

Urban Growth Modeling Considering Educational Institutions: An Investigation Into the Performance of Three Anfis Methods

Yousef Ghobadiha

Islamic Azad University South Tehran Branch

Hamid Motieyan (✉ H.motieyan@nit.ac.ir)

Babol Noshirvani University of Technology <https://orcid.org/0000-0002-3733-7147>

Research Article

Keywords: Adaptive Network-based Fuzzy Inference System, Fuzzy c-means, Markov Chain, Subtractive clustering, urban expansion, urban sprawl

Posted Date: September 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-923943/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Urban growth modeling considering educational institutions: An investigation into the performance of three ANFIS methods

Yousef Ghobadiha^a, Hamid Motieyan^{b,*}

a. Department of Civil Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran

b. Department of Civil Engineering, Babol Noshirvani University of Technology, Babol, Iran

* Corresponding author: Hamid Motieyan

Full postal address: Department of Civil Engineering, Babol Noshirvani University of Technology, Babol, Iran

Email: H.motieyan@nit.ac.ir

Tel: +98 9122405512

ORCID ID: <https://orcid.org/0000-0002-3733-7147>

Abstract

Due to increasing urbanization, the rapid expansion of urban spaces has become a major environmental concern over the last few decades. Therefore, modeling the urban expansion as a complex system has been scrutinized in recent years; however, determining the rules that lead to the expansion of urban areas has always been a challenging factor in this field, especially for disaggregated models like cellular automata (CA). To overcome this issue, in this research, an Adaptive Network-based Fuzzy Inference System (ANFIS) is proposed to enhance the simulation of urban growth through the automatic production of transition rules. The ANFIS can be associated with several inputs division methods, such as ANFIS accompanied by grid partitioning (ANFIS-GP), subtractive clustering (ANFIS-SC), and fuzzy c-means clustering (ANFIS-FCM). Hence, twenty-two ANFIS models based on Landsat images for the time interval from 2000 to 2010 and using different division methods were trained to investigate their effect on the efficiency of ANFIS in urban growth modeling. To examine the efficiency, the Cellular Automata-based Markov Chain (CA-MC) as a popular method was developed, and the simulation accuracy of CA-MC and the most accurate ANFIS models were obtained through comparison with observed data. The most accurate ANFIS-SC model had a Kappa of 0.76 and an overall accuracy of 93.41% for the 2019 simulated map. The results from this study reveal that the ANFIS model is effective at simulating urban expansion and the ANFIS-SC is superior to CA-MC, ANFIS-GP, and ANFIS-FCM models in urban expansion modeling.

Keywords Adaptive Network-based Fuzzy Inference System, Fuzzy c-means, Markov Chain, Subtractive clustering, urban expansion, urban sprawl

1 Introduction

The proportion of the world's population that resides in urban areas has grown from approximately 5% to over 50% in recent centuries (Ezeh et al., 2017), and it is expected that this proportion will rise to 70%, or even more, by 2050 (Broere, 2016; Hegazy et al., 2017; Kabisch & Haase, 2011). This issue causes a more serious problem to emerge, that is, the accelerating rate of urban expansion in most parts of the world, insofar as it is predicted that the world's urban areas will have tripled between 2000 and 2030 (Khamis et al., 2018; McDonnell & Hahs, 2015; Southon et al., 2017). The rapid urban expansion has led to a variety of complex problems, such as increasing informal settlements, air pollution, deforestation, soil erosion, and climate change (Kamusoko & Gamba, 2015; Wang et al., 2020; Xu et al., 2019). Additionally, the rapid growth of the cities has strained their capacity to provide services, such as education, health care, and transportation (Ewing et al., 2002). Moreover, the level of urbanization in Iran increased from 33.7% in 1960 to 74.9% in 2018 (World Bank, 2018), and this high rate of urbanization has led to the fast expansion of built-up areas in recent years (Bihamta et al., 2015; Taravat et al., 2017). Such issues can be prevented from occurring by modeling and studying the future urban expansion (Xu et al., 2019). Models provide a better understanding of the mechanisms of urban expansion and enable urban planners and policymakers to manage the continued growth of cities (Li et al., 2017; Samardžić-Petrović et al., 2017).

Over the last three decades, many models have been developed to simulate urban growth processes and land use/land cover (LULC) changes (Karimi et al., 2019; van Vliet et al., 2016). Three typical classes of urban growth models have been introduced so far: the land use transport interaction (LUTI), the cellular automata (CA), and the agent-based models (ABMs) (Li & Gong, 2016). Among these models, the CA is a popular (Liu et al., 2017; Musa et al., 2017) and strong technique for simulating the pattern of built-up areas (Aburas et al., 2016; Mustafa et al., 2018). Nevertheless, determining the transition rules, which leads to the calibration of the CA model, has always been a key challenge in this model because of the complex relationship between urban expansion and its multiple driving factors (Berberoğlu et al., 2016; Blečić et al., 2015; Li et al., 2017). To handle this challenge, Artificial Neural Networks (ANN) have been extensively utilized in urban growth/LULC change simulation (Lu et al., 2019) due to automatic determination of transition rules and less computation time (Mondal et al., 2020).

Li and Yeh (2001), Li and Yeh (2002), and Li et al. (2013) automatically determined the transition rules of CA using neural network. Furthermore, Xu et al. (2019) integrated ANN with the Cellular Automata-based Markov Chain (CA-MC) model. In these studies, due to associating the ANN with CA, the capability of the CA model has improved in handling nonlinear transition rules (Yang et al., 2008). Nevertheless, neural networks have some disadvantages, such as difficult calibration, the tendency to overfitting (Kamusoko & Gamba, 2015), and incompatibility with the uncertainty of transition rules (Azari et al., 2016). Azari et al. (2016) utilized a fuzzy set theory to tackle the uncertainty of transition rules. However, fuzzy models require a justifiable approach for calibrating (finding locations of the Membership Functions (MFs) for the model inputs) their parameters (Al-Ahmadi et al., 2009). Several works have demonstrated that the weaknesses of fuzzy and ANN can be overcome by using an Adaptive Network-based Fuzzy Inference System (ANFIS). An ANFIS is a hybrid intelligent system that carries the benefits and learning potential of neural networks and fuzzy logic (Parvinnezhad et al., 2019) and eliminates many limitations of these two algorithms (Ahmadlou et al., 2019). Mohammady et al. (2013) utilized ANFIS for urban growth modeling from 2000 to 2006 in Sanandaj, Iran. They employed Figure of Merit (FOM) and Percent Correct Match (PCM) for model evaluation, which obtained 64% and 94%, respectively. Mohammady and Delavar (2016) and Mohammady (2016) have demonstrated that the ANFIS model has more accurate results in the simulation of urban growth compared with the ANN and Logistic Regression (LR).

The ANFIS model employs and trains a Fuzzy Inference System (FIS) with distinct division methods, such as Grid Partitioning (GP), Subtractive Clustering (SC), and Fuzzy C-means (FCM) to divide the input data space. It has been reported in some scientific studies that clustering type can play an important role in ANFIS model performance, such as Adedeji et al. (2020) and Moradi et al. (2019), whereas there is a paucity of high-quality research investigating this subject in urban development modeling. It is explicit that calibrating various kinds of ANFIS with urban parameters, comparing the accuracy of trained models, and finally, employing the most well-trained ANFIS model, lead to more precise transition rules and, as a result, more accurate forecasting of LULC changes. This study, therefore, aims to implement ANFIS accompanied by GP (ANFIS-GP), SC (ANFIS-SC), and FCM (ANFIS-FCM) models to assess the effect of clustering methods and their parameters on the accuracy of ANFIS in urban growth modeling. Additionally, it aims to simulate the future built-up areas and warn about the possible harmful effects of spontaneous urban growth on the environment. To fulfill these aims, we executed multiple scenarios for the ANFIS training to achieve the most

accurately calibrated model. Additionally, we developed the CA-MC (Rimal et al., 2018; Rimal et al., 2017; R. Wang et al., 2018), a popular (Wang et al., 2020) existing model that integrates the CA and Markov chain analysis. Afterward, we compared the simulation results of CA-MC with the most acceptable ANFIS model for the 2019 simulated map. We finally simulated the upcoming urban expansion of Qazvin city for the year 2029, utilizing the ANFIS model, for the use of urban planners and policymakers.

2 Concepts and materials

2.1 Urban expansion

The rate of urbanization has significantly increased over the last five hundred years and by 30% in 2010 (Gross & Ouyang, 2021). The increase in the level of urbanization, especially in less developed countries (Angel et al., 2016), has been recognized as one of the main reasons of the rapid growth of urban areas. As we mentioned in the introduction, the rapid urban expansion with no efficient planning has posed some serious risks to the environment and residents. Therefore, recognizing the nature and types of urban expansion can be useful to handle its negative consequences and achieve sustainable urban development. Generally, there are three spatial patterns of urban expansion, namely the edge-expansion, infilling, and outlying (Xian et al., 2019; Zhang et al., 2020). The urban edge-expansion denotes the development along or approximately in parallel to the board strip, the infilling has been defined as the development of vacant parcels located inside of built-up areas and eventually the outlying refers to development of areas placed outside of existing urban areas (Jiao et al., 2018). There is another distinct expansion pattern called sprawl, which in previous studies is commonly known as scattered development (Artmann et al., 2019). As a complex pattern of urban development, sprawl generates large and low-density suburban areas associating with some hazards to human health, such as air pollution (Frumkin, 2002).

Regardless of the type of urban expansion, when a city is developing rapidly, the requirement of careful urban planning and the design of development policies to provide facilities for citizens become essential. Precise urban planning can greatly reduce the disadvantages of some of the mentioned urban expansions. Therefore, this study does not seek to distinguish diverse types of urban development; however, it aims to determine any urban expansion that has occurred through LULC.

2.2 Study area

Qazvin Province, with a top urbanization rate of 73.1%, is located in the northern half of Iran, as one of the 31 provinces of the country. Qazvin city, as the central city of this province, lies at an altitude of 1278 m (4193 ft.) above sea level. The study area of this research includes the Qazvin city and its suburbs, which locate between 36.2266° and 36.3604° N, 49.9614° and 50.0555° E (128.63 km²). Fig.1 represents the Qazvin city and its position in the county and country.

The most significant reason for choosing this city as our study area was that the city is suffering from fast growth in both population and urban areas. According to the census of the Statistical Center of Iran, the population of Qazvin city was 402,748 in 2016, whereas in 2006 this number had been 349,821 (Statistical Center of Iran, 2019). This increase in population has led to the rapid outward city growth. The low population-density zones and some formal and informal settlements have appeared surrounding the city in the last decades. The emergence of these areas causes some serious damage to the environment and natural resources. Consequently, an investigation of the rapid growth of this city appears to be vital for urban planners.

2.3 ANFIS model

2.3.1 Layers and architecture of ANFIS

ANFIS was developed by Jang (1993) as an integrated hybrid neuro-fuzzy model that employs the Takagi–Sugeno fuzzy inference system (Abraham, 2005). To put it more simply, ANFIS carries the advantages of both ANN and fuzzy reasoning at the same time (Zhou et al., 2017) by training a FIS model. As it is displayed in Fig.2, this model is composed of five layers.

