

# Using exogenous variables to improve precipitation predictions of ANNs in arid and hyper-arid climates

Hossein Bari Abarghouei<sup>1</sup> · Seyed Zeynalabedin Hosseini<sup>2</sup>

Received: 20 June 2015 / Accepted: 19 September 2016 / Published online: 6 October 2016  
© Saudi Society for Geosciences 2016

**Abstract** Because of scarcity and high variability of rainfall in arid areas, from one hand, reliable prediction of precipitation in such regions is considerably difficult. Furthermore, in some cases, shortage of observation data and several other limitations may intensify complexity of the forecasting. On the other hand, these regions highly suffer from low availability of water which necessitates development of an appropriate modeling approach to provide as precise as possible predictions of precipitation. Artificial neural networks (ANNs) are expected to be a powerful tool in capturing and analyzing high interannual variability of precipitation in arid climates and, subsequently, in proper prediction of precipitation fluctuations in the future. The end of this paper is to improve ANN predictions of precipitation in arid climates using better training of the network. To this end, two approaches were applied. In the first one, just the rainfall monthly data were considered as input. In the second approach, in addition to precipitation, several exogenous variables of precipitation are considered as input to predict precipitation. The chosen exogenous parameters are either effective on or relevant to the precipitation patterns. Then, several lag times, hidden layer sizes, and training algorithms for different running sums are used in order to produce best forecasts. It was shown that the performance of networks increases significantly by importing more external factors as inputs. The bigger time scales also exhibited better performances. In all the five time scales, smaller lag times (especially one month), bigger hidden layer sizes (especially between 31 and 40), and GDX training algorithm

presented the best performance. The highest obtained performance was presented by the network with 10 inputs, 1 month lag, 36 hidden layers, and CGF training method in 18 months running sum with  $R^2$  of 0.93.

**Keywords** Arid climates · Artificial neural networks · Exogenous parameters · Water scarcity · Rainfall prediction

## Introduction

Since arid regions of the world seriously suffer from inadequate water available, planning for optimal utilization of the current and future available water resources is very important. Therefore, prediction of future precipitation is one of the important requirements of water resource management in such regions, which has several advantages and applications. For example, a flood warning system for fast responding catchments may require a quantitative rainfall forecast to increase the lead time for warning. Similarly, a rainfall forecast system provides advanced information for drought management. Therefore, providing an accurate quantitative rainfall forecast is an important challenge in many real-world applications such as catchment management (Luk et al. 2001), flood predictions (Yu et al., 2006; Kisi and Cimen 2011), agriculture (Wei et al. 2005), etc.

Seasonal variations in rainfall pattern may modify the hydrological cycle and environmental processes in addition to the vegetation and the entire ecosystem (Lazaro et al. 2001; Ni and Zhang 2000). Generally, although the scale of these variations differs in different regions, greater spatial and temporal variations of rainfall in the arid zone rather than more humid regions are confirmed (Asadi Zarch et al. 2015). Consequently, long sequences with very little rain are common in such regions. Long dry periods may cause changes in the vegetation

✉ Hossein Bari Abarghouei  
h\_abarghouei@yahoo.com

<sup>1</sup> Department of Agriculture, Payame Noor University, Yazd, Iran

<sup>2</sup> Faculty of Natural Resources and Desert Studies, Yazd University, Yazd, Iran

and structure of the soil, especially at its surface, which have a considerable effect on infiltration and the production of runoff. Therefore, extreme variability of rainfall in time and space not only causes limitations and complexities in its prediction in arid and hyper-arid areas (Batisani and Yarnal 2010; Srivastava et al. 2010), huge degradation of the ecosystem is also resulted.

Although several numerical prediction models, from one hand, have been developed currently, they are still unable to produce quantitative precipitation forecasts at proper spatial and temporal resolution (Dierer et al. 2009). Global climate model (GCM) projections, for example, require downscaling prior to be used in case studies. On the other hand, standard time series models such as autoregressive moving average (ARMA) are also extensively used for hydrological time series forecasting. These models are fundamentally linear and assume that data are stationary. Therefore, they are not able to capture non-stationarity and non-linearity in hydrologic data.

Artificial intelligence approaches with several advantages over numerical and classic time series models have been developed to overcome this problem. Artificial neural network is a branch of artificial intelligence developed in the 1950s. Artificial neural networks (ANNs) emulate the biological neural system, distributing the computation process to small and simple processing units, i.e., neurons (nodes). ANNs are generally presented as systems of interconnected “neurons,” which send messages to each other. The connections have numeric weights that can be tuned based on experience, making neural nets adaptive to inputs and capable of learning. Although meaningful physical interpretations are not usually provided to describe the weights obtained from neural networks, many studies have been successfully carried out in different fields, e.g., Zhang and Govindaraju (2000), Birikundavyi et al. (2002), Jain and Indurthy (2003), Jain and Srinivasulu (2004), Daliakopoulos et al. (2005), Mohammadi et al. (2005), Moosavi et al. (2013a, b), etc.

It can be claimed that neural networks which do not pose any restrictive hypothesis are useful models for data with high variations and cyclicity. Therefore, since difficulty of precipitation prediction in arid climates is mostly due to its high interannual variability, ANNs likely present most accurate rainfall predictions over arid regions. The three-layer feed-forward neural network as the most common ANN model is employed for the present study.

The simplest and most common types of neural networks are based on multilayer perceptrons (MLPs) that create static models, where the input–output map depends only on the present input. Several studies applied multilayer perceptrons to a set of predictive variables, carefully chosen to be related to rainfall, and with data from precipitation gauges (pluviometers) to obtain rainfall quantity (Hung et al. 2009). Hung et al. (2009) used a simple persistent model and a feed-forward neural network model for hourly rainfall data and concluded that a feed-

forward neural network model with tangent hyperbolic transfer function produces the best rainfall forecasts. Moustiris et al. (2011) compared ANNs with classical statistical models to forecast precipitation. They showed that the neural networks are satisfactorily able to forecast precipitation. Adya and Collopy (1998) showed that neural networks produced more accurate predictions when they are effectively implemented and validated. They confirmed that effectively implemented and validated neural network models outperformed linear regression, step-wise polynomial regression, multiple regression, discriminant analysis, logic models, and rule-based system. McCullagh et al. (1995) used neural networks to estimate the 6-h rainfall over the southeast coast of Tasmania. They showed that ANN models can produce acceptable results.

To put it simply, rain falls when vapor is saturated in the air. However, rainfall is an end product of a number of complex atmospheric processes which considerably differs both spatially and temporally. Therefore, several climatic parameters (exogenous variables or externals) are playing important roles in rainfall process. Even if the rainfall processes can be described concisely and completely, the volume of calculations involved in this type of precipitation simulation and prediction may be prohibitive. Also, the data available to assist in the definition of control variables for the process models, such as rainfall intensity, are usually limited in both the spatial and temporal dimensions (Luk et al. 2001). The accuracy of the ANN outputs depends on the ability of the network in simulating the rainfall procedure. Therefore, better training of the ANN network results in more precise predictions of precipitation. Regarding this, the aim of this research is improving the networks’ training process. To this end, exogenous variables of precipitation as well as special models are employed to the ANNs.

