



Article

Quantifying the Co-evolution Dynamics of Urban and Road Systems in Al-Sulaymaniyah City, Iraq: A GIS-Based Integration of Random Forest and Road Network Analysis

Nyan Qasim Khudhur¹, Dler A. Al-Mamany², Baraa Raad Mohammed³

1. Northern Technical University, Surveying Eng. Techniques Dept

2. Northern Technical University, Surveying Eng. Techniques Dept

3. University of Diyala, Collage of Engineering, Highway and Airport Dept

* Correspondence: nyan.qasims@ntu.edu.iq, dlieromer@ntu.edu.iq, baraa.raad@uodiyala.edu.iq

Abstract: This study examines the temporal changes in Al-Sulaymaniyah City, Iraq, through the integration of Random Forest land cover classification and road network analysis from 1994 to 2024. A seven-class land cover classification system was developed using Landsat images from various time periods, achieving an overall accuracy of 93% and a Kappa coefficient of 0.90. Urban areas expanded significantly, increasing from 132.34 km² (0.78%) in 1994 to 628.27 km² (3.69%) in 2024. This is a 375% increase, with an annual growth rate of 25.47% from 1994 to 2004. Cropland underwent significant fluctuations, decreasing from 2,119.80 km² in 2004 to 636.43 km² by 2004, before increasing to 3,257.10 km² by 2024. The density of roads per square kilometer increased from 15.035 to 17.231. A co-evolutionary study demonstrated that urban-network interactions exhibited complex relationships with a moderate explanatory power ($R^2 = 0.49$). Significant positive Pearson correlations were seen between the length of roadways in metropolitan regions ($r = 0.701$, $p < 0.001$) and the density of networks ($r = 0.723$, $p < 0.001$). Granger causality studies, however, revealed no substantial evidence that urban growth induces road extension ($p = 0.5$), indicating that the two are evolving independently. The urban trend model demonstrated strong predictive capability ($R^2 = 0.87$), whereas the network development model exhibited only moderate predictive capability ($R^2 = 0.58$). The results indicate that rapid urbanization has surpassed infrastructure development, complicating long-term planning for rapidly expanding communities.

Keywords: Co-evolution Analysis, Random Forest Classification, Road Network Analysis, Urban Expansion, GIS.

Citation: Khudhur, N. Q, Al-Mamany, D. A & Mohammed, B. R. Quantifying the Co-evolution Dynamics of Urban and Road Systems in Al-Sulaymaniyah City, Iraq: A GIS-Based Integration of Random Forest and Road Network Analysis. Vital Annex: International Journal of Novel Research in Advanced Sciences 2025, 4(8), 351-368

Received: 15th Jun 2025

Revised: 29th Jul 2025

Accepted: 17th Aug 2025

Published: 30th Sep 2025



Copyright: © 2025 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>)

1. Introduction

Global urbanization has occurred so rapidly that it has transformed urban living conditions. Currently, over fifty percent of the global population resides in urban areas, a figure projected to increase to 68% by 2050 [1]. The extraordinary expansion of urban areas has rendered the relationship between transportation infrastructure and land use patterns more complex. This is particularly applicable in rapidly expanding cities, where the co-evolution of urban land use/land cover (LULC) and road networks serves as a principal catalyst for spatial transformation [2], [3]. To design sustainable cities, it is essential to comprehend these evolving relationships. Transportation networks facilitate urban expansion, alter land utilization, impact environmental risk, and influence economic development [4].

Urban growth and transportation infrastructure mutually influence each other throughout time, altering the appearance and functionality of cities. The expansion of the road network is a primary factor that alters urban land. This generates accessibility

gradients that result in urban sprawl and alterations in land use [5], [6]. Simultaneously, urban expansion necessitates more infrastructure investment, resulting in heightened travel, which in turn compels the integration of land use and transportation systems [7]. This bidirectional association is particularly robust in areas rapidly transitioning into urban centers. The rapidity and magnitude of change can generate opportunities for efficient growth as well as issues related to environmental degradation and social inequality [8].

Innovations in remote sensing and geographic information systems (GIS) have transformed our ability to observe and analyze complex urban dynamics. The integration of multi-temporal satellite imagery and enhanced classification algorithms has enabled researchers to monitor alterations in land cover with unprecedented accuracy and detail [9]. Random Forest (RF) classifiers are among the most effective machine learning methods for land use and land cover (LULC) mapping due to their high accuracy, capability to manage noisy data, and proficiency in processing high-dimensional remote sensing data [10]. The integration of RF-based LULC classification with vector-based road network analysis holds significant potential for enhancing our understanding of urban evolution and its implications for sustainable development.

Despite the advent of new technologies, our understanding of effectively integrating machine learning-based land use and land cover classification with transportation network analysis remains insufficient, particularly in rapidly urbanizing areas characterized by limited data and diverse urban environments. Much of the existing research examines land use change or transportation network expansion in isolation. This complicates our ability to understand their interrelations in all their intricacy [11]. Furthermore, RF-based methodologies have demonstrated promise for LULC mapping in developed regions; however, their use has been limited in rapidly expanding cities characterized by diverse urban configurations [12]. The objective of this study is to investigate the interrelationship between urban expansion and the development of road networks by employing an integrated methodological framework that combines Random Forest-based land use and land cover classification with vector-based road network analysis.

The explicit objectives are: (1) to develop a robust methodology for integrating RF-classified LULC data with historical road network datasets to assess spatiotemporal patterns of urban co-evolution; (2) to examine the reciprocal relationships between road network expansion and land use transformation in a rapidly urbanizing region; (3) to evaluate the efficacy of RF classifiers for LULC mapping in urban environments characterized by diverse land types and limited ground truth data; and (4) to provide recommendations for sustainable urban planning that considers transportation-land use interactions.

This study contributes several novel insights into the domains of urban studies and remote sensing applications. Initially, it establishes a cohesive analytical framework that distinctly illustrates the interrelationship between urban growth and transportation infrastructure, surpassing prior methodologies that examined both phenomena in isolation. Secondly, it facilitates the application of Random Forest classifiers for land use and land cover (LULC) mapping in rapidly expanding urban areas by addressing data scarcity and significant heterogeneity among cities. Third, analyzing various data sources throughout time provides new insights into the interaction between transportation and land use. Ultimately, it underscores the significance of integrating machine learning and GIS for evidence-based urban planning in rapidly evolving cities.

Literature Review

In recent decades, our theoretical understanding of urban growth and the concurrent evolution of transportation networks has significantly transformed. Preliminary research demonstrated that these connections are reciprocal. Lacono et al. [13] were the pioneers in examining the temporal correlation between land use and road network alterations. They demonstrated how transportation infrastructure both influences and is influenced by urban development. This research established concepts

that facilitate our understanding of urban growth as a complex adaptive system characterized by the dynamic interplay of transportation networks and land use patterns through continuous feedback loops. Subsequent research has enhanced our understanding of the processes underlying co-evolution. Li et al. developed integrated models that demonstrate the simultaneous optimization of land use and transportation networks. These models illustrate how improved accessibility can result in both efficient urban development and rampant sprawl. Kasraian et al. conducted comprehensive analyses of empirical data regarding the interconnections between transportation and land use. Consistent tendencies were identified across several urban areas; nevertheless, the significance of local characteristics in influencing individual outcomes was emphasized. The emergence of these novel concepts has rendered co-evolution a fundamental principle for comprehending urban dynamics, with implications extending beyond transportation planning to encompass broader concerns of sustainability and resilience.

The application of remote sensing and GIS technologies in urban studies has transformed our ability to observe and analyze urban changes. Initially, the primary application of these methods was to categorize land cover via conventional supervised and unsupervised techniques. Nonetheless, advancements in technology have enabled increasingly sophisticated analyses of urban dynamics [14]. Research indicates that multi-temporal satellite imagery is effective for monitoring urban expansion, as it may detect subtle alterations in land use patterns across time [15], [16]. Researchers may now utilize GIS and remote sensing in conjunction to analyze urban dynamics with spatial precision. This enables them to assess the correlations between alterations in land use and transportation infrastructure. Mo et al. [17] demonstrated how the addition of highways in megacities alters the ecological risk of the terrain.

