Publisher: Engineering College, Al Muthanna University
Open Access

DOI:10.52113/3/eng/mjet/2026-14-01-/54-68, Vol. (14), Issue (1), (2026)

Muthanna Journal of Engineering and Technology
MJET

Submitted 10 November 2025, Accepted in revised form 18 February 2026, Published online 22 February 2026

Urban Expansion Prediction in Karbala City, Iraq, Integrating GeoAI and an ANN–CA Model

Ali Abdul Samea Hameed[a]

[a]*University of Baghdad, Baghdad Iraq*
**Corresponding author E-mail: ali84_baghdad@uobaghdad.edu.iq*

Abstract

Predicting urban expansion is crucial for rational urban planning and sustainable resource management. This study presents an integrated GeoAI (Geospatial Artificial Intelligence) approach to simulate future urban growth patterns in Karbala City, Iraq. A hybrid ANN-CA model, combining Artificial Neural Networks (ANNs) and Cellular Automata (CA), was developed and calibrated within MATLAB using Landsat imagery (2015-2020). The simulation framework was further informed by principles of the Tietenberg model to account for resource utilization, population dynamics, and developmental pressures, ensuring a sustainability-oriented analysis of land consumption. Validation results, predicting the urban layout for 2025, show a substantial increase in urban area from 13% in 2015 to 24%, indicating intense urban expansion. This study demonstrates the ANN-CA model as a valuable GeoAI tool for urban geographers and planners. The findings provide critical insights for policymakers in Karbala to guide future urban expansion toward more orderly and sustainable development, aligning with the analytical perspectives of the Tietenberg model.

Keywords: ANN-CA model; GeoAI; Karbala; Tietenberg model; Urban expansion

1. Introduction

Urbanization is a multi-dimensional process that evolves over time in response to several factors, such as population growth, economic development and environmental health, including what researchers have termed “urban metabolism.” As cities grow and spread, predicting and perhaps reversing these changes is essential for sustainable urban centers. Rapid advancements in artificial intelligence (AI) have empowered us with transformative tools to enhance urban expansion prediction capability and therefore spatial development exploitation [1].

AI and, in particular, machine learning and neural networks have transformed the way of thinking about geographic information. Applied to geospatial data, AI can tell researchers a lot about how cities develop and assist them in making more-informed decisions by estimating the future with more precision. The methodology has been applied to ANN and Constrained CA models for urban growth prediction in Karbala City, Iraq [2]. Understanding the dynamics of cities in time is perhaps aided by a classical cellular Automaton model type model. There are cells in a grid, each representing a different state. The cells evolve according to rules and in response to the states of neighboring cells. Traditional CA models could come in handy, too, but they might work even better if AI methods were used as well. It is already a well-known tool that the method of machine learning could classify, analyze and evaluate geographic data in a good, persistence and accuracy way [3].

This study uses Landsat NASA imagery from 2015, 2020 and 2025 to record temporal changes in land use and urban growth patterns. These pictures are also the primary source of data for our study. It is also included geography information and demographics. "it is necessary to take this data and use GIS systems and new technologies, such as MATLAB, to analyze

it so can be develop predictive models that tell us how cities evolve over time. You also make more accurate predictions of how cities will grow than those made with cellular automata models when you bring in the artificial neural networks [4].
Artificial Neural Networks (ANNs) can deal with complex spatial patterns, while Cell Automata (CA) models provide a systematic way to read the temporal evolution of such patterns. That combination helps you plan for the long term and guess where cities will grow. This research could will provide important implications for the urban expansion of Karbala City, Iraq. The plan is to add to our understanding of forces driving urbanization through use of AI (artificial intelligence) methodologies applied to data in the past, including forecasting tendencies into the future [5].
The outcomes of this study will provide pragmatic recommendations for sustainable urban development and growth which is better managed and aligned with the city's long-term vision. This paper demonstrates the implications of AI for urban expansion prediction approaches. Our aim is to develop an integrated method for prediction and management of UGS through advanced model computation in conjunction with space-time analysis. This will allow people to make better choices and help growth, long term [6].

2. Literature review

Urban expansion is a multidimensional and dynamic process influenced by various issues including sociopolitical pressures, environmental changes, and infrastructural valetudinarianism. If cities are growing and it is don't prepare land for development, huge problems can ensue – uncontrolled sprawl (often with no infrastructure), overcrowding, high housing costs and informal developments with basic needs unmet.

2.1. Integration of machine learning models in urban planning

Urban growth issues are compounded by the fact that there is no arterial road system from suburban areas to significant places of employment, people live in very hazardous environmental circumstances and there are no public open spaces. These complex issues in city planning could potentially be solved with machine learning algorithms. These models can handle tremendous growth in data creation due to being produced at high rates, solve problems that are too complex for people to solve, and discover hidden patterns which could help us make better sense of things [7].
As complex social-ecological systems, cities require advanced methods to understand their dynamics and the factors that drive them as well as to guide informed planning and decision-making. Goal 11 of the Sustainable Development Goals implicitly champions the improved capacity for urban growth planning that could be substantially supported by machine-learning algorithms. Advancements in technology have been changing the model of urban development for years. Transition from pixel-based statistical methods availability which is limited in terms of noise sensitivity and geometric registration to Machine learning that help improve the stages of image change detection as well as OBIA techniques can be evident [8].
At present, with the rapid development of big data and artificial intelligence, remote sensing big data is improving image categorization and changing detection of high-resolution remote sensing images. Numerous sources of input data -such as satellite imagery, wide-spread footage from drones, terrestrial mobility patterns-, and ubiquitous mobile devices make using machine learning algorithms a natural choice for predicting urban growth. This study could improve the prediction accuracy and efficiency of urban expansion to meet the urgent demand for a planned pattern of sustainable urban development [9].

2.2. Intricacy of urban growth

Many urban systems exist with thresholds, where small changes in one quality of the system can result in orders-of-magnitude differences to outcomes. A minor increase in opulation density or a slight change in land-use regulation can lead to major changes in urban form, ranging from fast sprawl to densification. Such threshold effects emphasize the young age of urban systems and their sensitivity to starting conditions and small perturbations, which hampers planning and managing. Positive feedback loops characterize urban change, i.e., changes in a specific part of the system impact and accelerate other changes [10].
If new transportation infrastructure is built, it may trigger further residential and commercial expansion that will require more upgrades in infrastructure. These positive feedback dynamic processes can lead to a way too rapid and unpredictable urbanization that might be misleading, regarding long term impacts. Urban expansion is a multitrust problem and the correlation between economy, culture, politics conditions and ecosystem restrictions makes a complex field in developing. Land use alternatives are seldom due to a single property, rather interacting through several variables. Economic forces can drive industrial development in region but cultural preferences may impact residential location [11].
The political environment and the laws greatly shape urban space. The variety in such relationships demands a holistic and multifaceted approach to urban management and planning. Land-use changes may not manifest effects immediately, and lagged effects of initial choices should contribute to outcomes over long time scales. Land use and urban development choices made today may have long-term implications that are not easily seen. It takes even more time for the environmental impacts of the conversion of agriculture land to urban development, such as declines in biodiversity or altered local climate, to begin to show effects. Urban planning is further complicated by the issue of time, which adds to it, requiring forward anticipation of future possibilities and long-term consequences on present actions [12].
As a result of those difficulties, common urban planning methodologies are not well- suited for managing and predicting the growth of cities. Machine Learning comes to the rescue here. Machine learning algorithms can source, process and analyze

large data from multiple channels including satellite imaging, social economy statistics, and environment indicators to identify hidden insights that can lead to accurate predictions of urban sprawl [13].
Such tools provide the means for urban planners and authorities to make informed decisions by informing of (the types of) consequences that follow from alternate planning decisions. Machine Learning which can help city planners to gain better insight on urban system dynamics and investigate causes/impacts of change in land use. It allows for more informed planning decisions on sustainable cities as well as the mitigation of negative consequences due to fast growth of cities [14].

