International Journal of Civil Engineering (IJCE)
ISSN(P): 2278-9987; ISSN(E): 2278-9995
Vol. 5, Issue 3, Apr - May 2016; 1-10

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

PREDICTION OF PM10 AND SO_2 CONCENTRATIONS IN AMBIENT AIR USING ARTIFICIAL NEURAL NETWORKS FOR HYDERABAD

V. SAMPATH KUMAR REDDY[1], M. SRIMURALI[2] & B. POLAIAH[3]

[1]M. Tech Scholar, Department of Civil Engineering, Sri Venkateswara University College of Engineering, Andhra Pradesh, India

[2]Professor, Department of Civil Engineering, Sri Venkateswara University College of Engineering, Andhra Pradesh, India

[3]Professor, Department of Electronics and Communication Engineering, Sree Vidyanikethan Engineering College, Andhra Pradesh, India

ABSTRACT

An Artificial Neural Networks (ANNs) models are constructed to predict PM10 and SO_2 concentrations for Hyderabad. The model uses meteorological variables like wind speed, wind direction, temperature, relative humidity and atmospheric pressure as input variables. Three models have been developed one is for the prediction of PM10 using meteorological parameters, second one is for the prediction of SO_2 using meteorological parameters and particulate matter concentrations and the third one is for the prediction of PM10 and SO_2 using meteorological parameters as input variables. The correlation coefficient between observed and predicted concentrations are in the range of 0.982 to 0.962.The evaluation of models results shows that the degree of success in PM10 and SO_2 concentration are seems to be good.

KEYWORDS: Artificial Neural Networks

INTRODUCTION

Air pollution is a major threat to health and is generated by rapid urbanization, population growth and industrialization. Air quality forecasting tools are necessary to take precautionary measures, such as reducing the effect of a predicted pollution peak on the surrounding population and ecosystem.

Many factors influence the concentration of air pollutants. Among the most important are metrological conditions, topology and population density. This makes air pollution difficult to model. Many air pollution prediction models have been studied such as, mathematical emission models, linear models, artificial neural networks-based models, and hybrid models. The purpose is to design air quality prediction systems, moderate air pollution and limit the influence of peak periods by informing the community so that they may take the necessary precautions. Air pollutants can be of gaseous form such as SO_2, O_3, NO_X and CO_X, or solid such as PM_{10}.

In a developing country like India vehicular pollution is no longer intangible threat. It contributes to a shocking 64% of the total pollution in Delhi, 52% in Mumbai and 30% in Calcutta. In Indian urban life style, the atmosphere at traffic junctions and intersections of any urban center receive maximum input of traffic exhausts pollutants and thereby they are converted into localized high pollution episodes. Among the pollutants PM10 is of great concern and the major air pollutant in all the urban centers and vehicles are the main sources.

From the Figure 1.1, it is observed that the major sources for PM10 is the vehicular source next to that second one is the road dust, increase in the vehicular population and also usage of road which creates road dust leads to the increase in the emissions of PM10.Recent articles also shown that the levels of PM10 in Hyderabad are beyond the permissible limits.

