., & Dubovyk, O. (2017). Towards improved land use mapping of irrigated croplands: Performance assessment of different image classification algorithms and approaches. European Journal of Remote Sensing, 50(1), 187-201. https://doi.org/10.1080/22797254.2017.1308235

5. Belgiu, M., & Csillik, O. (2018). Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. Remote Sensing of Environment, 204, 509-523. https://doi.org/10.1016/j.rse.2017.10.005

6. Jaloliddinov, J., Tian, X., Bai, Y., Guo, Y., Chen, Z., Li, Y., & Wang, S. (2024). Large-scale cotton classification under insufficient sample conditions using

an adaptive feature network and Sentinel-2 imagery in Uzbekistan. *Agronomy*, *14*(1), 75. https://doi.org/10.3390/agronomy14010075

7. Erdanaev, E., Kappas, M., & Wyss, D. (2022). Irrigated crop types mapping in Tashkent Province of Uzbekistan with remote sensing-based classification methods. *Sensors*, *22*(15), 5683. https://doi.org/10.3390/s22155683

8. Otokhonov, O. D., & Ashirov, Y. R. (2025). SHORT-TERM VEGETATION COVER CHANGE ANALYSIS (2020-2025) IN THE LOWER AMU DARYA STATE BIOSPHERE RESERVE USING SENTINEL-2 DATA. Экономика и социум, (12-1 (139)), 356-360. https://doi.org/10.24412/2225-1545-2025-12-1139-356-360

9. Isaev, S., Mambetnazarov, A., Khalmuratova, B., Goziev, G., & Ashirov, Y. (2022). Efficiency of appropriate irrigation system of cotton and winter wheat in water scarce conditions of Uzbekistan. IOP Conference Series: Earth and Environmental Science, 1068(1), 012044. https://doi.org/10.1088/1755-1315/1068/1/012044