Monthly Precipitation Prediction by Artificial Neural Networks (Case study: Mashhad synoptic station)

Prévision de précipitations mensuelles à l'aide de réseaux de neurones artificiels (étude de cas de la station synoptique de Mashhad)

Saeed R. Khodashenas[1], N. Khalili[2], K. Davari[3]

[1] Assistant Professor, Ferdowsi University of Mashhad, Mashhad, Iran
saeedkhodashenas@yahoo.fr
[2] Ph.D. student, Ferdowsi University of Mashhad,
[3] Associated Professor, Ferdowsi University of Mashhad

RÉSUMÉ

Plusieurs modèles de réseaux de neurones artificiels (ANN) ont été développés pour la prévision des données de précipitations mensuelles en station synoptique de Mashhad. Sur les 636 données de précipitations mensuelles existantes (de 1958 a 2008), 580 données ont été utilisées pour le calage des réseaux neuronaux et les autres données, sélectionnées aléatoirement, ont été utilisées pour la validation des modèles. Pour extraire les précipitations dynamiques de cette station, le modèle ANN utilise une nouvelle approche avec un réseau composé de trois couches d'apprentissage itératif avec un algorithme retour de propagation. La sensibilité de l'exactitude de prédiction du contenu et la longueur de la couche d'entrée ont été étudiées. Basé sur les paramètres les plus appropriés, les deux structures M_{531} et M_{741} ont été sélectionnées. Les propriétés statistiques ont été calculées pour préciser les performances des modèles. Dans le meilleur modèle de prédiction mensuelle, le coefficient de corrélation (R), l'erreur quadratique moyenne (RMSE) et l'erreur moyenne absolue (EMA) sont respectivement 0.93, 0.99 mm et 6.02 mm,

ABSTRACT

Several ANN models were developed to prediction of monthly precipitation data in Mashhad synoptic station. From the total 636 monthly precipitation data (from 1958 to 2008), 580 data has been used for training networks and the rest selected randomly has been used for validation of the models. To extract the precipitation dynamic of this station by ANN, a new approach of three-layer feed-forward perceptron network with back propagation algorithm was used. The sensitivity of the prediction accuracy to the content and length of input layer was investigated. Based on the most suitable parameters, two structures M_{531} and M_{741} have been selected. Statistical properties were calculated to examine the performance of the models and it was found that in the best model of monthly prediction, the correlation coefficient (R), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), 0.93, 0.99 mm, 6.02 mm, respectively.

KEYWORDS

Artificial neural networks, precipitation, prediction

1   INTRODUCTION

The precipitation as meteorological parameters is very complex nonlinear phenomena and varies along with time and place. Nevertheless, literatures show that precipitation is predictable.

In the last years, many models of ANNs have been developed for precipitation prediction. Luk (2001) predicted precipitation in catchment's upper Parramatta River in Australia using Multi Layer Feed-forward Neural Network (MLFN). Their results showed that MLFN has more accuracy in precipitation modeling in comparison to Time Delay Neural Network (TDNN) and Recurrent Neural Network (RNN) . While TDNN anticipated to RNN and MLFN through precipitation prediction using large scale continental signals in west of Iran. Ramirez et al. (2005) also used a Multi Layer Feed-forward Perceptron (MLFP) neural network for daily precipitation prediction in the region of Sao Paolo State, Brazil. In their research, potential temperature, vertical component of the wind, specific humidity, air temperature, perceptible water, relative vorticity and moisture divergence flux are used as input data for training of networks. Results of ANN were superior to the ones obtained by the linear regression model, which is revealing a great potential for suitable performance . In a similar study, Saplioglu et al (2010) employed a three layer feed-forward neural network for daily precipitation prediction in the meteorological stations of Burdur, Egirdir, and Isparta cities in Turkey. The results indicate ANN models are superior to the commonly used weighted average and harmonic average methods in this study. Generally, there isn't any fixed ANN model as a suitable network for all of the problems, instead any ANN model using different criterion should be tested until one obtain the appropriate model for desirable purpose (Khalili, 2008). However, according to the literature, it seems that for precipitation prediction, Multi Layer Feed-forward Perceptron (MLFP) has more reasonable outputs in comparison to other ANN types.