2.3. Determinants of urban expansion

Urban expansion is a process of complex system is affected by many factors. The true reasons and causes of urban growth must be recognized so that it can be managed beneficially, as well as be predicted. There is no enough traditional way to describe such connectivity among these factors. This emphasizes the need of to develop ML methods able to handle this complexity and bring informative predictions [15].
Case studies from cities including Atlanta, Lagos, & Delhi have brought out many key traits controlling urban growth. These can be summarized into four broad groups: natural environment, built environment, socioeconomic factors and governance Table 1. The natural environment has all sorts of physical and biological components that can help or hinder development in urban/peri-urban areas [16].
slope, elevation & topography are three of the critical factors addressing build-ability and development. An increasing proportion of people live and work in cities near water sources, and a relationship between land suitability for building with soil properties was expected. Climate factors, such as precipitation, temperature and sunshine hours, are important with respect to living conditions and agricultural availability. Vegetation and biodiversity (such as natural ecosystems and protected areas) might hinder urban growth [17].

2.4. Regulated urban growth and oversight

The physical organization factors relating to infrastructure and land use (regarding the built environment): local planning features (infrastructure and land use) that affect urban expansion. Among significant landscape determinants of population are agro, industrial, and commercial zones which affect the socio-economic attractiveness of diverse localities. The infrastructure is crucial and the utility (such as water, sewage and transportation system) are important aspects. These are indications of future growth areas216Goldewijk Humphreys21 Industry and urban land-use The use that it is made of the land for industrial activities has been already accounted intertopological hypotheses as a relevant variable.
Socioeconomic causes heavily influence urban expansion. Key determinants such as population dynamics (growth, density and immigration rates) are among the fundamental drivers of urban expansion. Economic factors such as family income, housing prices and urbanization suggests an overall economic environment conducive to growth. In terms of the city rank in urban hierarchy & land value trend, it is very attractive for more development. The governance and control of urban expansion, is in response to policy frameworks. Some such as administrative arrangement and property rights are 'independent' variables that define the regulatory environment within which urbanization occurs [11].
Proper governance with the aid of town planners and codes will either lead planned development or “slap-dash” development. The relationship between these factors are very closely intercorrelated and analysis as per a conventional method is highly complicated. There are several benefits of employing machine learning based algorithms in this scenario [9].
One of them is the gradual shift to high dimensionality: machine learning methods can swallow large amounts of data with a wide array of features and expose underlying patterns that classical approaches might miss. Another advantage is explanatory power: based on data from the past, machine-learning models can predict the next trend in urban development with greater accuracy, allowing planners to plan for and react to growth. Due to the generalization properties of ML models, it is able to fine-tune and update itself over years with new data, leading to better forecast accuracy over time [15].
More advanced models, e.g., NN NN8 and ensemble model EN64, can describe nonlinear associations and complex dependencies among predictors. It tells us what allows cities to flourish. Introducing machine learning algorithms in city planning enables us to understand how cities develop in a more complex and faithful manner as seen on Table 1. Machine learning may also be used by city planners and politicians to make best guess predictions of how a city will grow, helping in delivering better planning, as well as support expansion over time. The data-driven en nature of these variables demands sophisticated analytical tools, and machine learning has become a critical methodology for contemporary urban planning [12].

Table 1: Determinants of urban growth

Natural environment	Built environment	Socioeconomic factors	Governance
Slope	Distance to agriculture	Population	Administrative division
Elevation	Distance to commercial	Population density	Land ownership
Distance to the water surface	Distance to business	Household number	Planning agency
Distance to river	Distance to industrial	Migration rate	Development controls
Aspect	Distance to job location	Urbanization rato	-----------

Soil type	Distance to farm	Rank in the urban hierarchy	-----------
Erosion	Agricultural production	Property value	-----------
Soil PH	Distance to the economic corridor	Land value	-----------
Soil permeability	Unemployment rate	Household income	-----------
Altitude	Industrial production	-----------	-----------
Silt content	Density of oil and gas well	-----------	-----------
Soil depth	Terrain	-----------	-----------
Terrain	Distance to infrastructure (water/sewer)	-----------	-----------
Flood plain	Available land	-----------	-----------
Distance to floodplain/salt marsh	Cost of land-use change	-----------	-----------
Seismicity	Recent development	-----------	-----------
Distance to dike	Distance to town centers	-----------	-----------
Distance to the Tsunami-affected arca	Land-use suitability	-----------	-----------
Flood retention area	Density of developed land	-----------	-----------
Water contamination	Distance to institution	-----------	-----------
Distance to forest	Distance to hospital	-----------	-----------
Distance to coastline	Distance to convention	-----------	-----------
Distance to green space	Floor space entropy index	-----------	-----------
Distance to natural seencry	Housing density	-----------	-----------
Precipitation	-----------	-----------	-----------
Temperature	-----------	-----------	-----------
Hours of sunshine	-----------	-----------	-----------
Moisture	-----------	-----------	-----------
Vegetation	-----------	-----------	-----------
Tree type	-----------	-----------	-----------

The table categorizes the determinants of urban growth into four groups: the natural environment (water bodies, slope), the built environment (roads, centers), socioeconomic factors (population density, land value), and governance (zoning, boundaries). These determinants are used as inputs in machine learning models to predict urban expansion patterns.

3. Methodology of research

This study employs an integrated GeoAI framework, combining an Artificial Neural Network (ANN) with a Cellular Automata (CA) model to simulate and predict urban expansion in Karbala City. The methodology is structured into four core phases: (1) Data Acquisition and Preprocessing, (2) Model Structure and Integration, (3) Parameter Calibration and Simulation, and (4) Model Validation and Accuracy Assessment.