Therefore, this paper is motivated by this hypothesis that in arid and hyper-arid regions, under extreme spatial and temporal variation of precipitation occurrences, the ANN outputs can be enhanced by providing more relevant input data. Selecting input variables is an important and crucial consideration in recognizing the best functional form of statistical models. The task of selecting input variables is common to the development of all statistical models and is largely dependent on the discovery of relationships within the available data to identify suitable predictors of the model output.

This article is aimed to discuss such the enhancement in providing input data in one of the main strategic hyper-arid regions of Iran in Yazd province. To enhance the precision of predicted precipitation by ANNs, different sets of inputs are tested. In the first test, only precipitation time series is used as inputs. Then, several different exogenous inputs, such as temperature, relative humidity, etc., are imported step by step to the models as inputs to select the best set of input data.

Generally, the objective of this research is to predict precipitation in arid regions using feed-forward artificial neural networks using different input sets (with and without exogenous variables), lag times, hidden layer sizes, and train algorithms in different running sums. The rest of the paper is as follows: a description of the study area, implementation of the method, presenting the results and interpretation of the results, and conclusion.

**Materials and methods**

**Study area**

Iran is situated in the mid-latitude belt of arid and semi-arid areas of the Earth. The arid and semi-arid regions cover more than 60 % of the country. In these regions, the rains are highly variable in time, space, amount, and duration, and water is the most important restrictive factor for biological and agricultural activities. The study is performed in Yazd province, Iran. Yazd province is located beside the central mountains, adjacent to the kavir. The climate varies from cold steppic to semi-desert. Average annual precipitation of the study area ranges from 250 mm in Shirkouh Mountain to 80 mm in the margin of Kavir-e-Abarkouh. Minimum temperature is recorded in December (8 °C) while the highest temperature touches

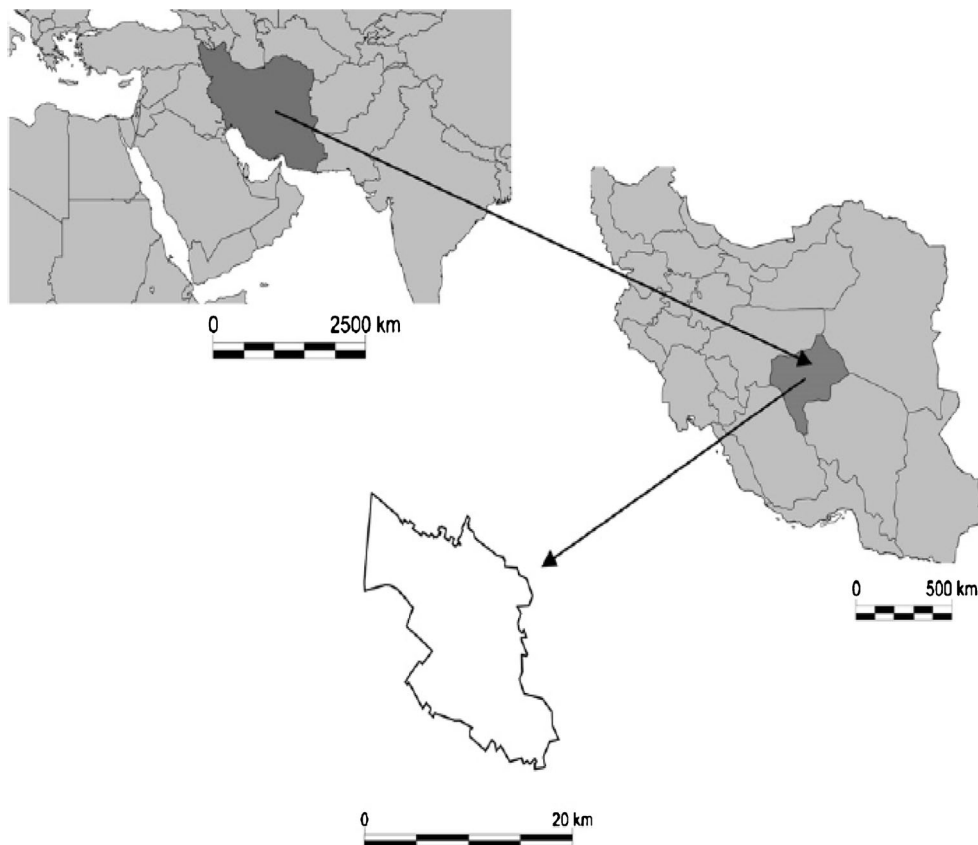
+45 °C in June (Fattahi 1998). The city of Yazd, one of the more ancient cities of Iran, is the capital of the Yazd province, (Fig. 1). Yazd is located in a desert environment with an annual precipitation of 60–70 mm while the rate of potential evapotranspiration is around 1750 mm (Asadi Zarch et al. 2011). Based on the UNESCO aridity index (Asadi Zarch et al. 2015), the climate of Yazd can be classified to arid.

In this study, monthly precipitation amounts for Yazd city are predicted using neural network models. Monthly time series data are obtained from Iran’s meteorological organization site ([www.weather.ir](http://www.weather.ir)) for the period of 1952 to 2010. Total data sets (59 years) are divided into three data sets: training, validation, and test sets. Training data (65 % of the total data) are presented to the network during training, and the network is adjusted according to its error. Validation data (15 %) are presented to the network generalization and to halt training when generalization stops improving. Testing data (20 %) have no effect on training and so provide an independent measure of network performance during and after training (MathWorks 2015).

**Model implementation**

Artificial neural networks (ANN) have been successfully used as a tool for time series prediction and modeling in a variety of application domains. In particular, when the time series is

**Fig. 1** The study area of the research



noisy and the underlying dynamical system is non-linear and not easily approached through analytical means, ANN models frequently outperform standard techniques (Huon and Poo 2013). A neural network model is composed of many artificial nodes that are linked together. The objective is to transform the inputs into meaningful outputs. In this study, the most common ANN model, i.e., the three-layer feed-forward neural network model, which was trained with back propagation method (BPANN), is used for precipitation prediction. In this model, input values are first imported to the nodes in the input and the hidden layers; then, they are processed and passed to the next layer. The number of nodes in the input layer and the output layer are considered equal to the number of input and output parameters (Mar and Naing 2008). Feed-forward neural network architecture and the corresponding learning algorithm can be viewed as a generalization of the popular least mean square (LMS) algorithm (Haykin 1999).

### Background of the model

All the selected parameters affect precipitation amount but not in a quite direct way. For example, if the humidity is more or relative humidity (the amount of water vapor in the air as a percentage of its total capacity) becomes maximum, it leads to the formation of clouds and subsequent precipitation. Since warm air will hold more moisture than cold air would, the percentage of relative humidity must change with changes in air temperature. For example, at 30° air temperature, an air parcel may be saturated and will hold no more molecules of water vapor and the relative humidity can be 100 %. If the temperature of the air parcel is raised to 35°, it can hold more water molecules to reach saturation. So, the air temperature can change the relative humidity. Thus, a cool, dry air mass may actually have a higher relative humidity than a warm, moist air mass. Relative humidity alone can be misleading when comparing atmospheric moisture conditions. Therefore, importing both temperature and humidity to the ANN network may result in a better performance.

