



Comparative evaluation of classical and SARIMA-BL time series hybrid models in predicting monthly qualitative parameters of Maroon river

Abbas Ahmadpour¹ · Seyed Hassan Mirhashemi¹ · Mehdi Panahi²

Received: 15 November 2022 / Accepted: 17 January 2023 / Published online: 3 February 2023
© The Author(s) 2023

Abstract

In the present study, seasonal autoregressive integrated moving average (SARIMA) time series models, nonlinear BL model, multi-layer perceptron artificial neural network and SARIMA-bilinear hybrid models were employed to predict the quality parameters of total dissolved solids (TDS), sodium adsorption ratio (SAR), electrical conductivity (EC) in the Maroon basin in Khuzestan Province. For fitting the mentioned models, the monthly data were used for the calibration (1970–2000), confirmation (2001–2011) and prediction of the model (2012–2018). The appropriate models of SARIMA, the bilinear BL and SARIMA-BL hybrid models of the rationalized parameters of the above-mentioned quality parameters were selected based on the adequacy tests, such as the Akaike criterion and the independence test of the model residuals (Ljung–Box). To determine the effective input parameters of the network, the partial mutual information (PMI) algorithm was used to model the three monthly EC, TDS, and SAR parameters. Also, for input and output layers linear transfer function and for hidden layer various active functions with the back-propagation learning algorithm were used for modeling and predicting this water quality parameters. Comparison of the models showed the tangible superiority of SARIMA-bilinear hybrid models than the artificial neural network with effective input parameters based on PMI algorithm, SARIMA and bl models in performed prediction of the all three monthly qualitative parameters in all three stages, training-validation-test in Maroon basin.

Keywords Akaike statistics · Artificial neural network · Bilinear model SARIMA · Hybrid models · Ljung–Box statistics · Water saturation

Introduction

One of the most important issues in hydrological science is water quality issue as it is valuable in supplying water for the agricultural or drinking and industrial purposes. Today, most of the researchers study on water quality and focus on the surface and groundwater contamination issues. The pollution problem not only in industrialized countries but also in developing countries is important. In the field of water quality

management, several models such as artificial intelligence models and artificial neural networks, hydrological models such as MIKEH11 and QUAL2K, WASP, QUAN2E, WAT-EVAL, hec-5Q and time series models have been developed. Also due to the complexity and lack of sufficient knowledge about the physical processes in the hydrological cycle, the production of statistical models and their extension to processes has always been an engineer's concerns. The basis of many decisions in hydrological processes is the exploitation of water resources, which is based on the prediction and analysis of time series. Accordingly, some international researches have been carried on to predict the river water quality with different models. Singh et al. (2009) developed a neural network model for estimating the amount of dissolved oxygen (DO) and biochemical oxygen demand (BOD) levels in the Gomti River (India). In this study, both the models employed eleven input water quality variables measured in river water over a monthly period of 10 years at eight different sites. Jiang et al. (2018) studied the temperature, electrical conductivity and pH of groundwater as well as microbial and hydrochemical

✉ Mehdi Panahi
panahi40@yahoo.com

Abbas Ahmadpour
aahmadpour67@yahoo.com

Seyed Hassan Mirhashemi
hassan.mirhashemi@yahoo.com

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

² Department of Water Engineering, Faculty of Agriculture, University of Zanjan, Zanjan, Iran

analyses in a Karst valley in northwestern China. Dieng et al. (2017) revealed the presence of estuary in the entrance of the Saloum River in Senegal, significantly effecting the water quality, which caused the intrusion of saline water in the aquifers of the region. Narany et al. (2017) based on 25-year data investigated the quantitative effect of deforestation on the groundwater quality in northwestern Kelantan, Malaysia, and concluded that the amount of nitrate annually increased by 8.1%. Nosrati and Eeckhaut (2012) in their research conducted multivariate statistical analyses to assess groundwater quality using multiple statistical techniques in Hashtgerd Plain and concluded that the main factors affecting the quality of groundwater associated with natural pollutants are household sewage, industrial and agriculture materials. Ahmad and Ali Shah (2017) used an adaptive neural fuzzy inference system to estimate the need for BOD in the Surma River in Bangladesh. The results showed that the ANFIS model can predict water quality with high precision. Asadollahfardi et al (2011) reviewed the ability of two models of artificial neural network and recurrent neural network to estimate the water quality index of the Talkhe River in East-Azerbaijan Province to estimate the TDS parameter. Chowdhury et al. (2010) evaluated the performance of artificial neural network and ordinary Kriging for interpolation of arsenic values. The findings showed that the artificial neural network model estimates 15% better than Kriging method of arsenic values. Zou et al (2010) compared two models of ARIMA and artificial neural networks to predict soil salt and water content, and the results showed the superiority of ARIMA model. Finally, the performance of the comparative neural fuzzy inference model in comparison with the neural network indicates the superiority of the fuzzy inference model compared to the neural network. Sandhu and Finch (1995) have also confirmed the ability of artificial neural networks to predict daily and actual salinity levels in different waters of the watersheds and the ability to estimate the concentration of cations, anions, EC and TDS in these basins. The main purpose of this study is to compare the performance of the classic multi-layer perceptron neural network, linear (SARIMA), nonlinear (BL) time series and SARIMA-BL time series hybrid model in predicting and modeling the monthly qualitative parameters such as TDS, SAR and EC of the Maroon river. In addition, in this research, PMI algorithm was used to introduce the input parameters to the neural networks.

