Estimation and filling of missing runoff data at Al-Jawadiyah station using artificial neural networks

Alaa Ali slieman[*], and *Dmitry* Kozlov

Moscow State University of Civil Engineering, Yaroslavskoe shosse, 26, Moscow, 129337, 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 aims to estimate the runoff at Al-Jawadiyah hydrometric station using artificial neural networks. 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, and the network (2:12:1) was adopted in the process of filling the gaps in the runoff time series at Al-Jawadiyah station during the study period. This study recommends working on preparing 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, as 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, 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

---

* Corresponding author: alaa-slieman@hotmail.com

Saghafian B, 2009) used ANN model for runoff estimation in Plaszjan 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 between 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 found to be more successful in relation to 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 of rainfall – runoff relation, and the results showed the high reliability of the artificial models according to various scenarios [5-8].

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, but the possibilities of artificial intelligence in this field were not used [9,10]. This study aims to estimate the runoff at Al-Jawadiyah hydrometric station using artificial neural networks using only the runoff data at Al-Jawadiyah and Al-Amiri hydrometric stations on the Syrian-Lebanese border.

2 Methods

2.1 Study site & data availability

In this study were 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 [11].

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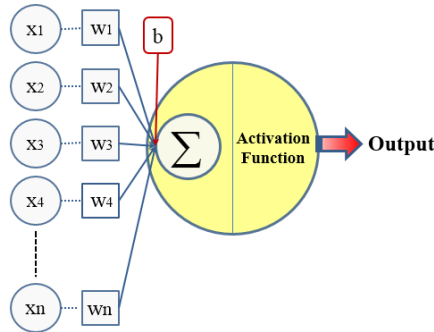
Fig. 1. The figure presents the stages of the work of the artificial neuron.

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, and 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, and the figure (2) presents A model of feedforward artificial neural network.
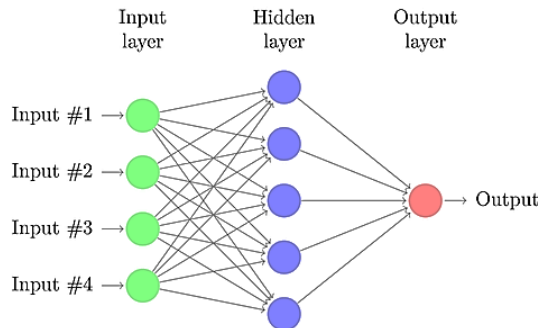

Fig. 2. The figure presents A model of feedforward artificial neural network.

2.3 Data Normalization

In this stage, the field of change of data used in the artificial neural network is standardized, which facilitates 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. (2) [12]:

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

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 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) [13]:

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

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

Where n is the number of observations, y_i is the estimated using the artificial neural networks, $\hat{y}_i$ is the observed ET_0 (calculated by the PM method), $\bar{y}$ and $\bar{\hat{y}}$ are the average value for y_i and $\hat{y}_i$.

3 Results

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.

A large number of 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 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 the table (1) presents values of (R) and root mean square error (RMSE) obtained by the best ANNs models.

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

	Network architecture	Train		Validation		Test	
		R %	RMSE m³/sec	R %	RMSE m³/sec	R %	RMSE m³/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 (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. The 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. The 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 to the lowest value of MSE during the validation period.
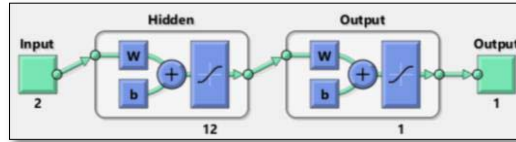

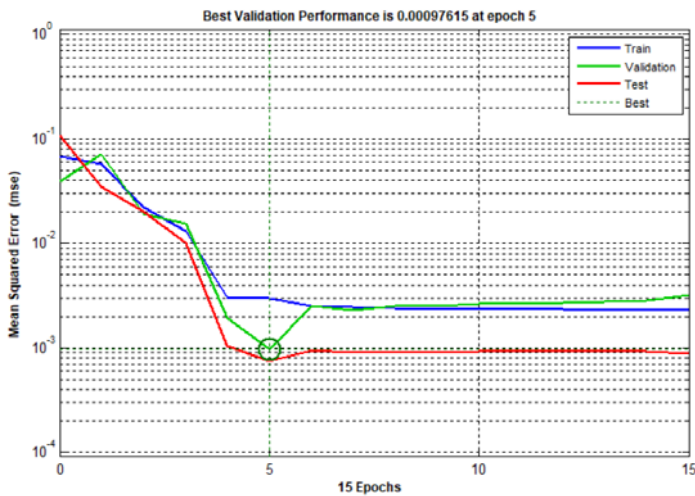**Fig. 3.** The figure presents the (2,12,1) ANN model.

