IRACST – Engineering Science and Technology: An International Journal (ESTIJ), ISSN: 2250-3498,
Vol.3, No.4, August 2013

Modelling of Respirable Suspended Particulate Matter Concentration Using Artificial Neural Networks in an Urban Area

Sonaje N P
Deputy Registrar
Shivaji University, Kolhapur
nitinsonaje@yahoo.co.in

Mane S J
Assistant Professor
Department of Civil Engineering
D.Y. Patil College of Engg. Akurdi, Pune

Kote A S
Professor
Department of Civil Engineering
D.Y. Patil College of Engg. Akurdi, Pune

***Abstract*- This paper describes the development of an artificial neural network (ANN) model for forecasting respirable suspended particulate matter (RSPM) concentration in a major urban area Pune Maharashtra. This study is targeted because the small particles affect the health of people and it constitutes the major concern for air quality of urban area. The data set consisting of RSPM concentrations for 6 yrs (2004-2010) used in the study was collected from Maharashtra pollution control board (MPCB) monitoring station. In order to provide with an operational air quality forecasting module for RSPM, ANN models based on a three-layered feed forward neural network trained by back-propagation algorithm were investigated and applied. The developed model efficiently models the trend of RSPM concentrations (R=0.902). The results obtained from the proposed models showed that ANN can be efficiently used in forecasting of air quality.**

Keywords- Artificial neural network, Respirable suspended particulate matter, Multilayer perceptron, Generalized feed forward network.

I. INTRODUCTION

In India many epidemiological studies have demonstrated that chronic non-communicable diseases such as cancer, cardiovascular diseases and respiratory disorders are more dominant than old infectious diseases [1]. The pattern of economic growth that we are adopting is increasingly associated with environmental pollution [2]. It is clear that in the developing or developed countries increased mechanization, transportation, populations, etc. causes air pollution which is a major and effective environmental problem. Study showed that due to air pollution there is reduction in pulmonary function and increased respiratory symptoms [3-4]. The increased sensibility of populations to environmental problems has obliged the National administration to form laws that govern whether the pollutant emission of industries for the maximum values of pollutant’s concentration on the ground. An accurate pollutant concentrations forecasting by means of time and location is necessary for designing an adequate health warning system. On this basis, computational methods that allow for domain modelling and quality of life parameter forecasting are becoming more and more important.

In recent years the trend has been to use statistical methods instead of traditional deterministic methods to forecast air pollution. ANN models have been used for forecasting of a wide range of pollutants and their concentrations at various time scales, with very good results [5]. Many researchers have successfully used ANN models for prediction of SO_2[6-7] NO_2[8] Ozone[9], and particulate matter level[10-11] and found that the ANN model is suitable for predicting air pollutant. In this study, ANN models with back propagation algorithm are used to forecast respirable suspended particulate matter (RSPM) concentration observed in Pune area using historical data.

II. MATERIAL AND METHOD

The potential of three-layered feed forward neural network trained by back-propagation algorithm in forecasting RSPM concentration has been illustrated by developing seasonal time series model for Pune city. Pune city located in the western region of Maharashtra, India at 18°25'0"N and 73°44'0"E (Figure 1) has a population of about 37 lakhs. It has large industrial area covering nearly 10000 industries that include TATA, Bajaj, Thermax, L&T, etc. An adverse consequence of the rapid rise in population and industrialization is the increase in environmental degradation, particularly air quality in Pune city. According to national ambient air quality standards the limit for RSPM is 100 µg/m^3, however some observed values exceed the limit. The maximum concentration of RSPM observed are 332 µg/m^3 and 300 µg/m^3 which is dangerous to human health. Thus prediction of air quality is highly necessary.

The daily data of RSPM from September 2004-December2010 *ie* 6yrs has been obtained for Pune from MPCB [12]. The daily RSPM concentration consisting of 2285 data points are used for modelling.

