Asian Journal of Water, Environment and Pollution, Vol. 9, No. 1 (2012), pp. 71–81.

Artificial Neural Network Model to Forecast the Concentration of Pollutants Over Delhi: Skill Assessment of Learning Rules

Sutapa Chaudhuri* **and Rajashree Acharya**

Department of Atmospheric Sciences, University of Calcutta
51/2 Hazra Road, Kolkata – 700019, India
✉ chaudhuri_sutapa@yahoo.com

Received May 21, 2010; revised and accepted November 7, 2011

Abstract: Air pollution has been reported to persuade climate as well as health significantly and is a matter of concern. The scientific endeavour should thus, be to develop forecast or warning system to predict the concentration of pollutants with considerable accuracy so that the calamities associated with pollution can be minimized, if not eradicated. The purpose of the study is to develop an Artificial Neural Network (ANN) model with different learning rules to predict the concentration of pollutants over Delhi (28° 38′N, 77° 12′E), India for the year 2009. Two types of learning rules are implemented in this study to forecast the concentration of different pollutants. The result reveals that the forecast accuracy of a particular pollutant depends on the type of the learning rule of the ANN model. The result of the study further reveals that the non-linear perceptron is better for forecasting the concentration of sulphur dioxide (SO_2), carbon monoxide (CO), suspended particulate matter (SPM) and ozone (O_3) whereas delta learning is better for forecasting nitrogen dioxide (NO_2). The percentage errors in forecast with different learning rules of the ANN model are compared for all the pollutants. The result shows that the concentration of SO_2 can be predicted over Delhi with maximum accuracy using nonlinear perceptron.

Key words: Concentration, air pollutants, prediction, artificial neural network, non-linear perceptron, delta learning.

Introduction

Air pollution is a global problem, starting from the crowded city to top of the Mount Everest. The air is polluted by thousands of substances. Air pollution threatens with long-term changes caused by depletion of the ozone layer and global warming causing changes in climate, health hazards, water pollution and damaging ecosystems. Pollutions are extremely harmful to the environment and could result in serious damages. The only solution of air pollution is to prevent it (Perez et al., 2000).

There are several types of pollution and well-known consequences which are commonly discussed in literature. Pollution is the release of noxious gases, such as sulphur dioxide, carbon monoxide, oxides of nitrogen, and chemical vapours. These can take part in further chemical reactions once they are in the atmosphere, forming smog and acid rain (Comrie, 1997).

Air pollutants including oxides of sulphur (SO_x), oxides of nitrogen (NO_x) and hydrocarbons are puffing up from various sources and their effect on human become a serious problem. SO_x causes nose, eye irritation with respiratory illness (Kampa and Castanas, 2008). The principal sources of SO_x are industry and transportation. It is released in the atmosphere due to combustion of sulphur containing fuels. The main sources of nitrogen oxides are transportation and industrial burning. Carbon monoxide is the most abundant pollutant in the lower atmosphere. The sources of carbon monoxide are motor

*Corresponding Author

vehicle exhaust. Photochemical smog is the mixture of reactants and products when hydrocarbon and nitrogen oxides are combined together in the presence of sunlight. Ozone mainly forms by the chemical reactions of various primary pollutants. These pollutants affect health of human beings and animals and damage plants. An accurate estimation of the concentration of various pollutants is, therefore, very important for estimating the climate variability, social planning and industrial growth.

Most of the studies involving pollution and meteorological parameters were statistical in nature (Elsom and Chandler, 1978; Annand and Hudson, 1981). However, pollution and weather relationship is typically complex and non-linear and therefore the association might be established in a sensible manner with artificial neural network technique (Patterson, 1996). In the recent times, the interrelationship between air pollution and atmospheric events are observed using artificial neural network (Díaz-Robles et al., 2008). The present study, thus, aims at developing artificial neural network model to forecast the concentration of pollutants over Delhi.

Data Analysis

The dataset explored in this study consist of the concentration of five pollutants over Delhi. The pollutants are sulphur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), ozone (O_3) and suspended particulate matter (SPM).

Delhi is selected for this study because Delhi has been considered to be the fourth most polluted city in the world (World Health Organization, 2001). The datasets considered in the present study are collected from the website of the Central Pollution Control Board, Delhi (www.cpcb.co.in). Website provides the daily concentrations of the pollutants in $\mu g/m^3$. The data of sulphur dioxide, nitrogen dioxide, carbon monoxide, ozone and SPM are collected during the period from 2000 to 2009. The data are arranged as monthly mean for all the months of the year. The ANN model is trained with the data of 2000 to 2008 and tested or validated with the data of 2009.

Methodology

Artificial neural network is simplified imitations of the central nervous system and obviously has been motivated by the kind of computing performed by the human brain. The structural constituents of a human brain termed as neurons are the entities, which perform computations such as cognition, logical inference, pattern recognition and so on. ANN is a massively parallel distributed processing system made up of highly interconnected neural computing elements that have the ability to learn and thereby acquire knowledge and make it available for use (Hornik, 1991). Hence the technology, which has been built on a simplified imitation of computing by neurons of a brain, has been termed artificial neural network or simply neural network. Plethora of literature is available where the ANN is extensively discussed (Chaudhuri, 2006).

Two types of learning methods are considered to forecast the concentration of pollutants. These are the ‘perceptron’ learning rule and the ‘delta’ or ‘LMS’ rule. Both methods are iterative procedures that adjust the weights.

Multilayer or non-linear feed forward network is made up of multiple layers (Chaudhuri, 2007). Thus the network besides possessing an input and an output layer also have one or more intermediary layers called hidden layers (Jacobson, 1998). The optimal weights are obtained by using the back propagation learning method (Rajaskaran and Vijayalakshmi, 2003) which is built on high mathematical foundation and has very good application potential.

Delta rule works with a linear activation function and this type of network comprises mainly two layers, namely the input layer and the output layer. The input layer neurons receive the input signals and the output layer receives the output signals. The synaptic links carrying the weights connect every neuron to the output neuron but not vice-versa. Such a network is fed forward in type and the name is single layer or linear feed forward network.

Implementation Procedure

The initial step to construct the ANN model is the selection of the input data set. Concentrations of five different pollutants are considered here to form the input matrix of the ANN model. The first nine years (2000–2008) data is used to train the model. The model is tested/validated with the data of 2009. The input data set is arranged in a matrix of order (12 × 9) and the order of the output matrix is (12 × 1). The next step is to standardize the input data between 0 and 1 because the sigmoid function is used as the activation function which squash the output value in the range between 0 and 1 (Hagan, 1996). Scaling is performed to avoid the asymptotes of the sigmoid function (Comrie, 1997);

$$Z_i = 0 + \frac{(X_i - X_{\min})}{(X_{\max} - X_{\min})} \quad (1)$$

where X_i = ith input value and X_{min} and X_{max} are the minimum and maximum values respectively of the input matrix. A particular weight within the range (0 to +1) (www.learnartificialneuralnetworks.com) is chosen for a particular input data in the form of a matrix of order (9 × 1) and the weights are updated with learning rules in every epoch to minimize errors.