Materials and methods

Area of study

Maroon basin with an area of about 3824 km² is located in the geographic coordinates of longitude 49° 50' to 51° 10' N and latitude of 30° 30' to 31° 20' E at the mountains of Behbahan city in Khuzestan Province. The Maroon drainage

basin is surrounded by watersheds of Zohreh and Karun rivers in Khuzestan and Kohgiluyeh and Boyer-Ahmad Provinces. The average annual rainfall in the Maroon basin varied from about 150 mm in the coastal plain to about 900 mm in the northern mountains, and the rainfall is Mediterranean, with the maximum precipitation occurring between December and March. Table 1 presents the statistical characteristics of the time series of the monthly quality parameters of the Idanak hydrometric station used in this study. Also, Fig. 1 shows the main rivers of maroon basin.

SARIMA model

Box et al (1994) developed the ARIMA model for seasonal time series. If, in a time series, periodical behavior is seen in definite interval (S), this time series is seasonal, and the SARIMA model is used to model it. This model is shown in the form of SARIMA (p, d, q) (P, D, Q) s in which their structure is P, D, Q , the seasonal component of the model and (p, d, q) nonseasonal component of the model and S is the length of the season. Using the backward shift operator B , the general form of the model is shown as follows:

$$\phi_p(B)\phi_P(B^S)\nabla^d\nabla_s^D Z_t = \theta_q(B)\Theta_Q(B^S)a_t \quad (1)$$

where $\phi(B)$ and $\theta(B)$ are polynomials of order p and q , respectively. $\Phi(B^S)$ and $\Theta(B^S)$ are polynomials of order P and Q in B^S , respectively. p is nonseasonal auto-correlation order, d is nonseasonal difference, q is nonseasonal moving mean order, P is seasonal auto-correlation order, D is the number of seasonal differences, Q is seasonal moving mean order, and S is the season duration. ∇^d is nonseasonal operator and ∇_s^D is seasonal operator.

Time series models consist of this step process:

Model identification This step is started by plotting the autocorrelation function (ACF) and the partial autocorrelation function (PACF) charts, and the stationary in the

Table 1 The statistical features of time series of monthly qualitative parameters of Idanak hydrometry station

Qualitative parameter	Mean	Max	Min	SD
T (temperature)	25.80	40	10	10.25
EC	910	3777	350	378
TDS	583	10,105	6.6	463
Mg	1.8	12	0.10	1.7
SAR	1.85	11.36	0.16	1.17
Ca	5.19	33	1.30	2.82
Na	3	12.5	0.22	1.96
pH	7.8	8.7	6.5	0.32
SO ₄	6.75	20,006	0.07	79
Q (discharge)	52.15	377.17	4.25	58

