Evaluating Different Machine Learning Models for Runoff Modelling

Alaa Ali Slieman [1], and *Dmitry* V. Kozlov

Moscow State University of Civil Engineering, 26, Yaroslavskoye shosse, Moscow, 129337, Russia

Abstract. Estimation and forecasting of hydrological factors are of particular importance in hydrological modelling, and surface runoff is one of the most important of these factors. Machine learning (ML) models have attracted the attention of researchers in this field. So, this article aims to evaluate several types of ML models such as autoregressive integrated moving average (ARIMA), feed forward back propagation artificial neural network (FFBP-ANN), and adaptive neuro-fuzzy inference system (ANFIS) models in order to estimate runoff values at Al-Jawadiya meteostation in the Orontes River basin in Syria. A large number of ARIMA models were built and the seasonal effect on the models also verified. After that, FFBP-ANN models were used with the change in the number of inputs, the number of hidden layers, and the number of neurons in the hidden layer. Also, a large number of FIS models have been built and artificial neural algorithms have been used in the process of model parameters optimization. The results showed a preference for artificial intelligence models in general over ARIMA models, as well as a slight preference for FFBP-ANN models over ANFIS models. This study recommends expanding the use of ML models to reach the best models for forecasting hydrological factors.

Keywords. Surface runoff, Machine learning, ARIMA, ANN, ANFIS, Estimation.

1 Introduction

The ability to hydrological modelling is of great importance because it helps in answering many practical questions. It also contributes to the process of water resources management and planning, and the ability to estimate and predict hydrological factors [1]. Surface runoff is one of the most important elements of the hydrological cycle and one of the most important aspects of water resource planning [2].

The topic of estimation, prediction and modeling of surface runoff has gained the interest of many researchers, especially using machine learning which is a kind of artificial intelligence that can make accurate predictions by training and testing datasets [3]. Where (Valipour M., 2015) used autoregressive integrated moving average models (ARIMA) to

[1] Corresponding author: alaa-slieman@hotmail.com

model the long-term runoff series in the United States of America, and the results were good, and the results also showed a correlation between seasonal data between 20 years to a quarter of a century [4]. While (Ghanbarpour M.R. *et al.,* 2010) used the ARIMA models to analyze the surface runoff data series in Sangsoorakh karst drainage basin in the Karkheh subbasin of southwest Iran, and the results showed the ability of these models to weekly, monthly and bimonthly flow forecasting applications in the study area [5].

Also (Fereydooni M. *et al.,* 2012) compared between the artificial neural network models and the ARIMA models in predicting the monthly runoff in Ghara –Aghaj River in the southwest of Iran, and the results showed the preference of the artificial neural networks and their reliability in the estimation and prediction process [6], and (Chen et al., 2013) used feed forward back propagation artificial neural networks (FFBP ANN) and conventional regression analysis (CRA) to model surface runoff using rainfall data, and fill in the gaps in the timeseries, and comparison results showed the preference of artificial neural network models [7].

On the other hand (Lohani A.K. et al., 2011) compared artificial neural network (ANN), fuzzy logic (FL) and linear transfer function (LTF) 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 [8]. Likewise, (Tayfur G. and Singh V. P., 2006) demonstrated the preference of artificial neural network models and fuzzy logic models against kinematic wave approximation (KWA) in predicting event-based rainfall runoff [9].

While no research was conducted in the Upper Orontes Basin region in Syria using machine learning models. So, this article aims to evaluate several types of ML models such as autoregressive integrated moving average (ARIMA), feed forward back propagation artificial neural network (FFBP ANN), and adaptive neuro-fuzzy inference system (ANFIS) models in order to estimate runoff values at Al-Jawadiya meteostation in the Orontes River basin in Syria.

2 Methods

2.1 Study site & data availability

The study area in this article is the Upper Orontes River Basin in the Syrian Arab Republic, which is located between the Syrian-Lebanese border and Lake Qatina, and the surface runoff data were used in it at Al- Jawadiya meteostation at the entrance to Lake Qatina and Al-Amiri meteostation on the Syrian-Lebanese border.

