ORIGINAL ARTICLE

Comparison of classic time series and artificial intelligence models, various Holt-Winters hybrid models in predicting the monthly flow discharge in Marun dam reservoir

Abbas Ahmadpour[1] · Parviz Haghighat Jou[1] · Seyed Hassan Mirhashemi[1]

Abstract

In this study, the data at Idenak hydrometric station were used to predict the inflow to Maroun Dam reservoir. For this purpose, different models such as artificial intelligence, Holt-Winters and hybrid models were used. Partial mutual information algorithm was used to determine the input parameters affecting modeling the monthly inflow by artificial intelligence models. According to the Hempel and Akaike information criterion, we introduced the monthly inflow with a 3-month lag, and the temperature with a 1-month lag, with respect to the lowest values of Akaike and the highest values of Hempel as input parameters of artificial intelligence models. The results showed the weak performance of the Holt-Winters model compared to other models and confirmed the superiority of the Holt-adaptive network-based fuzzy inference system (ANFIS) hybrid model with the root-mean-square error of 54 and the coefficient of determination (R^2) of 0.83 in the testing process compared to other mentioned models. In addition, the above hybrid models performed better than other models in the test process.

Keywords Artificial intelligence · Neural network model · Holt-Winters and hybrid models · Maroun basin

Introduction

Modeling hydrological factors, such as runoff, has a great effect on reducing droughts and managing water resources (Nourani et al. 2011). Several models have been developed to simulate complex hydrological processes. Intelligent models such as ANFIS and neural network have shown an almost acceptable ability to model and forecast nonlinear hydrological time series (Nourani et al. 2009). Rainfall-runoff models have been used since the late nineteenth century; besides, there are currently several hydrological models to simulate the rainfall-runoff process (Wu and Chau 2011). Rainfall-runoff models include conceptual, physical, empirical, and artificial intelligence models (Jothiprakash and Magar 2009). Many studies have been conducted on the effect of the runoff on rivers and the resulting problems (Zhou et al. 2015; Liu et al. 2014). Approximately 40% of the natural hazards are related to floods, in addition 20–300 million people are affected by the problems caused by floods yearly (Dewan 2013). The flood in Pakistan caused many deaths and financial losses to the agriculture and infrastructure in 2010 (Ghumman et al. 2011).

Data driven models (DDMs) are among the modern methods to model hydrological parameters and the widely used methods by researchers (Kisi et al. 2019). The DDMs are used to model hydrological processes such as linear and multilinear regression (LMR), due to the high accuracy (Abdulelah Al-Sudani et al. 2019). The artificial neural network (ANN) and autoregressive integrated moving average (ARIMA), monthly flow modeling using ANFIS and ANN–ARIMA using snow telemetry data were performed in Elephant Butte reservoir in Mexico city (Zamani Sabzi et al.2017). Linear time series models are widely used to model hydrological time series (Mohammadi et al. 2006). Also, nonlinear time series approaches can be used for hydrological modeling (Xie et al. 2016).

✉ Seyed Hassan Mirhashemi
hassan.mirhashemi@yahoo.com

Abbas Ahmadpour
Aahmadpour67@yahoo.com

Parviz Haghighat Jou
parvizhjou@uoz.ac.ir

1   Department of Water Engineering, Faculty of Water and Soil, University of Zabol, Zabol, Iran

The study of Altunkaynak (2007) showed that ANN can well model the relationship between water and rainfall levels. The study of Nayak et al. (2005) showed that ANFIS performed better than ANN and ARIMA in estimating the discharge of the Bantarani river in Odisha, India. ANFIS was used by Nourani and Komasi (2013) to simulate the rainfall-runoff process of Eel river in California. One of the most important factors affecting water resources planning and management of the basin, is the response of the basin to rainfall-runoff versus infiltration and evaporation (Uhlenbrook et al. 2004). The innovation of this study is using PMI algorithm to introduce effective input parameters in forecasting and modeling the monthly flow of the Maroon basin with artificial intelligence models (ANFIS–support vector machines–and artificial neural network). This algorithm has been rarely used by previous researchers for this purpose. In addition, another innovation of this research is using different Holt-Winters hybrid models and the comparison of their performances with monthly classical artificial intelligence models. Therefore, the objective of this study is to compare the performance of classical artificial intelligence models with Holt-Winters hybrid models to predict monthly inflow.

