Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2021.26987

218 (2021) 193–209
April

Comparison of Box–Jenkin time series and radial basis function for sodium adsorption rate forecasting; a case study Aras, Sepidrud, Karun, and Mond Rivers

Elham Rahnama[a], Omolbanin Bazrafshan[a], Gholamreza Asadollahfardi[b,*], S. Yaser Samadi[c]

[a]*Departement of Natural Resources Engineering, Faculty of Agricultural Engineering and Natural Resources, Hormozgan University, Bandar Abbas, Iran, emails: rahnamae57@gmail.com (E. Rahnama), O.bazrafshan@hormozgan.ac.ir (O. Bazrafshan)*
[b]*Civil Engineering Department, Faculty of Engineering, Kharazmi University, Tehran, Iran, email: fardi@khu.ac.ir*
[c]*School of Mathematical and Statistical Sciences, Southern Illinois University, Carbondale, IL, USA, email: ysamadi@siu.edu*

Received 13 July 2020; Accepted 24 Decexmber 2020

ABSTRACT

Regarding water consumption for various usage, water quality management is a significant part of water management. Future prediction of water quality parameters is necessary for the planning of water quality management. Sodium adsorption rate (SAR) is a parameter, which has a significant role in irrigation. In the present study, we compared the forecasting of SAR of water in Aras, Sepidrud, Karun, and Mond Rivers, Iran, using autoregressive integrated moving average (ARIMA) time series and radial basis functions (RBF) neural network. We found ARIMA (0,1,1) × (0,1,1)12, ARIMA (0,1,1) (0,1,1)12, ARIMA (0,1,1), and ARIMA (0,1,2) × (1,1,1)12 with minimum Akaike's Information Criterion (AIC) of –1.7127, –1.8177, 2.9317, and 12.44 for SAR prediction of Aras, Sepidrud, Karun, and Mond Rivers, respectively. The residual of the mentioned ARIMA models was independent (p-value > 0.05). Using RBF neural network for SAR forecasting of Aras, Sepidrud, Karun, and Mond Rivers with normalized data, we obtained proper training and testing. Mean squared error (MSE are between 0.00026 and 0.0255) and mean bias error (MBE are between 0.01566 and 0.0612) for training is very low. Likewise, coefficient of determination (R^2 are between 0.907 and 0.960), index of agreement (IA are between 0.981 and 0.999), and the Nash–Sutcliffe efficiency (E are between 0.964 and 0.999) approach to 1, which describes the reliability of the model's performance. Thirty-six months RBF neural network and 12 months of ARIMA time series of SAR forecasting for Aras River comparatively match to the measured data and forecast error of both RBF and ARIMA were comparable. We compared forecast errors of the ARIMA time series and RBF neural network for SAR forecasting of Sepidrud, Karun, and Mond Rivers; the results presented that RBF neural network is more reliable than ARIMA for the predicting of SAR.

Keywords: Sodium adsorption rate forecasting; Autoregressive integrated moving average; Radial basis function neural network

* Corresponding author.

1. Introduction

Predicting water quality in rivers is significant in the planning of water quality management. Since for instance, any management decision for crop cultivation and planting, the future water quality level of the river is necessary. One of the important water quality parameters for agricultural purposes is the sodium adsorption rate (SAR). Because it affects soil infiltration. Numerous researchers have worked to determine the effect of SAR on soil infiltration. The rate of infiltration usually reduces with either sodium content associated with calcium and magnesium [1]. Shainberg and Letely [2] also stated that rises in sodium level will decrease the amount of infiltration. Likewise, numerous other studies described that an increase of SAR reduced infiltration in the soil [3–5]. Consequently, future predicting of SAR is significant.

Mathematical methods have been used often to forecast water quality [6]. Some techniques are available for projecting SARs in the water of the river, for instance, the Box–Jenkins time series, artificial neural network (ANN), and Bayesian time series. Several usages of the first technique are: Jayawardena and Lai [7] studied the water quality of the Pearl River applying the Box–Jenkins time series. Sun and Koch [8] employed the Box–Jenkins time series, containing autoregressive integrated moving average (ARIMA) time series, to project the Apalachicola Gulf water quality. They determined water level variations in tide conditions, inflowing streams water quality, local precipitation, the velocity of wind, and wastewater discharge into the water body was as the influential elements in the Gulf water quality. Asadollahfardi [9] used different Box–Jenkins time series to investigate Tehran's surface water quality, Iran. Kurnc et al. [10] applied Thomas–Fiering and ARIMA to forecast water quality and discharge of the Yeşilırmak River in Turkey. The mean absolute error and root mean square error (RMSE) described that the Thomas–Fiering model forecasting was more reliable than ARIMA. Asadollahfardi et al. [11] investigated upstream and downstream of the Latian Dam water quality, Iran, applying the ARIMA time series. Abudu et al. [12] used ARIMA, transfer function-noise (TFN), and ANNs methods to predict the monthly TDS of the Rio Grande in El Paso, Texas. Ranjbar and Khaledian [13] used the ARIMA method to predict water quality parameters of the Sefīd-Rūd River. They reported satisfactory outcomes. Arya and Zhang [14] applied ARIMA models to predict dissolved oxygen and temperature in four water quality monitoring stations at Stillaguamish River Washington, US. They showed the suitability of the Box–Jenkins time series in the predicting. Salmani and Salmani Jajaei [15] studied TDS of Karun River in the southeast of Iran using ARIMA, and they also used a transfer function model to formulate TDS as a function of water flow volume. Some researchers have performed using different types of ANNs to forecast SAR in rivers such as Asadollahfardi et al. [16], Azad et al. [17], Sattari et al. [18], Singh [19], and Al-Obaidi et al. [20]. However, they did not compare ARIMA and ANN in predicting SAR.

There are a few studies, which compared the ANN with ARIMA. However, they have not related to water quality such as Aslanargun et al. [21], that they forecasted tourist arrival, Dı´az-Robles et al. [22] projected particulate matter in urban areas, Yassen [23] predicted economy, Adebiyi et al. [24] worked about the stock market, Ighravwea and Anyaeche [25] predicted port productivity and berth effectiveness and Li and Li [26] studied the incidence of AIDS.

Only Asadollahfardi et al. [27] used the ARIMA time series model and ANN in the framework of a multilayer perceptron (MLP) to project the total dissolved solids (TDS) in the Zāyandé-Rūd River, Esfahan province, Iran. They stated the results of MLP were more reliable than the Box–Jenkins time series to predict TDS in the river. The difference between the present study and the work of Asadollahfardi et al. [27] is we applied different ANFIS methods and ARIMA time series for four rivers, which cover most of the surface water of Iran. The application of the two mentioned methods and comparison between them to forecast SAR have not been reported in the literature.

Therefore, the first objective of the present study was to apply the ARIMA time series and RBF neural network to predict the SAR of Aras, SipidRud, Karun, and Mond Rivers. Second, we compared the result of the forecasting of the two mentioned models to determine the reliability of SAR prediction.

1.1. Study area

Fig. 1 presents the places of four rivers and water quality monitoring stations. Aras, Sefid-rud, Karun, and Mond Rivers are the four rivers, which place in the north, northwest, southwest of Iran. A major part of the surface water and groundwater of Iran flow in the four rivers. The Aras is a river that originates from Turkey and then discharges along the boundaries of Turkey, Armenia, the Nakhichevan part of Azerbaijan, and Iran and it discharges to the Kura River in Azerbaijan. The length of the river is 1,072 km, covering a basin of 102,000 km^2.

Sefid-Rud River is the lengthiest river of northern Iran, which its water enters into the Caspian Sea. The Sefid-Rud River is about 1,000 km long, and its basin covers 56,200 km^2. The mean annual water discharge is 130 m^3/s; nevertheless, it is subject to substantial fluctuation from 600 m^3/s in the spring during the melting of the snow and throughout heavy precipitation. In winter, it decreased to 70–80 m^3/s. The weather condition is wet and subtropical. The mean temperature is between 11°C in January and 26°C in July. Rainfall is around 1,000 mm/y between Elburz mountain and the Caspian Sea, and precipitation in the southern area is between 200 and 400 mm.

The Karun River is placed in the southwestern of Iran with an area of around 67,500 km^2. The longitude of the basin is 48° to 52° east and the latitude of the basin is 30° to 34°15′N. The yearly precipitation is around 550 mm [20]. The Mond catchment area is placed between the provinces of Bushehr and Fars in Iran. The catchment area covers 47,654 km^2. The geographical coordination is 54°10′ to 54°38′E and 27°20′ to 30°05′N. The mean precipitation in the catchment area is 350 mm, near 60% of which happen in winter. As stated by Ministry of Energy data, the mean yearly flow of the river at the outlet of the catchment area is 1,720 million m^3, of which the Firuzabad River, with an average flow of 400 million m^3, is the greatest important

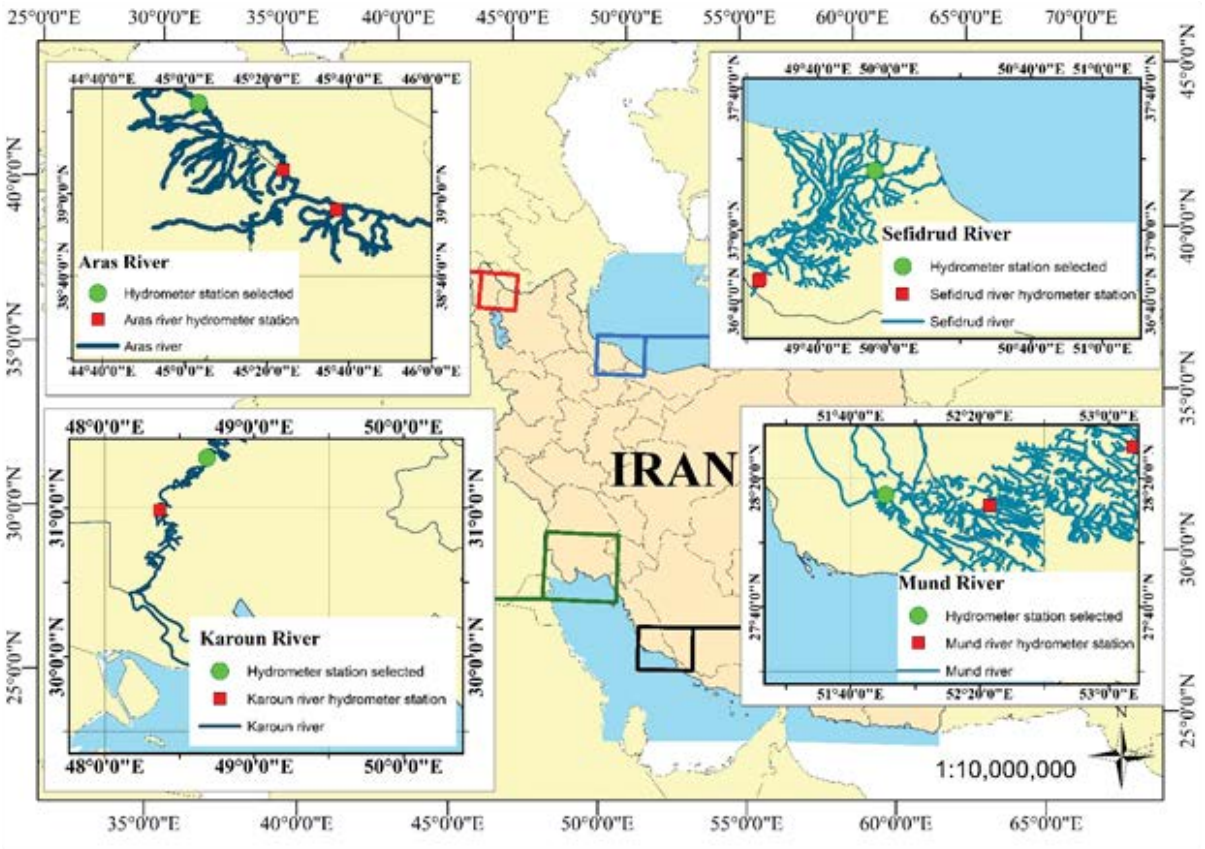

Fig. 1. Study area of the four rivers, including Aras, Sefid-Rud, Karun, and Mond [29].

water supply flow in the basin [28,29]. These four catchment areas could be representing the main area of Iran's surface water supply.