The parameters of the first and fourth layers are recognized as premise and consequence, respectively. Unlike the parameters of the other layers, they are adjusted throughout the training process. The first layer has come to be known as fuzzification layer. Nodes in this layer take the numeric inputs and express the membership grade of the inputs in a fuzzy set as the output by fuzzy Linguistic terms. For example, for the input x and fuzzy set A_i , the output of the first layer is defined as follows.

$$O_i^1 = \mu_{A_i}(x), \quad i = 1: \text{number of MFs} \quad (1)$$

Where, $\mu_{A_i}(x)$ represents the membership grade of the input x in fuzzy set A_i . The second layer is known as the rule layer. The general structure of a fuzzy If-Then rule in the ANFIS network is as follows.

$$\text{Rule } i: \text{ IF } x \text{ is } A_i \text{ AND } y \text{ is } B_i \text{ THEN } F_i = p_i x + q_i y + r_i, \quad i = 1: \text{ number of rules} \quad (2)$$

Where, x and y are ANFIS inputs, A and B are fuzzy sets, F_i is an output function, and p_i , q_i , and r_i are consequence parameters belonging to the output function. Any Triangular norm (T-norm) operator that performs fuzzy AND belonging to the IF part of the fuzzy rule can be deployed in the second layer as the fuzzy implication function. Hence, this layer is utilized the minimum T-norm and multiplies the membership grades coming from the first layer in order to obtain the weight of each rule. Eq.3 represents the functionality of this layer for two x and y inputs.

$$O_1^2 = w_i = \mu_{A_i}(x) \cdot \mu_{B_i}(y), \quad i = 1: \text{ number of MFs} \quad (3)$$

Where, w_i is the weight of i th rule and $\mu_{A_i}(x)$ and $\mu_{B_i}(y)$ represent the membership grades of the x and y . The third layer computes the normalized firing strength of i th rule (\bar{w}_i) using Eq.4.

$$O_1^3 = \bar{w}_i = \frac{w_i}{w_1 + \dots + w_j}, \quad i = \text{ number of rule and } j = \text{ total number of rules} \quad (4)$$

The fourth layer, which is denoted as the defuzzification layer, multiplies the normalized weights from the previous layer by Then part of fuzzy If-Then rule (output function). Eq.5 exhibits the operation of this layer.

$$O_1^4 = \bar{w}_i F_i = \bar{w}_i (p_i x + q_i y + r_i), \quad i = \text{ number of rule} \quad (5)$$

The only node in the fifth layer finally adds the outputs of the previous layer and generates the conclusive output of ANFIS (Karaboga & Kaya, 2019).

Generally, there are two learning algorithms for ANFIS training, namely back-propagation and hybrid in order to adjust the premise and consequence parameters in every epoch. The hybrid method consists of two steps, forward and backward pass. The consequent parameters are specified by least-squares method over the forward pass and the premise parameters are adjusted by gradient descent over the backward pass.

2.3.2 Division methods in ANFIS

ANFIS can be categorized into four types, namely ANFIS-GP, ANFIS-SC, ANFIS-FCM, and ANFIS accompanied by a context-based fuzzy c-means (ANFIS-CFCM) according to the division method of inputs (Yeom & Kwak, 2018). The division methods, such as GP and SC cluster the input data to fuzzy clusters. Afterward, an MF is allocated to each cluster and a FIS is generated (Adeleke et al., 2020).

The GP method divides the entire input data space into rectangular partitions and every grid is used to produce fuzzy rules. The accuracy of clustering in this method is greatly relying on the size of inputs and grids (Benmouiza & Cheknane, 2019). The SC method, as the first step, calculates the potential of being the center of cluster based on the density of the data surrounding it for each data and selects the data with the greatest potential as the center of the first cluster. Afterward, all data that falls within the center's range of influence of the first cluster is deleted. This process iterates for the remaining data so that all data are clustered. The FCM clustering assigns each data to clusters using a membership grade. As the first step of this algorithm, the number of clusters is determined and the initial locations of clusters centroids are discovered by an initial guess. Afterward, the membership grades of each data for being in each cluster are calculated. The FCM algorithm adjusts the centroid location and membership grades in an iterative process.

2.4 Markov chain analysis

The Markov chain analysis (MCA) is a foreseeing approach that event occurring probabilities are varying depending on the values at the previous time interval (Fan et al., 2008). In fact, the future values of a particular variable depend only on its current (not on the past) value. Generally, according to the Markov chain theory, if the sequence of the random variables of x_n takes the distinct values of a_1, a_2, \dots, a_n , then the Markov chain can be represented by Eq.6 (Fan et al., 2008).

$$P(x_n = a_{i_n} | x_{n-1} = a_{i_{n-1}}) \quad (6)$$

With this assumption that the values of the variables were changed from time t to time $t-1$, the transition probability in time t can be computed by the following relation (Balzter, 2000).

$$P(t) = P(t - 1). p \quad (7)$$

The term p in Eq.7, which is named transition probability matrix, represents the probabilities of transitioning. Assuming that the states are $1, \dots, m$, the p can be displayed as the following matrix.

$$p = \begin{bmatrix} p_{11} & \dots & p_{1m} \\ \vdots & \dots & \vdots \\ p_{m1} & \dots & p_{mm} \end{bmatrix}$$

Where, the p_{1m} represents the probability of transition from state 1 to state m .

2.5 Driving factors

Up to this time, various driving factors of urban expansion, such as distance to primary and/or secondary roads, distance to previous urban, slope, elevation, population, and access to jobs have been considered in diverse investigations (Mustafa et al., 2017; Pijanowski et al., 2014; Radwan et al., 2019; Shafizadeh-Moghadam, Asghari, Taleai, et al., 2017; Shafizadeh-Moghadam, Asghari, Tayyebi, et al., 2017). In determining driving variables, it is highly important to pay attention to the particular characteristics of the local area. Qazvin, as our study area, possesses several colleges and universities at the present. Most of these institutions have been built outside of the city limits, which makes the value of nearby land grow and the chance of LULC change increase. Additionally, the past researches have not focused on the distance from educational institutions as a driving force in urban growth or sprawl modeling (Aburas et al., 2016; Hosseini & Hajilou, 2019; Santé et al., 2010), except for a very limited number of studies such as Triantakou et al. (2014). Hence, distance to the universities by five other spatial variables, namely slope, aspect, distance to the urban areas, distance to the urban transportation network, and distance to the urban squares, were selected as the driving factors of this study.

3 Methodology

Satellite imageries are qualified primary data sources for studying LULC changes. Hence, for this research, we adopted the Landsat images (from 2000, 2010, and 2019) due to their accessibility and appropriate spatial resolution. Fig.3 provides a general overview of the principal performed steps of this study. As it can be observed from this figure, image pre-processing was performed on the data because the received images were accompanied by errors. Afterward, images were classified and the six driving factors of urban growth as input layers of the ANFIS and CA-MC models were generated. This study developed and implemented a two-step model that comprised two phases of calibration for 2000 to 2010 and prediction for 2010 to 2019. The calibration phase consisted of three parts. In the first and second parts, ANFIS-GP, ANFIS-SC, and ANFIS-FCM models were calibrated and compared in terms of performance and accuracy of training in pairs using Mean Squared Error (MSE) and Root Mean Square Error (RMSE), and the most appropriate one was selected. In the third part, a CA-MC model with WLC for the generation of transition suitability images and AHP for factor weighting was developed. In the prediction phase, the built-up areas of 2019 were predicted using the CA-MC and most well-trained ANFIS models. Eventually, the predicted map of CA-MC was evaluated using overall accuracy, Kstandard (Kappa standard), Kno

(Kappa for no information), Klocation (Kappa for location), and KlocationStra (Kappa for stratum-level location), and the predicted map of ANFIS was assessed utilizing overall accuracy and Kappa.

3.1 Data sources and imageries preparation

Landsat images have been largely used for urban planning due to aspects of appropriate spatial resolution and shorter revisit time (S. Wang et al., 2018). Hence, for this research, we adopted the Landsat 7 ETM+ (Enhanced Thematic Mapper) images on May 22, 2000, and June 3, 2010, and Landsat 8 OLI/TIRS (Operational Land Imager/Thermal Infrared Sensor) image on June 20, 2019. The images were downloaded from the United States Geological Survey (USGS) database (<https://earthexplorer.usgs.gov/>) as the primary data sources of this study. As each sensor had a distinct spatial resolution, all images were resampled to 30m resolution, and then they projected to Universal Transverse Mercator (UTM) 39 N coordinate system.

Before utilizing the satellite images, pre-processing steps must be performed on the imageries. Pre-processing refers to those processes that apply to the raw satellite images to reduce or eliminate the errors of sensor and platform and includes radiometric and geometric corrections (Asubonteng et al., 2018). The obtained images have georeferenced and orthorectified with high geometric accuracy (fewer than 6.8m RMSE), and no geometric corrections were required to add to the imageries accordingly.

Atmospheric correction, as one of the most important steps in pre-processing, eliminates the influences of Atmospheric phenomena, such as absorption and scattering (Jha et al., 2021). Hence, we applied an absolute atmospheric correction to all imageries by using the fast line-of-sight atmospheric analysis of hypercubes (FLAASH) algorithm in the ENVI platform.

The Scan Line Corrector (SLC) of Landsat 7 ETM+ has failed since May 31, 2003, which made a loss of about 22% of data in the SLC-off images (Masoumi et al., 2017). Therefore, in order to fill the gap areas of the image of 2010, we used the Local Linear Histogram Matching (LLHM) approach by adding an add-on to the ENVI software and adopting another Landsat 7 image (image on June 19, 2010) from the same area of the initial image.

We selected the suitable three-band combination by applying the Optimum Index Factor (OIF) technique for improvement in classification accuracy (Patel & Kaushal, 2011). Table 1 displays the implementation results of OIF method in the ILWIS software. As a preferable combination, the one with the highest OIF represents the most details in the images.

Table 1 the OIF values belonging to the best band combinations of satellite images

year	band combination	OIF
2000	7-4-3	1778.67
2000	7-4-1	1712.85
2000	4-3-1	1679.62
2010	5-4-3	1298.48
2010	7-5-4	1294.82
2010	7-4-3	1272.70
2019	7-5-4	1162.31
2019	7-5-1	1067.71
2019	7-6-5	1061.95

Afterward, we carried out a supervised method, namely maximum likelihood, to classify the imageries in the ArcMap software. Urban, agricultural, barren, and transport were selected as the land use classes based on study objectives. For assessing the classification performance, we collected the land use of random sample points by GPS and generated a confusion matrix. An overall accuracy of 80% and a kappa coefficient of 0.72 were received for the classification. According to Pijanowski et al. (2005), the accuracy of classification was acceptable.