Materials and methods

Study area

Maroun basin with an area of about 3824 km^2 is located in the geographical coordinates of 49°50′ to 51°10′ east longitude, 30°30′ to 31°20′ north latitude and the heights of Behbahan city of Khuzestan province of Iran. Maroun basin is surrounded by the basins of the Zohreh and Karoun rivers in Khuzestan and Kohgiluyeh and Boyer Ahmad provinces. The climate of Maroun basin is affected by the low latitude, changes in altitude in different areas (from zero to over 3600 m above sea level) and finally the proximity to the Persian Gulf in its southern parts. The mean annual rainfall in the Maroun basin varies from about 150 mm in the low coastal plains to about 900 mm in the northern highlands, and the regime of such rainfall is Mediterranean. The maximum rainfall occurs between December and March. The dry season of the basin is long, and in lower areas lasts from May to late October. Most of the catchment area of Maroun river is made up of mountainous area. The northern and eastern parts are higher than the other parts.

Support vector machine

If the data is discrete, support vector machine (SVM) trains linear machines to produce an optimum level that separates the data without error and with the maximum distance between the plane and the nearest training points (support vectors). If we define the training points as $[X_i, V_i]$ and the input vector $X_i \in R^n$ and the class value $y_i \in \{-1, 1\}$, $i = 1, \ldots, N$ in case of linearly separable data, the decision rules based on Eq. (1) are defined in such a way that they separate the binary decision classes by an optimal plane:

$$y = \text{sign}\left(\sum_{i=1}^{N} y_i a_i(X.X) + b\right) \tag{1}$$

where Y is the output of the equation, y_i is the class value of the training sample X_i and represents the internal multiplication.

The vector $x = (X_1, X_2, \ldots, X_n)$ represents an input data and the vectors $i = 1, \ldots, N$, X_i are support vectors. In Eq. (1), the parameters a_i and b determine the super-plane. If the data are not linearly separable, Eq. (1) is transformed to Eq. (2).

$$y = \text{sign}\left(\sum_{i=1}^{N} y_i a_i K(X.X) + b\right) \tag{2}$$

The function $K(X, X_i)$ is a kernel function that generates internal multiplications to create machines with different types of nonlinear decision levels in the data space (Vapnik 2000).

Adaptive network fuzzy inference system

Adaptive network fuzzy inference system (ANFIS) uses available input–output data pairs during training. Then, IF–THEN fuzzy rules are obtained to connect these parts to each other. ANFIS training means determining the parameters belonging to two parts, introduction and consequence, using the optimization algorithm (Jang 1993).

If the output of each layer of neurophasic network is shown as $Q_{1,i}$ which is the output of the i group in the 1st layer, then the performance of different layers can be expressed as follows:

Layer 1: Each node in this layer is equivalent to a fuzzy set (Eq. 3):

$$Q_i^1 = \mu_{Ai}(x) \tag{3}$$

Layer 2: Each node of this layer is shown by II in which the input signals are multiplied by each other and produce output (Eq. 4):

$$W_i = \mu_{Ai}(x) \times \mu_{Bi}(y), \quad i = 1, 2 \tag{4}$$

Layer 3: Each node in this layer is represented by N, which calculates the ratio of the activity degree of rule

provision to the sum of the activity degrees of all rules (Eq. 5):

$$\overline{W_i} = \frac{W_i}{W_1 + W_2}, \quad i = 1, 2 \tag{5}$$

Layer 4: The output of each node in this layer is calculated by Eq. (6):

$$Q_i^4 = \overline{W_i} f_i = \overline{W_i}(p_i x + q_i y + r_i) \tag{6}$$

where $\overline{W_i}$ is the output of Layer 3 and (p_i, q_i, r_i) is the set of adaptive parameters of this layer. These parameters are called result parameters.