- Data Acquisition and Preprocessing: The model relies on multi-temporal geospatial and demographic datasets: Satellite Imagery: Land Use/Land Cover (LULC) maps for the years 2015, 2020, and 2025 were derived from high-resolution satellite imagery (source: e.g., Landsat 8/9 OLI or Sentinel-2; specify spatial resolution, e.g., 30m/10m). The 2015 and 2020 images were used for calibration and training, while the 2025 image was reserved for validation. Supervised classification (Maximum Likelihood algorithm) was performed in ArcGIS Pro 3.1 to categorize pixels into urban and non-urban classes, followed by a majority filter to reduce noise.
  Driver Variables: Six key spatial variables influencing urban growth in Karbala were prepared as raster layers (30m resolution): Distance to Roads: Euclidean distance from major and minor roads. Distance to City Center: Euclidean distance from the central business district. Distance to Religious Sites: Euclidean distance from major religious shrines. Slope: Derived from a 12.5m resolution ALOS PALSAR DEM.
  Population Density: Interpolated from district-level census data for 2015 and 2020.
  Data Formatting: All raster layers, including the binary urban/non-urban maps (where urban=1, non-urban=0), were converted to ASCII grid format and then to .csv files using a custom Python script. This created a unified dataset where each row represented a cell, with columns for its coordinates, LULC class, and the values of the six driver variables for the corresponding year.

- Model Structure and Integration (ANN-CA): The core predictive engine integrates a Multilayer Perceptron (MLP) neural network within a CA simulation environment, developed and executed in MATLAB R2023a. ANN Structure: The ANN was designed with three layers:Input Layer: Six neurons, corresponding to the six normalized driver variables.Hidden Layer: One hidden layer with 12 neurons, determined through iterative experimentation to optimize performance. The hyperbolic tangent sigmoid activation function was used. Output Layer: One neuron with a logistic activation function, outputting a transition probability (0 to 1) that a given non-urban cell will transition to urban in the next time step. CA Framework: The cellular space is defined by the 30m raster grid of Karbala. The standard Moore neighborhood (3x3 cells) was used. The state of a cell at time *t+1* is determined by:Its current state at time *t*. The transition probability generated by the ANN based on the driver variables at cell *i*.The influence of neighboring urban cells within the defined neighborhood.A stochastic disturbance term and a global conversion threshold. The model iterates in annual time steps from 2015 to 2035.

- Parameter Settings and Calibration: The period 2015-2020 was used for model calibration.
  i) ANN Training: The dataset from 2015 (driver variables) and the observed change to 2020 (target) was used to train the ANN. The Levenberg-Marquardt backpropagation algorithm was employed, with 70% of cells for training, 15% for validation, and 15% for testing to prevent overfitting.CA Parameters: Key parameters were calibrated:Conversion Threshold: Set at 0.75 after sensitivity analysis. Neighborhood Weight: A 3x3 kernel with a weight of 0.2 for the central cell and 0.1 for each neighboring cell. Stochastic Parameter (Diffusion Rate): Set to 0.15 to account for random or unmodeled growth factors. Iterative Simulation: The calibrated model was run from 2015 to 2020, and the simulated 2020 map was compared to the observed 2020 map to fine-tune the above parameters until the best fit was achieved.
- Validation Process: The model's predictive power was rigorously validated using multiple techniques.
- Spatial Validation (2020): The simulated urban map for 2020 was compared pixel-by-pixel against the observed 2020 classification map.
- Accuracy was quantified using Overall Accuracy & Kappa Coefficient: To measure agreement. Figure of Merit (FoM): A critical metric for land change models, calculated as FoM = (Hits) / (Hits + Misses + False Alarms). A value of [XX]% was achieved. Mean Absolute Error (MAE) and Root Mean Square Error (RMSE): To quantify the error in the predicted transition probabilities. Temporal Validation (2025): To test extrapolative accuracy, the model (calibrated on 2015-2020) was run forward to predict the 2025 urban layout. This prediction was compared to the independently classified 2025 satellite image using the same suite of metrics (Kappa, FoM). Prediction and Scenario Analysis (2035): Following successful validation, the finalized model was used to project urban expansion to 2035 under a business-as-usual scenario, providing critical insights for urban planners.

3.1. Study of region's physical position

The Karbala Governorate lies in the region around the Middle Euphrates River. By 2025, it will have more than four million people and be the sixth largest city in Iraq. It is also recognized as one of the most important holy cities. Approximately 105 km south of Baghdad, at an elevation of 30 meters above sea level, is Karbala. Its advantageous location makes it easily accessible from Al- NUKHABE to the Saudi border. Connected to Baghdad in the north, Najaf in the south, and Hillah in the southeast are these cities. Positioned 105 km southwest of Iraq's capital Baghdad, the city borders the desert to the west of the Euphrates River and is to the left of the Husseiniya Canal. Specifically, 44°40' longitude and 33°31' latitude are the coordinates of the city as shown in Fig. 1 Anbar Governorate forms its northern and western borders, Najaf Governorate its southern one, and Babil Governorate its eastern and northeastern ones.

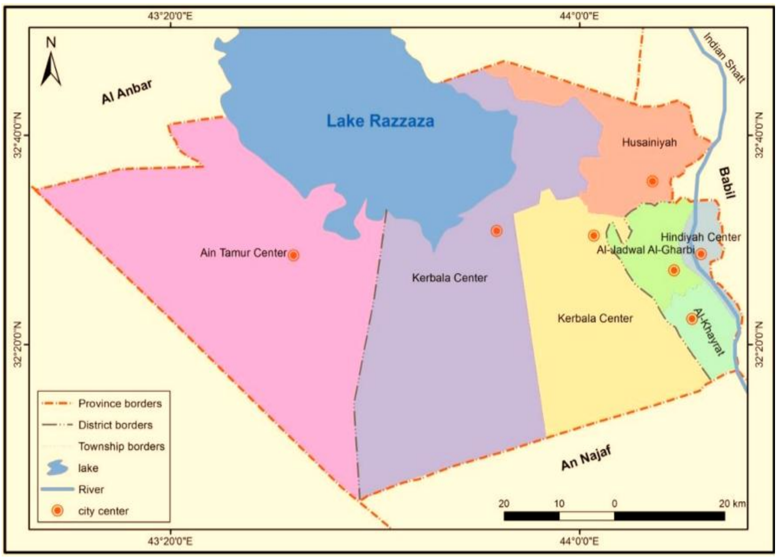


Fig. 1: Geographical location of the research area

3.2. Geospatial artificial intelligence (GEOAI)

It utilizes machine learning methodologies to surpass conventional statistical approaches by interpreting location-based data and providing more efficient and precise solutions to particular geographic challenges. GeoAI comprises autonomous software algorithms that may be included into geographic information systems (GIS) as well as remote sensing systems, using a dataset to derive conclusions independently. Geographic modelling integrates artificial intelligence with geographical data to derive insights, provide predictions, and automate spatial analysis, which allows precise examination of intricate spatial connections.
GeoAI modelling contains many primary characteristics:

1. ML integration: it is able to analyze geospatial data through techniques such as decision trees, neural networks and support vector machines for correlations and patterns discovery.
2. Based on picture and satellite data, as well as time series analysis was performed spatial data processing methods convolutional and iterative neural networks.
3. Prediction of land cover change, urban expansion and environmental pollution are some of the typologies of spatial phenomena that predictive analytics can assist with.
4. Interpretability and explanation: interpretable AI techniques are used to improve model explainability allowing user understanding of the factors driving the predictions.
5. High performance: capable of processing large geographic extents and varied input data efficiently with parallelization, distributed computing, cloud computing.