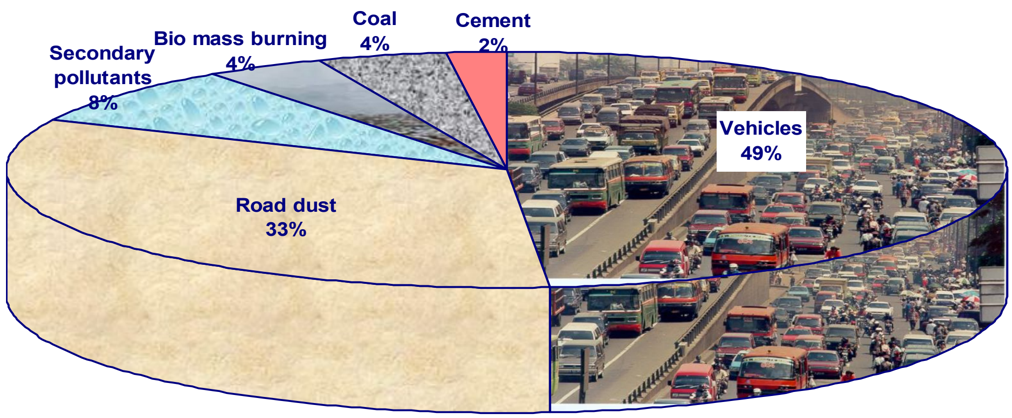

Figure 1.1: Sources that Contribute for the Production of PM10

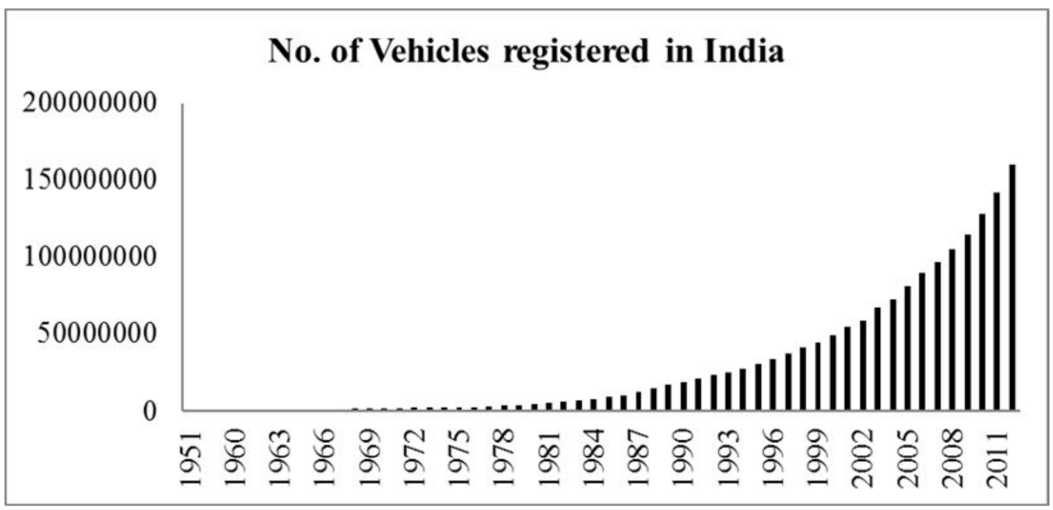

Figure 1.2: No. of Vehicles Registered in India from 1951 to 2011

The vehicular pollution is the major source for the PM10 emission if we observe from the Figure 1.2, it is observed that there is a gradual increase in number of vehicles registered in India from 1951 to 2011 which leads to the increase in the emissions of the PM10 concentrations. In the past several studies has been done to predict the pollutant concentration by using meteorological parameters by using various statistical tools among that Artificial Neural Networks is one of the most used tool for the prediction. *Boznar et al (1993)* used neural networks to predict short term SO_2 concentrations in highly polluted industrial areas of complex terrain around the Slovenian Thermal Power Plant at Sostanj, India. *Zickus et al (2000)* evaluated the variable section and prediction performance of several machine learning techniques. The techniques were applied to a PM10 data set in Helsinki, Finland. *Ojha et al. (2002)* presented a compendium of available methods and software for ozone and PM10 forecasting. *Sofuoglu et al. (2006)* constructed an Artificial Neural Networks (ANNs) model to forecast SO_2 concentrations in Izmir air. The model uses meteorological variables (wind speed and temperature) and measured particulate matter concentrations as input variables.

ARTIFICIAL NEURAL NETWORKS

The human information processing system consists of the biological brain. The basic building block of the nervous system is biological neuron, the cell that communicates information to and from the various parts of the body. The neuron consists of a cell body called a soma, several spine like extensions of the body cell body called dendrites, a single nerve fiber called the axon that branches out from soma and connects to many other neurons, and the junctions by which connections between neurons occur either on the cell body or on the dendrites called synapses. A simple structure of biological neuron is shown in Figure 2.1.

