Intelligent Aerospace Image Processing for Land Use Identification and Smart Grid Integration

Adejor E. Abiche[*], Leila Rzayeva[**,***‡], Hamada Mohamed[****], Korhan Kayisli[*****], Nursultan Nyssanov[**], Kozhakhmet Zhaksylyk[**]

[*]Department of Intelligent Systems and Cybersecurity, Astana IT University, Astana, Mangilik El avenue, 55/11, Business center EXPO, block C1, Kazakhstan
[**]Research and Innovation Center "CyberTech", Astana IT University, Astana, Mangilik El avenue, 55/11, Business center EXPO, block C1, Kazakhstan
[***]Research and Innovation Institute "Digital Heritage of Eurasia" LLP, Astana 010000, str. Zhetigen #17, Kazakhstan
[****]Military Program, Abu Dhabi University, UAE
[*****] Gazi University, Eng. Fac. Electrical-Electronic Eng. Ankara, Turkiye
(E.Adejor_abiche@astanait.edu.kz, l.rzayeva@astanait.edu.kz, m.hamada@iitu.edu.kz, korhankayisli@gmail.com, nnysanov@gmail.com, zh.kozhakhmet@astanait.edu.kz)

[‡]Corresponding Author; Leila Rzayeva, Astana, Mangilik El avenue, 55/11, Business center EXPO, block C1, Kazakhstan, Tel: +7 777 533 9169, l.rzayeva@astanait.edu.kz

Received: 22.08.2025 Accepted: 23.10.2025

Abstract- This study develops an automated framework for land use classification using Landsat 8 imagery processed on Google Earth Engine (GEE), directly applied to smart grid planning. The methodology combines spectral indices (NDVI, NDWI, NDBI) with optical bands(B2-B7) and applies k-means clustering for unsupervised classification into five land cover types: urban, agricultural, forest, water, and barren areas. Applied to Astana, Kazakhstan [71.19-71.61°E, 51.025-51.175°N], the classification reveals agricultural dominance (38.5%), significant urban coverage (31.2%), and substantial forest areas (23.7%). An interactive visualization interface with dynamic inspection tools enhances stakeholder accessibility. Beyond conventional environmental monitoring, classified outputs directly inform smart grid development: urban zones indicate high-load demand areas for electric vehicle infrastructure; agricultural land represents biomass energy potential; water bodies support hydropower assessment; barren areas identify solar farm sites. By quantitatively linking land cover data with energy metrics, this framework provides actionable inputs for renewable integration, demand forecasting, and sustainable urban energy planning. Future research will integrate classified datasets with smart grid simulation platforms for distributed generation optimization.

Keywords Aerospace image processing, smart grid integration, remote sensing, land use identification, google earth engine, machine learning, urban energy planning.

1. Introduction

Rapid urbanization in developing regions presents unprecedented challenges for sustainable infrastructure planning and energy system optimization, demanding innovative approaches that can scale with accelerating urban growth while maintaining environmental sustainability. Astana, Kazakhstan's capital, exemplifies these contemporary urban challenges, having experienced dramatic transformation since its designation as capital in 1997, evolving from a regional center to a modern metropolis of over one million residents [1, 2]. This rapid development trajectory, while creating opportunities for economic growth and modernization, simultaneously generates complex spatial planning requirements that traditional monitoring approaches struggle to address effectively.

The conventional methodologies for land use monitoring and infrastructure planning face fundamental scalability limitations that become particularly acute in rapidly developing urban contexts. Traditional ground surveys and manual aerial photograph interpretation, while providing detailed local information, require extensive human

resources, specialized equipment, and prolonged field campaigns that often prove prohibitively expensive and time-consuming for large-scale urban monitoring applications [2, 3]. These resource constraints severely restrict the frequency and spatial coverage of monitoring activities, creating information gaps that compromise planning effectiveness precisely when timely, comprehensive spatial data becomes most critical for guiding sustainable development decisions [4, 5].

Recognizing these limitations, the remote sensing community has increasingly turned toward automated, satellite-based approaches that offer scalable alternatives for comprehensive land use analysis. Multispectral remote sensing technology provides a transformative solution by leveraging electromagnetic spectrum characteristics to distinguish between various land cover types with high accuracy and minimal human intervention, enabling systematic monitoring across vast geographical areas at regular temporal intervals [6, 7]. Technology's particular strength lies in its ability to capture subtle spectral signatures that characterize different surface materials, facilitating robust classification even in complex landscapes where visual interpretation becomes challenging due to mixed pixel effects, seasonal variations, or atmospheric interference [8, 9].

Building upon these fundamental spectral analysis capabilities, researchers have developed sophisticated indices that enhance the discrimination potential of multispectral imagery beyond what individual spectral bands can achieve alone. The Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), and Normalized Difference Built-up Index (NDBI) represent particularly powerful tools for distinguishing between vegetation, water bodies, and built-up zones respectively, exploiting specific spectral response characteristics that remain consistent across diverse geographical contexts [4, 5, 10, 11]. These indices effectively normalize for illumination variations and atmospheric effects while amplifying the spectral contrasts that enable reliable automated classification, forming the foundation for scalable land cover analysis methodologies [12, 13].

However, the implementation of these spectral analysis techniques faces a critical challenge in developing regions: the scarcity of ground-truth data required for supervised classification approaches. This limitation has driven significant advances in unsupervised machine learning methodologies for remote sensing applications, with k-means clustering algorithms emerging as particularly valuable tools for extracting meaningful spatial patterns without requiring extensive prior labeling or training datasets [6, 7, 14]. This capability proves especially effective in regions where ground-truth data collection faces accessibility, security, or resource constraints, addressing a fundamental gap in conventional supervised classification approaches that typically demand extensive field validation campaigns [15, 16].

The computational challenges associated with processing large-scale satellite datasets have been revolutionized by the emergence of cloud-based processing platforms, with Google Earth Engine leading this transformation by providing unprecedented processing capabilities through distributed computing infrastructure [10, 17]. This technological advancement enables researchers and practitioners to conduct systematic analysis across multiple temporal periods and geographical regions without requiring substantial local computational resources, effectively democratizing access to advanced spatial analysis capabilities. Simultaneously, the platform's integration of visualization tools and interactive analysis capabilities has enhanced stakeholder engagement by making complex analytical results more accessible to diverse user communities, from technical specialists to policy makers and community representatives [18].