Layer 5: Each node in this layer, which is shown as Σ, calculates the final output value by Eq. (7) (Jang 1993):

$$Q_i^5 = \sum_i \overline{W_i} f_i = \frac{\sum_i W_i f_i}{\sum_i W_i} \tag{7}$$

Holt-Winters model

To predict the time series possessing seasonal or cyclical changes in addition to the trend, Holt-Winters model is used, which has been used in the studies on changes and the prediction of some meteorological elements in the country. To use this model, it is necessary to estimate three components of level or average (x), trend (T) and seasonal component (S). The three factors are calculated by the following equations.

$$F_t = \alpha\left(F_{t-1} - T_{t-1}\right) + (1 - \alpha)\frac{Y_{t-1}}{S_{t-k}} \tag{8}$$

$$S_t = Y S_{t-k} + (1 - Y)\frac{Y_t}{F_t} \tag{9}$$

$$T_t = \beta T_{t-1} + (1 - \beta)\left(F_t - F_{t-1}\right) \tag{10}$$

where F_t: smoothed surface at time t,

F_{t-1}: smoothed surface at time $t-1$,

Y_t-_1: actual data value at time $t-1$,

T_t: estimated trend,

S_t: estimated seasonal value, α: smoothing constant for the data, β: smoothing constant for trend estimation,

Y: smoothing constant for estimating seasonal changes and k periods in each year. Holt-Winters exponential smoothing coefficients are always between zero and one (Ord et al. 1997; Hyndman et al. 2002).

PMI-based input selection (PMIS) algorithm

PMI-based input selection (PMIS) algorithm introduced in this article was first developed by Sharma (2000) to identify input variables affecting hydrological models. PMI algorithm performs each iteration by finding the Cs that maximize the PMI value with respect to the output variable (with respect to the inputs that are pre-selected) and an input (C) and output (Y). The statistical concept that PMI estimates for Cs is based on a confidence interval determined by the distribution formed by a self-starting loop. If the input is significant, Cs is added to S and the selection continues until no significant input remains, then subsequently the algorithm stops. Given a random output variable Y, there is some uncertainty about an observation y that is a member of Y, which can be defined according to the Shannon entropy H, (Shannon 1948). However, assuming a random input variable X on which Y depends, cross-observation (x, y) reduces such uncertainty, since possessing x allows the value of y to be inferred, and vice versa. According to the definition of mutual information $I(X; Y)$, the reduction in uncertainty of variable Y is due to the observation of X (Cover and Thomas 1991). This problem is displayed as a common part between two circles. This common part is where the reduced uncertainty around X and Y is specified through the conditional entropy $H(X|Y)$ and $H(Y|X)$, respectively. The mutual information (MI) can be obtained directly from the following Eq. 11:

$$I(X;Y) = \iint p(x,y)\log\frac{p(x,y)}{p(x)p(y)}dxdy \tag{11}$$

$p(y)$ and $p(x)$ are the marginal probability density functions (pdfs) of X and Y, respectively, and $p(x, y)$ is the joint probability density function. In any case, practically, the correct form of probability density functions in Eq. (11) is unknown. Hence, probability density estimation is used instead. By replacing the probability density estimates with the numerical approximation of the integral in Eq. (12), there will be:

$$I(X;Y) \approx \frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{f(xi, yi)}{f(xi)f(yi)}\right] \tag{12}$$

where f represents the estimated density based on a sample of n observations (x, y). It should be noted that different bases are used for the logarithm, but usually 2 or e is used. If the base of the logarithm is not mentioned, the natural logarithm will be considered. Assuming Eq. (6), the accurate and effective estimation of MI depends to some extent on the method used to estimate the marginal and joint probability density functions. In general, there are three criteria for stopping PMI algorithm, (1) tabulated critical values, (2) AIC and (3) Hempel, as explained below. Tables of critical values

Table 1 Optimum values of Holt-Winters model coefficients for the monthly flow at Idenak hydrometric station