Moreover, on one hand, intensifying a variable like wind speed can accelerate evaporation and increase relative humidity. On the other hand, more wind may carry out the water vapor in the air and decrease the humidity. Therefore, the chosen factors affect the amount of precipitation in different ways. So, it is hypothesized that importing exogenous variables leads to better performance of ANN. Also, adding more of these external variables can help our model to simulate precipitation procedure better and predict precipitation more precisely. These assumptions are tested for verification. To this end, two ANN models, one just based on precipitation and the other one based on precipitation and its exogenous variables, are employed.

### NAR model

Various non-linear dynamic models have been proposed in the literature and non-linear autoregressive or NAR is one of them. In this model, the future values of a time series  $y(t)$  are predicted only from the past values of that series. Therefore, there is only one series involved to train the network. The model can be shown as follows:

$$y(t) = f(y(t-1), \dots, y(t-d)) \quad (1)$$

where  $d$  is assigned delay time. In fact in this approach, the following equation is used as the main function. After configuration of input parameters of neural network models, the next step is to train the neural network models with these settings. In the present case, the models are trained by using the following train data format:

$$P_t = f(P_{t-1}, \dots, P_{t-n}) \quad (2)$$

where  $P$  shows the precipitation,  $n$  is the number of inputs based on the lag times (month), and  $t$  shows the time. In this approach, the neural network is used as a time series model such as ARIMA. In fact, both inputs and outputs are precipitation time series. In other words, every precipitation data before time  $t$  could be applied as inputs for output precipitation at time  $t$ .

### NARX model

The non-linear autoregressive network with exogenous inputs (NARX) is an important class of discrete-time non-linear systems (Huon and Poo 2013). This is a powerful class of models which has been demonstrated to be well suited for modeling non-linear systems and specially time series (Diaconescu 2008). This means that the model predicts the current value of a time series based on its relation to the past values of the series and current and past values of the exogenous series (Safavieh et al. 2007). The defining equation for the NARX model is

$$y(t) = f(y(t-1), y(t-2), \dots, y(t-d_y), u(t-1), u(t-2), \dots, u(t-d_u)) \quad (3)$$

where  $y$  is the output,  $u$  is the input, and  $d_u$  and  $d_y$  are the delays of the input and output respectively. In fact, in this approach, more than one climatic parameter is used for precipitation prediction and the network is trained using various climate monthly data such as precipitation, average minimum temperature, average maximum temperature, mean temperature, relative humidity, average wind speed, number of days with storm, number of snowy days, and number of cloudy days. Table 1 indicates the applied climatic data and their symbols for network modeling. Figure 2 illustrates the NARX model used in this research. In the used neural network

**Table 1** The climatic parameters used in the study and their symbols

Climatic data	Symbols
Precipitation	$P$
Average minimum temperature	$T_{min}$
Average maximum temperature	$T_{max}$
Mean temperature	$T_{mean}$
Relative humidity	$RH$
Average vapor pressure	$VP$
Average wind speed	$WS$
Number of days with storm	$Dstorm$
Number of snowy days	$Dsnow$
Cloudy days	$CD$

models, the current time precipitation ( $P_t$ ) is considered as output and the  $V_{t-1}$ ,  $V_{t-2}$ , etc. ( $V$  stands for variables such as precipitation, temperature, wind speed, etc.) are considered as

$$P_{SUM_{M,t}} = f \left( \begin{matrix} P_{t-1}, T_{min_{t-1}}, T_{max_{t-1}}, T_{mean_{t-1}}, RH_{t-1}, VP_{t-1}, WS_{t-1}, Dstorm_{t-1}, Dsnow_{t-1}, CD_{t-1}, \\ P_{t-2}, T_{min_{t-2}}, T_{max_{t-2}}, T_{mean_{t-2}}, RH_{t-2}, VP_{t-2}, WS_{t-2}, Dstorm_{t-2}, Dsnow_{t-2}, CD_{t-2}, \end{matrix} \right) \quad (4)$$

**ANN architecture**

Choosing the right network architecture is an important task of ANN-based studies. A neural network in general consists of highly interconnected layers of neuron-like nodes in which the input, output, and hidden layers are placed between them. The numbers of nodes in input and output layers correspond to the input and output variables of the process, respectively (Dhussa et al. 2014). In this study, for each 3, 6, 9, 12, and 18 running sums, the networks are trained up to 12 months of lag time. According to the above statement, for a 2-month lag time, 20 input variables; for a 3-month lag time, 30 inputs; and for a 12-month lag time, 120 input values were used for each output.

The number of neurons in the hidden layer allows neural networks to determine the patterns and to perform complex non-linear mapping between the input and output variables and plays very important roles for many successful

inputs. So, when the lag time is increased, the numbers of inputs are also increased.

As mentioned before, the precipitation parameter in dry lands such as Yazd city has a considerable variance and also crenulations that reduce the accuracy of the predicted precipitation. Other relevant parameter relations in precipitation are moving average and running sum on precipitation. So, both running sum and moving average could be used. Since the amount of precipitation is very low in this region (like other arid zones), to show results better, running sum is selected as output.

Three, 6, 9, 12, and 18 running sums were selected as output. As the running sum performs the same to different types of monthly precipitation, it has not been normalized. Different lag times were used as input in the second approach. In the first lag time, each of the abovementioned variables in time  $t - 1$  was considered as input. Equation 4 shows the inputs at 2 months of lag time as an example.

applications of neural networks. The number of hidden layers and the number of nodes in each of them are decided by the user and can vary from one to a finite number. It has been verified that only one hidden layer is enough for ANNs to approximate any complex non-linear function with any desired accuracy (Horn et al. 1989). In the case of the popular one hidden layer networks, several practical guidelines exist. These include using “ $2n + 1$ ” (Hect-Nielsen, 1990), “ $2n$ ” (Wong 1991), and “ $n$ ” (Tang and Fishwick 1993) hidden neurons for better forecasting accuracy, where  $n$  is the number of input nodes. As it is confirmed by Mishra and Desai (2006), Rahimikhoob (2014), Moosavi et al. (2013a,b) and Shirmohammadi et al. (2013), the optimum number of neurons in hidden layer cannot be always determined using a specific formula and should be investigated by a trial-and-error method. For this study, up to 40 hidden layers are performed to obtain the best ANN architecture.