2.2 Autoregressive integrated moving average (ARIMA)

The Box-Jenkins ARIMA methodology is one of the most popular applications of time series analysis and forecasting. It originates from the autoregressive model (AR), the moving average model (MA), and the combination of autoregressive and moving average models (ARMA) [10].

In this study, ARIMA (p,d,q) models were used, which depend on the moving average model, number of differences, autoregressive model, which are p,d,q respectively. SARIMA (p,d,q) (P,D,Q)s models were also used, which are used in addition to the seasonal components of the autoregressive, differences and moving average, and s is the periodic term [11].

2.3 Feed forward back propagation artificial neural network (FFBP ANN)

Artificial neural networks are considered one of the most important models of artificial intelligence, which reflect great interest in the human way of thinking. Artificial neural networks consist of simple, parallel components called neurons. These items are inspired by the biological nervous system [12]. Artificial neural networks generally consist of an input layer, an output layer, in addition to one or more hidden layers, through which the deep learning process and the link between the input and output layers [13]. Feed forward back propagation artificial neural network FFBP ANN is the most commonly used ANN approach in hydrological predictions and in approximating nonlinear functions. The FFBP is a supervised learning technique used for training artificial neural network, and it is a gradient descent technique used to reduce the error criteria in the network because of the method used in the training process [7]. In this article, the data were standardized within the range between 0 and 1, which facilitates the process of training the model and speeds up access to the best results [14].

2.4 Adaptive neuro-fuzzy inference system (ANFIS)

The fuzzy logic theory is widely used to simulate ambiguity and uncertainty in engineering problems. Using of artificial neural networks in fuzzy models increases the capabilities of these models and their ability to solve engineering problems [15]. The fuzzy model is built according to the following three main stages (Fuzzification, Fuzzy Inference Operations, Defuzzification) [16]. Whereas, each of the inputs and outputs of the fuzzy model must be transformed from the classical form to the fuzzy form using the membership functions as a first stage, and then the fuzzy inference process takes place through the fuzzy rules (if-then) to derive the fuzzy outputs, and those outputs are arranged in a table called the search table where the core of the fuzzy inference processes [17].

3 Results

3.1 Autoregressive integrated moving average (ARIMA)

In the beginning, we searched for the longest time series in the Al-Jawadia station without any missing values, because this is the basic condition in ARIMA models, as the ARIMA model cannot be built on a time series that contains missing values. Therefore, the monthly flow data was used from February 1989 until October 2008, 237 monthly values.

Then the stability of this series was verified in the mean and the standard deviation, and an attempt was made to increase the stability by making several changes such as the first and second differentiation, the square root, and the squared, but these changes did not improve the stability of the series, so the existing basic data was used.

A large number of ARIMA models were prepared with the change in the moving average, autoregressive components and differences. The table 1 shows the results of the best models that have been reached.

Table 1. Root mean square error (RMSE) obtained by the best ARIMA models.

Model	RMSE (m^3/sec)
(1,0,1)	**1.851**
(1,1,1)	1.898

(2,1,2)	1.872
(3,1,2)	1.881
(4,1,1)	1.853

As shown in the table, the models gave similar results, and the model (1,0,1) was the best, as it gave a root mean square error value equal to 1.851 m^3/sec.

After that, the effect of adding seasonality on ARIMA models was investigated, as a number of ARIMA models were prepared by introducing seasonal components, and the table shows the best models that have been reached.

Table 2. Root mean square error (RMSE) obtained by the best SARIMA models.

Model	RMSE (m^3/sec)
(2,1,1) (0,1,2)12	1.747
(2,1,2) (1,1,1)12	1.758
(3,1,2) (0,1,1)12	1.738
(4,1,1) (0,1,1)12	**1.726**
(4,1,1) (0,1,2)12	1.745

As shown in the table, adding the effect of seasonality led to a slight improvement in the results, and the model (4,1,1) $(0,1,1)_{12}$ is the best. It gave a root mean square error value equal to 1.726 m^3/sec. The figure 1 shows a comparison between the measured values and the values generated by the ARIMA model.