3.2 Generation and preparation of Input layers

In this study, six driving factors were selected and, as a result, six raster layers related to them were created in the ArcMap environment. Two layers of slope and aspect were achieved using Aster Digital Elevation Model (DEM) v.3. Four distance raster layers were generated by Euclidean distance ArcMap tool. Furthermore, the layer of areas with unchangeable LULC, such as military bases and parks, were generated by the LULC map. The value of these pixels remained 0 (unchangeable land use) in all study stages. All accomplished raster layers for 2019 with 30m pixel resolution and UTM 39N coordinate system are presented in Fig.4.

Since pixels in every layer had a distinct range of quantities and units, we normalized all of them in the range of 0 to 1 by the Eq.8 in the MATLAB (version 2016b) software.

$$X(i, j) = ((X(i, j) - \min(X(i, j))) / (\max(X(i, j)) - \min(X(i, j)))) \quad (8)$$

Where, $X(i, j)$ is the normalized value of each pixel, $\min(X(i, j))$ is the minimum value of all pixels in a specific layer, and $\max(X(i, j))$ is the maximum value of them.

3.3 ANFIS models implementation

As the common ANFIS methods, we implemented ANFIS-GP, ANFIS-SC, and ANFIS-FCM using GenFIS1 (generate initial FIS), GenFIS2, and GenFIS3 MATLAB functions respectively. All models were accomplished with various input parameters and Gaussian MF. Fig.5, which achieved by ANFIS editor Graphical User Interface (GUI), represents our ANFIS-SC structure with the six driving factors of this study as inputs of the model and one output representing the development probability of each pixel.

The number of generated rules in ANFIS-GP is equal to the number of MFs to the power of the number of input layers, whereas the number of generated rules in ANFIS-SC and ANFIS-FCM is equal to the number of MFs. Hence, if we had only five MFs, GenFIS1 would have generated 15625 fuzzy if-then rules with our six input layers. It takes weeks or months for a common computer to run an ANFIS model with this amount of rules. Therefore, we calibrated both ANFIS-GP and ANFIS-SC in a separate part decreasing the number of MFs (for ANFIS-GP), input layers, and training dataset to 3, 4, and 50000 respectively. Consequently, the ANFIS-GP model was implemented and the ability to compare the results of these two models was provided. Due to the better performance of ANFIS-SC in the first part, we implemented ANFIS-SC and ANFIS-FCM as the second part of the calibration. We accomplished ten ANFIS-SC and eight ANFIS-FCM considering 6 input layers, zero error tolerance, and 200 epochs in ANFIS Architecture. It is worth mentioning that we examined various radius of influence for the SC method in ANFIS-SC and the number of generated clusters for FCM clustering in ANFIS-FCM using both hybrid and back-propagation optimization methods.

We predicted the map of built-up areas for 2019 by the most accurate ANFIS model in the prediction phase, in order to comparison of simulation accuracy with the CA-MC model. Finally, we computed the kappa coefficient and overall accuracy for the predicted map by comparing this map with the classified land use map that we had already obtained.

3.3.1 Training dataset

Regarding there is no accepted rule in ANFIS to determine the number of training dataset, we randomly created a 100000*7 training dataset matrix using 100000 pixels (from a total of 142926) with nonresidential land use from the map of the year 2000. We used 70% of these pixels for training and the other 30% to validate the quality of the training process. The matrix below illustrates our training dataset matrix acceptably.

$$\text{Training dataset matrix} = \begin{bmatrix} X_{1,1} & \dots & X_{1,6} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ X_{100000,1} & \dots & X_{100000,6} & 1 \end{bmatrix}$$

The first six columns were created by normalized input layers, whereas the seventh column was the ANFIS output illustrating the pixel development position in the map of the year 2010. Zero indicates undeveloped and one indicates developed pixel in this column.

3.4 CA-MC model implementation

As a stochastic model, the Markov chain can express the probability of the LULC changing in simulating land use changes (Wu et al., 2019); however, it cannot obtain the spatial distribution of projected LULC (Nguyen et al., 2019). In order to resolve this issue, the CA-MC model, which integrates the CA with the MCA model, has been employed in various urban growth and LULC changes studies. This integrated model includes a spatial dimension considering the effects of neighboring cells by CA neighborhood.

To implement the CA-MC mode, we accomplished the MCA model in the Idrisi software by introducing land use maps of 2000 and 2010 and two time intervals from 2000 to 2010 and 2010 to 2019 into the model. As a result, the transition probability matrix (reveals the probabilities of transitioning from one land use to another one) was obtained. One of the inputs of CA-MC is the transition suitability image which reveals the suitability of converting to another land use. We used weighted linear combination (WLC) as one of the multicriteria evaluation (MCE) methods to obtain these maps. To achieve this aim, first, we obtained transition potential images (express the potential of change in the specific time) for three urban, agricultural, and barren land uses and then converted all three maps to transition suitability images in the Idrisi. It is worth mentioning that we normalized all pixels between 0 to 255 by fuzzy reasoning due to differences in units and pixel values, and utilized the AHP for factor weighting in the MCE method.

The CA-MC model with an iterative process was employed a kernel size of 5×5 as a filter. With every iteration of the model, every land use map is weighted again through the filter to finally simulate the transferred areas of each land use for the 2019 year.

3.4.1 CA-MC model Validation

We assessed the CA-MC model performance by comparing the output map of the model and the land use map of 2019 and obtaining the kappa coefficients in the Idrisi. To fulfill this aim, we computed various Kappa coefficients, namely Kstandard, Kno, Klocation, and KlocationStra to validate the model in both quantification and location. The standard Kappa (denoted as kstandard) reveals the ability of classification to attain perfect classification and uses for quantity (not location) assessment (Pontius Jr, 2000). As Eq.9 indicates, Kstandard signifies the ratio of the categories that correlated correctly to categories that correlated correctly by chance (Khwarahm et al., 2021).

$$K_{\text{standard}} = \frac{M(m) - N(m)}{P(p) - N(m)} \quad (9)$$

Where, M(m), N(m), and P(p) represent the resemblance between the reference map and: the simulated map, a map with random sites of the cells, and a map that has perfect data, respectively. klocation (kappa for location) and klocationstra (Kappa for stratum-level location) reveal what extent the cells are located correctly on the landscape and within the strata, respectively. Eq.10 and Eq.11 calculate these two coefficients.

$$K_{\text{location}} = \frac{M(m) - N(m)}{P(m) - N(m)} \quad (10)$$

$$K_{\text{locationstra}} = \frac{M(m) - H(m)}{K(m) - H(m)} \quad (11)$$

Where, H(m), P(m), and K(m) reveal the resemblance between the reference map and: a map with random sites of the cells in each stratum, a map with regenerate sites of the cells as well as possible in the entire map, and a map with regenerate sites of the cells as well as possible in each stratum, respectively. Finally, Eq.12 states the formula for calculating the kno (kappa for no information) that is defined as the ratio of pixels with correct classification to expected correct classification with no ability to specify quantity or location (Pontius Jr, 2000). The N(n) in Eq.12 can be defined as accidental agreement.

$$K_{\text{no}} = \frac{M(m) - N(n)}{P(p) - N(n)} \quad (12)$$

All mentioned kappas can vary from zero (completely classified by chance) to one (perfect agreement between maps) (Christensen & Jokar Arsanjani, 2020).

In addition to the kappa coefficients, we obtained the overall accuracy in MATLAB to compare the results of the two models extensively.

4 Results and discussion

All input parameters and execution results of ANFIS-GP and ANFIS-SC models, executed in the first part of the calibration section, are provided in table 2. According to this table, although the RMSE of ANFIS-GP was approximately equal to the RMSE of ANFIS-SC, its MSE was greater than the MSE of ANFIS-SC. Besides, it is apparent that the execution time of the ANFIS-GP was longer than ANFIS-SC due to more number of generated rules. On the other hand, ANFIS-SC was able to approximately obtain an equal accuracy to ANFIS-GP with more MFs and fewer rules. More MFs and fewer generated rules lead to more coverage of the input space and less execution time of the model, respectively.

Table 2 execution results of ANFIS-GP and ANFIS-SC in the first part of calibration

clustering method	center's range of influence	Number of MFs	Number of rules	Optimization method	execution time (second)	MSE		RMSE	
						training	test	training	test
ANFIS-GP	-	3	81	hybrid	12355	9.07×10^{-9}	4.47×10^{-3}	0.2325	0.2350
ANFIS-SC	0.1	29	29	hybrid	4024	1.13×10^{-10}	4.06×10^{-3}	0.2327	0.2363
ANFIS-SC	0.2	19	19	hybrid	1997	5.24×10^{-11}	4.33×10^{-3}	0.2369	0.2381
ANFIS-SC	0.3	11	11	hybrid	982	3.79×10^{-11}	4.46×10^{-3}	0.2442	0.2433

Additionally, the execution results and characteristics of ANFIS-SC and ANFIS-FCM in the second part are presented in tables 3 and 4, respectively. The stopping criteria was considered for algorithm as the maximum execution time of 50000 seconds. All models were successfully executed within the time limit, except for the ANFIS-SC with the radius of 0.1. By comparing all RMSE and MSE errors of the training and test data, it was found that the ANFIS-SC with the radius of influence of 0.2 (the hybrid one) was the most accurate for training data. The ANFIS-SC with the radius of 0.4 had the least MSE for test data; however, it had more RMSE and MSE for training data in comparison with the ANFIS-SC with the radius of 0.2. The results were generally indicated the superiority of the ANFIS-SC to the ANFIS-FCM and ANFIS-GP

in urban growth modeling. This finding is consistent with the results of Adedeji et al. (2020), Moradi et al. (2019), and Akkaya (2016) studies accomplished on different areas.

