

A Comparative Study Between Artificial Neural Networks and Fuzzy Inference System for Estimation and Filling of Missing Runoff Data at Al-Jawadiyah Station

Alaa Ali Slieman^{1*}, and Dmitry Kozlov²

¹Moscow State University of Civil Engineering (National Research University), Moscow, Russia

²Moscow State University of Civil Engineering (National Research University), Moscow, Russia

Abstract. Runoff is one of the most important components of the hydrological cycle, and having complete series of runoff data is essential for any hydrological modelling process. This study compares artificial neural networks and fuzzy inference systems for estimation of runoff data at the Al-Jawadiyah hydrometric station. This study used only the runoff data at Al-Jawadiyah station in addition to the runoff values measured at Al-Amiri station on the Syrian-Lebanese border. Many experiments were conducted, and a very large number of artificial neural networks were trained with changing the number of hidden layers, the number of neurons, and the training algorithms until the best network was reached according to the regression criteria and the root mean of the error squares between the measured values and the predicted values. Also, many fuzzy inference models have been prepared, changing the number and type of membership functions until the most accurate model has been reached. The results showed the high reliability of both the artificial neural network models and the fuzzy inference models in estimating runoff in the study area, and the comparison between the results showed the great convergence of the two models with a slight preference for the fuzzy models. This study recommends using the rest of the artificial intelligence models and comparing them to arrive at the most accurate model. This helps prepare a complete series of hydrological and climatic measurements that form a basis for preparing an accurate hydrological model for the study area.

1 Introduction

Water resources are of distinct importance that may exceed other economic resources. They are the cornerstone of the development of various agricultural, industrial, and economic activities, and hydrological information is the basis for assessing the water situation and developing detailed water budget Studies in any region. Surface runoff in arid and semi-arid environments is one of the most important nutritious resources for the waters of rivers,

*Corresponding author: alaa-slieman@hotmail.com

lakes, and streams, which requires decision-makers to follow mathematical formulas and solutions related to calculating surface runoff and its other characteristics [1].

Many researchers have been interested in using artificial intelligence and artificial neural networks in the field of surface runoff estimation, for example (Zhang B. and Govindaraju R. S, 2003) developed an artificial neural network based on the morphological characteristics of the watershed to estimate runoff in spillways, such a geomorphology-based artificial neural network (GANN) is utilized to estimate runoff hydrographs from several storms over two Indiana watersheds. Comparisons of validation results from the GANN model with observed hydrographs over several events for two watersheds are presented and revealed GANNs to be promising tools for estimating direct runoff [2]. Also, (Yazdani M. R, and Saghafian B, 2009) used the ANN model for runoff estimation in the Plasjjan River basin in the central part of Iran. The models used are Multiple Perceptron (MLP) and Recurrent Neural Network (RNN). Inputs include data obtained from 5 rain gauges as well as from 2 temperature recording gauges, the output of the model being the monthly flow in Eskandari Hydrometric Station, and the results indicated that ANN is an appropriate technique for monthly runoff estimation in the selected basin with these networks being also of the capability to show basin response to rainfall events [3]. And (Jimeno-s P. *et al.*, 2018) compared the Soil and Water Assessment Tool (SWAT) and Artificial Neural Network (ANN) models have been evaluated to find a method to improve streamflow estimation; the results indicate that SWAT and ANNs were generally good tools in daily streamflow modelling. However, SWAT was more successful in better simulation of lower flows, while ANNs were superior at estimating higher flows in all cases [4]. Other researchers used artificial neural network models to estimate rainfall-runoff relations, and the results showed the high reliability of the artificial models according to various scenarios [5-8].

