

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

1Mr.Bobe R.R., 2Mr.Mane S J

1M.E. Student (Env. Engg), Pune University.

2Assistant Professor Department of Civil Engineering D.Y. Patil College of Engg. Akurdi, Pune

Abstract- The present work aims at developing an efficient and reliable module, for forecasting concentration levels of respirable suspended particulate matter (RSPM) for the Pune city. There is dense vehicular traffic, due to increasing motorization & industrial activities, this has led to increasing population parameters. This present work defines efficient program for operational concentration level of RSPM. The daily RSPM dataset are collected from monitoring station operated by Maharashtra pollution control board for a period of three years (2010-2013). For determination and prediction of RSPM, multilayer perceptron, recurrent, generalized feed forward networks of artificial neural networks are used as network structure. Developed ANN model using multilayer perceptron network efficiently predict RSPM concentration ($r = 0.96$). The trial and error method is adopted to compare network output with desired output in terms of error statistics. The result obtained through proposed models show that ANN can be efficiently used in the analysis and prediction of air quality.

Keywords- Artificial neural network, Respirable suspended particulate matter, Multilayer perceptron, generalized feed forward network, Graphical user interface.

1. INTRODUCTION

Air pollution in large cities has been a major and a serious environmental problem all over the world; hence, many countries initiated air quality management systems to monitor and control pollution rates around big cities. It was found that harmful emission into the air is a symbol for environmental force that affects seriously man's health, natural life and agriculture; thus leading to major loss on the nation's economy^[1]. 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^[2]. The pattern of economic growth that we are adopting is increasingly associated with environmental pollution^[3]. 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.

Several methods have been used for modeling air quality. In recent years, artificial neural networks (ANNs) have become a popular and useful tool for modeling environmental systems^[5]. Many researchers have successfully used ANN models for prediction of SO₂^[6-7] NO₂^[8] Ozone^[9], and particulate matter level^[10-11] and found that the ANN model is suitable for predicting air pollutant. Computing with neural networks is one of the fastest growing fields in the history of artificial intelligence (AI), largely because NNs can be trained to identify nonlinear patterns between input and output values and can solve complex problems much faster than digital computers. This method is reliable for instantaneous prediction of pollutant^[12]. Owing to their wide range of applicability and their ability to learn complex and non-linear relationships – including noisy or less precise information – NNs are very well suited to solving problems in environmental engineering and in particularly towards the analysis of air pollution.

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.

2. DATA USED

- a. The daily data of RSPM from January 2010 - December 2013 *i.e.* 4yrs has been obtained for Pune from MPCB [13].
- b. Indian Metrological Department, Shivaji Nagar Pune.

3. MATERIAL AND METHOD

For this study potential feed forward three layered network been used. This network has been trained by back-propagation algorithm in forecasting RSPM concentration. For the illustration, by developing seasonal time series model with daily data for Pune city. Pune city located in the western region of Maharashtra, India at Coordinates: 18°31'13"N 73°51'24"E (Figure 1) has a population of about 8.24 lakhs. Pune is facing the problem of exponential population growth, with wide job opportunities starting from automotive, IT & bio-technology. Pune has a twin city *i.e.* Pimpri-Chinchwad. The growth of Pune is not limited to only PMC or PCMC, but all the circumferential area with industries, universities, institutes coming up in that area^[6].

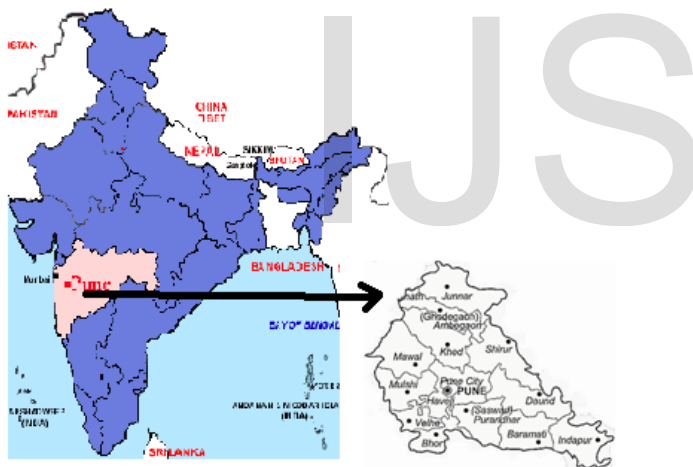


Figure 1: Location map of Pune City

A. Artificial Neural Network Model

ANN is based on the neurons connected or functionally-related to each other; imitate the behavior of human biological neurons. The extensive applications of neural networks are pattern recognition and classification, time series prediction and modeling^[14]. Neurons which are the basic components of the neural network are interconnected through different layers such as input, hidden and output layers. The degree of interconnection is defined by the weight which means the impact of neuron on neuron^[15].