Table 3 execution results of ANFIS-SC in the second part of calibration

clustering method	center's range of influence	Number of MFs	Number of rules	Optimization method	execution time (second)	MSE		RMSE	
						training	test	training	test
ANFIS-SC	0.1	-	-	hybrid	unfinished	-	-	-	-
ANFIS-SC	0.2	48	48	hybrid	14466	1.19×10^{-11}	8.86×10^{-4}	0.2296	0.2366
ANFIS-SC	0.4	14	14	hybrid	1954	7.88×10^{-11}	1.09×10^{-4}	0.2372	0.2395
ANFIS-SC	0.6	6	6	hybrid	899	1.84×10^{-10}	3.02×10^{-4}	0.2429	0.2447
ANFIS-SC	0.8	3	3	hybrid	701	2.21×10^{-10}	1.61×10^{-4}	0.2439	0.2456
ANFIS-SC	0.1	154	154	back-propagation	8458	3.48×10^{-4}	6.02×10^{-4}	0.2338	0.2446
ANFIS-SC	0.2	48	48	back-propagation	2427	1.32×10^{-3}	1.2×10^{-3}	0.2456	0.2480
ANFIS-SC	0.4	14	14	back-propagation	990	2.55×10^{-3}	2.8×10^{-3}	0.2486	0.2499
ANFIS-SC	0.6	6	6	back-propagation	709	4.59×10^{-4}	4.35×10^{-4}	0.2510	0.2524
ANFIS-SC	0.8	3	3	back-propagation	629	1.69×10^{-3}	1.11×10^{-3}	0.2549	0.2557

Table 4 execution results of ANFIS-FCM in the second part of calibration

clustering method	Number of clusters	Number of MFs	Number of rules	Optimization method	execution time (second)	MSE		RMSE	
						training	test	training	test
ANFIS-FCM	2	2	2	hybrid	245	2.53×10^{-11}	2.18×10^{-4}	0.2497	0.2505
ANFIS-FCM	4	4	4	hybrid	341	2.28×10^{-11}	5.2×10^{-4}	0.2470	0.2482
ANFIS-FCM	6	6	6	hybrid	497	1.71×10^{-11}	9.16×10^{-2}	0.2586	0.3682
ANFIS-FCM	8	8	8	hybrid	715	3.58×10^{-11}	8.87×10^{-2}	0.2524	0.3601
ANFIS-FCM	2	2	2	back-propagation	201	3.79×10^{-3}	3.09×10^{-3}	0.2573	0.2581

ANFIS-FCM	4	4	4	back-propagation	311	8.81×10^{-4}	8.6×10^{-4}	0.2563	0.2569
ANFIS-FCM	6	6	6	back-propagation	439	6.68×10^{-3}	9.59×10^{-2}	0.2662	0.3531
ANFIS-FCM	8	8	8	back-propagation	698	4.12×10^{-4}	6.32×10^{-4}	0.2641	0.2668

The line charts in Fig.6 reveal the relation between the RMSE error of model training and the number of clusters/center's range of influence in both back-propagation and hybrid methods. From these line charts it can be observed that the accuracy of training in ANFIS-FCM has not necessarily increased with the increase of clusters; however, the accuracy of ANFIS-SC has increased with decreasing radius (increasing clusters). Furthermore, it is quite obvious that the models that have used the hybrid optimization algorithm have always had less error than the back-propagation algorithm. Hence, utilizing an ANFIS-SC with a hybrid algorithm and smaller radius of influence can lead to more accurate modeling of urban expansion.

As it is indicated in Fig.7, the target-output plot, errors plot, and the error histogram chart of the ANFIS-SC with the radius of 0.2 were generated in order to further examine the outputs of ANFIS model. The target-output plot represents the ideal output of ANFIS (0 and 1) as target and the predicted output as output at the same time, whereas the errors plot reveals the difference between them, and the error histogram chart illustrates the distribution of errors.

Looking at the error histogram chart, it is apparent that more errors were close to zero, which means the predicted outputs of the ANFIS-SC have been close to the ideal output. Consequently, the results of the model have been satisfactory.

The predicted map of the ANFIS-SC (radius of 0.2) model for 2019 had the Kappa coefficient of 0.76 and overall accuracy of 93.41, whereas the CA-MC model had the Kappa of 0.66 and overall accuracy of 77.96. The kno of 0.66, Klocation of 0.79, and KlocationStra of 0.79 for the CA-MC model indicated that this model can have satisfactory results; however, the ANFIS model had much better results than CA-MC due to better Kappa and overall accuracy. Besides, the Kappa between 0.6 and 0.8 is considered a very good kappa (Pijanowski et al., 2005). Hence, as it is presented in Fig.8, the predicted map for 2029 was achieved by ANFIS-SC model.

Comparing this map with the 2019 land use map, it is apparent that the projected development will occur in the north more than in other directions. A possible explanation for this may be the existence of two large universities by enrollment in the north of the area in addition to the suitability of other driving factors. These universities with more than 36000 enrolled students cause a large number of trips to the north of the city during the day. The examination of the road land use changes over the last few years reveals that urban decision-makers have improved the road transport network in the northern areas to meet the requirements of the students. The built-up roads to the universities have improved access into the northern barren lands; consequently, they have increased the probability of LULC change of these lands. Hence, large educational institutions can generally play an important role in increasing residential land use in the region, especially in small and medium-sized cities.

5 Conclusion

This study was established to develop a powerful model based on ANFIS, and to investigate the role of clustering method in the accuracy of ANFIS model training and simulation power in urban expansion modeling. The most obvious finding to emerge from this study was that the proposed ANFIS model had a high accuracy in modeling the complex phenomenon of urban expansion and was able to be an acceptable alternative to the traditional model of cellular automata. Besides, according to the results of this study, the ANFIS model offered a more precise simulation of LULC in comparison with the CA-MC model. According to the results, we can infer that the ANFIS-SC model will lead to more precise urban expansion modeling in comparison with the ANFIS-FCM. Moreover, although the ANFIS-GP has precise and acceptable results, it is not a great option for modeling urban expansion due to longer execution time and inability to execute and high complexity in handling high-dimensional problems (Adedeji et al., 2020).

An important aim of the present research is to warn about the harmful effects of spontaneous urban expansion on the environment, using a functional model and predicting the land use changes from a ten-year perspective. The valuable finding of our study is that the city expansion will occur to the north toward the barren lands, and expansion to the other directions will not approximately appear. The sprawl development in these barren lands can pose a significant threat to the environment of this area that is a habitat for various animals. Contrary to expectations, the predicted map of 2029 indicates that agricultural lands that are located in the south, west, and southwest of the city, will not be damaged in the future due to land use

changes. This finding is likely to be attributed to the lower price of barren land compared with agricultural land, the fertility of agricultural land, and supportive government policies protecting agricultural land, and it will be of broad use to policymakers.

The most important limitation of this study lies in the fact that we are not able to utilize all driving factors influencing the expansion of the city due to some factors like the price of land parcels of the area in the past is not available now. Notwithstanding this limitation, this research employs as many important factors as possible to study urban growth. Another limitation of this study is the need for computers with powerful hardware to run more complex ANFIS models, especially the ANFIS-GP or ANFIS-SC with a smaller radius of influence. According to the results of training and test dataset, the key strength of this study has been high-precision training of the model (without overfitting) by a high number of training dataset (100000 pixels out of a total of approximately 142000 pixels).

ANFIS models were implemented based on cellular spaces in this study. Hence, it might be possible to use a parcel-based land use classification in further investigation, and instead of using cells in the model, parcels be used as the smallest unit for studying land use changes. Additionally, the center's range of influence between 0.2 and 0.8 in ANFIS-SC was investigated due to the mentioned limitations, whereas it revealed that the model results improved with decreasing radius of clusters. Hence, further studies need to be carried out to achieve a more accurate training of the model using a radius smaller than 0.2, and cautiously investigate the relationship between decreasing radius and increasing model accuracy. Moreover, a further study on implementing ANFIS-CFCM in urban growth modeling and comparing the results with the other three ANFIS models are suggested.

Figures caption

Fig.1 the map and satellite image of the study area: **a** location of Qazvin province in Iran, **b** location of Qazvin city in the county, and **c** satellite image of Qazvin city (Source: Landsat 8 satellite imagery)

Fig.2 the ANFIS architecture with three inputs, three rules and one output

Fig.3 principal performed steps of the study

Fig.4 raster layers belonging to six expansion driving variables: **a** slope, **b** aspect, **c** distance to the urban areas, **d** distance to the urban transportation network, **e** distance to the urban squares, and **f** distance to the universities

Fig.5 ANFIS-SC structure generated in MATLAB GUI

Fig.6 RMSE error based on the number and radius of the clusters: **a** for ANFIS-FCM and **b** for ANFIS-SC

Fig.7 target-output plot, errors plot, and the error histogram chart of ANFIS-SC: **a** target-output plot, **b** errors plot, and **c** error histogram chart

Fig.8 2029 projected map produced by ANFIS-SC

Acknowledgments

The authors would like to acknowledge the GIS bureau of Qazvin municipality for their assistance with the collection of data and the United States Geological Survey for providing satellite images.

Declarations

Funding

No funds, grants, or other support was received.

Conflicts of interest/Competing interests

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Availability of data and material

The datasets generated during the current study are available from the corresponding author on reasonable request.

Code availability

The custom computer code used in this study to generate results are available from the corresponding author on reasonable request. No custom-developed software was used in this study.

Authors' contributions

Both authors contributed to the study stages, such as data curation, construction and development of methodology, and implementation and evaluation of the computer code. Yousef Ghobadiha has written the initial draft, and Hamid Motieyan commented on the early draft of the manuscript. Both authors study and confirmed the eventual manuscript.

Ethics approval

Not applicable

Consent to participate

Not applicable

Consent for publication

Not applicable

References

- Abraham, A. (2005). Adaptation of fuzzy inference system using neural learning. In N. Nedjah & L. Macedo Mourelle (Eds.), *Fuzzy systems engineering* (pp. 53-83). Springer. https://doi.org/10.1007/11339366_3
- Aburas, M. M., Ho, Y. M., Ramli, M. F., & Ash'aari, Z. H. (2016). The simulation and prediction of spatio-temporal urban growth trends using cellular automata models: A review. *International Journal of Applied Earth Observation and Geoinformation*, 52, 380-389. <https://doi.org/10.1016/j.jag.2016.07.007>
- Adedeji, P. A., Akinlabi, S., Madushele, N., & Olatunji, O. O. (2020). Wind turbine power output very short-term forecast: a comparative study of data clustering techniques in a PSO-ANFIS model. *Journal of Cleaner Production*, 254, 120-135. <https://doi.org/10.1016/j.jclepro.2020.120135>
- Adeleke, O., Akinlabi, S. A., Jen, T.-C., & Dunmade, I. (2020). Prediction of municipal solid waste generation: an investigation of the effect of clustering techniques and parameters on ANFIS model performance. *Environmental Technology*, 1-14. <https://doi.org/10.1080/09593330.2020.1845819>
- Ahmadlou, M., Karimi, M., Alizadeh, S., Shirzadi, A., Parvinnejhad, D., Shahabi, H., & Panahi, M. (2019). Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA). *Geocarto International*, 34(11), 1252-1272. <https://doi.org/10.1080/10106049.2018.1474276>
- Akkaya, E. (2016). ANFIS based prediction model for biomass heating value using proximate analysis components. *Fuel*, 180, 687-693. <https://doi.org/10.1016/j.fuel.2016.04.112>
- Al-Ahmadi, K., See, L., Heppenstall, A., & Hogg, J. (2009). Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia. *Ecological complexity*, 6(2), 80-101. <https://doi.org/10.1016/j.ecocom.2008.09.004>
- Angel, S., Blei, A. M., Parent, J., Lamson-Hall, P., Sánchez, N. G., Civco, D., Lei, R. Q., & Thom, K. (2016). *Atlas of urban expansion—2016 Edition* (Vol. 1: Areas and densities). New York: New York University, Nairobi: UN-Habitat, and Cambridge, MA: Lincoln Institute of Land Policy.
- Artmann, M., Inostroza, L., & Fan, P. (2019). Urban sprawl, compact urban development and green cities. How much do we know, how much do we agree?. *Ecological Indicators*, 96, 3-9. <https://doi.org/10.1016/j.ecolind.2018.10.059>
- Asubonteng, K., Pfeffer, K., Ros-Tonen, M., Verbesselt, J., & Baud, I. (2018). Effects of tree-crop farming on land-cover transitions in a mosaic landscape in the eastern region of Ghana. *Environmental management*, 62(3), 529-547. <https://doi.org/10.1007/s00267-018-1060-3>
- Azari, M., Tayyebi, A., Helbich, M., & Reveshty, M. A. (2016). Integrating cellular automata, artificial neural network, and fuzzy set theory to simulate threatened orchards: application to Maragheh, Iran. *GIScience & Remote Sensing*, 53(2), 183-205. <https://doi.org/10.1080/15481603.2015.1137111>
- Balzter, H. (2000). Markov chain models for vegetation dynamics. *Ecological modelling*, 126(2-3), 139-154. [https://doi.org/10.1016/S0304-3800\(00\)00262-3](https://doi.org/10.1016/S0304-3800(00)00262-3)
- Benmouiza, K., & Cheknane, A. (2019). Clustered ANFIS network using fuzzy c-means, subtractive clustering, and grid partitioning for hourly solar radiation forecasting. *Theoretical and Applied Climatology*, 137(1), 31-43. <https://doi.org/10.1007/s00704-018-2576-4>