Other researchers used the fuzzy inference system to study runoff and compared the results with other methods such as statistical methods or even with artificial neural network models. For example (Mahabir C. F. E. H. and Fayek A. R.,2003) investigated the applicability of fuzzy logic modelling techniques for forecasting water supply. The results showed that the fuzzy logic has a promising potential for providing reliable water supply forecasts [9]. On the other hand, (Tayfur G. and Singh V. P., 2006) presented the development of artificial neural network and fuzzy logic models for predicting event-based rainfall runoff and tests these models against the kinematic wave approximation [10], and (Şen Z. and Altunkaynak A., 2006) performed a study to estimate runoff using fuzzy logic based on the given rainfall measurements [11]. And also (Lohani A. K. *et al.*, 2011) compared artificial neural network (ANN), fuzzy logic (FL), and linear transfer function (LTF)-based approaches for daily rainfall-runoff modelling, and the results show that the fuzzy modelling approach is uniformly outperforming the LTF and also always superior to the ANN-based models [12]. Also (Wang K. H. and Altunkaynak A., 2012) presented a comparative case study between the storm water management model (SWMM) and fuzzy logic model for the predictions of total runoff within the watershed of Cascina Scala, Pavia in Italy. The results showed that the predicted total runoffs from either the SWMM or the fuzzy logic model are found to agree reasonably well with the measured data. For large rainfall events, the fuzzy logic model generally outperforms the SWMM unless the modification of the impervious ratio is applied to improve the SWMM results [13].

In the study area in the Upper Orontes Basin in the Syrian Arab Republic, many studies were also conducted to estimate the runoff and the rainfall-runoff relationship using mathematical and statistical models. Still, the possibilities of artificial intelligence in this field were not used [14,15]. This study compares artificial neural networks and fuzzy inference systems for estimation of runoff data at Al-Jawadiyah hydrometric station using

only the runoff data at Al-Jawadiyah and Al-Amiri hydrometric stations on the Syrian-Lebanese border.

2 Materials and Methods

2.1 Study site and data availability

This study used 266 monthly values of runoff in River Al-Asi at Al-Jawadiyah station (entrance to Lake Qattinah) and al-Amiri station on the Syrian-Lebanese border, as the data series extends from December 1978 to September 2011 with some missing data in the time series.

2.2 Artificial Neural Network

Artificial neural networks are a kind of black box; this means we do not know its structure but just regard its behaviour in practice [16].

These networks consist of neural processing elements called neurons, which act as local memory used in various processing operations, and the figure (1) presents the stages of the work of the artificial neuron.

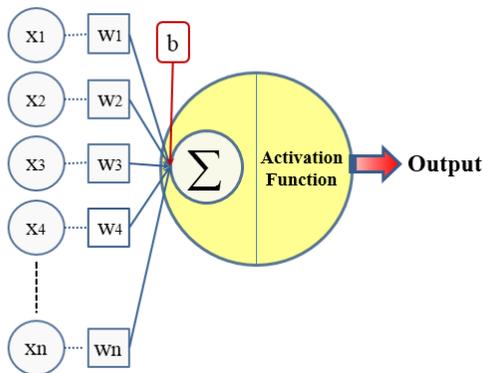


Fig. 1. The figure presents the stages of the work of the artificial neuron.

Feedforward artificial neural networks are one of the most used types of artificial neural networks, as this type of network consists of at least two layers. There is often one or more hidden layers between the input and output layers [6], so named because the transmission of data and the conduct of computational operations are Forward from the input layer to the output layer through the hidden layers of the network. The figure (2) presents A model of feedforward artificial neural network.

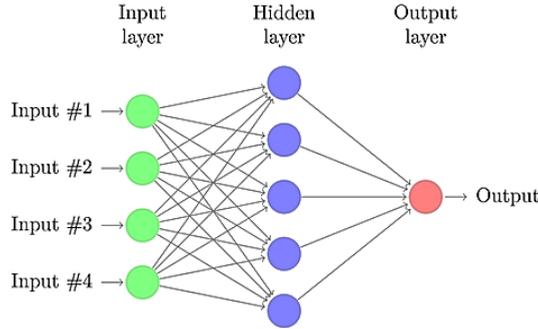


Fig. 2. The figure presents a model of feedforward artificial neural network.