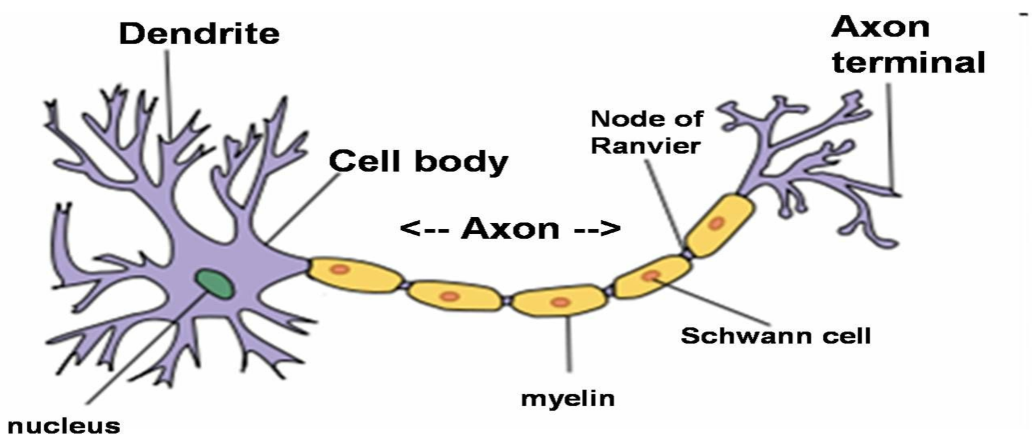

Figure 2.1: Structure of Biological Neuron

An Artificial Neural Network (ANN), often called a "Neural Network" (NN), is interconnected group artificial neuron that uses a mathematical model or computational model for information processing based on a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network. In more practical terms neural networks are non-linear statistical data modeling tools.

ANNs have been widely used for modeling, control, pattern recognition, signal processing, prediction, etc. (Zurada 1992). The ANN is taught to model a relationship during a supervised training procedure by using series of input and associated output data. The ANN consists of several layers: the first layer is the input layer, and the final one is the output layer. The layers between the first and the last layers are the hidden or intermediate layers. The general structure of an ANN is well known and can be found in numerous publications (Baughman and Liu 1995; Bulsari 1995; Zurada 1992). Basically, the learning in the network is achieved through an iterative algorithm that minimizes the mean-square errors between the desired and actual outputs.

