American Journal of Innovative Research and Applied Sciences. ISSN 2429-5396 I www.american-jiras.com

OGRIGINAL ARTICLE

AN ARTIFICIAL NEURAL NETWORK MODEL FOR MONTHLY PRECIPITATION FORECASTING INHOMS STATION, SYRIA

| Ghatfan Abdalkareem Ammar [1] | Badia, Youcef Haidar [2*] | and | Amer Qousai Al Darwish [3] |

[1.] Tishreen University | department of water engineering and irrigation | Lattakia | Syria |
[2.] Tishreen University | department of structuralengineering | Lattakia | Syria |
[3.] Tishreen University | department of water engineering and irrigation | Lattakia | Syria |

|Received | 16 May 2016| |Accepted | 23 May 2017| |Published 30 May 2017 |

ABSTRACT

Background: One of the major problems in water resources management is the precipitation forecasting. Given the effect of precipitation on water resources, it is found that a more accurate prediction of precipitation would enable more efficient utilization of water resources and flood management. On the other hand, climate and precipitation are highly non-linear and complicated phenomena, which require non-linear mathematical modeling and simulation for accurate prediction. **Objectives**: The objective of this study is to predict the precipitation for one month ahead using feed forward neural network FFNN model with the Back Propagation algorithm, with the use of a combination of meteorological parameters (evaporation, air temperature, relative humidity and monthly index) from year 1976 to 2009 for Homs Meteorological Station, Syria. **Methods**: A feed forward neural network FFNN model was applied to predict the precipitation on a monthly basis. The models were trained based on the Levenberg-Marquardt algorithm. They investigated the effect of the number of hidden neurons and activation function on ANN modeling. The effect of climatic parameters was also examined by training three separate ANN models with three different input data sets. **Results**: Evaluation of the model performance using the last four years showed that there was a good agreement between the ANN estimations and the measured precipitation values. Comparison of correlation coefficient (R) and Root Mean Square Errors (RMSE) showed that the neural Network architecture (4-35-1) is of the best performance with logsigmoid activation function for the hidden layer and tansigmoid activation function for the output layer. Sensitivity analysis performed to determine the most important parameter indicated that the most important input parameter in forecasting precipitation is relative humidity, especially when it is accompanied by air temperature. **Conclusions**: The study reveals that FFNN model with Levenberg-Marquardt Back Propagation algorithm can be used as an appropriate forecasting tool in predicting the monthly precipitation in Homs Meteorological Station.
Keywords: *Precipitation, Prediction, Artificial neural network, Back Propagation algorithm.*

1. INTRODUCTION

Weather forecasting (and precipitation prediction in particular) is one of the most important imperatives, necessities and demanding tasks to be performed, as well as a challenge posed by meteorological services all around the world. It is a complicated procedure, which involves several specialized fields of knowledge [1]. Precipitation prediction is very important for countries whose economy depends mainly on agriculture, like many of the third World countries [2].

A large number of attempts have been made on a global scale by various researchers with the aim of predicting the amounts of rainfall using different techniques [3]. However, because of the nonlinear nature of precipitation and the level of accuracy of prediction that has been obtained and which is still below satisfaction in the light of the development taking place, neural networks have become at present an inductive method attracting researchers due to their high understanding of the nonlinear phenomena and their ability to establish the complex relationship between influential input and output [4].

Shafie et al. (2001) compared two models for predicting yearly and monthly rainfall (January, December) at Alexandria, Egypt, and the results of the network testing of the recent 10 years demonstrated the preeminence of the feed forward Back Propagation neural networks model over the Multi Linear Regression model [5]. Bustami et al. (2007) used Artificial Neural Network models along with the Back Propagation algorithm in order to predict the amount of rainfall and the level of Bedup River at Sarawak state, Malaysia [6]. Hung et al. (2008) forecasted the falling of rain for 1-3 hours in Bangkok using feed forward Neural Networks, relying on hourly data for 4 years [7]. Vamsidhar et al. (2010) put a Back Propagation Neural Network model to forecast the monthly rainfall in India relying on humidity, dew point and pressure as the inputs of the model [8]. Naik and Pathan (2012) proposed a new method for weather forecasting using feed forward ANN with Levenberg Back Propagation Algorithm for training. The results showed that FFNN is appropriate for weather forecasting [9]. Moustris et al (2011) considered the possibility of forecasting weather on the long term (4 straight months) through applying Neural Networks relying on a long chain of monthly rainfall from 4 weather forecasting stations in Greece [10]. Alhashimi (2014) put three different models for forecasting monthly rainfall. Her study included monthly measurements of precipitation, mean temperature, wind speed and relative humidity from year 1970 to 2008. The results showed the preeminence of the ANN models over other models [11].