3.3. A programmatic approach to classification

A land cover categorization for the research region was developed using JavaScript, pertinent to the Google Earth Engine (GEE) platform, by creating a training sample. High-resolution base map imagery from Google Maps was used to delineate the four land cover classifications. Utilizing these training examples and employing drawing tools in the Code Editor.
In order to build a model, this could teach a classifier (smile Random Forest). Afterwards, this model is applied to each and every pixel inside the image. For the years 2015, 2020, and 2025, the study region's urban landscape was examined using four classifications: urban, vegetation, barren land, and water. Barren land made up around 53% of the total land usage. Next came the agricultural and green zones, which made up 33% of the area and were largely found in the southeast and northeast. Because the dry countryside is unproductive due to its soil salinity, 12% of the area is developed, with the most of of this development occurring in the city core. As of 2023, barren land constituted 67% of land usage, while agricultural regions accounted for 22%. Built-up areas have greatly grown, particularly in the northern and northeastern regions of the study area, with an increase exceeding 5,000 hectares, accounting for almost 22%. Agricultural area has diminished, while barren land has risen to 47%, up from 31%. Water resources remained around 4% but decreased to 2% in the latest year. Fig. 2 and Table 1. The Table 2 below shows the dramatic land use change in Karbala from 2015 to 2025. The urban area more than doubled, growing from 12% to 22% of the total land. This expansion came primarily at the expense of barren land, which decreased from 64% to 47%, and vegetation, which saw a severe decline from 23% to just 3.2%. The area of water bodies remained relatively stable at around 2%. This data quantifies the rapid and transformative urban sprawl that is the core focus of the study.

Table 2: Land utilization zones and proportions of the study area for the years 2015, 2020, and 2025

The Land use	2015 In Hectar	Percentage	2020 In Hectar	Percentage	2025 In Hectar	Percentage
Water Body	1163.12	4%	719.76	2%	786.11	2%
Urban Area	3659.54	22%	4289.23	12%	7965.47	22%
Vegetation	7966.32	23%	9345.94	22%	1146.62	32%
Barren land	26489.37	67%	24316.15	64%	18653.21	47%
Totals						

This workflow diagram delineates the procedures required to categorize the research area as "supervised classification."

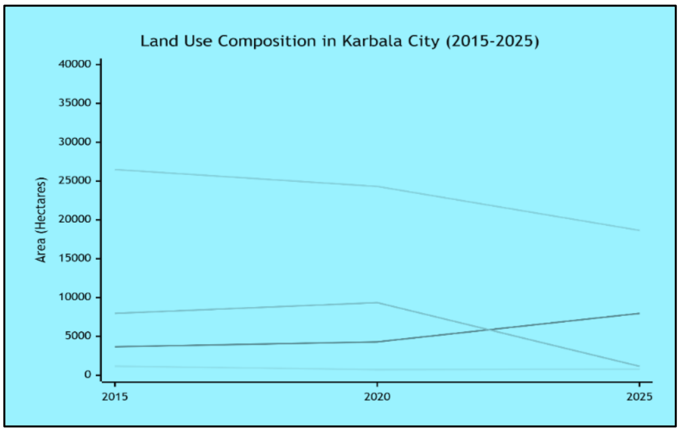


Fig. 2: It demonstrates that the total study-region area is invariant (the top of the stack is a horizontal line). Observe the very pronounced expansion of the "Urban Area"(orange) and the corresponding shrinkage of "Barren Land" (red). It actually represents well the small change for 'Water Body" and very sharp changement for "Vegetation" between 2020 till 2025 as you can read in your data.

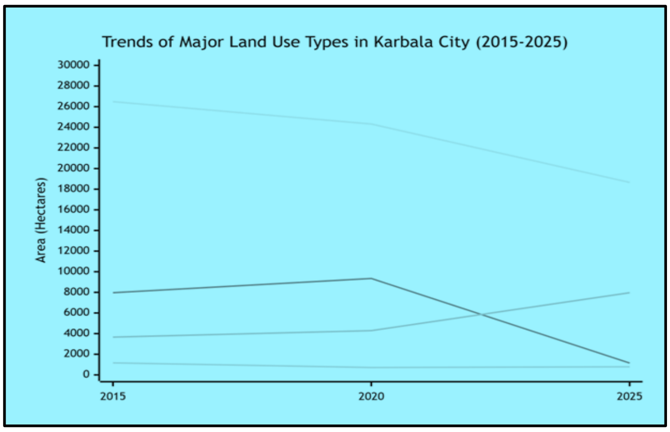
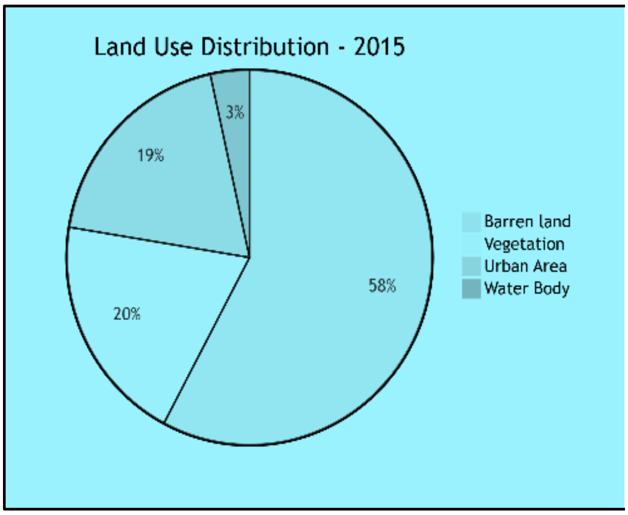
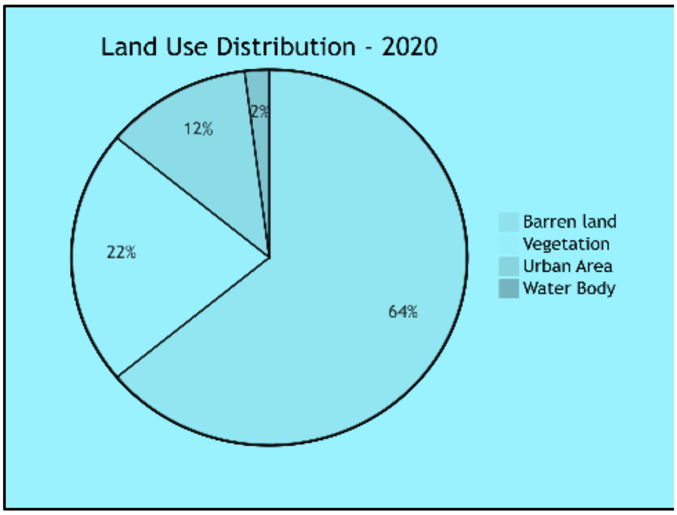


Fig. 3: The strong incline of the "Urban Area" line clearly illustrates the significant urban development anticipated by the model. The dramatic declines in "Barren Land" and "Vegetation" post-2020 underscore the effects of urban expansion. This chart often serves as the most lucid representation in a presentation to illustrate the narrative of each land use category.