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

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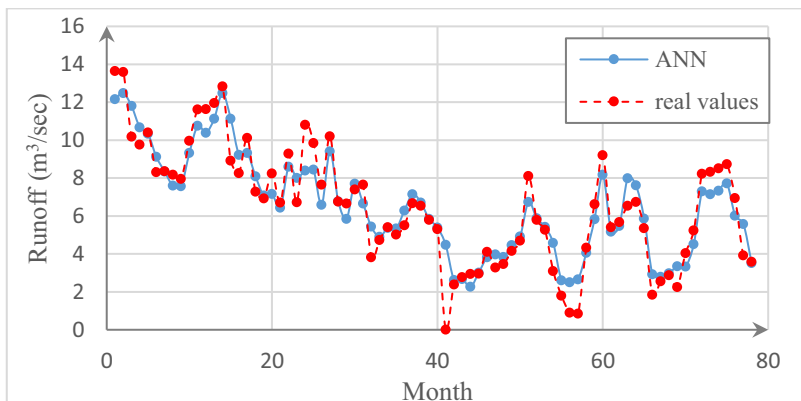

Fig. 5. The figure presents comparison between estimated and values computed with ANN model during the validation and testing periods.
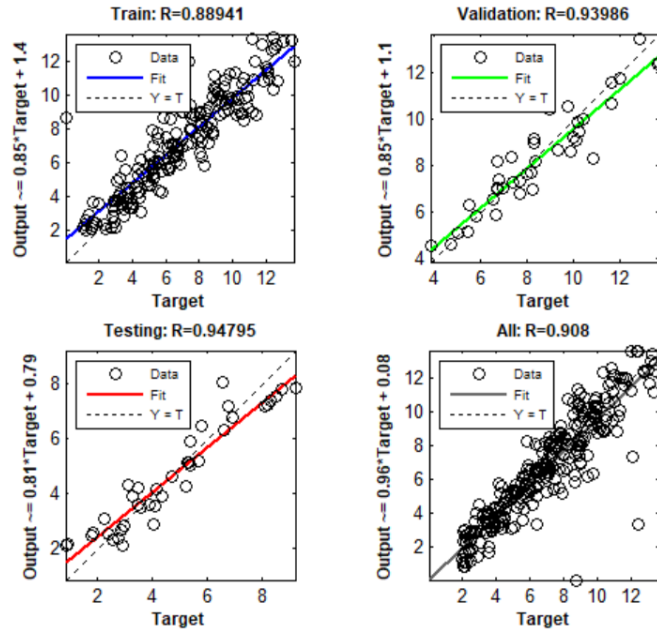
Fig. 6. The figure presents correlation between measured runoff and runoff computed with ANN model during all periods.

4 Discussion

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. A large number of 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) m^3/sec.

5 Conclusions

Artificial neural networks have shown the ability to estimate the values of surface runoff at Al-Jawadiyah station with high reliability, with high values of correlation coefficients and low values of root mean squares of errors, and it was found that feedforward artificial neural networks gave the best and fastest results in reaching the optimal model.

This study recommends expanding the use of artificial intelligence models and a comparison between them to find accurate estimation and prediction models for missing data in time series, which forms the basis for preparing hydrological models and studying the water balance in watersheds.

References

1. D. Khalaf Hamid, *Spatial Analysis to Estimate Runoff Using SCS(CN) to Wadi Almur North of Iraq*, Tikrit Journal of Pure Sciences, **21(5)**, 110-121, (2016).

2. B. Zhang and R. S. Govindaraju, *Geomorphology-based artificial neural networks (GANNs) for estimation of direct runoff over watersheds*, Journal of Hydrology, **273**, 18–34, (2003).

3. M. R. Yazdani, B. Saghafian, M. H. Mahdian, and S. Soltani, *Monthly Runoff Estimation Using Artificial Neural Networks*, J. Agric. Sci. Technol., **11**, 355–362, (2009).

4. P. Jimeno-s, J. Senent-aparicio, and D. Pulido-velazquez, *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. S. Srinivasulu and A. Jain, *A comparative analysis of training methods for artificial neural network rainfall – runoff models*, Applied Soft Computing, **6**, 295–306, doi: 10.1016/j.asoc.2005.02.002, (2006).

6. 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).

7. S. M. Chen, Y. M. Wang, and I. Tsou, *Using artificial neural network approach for modelling rainfall – runoff due to typhoon*, J. Earth Syst. Sci., **2**, 399–405, (2013).

8. A. Chakravarti, N. Joshi, and H. Panjiar, *Rainfall Runoff Analysis Using Artificial Neural Network*, Indian Journal of Science and Technology, **8(14)**, doi: 10.17485/ijst/2015/v8i14/54370, (2019).

9. M. Issa, *The relationship between river flow and precipitation in the Orontes Basin*, Damascus University Journal, **31(2)**, (2015).

10. Y. Hamdan, E. Layos and I. Mohammed, *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).

11. N. Ghahreman and M. Sameti, *Comparison of M5 Model Tree and Artificial Neural Network for Estimating Potential Evapotranspiration in Semi-arid Climates*, DESERT, **1**, 75–81, (2014).

12. M. Kumar, N. S. Raghuwanshi, R. Singh, W. W. Wallender, and W. O. Pruitt, *Estimating Evapotranspiration using Artificial Neural Network*, J. Irrig. Drain Eng., **128**, 224–233, (2002).

13. T. Razavi and K. Province, *Estimating of Reference Evapotranspiration by Using Artificial Neural Networks*, International Conference on Transport, Environment and Civil Engineering (ICTECE'2012) August, Kuala Lumpur (Malaysia), 80–84, (2012).