2. Materials and methods

2.1. Box–Jenkins methodology of time series modeling

Decomposition of time series data into its components, while being instructive and revealing, needs considerable time and effort is a hard job. Besides, it sources more errors by the accumulation of component errors [9]. To avoid these problems, Box and Jenkins [30] developed a new method, which basically, does the same task. It joins all concepts mentioned above. In this technique, we apply some transformations such as simple and seasonal differences, the trends, and seasonal and cyclical components on existing data. Next, a family of models is entertained for the transformed data that is supposed to be as simple as possible.

The Box–Jenkins method is based on the idea of a stationary time series shortly described in the next section.

2.2. Classification of time series models

The general seasonal ARIMA model of order (P, Q, D, d, p, q) is:

$$\varnothing_P(B)\varnothing_P(B^L)\nabla_L^D\nabla^d y_t^* = \delta + \theta_q(B)\theta_q(B^L)\alpha_t \quad (1)$$

$$\varnothing_P(B) = \left(1 - Q_1 B - Q_2 B^2 - \ldots - Q_P B^P\right) \quad (2)$$

$$\varnothing_P(B^L) = \left(1 - \varnothing_{1,L} B^L - \varnothing_{2,L} B^{2L} - \ldots - \varnothing_{P,L} B^{PL}\right) \quad (3)$$

$$\theta_q(B) = \left(1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q\right) \quad (4)$$

$$\theta_Q(B^L) = \left(1 - \theta_{1,L} B^L - \ldots - \theta_{Q,L} B^{QL}\right) \quad (5)$$

where $\varnothing_P(B)$, $\varnothing_P(B^L)$, $\theta_q(B)$ and $\theta_Q(B^L)$ are the non-seasonal autoregressive operator of order p, seasonal autoregressive operators order P, non-seasonal moving average operator q, and seasonal moving average operator order Q, respectively. The δ is a constant of the model, which μ is the real mean of stationary time series. The δ can be calculated by Eq. (6):

$$\delta = \mu\varnothing_P(B)\varnothing_P(B^L) \quad (6)$$

where B is the backward shift operator $B^k y_t = y_{t-k}$, ∇^d equals to the backward difference operator and $\varnothing\varnothing_1, \varnothing_2, \ldots, \varnothing_P, \varnothing_{1,L}, \varnothing_{2,L}, \ldots, \varnothing_{P,L}, \theta_1, \theta_2, \ldots, \theta_q, \theta_{1,L}, \theta_{2,L}, \ldots, \theta_{Q,L}$ are autoregressive and moving average orders, which are unknown that it should be estimated by the sample data; α_t is a random variable with mean zero and constant

variance. The a_{ts} are assumed to be independent and represent random error or random shocks.

The number of lags for the ARIMA(p,d,q) models is determined based on the Akaike information criterion (AIC) or the Bayesian information criterion (BIC). First, the integrated parameter d is specified by an adequate number of differencing to make the series stationary. Then, the ACF, PACF, and EACF plots are utilized to select the appropriate number of lags for the model. In the end, the minimum values of AIC and BIC are used to determine the number of lags of the model from the list of lag orders suggested by ACF, PACF, and EACF plots.

Fig. 2 indicates a flowchart of the description of the predictive of ARIMA models.

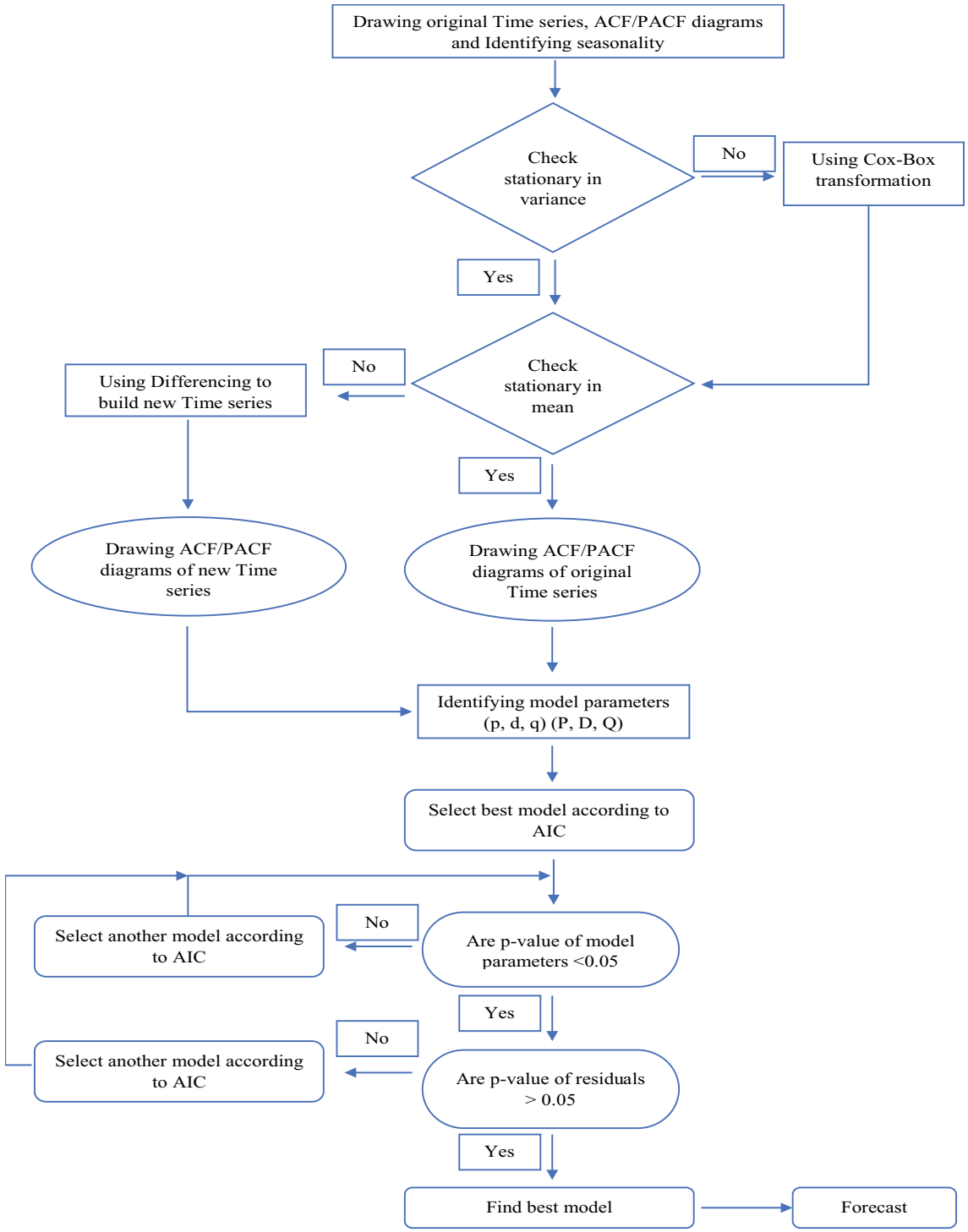

Fig. 2. Flowchart of the description of the predictive of ARIMA models.

For more detail refer to Box and Jenkins [30]. Because the objective of the present study was to develop an ARIMA time series to predict the future of the SAR of the four rivers. Therefore, we applied the AIC to examine future prediction instead of the coefficient of determination (R^2), which well fits the past value. AIC is the basic formula, which is defined by the following equation:

$$\mathrm{AIC} = -2\left(\log - \mathrm{likelihood}\right) + 2K \tag{7}$$

where K is the number of model parameters. Log-likelihood is a measure of model fit. The higher the number, the better the fit. We used Minitab 19 software version 19.1.1 for analysis of the ARIMA time series. We used Mintab19.1.1 to analyze the ARIMA time series.

2.3. Radial basis function

The radial basis function (RBF (neural network contains a simple architecture. Their structure includes an input layer, one hidden layer, and one output layer, which at each output node permits a linear combination of the outputs of the hidden-layer nodes. Each neuron in the hidden layer is made of RBF centered on a point the same dimensions as forecaster parameters.

2.4. Selecting input parameters

To choose the input parameters to the RBF, we attempted to come after the protocol in the paper of Wu et al. [31], which expresses the input parameters should be independent and model validation considering replicative and structural validity. Consequently, considering the accessibility and existing data, we selected six parameters as input to RBF neural network, which were discharge, pH, Cl, SO_4, HCO_3, and electrical conductivity (EC); and output parameters were SAR.

We selected the number of data required for neural network training from the available data using trial and error [32]. Earlier investigators have chosen between 70% and 80% of the data for the training of the ANN [33–37,17]. From 180 monthly available data, we applied 80% for training 20% for testing of the network.

2.5. Data preparation

Before using data as input to the RBF neural network, data can be scaled between zero and one. We applied Eq. (8) to standardize the input data. Lastly, the output data was altered to the primary scale for comparison with the measured data:

$$O_S = \frac{O_i - O_{\min}}{O_{\max} - O_{\min}} \times 2 - 1 \tag{8}$$

where O_S and O_i are the scaled and the measured amounts of the parameters, respectively. $O_{\min}$ and $O_{\max}$ are the minimum and maximum amounts of the parameters, respectively.

By applying a Gaussian basis function network in the hidden layer (Eq. (9)) of the RBF neural network, the RBF simulates the unknown water quality parameter and linear activation functions in the output layer:

$$f\left(x\right) = e^{-x^2/2\sigma^2} \tag{9}$$

where x is the weighted sum of inputs to the neuron; ƀ is the sphere of influence or the basis function width, and $f(x)$ = the corresponding output of the neuron [38].