In this paper, ANNs have been used to obtain a prediction model for the monthly precipitation of Mashhad's synoptic station in Iran. This is due to the intelligent capability of the neural networks in the extraction of the features of the systems, even in the cases that there is not much information about the system dynamics. In these cases, it will suppose to use a black box model for the system and capture system's dynamics through its memory. Having more information about the affective factors on the system, we can use a gray box model and utilize the additional information in our modeling. However, Hurst's Rescaled range statistical (R/S) analysis shows that the collected data in Mashhad station is predictable using only the past information of system and only a black box modeling of the system could capture its dynamic.

2- MATERIALS AND METHODS

Monthly precipitation data has collected from Mashhad's synoptic weather station (located in the north east of Iran. Using the collected data for the mentioned case, some tests such as Mackus and Run test were carried out for determination of data sufficiency and homogenous. In addition, Hurst's Rescaled range statistical (R/S) analysis test was used for assessment of data predictability. Indeed, this statistical index captures the existence of memory effect in the given data. Finally, precipitation was predicted using MLFP that is a suitable type of ANNs for meteorological predictions

2-1-Data and Location of the study

In Mashhad's special geographical situation, interfacing different weather fronts make it such a region with a special different continental climate. Overall, its climate is as mid dry and cold with dry-hot summers and wet-cold winters. The maximum annual temperature is about +35 and the minimum is about -15. The annual average precipitation is about 253 mm in Mashhad. The synoptic station is located at $36^o16'$ Northern longitude, $59^o38'$ Eastern latitude and 999.2 meter elevation.

The main objective of the research is to develop an artificial neural network for the purpose of monthly precipitation prediction. To achieve this objective, we used the monthly precipitation data from 1951 to 2003. All data arranged as a set of time series. Table (1) shows the minimum, maximum, mean and coefficients of variations for monthly precipitation data.

statistical characteristic	Value
Coefficient of variations of data	58(%)
Average of data	21.24 (mm)
Maximum of data	132 (mm)
Minimum of data	0 (mm)

Table 1- Summary of statistical characteristics of data

2.2. Structure of ANNs

After assessment of predictability of available data (R/S analysis), performance of different neural network types was checked roughly. Upon the gained results and based on the literature, Multi Layer Feed forward Perceptron (MLFP) neural network with back propagation training method was selected. Fig (2) shows the three layer topology that was used in this research. The input monthly data for 53 years period (53×12) were arranged as a time series with length of 638 monthly data. Consequently, from this time series, 580 data was used for the training phase randomly, and the rest were used for the validation phase. The number of hidden layer neurons, in a trial-error process, was chosen in a way to obtain the best output from the networks. It should also be noted that epochs was selected equal to 1000 and η was held at constant 0.5. The selected activation functions in hidden and output layer of the networks were sigmoid (Eq (1)) and linear (Eq (2)), respectively.

$$f_{(x)} = \frac{1}{1 + e^{-ax}} \tag{9}$$

$$a = f(n) = n \tag{10}$$

All applied variants of MLFP structures, for the purpose of monthly predictions, were designed by input layers related to precipitation data of the previous time period, and output layers predict monthly precipitation.

3- RESULTS AND DISCUSSIONS

R/S time series analysis test was performed in order to determine the predictability of available precipitation data. Hurst exponent of monthly data was calculated 0.9621. This value implies that the monthly set of given time series is predictable. Afterwards, using MATLAB, several different structures of MLFP were designed with different numbers of neurons in the input as well as hidden layers. Superior models for monthly prediction (such as, M_{531} and M_{741}) are explained in the first and second following subsections. The three indexes of models are: the number of neurons in input (i), hidden (j) and output (k) layers.