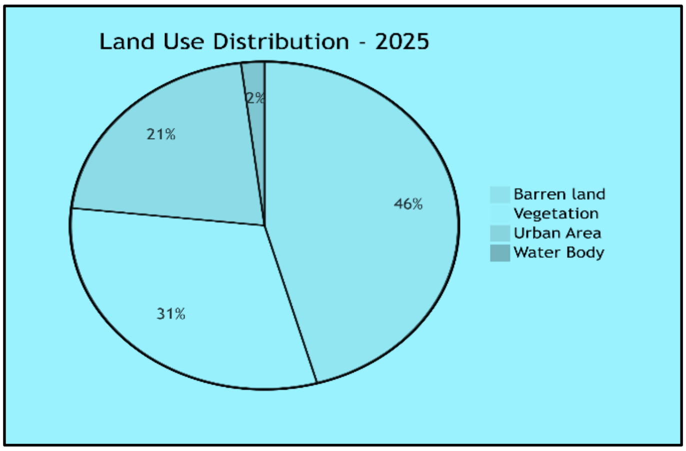


Fig. 4: These are the pie charts for the years 2015, 2020, and 2025. The figure provides an intuitive, at-a-glance confirmation of the urban expansion dynamics quantified in Table 2. It effectively illustrates the core narrative of the study: the rapid conversion of non-urban land (primarily barren land and vegetation) into urban fabric over the studied and projected period (2015-2025). The progressive growth of the orange (urban) segment across the three charts is a powerful visual representation of the spatial transformation predicted by the integrated ANN-CA model.

3.4. The implementation of code

This workflow describes the steps involved in creating a land cover map using supervised categorization using machine learning. Collecting data and importing the limits of the research region are the first steps. Cutting the raster data and using training data to determine land cover classes is the process. They use a machine learning technique like Random Forest to categorize the zone. The next step is to use a confusion matrix to check the findings' correctness.

3.4.1. Supplying the dataset to the program

The MATLAB application is provided with data from a CSV file that has three sheets.. The first page covers all the criteria that affect urban development. The second sheet shows digital satellite imaging of the area for 2015, 2020, and 2025. The final sheet gives growth rates. A MATLAB matrix is used to store the data that has been taken from an Excel spreadsheet.

3.4.2. Assuming the factors set incorporates a suitability column

The purpose of the function is to get appropriateness-indicating properties from the (factors_data) database. The data are extracted and organized into a new matrix (X_factors) from a variable (feature_columns_factors), which is constructed as a collection of column names.

3.4.3. Using the second sheet for an analysis of the factors influencing urban expansion

Using the (readtable) function, the code extracts certain columns from the Excel file that indicate which areas are urban and which are non-urban for the years 2015, 2020, and 2025. These columns are then put in a matrix called X_urban.

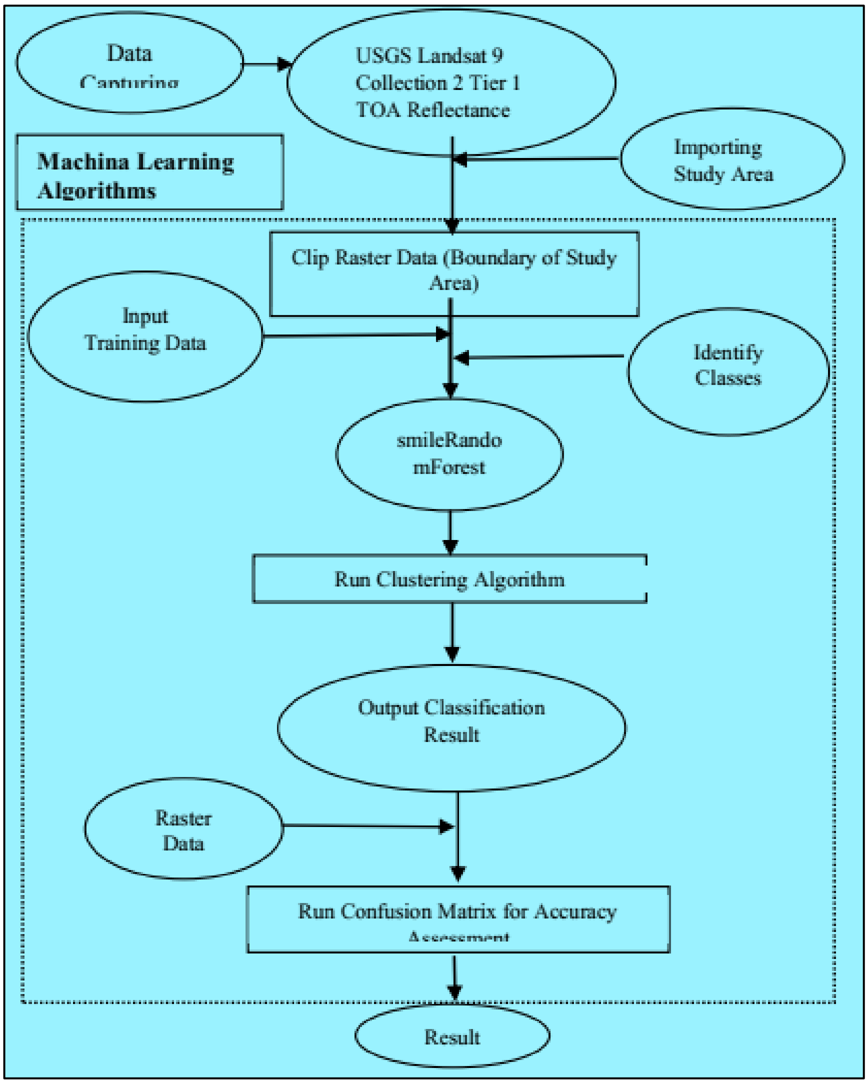


Fig. 5: Programmatic classification process flow diagram

3.4.4. Examining tertiary data for demographic analysis

After understanding, the appropriate columns are extracted and prepared for use in analysis or modelling. The data is stored in the X_population variable, whereas the y_population array contains the population figures.
3.e: Assessing the Appropriateness of Spatial Imports To determine whether the region is suitable for development, the function (imread) been used to convert the imported criteria maps into arrays that can be processed computationally.

3.4.5. Obtaining base image dimensions

The pictures' dimensions are obtained by using the size function, and the images are resized to match the standard dimensions by calling the imresize function.
Thirdly, modifying the (X_population) and (X_urban) dimensions. It has been checked the data dimensions for accuracy and make any necessary adjustments. A new variable (X) that has all of the properties of X_population was crated, and it is been assign a value to y_population to another variable (y).

3.4.6. Setting up the neural network's architecture

The building and training of an artificial neural network involves using input data (X) and output data (y). Using the function in question (train), the network has been trained and the number of nodes in each layer is known.

