



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 8, August 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.569

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



Prediction of Air Pollutants Concentration by Using Artificial Neuron Network

Rutuja Kasbe¹, Dr Sachin J. Mane², Dr. Ashok B. More³

D. Y. Patil College of Engineering, Akurdi, India^{1,2,3}

ABSTRACT: In this paper, an artificial neural network is anticipated for prediction of PM_{2.5} and PM₁₀ concentrations for Nagpur city, the major industrial urban area of Maharashtra, India. The developed artificial neural network models involve the parameter considered for the maximum and the average value of pollutants and past data for 4 yrs (2017-2021) as an input. The following concentration of pollutants for one day ahead being the output parameter of this study is probable. The developed model is based on three-layer neural network trained by a multilayer perceptron and generalized feed forward network with number of epoch. The models accurately go with the drift of pm_{2.5} and pm₁₀ concentrations for one day ahead (R=0.790).

KEYWORDS: Artificial neural network, Generalized Feed Forward Network, PM_{2.5}, PM₁₀, Multilayer Perceptron, Artificial Neuron Network.

I. INTRODUCTION

Air pollution causes uneasiness to the health of humans and animals and causes damage to plants and materials. With increasing urbanization and industry, humans began to unharness additional wastes into the atmosphere than nature might touch upon with additional pollution has been supplementary to the air by industrial, business and domestic sources. More attention is given to pm_{2.5} and pm₁₀ as it weakens the respiratory system of humans and corrupting air quality. These particles once breathed in, dwell our respiratory organ tissues and cause respiratory organ injury and metabolic process issues. Therefore, its necessary and needful to carry out prediction for the prevention of further loss and damages caused by the air pollution and thus, to create the health awareness among the masses. The forecasting tool like ANN has been in demand and it is used for the prediction of pm_{2.5} and pm₁₀.

The Artificial intelligence and machine learning has become key focus. Procedure of machine learning is that the machine first senses the environment and the factors related to it. It then gathers information and makes appropriate decisions according to it. One of the main reasons why machine learning was chosen to predict air quality index and other harmful pollutants was this ability of adapting of machine learning (ML) algorithms.

The basic features of ANN are its neutrally implemented mathematical model which includes bulky number of interrelated processing elements called neurons for all the process. The information is stored in neurons which are basically linkage of neurons. For processing elements the input signals arrive by connections and connecting weights. The following fig.1 and fig.2 explains biological neural network which is inspired by the human brain, These biologically inspired methods of computing. Computers do rote things well, like keeping ledgers or performing multifarious math. But they have difficulty in recognizing even simple patterns much less generalizing those patterns of the past into actions of the future.(13) Now, advances in biological research promise an initial understanding of the natural thinking mechanism. This research shows that brains store information as patterns. Some of these patterns are very complicated and allow us the ability to recognize individual faces from many different angles(10). This process of storing information as patterns, utilizing those patterns, and then solving problems encompasses a new field in computing. This field, involves the making of extremely parallel networks and the training of those networks to solve specific problems. This field also utilizes words very different from traditional computing, words like behave, react, self-organize, learn, generalize, and forget. Following are the diagrams which represents the biological neuron and working of ANN.

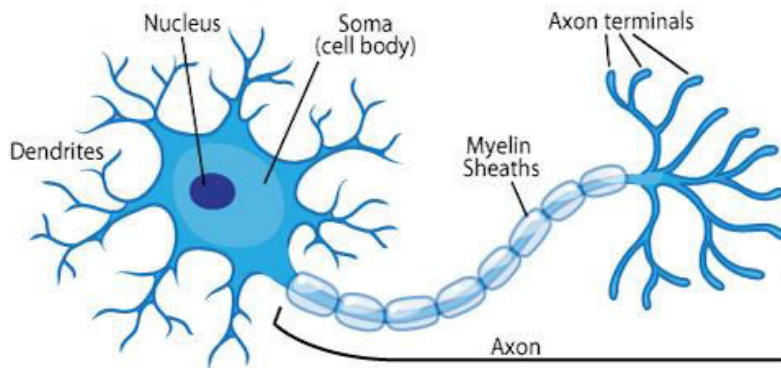


Fig. 1 :- Sketch of biological neuron

(Source: <https://support.google.com/legal/answer/3463239?hl=en-US>)