2.6. Training

Training of RBF has two steps. First, the basic functions are established using an algorithm to cluster data in the training set. Kohohen self-organizing maps (SOMs) or a *k*-means clustering algorithm is typically used. Kohohen SOMs is a method of "self-organizing" RBF neural network, which learns to differentiate patterns within input data [39]. Then a SOM will cluster input data based on perceived patterns without taking to be accorded a corresponding output response. *K* describes clustering as comprises the organization of all objects into a preset number of groups by minimizing the total squared Euclidean distance for each object regarding its adjacent cluster center. Other methods, for instance, orthogonal least squares and Maxi Min algorithms were used [40]. Next, the weights joining the hidden and the output layer are calculated directly using the simple matrix inversion and multiplication. The straight computing of weights in an RBF makes it far faster to train than an equivalent MLP neural network [38].

We trained a large network and after that prune with validation [41] applying orthogonal least squares [42,43]. Firstly, we performed a growing scheme that iteratively adds new hidden nodes to full-trained neural networks. After that, a non-heuristic one-pass pruning method was carried out, which used orthogonal least squares. Derived from pruning, a one-pass method was developed for generating the validation error vs. network size curve. A combined technique was defined in which networks were continually pruned during the growing process. Consequently, the hidden nodes were ordered based on their usefulness, and the least useful nodes were eliminated [41].

We selected the Gaussian activation function with variable *r* (the radius or standard deviation) and *C* (the center or average is taken from the input space) for each RBF node in the hidden layer.

For each node in the hidden layer, we obtained the parameter of *r* and *c*. Then we calculate the Euclidean distance from the point being evaluated to the center *c* of each neuron. While the RBF center was obtained, we computed the width of each RBF unit applying *k*-nearest neighbors' algorithm. Some *K* was selected, and for each center, the *k* nearest center was obtained.

The RMSE distance between the current cluster center and its *k*-nearest neighbors was computed.

2.7. Error assessment in training

To control the usability of the RBF neural network, two statistical criteria were applied, that is, RMSE and mean

bias error (MBE). The MBE indicates if the model forecasting in the training process overestimate (MBE > 0) or underestimates (MBE < 0). The best score is MBE = 0. Eqs. (10) and (11) present RMSE and MBE [44]:

$$\mathrm{RMSE} = \sqrt{n^{-1}\sum_{i=1}^{n}\left(F_i - M_i\right)^2} \tag{10}$$

$$\mathrm{MBE} = n^{-1}\sum_{i=1}^{n}F_i - M_i \tag{11}$$

where F_i and M_i are the forecasted and the measured amount of the parameters respectively, and n is the number of data.

2.7.1. Model efficiency

To check the performance of the RBF neural network, the coefficient of determination (R^2) [45] was computed as stated by Eqs. (12) and (13):

$$R^2 = \left\{\frac{\Sigma\left(M_i - \bar{M}\right)\left(F_i - \bar{F}\right)}{\sqrt{\Sigma\left(M_i - \bar{M}\right)^2 \times \left(F_i - \bar{F}\right)^2}}\right\}^2 \tag{12}$$

where M_i and F_i are the measured and the projected data of the parameters, respectively. $\bar{M}$ and $\bar{F}$ equal to the mean of measured and the projected data. Likewise, we calculated the index of agreement (IA) to find how nearby the forecasted data was to the observed data [45]:

$$\mathrm{IA} = 1 - \frac{\sum\overline{\left(F_i - M\right)^2}}{\left[\left|F_i - \bar{M}\right| + \left|M_i - \bar{M}\right|\right]^2} \tag{13}$$

The index of agreement (IA) varies from 0.0 (theoretical minimum) to 1.0 (perfect agreement).

The efficiency E, suggested by Nash and Sutcliffe [46], is obtained as one minus the sum of the absolute squared differences between the projected data and observed data normalized by the variance of the observed amounts in the period under study. The variety of E is between 1.0 (perfect fit) and −∞. Eq. (14) describes E [45]. We applied Mathworks Matlab R2020a software for analysis of the RBF ANN:

$$E = 1 - \frac{\sum_{i=1}^{n}\left(M_i - F_i\right)^2}{\sum_{i=1}^{n}\left(M_i - \bar{M}\right)^2} \tag{14}$$

3. Results

3.1. Aras River

From 180 monthly SAR measurements, we selected 168 measured SAR for building the ARIMA model and 12 monthly measured SAR for testing of predicting the model. Fig. 2 indicates the results of the trend analysis plot, Box–Cox plot, time series plot (difference = 1), and trend analysis plot (difference = 1) for SAR of the Aras River.

Since the time-series observations are non-stationary in the mean. First, we evaluated the stationary of time series related to variance, and draw the Box–Cox plot (Fig. 2). According to Fig. 2, the optimal value for the parameter λ was equal to one. Therefore, there was no need to convert the series to stationary, and the time series is stationary in the variance. Thus, using the difference method, we made the data stationary on the mean. For this purpose, we converted the time series to stationary by taking a one-time difference and draw its diagram (Fig. 2). The time series diagram after one-time difference indicates the elimination of the nonstationary in the mean; however, for insurance, we also draw the series trend Fig. 2. After that, the autocorrelation function and partial autocorrelation function plots for SAR of Aras River were plotted (Fig. 3).

According to Fig. 3, after the series was stationary, there was still a seasonal parameter, and for this purpose, we eliminated the seasonal parameter by using the differences operation and possible models examined on the observed data. However, according to Fig. 3, the PACF figure tends to zero earlier than the ACF diagram, which may indicate the absence of AR and SAR factors in the sample.

3.2. Model development

As mentioned previously, the number of SAR monthly observations data for the building model was 168, and the length of seasonality was 12. Forecasting SAR was performed for12 months. As indicated in Table 1, we examined six different models and computed AIC for them.

The residuals examination (modified Portmanteau test).

We carried out modified Portmanteau test for ARIMA (0,1,1) × (0,1,1)12 model with constant for residuals correlation. According to Table 2, the p-value obtained for residual ARIMA (0,1,0) × (0,1,1)12 is less than 0.05, which confirms the hypothesis of the residuals dependent.

Despite minimum AIC result, we cannot select ARIMA (0,1,0) × (0,1,1)12 model. Therefore, the residuals of ARIMA (0,1,1) × (0,1,1)12 with constant was checked. According to Table 2, the p-value obtained for residual in all delays is more than 0.05, which shows the residual of the model is independent.

3.3. Sefid-Rud river

The number of SAR data for the Sefid-Rud river was equal to to180. We opted for 168 measured SAR for building the ARIMA model and 12 monthly measured SAR for testing of model forecasting.

Fig. 4 indicates the time series trend of the Sefid-Rud River. Due to the declining trend of the main data, the time series indicates non-stationary in the mean. To change data to stationary, Box–Cox of data was drawn, then a suitable value of λ was computed (Fig. 4). According to the Cox–Box figure, we achieved λ equal to 0.5 and failure to place the insignificant amount 1 in the 95% confidence interval. The time series should be changed to stationary in the variance proportional to the value obtained for the

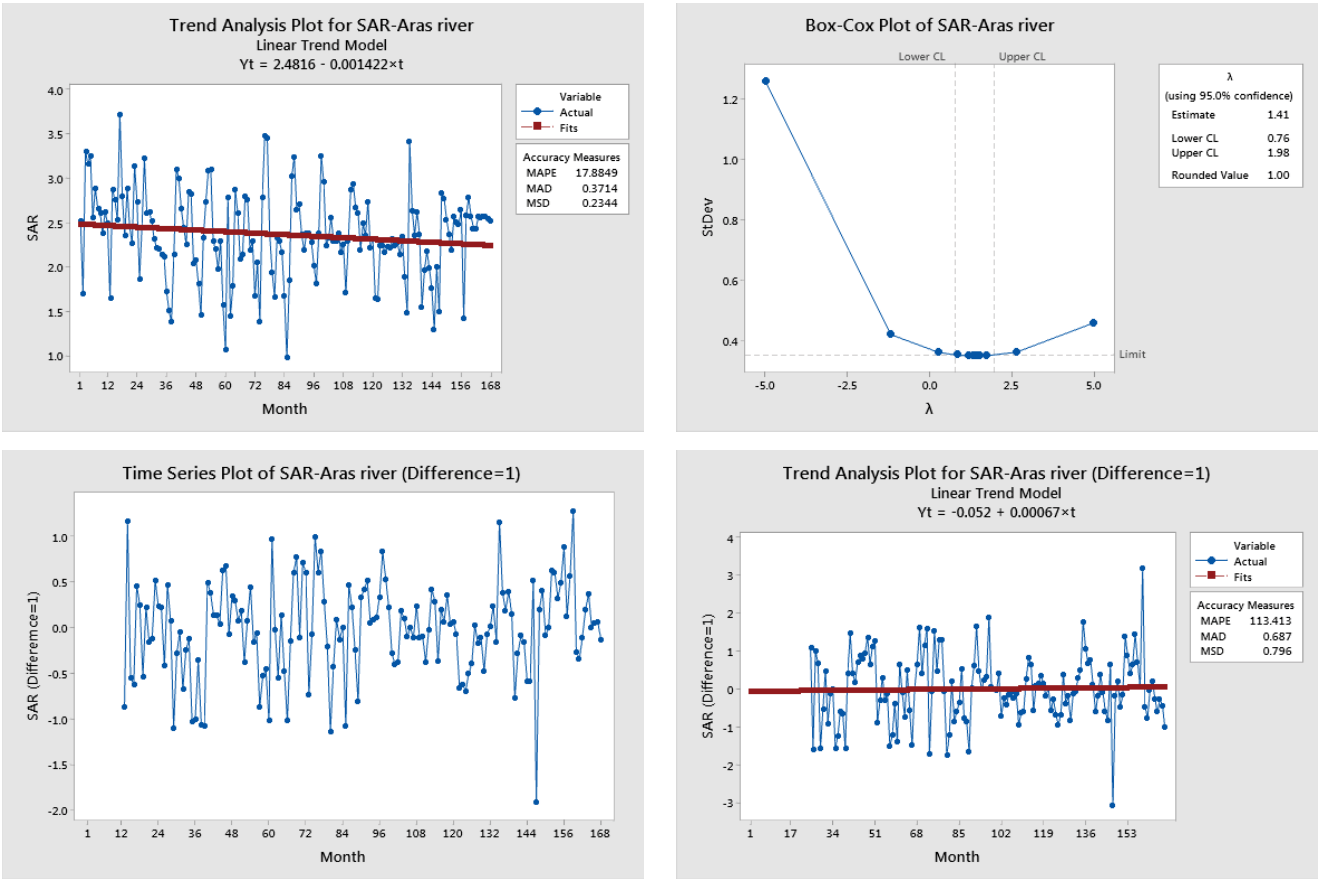

Fig. 2. Trend analysis plot, Box–Cox plot, time series plot (difference = 1) and trend analysis plot for SAR of the Aras river (difference = 1).

