ARTIFICIAL NEURAL NETWORK MODELS COMBINED WITH SIMPLE STATISTICAL HYDROLOGY TOOLS

Hafzullah Aksoy* & Ahmad Dahamsheh**

***Istanbul Technical University, Department of Civil Engineering, Hydraulics Division, 34469 Maslak, Istanbul, Turkey**

****Al-Hussein Bin Talal University, College of Engineering, Department of Civil Engineering, Ma'an, Jordan**

haksoy@itu.edu.tr

IUGG MELBOURNE Australia 2011

1 INTRODUCTION

Forecasting precipitation in arid and semi-arid regions has particular importance since precipitation is the unique source of water in such regions. In this study, one-month ahead precipitation forecasts are made using different type artificial neural network (ANN) models (only Feed Forward Back Propagation – FFBP - ANN model is shown in this poster). The models are tested on monthly total precipitation data. Markov chain is incorporated into the model such that dry months can be forecasted and negative forecasts which has no physical meaning are prevented in those particular months. Due to the importance of data length to better calibrate the models synthetic data statistically similar to observed data are generated by using Thomas-Fiering model known in the hydrological literature. At the same time a simple empirical relation is incorporated into the models where inputs are treated depending on the season of the year. It is found that Markov chain and synthetic data improve the forecast although not substantially. However, the performance of Markov chain in forecasting zero-precipitation in dry months is remarkable. Synthetic data do not increase the performance of the models considerably although a slight improvement is observed. The main improvement is obtained by the empirical conditional model. By using this empirical concept, the models become much better with respect to all performance criteria used in the study, including determination coefficient, mean square error, mean absolute error, the slope and the intercept in the best-fit linear line of the scatter diagram (not all depicted in this poster).

2 STUDY AREA & DATA

Three meteorological stations (Baqura, Amman, and Safawi of Jordan Meteorological Department) from different climatological regions in Jordan in the Middle East are selected for investigation in this study (only Amman is presented). Fig. 1 shows the locations of the stations. Table 1 summarizes some details of the Amman station together with statistical characteristics of the precipitation data in this particular station. Table 2 shows that in Amman, the monthly total precipitation in the seven-month period extending from April to October remains well under the long-term average. This time span is, therefore, considered as the dry period within which four months (June-September) are completely dry with no precipitation at all (when monthly total precipitation less than 1 mm is ignored).

Table 1

Station	Observation Period	Latitude N	Longitude E	Elevation (m)	Mean (mm)	St.Dev (mm)	Median (mm)	C_v	C_s	r_1	Maximum (mm)	Year	Minimum (mm)	Year
Amman	1923-2005	31.98	35.98	772	270.4	90.0	263.3	0.33	0.30	-0.133	476.5	1938	98.0	1995

Table 2

Month	1	2	3	4	5	6	7	8	9	10	11	12
Mean (mm)	64.1	62.3	42.8	13.7	3.4	0.0	0.0	0.0	0.3	6.3	28.0	49.5
St.Dev (mm)	38.2	44.5	35.6	20.7	6.4	0.2	0.0	0.0	1.8	9.9	31.2	43.6
C_v	0.60	0.71	0.83	1.51	1.88	4.70	9.11	*	6.02	1.58	1.11	0.88
C_s	1.18	0.96	1.45	4.12	2.50	4.94	9.11	*	7.40	2.44	1.88	1.23
Max. (mm)	235.2	199.8	168.7	151.3	30.6	1.2	0.2	0.0	15.4	54.6	136.8	179.8
Min (mm)	1.7	5.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4
r_1	0.07	-0.05	0.12	0.06	0.11	-0.05	-0.01	*	-0.03	0.06	-0.06	-0.20
Zero (%)	0.0	0.0	1.2	7.2	37.3	95.2	98.8	100.0	91.6	22.9	3.6	0.0

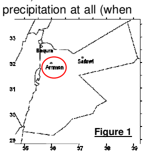


Figure 1

3 METHOD

The FFBP ANN consists of three or more layers, namely input layer, hidden layer(s) and output layer, each with a certain number of neurons. The number of input (independent) variables directly gives the number of neurons in the input layer. Similarly, the number of dependent variables is equal to the number of neurons in the output layer. The number of neurons in the hidden layer is subject to determination by a trial-and-error procedure with which the error between the observed variable and the model output is minimized. Using a proper function *f*, the input variables (x_i, i = 1, ..., n) are processed in the hidden layer as:

Equation 1

$$z_j = f\left(\sum_{i=1}^{n} x_i w_{ij} + b_j\right), \; j = 1,\ldots,h$$

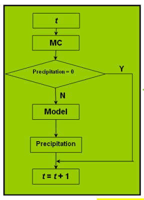
where w_{ij} is the weight of the connection from the ith input neuron to the jth hidden neuron, and b_j is the bias for the jth hidden neuron.