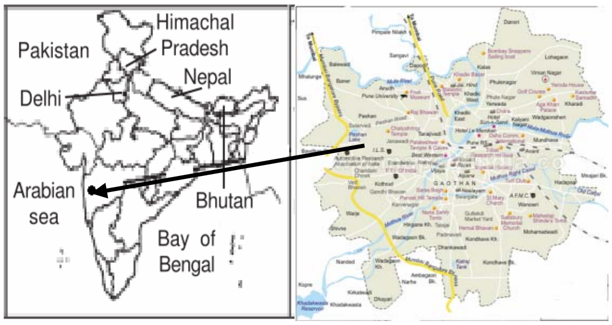

Figure 1. Location map of Pune City

A. Artificial Neural Network Model

The detailed coverage of ANN can be found in many books (13,14). ANN is a powerful machine learning method widely used in the problems of numerical prediction and classification. The network is made up of number of interconnected nodes (processing elements), arranged into three basic layers: input, hidden, and output. (Figure 2) The links represent weighted connections between the nodes. A processing element simply multiplies input by a set of weights, and linearly or non-linearly transforms the result into an output value. By adapting its weight, the neural network works towards generating an output that would be close to the measured (target) output. Trial and error procedure would be adopted with respect to input determination; pre-processing, percentage training and testing data sets; network architecture; algorithms; learning rules; learning rates; and training cycles.

Multilayer perceptrons (MLP) are layered feedforward networks typically trained with static back-propagation. These networks have found their way into countless applications requiring static pattern classification. Their main advantage is that they are easy to use, and that they can approximate any input/output map. The key disadvantages are that they train slowly, and require lots of training data (typically three times more training samples than network weights).

Generalized feedforward networks (GFN) are a generalization of the MLP such that connections can jump over one or more layers. In theory, a MLP can solve any problem that a generalized feedforward network can solve. In practice, however, generalized feedforward networks often solve the problem much more efficiently. The methodology adopted in developing the ANN model in the present study is shown in Figure 3.

The model used for predicting RSPM values of current day using four previous values of RSPM represented as follows.

RSPM(t+1)=f[RSPM(t), RSPM(t-1), RSPM(t-2), RSPM(t-3)]

Where ‘t’ is day for which prediction is done.

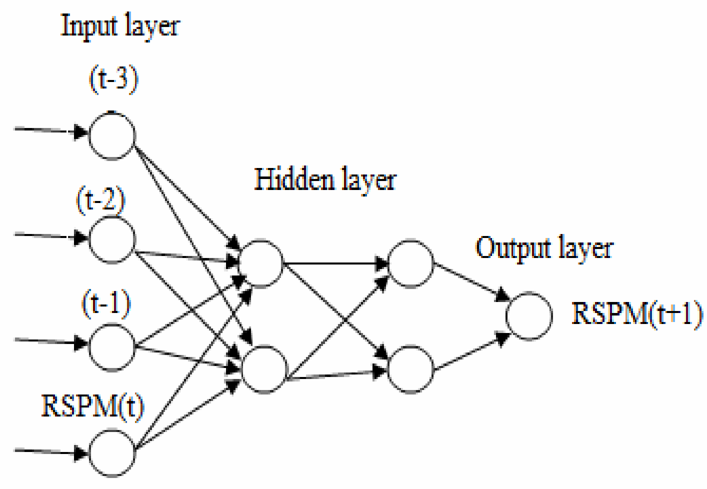

Figure 2. A typical neural network architecture

Out of large number of ANN models developed for the data sets, best ANN model is investigated in detail. The goodness-of-fit measures considered in the present study to evaluate the developed models are mean square error (MSE), mean absolute error (MAE), and coefficient of correlation (R) between the forecasted and observed RSPM concentration. The MAE and MSE are good measures for indicating the goodness-of-fit at moderate and high output values, respectively and the values equal to zero indicates perfect fit. The R value quantifies the efficiency of a model in capturing the complex, dynamic and non-linear nature of the physical process being modeled and the value equal to one shows perfect fit.

In this study two network *viz* multilayer perceptron (MLP) and generalized feed forward neural network (GFN) was used as per the methodology given in figure 3.

III. RESULTS AND DISCUSSION

After a number of trials by changing the length of data set for training, testing and also the number of neuron in the hidden layer the maximum value of ‘R’ was obtained. The MLP neural network consists of TanAxon transfer function and momentum as learning rule where the value of threshold was taken as 0.001. The number of epochs was set to 1000 throughout the trial and error process. The momentum training algorithm was adopted as it is a faster algorithm. After carrying out trial and error for different networks (4-2-1) the neural network architecture was finalized *i.e.* four neurons in input layer, two hidden layers and one neuron in output layer. The maximum value of R is 0.902, MSE=605.454 $(\mu g/m^3)^2$ and MAE=17.45 $\mu g/m^3$ as shown in Table1. As a factor of highest accuracy this value of ‘R’ has to be 0.95 or even more to achieve the highest accuracy. In this case the value of 'R' obtained is quite high and reliable for short term prediction. Numbers of ANN models were developed with MLP and GFN. The hidden layers were varied from 1 to 3. The training and testing data set were varied as shown in the Table1 in order to achieve better accuracy.