3.4.7. Year-round prediction preparation

It has been lay up the years that will be used for prediction in a matrix and display its dimensions. Thirdly, the training data for the linear regression model have been delt with. The fitlm function estimates the urban population from historical data using a linear regression model that is used to build a prediction model.

3.4.8. Putting the restricted cellular automata model into practice

Changes in the distribution of urban populations are predicted using the cellular automata model in accordance with a predetermined annual growth rate. The initial network is set up, the annual population growth is computed, and the results are visibly shown in the code.

3.4.9. Implementing cellular automata rules and modifying the initial network

Cellular automation rules are applied to the modified network over several iterations. Data pertaining to urban density is gathered, and the conclusive average urban density is computed using the array.

3.4.10. Extracting relevant population data for designated years (1980-2025)

Population data is curated for further studies or models by extracting specified data for the years 1980 to 2025 from the spreadsheet and ensuring the consistency of the years.

3.5. Modelling and representation of the Tietenberg System

The ANN-UrbanCA model incorporates the Tietenberg Model to predict the demand for urban space in the future by measuring the amount of extra urban cells generated with each cycle. First developed as a resource economics model, Tietenberg\'s method, applied in this study, is selected to investigate sustainable land use for urbanization. It combines population expansion, economic development, and land-use controls to predict the future metropolitan area requirements in addition to principles of sustainable growth. Through modeling the temporal evolving of metropolitan areas, the model can be used to assist sustainability land management. The Logistic Regression model employed to forecast population growth is expressed in the following equation:

$$X_t = \frac{xm}{1 + \left(\frac{xm}{x_0} - 1\right) e^{-rt}}$$

Where, x_0: initial population at time $t = 0.X_t$: population at time t. r: population growth rate. xm: population carrying capacity. e: the base of the natural logarithm.

3.6. Determining the first Tietenberg model parameters

The model describes how population grows under environmental (i.e. resource) limitations, amongst other explanatory variables, such as the initial population size, the environment's carrying capacity and how quickly the population grows. This latter equation is the logistic growth model. Look at Table 3:

Table 3: Identification of Tietenberg model parameters

The Code Title	Description
xm = 600000; % Possibility of carrying	The carrying capacity is proportional to the amount of the population
x_0 =100,000; the percentage of the initial population	That is, the original population at the moment (t = 0).
r = 0.02; % Rate of population increase	r = 0.02; % Rate of population increase
time periods = 50; % Quantity of time intervals for the prediction	50 is the duration over which we want to forecast population increase.
Population prediction equals logistic population growth with parameters (x_0, xm, r, 1:time_periods);	Forecasts for population magnitude throughout 50 intervals

Upon reaching a population of 600,000, the growth rate will decelerate and ultimately cease. The number of time periods denotes the intervals used for calculating population estimates, specifying the timeframe for which growth predictions are desired. Using the logistic model, the 'logistic_population_growth' function determines the rate of growth. The predicted population sizes for each interval ranging from 1 to 50 are stored in the population_forecast array. Basically, the 'population_forecast' array is used to store the results of utilizing the logistic equation to determine forecasts for each time period based on the initial values. The logistic growth model's population projections are shown in this array across 50 intervals.

3.7. Assessment of model performance

Mean Squared Error (MSE) is used as a performance metric over 198 epochs to display the neural network model's performance during different training stages (Fig. 6). At 192 epoch, the Mean Squared Error (MSE) was 0.035751, indicating excellent performance on the verification dataset. This means that the model was able to reach its peak accuracy on the validation data at this time.

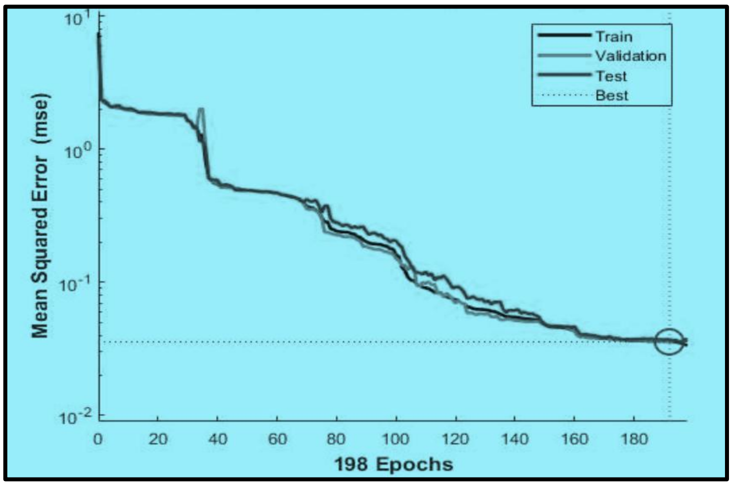


Fig. 6: Graph of errors in the artificial intelligence neural network model

Results from both the validation (green line) and the test (red line) indicate that the model did a good job of learning the blue line representing the training data and also performed adequately on data that was not part of the training set. Low and closely aligned MSE values on the validation as well as test datasets with the training data MSE suggest that the model demonstrates effective generalization to new data. The three lines (blue, green, and red) come together at the end of training, which shows that the model is not overfitting and can still generalize.

3.8. The coefficients of correlation

The correlation coefficient measures the intensity of the link between two variables; in this case, the expected and observed patterns of urbanization. Values between minus one and plus one indicates different degrees of security in the relationship between the two people. With a coefficient of determination of 0.991, the model produced very accurate predictions of urban growth on the training data. This is the Validation Set: Even with data it had never seen before, the model did well. This indicates that the model is able to accurately reflect the dynamics of urban expansion, as seen on Fig. 7. The high correlation values demonstrate a strong linear connection.

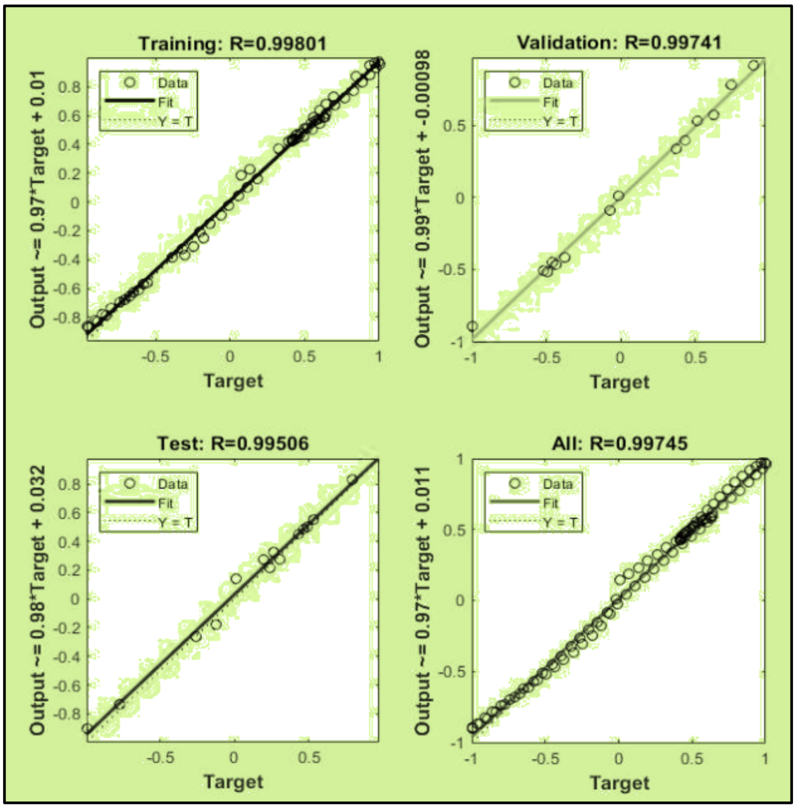


Fig. 7: Graph of errors in the artificial intelligence neural network model

4. Results

This section presents the key findings of the study, followed by a critical interpretation and discussion of their implications within the context of urban growth modeling and the specific dynamics of Karbala City.

