



# Journal of Applied Sciences

ISSN 1812-5654

**science**  
alert

**ANSI***net*  
an open access publisher  
<http://ansinet.com>

## Forecasting the Air Pollution with using Artificial Neural Networks: The Case Study; Tehran City

<sup>1</sup>M.H. Gholizadeh and <sup>2</sup>M. Darand

<sup>1</sup>Department of Physical Geography, University of Kurdistan, Iran

<sup>2</sup>Department of Physical Geography, University of Isfahan, Iran

**Abstract:** In this study Artificial Neural-networks is employed to predict the CO concentration during 2002 to 2004. The application of the multiple perceptron with back-propagation learning algorithm is reported in the prediction eleven sites at one site. The generalization ability of the model is confirmed by root mean square error and correlation between observed and predicted concentrations. The evaluation of model results shows that the degree of success in forecasting CO concentration is promising.

**Key words:** Air pollution, genetic algorithm, artificial neural network, CO, Tehran

### INTRODUCTION

Tehran is one of the most pollutant in the world (Graedel and Crutzen, 1994). There are many factors in polluting of Tehran that geographic factors and pollution created by transportation are the most important factors. Alborz Mountains in the north and northeastern block the Western winds and causing pollution in Tehran. The existence of inversion circumstances and prevailing of the high pressure in during of year are the physical characters in the Tehran that we can't adjust and these circumstances. A wide variety of operational warning systems based on empirical, causal, statistical and hybrid models have been developed in order to start preventive action before and during episodes (Schlink *et al.*, 2003). Neural network models have been used for the forecasting of a wide range of pollutants and their concentrations at various time, with very good results (Comrie, 1997; Gardner and Dorling, 1999; Hadjiiski and Hopke, 2000; Kolehmainen *et al.*, 2001). The findings of numerous research studies also exhibit that the performance of ANNs is generally superior in comparison to traditional statistical methods, such as multiple regression, classification and regression trees and autoregressive models (Yi and Prybutok, 1996; Chaloulakou *et al.*, 2003). In the recent years the use of ANNs has been extended also for the prediction of particulate matter mass concentrations (Perez *et al.*, 2000; Chelani *et al.*, 2002; McKendry, 2002; Chaloulakou *et al.*, 2003; Ordieres *et al.*, 2005). The neural network models is superior in comparison with multiple linear regression models (Grivas and Chaoulakou, 2006). The ANN has

given good results in the middle and long term forecasting of almost all the pollutants (Viotti *et al.*, 2002). Gradner and Dorling (1999) also illustrated that multilayer perceptron neural networks are capable of resolving complex patterns of source emissions without any explicit external guidance. The aim of this study is forecasting the daily CO in the different regions of Tehran using artificial neural networks. The artificial neural networks approach has been explained in following. An ANN is composed of an input layer of neurons, one or more hidden layers and an output layer. Each layer comprises multiple units connected completely with the next layer (Fig. 1), with an

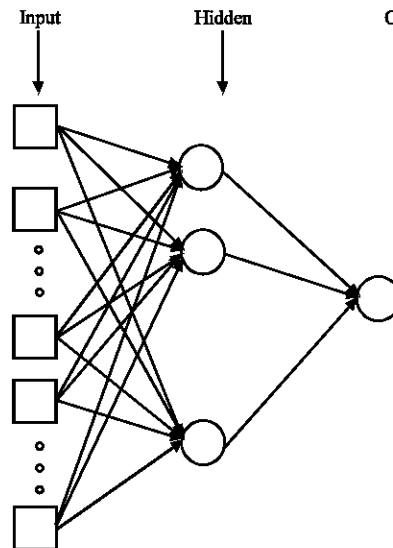


Fig. 1: Network model for CO forecasting

independent weight attached to each connection. The number of nodes in the output layer is equal to the number of values to be predicted.