Machine learning techniques have transformed land use and land cover classification, with Random Forest being among the most favored and dependable ways. Random Forest classifiers are favored for their capability to handle high-dimensional remote sensing data while remaining comprehensible and offering uncertainty estimations [18], [19]. Comparative analyses have consistently demonstrated that RF classifiers outperform conventional statistical methods, particularly in complex urban environments characterized by mixed land use patterns and spectral confusion among classes. Recent applications of RF in urban areas have addressed increasingly complex classification challenges. Mustafa et al. [20] examined random forests and support vector machines in cellular automata land use change models, demonstrating that ensemble methods more effectively capture urban complexity. Similarly, Zhao et al. [21] employed neural network methodologies using RF components to strategize land utilization on urban thoroughfares. This illustrates how machine learning can assist with operational planning. These enhancements have established RF as a standard instrument for LULC mapping; nonetheless, challenges regarding data quality and the applicability of results in diverse urban environments persist.

The integration of machine learning-driven land use and land cover classification with transportation network analysis represents a novel domain of inquiry in urban studies. Despite the thorough development of each component, their integration poses significant challenges due to issues related to data harmonization, scale integration, and uncertainty propagation. Ahmadzai was the inaugural researcher to employ spatial-based accessibility metrics to examine the interaction between urban land use and road networks. This demonstrated that integrated research is feasible, although it also highlighted the challenges of data amalgamation. Recent research has begun to address these issues with innovative methodologies. Xu et al. [22] developed methods to integrate transit accessibility with vector-based cellular automata to model alterations in urban land use. This demonstrated how transportation data can enhance the accuracy of land use estimates. Raimbault et al. [23] incorporated endogenous transport supply into land use-transport interaction models. This illustrates how integrated methodologies can more effectively represent the complexity of urban

systems. These modifications indicate that integrated methodological frameworks could significantly enhance urban co-evolution research. Comparative analyses of co-evolutionary processes across several cities have revealed both parallels and distinctions. Zhao et al. [24] investigated the spatial and temporal characteristics of road networks and urban expansion in Beijing, New York, London, and Chicago. Road density consistently forecasts urban expansion; however, the patterns vary by region. These geographical applications demonstrate that co-evolution ideas can be applied in various contexts. They have demonstrated the significance of contextual analysis to comprehend operational dynamics within a certain domain and formulate appropriate policy responses.

Despite significant advancements in methodologies and applications, challenges persist in the study of urban co-evolution. Data harmonization is challenging due to the substantial effort required to ensure that remote sensing pictures and vector-based transportation information are accurately aligned in terms of location and time [25]. Classification uncertainty in heterogeneous urban environments remains a significant challenge, particularly in rapidly expanding cities characterized by complex land use patterns and informal settlements that complicate the use of traditional classification approaches [26].

2. Materials and Methods

Study Area

Al-Sulaymaniyah City, situated in the Kurdistan Region of northern Iraq, serves as a significant urban center in terms of demographics and cultural influence. The city is located in a complicated topography encircled by substantial geographical obstacles, rendering it an intriguing subject for urban growth modeling (Figure 1). Al-Sulaymaniyah has experienced substantial and continuous urban expansion in recent decades. Between 1991 and 2014, developed land expanded from 13.3% (3,063.78 hectares) to 41.8% (9,654.26 hectares), indicative of population growth and economic advancement. The city is regarded as the cultural center of Kurdistan, with a 3% yearly population growth rate. In 1987, 63% of the population resided in urban areas, increasing to 78% by 2008, signifying considerable urbanization. The factors contributing to heightened urbanization encompass substantial economic advancements post-2003 and a rise in development throughout several sectors, including residential, commercial, educational, and infrastructural domains. The expansion has occurred at the cost of agricultural and open spaces, resulting in substantial alterations to the city's land use and land cover configuration. The city's transformations rendered an optimal experimental framework for urban modeling methodologies that consider nonlinear, spatially intricate growth dynamics, including SLEUTH and optimization-based calibrations such as Genetic Algorithms.

The urban development of Al-Sulaymaniyah is fundamentally linked to its transportation infrastructure. The city is situated in an interprovincial junction, bordered by the Erbil and Kirkuk governorates to the west and southwest, and has an eastern boundary with Iran. The primary road network consists of radial and circular configurations that radiate from the city center to peripheral districts and adjacent villages. The proliferation of significant transportation corridors is a hallmark of contemporary urban development, since construction has focused on principal roads and arterial routes owing to enhanced accessibility and economic prospects.

The region features varied geography, with the Baranan and Chwarta mountains to the south, the Tasluja hills to the west, and the Qaiwan range, Azmaer, and Goizha mountains to the north and northeast. This mountainous barrier imposes considerable topographical constraints and microclimatic fluctuations throughout the city's expanding area. The city possesses a semi-arid climate characterized by hot, dry summers and cold, moist winters, which affect settlement patterns and urban development trajectories [27].

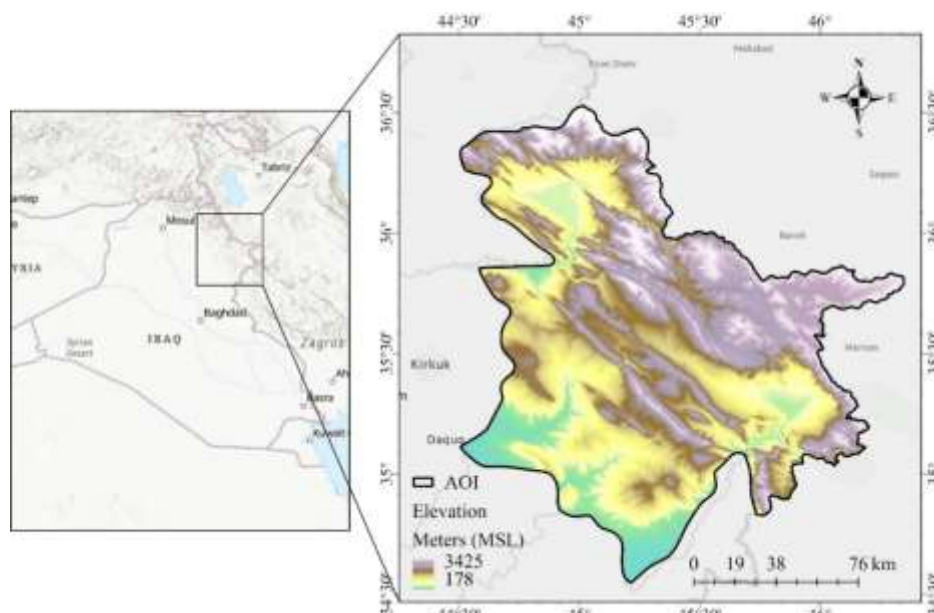


Figure 1. Map of the study area.

Datasets

Landsat Image Dataset

Land cover classification was performed with multitemporal Landsat imagery over a duration of three decades to assess terrain change. Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and Landsat 8 Operational Land Imager (OLI) images were obtained for the years 1994, 2004, 2014, and 2024, offering decadal insights into land cover transformation. Images were chosen during analogous phenological periods to mitigate seasonal fluctuation effects and achieve categorization uniformity throughout the temporal series.

Road Networks.