3.1. M_{531} neural network:

The input layer of this network consists of 5 neurons for the last 5 bi-monthly precipitation moving-averages. For this network, trial and error yielded 3 neurons as the best numbers of neurons for the hidden layer. The structure of this network has been shown in Fig (1).
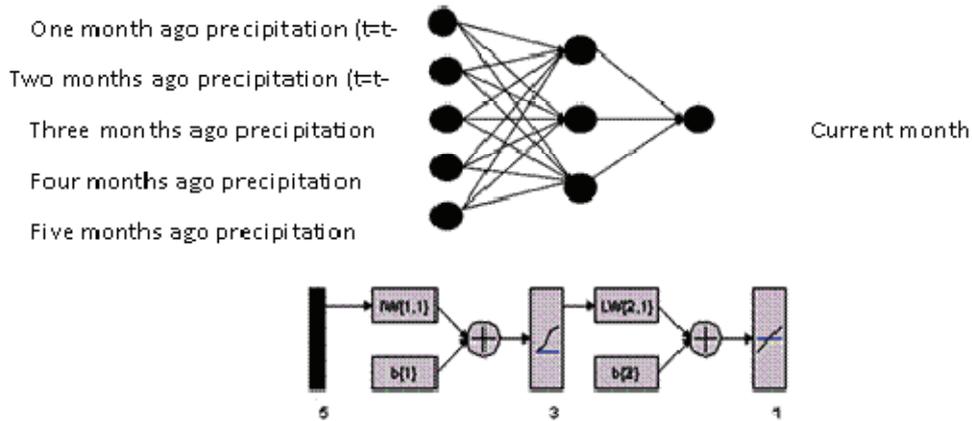
One month ago precipitation (t=t-

Two months ago precipitation (t=t-

Three months ago precipitation

Four months ago precipitation

Five months ago precipitation

Current month

Figure 1- M$_{531}$ neural network structure

3.2. M$_{741}$ neural network

In this network, 7 neurons were used in the input layer. The seven inputs are long term average of precipitation for the estimated month, last year precipitation for the estimated month, and the last 5 bi-monthly precipitation moving-averages, respectively. After some trial and error, seeking optimal number of neurons in the hidden layer, 4 neurons were selected. Moreover, for this network, all input data were initially normalized to decrease range of variables variations, and consequently to increase the algorithm learning rate. Obviously, the network output should be de-normalized. Fig (2) shows the topology of this network.
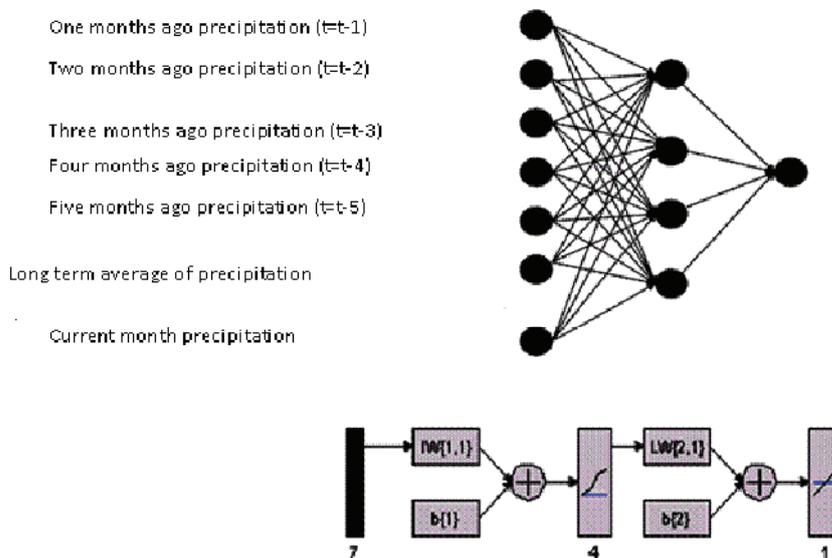
One months ago precipitation (t=t-1)

Two months ago precipitation (t=t-2)

Three months ago precipitation (t=t-3)

Four months ago precipitation (t=t-4)

Five months ago precipitation (t=t-5)

Long term average of precipitation

Current month precipitation

Figure 2- M$_{741}$ neural network structure

Table (2) shows results of models, including Correlation Coefficient (R), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Variance Accounted for between two signals (VAF) and Non Dimensional Error Index (NDEI) for monthly prediction using M$_{531}$ and M$_{741}$.

ANN structures	R	RMSE (mm)	MEA	VAF	NDEI
M_{531}	0.89	1.10	6.67	81.00	0.053
M_{741}	0.92	0.99	6.02	85.27	0.048

Table 2- results two ANNs structures

Table(2) shows a strong correlation ($R^2 = 0.79$) with slope of the trend line (0.8961) indicate that the M_{531} model provided good predictions of monthly precipitation values for Mashhad's synoptic station. However, there is a slight tendency to under predict of precipitation and this can be attributed largely to the input layer. That means in this MLFP structure only the last 5 bi-monthly precipitation moving-averages precipitation values were employed for future precipitation prediction and it seems that because of precipitation values in the last five time steps and insufficient training, precipitation values in the next time was under predicted. However, certainly this input layer always doesn't result in under prediction and perhaps in the other problems with different data and locations this one causes different results and even over prediction.