2. MATERIALS and METHODS

2.1 The Study Region and Data

Homs is a middle city of Syria, 162 km to the north of the capital Damascus, lying on latitude 34° 43' 45" N., and longitude 36° 43' 30" E. The data needed were collected from Homs Forecasting Station, 487 m above the sea level.

For the purpose of forecasting rainfall in Homs, monthly rainfall data were used in addition to evaporation and dry air temperature and relative average humidity of 34 years starting on 1st. of January 1976 to 12th. of December 2009. The data of the first 30 years (90 % of the total data) were used to train and calibrate the model, and the data of the last 4 years (10 % of the total data) to test and evaluate the performance of the models.

Figure 1 shows the average monthly precipitation in Homs between the years 1976 and 2009, where the average annual rainfall is about 425.3 mm, and the highest average of monthly rainfall is 85.6 mm observed in January, and the lowest average of monthly rainfall is about 0 mm observed in July and August.

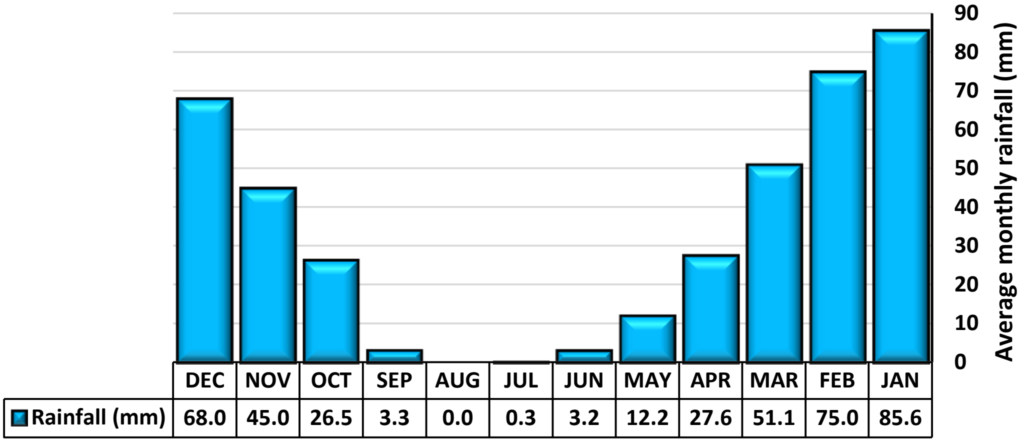

Figure1: Average monthly precipitation in Homs.

2.2 Data Preprocessing and Evaluation of Model Performances: The input and the output data obtained have to be normalized because they are of different units, otherwise there will be no correlation between the input and the output values. We normalize the collected data in the range [0, 1]. Normalization is done using the following Eq. (1) [12]:

$$P_{norm} = 0.5 * \left(\frac{P_{obs} - \bar{P}_{obs}}{P_{max} - P_{min}}\right) + 0.5 \qquad (1)$$

Where P_{norm} is the normalized value, P_{obs} is the original value, $\bar{P}_{obs}$ is the average value, P_{max} is the maximum value, P_{min} is the minimum value.

After obtaining the normalized data, the next step is to train the input data using Matlab Back-propagation Algorithm. The proposed ANN model is basically a three layered ANN Back Propagation learning. To evaluate the generalization capability of all models, the data has been divided into three data sets: training set (312 months data), validation set (48 months data) and testing set (48 months data). Training set is used in the training phase of neural network, while validation set is used to validate the neural network performance during the training. Testing set is used to test the performance of neural network after the training has been completed. Researchers have used different data division among training, validation and testing data sets, and it generally varies according to problem type. There is no certain rule for data division. In this study, adopted data division among training, validation, and testing sets is determined as (10%, 10%, 80%), respectively. This data division gave the best results for the different performance measures (e.g., RMSE, R).