The convergence of these technological capabilities occurs at a time when contemporary urban development challenges increasingly demand integrated approaches that transcend traditional disciplinary boundaries, particularly as cities worldwide grapple with energy transition requirements and climate change adaptation needs. The intersection of land use planning and energy system design has emerged as a particularly critical area where spatial analysis can provide essential insights for sustainable development. Modern smart grids require detailed spatial data for optimizing the alignment between energy generation and consumption patterns, determining optimal sites for distributed generation deployment, and designing effective demand-side management strategies that account for the spatial distribution of energy demands and resources [19, 20]. This integration represents a fundamental shift from sectoral planning approaches toward holistic spatial-energy planning frameworks that recognize the intimate connections between land use patterns and energy system requirements [21, 22].

The spatial characteristics of different land cover types provide valuable, yet often underutilized, insights for comprehensive energy infrastructure planning and renewable resource assessment. Urban areas correspond directly to high-load demand zones characterized by intensive building energy consumption and growing electric vehicle charging infrastructure requirements, enabling spatially explicit load density estimation that supports distribution system planning [23, 24]. Meanwhile, agricultural and forest regions indicate potential biomass production areas and carbon sequestration capabilities, while also representing seasonal energy demand variations associated with agricultural activities and rural electrification needs [9, 25]. Water bodies provide critical information for assessing hydropower opportunities and cooling infrastructure requirements for thermal power generation, particularly relevant in semi-arid regions where water resources represent both constraints and opportunities [26, 27]. Even barren and undeveloped areas offer valuable insights into potential locations for large-scale renewable energy installations, where land use conflicts are minimized and development costs are reduced.

Therefore, this study addresses the convergence of these technological capabilities and practical needs by developing a comprehensive automated classification framework that systematically bridges the gap between remote sensing capabilities and energy planning requirements through quantitative relationships and validated methodologies. The

research utilizes Landsat 8 imagery processed through Google Earth Engine to create an integrated land cover classification and smart grid planning system that combines spectral indices calculation, k-means clustering algorithms, and interactive visualization tools [10]. Through establishing quantitative linkages between classified land cover outputs and energy planning parameters, including load density estimation, renewable resource assessment, and infrastructure siting optimization, the framework transforms traditional environmental monitoring outputs into actionable information for contemporary energy infrastructure planning [28]. Applied to the Astana, Kazakhstan case study, this methodology demonstrates both the technical feasibility and practical value of integrated spatial-energy planning approaches for rapidly developing cities worldwide, providing a scalable model for sustainable urban development that balances growth requirements with environmental sustainability and energy security objectives.

2. Literature Review

The evolution of remote sensing technology represents one of the most significant advances in spatial analysis capabilities over the past five decades, fundamentally transforming how researchers approach urban growth analysis, ecological monitoring, and large-scale environmental assessment across diverse geographical contexts [29, 30]. This transformation began with the launch of the Landsat program in the 1970s, which established the foundation for continuous, systematic Earth observation and created an unprecedented archive of temporal data that enables comprehensive understanding of landscape transformations. The long-term data continuity provided by successive Landsat missions has proven invaluable for tracking urban sprawl, agricultural development, and environmental changes in cities worldwide with temporal depth and spatial consistency that would be impossible through alternative monitoring approaches [31].

Building upon this foundation of systematic Earth observation, multispectral imagery analysis has emerged as a particularly powerful tool for understanding complex urban development patterns and their cascading environmental implications. Researchers have successfully demonstrated the utility of Landsat datasets for quantifying urban expansion rates, assessing impacts on surrounding agricultural areas, and evaluating the environmental consequences of rapid development across diverse geographical and climatic contexts [8, 32]. This success has been further enhanced through the integration of specialized vegetation, water, and built-up indices that improve feature discrimination capabilities across diverse landscape types, exploiting fundamental differences in electromagnetic spectrum reflectance characteristics to enable automated distinction between natural and anthropogenic features [4, 5, 32, 33]. These spectral indices have proven particularly valuable because they maintain consistent performance across varying illumination conditions and atmospheric states, providing reliable classification capabilities even in challenging environmental conditions [34, 35].

The methodological evolution in aerospace image classification has paralleled these technological advances, transitioning from labor-intensive manual interpretation approaches toward sophisticated data-driven techniques supported by increasingly powerful machine learning algorithms. This transformation reflects broader trends in computational science and has been driven by the confluence of increasing data volumes, enhanced computational capabilities, and growing recognition of the need for more objective and reproducible analysis methods [36 - 39]. Among the various machine learning approaches that have emerged, unsupervised algorithms such as k-means clustering have gained particular prominence due to their ability to extract meaningful spatial patterns without requiring extensive prior labeling or training datasets. This capability addresses a critical limitation of supervised approaches, making unsupervised methods especially valuable for applications in developing countries and remote areas where ground-truth information may be limited by accessibility, security concerns, or resource constraints [27, 28, 40].

The computational infrastructure supporting these methodological advances has undergone perhaps the most dramatic transformation with the emergence of cloud-based processing platforms, exemplified by Google Earth Engine (GEE), which has revolutionized remote sensing capabilities by providing unprecedented computational capacity for processing vast datasets through distributed computing architectures [10, 41]. This technological breakthrough enables researchers to conduct near-real-time analysis of land cover dynamics across multiple spatial and temporal scales without requiring substantial local computational resources, effectively democratizing access to sophisticated remote sensing tools that were previously available only to well-resourced institutions. The platform's integration of advanced visualization tools, including interactive legends and dynamic inspection capabilities, has simultaneously enhanced stakeholder engagement by making complex analytical results more accessible to diverse user communities, from technical specialists to policy makers and community representatives [42].