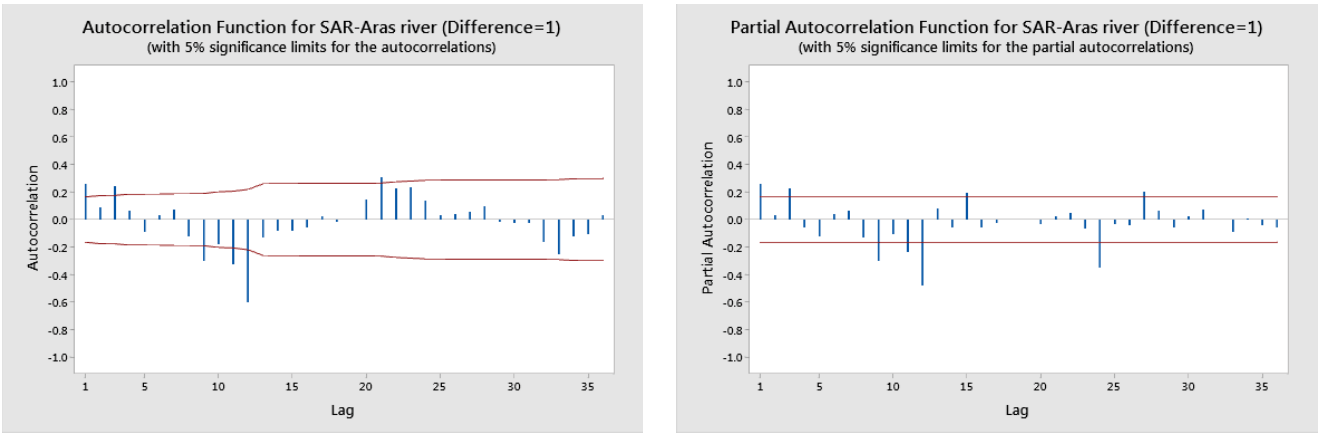

Fig. 3. Autocorrelation function and partial autocorrelation function plots for SAR-Aras river (difference = 1).

parameter λ using the transformation $\sqrt{(x\_t)}$. Then, it went through the next steps of identifying the model.

As shown in Fig. 4, the series is stationary in variance after the performance of the Box–Cox transformation, and we reached λ equal to 1. We performed one difference in the SAR1 time series to reach the stationary time series in the mean. As indicated in Fig. 4, finally the time series is stationary related to mean and variance. Fig. 5 presents the autocorrelation function and partial autocorrelation function of SAR2 of the Sefid-Rud River.

Considering Fig. 5, We examined different nonseasonal ARIMA time series. However, the residual of the models was dependently distributed. Therefore, we selected different types of seasonal ARIMA time series (Table 3), and

Table 1
Comparison of the results of fitting different models to the data

Model	AIC	Model parameters
ARIMA (2,1,1) × (2,1,2)12 with constant	2.0818	AR (2) is not OK or it is not statistically significant
ARIMA (1,1,1) × (2,1,2)12 with constant	0.2051	SAR (2), SMA (2) are not OK or it is not statistically significant
ARIMA (0,1,1) × (1,1,1)12 with constant	–1.7013	SAR (1) is not OK or it is not statistically significant
ARIMA (0,1,1) × (0,1,1)12 with constant	–1.7127	OK
ARIMA (0,1,1) × (0,1,0)12 with constant	–0.6202	OK
ARIMA (0,1,0) × (0,1,1)12 with constant	–2.7277	OK

Table 2
Modified Box–Pierce (Ljung–Box) chi-square statistic

ARIMA (0,1,0) × (0,1,1)12				
Lag	12	24	36	48
Chi-square	49.29	58.75	65.06	87.25
DF	10	22	34	46
p-value	0.000	0.000	0.001	0.000
ARIMA (0,1,1) × (0,1,1)12				
Lag	12	24	36	48
Chi-square	13.60	32.41	38.68	52.48
DF	9	21	33	45
p-value	0.137	0.053	0.229	0.207

the residuals are independently distributed. According to Table 3, we selected ARIMA (0,1,1) (0,1,1)12 with a minimum of ACI of –1.8177 for future predictions.

3.4. Karun River

Fig. 6 indicates the process of converting SAR data of the Karun River to be stationary in mean and variance. Due to the upward trend of the time series of SAR on the Karun River (Fig. 6), the series is not stationary in mean; therefore, it was necessary to check the stationary of variance. The Box–Cox diagram of the time series was drawn and obtained a suitable λ (Fig. 6).

As presented in Fig. 6, the optimal value for the parameter λ is equal to 0.5. Therefore, it is necessary to convert the series using the operator $\sqrt{(x\_t)}$. However, considering the confidence limit of 0.95% and the value of $\lambda = 1$ is also in this range, we concluded that no transformation was necessary for the stationary of the variance of this series, and the time series was stationary in the variance. To reach time series to be stationary in mean, the first difference was performed. Fig. 6 also indicates the SAR time series of Karun River are stationary in mean and variance.

After the time series reached stationary, the ACF and PACF diagram was drawn for SAR2 to identify model parameters (Fig. 7). As indicated in Fig. 7a a seasonal factor was observed.

To determine the dependence of residual, the chi-square statistic Box–Pierce (Ljung–Box) test was used. If the *p*-value of the residues is higher than 0.05, it means that the residues are independent and the model is appropriate. Also, if the *p*-value of the model parameters is less than 0.05; the model is not acceptable. First, the non-seasonal ARIMA time series model was examined. The *p*-value was greater than 0.05, and the model ARIMA (0,1,1) with AIC of 2.9317 is acceptable (Table 4).

3.5. Mond River

Fig. 8 indicates the process of converting SAR data of the Karun River to be stationary in mean and variance. The time series trend diagram describes the non-stationary in the mean of the main data. For stationary of SAR time series, we checked stationary of time series of variance and plot Box–Cox series to determine λ. As indicated in the Cox–Box transformation λ equal to one. Therefore, the time series is stationary of variance. To reach the time series to be stationary in mean, the first difference was performed (Fig. 8). Fig. 8 indicates the SAR time series of Mond River are stationary in mean and variance.

Fig. 9 presents the autocorrelation and partial autocorrelation functions of SAR data for Mond River. According to Fig. 9, we identify a suitable model.

3.6. Final model

Considering Fig. 9, autocorrelation function, and partial autocorrelation function, we applied the different models to the SAR data, and the results are indicated in Table 5. According to Table 5, ARIMA (0,1,2) (1,1,1)12 was selected because of the less value of AIC, and residual of data according to Box–Pierce (Ljung–Box) chi-square statistic is also independent.

3.7. RBF neural network results

Table 3 presents errors in the training and performance of the models for the Aras, Sefid-Rud, Karun, and Mond River using RBF. MSE of Karun and Mond Rivers are lower than MSE of Aras and Sefid-Rud Rivers. However, the rate of MSE is reasonable for the four rivers, which indicates the accuracy of training of the RBF neural network for the SAR. MBE describes the adequacy of training of the RBF neural network of SAR in the four rivers, which are low and is close to zero (Table 6). The results approximately indicate the adequacy of training of the RBF neural network in the four rivers. R^2, IA, and E of Sefid-Rud and Karun rivers are higher than Aras and Mond Rivers, which presents the better model performance of the first

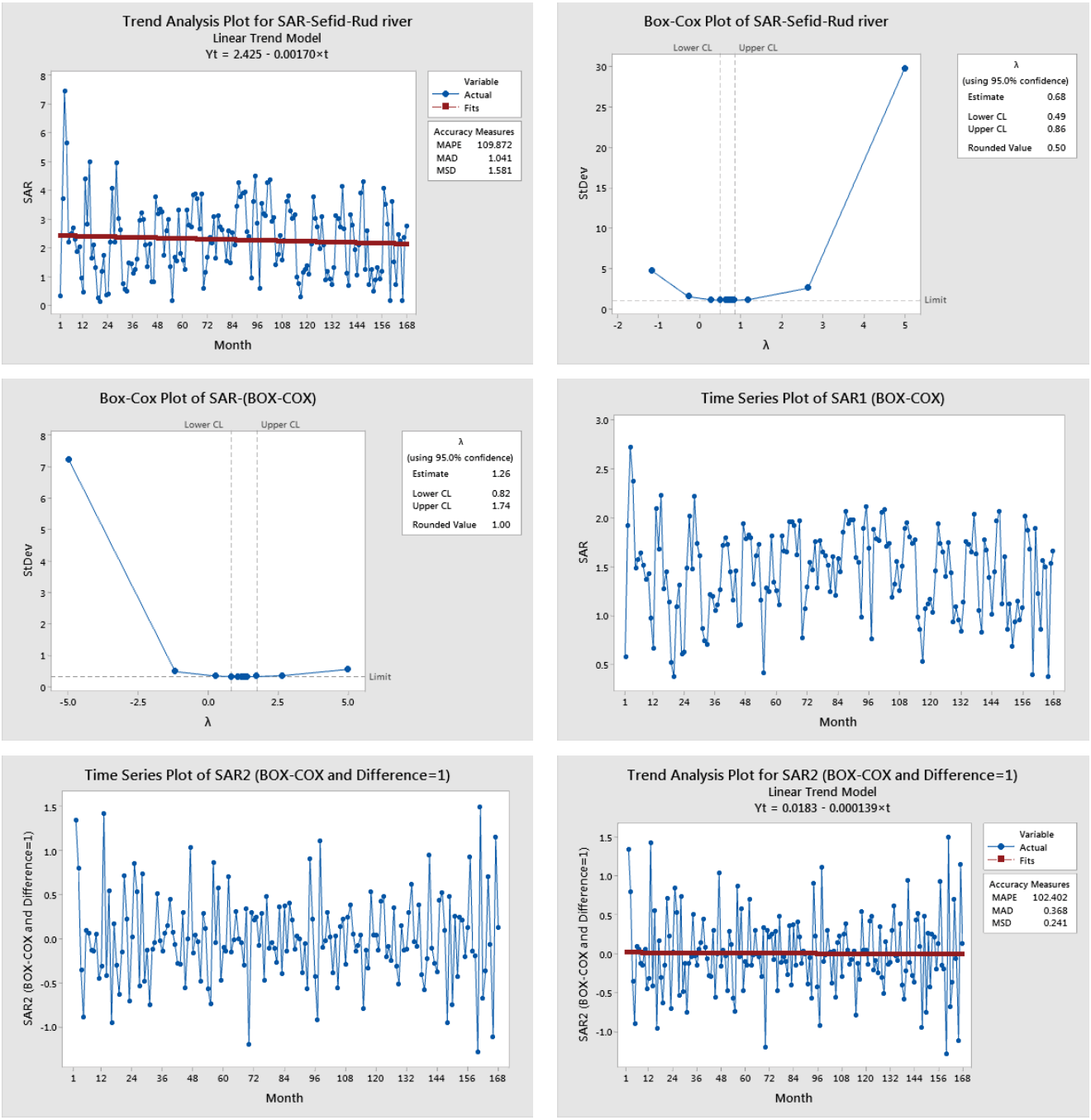

Fig. 4. Trend analysis plot, Box–Cox plot, time series plot (difference = 1) and trend analysis plot for SAR of the Sefid-Rud river (difference = 1).

two rivers. As a general conclusion, the RBF neural network of the four rivers for SAR prediction has good performance because of R^2, IA, and E re close to 1.