2.3 Data Normalization

In this stage, the field of change of data used in the artificial neural network is standardized, facilitating the training process and speeds up access to the optimal network. In this article, they were scaled between the interval 0 and 1 using the following Eq. (1) [17]:

$$x_{norm} = 0.5 \cdot \left(\frac{x_0 - \bar{x}}{x_{max} - x_{min}} \right) + 0.5 \quad (1)$$

Where x_{norm} is the normalized value, x_0 is the original value, \bar{x} is the average value, x_{max} is the maximum value, x_{min} is the minimum value.

These values are used during the construction phases of the artificial neural network, and then the values produced after the simulation process are returned to their original state before the normalization.

2.4 Fuzzy Inference System

Fuzzy logic starts with the concept of a fuzzy set. Fuzzy set theory has been developed for modelling complex systems in an uncertain and imprecise environment. A fuzzy set is an extension of a classical set whose elements may partially belong to that set [18].

To build a Fuzzy Model, there are three main steps (Fuzzification, Fuzzy Inference Operations, Defuzzification) whereas, in the beginning, the classic sharp inputs are transformed into fuzzy inputs by using a set of Membership Functions (MFs), which can take different shapes such as triangles, trapezoids, bells and (Fuzzification). After that, a set of operations are performed based on a set of (If-Then) fuzzy rules, which are the basis for fuzzy inference operations; these rules are combined with the fuzzy input to derive the fuzzy output of the model (Fuzzy Inference Operations).

And after that, the fuzzy inputs are transformed into classic inputs (Defuzzification) [19],

2.5 Models Evaluation

The comparison between different ANNs models was done by using two statistical indices: correlation coefficient (R) and root mean square error (RMSE), which are defined as Eq. (2) and Eq. (3) [20]:

$$R = \frac{\sum_{i=1}^n (y_i - \bar{y}) \cdot (\hat{y}_i - \bar{\hat{y}})}{\left[\sum_{i=1}^n (y_i - \bar{y})^2 \cdot \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2 \right]^{\frac{1}{2}}} \tag{2}$$

$$RMSE = \left[\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \right]^{0.5} \tag{3}$$

Where n is the number of observations, y_i is the estimated using the artificial neural networks, \hat{y}_i is the observed ET₀ (calculated by the PM method), \bar{y} and $\bar{\hat{y}}$ are the average value for y_i and \hat{y}_i .

3 Results and Discussion

3.1 Artificial Neural Network

In the beginning, a statistical study of the data used was conducted, and then the data was normalized using Equation No. 1, then the data was split into three datasets for training and validation and testing in the ratios 70:15:15 respectively, by using the dividing function (divide block), which maintains the same values in groups and thus increases the accuracy of the comparison process between models.

Many artificial neural networks were built and trained with the change in the number of inputs, the number of neurons in the hidden layer, the activation function, and the training algorithms. The input layer for networks contained two neurons, which are the runoff at Al-Jawadiyah station in time (t-1) and the runoff at the Al-Amiri station in time (t), while the output layer contained one neuron representing the runoff at Al-Jawadiyah station in time (t).

Comparison of models was dependent on correlation coefficient (R), and root mean square error (RMSE), and table (1) presents values of (R), and root mean square error (RMSE) obtained by the best ANNs models.

Table 1. The table presents the correlation coefficient (R) values and root mean square error (RMSE) obtained by the best ANNs models.