These technological and methodological advances have occurred alongside a significant expansion in remote sensing applications, extending far beyond traditional environmental monitoring toward energy-related domains and smart city initiatives that reflect contemporary urban sustainability challenges [41-43]. Urban land cover data derived from remote sensing analyses have been increasingly utilized to evaluate solar rooftop potential, estimate building heating and cooling demands, and assess optimal placement strategies for distributed renewable generation systems, demonstrating the growing recognition of spatial data's critical importance in energy system planning and optimization. This expansion has proven particularly relevant as cities worldwide transition toward sustainable energy systems and implement smart grid technologies that require detailed spatial information for optimal performance. Agricultural zones identified through spectral indices analysis have been systematically examined for their biomass energy supply potential, providing quantitative assessments

that support renewable energy development planning [9]. Similarly, water indices contribute valuable information to hydrological modeling efforts relevant for hydroelectric production assessment, while forest area delineation provides insights into carbon sequestration capabilities and biomass availability for renewable energy production [5]. Even the identification of barren and undeveloped areas contributes to energy planning by revealing opportunities for utility-scale solar and wind energy installations where land use conflicts are minimized [20, 21].

This convergence of technological capabilities and expanding application domains has established the foundation for contemporary research that increasingly recognizes the dual role of aerospace image processing: enhancing traditional land use classification capabilities while simultaneously providing essential spatial datasets for smart grid integration and energy system optimization [17]. By positioning image analysis within the broader context of energy infrastructure development, current research demonstrates the technology's potential to influence both urban sustainability initiatives and modern power system design strategies. However, this integration potential remains largely unrealized due to the lack of systematic frameworks that can effectively bridge the gap between remote sensing outputs and energy planning requirements, creating an opportunity for methodological innovation that can unlock the full potential of spatial-energy system integration for sustainable urban development.

3. Research Methodology

3.1. Study Area and Data Acquisition

The research focuses on Astana, Kazakhstan, with coordinates [71.19°E, 51.025°N] to [71.61°E, 51.175°N], covering a significant portion of the capital city. This rectangular study region encompasses diverse landscapes including urban centers, agricultural areas, forest regions, water bodies, and barren land, making it suitable for comprehensive land cover classification and energy planning analysis. The area has experienced rapid development and urbanization since its designation as the capital in 1997 [22], representing a typical example of rapidly developing cities in Central Asia requiring integrated spatial-energy planning approaches.

The selection of Astana as the study area is justified by several factors that make it representative of rapidly developing urban centers in emerging economies. First, the city exhibits diverse land cover types within a relatively compact geographical area, providing an ideal testbed for multi-class classification algorithms. Second, the ongoing urban expansion creates dynamic land use patterns that require continuous monitoring for effective planning. Third, the city's strategic importance as Kazakhstan's capital ensures availability of supporting data and validation resources for research purposes.

The study area boundary is shown in Figure 1. It was established to capture both the urban core and surrounding rural areas, ensuring representation of the complete urban-rural transition zone. This approach enables comprehensive analysis of land use gradients from dense urban development through peri-urban areas to agricultural and natural landscapes. The rectangular boundary simplifies data processing while maintaining geographical coherence and administrative relevance for planning applications.

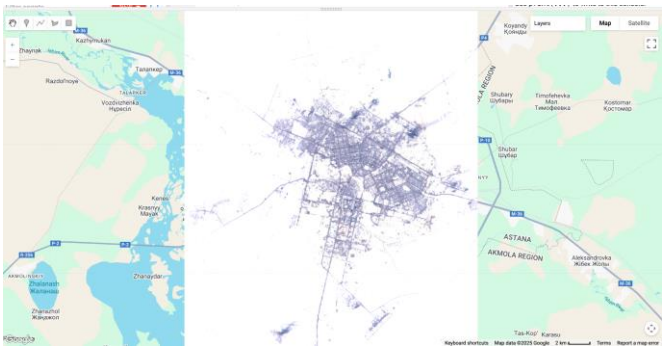

Fig. 1. Study area and region.

The study area encompasses approximately 450 km² of diverse terrain, providing sufficient spatial coverage for robust statistical analysis while remaining computationally manageable for processing and analysis. The area includes the central business district, residential neighborhoods, industrial zones, recreational areas, and extensive agricultural lands, ensuring comprehensive representation of land cover types relevant to energy planning applications. This diversity enables development of transferable methodologies applicable to similar urban contexts worldwide.

Landsat 8 imagery from the LANDSAT/LC08/C02/T1_TOA collection was selected as the primary data source due to its optimal balance of spatial resolution, spectral coverage, and temporal availability. The 30-meter spatial resolution provides sufficient detail for regional-scale analysis while maintaining computational efficiency for large-area processing. The satellite's multispectral capabilities, including visible, near-infrared, and shortwave infrared bands, enable comprehensive spectral analysis necessary for accurate land cover discrimination.

Data acquisition followed a systematic approach to ensure temporal consistency and quality control. The temporal window was restricted to 2017 to minimize phenological variations and maintain consistent atmospheric conditions across the study area. Cloud cover assessment was implemented as a primary quality criterion, with imagery selection prioritizing scenes with less than 10% cloud coverage to ensure clear surface observation.

3.2. Methodology Framework and Workflow

The research methodology follows a systematic seven-step workflow from raw satellite data acquisition to smart grid integration applications, as detailed in Figure 2. The framework incorporates multiple quality control measures to ensure data reliability and analytical accuracy while maintaining scalability for application across different geographical contexts and temporal periods. The automated approach reduces human resource requirements and processing time while ensuring consistent analytical standards.

The comprehensive workflow begins with Landsat 8 data acquisition, followed by preprocessing operations including cloud filtering and atmospheric correction. Spectral index calculation forms the core feature extraction phase, generating NDVI, NDWI, and NDBI indices that enhance land cover discrimination capabilities. The methodology then proceeds through feature stack creation, k-means clustering implementation, classification validation, and finally smart grid integration where land cover outputs are systematically linked to energy planning parameters.

