Hybrid MARMA-NARX model for flow forecasting based on the large-scale climate signals, sea-surface temperatures, and rainfall

Mohammad Ebrahim Banihabib, Arezoo Ahmadian
and Mohammad Valipour

ABSTRACT

In this study, to reflect the effect of large-scale climate signals on runoff, these indices are accompanied with rainfall (the most effective local factor in runoff) as the inputs of the hybrid model. Where one-year in advance forecasting of reservoir inflows can provide data to have an optimal reservoir operation, reports show we still need more accurate models which include all effective parameters to have more forecasting accuracy than traditional linear models (ARMA and ARIMA). Thus, hybridization of models was employed for improving the accuracy of flow forecasting. Moreover, various forecasters including large-scale climate signals were tested to promote forecasting. This paper focuses on testing MARMA-NARX hybrid model to enhance the accuracy of monthly inflow forecasts. Since the inflow in different periods of the year has in linear and non-linear trends, the hybrid model is proposed as a means of combining linear model, monthly autoregressive moving average (MARMA), and non-linear model, nonlinear autoregressive model with exogenous (NARX) inputs to upgrade the accuracy of flow forecasting. The results of the study showed enhanced forecasting accuracy through using the hybrid model.

Key words | forecasting, hybrid model, Iran, large-scale climate signals, monthly flow, sea-surface temperature

Mohammad Ebrahim Banihabib (corresponding author)
Department of Irrigation and Drainage Engineering,
University of Tehran,
Aburaihan Campus, Pakdasht, Tehran,
Iran
E-mail: banihabib@ut.ac.ir

Arezoo Ahmadian
Water Resources Engineering,
University of Tehran,
Aburaihan Campus, Tehran,
Iran

Mohammad Valipour
Department of Crop, Soil and Environmental
Sciences,
Auburn University,
Auburn, Alabama,
USA

INTRODUCTION

Improving accuracy of hydrological forecasting, especially long-term forecasting, is an essential yet often difficult task facing modern water resources management. Much recent literature has focused on real-life case studies of contemporary soft computing techniques in hydrological engineering (Chau & Wu 2010; Wu et al. 2010; Chen et al. 2015; Gholami et al. 2015; Taormina et al. 2015; Wang et al. 2015). Long-term forecasting is of high importance in the effective management of water resources. Since using the existing linear models cannot lead to accurate forecasts (Khashei & Bijari 2011), it is proposed to employ hybrid models for flow forecasting to consider both linear and non-linear components in modeling. Autoregressive moving average

(ARMA) and artificial neural network (ANN) have been used in flow forecasting models. Mohammadi et al. (2006) and Yurekli et al. (2005) applied these models for monthly inflow forecasting. Moreover, many researchers have accomplished forecasting using ANN models (Kilinç & Ciğizoğlu 2003; Banihabib et al. 2008; Valipour et al. 2013), and Bazartseren et al. (2003) compared neural network and neuro-fuzzy models with ARMA and autoregressive with exogenous terms (ARX) to evaluate the models' accuracy in forecasting water levels of two rivers. The most critical point of these studies is the clarification of the advantage and disadvantage of using linear time series and ANN models. For instance, Bazartseren et al. (2003) reported

that ARMA and ARX models have performed better than neural network and neuro-fuzzy models in short-term forecasts (up to 5 hours) of water level; however, by increasing the forecasting horizon, neural networks and neuro-fuzzy models performed better. Therefore, it is necessary to develop a method which benefits from the advantages of both linear time series and ANN models.

In various research fields excluding water science, many researchers have employed hybrid models to use the benefits of the models simultaneously. For instance, the hybrid model of autoregressive integrated moving average (ARIMA) and ANN was developed by Koutroumanidis et al. (2009) in forecasting annual petrol price as well as in hourly and daily forecasting of air quality and meteorological data by Aladag et al. (2009) and Diaz-Robles et al. (2008). Khashei & Bijari (2011) applied hybrid ARIMA and ANNs to forecast annual sunspot series, and Chau & Wu (2010) developed a hybrid model integrating ANNs and support vector regression for daily rainfall prediction. However, rarely, have some researchers like Banihabib et al. (2014) used hybrid models for inflow forecasting. Therefore, there is the possibility of using a hybrid model to benefit from the strengths of ARMA model in forecasting the linear component and ANN model in forecasting the non-linear component to increase forecasting accuracy.

Since studies on large-scale climate signals in Iran are a relatively rare topic (Rivandi et al. 2013), and some reports in recent years have expressed the effect of these signals in long-term forecasting (Karamouz & Araghinejad 2005), employing these signals along with rainfall data (as the most important recognized local factor on flow) can enhance flow forecasting. In recent decades, the effect of large-scale climate signals on flow and hydrologic variables have been studied worldwide (Bradbury et al. 2002; Fatehi Marj & Borhani Darian 2005; Ashoori et al. 2008; Ryu et al. 2010). Some researchers like Chandimala & Zubair (2007), Azimi et al. (2010), and Rahimi Khoob (2011) have used correlation analysis, data mining method, and regression equations, and others like Fallah Ghalhari & Khoshhal (2010), Borhani & Fatehi (2008), and Wang et al. (2006) have forecasted rainfall and river flow. However, few studies have assessed the effect of large-scale climate signals and rainfall on stream flow. According to studies in Iran (Rivandi et al. 2013), it is clear that signals such as the

El Niño–Southern Oscillation (ENSO), the North Atlantic Oscillation (NAO), and the Persian Gulf Sea-surface temperature (PGSST) and Red Sea-surface temperature (RSST) impacted on the hydrologic variables in some parts of Iran. However, these studies are generally accomplished for a specific season during a year or merely for a three- or six-month period, while no forecasting is accomplished for the horizon of the whole year and on a monthly basis.

Hybridization of ARMA and ANN models can be applied for increasing forecasting accuracy of the variables with linear and non-linear components. Thus, ARMA as a time series model for estimating the linear component of discharge and NARX ANN model for estimating the non-linear component of discharge were used for flow forecasting in this study. The previous investigations indicate advantages of the NARX compared with the non-linear autoregressive neural network (NARNN) and non-linear input-output (NIO) models (Valipour 2016). Moreover, choosing the proper forecasters was another purpose of this study. Note that using hybrid ARMA-ANN models, in recent investigations, represents a reliability aspect of this technique in various fields (Benmouiza & Cheknane 2016; Gairaa et al. 2016).

Banihabib et al. (2008) aimed to find the relation between some stations in order to forecast the flow, but it was used only from the ANN. The research goal of Valipour et al. (2013) is forecasting the inflow of Dez dam reservoir by using ARMA and ARIMA models while increasing the number of parameters to increase the forecast accuracy to four parameters and comparing them with the static and dynamic ANNs.

By investigating the results, it showed that the ARIMA model could be used for forecasting an appropriate monthly inflow for the next 12 months, and dynamic autoregressive ANN model with sigmoid activity function could be used for forecasting discharge of the next five years.

All of the studies which were done showed that ANN and time series models have excellent ability to forecast the flow, but each of the studies used the models singly and did not use the abilities of these models for increasing the ability of forecasting. The data which were used in these models were streamflow data while in this study we attempted to use rainfall and precipitation data and also large-scale signals to reduce the uncertainty of forecasting.

Accordingly, in this study, the efficiency of using some large-scale climate signals such as ENSO, NAO, SST-PG, and SST-RS as global forecaster along with rainfall as local forecaster were examined to forecast the inflow to Dez reservoir accurately.

Selection of these signals is according to the results of research which have shown the impact of these signals and water surface temperature on the precipitation and streamflow in Iran, spatially in the studied area and its neighbor areas. These studies are such as those of Azimi *et al.* (2010) and Gheiby & Noorafshan (2013) which are mentioned in the sources or are the articles which have presented and published in conferences and journals of Iran.

There were other indices for investigating for this purpose, such as water temperature of the Caspian Sea, but we did not find any research about the impact of it on the study area. Moreover, the Pacific Decadal Oscillation (PDO) is another large-scale index, but we did not use it because past investigations did not show the impact of it in the study area. Indeed, we are trying to use indices that previous reports have declared their impact on in the study area on some hydrological factors, i.e., rainfall.

Over the past year, in advance forecasting of reservoir inflows has offered essential data to operate reservoirs optimally. Reviewing the studies indicates we still require more accurate models which can include active parameters such as large-scale climate signals to achieve more forecasting accuracy than traditional linear models (ARMA and ARIMA). Consequently, the effect of using large-scale climate signals was studied for improving the monthly inflow forecasts with ARMA-NARX hybrid model for the following one-year horizon.