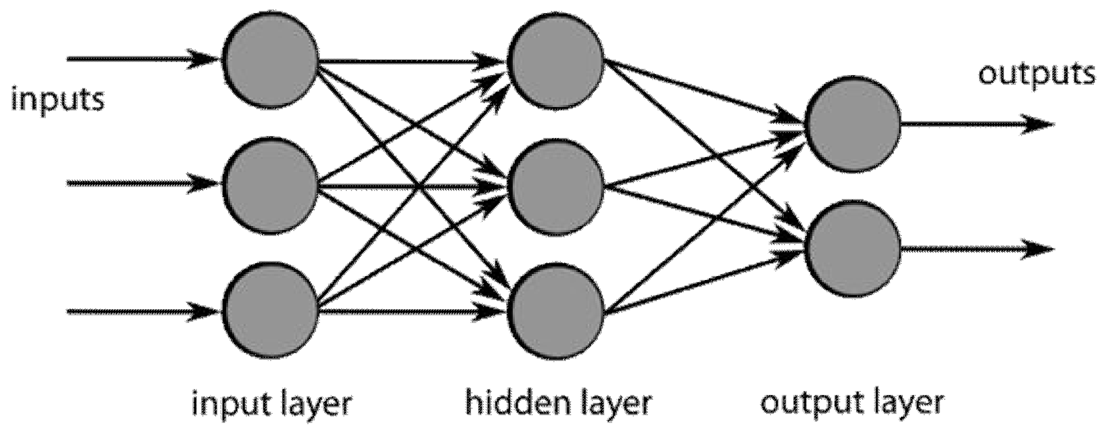

Figure 2.2: Structure of ANN

AMBIENT AIR QUALITY DATA COLLECTION

Hyderabad, a 400-year-old city, is the capital of Andhra Pradesh and Telangana, India. Occupying 625 square kilometers along the banks of the Musi River, it has a population of about 6.8 million and a metropolitan population of about 7.75 million. It is the fourth largest and one of the fastest growing cities in India, with a population density of 18,480 persons / km^2. A booming information technology industry has led to expansion of the city, which now includes the satellite districts, collectively known as the Hyderabad Urban Development Area (HUDA).

The hourly concentrations of air pollutants like Particulate Matter (PM10) and Sulfur Dioxide (SO_2) and hourly meteorological parameters like Wind Speed (WS), Wind Direction (WD), Relative Humidity (RH), Solar Radiation (SR), Atmospheric Temperature (AT) & Atmospheric Pressure (AP) were collected simultaneously from Ambient Air Quality Stations established by the Andhra Pradesh Pollution Control Board at Sanath Nagar station, Hyderabad from 2009- 2013.

NEURAL NETWORK MODEL DEVELOPMENT

In this study, a multilayer Feed- Forward Back propagation type of ANN was considered to forecast ambient air PM10 and SO_2 concentrations based on meteorological parameters like Wind Speed, Wind Direction, Relative Humidity, Solar Radiation, Atmospheric Temperature and Barometric Pressure. In a Feed- Forward Back propagation network, the input quantities are fed into input layer neurons, which in turn pass them onto hidden layer neurons after multiplying by a weight. The weights are adaptive coefficients within the network that determine the intensity of the input signal. A hidden layer neuron adds up the weighted input received from each input neuron, associates it with a bias, and then passes the result on through a nonlinear transfer function. The output neurons do the same operation as that of a hidden neuron. The bias neurons are connected to the all neurons in the next hidden and output layer neurons to improve the convergence properties of the network. Each bias neuron is assigned a constant random number. The performance of the network was evaluated using two criteria. The first one is the Coefficient of Determination, R^2 value (correlation coefficient); it denotes the level of correlation between the observed and forecasted concentrations, and it is preferred due to comparability with conventional studies. The whole dataset consist of 1340 samples of daily average values of each and every paramater which are divided into three parts: one part i.e., about 60% of data for study area is for training of models, second part i.e., about 20% of data for study area is for evaluate the performance of model developed and the remaining 20% of the data is to test the network developed in this study with different combination of no. of neurons and hidden layers, the data used for development of neural network. The no. of hidden layers has been chosen from the following table given by various authors and the no. of neurons is based on trial and error method.

Table 1: Different Approaches for Deciding No. of Hidden Layers

Author	Hidden Layer
Maren et al (1990); Masters (1993); Rojas (1996)	Trial and error method
Berke and Hajela (1991)	(Input + Output)/2
Hecht-Nielsen (1990); Caudill (1989)	$(2I^*+1)$
Masters (1993)	(No. of training samples)/No. of layers = 2
	(Max value of tr sample)/No. of layers = 4
Hush and Horne (1993)	(Max value of tr sample)/No. of layers = 10
Amari et al (1997)	(Max value of tr sample)/No. of layers = 30

In these present study three architectures has been derived for the prediction of air pollutant concentration:

- Architecture – 1 for the prediction of PM10 by using meteorological parameters as inputs.
- Architecture – 2 for the prediction of SO_2 by using meteorological parameters including PM10 as inputs.
- Architecture – 3 for the prediction of PM10 & SO_2 by using meteorological parameters as inputs.