The performance of various models during calibration (training set, validation set) and test were evaluated by using the statistical indices:

- Root Mean Squared Error (RMSE) is defined as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_{obs} - P_{pre})^2}{n}} \quad (2)$$

- Correlation Coefficient: R defined as

$$R = \frac{\sum_{i=1}^{n}(P_{obs} - \bar{P}_{obs})(P_{pre} - \bar{P}_{pre})}{\sqrt{\sum_{i=1}^{n}(P_{obs} - \bar{P}_{obs})^2 * \sum_{i=1}^{n}(P_{pre} - \bar{P}_{pre})^2}} \quad (3)$$

Where n is the number of training or testing samples, P_{obs} is the observed precipitation, P_{pre} is the simulated value of precipitation, $\bar{P}_{obs}$ and $\bar{P}_{pre}$ is the average value of the observed precipitation and the simulated precipitation.

2.3 Artificial Neural Network: Neural Network is an algorithm that dynamically inherits human neuron information processing capability [13]. This capability enables Neural Network to perform a brain-like function such as forecasting, classification, and pattern matching. The neural network model is used to predict the consequence of a given action. Neural network can be categorized into single and multi-layer networks. Single layer network is a model that consists of input and output layers, whereas a multi-layer network consists of at least one hidden layer between the input and the output layers.

Figure 2 shows a simple neural network model. The input layer represents the action which is fed into the next layer until the output layer.

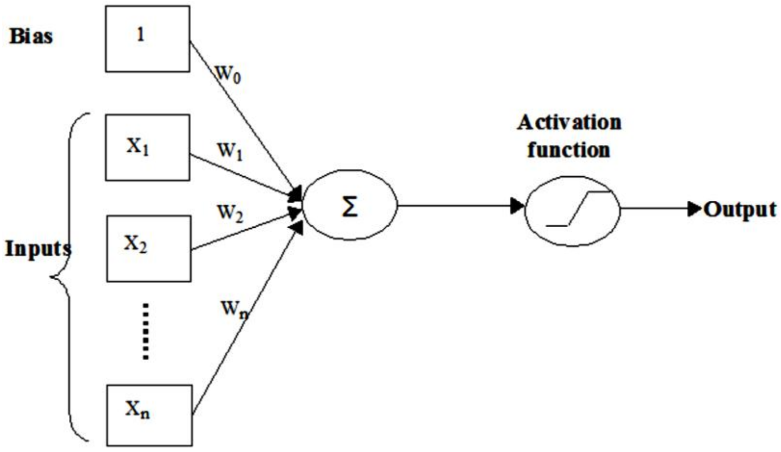

Figure 2: The figure presents the simple Neural Network Model.

In order to perform weather forecasting using neural network and compare the performance of neural network models with different learning rates and by setting different number of neurons in hidden layer, we are using Back Propagation algorithm and supervised learning [14] . The following Steps are followed in the process:

1. Data preprocessing
2. Defining the ANN model
3. Training of ANN
4. Testing of data

To obtain the optimal weights (parameters) of the neural network structure, Levenberg–Marquardt Back Propagation algorithm was used to train the network. This algorithm is a second-order non-linear optimization technique that is usually faster and more reliable than any other Back Propagation techniques [15]. The neural network model is trained to minimize the error between the actual and predicted output. Back Propagation requires that the activation function used by the artificial neurons be differentiable. Different activation functions for hidden and output layers were used to find the best ANN structure for this study. We use two activation functions for the hidden layer that are either tansigmoid or logsigmoid, and one activation function for the output layer that is tansigmoid one.

Researches about the neural networks refer to the nonexistence of a standard style for the determination of the number of hidden layers or the number of neurons. Instead, this number is determined according to the view of the designer of the model [16]. In this study, one hidden layer is adopted, and the number of neurons in it was determined by trial and error procedure.The number of hidden neurons was determined by trial and error procedure. The output neuron will be the original value of one step ahead.