	Network architecture	Train		Validation		Test	
		R %	RMSE m ³ /sec	R %	RMSE m ³ /sec	R %	RMSE m ³ /sec
(A)	2-12-1	88.941	1.4833	93.986	0.8383	94.795	0.7331
B	2-18-1	89.022	1.4277	92.494	0.9683	92.311	0.8738
C	2-6-1	88.099	1.5455	93.595	0.9670	94.4158	0.9867
D	2-10-1	87.048	1.5783	91.756	0.9631	93.312	0.8619

As shown in table (1), the network 2.12.1, which contains 12 neurons in the hidden layer, is the best as it gave correlation coefficients (88.941%, 93.986%, 94.795%) during the training, validation, and testing periods, respectively, as well as the root mean square error values during the same periods respectively (1.4833, 0.8383, 0.7331) m³/sec. Figure (3) presents the (2,12,1) ANN model, which depends on the activation function (tan-sigmoid) in both the hidden layer and the output layer. Figure (4) presents the performance of the

(2,12,1) ANN model during the training, validation, and test period and how to stop the training when reaching the lowest value of MSE during the validation period.

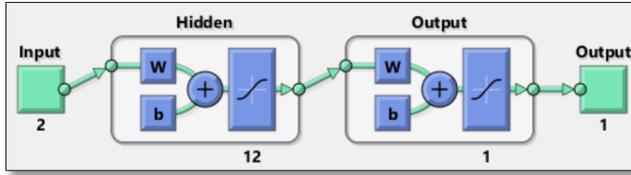


Fig. 3. The figure presents the (2,12,1) ANN model.

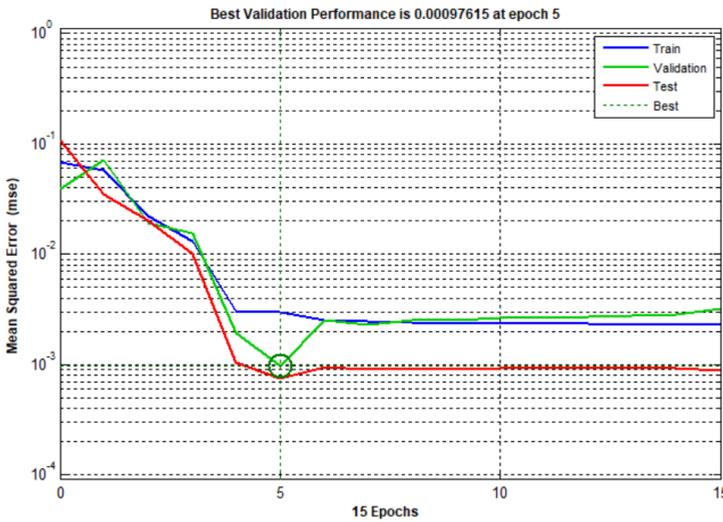


Fig. 4. The figure presents the performance of the (2,12,1) ANN model during the training, validation, and test period.

Figure (5) presents the comparison between real values and values computed with the ANN model during the validation and testing periods, and figure (6) presents the correlation between measured runoff and runoff computed with the ANN model during all periods.

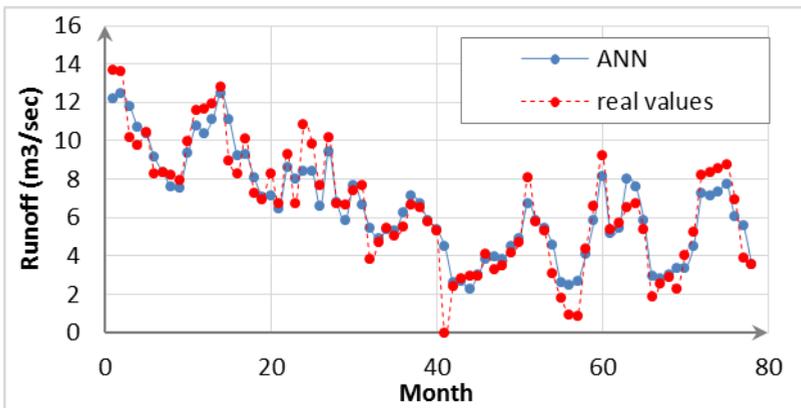


Fig. 5. The figure compares real values and values computed with the ANN model during the validation and testing periods.