**Fig. 2** NARX network used for this study with delay of 1 to 4 months

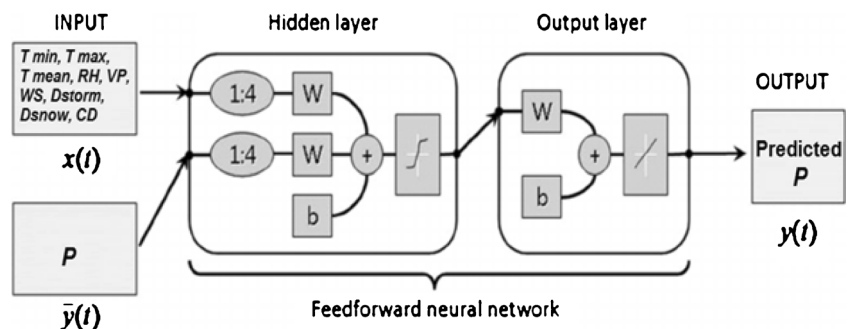


Figure 3 shows the applied feed-forward network with three hidden layers and 1-month lag time.

**Training methods**

In order to train the models, the backpropagation approach is used. The aim is to create a network that gives an optimum result. There are many variations of the backpropagation algorithm. For example, the gradient descent with momentum and adaptive learning rate backpropagation (GDX) has widely been considered as an effective backpropagation learning algorithm. Also, other backpropagation learning algorithms have been surveyed in different ANN architects. In this study, to make sure that the best selected architecture is trained by the most efficient training method, nine popular training algorithms are employed to the data and the most capable method is ascertained. The applied training algorithms are (1) Levenberg-Marquardt (LM), (2) BFGS quasi-Newton (BFG), (3) resilient backpropagation (RP), (4) scaled

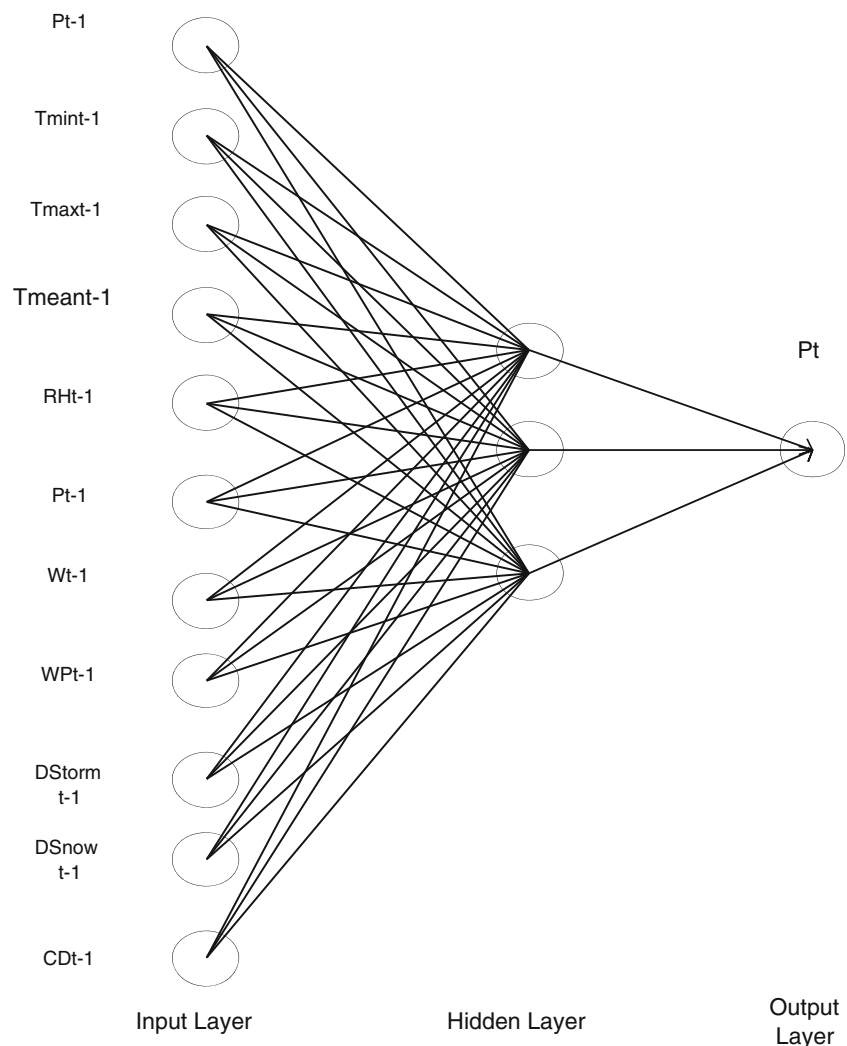
conjugate gradient (SCG), (5) conjugate gradient with Powell/Beale restarts (CGB), (6) Fletcher-Powell conjugate gradient (CGF), (7) Polak-Ribière conjugate gradient (CGP), (8) one-step secant (OSS), and (9) variable learning rate gradient descent (GDX).

**Performance comparison of models**

Correlation coefficient (R) and root mean squared error (RMSE) were used to compare the performance of models and to select the best one (Sreekanth et al. 2009).

$$R = \left( \frac{\sum_{i=1}^n (o_i - o)(e_i - e)}{\sqrt{\sum_{i=1}^n (o_i - o)^2} \sqrt{\sum_{i=1}^n (e_i - e)^2}} \right) \tag{5}$$

**Fig. 3** The feed-forward network used for the study with lag time of 1 and hidden layer size of 3



$$RMSE = \sqrt{\frac{\sum_{i=1}^n (o_i - e_i)^2}{n}} \tag{6}$$

where  $o$ ,  $e$ , and  $n$  are observed precipitation, estimated precipitation, and number of data, respectively.

## Results and discussion

### The performance without and with importing the externals

As mentioned earlier, ten climatic variables (precipitation and nine other parameters as exogenous factors of precipitation) are used as data for this study. To compare the performance of the two employed models, NAR and NARX, precipitation is predicted by applying both the models. As explained before, to run NAR, just precipitation data are required. Performance result of NAR is described in the third row of Table 2. Regarding existing nine externals and to show results of NARX more smoothly and efficiently, the model is run nine times with nine different groups of data as input of the model.

As presented in Table 2, in the first run, precipitation (as target input) and minimum temperature (as an exogenous variable of the target input) are imported to the ANN network as inputs. In the second run, in addition to precipitation and minimum temperature, maximum temperature is also employed to NARX as second exogenous. Finally, in the ninth run, all the nine externals are inserted into the model. As the information of the table shows, based on both the performance indexes ( $R^2$  and RMSE), all the nine types of the NARX model perform better than

did those of NAR in all the five time series. It should be noted that lower values of RMSE represent higher correlations. Among the NARX models, with increasing the number of imported exogenous variables, the performance is clearly and regularly rising. Table 2 also presents that the network performance is improving by increasing the size of time scales. Therefore, the networks with ten inputs and 18-month time scale are the most accurate ones.

### The most effective exogenous variable

As explained earlier, although all the selected variables affect the amounts of precipitation in a region, undoubtedly, the control of some of them on rainfall procedure is higher than of the other ones. Knowing the more effective exogenous parameters has lots of benefits. It helps meteorologists to better understand the procedure of falling rain. So, it contributes them to model rainfall more efficiently and accurately. Therefore, determining efficiency of any of the nine considered exogenous parameters in rainfall forecasting in arid and hyper-arid areas is ended in this research. The higher performance, recognized by the results of performance criteria, resulted by any of the externals in precipitation prediction shows more effectiveness on precipitation patterns. To this end, the mean correlation of any of the exogenous variables for any of the five time series is ascertained and presented in Table 3.