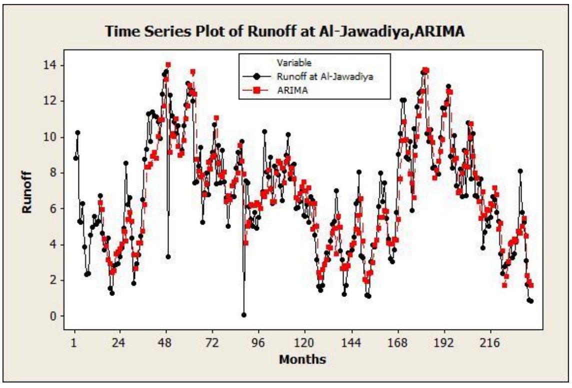

Fig. 1. Comparison between the measured values and the values generated by the ARIMA model.

3.2 Feed forward back propagation artificial neural network (FFBP ANN)

In this model, we were able to use a longer data series consisting of 266 months and containing missing values of surface runoff at Al-Jawadiya station at the entrance to Lake Qattina and Al- Amiri station on the Syrian-Lebanese border, and this data was divided into three groups for training, validation and testing according to the following ratios: 70% for the training dataset, 15% for the validation dataset, and 15% for the test dataset, where a large number of artificial neural network models were built using the feed-forward back-propagation algorithm with changing the number of neurons in the input layer, the input layer, changing the number of hidden layers and the number of neurons, as well as activation functions.

The results showed that the best number of neurons in the input layer is two. It is the surface runoff at the Amiri station at time t and the surface runoff at the Al-Jawadiyah station from time t-1, while the surface runoff at the Al-Jawadiya station at time t is the neuron in the output layer. Table 3 shows the results of the best models that have been reached.

Table 3. Results of the best obtained FFBP ANN 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 the table, the network 2-12-1 is the best, which contains 12 neurons in the hidden layer, and it gives a root mean square error value of 1.483 m^3/sec for the train dataset, 0.838 m^3/sec for the validation dataset and 0.733 m^3/sec for the test dataset. Figure 2 shows a comparison between the values measured at Al-Jawadiyah station and the values generated by FFBP ANN model during the verification and testing datasets.

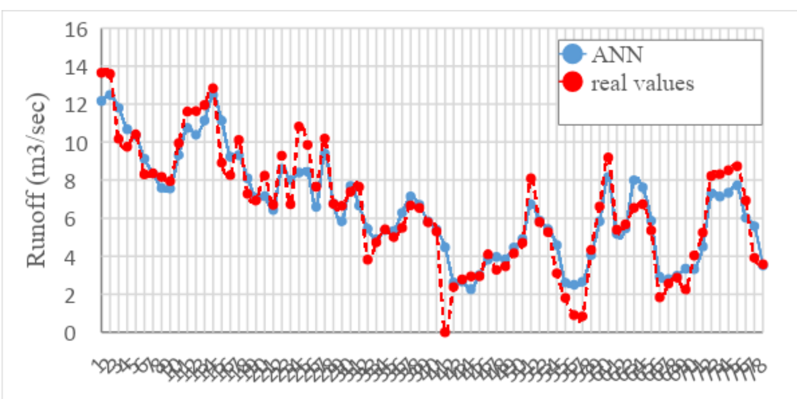

Fig. 2. comparison between the values measured at Al-Jawadiyah station and the values generated by FFBP ANN model during the verification and testing datasets.

3.3 Adaptive neuro-fuzzy inference system (ANFIS)

For the hybrid fuzzy neural models, the same data used for the artificial neural networks in the previous paragraph were used, using the same method of dividing into three groups for training, verification and testing, and with the same proportions of division, for the ease and reliability of the comparison between the results.