MATERIALS AND METHODS

Case study

Dez basin is a part of the middle Zagros Mountains. Regarding the general hydrological classification of Iran, it is considered as a part of the Persian Gulf Basin. The basin is located between 32° 35′ to 34° 07′ northern latitude and 48° 20′ to 50° 20′ eastern longitude in the southwest of Iran.

This area was selected because Dez dam, which is one of the important dams of Iran, has an essential role in the control of the floods of Dez River, irrigating the extended area of the Khuzestan plain, and producing electricity, and so, forecasting the flow of this river can be useful in making management decisions. Also, much research on flow forecasting in this area has been carried out, and because of the importance of this region, we were trying to increase its accuracy. For example, Azimi *et al.* (2010) predicted seasonal flow, and we reduced its timescale to the monthly flow. Other studies such as those of Banihabib *et al.* (2008) and Valipour *et al.* (2013) presented the advantage of time series models such as ARMA or ANN comparing with regression models in order to forecast the flow. Valipour *et al.* (2013) compared ARMA, ARIMA, and ANN models in forecasting monthly streamflow but they did not use the advantages of these models in promoting the forecasting. Also, another reason that encouraged us to select this region was the data of this dam which were accessible for more than 30 years.

Previous investigations show NAO and Southern Oscillation Index (SOI) are useful methods for rainfall and flow forecasting at different sites in Iran (Fatehi Marj & Borhani Darian 2005; Ashoori *et al.* 2008; Borhani & Fatehi 2008) and Dez basin (Azimi *et al.* 2010). As well, the effect of sea surface temperature of the Persian Gulf and the Red Sea on rainfall near the basin was observed (Rahimi Khoob 2011). However, the most effective factor on the inflow had not been evaluated. Therefore, this paper applies these data as a forecaster and assesses the effect of them singularly or in sets of two, three, and four in improving the accuracy of flow forecasting.

In this study, to forecast the reservoir inflow, the flow data of hydrometric stations in the upstream of Dez reservoir (Telezang station) and rainfall data of three stations (Keshvar Sorkhab, Sepiddasht-Sezar, and Gelerood) were used. The data were obtained from the Water Resources Management Organization (WRMO) of Iran. Unit of flow data is cubic meter per second and unit of precipitation is millimeter.

Figure 1 shows Dez dam basin, the location of the dam, Telezang hydrometric station, and the rainfall gauges. As well as rainfall data, SOI, NAO, PGSST, and RSST were used as forecasters of reservoir inflow. The information of large-scale climate indices was gained from the National Oceanic and Atmospheric Administration (NOAA; http://www.noaa.gov/). These data are dimensionless.
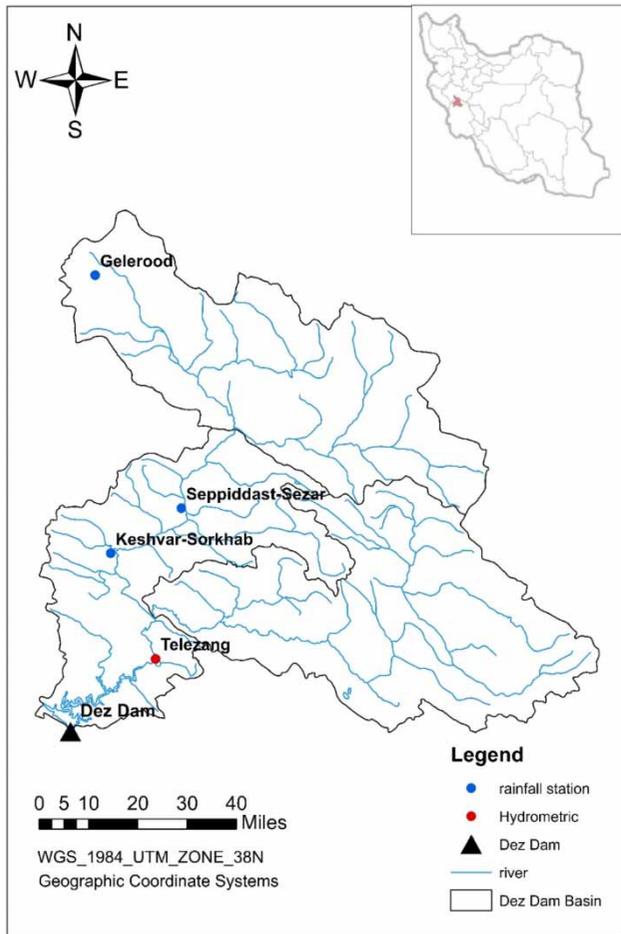
Figure 1 | Dez basin and the location of the dam.

Research method

In this study, first, the flow discharge is forecasted through ARMA time series model. Then, the residuals of the model and other input data (rainfall, SOI, NAO, PGSST, and RSST indices) were entered into the NARX neural network model to forecast in the final stage.

The flowchart shows the research procedures (Figure 2).

The data should be normalized before being employed as input for ARMA. In this article, the data were normalized by using the Neper number. Autocorrelation function (r_k) and partial autocorrelation function ($\varnothing_k$) were used for recognizing stationary and nonstationary data with the k delay and average $\bar{x}$.

$$r_k = \frac{\sum_{i=1}^{n-k}(x_i - x)(x_{i+k} - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \; ; \; -1 \le r_k \le 1 \tag{1}$$
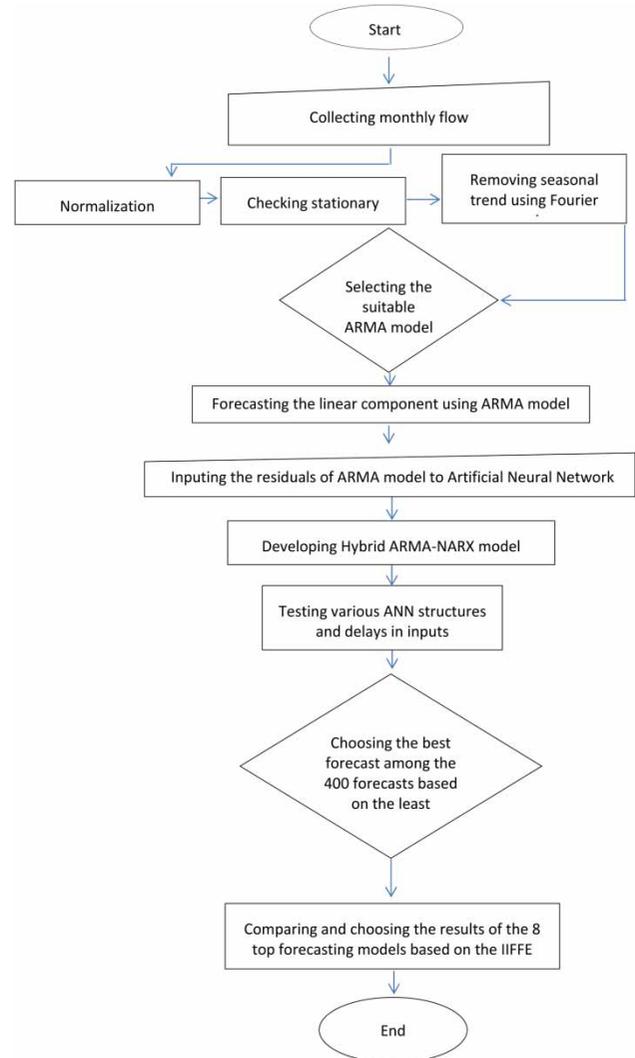

Figure 2 | The research procedure.

where x_{i+k} is the values of the parameter in temporal stage i and time delay k (Cryer & Chan 2008).

If r_k is attenuating with the proper trend and remains within limits in the ACF graph, the time series will be considered stationary; else, it will be considered nonstationary:

$$\varnothing_k = \frac{r_k - \sum_{i=1}^{k-1}\varnothing_i(k-1)r_{k-1}}{1 - \sum_{i=1}^{k-1}\varnothing_i(k-1)_i r_i} \tag{2}$$

The ARIMA modeling approach involves the following three steps: model identification, parameter estimation,

diagnostic checking. Identification of the general form of a model includes two stages:

1. If it is necessary, appropriated differencing of the series is performed to achieve stationary and normalit.
2. The temporal correlation structure of the transformed data is identified by examining its autocorrelation (ACF) and partial autocorrelation (PACF) functions (Mishra & Desai 2005). For determining the number of ARMA and ARIMA models parameters perform by partial auto-correlation function (PACF) and auto-correlation function (ACF) curves (Balaguer et al. 2008; Cryer & Chan 2008; Valipour et al. 2013).

ACF is a useful statistical tool that measures the relation between the earlier values in the series to following values. PACF is the amount of correlation between a variable and a lag of itself that is not explained by correlations at all low order lags.