In non-linear perceptron model and delta learning method the output is computed using the equation

$$y = \varphi\left(\sum_{i=1}^{n} \omega_i x_i + b\right) = \varphi\left(w^T x + b\right) \tag{2}$$

where w denotes the vector of weights, x is the vector of inputs, ω_i is a weighing factor, b is the bias and φ is the activation function. Generally in non-linear perceptron the activation function φ is considered as logistic sigmoid:

$$\varphi = \frac{1}{1+\exp(-x)} \tag{3}$$

In non-linear perceptron, error in each set of test sample is computed as

$$E = |y_a - y_p| \tag{4}$$

where E denotes the error, y_a is the actual value of the output and y_p is the predicted value.

The weights are updated as:

$$w_{new} = w_{old} + E_i \times X_i' \tag{5}$$

where E_i is the ith error and X_i' is the inverse of ith input.

In delta learning method, that is, in the linear feed forward network, the error in each set of test sample is computed as

$$E = \frac{1}{2}(y_a - y_p)^2 \tag{6}$$

The weights are updated as

$$w_{new} = w_{old} + \mu(y_a - y_p) f'(w_i x_i) \times x_i \tag{7}$$

where x_i is the ith input, y_a is the actual value of the output and y_p is the predicted value, μ is the learning rate generally taken as 0.9 (Pal and Mitra, 1999) and $f'(x_i)$ is the activation function φ. The quality of the prediction is obtained from the performance with the test set of data. Percentage of errors of prediction (PE) is computed as

$$\mathrm{PE} = \frac{\langle |y_a - y_p| \rangle}{\langle y_a \rangle} \tag{8}$$

where < > implies average overall inputs for every input data sets (Patterson, 1996).

Result and Discussion

The ANN model forecast with non-linear perceptron and delta learning is compared for the concentration of different pollutants over Delhi (Table 1). It is apparent from the table that the per cent error between the target and the ANN model forecast is very low with non-linear perceptron for the concentration of sulphur dioxide, ozone, carbon monoxide, and suspended particulate matter, whereas the per cent error between the target and the ANN model forecast is quite high with non-linear perceptron for nitrogen dioxide. However, the per cent error is low with delta learning for nitrogen dioxide.

Table 1: The percentage error and number of epochs obtained for forecasting the concentration of different pollutants with ANN model for different learning rules

Pollutants	*Learning rules*	*Percentage error*	*Number of iterations*
Sulphur dioxide (SO_2)	Multi-layer perceptron	0.074	30
	Delta learning	22.5	12
Ozone (O_3)	Multi-layer perceptron	7.7	3
	Delta learning	21	14
Carbon monoxide (CO)	Multi-layer perceptron	1	50
	Delta learning	42.7	16
Suspended particulate matter (SPM)	Multi-layer perceptron	11.2	16
	Delta learning	28.3	6
Nitrogen dioxide (NO_2)	Multi-layer perceptron	77	28
	Delta learning	25.2	30

The difference between the target and the ANN model forecast is very low for the concentration of SO_2 over Delhi for the year 2009 with non-linear perceptron (Figure 1). However, the difference is quite high with delta learning (Figure 2). Moreover, the concentration of SO_2 is observed to be highest in the month of April (pre monsoon month) and quite high in the months of October and November (post monsoon months) whereas it is comparatively low during the monsoon (May–June) (Figure 1). The pre monsoon and post monsoon in India are prone to severe thunderstorms associated with huge cumulonimbus clouds. Sulphur dioxide acts as pollutant but when it is oxidized, the sulphur compounds produce non-sea-salt sulphate (n.s.s.) (SO_4^{2-}) aerosols (Wigley, 1989), which acts as cloud condensation nuclei (CCN). Changes in the flux of SO_2 into the atmosphere can change the concentration of CCN (Schwartz, 1988). The coalescence process may help the CCN to lead to the formation of cumulonimbus clouds. It may thus be surmised that the increase in the flux of SO_2 into the atmosphere during pre and post monsoon period may be instrumental for the genesis of thunderstorms if other meteorological conditions are also favourable. The concentration of SO_2 is low during the monsoon season. The reason behind this may be that the monsoon brings significant amount of precipitation, high wind velocities and changes in the wind direction. The precipitation helps to washout the pollution and the high wind velocities allow the pollutants to move away from the source, increase mixing processes and hence decrease the concentration of pollutants (Aneja, 2001). The percentage error of sulphur dioxide is observed to decrease continuously using non-linear perceptron and finally it acquires almost a constant value after twenty-five iterations (Figure 3). The decrease in percentage error is less with delta learning and it becomes stable after twelve iterations (Figure 4).

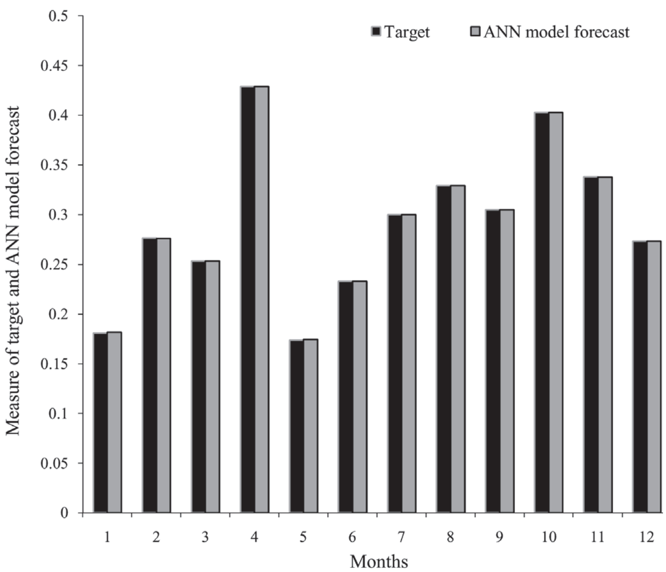

Figure 1: The target and the ANN model forecast with non-linear perceptron in forecasting the concentration of sulphur dioxide (SO_2) over Delhi for the year 2009.