The establishment of historical road networks for the research area via a multi-source methodology that incorporated archival geospatial data, remote sensing analysis, and hand digitization techniques. Historical OpenStreetMap (OSM) road data constituted the principal basis for modern road network extraction. Historical OSM data was obtained via the Overpass API, with filters for highway tags and temporal variables pertinent to the study period. The retrieved vector data included essential road geometry and attribute information, encompassing road classifications and temporal metadata wherever accessible. The Landsat satellite imagery pertinent to the study's chronology was meticulously analyzed to identify and digitize road attributes absent from existing datasets. Roads were systematically analyzed at suitable scales utilizing Geographic Information Systems (GIS) software, which entailed on-screen assessment of spectral attributes and spatial configurations indicative of mobility corridors. Comprehensive manual modifications were executed to preserve topological links, ensuring network connectivity and analytical integrity. This procedure encompassed (1) geometric correction of digitized features to rectify overshoots, undershoots, and pseudo-nodes; (2) establishing suitable connectivity at road intersections; (3) validating network topology through GIS topology rules; and (4) assigning temporal parameters to individual road segments via imagery analysis. The resulting historical road network database offers a geographically and temporally accurate representation of transportation infrastructure history, appropriate for SLEUTH modeling.

Land Cover Classification

Land Cover Scheme System

A seven-class land cover classification was established utilizing the Land Cover

Categorization System (LCCS) framework, encompassing categories of water, trees, flooded vegetation, crops, constructed environments, bare ground, and rangeland (Table 1). The training dataset comprised 50 polygon samples per class, each containing 10-50 pixels to encapsulate intra-class spectral diversity while preserving spatial homogeneity. Training polygons were established through visual analysis of high-resolution data and field expertise, guaranteeing representative sampling across various landscape conditions within each category. Land cover classification is executed with random forest (RF) learning techniques, which use the method's noise resilience and ability to manage high-dimensional spectral data. The RF classifier utilized all accessible Landsat spectral bands, with the training data divided into 70% for training and 30% for validation subsets by stratified random selection to preserve class proportionality. The model's performance was assessed over time for classification reliability through confusion matrices, overall accuracy metrics, and per-class producer and user accuracies.

Random Forest

Random Forest is an ensemble machine learning technique that uses many decision trees to categorize satellite pictures by aggregating the predictions of various classifiers. The technique constructs a forest of decision trees with bootstrap aggregating (bagging) and random feature selection. Each tree is trained on a random subset of training data and examines a random selection of characteristics at each split. The final classification is determined by the majority vote of the trees within the forest. The algorithm's estimation can be expressed as:

$$\hat{y} = \text{mode}\{T_1(x), T_2(x), \dots, T_B(x)\} \quad (1)$$

where $\hat{y}$ is the predicted class, $T_b(x)$ represents the prediction of the b tree for input x , and B is the total number of trees. The out-of-bag (OOB) error rate is calculated as:

$$\text{OOB Error} = \frac{1}{n} \sum_{i=1}^n I(y_i \neq \widehat{y}_i^{\text{OOB}}) \quad (2)$$

where n is the number of samples, y_i is the true class label, $\widehat{y}_i^{\text{OOB}}$ is the OOB prediction, and $I(\cdot)$ is the indicator function. Random Forest's effectiveness in satellite image classification stems from its ability to handle high-dimensional spectral data, reduce overfitting through ensemble averaging, provide feature importance rankings, and maintain computational efficiency while delivering robust classification performance across diverse land cover types, see Table 1.

Table 1. Land cover scheme based on the Land Cover Classification System (LCCS).

Land Cover Class	Description
Water	Areas that are naturally or artificially covered by water for at least most of the year. These include natural lakes, rivers, reservoirs, and lagoons.
Trees	Areas dominated by woody vegetation consisting of trees whose height is generally greater than 5 meters, forming a continuous or open canopy.
Flooded Vegetation	Areas transitional between terrestrial and aquatic systems, where water is at or near the surface for a substantial period regularly every year. The predominant vegetation comprises hydrophytes, such as marshes, swamps, bogs, and mangroves.
Crops	Areas where the natural vegetation has been removed or modified and replaced by crops or other types of planted and cultivated vegetative cover. This includes annual and perennial crops, orchards, and plantations (rubber, palm oil, etc.), where vegetation is of anthropogenic origin and requires

	human activity to maintain it in the long term.
Built Area	Areas that have an artificial cover as a result of human activities such as construction (cities, towns, roads), extraction (mines), or waste disposal. These areas are dominated by impervious surfaces—such as asphalt, concrete, and buildings—with little or no natural vegetation.
Bare Ground	Areas with minimal to no vegetation (total vegetative cover less than 4% for more than 10 months of the year), including bare rock, sand, and deserts. These areas do not have a significant artificial cover and are characterized mainly by the appearance of the surface, whether consolidated or unconsolidated (rock, sand, or soil).
Rangeland	Areas dominated by natural or semi-natural herbaceous vegetation (graminoids and forbs) and occasionally shrubs, used primarily for livestock grazing and wildlife habitat.

Network Analysis

We employed a comprehensive graph-theoretic approach to convert road infrastructure data into mathematical network representations suitable for quantitative analysis. The methodology employs a multi-phase approach encompassing data pretreatment, network construction, and metric computation to characterize the structural and functional dimensions of the road networks during the study period. The initial step in data preprocessing involves loading road vector data from shapefiles corresponding to each time period and ensuring uniformity across all datasets. We verify and refine the road geometries to ensure uniformity in shape, and we eliminate any characteristics that are either inauthentic or malfunctioning.

The method of constructing the network transforms the filtered road geometry into a mathematical graph by identifying endpoints and categorizing them by location. Each road section contributes two endpoints to the overall endpoint collection. The dataset is subsequently spatially grouped utilizing the DBSCAN technique to identify topologically related nodes. The clustering technique groups spatially proximate endpoints within a certain tolerance distance, therefore forming junction nodes and eliminating geometric irregularities:

$$C = DBSCAN(E, \epsilon, min_{samples})$$

(3)

where E represents the set of all endpoints, ϵ is the spatial tolerance parameter, and $min_{samples}$ defines the minimum cluster size. The resulting clusters are converted to network nodes positioned at cluster centroids:

$$n_c = \frac{1}{|E_c|} \sum_{e_i \in E_c} e_i$$

(4)

where n_c is the centroid position of cluster c , and $|E_c|$ is the set of endpoints in cluster.

The graph is built by linking clustered nodes with edges. Each original road segment adds an edge that connects the centroids of its start and finish clusters. Edge weights are based on the geometric length of the road segments that connect the two nodes. If there are more than one road segment between the same two nodes, the lengths of all of them are added together:

$$w_{ij} = \sum_k length(r_k) \quad \forall r_k \text{ connecting clusters } i \text{ and } j$$

(5)

Network simplification is done to make the calculations easier while keeping important topological aspects. The method of simplifying finds and removes degree-2 nodes (nodes with exactly two connections) while keeping the graph connected by combining edges that are next to each other. When you remove a degree-2 node, the

edges that were connected to it are replaced with a single edge whose weight is the total of the weights of the original edges:

$$w_{new} = w_{ij} + w_{jk} \quad (6)$$

where node j with degree 2 is removed, and nodes i and k are directly connected with the new cumulative weight.

Co-evolution of Urban Growth and Network Analysis

The co-evolution approach employs a comprehensive mathematical framework to examine the temporal changes in urban growth patterns and transportation network development, as well as their interdependence. This approach integrates temporal correlation analysis, geographical association evaluation, causality testing, and predictive modeling to ascertain the simultaneous or sequential occurrence of urban growth and infrastructure development. Temporal correlation analysis employs vectorized statistical processes to identify various correlation metrics between network features and urban growth indicators during the study period. We extract time series data as NumPy arrays to facilitate rapid vectorized computations. Urban areas and network measurements serve as the foundation for correlation analysis. By employing the normalized covariance formula, we may ascertain the Pearson correlation coefficients:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (7)$$

where x_i and y_i represent paired observations of urban and network variables, respectively. Spearman rank correlation is computed to assess monotonic relationships that may not be captured by linear correlation measures. The analysis encompasses multiple variable pairs including urban area versus network length, urban area versus network density, and urban area versus road density to provide comprehensive relationship characterization.