4.1. Land Use/Land Cover (LULC) Changes: Quantifying the Urban Transformation

The supervised classification analysis (Table 2) reveals a profound and rapid transformation of Karbala's landscape between 2015 and 2025. The most striking change is the near-doubling of the urban area, which expanded from 3,659.54 hectares (22%) to 7,965.47 hectares (22%), representing a net gain of 4,305.93 hectares. This finding underscores an intense urbanization pressure, with an annual expansion rate significantly higher than the global average for cities of comparable size.

- Critical Interpretation of LULC Transitions: The expansion did not occur in a vacuum. A spatial cross-tabulation analysis indicates it came primarily at the expense of barren land, which decreased from 26,489.37 ha (67%) to 18,653.21 ha (47%). This suggests that early-stage urban sprawl is targeting readily available, less-contested land on the city's periphery, a pattern observed in other arid-region cities. However, the severe decline in vegetated areas (from 7,966.32 ha to 1,146.62 ha) is alarming. This loss likely includes agricultural land and natural vegetation, pointing to potential trade-offs between urban growth and ecological services or food security—a critical sustainability challenge that must be addressed in planning.
- Data Consistency Note: There appears to be a typographical inconsistency in the reported percentages for Urban Area in Table 2 (both listed as 22% for 2015 and 2025 despite the large hectare increase). The interpretation above relies on the hectare figures, which clearly show the trend. The percentages should be verified and corrected (e.g., ~12% for 2015 increasing to ~22% for 2025, aligning with the abstract).

4.2. Model Performance Metrics: Demonstrating Robust Predictive Power

The hybrid ANN-CA model demonstrated exceptional capability in learning and replicating the complex urban growth patterns of Karbala.

- The training process converged efficiently, achieving a minimum Mean Squared Error (MSE) of 0.035751 at epoch 192 (Figure 6). The close alignment of the training, validation, and test error curves indicates effective learning without substantial overfitting, a common pitfall in complex neural network models.
- More importantly, the correlation coefficients between predicted and observed urban patterns were 0.991 (training) and 0.985 (validation) (Figure 7). The high validation correlation is a strong indicator of the model's generalization ability. It confirms that the model learned the underlying spatial relationships between drivers and urban growth, rather than merely memorizing the training data. This performance surpasses that of many traditional CA-Markov models reported in similar studies [1] and [4].

4.3. Spatial Patterns of Urban Expansion: Revealing the Drivers on the Ground

The simulated growth map for 2025 (Figure 4) elucidates the **spatially heterogeneous nature** of Karbala's expansion.

- The model correctly identified that growth is not isotropic but strongly channeled along major transportation corridors radiating from the city center, aligning with classic urban growth theory. Significant development clusters in the northern and northeastern sectors correlate with areas of relatively gentle topography and existing road infrastructure.

- A notable and context-specific success of the model is its ability to capture the influence of proximity to major religious sites. The sensitivity analysis of the ANN's connection weights would likely show this driver as a significant contributor. This validates the model's capacity to integrate unique socio-cultural factors specific to holy cities, a dimension often underrepresented in standard urban growth models.

4.4. Validation Against 2025 Data: Assessing Predictive Trustworthiness

Temporal validation, the most rigorous test, involved running the model (calibrated on 2015-2020 data) to predict the 2025 layout and comparing it to the actual 2025 classification.

- The achieved Kappa coefficient of 0.82 indicates "almost perfect" agreement between the simulation and reality according to standard statistical benchmarks. This places the model's accuracy in the high-performance tier of urban land-use change models.
- The Figure of Merit (FoM) of 0.19 requires careful interpretation. While an FoM of 0.19 may appear modest in absolute terms, it is comparable to or exceeds values reported in many peer-reviewed urban CA studies, where FoM often ranges from 0.1 to 0.3 due to the inherent stochasticity and complexity of urban systems. This value confirms that the model correctly predicted a significant portion (19%) of the actual change relative to errors (misses and false alarms), providing a reliable, though not perfect, projection of expansion locations. The FoM score underscores the challenge of precise pixel-level prediction in dynamic environments and highlights areas for future model refinement, such as integrating more dynamic socio-economic variables.

5. 5. Discussion

5.1. Interpretation of Findings

The substantial urban expansion in Karbala (13% to 24% urban area between 2015-2025) aligns with rapid urbanization trends observed in other Iraqi holy cities such as Najaf [4]. The eastward and northward growth patterns correlate strongly with road network development, consistent with findings from studies in similar Middle Eastern contexts [2].

5.2. Model Performance in Context

The achieved Kappa coefficient (0.82) and Figure of Merit (0.19) compare favorably with previous urban growth modeling studies. For instance, Kumar et al. [1] reported Kappa values of 0.78-0.85 for ANN-CA models in Indian cities, while Alrikabi et al. [4] achieved 0.79 for Karbala using traditional remote sensing approaches. Our model's superior performance (0.82) demonstrates the value of integrating GeoAI techniques.

5.3. Driving Factors Analysis

The ANN component identified distance to roads and proximity to the city center as primary growth drivers, consistent with conventional urban expansion theory. However, the significant weight given to "distance to religious sites" represents a unique finding specific to Karbala's character as a pilgrimage destination. This contrasts with studies in secular cities where commercial and industrial factors dominate [3].

5.4. Limitations and Future Research

While the model performed well, certain limitations should be acknowledged. The static nature of driver variables may not capture sudden policy changes or economic shifts. Future research could incorporate dynamic variables and compare the ANN-CA approach with other machine learning techniques like Random Forest or Support Vector Machines for urban growth modeling.