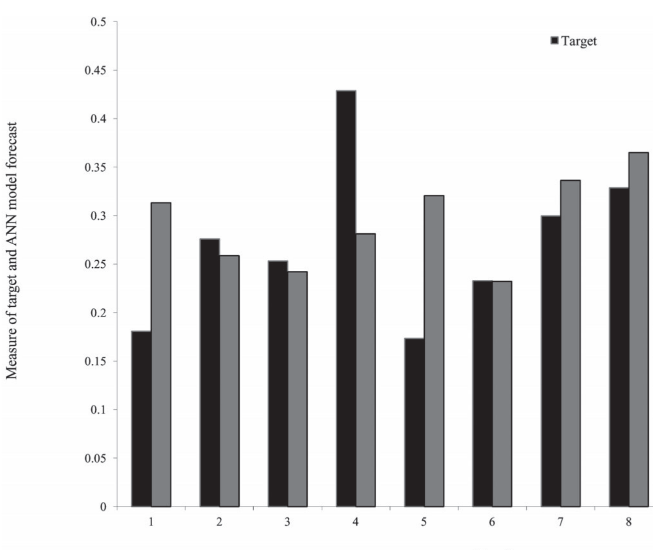

Figure 2: The target and the ANN model forecast with delta learning rule in forecasting the concentration of sulphur dioxide (SO_2) over Delhi for the year 2009.

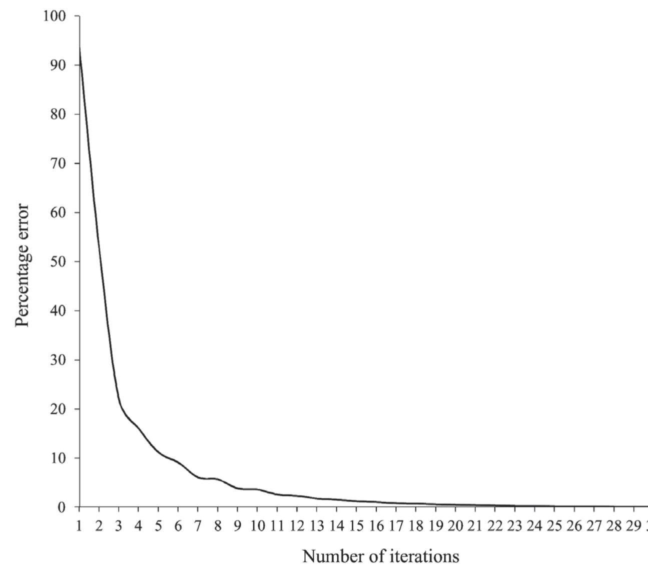

Figure 3: The percentage error and the number of iterations in forecasting the concentration of sulphur dioxide (SO_2) using ANN model with non-linear perceptron.

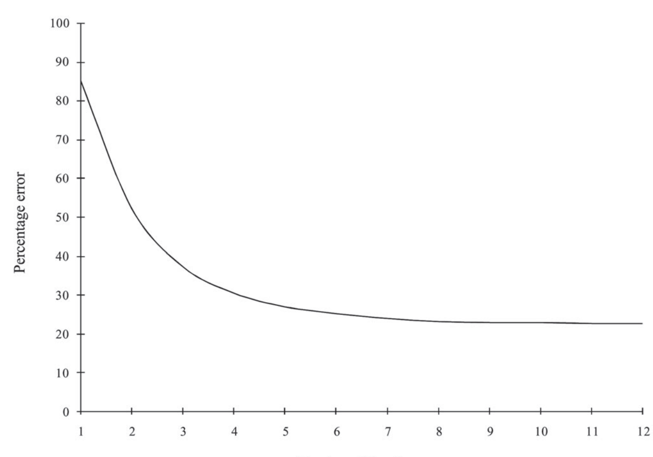

Figure 4: The percentage error and the number of iterations in forecasting the concentration of sulphur dioxide (SO_2) using ANN model with delta learning rule.

The difference between target and ANN model forecast is low for the concentration of ozone using non-linear perceptron (Figure 5) and is high when delta learning is used (Figure 6). The concentration of ozone attains an average value throughout the year; however, the minimum is observed in the months of November and December (post monsoon) and maximum is observed in the month of April (pre monsoon) over Delhi (Figure 5). The maximum concentration during the pre monsoon time can be explained on the basis of large photochemical ozone production due to high concentration of pollution during the period. The minimum concentration of ozone during the monsoon may be due to the non-availability of sufficient solar radiation and washout of pollutants as well as consumption of O_3 by HO_x radicals (Jain et al., 2005). The percentage error in forecasting the concentration of ozone with ANN model is observed to be less with non-linear perceptron (Figure 7) than delta learning method (Figure 8).

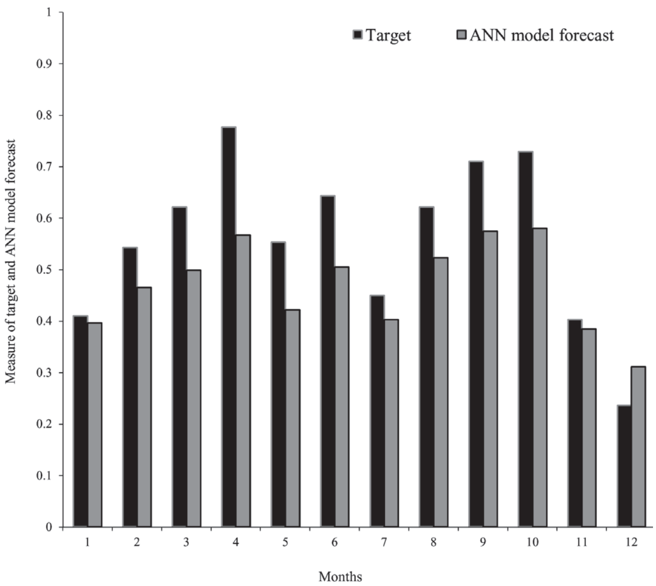

Figure 5: The target and the ANN model forecast with non-linear perceptron in forecasting the concentration of ozone (O_3) over Delhi for the year 2009.

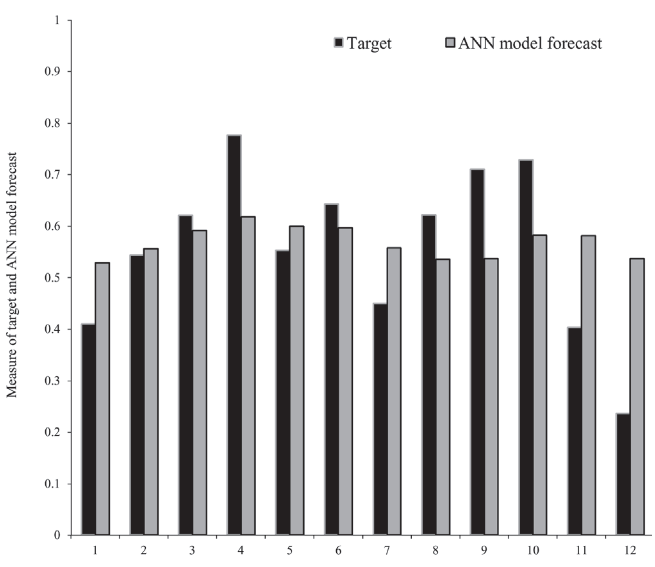

Figure 6: The target and ANN model forecast with delta learning in forecasting the concentration of ozone (O_3) over Delhi for the year 2009.