The proposed ANN models are developed using "Graphical User Interface (GUI)" using NN tool in MATLAB software. About 42 networks have been developed by taking different combinations of no. of neurons and no. of hidden layers and the results are given below:

Table 2: Performance Statistics of a Neural Network Model for Architecture-1 for PM10 Concentration for Hyderabad Station

S. No	Models	No. of Neurons	No. of Hidden Layers	R (Correlation)		
				Training	Testing	All
1	6-5-5-1	5	5	0.696	0.559	0.679
2	6-5-4-1	5	4	0.660	0.772	0.666
3	6-5-13-1	5	13	0.600	0.669	0.648
4	6-5-670-1	5	670	0.760	0.791	0.765
5	6-5-335-1	5	335	0.795	0.780	0.789
6	6-5-134-1	5	134	0.800	0.776	0.791
7	6-5-45-1	5	45	0.789	0.825	0.803
8	6-6-5-1	6	5	0.822	0.856	0.827
9	6-6-4-1	6	4	0.808	0.814	0.808
10	6-6-13-1	6	13	0.819	0.772	0.809
11	6-6-670-1	6	670	0.806	0.831	0.811
12	6-6-335-1	6	335	0.803	0.838	0.811
13	6-6-134-1	6	134	0.811	0.824	0.814
14	6-6-45-1	6	45	0.822	0.824	0.820
15	6-7-5-1	7	5	0.827	0.829	0.831
16	6-7-4-1	7	4	0.818	0.820	0.819
17	6-7-13-1	7	13	0.832	0.813	0.829
18	6-7-670-1	7	670	0.816	0.841	0.820
19	6-7-335-1	7	335	0.832	0.783	0.827

20	6-7-134-1	7	134	0.827	0.834	0.828
21	6-7-45-1	7	45	0.814	0.856	0.823
22	6-8-5-1	8	5	0.809	0.836	0.805
23	6-8-4-1	8	4	0.833	0.848	0.830
24	6-8-13-1	8	13	0.775	0.721	0.764
25	6-8-670-1	8	670	0.832	0.786	0.821
26	6-8-335-1	8	335	0.783	0.848	0.801
27	6-8-134-1	8	134	0.819	0.859	0.826
28	6-8-45-1	8	45	0.853	0.803	0.843
29	6-9-5-1	9	5	0.842	0.844	0.841
30	6-9-4-1	9	4	0.839	0.829	0.840
31	6-9-13-1	9	13	0.846	0.835	0.843
32	6-9-670-1	9	670	0.838	0.847	0.837
33	6-9-335-1	9	335	0.823	0.863	0.831
34	6-9-134-1	9	134	0.842	0.837	0.837
35	6-9-45-1	9	45	0.822	0.820	0.823
36	6-10-5-1	10	5	0.856	0.901	0.893
37	6-10-4-1	10	4	0.892	0.904	0.898
38	6-10-13-1	10	13	0.943	0.982	0.982
39	6-10-670-1	10	670	0.912	0.959	0.962
40	6-10-335-1	10	335	0.893	0.943	0.914
41	6-10-134-1	10	134	0.885	0.918	0.894
42	6-10-45-1	10	45	0.875	0.893	0.882

Table 3: Performance Statistics of a Neural Network Model for Architecture-2 for SO_2 Concentration for Hyderabad Station