- Berberoğlu, S., Akın, A., & Clarke, K. C. (2016). Cellular automata modeling approaches to forecast urban growth for adana, Turkey: A comparative approach. *Landscape and Urban Planning*, 153, 11-27. <https://doi.org/10.1016/j.landurbplan.2016.04.017>
- Bihamta, N., Soffianian, A., Fakheran, S., & Gholamalifard, M. (2015). Using the SLEUTH urban growth model to simulate future urban expansion of the Isfahan metropolitan area, Iran. *Journal of the Indian Society of Remote Sensing*, 43(2), 407-414. <https://doi.org/10.1007/s12524-014-0402-8>
- Blecic, I., Cecchini, A., & Trunfio, G. A. (2015). How much past to see the future: a computational study in calibrating urban cellular automata. *International Journal of Geographical Information Science*, 29(3), 349-374. <https://doi.org/10.1080/13658816.2014.970190>
- Broere, W. (2016). Urban underground space: Solving the problems of today's cities. *Tunnelling and Underground Space Technology*, 55, 245-248. <https://doi.org/10.1016/j.tust.2015.11.012>
- Christensen, M., & Jokar Arsanjani, J. (2020). Stimulating implementation of sustainable development goals and conservation action: predicting future land use/cover change in Virunga National Park, Congo. *Sustainability*, 12(4), 1570. <https://doi.org/10.3390/su12041570>
- Ewing, R. H., Pendall, R., & Chen, D. D. (2002). *Measuring sprawl and its impact* (Vol. 1). Washington, DC: Smart Growth America.
- Ezeh, A., Oyeboode, O., Satterthwaite, D., Chen, Y.-F., Ndugwa, R., Sartori, J., Mberu, B., Melendez-Torres, G. J., Haregu, T., & Watson, S. I. (2017). The history, geography, and sociology of slums and the health problems of people who live in slums. *The lancet*, 389(10068), 547-558. [https://doi.org/10.1016/S0140-6736\(16\)31650-6](https://doi.org/10.1016/S0140-6736(16)31650-6)
- Fan, F., Wang, Y., & Wang, Z. (2008). Temporal and spatial change detecting (1998–2003) and predicting of land use and land cover in Core corridor of Pearl River Delta (China) by using TM and ETM+ images. *Environmental monitoring and assessment*, 137(1), 127-147. <https://doi.org/10.1007/s10661-007-9734-y>
- Frumkin, H. (2002). Urban sprawl and public health. *Public health reports*, 117(3), 201-217. <https://doi.org/10.1093/phr/117.3.201>
- Gross, J., & Ouyang, Y. (2021). Types of urbanization and economic growth. *International Journal of Urban Sciences*, 25(1), 71-85. <https://doi.org/10.1080/12265934.2020.1759447>
- Hegazy, I., Seddik, W., & Ibrahim, H. (2017). Towards green cities in developing countries: Egyptian new cities as a case study. *International Journal of Low-Carbon Technologies*, 12(4), 358-368. <https://doi.org/10.1093/ijlct/ctx009>
- Hosseini, S. H., & Hajilou, M. (2019). Drivers of urban sprawl in urban areas of Iran. *Papers in Regional Science*, 98(2), 1137-1158. <https://doi.org/10.1111/pirs.12381>
- Jang, J.-S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, 23(3), 665-685. <https://doi.org/10.1109/21.256541>
- Jha, S. S., Manohar Kumar, C., & Nidamanuri, R. R. (2021). Flexible atmospheric compensation technique (FACT): a 6S based atmospheric correction scheme for remote sensing data. *Geocarto International*, 36(1), 28-46. <https://doi.org/10.1080/10106049.2019.1588391>
- Jiao, L., Liu, J., Xu, G., Dong, T., Gu, Y., Zhang, B., Liu, Y., & Liu, X. (2018). Proximity Expansion Index: An improved approach to characterize evolution process of urban expansion. *Computers, Environment and Urban Systems*, 70, 102-112. <https://doi.org/10.1016/j.compenvurbsys.2018.02.005>

- Kabisch, N., & Haase, D. (2011). Diversifying European agglomerations: Evidence of urban population trends for the 21st century. *Population, space and place*, 17(3), 236-253. <https://doi.org/10.1002/psp.600>
- Kamusoko, C., & Gamba, J. (2015). Simulating urban growth using a Random Forest-Cellular Automata (RF-CA) model. *ISPRS International Journal of Geo-Information*, 4(2), 447-470. <https://doi.org/10.3390/ijgi4020447>
- Karaboga, D., & Kaya, E. (2019). Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey. *Artificial Intelligence Review*, 52(4), 2263-2293. <https://doi.org/10.1007/s10462-017-9610-2>
- Karimi, F., Sultana, S., Babakan, A. S., & Suthaharan, S. (2019). An enhanced support vector machine model for urban expansion prediction. *Computers, Environment and Urban Systems*, 75, 61-75. <https://doi.org/10.1016/j.compenvurbsys.2019.01.001>
- Khamis, K., Bradley, C., & Hannah, D. (2018). Understanding dissolved organic matter dynamics in urban catchments: insights from in situ fluorescence sensor technology. *Wiley Interdisciplinary Reviews: Water*, 5(1), e1259. <https://doi.org/10.1002/wat2.1259>
- Khwarahm, N. R., Qader, S., Ararat, K., & Al-Quraishi, A. M. F. (2021). Predicting and mapping land cover/land use changes in Erbil/Iraq using CA-Markov synergy model. *Earth Science Informatics*, 14(1), 393-406. <https://doi.org/10.1007/s12145-020-00541-x>
- Li, X., Chen, Y., Liu, X., Xu, X., & Chen, G. (2017). Experiences and issues of using cellular automata for assisting urban and regional planning in China. *International Journal of Geographical Information Science*, 31(8), 1606-1629. <https://doi.org/10.1080/13658816.2017.1301457>
- Li, X., & Gong, P. (2016). Urban growth models: progress and perspective. *Science bulletin*, 61(21), 1637-1650. <https://doi.org/10.1007/s11434-016-1111-1>
- Li, X., Lao, C., Liu, Y., Liu, X., Chen, Y., Li, S., Ai, B., & He, Z. (2013). Early warning of illegal development for protected areas by integrating cellular automata with neural networks. *Journal of Environmental Management*, 130, 106-116. <https://doi.org/10.1016/j.jenvman.2013.08.055>
- Li, X., & Yeh, A. G.-O. (2001). Calibration of cellular automata by using neural networks for the simulation of complex urban systems. *Environment and Planning A*, 33(8), 1445-1462. <https://doi.org/10.1068%2Fa33210>
- Li, X., & Yeh, A. G.-O. (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 16(4), 323-343. <https://doi.org/10.1080/13658810210137004>
- Liu, Y., Hu, Y., Long, S., Liu, L., & Liu, X. (2017). Analysis of the effectiveness of urban land-use-change models based on the measurement of spatio-temporal, dynamic urban growth: A cellular automata case study. *Sustainability*, 9(5), 796. <https://doi.org/10.3390/su9050796>
- Lu, Q., Joyce, J., Imen, S., & Chang, N.-B. (2019). Linking socioeconomic development, sea level rise, and climate change impacts on urban growth in New York City with a fuzzy cellular automata-based Markov chain model. *Environment and Planning B: Urban Analytics and City Science*, 46(3), 551-572. <https://doi.org/10.1177%2F2399808317720797>
- Masoumi, F., Eslamkish, T., Honarmand, M., & Abkar, A. A. (2017). A comparative study of landsat-7 and landsat-8 data using image processing methods for hydrothermal alteration mapping. *Resource Geology*, 67(1), 72-88. <https://doi.org/10.1111/rge.12117>

- McDonnell, M. J., & Hahs, A. K. (2015). Adaptation and adaptedness of organisms to urban environments. *Annual review of ecology, evolution, and systematics*, 46, 261-280. <https://doi.org/10.1146/annurev-ecolsys-112414-054258>
- Mohammady, S. (2016). Optimization of adaptive neuro fuzzy inference system based urban growth model. *City, Territory and Architecture*, 3(1), 1-15. <https://doi.org/10.1186/s40410-016-0039-8>
- Mohammady, S., Delavar, M., & Pijanowski, B. (2013). Urban growth modeling using ANFIS algorithm: a case study for Sanandaj city, Iran. *Int Arch Photogramm Remote Sens Spatial Inf Sci*, 3, 493-498. <https://doi.org/10.5194/isprsarchives-XL-1-W3-493-2013>
- Mohammady, S., & Delavar, M. R. (2016). Urban sprawl assessment and modeling using landsat images and GIS. *Modeling Earth Systems and Environment*, 2(3), 1-14. <https://doi.org/10.1007/s40808-016-0209-4>
- Mondal, B., Chakraborti, S., Das, D. N., Joshi, P. K., Maity, S., Pramanik, M. K., & Chatterjee, S. (2020). Comparison of spatial modelling approaches to simulate urban growth: a case study on Udaipur city, India. *Geocarto International*, 35(4), 411-433. <https://doi.org/10.1080/10106049.2018.1520922>
- Moradi, F., Bonakdari, H., Kisi, O., Ebtahaj, I., Shiri, J., & Gharabaghi, B. (2019). Abutment scour depth modeling using neuro-fuzzy-embedded techniques. *Marine Georesources & Geotechnology*, 37(2), 190-200. <https://doi.org/10.1080/1064119X.2017.1420113>
- Musa, S. I., Hashim, M., & Reba, M. N. M. (2017). A review of geospatial-based urban growth models and modelling initiatives. *Geocarto International*, 32(8), 813-833. <https://doi.org/10.1080/10106049.2016.1213891>
- Mustafa, A., Cools, M., Saadi, I., & Teller, J. (2017). Coupling agent-based, cellular automata and logistic regression into a hybrid urban expansion model (HUEM). *Land Use Policy*, 69, 529-540. <https://doi.org/10.1016/j.landusepol.2017.10.009>
- Mustafa, A., Heppenstall, A., Omrani, H., Saadi, I., Cools, M., & Teller, J. (2018). Modelling built-up expansion and densification with multinomial logistic regression, cellular automata and genetic algorithm. *Computers, Environment and Urban Systems*, 67, 147-156. <https://doi.org/10.1016/j.compenvurbsys.2017.09.009>
- Nguyen, T. A., Le, P. M. T., Pham, T. M., Hoang, H. T. T., Nguyen, M. Q., Ta, H. Q., Phung, H. T. M., Le, H. T. T., & Hens, L. (2019). Toward a sustainable city of tomorrow: a hybrid Markov–Cellular Automata modeling for urban landscape evolution in the Hanoi city (Vietnam) during 1990–2030. *Environment, Development and Sustainability*, 21(1), 429-446. <https://doi.org/10.1007/s10668-017-0046-2>
- Parvinnezhad, D., Delavar, M. R., Claramunt, C., & Pijanowski, B. C. (2019). A modified spatial entropy for urban sprawl assessment. *Geocarto International*, 1-16. <https://doi.org/10.1080/10106049.2019.1678676>
- Patel, N., & Kaushal, B. (2011). Classification of features selected through Optimum Index Factor (OIF) for improving classification accuracy. *Journal of Forestry Research*, 22(1), 99-105. <https://doi.org/10.1007/s11676-011-0133-4>
- Pijanowski, B. C., Pithadia, S., Shellito, B. A., & Alexandridis, K. (2005). Calibrating a neural network-based urban change model for two metropolitan areas of the Upper Midwest of the United States. *International Journal of Geographical Information Science*, 19(2), 197-215. <https://doi.org/10.1080/13658810410001713416>
- Pijanowski, B. C., Tayyebi, A., Doucette, J., Pekin, B. K., Braun, D., & Plourde, J. (2014). A big data urban growth simulation at a national scale: configuring the GIS and neural network based land transformation model to run in a high performance computing (HPC) environment. *Environmental Modelling & Software*, 51, 250-268. <https://doi.org/10.1016/j.envsoft.2013.09.015>