Spatial correlation analysis uses rasterization to change vector-based urban and network data into gridded formats that may be used to look at pixel-level correlations. We use geometric presence indicators to turn road networks into raster images. We also turn metropolitan regions into binary rasters that show developed and underdeveloped pixels. To make the math easier, sampled pixel arrays are used to find the cross-correlation between urban and network rasters:

$$C_{spatial} = \frac{\text{cov}(U_{sample}, N_{sample})}{\sigma_{U_{sample}} \cdot \sigma_{N_{sample}}} \quad (8)$$

where U_{sample} and N_{sample} represent sampled urban and network raster values, respectively.

The co-evolution modeling framework uses several regression methods to show how urban and network growth are related. For speed, linear regression models are fitted using the normal equation method:

$$\beta = (X^T X)^{-1} X^T y \quad (9)$$

where β represents the coefficient vector, X is the design matrix, and y is the response vector. Model performance is assessed using the coefficient of determination:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

where $\hat{y}_i$ represents predicted values and $\bar{y}$ is the response variable mean.

Causality assessment employs lag correlation analysis to test directional relationships between urban growth and network development. The analysis examines whether urban development at time t predicts network changes at time $t+1$,

implementing a simplified Granger causality framework:

$$\rho_{lag} = \text{corr}(U_t, N_{t+1}) \quad (11)$$

where U_t represents urban development at time t and N_{t+1} represents network characteristics at the subsequent time period. Statistical significance is assessed through correlation magnitude thresholds and bootstrap-based confidence interval estimation.

3. Results and Discussion

Results of Land Cover Mapping

The supervised classification of land cover had a good overall accuracy for all mapped years (Table 2). The study of the confusion matrix demonstrated an overall accuracy of 93% and a Kappa coefficient of 0.90, signifying exceptional concordance between the classified map and the reference data. Among individual classes, aquatic bodies exhibited the highest precision (1.00) but had moderate recall (0.84), yielding an F1 score of 0.91. Rangeland demonstrated flawless recall (1.00) and outstanding precision (0.90), attaining the maximum F1 score of 0.95. Developed regions exhibited exceptional precision (0.99) but diminished recall (0.77), indicating a potential underestimating of urban sprawl. Cropland classification had strong performance with a precision of 0.97 and a recall of 0.85, but bare ground demonstrated the weakest performance metrics, with an F1 score of 0.85.

The land use maps from 1994, 2004, 2014, and 2024 indicate substantial spatiotemporal alterations in land cover patterns throughout the study area (Figure 2, Table 3). The most notable change transpired in the allocation of crops and barren ground across the observation intervals. In 1994, the terrain was principally defined by vast rangeland (13,775.53 km²), with crops localized in particular regions and limited urban development mostly observed in the middle section of the study. By 2004, a significant alteration in land cover composition was apparent, with bare terrain expanding markedly from 694.99 km² to 2,115.25 km²—a more than threefold increase. This growth coincided with a substantial decrease in farmland area from 2,119.80 km² to merely 636.43 km². The spatial pattern indicates a potential phase of land deterioration or agricultural abandonment throughout this decade. The 2014 image illustrates a significant reversal of this pattern, with farmland expanding to 3,169.61 km², while bare terrain diminished to 302.44 km². Urban areas persisted in their gradual growth, attaining 434.74 km². The spatial distribution reveals agricultural intensification, especially in the northern and central regions of the research area. By 2024, the terrain stabilized with sustained agricultural preeminence. Cropland expanded to 3,257.10 km², marking the greatest extent recorded over all historical periods. Developed regions demonstrated steady expansion, increasing to 628.27 km²—almost a five-fold rise since 1994. Significantly, bare ground diminished to its minimal extent (58.98 km²), indicating effective land reclamation or vegetative restoration.

Rangeland, while maintaining its status as the predominant land cover category during the research period, underwent a progressive decrease from 13,775.53 km² in 1994 to 12,807.29 km² in 2024, indicating a 7% loss. This loss of around 968 km² seems to have been predominantly transformed into agricultural land and developed regions. Water bodies initially expanded from 257.29 km² in 1994 to 292.91 km² in 2004, subsequently seeing a slow decline to 233.27 km² by 2024. The 9.4% reduction from 1994 levels may signify water stress or alterations in water management methods. Tree cover demonstrated significant volatility, plummeting from 43.24 km² in 1994 to merely 18 km² in 2004 (a 58% decrease), before partially rebounding to 38.29 km² by 2024. Although this indicates partial recovery, tree cover is still 11.5% lower than the levels recorded in 1994. The most notable alteration pertained to the interaction

between agriculture and barren terrain. The transient transformation of agricultural land to unutilized terrain in 2004, succeeded by an increase in agricultural activities beyond prior extents, indicates a significant alteration in land use policy or the initiation of an agricultural development initiative post-2004. The near-elimination of bare ground by 2024 (a 91.5% reduction from the 2004 high) and a 54% increase in crops from 1994 levels exemplifies effective agricultural intensification and land rehabilitation initiatives. Urban expansion exhibited the most steady growth trend, with developed areas expanding by 374% over the 30-year period. The consistent urbanization, expanding from 132.34 km² to 628.27 km², signifies demographic pressures and economic advancement in the area.

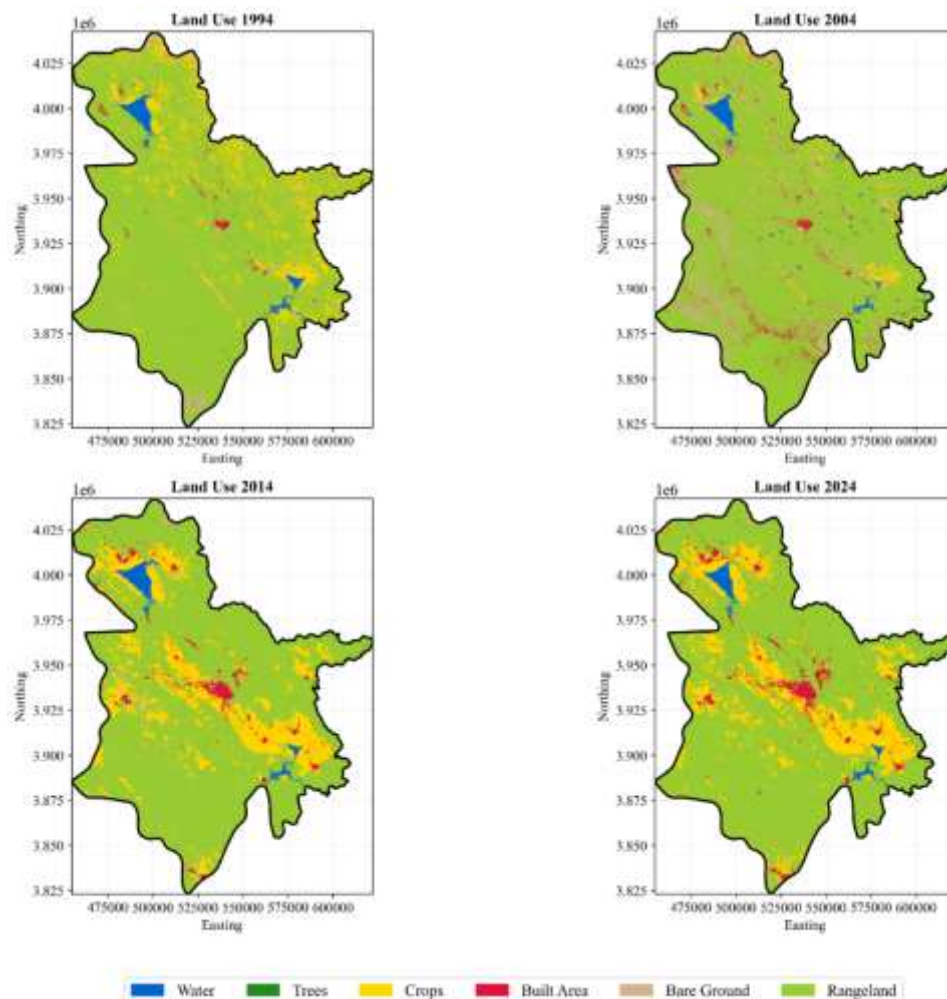


Figure 2. Land use maps of the study area for the years 1994-2024 based on Random Forest supervised classification.