Station	Alpha	Beta	Gamma
Idenak	0.38	0.02	0.23

of the correlation coefficient (R) are available, based on an analytical formula for the error distribution of an estimate for an assumed sample size. In the case of the linear R, the distribution of the sample estimate follows a t distribution. The tables of critical values of R based on the t distribution are simply constructed (David 1966) which provide a critical value for R for a given sample size and a given confidence interval. However, unlike the linear R, an equivalent analytical definition for I cannot be deduced according to Eq. (6) (Goebel et al. 2005). Hence, researchers should use self-starting to calculate $f(I)$ (Granger et al. 2004; Sharma 2000). A study by Granger et al. (2004) investigated the distribution of the information estimator based on $p\ \hat{S}$ for a number of time series models. Finally, a practical alternative to the self-starting system was proposed. Monte Carlo simulation was used to practically determine the distribution of MI estimator, which was used in the first step to develop a stopping criterion based on approximate critical values. In each simulation, MI will be calculated for a series of data compared to Gaussian-white noise data (with sample size n, between 50 and 5000 values), which is used to obtain data with critical values that can be used for the independent test based on MI. For each sample size, first a series $\varepsilon_y \sim N(0,1)$ is constructed and then the marginal probability density functions $f\varepsilon_y$ are calculated. A total of one hundred thousand independent iterations of the series $\varepsilon_x \sim N(0,1)$ were made, independent of $\varepsilon_y \sim N(0,1)$. For each sample of ε_x, the marginal probability density functions f_{ε_y} and $f_{\varepsilon_x \varepsilon_y}$ were estimated and subsequently $\hat{I}(\varepsilon_x \varepsilon_y)$ was evaluated. The critical values of $\hat{I}(\varepsilon_x \varepsilon_y)$ are listed in Table 1 at different confidence intervals. Two alternative stopping rules were formulated by which, in each iteration, I'CsYs was compared with the corresponding critical values of Ib (95) and Ib (99) used instead of direct calculations of the self-starting system to determine which variable should be selected or when the algorithm should be stopped. Computational elimination of the self-starting system loop makes the selection of input variables much faster.

AIC (Akaike 1974), as a measure of the relationship between the accuracy of the regression filter and the size of the input set S, is adopted to formulate this stopping criterion. The criteria such as AIC are usually used as the basis for evaluation in model selection. AIC Eq. (13) is as follows:

$$AIC = n \log_e \left(\frac{1}{n} \sum_{i=1}^{n} u_i^2 \right) + 2p \tag{13}$$

where n is the number of observations, ui is the n residuals and p is the number of model parameters. For linear regression, the number of parameters is equal to $K+1$, where K is equal to the number of variables. To select effective input variables using AIC, variables that have the minimum AIC value are selected. The outlier selection method is a powerful statistical method to determine whether a given x value significantly deviates from other values in a set of x values. This test evaluates the deviation of a single observation from the mean of all observations. An observation value with a Z-score higher than 3 based on the 3σ rule for normal distribution is usually considered as outlier (outlier has a value higher than 3 times the standard deviation plus the mean of the data set). For the formulation of a stopping rule based on the identification of outliers for PMIS algorithm, the basic assumption is that the data set has a proportion of additional and irrelevant variables and important variables. However, the potential masking of outliers is important and should be considered, assuming that the data set has more than one dependent variable. For this reason, a modified Z value, which uses Hempel's distance, was used instead of increasing the efficiency of the method (Davies and Gather 1993). Hempel's distance is based on the median of the set of inputs. The failure point of test is $n/2$ and it is known as one of the most powerful tests to identify outlier (Davies and Gather 1993; Pearson 2002). The test starts by calculating the absolute deviation from the PMI mean for all inputs as the following equation:

$$d_j = \left| I_{C_jY.S} - I_{C_jY.S}^{(50)} \right| \tag{14}$$

where d_j represents the absolute deviation, and $I_{C_jY.S}^{(50)}$ represents the median PMI for the set of C inputs. Therefore, Hempel's distance can be determined as follows (15):