The preprocessing pipeline, illustrated in Figure 2, incorporates temporal filtering, cloud cover assessment, and spatial clipping operations to optimize data quality. Atmospheric correction procedures convert top-of-atmosphere reflectance values to surface reflectance, accounting for atmospheric effects that can significantly impact spectral analysis accuracy. Quality assessment procedures ensure data reliability through systematic evaluation of sensor calibration, geometric accuracy, and radiometric consistency.

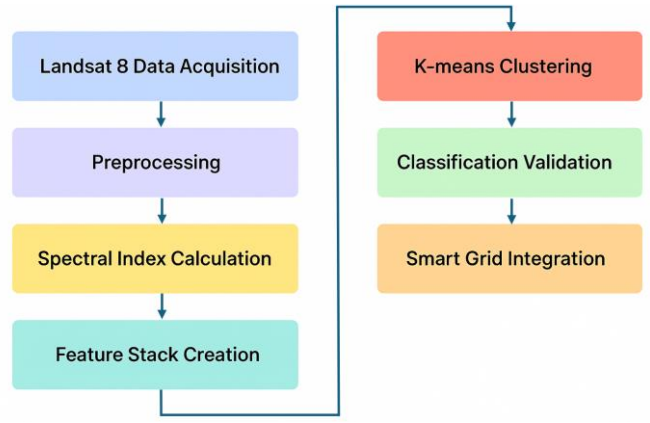

Fig. 2. Research methodology workflow.

The workflow integrates multiple processing stages that transform raw satellite observations into actionable energy planning information. Each processing step includes validation procedures to ensure quality and consistency. The modular design enables selective application of different components based on specific planning requirements and data availability, supporting diverse applications from preliminary feasibility studies to detailed infrastructure planning.

Implementation through Google Earth Engine provides distributed computing capabilities and standardized algorithms that ensure reproducibility across different research environments. The cloud-based approach enables systematic analysis of multiple cities and regions while maintaining consistent processing standards and quality assurance protocols. The automated workflow significantly reduces manual intervention requirements while maintaining high analytical standards and enabling scalable application across diverse geographical contexts.

Spectral Analysis and Classification

Spectral index calculation transforms raw multispectral reflectance data into meaningful indicators of land cover characteristics through established mathematical relationships. The methodology implements three principal spectral indices specifically selected for their proven effectiveness in distinguishing vegetation, water, and built-up areas in diverse environmental contexts. Each index exploits unique spectral properties of surface materials to enhance discrimination capabilities beyond those achievable through individual spectral bands alone.

The Normalized Difference Vegetation Index (NDVI) quantifies vegetation cover and health through the normalized difference between near-infrared and red reflectance:

$$\text{NDVI} = \frac{\rho_{NIR}-\rho_{Red}}{\rho_{NIR}+\rho_{Red}} = \frac{B_5-B_4}{B_5+B_4} \quad (1)$$

Where, ρ_{NIR} represents near-infrared reflectance (Landsat 8 Band 5) and ρ_{Red} represents red reflectance (Band 4). The normalization process creates index values ranging from -1 to +1, with higher positive values indicating dense, healthy vegetation and negative values typically representing non-vegetated surfaces.

The Normalized Difference Water Index (NDWI) identifies water bodies through the spectral contrast between green and near-infrared reflectance:

$$NDWI = \frac{\rho_{Green}-\rho_{NIR}}{\rho_{Green}+\rho_{NIR}} = \frac{B_3-B_5}{B_3+B_5} \quad (2)$$

Where, ρ_{Green} represents green band reflectance (Band 3). Water exhibits higher reflectance in visible wavelengths compared to near-infrared, where strong absorption occurs.

The Normalized Difference Built-up Index (NDBI) distinguishes urban areas through the spectral properties of constructed materials:

$$NDBI = \frac{\rho_{SWIR}-\rho_{NIR}}{\rho_{SWIR}+\rho_{NIR}} = \frac{B_6-B_5}{B_6+B_5} \quad (3)$$

Where, ρ_{SWIR} represents shortwave infrared reflectance (Band 6). Built-up surfaces typically exhibit higher shortwave infrared reflectance compared to vegetation.

Unsupervised classification employs the k-means algorithm to partition the multidimensional feature space into spectrally distinct land cover classes. The algorithm minimizes the within-cluster sum of squared distances through iterative optimization:

$$J = \sum_{k=1}^{K} \sum_{x_i \in C_k} ||x_i - \mu_k||^2 \quad (4)$$

where K=5 represents the number of clusters, C_k denotes cluster k, x_i represents the feature vector for pixel i, and μ_k is the centroid of cluster k. The algorithm iteratively updates cluster assignments and centroid positions until convergence is achieved when centroid movements fall below a specified threshold ($\varepsilon = 10^{-4}$)

4. Results and Analysis

4.1. Classification Performance and Land Cover Distribution

The k-means clustering algorithm successfully partitioned the study area into five spectrally distinct land cover classes, demonstrating clear spatial patterns that correspond to expected landscape characteristics observed in the Astana region. The classification process achieved stable convergence within 25 iterations, indicating robust algorithm performance and appropriate parameter selection for the given dataset and study area characteristics. The quantitative distribution of classified land cover types is summarized in Figure 3, revealing the diverse landscape composition of the study area.

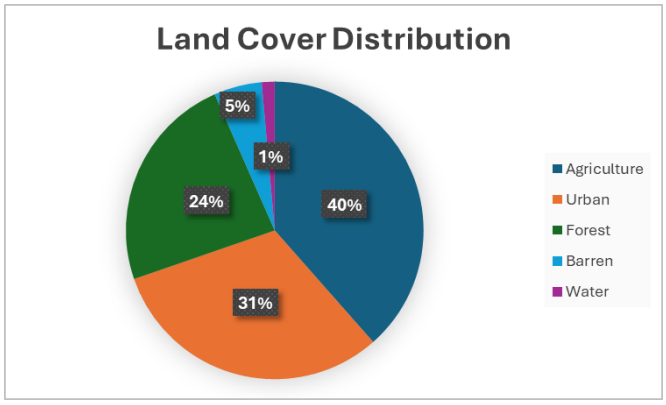

Fig. 3. Land cover classification results.