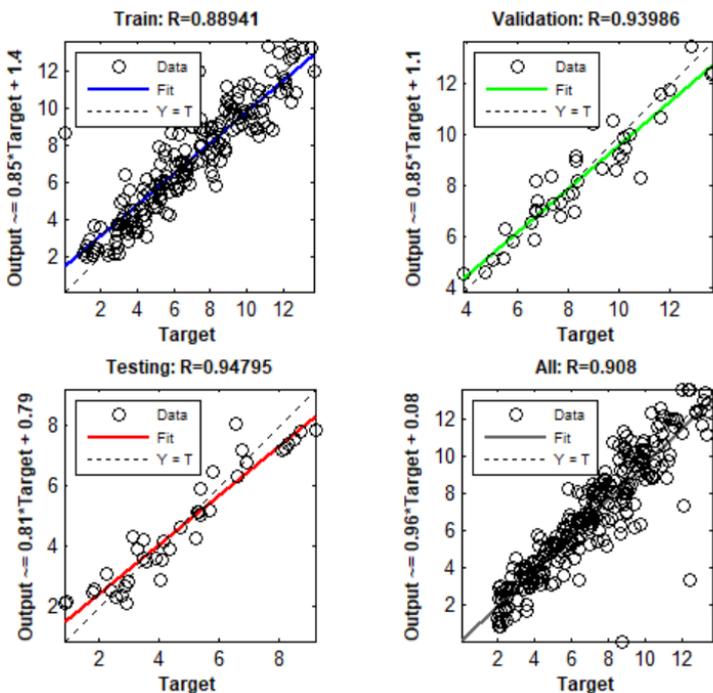


Fig. 6. The figure presents correlation between measured runoff and runoff computed with ANN model during all periods.

3.2 Fuzzy Inference System

Many fuzzy inference models have been prepared, changing the number and type of membership functions until the most accurate model has been reached. Artificial neural networks have been used in the process of building the fuzzy inference system, which greatly facilitates and speeds up the construction process and access to the best model and the table (2) presents values of the correlation coefficient (R), and root mean square error (RMSE) obtained by the best FIS models.

Table 2. The table presents values of the correlation coefficient (R), and root mean square error (RMSE) obtained by the best FIS models.

Number of membership functions	Type of membership functions	Train		Validation		Test	
		R %	RMSE m ³ /sec	R %	RMSE m ³ /sec	R %	RMSE m ³ /sec
(4)	Gauss mf	91.494	1.2368	87.150	1.2770	94.789	0.7791
3	Gauss mf	91.131	1.2617	90.772	1.1222	95.230	0.8022
4	Tri mf	91.488	1.2372	88.336	1.2367	95.027	0.7928
3	Tri mf	90.995	1.2709	90.975	1.110	95.044	0.7813

As shown in table (2), the model which based on four Gauss membership functions is the best as it gave correlation coefficients (91.494%, 87.150%, 94.789%) during the training,

validation, and testing periods, respectively, as well as the root mean square error values during the same periods respectively (1.2368, 1.2770, 0.7791) m³/sec.

Figure (7) presents a structure of the best fuzzy inference model, and figure (8) presents the comparison between real values and values computed with the FIS model during the validation and testing periods, and figure (9) presents the correlation between measured runoff and runoff computed with FIS model during all periods.