Biological Neuron versus Artificial Neural Network

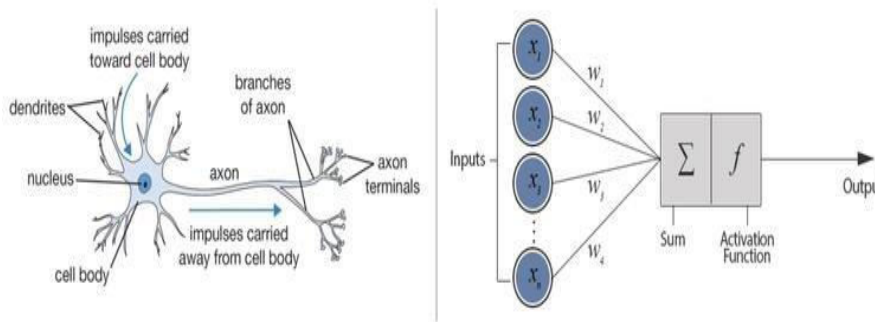


Fig. 2 :- Schematic representation of biological neuron versus artificial neural network working.

(Source: <https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.researchgate.net%2Ffigure%2Ffigure%2FBiological-Neuron-versus-Artificial-Neural-Network>)

The two types of network are used for prediction of air pollutants concentration which includes MLP and GFF. The MLP it includes three layers of nodes i.e input layer ,hidden layer and output layer. Each node is a neuron that uses a nonlinear activation function exception is only for input layer. It utilizes a supervised learning technique called backpropagation for training is a feedforward neural network, which means that the data is transmitted from the input layer to the output layer in the forward direction. The connections between the layers are assigned weights. The weight of a connection specifies its importance. This concept is the backbone of an MLP’s learning process.

Backpropagation is a technique used to optimize the weights of an MLP using the outputs as inputs. In a conventional MLP, random weights are assigned to all the connections. These random weights propagate values through the network to produce the actual output. Naturally, this output would differ from the expected output(8). The difference between the two values is called the error. Backpropagation refers to the process of sending this error back through the network, readjusting the weights automatically so that eventually, the error between the actual and expected output is minimized. In this way, the output of the current iteration becomes the input and affects the next output. This is repeated until the correct output is produced(7). The weights at the end of the process would be the ones on which the neural network works correctly.

A GFF neural network is an ANN where the unit connections do not form a cycle. Feed forward neural network was the first and debatably most simple type of artificial neural network devised. In this network the information moves in only one direction — forwards: From the input nodes data goes through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network.(12) Feed forward networks can be constructed from different types of units, e.g. binary McCulloch-Pitts neurons, the



simplest example being the perceptron. Continuous neurons, frequently with sigmoidal activation, are used in the context of back propagation of error(9).

Feed-forward networks have the following characteristics: 1. Perceptrons are arranged in layers, with the first layer taking in inputs and the last layer producing outputs. The middle layers have no connection with the external world, and hence are called hidden layers. Each perceptron in one layer is connected to every perceptron on the next layer(11).Hence information is constantly "fed forward" from one layer to the next., and this explains why these networks are called feed-forward networks. 3. There is no connection among perceptrons in the same layer.

II. MATERIAL AND METHOD

Nagpur is a large city in the central Indian state of Maharashtra and it is called as ‘Orange city’ It is the winter capital, a sprawling metropolis and third largest city of the Indian state of Maharashtra after Mumbai and Pune .Nagpur is the 13th largest Indian city in terms of population. Due to metropolitan city there is rapid ascend in population, vehicular traffic and urbanisation which leads to environmental deprivation, particularly air quality in Nagpur city. Thus ,controlling air pollutant parameter is needed and its prediction is required .The potential of three-layered MLP, GFN,networks trained by back-propagation algorithm in forecasting pm2.5 and pm10 concentration has been utilize to develop a time series model and scatter plot for Nagpur city. The daily hourly basis data of pm2.5 and pm10 from (1-1-2017) and (30-4-2021) was collected from CPCB (Central Pollution Control Board).