3. RESULTS

Table 1 demonstrates the effect of climatic factors all together (evaporation, air temperature, relative humidity) along with monthly index on a one-month-ahead rainfall forecast by using different activation functions and gradually increasing the number of neurons inside the hidden layer on the performance of the network. It is observed that the use of an activation function of the type (logsigmoid) inside the hidden layer with 35 neurons followed with an activation function of the type (tansigmoid) in the output layer gave the best performance of the neural network. It gave the lowest value of the adopted standard (RMSE) during the training and testing of the network in repetitious constant-in-number turns (1000 epochs) and initial constant weights. Therefore, the structure of the artificial neural network adopted in this study is (4- 35- 1) i.e. 4 neurons inside the input layer, 35 neurons inside the hidden layer, and one neuron inside the output layer. The results show good accuracy in a one-month-ahead rainfall forecast during the four-recent-year test period, with the values of R and RMSE during this period were 0.994 and 3.684mm, respectively.

Table1: The table presents the statistical performance evaluation criteria for each ANN model.

Model	$R_{t+1} = \mathrm{f}\,(\mathrm{Evap} + \mathrm{Temp} + \mathrm{Hum})\mathrm{t}$						
Activation Function of Hidden	**Activation Function of Output**	**Number of Neurons**	**Statistical Indicators**	**All Data (408 month)**	**Training Set (312 month)**	**Validation Set (48 month)**	**Testing Set (48month)**
tansig	tansig	15	R	0.928	0.914	0.980	0.842
			RMSE (mm/month)	15.429	15.507	12.750	17.264
logsig	tansig	20	R	0.945	0.927	0.986	0.944
			RMSE (mm/month)	13.596	14.278	10.758	11.415
logsig	tansig	25	R	0.966	0.958	0.988	0.952
			RMSE (mm/month)	10.948	11.233	9.820	10.105
logsig	tansig	30	R	0.983	0.981	0.990	0.980
			RMSE (mm/month)	7.725	7.654	8.977	6.772
logsig	tansig	35	R	**0.991**	**0.992**	**0.989**	**0.994**
			RMSE (mm/month)	**5.628**	**4.984**	**9.706**	**3.684**
where R_{t+1} represents a next time-step monthly rainfall of day t.							

Table 2 shows the effect of the absence of any of the network's inputs (evaporation, air temperature, relative humidity) on its performance, while maintaining the monthly index as a fundamental input. This analysis was based on two stages. The first stage involved testing the performance of the artificial neural network in the presence of a single variable and the absence of other variables. It was noted that the most influential element was the relative humidity, with a total correlation coefficient of 0.894, then came the elements of average dry temperature and evaporation with correlation coefficient of 0.865 and 0.847, respectively. The second stage involved testing the accompaniment of each climatic variable with the other variables, which elucidated that the accompaniment of relative humidity with temperature gives the best performance of the proposed neural network. This goes along with the conclusion of the first stage.

Table 2: sensitivity analysis of input variables of the proposed neural network.

The first stage: the input of the network one variable					The second stage: the input of the network tow variable				
model	$R_{t+1} = \mathrm{f}\,(\mathrm{Evap})\mathrm{t}$				model	$R_{t+1} = \mathrm{f}\,(\mathrm{Evap} + \mathrm{Temp})\mathrm{t}$			
Statistical indicators	Training set	Validation set	Testing set	All data	Statistical indicators	Training set	Validation set	Testing set	All data
R	0.858	0.870	0.862	0.847	R	0.968	0.978	0.938	0.968
RMSE (mm/month)	19.730	35.913	17.413	22.025	RMSE (mm/month)	9.694	13.864	11.396	10.476
model	$R_{t+1} = \mathrm{f}\,(\mathrm{Hum})\mathrm{t}$				model	$R_{t+1} = \mathrm{f}\,(\mathrm{Evap} + \mathrm{Hum})\mathrm{t}$			
R	**0.874**	**0.954**	**0.837**	**0.894**	R	0.958	0.989	0.943	0.966
RMSE (mm/month)	**18.567**	**19.324**	**17.642**	**18.552**	RMSE (mm/month)	11.023	9.706	11.017	10.876
model	$R_{t+1} = \mathrm{f}\,(\mathrm{Temp})\mathrm{t}$				model	$R_{t+1} = \mathrm{f}\,(\mathrm{Temp} + \mathrm{Hum})\mathrm{t}$			
R	0.848	0.898	0.927	0.865	R	**0.960**	**0.972**	**0.961**	**0.963**
RMSE (mm/month)	20.235	28.964	12.629	20.750	RMSE (mm/month)	**10.776**	**15.458**	**9.367**	**11.280**