$$Z_j = \frac{d_j}{1.4826 d_j^{(50)}} \tag{15}$$

where Z_j represents Hempel's distance (modified Z-score) for the input set C_j and $d_j^{(50)}$ represents the median absolute deviation (MAD), d_j. A coefficient of 1.4826 modifies the distance so that the $Z > 3$ rule can be applied, as applied to the conventional Z-test (Pearson 2002). Given this stopping criterion, the PMI algorithm based on input selection no longer contains a self-starting loop, and PMI is not compared with any critical value of I. Instead, the value of Zs is determined to select Cs, and if $Zs > 3$, the input is selected and added to S, otherwise, the operation of the input variable selection algorithm is stopped. This study has two separate parts, which are described below. First, the input parameters for modeling and predicting the effective flow rate of Maroun river were selected based on PMI algorithm and
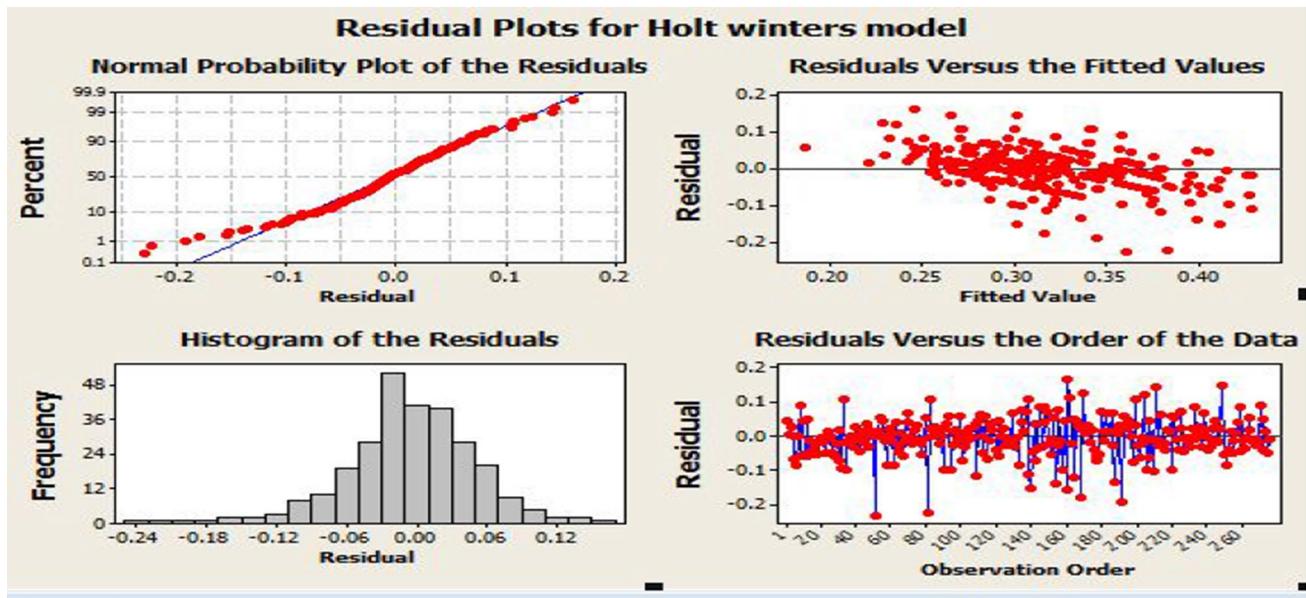
Fig. 1 Histogram of the residuals of Holt-Winters fitted to the monthly flow rate at Idenak hydrometric station

introduced as the input to ANFIS–ANN–SVM. Second, for the construction of hybrid models, after confirming the accuracy of the fitting of the above model, the residuals from the fitting of the univariate Holt-Winters model were introduced as the input to ANFIS–ANN–SVM for modeling and predicting the monthly inflow to Maroun dam reservoir.

Discussion and results

In this study, to model and predict the monthly flow rate of Maroun river through the proposed models, the data are divided into two series of training and test data. In addition, the monthly data of 2002–2012 was used in the test phase and the monthly data of 2013–2021 was used in the training phase.

Figure 1 shows the histogram of residuals, variance of residuals and residuals on probability paper. The histogram of the residuals of the fitted Holt-Winters model shows the monthly flow rate at Idenak hydrometric station by Minitab. As it is shown, the error distribution of the residuals resulting from the fitting of the above model is normal along a straight line. The probability curve of the data confirms the normality of the residuals of the model, the variance of the

residuals is stationary. Also, in the curve of the residuals versus time around the zero horizontal level, the rectangular distribution confirms no trend and the stationarity of the residuals resulting from the fitting is obvious. One of the important points in this study for the Holt-Winter models is the estimation of coefficients of the model. Given that no rule or algorithm was embedded by Minitab to estimate the coefficients of Holt-Winters model, trial and error was used to estimate these coefficients. On the other hand, it was found that by reducing the beta coefficient, the predictive power of the model increased to some extent. But the accuracy of the fit of the model is reduced. This is opposite in the case of gamma and alpha coefficients, i.e. by increasing the values of the alpha and gamma coefficients, the predictive power and accuracy of the Holt-Winters model increase to some extent (Table 1). To select the effective input parameters for modeling and predicting the monthly flow rate of Maroun through the above artificial intelligence models (ANN, ANFIS and SVM), PMI algorithm was used. According to Hempel and AIC, the monthly flow rate with a 3-month lag, and both parameters precipitation and temperature with a 1-month lag, with the lowest values of AIC and the highest values of Hempel, were introduced as the input to all artificial intelligence models (Table 2). To determine