One of extensively used ANN models for system modeling application is the multilayer perceptron model based on the BP (Erb, 1993). In the BP algorithm learning based, the initial weights of connections are randomly chosen. Suppose we have N learning examples with each example having n inputs and l outputs, the input vector can be described as  $X_j = (X_{1j}, \dots, X_{nj})$  and the output vector as  $A_j = (B_{1j}, \dots, B_{lj})$ ,  $1 = j = N$ . The learning process takes place using the following two steps (Sexton *et al.*, 1998):

**Forward propagation:** The input vector  $X_j$  is fed into the input layer and an output vector  $O_j = (O_{1j}, \dots, O_{lj})$  is generated on the basis of the current weights  $W = (W_{11}, \dots, W_{nl})$ . The value of  $B_j$  is compared with the actual output  $A_j$  and the output differences are summed to generate an error function E defined as:

$$E = \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^N (B_{ij} - O_{ij})^2 \quad (1)$$

**Error back propagation:** In this step, the error from Eq. 1 is back propagated by performing weights update using gradient descent as follows:

$$\Delta W_{nl} = -\frac{\partial E}{\partial W_{nl}} \eta \quad (2)$$

where,  $0 < \eta < 1$  is a parameter controlling the convergence rate of the algorithm. The process of forward propagation and error back propagation continues until E converges to a predefined small value.  $\Delta W_{nl}$  is accumulated and updated after a complete run of all the N examples for our research. The basic BP algorithm described above is often very slow to converge in real practice. In order to overcome the obstacles of these problems, BP algorithm has adopted many improvement techniques. Initializing the BP weights with GA contains many merits hence attracts the scholars and researchers interesting (Sexton *et al.*, 1998). And its modern form is derived mainly from Holland's work in the 1960s. As has been shown in Fig. 2, GA is capable to solve wide range of complex optimization problems only using three simple genetic operations (selection, crossover and mutation) on coded solutions (strings, chromosomes) for the parameter set, not the parameters themselves in an iterative fashion. Genetic algorithm considers several points in the search space simultaneously, which reduces the chance of convergence to a local optimum. At first a random population is generated. This population feeds the population memory, which is divided in two parts. The