In the next step, to remove the seasonal trend, the mean and standard deviation of the data were calculated using Fourier series. Fourier series is defined as Equation (3). Assume that u_τ is the estimation of a sample from alternating factors of population v_τ. In the formula, $\bar{u}$ is the mean of u_τ (Equation (6)). A_j and B_j are Fourier series ratios (Equations (4) and (5)). j is harmony and h is the sum of harmonies that equals $w/2$ alternatively, $(w-1)/2$, based on w being an odd or an even value. The variable w is correlated to the type of temporal scale. In monthly scale, $w = 12$ and $h = 6$ (Salas 1980).

$$\widehat{u_T} = \bar{u} + \sum_{j=1}^{h}\left[A_j\cos\left(\frac{2\pi j\tau}{w}\right) + B_j\sin\left(\frac{2\pi j\tau}{w}\right)\right] \quad \tau = 1.2.3\ldots w \tag{3}$$

$$B_j = \frac{2}{w}\sum_{\tau=1}^{w} u_\tau \sin\left(\frac{2\pi j\tau}{w}\right); \quad j = 1.2.3\ldots h \tag{4}$$

$$A_j = \frac{2}{w}\sum_{\tau=1}^{w} u_\tau \cos\left(\frac{2\pi j\tau}{w}\right); \quad j = 1.2.3\ldots h \tag{5}$$

$$\bar{u} = \frac{1}{w}\sum_{\tau=1}^{w} u_\tau \tag{6}$$

The Fourier series formula was used for the average and also the standard deviation calculations. Forecasting was done using time series models, after preparing the data. Autoregressive and moving average models are among the most important models derived from the total linear process, and by combining these models, ARMA model is produced. ARMA is formed by combining regressive and moving average process and includes p sentences of autoregressive average and q sentences of moving average, as shown in Equation (7):

$$(1 - \alpha_1 B - \ldots - \alpha_p B^p)(x_t - \mu) = (1 + \beta_1 B + \ldots + \beta_1 B^q)z_t \tag{7}$$

where x_t is the value of the variable in time t, z_t expresses the pure random process with an average of 0, and σ_z^2 is variance, and β_{t-q} is the moving average parameter in delay step q and σ_p is the autoregressive coefficient in delay step p. The maximum likelihood, moment or other methods are used to calculate the model parameters. In this paper, the ARMA model parameters were calculated using maximum likelihood method and log-likelihood, shown in Equation (9) (Wang 2006). When ε_t is the forecasting error in every temporal step, the variance of forecasting error using σ^2, the variance of model errors (residuals), and f_t, the scale parameter can be calculated by Equation (8) (Wang 2006):

$$var(\varepsilon_t) = \sigma^2 f_t; \quad t = 1.2.3\ldots N \tag{8}$$

$$\log L\left(x_{p+1}.\ldots.x_N\right) = -\frac{N}{2}\log(2\pi\sigma^2) - \frac{1}{2}\sum_{t=1}^{N}\log f_t$$
$$- \frac{1}{2\sigma^2}\sum_{t=1}^{N}\frac{\varepsilon_t^2}{f_t} \tag{9}$$

Maximum likelihood or least square error were used for estimating the parameters of time series model and using each of them is permitted (Van Calster et al. 2017). The use of ML is common in studies such as Amiri (2015).

Among the fitted models, the model with the least value of the Akaike information criterion (AIC) was selected. When this criterion was similar for several models, the most suitable model, based on parsimony principle, was the one with the least number of parameters. Equation (10) shows the parsimony principle (Wang 2006):

$$AIC = -2 \times \log ML + 2 \times k \tag{10}$$

where ML is the maximum likelihood, and k is the number of used parameters in the model.

To investigate the independency and stationarity of residuals, the goodness test was calculated with the statistics of Ljung–Box, Equation (11), was used to ensure selecting the model:

$$Q_k = n(n+2) \sum_{i=1}^{k} \frac{r_i}{n-i} \tag{11}$$

Using the Ljung-Box test in time series models is common and has been reported by previous researchers (Kamil & Razali 2015; Fathian et al. 2017). Where k is the number of delays, r_i equals the ith ACF of the residuals and n is the number of samples. The null hypothesis (H_0) is expressed as the ACF of all residuals equals zero and all data are random. Under these circumstances, if the p-value is less than or equal to the level of the meaningfulness of the test (5%), the H_0 is rejected (Cryer & Chan 2008).

While goodness-of-fit criteria such as the sum-squared error (SSE) or coefficient of determination (R^2) may be appropriate for these purposes, it is often better to use more parsimonious model selection criteria such as Akaike's information criterion (AIC) or the Bayesian information criterion (BIC), which penalize unnecessary model complexity (Humphrey et al. 2016).

For in-sample model comparison usually, three criteria, namely, loglikelihood, AIC, and BIC are calculated. AIC and BIC criteria are objective measures of model suitability that balance model fits and model complexity. In another word, AIC and BIC are fitness indexes for trading off the complexity of a model against how well the model fits the data. Since AIC and BIC attempt to find the model that best explains the data with a minimum number of parameters, they are considered an approach favoring simplicity (Amiri 2015). In finite sample sizes, the BIC can perform worse than the AIC, even when the true model is under consideration. The choice between the AIC and BIC depends on one's notion of the true model. If the true model is assumed to be complicated, with large, moderate, and small effects, and the candidate models oversimplifications, then the AIC may be preferred to the BIC (Vrieze 2012). Due to the reasons mentioned above, between AIC and BIC, we have used AIC.

After ARMA forecasting, the results enter NARX neural network. The NARX dynamic ANNs shape an important group of non-linear back propagation dynamic neural networks in which the equation of NARX model is as Equation (12) (Mandal & Prabaharan 2006):

$$y(t) = f\{y(t-1). \, y(t-2). \, \ldots . \, y(t-d_y). \, u(t)\} \tag{12}$$

where $d_y \geq 1$ equals the number of delays in the output, $u(t)$ is the value of the input to the network in time t, $y(t)$ is the forecasted output in time t, which is dependent on previous values of output or input. Figure 3 shows the structure of NARX dynamic ANN with d_y output delays and d_u number of input delays. NARX ANN has two different series and series-parallel structures. In the parallel structure, the estimated output is back propagated with a specified number of delays to the input of the feed forward neural network, shown in Equation (13) (Mandal & Prabaharan 2006):

$$\begin{aligned} \hat{y}(t+1) &= f[y_p(t) \, ; \, u(t) \, ; \, w] \\ &= f[\hat{y}(t). \, \hat{y}(t-1). \, \ldots . \, \hat{y}(t-n_y+1); \\ & \quad u(t). \, u(t-1). \, \ldots . u(t-n_u+1); \, w] \end{aligned} \tag{13}$$

Moreover, in the series-parallel structure, the NARX neural network output is calculated based on observed output with specified number of delays as well as the input
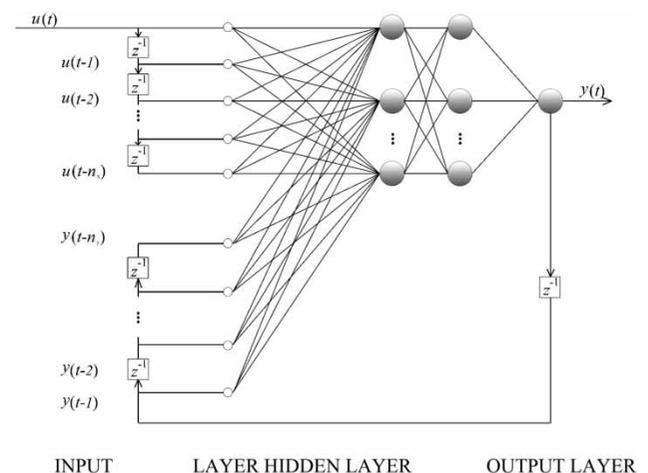

Figure 3 | NARX dynamic ANN with d_y output delays and d_u input delays.

delays (Equation (14)) (Mandal & Prabaharan 2006):

$$\hat{y}(t+1) = f[y_{sp}(t) \ ; \ u(t) \ ; \ w]$$
$$= f[y(t). \ y(t-1). \ \ldots. \ y(t - n_y + 1);$$
$$u(t). \ u(t-1). \ \ldots. u(t - n_u + 1); w] \qquad (14)$$

In this paper, the training stage had series-parallel structure and forecasting stage had parallel structure. Moreover, the transfer function was tangent-sigmoid, and learning function was Levenberg–Marquardt.

NARX ANN's training function is Levenberg–Marquardt by default, because the Levenberg–Marquardt algorithm has high efficiency and high speed in most problems and also its accuracy is noticeable. Since it has the lowest error in most of the cases, the Levenberg–Marquardt algorithm was selected in this study because of its faster convergence in network training with medium size and also it has less epoch than other algorithms. Use of this algorithm has been reported by several studies (Abdul Jaleel & Aparna 2015; Raj & Sivanandan 2017).