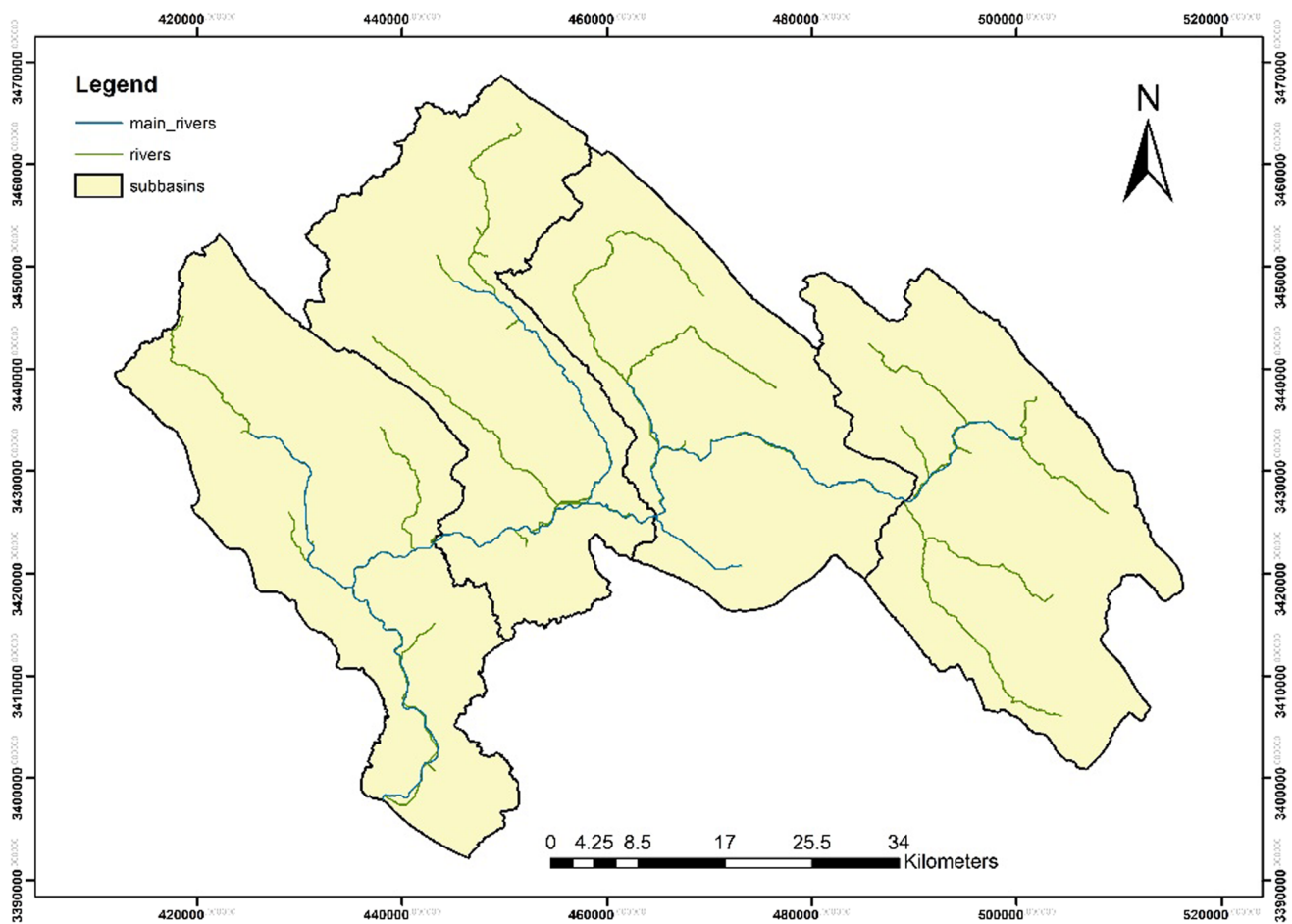


Fig. 1 The main rivers of Maroon basin

mean and variance of the data is evaluated. Autocorrelation function (ACF) is one of the most important tools for testing data dependency. This function often gives us an insight into the probabilistic pattern that produces the data. This is used to identify and fit the appropriate stochastic model for data.

Model fit (estimation of parameters): At this step, by identifying the appropriate models in the previous step, to compare several models and select the best ones, the following equation is used to calculate Akaike criterion:

$$AIC = N \left(\ln \left(\frac{2\pi RSS}{n} \right) \right) + 1 + 2m \tag{2}$$

where n is the total number of data, $m = (p + q + P + Q)$, RSS is the root of the sum of the squares of the residuals. A model is chosen that has the lowest value of AIC.

Model recognition: To check the accuracy of the model, the residual chart is evaluated for normal and stationary aspects. Therefore, by plotting the ACF and PACF charts of the residuals of fitted models, if the values of the auto-correlation coefficients and partial auto-correlation coefficients

are within the 95% confidence level, the fitted model is adequate.

Bilinear model (BL)

The bilinear model was introduced by Granger and Andersen (1978). Linear models of time series are Taylor’s first-order extension, and the main idea of the bilinear model is the Taylor’s second rank extension, which is defined as follows:

$$BL(p, q, r, s): Z_t = \sum_{i=1}^p (\varphi_i \cdot Z_{t-i}) - \sum_{j=1}^q (\theta_j \cdot \varepsilon_{t-j}) + \sum_{i=0}^r \sum_{j=1}^s (\beta_{ij} \cdot Z_{t-i-j} \cdot \varepsilon_{t-i}) + \varepsilon_t \tag{3}$$

where Z_t is the required time series, and $p, q, r,$ and s are integers that represent the rank of the bilinear model, which is denoted by the form $BL(p, q, r, s)$. The bilinear model is actually the same as the extended linear ARMA model, which the nonlinear term is added to its right side. In this

term, the product of two variables $Z_{t-i-j}, \epsilon_{t-j}$, which are varied with time, eliminates the linear state and the model is turned to be nonlinear. There are two steps to fit the bilinear model; the first is to determine the order of the model (p, q, r, s) and the other is to estimate β, φ, θ coefficients. Determining the model parameters is performed the Akaike method. Each rank with a smaller acacia value is selected as the appropriate model order (Akaike 1974). One of the tests used to express the model's adequacy of linear (SARIMA) and nonlinear (BL) and time series hybrids models' (SARIMA-BL) is Ljung–Box test. This test is applied by computing Q statistics that follows (χ_{k-m}^2) Chi-squared distribution:

$$Q = N(N + 2) \sum_{k=1}^L (N - k)^{-1} r_k^2(\epsilon) \quad (4)$$

According to this equation, N is the number of samples, L is the number of delays of auto-correlation function, and r_k^2 is the square of ϵ residual time series auto-correlation in the delay k . Q statistic was compared with the value of Chi-squared extracted from table at a significant level of 5%. The adequacy of the model was confirmed if Q value was lower than table Chi-squared value.

Artificial neural networks

An artificial neural network is a computational mechanism that can produce a series of new information by gathering information and calculating them. One of the most commonly used neural networks used in the hydrological process is the multilayer perceptron neural network. This network consists of an input layer, an output layer and one or more layers between them that are not directly connected to the input data and output results. Input layer units only have the task of distributing the input levels of the next layer, and the output layer also responds to the output signals. In these two layers, the number of neurons is equal to the inputs and outputs, and the hidden layer is responsible for the relationship between them. One of the steps in numerical preparation and calculations for feeding the neural networks of data normalization is to increase the sensitivity of data and increase the learning power, which increases the predictive ability. In this research, input and output vectors of the network were standardized using (6) in the interval (1 and 0).

$$y = 0.5 + \left(0.5 \times \left(\frac{x - \bar{x}}{x_{\max} - x_{\min}} \right) \right) \quad (5)$$

In the above equation, X is the given data, $\bar{X} \rightarrow$ mean of data, X_{\max} : maximum data, X_{\min} : minimum data and y : normalized data.