3.8. Comparison ARIMA and RBF for SAR prediction of the four rivers

Tables 7 and 8 indicate the forecast error for predictions of SAR from ARIMA and RBF neural network for Aras and Sepid rud river. As presented in Table 7, both ARIMA and RBF models reached acceptable forecast errors for Aras River. However, in the Sefid-Rud river, the forecast error of RBF much lower than ARIMA, which means the RBF neural network is more suitable than ARIMA for the prediction of SAR (Tables 7 and 8). Another advantage of the RBF neural network is in model building 144 monthly data were used, and future predicting was for 36 months. However, for the ARIMA time series 168 monthly data were used for building, and only a 12-month prediction was predicted. For Karun and Mond rivers, we reached similar results as

Table 3
Different ARIMA time series models of SAR for the Sepidrud River

Model	AIC	Box–Pierce (Ljung–Box) chi-square statistic				
ARIMA (0,1,1) (0,1,1)12	–1.8177	Lag	12	24	36	48
		Chi-square	12.99	18.63	28.64	39.52
		DF	10	22	34	46
		p-value	0.224	0.668	0.727	0.739
ARIMA (1,1,1) (0,1,1)12	0.1344	Lag	12	24	36	48
		Chi-square	8.16	13.42	20.43	32.09
		DF	9	21	33	45
		p-value	0.518	0.893	0.957	0.926
ARIMA (0,1,2) (1,1,0)12	0.5793	Lag	12	24	36	48
		Chi-square	8.25	18.24	27.13	38.42
		DF	9	21	33	45
		p-value	0.510	0.634	0.754	0.745
ARIMA (0,1,2) (0,1,1)12	0.1284	Lag	12	24	36	48
		Chi-square	9.23	14.14	21.48	34.71
		DF	9	21	33	45
		p-value	0.417	0.864	0.939	0.866
ARIMA (0,1,2) (2,1,0)12	0.3582	Lag	12	24	36	48
		Chi-square	10.90	16.60	27.75	37.36
		DF	8	20	32	44
		p-value	0.208	0.679	0.682	0.750
ARIMA (1,1,2) (1,1,0)12	2.5761	Lag	12	24	36	48
		Chi-square	7.01	17.92	24.82	34.14
		DF	8	20	32	44
		p-value	0.536	0.592	0.813	0.857
ARIMA (1,1,1) (1,1,0)12	0.5740	Lag	12	24	36	48
		Chi-square	7.18	18.26	25.16	34.03
		DF	9	21	33	45
		p-value	0.618	0.632	0.834	0.884
ARIMA (3,1,0) (0,1,1)12	2.2614	Lag	12	24	36	48
		Chi-square	13.09	18.24	25.40	35.33
		DF	8	20	32	44
		p-value	0.109	0.571	0.789	0.821
ARIMA (3,1,0) (1,1,0)12	1.7658	Lag	12	24	36	48
		Chi-square	8.79	21.51	28.06	37.89
		DF	8	20	32	44
		p-value	0.360	0.368	0.666	0.730

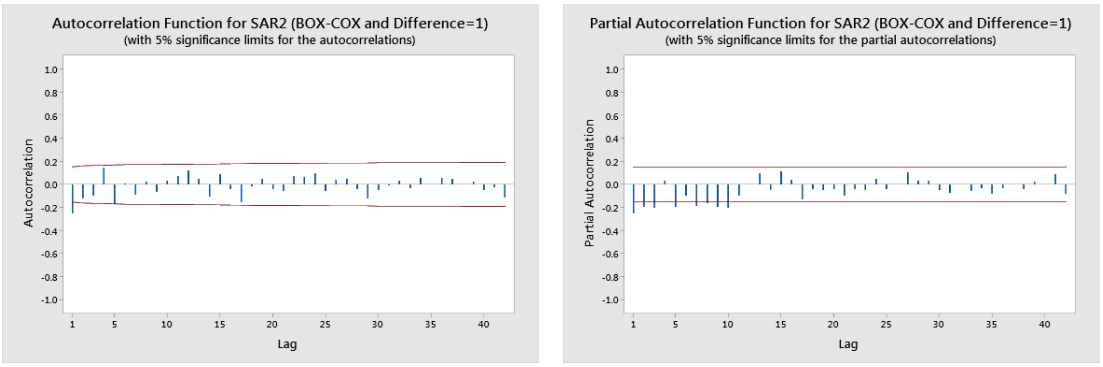

Fig. 5. Autocorrelation function and partial autocorrelation function plots for SAR-Aras river (difference = 10).

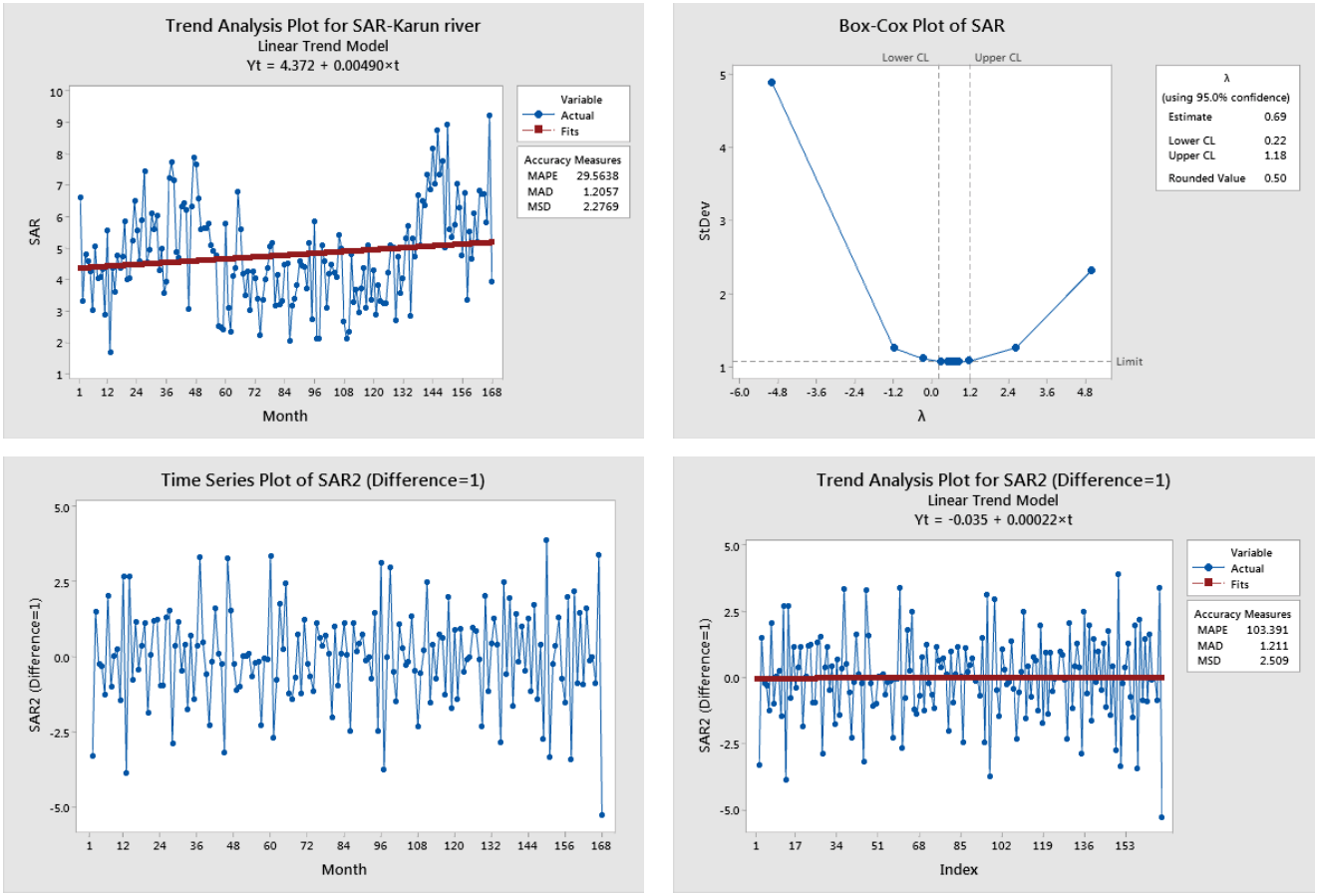

Fig. 6. Process of converting SAR data of the Karun River to be stationary in mean and variance.

Table 4
different nonseasonal ARIMA time series models for Karun River

Model	ACI	Box–Pierce (Ljung–Box) chi-square statistic				
ARIMA (0,1,1)	2.9317	Lag	12	24	36	48
		Chi-square	10.97	19.44	32.97	40.82
		DF	11	23	35	47
		p-value	0.446	0.675	0.566	0.725
ARIMA (3,1,0)	6.9444	Lag	12	24	36	48
		Chi-square	11.88	20.01	36.16	48.27
		DF	9	21	33	45
		p-value	0.220	0.521	0.323	0.342

the Sefid-Rud river. Therefore, RBF neural network much better for future prediction than the ARIMA time series for the Sefid-Rud river, Karun, and Mond rivers.

Fig. 10 presents the comparison of SAR forecasting using the ARIMA time series and RBF neural network for Aras, Sefid-Rud, Karun, and Mond rivers. According to Fig. 10, 36 months RBF neural network and 12 months forecasting of the ARIMA time series of SAR predictions for the Aras river relatively matches the measured data. However, some of the measured data are beyond the RBF model prediction. RBF neural network prediction of SAR for Sefid-Rud River is very match to the measured data, which indicates the reliability of RBF forecasting of SAR for Sefid-Rud River. However, the reliability of the 12 month SAR projection of the ARIMA time series for the Sefid-Rud river is not the same as the RBF neural network. Thirty-six month SAR prediction of RBF neural network of Karun River as indicated in Fig. 10 except a few points match to the measured data. However, the forecasting of the non-seasonal ARIMA time series of Karun River is a horizontal line, which does not describe the seasonality of measured data. Therefore, the RBF neural network perdition more reliable than ARIMA

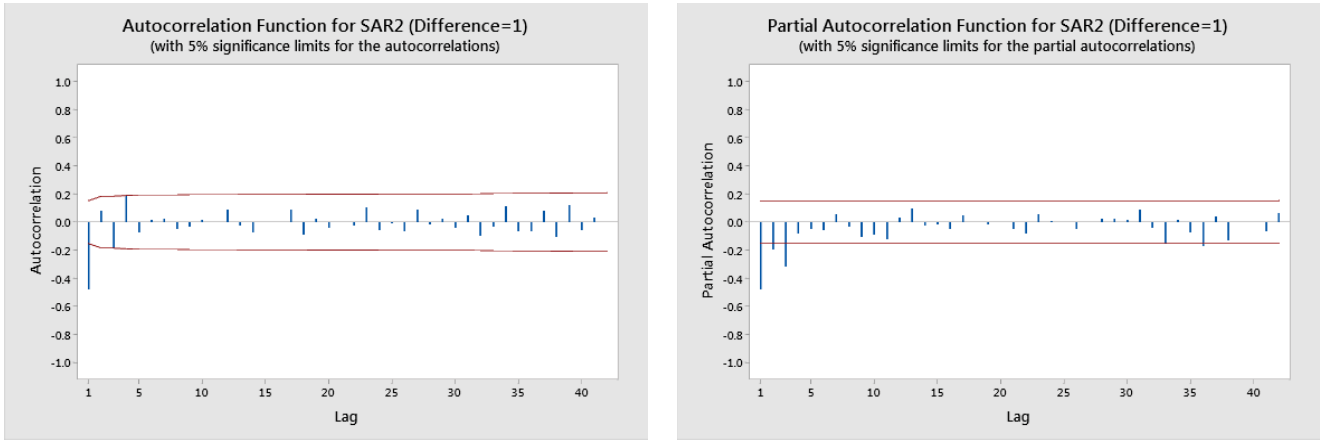

Fig. 7. Autocorrelation and partial autocorrelation plots for SAR of the Karun River.