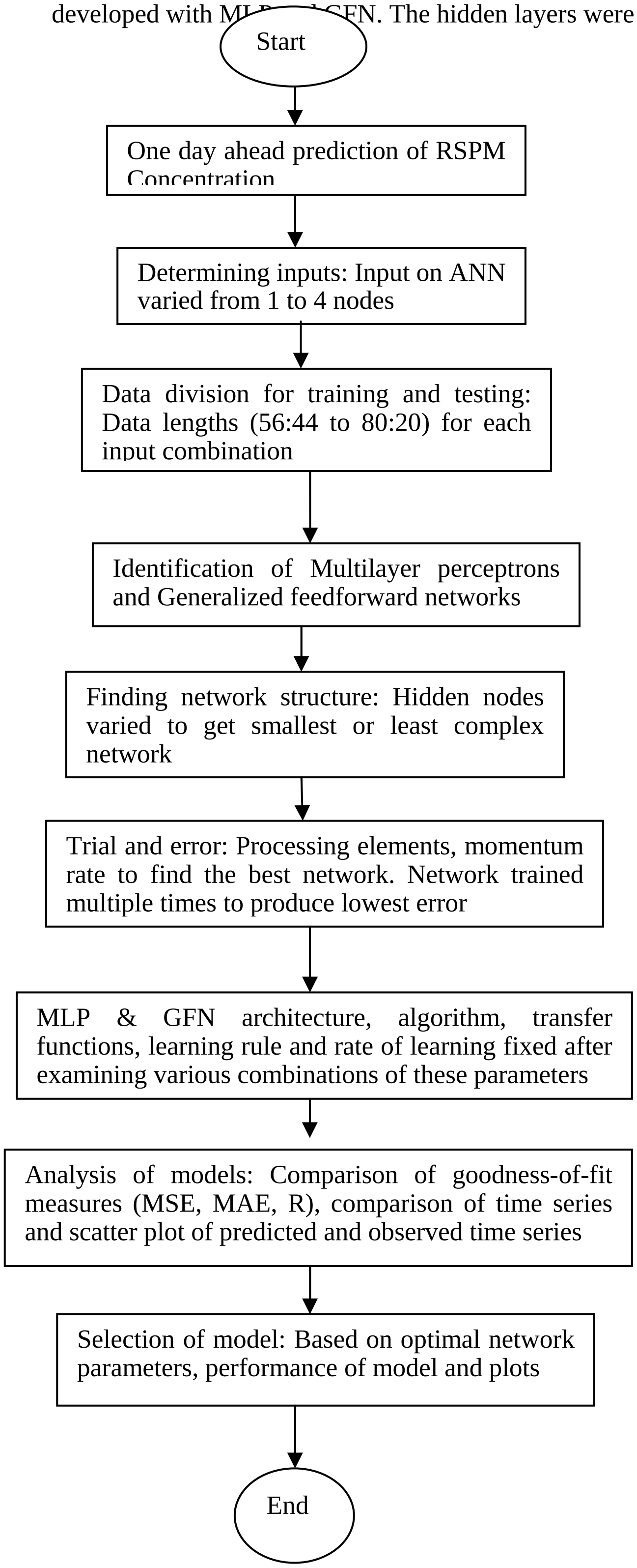

Figure 3 Flow chart of methodology

The graph of the model output verses desired output as seen in Figure 4 and Figure 5 showed that most of the forecasted values of RSPM were close to the expected values, which indicate that the modelling has been quite successful

IV. CONCLUSION AND FUTURE SCOPE

The use of MLP method of ANN for short term instantaneous prediction of RSPM values provided reliable result, although the forecasting has been done for real source. The value of R (0.902) is highly encouraging i.e. 91% accuracy in forecasting. This method has been applied for RSPM values for one station only. Whether it is applicable to other stations and forecasting of other parameters (NO_X and SO_X) has to be tested and verified. The concept of time series has been assumed to follow a certain pattern. However the effect of humidity, temperature on RSPM values needs to be investigated