Figure 3 shows linear compatibility by analyzing the correlation between the outputs of the model and the actual values in order to evaluate the performance of the neural network (4-35-1) of the three sets (validation, training and testing) as well as the 'all data' set. It should be known that the validation and testing sets are samples that were not recognized by the network during training stage, and that performing the validation and testing processes was to ensure the compatible performance and to prevent over training of the proposed network, with the aim to reach the weights which correlate the inputs and the outputs of the network in the optimum way. The scatter plot of the values shows the accuracy of the model since they converge the line 45 of the three sets and the set of all data.
Figure 4 illustrates the ability of the model ANN (4-35-1) to forecast one-month-ahead rainfall with high accuracy, with R= 0.994 and RMSE= 3.68 mm during the last 4 years of the study.

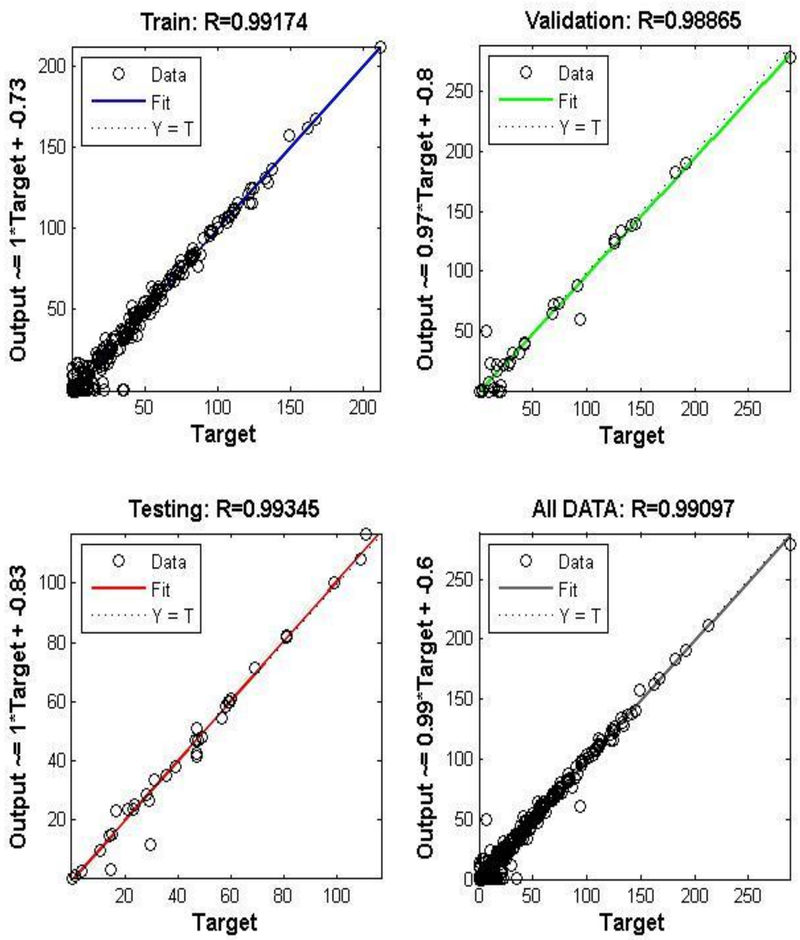

Figure3: The figure presents the scatterplots comparing observed and forecasted precipitations using ANN (4-35-1).

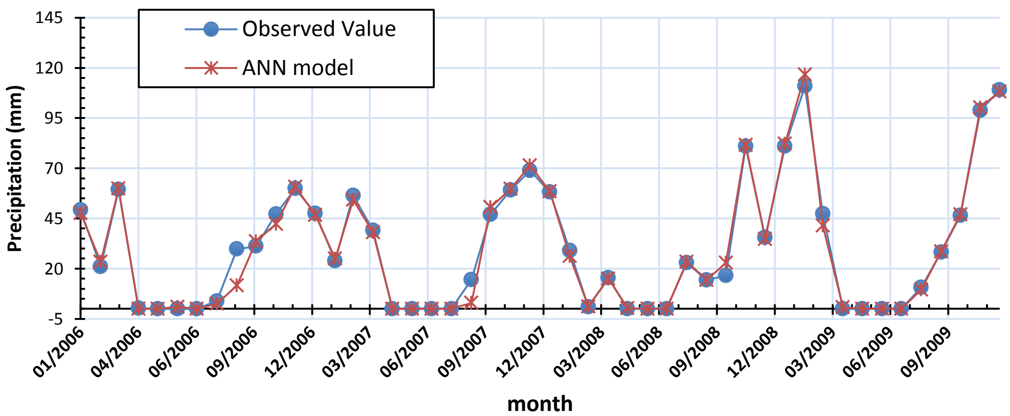

Figure 4: The figure presents the comparison between simulated and actual precipitation using ANN (4-35-1).