S. No	Models	No. of Neurons	No. of Hidden Layers	R (Correlation)		
				Training	Testing	All
1	7-5-5-1	5	5	0.801	0.771	0.790
2	7-5-4-1	5	4	0.803	0.788	0.797
3	7-5-15-1	5	15	0.794	0.811	0.800
4	7-5-670-1	5	135	0.775	0.721	0.764
5	7-5-335-1	5	67	0.832	0.786	0.821
6	7-5-134-1	5	27	0.783	0.848	0.801
7	7-5-45-1	5	9	0.809	0.836	0.805
8	7-6-5-1	6	5	0.811	0.830	0.818
9	7-6-4-1	6	4	0.786	0.744	0.799
10	7-6-15-1	6	15	0.818	0.791	0.800
11	7-6-670-1	6	135	0.819	0.808	0.808
12	7-6-335-1	6	67	0.814	0.834	0.819
13	7-6-134-1	6	27	0.789	0.749	0.789
14	7-6-45-1	6	9	0.818	0.798	0.808
15	7-7-5-1	7	5	0.803	0.788	0.797
16	7-7-4-1	7	4	0.765	0.868	0.874
17	7-7-15-1	7	15	0.865	0.858	0.884
18	7-7-670-1	7	135	0.724	0.689	0.704
19	7-7-335-1	7	67	0.788	0.767	0.798
20	7-7-134-1	7	27	0.797	0.797	0.769
21	7-7-45-1	7	9	0.804	0.809	0.797
22	7-8-5-1	8	5	0.919	0.933	0.931
23	7-8-4-1	8	4	0.919	0.934	0.948

24	7-8-15-1	8	15	0.937	0.955	0.938
25	7-8-670-1	8	135	0.927	0.953	0.932
26	7-8-335-1	8	67	0.928	0.894	0.932
27	7-8-134-1	8	27	0.931	0.958	0.940
28	7-8-45-1	8	9	0.951	0.944	0.959
29	7-9-5-1	9	5	0.915	0.913	0.891
30	7-9-4-1	9	4	0.909	0.904	0.898
31	7-9-15-1	9	15	0.935	0.925	0.928
32	7-9-670-1	9	135	0.919	0.953	0.922
33	7-9-335-1	9	67	0.921	0.894	0.912
34	7-9-134-1	9	27	0.929	0.958	0.940
35	7-9-45-1	9	9	0.912	0.944	0.939
36	7-10-5-1	10	5	0.931	0.958	0.940
37	7-10-4-1	10	4	0.952	0.982	0.971
38	7-10-15-1	10	15	0.945	0.947	0.943
39	7-10-670-1	10	135	0.940	0.939	0.935
40	7-10-335-1	10	67	0.933	0.931	0.939
41	7-10-134-1	10	27	0.938	0.926	0.946
42	7-10-45-1	10	9	0.931	0.924	0.934

Table 4: Performance Statistics of a Neural Network Model for Architecture-3 for PM10 & SO_2 Concentration for Hyderabad Station

S. No	Models	No. of Neurons	No. of Hidden Layers	R (Correlation)		
				Training	Testing	All
1	6-5-5-2	5	5	0.724	0.689	0.704
2	6-5-4-2	5	4	0.788	0.767	0.798
3	6-5-13-2	5	13	0.797	0.797	0.769
4	6-5-670-2	5	670	0.804	0.809	0.797
5	6-5-335-2	5	335	0.819	0.814	0.816
6	6-5-134-2	5	134	0.775	0.721	0.764
7	6-5-45-2	5	45	0.764	0.775	0.771
8	6-6-5-2	6	5	0.724	0.699	0.706
9	6-6-4-2	6	4	0.781	0.761	0.779
10	6-6-13-2	6	13	0.790	0.792	0.785
11	6-6-670-2	6	670	0.802	0.804	0.789
12	6-6-335-2	6	335	0.809	0.814	0.804
13	6-6-134-2	6	134	0.765	0.758	0.774
14	6-6-45-2	6	45	0.794	0.811	0.800
15	6-7-5-2	7	5	0.823	0.835	0.822
16	6-7-4-2	7	4	0.809	0.814	0.804
17	6-7-13-2	7	13	0.724	0.689	0.704
18	6-7-670-2	7	670	0.788	0.767	0.798
19	6-7-335-2	7	335	0.797	0.797	0.769
20	6-7-134-2	7	134	0.795	0.780	0.793
21	6-7-45-2	7	45	0.802	0.786	0.795
22	6-8-5-2	8	5	0.881	0.861	0.879
23	6-8-4-2	8	4	0.890	0.892	0.885
24	6-8-13-2	8	13	0.822	0.884	0.889
25	6-8-670-2	8	670	0.809	0.814	0.804
26	6-8-335-2	8	335	0.809	0.839	0.897
27	6-8-134-2	8	134	0.828	0.859	0.844
28	6-8-45-2	8	45	0.959	0.929	0.941