Determining effective input variables in data-based models

There are various methods to determine input variables in models based on data processing, especially in terms of artificial neural networks. These include trial-and-error method, exploration method, expert knowledge-based method, statistical-based method of analysis and combination of different methods (Bowden et al. 2005). The trial-and-error method is two major flaws:

1. Although many combinations of variables are considered as inputs in models based on data processing, there is no guarantee of choosing the best combination of input variables, especially for modeling phenomena that include many input variables.
2. Given that, different models of ANNs should be trained and tested for each combination of input variables which takes a lot of time and complex computing. In the heuristic, the word variables are selected step by step. This method is often used in order to avoid over-estimating all subsets of input variables. The standard step-by-step approach is forward selection and backward selection. In the forward method, in each successive step with a set of input variables selected, the variable that improves the model performance is added as the best final input variable. In the backward approach, the program starts by selecting all input variables and successively input variables that reduce the model performance. The main drawback of the exploration methods is that there are no guarantees to choose the best combination of input variables like trial and error. Used to determine the initial set of input variables, this method has been approved by the American Society of Civil Engineers (ASEC) Committee for the Application of ANN to Hydrology Studies (Sharma 2000). Well understanding of the hydrological system is important to proper select the input variables in the model. If effective input variables not be selected and considered correctly, some information about the system may be lost which led to a confusion in the ANN model training process. In the combination of different methods, the various methods mentioned above are used to determine the input variables in the modeling of the system studied with the ANN models. Although each of these methods has weaknesses for determining effective input variables, the statistical method is a suitable method for determining effective input variables. Since that it is not limited to a specific application, so there is a good potential for providing appropriate algorithms to be added to the ANN model structure as a component, based on data

analysis rather than expert knowledge or heuristic only nonlinear variables selection algorithm. The input to determine the effective input variables in the models is based on the data processing PMI algorithm.

Introduction of PMI algorithm based on input selection (PMIS)

The PMI-based input selection algorithm was initially developed by Sharma (2000) to identify the effective input variables in hydrology models. The PMI algorithm performs each iteration by considering an input (*C*) and an output (*Y*) and finding the *C*s that maximizes the PMI value according to the output variable (according to inputs that are pre-selected). The statistical concept that PMI estimates for *C*s is based on the confidence ranges defined by the distribution generated by a bootstrap loop. If the input is significant, the *C*s is added to *S* and the selection will continue, so that no significant input remains and the algorithm is stopped subsequently.

Estimation of partial mutual information (PMI)

Given a stochastic variable *Y*, there is some uncertainty about *y* observation which can be defined according to the Shannon *H* entropy (Shannon 1948). Assuming a random input variable *X*, where *Y* is dependent on it, mutual observations (*x*, *y*) reduce this uncertainty, since knowing *x* allows the derivation of *y*, and vice versa. According to the definition of mutual information *I* (*X*; *Y*), the reduction in the uncertainty of the variable *Y* is due to the observation of *X* (Cover and Thomas 1991). This problem is represented as a common part between two circles in Fig. 2. This common part is where the reduced uncertainty around *X* and *Y* is specified by the conditional entropy *H* (*X* | *Y*) and *H* (*Y* | *X*), respectively. Mutual information (MI) can be calculated directly by the following formula:

$$I(X;Y) = \iint p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dx dy \tag{6}$$

where *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. However, in practice, the correct form of probability density functions in (6) is unknown. Hence, estimates of probability densities are used instead of it. By placing estimates of probability density with the numerical approximation of the integral in (5) we have:

$$I(X;Y) \approx \frac{1}{n} \sum_{i=1}^n \log \left[\frac{f(x_i, y_i)}{f(x_i)f(y_i)} \right] \tag{7}$$

where *f* denotes the estimated density based on the sample of *n* observations of (*x*, *y*). It is noteworthy that various

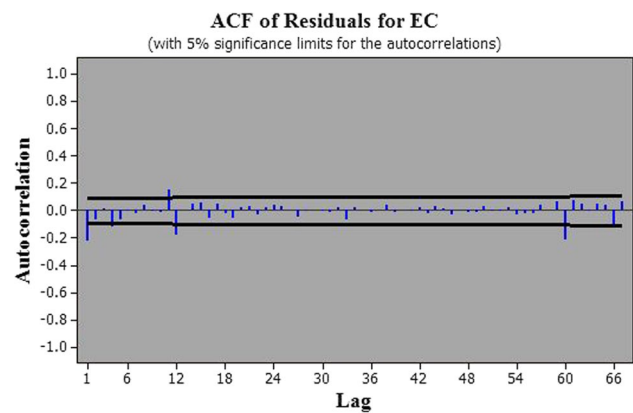


Fig. 2 Autocorrelation function chart of residuals of fitted SARIMA (1,1,2) (2,1,1) 12 time series EC quality parameter

logarithms are used, but usually 2 or *e* is used. If the basis of logarithm is not mentioned, then the logarithm is natural. As shown in Eq. (6), it could be said that the accurate and effective estimation of MI depends on the method used to estimate the marginal and joint probability density functions. Totally, there are three criteria of PMI algorithm stopping as: 1. tabulated critical values, 2. the criterion based on Akaike information criterion (AIC), 3. the Hampl test criterion (Shannon 1948; David 1966; Davies and Gather 1993; Maier and Dandy 2000; Sharma 2000; Pearson 2002; Goebel et al. 2005).

SARIMA-BL hybrid models

To construct a hybrid SARIMA-bilinear models first seasonal ARIMA model, the time series data are processed using the very well-known Box–Jenkins method, which consists of three steps: identification, estimation and verification. The details regarding Box–Jenkins method are given in the model description. After verification, the residuals of the best SARIMA fitted model are used as inputs for the bilinear model. The purpose of this final step is to forecast the error sequence of the fit best SARIMA model, which cannot be defined using a linear method like SARIMA. The construction of the SARIMA-bilinear hybrid model includes six steps that are referred to below.

1. Determine the adequacy of data: One way to test the adequacy of data is to use the Hurst coefficient. The Hurst coefficient is a criterion for measuring the adequacy of the data information in terms of the length of the statistical period.
2. Data normalization test: The monthly electrical conductivity of the Maroon basin, in which the Landa coefficient of Box–Box conversion was used after normalizing the data.