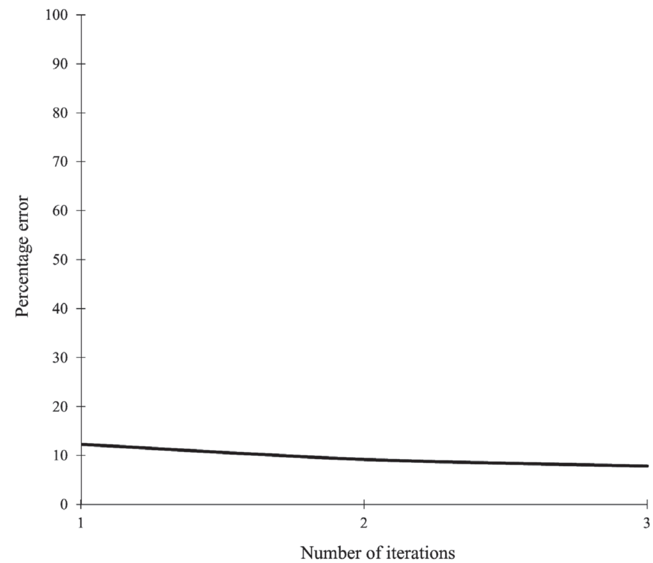

Figure 7: The percentage error and the number of iterations in forecasting the concentration of ozone (O_3) using ANN model with non-linear perceptron.

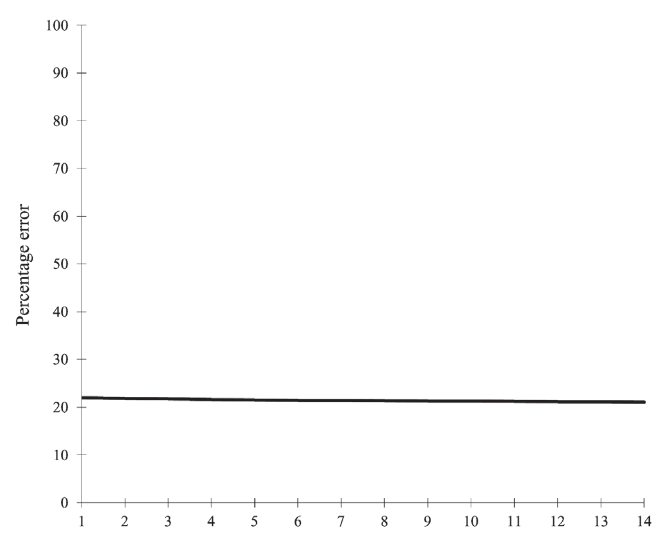

Figure 8: The percentage error and the number of iterations in forecasting the concentration of ozone (O_3) using ANN model with delta learning rule.

The difference between the target and ANN model forecast is observed to be very low for the concentration of CO over Delhi using non-linear perceptron (Figure 9) but the difference is quite high when delta learning rule is used (Figure 10). The concentration of carbon monoxide is observed to be high in the months of November and December and it is minimum in the month of February (Figure 9). The percentage error is observed to decrease continuously with non-linear perceptron and becomes almost constant after fifty iterations (Figure 11). However, while the delta learning is used it is observed that the percentage error decreases sharply initially but quickly it becomes stable after sixteen iterations (Figure 12).

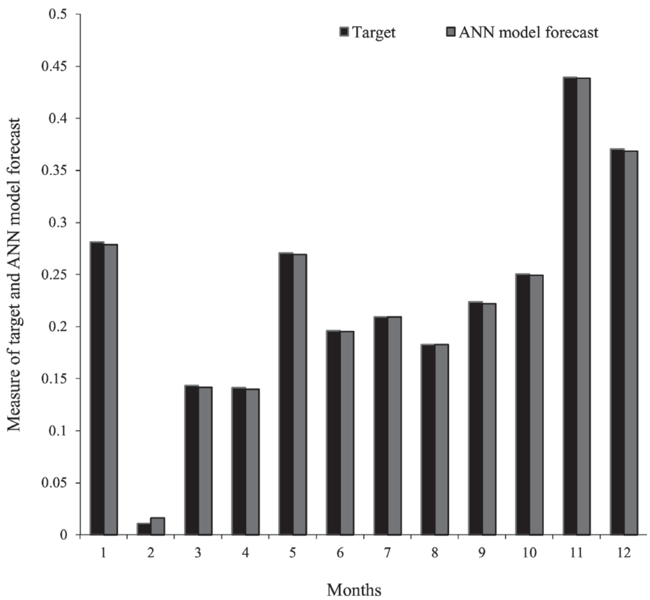

Figure 9: The target and the ANN model forecast with non-linear perceptron in forecasting the concentration of carbon monoxide (CO) over Delhi for the year 2009.

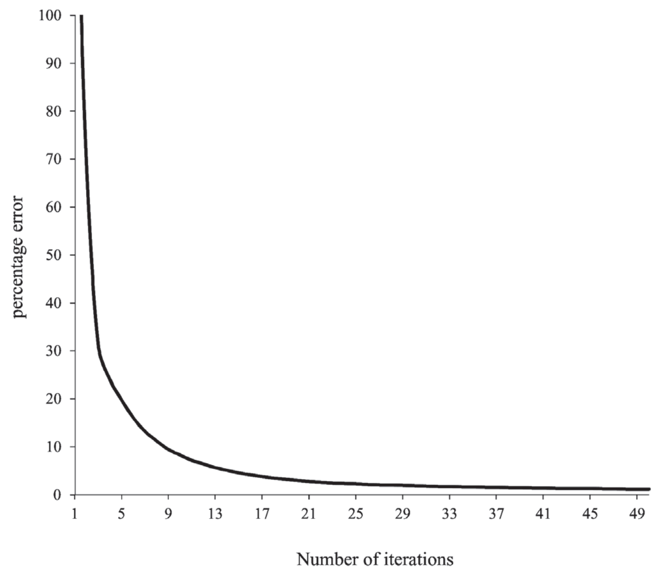

Figure 11: The percentage error and the number of iterations in forecasting the concentration of carbon monoxide using ANN model with non-linear perceptron.