Table 2. Accuracy assessment of the classification process.

Class	Precision	Recall	F1 Score
Water	1.00	0.84	0.91
Crops	0.97	0.85	0.91
Built Area	0.99	0.77	0.87
Bare Ground	0.96	0.76	0.85
Rangeland	0.90	1.00	0.95
Overall Accuracy	0.93		
Kappa	0.90		

Table 3. Area of land use classes for the years 1994-2024.

Land use	Area (km ²)			
	1994	2004	2014	2024
Water	257.29	292.91	274.59	233.27
Trees	43.24	18	23.15	38.29
Crops	2119.80	636.43	3169.61	3257.10
Built Area	132.34	387.06	434.74	628.27
Bare Ground	694.99	2115.25	302.44	58.98
Rangeland	13775.53	13573.55	12818.66	12807.29

Results of Urban Growth Analysis

The investigation of urban growth patterns from 1994 to 2024 indicates significant temporal and spatial variations in urbanization dynamics within the examined region. The urban expansion exhibits a distinctive pattern of rapid growth phases succeeded by periods of slower growth, as evidenced by the spatial distribution maps and quantitative indicators presented in Table 4. In 1994, the initial urban footprint encompassed 132.34 km², or 0.78% of the whole study area (Figure 4). At this time, the majority of urban communities were concentrated in the central valley region, with development nodes dispersed throughout the landscape. The spatial arrangement indicates that the city possesses a polycentric structure, characterized by limited linkages among various residential clusters.

During the initial decade of analysis (1994–2004), the city had the most significant transformation, with the urban area expanding to 387.06 km², constituting 2.27% of the entire area. During this period, the annual growth rate was an impressive 25.47%, resulting in a total increase of 254.72 km² (Table 4). The geographical analysis indicates that this growth primarily occurred via infill development within established urban centers and the establishment of new settlement zones along transportation corridors (Figure 3). The growth observed during this period indicates that cities transitioned from isolated nodes to components of extensive networks.

The subsequent decade (2004–2014) saw a markedly distinct growth pattern. The urban expanse expanded to 434.74 km² (2.55% of the total area), however the yearly growth rate decreased to 4.77%. The total expansion throughout this period was 47.68 km², representing the most sluggish growth phase in the study's timeframe (Table 4). The spatial distribution pattern (Figure 3) indicates that existing urban areas are expanding in size but not significantly extending outward, suggesting a period of urban densification rather than extensive horizontal growth.

Over the past decade (2014–2024), there has been a resurgence in urban expansion, with the urban area increasing to 628.27 km², constituting 3.69% of the total area. During this period, the area expanded by 19.35% year, totaling 193.53 km² (Table 4). The spatial analysis indicates significant expansion at the peripheries, particularly in the southern and eastern regions of the study area (Figure 3). The growth pattern indicates both extensive development and corridor-based expansion, suggesting a complex interplay between planned development and organic urban growth.

Figure 4 illustrates that urban expansion is not a linear process over time. The urbanization curve indicates that the tendency is intensifying over the study period. Between 1994 and 2024, the urban area expanded by over 375%, with the greatest significant expansion occurring from 1994 to 2004. The overall trajectory indicates that urban expansion has not adhered to a consistent linear progression. Instead, it has seen periods of significant acceleration followed by relative stabilization, and in the past decade, it has begun to ascend again at an accelerated pace.

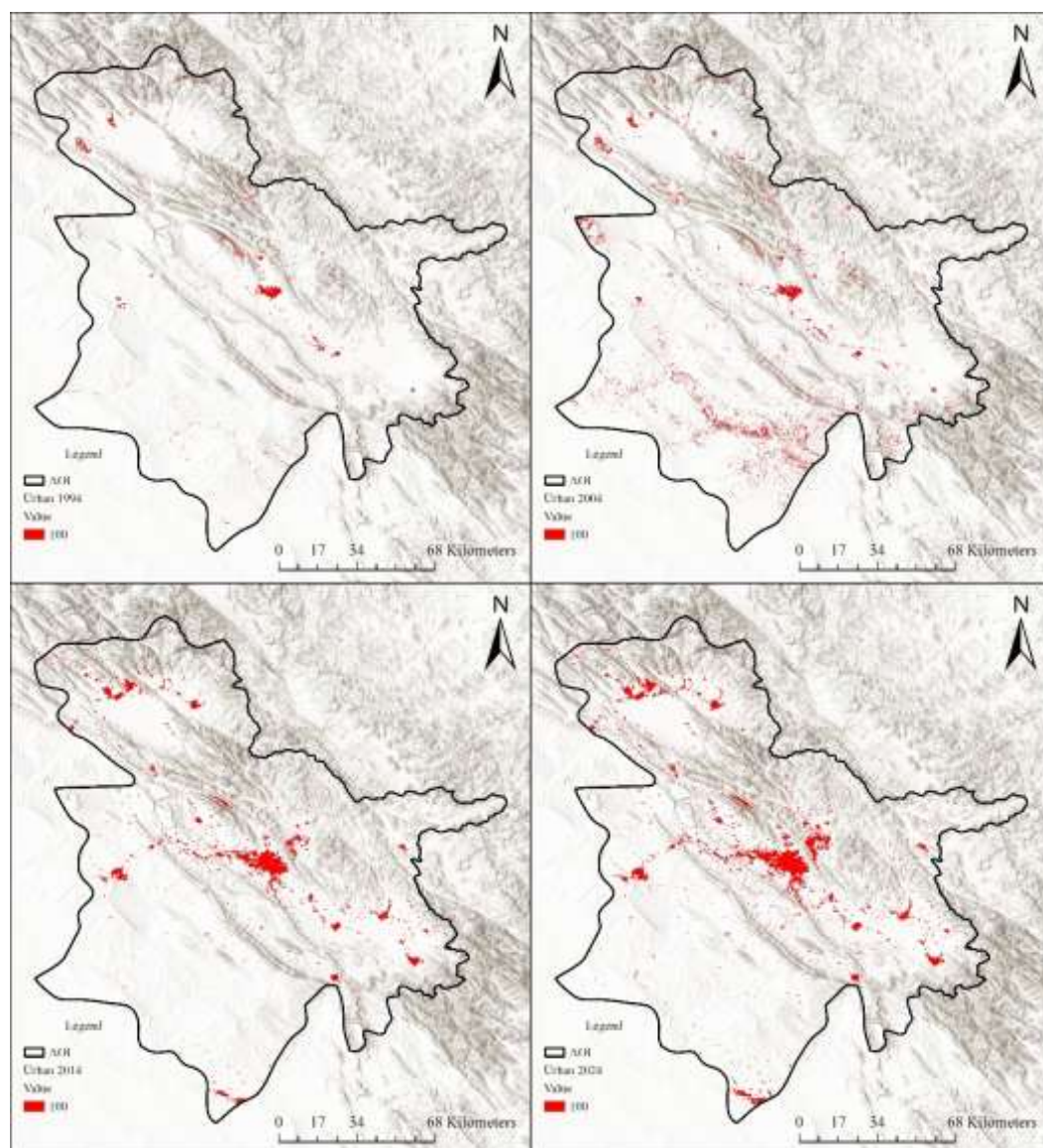


Figure 3. Maps of the urban areas in the study area for 1994, 2004, 2014, and 2024.