4. DISCUSSION

This study used the model FFNN with Levenberg-Marquardt algorithm to forecast one-month-ahead rainfall utilizing monthly data taken from Homs Forecasting Station for a 34-year period but leaving the last 4 years as a model-testing set. These data involved evaporation, air temperature and relative humidity, with the monthly index as a fundamental input of the model. It turned out that the (4-35-1) model, using a logsigmoid activation function in the hidden layer and a tansigmoid activation function in the output layer, presents the best performance evaluating standards used in the study, with R= 0.994 and RMSE= 3.684 mm for the testing set, which is considered a satisfactory result if compared with other researches similar to this study. For instance, Vamsidhar et al. (2010) developa Back Propagation Neural Network (BPNNs) model with a (1-7-3) structure to forecast monthly rainfall in India. The input data involved humidity, dew point and pressure. He got an acceptable degree of accuracy (Accuracy = 100-MSE) with a value of 94.28 % for the testing set [8]. In a different study, Alhashimi (2014) used monthly measurements of rainfall, wind speed, average temperature and relative humidity, and she found that the best feed forward performance ANNs with a proposed (1-8-4) structure for the testing set was of the values R=0.91 and RMSE = 27.278mm [11].

The results of sensitivity analysis demonstrate that the most influential climate variable, among other variables, on rainfall is average relative humidity, and that it becomes more and more influential when it is accompanied with average dry air temperature. This goes along with what has been achieved by previous studies; for instance, Hung et al. (2008) found, through sensitivity analysis, that the most important input, in addition to rainfall values, was wet air temperature [7].

5. CONCLUSION

Predictions of rainfall are affected by variations of geographical and regional characteristics of each region. The models presented in the literatures adapt the characteristics of a certain area and thus cannot be applied directly to a different area. Therefore, several FFNN models are used in Homs Meteorological Station in order to forecast the amount of rainfall for one month ahead, and the data used involve monthly measurements of evaporation, air temperature, and relative humidity with a monthly index. It turned out, from the testing set (the last 4 years), that the FFNN model with an LM Back Propagation algorithm could be a pre-eminent tool in rainfall forecasting. It was found that the ANN (4-5-1) model presents the best performance evaluating standards. In addition, the effect of the absence of one of the model inputs on its performance was investigated, and it was found that the most important variable was the relative humidity, and that its effect was greater when accompanied with air temperature.

If more studies are done utilizing more data such as air pressure and wind speed from other adjacent stations after performing the suitable homogeneity tests, this conclusion may be maintained, and the accuracy of the model may be raised especially in the case, for instance, where other types of neural networks or another combination of both neural networks and fuzzy logic are used.

6. REFERENCES

[1]. Lynch P. The origins of computer weather prediction and climate modeling. Journal of Computational Physics 227, 2008, 3431–3444, Available on:http://www.dca.iag.usp.br/www/material/adwgandu/DESATIVADOS/agm5801_2010/ARTIGOS/HISTORIA/Lynch_ JCP_2007.pdf.

[2]. Abraham A, Philip N. S, Joseph K. B.Will we have a wet summer? Long-term rain forecasting using soft computing. 2001, 1044–1048, Available on: https://www.researchgate.net/publication/2330379

[3]. Bushara N O, Abraham A. Computational Intelligence in Weather Forecasting: A Review. Journal of Network and Innovative Computing ,2013, 320-331, Available on: http://www.softcomputing.net/jnic1.pdf

[4]. Petrovic I, Baotic P, Peric N. Model structure selection for nonlinear system identification using feed forward neural networks. Department of Control and Computer Engineering in Automation, Croatia, 2000, 53-57, Available on: http://ieeexplore.ieee.org/document/857813/

[5]. El-Shafie, A. H, El-Mazogh H. G, Shehata A, Taha, M. R. Artificial neural network technique for precipitation forecasting applied to Alexandria, Egypt, International Journal of the Physical Sciences, 2001, 6 (6), 1306-1316, Available on:http://www.academicjournals.org/journal/IJPS/article-full-text-pdf/A8043DC28734.