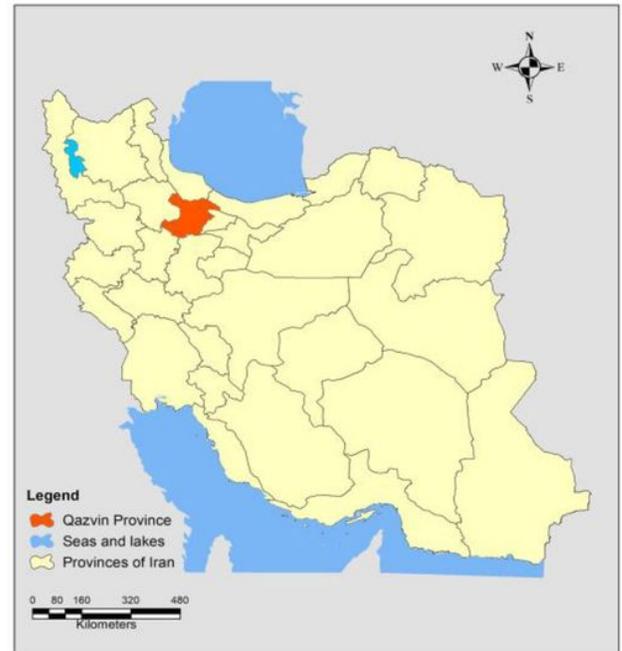
- Pontius Jr, R. G. (2000). Comparison of categorical maps. *Photogramm. Eng. Remote Sens.*, 66, 1011-1016. http://www2.clarku.edu/~rpontius/pontius_2000_pers.pdf
- Radwan, T. M., Blackburn, G. A., Whyatt, J. D., & Atkinson, P. M. (2019). Dramatic loss of agricultural land due to urban expansion threatens food security in the Nile Delta, Egypt. *Remote Sensing*, 11(3), 332. <https://doi.org/10.3390/rs11030332>
- Rimal, B., Zhang, L., Keshtkar, H., Haack, B. N., Rijal, S., & Zhang, P. (2018). Land use/land cover dynamics and modeling of urban land expansion by the integration of cellular automata and markov chain. *ISPRS International Journal of Geo-Information*, 7(4), 154. <https://doi.org/10.3390/ijgi7040154>
- Rimal, B., Zhang, L., Keshtkar, H., Wang, N., & Lin, Y. (2017). Monitoring and modeling of spatiotemporal urban expansion and land-use/land-cover change using integrated Markov chain cellular automata model. *ISPRS International Journal of Geo-Information*, 6(9), 288. <https://doi.org/10.3390/ijgi6090288>
- Samardžić-Petrović, M., Kovačević, M., Bajat, B., & Dragičević, S. (2017). Machine learning techniques for modelling short term land-use change. *ISPRS International Journal of Geo-Information*, 6(12), 387. <https://doi.org/10.3390/ijgi6120387>
- Santé, I., García, A. M., Miranda, D., & Crecente, R. (2010). Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landscape and Urban Planning*, 96(2), 108-122. <https://doi.org/10.1016/j.landurbplan.2010.03.001>
- Shafizadeh-Moghadam, H., Asghari, A., Taleai, M., Helbich, M., & Tayyebi, A. (2017). Sensitivity analysis and accuracy assessment of the land transformation model using cellular automata. *GIScience & Remote Sensing*, 54(5), 639-656. <https://doi.org/10.1080/15481603.2017.1309125>
- Shafizadeh-Moghadam, H., Asghari, A., Tayyebi, A., & Taleai, M. (2017). Coupling machine learning, tree-based and statistical models with cellular automata to simulate urban growth. *Computers, Environment and Urban Systems*, 64, 297-308. <https://doi.org/10.1016/j.compenvurbsys.2017.04.002>
- Southon, G. E., Jorgensen, A., Dunnett, N., Hoyle, H., & Evans, K. L. (2017). Biodiverse perennial meadows have aesthetic value and increase residents' perceptions of site quality in urban green-space. *Landscape and Urban Planning*, 158, 105-118. <https://doi.org/10.1016/j.landurbplan.2016.08.003>
- Statistical Center of Iran. (2019). POPULATION. In *IRAN STATISTICAL YEARBOOK 2018-2019* (pp. 127-174). <https://www.amar.org.ir/Portals/1/yearbook/1397/03.pdf>
- Taravat, A., Rajaei, M., & Emadodin, I. (2017). Urbanization dynamics of Tehran city (1975–2015) using artificial neural networks. *Journal of Maps*, 13(1), 24-30. <https://doi.org/10.1080/17445647.2017.1305300>
- Triantakoustantis, D., Prastacos, P., & Tsoukala, A. (2014). Analyzing urban sprawl in Rethymno, Greece. *Journal of the Indian Society of Remote Sensing*, 42(3), 601-610. <https://doi.org/10.1007/s12524-013-0350-8>
- van Vliet, J., Bregt, A. K., Brown, D. G., van Delden, H., Heckbert, S., & Verburg, P. H. (2016). A review of current calibration and validation practices in land-change modeling. *Environmental Modelling & Software*, 82, 174-182. <https://doi.org/10.1016/j.envsoft.2016.04.017>
- Wang, R., Derdouri, A., & Murayama, Y. (2018). Spatiotemporal simulation of future land use/cover change scenarios in the Tokyo metropolitan area. *Sustainability*, 10(6), 2056. <https://doi.org/10.3390/su10062056>
- Wang, S., Jiao, X., Wang, L., Gong, A., Sang, H., Salahou, M. K., & Zhang, L. (2020). Integration of Boosted Regression Trees and Cellular Automata—Markov Model to Predict the Land Use Spatial Pattern in Hotan Oasis. *Sustainability*, 12(4), 1396. <https://doi.org/10.3390/su12041396>

- Wang, S., Ma, Q., Ding, H., & Liang, H. (2018). Detection of urban expansion and land surface temperature change using multi-temporal landsat images. *Resources, Conservation and Recycling*, 128, 526-534.
<https://doi.org/10.1016/j.resconrec.2016.05.011>
- World Bank. (2018). *Urban population (% of total population) - Iran, Islamic Rep.* Retrieved April 21, 2021, from
<https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?end=2018&locations=IR&start=1960&view=chart>
- Wu, H., Li, Z., Clarke, K. C., Shi, W., Fang, L., Lin, A., & Zhou, J. (2019). Examining the sensitivity of spatial scale in cellular automata Markov chain simulation of land use change. *International Journal of Geographical Information Science*, 33(5), 1040-1061. <https://doi.org/10.1080/13658816.2019.1568441>
- Xian, S., Li, L., & Qi, Z. (2019). Toward a sustainable urban expansion: A case study of Zhuhai, China. *Journal of Cleaner Production*, 230, 276-285.
<https://doi.org/10.1016/j.jclepro.2019.05.087>
- Xu, T., Gao, J., & Coco, G. (2019). Simulation of urban expansion via integrating artificial neural network with Markov chain–cellular automata. *International Journal of Geographical Information Science*, 33(10), 1960-1983.
<https://doi.org/10.1080/13658816.2019.1600701>
- Yang, Q., Li, X., & Shi, X. (2008). Cellular automata for simulating land use changes based on support vector machines. *Computers & geosciences*, 34(6), 592-602.
<https://doi.org/10.1016/j.cageo.2007.08.003>
- Yeom, C.-U., & Kwak, K.-C. (2018). Performance comparison of ANFIS models by input space partitioning methods. *Symmetry*, 10(12), 700.
<https://doi.org/10.3390/sym10120700>
- Zhang, Y., Shen, W., Li, M., & Lv, Y. (2020). Assessing spatio-temporal changes in forest cover and fragmentation under urban expansion in Nanjing, eastern China, from long-term Landsat observations (1987–2017). *Applied Geography*, 117, 102190.
<https://doi.org/10.1016/j.apgeog.2020.102190>
- Zhou, Q., Herrera-Herbert, J., & Hidalgo, A. (2017). Predicting the risk of fault-induced water inrush using the adaptive neuro-fuzzy inference system. *Minerals*, 7(4), 55.
<https://doi.org/10.3390/min7040055>

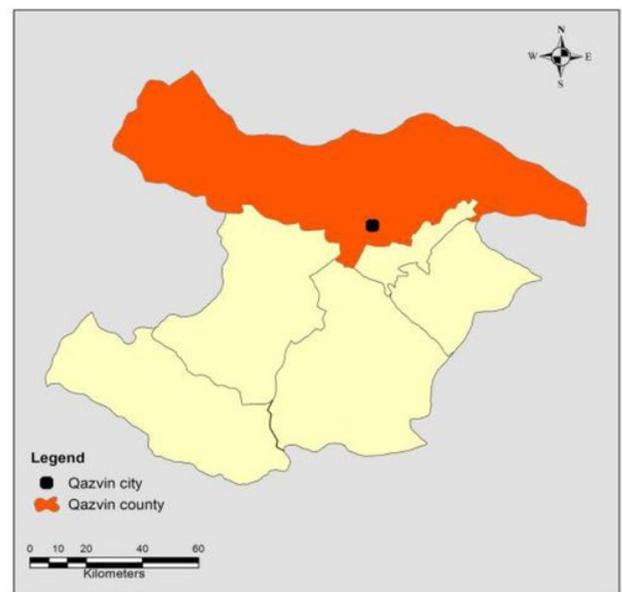
Figures



(c)