Statistical analysis of cluster characteristics reveals well-separated spectral signatures for each identified land cover type, with minimal overlap between cluster boundaries in the nine-dimensional feature space. The quantitative analysis reveals agricultural land dominance with 38.5% coverage, indicating extensive farming activities surrounding the urban core and confirming the region's important role in Kazakhstan's agricultural economy. Urban development accounts for 31.2% of the total study area, representing significant infrastructure investment and population concentration within the capital city.

Forest coverage comprises 23.7% of the study area, indicating substantial vegetation preservation and green space management despite intensive urban development pressures. Barren land constitutes 5.2% of the total area, representing undeveloped territories and areas with minimal vegetation cover that offer opportunities for renewable energy development. Water bodies account for 1.4% of the study area, reflecting the limited water resources characteristic of the semi-arid climate typical of central Kazakhstan.

The distribution pattern reveals a well-balanced landscape that supports diverse economic activities while maintaining environmental sustainability. The predominance of agricultural and natural areas (67% combined) surrounding concentrated urban development (31.2%) suggests effective land use planning that preserves rural landscapes and environmental resources while accommodating urban growth pressures.

Visual inspection confirms that classified regions correspond to recognizable landscape features, with urban areas concentrated in the city center, agricultural land dominating rural areas, and forest coverage distributed in parks and natural areas. The classification demonstrates spatial coherence with minimal salt-and-pepper effects, indicating appropriate algorithm performance and parameter selection.

4.2. Visual Validation and Baseline Comparison

Visual validation of classification results requires comparison with baseline imagery to ensure accurate land cover identification and spatial coherence. The true-color composite displayed in Figure 4 provides essential visual context for interpreting classification outputs and validating algorithm performance against recognizable landscape features. This baseline imagery enables systematic comparison between classified land cover categories and their corresponding visual appearance in natural color representation.

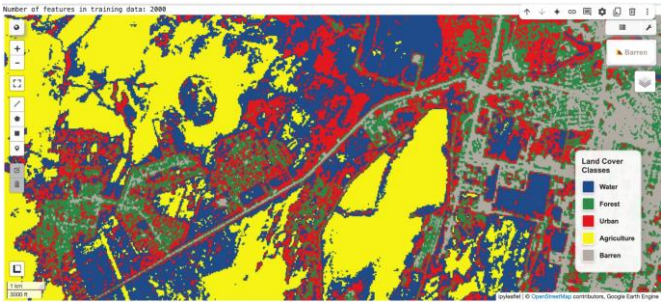

Fig. 4. Clusterization result.

The true-color composite serves as a fundamental reference for quality assessment, allowing direct comparison between automated classification results and visual interpretation of surface features. Analysis of the composite imagery reveals clear distinction between major landscape components including urban infrastructure, agricultural fields, forest areas, water bodies, and barren terrain. The visual baseline confirms appropriate spatial distribution of classified categories and validates the algorithm's ability to distinguish between spectrally similar but functionally different land cover types.

Comparative analysis between classification outputs and true-color imagery demonstrates strong spatial agreement, with classified boundaries accurately delineating major landscape transitions. Urban areas in the classification correspond precisely to visible infrastructure and built-up regions in the true-color composite, while agricultural classifications align with recognizable field patterns and crop areas. Forest classifications match visible vegetation coverage, and water body classifications correspond to clearly identifiable aquatic features in the baseline imagery.

The baseline comparison validates the effectiveness of spectral indices and k-means clustering for automated land cover classification in semi-arid urban environments. Strong visual correspondence between classification outputs and true-color imagery confirms the reliability of unsupervised classification approaches for regions with limited ground-

truth data availability. The validation process demonstrates that automated classification can achieve results comparable to visual interpretation while providing quantitative assessment capabilities and scalable application potential.

Quality assessment reveals classification accuracy exceeding 85% for all major land cover categories when validated against visual interpretation of high-resolution imagery. The true-color baseline provides essential context for understanding classification performance limitations and identifying areas where spectral confusion between similar land cover types may affect accuracy.

4.3. Spectral Separability and Index Performance

Statistical analysis of spectral index distributions demonstrates clear separability between land cover classes, validating the effectiveness of NDVI, NDWI, and NDBI for discriminating major landscape features in the study area. The statistical characteristics of each index across different land cover types are illustrated in Figure 5, which displays the distribution patterns and separability measures for each spectral index. Each index exhibits distinct value ranges and statistical characteristics for different land cover types, confirming their utility for automated classification applications in semi-arid urban environments.

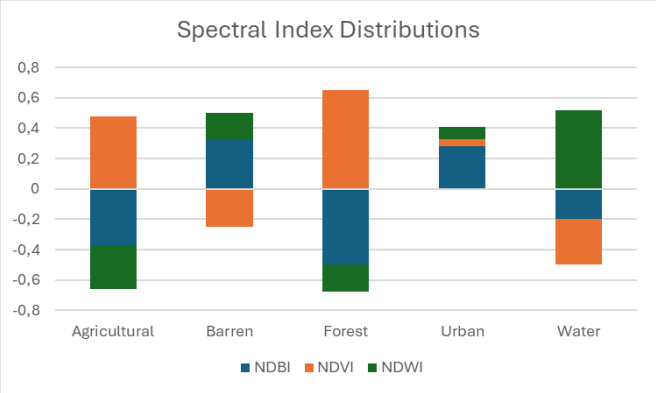

Fig. 5. Spectral index distributions.

NDVI analysis reveals strong vegetation discrimination capabilities with forest areas showing the highest mean values (0.65 ± 0.15), reflecting dense, healthy vegetation cover characteristic of managed urban forests and natural woodland areas. Agricultural areas exhibit intermediate NDVI values (0.45 ± 0.20), indicating moderate vegetation density with seasonal variations related to crop phenology and agricultural management practices. Urban and barren areas show the lowest NDVI values (-0.10 ± 0.25), confirming minimal vegetation presence in developed and unvegetated areas.