The structure of neural network is shown in Fig. 2.

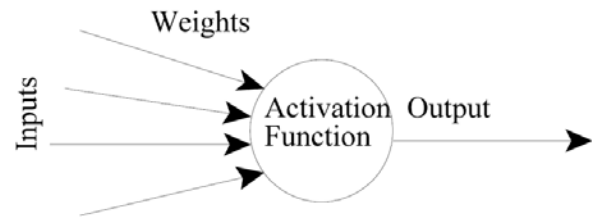


Figure 2: Schematic of a neural network

The important factor on performance is the topology of ANN. ANN is classified into major groups based on the pattern of interconnection of neurons to propagate data. The main categories of neural networks are Feed Forward and Recurrent Neural Networks. In Feed Forward Networks such as Multi-Layer Perceptron (MLP) and Feed Forward Network (FNN), data enter into input layer for processing layer by layer to achieve the output layer. Recurrent Neural Networks such as Elman and Hopfield have the recurrent path through feed backing from output layer to hidden layer for the detection and generation of time-varying patterns. ANN should be configured to produce the desired output by adjusting the weights of interconnections among all neuron pairs. This process is called as training which is categorized into two main groups called supervised and unsupervised learning. In supervised learning, ANN feeds with the learning patterns and adjusts the weights by comparison of the desired output with the actual output obtained from the input variables to achieve the minimum error. On the contrary, unsupervised learning is based on discovering the features of input data in a statistical manner by input irritant classification^[14]. In this paper, we focus on the supervised learning. The most popular learning used in the prediction purpose is the Back Propagation Algorithm (BPA) which aims to reduce the overall system error.

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.

$$\text{RSPM}(t+1) = f[\text{RSPM}(t), \text{RSPM}(t-1), \text{RSPM}(t-2), \text{RSPM}(t-3)]$$

Where 't' is day for which prediction is done.

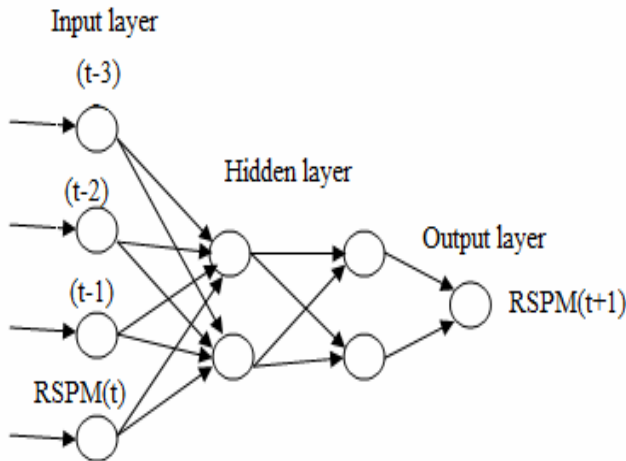


Figure 3: A typical neural network architecture

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 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.

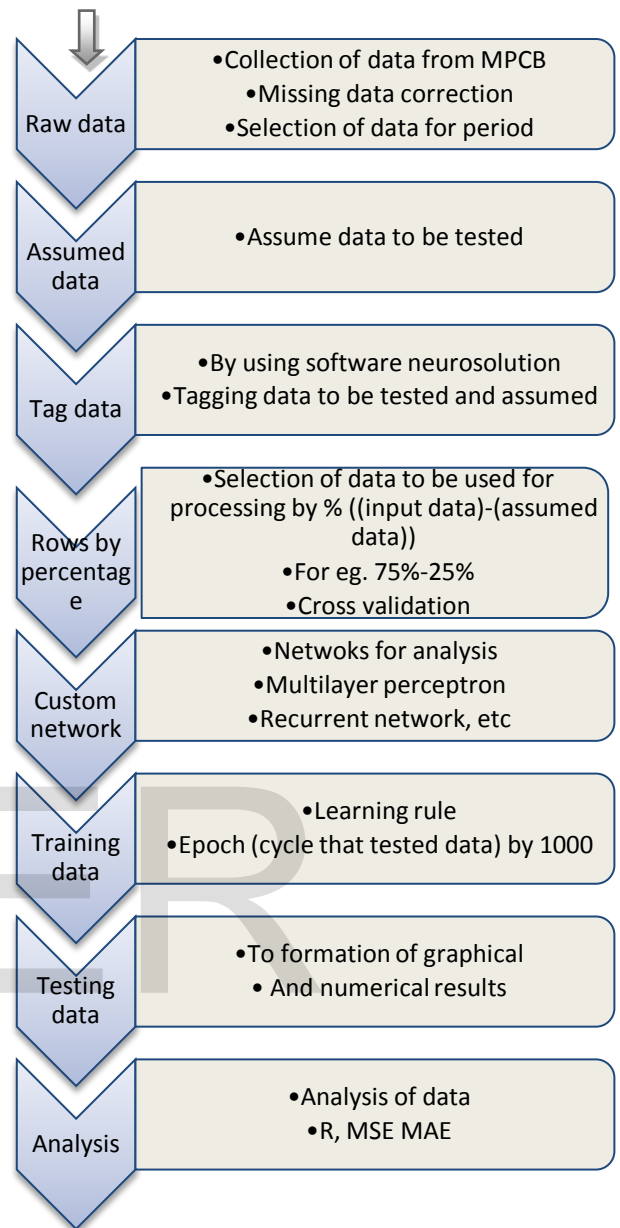
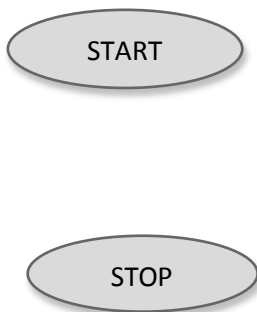