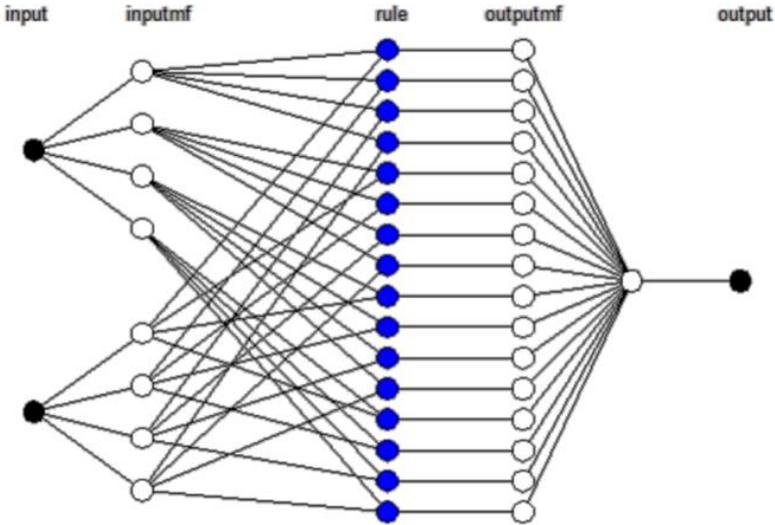


Fig. 7. structure of the best fuzzy inference model.

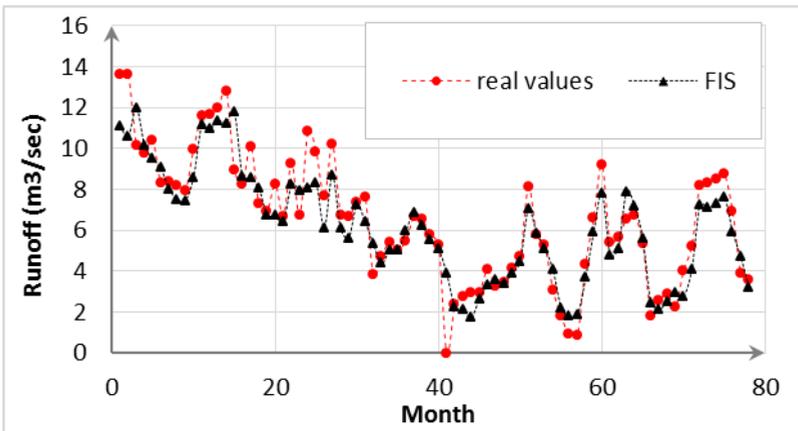


Fig. 8. The figure presents the comparison between real values and values computed with FIS model during the validation and testing periods.

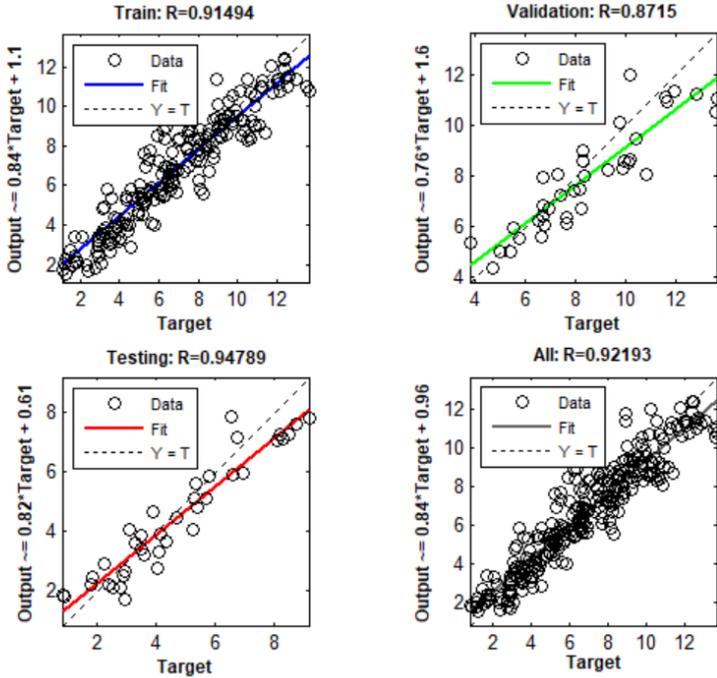


Fig. 9. The figure presents correlation between measured runoff and runoff computed with FIS model during all periods.

3.3 Comparison between ANN & FIS Models

By comparing the results, the high reliability of both artificial neural networks and fuzzy inference models is revealed, with a slight preference for fuzzy models, whether in terms of model evaluation criteria or in terms of speed of access to the optimal model and the figure (10) presents the comparison between real runoff values and values which computed with ANN & FIS model during the validation and testing periods.