A large number of fuzzy models were built with the change in the model's parameters, such as the change in the number and type of membership functions and fuzzing methods. Artificial neural networks were also used in the process of improving the model's parameters and accelerating access to the best structure for it. The table 4 shows the best models obtained during the training and verification and test datasets.

Table 4. Root mean square error (RMSE) obtained by the best ANFIS 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 the table, the model that contains four Gaussian membership functions is the best, and it gives a root mean square error value of 1.237 m^3/sec for the train dataset, 1.277 m^3/sec for the validation dataset and 0.779 m^3/sec for the test dataset. And the figure 3 shows comparison between the measured values and the estimated values of surface runoff using ANFIS model during the verification and testing periods.

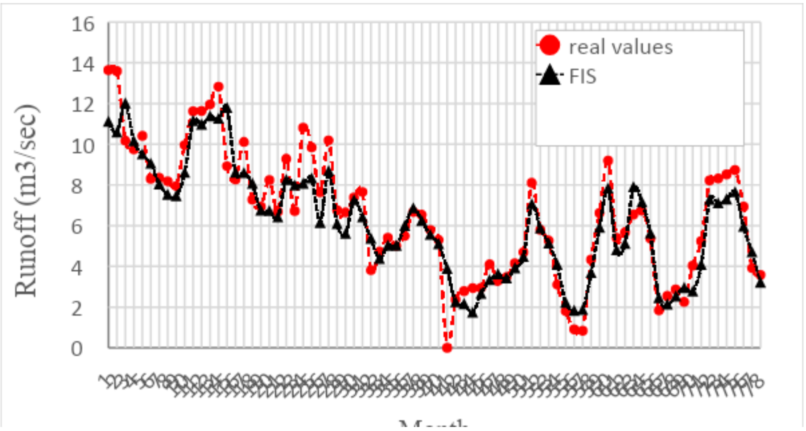

Figure 3. The figure presents comparison between the measured values and the estimated values of surface runoff using ANFIS model during the verification and testing periods.

4 Discussion

A large number of ARIMA models were built with changing the components of the moving average, auto regressive, and complementarity. The effect of seasonality was also tested, and the results showed that the model (4,1,1) (0,1,1)12 is the best. It gave the best results with root mean square errors equal to 1.726 m^3/sec. It takes into account the seasonality of 12 months. Also, a large number of artificial neural network models were also built with changing training algorithms, activation functions, the number of hidden layers, and the number of neurons in them. The results showed that the model 2-12-1 is the best, which gives a very value of mean square errors equal to 0.733 m^3/sec for the test dataset. Likewise, a large number of hybrid neural fuzzy logic models were built with the change in the model parameters, the number and shape of membership functions, and artificial neural networks are used in the training process. The results showed that the model which contains four Gaussian membership functions is the best, and it gives a root mean square error value of 0.779 m^3/sec for the test dataset. In comparison, we find Artificial neural networks and fuzzy models are clearly better than ARIMA models, as shown by the results of convergence between the results of artificial neural networks and hybrid fuzzy models with a slight advantage of artificial neural networks. Also, it should be noted the preference of artificial intelligence models such as artificial neural networks and hybrid fuzzy models over ARIMA models in terms of the possibility of building the model even if there are gaps in the time series, and this is not possible in ARIMA models.

5 Conclusions

In this study, different types of machine learning models were used to predict surface runoff at Al-Jawadiya meteostation in the upper Orontes River basin in the Syrian Arab Republic, where Autoregressive integrated moving average (ARIMA) models, feed-forward back propagation artificial neural networks (FFBP ANN) models and Adaptive neuro-fuzzy inference system (ANFIS) models were used, and the results showed the preference of neural networks artificial and fuzzy models over ARIMA models in general, while the comparison results also showed a slight preference for artificial neural networks over fuzzy models. This study recommends expanding the use of machine learning models in the hydrological modeling process, as well as in predicting and estimating the various elements of the hydrological cycle.