A stronger correlation ($R^2 = 0.84$) between predicted and actual precipitation values is observed for validation phase of M_{741} in comparison to M_{531}. Due to using input neurons included long term average of precipitation so as to estimate monthly, last year precipitation for the estimated month beside the last 5 bi-monthly precipitation moving-averages, M_{741} has a better result than M_{531}. In other words, it seems that these results because of quantities and quality of input layer and better network training in M_{741}.

Fig (3) shows powerfully ability of M531 and M741 validation for monthly precipitation. Obviously, there is agreement between estimated and observed precipitation in M_{531}. Nevertheless the performance of M741 is higher than M_{741} for above reasons.
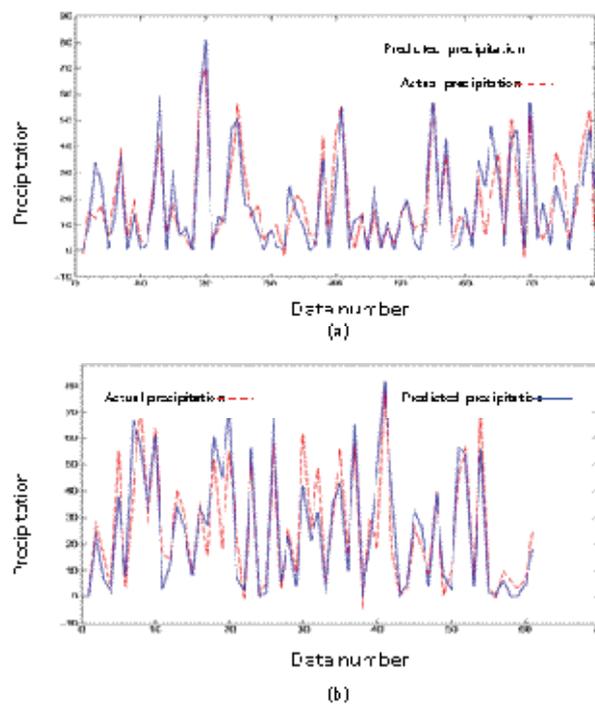

Figure 3- Comparison between predicted and actual monthly precipitation for validation phase; (a) M_{741}, and (b) M_{531}

4- CONCLUSION

The following consequences were derived from the test results of this study:

- ANN types, training algorithms, activation functions, epochs, etc for determination of the best models are different for various problems and different type of raw data. The selection of appropriate structure could be specified by several trial and errors.
- Regarding to training and validation phases for available rainfall data in this study, the best-achieved structure was a three-layer feed-forward perceptron with back propagation algorithm in the form of M_{741}. The inputs also selected from the monthly data in the past time steps with a special pattern. In this structure of ANN, the prediction performance was assessed by different statistical criteria like R, RMSE, MAE, VAF and NDVI. For this case (M_{741}) these parameters obtained as 0.93, 0.99 mm, 6.02 mm, 85.27 and 0.048, respectively.
- Using long-term average of precipitation for the current month along with precipitation for the corresponding month in the previous year and the last five bi-monthly precipitation moving-averages, results in the best acceptable achievable prediction.

Acknowledgment:

This research is supported in data by meteorology organization of Iran.

LIST OF REFERENCES

Ramirez, M.C.V., N.J., Ferreira, and H.F. Velho,(2005). Artificial *neural network technique for rainfall forecasting applied to the Sa˜o Paulo region* , Journal of Hydrology,301, 146–162.

Luk, K.C.,(2001). *An application of artificial neural networks for rainfall   forecasting*, Mathematical and Computer Modeling Journal, 33, 883-699.

Saplioglu K, Cimeny M, Akman B.(2010) Daily Precipitation Prediction in Isparta Station by Artificial Neural Network. Proceedings of the 4th International Scientific Conference on Water Observation and Information System for Decision Support (BALWOIS). Ohrid, Republic of Macedonia

Khalili N. (2008) Rainfall prediction using artificial neural networks. Thesis of Master degree of Irrigation and drainage engineering. Faculty of Agriculture, Ferdowsi university of Mashhad (In Persian).