Table 2 Effective input variables in modeling and forecasting the monthly flow based on the results of the PMI algorithm

Iteration	Variable	$I(x; y)$	MC-I*(95)	MC-I*(99)	AIC (k)	AIC (p)	Hempel
1	Log Q (t-3)	0.0554575	0.0683791	0.0468845	−3.48431	7.519	2.63321
2	Log T (t−1)	0.053388	0.0683791	0.0768845	5.4374	12.261	1.47619
3	Log p (t−1)	0.04732	0.06381	0.07677	4.3214	9.312	1.8312

the neurons of all three layers in the neural network and the hybrid model (Holt-Winters-ANN), because no algorithm is considered to determine the optimum number of neurons in the neural networks, the trial and error was used. Among different active functions in the neural network, the hyperbolic tangent function was selected as the optimum active function for neural networks while sigmoid function was selected as the optimum function for Holt-Winters hybrid model. Therefore, the function with the error backpropagation with an optimum neural structure for the (5–2–3) ANN and the (6–2–3) Holt-Winters hybrid model was selected. The number and type of membership functions in ANFIS and Holt-Winters-ANFIS hybrid model were selected by trial and error. Since the number of membership functions has a negative impact on the efficiency of the model, the principle of using the lower membership function was used for fitting both ANFIS and Holt-Winters-ANFIS hybrid models. Also, input Gaussian and output Triangular membership function was used as the optimum function in ANFIS and in Holt-Winters-ANFIS hybrid models. The input Triangular and output Triangular membership function was determined for Holt-Winters-ANFIS hybrid models. In this study, RBF kernel function was used to model the monthly flow rate for both SVM and the Holt-Winters-SVM hybrid model. One of the principles of using the kernel function is to find the optimum values of the parameters C, E and Γ for SVM and for the Holt-Winters-Support Vector Machine (HSVM) hybrid

model. For this purpose, the network search method was used. Also the optimal values of the mentioned parameters were determined, ($C = 2.4$, $E = 0.0018$, $\Gamma = 3.8$) For SVM and ($C = 2.9$, $E = 0.0016$, $\Gamma = 4.1$) for Holt-SVM hybrid models. In the testing (94 root mean square error of 94 and coefficient of explanation of 0.76) and train stage (root mean square error of 102 and coefficient of explanation of 0.71), the Holt Winters models have the weakest performance. Also among the three mentioned artificial intelligence models, the support vector machine model performs better than the other two models in both testing (with root mean square error of 78 and coefficient of explanation of 0.83) and training stages. With root mean square error of 86 and coefficient of explanation of 0.85. Table 3 shows the evaluation criteria of the performance of pre-institutional models for predicting the monthly flow rate at Idenak hydrometric station. Figure 2 shows the distribution of observed monthly flow rate and monthly flow rate predicted by the proposed models.

Conclusion

The evaluation criteria of the performance of the models in Table 3, such as the root mean square error and R^2 show the poor performance of the seasonal model of the Holt-Winters time series with root mean square error of 102 and coefficient of explanation of 0.63 in train stage and with root

Table 3 Evaluation criteria of the performance of pre-institutional models for predicting the monthly flow rate at Idenak hydrometric station