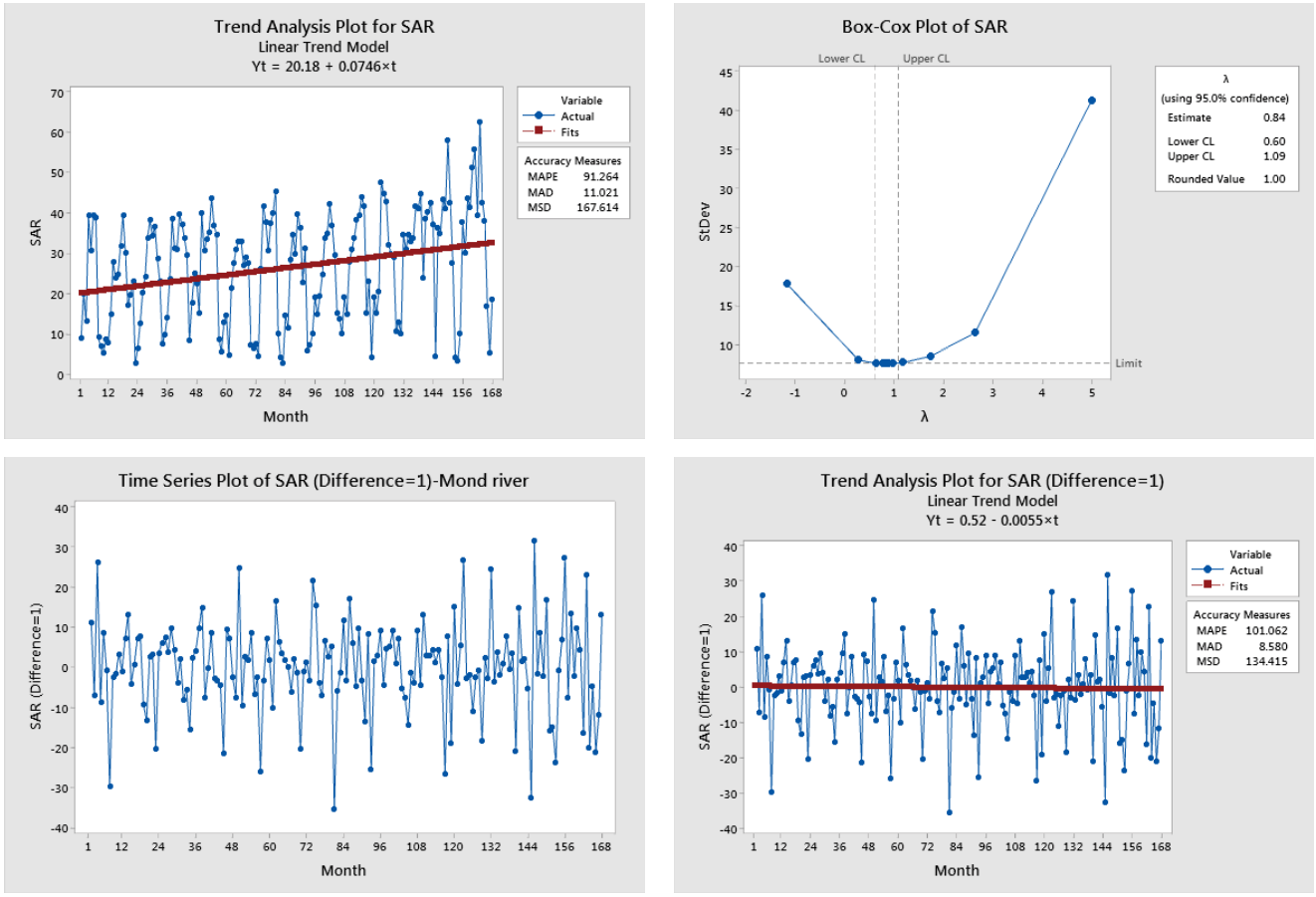

Fig. 8. Process of converting SAR data of the Mond River to be stationary in mean and variance.

prediction in this river. The SAR prediction of Mond River using the RBF neural network is very well-matched to the measured data, which describe the reliability of RBF neural network prediction. ARIMA time series prediction for 12 months in Mond River reaches relatively good prediction, but not the same as RBF neural network projection. In conclusion, the RBF neural network reached a better prediction related to ARIMA time series in the four rivers.

4. Discussion

Several mathematical models have been developed for forecasting water quality in surface water to help water quality management for planning and performance. The limited studies have carried out using different types of ANN for the SAR predictions of water in rivers, which include the working of Asadollahfardi et al. [16], Azad et

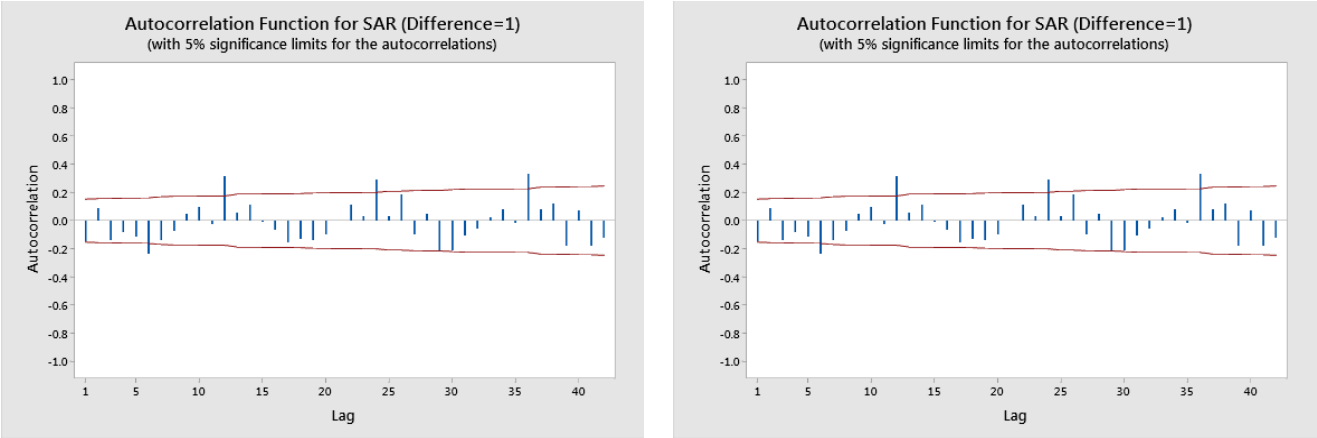

Fig. 9. Autocorrelation function and partial plots of for SAR-Karun river.

Table 5
Seasonal Box–Jenkins time series for the SAR

Model	ACI	Box–Pierce (Ljung–Box) chi-square statistic				
ARIMA (0,1,2) × (1,1,0)12	12.91	Lag	12	24	36	48
		Chi-square	13.72	29.80	45.71	54.65
		DF	9	21	33	45
		p-value	0.133	0.096	0.070	0.154
ARIMA (0,1,2) × (1,1,1)12	12.44	Lag	12	24	36	48
		Chi-square	13.71	21.92	36.12	47.36
		DF	8	20	32	44
		p-value	0.090	0.345	0.282	0.337
ARIMA (1,1,1) × (0,1,2)12	12.47	Lag	12	24	36	48
		Chi-square	13.56	20.97	33.80	43.70
		DF	8	20	32	44
		p-value	0.094	0.399	0.380	0.484
ARIMA (1,1,1) × (1,1,1)12	12.45	Lag	12	24	36	48
		Chi-square	13.65	21.11	35.06	45.59
		DF	8	20	32	44
		p-value	0.091	0.391	0.325	0.406
ARIMA (2,1,1) × (0,1,2)12	14.47	Lag	12	24	36	48
		Chi-square	13.38	21.69	34.91	45.52
		DF	7	19	31	43
		p-value	0.063	0.300	0.287	0.368
ARIMA (2,1,1) × (1,1,1)12	14.23	Lag	12	24	36	48
		Chi-square	12.45	20.81	35.37	47.16
		DF	6	18	30	42
		p-value	0.053	0.289	0.229	0.270

al. [17], Sattari et al. [18], and Singh [19], Al-Obadi et al. [20], and Rahnama et al. [29]. The present study differs from the mentioned researches because we compared the results of the forecasting of SAR of the four rivers using ARIMA and RBF neural networks, which have not been reported in the literature. As mentioned in the introduction section, a few studies compared ARIMA and ANNs for predictions. However, none of the studies was about water quality. There were about forecasting tourist arrival, particulate matter in urban areas, port productivity, and berth effectiveness, the incidence of acquired immunodeficiency syndrome (AIDS), and the economy.

Jayawardena and Lai [7], Sun et al. [8], Asadollahfardi [9], Kurnc et al. [10], Asadollahfardi et al. [11], Abudu et al. [12], Ranjbar and Khaledian [13], Arya and Zhang [14], and Salmani and Salmani Jajaei [15] applied ARIMA time

Table 6
Error in training and performance of RBF models for the SAR prediction of Aras, Sefid-Rud, Karun and Mond River

Rivers	Model	MSE	MBE	R^2	IA	E
		Training			Testing	
Aras	RBF (normalized)	0.00175	0.06012	0.907	0.989	0.964
Sefid-Rud	RBF (normalized)	0.0255	0.01566	0.96	0.9975	0.990
Karun	RBF (normalized)	0.00026	0.047	0.95	0.999	0.999
Mond	RBF (normalized)	0.0009	–0.06	0.92	0.981	0.995

Table 7
Comparison of forecast error for ARIMA time series and RBF neural network for Aras and Sefid-Rud rivers