29	6-9-5-2	9	5	0.919	0.933	0.931
30	6-9-4-2	9	4	0.919	0.934	0.948
31	6-9-13-2	9	13	0.937	0.955	0.938
32	6-9-670-2	9	670	0.927	0.953	0.932
33	6-9-335-2	9	335	0.928	0.894	0.932
34	6-9-134-2	9	134	0.931	0.958	0.940
35	6-9-45-2	9	45	0.951	0.944	0.959
36	6-10-5-2	10	5	0.917	0.937	0.924
37	6-10-4-2	10	4	0.955	0.966	0.955
38	6-10-13-2	10	13	0.935	0.947	0.928
39	6-10-670-2	10	670	0.930	0.919	0.925
40	6-10-335-2	10	335	0.923	0.941	0.930
41	6-10-134-2	10	134	0.918	0.826	0.906
42	6-10-45-2	10	45	0.943	0.982	0.962

RESULTS AND DISCUSSIONS

The networks model are developed by taking Wind Speed, Wind Direction, Relative Humidity, Atmospheric Temperature, and Atmospheric Pressure and Solar radiation as input variables. Among the 42 neural networks developed the best neural networks have been selected in such a way that the correlation value of the network should be >0.9. From the above tabular columns the best network models proposed for prediction of PM10 and SO_2 are given in Table 4.5. The Regression values of these proposed models are good.

Table 5: Proposed Neural Network Models for Three Architectures

S. No	Architecture	Hyderabad				
		Network	Training	Validation	Testing	Overall
1	Architecture-1	6-10-13-1	0.943	0.993	0.982	0.982
2	Architecture-2	7-10-4-1	0.952	0.971	0.982	0.971
3	Architecture-3	6-10-45-2	0.943	0.972	0.982	0.962

For the best neural network selected, the comparison is made between the predicted values and the measured values and the represented in the form graphs which are shown below:

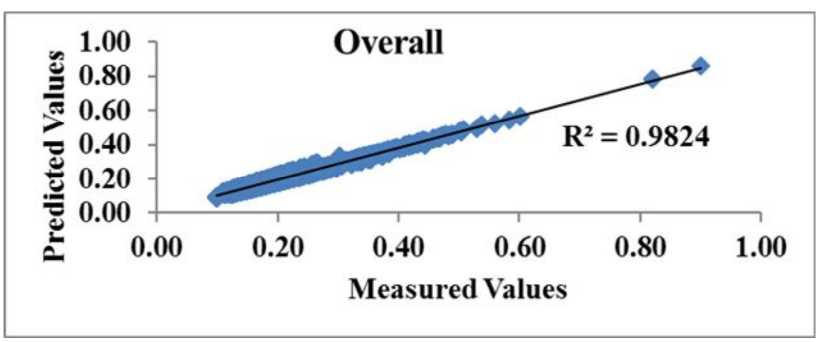

Figure 5.1: Overall Regression Analysis of Measured Values Vs Predicted Values for Architecture – 1 (Prediction of PM10)

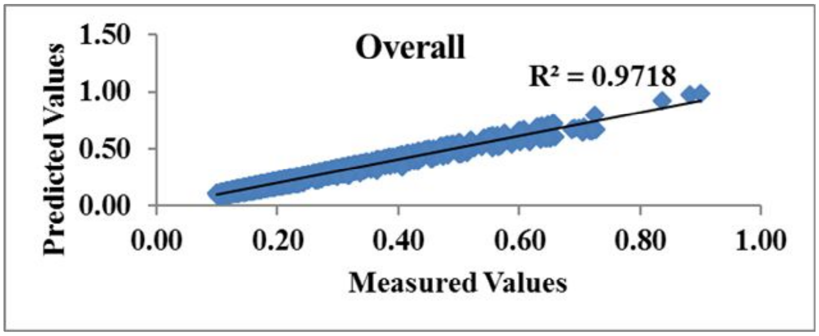

Figure 5.2: Overall Regression Analysis of Measured Values Vs Predicted Values for Architecture – 2 (Prediction of SO_2)