The table shows that in the 3-month time series, the temperature parameters (minimum, maximum, and mean) are equally the most effective exogenous variables to increase accuracy of precipitation prediction. In the case of the 6-month time scale, in addition to the temperature parameters, importing wind speed also remarkably causes better rainfall forecasts. In time series of 9 and 12, employing wind speed as an exogenous parameter results

**Table 2** Performance of the networks using different input sets for different time series

Running sum time scale	3		6		9		12		18	
Performance indexes	$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$	RMSE
Exogenous parameters										
$P$	0.19	0.43	0.20	0.34	0.18	0.28	0.30	0.22	0.27	0.22
$P, T_{min}$	0.30	0.37	0.29	0.30	0.24	0.27	0.35	0.21	0.37	0.18
$P, T_{min}, T_{max}$	0.33	0.33	0.34	0.26	0.31	0.21	0.43	0.18	0.41	0.17
$P, T_{min}, T_{max}, T_{mean}$	0.33	0.32	0.34	0.26	0.30	0.21	0.43	0.17	0.42	0.16
$P, T_{min}, T_{max}, T_{mean}, RH$	0.34	0.33	0.36	0.24	0.34	0.19	0.49	0.17	0.47	0.16
$P, T_{min}, T_{max}, T_{mean}, RH, VP$	0.36	0.33	0.42	0.24	0.39	0.19	0.56	0.16	0.53	0.14
$P, T_{min}, T_{max}, T_{mean}, RH, VP, WS$	0.42	0.30	0.50	0.22	0.51	0.17	0.66	0.14	0.62	0.13
$P, T_{min}, T_{max}, T_{mean}, RH, VP, WS, Dstorm$	0.43	0.30	0.52	0.22	0.52	0.17	0.68	0.14	0.68	0.12
$P, T_{min}, T_{max}, T_{mean}, RH, VP, WS, Dstorm, Dsnow$	0.43	0.31	0.52	0.22	0.52	0.17	0.68	0.14	0.70	0.11
$P, T_{min}, T_{max}, T_{mean}, RH, VP, WS, Dstorm, Dsnow, CD$	0.45	0.30	0.55	0.22	0.55	0.17	0.69	0.14	0.72	0.11

**Table 3** Mean  $R^2$  values for different externals for the five time series

No.	Time scale Variables	3	6	9	12	18
1	<i>P</i>	0.19	0.20	0.18	0.30	0.27
2	<i>P + T min</i>	0.30	0.29	0.24	0.35	0.37
3	<i>P + T max</i>	0.30	0.29	0.26	0.36	0.38
4	<i>P + T mean</i>	0.30	0.29	0.25	0.33	0.38
5	<i>P + RH</i>	0.27	0.28	0.27	0.38	0.33
6	<i>P + VP</i>	0.25	0.27	0.22	0.36	0.33
7	<i>P + WS</i>	0.27	0.29	0.28	0.42	0.37
8	<i>P + Dstorm</i>	0.25	0.28	0.22	0.34	0.32
9	<i>P + Dsnow</i>	0.21	0.25	0.22	0.32	0.34
10	<i>P + CD</i>	0.21	0.26	0.26	0.37	0.33

with the highest correlation between observed and simulated values. Finally, in the 18-month series, the predicted values based on maximum and mean temperatures are (equally) highly correlated with the corresponding real precipitation amounts.

Figure 4 depicts real observations and their corresponded ANN simulated values for the 12-month time series in ten situations of the imported data, which is based on the order showed in Table 3. As the plots show, considering wind speed as an external results in a relatively good match of precipitation predictions and real values. However, it can be claimed that this model is unable to exactly simulate the peaks. As the figure shows, taking some other variables (i.e., vapor pressure) into account as an exogenous can lead to better prediction of peaks. This is of importance for the studies dealing with extremes like floods or droughts and dam operation.

### The architecture of the applied networks

As mentioned earlier, the 3-, 6-, 9-, 12-, and 18-month time series of precipitation and its nine external parameters are employed to the networks with lag time varying from 1 to 12 months and hidden layer size from 1 to 40. The networks are trained by applying nine popular algorithms. The performance of any network is calculated. The information of the network with highest performance for each of the ten groups of inputs for any of the five time series is presented in Table 4. It should be mentioned that importing the variables to the networks follows the order presented in Table 3. The italic rows show the best performance among the ten groups of data in each of the time series.

As it can be clearly seen from the table, the network with all the ten input variables show the highest performance in any time series. It should be noted that in the 3-month time series,

**Fig. 4** Efficiency of each external variable in 12 months of precipitation forecasting (ordered based on the first column of Table 2). Blue and red lines present real observations and ANN predictions, respectively

the correlation of the networks with nine and ten inputs are the same at 0.89. In all the time series, the performance presented by the network with one input toward ten inputs has generally an increasing trend, even though there are some fluctuations. The table shows that 1 month is the best lag time assigned to the networks in 6, 9, 12, and 18 months. In 3 months, both 1 and 12 months are the most efficient lags. It may be because of the high temporal variability of rainfall in arid and hyper-arid lands (Asadi Zarch et al. 2015). Hence, high irregularity of precipitation mounts in these regions may cause higher efficiency of lower lag times.

About hidden layer size, all the best networks have a size between 31 and 40. Therefore, higher sizes of hidden layer significantly result in better performance of the networks. As mentioned before, nine train algorithms are used in this research. Table 4 shows that method 9 (GDX) performs by far better than the other algorithms did. In 3, 6, 9, and 12 months of time series, the best networks are trained by GDX. GDX uses backpropagation to calculate derivatives of performance the function with respect to the weight and bias variables of the network. Each variable is adjusted according to the gradient descent with momentum. Just in 18 months, the CGF method shows a higher performance. Therefore, it can be strongly concluded that GDX training algorithm is a capable method in training networks to predict precipitation in arid and hyper-arid climates.

Figure 5 compares the predicted values and real observations for the best networks of the five time series (shown in Table 4). As the figure shows, matching of the simulated and real values is clearly increasing from 3 toward the 18 months time series. The performance criteria (presented in Table 4) also show this trend quite clearly. The figure presents that in the 6 and 18 months time series, the networks are able to simulate high and low points accurately. Therefore, if a short time scale prediction is desired, the 6 months time series can be employed. However, if a longer time is preferred, 18 months of prediction would be best choice.