3. Fitting SARIMA models on monthly qualitative flow data of the Idenak hydrometer station.
4. Confirmation of the fitting of SARIMA models based on the behavior of the residual. Auto-correlation function (ACF), partial correlation (PACF) and the Manteo port statistics (Ljung–Box) are common steps in the modeling of both the SARIMA and the hybrid model.
5. Selection of the best SARIMA model based on the minimum values of the AIC and Schwartz statistics
6. Introduction of the residual SARIMA models as inputs to the bilinear time series models to build a hybrid model.

In addition, considering the nonlinear structure of the data in the SARIMA model, in the condition where residual data of the model contain nonlinear data, the use of the bilinear time series models model would be appropriate to describe the residual SARIMA models. Therefore, the results of bilinear models can be considered as a substitute for SARIMA model errors. Then, the sum of the results of both SARIMA and bilinear forecasts is considered to be the final forecast. To fit the qualitative data with the SARIMA model, the graphical beauty of the software output mini tab software was used. Then, to construct the hybrid model, the residuals obtained from fitting the SARIMA models to the monthly qualitative data were fitted with the bilinear time series R software.

Model evaluation criteria

In this paper, in order to evaluate the performance of each mentioned model, a number of numerical criteria were used to determine the performance of the models. These criteria include:

- Average absolute error

$$MAE = \frac{1}{n} \sum |Q_i - \hat{Q}_i| \tag{8}$$

- Root-mean-squared error

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Q_i - \hat{Q}_i)^2}{n}} \tag{9}$$

- The coefficient of determination

$$R^2 = \left[\frac{\sum_{i=1}^n (Q_i - \bar{Q})(\hat{Q}_i - \bar{\hat{Q}})}{\sqrt{\sum_{i=1}^n (Q_i - \bar{Q})^2 \sum_{i=1}^n (\hat{Q}_i - \bar{\hat{Q}})^2}} \right] \tag{10}$$

In the above equations, Q_i is the observed monthly qualitative parameter values, \hat{Q}_i the predicted monthly qualitative parameter values, \bar{Q} the average monthly qualitative parameter values, and $\bar{\hat{Q}}$ the average of the predicted monthly qualitative parameter values.

Results and discussion

As mentioned, one of the methods for testing the adequacy of the fitted model for the monthly discharge time series is to investigate the residual auto-correlation function. Figures 2, 3 and 4 represent the autocorrelation function and partial correlation of the residuals of fitted SARIMA models based on EC, SAR and TDS monthly qualitative parameters. According to Figs. 2, 3 and 4, the ACF values of the residuals are located in the permitted range of 95% confidence interval ($\pm 1.96/\sqrt{n}$), so the residuals were stationary and did not show any trend. Also, the Q values of the Ljung–Box test confirmed the zero hypothesis that the residuals series is stochastic (Table 2). Also, the best fitted SARIMA, BL and SARIMA-BL Hybrid models were selected based on the lowest AIC information criterion, the Portmanteau statistics and the model residuals test (in order to evaluate the adequacy of the residuals of the best model).

Table 3 shows a summary of the parameters of the fitted bilinear models on the monthly data of TDS, EC, SAR. Given the values of the Ljung–Box statistics, BL bilinear models and time series hybrid models (SARIMA-bilinear) comparing them with the corresponding values of the Chi-square of table which handle to be sure of the validity of fitting bilinear models, finally, bilinear model BL (5,1,1,1) with an Akaike value of 379.18 as the best model for modeling and predicting the SAR parameter, the bilinear model (12,0,1,1) with an Akaike value of 396.24 for modeling and predicting

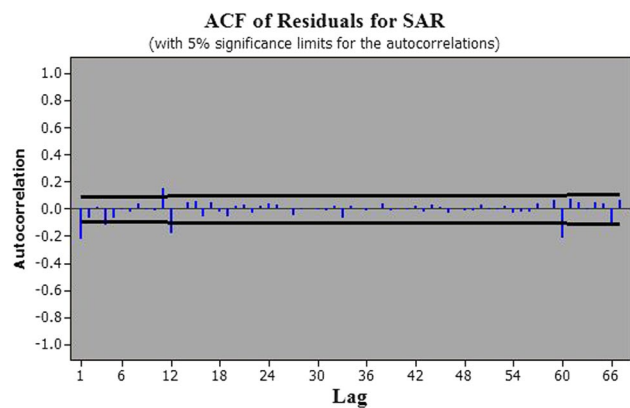


Fig. 3 Autocorrelation function chart of residuals of fitted SARIMA model (1.1.1) (1.1.1) 12 SAR time series quality parameter

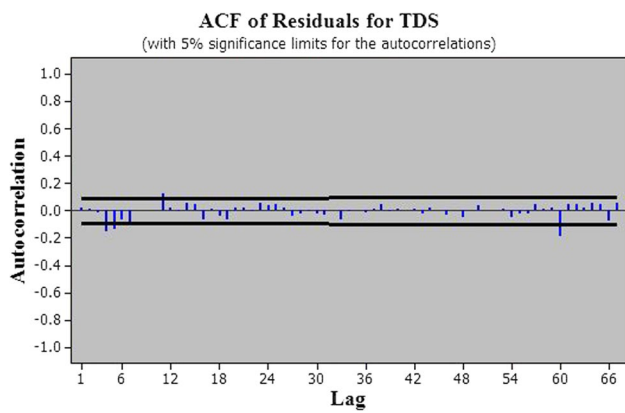


Fig. 4 Autocorrelation function chart of residuals of fitted SARIMA model (2.1.1) (1.1.2) 12 time series TDS qualitative parameter

the TDS qualitative parameter... and the bilinear model BL (7,2,1,1) having Akaike value of 385.97 were used to predict and model the qualitative parameter of EC. It is worth to mention that, by increasing the moving average of bilinear models, their ability becomes weaker for model and predict. On the other hand, in the SARIMA linear model, as the order of the model increases, the model performance in forecasting as well as the accuracy of the model greatly reduces. In the nonlinear BL and SARIMA-BL hybrid models, it is true about the models with higher order of p and lower values of

q , r and s which have better performance and more fitting accuracy.