Both ARMA and ANN models in their linear and non-linear ranges are successful models; however, they cannot be singularly used for all conditions. Therefore, using hybrid models that have the capability of both linear and non-linear components can be a suitable solution to improve the forecasting (Aladag et al. 2009). In hybrid modeling, a time series is considered as a combination of linear auto-correlated and a non-linear component structure, where L_t is the linear component and N_t is the non-linear component which should be estimated by the data. Thus, by first using the ARMA model, the linear component was modeled. Accordingly, the linear model residuals merely include non-linear parts. Then, the residuals are input to NARX neural network, which models the residuals and the hybrid model results were created by summing up the linear and non-linear components (Equation (15)):

$$Y_t = N_t + L_t \qquad (15)$$

Indices for evaluating the models

The following indices are used to select the best structure between ARMA and hybrid ARMA and ANN (ARMA-NARX). There are many investigations in which these indices have been employed to evaluate the results (Yurekli et al. 2005; Mohammadi et al. 2006; Wang 2006; Wang et al. 2006; Rahimi Khoob 2011; Rivandi et al. 2013).

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (Q_{ci} - Q_{oi})^2}{\sum_{i=1}^{n} (Q_{oi} - \overline{Q_{obs}})^2} \qquad (16)$$

$$MARE = \frac{\left(\sum_{i=1}^{n} (|Q_{oi} - Q_{ci}|/Q_{oi})\right)}{n} \qquad (17)$$

where Q_{ci} is the calculated discharge in the month i (m^3/day), Q_{oi} is the observed discharge in the month i (m^3/day) and n is the number of data and MARE is the mean absolute relative error. The MARE is a suitable index for showing the trend in forecasting. However, similar to other indices, this index has deficiencies. For instance, it increases the error in peak flow, and it does not consider a delay in the forecasting peak flow. Therefore, in this paper, although considering the advantages of MARE error index and considering the delay errors in estimating the time and the amount of the peak flow, a new index called Integrated Index for Flow Forecasting Evaluation (IIFFE) is introduced. This index is composed of three indices of forecasting MARE, temporal delay relative error (TOPRE) and maximum relative error (MAXRE) with equal weighting for all the three indices. TOPRE is calculated by subtracting the number of the month of observed and forecasted peak flow, divided by the number of the month of observed peak flow. In other words, it shows a temporal delay in calculating peak flow. Moreover, MAXRE is calculated by subtracting the observed and forecasted discharge values, divided by the observed peak flow. In other words, it is the relative error of calculating peak flow. In calculating IIFFE, each of the indices above is normalized through being divided to its maximum value (Equation (20)):

$$MAXRE = \frac{|MAX_O - MAX_C|}{MAX_O} \qquad (18)$$

$$TOPRE = \frac{|TOP_O - TOP_C|}{TOP_O} \qquad (19)$$

$$IIFFE = w_1 * \frac{MARE}{MARE_{MAX}} + w_2 * \frac{TOPRE}{TOPRE_{MAX}}$$
$$+ w_3 * \frac{MAXRE}{MAXRE_{MAX}} \qquad (20)$$

where MAX_o is the observed peak flow (m³/day), MAX_c is the calculated peak flow (m³/day), TOP_o is the number of the month of observed peak flow, TOP_o is the number of the month of calculating peak flow, $MAXRE$ is the value of relative peak flow error, and $TOPRE$ is the relative peak flow temporal delay.

RESULTS AND DISCUSSION

In this paper, forecasting was primarily accomplished with ARMA model, then it enters the neural network and afterwards, the effect of adding rainfall, SOI, NAO, PGSST, and RSST indices on forecasting accuracy were evaluated. The impact of rainfall data, SOI, NAO, PGSST, and RSST on the accuracy of the model was measured in single, binary, and triplet and its result was the comparison of 80 models. In Table 1, ten choices of models are presented in order to show how the indicators are introduced into the model.

In the time series analysis, it is necessary to investigate the normality of the data. The importance of normalized data is because of the time series theory which is developed by normality and when the data are not standard it is necessary to normalize them by some methods like the Log of them. According to the qq-plot diagram, and also fitting

unsuitable data on the normal index line, we decided to normalize these abnormal data, and because of the fluctuations of current data, logarithmic based on the number of Neper was used.

After normalizing the data and removing the seasonal trend, the ACF and PACF graphs were examined. Since ACF and PACF graphs were both attenuating within the allowed limits, ARMA (p,q) was the proper model. The ARMA model with 28-year flow discharge (1982–83 to 2009–10) for Telezang station was trained in monthly scale and forecasting for 12 months of 2010–11 was examined using various inputs. Using this period is due to lack of access to the newer data. However, the changes in the new data can be a component of uncertainty in results for future research.

Among forecasting models, ARMA(1,1) with the least amount of AIC and the number of parameters was selected. Then the auto-regressive parameter in delay step one is 0.8396 (α_1) and the moving average parameter in delay step one is −0.2851 (β_1).

After selecting the best model, the goodness of fit test was tested to demonstrate independency and stationarity of residuals through investigating the Ljung–Box statistics. The results of the statistics for ARMA(1,1) are shown in Figure 4. Considering that the results were within the allowed limits, the residuals were steady, and since the

Table 1 | A selection of the simulated models

| Model name | Model input | | | Error indices | | | |
	Discharge forecasted by ARMA	NAO, SOI/PGSST, RSSST	Rainfall station	MARE	TOPRE	MAXRE	IIFFE
MARMA-NARX(1)	Yes	–	–	0.114	0.167	0.227	0.805
MARMA-NARX(2)	Yes	–	Keshvar Sorkhab	0.132	0.167	0.005	0.599
MARMA-NARX(6)	Yes	NAO		0.16	0.17	0.11	0.77
MARMA-NARX(7)	Yes	NAO	Keshvar Sorkhab	0.16	0.17	0.02	0.67
MARMA-NARX(30)	Yes	NAO,SOI	Sepid Dasht, Keshvar Sorkhab, Gelerood	0.14	0.17	0.11	0.72
MARMA-NARX(62)	Yes	RSSST,NAO,SOI	Keshvar Sorkhab	0.13	0.17	0.10	0.70
MARMA-NARX(60)	Yes	PGSST,NAO,SOI	Sepid Dasht, Keshvar Sorkhab, Gelerood	0.16	0.17	0.24	0.92
MARMA-NARX(66)	Yes	PGSST,RSSST,NAO	–	0.15	0.17	0.08	0.72
MARMA-NARX(79)	Yes	PGSST,RSSST,SOI,NAO	Gelerood	0.16	0.17	0.07	0.72
MARMA-NARX(80)	Yes	PGSST,RSSST,SOI,NAO	Sepid Dasht, Keshvar Sorkhab, Gelerood	0.15	0.17	0.03	0.67