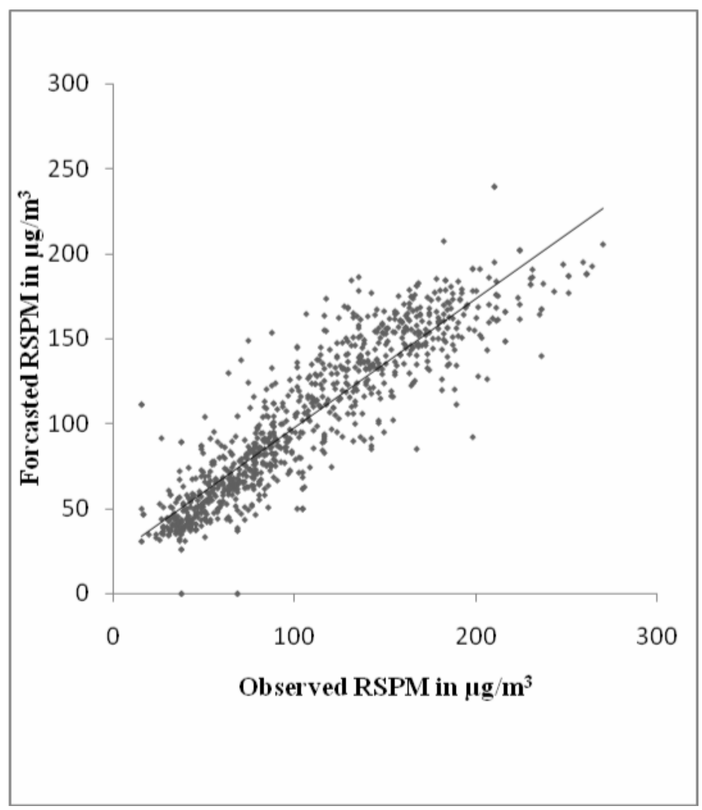

Figure 4 Scatter plot MLP (4-2-1) model

TABLE 1. SUMMERY OF ANN MODELS

Model	Training Data%	Testing Data %	Epochs	R	MSE	MAE
MLP (4-2-1)	70	30	1000	0.881	691.506	18.088
	60	40	1000	**0.902**	605.454	17.450
	58	42	1000	0.833	932.335	19.134
	56	44	1000	0.879	683.672	17.756
GFN (4-2-1)	75	25	1000	0.855	908.289	20.536
	70	30	1000	0.701	1815.925	24.045
	65	35	1000	0.668	2599.384	26.757
	60	40	1000	0.759	1479.2179	21.773