In this study, the network has only one hidden layer. It is noted, from Equation (1), that *h* neurons exist in the hidden layer. In order to obtain the model output, the outputs of the hidden layer (z_j, j = 1, ..., h) are transformed by

Equation 2

$$y_k = f\left(\sum_{j=1}^{h} z_j w_{jk} + b_k\right), k = 1,\ldots,m$$

where w_{jk} is the weight of the connection from the j^{th} hidden neuron to the k^{th} output neuron, and b_k is the bias for the k^{th} output neuron. It is noted, from Equation (2), that m neurons exist in the output layer.



Model Inputs

$$x_t = f(x_{t-1}, x_{t-2}, x_{t-12}, p_t) \Longleftarrow P(t) = sin\left(\frac{2\pi t}{12}\right) + cos\left(\frac{2\pi t}{12}\right)$$

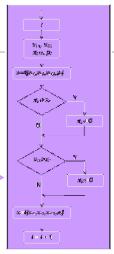
$$(x_{t-1} > x_t \Rightarrow x_{t-1} = 0) \cap (x_{t-2} > x_t \Rightarrow x_{t-2} = 0) \quad for \;\; \forall t$$

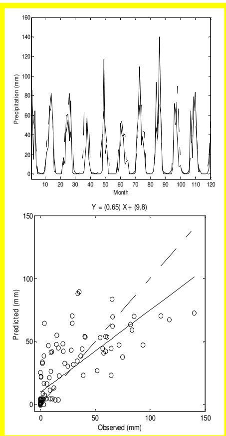
Markov Chain

Synthetic Series

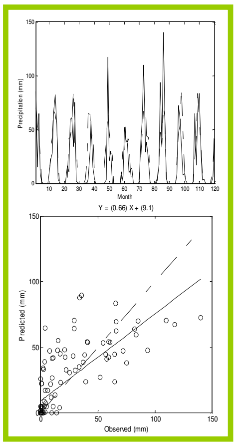
Conditional Statement

ANN ⟹ **MC-ANN** ⟹ **MC-sANN** ⟹ **MC-CsANN**

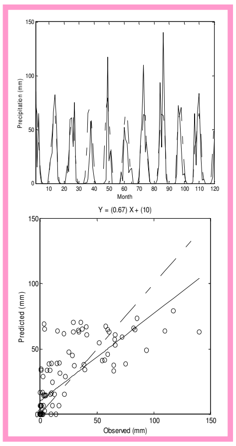




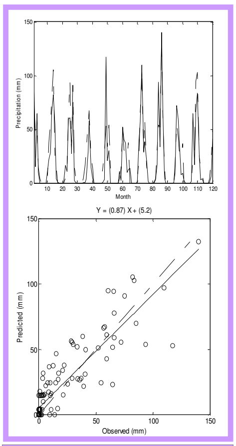
ANN Model



Markov chain incoparated ANN Model



Markov chain incoparated ANN Model trained with synthetic data



Markov chain incoparated conditional ANN Model trained with synthetic data

4 CONCLUSION

Based on the present results and observations, the following conclusions may be drawn:

- Intermittent monthly precipitation data of semi-arid to arid regions should be analyzed with more care than perennial precipitation time series of the humid regions.
- None of the ANN models in this study turned out to be a best choice, as there may always exist a better model that can be constructed on the basis of a trial-and-error procedure, with a different architecture, different activation functions or spread coefficients (when applicable).
- Higher variability results in worse performance of ANN models, i.e. increasing variability reduces the success of the models in approaching the observed precipitation sequence.

Ideas for the future

Some of the potential limitations of the models can be eliminated, and improvements in model performance may be achieved by considering the following:

- Monthly analysis can be performed and monthly models can be proposed at the expense of increase in the model parameters, weights and biases.
- If a parsimonious model is desired, and if the monthly analysis is found too costly with respect to model parameters, then the dry and wet periods in the year can be modeled separately. The observed data can simply be divided into two periods; for instance above and below the long-term average precipitation to be taken as a threshold. Models can then be constructed separately for each period.
- It is noted that the analysis of the intermittent monthly precipitation requires some helping tools to be incorporated into the forecasting models. For instance; Markov chains and synthetic data generation tehniques are simple statistical tools used for decades in hydrological applications and can be used together with ANNs. Promising results are already seen with forecasting precipitation, particularly in dry months of the year, and eliminating negative precipitation forecasted by ANN models.

Published Works

1. Dahamsheh A & Aksoy H (2007) Structural characteristics of annual precipitation data in Jordan. *Theoretical and Applied Climatology*, 88, 201-212.
2. Dahamsheh A (2008) *Forecasting monthly precipitation for arid regions using conditional artificial neural networks combined with Markov chain.* PhD Thesis, Istanbul Technical University, Institute of Science and Technology, Istanbul, Turkey, 235 p.
3. Dahamsheh A & Aksoy H (2009) Artificial neural network models for forecasting intermittent monthly precipitation in arid regions. *Meteorological Applications*, 16, 325-337.
4. Aksoy H & Dahamsheh A (2009) Artificial neural network models for forecasting monthly precipitation in Jordan. *Stochastic Environmental Research and Risk Assessment*, 23, 917-931.