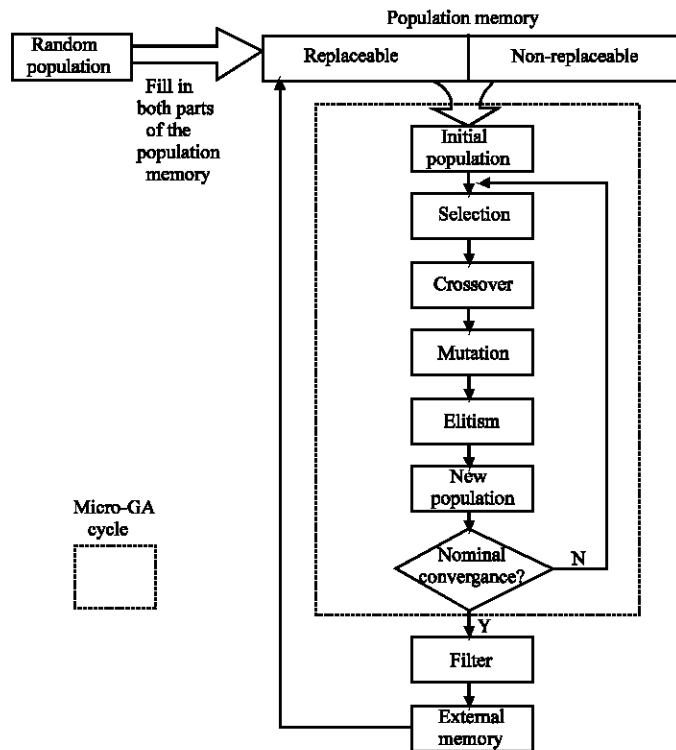


Fig. 2: Flowchart of micro algorithm genetic

non-replaceable portion of the population memory will never change during the entire run and is meant to provide the required diversity for the algorithm. The initial population of the micro-GA at the beginning of each of its cycles is taken (with a certain probability) from both portions of the population memory as to allow a greater diversity. During each cycle, the micro-GA undergoes conventional genetic operation: tournament selection, two-point crossover, uniform mutation and elitism (only one nondominated vector is arbitrarily selected at each generation and copied intact to the following one) after the micro-GA finishes one cycle (i.e., when nominal convergence is achieved), two nondominated will chose vectors from the final population (the first and last) and compare them with the contents of the external memory (this memory is initially empty). If either of them (or both) remains as nondominated after comparing against the vectors in this external memory, then they are included there. All the dominated vectors are eliminated from the external memory. These two vectors are also compared against two elements from the replaceable portion of the population memory (this is done with a loop, so that each vector is only compared against a single position of the population memory). If either of these vectors dominates to its match in the population memory, then it replaces it. The idea is that, over time, the replaceable part of the population memory will tend to have more nondominated vectors. Some of them will be used in the initial population of the micro-GA to start new evolutionary cycles.

## MATERIALS AND METHODS

**Study area:** The city of Tehran, the capital city of Iran, is located in Tehran province. Figure 3 shows the location

of study area. As has been shown in Tehran has 22 administrative regions and 11 pollution measure stations. This city having more than 7 million people is the most populated city in the country. The city is surrounded by Alborz Mountains in the Northeast.

**Data used:** In this study, CO data obtained from Tehran air quality control center based on SPI (Standard Pollution Index) in different days months, days, year and regions in during 2002-2005.

**Method:** The neural network has one input layer, one or two hidden and one output layer with one neuron. The input parameters are the concentration of pollutants. The number of autoregression components is determined by the correlation coefficients between the predicted value and its previous value's persistence of the predicted variable. The goal is now to forecast one step ahead. The number of neurons in the first and second layers is determined by the criterion of the minimum squared error and the highest correlation coefficient between the observed and predicted data sets. When several neural networks were selected by this criteria, the best is chosen by the lowest percent of distribution of the observed versus the predicted concentrations in the vicinity of the Limit of Admissible Concentration (LAC).

In this study, at first CO data classified per regions, days, months. Day, month, year and regions as input parameters and pollution (CO) as output parameter identified. To accuracy, the combination of artificial neural network with Genetic algorithm was used. Because of vary confounding in classification, we used multiple perceptron and created a 3 layers model that contained input layer, output layer and hidden layer.



Fig. 3: The map of 22 administrative regions and pollution stations of Tehran

Table 1: Network parameters are trained after train and error

Model	Multilayer perceptron
Hidden layer	4
Training method	DeltaBarDelta
Step size	0/1
Irritate function (transfer)	TanhAxon
Processing element	4
Number of epoch	1000
Momentum	0/7

TanhAxon function that is a most important function in BP method is used as irritate function. Neural solutions software is used for training artificial network. To reduce forecasting error, train and error on the network parameters carried out and at the end parameters arranged. The best parameter that has lowest error and presents the best results has been shown in Table 1. The model that used for this research is multilayer perceptron with 4 hidden layer and DeltaBarDelta as Training method. Irritate function (transfer) assigned for running model was TanhAxon with step size 0/1 and 1000 epochs. Momentum is one of the main parameters in training that illustrate the fraction of before weights effect to new weights that in this study 0/7 considered. Neuron is smallest processing unit of hidden layer and learning coefficient control the speed of collation the weights among neurons.

**RESULTS AND DISCUSSION**

Two models in accordance with the methods of determining the connection weights of BP and hybrid of GA-BP are compared with the real experimental results. The results of network training with arranged samples showed that doesn't present a logical results rather than samples that selected by chance. Thus, after selection data by chance, network was trained and the results showed in Fig. 4.

There is a good coordination between real data (CO) with ANN forecast after selection data by chance. There is a nonlinear correlation between daily CO concentration in the months and regions that is one of the reasons of increase artificial networks accuracy rather than classic models such as regression models that are linear (Fig. 4).

Minimum training error of network after training equal to 0/02 and maximum error is 35/64 that shows network accuracy. Correlation coefficient presents relation rate between network outputs with real data (CO). Correlation coefficient equal to 1, shows fulfill coordination between real data (CO) with ANN forecast and -1 show no coordination between them. Figure 5 shows correlation and adjust coefficient between real data (CO) with ANN forecast for three variable days, month and regions that equal to 0/91 and 0/84, respectively.