(a)



(b)

Figure 1

the map and satellite image of the study area: a location of Qazvin province in Iran, b location of Qazvin city in the county, and c satellite image of Qazvin city (Source: Landsat 8 satellite imagery)

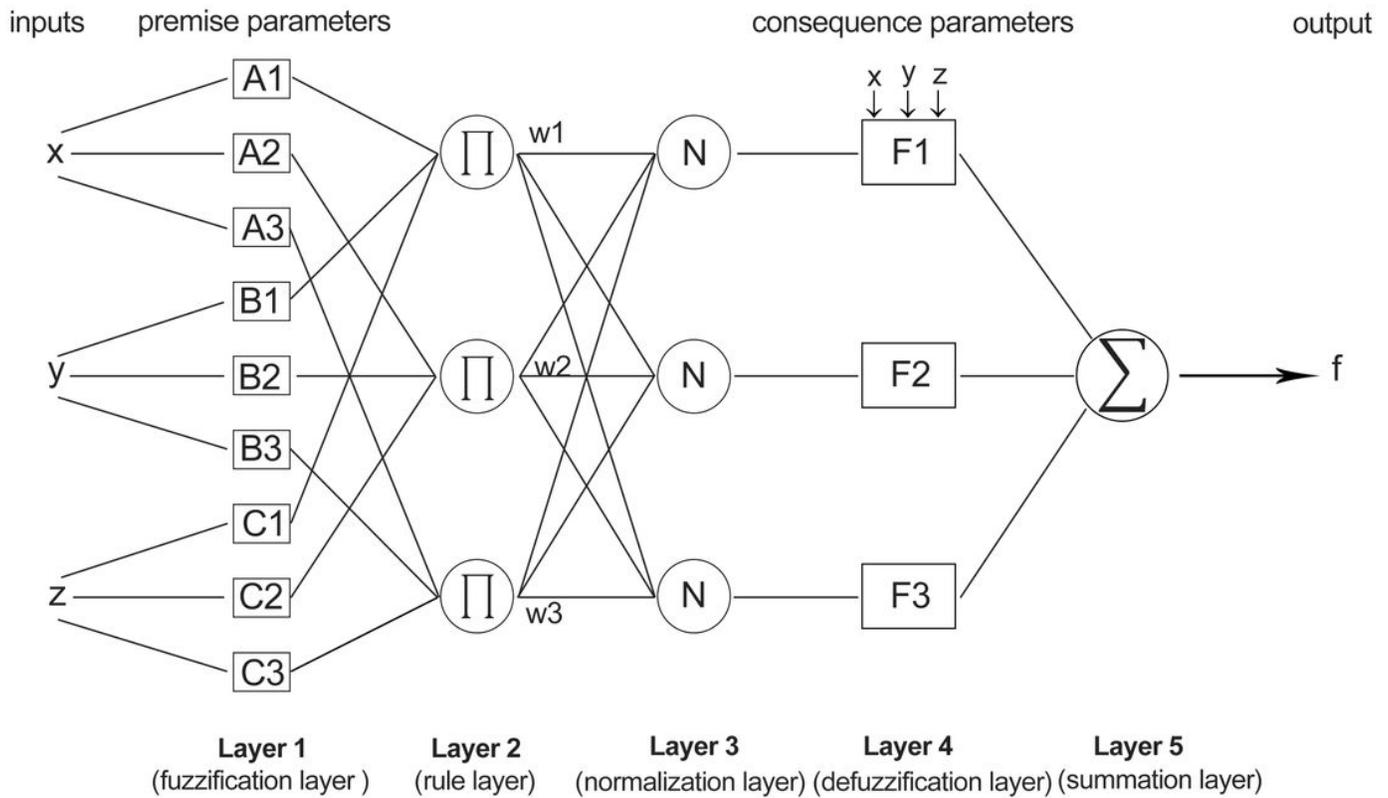


Figure 2

the ANFIS architecture with three inputs, three rules and one output

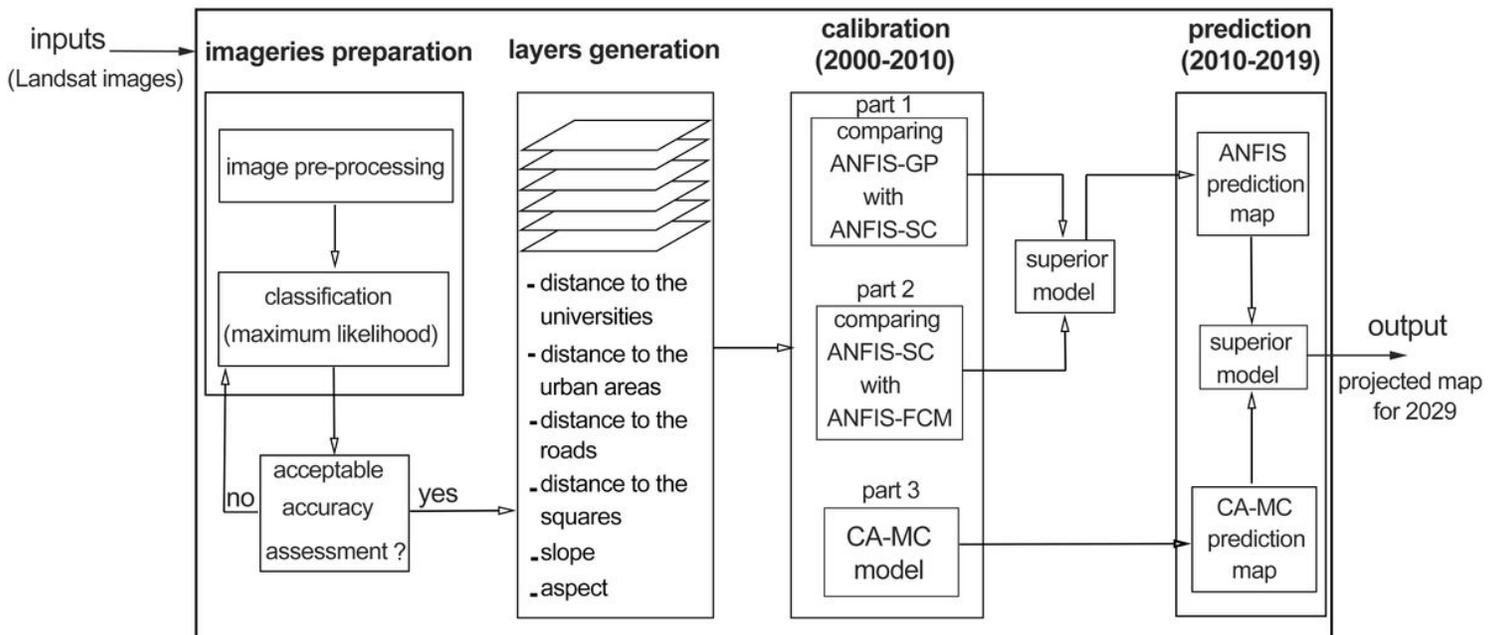


Figure 3

principal performed steps of the study

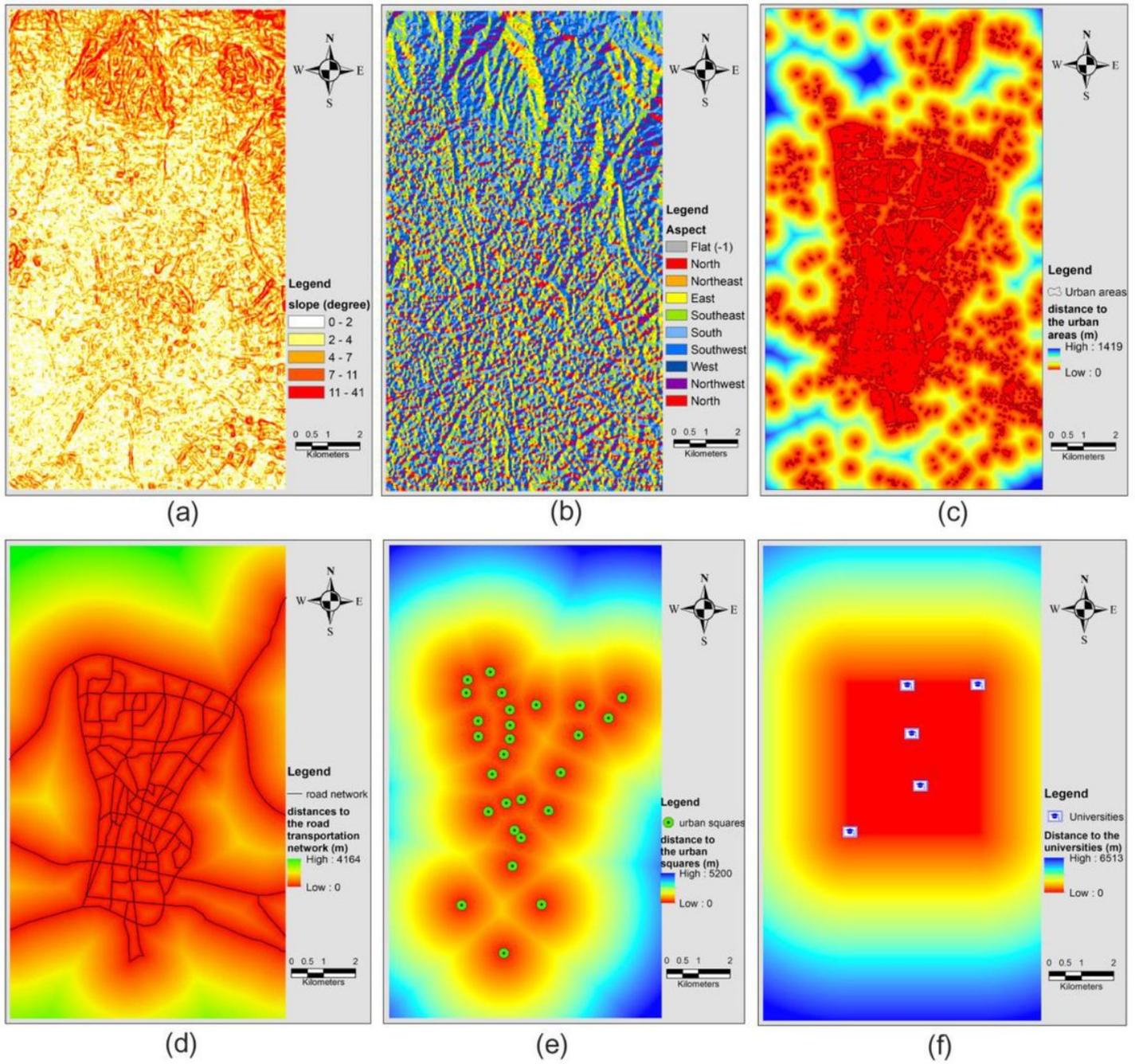


Figure 4

raster layers belonging to six expansion driving variables: a slope, b aspect, c distance to the urban areas, d distance to the urban transportation network, e distance to the urban squares, and f distance to the universities