Determination of effective input variables on qualitative variables in artificial neural network

To estimate the qualitative parameters of EC, TDS and SAR flow at the Idanak hydrometric station with artificial neural network, we initially attempt to determine the effective input variables. To provide the potential set of input variables, including sodium, magnesium, calcium, sulfate, HC and PH, each variable was considered up to three months ago. Then input variables affecting output variables including EC, TDS and SAR were obtained using PMI algorithm. Table 4 presents the qualitative parameters affecting the monthly electrical conductivity parameter based on the Hampel and Akaike criterion, where the sodium quality parameter (with a delay of one month), temperature (with a delay of two months), flow rate (with a delay of two months), magnesium (with two months delay) and acidity (with a delay of one month) were supposed as effective parameters. Table 5 shows the qualitative parameters affecting the monthly TDS parameter based on the Hampel and Akaike criterion, in which the parameters of sodium, magnesium, calcium, flow rate, temperature (with a delay of one month) were introduced as an effective input parameter to the neural network for modeling and advance monthly TDS parameters. Also parameters of

Table 2 A summary of the statistical parameters of fitted SARIMA model of the qualitative vectors of SAR, EC, TDS of Idanak station

Model	Parameter	ϕ_1	ϕ_2	θ_1	θ_2	Φ_1	Φ_2	Θ_1	Θ_2	Q	λ^2	AIC
SARIMA(2,1,1)(2,1,1) ₁₂	TDS	0.02	0.02	0.97	-	0.01	0.02	0.46	-	5.1	18	374
SARIMA(1,1,2)(2,1,1) ₁₂	EC	0.08	-	0.07	0.02	0.07	0.03	0.02	-	21	37	283
SARIMA(1,1,1)(1,1,1) ₁₂	SAR	0.04	-	0.86	-	0.02	-	0.28	-	25	48	317

Table 3 Results of the goodness of fit test of fitted BL bilinear and SARIMA-BL hybrids models based on monthly qualitative parameter of Idanak station

Model	Monthly qualitative parameters	Lag	Q test	Chi-square value of Table	Test result
BL(7,2,1,1)	EC	85	79	107	Acceptable
BL(12,0,1,1)	TDS	75	92	101	Acceptable
BL(5,1,1,1)	SAR	95	94	118	Acceptable
SARIMA(2,1,1)(2,1,1) ₁₂ *BL(2,1,1,1)	TDS	88	83	105	Acceptable
SARIMA(1,1,2)(2,1,1) ₁₂ *BL(4,1,1,1)	EC	83	78	98	Acceptable
SARIMA(1,1,1)(1,1,1) ₁₂ *BL(2,0,1,1)	SAR	98	90	109	Acceptable

Table 4 Results of the PMI algorithm for EC output variable

Iteration	Variable	$I(x;y)$	MC-I*(95)	MC-I*(99)	AIC(k)	AIC(p)	Hampel
0	logMg($t-2$)	0.0564575	0.0683791	0.0768845	-1.48833	10.519	2.83221
1	logNa($t-1$)	0.0568879	0.0683791	0.0768845	-2.41602	10.8851	2.67548
2	LogT($t-1$)	0.0552132	0.0683791	0.0768845	2.6141	10.9821	2.7641
3	LogQ($t-2$)	0.055312	0.068371	0.0768841	3.162	11.3612	2.8862
9	PH($t-1$)	0.053388	0.0683791	0.0768845	6.4374	18.261	1.47619

sodium, magnesium, calcium and acidity (with one month delay) parameters were introduced as effective input parameters in the modeling and prediction of the monthly parameter based on the Hampel and Akaike criterion. Also Based on the results in Table 6, monthly parameters of calcium, sodium Magnesium without time delay and ph parameter with one time delay based on Hampel and Akaike criteria were introduced as effective input parameters for modeling and forecasting the monthly qualitative parameters of SAR by PMI algorithm to the artificial neural networks. Unlike Zou et al. (2010) who found the superiority of ARIMA model than ANN in predicting EC parameter, our results are agreed with Sandhu and Finch (1995) results, who found the ability of neural network in predicting the hydrological parameters.

Different models were evaluated to determine the optimum levels of hidden layers and transmission functions in artificial neural networks. The number of layers as well as the number of hidden layer neurons was determined based on the trial-and-error test. In the artificial neural network model, various scenarios were tested for the three parameters TDS, SAR, EC. Then, according to the model's studied criteria, the scenarios with the least error and the highest correlation with the observational data of the qualitative parameter related to the optimal structure were selected. For modeling of the neural network, data are usually divided into train and test groups. Therefore, in this study, to avoid over-fitting which results in non-generalizability for new data, the data were divided into three train data (1970–2000), validation data (2001–2011) and test data (2012–2018). Among the available transfer functions, the function for the input and output layer of the linear function and in the hidden layer of various functions such as sigmoid tangent function, sigmoid log, etc., error back-propagation train algorithm was applied. The calculations related to optimal models showed that