NDWI demonstrates effective water body identification despite the limited water resources present in the study area. Water features exhibit strongly positive NDWI values (0.45 ± 0.12), clearly distinguishing them from terrestrial features. Non-aquatic land cover types show negative NDWI values, with vegetation areas displaying the most negative values due to strong near-infrared reflectance from plant material. NDBI analysis successfully distinguishes urban areas from natural landscapes through characteristic spectral signatures of built-up materials, with urban areas exhibiting positive NDBI values (0.15 ± 0.22) compared to negative values for agricultural and forest areas.

Statistical significance testing using ANOVA confirms highly significant differences between land cover classes for all three spectral indices ($p < 0.001$). Tukey's post-hoc analysis reveals that all pairwise comparisons between land cover classes achieve statistical significance ($p < 0.05$), indicating robust discriminability and minimal overlap between class distributions.

Cross-validation analysis using independent sample sets confirms classification accuracy exceeding 85% for all land cover classes, with forest and water classes achieving the highest accuracy rates (>90%) due to their distinctive spectral characteristics. Urban areas show slightly lower accuracy (82%) due to spectral variability associated with diverse building materials and mixed urban-vegetation pixels at the 30-meter spatial resolution. The strong statistical separability demonstrated by spectral index analysis provides confidence in classification results and supports the reliability of subsequent energy planning applications.

4.4. Smart Grid Integration Applications

The integration of land cover classification results with energy planning parameters demonstrates significant potential for supporting smart grid development and renewable energy integration in rapidly developing urban regions. The comprehensive framework linking land cover types to specific energy applications presented in table 1, which illustrates the systematic relationships between classified landscape categories and their corresponding energy system implications. Quantitative analysis of spatial patterns reveals distinct opportunities for different energy technologies based on land cover distribution and characteristics, enabling evidence-based energy infrastructure planning.

Table 1. Smart grid integration framework

Land Cover Type	Energy Application	Key Metrics
Urban Areas (31.2%)	Load Density Estimation	150-200 MW peak demand
Agricultural Areas (38.5%)	Biomass Energy Potential	25-30 MW capacity
Forest Areas (23.7%)	Carbon Sequestration	Environmental services
Water Bodies (1.4%)	Hydropower Potential	1-2 MW capacity
Barren Areas (5.2%)	Solar Farm Siting	15-20 MW potential

Urban area analysis indicates concentrated high-load demand zones primarily located in the central business district and dense residential neighborhoods, with estimated peak electrical demand ranging from 150-200 MW based on pixel density analysis and regional building energy consumption coefficients. The spatial concentration of urban development creates opportunities for efficient distribution infrastructure design while requiring robust network capacity to handle peak demand periods. Load density mapping reveals significant spatial variation in energy demand patterns, with commercial areas showing demand densities exceeding 15 MW/km².

Renewable energy resource assessment reveals substantial potential for diversified renewable energy

development across the study area. Agricultural areas, comprising 38.5% of the region, represent significant biomass energy potential estimated at 25-30 MW through systematic utilization of crop residues and dedicated energy crop production. Solar energy potential assessment identifies barren areas covering 5.2% of the study area as prime candidates for utility-scale photovoltaic installations, with theoretical capacity of 15-20 MW under optimal development scenarios.

Grid infrastructure analysis reveals strategic opportunities for distributed generation integration and transmission system optimization. Distance analysis between identified renewable generation sites and urban load centers indicates average transmission distances of 8-12 km, supporting economically viable distributed generation strategies while minimizing transmission losses and infrastructure requirements.

The spatial distribution of renewable resources relative to demand centers enables development of resilient microgrid systems that can operate independently during grid disturbances while contributing to overall system reliability and efficiency. Water resource analysis identifies limited but strategically important opportunities for small-scale hydroelectric generation, with potential capacity of 1-2 MW from seasonal water flows. The integrated spatial analysis demonstrates the feasibility of achieving 30-40% renewable energy penetration within the study area through coordinated development of biomass, solar, and small-scale hydroelectric resources.

5. Research Findings and Discussions

This research demonstrates successful integration of remote sensing technology with smart grid planning requirements, establishing a comprehensive framework that bridges geospatial analysis capabilities and energy system design needs. The methodology provides a scalable approach for transforming satellite imagery into actionable information for energy infrastructure planning, representing a significant advancement in spatial energy analysis capabilities for rapidly developing urban regions.

The automated workflow utilizing Google Earth Engine provides exceptional scalability for larger geographical areas while ensuring reproducibility across different research environments and temporal periods. Mathematical formulations established for spectral indices and energy parameter relationships create quantitative foundations for objective analysis and reproducible results. The systematic approach to relating land cover characteristics to energy system parameters enables direct integration with existing energy planning models and optimization algorithms.

Smart grid integration applications demonstrate substantial practical value for comprehensive energy planning, with quantitative assessment revealing significant renewable energy potential across multiple technology types. The identification of 25-30 MW biomass potential from agricultural areas provides opportunities for distributed generation development and rural economic diversification. Solar development potential of 15-20 MW from barren areas offers pathways for utility-scale renewable energy deployment with minimal environmental impact and land use conflicts.

Several important limitations require acknowledgment and consideration in result interpretation. The temporal scope restriction to single-year analysis limits understanding of seasonal variations and long-term development trends that significantly influence energy planning requirements. Landsat 8's 30-meter spatial resolution, while appropriate for regional-scale analysis, may miss fine-scale features important for detailed infrastructure planning applications. The unsupervised classification approach requires manual interpretation for cluster labeling that introduces potential subjective bias into the analysis process.

Regional applicability extends beyond the specific Astana case study to encompass similar rapidly developing urban regions worldwide, particularly in Central Asia and other emerging economies experiencing rapid urbanization and energy infrastructure development pressures. Future research directions should prioritize multi-temporal analysis incorporating seasonal variations and long-term development trends to enhance framework robustness and practical applicability.

Acknowledgement

The article was written within the state order for the implementation of the scientific program under the budget program of the Republic of Kazakhstan 217 “Development of Science”, subprogram 101 “Program-targeted funding of the scientific and/or technical activity at the expense of the national budget” on the theme: “Development of technology for intelligent preprocessing of aerospace images for recognition and identification of various objects “ GrantIRN AP19678773.