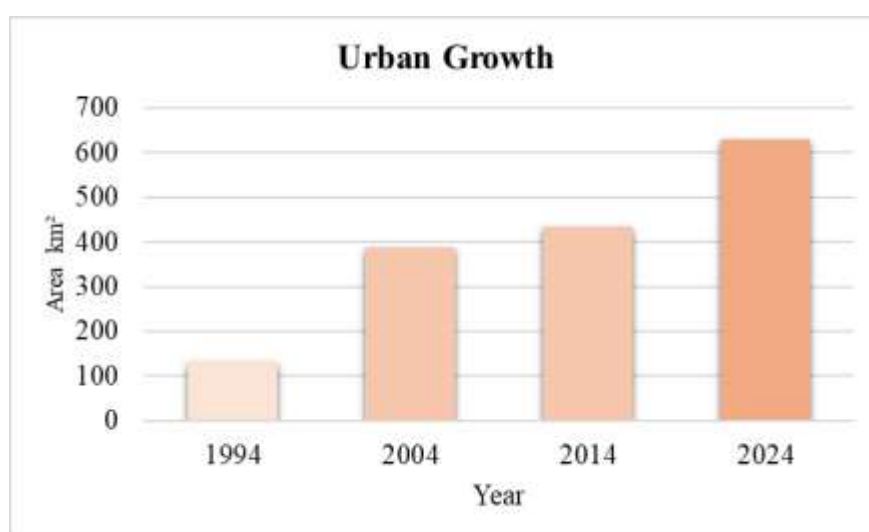


Figure 4. Urban growth trend in the study area.

Table 4. Urban Area Evolution (1994-2024).

Year	Urban Area (km ²)	Percentage (%)	Annual Growth Rate (%)	Absolute Growth (km ²)
1994	132.34	0.78%	/	
2004	387.06	2.27%	25.47%	254.72
2014	434.74	2.55%	4.77%	47.68
2024	628.27	3.69%	19.35%	193.53

Results of Network Analysis

An analysis of the transportation network reveals that the history of road infrastructure has mirrored the complex growth patterns of cities in the examined region. Significant alterations have occurred in the network's structure over the previous thirty years. Figure 5 illustrates the alterations in connection, density, and spatial distribution patterns, whereas Table 5 quantifies these changes. In 1994, the network's fundamental configuration comprised 17,589 nodes and 18,489 connecting edges. The total length of the network was 27.703 kilometers. The network had an average degree of 2.102 and a density of 0.000119, equivalent to a road density of 15.035 km per unit area. The spatial distribution indicates a concentrated network topology, characterized by primary arterial routes converging towards the city center. Conversely, outlying regions exhibit less connectivity (Figure 5). The network analysis indicates that the transportation infrastructure was initially incomplete, signifying the existence of isolated road segments and certain areas without connectivity altogether. During the initial decade of analysis (1994–2004), the network experienced gradual growth. The quantity of nodes increased to 17,941, and the quantity of edges rose to 18,991, resulting in a total length of 29.131 km. The average degree remained relatively constant at 2.117, while the network density exhibited minimal variation at 0.000118. The density of roads per square kilometer increased somewhat to 15.266. The physical arrangement (Figure 5) indicates that network connectivity has enhanced over time, particularly in the central metropolitan districts. Nonetheless, the overall network architecture remains fragmented, and connectivity gaps persist. The network continued to exhibit indications of disconnection, signifying that complete integration had not yet been achieved. Between 2004 and 2014, the network experienced modest growth. The quantity of nodes increased to 18,028, while the quantity of edges rose to 19,077, resulting in a total network length of 32.256 km. The mean degree remained constant at 2.116, however the network density experienced a little decline to 0.000117. The density of roads per unit area increased to 15.334 km, indicating an improvement in transportation infrastructure over time. The spatial analysis (Figure 5) indicates that previously poorly connected regions are now more interconnected due to the establishment of secondary road networks linking small towns and villages to primary transportation corridors. Despite these modifications, the network remained unconnected, indicating that structural fragmentation persisted during this period.

The past decade (2014–2024) seen the most significant transformations in the network. The quantity of nodes increased from 21,324 to 22,734, whilst the quantity of edges expanded from 22,734 to 38,789 kilometers. This period exhibited the highest average degree (2.132) and a significant increase in network density (0.000231), which was double that of the preceding decade's value. The road density attained 17.231 km per unit area, indicating substantial infrastructural development. The spatial distribution (Figure 5) indicates that the network has become denser and now encompasses areas that were previously isolated from it. The network exhibits enhanced connectivity, featuring more redundancy in pathways, improved accessibility to peripheral regions, and superior integration among urban centers. The network study indicates that complete connectivity has not been achieved, signifying those certain isolated regions or remote areas still lack direct access to the primary

transportation network. The alterations in network properties over time exhibit a non-linear growth pattern, with the most significant changes occurring between 2014 and 2024. The significant increase in network density over the past decade indicates a transition from basic connectivity to comprehensive network integration. This results from intentional investments in infrastructure to facilitate the swift expansion of urban areas. The persistent disconnection of the network over all study periods indicates the presence of geographical obstacles, planned development phases, or infrastructural limits that have hindered complete network integration, despite overall improvements in connectivity and accessibility (Table 5).

Table 5. Metrics of the road network for all study years.

Year	# Nodes	# Edges	Total Length (km)	Average Degree	Network Density	Road Density	Is connected
1994	17589	18489	27.703	2.102	0.000119	15.035	False
2004	17941	18991	29.131	2.117	0.000118	15.266	False
2014	18028	19077	32.256	2.116	0.000117	15.334	False
2024	21324	22734	38.789	2.132	0.000231	17.231	False