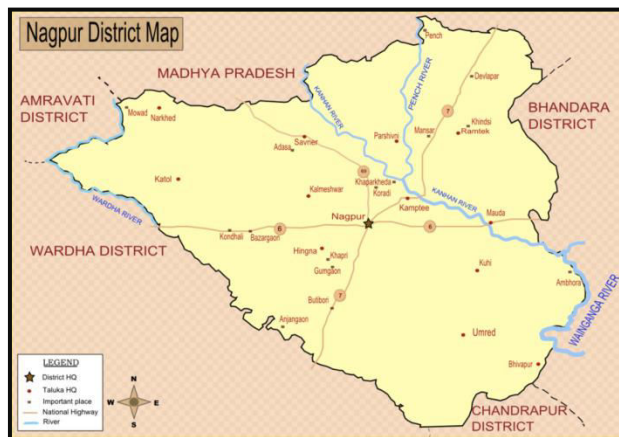


Fig.3 : District map of Nagpur

(Source: <https://images.app.goo.gl/6AXUEMwwXsj4y7tZ8>)

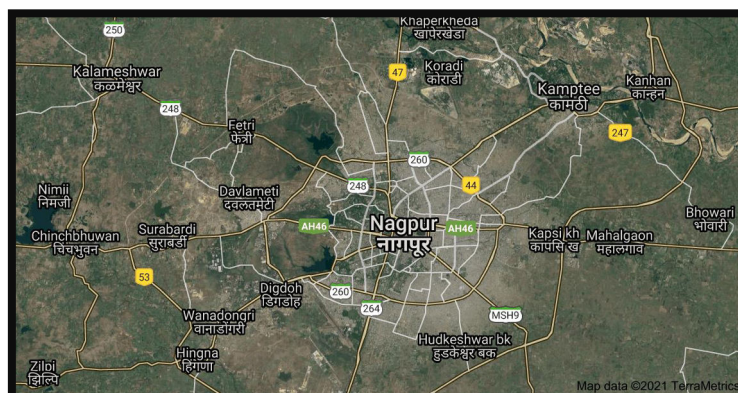


Fig.4 : Geographical map of Nagpur

(Source:<https://cdn.s3.waas.gov.in/s3d1f491a404d6854880943e5c3cd9ca25/uploads/2018/03/2018030559.jpg>)

Two types of networks used to develop model by using Tanhaxon as transfer function and learning rule. It has four types of architecture viz., 3-1-1-1, 3-1-2-1, 3-2-1-1, 3-2-2-1 with training, testing data and validation in the ratio of 50-



15-35, 60-15-25, 70-15-15 an epoch values of 500,1000 and 2000 are utilized for analysis. For developing of ANN models the neurosolutions trail software is used. The developed network consists of number of interconnected nodes (processing elements), arranged into three basic layers viz., input, hidden and output layers. Hence, in the present study ANN has been used along with statistical measures for short term prediction of pm2.5 and pm10 considering its average and maximum values. For the purpose of prediction study, Multi Layer Perception (MLP) Network, Generalized Feed Forward (GFF) Network, Network models of ANN are used to forecast one day ahead pm2.5 and pm10 concentration observed in Nagpur city of Maharashtra, India by using past data of four years on hourly basis from the period 1-1-2017 to 30-4-2021 was collected from CPCB (Central Pollution Control Board).

The sample network of 3-2-2-1 architecture is as shown in Fig 3 below.

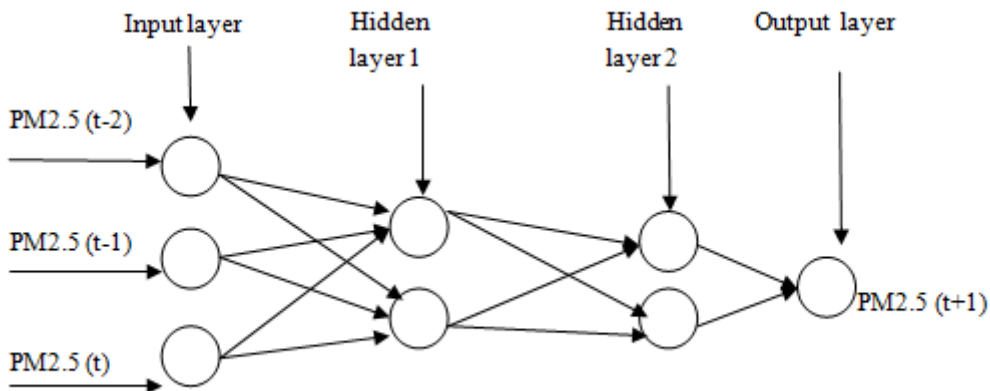


Fig 5. Typical Neural Network

The model used for predicting PM2.5 and PM10 values of current day using three previous values are represented in the above fig.5 $PM2.5(t+1) = f [PM2.5(t), PM2.5(t-1), PM2.5(t-2)]$ Where 't' is day for which prediction is done.