Standardized Residuals

ACF of Residuals

p values for Ljung-Box statistic
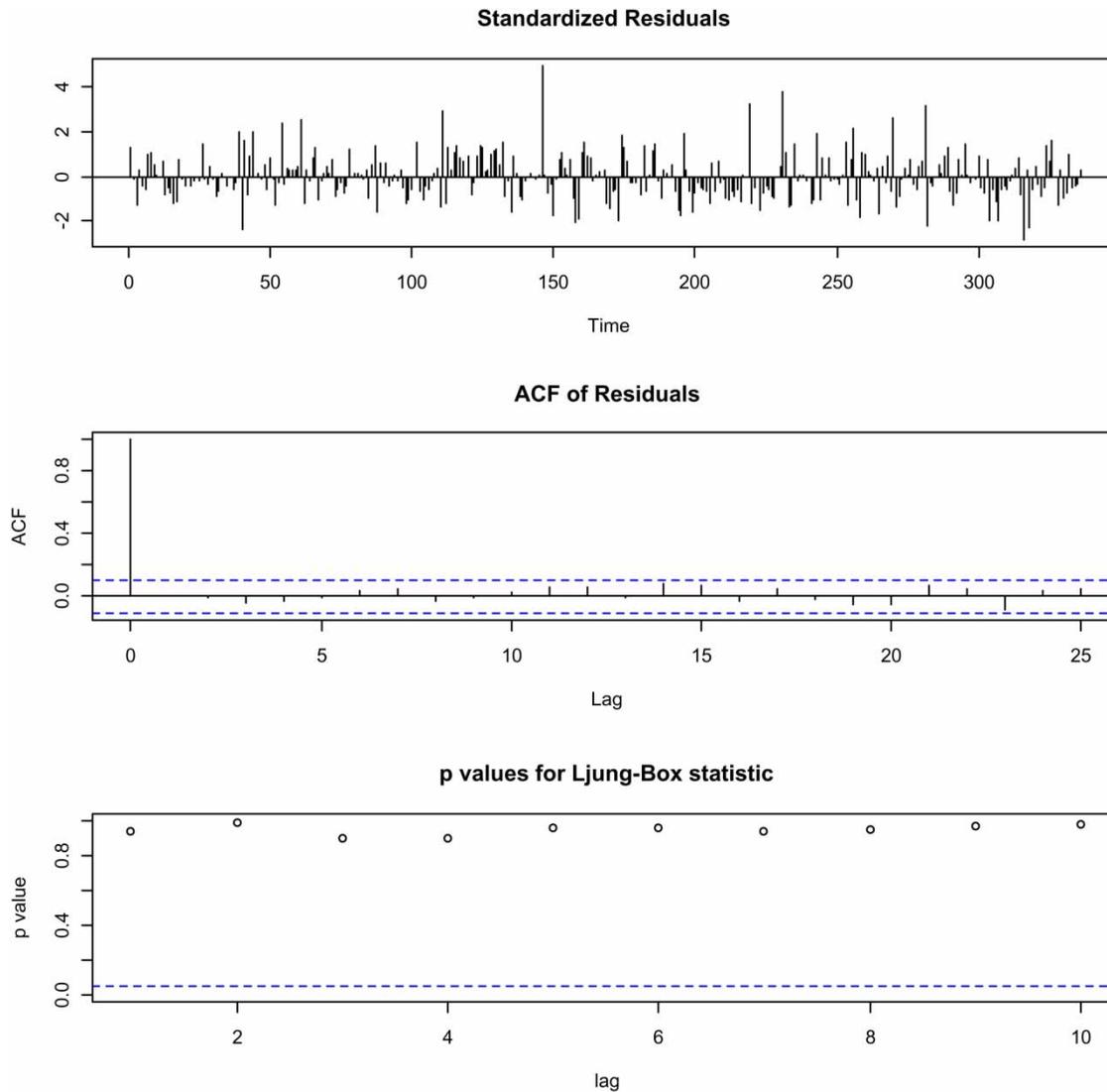
Figure 4 | The results of tests on ARMA(1,1) residuals.

residuals did not have any trends and *p*-value more signifi-
cant than the test meaningfulness level (5%), the H_0 is
accepted, and the residuals were confirmed to be random.
Accordingly, ARMA (1,1) was an appropriate model.

Successful applications of time series models have
been reported in hydrological forecasting (Tyralis &
Papacharalampous 2017; Papacharalampous *et al.* 2018a).
Papacharalampous *et al.* (2018a) examined how various
handling of the seasonality and non-normality affects the
outcomes of the automatic univariate time-series models in
forecasting monthly temperature and precipitation, and
they showed that monthly temperature and precipitation

could be forecasted acceptably which could be enhanced
applying other methods.

The result of comparing the forecasting with the
observed hydrograph is shown in Figure 5. It is observed
that the forecasted values in most points are significantly dis-
tant and peak flow has been estimated with an error in the
amount and with a one-month delay. To find the amount
of error, *MARE* index is calculated, that is 0.55 for the fore-
cast when using ARMA.

In the following step, the ARMA model residuals were
inputted to NARX neural network to forecast the inflow to
Dez reservoir using hybrid ARMA-NARX model. As the
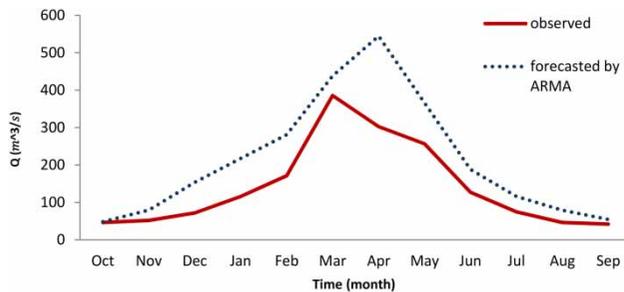
Figure 5 | The forecasting results of ARMA(1,1) model.

forecasting was carried out on a monthly basis, the name of monthly ARMA-NARX model here has been abbreviated to MARMA-NARX. Figure 6 shows the comparison between ARMA and the hybrid model (in which the mere input is the ARMA forecasting residuals) forecasts. Accordingly, the results of the hybrid model have improved concerning the *MARE* from 0.55 to 0.114, for ARMA (1,1) and MARMA-NARX (1), respectively. There are a great many investigations for which successful applications of machine learning models have been reported in hydrological forecasting (Papacharalampous *et al.* 2017a, 2017b, 2017c, 2018b). Papacharalampous *et al.* (2017a) found that some forecasting models are more advantageous than others. Similarly, in this study, we found some models forecasted reservoir inflow better than others did. As Papacharalampous *et al.* (2017a) also indicated, the forecasting errors at each time step of a forecast horizon within a particular case study intensely depend on the tested case regardless of the forecasting model.

Similarly, other inputs, including rainfall data, SOI, NAO, PGSST, and RSST with 80 different compositions are added to the ARMA residuals. These compositions
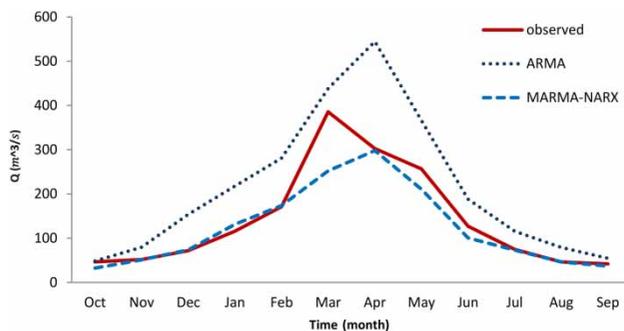

**Figure 6** | Comparison between ARMA and MARMA-NARX(1).

include the singular input of large-scale climate signals and rainfall data, then, the sets of two, three, and four of these indices and consequently considers the large-scale climate signals with rainfall data. Here, it is necessary to explain the architecture of the hybrid network. The numbers of output delays were selected from 1 to 20 months that each delay was fitted with 20 neurons and entirely for each input, 400 forecasts were produced in monthly scale. The number of neurons in the output layer for all models was 1. In this stage, the best model for each input is selected based on the least *MARE*. The operations were accomplished for different inputs to ultimately reach 80 selected models with different inputs. The hydrographs of the 80 models and the observed flow were compared. To compare the results of the 80 models, a new index, *IIFFE*, is defined in this paper that is calculated from the average of *MARE*, the error in forecasting peak and the error in estimating the amount of peak flow (shown in Equation (19)). *MARE* is a suitable index in calculating the amount of error. However, it has not been able to consider time delay error and peak flow value. The index *IIFFE* was calculated for the 80 models, and the eight top models with the least *IIFFE* are shown in Table 2, and they are compared with MARMA-NARX(1) with the mere input of ARMA(1,1) residuals. In this table, the four top models with large-scale climate signals input (without rainfall data) are shown (MARMA-NARX(31), MARMA-NARX (36), MARMA-NARX(56), and MARMA-NARX(76)). Accordingly, the number of input large-scale climate signals to the models is more than two and the models with singular large-scale climate signals input have not been able to produce less *IIFFE* to be among the top four models. Therefore, the most important point is the positive effect of considering multiple large-scale climate signals in improving forecasting. In this stage, none of the models have decreased forecasting peak flow time delay and *MARE*, and top models are selected based on decreasing the amount of peak flow. Considering the least *IIFFE* among the four models (0.59), MARMA-NARX(36) has the best forecast using NAO and PGSST. These four models have almost similar *IIFFE* and also they have small differences in their hydrographs, shown in Figures 7–10. The second four top models, MARMA-NARX(27), MARMA-NARX (48), MARMA-NARX(55), and MARMA-NARX(58) are