To better understand how hidden layer size affects the network, Fig. 6 illustrates how performance of the networks respond to changes in the number of hidden layers. To this end, all the performance values for any of 1 to 40 sizes (which changed based on lag times between 1 and 12 and nine train algorithms in situation of entering all the ten externals into the networks) are averaged and presented. The figure shows that in all the time series, there is a significant rising trend of performance, when the size of hidden layer increases. Therefore, higher number of hidden layers definitely causes higher performance. It also can be claimed that increasing the





**Table 4** Performance of the best networks and their architecture for different groups of exogenous variables in the five time series

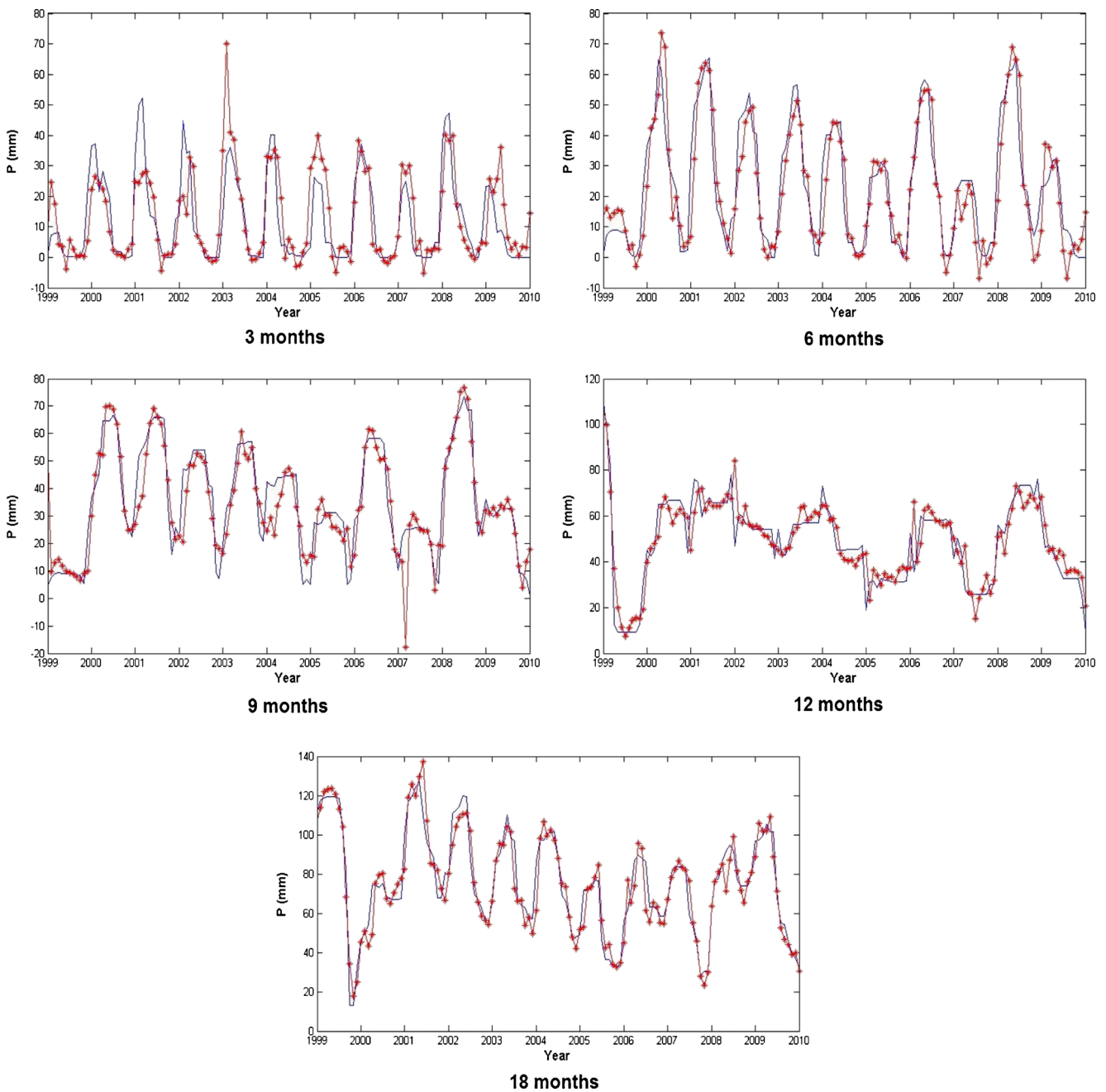
Time scale	Number of externals	Lag time (months)	Hidden layer size	Train algorithm	$R^2$	Time scale	Number of externals	Lag time (months)	Hidden layer size	Train algorithm	$R^2$
3	1	1	23	3	0.62	9	6	1	34	9	0.85
3	2	1	29	3	0.70	9	7	1	31	1	0.85
3	3	1	40	7	0.76	9	8	12	39	9	0.90
3	4	1	39	9	0.82	9	9	1	37	9	0.87
3	5	1	29	9	0.85	9	10	1	38	9	0.91
3	6	1	34	9	0.85	12	1	1	26	9	0.85
3	7	1	31	4	0.87	12	2	1	23	9	0.70
3	8	1	26	9	0.88	12	3	1	35	5	0.77
3	9	12	31	8	0.89	12	4	1	38	9	0.82
3	10	1	37	9	0.89	12	5	1	29	9	0.83
6	1	1	39	9	0.74	12	6	1	37	9	0.86
6	2	1	40	9	0.69	12	7	1	26	9	0.86
6	3	1	31	4	0.74	12	8	12	34	9	0.90
6	4	1	35	9	0.82	12	9	1	40	9	0.87
6	5	1	37	9	0.86	12	10	1	38	9	0.91
6	6	1	39	6	0.84	18	1	1	38	9	0.85
6	7	1	32	9	0.88	18	2	1	21	9	0.70
6	8	1	37	9	0.89	18	3	1	39	9	0.80
6	9	1	39	9	0.86	18	4	1	34	9	0.85
6	10	1	40	9	0.90	18	5	1	25	8	0.84
9	1	1	28	9	0.80	18	6	1	31	6	0.87
9	2	1	36	1	0.70	18	7	1	38	9	0.86
9	3	1	28	9	0.75	18	8	12	35	9	0.90
9	4	1	31	9	0.82	18	9	1	39	9	0.88
9	5	1	35	9	0.87	18	10	1	36	6	0.93

size of hidden layer beyond of 40 may result in even higher accuracy. However, it should be considered that the computation time is also increased when the number of neurons in the hidden layer rises.

The performance of different lag times in different time series for all the ten situations of input data (presented in Table 2) is illustrated in Fig. 7. It should be noted that the size of delay and feedback delay are set as the same. As mentioned previously, 1 month of lag time shows the highest accuracy in all the time series. In spite of some fluctuations in 3 and 6 months, from 1 to 12 lags, generally first, a decreasing and then an increasing trend can be seen in almost all the time series. Although performance of 1 month lag is relatively high in all situations, the performance of 12 months of lag depends on the group of data considered as input. It can be clearly seen that with increasing the number of exogenous variables to 10, the performance of 12 lag time remarkably grows.