Out of large number of ANN models developed for the data sets, best ANN model is investigated in detail. Present study the goodness-of-fit measures considered to evaluate the developed models are mean square error (MSE), mean absolute error (MAE), and coefficient of correlation (R) between the forecasted and observed 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 modelled and the value equal to one shows perfect. In this study two network viz multilayer perceptron (MLP) and generalized feed forward neural network (GFN) was used .

III. RESULT

The number of trials has been performed on the observed data and finally it is concluded that following set of various parameters has been considered for detailed experimentation, Best Models for average values of pm2.5 with each set of epoch. Out of which the network GFF has given the highest value of R 0.79 and the second highest is given by the network of MLP which has given the value of R 0.78.

Table no 1: Summary of ANN model developed for average PM2.5 Concentration

Network	Epoch	Architecture	Training: Testing: Validation	MSE	NMSE	MAE	R
GFF	1000	3-1(1)-1	50-15-35	12.23	0.211	77.45	0.79
MLP	500	3-2(2)-1	50-15-35	425.45	0.229	19.22	0.52
GFF	2000	3-2(1)-1	60-15-25	523.41	0.130	19.29	0.74
MLP	500	3-2(2)-1	60-15-25	98.69	0.123	21.43	0.78

The table no 1 shows us the best model obtained for average pm2.5 by network of MLP and GFF with every architecture used 3-1-1-1,3-2-2-1,3-2-1-2,3-2-2-1 and training :Testing: validation values for each epoch is different. Three different epochs were used that is 500,1000 and 2000 out of which 1000 and 500 gives the highest value of R fig 6 and fig7 shows timeseries and scatter plot of observed vs predicated value of max R value model respectively.

Table no 2: Summary of ANN model developed for Max PM2.5 Concentration

Network	Epoch	Architecture	Training: Testing: Validation	MSE	NMSE	MAE	R
MLP	1000	3-1(1)-1	60-15-25	210.45	0.181	14.79	0.79
GFF	500	3-1(2)-1	60-15-25	598.69	0.123	20.14	0.78
MLP	500	3-2(1)-1	60-15-25	318.69	0.113	19.24	0.78
GFF	500	3-2(2)-1	60-15-25	598.69	0.123	20.14	0.78

In this table no .2 best model are obtained for maximum pm2.5 by network of MLP and GFF with every architecture used 3-1-1-1,3-2-2-1,3-2-1-2,3-2-2-1 and training :Testing: validation values for each epoch is different. Three different epochs were used that is 500,1000 and 2000 out of which 1000 epoch model of MLP network gives the highest R 0.79 value.

Table no 3 : Summary of ANN model developed for average PM10 Concentration

Network	Epoch	Architecture	Training: Testing: Validation	MSE	NMSE	MAE	R
MLP	500	3-1(1)-1	50-15-35	29.45	0.129	18.72	0.62
GFF	500	3-1(2)-1	60-15-25	73.63	0.271	19.51	0.77
MLP	500	3-2(1)-1	50-15-35	22.23	0.226	72.39	0.77
GFF	500	3-2(2)-1	50-15-35	10.45	0.109	19.73	0.62

In this table no.3 best model are obtained for average pm10 by network of MLP and GFF with every architecture used 3-1-1-1,3-2-2-1,3-2-1-2,3-2-2-1 and training :Testing: validation values for each epoch is different. Three different epochs were used that is 500,1000 and 2000. The highest value of R is given by network GFF with 500 epoch i.e 0.77 and the same R value is given by network MLP with 500 epoch but less MSE 22.23.

Table no 4 : Summary of ANN model developed for Max PM10 Concentration

Network	Epoch	Architecture	Training: Testing: Validation	MSE	NMSE	MAE	R
MLP	500	3-1(1)-1	50-15-35	11.45	0.119	19.71	0.52
MLP	500	3-1(2)-1	50-15-35	63.92	0.121	18.75	0.52
GFF	500	3-2(1)-1	60-15-25	78.63	0.737	18.49	0.52
MLP	1000	3-2(2)-1	60-15-25	21.05	0.178	19.79	0.79

In this table no .4 best model are obtained for maximum pm10 by network of MLP and GFF with every architecture used 3-1-1-1,3-2-2-1,3-2-1-2,3-2-2-1 and training :Testing: validation values for each epoch is different. In this case also MLP network developed model gives maximum R 0.79. fig 8 and fig 9 shows timeseries and scatter plot of observed vs predicated value of max R value model respectively.