**Table 2** | The eight top models with the least *IIFFE* error index

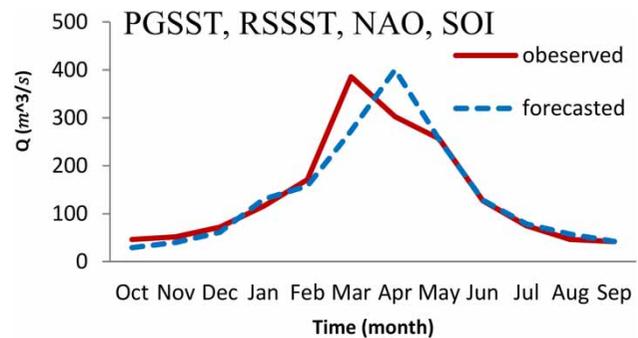| Model	Large-scale climate signals	Rainfall station	MARE	TOPRE	MAXRE	IIFFE
MARMA-NARX(1)			0.11	0.17	0.23	0.80
MARMA-NARX(31)	PGSST, RSST		0.11	0.17	0.14	0.71
MARMA-NARX(36)	PGSST, NAO		0.13	0.17	0.00	0.59
MARMA-NARX(56)	PGSST, NAO, SOI		0.15	0.17	0.03	0.65
MARMA-NARX(76)	PGSST, RSST, NAO, SOI		0.16	0.17	0.04	0.68
MARMA-NARX(27)	NAO, SOI	Keshvar Sorkhab	0.17	0.00	0.03	0.37
MARMA-NARX(48)	RSST, NAO	Sepid Dasht	0.13	0.00	0.11	0.37
MARMA-NARX(55)	RSST, SOI	Sepid Dasht, Keshvar Sorkhab, Gelerood	0.15	0.00	0.05	0.35
MARMA-NARX(58)	PGSST, NAO, SOI	Sepid Dasht	0.14	0.00	0.03	0.32
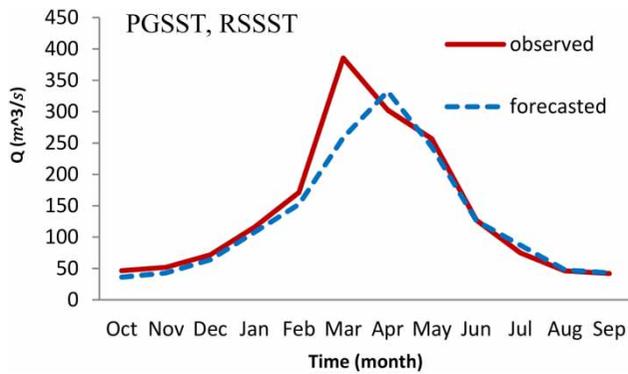
Figure 7 | Forecasting with MARMA -NARX(31).
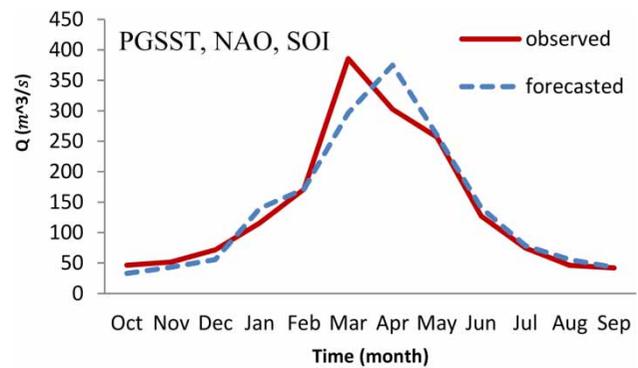
Figure 9 | Forecasting with MARMA -NARX(56).
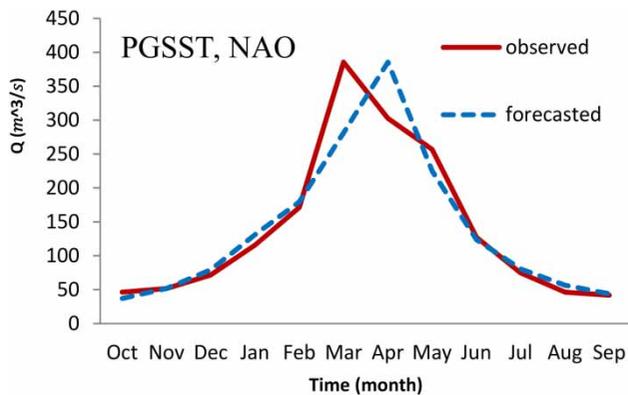
Figure 8 | Forecasting with MARMA -NARX(36).

Figure 10 | Forecasting with MARMA-NARX(76).

the models with large-scale climate signals and rainfall data. Considering Table 2, the assemblage of large-scale climate signals and rainfall data had a positive effect on decreasing *IIFFE* and time delay error in forecasting peak flow, which can be observed in Figures 7–14 (the inputs which have

been added to the hybrid model in addition to the ARMA residuals are shown in each graph). Moreover, similar to what was observed in the previous four models (MARMA-NARX(31), MARMA-NARX(36), MARMA-NARX(56), and MARMA-NARX(76)), these four models had the best results while using multiple large-scale climate signals and rainfall data, not the singular use of large-scale climate signals and
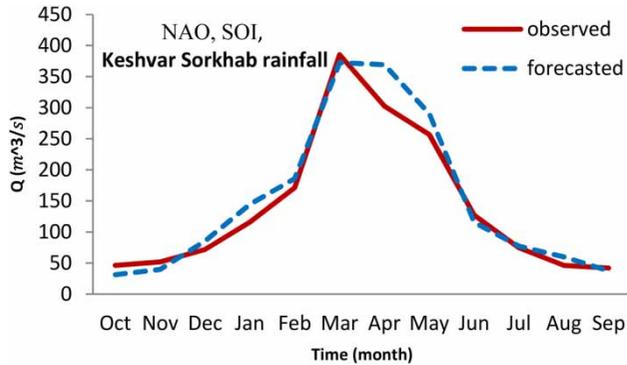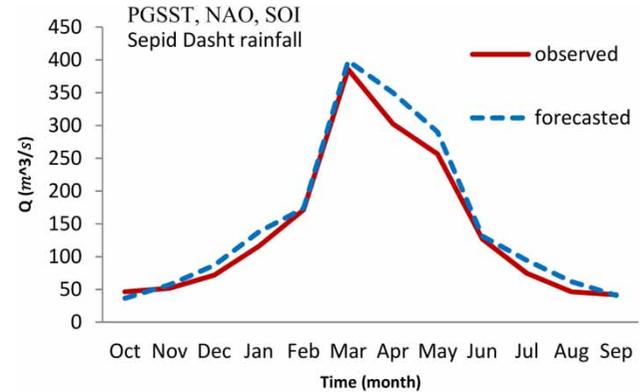
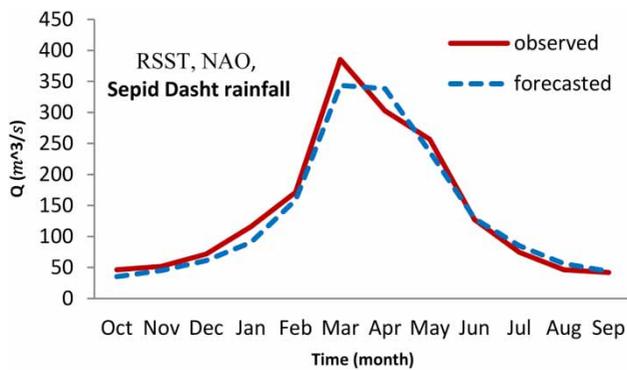**Figure 11** | Forecasting with MARMA-NARX(27).

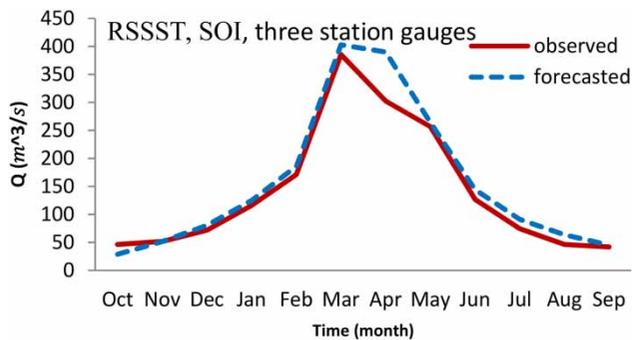
Figure 12 | Forecasting with MARMA-NARX(48).

Figure 13 | Forecasting with MARMA-NARX(55).

rainfall data. Another point shown in the table is that the input of the rainfall data singularly and without using large-scale climate signals is not among the eight top models with the least *IIFFE*.

Considering the aforementioned points, the singular input of large-scale climate signals or rainfall data did not have the least *IIFFE*, and they are not among the eight top

Figure 14 | Forecasting with MARMA-NARX(58).

models. In contrast, multiple large-scale climate signals, especially with rainfall data as input, have led to decreasing the *IIFFE*. The most successful models in decreasing *IIFFE* were the models with multiple large-scale climate signals and rainfall data that have decreased *IIFFE* by decreasing the time delay of peak flow (*TOPRE*). Ultimately, MARMA-NARX(58) had the least *IIFFE* (0.32) with SOI, NAO, and PGSST along with rainfall data of Sepiddasht-Sezar station as input. The most effective climate index is PGSST and has been the input to five of eight selected models. It is necessary to mention that the PGSST is spatially the nearest index for the study area (Dez dam). Therefore, the spatial proximity of the selected indices can cause a positive effect on increasing the forecasting accuracy. In what follows, the best forecast is discussed based on comparing ARMA and MARMA-NARX models.