[6]. Bustami R, Bessaih N, Bong C, Suhaili S. Artificial Neural Network for Precipitation and Water Level Predictions of Bedup River. International Journal of Computer Science, 2007, 34 (2), Available on: http://www.iaeng.org/IJCS/issues_v34/issue_2/IJCS_34_2_10.pdf

[7]. Hung N. Q, Babel M. S, Weesakul S, Tripathi N. K. An articial neural network model for precipitation forecasting in Bangkok, Thailand. Hydrology and Earth System Sciences Discussions, 2008, 183-218, Available on: https://hal.archives-ouvertes.fr/hal-00298924/document.

[8]. Vamsidhar E, Varma K. V. S. R. P, Rao P. S, Satapati R. Prediction of Precipitation Using Backpropagation Neural Network Model. International Journal on Computer Science and Engineering, 2010, 2 (4), 1119-1121, Available on: https://pdfs.semanticscholar.org/ac71/e1c890b686592ed14994e2d60e72c1092844.pdf .

[9]. Moustris K.P, Larissi I.K, Nastos, P.T, Paliatsos, A. G. Precipitation forecast using artificial neural networks in specific regions of Greece. Water Resour Manag, 2011, 1979–1993, Available on: http://www.academia.edu/5679548/Precipitation_Forecast _Using_Artificial_Neural_Networks_in_Specific_Regions_of_Greece .

[10]. Naik A.R, Pathan S.K. Weather Classification and forecasting using Back Propagation Feed-forward Neural Network. International Journal of Scientific and Research Publications, 2 (12), 2012, 2250-3153, Available on: http://www.ijsrp. org/research-paper-1212/ijsrp-p1211.pdf

[11]. Alhashimi S. A. Prediction ofMonthly Precipitationin Kirkuk Using Artificial Neural Network and Time Series Models. Journal of Engineering and Development, 2014, 18 (1), 129-143, Available on: http://www.iasj.net /iasj?func =fulltext&aId=85274 .

[12]. Solgi A, Nourani V, Pourhaghi A. Forecasting daily precipitation using hybrid model of wavelet-artificial neural network and comparison with adaptive neurofuzzy inference system (case study: verayneh station, nahavand). Hindawi Publishing Corporation .Advances in Civil Engineering, 2014, Available on:http://downloads.hindawi.com/journals/ ace/2014/279368.pdf

[13]. Graupe D. Principles of Artificial Neural Networks. Singapore: World Scientific Publishing, 1997, Available on: http://www.worldscientific.com/doi/pdf/10.1142/9789812770578_fmatter
[14]. Zurada M.J. Introduction to Artificial Neural Network. Jaico Publishing House, 1997, P764, Available on:https://anuradhasrinivas.files.wordpress.com/2013/08/29721562-zurada-introduction-to-artificial-neural-systems-wpc-1992.pdf
[15]. Hagan M.T,Menhaj M.B. Training feed forward networks with Marquardt algorithm. 1994, 5(6), 989–993. Available on: http://134.208.26.59/AdvancedNA/ConjugateGradientMethod/Marquardt%20algorithm%20for%20MLP.pdf
[16]. LopezG, RubioM. A, MartinezM, BahlesF. J.Estimation of Hourly Global Photosynthetically Active RadiationUsing Artificial Neural Network Models." Agricultural and ForestMeteorology 107, 2001, 279-291.https://www.researchgate.net/profile/Gabriel_Lopez12/publication/223673797_Estimation_of_hourly_global_photosynthetically_active_radiation_using_artificial_neural_network_models/links/00b7d53bd781b665a7000000.pdf .

Cite this article: Ammar, G.A., Haidar, B.Y. and Al Darwish, A.Q. An Artificial Neural Network Model for Monthly Precipitation Forecasting in Homs Station, Syria. *American Journal of Innovative Research and Applied Sciences*. 2017; 4(6): 240-246.