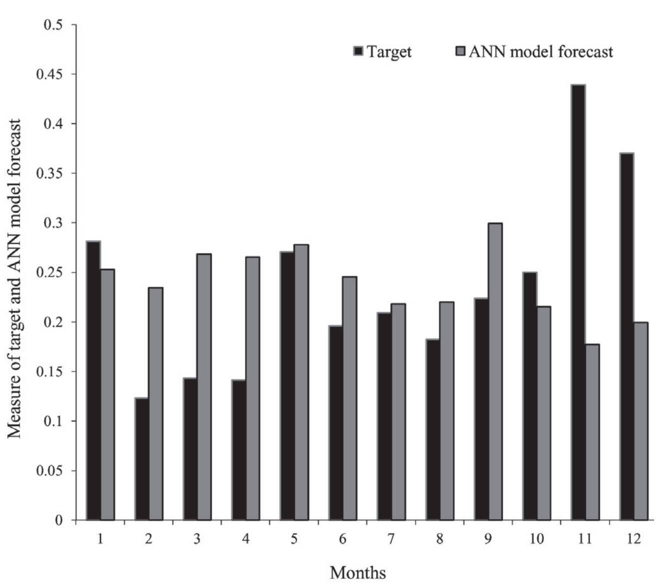

Figure 10: The target and the ANN model forecast with delta learning rule in forecasting the concentration of carbon monoxide (CO) over Delhi for the year 2009.

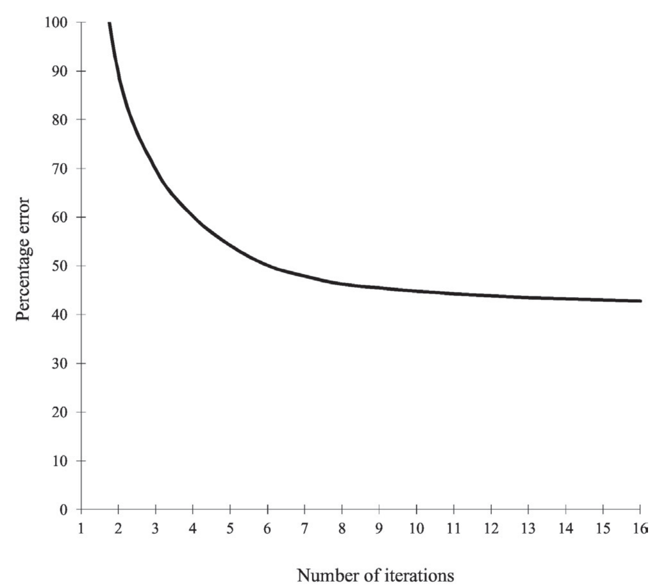

Figure 12: The percentage error and the number of iterations in forecasting the concentration of carbon monoxide (CO) using ANN model with delta learning.

The difference between the target and ANN model forecast is less with non-linear perceptron for suspended SPM (Figure 13) while it is high using delta learning (Figure 14). The concentration of SPM is quite high during the pre monsoon (April–May) and post monsoon (November–February). SPM also supports CCN formation. It can thus be stated that the increase in the concentration of SPM may lead to the formation of thunderclouds. The percentage error with non-linear perceptron shows a chaotic pattern (Figure 15). The error first decreases and then increases and again decreases and finally shows a decreasing tendency after sixteen iterations. However, when delta learning is used (Figure 16), the error is observed to decrease sharply initially and then slightly increases and finally becomes stable.

The situation is completely different for the concentration of nitrogen dioxide. The difference between the target and ANN model forecast is very high when non-linear perceptron (Figure 17) is used whereas it is comparatively low using delta learning method (Figure 18). The concentration is observed to be highest during post monsoon to winter (October to February). The region of study during the said period is dominated by high pressure usually centred over western China causing increased atmospheric stability which in turn allows for less general circulation and thus more stagnant air masses which allows accumulation of pollutants. The atmospheric dispersion is minimum and therefore the pollutants cannot be widely dispersed throughout the planetary boundary level (PBL). However, during the pre monsoon to summer the average PBL height is high resulting in increased mixing through a greater volume of the troposphere and hence lower concentration of pollutants.

No stability is observed in the variation of percentage error in forecasting the concentration of nitrogen dioxide with every epoch using non-linear perceptron (Figure 19). Initially the error is observed to increase, then