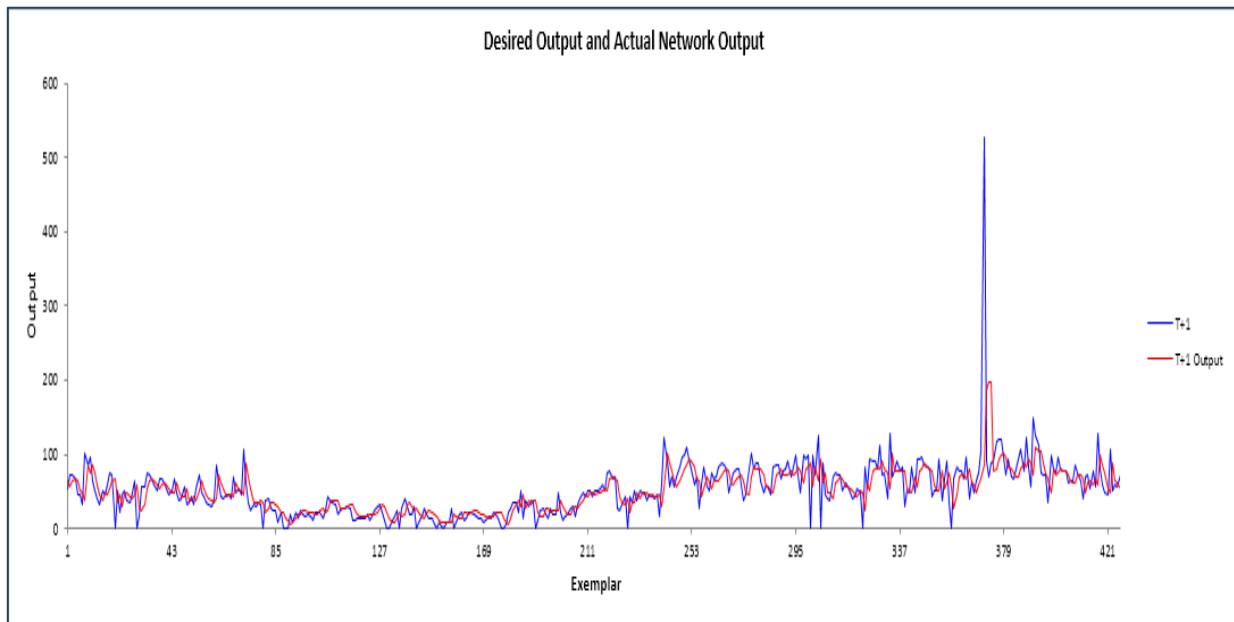


fig 6: Time series plot MLP 3-2(1)-1model For Avg 2.5PM value

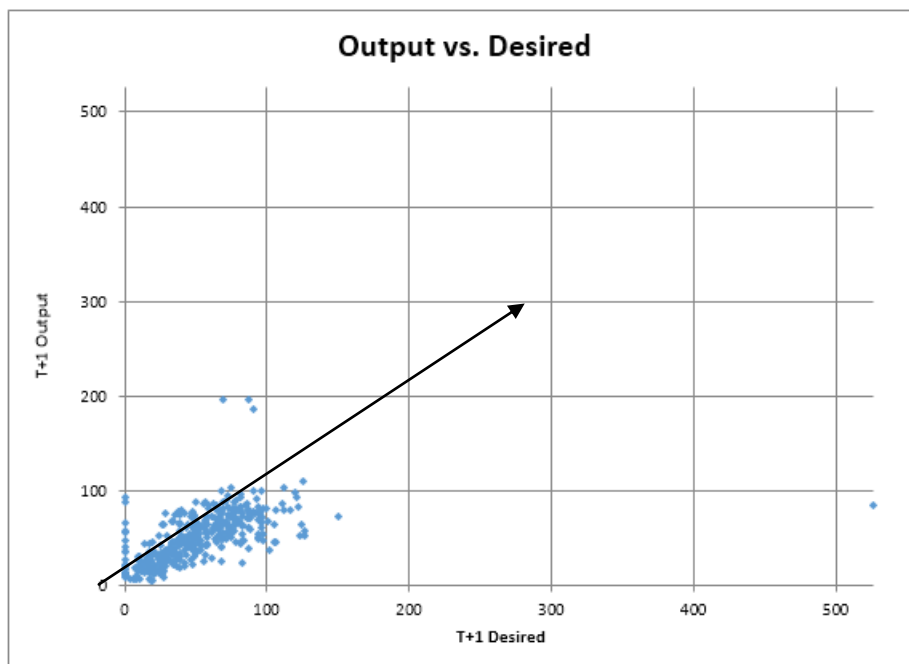


Fig7: Scatter plot MLP 3-2(1)-1model For Avg 2.5PM value

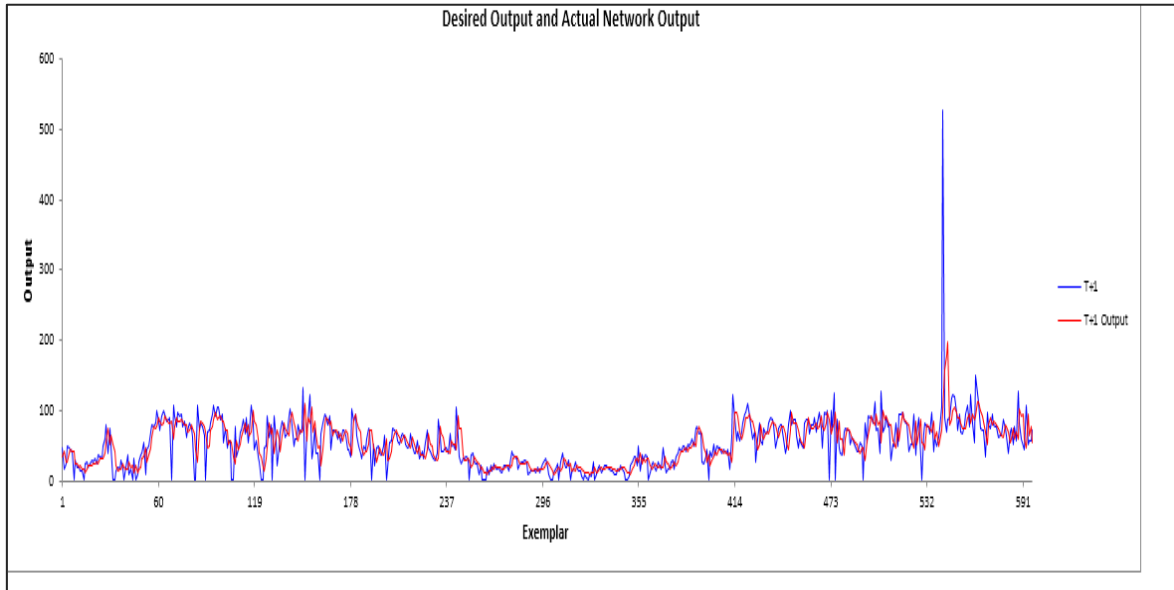


Fig8: Time series plot MLP 3-2(2)-1 model for Max PM10 Value

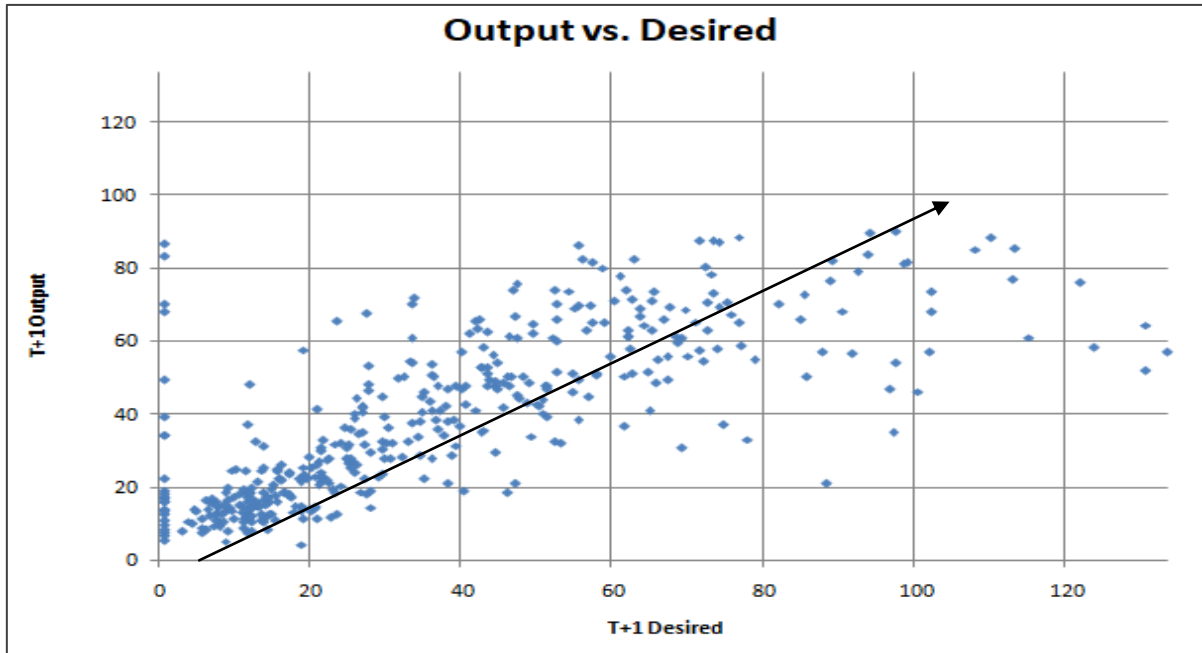


Fig9: Scatter plot MLP 3-2(2)-1 model for Max PM10 value

IV. CONCLUSION

The system for forecasting ANN modeling tool helps to predict next date pollution details based on basic parameters and analyzing pollution details and forecast future pollution. This system gives solution for air pollution prediction. However, it also provides relevant information about the areas people are actually interested in. The user friendly



approaches like ANN techniques is the attractive alternate tool to numerical modelling and real-time prediction. This is very important to get RSPM values for real time as application in following:

REFERENCES