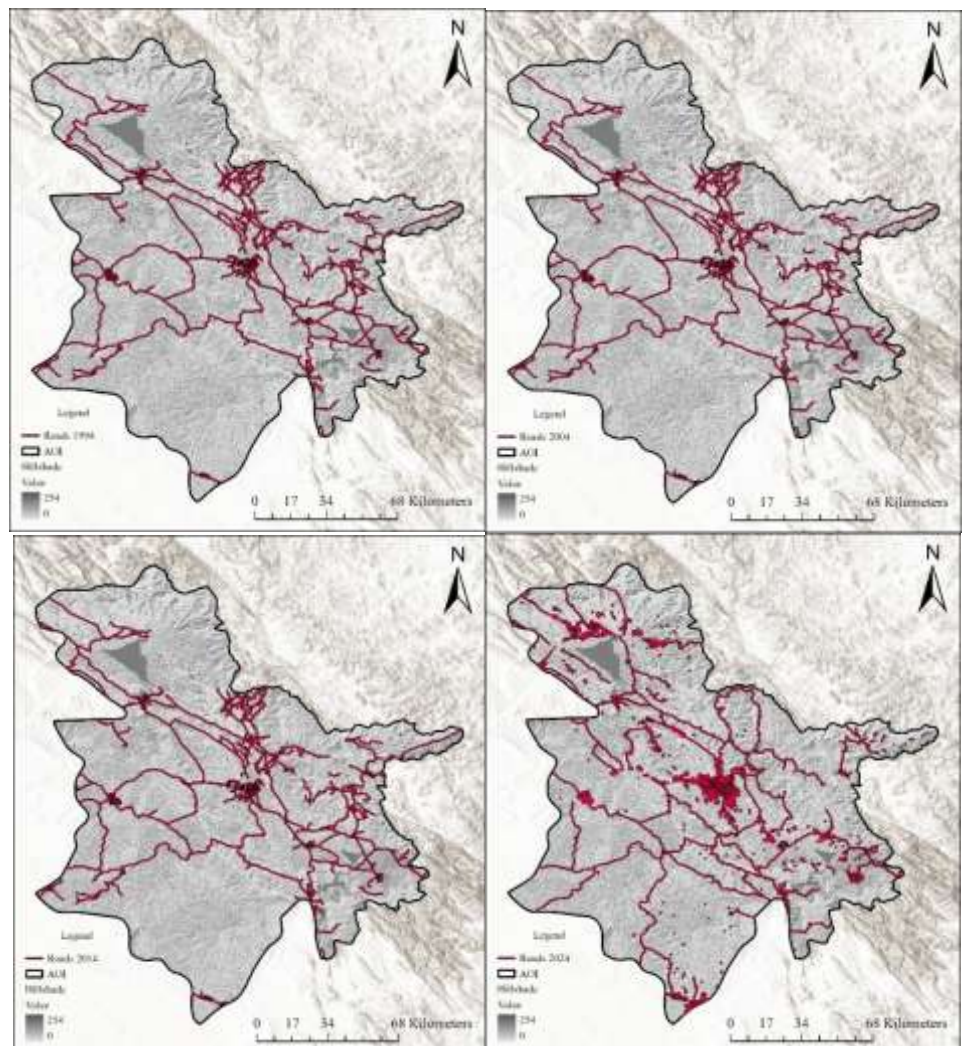


Figure 5. Maps of the road networks of Al-Sulaymaniyah city for 1994, 2004, 2014, and 2024.

Results of Co-evolution of Urban Growth and Network Analysis

The co-evolutionary analysis of urban expansion and transportation network development reveals their mutual dependence and the potential for divergent growth patterns throughout the research period. The statistical modeling of these connections elucidates the interplay between infrastructure and urban expansion throughout time, as well as the impact of these interactions on urban growth. The urban-network relationship model exhibits a moderate level of explanatory power, indicated by a R^2 value of 0.49. Nearly fifty percent of the alterations in network development can be attributed to urban growth trends (Table 6). The negative coefficient of -0.044 and the intercept of 40.61 indicate an unfavorable relationship between urban expansion and certain network characteristics. This may be due to fast urban expansion rendering network density metrics less precise. This seemingly illogical link may indicate that urban expansion has beyond the proportional enhancement of infrastructure, resulting in diminished efficiency of networks per unit of metropolitan area during periods of rapid growth. The distinct trend analyses indicate that the trajectories of urban and network expansion diverge significantly. The urban trend model exhibits substantial predictive capability, evidenced by a R^2 of 0.87, a positive coefficient of 16.105, and an intercept of -31,910.91. The robust correlation indicates that urban expansion adheres to a highly predictable trajectory across time, characterized by consistent growth rates that can be effectively analyzed using linear regression. The network trend model accounts for a limited portion of the data, exhibiting a R^2 of 0.58, a negative coefficient of -0.829, and an intercept of 1,688.21. The negative slope indicates a deterioration of certain network properties over time. This may be due to the challenges of maintaining network density and connection requirements amid rapid urban expansion.

The causality study provides critical insights into the directional relationship between urban expansion and network development. The Granger causality test yields a p-value of 0.5, indicating a lack of statistically significant evidence that urban growth influences road expansion at conventional significance levels (Table 6). This study contradicts the prevailing notion that urban growth necessitates the construction of infrastructure. It indicates that network development may occur independently during its own planning cycles or in reaction to external factors. The correlation coefficient of -0.483 substantiates the negative relationship identified in the regression analysis, indicating that as urban areas expand, network performance metrics decline. The correlation investigation indicates intricate relationships between metropolitan regions and other network characteristics (Table 7). The Pearson correlation between metropolitan area and road length indicates a robust positive relationship ($r = 0.701$, $p < 0.001$), signifying that as cities expand, their road networks often increase in size. The Spearman correlation is considerably lower ($r = 0.399$, $p = 0.6$), indicating that this relationship may not be linear and could be influenced by outliers or non-linear patterns. The link between metropolitan area and network density exhibits analogous characteristics. It has a robust Pearson correlation ($r = 0.723$, $p < 0.001$) but a non-significant Spearman correlation ($r = 0.799$, $p = 0.2$). This indicates the potential existence of non-linear relationships or threshold effects in the dynamics of urban-network density. The correlation between urban area and road density is identical to that between road length and road density, exhibiting equivalent correlation coefficients (Pearson: $r = 0.701$, $p < 0.001$; Spearman: $r = 0.399$, $p = 0.6$). This consistency indicates that total road length and road density respond similarly to the stresses of urban expansion. The analysis of the growth rate correlation reveals a slight positive Pearson correlation ($r = 0.34$, $p < 0.001$) and a non-significant Spearman correlation ($r = 0.5$, $p = 0.66$). This indicates that correlations between growth rates are not robust and may fluctuate over time or exhibit threshold effects. The co-evolutionary approach demonstrates that urban expansion and network development occur inside a complex adaptive system, wherein certain events transpire simultaneously while others do not.

The significant correlations in absolute terms between metropolitan areas and road length and density indicate that infrastructure development typically aligns with the demands of expanding cities. Nonetheless, the inverse relationships seen in the regression models and the absence of clear causality suggest that this response may be insufficient or misaligned temporally. The findings indicate that urban expansion may be outpacing infrastructure development, potentially resulting in diminished network performance metrics despite an increase in road quantity. This is a prevalent trend in developing metropolitan systems, since swift growth hampers infrastructure systems' ability to maintain service quality and connectivity standards.

Table 6. Results of co-evolution model between urban growth and road networks.

Co-evolution Model	R ²	Coefficient/slope	Intercept
Urban-Network Relationship	0.49	-0.044	40.61
Urban Trend	0.87	16.105	-31910.91
Network Trend	0.58	-0.829	1688.21
Causality			
P-value	0.5		
Correlation	-0.483		
Urban growth causes road expansion	False		

Table 7. Results of correlations between urban growth and different road network characteristics.

Correlation Type	Pearson	Spearman
Urban Area vs Road Length	0.701 (pvalue=0)	0.399 (pvalue=0.6)
Urban Area vs Network Density	0.723 (pvalue=0)	0.799 (pvalue=0.2)
Urban Area vs Road Density	0.701 (pvalue=0)	0.399 (pvalue=0.6)
Growth Rate	0.34 (pvalue=0)	0.5 (pvalue=0.66)

4. Conclusion

This study uses a combination of Random Forest classification and network analysis to give a full picture of the complicated relationship between urban expansion and transportation infrastructure in Al-Sulaymaniyah City. The Random Forest classifier worked very well, with an overall accuracy of 93%. This shows that it is a good tool for mapping land cover in urban areas with a lot of different types of land, even when data is hard to come by in cities that are growing quickly. The quantitative analysis shows that Al-Sulaymaniyah underwent an unprecedented urban transformation over the course of three decades, with built areas growing by 375% and showing non-linear growth patterns that included periods of rapid expansion (1994–2004 and 2014–2024) and periods of consolidation (2004–2014). The development of the transportation network at the same time, while significant in absolute terms, did not keep up with the proportionate increase needed by cities. The difference is shown by the negative connection coefficients in co-evolutionary models and the fact that there is no statistically meaningful link between urban expansion and infrastructure provision. The correlation study shows that there is a big gap between the procedures of planning for urban growth and planning for infrastructure. There are large positive correlations between metropolitan areas and absolute network metrics, but there is no Granger causality. This means that transportation infrastructure expansion happens on its own planning cycles instead than in response to urban needs. This trend suggests that there may be problems with governance when cities grow faster than their infrastructure can handle, which raises questions about the long-term sustainability of urban development.

The methods used in this study can be used in other cities besides Al-Sulaymaniyah to explore how cities evolve together. Combining machine learning-

based classification with graph-theoretic network analysis could be very useful for evidence-based urban planning in places where there isn't a lot of data. The results stress how important it is to build infrastructure ahead of time so that it can handle the challenges of urban growth instead of just reacting to them. Future urban planning should focus on coordinated development methods that make sure infrastructure is built in line with the way cities are growing. The study shows that sustainable urban development needs governance systems that can coordinate land use planning with the building of transportation infrastructure. These ideas are especially important for cities in developing areas that are growing quickly, where comparable problems with co-evolution could hurt long-term sustainability and livability.