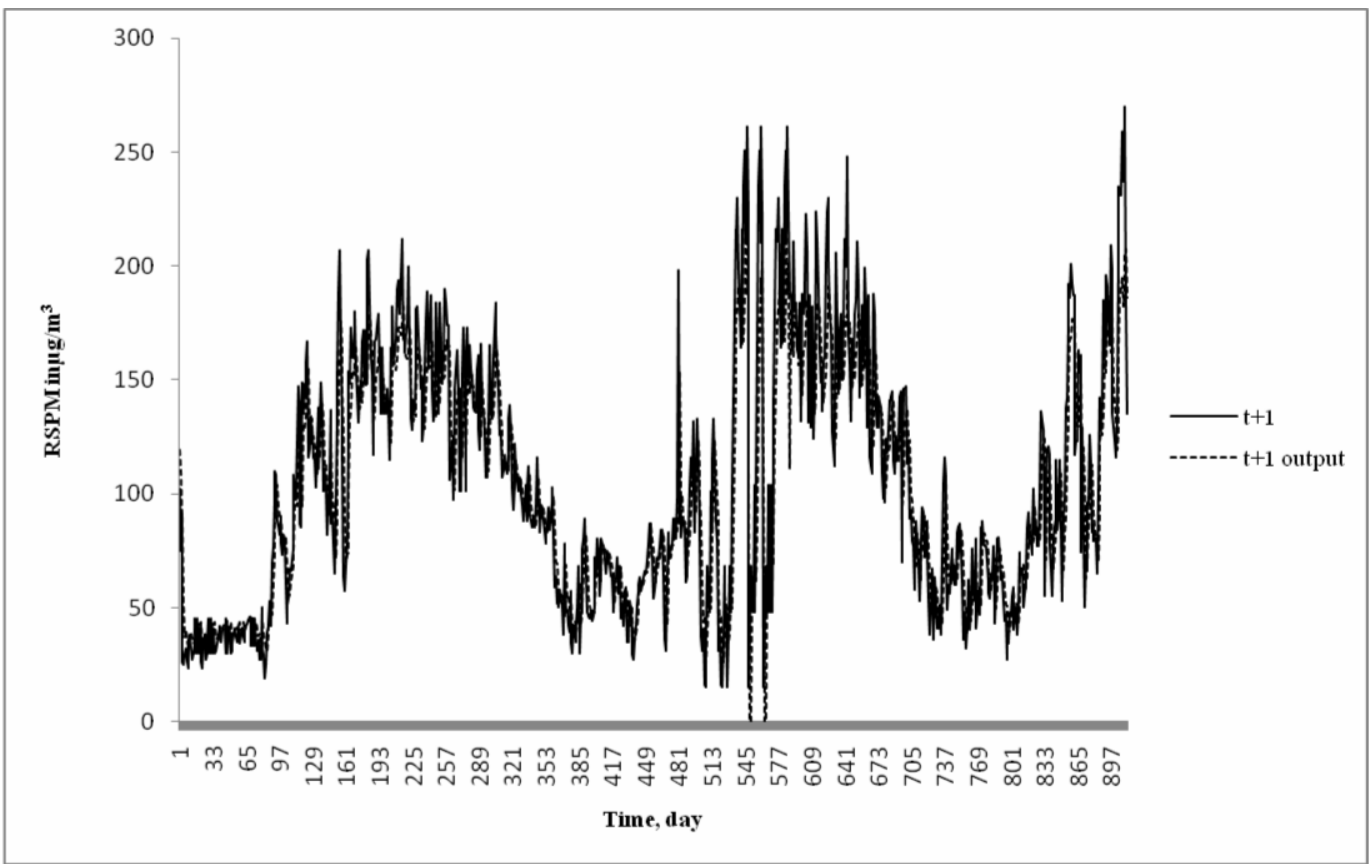

Figure 5 Time series MLP (4-2-1) model

REFERENCES

[1] Bacharova L., Fandakova K., Bratinka, J., Budinska M., Bachar J., and Gudaba, M., “The Association Between Air Pollution and the Daily Number of Deaths: Findings from the Slovak Republic Contribution to the APHEA Project, "Journal of Epidemiology Community Health, vol. 50(S1), pp. S19-S21, 1996.

[2] Lvovsky K., “Economic Costs of Air Pollution with Special Reference to India” National Conference on Health and Environment Delhi, India, 7-9 July 1998.

[3] Singh G., Pal A. K., and Tiwari A., “Air Pollution and its Impact on Social Spectrum with Special References to KORBA coal field on Chhattisgarh India”, First International conference on MSECCMI, New Delhi, pp. 549-558, 2007.

[4] Timonen K.L., Pekkanen J., Tiittanen P. and Salonen R.O., “Effects of Air Pollution on Changes in Lung Function Induced by Exercise in Children with Chronic Respiratory Symptoms”, Occupational Environmental Medicine, vol. 59(2), pp. 129-134, 2002.

[5] Maier H. R. and Dandy, G. C., “Neural Network Based Modelling of Environmental Variables a Systematic Approach”, Mathametical and Computer Modelling, vol. 33, pp. 669-682, 2001.

[6] Tecer L. H., “Prediction of SO_2 and PM Concentrations in a Coastal Mining Area (Zonguldak, Turkey) Using an Artificial Neural Network.” Environmental Studies, vol. 16(4), pp. 633-638, 2007.

[7] Brunelli U., Piazza V., Pignato L., Sorbello F., and Vitabile S., “Three Hours Ahead Prediction of SO_2 Pollutant Concentration Using an Elman Neural Based forcaster”, Building and Environment, vol. 43, pp. 304-314, 2008.

[8] Agirre-Basurko E., Ibarra-Berastegi G., and Madariaga I., “Regression and Multilayer Perceptron-based Models to Forecast Hourly O_3 and NO_2 Levels in the Bilbao Area”, Environmental Modelling & Software, vol. 21, pp. 430-446, 2006.

[9] Al-Alawi S.M., Abdual-Wahab S.A. and Bakheit C. S., “Combining Principal Component Regression and Artificial Neural Network for More Accurate Predictions of Ground Level Ozone”, Environmental Modelling & Software, vol. 23, pp. 396-403, 2008.

[10] Patricio P. and Jorge R., “An Integrated Neural Network Model for PM10 Forecasting”, Atmospheric Environment, vol. 40, pp. 2845-2851, 2006.

[11] Slini T., Kaprara A., Karatzas K. and Moussiopoulos N., “PM_{10} Forecasting for Thessaloniki, Greece”, Environmental Modelling & Software, vol. 21, pp. 559-565, 2006.

[12] http://mpcb.mah.nic.in/envtdata/envtair.php

[13] J A Freeman and D M Skapura, 'Neural Networks-algorithms, Applications, and Programming Techniques'. Addison-Wesley Publishing Company, New York, 1991.

[14] S Haykin, 'Neural Network: A Comprehensive Foundation'. Prentice-Hall, Upper Saddle River, New Jersey, 1999.