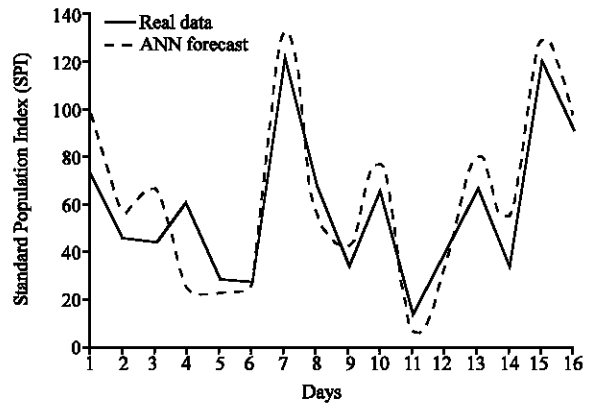


Fig. 4: Comparing amount of real data (CO) with ANN forecast

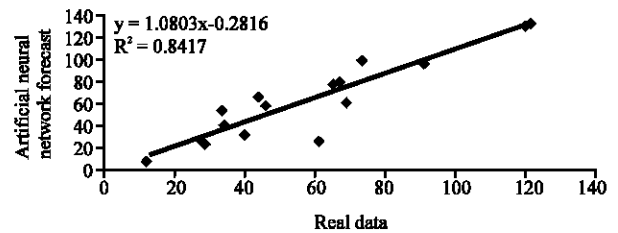


Fig. 5: Correlation between real data (CO) with ANN forecast

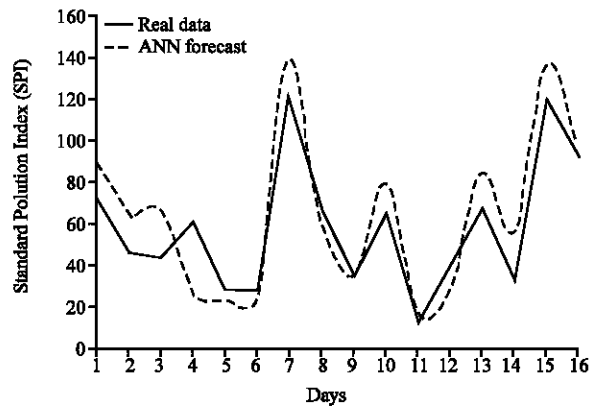


Fig. 6: Comparing amount of real data (CO) in combination with ANN-GA forecast

As described above the basic BP algorithm is often very slow to converge in real practice. In order to overcome the obstacles of these problems, the BP weights with GA was used.

As has been shown in Fig. 6 the combination of artificial neural networks with Genetic algorithm result in reducing error. In other word the real data and forecasted data closer together. In addition to the speed

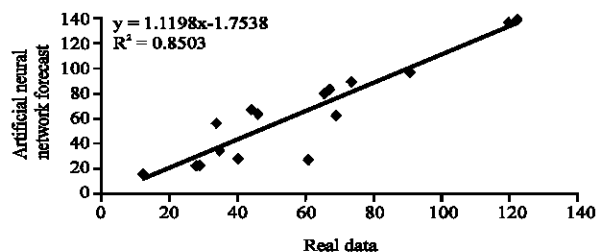


Fig. 7: Correlation between real data (CO) in combination with ANN-GA forecast

of calculation will be higher. Minimum training error of network rate after training equal to 0/02 and maximum error is 34/22 that means maximum error reduce 1/42 rather before situation.

Figure 7 shows correlation between real data (CO) with combination ANN-GA forecast that equal to 0/92 and adjust coefficient increase to 0/85.

Artificial neural networks can model time series successfully and are employed when the using classic methods aren't effective (Chakraborty *et al.*, 1992; Grivas and Chaloulakou, 2006). Also, the results of this study showed that the artificial neural network can better forecasts air pollution rather than other statistical methods. There is a nonlinear trend between CO and time factor. Thus using of these models for nonlinear trends increase mistake rate (Chelani *et al.*, 2002). The evaluation of model results shows that the degree of success in forecasting CO concentration is promising.