Sample period	Aras river						Sefid-Rud river					
	Actual values		Predict values		Forecast error		Actual values		Predict values		Forecast error	
	RBF	ARIMA	RBF	ARIMA	RBF	ARIMA	RBF	ARIMA	RBF	ARIMA	RBF	ARIMA
1	1.296	–	1.472	–	–0.135	–	2.108	–	1.989	–	0.057	–
2	2.004	–	2.052	–	–0.024	–	3.893	–	3.774	–	0.031	–
3	1.492	–	1.632	–	–0.094	–	4.294	–	4.175	–	0.028	–
4	2.83	–	2.729	–	0.036	–	1.259	–	1.140	–	0.095	–
5	2.774	–	2.684	–	0.033	–	2.59	–	2.471	–	0.046	–
6	2.535	–	2.488	–	0.019	–	0.732	–	0.613	–	0.163	–
7	2.366	–	2.349	–	0.007	–	1.248	–	1.129	–	0.096	–
8	2.187	–	2.563	–	–0.172	–	0.473	–	0.354	–	0.252	–
9	2.564	–	2.511	–	0.021	–	0.876	–	0.757	–	0.136	–
10	2.507	–	2.465	–	0.017	–	1.324	–	1.205	–	0.090	–
11	2.479	–	2.442	–	0.015	–	0.907	–	0.788	–	0.132	–
12	2.639	–	2.573	–	0.025	–	1.177	–	1.058	–	0.101	–
13	1.423	–	1.988	–	–0.397	–	4.069	–	3.950	–	0.029	–
14	2.579	–	2.524	–	0.021	–	3.52	–	3.401	–	0.034	–
15	2.78	–	2.688	–	0.033	–	2.833	–	2.714	–	0.042	–
16	2.566	–	2.156	–	0.160	–	0.155	–	0.036	–	0.770	–
17	2.432	–	1.423	–	0.415	–	3.608	–	3.489	–	0.033	–
18	2.432	–	1.403	–	0.423	–	1.498	–	1.379	–	0.080	–
19	2.566	–	1.512	–	0.411	–	1.73	–	0.614	–	0.645	–
20	2.558	–	1.507	–	0.411	–	2.444	–	2.325	–	0.049	–
21	2.566	–	1.515	–	0.410	–	2.24	–	2.121	–	0.053	–
22	2.566	–	2.513	–	0.021	–	0.144	–	0.025	–	0.829	–
23	2.542	–	2.493	–	0.019	–	2.361	–	2.242	–	0.051	–
24	2.512	–	2.469	–	0.017	–	2.766	–	2.647	–	0.043	–
25	1.703	1.703	1.805	1.826	–0.060	–0.072	3.517	3.517	3.398	1.360	0.034	0.613
26	1.84	1.84	1.918	2.091	–0.042	–0.137	4.271	4.271	4.015	1.735	0.060	0.594
27	2.491	2.491	2.451	2.776	0.016	–0.115	3.522	3.522	3.403	1.804	0.034	0.488
28	2.614	2.614	1.553	2.935	0.406	–0.123	4.068	4.068	3.949	1.405	0.029	0.655
29	2.599	2.599	2.540	2.837	0.023	–0.092	2.596	2.596	2.135	1.752	0.178	0.325
30	2.512	2.512	2.469	2.662	0.017	–0.060	2.424	2.424	2.305	1.435	0.049	0.408
31	2.512	2.512	2.469	2.520	0.017	–0.003	2.512	2.512	2.393	1.165	0.048	0.536
32	2.614	2.614	1.954	2.451	0.252	0.062	2.057	2.057	1.395	1.177	0.322	0.428
33	2.432	2.432	1.404	2.561	0.423	–0.053	2.156	2.156	2.037	1.147	0.055	0.468
34	2.556	2.556	2.505	2.578	0.020	–0.009	3.037	3.037	2.918	1.128	0.039	0.629
35	2.432	2.432	2.403	2.462	0.012	–0.012	1.952	1.952	1.833	1.248	0.061	0.361
36	2.516	2.516	2.472	2.340	0.017	0.070	1.186	1.186	1.067	1.195	0.101	–0.008

Table 8
Comparison of forecast error for the ARIMA time series and RBF neural network for Karun and Mond rivers

Sample period	Karon river						Mond river					
	Actual values		Predict values		Forecast error		Actual values		Predict values		Forecast error	
	RBF	ARIMA	RBF	ARIMA	RBF	ARIMA	RBF	ARIMA	RBF	ARIMA	RBF	ARIMA
1	7.050	–	6.584	–	0.066	–	4.593	–	3.736	–	0.187	–
2	8.782	–	8.624	–	0.018	–	36.409	–	38.733	–	–0.064	–
3	7.358	–	7.286	–	0.010	–	34.935	–	37.112	–	–0.062	–
4	7.777	–	7.452	–	0.042	–	43.467	–	46.497	–	–0.070	–
5	5.033	–	5.100	–	–0.013	–	41.352	–	44.170	–	–0.068	–
6	8.946	–	8.778	–	0.019	–	58.291	–	62.803	–	–0.077	–
7	5.603	–	5.636	–	–0.006	–	42.618	–	45.563	–	–0.069	–
8	5.371	–	5.418	–	–0.009	–	27.770	–	29.230	–	–0.053	–
9	5.754	–	5.778	–	–0.004	–	4.116	–	3.211	–	0.220	–
10	7.054	–	7.000	–	0.008	–	3.307	–	2.321	–	0.298	–
11	6.308	–	6.299	–	0.001	–	10.140	–	9.837	–	0.030	–
12	4.790	–	4.872	–	–0.017	–	37.689	–	40.141	–	–0.065	–
13	6.775	–	6.738	–	0.005	–	30.196	–	31.899	–	–0.056	–
14	3.341	–	3.510	–	–0.051	–	43.804	–	46.868	–	–0.070	–
15	5.531	–	5.568	–	–0.007	–	41.557	–	44.396	–	–0.068	–
16	4.658	–	4.748	–	–0.019	–	51.502	–	55.335	–	–0.074	–
17	6.113	–	5.116	–	0.163	–	55.918	–	60.193	–	–0.076	–
18	5.197	–	5.254	–	–0.011	–	39.630	–	42.276	–	–0.067	–
19	6.829	–	6.523	–	0.045	–	62.702	–	67.655	–	–0.079	–
20	6.723	–	6.123	–	0.089	–	42.741	–	45.698	–	–0.069	–
21	6.726	–	6.692	–	0.005	–	38.100	–	40.593	–	–0.065	–
22	5.844	–	5.863	–	–0.003	–	17.040	–	17.427	–	–0.023	–
23	9.231	–	8.456	–	0.084	–	5.360	–	4.579	–	0.146	–
24	3.940	–	4.725	–	–0.199	–	18.630	–	20.176	–	–0.083	–
25	4.994	4.994	5.064	6.150	–0.014	–0.232	21.475	21.475	22.306	19.43	–0.039	0.095
26	3.765	3.765	3.908	6.150	–0.038	–0.634	13.964	13.964	15.042	36.12	–0.077	–1.587
27	3.960	3.960	4.092	6.150	–0.033	–0.553	24.695	24.695	25.848	40.55	–0.047	–0.642
28	3.960	3.960	4.258	6.150	–0.075	–0.553	5.965	5.965	5.245	43.70	0.121	–6.327
29	5.630	5.630	4.635	6.150	0.177	–0.092	39.159	39.159	41.758	45.73	–0.066	–0.168
30	4.055	4.055	4.181	6.150	–0.031	–0.517	39.630	39.630	42.276	49.07	–0.067	–0.238
31	5.096	5.096	5.159	6.150	–0.012	–0.207	27.706	27.706	29.160	44.54	–0.052	–0.608
32	4.538	4.538	4.035	6.150	0.111	–0.355	26.120	26.120	27.415	36.97	–0.050	–0.416
33	3.064	3.064	3.954	6.150	–0.291	–1.007	10.973	10.973	10.754	22.53	0.020	–1.054
34	5.184	5.184	5.521	6.150	–0.065	–0.186	15.030	15.030	15.216	22.23	–0.012	–0.479
35	5.259	5.259	5.313	6.150	–0.010	–0.169	5.611	5.611	3.853	21.46	0.313	–2.826
36	7.785	7.785	7.012	6.150	0.099	0.210	15.146	15.146	15.344	32.82	–0.013	–1.167

series to predict different water quality in surface water. None of the mentioned studies forecasting the SAR of rivers using the ARIMA time series.

5. Conclusions

The results of developing the Box–Jenkins time series (ARIMA) and the RBF neural network for Aras, Sepidrud, Karun, and Mond River summarizes as follows:

- Between different ARIMA models for SAR prediction of Aras, Sepidrud, Karun, and Mond Rivers, ARIMA (0,1,1) × (0,1,1)12, ARIMA (0,1,1) (0,1,1)12, ARIMA (0,1,1), and ARIMA (0,1,2) × (1,1,1)12, with minimum AIC of –1.7127, –1.8177, 2.9317, and 12.44 were selected, respectively. The residual of all the selected ARIMA time series was independent (p-value > 0.05), which indicates the accuracy of the selected models.
- Using RBF neural network for SAR prediction of Aras, Sepidrud, Karun, and Mond Rivers with normalized data reached suitable training, which means squared error (MSE are between 0.00026 and 0.0255) and MBE (MBE are between 0.01566 and 0.0612) for training is very low.

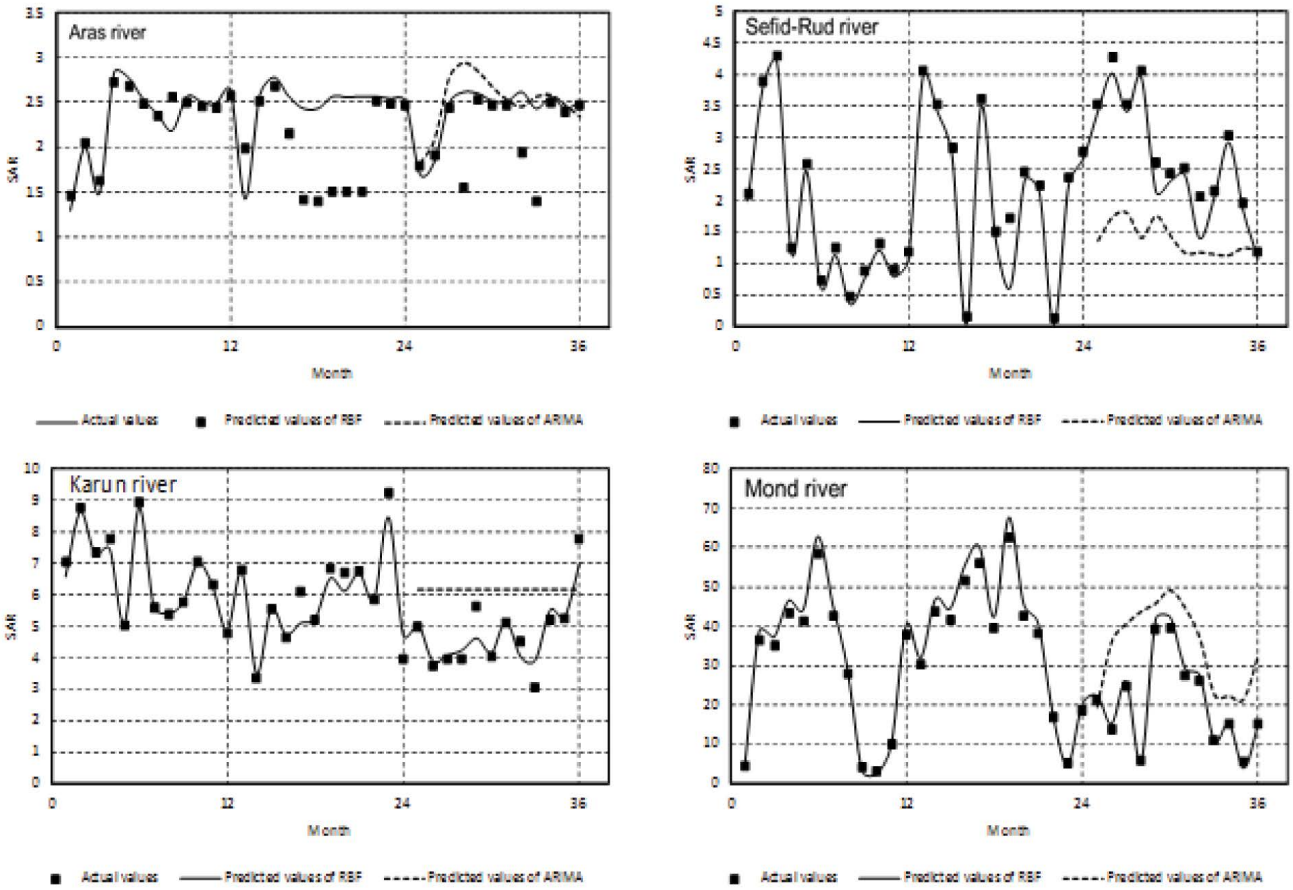

Fig. 10. Comparison of SAR forecasting using ARIMA time series and RBF neural network for Aras, Sefid-Rud, Karun, and Mond rivers.