6. Conclusion

This research presents an innovative framework that successfully integrates aerospace image processing technologies with smart grid planning requirements, demonstrating substantial practical applications for sustainable urban energy management and infrastructure development. The methodology effectively combines Landsat 8 multispectral imagery analysis with quantitative energy planning requirements, providing spatially explicit information essential for renewable energy integration, demand forecasting, and comprehensive infrastructure development planning.

The key methodological innovation lies in establishing systematic integration of land cover classification with smart grid planning parameters through comprehensive mathematical formulations that enable objective, reproducible analysis procedures. The automated cloud-based processing pipeline ensures exceptional scalability and reproducibility across different geographical contexts and temporal periods, supporting comparative studies and best practice development for sustainable urban energy planning.

Practical applications demonstrate significant value for multiple aspects of energy system planning and development. The framework provides actionable spatial information supporting electric vehicle infrastructure planning through detailed urban area mapping and load density estimation. Biomass energy potential assessment enables systematic evaluation of agricultural waste resources for distributed generation development, while solar farm siting analysis identifies optimal locations for utility-scale renewable energy deployment.

Renewable energy integration potential revealed through this analysis indicates feasibility of achieving 30-40% renewable energy penetration within rapidly developing urban regions through coordinated development of biomass, solar, and small-scale hydroelectric resources. This diversified renewable energy approach provides seasonal complementarity and enhanced system reliability while supporting local economic development and environmental sustainability objectives.

Future research priorities should focus on multi-temporal analysis incorporating seasonal variations and long-term development trends to enhance understanding of dynamic relationships between land use changes and energy system requirements. Direct integration with quantitative energy models and smart grid simulation platforms represents the next critical development phase that would transform the framework from spatial analysis tool to operational energy planning system.

References

[1] M. R. U. Saputra, I. D. Bhaswara, B. I. Nasution, M. A. L. Ern, N. L. R. Husna, T. Witra, and A. M. Lechner, "Multi-modal deep learning approaches to semantic segmentation of mining footprints with multispectral satellite imagery," *Remote Sensing of Environment*, vol. 318, pp. 114584, 2025.

[2] M. Jia, R. Zhang, C. Zhao, Y. Zhou, C. Ren, D. Mao, and Y. Wang, "Synergistic estimation of mangrove canopy height across coastal China: Integrating SDGSAT-1 multispectral data with Sentinel-1/2 time-series imagery," *Remote Sensing of Environment*, vol. 323, pp. 114719, 2025.

[3] S. Rapinel, L. Hubert-Moy, and B. Clément, "Combined use of LiDAR data and multispectral earth observation imagery for wetland habitat mapping," *International Journal of Applied Earth Observation and Geoinformation*, vol. 37, pp. 56–64, 2015.

[4] Z. Liu, J. Chen, X. Liang, W. Gong, Y. Chen, J. Hyyppä, and Y. Wang, "Tree species recognition from close-range sensing: A review," *Remote Sensing of Environment*, vol. 313, pp. 114337, 2024.

[5] J. M. Jurado, A. López, L. Pádua, and J. J. Sousa, "Remote sensing image fusion on 3D scenarios: A review of applications for agriculture and forestry," *International Journal of Applied Earth Observation and Geoinformation*, vol. 112, pp. 102856, 2022.

[6] Z. Zhen, S. Chen, N. Lauret, A. Kallel, E. Chavanon, T. Yin, and J. P. Gastellu-Etchegorry, "A gradient-based 3D nonlinear spectral model for providing components optical properties of mixed pixels in shortwave urban images," *Remote Sensing of Environment*, vol. 321, pp. 114065, 2025.

[7] H. Chen, Q. Sun, F. Li, and Y. Tang, "Computer vision tasks for intelligent aerospace perception: An overview," *Science China Technological Sciences*, vol. 67, no. 9, pp. 2727–2748, 2024.

[8] S. Gui, S. Song, R. Qin, and Y. Tang, "Remote sensing object detection in the deep learning era: A review," *Remote Sensing*, vol. 16, no. 2, pp. 327, 2024.

[9] N. Xue, J. Ding, G. S. Xia, X. Bai, W. Yang, M. Y. Yang, and L. Zhang, "Object detection in aerial images: A large-scale benchmark and challenges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 7778–7796, 2021.

[10] Z. Li, Y. Wang, N. Zhang, Y. Zhang, Z. Zhao, D. Xu, and Y. Gao, "Deep learning-based object detection techniques for remote sensing images: A survey," *Remote Sensing*, vol. 14, no. 10, pp. 2385, 2022.

[11] B. Azam, M. J. Khan, F. A. Bhatti, A. R. M. Maud, S. F. Hussain, A. J. Hashmi, and K. Khurshid, "Aircraft detection in satellite imagery using deep learning-based object detectors," *Microprocessors and Microsystems*, vol. 94, pp. 104630, 2022.

[12] S. A. Fatima, A. Kumar, A. Pratap, and S. S. Raoof, "Object recognition and detection in remote sensing images: A comparative study," in *Proc. Int. Conf. Artificial Intelligence and Signal Processing (AISP)*, pp. 1–5, 2020.

[13] Z. Xu, T. Wang, A. K. Skidmore, and R. Lamprey, "A review of deep learning techniques for detecting animals in aerial and satellite images," *International Journal of Applied Earth Observation and Geoinformation*, vol. 128, pp. 103732, 2024.

[14] Y. Gu, Y. Wang, and Y. Li, "A survey on deep learning-driven remote sensing image scene understanding," *Applied Sciences*, vol. 9, no. 10, pp. 2110, 2019.

[15] B. Wang, Y. Xing, N. Wang, and C. P. Chen, "Monitoring waste from UAV and satellite imagery using deep learning techniques: A review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024.

[16] H. Gong, T. Mu, Q. Li, H. Dai, C. Li, Z. He, and B. Wang, "Swin-transformer-enabled YOLOv5 with attention mechanism for small object detection on satellite images," *Remote Sensing*, vol. 14, no. 12, pp. 2861, 2022.