1. Adharsh K C1, Muhammad Fayis, "Prediction Of Water Quality Parameters Of Kallai River Using Artificial Neural Network" international Research Journal Of Engineering And Technology, E-issn, pp125-137, Vol: 07 Issue: 07 | July 2020
2. Sachin Mane, Sonaje N. P., 2018, "Development of SO₂ Prediction Model with
3. Artificial Neural Network for Pune City" International Journal of Application or Innovation in Engineering & Management Vol7, Issue 6, p.g 69-75, June 2018
4. Sachin Mane, Sonaje N. P., 2018, "Artificial Neural Network model for Respirable Particulate Matter" International Journal of Technical Innovation in Modern Engineering & Science, Impact Factor: 5.22 (SJIF-2017), e-ISSN: 2455-2585 Vol 4, Issue 6, June-2018, pp 1479-1485.
5. Madhavi Anushka Elangasinghe, "development Of An Ann-based Air Pollution Forecasting System With Explicit Knowledge Through Sensitivity Analysis, atmospheric Pollution Research 5(2014), pp 696 -708 11-lucica Barbes *, Corneliu Neagu, the Use Of Artificial Neural Network (Ann For Prediction Of Some Airborne Pollutants Concentratio In Urban Areas, journal Of Chemistry, Special Issue "Challenges And Opportunities In The Application Of Chemo metrics In The Pharmaceutical And Food Science Industry
6. <https://cpcb.nic.in/>
7. <https://www.educative.io/>
8. <https://www.ijert.org/a-comparative-study-of-lenn-and-chflann-algorithm-for-chronic-kidney-disease-data-classification>
9. <https://www.educative.io/edpresso/what-is-a-multi-layered-perceptron>
10. <https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/Architecture/feedforward.html>
11. <http://www2.psych.utoronto.ca/users/reingold/courses/ai/cache/neural2.html>
12. <https://cs.stanford.edu/people/eroberts/courses/soco/projects/2000-01/neural-networks/Architecture/feedforward.html>
13. <https://www.vskills.in/certification/tutorial/topologies-and-models/>
14. <http://www2.psych.utoronto.ca/users/reingold/courses/ai/cache/neural2.html>



**Impact Factor:
7.569**



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details