the number of neurons 8, 9, 8 in the hidden layer was for modeling and predicting EC, TDS and SAR, respectively, as being applied as optimal values. It should be noted that by increasing or decreasing the number of neurons in the three layers (the input, middle and output) the performance of the network was not significantly changed. In addition, in order to evaluate the performance of appropriately fitted models, different statistics were used. In Table 7, the numerical criteria of performance evaluation of the models were used to predict monthly qualitative parameters of EC, SAR and TDS. Also, among various active function the hyperbolic tangent active function was selected as the superior active function of the artificial neural network for predicting and modeling monthly qualitative parameters for EC, and the best active function for SAR was log sigmoid; similarly, tangent sigmoid was the best active function determined for the TDS parameters. Evaluation of the performance criteria of the models in Table 7 shows the gradual decline in the performance of the SARIMA models used for predicting and modeling the TDS and SAR monthly qualitative parameters from the training stage to the test stage so that the model performance in TDS prediction. The coefficient of determination was 0.81 at the calibration stage and reached to 0.76. But the performance of the SARIMA model in predicting EC parameterization shows a gradual improvement of the model performance. So that, in the validation stage the model performance was better than in the calibration stage; however, in the reduced model performance testing phase, this is due to the numerical values of the model performance evaluation. On the other hand, artificial neural network model with passive structures and active functions is used to predict monthly qualitative parameters and showed the highest coefficient of determination and lowest root-mean-square error and mean error compared to SARIMA time series models and binary BL models have better performance.

Table 5 The results of PMI algorithm for TDS output variable

Iteration	Variable	$I(x;y)$	MC-I*(95)	MC-I*(99)	AIC(k)	AIC(p)	Hampel
0	logNa(t)	0.507517	0.0683791	0.0768845	-201.039	-187.064	12.4131
1	logCa(t)	0.269758	0.0683791	0.0768845	-212.635	-153.945	10.0952
2	logMg(t)	0.107627	0.0683791	0.0768845	-180.195	-46.3449	3.05574
3	logT(t-1)	0.0944237	0.0683791	0.0768845	-226.098	-18.1274	2.65314
4	Log Q(t)	0.0914213	0.0683791	0.0768845	-271.021	-14.3614	2.43612

Table 6 The results of PMI algorithm for SAR output variable

Iteration	Variable	$I(x;y)$	MC-I*(95)	MC-I*(99)	AIC(k)	AIC(p)	Hampel
0	logNa(t)	1.015	0.0683791	0.0768845	-537.902	-523.927	4.451
1	logCa(t)	0.171907	0.0683791	0.0768845	-698.673	-639.983	2.03252
2	logMg(t)	0.122974	0.0683791	0.0768845	-649.99	-516.139	2.21644
3	PH(t-1)	0.0912921	0.0683791	0.0768845	-657.061	-415.829	1.70985

Table 7 Numerical criteria for evaluating the predicted values of monthly qualitative parameters using linear models of SARIMA and nonlinear bilinear time series, neural network, SARIMA-bilinear hybrid models (test stage)

Model	Parameter quality	Stage	R2	RMSE	MAE
ANN(5-9-1)	TDS	Test	0.84	111.56	92.54
		Validation	0.94	96.3	70
		Train	0.96	88.3	55.3
BL(12,0,1,1)		Test	0.79	149.96	131.79
		Validation	0.81	139	121
		Train	0.84	121	102
SARIMA(2,1,1) (1,1,2) ₁₂		Test	0.76	178.36	160.42
		Validation	0.79	163	148
		Train	0.81	138	113
SARIMA-BL		Test	0.86	105.6	102
		Validation	0.83	127	114
		Train	0.89	91	82
ANN(6-8-1)	EC	Test	0.87	180.81	153.67
		Validation	0.89	164	138
		Train	0.91	148	127
SARIMA(1,1,2) (2,1,1) ₁₂		Test	0.83	290.02	264.79
		Validation	0.85	217	214
		Train	0.79	304	281
BL(7,2,1,1)		Test	0.80	245.02	219.33
		Validation	0.83	237	206
		Train	0.86	221	194
SARIMA-BL		Test	0.89	171.3	143.6
		Validation	0.92	150	126
		Train	0.95	127	101
SARIMA(1,1,1) (1,1,1) ₁₂	SAR	Test	0.73	127	186
		Validation	0.70	142	196
		Train	0.79	173	114
ANN(3-8-1)		Test	0.81	104	156
		Validation	0.87	89	113
		Train	0.90	75	98
BL(5,1,1,1)		Test	0.78	132	172
		Train	0.81	116	151
		Validation	0.83	105	147
SARIMA-BL		Test	0.84	95	146
		Validation	0.86	89	138
		Train	0.88	0.78	127

Table 7 shows numerical criteria for evaluating the predicted values of monthly qualitative parameters using linear models of SARIMA and nonlinear bilinear time series, neural network, SARIMA, bilinear hybrid models. On the other hand, the evaluation of the performance of the models in Table 7 according to the coefficient of determination and average absolute error in the training-validation-test

stages was investigated. Besides, the results showed that the values of the corresponding tests decreased and the root-mean-square error increased. In addition, determining both the effective input parameters to the neural network by PMI algorithm and the optimal structure along with active function is very time-consuming and exhausting task. Therefore, using the SARIMA-BL hybrid model is a suitable solution to this problem.