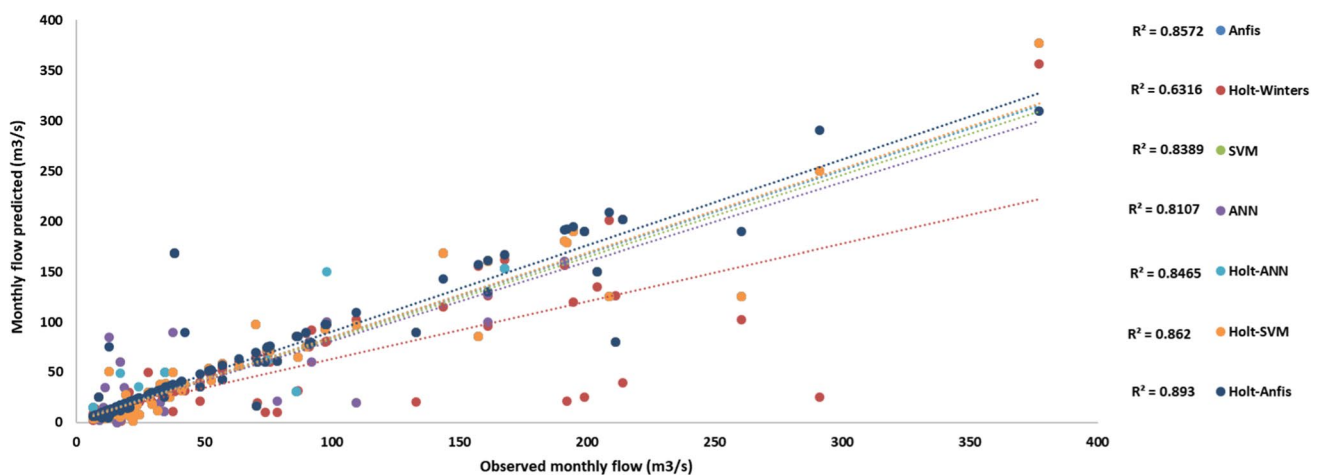|		Holt-Winters	ANFIS	SVM	ANN	Holt-ANN	Holt-SVM	Holt-ANFIS
R^2	Train	0.63	0.85	0.83	0.81	0.84	0.86	0.89
	Test	0.58	0.83	0.78	0.76	0.80	0.79	0.83
RMSE	Train	102	86	85	82	79	75	68
	Test	94	78	80	73	69	67	54

Fig. 2 Distribution of observed and predicted monthly flow by the proposed models

mean square error of 94 and coefficient of explanation of 0.58 in test stage. While all the artificial intelligence models performed much better at both stages than the Holt-Winters model, among the artificial intelligence models, ANN with optimal neural structure (3–5–2) showed some weak performance for modeling the monthly flow at Idenak station with root mean square error of 82 and coefficient of explanation of 0.81, in train stage and with root mean square error of 73 and coefficient of explanation of 0.76 in test stage In general, the performance of the hybrid models is better than all the artificial intelligence models despite the application of the PMI algorithm for improving the performance of the aforementioned models in simulating and predicting river flow. In addition, increasing or decreasing the number of neurons do not improve the performance of the neural artificial model and the hybrid Holt-Winters-neural artificial model with optimal neural structure (2–4–2) in predicting river flow. Also increasing or decreasing the number of the membership function do not improve the performance of ANFIS models and Holt-Winters ANFIS hybrid model in predicting river flow discharge. Among the various hybrid models used in this study, the Holt-ANFIS hybrid model showed a good ability to simulate peak flow rate with root mean square error of 68 and coefficient of explanation of 0.89 in train stage and with root mean square error of 54 and coefficient of explanation of 0.83 in test stage. In this research, since the three artificial intelligence models, compared to the Holt-Winters model and the aforementioned hybrid models, have many input parameters for modeling and forecasting the river flow, it seems that using hybrid models is a suitable solution to prevent wasting time by using less input parameters.

Conflict of interest The corresponding author states that there is no conflict of interest.

References

Abdulelah Al-Sudani Z, Salih SQ, Sharafati A, Yaseen ZM (2019) Development of multivariate adaptive regression spline integrated with differential evolution model for streamflow simulation. J Hydrol 573:1–12. https://doi.org/10.1016/j.jhydrol.2019.03.004

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723

Altunkaynak A (2007) Forecasting surface water level fluctuations of Lake Van by artificial neural networks. Water Resour Manage 21(2):399–408

Cover TM, Thomas JA (1991) Elements of information theory. Wiley Series in Telecommunications, New York.