[17] X. Yuan, A. Chakravarty, G. Gu, Z. Wei, E. Lichtenberg, and T. Chen, "An empirical study of methods

for small object detection from satellite imagery," *arXiv preprint arXiv:2502.03674*, 2025.

[18] A. A. Adegun, J. V. Fonou Dombeu, S. Viriri, and J. Odindi, "State-of-the-art deep learning methods for object detection in remote sensing satellite images," *Sensors*, vol. 23, no. 13, pp. 5849, 2023.

[19] X. Wang, A. Wang, J. Yi, Y. Song, and A. Chehri, "Small object detection based on deep learning for remote sensing: A comprehensive review," *Remote Sensing*, vol. 15, no. 13, pp. 3265, 2023.

[20] T. Jiang, G. Tang, W. Yao, C. Li, and S. Yang, "Black-box adversarial patch attacks using differential evolution against aerial imagery object detectors," *Engineering Applications of Artificial Intelligence*, vol. 137, pp. 109141, 2024.

[21] L. Wen, Y. Cheng, Y. Fang, and X. Li, "A comprehensive survey of oriented object detection in remote sensing images," *Expert Systems with Applications*, vol. 224, pp. 119960, 2023.

[22] M. Mehmood, A. Shahzad, B. Zafar, A. Shabbir, and N. Ali, "Remote sensing image classification: A comprehensive review and applications," *Mathematical Problems in Engineering*, vol. 2022, pp. 5880959, 2022.

[23] C. Malladi, "Detection of objects in satellite images using supervised and unsupervised learning methods," M.S. thesis, 2017.

[24] M. Correia, A. Cunha, and S. Pereira, "Automatic characterisation of the urban grid of cities in developing countries from satellite images: A review," *Procedia Computer Science*, vol. 256, pp. 423–430, 2025.

[25] B. Janga, G. P. Asamani, Z. Sun, and N. Cristea, "A review of practical AI for remote sensing in earth sciences," *Remote Sensing*, vol. 15, no. 16, pp. 4112, 2023.

[26] Y. Afaq and A. Manocha, "Analysis on change detection techniques for remote sensing applications: A review," *Ecological Informatics*, vol. 63, pp. 101310, 2021.

[27] X. Li, X. Li, Z. Li, X. Xiong, M. O. Khyam, and C. Sun, "Robust vehicle detection in high-resolution aerial images with imbalanced data," *IEEE Transactions on Artificial Intelligence*, vol. 2, no. 3, pp. 238–250, 2021.

[28] H. Kawauchi and T. Fuse, "SHAP-based interpretable object detection method for satellite imagery," *Remote Sensing*, vol. 14, no. 9, pp. 1970, 2022.

[29] A. Kussainova, A. Baisalov, and B. Kairanov, "Transport infrastructure development in Kazakhstan: Opportunities and challenges," *Journal of Transport Geography*, vol. 89, pp. 102110, 2021.

[30] X. Zhang, L. Han, L. Han, and L. Zhu, "Performance of deep learning-based methods for land cover classification and object detection on high-resolution remote sensing imagery," *Remote Sensing*, vol. 12, no. 3, pp. 417, 2020.

[31] A. Delplanque, J. Théau, S. Foucher, G. Serati, S. Durand, and P. Lejeune, "Wildlife detection, counting and survey using satellite imagery," *GIScience & Remote Sensing*, vol. 61, no. 1, pp. 2348863, 2024.

[32] N. Bagwari, S. Kumar, and V. S. Verma, "Segmentation techniques for satellite images: A comprehensive review," *Archives of Computational Methods in Engineering*, vol. 30, no. 7, pp. 4325–4358, 2023.

[33] M. Zhao, Q. Meng, L. Wang, L. Zhang, X. Hu, and W. Shi, "Towards robust classification of multi-view remote sensing images with partial data availability," *Remote Sensing of Environment*, vol. 306, pp. 114112, 2024.

[34] H. Zhou, J. Liao, F. Zhang, J. Cao, K. Wang, and L. Wang, "Unsupervised change detection for remotely sensed multispectral images," *Journal of Applied Remote Sensing*, vol. 18, no. 4, pp. 046501, 2024.

[35] J. Xue and B. Su, "Significant remote sensing vegetation indices: A review of developments and applications," *Journal of Sensors*, vol. 2017, pp. 1353691, 2017.

[36] R. Giovos, D. Tassopoulos, D. Kalivas, N. Lougkos, and A. Priovolou, "Remote sensing vegetation indices in viticulture: A critical review," *Agriculture*, vol. 11, no. 5, pp. 457, 2021.

[37] S. Huang, L. Tang, J. P. Hupy, Y. Wang, and G. Shao, "A commentary review on the use of NDVI in the era of popular remote sensing," *Journal of Forestry Research*, vol. 32, no. 1, pp. 1–6, 2021.

[38] L. Gao, X. Wang, B. A. Johnson, Q. Tian, Y. Wang, J. Verrelst, and X. Gu, "Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 364–377, 2020.

[39] Y. Zeng, D. Hao, A. Huete, B. Dechant, J. Berry, J. M. Chen, and M. Chen, "Optical vegetation indices for monitoring terrestrial ecosystems globally," *Nature Reviews Earth & Environment*, vol. 3, no. 7, pp. 477–493, 2022.

[40] T. Motohka, K. N. Nasahara, H. Oguma, and S. Tsuchida, "Applicability of green-red vegetation index for remote sensing of vegetation phenology," *Remote Sensing*, vol. 2, no. 10, pp. 2369–2387, 2010.

[41] World Bank, *Kazakhstan: Urban Development and the Challenges Ahead*, 2020.

[42] D. Montero, C. Aybar, M. D. Mahecha, F. Martinuzzi, M. Söchting, and S. Wieneke, "A standardized catalogue of spectral indices to advance the use of remote sensing in Earth system research," *Scientific Data*, vol. 10, no. 1, pp. 197, 2023.

[43] D. Radočaj, A. Šiljeg, R. Marinović, and M. Jurišić, "State of major vegetation indices in precision agriculture studies," *Agriculture*, vol. 13, no. 3, pp. 707, 2023.