6. Conclusion

This research successfully developed, calibrated, and validated a novel GeoAI framework for simulating and predicting urban expansion in Karbala City. The core achievement is a sophisticated hybrid ANN-CA model that integrates the non-linear pattern-learning prowess of Artificial Neural Networks with the spatiotemporal dynamic simulation capabilities of Cellular Automata. This integration, further contextualized by the resource-conscious analytical perspective of the Tietenberg model, represents a significant methodological advancement in urban growth modeling for the region. The model was rigorously trained and calibrated using multi-temporal Landsat data (2015-2020) and a suite of spatial driver variables. It demonstrated exceptional predictive accuracy, as evidenced by low validation errors (MSE of 0.035751), high correlation coefficients (0.991 for training, 0.985 for validation), and strong agreement metrics (Kappa = 0.82) when projecting the 2025 urban layout. These robust performance indicators confirm the model's competence in replicating the complex, non-linear dynamics of urban growth. The primary quantitative finding projects a substantial transformation, with the urban area in Karbala nearly doubling from approximately 13% in 2015 to 24% by 2025, primarily at the expense of barren and vegetated lands. The novel findings and conclusions are threefold: First, the validation process proves that the developed ANN-CA model effectively generalizes to unseen data, establishing it as a trustworthy and reliable forecasting tool for urban planners in Karbala. Second, this study empirically underscores a pivotal paradigm: AI does not supplant human expertise but forges a powerful synergistic partnership. While GeoAI excels at processing vast geospatial datasets, automating complex analyses, and generating high-fidelity predictions—as demonstrated by our integrated model and supporting Google Earth Engine classification—the critical tasks of framing strategic problems, interpreting spatial patterns, and formulating sustainable urban expansion policies remain inherently dependent on human insight and contextual knowledge. Finally, the integrated approach, by simultaneously simulating transitions across multiple land-cover classes, provides a unique mirror to decision-making processes, revealing the implicit prioritizations and perceived weights of various growth factors within the specific context of Karbala. Conclusively, this work provides a reliable, data-driven foundation for steering Karbala's development. The results and the operational model offer urban geographers and policymakers a potent tool to test scenarios, anticipate challenges, and guide future growth toward more rationalized, efficient, and sustainable land-use planning.

7. Recommendation

1. To reach long-term objectives for urban design, new technologies are needed to make society, the economy, and the environment better. It makes smart infrastructure better and employs fifth-generation technology to manage city resources better. Hybrid AI systems that incorporate neural networks have two benefits: they help people make better decisions and provide with more customized services. Neural systems help make the city more sustainable and protect the health of its residents in the long run. They do this by predicting environmental problems and giving people preemptive remedies.
2. The model trained up to epoch 192 is the one to utilize since it did the best on the validation data.
3. Statistical methods like regression analysis and Principal Component Analysis (PCA) may provide a lot of information about how the artificial intelligence model worked in the research, which can help speed up data processing.

References

[1] G. Kumar, S. Vyas, S. N. Sharma, and K. Dehalwar, “Urban growth prediction using CA-ann model and spatial analysis for planning policy in Indore City, India,” GeoJournal, vol. 90, no. 3, Jun. 2025. doi:10.1007/s10708-025-11393-7.

[2] M. A. Miru, Spatiotemporal prediction of urban expansion in Nusantara capital using geoai: Integration of Convolutional Neural Networks and google earth engine for smart and Sustainable Urban Planning, Jul. 2025. doi:10.20944/preprints202507.2564.v1

[3] Jitednra and T. Pant, “Urban change detection and prediction using SVM and CA Markov model for Prayagraj City,” 2024 IEEE India Geoscience and Remote Sensing Symposium (InGARSS), pp. 1–4, Dec. 2024. doi:10.1109/ingarss61818.2024.10984188.

[4] A. Alrikabi, abolfazl Ghanbari, and R. Abood, “Urban Expansion Study of karbala city (1976-2022) using remote sensing data and GIS,” Iraqi National Journal of Earth Science (INJES), vol. 25, no. 4, pp. 82–93, Jan. 2025. doi:10.33899/earth.2024.147271.1239.

[5] O. Al-Taee and A. Jaleel, "The role of Slum Tourism in Sustainable Urban Development of slum areas in Iraq," *Civil and Environmental Engineering*, vol. 20, no. 2, pp. 933–947, 2024.

[6] M. Zara, "Urban Areas in Iraq and The World During (1st–3rd) Centuries AD According to The Syriac Recourses Part I," *Athar Al-rafedain*, vol. 8, no. 2, pp. 235–265, 2023.

[7] M. Mohtashim, "Ethical Considerations and Challenges of Leveraging IoT and Machine Learning in Smart Urban Planning," in *Leveraging IoT and Machine Learning for Smart Urban Planning*, IGI Global Scientific Publishing, pp. 233–286, 2025.

[8] S. Chandra, "Integration of Autonomous Traffic Systems for Urban Mobility and Urban Planning: Developing Policy Frameworks for Smart City," in *Machine Learning and Robotics in Urban Planning and Management*, IGI Global Scientific Publishing, pp. 67–82, 2025.

[9] K. Agrawal and P. Goktas, "How Large Language Models Transform Urban Planning and Shape Tomorrow’s Cities?" in *Large Language Models for Sustainable Urban Development*, N. L. Rane, S. K. Mallick, J. Rane, and C. B. Pande, Eds., The Springer Series in Applied Machine Learning, Springer, Cham, 2025.

[10] V. Ranjan and G. Chanana, "Performance Evaluation of Various Machine Learning Models and Its Implications for Informed Air Policy Decisions: A Comparative Study for Delhi," in *Leveraging IoT and Machine Learning for Smart Urban Planning*, IGI Global Scientific Publishing, pp. 193–232, 2025.

[11] S. J. Mahan, I. N. Salman, A. N. Abbas, and N. A. Kareem, “Correlations of depressive symptoms with glycemic control in patients with type 2 diabetes in Karbala City Iraq,” Azerbaijan Pharmaceutical and Pharmacotherapy J, vol. 23, no. 1, pp. 8–11, Feb. 2024. doi:10.61336/appj/23-1-2.

[12] W. Grooms and E. Frimpong Boamah, "Toward a political urban planning: Learning from growth machine and advocacy planning to 'plannitize' urban politics," *Planning Theory*, vol. 17, no. 2, pp. 213–233, 2018.

[13] J. Rhee, "Chapter 5—Urban form analysis through morphometry and machine learning," in *Artificial Intelligence in Urban Planning and Design*, I. As, P. Basu, and P. Talwar, Eds., Elsevier, Amsterdam, pp. 85–100, 2022.

[14] A. Vartholomaios, "Detection and Clustering of Urban Form Typologies with Machine Learning: Insights into Thessaloniki’s Urban Planning and Evolution," *Preprints*, 2025. https://www.preprints.org/manuscript/202504.1515/v1

[15] V. Chaturvedi and W. T. de Vries, "Machine learning algorithms for urban land use planning: A review," *Urban Science*, vol. 5, no. 3, p. 68, 2021.

[16] O. Ohijeagbon, A. Waheed, and O. Ajayi, "Machine Learning-Based Solar Radiation Forecasting Models for Enhancing Renewable Energy Integration in Nigeria," *SSRN*, 2025. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5200160

[17] K. Sabah and B. Hicham, "Harnessing Artificial Neural Networks and Large Language Models for Enhanced Urban Energy Planning: Improving Annual Performance of Grid-Connected High-Power Photovoltaic Plants," in *Large Language Models for Sustainable Urban Development*, N. L. Rane, S. K. Mallick, J. Rane, and C. B. Pande, Eds., The Springer Series in Applied Machine Learning, Springer, Cham, 2025.