David FN (1966) Tables of the correlation coefficient. In: Pearson ES, Hartley HO (eds) Biometrika tables for statisticians, 3rd edn, vol 1. Cambridge University Press, Cambridge

Davies L, Gather U (1993) The identification of multiple outliers. J Am Stat Assoc 88(423):782–792

Dewan A (2013) Floods in a megacity: geospatial techniques in assessing hazards, risk and vulnerability. Springer, Dordrecht, pp 119–156

Ghumman AR, Ghazaw YM, Sohail AR, Watanabe K (2011) Runoff forecasting by artificial neural network and conventional model. Alex Eng J 50(4):345–350

Goebel B, Dawy Z, Hagenauer J, Mueller JC (2005) An approximation to the distribution of finite sample size mutual information estimates. In: IEEE International conference on communications, 2005. ICC 2005, vol 2, pp 1102–1106. IEEE, New York.

Granger CW, Maasoumi E, Racine J (2004) A dependence metric for possibly nonlinear processes. J Time Ser Anal 25(5):649–669

Hyndman RJ, Koehler AB, Snyder RD, Grose S (2002) A state space framework for automatic forecasting using exponential smoothing methods. Int J Forecast 18:439–454

Jang JSR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23:665–685. https://doi.org/10.1109/21.256541

Jothiprakash V, Magar R (2009) Soft computing tools in rainfall-runoff modeling. ISH Journal of Hydraulic Engineering 15(sup1):84–96

Kisi O, Choubin B, Deo RC, Yaseen ZM (2019) Incorporating synoptic-scale climate signals for streamflow modelling over the Mediterranean region using machine learning models. Hydrol Sci J 64(10):1240–1252. https://doi.org/10.1080/02626667.2019.1632460

Liu R, Wang J, Shi J, Chen Y, Sun C, Zhang P, Shen Z (2014) Runoff characteristics and nutrient loss mechanism from plain farmland under simulated rainfall conditions. Sci Total Environ 468:1069–1077

Mohammadi K, Eslami HR, Kahawita R (2006) Parameter estimation of an ARMA model for river flow forecasting using goal programming. J Hydrol 331(1–2):293–299

Nayak PC, Sudheer KP, Rangan DM, Ramasastri KS (2005) Short-term flood forecasting with a neuro-fuzzy model. Water Resources Res 41(4).

Nourani V, Komasi M (2013) A geomorphology-based ANFIS model for multi-station modeling of rainfall–runoff process. J Hydrol 490:41–55

Nourani V, Alami MT, Aminfar MH (2009) A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. Eng Appl Artif Intell 22(3):466–472

Nourani V, Kisi Ö, Komasi M (2011) Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. J Hydrol 402(1–2):41–59

Ord JK, Koehler AB, Snyder RD (1997) Estimation and prediction for a class of dynamic nonlinear statistical models. J Am Stat Assoc 92:1621–1629

Pearson RK (2002) Outliers in process modeling and identification. IEEE Trans Control Syst Technol 10(1):55–63

Shannon CE (1948) A mathematical theory of communication. Bell Syst Techn J 27(3):379–423

Sharma A (2000) Seasonal to inter-annual rainfall probabilistic forecasts for improved water supply management: Part 1—A strategy for system predictor identification. J Hydrol 239(1–4):232–239

Uhlenbrook S, Roser S, Tilch N (2004) Hydrological process representation at the meso-scale: the potential of a distributed, conceptual catchment model. J Hydrol 291(3–4):278–296

Vapnik V (2000) The nature of statistical learning theory. Springer, New York

Wu CL, Chau KW (2011) Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. J Hydrol 399(3–4):394–409

Xie H, Li D, Xiong L (2016) Exploring the regional variance using ARMA-GARCH models. Water Resour Manage 30(10):3507–3518

ZamaniSabzi H, King JP, Abudu S (2017) Developing an intelligent expert system for streamflow prediction, integrated in a dynamic decision support system for managing multiple reservoirs: a case study. Expert Syst Appl 83:145–163. https://doi.org/10.1016/j.eswa.2017.04.039

Zhou ZZ, Huang TL, Ma WX, Li Y, Zeng K (2015) Impacts of water quality variation and rainfall runoff on Jinpen Reservoir, Northwest China. Water Sci Eng 8(4):301–308