As seen in Figure 15, the results of the best ARMA forecast have a significant difference with observed hydrograph, and considerable error is seen on forecasting the peak
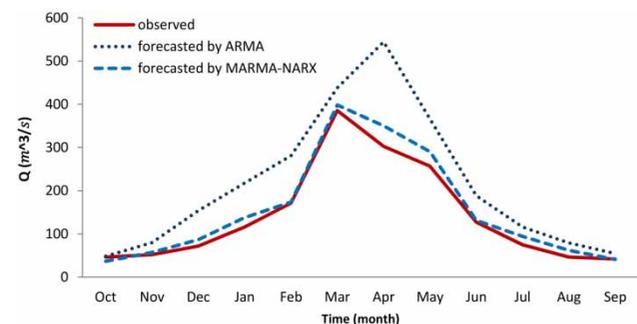

Figure 15 | The comparison between forecasts of ARMA and MARMA-NARX.

discharge and its time. Compared to it, the hybrid MARMA-NARX(58), the best MARMA-NARX model, has good forecasting except in April and the *IIFFE* of the best ARMA model (7.29) is improved to 0.32 in the best MARMA-NARX model.

Considering the results, ARMA forecasted a good trend for flow discharge, but it is considerably distant from the observed value, while the hybrid MARMA-NARX has increased the forecasting accuracy by 96%. Therefore, separating the linear and non-linear components of flow increases the forecasting accuracy. Applying ARMA to forecast the linear component of flow and using the capabilities of the ANN model to forecast the non-linear component of discharge led to the best forecasting results of the hybrid model MARMA-NARX(58) which have *IIFFE* of 0.32 compared to 7.29 of ARMA, representing 96% enhancement. It is possible to conclude that where the linearity or non-linearity of the existing patterns in the system is not clear or the studied system has both patterns, the hybrid model can yield better results.

There are many investigations in which the influence of the climatic indexes on the characterization of the flow regime have been analyzed (Hosseinzadeh Talaee *et al.* 2012; Tabari *et al.* 2013; Meidani & Araghinejad 2014). Time series and ANN models are the models which are used in flow forecasting. These models will work strongly in linear forecasting or non-linear forecasting, but not in both patterns simultaneously. The hybrid model is the suggested model which allowed using benefits of both linear and non-linear models simultaneously.

There are a great many linear pattern and non-linear models used for monthly flow forecasting. ARMA is an example for linear and NARX is a good example for non-linear. Each of these models has advantages and disadvantages which we mentioned in the Introduction. There are many studies which show these models' abilities for flow forecasting, but because the flow is one of the components of the hydrological cycle which follows linear and non-linear patterns during the different periods of the year, it is necessary to use the advantages of both models for flow forecasting. Thus, the hybrid model has been used to increase the accuracy of the flow.

In this article, we have intended to enhance the accuracy of flow forecasting in time series models and also ANN with a hybrid model.

In this research, we used all the data that were available for our research. However, to improve the accuracy of the estimations, adding more data is suggested for future works.

CONCLUSION

This paper focused on forecasting the following year inflow to Dez reservoir, using hybrid MARMA-NARX model and comparing the results with ARMA model. The significant difficulties of this work are providing data and information of runoff and precipitation after solving the programming problems. If the data are more accessible and more extended, accuracy for prediction will be increased, and uncertainty will be decreased.

Eighty models with different input including singular input of large-scale climate signals of SOI, NAO, PGSST, and RSST along with data of three rainfall stations, Keshvar Sorkhab, Sepid Dasht, and Gelerood, and sets of two, three, and four input of the large-scale climate signals were used, and then, these indices with singular and joint rainfall station were examined, and ultimately the most successful model was determined by using proposed IIFFE error index. The hybrid MARMA-NARX(58) with SOI, NAO, PGSST, and rainfall data as input has the best forecast among MARMA-NARX models, which has improved *IIFFE* from 7.29 (for the best ARMA model) to 0.32. Comparing the models shows that both ARMA and hybrid MARMA-NARX have been able to forecast the trend of long-term monthly flow for a year. In other words, both models have had the ability to forecast the flow trend in the following year with different accuracies. However, considering the 96% improvement in the results of the hybrid model, we can determine better flow forecasting when the linearity or non-linearity of the patterns in the system is unknown, or the system has both patterns.

According to the results above, the large-scale climate signals in monthly scale were useful in flow forecasting and using multiple indices yields better performance of the model compared to the model with singular input. Rainfall data as input had an important role in improving the forecasting results, especially in decreasing the forecasting delay error. The joint input of rainfall and large-scale climate signals shows better results than the separate input of the

data to the model. The effect of the large-scale climate signals on Dez dam basin is clear; however, since forecasting is solely accomplished in a year and the impact of these indices varies in different years and seasons, this paper is not able to determine the dominant index on the climate of the basin. Merely based on the results that the PGSST is the input of five models among the eight top models, it is possible to consider the impact of this index to be more effective.

Lack of more data is the limitation of this study. Hence, using more and recent data could improve the accuracy of the models. Moreover, employment of modern hybrid models such as ANN-Fuzzy-ARIMA and quantum models is suggested for future work.

REFERENCES

Abdul Jaleel, E. & Aparna, K. 2015 Identification of ethane-ethylene distillation column using neural network and ANFIS. In: *2015 Fifth International Conference on Advances in Computing and Communications*, 2–4 September, Kochi, India.

Aladag, C. H., Egrioglu, E. & Kadilar, C. 2009 Forecasting nonlinear time series with a hybrid methodology. *Applied Mathematics Letters* **22** (9), 1467–1470.

Amiri, E. 2015 Forecasting daily river flows using nonlinear time series models. *Journal of Hydrology* **527**, 1054–1072.

Ashoori, H., Abrishamchi, A., Moradkhani, H. & Tajrishy, M. 2008 Effect of El Nino-Southern oscillation and Pacific decadal oscillation on inflow volume to the Zaiandeh-Rood Dam. In: *3rd Iran Water Resources Management Conference*, Tabriz, Iran.

Azimi, M., Tajrishi, M. & Abrishamchi, A. 2010 Seasonal prediction of Dez streamflow using larg-scale climate indices. In: *9th Iranian Hydraulic Conference*. Modares University, Tehran, Iran.

Balaguer, E., Palomares, A., Sorie, E. & Martin-Guerrero, J. D. 2008 Predicting service request in support centers based on nonlinear dynamics, ARMA modeling and neural networks. *Expert Systems with Applications* **34**, 665–672.

Banihabib, M. E., Mousavi, S. M. & Jamali, F. S. 2008 Artificial neural network model to study the spatial and temporal correlation between stations in reservoir inflow forecasting. In: *3rd Iran Water Resources Management Conference*, Tabriz, Iran.

Banihabib, M. E., Bandari, R. & Peralta, R. 2014 NARX recurrent neural-network model for long-term forecasting of daily flow. In: *Spring Runoff Conference*, Utah State University, Logan, UT, USA.

Bazartseren, B., Hildebrandt, G. & Holz, K.-P. 2003 Short-term water level prediction using neural networks and neuro-fuzzy approach. *Neurocomputing* **55** (3), 439–450.

Benmouiza, K. & Cheknane, A. 2016 Small-scale solar radiation forecasting using ARMA and nonlinear autoregressive neural network models. *Theoretical and Applied Climatology* **124** (3–4), 945–958.

Borhani, D. A. & Fatehi, M. A. 2008 Application of artificial neural network in stream flow forecasting using climatic indices (case study: Nazloochay River basin). *Journal of Faculty of Engineering (University of Tabriz)* **35**, 25–36.

Bradbury, J. A., Dingman, S. L. & Keim, B. D. 2002 New England drought and relations with large scale atmospheric circulation patterns. *Journal of the American Water Resources Association* **38** (5), 1287–1299.

Chandimala, J. & Zubair, L. 2007 Predictability of stream flow and rainfall based on ENSO for water resources management in Sri Lanka. *Journal of Hydrology* **335** (3), 303–312.

Chau, K. W. & Wu, C. L. 2010 A hybrid model coupled with singular spectrum analysis for daily rainfall prediction. *Journal of Hydroinformatics* **12** (4), 458–473.

Chen, X. Y., Chau, K. W. & Busari, A. O. 2015 A comparative study of population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model. *Engineering Applications of Artificial Intelligence* **46** (A), 258–268.

Cryer, J. D. & Chan, K.-S. 2008 *Time Series Analysis With Applications in R*. Springer, New York, USA.

Diaz-Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G. & Moncada-Herrera, J. A. 2008 A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: the case of Temuco, Chile. *Atmospheric Environment* **42** (35), 8331–8340.