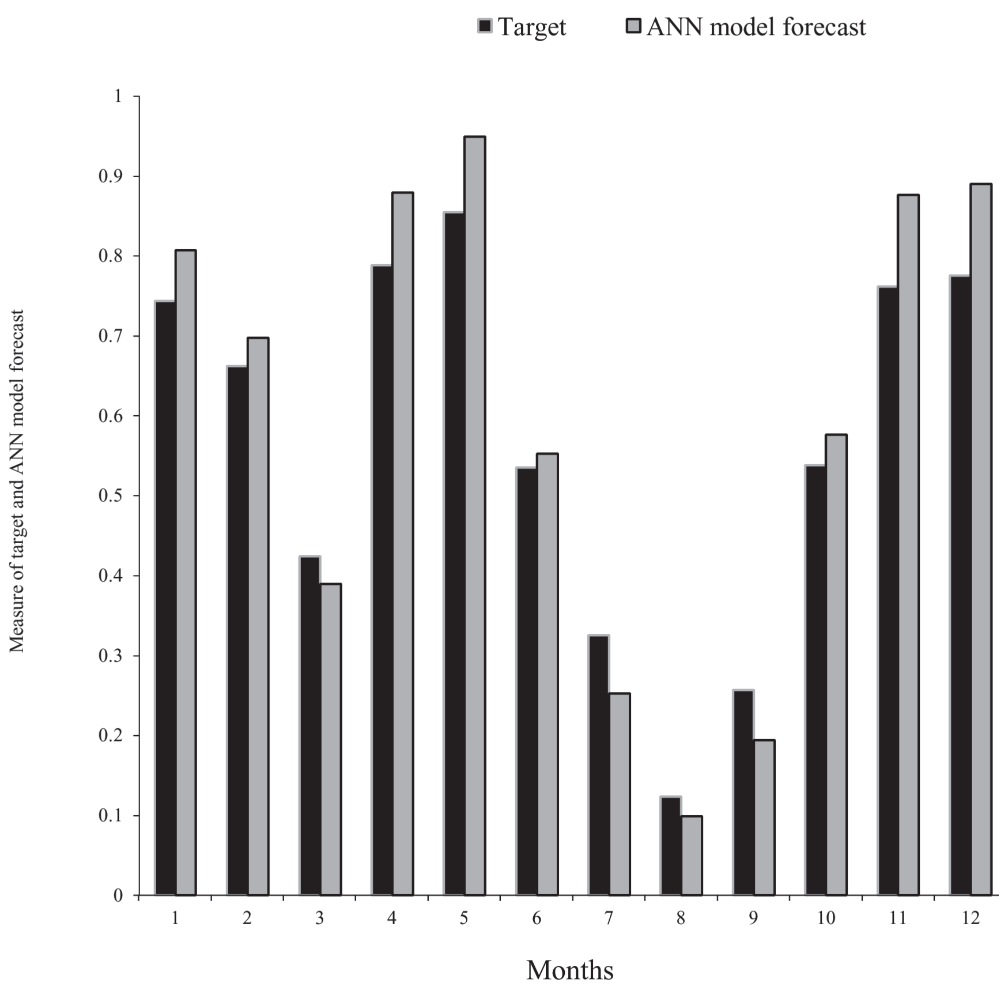

Figure 13: The target and the ANN model forecast with non-linear perceptron in forecasting the concentration of suspended particulate matter (SPM) over Delhi for the year 2009.

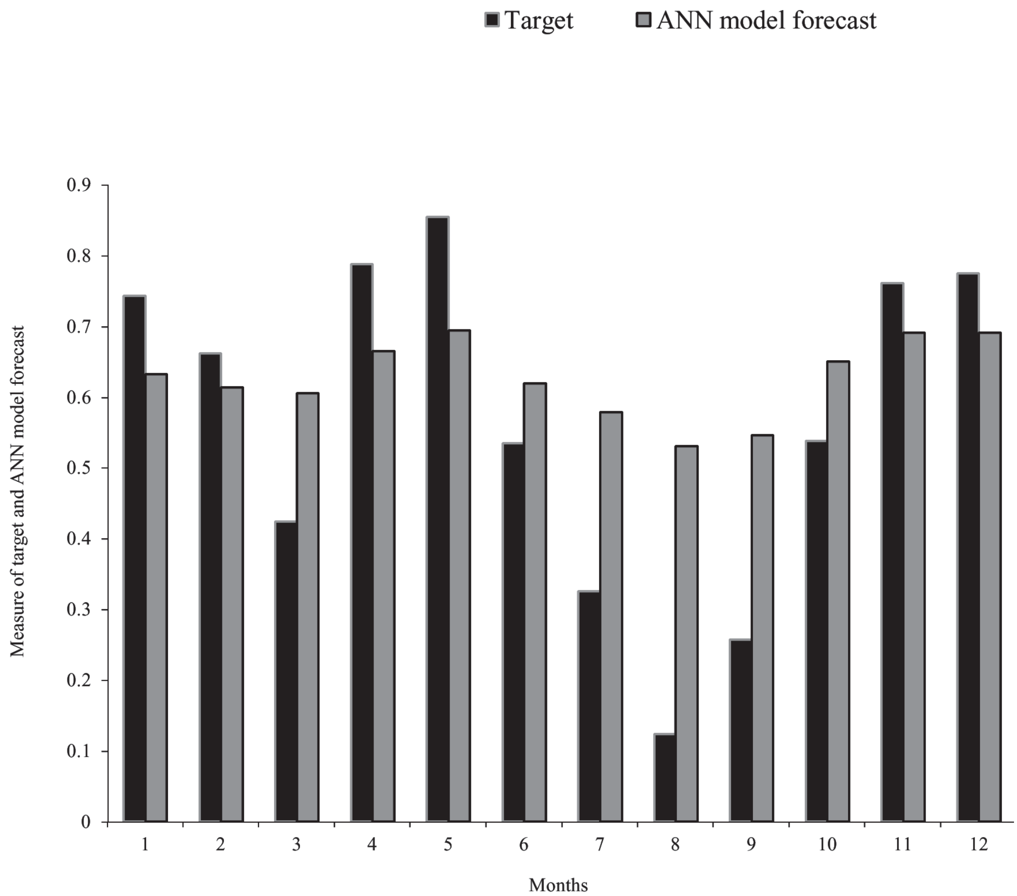

Figure 14: The target and the ANN model forecast with delta learning in forecasting the concentration of suspended particulate matter (SPM) over Delhi for the year 2009.

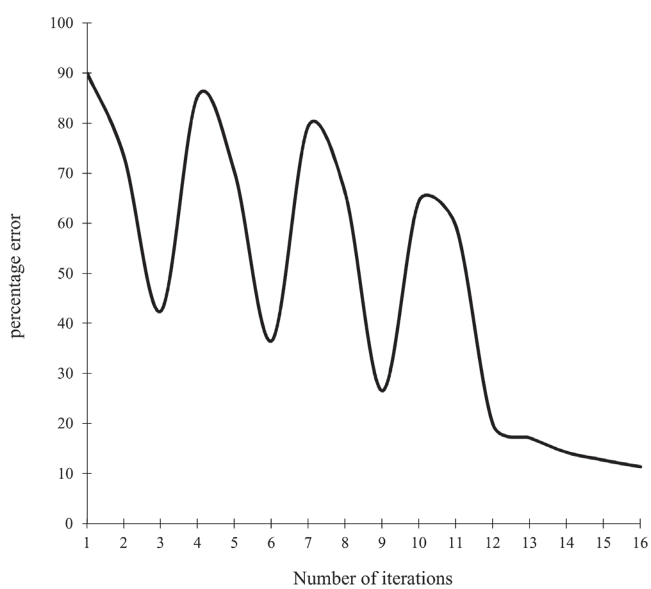

Figure 15: The percentage error and the number of iterations in forecasting the concentration of suspended particulate matter (SPM) using ANN model with non-linear perceptron.

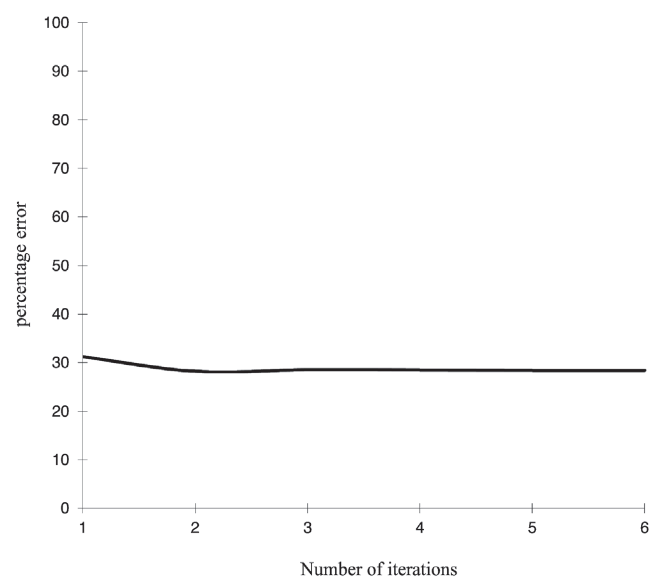

Figure 16: The percentage error and the number of iterations in forecasting the concentration of suspended particulate matter (SPM) using ANN model with delta learning.