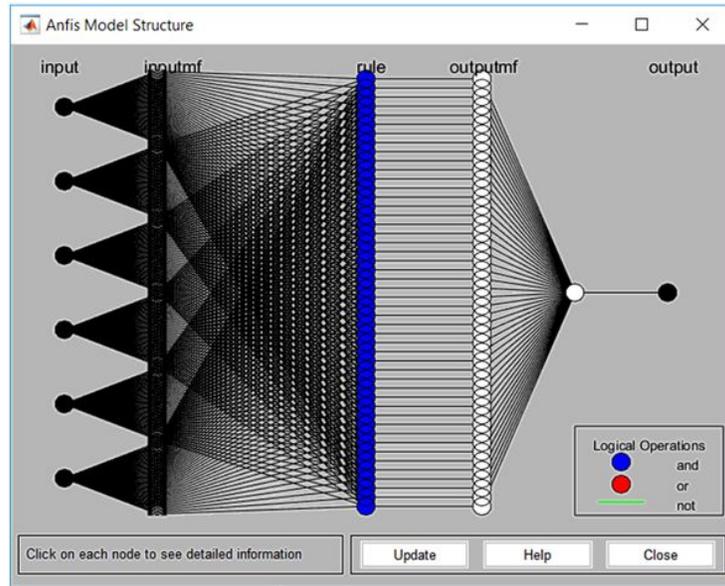


Figure 5

ANFIS-SC structure generated in MATLAB GUI

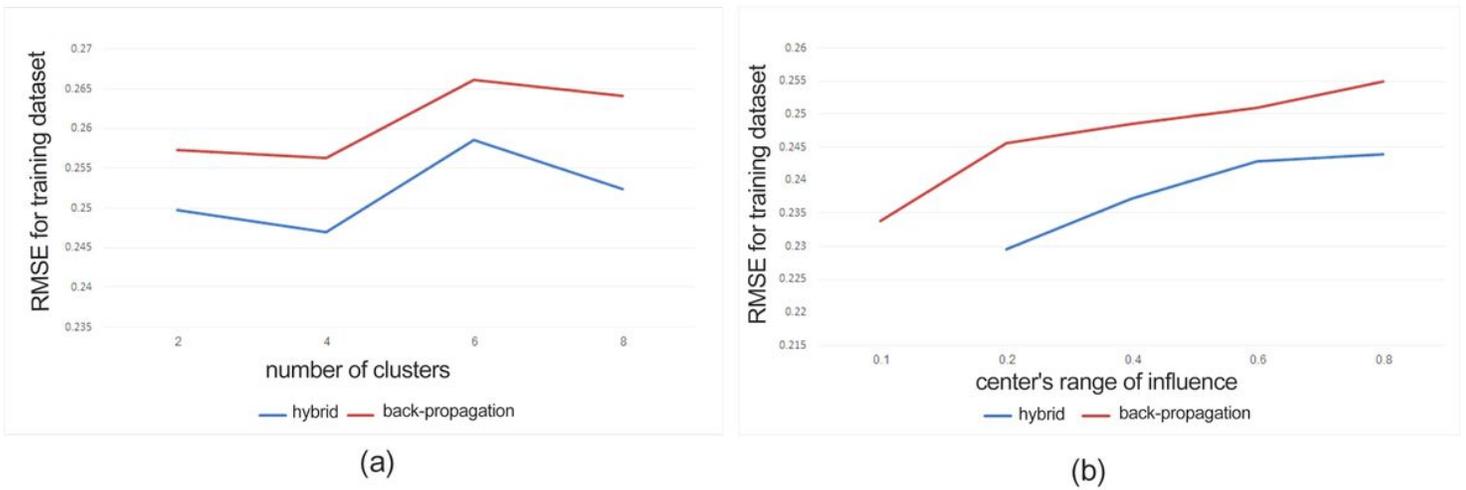


Figure 6

RMSE error based on the number and radius of the clusters: a for ANFIS-FCM and b for ANFIS-SC

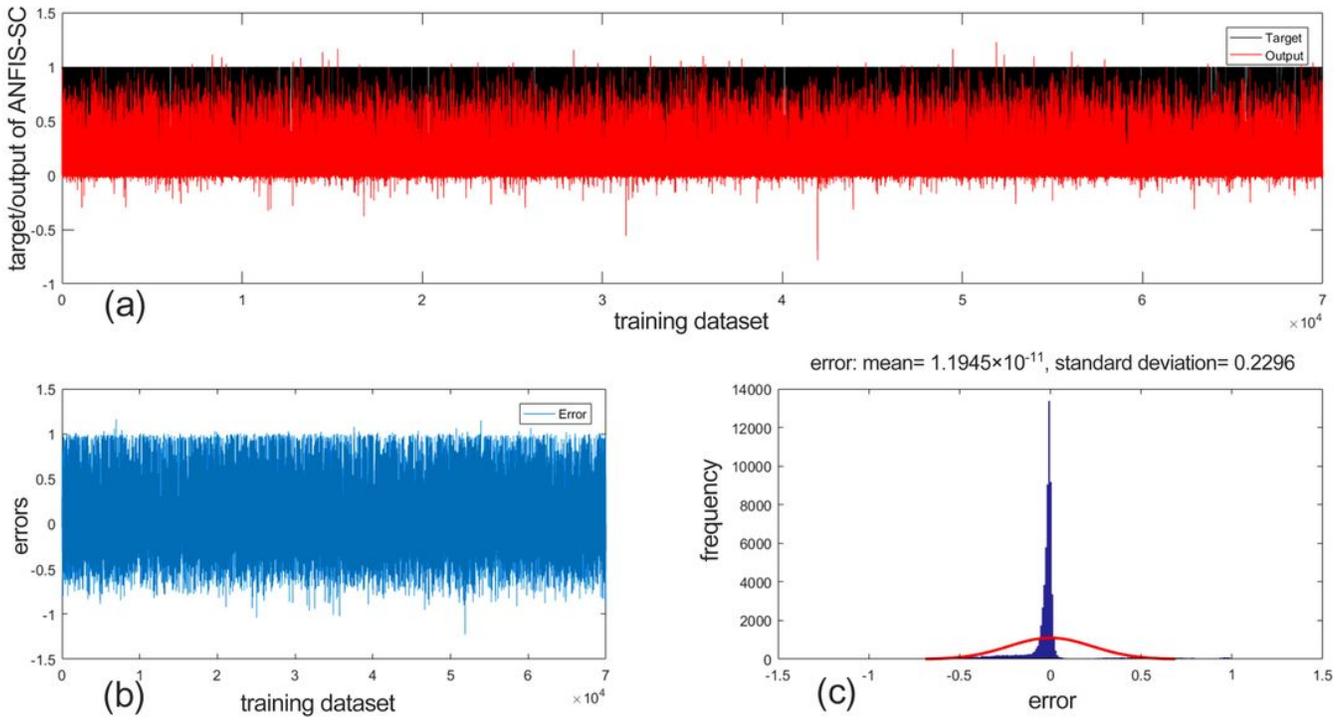


Figure 7

target-output plot, errors plot, and the error histogram chart of ANFIS-SC: a target-output plot, b errors plot, and c error histogram chart

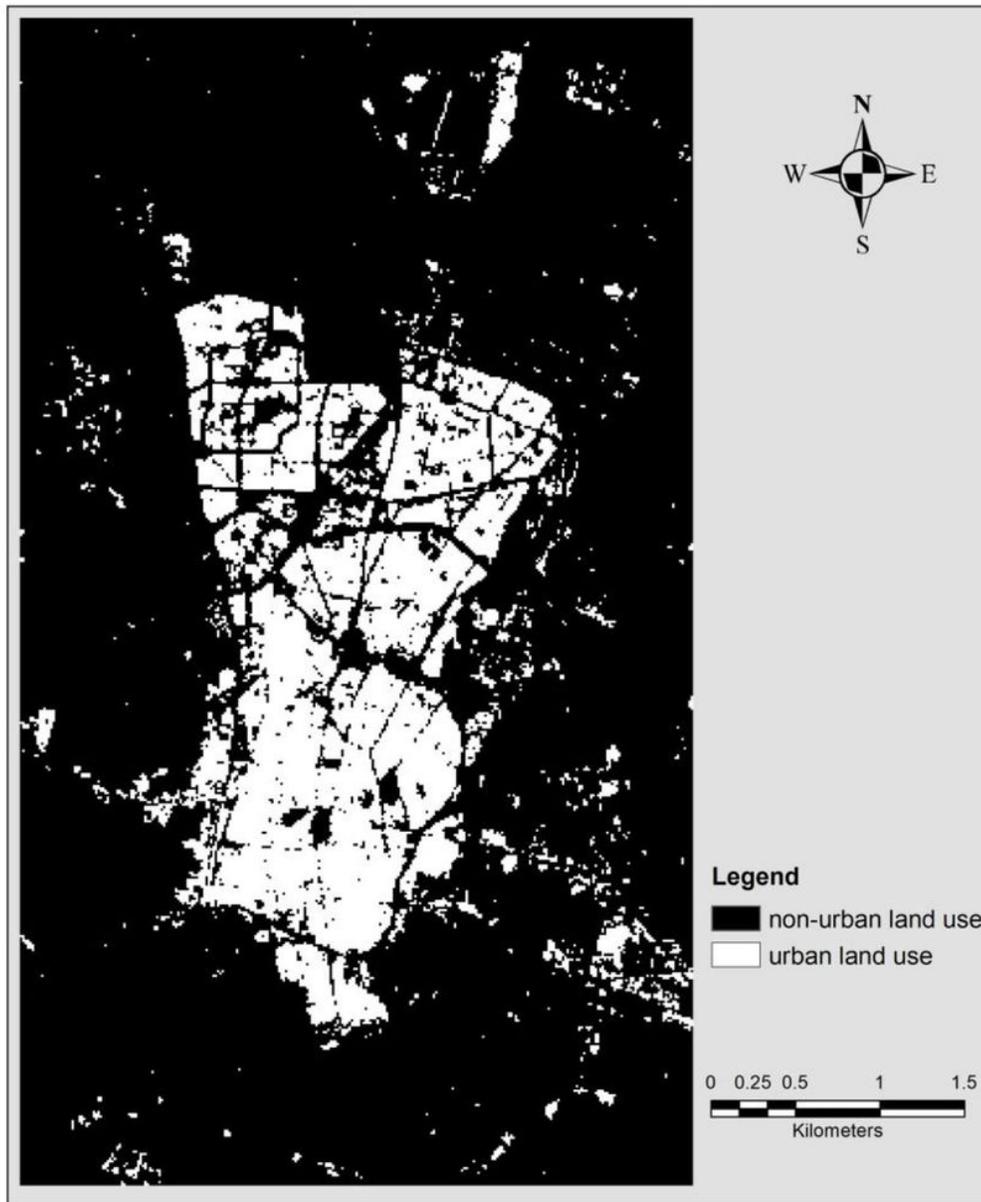


Figure 8

2029 projected map produced by ANFIS-SC

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [graphicalabstract.tif](#)