## Conclusions

Regarding high variability of rainfall in arid and hyper-arid regions, ANNs were assumed to be a capable tool to simulate and forecast precipitation. In this paper, traditional feed-forward, multilayer perceptron (MLP) networks were used. The aim of this research was improving accuracy of ANNs in precipitation simulations in arid climates. The idea to reach this end was importing exogenous variables to the networks. Nine climatic parameters were employed as most potentially effective factors on rainfall procedure. Then, NARX and NAR (the model with and with externals, respectively) were employed. To obtain the best architecture of the network, 1 to 12 months of lag times, 1 to 40 hidden layers, and 9 train algorithms are also applied. The results showed that the NARX model performs by far better than NAR. NARX model with more externals also presented higher performances. Less lags, larger hidden



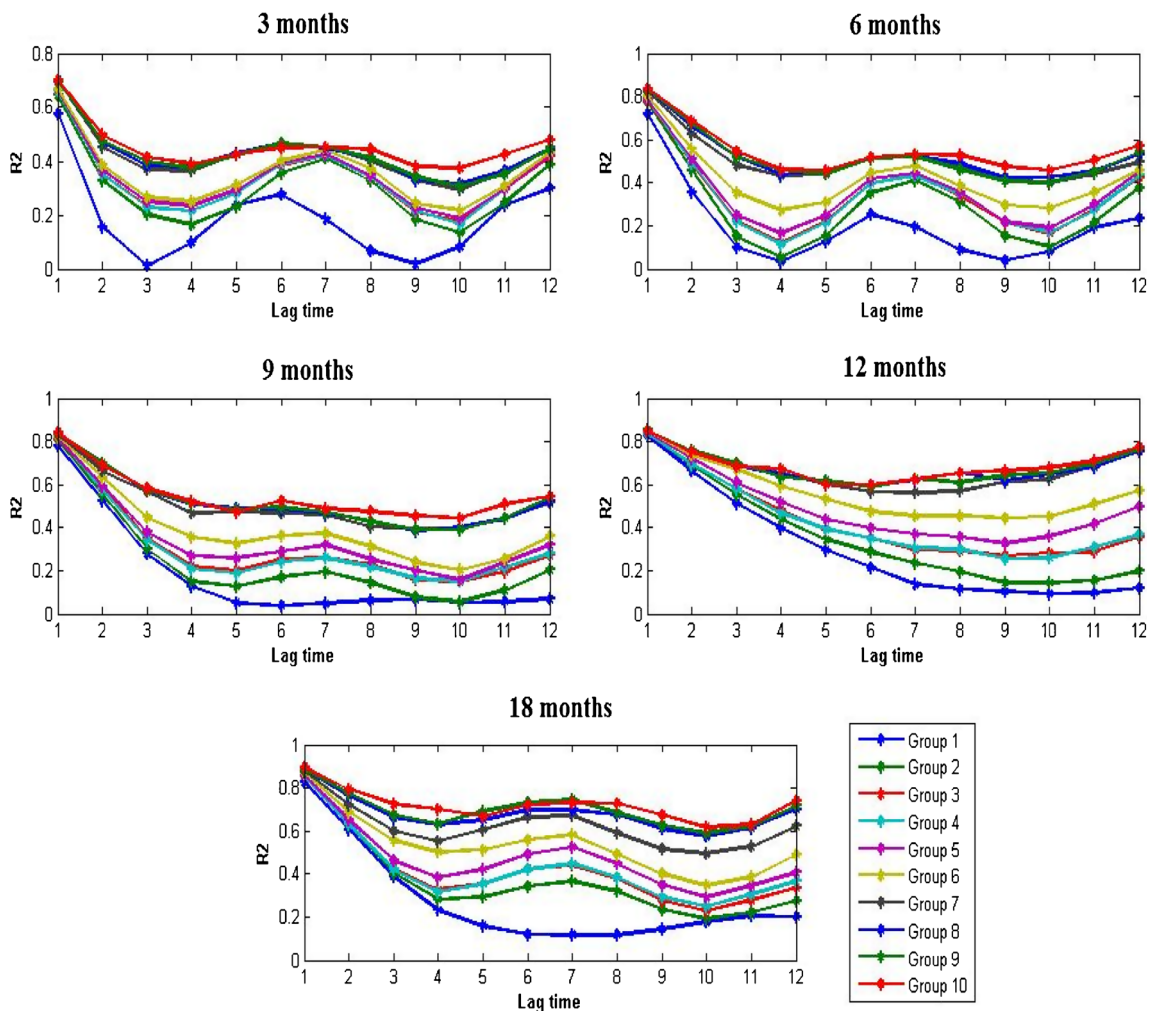
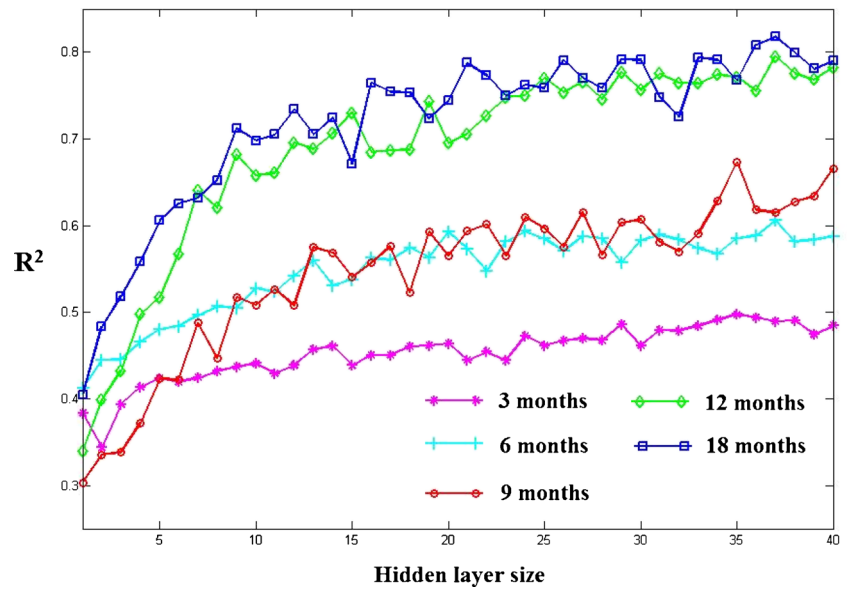
**Fig. 5** Comparison of real values of rainfall (blue lines) and predicted values (red lines) of the best network in different time series (presented in Table 4)

layer sizes, and GDX training algorithm also indicated more accurate simulations in almost all the time series.

As shown before, the ideal prediction is produced by including all these exogenous parameters in the ANN network. However, it can be claimed that determining the most effective exogenous on rainfall is significantly of importance, especially to better understand rainfall patterns in arid and hyper-arid zones. To reveal the most

effective exogenous variable in rainfall prediction, precipitation is predicted by separately adding any of the exogenous parameters (i.e.  $P$  and  $T$  min,  $P$  and  $T$  max, etc.) to the network for each time series of 3, 6, 9, 12, and 18. Then, the correlation of the predicted values and real observations were estimated. The results showed in different time scales that various externals are more effective.