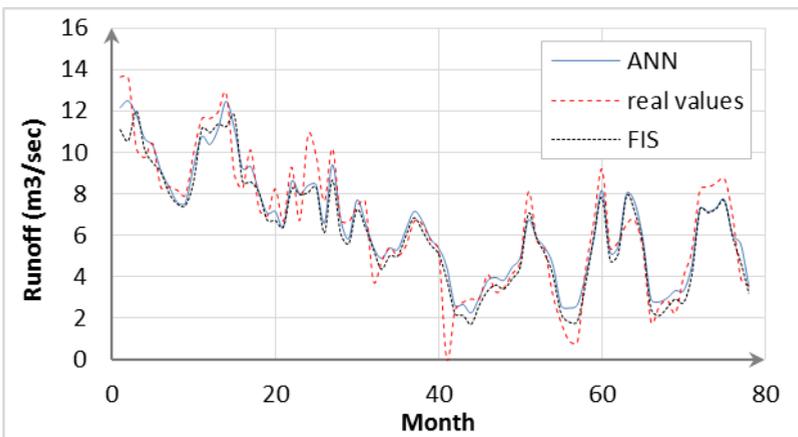


Fig. 10. The figure presents the comparison between real runoff values and values which computed with ANN & FIS model during the validation and testing periods.

In this study, ANNs models were developed for estimating the monthly runoff at Al-Jawadiyah station by using runoff at Al-Jawadiyah station in time (t-1) and runoff at Al-Amiri station in time (t) as inputs. Many artificial neural networks were prepared and trained with the change in the number of inputs, the number of neurons in the hidden layer, the activation function, and the training algorithms. Each neural network was trained with more than 1000 iterative cycles, and the best networks were chosen based on the correlation coefficient, and root mean squares errors, and the results showed that the network 2.12.1, which contains 12 neurons in the hidden layer, is the best as it gave correlation coefficients (88.941%, 93.986%, 94.795%) during the training, validation and testing periods, respectively, as well as the root mean square error values during the same periods respectively (1.4833, 0.8383, 0.7331) m³/sec. Also, many fuzzy inference models have been prepared, changing the number and type of membership functions until the most accurate model has been reached. The model, which based on four Gauss membership functions, is the best as it gave correlation coefficients (91.494%, 87.150%, 94.789%) during the training, validation, and testing periods, respectively, as well as the root mean square error values during the same periods respectively (1.2368, 1.2770, 0.7791) m³/sec, and the comparison between the results showed the great convergence of the two models with a slight preference for the fuzzy models whether in terms of model evaluation criteria or in terms of speed of access to the optimal model.

4 Conclusions

Artificial neural networks and fuzzy inference models have shown the ability to estimate surface runoff values at Al-Jawadiyah station with high reliability, with high values of correlation coefficients and low values of root mean squares of errors with a slight preference for the fuzzy models. This study recommends using the rest of the types of artificial intelligence models and comparing them also to arrive at the most accurate model, as this helps to prepare complete series of hydrological and climatic measurements that form a basis for preparing an accurate hydrological model for the study area.

References

1. Khalaf Hamid D. *Spatial Analysis to Estimate Runoff Using SCS(CN) to Wadi Almur North of Iraq*, Tikrit Journal of Pure Sciences, **21(5)**, pp. 110-121 (2016)
2. Zhang B and Govindaraju R S., *Geomorphology-based artificial neural networks (GANNs) for estimation of direct runoff over watersheds*, Journal of Hydrology, **273**, pp.18–34 (2003)
3. Yazdani M R, Saghafian B, Mahdian M H and Soltani S. *Monthly Runoff Estimation Using Artificial Neural Networks*, J. Agric. Sci. Technol., **11**, pp. 355–362 (2009)
4. Jimeno-s p, Senent-aparicio J and Pulido-velazquez D. *A Comparison of SWAT and ANN Models for Daily Runoff Simulation in Different Climatic Zones of Peninsular Spain*, Water 2018, doi: 10.3390/w10020192 (2018)
5. Srinivasulu S and Jain A. *A comparative analysis of training methods for artificial neural network rainfall-runoff models*, Applied Soft Computing, **6**, pp. 295–306, doi: 10.1016/j.asoc.2005.02.002 (2006)
6. Solaimani K. *Rainfall-runoff Prediction Based on Artificial Neural Network (A Case Study: Jarahi Watershed)*, American-Eurasian J. Agric. & Environ. Sci., **5**, 6, pp. 856–865
7. Chen S M, Wang Y M and Tsou I. *Using artificial neural network approach for*