- The performance of the RBF neural network for the four rivers also are reliable because the coefficient of determination (R^2 are between 0.907 and 0.960), index of agreement (IA are between 0.981 to 0.999), and the Nash-Sutcliffe Efficiency (E are between 0.964 and 0.999) indicate the accuracy of the predictions.
- Thirty-six months RBF neural network and 12 months forecasting of ARIMA time series of SAR predictions for Aras river relatively match to the measured data and forecast error of both were similar.
- Although, we trained the RBF neural networks with fewer input data than ARIMA and the SAR forecast of RBF was for 3 y, the RBF has a more accurate prediction than the ARIMA time series model with more input data and 1 y forecast. In general, we could report that the RBF neural network for SAR prediction in the rivers is more reliable than the ARIMA time series.

References

[1] A.M. Aboukarima, M.A. Al-Sulaiman, M.S.A. EI Marazky, Effect of sodium adsorption ratio and electric conductivity of the applied water on infiltration in a sandy-loam soil, Water SA, 44 (2018) 105–110.
[2] I. Shainberg, J. Letey, Response of soils to sodic and saline conditions, Hilgardia, 52 (1984) 1–57.
[3] E.A. El-Morsy, M. Malik, J. Letey, Interactions between water quality and polymer treatment on infiltration rate and clay migration, Soil Technol., 4 (1991) 221–231.
[4] M.R. Emdad, R. Steven, R.J. Smith, H. Fardad, Effect of water quality on soil structure and infiltration under furrow irrigation, Irrig. Sci., 23 (2004) 55–60.
[5] B.B. Patel, R.S. Dave, Studies on the infiltration of saline-alkali soils of several parts of Mehsana and Patan districts of North Gujarat, J. Appl. Technol. Environ. Sanitation, 1 (2011) 87–92.
[6] D.Ö. Faruk, A hybrid neural network and ARIMA model for water quality time series prediction, Eng. Appl. Artif. Intell., 23 (2010) 586–594.
[7] A.W. Jayawardena, F. Lai, Time series analysis of water quality data in Pearl River, China, J. Environ. Eng., 115 (1989) 590–607.
[8] H. Sun, M. Koch, Case study: analysis and forecasting of salinity in Apalachicola Bay, Florida, using Box-Jenkins ARIMA models, J. Hydraul. Eng., 127 (2001) 718–727.
[9] G. Asadollahfardi, Analysis of surface water quality in Tehran, Water Qual. Res. J., 37 (2002) 489–511.
[10] A. Kurnc K. Yürekli, O. Cevik, Performance of two stochastic approaches for forecasting water quality and streamflow data from Yeşilırmak River, Turkey, Environ. Modell. Software, 20 (2005) 1195–1200.
[11] G. Asadollahfardi, M. Rahbar, M. Fatemiaghda, Application of time series models to predict water quality of upstream and downstream of the Latian Dam in Iran, Univ. J. Environ. Res. Technol., 2 (2012) 26–36.
[12] S.J. Abudu, P. King, Z. Sheng, Comparison of the performance of statistical models in forecasting monthly total dissolved

solids in the Rio Grande, J. Am. Water Resour. Assoc., 48 (2012) 10–23.
[13] M. Ranjbar, M. Khaledian, Using ARIMA time series model in forecasting the trend of changes in qualitative parameters of Sefid-Rud River, Int. Res. J. Appl. Basic Sci., 8 (2014) 346–351.
[14] F.K. Arya, L. Zhang, Time series analysis of water quality parameters at Stillaguamish River using order series method, Stochastic Environ. Res. Risk Assess., 29 (2015) 227–239.
[15] M.H. Salmani, E. Salmani Jajaei, Forecasting models for flow and total dissolved solids in Karoun river-Iran, J. Hydrol., 535 (2016) 148–159.
[16] G. Asadollahfardi, A. Hemati, S. Moradinejad, R. Asadollahfardi, Sodium adsorption ratio (SAR) prediction of the Chalghazi river using artificial neural network (ANN) Iran, Curr. World Environ., 8 (2013) 169–178.
[17] A. Azad, H. Karami, S. Farzin, A. Saeedian, H. Kashi, F. Sayyahi, Prediction of water quality parameters using ANFIS optimized by intelligence algorithms (case study: Gorganrood River), KSCE J. Civ. Eng., 22 (2018) 2206–2213.
[18] M.T. Sattari, A. Farkhondeh, J. Patrick Abraham, Estimation of sodium adsorption ratio indicator using data mining methods: a case study in Urmia Lake basin, Iran, Environ. Sci. Pollut. Res., 25 (2018) 4776–4786.
[19] B. Singh, Prediction of the sodium absorption ratio using data-driven models: a case study in Iran, Geol. Ecol. Landscapes, 4 (2020) 1–10.
[20] B.H.K. Al-Obaidi, B.H. Khudhair, R.S. Mahmood, R.A. Kadhim, Water quality assessment and sodium adsorption ratio prediction of Tigris River using artificial neural network, J. Eng. Sci. Technol., 15 (2000) 3055–3066.
[21] A. Aslanargun, M. Mammadov, B. Yazici, S. Yolacan, Comparison of ARIMA, neural networks and hybrid models in time series: tourist arrival forecasting, J. Stat. Comput. Simul., 77 (2007) 29–53.
[22] L.A. Dı´az-Robles, J.C. Ortega, J.S. Fu, G.D. Reed, J.C. Chow, J.G. Watson, J.A. Moncada-Herrera, A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: the case of Temuco, Chile, Atmos. Environ., 42 (2008) 8331–8340.
[23] A.A. Yassen, Comparative Study of Artificial Neural Network and ARIMA Models for Economic Forecasting, Mater Thesis, Al-Azhar University, Gaza, 2011.
[24] A.A. Adebiyi, A.O. Adewumi, C.K. Ayo, Comparison of ARIMA and artificial neural networks models for stock price prediction, J. Appl. Math., 1 (2014) 1–7.
[25] D.E. Ighravwea, C.O. Anyaeche, A comparison of ARIMA and ANN techniques in predicting port productivity and berth effectiveness, Int. J. Data Network Sci., 3 (2019) 13–22.
[26] Z. Li, Y. Li, A comparative study on the prediction of the BP artificial neural network model and the ARIMA model in the incidence of AIDS, BMC Med. Inf. Decis. Making, 143 (2020) 1–13.
[27] G. Asadollahfardi, H. Zangooi, M. Asadi, M. Tayebi Jebeli, A. Meshkat-Dini, N. Roohani. Comparison of Box–Jenkins time series and ANN in predicting total dissolved solid at the Zāyandé-Rūd River, Iran, J. Water Supply Res. Technol. AQUA, 67(2018) 673–684.
[28] G. Asadollahfardi, N. Heidarzadeh, A. Sekhavati, M. Asadi, Optimization of water quality monitoring stations using dynamic programming approach, a case study of the Mond Basin Rivers, Iran, Environ. Dev. Sustainability, 23 (2021) 2867–2881, doi: 10.1007/s10668–020–00693–2.
[29] E. Rahnama, O. Bazrafshan, G. Asadollahfardi, Application of data-driven methods to predict the sodium adsorption rate (SAR) in different climate in Iran, Arabian J. Geosci., 13 (2020), doi: 10.1007/s12517–020–06146–4.
[30] G.E.P. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, 5th ed., Holden Day, San, Francisco, 1976.
[31] W. Wu, G.C. Dandy, H.R. Maier, Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling, Environ. Modell. Software, 54 (2014) 108–127.
[32] H.R. Maier, A. Jain, G.C. Dandy, K.P. Sudheer, Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions, Environ. Modell. Software, 25 (2010) 891–909.
[33] M. Cakmakci, Adaptive neuro-fuzzy modelling of anaerobic digestion of primary sedimentation sludge, Bioprocess Biosyst. Eng., 30 (2007) 349–357.
[34] H. Abu Qdais, K. Bani Hani, N. Shatnawi, Modeling and optimization of biogas production from a waste digester using artificial neural network and genetic algorithm, Resour. Conserv. Recyl., 54 (2010) 359–363.
[35] G. Asadollahfardi, A. Taklify, A. Ghanbari, Application of artificial neural network to predict TDS in Talkheh Rud River, J. Irrig. Drain. Eng., 138 (2012) 363–370.
[36] T. Beltramo, C. Ranzan, J. Hinrichs, B. Hitzmann, Artificial neural network prediction of the biogas flow rate optimised with an ant colony algorithm, Biosyst. Eng., 143 (2016) 68–78.
[37] B. Najafi, S. Faizollahzadeh Ardabili, Application of ANFIS, ANN, and logistic methods in estimating biogas production from spent mushroom compost (SMC), Resour. Conserv. Recycl., 133 (2018) 169–178.
[38] C.W. Dawson, R.L. Wilby, Hydrological modelling using artificial neural networks, Prog. Phys. Geogr., 25 (2001) 80–108.
[39] T. Kohonen, Self-Organization and Associative Memory, Springer, New York, NY, Berlin, Heidelberg, 1984.
[40] X.M. Song, Radial Basis Function Networks for Empirical Modeling of Chemical Process, MSc Thesis, University of Helsinki, 1996.
[41] P.L. Narasimha, W.H. Delashmit, M.T. Manry, J. Li, F. Maldonado, An integrated growing-pruning method for feedforward network training, Neurocomputing, 71 (2008) 2831–2847.
[42] S. Chen, S.A. Billings, W. Luo, Orthogonal least squares methods and their application to nonlinear system identification, Int. J. Control, 50 (1989) 1873–1896.
[43] S. Chen, C.F.N. Cowan, P.M. Grant, Orthogonal least squares learning algorithm for radial basis function networks, IEEE Trans. Neural Networks, 2 (1991) 302–309.
[44] C.J. Willmott, K. Matsuura, Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Clim. Res., 30 (2005) 79–82.
[45] P. Krause, D.P. Boyle, F. Bäse, Comparison of different efficiency criteria for hydrological model assessment, Adv. Geosci., 5 (2005) 89–97.
[46] J.E. Nash, J.V. Sutcliffe, River flow forecasting through conceptual models' part I—a discussion of principles, J. Hydrol., 10 (1970) 282–290.