Figure 4: Flow chart of methodology

Methodology of getting RSPM from software. As this describes, analysis of R, MSE, MAE is done. This will form breadboard in neurosolutions software. This bread board having formulation of this data trained. This formulation useful for the preparation of Graphical User Interface (GUI)

3. RESULTS AND DISCUSSION

3.1 Result

By taking number of trials by changing the length of data set for training, testing and 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 (5-3-2-1) the neural network architecture was finalized *i.e.* five neurons in input layer, two hidden layers and one neuron in output layer. The maximum value of R is 0.96, MSE=605.454 ($\mu\text{g}/\text{m}^3$) and MAE=17.45 $\mu\text{g}/\text{m}^3$ as shown in Table1.

TABLE 1. SUMMERY OF TEST REPORT

Performance	T+1
MSE	282.5512679
NMSE	0.078088254
MAE	12.60240346
Min Abs Error	0.004001789
Max Abs Error	93.72074686
r	0.960163059

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.

The graph of the model output verses desired output as seen in Figure 5 and Figure 6 showed that most of the forecasted values of RSPM were close to the expected values, which indicate that the modeling has been quite successful.

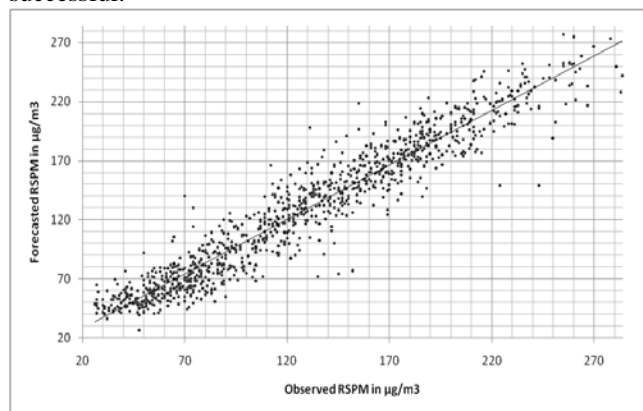


Figure 5: Scatter plot MLP (5-3-2-1) model

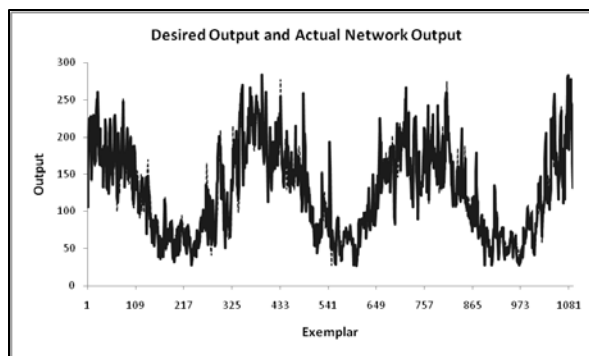


Figure 6: graph of ANN models

3.2 Discussion

As this result is tending to efficiency of 100%, it needs to be added some improvement as follows:

- Data input having missing data. Precise data cl be helpful to get better results.
- Data set for training added with real-time humidity, temperature, vehicular passes and daily trends and lifestyle of that area.
- These parameters can improve the results and can be predict maximum efficiency.

CONCLUSION

This is instantaneous prediction method for RSPM. The use of MLP method of ANN values provides reliable result, although the forecasting has been done for real source.

The value of R (0.961) is highly encouraging *i.e.* 96.2% 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.

Future Scope:

For ease of use this result, further development of graphical user interface (GUI) need to be done. This GUI helps to people get predication easily, and not need to know entire software and its methodologies.

This GUI can also useful to automated air pollution control unit. Figure 7 shows the wire frame of the GUI. Figure 8 (a) is the input data window. Figure 8 (b) shows the result of the next day.