**Fig. 6** Performance of different sizes of hidden layer in different time series, when all the externals are considered in the network



**Fig. 7** Performance of 1 to 12 months of lag times in 3, 6, 9, 12, and 18 months time series

**Acknowledgments** This paper was financially supported by the Payame Noor University (PNU), Iran, under grant no. 48910/7. The authors appreciate the Associate Editor and two anonymous reviewers for their constructive and helpful comments.

## References

- Adya M, Collopy F (1998) How effective are neural networks at forecasting and prediction? A review and evaluation. *J Forecasting* 17(5–6):481–495
- Asadi Zarch MA, Mobin MH, Malekinezhad H, Dastorani MT, Kousari MR (2011) Drought monitoring by reconnaissance drought index (RDI) in Iran. *Water Recour. Manage* 25: 3485–3504
- Asadi Zarch MA, Sivakumar B, Sharma A (2015) Assessment of global aridity change. *J Hydrol* 520:300–313
- Batisani N, Yarnal B (2010) Rainfall variability and trends in semi-arid Botswana: implications for climate change adaptation policy. *Appl Geogr* 30:483–489
- Birikundavyi S, Labib R, Trung HT, Rousselle J (2002) Performance of neural network in daily streamflow forecasting. *J Hydrol Eng* 7(5): 393–398
- Daliakopoulos I, Coulibalya P, Tsani IK (2005) Groundwater level forecasting using artificial neural network. *J Hydrol* 309:229–240
- Dhussa AK, Sambhi SS, Kumar S, Kumar S, Kumar S (2014) Nonlinear autoregressive exogenous modeling of a large anaerobic digester producing biogas from cattle waste. *Bioresour Technol* 170: 342–349
- Diaconescu E (2008) The use of NARX neural networks to predict chaotic time series. *WSEAS transactions on computer research*. 3(3 ):182–191
- Dierer S, Arpagaus M, Seifert A, Avgoustoglou E, Dumitrache R, Grazzini F (2009) Deficiencies in quantitative precipitation forecasts: sensitivity studies using the COSMO model. *Meteorol Z* 18: 631–645
- Fattahi M (1998) Survey of qualitative and quantitative trend in vegetation variation of Poshtkouh rangelands in the period of 1986–1988. MSc. Thesis, Natural Resources College of Tehran University (in Persian)
- Haykin S (1999) *Neural Network: A Comprehensive Foundation*. Prentice Hall
- Hect-Nielsen R (1990) *Neurocomputing*. Addison-Wesley, Reading, MA
- Horn K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. *Neural Net* 2:359–366
- Hung NQ, Babel MS, Weesakul S, Tripathi NK (2009) An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrol Earth Syst Sci* 13:1413–1425
- Huon F, Poo AN (2013) Nonlinear autoregressive network with exogenous inputs based contour error reduction in CNC machines. *Int J Mach Tool Manu* 67:45–52
- Jain A, Indurthy SKVP (2003) Comparative analysis of event based rainfall-runoff modeling techniques-deterministic, statistical and artificial neural networks. *J Hydrol Eng* 8:93–98
- Jain A, Srinivasulu S (2004) Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms and artificial neural network techniques. *Water Resour Res* 40(4): doi:10.1029/ 2003 WR002355
- Kisi O, Cimen M (2011) A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *J Hydrol* 399:132–140
- Lazaro R, Rodrigo FS, Gutierrez L, Domingo F, Puigdefabregas J (2001) Analysis of a 30-year rainfall record (1967–1997) in semi-arid SE Spain for implications on vegetation. *J Arid Environ* 48:373–395
- Luk KC, Ball JE, Sharma A (2001) An application of artificial neural networks for rainfall forecasting. *Math Comput Model* 33:883–699
- Mar KW, Naing TT (2008) Optimum neural network architecture for precipitation prediction of Myanmar. *World Acad Sci Eng Technol* 2:12–20
- MathWorks (2015) Neural network time series prediction and modelling. Retrieved from <http://www.mathworks.com/help/nnet/gs/neural-network-time-series-prediction-and-modeling.html>.
- McCullagh J, Bluff K, Ebert E (1995) A Neural Network Model for Rainfall Estimation. *Second New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems*. Pages 389–392
- Mishra AK, Desai VR (2006) Drought forecasting using feed forward recursive neural network. *Ecol Model* 198:127–138
- Mohammadi K, Eslami HR, Dayyani Dardashti S (2005) Comparison of regression, ARIMA and ANN models for reservoir in flow forecasting using snow melt equivalent (acase study of Karaj). *J Agric Sci Technol* 7:17–30
- Moosavi V, Vafakhah M, Shirmohammadi B, Behnia N (2013a) A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods. *Water Resour Manag* 27: 1301–1321
- Moosavi V, Vafakhah M, Shirmohammadi B, Ranjbar M (2013b) Optimization of wavelet-ANFIS and wavelet-ANN hybrid models by Taguchi method for groundwater level forecasting. *Arab J Sci Eng* 39:1785–1796
- Moustris KP, Larissi IK, Nastos PT, Paliatsos AG (2011) Precipitation forecast using artificial neural networks in specific regions of Greece. *Water Resour Manag* 25:1979–1993
- Ni J, Zhang XS (2000) Climate variability, ecological gradient and the Northeast China transect (NECT). *J Arid Environ* 46:313–325
- Rahimikhoob A (2014) Estimating sunshine duration from other climatic data by artificial neural network for ET0 estimation in an arid environment. *Theor Appl Climatol* 118(1–2):1–8
- Safavieh E, Andalib S, Andalib A (2007) Forecasting the unknown dynamics in NN3 database using a nonlinear autoregressive recurrent neural network. *International Joint Conference on Neural Networks*: 2105–2109
- Shirmohammadi B, Moradi HR, Moosavi V, Taie Semiromi M, Zeinali A (2013) Forecasting of meteorological drought using wavelet- ANFIS hybrid model for different time steps (case study: southeastern part of East Azerbaijan province, Iran. *Nat Hazards* 69:389–402
- Sreekanth P, Geethanjali DN, Sreedevi PD, Ahmed S, Kumar NR, Jayanthi PDK (2009) Forecasting groundwater level using artificial neural networks. *Curr Sci* 96:933–939
- Srivastava G, Panda SN, Mondal P, Liu J (2010) Forecasting of rainfall using ocean-atmospheric indices with a fuzzy neural technique. *J Hydrol* 395:190–198
- Tang Z, Fishwick PA (1993) Feed forward neural nets as models for time series forecasting. *ORSA. Journal on. Computing* 5(4):374–385

- Wei H, Li JL, Liang TG (2005) Study on the estimation of precipitation resources for rainwater harvesting agriculture in semi-arid land of China. *Agric Water Manag* 71:33–45
- Wong FS (1991) A 3D Neural Network For Business Forecasting. The 24th Annual Hawaii International Conference on System Science. Pages 113–123
- Yu PS, Chen ST, Chang IF (2006) Support vector regression for real-time flood stage forecasting. *J Hydrol* 328:704–716
- Zhang B, Govindaraju RS (2000) Prediction of watershed runoff using Bayesian concepts of modular neural networks. *Water Resour Res* 36(3):753–762