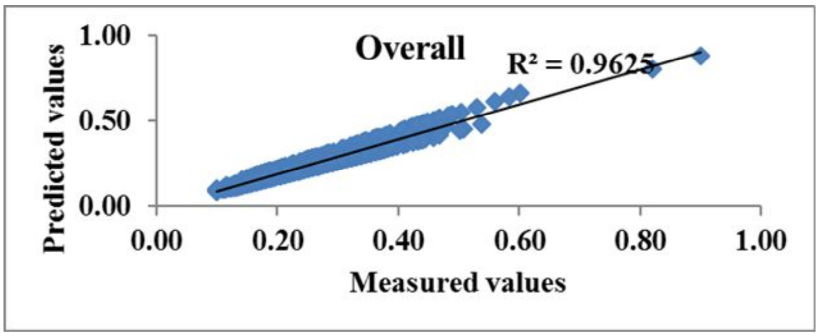

Figure 5.3: Overall Regression Analysis of Measured Values Vs Predicted Values for Architecture – 3 (Prediction of PM10 and SO_2)

CONCLUSIONS

ANN modeling is shown to be a successful method to forecast SO_2 and PM10 concentrations in the ambient air as correlation coefficients between predicted and measured concentrations of the three architecture networks are >0.90. For the prediction of PM10 for Hyderabad, Architecture – 1 is most suitable rather than Architecture – 3 because the regression value of Architecture – 1 is 0.982 whereas for Architecture – 3 is 0.962; whereas for the prediction of SO_2 Architecture – 2 is most suitable rather than Architecture – 3 because the regression value of Architecture – 2 is 0.971 whereas for Architecture – 3 is 0.962.

ACKNOWLEDGEMENTS

The authors are grateful to the Coordinator, Center of Excellence for providing necessary funding and support for the completion of project and also the Director, National Atmospheric Research Center for providing the data on Ambient Air Quality and also Andhra Pradesh Pollution Control Board for providing the Air Quality Data.

REFERENCES

1. Boznar M, Lesjak M, Mlakar P (1993), "A neural network-based method for short-term predictions of ambient SO2 concentrations in highly polluted industrial areas of complex terrain". Atmospheric Environment 27B: 221–230.
2. Ojha S., Coutinho J. and Kumar A., (2002), "Developing System to Forecasting ozone and Particulate Matter

Levels", Environmental Progress, 21(2), PP. J5 – J10.

3. Sait C. Sofuoglu, Aysun Sofuoglu, Savas Birgili and Gokmen Tayfur, (2006), "Forecasting Ambient Air SO_2 Concentrations Using Artificial Neural Networks". Energy Sources, Part B: Economics, Planning, and Policy, Taylor & Francis Group, LLC ISSN: 1556-7249 print/1556-7257 online DOI: 10.1080/009083190881526.

4. Zickus M., Greig A.J., Niranjan M., (2002) "Evaluating the Variable Selection and Prediction Performance of Several Machine Learning Algorithms Applied to a PM10 Data Set", http://www.mrec.ugam.ca/v-f/Publications/Publicatins.html.

5. Baughman, D.R., Liu, Y.A., (1995) Neural networks in bioprocessing and chemical engineering, Academic Press, New York.

6. Bulsari A.B., Neural Networks for Chemical Engineers, Computer -Aided Chemical Engineering, 6, (ed.) A. B. Bulsari, Elsevier Sci., Amsterdam (1995).

7. Zurada, J.M. (1992). "Introduction to Artificial Neural Systems", West Publishing Company, St. Paul, U.S.A.