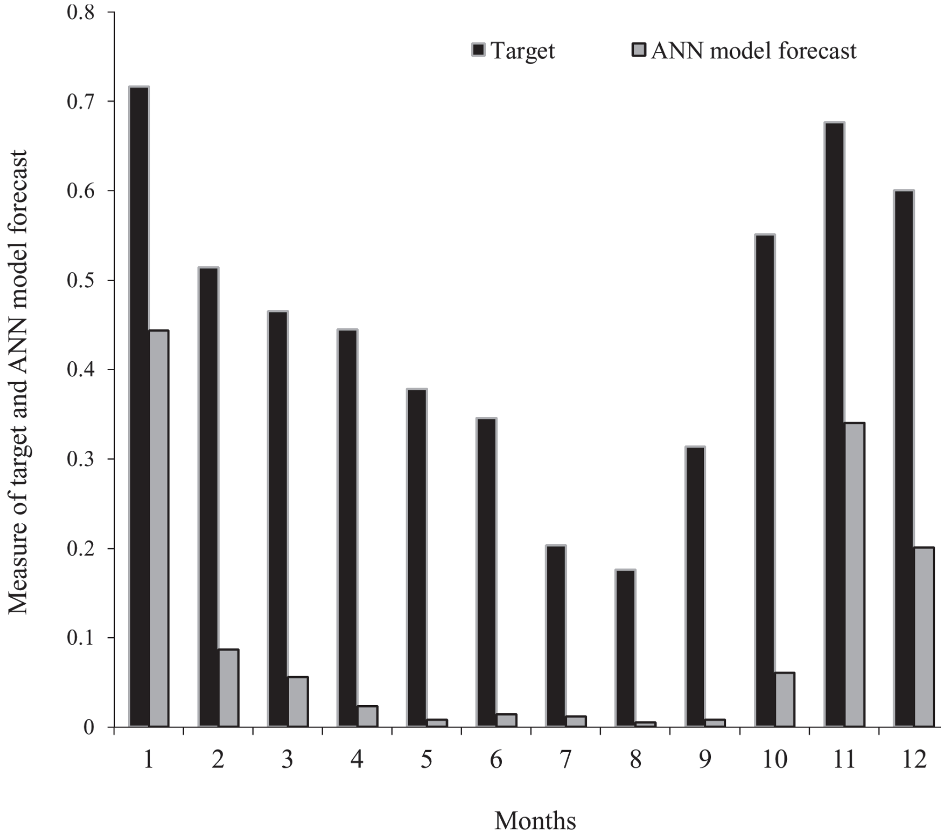

Figure 17: The target and ANN model forecast with non-linear perceptron in forecasting the concentration of nitrogen dioxide (NO_2) over Delhi for the year 2009.

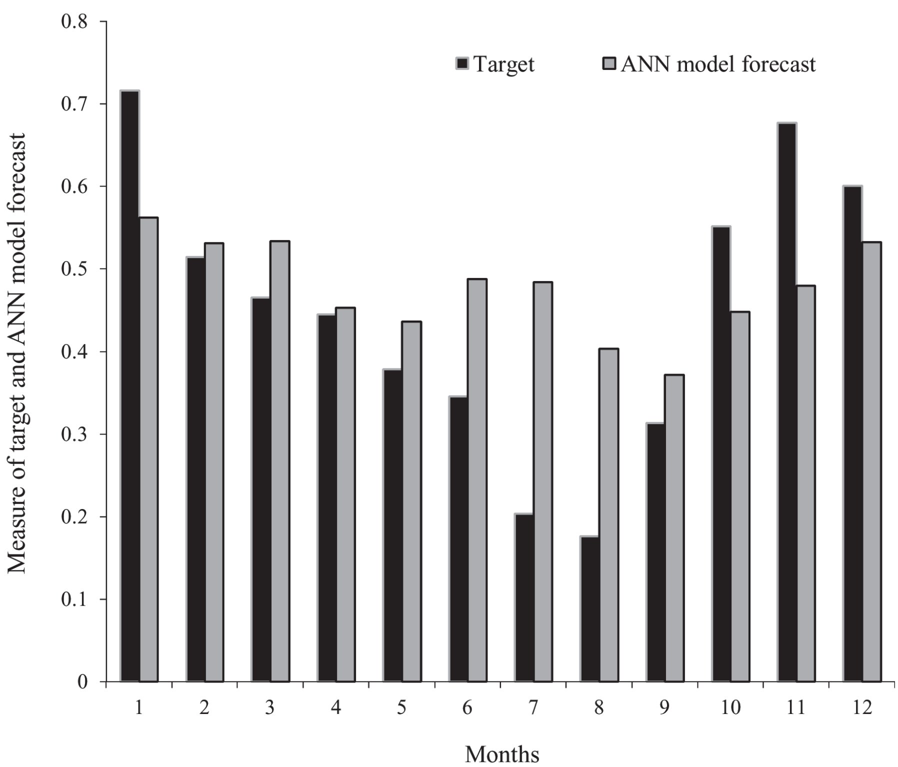

Figure 18: The target and ANN model forecast with delta learning in forecasting the concentration of nitrogen dioxide (NO_2) over Delhi for the year 2009.

decreases and so on till twenty iterations but when delta learning is used the percentage error is observed to decrease slightly at first and then it becomes stable just after four iterations (Figure 20).

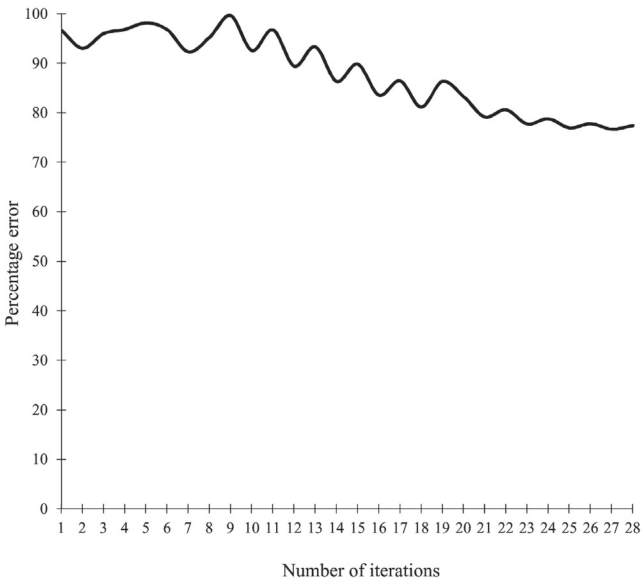

Figure 19: The percentage error and the number of iterations in forecasting the concentration of nitrogen dioxide (NO_2) using ANN model with non-linear perceptron.