Fallah Ghalhari, G. A. & Khoshhal, J. 2010 Spring rainfall prediction of Khorasan-e Razavi Province, based upon climatic large scale signals by using artificial neural network. *Physical Geography Research Quarterly* **41** (69), 115–133.

Fatehi Marj, A. & Borhani Darian, A. 2005 Forecasting of seasonal streamflow using climate indicators. Case study: the Simineh Rood. In: *5th Iranian Hydraulic Conference*, Kerman, Iran.

Fathian, F., Fakheri Fard, A., Dinpashoh, Y. & Mousavi Nadoushani, S. S. 2017 Performance evaluation of linear (ARMA) and threshold nonlinear (TAR) time series models in daily river flow modeling (case study 58; Upstream basin rivers of Zarrineh Roud Dam). *Majallah-i āb va Khāk* **30** (5), 1440–1460.

Gairaa, K., Khellaf, A., Messlem, Y. & Chellali, F. 2016 Estimation of the daily global solar radiation based on Box–Jenkins and ANN models: a combined approach. *Renewable and Sustainable Energy Reviews* **57**, 238–249.

Gheiby, A. & Noorafshan, M. 2013 Case study: ENSO events, rainfall variability and the potential of SOI for the seasonal precipitation predictions in Iran. *Scientific Research* **2**, 34–45.

Gholami, V., Chau, K. W., Fadaee, F. & Torkaman, J. 2015 Modeling of groundwater level fluctuations using

dendrochronology in alluvial aquifers. *Journal of Hydrology* **529** (3), 1060–1069.

Hosseinzadeh Talaee, P., Tabari, H. & Sobhan Ardakani, S. 2012 Hydrological drought in the west of Iran and possible association with large-scale atmospheric circulation patterns. *Hydrological Processes* **28** (3), 764–773. DOI: 10.1002/hyp. 9586.

Humphrey, G. B., Gibbs, M. S., Dandy, G. C. & Maier, H. R. 2016 A hybrid approach to monthly streamflow forecasting: integrating hydrological model outputs into a Bayesian artificial neural network. *Journal of Hydrology* **540**, 623–640.

Kamil, M. S. & Razali, A. M. 2015 Time series SARIMA models for average monthly solar radiation in Malaysia. In: *2015 International Conference on Research and Education in Mathematics*, Kuala Lumpur, Malaysia.

Karamouz, M. & Araghinejad, S. 2005 *Advanced Hydrology*. Polytechnique University Publication, Tehran, Iran.

Khashei, M. & Bijari, M. 2011 A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied Soft Computing* **11** (2), 2664–2675.

Kilinç, I. & Ciğizoğlu, K. 2003 *Reservoir Management Using Artificial Neural Networks*. State Hydraulic Works, Istanbul, Turkey.

Koutroumanidis, T., Ioannou, K. & Arabatzis, G. 2009 Predicting fuelwood prices in Greece with the use of ARIMA models, artificial neural networks and a hybrid ARIMA–ANN model. *Energy Policy* **37** (9), 3627–3634.

Mandal, S. & Prabaharan, N. 2006 Ocean wave forecasting using recurrent neural networks. *Ocean Engineering* **33** (10), 1401–1410.

Meidani, E. & Araghinejad, S. 2014 Long-lead streamflow forecasting in the southwest of Iran by sea surface temperature of the Mediterranean Sea. *Journal of Hydrologic Engineering* **19** (8). https://doi.org/10.1061/(ASCE)HE.1943-5584.0000965.

Mishra, A. K. & Desai, V. R. 2005 Drought forecasting using stochastic models. *Stochastic Environmental Research and Risk Assessment* **19** (5), 326–339.

Mohammadi, K., Eslami, H. & Kahawita, R. 2006 Parameter estimation of an ARMA model for river flow forecasting using goal programming. *Journal of Hydrology* **331** (1), 293–299.

Papacharalampous, G. A., Tyralis, H. & Koutsoyiannis, D. 2017a Comparison of stochastic and machine learning methods for multi-step ahead forecasting of hydrological processes. In: *19th EGU General Assembly*, Vienna, Austria. doi:10.20944/preprints201710.0133.v1.

Papacharalampous, G. A., Tyralis, H. & Koutsoyiannis, D. 2017b Forecasting of geophysical processes using stochastic and machine learning algorithms. In: *Proceedings of the 10th World Congress of EWRA on Water Resources and Environment*, Athens, Greece. doi:10.13140/RG.2.2.30581. 27361.

Papacharalampous, G., Tyralis, H. & Koutsoyiannis, D. 2017c Error evolution in multi-step ahead streamflow forecasting

for the operation of hydropower reservoirs. *Preprints*, 2017100129. doi: 10.20944/preprints201710.0129.v1.

Papacharalampous, G., Tyralis, H. & Koutsoyiannis, D. 2018a Predictability of monthly temperature and precipitation using automatic time series forecasting methods. *Acta Geophysica* **66** (1), 1–25.

Papacharalampous, G., Tyralis, H. & Koutsoyiannis, D. 2018b One-step ahead forecasting of geophysical processes within a purely statistical framework. *Geoscience Letters* **5** (1), 12.

Rahimi Khoob, A. 2011 A prediction of maximum monthly precipitation recorded at Ilam meteorological station based on Persian Gulf and Red Sea surface temperature through recordings data mining method. *Iranian Journal of Soil and Water Research* **42** (1), 1–7.

Raj, R. & Sivanandan, K. S. 2017 Elbow joint angle and elbow movement velocity estimation using NARX-multiple layer perceptron neural network model with surface EMG time domain parameters. *Journal of Back and Musculoskeletal Rehabilitation* **30** (3), 515–525.

Rivandi, A., Nasr-Esfahany, M.-A. & Memarian, M.-H. 2013 Effects of the North Atlantic Oscillation (NAO) on the dominant patterns of geopotential height variability in the Mediterranean region by using empirical orthogonal functions. *Iranian Geophysical Society* **7** (3), 66–77.

Ryu, J. H., Svoboda, M. D., Lenters, J. D., Tadesse, S. T. & Knutson, C. L. 2010 Potential extents for ENSO-driven hydrologic drought forecasts in the United States. *Climatic Change* **101** (3–4), 575–597.

Salas, J. D. 1980 *Applied Modeling of Hydrologic Time Series*. Water Resources Publication, LLC, Mississippi, USA.

Tabari, H., Abghari, H. & Hosseinzadeh Talaee, P. 2013 Impact of the North Atlantic Oscillation on streamflow in Western Iran. *Hydrological Processes* **28**, 4411–4418. DOI: 10.1002/ hyp.9960/.

Taormina, R., Chau, K. W. & Sivakumar, B. 2015 Neural network river forecasting through baseflow separation and binary-coded swarm optimization. *Journal of Hydrology* **529** (3), 1788–1797.

Tyralis, H. & Papacharalampous, G. 2017 Variable selection in time series forecasting using random forests. *Algorithms* **10** (4), 114.

Valipour, M. 2016 Optimization of neural networks for precipitation analysis in a humid region to detect drought and wet year alarms. *Meteorological Applications* **23** (1), 91–100.

Valipour, M., Banihabib, M. E. & Behbahani, S. M. R. 2013 Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *Journal of Hydrology* **476**, 433–441.

Van Calster, T., Baesens, B. & Lemahieu, W. 2017 ProfARIMA: a profit-driven order identification algorithm for ARIMA models in sales forecasting. *Applied Soft Computing* **60**,

775–785. DOI: http://dx.doi.org/doi:10.1016/j.asoc.2017.
02.011.

Vrieze, S. I. 2012 Model selection and psychological theory: a
discussion of the differences between the Akaike Information
Criterion (AIC) and the Bayesian Information Criterion
(BIC). *Psychological Methods* **17** (2), 228–243.

Wang, W. 2006 *Stochasticity, Nonlinearity and Forecasting of
Streamflow Processes*. Ios Press, Amsterdam, The Netherlands.

Wang, W., Van Gelder, P. H., Vrijling, J. & Ma, J. 2006 Forecasting
daily streamflow using hybrid ANN models. *Journal of
Hydrology* **324** (1), 383–399.

Wang, W. C., Chau, K., Xu, D. & Chen, X. Y. 2015 Improving
forecasting accuracy of annual runoff time series using
ARIMA based on EEMD decomposition. *Water Resources
Management* **29** (8), 2655–2675.

Wu, C. L., Chau, K. W. & Fan, C. 2010 Prediction of rainfall time
series using modular artificial neural networks coupled with
data-preprocessing techniques. *Journal of Hydrology*
389 (1–2), 146–167.

Yurekli, K., Kurunc, A. & Ozturk, F. 2005 Application of linear
stochastic models to monthly flow data of Kelkit Stream.
Ecological Modelling **183** (1), 67–75.