Conclusion

So in this study, SARIMA, bilinear time series models, multi-layer perceptron neural network and SARIMA-BL hybrid models were used to predict monthly qualitative parameters of SAR, TDS and EC. The validity of the SARIMA fit models for modeling the aforementioned monthly qualitative parameters was confirmed based on the Portmanteau statistic, the residual of autocorrelation and partial autocorrelation. The validity of the SARIMA-BL hybrid and bi-linear models was verified using numerical values of Portmanteau statistics. Finally in this study, PMI algorithm was used to improve the performance of artificial neural network based on the results of PMI algorithm with different time delay as input to network to predict monthly qualitative parameters. In addition, evaluated criteria of model showed that BL models performed better than SARIMA models in predicting monthly EC-TDS-SAR quality parameters. So that, BL models have higher coefficient of determination and lower mean square root. They had training-validation-testing in all three phases compared to the SARIMA models. On the other hand, the evaluation of model for the studied criteria shows a gradual decrease in the performance of neural network models. BL models and hybrid models are from training to testing, respectively; the comparison of four models showed the tangible superiority of SARIMA-bilinear hybrid models in comparing to artificial neural network with effective input parameters based on PMI algorithm in prediction of the all there monthly qualitative parameters in Maroon basin. SARIMA-BL hybrid models performed better in all three stages of training-validation-test despite using lower input parameters than artificial neural networks with the least square root mean and the coefficient of determination.

Funding The author(s) received no specific funding for this work.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahmad AAM, Shah SMA (2017) Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River. *J King Saud Univ Eng Sci* 29(3):237–243. <https://doi.org/10.1016/j.jksues.2015.02.001>
- Akaike H (1974) A new look at statistical model Identification. *IEEE Trans Autom Control* 19:716–723
- Asadollahfardi G, Taklify A, Ghanbari A (2011) Application of Artificial Neural network to predict TDS in Talkheh Rud River. *J Irrig Drain Eng (ASCE)* 138(20):363–370. <https://doi.org/10.4491/ieer.2015.096>
- Bowden GJ, Maier HR, Dandy GC (2005) Input determination for neural network models in water resources applications. Part 2. Case study: forecasting salinity in a river. *J Hydrol* 301(1–4):93–107. <https://doi.org/10.1016/j.jhydrol.2004.06.020>
- Box GEP, Jenkins GM, Reinsel GC (1994) Time series analysis: forecasting and control, 3rd edn. Englewood Cliffs Inc, New York, p 598
- Chowdhury M, Alouani A, Hossain F (2010) comparison of ordinary kriging artificial neural network for spatial mapping of arsenic contamination of groundwater. *Stoch Environ Res Risk Assess* 24(1):1–7. <https://doi.org/10.1007/s00477-008-0296-5>
- Cover TM, Thomas JA (1991) Elements of information theory. Wiley series in telecommunications. Wiley, New York
- David FN (1966) Tables of the correlation coefficient. In: Pearson ES, Hartley HO (eds) *Biometrika tables for statisticians*, vol 1, 3rd edn. Cambridge University Press, Cambridge
- Davies L, Gather U (1993) The identification of multiple outliers. *J Am Stat Assoc* 88(423):782–792
- Dieng NM, Orban P, Otten J, Stumpp C, Faye S, Dassargues A (2017) Temporal changes in groundwater quality of the Saloum coastal aquifer. *J Hydrol* 9:163–182. <https://doi.org/10.1016/j.ejrh.2016.12.082>
- 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 (ICC-05)*, Seoul, South Korea. <https://doi.org/10.1109/icc.2005.1494518>
- Granger CWJ, Andersen AP (1978) An introduction to bilinear time series models. Vandenhoeck and Ruprecht, Göttingen
- Jiang Y, Cao M, Yuan D, Zhang Y, He Q (2018) Hydrogeological characterization and environmental effects of the deteriorating urban karst groundwater in a karst trough valley: Nanshan, SW China. *Hydrogeol J* 26(5):1487–1497. <https://doi.org/10.1007/s10040-018-1729-y>
- Maier HR, Dandy GC (2000) Artificial neural networks for the prediction and forecasting of water resources variable: a review of modeling issues and application. *Environ Model Softw* 15(1):101–124. [https://doi.org/10.1016/s1364-8152\(99\)00007-9](https://doi.org/10.1016/s1364-8152(99)00007-9)
- Narany TS, Aris AZ, Sefie A, Keesstra S (2017) Detecting and predicting the impact of land use changes on groundwater quality, a case study in Northern Kelantan, Malaysia. *Sci Total Environ* 599:844–853. <https://doi.org/10.1016/j.scitotenv.2017.04.171>
- Nosrati K, Van Den Eeckhaut M (2012) Assessment of groundwater quality using multivariate statistical techniques in Hashtgerd Plain. Iran. *Environ Earth Sci* 65(1):331–344. <https://doi.org/10.1007/s12665-011-1092-y>
- Pearson RK (2002) Outliers in process modeling and identification. *IEEE Trans Control Syst Technol* 10(1):55–63. <https://doi.org/10.1109/87.974338>
- Sandhu N, Finch R (1995) Methodology for flow and salinity estimation in the sacramento-san Joaquin Delta and Suisun marsh chapter 7. In: *Artificial neural networks and their application*, 16th Annual progress Report, New York, p 85
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–423
- Sharma A (2000) Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: part 1—a strategy for system predictor identification. *J Hydrol* 239:232–239. [https://doi.org/10.1016/s0022-1694\(00\)00346-2](https://doi.org/10.1016/s0022-1694(00)00346-2)
- Singh KP, Basant A, Malik A, Jain G (2009) Artificial Neural Network modelling of the river water quality—a case study. *J Ecol Model* 220:888–895. <https://doi.org/10.1016/j.ecolmodel.2009.01.004>
- Zou P, Jing Song Y, Jianrong F, Guangming L, Dongshun L (2010) Artificial neural network and time series models for predicting soil salt and water content. *Agric Water Manag* 97:2009–2019. <https://doi.org/10.1016/j.agwat.2010.02.011>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.