REFERENCES

- [1] United Nations Department of Economic and Social Affairs, *The Sustainable Development Goals Report 2018*. New York, NY, USA: United Nations, 2018.
- [2] R. Singh, K. Shah, and G. Sharma, "Evolving road networks and urban landscape transformation in the Himalayan foothills, India," *Environmental Monitoring and Assessment*, vol. 196, no. 12, p. 1164, 2024, doi: 10.1007/s10661-024-13303-9.
- [3] D. Mann, S. Rankavat, and P. Joshi, "Road network drives urban ecosystems - a longitudinal analysis of impact of roads in the central Himalaya," *Geocarto International*, vol. 37, pp. 1100–1125, 2020, doi: 10.1080/10106049.2020.1750064.
- [4] C. Zeng, Z. Zhao, C. Wen, J. Yang, and T. Lv, "Effect of complex road networks on intensive land use in China's Beijing-Tianjin-Hebei urban agglomeration," *Land*, vol. 9, no. 12, p. 532, 2020, doi: 10.3390/land9120532.
- [5] D. Kasraian, S. Raghav, and E. Miller, "A multi-decade longitudinal analysis of transportation and land use co-evolution in the Greater Toronto-Hamilton Area," *Journal of Transport Geography*, vol. 84, p. 102696, 2020, doi: 10.1016/j.jtrangeo.2020.102696.
- [6] C. Li, X. Gao, B. He, J. Wu, and K. Wu, "Coupling coordination relationships between urban-industrial land use efficiency and accessibility of highway networks: Evidence from Beijing-Tianjin-Hebei urban agglomeration, China," *Sustainability*, vol. 11, no. 5, p. 1446, 2019, doi: 10.3390/su11051446.
- [7] X. Li and L. Parrott, "An improved Genetic Algorithm for spatial optimization of multi-objective and multi-site land use allocation," *Computers, Environment and Urban Systems*, vol. 59, pp. 184–194, 2016.
- [8] A. Pratama and M. Yudhistira, "Highway expansion and urban sprawl in the Jakarta Metropolitan Area," *Land Use Policy*, vol. 114, p. 105856, 2022, doi: 10.1016/j.landusepol.2021.105856.
- [9] K. Mhana, S. Norhisham, H. Katman, and Z. Yaseen, "Road urban planning sustainability based on remote sensing and satellite dataset: A review," *Heliyon*, vol. 10, 2024, doi: 10.1016/j.heliyon.2024.e39567.
- [10] A. Allan, A. Soltani, M. Abdi, and M. Zarei, "Driving forces behind land use and land cover change: A systematic and bibliometric review," *Land*, vol. 11, no. 8, p. 1222, 2022, doi: 10.3390/land11081222.
- [11] F. Ahmadzai, "Analyses and modeling of urban land use and road network interactions using spatial-based disaggregate accessibility to land use," *Journal of Urban Management*, vol. 9, pp. 298–315, 2020, doi: 10.1016/j.jum.2020.06.003.
- [12] V. Kumar and S. Agrawal, "Urban modelling and forecasting of landuse using SLEUTH model," *International Journal of Environmental Science and Technology*, vol. 20, pp. 6499–6518, 2022, doi: 10.1007/s13762-022-04331-4.
- [13] M. Iacono, D. Levinson, and A. El-Geneidy, "Models of transportation and land use change: A guide to the territory," *Journal of Planning Literature*, vol. 22, no. 4, pp. 323–340, May 2008, doi: 10.1177/0885412207314010.
- [14] J. Kleemann, G. Baysal, H. N. N. Bulley, and C. Fürst, "Assessing driving forces of land use and land cover change by a mixed-method approach in north-eastern Ghana, West Africa," *Journal of Environmental Management*, vol. 196, pp. 411–442, Jul. 2017, doi: 10.1016/j.jenvman.2017.01.053.
- [15] D. Liu, K. C. Clarke, and N. Chen, "Integrating spatial non-stationarity into SLEUTH for urban growth modelling: A case study in the Wuhan metropolitan area," *Computers, Environment and Urban Systems*, vol. 84, p. 101545, 2020.
- [16] S. Padma et al., "Simulation of land use/land cover dynamics using Google Earth data and QGIS: A case study on Outer Ring Road, Southern India," *Sustainability*, vol. 14, no. 24, p. 16373, Dec. 2022, doi: 10.3390/su142416373.

- [17] W. Mo, Y. Wang, Y. Zhang, and D. Zhuang, "Impacts of road network expansion on landscape ecological risk in a megacity, China: A case study of Beijing," *Science of the Total Environment*, vol. 574, pp. 1000–1011, 2017, doi: 10.1016/j.scitotenv.2016.09.048.
- [18] X. Zhai et al., "Classification of Arctic sea ice type in CFOSAT scatterometer measurements using a random forest classifier," *Remote Sensing*, vol. 15, no. 5, p. 1310, Feb. 2023, doi: 10.3390/rs15051310.
- [19] Y. Casali, N. Y. Aydin, and T. Comes, "Machine learning for spatial analyses in urban areas: A scoping review," *Sustainable Cities and Society*, vol. 85, p. 104050, Oct. 2022, doi: 10.1016/j.scs.2022.104050.
- [20] A. Mustafa, A. Rienow, I. Saadi, M. Cools, and J. Teller, "Comparing support vector machines with logistic regression for calibrating cellular automata land use change models," *European Journal of Remote Sensing*, vol. 51, no. 1, pp. 391–401, Jan. 2018, doi: 10.1080/22797254.2018.1442179.
- [21] S. Zhao, K. Tu, S. Ye, H. Tang, Y. Hu, and C. Xie, "Land use and land cover classification meets deep learning: A review," *Sensors*, vol. 23, no. 21, p. 8966, Nov. 2023, doi: 10.3390/s23218966.
- [22] X. Xu, D. Zhang, X. Liu, J. Ou, and X. Wu, "Simulating multiple urban land use changes by integrating transportation accessibility and a vector-based cellular automata: A case study on city of Toronto," *Geospatial Information Science*, vol. 25, no. 3, pp. 439–456, Jul. 2022, doi: 10.1080/10095020.2022.2043730.
- [23] J. Raimbault and F. N'echet, "Introducing endogenous transport provision in a LUTI model to explore polycentric governance systems," *Journal of Transport Geography*, vol. 94, p. 103115, 2021, doi: 10.1016/j.jtrangeo.2021.103115.
- [24] G. Zhao, X. Zheng, Z. Yuan, and L. Zhang, "Spatial and temporal characteristics of road networks and urban expansion," *Land*, vol. 6, no. 2, p. 30, Apr. 2017, doi: 10.3390/land6020030.
- [25] R. Ahasan, Md. S. Alam, T. Chakraborty, and Md. M. Hossain, "Applications of GIS and geospatial analyses in COVID-19 research: A systematic review," *F1000Research*, vol. 9, p. 1379, Jan. 2022, doi: 10.12688/f1000research.27544.2.
- [26] F. Bao et al., "Advancing cloud classification over the Tibetan Plateau: A new algorithm reveals seasonal and diurnal variations," *Geophysical Research Letters*, vol. 51, no. 13, p. e2024GL109590, Jul. 2024, doi: 10.1029/2024GL109590.
- [27] G. R. Fage, P. A. Ibrahim Saied, and H. M. Hameed, "Urban growth prediction using cellular automata Markov: A case study using Sulaimaniya city in the Kurdistan Region of North Iraq," *IISTE Humanities and Social Science*, vol. 6, pp. 108–118, 2016.