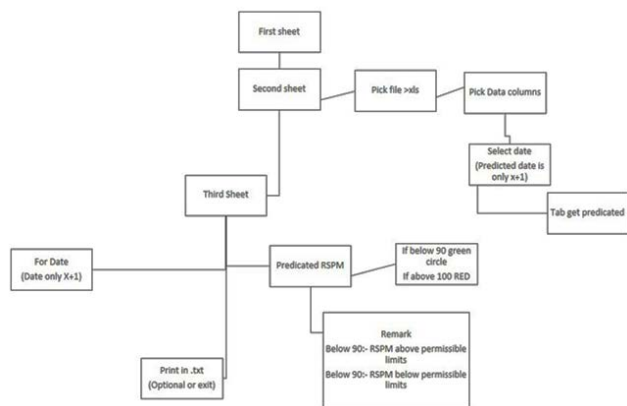
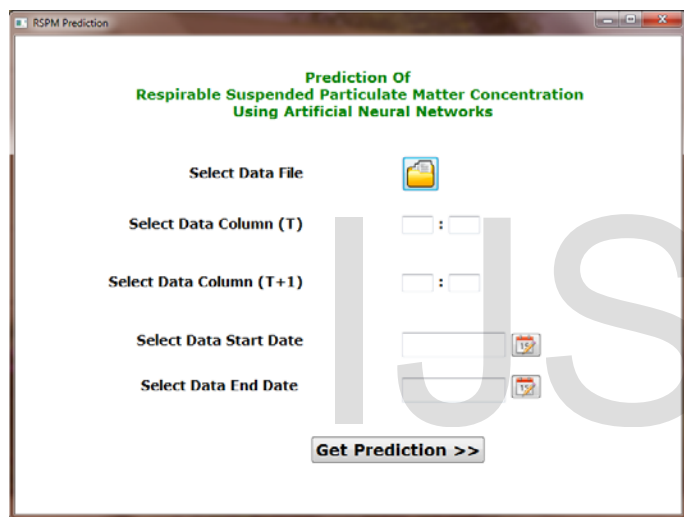
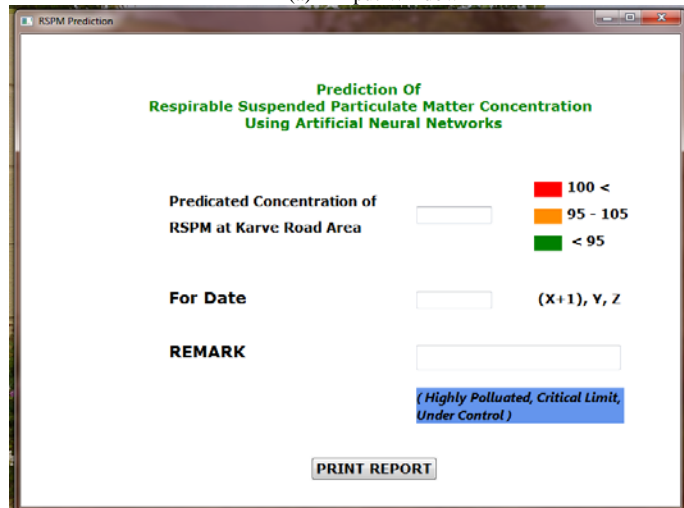


Figure 7: wire frame of GUI



(a) Input Window



(b) Predication Report window

Figure 8 (a & b): Graphical user interface

ACKNOWLEDGEMENT

I extend my sincere thanks to the H.O.D. Civil Department of D.Y. Patil College of Engineering, Akurdi, Pune, and Mr. Sunil Enadle, Soft Engg. Synerzip, Pune for the major roll preparation of GUI.

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 MSECMI, 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", *Mathematical and Computer Modelling*, vol. 33, pp. 669-682, 2001.
- [6] Tecer L. H., "Prediction of SO₂ 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 SO₂ 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 O₃ and NO₂ 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., "PM₁₀ Forecasting for Thessaloniki, Greece", *Environmental Modelling & Software*, vol. 21, pp. 559-565, 2006.
- [12] Sonaje N P Mane S J Kote A S, *Modelling of Respirable Suspended Particulate Matter Concentration Using Artificial Neural Networks in an Urban Area*, IRACST – Engineering Science and Technology: An International Journal (ESTIJ), ISSN: 2250-3498, Vol.3, No.4, August 2013

[13] Maharashtra Pollution Control Board
<http://mpcb.mah.nic.in/envtdata/envtair.php>

[14] M. A. Behrang, E. Assareh, A. Ghanbarzadeh, A. R. Noghrehabadi, "The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data," *Solar Energy*. 2010, 84: 1468-1480

[15] *International Journal of Environmental Science and Development*, Vol. 3, No. 1, February 2012 The Potential of Artificial Neural Network Technique in Daily and Monthly Ambient Air Temperature Prediction Mahboubeh Afzali, Afsaneh Afzali, and Gholamreza Zahedi.

IJSER