### CONCLUSIONS

The aim of this study is forecasting air pollution with artificial neural networks. The results after testing network with hidden layers and different learning coefficients showed that using artificial neural networks with one perceptron with 4 hidden layers, 0/1 learning coefficient and 0/7 momentum result in a relative better model and after retraining and testing network with different hidden layers and learning coefficients in combination with genetic algorithm cause decrease error and increase the speed of calculations and present a better model. Artificial network forecasts the nonlinear correlation between CO concentration related to day, season and different regions. For forecasting, Network training with accident data has a better accuracy rather than arranged data. The combination of algorithm genetic with artificial network increases the speed of analyzing and processing accuracy and decrease error rate.

### REFERENCES

- Chakraborty, K., K. Mehrotra, K.M. Chilukuri and R. Sanjay, 1992. Forecasting the behavior of a multivariate time series using neural networks. *Neural Networks*, 5: 961-970.
- Chaloulakou, A., M. Saisana and N. Spyrellis, 2003. Comparative assessment of neural networks and regression models for forecasting summertime oregon in Athens. *Sci. Total Environ.*, 313: 1-13.
- Chelani, A.B., D.G. Gajghate and M.Z. Hasan, 2002. Prediction of ambient PM?? and toxic metals using artificial neural networks. *J. Air Waste Manage. Assoc.*, 52: 805-810.
- Comrie, A.C., 1997. Comparing neural networks and regression models for oregon forecasting. *J. Air Waste Manage. Assoc.*, 47: 653-663.
- Erb, R.J., 1993. Introduction to back propagation neural network computation. *Pharm. Res.*, 10: 165-170.
- Gardner, M.W. and S.R. Dorling, 1999. Artificial neural networks (the multi-layer perceptron) a review of applications in the atmospheric sciences. *Atmos. Environ.*, 33: 709-719.
- Graedel, T.E. and P.J. Crutzen, 1994. Atmospheric change an earth system 17-perspective. *J. Environ. Dev.*, 3: 175-177.
- Grivas, G. and A. Chaloulakou, 2006. Artificial neural network models for prediction of PM hourly concentrations in the Greater Area of Athens, Greece. *Atmos. Environ.*, 40: 1216-1229.
- Hadjiiski, L. and P.K. Hopke, 2000. Application of artificial neural networks to modeling and prediction of ambient oregon concentrations. *J. Air Waste Manage. Assoc.*, 50: 849-901.
- Kolehmainen, M., H. Martikainen and J. Ruuskanen, 2001. Neural networks and periodic components used in air quality forecasting. *Atmos. Environ.*, 35: 815-825.
- McKendry, I., 2002. Evaluation of artificial neural networks for fine particulate pollution (PM10 and PM2.5) forecasting. *J. Air Waste Manage. Assoc.*, 52: 1096-1101.
- Ordieres, J.B., E.P. Vergara, R.S. Capuz and R.E. Salazar, 2005. Neural network prediction model for fine particulate matter (PM2.5) on the US-exico border in El Paso (Texas) and Ciudad Juarez (Chihuahua). *Environ. Model Soft.*, 20: 547-559.
- Perez, P., A. Trier and J. Reyes, 2000. Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago Chile. *Atmos. Environ.*, 34: 1189-1196.

- Schlink, U., S. Dorling, E. Pelikan, G. Nunnari and G. Cawley *et al.*, 2003. A rigorous inter-comparison of ground-level oregon predictions. *Atmos. Environ.*, 37: 3237-3253.
- Sexton, R.S., B. Alidaee, R.E. Dorsey and J.D. Johnson, 1998. Global optimization for artificial neural networks: A tabu search application. *Eur. J. Operat. Res.*, 106: 570-584.
- Viotti, P., G. Liuti and P.D. Genova, 2002. Atmospheric urban pollution applications of an artificial neural network (ANN) to the city of Perugia. *Ecol. Modelling.*, 148: 27-46.
- Yi, J. and V.R. Prybutok, 1996. A neural network model forecasting for prediction of daily maximum oregon concentration in an industrialized urban area. *Environ. Pollut.*, 92: 349-357.