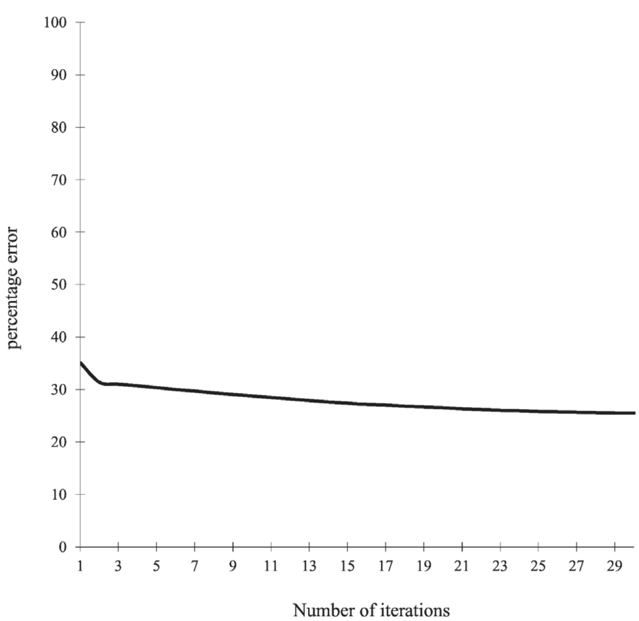

Figure 20: The percentage error and the number of iterations in forecasting the concentration of suspended particulate matter (SPM) using ANN model with delta learning.

Conclusion

ANN model is found to be suitable for forecasting the concentration of pollutant over Delhi. It is evident from the study that the non-linear perceptron is suitable for forecasting the concentration of different pollutants over Delhi, except nitrogen dioxide which can be predicted more accurately by delta learning method. In delta learning method small changes in the concentration with small time interval is considered. It can thus be stated that only those reactions will be supported by this method which are generally a first order reaction. In first order reaction the product obtained does not readily participate in any other side reaction or the product follows a linear relationship with the constituents. The multi-layer perceptron method is highly used to establish a link or a suitable pattern between the products and the constituents or in other words between the target and the output. Thus, it may be concluded that those reactions which are generally multi-order reaction or follows a non-linear relationship between product and constituent will be supported by non-linear perceptron.

NO_2 can thus be predicted more accurately by delta learning method because it is a quite stable compound in normal condition, whereas the perceptron model is suitable for forecasting the concentration of other pollutants as they generally follow a non-linear behaviour.

Acknowledgements

The authors thank the Central Pollution Control Board (CPCB) and India Meteorological Department (IMD), Government of India for making the data available for users in their respective website.

References

Aneja, V.P., Agarwal, A., Roelle, A.P., Phillips, S.B., Tong, Q., Watkinson, N. and Richard Yablonsky (2001). Measurements and analysis of criteria pollutants in New Delhi, India. *Environment International*, **27:** 35–42.

Annand, W.J.D. and A.M. Hudson (1981). Meteorological effects on smoke and sulphur dioxide concentrations in the Manchester area. *Atmos. Environ*., **115:** 799–806.

Chaudhuri, S. (2006). Artificial Neural Network Model to forecast maximum wind speed associated with severe thunderstorms. *Vatabaran*, *AFAC J. of Meteorol.*, **30(1):** 14–19.

Chaudhuri, S. (2007). Artificial Neural Network Model to Forecast Ozone Concentration over Kolkata. *Sci & Cult.*, **73:** 9–10, 303–307.

Comrie, A.C. (1997). Comparing neural network and regression models for RSPM forecasting. *Journal of Air and Waste Management Association*, **47:** 653–663.

Díaz-Robles, L.A., Ortega, J.C., Fu, J.S., Reed, G.D., Chow, J.C., Watson. J.G. and J.A. Moncada Herrera (2008). A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile, **42(35):** 8331–8340.

Elsom, D.M. and T.J. Chandler (1978). Meteorological controls upon ground level smoke and sulphur dioxide in two urban areas of the United Kingdom. *Atmos. Environ.*, **12:** 1543–1554.

Hagan, Martin T., Demuth Howard, B. and H. Beale Mark (1996). ISBN: 0-9717321-0-8. *Neural Network Design*, **25:** 2–6.

Hornik, K. (1991). Approximation capabilities of multi-layer neural networks. *Neural Networks*, **4:** 251–257.

Jacobson, M.Z. (1998). Fundamentals of Atmospheric Modelling, Cambridge University Press, NewYork, p. 656.

Jain, S.L., Arya, B.C., Kumar Arun, Ghude, S.D. and P.S. Kulkarni (2005). Observational study of surface ozone at New Delhi, India. *International Journal of Remote Sensing*, **26:** 3515–3524.

Kampa, M. and Castanas (2008). Human health effects of air pollution. *Environ. Pollution*, **151:** 362–367.

Pal, S.K. and S. Mitra (1999). Neuro–Fuzzy Pattern Recognition. Wiley Interscience Publication, New York.

Patterson, D.W. (1996). Artificial Neural Networks: Theory and Applications. Prentice Hall, Englewoods Cliffs.

Perez, P., Trier, A. and J. Reyes (2000). Prediction of $PM_{2.5}$ concentrations several hours in advance using neural networks in Santiago, Chile. *Atmospheric Environment*, **34:** 1189–1196.

Rajaskaran, S. and G.A. Vijayalakshmi Pai (2003). Neural Networks, Fuzzy Logic and Genetic algorithms. *In:* Synthesis and Applications, Prentice Hall, New Delhi.

Sang Hyun Sohn, Sea Cheon Oh and Yeong-Koo Yeo (1999). Prediction of air pollutants by using an artificial neural network. Dept. of Chem. Eng., Hanyang University, Seoul, Korea, 133–791.

Schwartz, S.E. (1988). Are global cloud albedo and climate controlled by marine phytoplankton? *Nature*, **336:** 441–445.

Wigley, T.M.L. (1989). Possible climate change due to SO_2-derived cloud condensation nuclei. *Nature*, **339:** 365–367.

Asian Journal of Water, Environment and Pollution

Volume 8 Number 3 **July – September, 2011**

Contents