References

1. J. Seibert, and M. J. P. Vis, *Teaching hydrological modeling with a user-friendly catchment-runoff-model software package*, Hydrol. Earth Syst. Sci., 16, 3315–3325, 2012, doi:10.5194/hess-16-3315-2012
2. A. Kumar, P. Kumar, V. K. Singh, *Evaluating Different Machine Learning Models for Runoff and Suspended Sediment Simulation*, Water Resour Manage 33, 1217–1231, 2019. https://doi.org/10.1007/s11269-018-2178-z
3. A. Bhusal, U. Parajuli, S. Regmi, A. Kalra, *Application of Machine Learning and Process-Based Models for Rainfall-Runoff Simulation in DuPage River Basin, Illinois*. Hydrology 2022, 9, 117. https://doi.org/10.3390/hydrology9070117
4. M. Valipour, *Long-term runoff study using SARIMA and ARIMA models in the United States*. Meteorol. Appl. 22: 592–598 (2015), DOI: 10.1002/met.1491
5. M.R. Ghanbarpour, K.C. Abbaspour, G. Jalalvand, and G.A. Moghaddam – *Stochastic modeling of surface stream flow at different time scales: Sangsoorakh karst basin, Iran*. Journal of Cave and Karst Studies, v. 72, no. 1, p. 1–10. (2010) DOI: 10.4311/jcks2007ES0017
6. M. Fereydooni, M. Rahnemaei, H. Babazadeh, H. Sedghi, M.R. Elhami, *Comparison of artificial neural networks and stochastic models in river discharge forecasting, (Case study:Ghara- Aghaj River, Fars Province, Iran),* African Journal of Agricultural Research, Vol. 7(40), pp. 5446-5458, 23 October, (2012). DOI: 10.5897/AJAR11.1091
7. S. M. Chen, Y. M. Wang, I. Tsou, *Using artificial neural network approach for modelling rainfall–runoff due to typhoon*, J. Earth Syst. Sci. 122, No. 2, pp. 399–405, 2013.
8. A. K. Lohani, N. K. Goel, and K. K. S. Bhatia, *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), 175–193, 2011, doi:10.1002/hyp.7831.

9. G. Tayfur and V. P. Singh, *ANN and Fuzzy Logic Models for Simulating Event-Based Rainfall-Runoff*, J. Hydraul. Eng., 132 (12), 1321–1330, 2006.
10. W. Wang, K. Chau, D. Xu, X. Chen, *Improving Forecasting Accuracy of Annual Runoff Time Series Using ARIMA Based on EEMD Decomposition*, Water Resour Manage, 29:2655–2675, 2015, DOI 10.1007/s11269-015-0962-6.
11. M. Valipour, Long-term runoff study using SARIMA and ARIMA models in the United States. Meteorol. Appl. 22: 592–598 (2015), DOI: 10.1002/met.1491
12. N. Ghahreman, M. Sameti, *Comparison of M5 Model Tree and Artificial Neural Network for Estimating Potential Evapotranspiration in Semi-arid Climates*, DESERT, 1, 75–81, 2014.
13. K. Solaimani, Rainfall-runoff Prediction Based on Artificial Neural Network (A Case Study: Jarahi Watershed), American-Eurasian J. Agric. & Environ. Sci., 5, 6, 856–865, 2009.
14. M. Kumar; N. S. Raghuwanshi, R. Singh, W. W. Wallender, W. O. Pruitt, *Estimating Evapotranspiration using Artificial Neural Network*, J. Irrig. Drain Eng., 128, 224–233, 2002.
15. J. Behmanesh, S. Ayashm, *Rainfall-runoff modeling in the Turkey River using numerical and regression methods*, Journal of Fundamental and Applied Sciences, Vol. 7 No. 1, 91-102, 2015, DOI:10.4314/jfas.v7i1.8.
16. M. H. Kazeminezhad, S. G. Mousavi, *Application of fuzzy inference system in the prediction of wave parameters*, Ocean Engineering, 32, 1709–1725, 2005, doi: 10.1016/j.oceaneng.2005.02.001.
17. S. K. H. Al Shalawi, *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), 23-32, 2011.