- modelling rainfall – runoff due to typhoon*, J. Earth Syst. Sci., **2**, pp. 399–405, (2013)
8. Chakravarti A, Joshi N and Panjiar H. *Rainfall Runoff Analysis Using Artificial Neural Network*, Indian Journal of Science and Technology, **8 (14)**, doi: 10.17485/ijst/2015/v8i14/54370, (2019)
 9. Mahabir C, Hicks F E, and Fayek A R. *Application of fuzzy logic to forecast seasonal runoff*, Hydrol. Process., **17 (18)**, pp. 3749–3762, doi: 10.1002/hyp.1359, (2003)
 10. Tayfur G and Singh V P. *ANN and Fuzzy Logic Models for Simulating Event-Based Rainfall-Runoff*, J. Hydraul. Eng., **132 (12)**, pp. 1321–1330, (2006)
 11. Şen Z and Altunkaynak A. *A comparative fuzzy logic approach to runoff coefficient and runoff estimation*. Hydrol. Process. **20 (9)**, pp. 1993–2009, doi: 10.1002/hyp.5992. (2006)
 12. Lohani A K, Goel N K, and Bhatia K K S. *Comparative study of neural network, fuzzy logic and linear transfer function techniques in daily rainfall-runoff modelling under different input domains*, Hydrol. Process., **25 (2)**, pp. 175–193, doi: 10.1002/hyp.7831. (2011)
 13. Wang K H and Altunkaynak A. *Comparative Case Study of Rainfall-Runoff Modeling between SWMM and Fuzzy Logic Approach*. J. Hydrol. Eng., **17 (2)**, pp. 283–291, doi: 10.1061/(asce)he.1943-5584.0000419. (2012)
 14. Issa M. *The relationship between river flow and precipitation in the Orontes Basin*, Damascus University Journal, **31(2)**, (2015)
 15. Hamdan Y, Layos E and Mohammed I. *Identify indicators of climate change through the analysis of the amount of rain on upper basin for Orontes River*, Al-Baath University Journal, **39(43)**, (2017)
 16. Ghahreman N and Sameti M. *Comparison of M5 Model Tree and Artificial Neural Network for Estimating Potential Evapotranspiration in Semi-arid Climates*, DESERT, **1**, pp. 75–81. (2014)
 17. Kumar M, Raghuwanshi N S, Singh R, Wallender W W, and Pruitt W O. *Estimating Evapotranspiration using Artificial Neural Network*, J. Irrig. Drain Eng., **128**, pp. 224–233. (2002)
 18. Kazeminezhad M H and Mousavi S G. *Application of fuzzy inference system in the prediction of wave parameters*, Ocean Engineering, **32**, pp. 1709–1725, doi: 10.1016/j.oceaneng.2005.02.001. (2005)
 19. Al Shalawi S K H. *Comparison of Artificial Neural Network and Fuzzy Logic System applications for estimating pan-evaporation for Mosul region*, Kufa Magazine for Mathematics and Computers, **1(3)**, pp. 23-32, (2011)
 20. Razavi T and Province K. *Estimating of Reference Evapotranspiration by Using Artificial Neural Networks*, International Conference on Transport, Environment and Civil Engineering (ICTECE'2012) August, Kuala Lumpur (Malaysia), pp. 80–84, (2012)