A neural network approach to forecast particulate matter concentration in Manali area of Chennai City

Nadeem Imran* and Sheik Uduman P.S.
Department of Mathematics, B.S Abdur Rahman Crescent Institute of Science and Technology, Chennai- 600048, INDIA
*malikimrannadeem218@gmail.com

Abstract
Air pollution is one of the threatening menaces confronting all over the globe in recent decades. Among the air pollutants, $PM_{2.5}$ is one of the major alarming components rising at a rapid pace in the metropolitan city of Chennai due to the fast expansion of industrialization and urbanization. An early observation for tackling the rise in particulate matter concentration ($PM_{2.5}$) levels requires precise prediction.

In this regard, the present study employs a Multilayer Perception Neural Network (MLPNN) technique using the Levenberg-Marquardt optimisation training algorithm for forecasting the one-day concentration $PM_{2.5}$. The model evaluation statistics R^2 and index of the agreement have been utilized for assessing the forecasting accuracy. The results confirm that the ANN_6 model is the most suitable for acquiring the almost error-free model to achieve real-time forecasting of $PM_{2.5}$ concentration.

Keywords: Air pollution, Levenberg-Marquardt algorithm, Neural network, $PM_{2.5}$, Time series

Introduction

Air pollution is one of the major environmental issues experienced by almost all the nations of the globe. In recent years, the air pollutants concentration level has been rising to a much higher level than the level prescribed by several global agencies all over the globe that are framed for safeguarding humanity from this drastic environmental threat. Air pollution has become a prime issue in Chennai city mainly due to the high rising level of $PM_{2.5}$ from the past decade. The main cause of the rise in $PM_{2.5}$ level in the city is road traffic, particularly the vehicles having poor standards. The environmental agencies have fixed an annual and daily limit to manage this pollutant.

A significant aspect to maintain public health is the quality of air which is mainly dependent on particulate matter.[1,7,12] The effect of suspended particulate less than or equal to 2.5 micrometres ($PM_{2.5}$) on the health factor has been well accepted. A rise in the level of $PM_{2.5}$ has influenced life expectancy by piercing the lungs and resulting in acute health issues of lung cancer, asthma and certain other respiratory diseases.[39] Another study declared that the $PM_{2.5}$ level rise can result in a rise in mortality rate on a succeeding day. [15,25]

The artificial neural network (ANN) model is one of the most commonly employed designs of artificial intelligence models in various fields of research including air pollution.[2,3,10,20] The ANN model can be employed to evaluate the sophisticated non-linear functions in high dimensional spaces compared to the classical statistical techniques. ANN model has been adopted as an effective technique to estimate air pollutants levels, particularly in those urban areas where surveillance centres assessed the concentration of pollutants in addition to metrological factors.[10,20]

Recent studies have depicted that the output of the ANN technique is better than the statistical linear techniques as these techniques are unable to bring a quantitative assessment of the link between pollution and air due to complex and non-linear characteristics.[4,11,26,40] Several statistical models are available to forecast the air pollutant concentration like principal component analysis with multiple linear regression (PCA-MLR) applied by Ul-Saufie et al[34], autoregressive integrated moving average (ARIMA) model employed by Diaz-Robles et al[8], nearest neighbour model (NNM) applied by Perez et al[24], hidden Markov model used by Sun et al[32] and artificial neural network (ANN).[6,16] ANN and multiple linear regression models were compared by Ceylan et al[5] to forecast CM_{10} concentration.

Among these models, ANNs have been justified to be quite efficient due to their competency in achieving non-linear mapping, self-adaption and robustness in forecasting $PM_{2.5}$.[17,19,23,33,35] The various designs of forecasting feed-forward neural network, backpropagation neural network, wavelet neural network, Elman network and general regression neural network make ANN modelling an exceptional technique for forecasting. The study conducted by Perez et al[23] employed an ANN model to assess the PM_{10} aggregation. They revealed that the multilayer neural network has a better capability of forecasting compared to the linear model.

Kukkonen et al,[14] conducted contrastive research on five neural networks (NN) models by employing the linear statistical model and deterministic model for forecasting PM_{10} and NO_2 aggregation levels. They observed that the non-linear ANN models provide marginally better results than both deterministic and linear statistical models.

One of the major issues in the case $PM_{2.5}$ is to forecast maximum concentration on the following day. This is essential as commuters can handle their routes or delay their trips if an intense $PM_{2.5}$ concentration is anticipated. The objective of the present study is to develop a relationship between $PM_{2.5}$ concentrations and metrological parameters by employing ANN modelling for achieving a good degree of forecasting of $PM_{2.5}$ concentrations on the following day. The one-day-ahead forecasting results of $PM_{2.5}$ attained in the present study can be quite beneficial for environmental protection authorities to acquire mandatory preventative measures for implementing an efficient policy to safeguard the environment from the excessive concentration of $PM_{2.5}$.

There is no evidence of forecasting the threatening pollutant like particulate matter ($PM_{2.5}$) in metropolitan cities of India like Chennai, Mumbai, Calcutta and Hyderabad by employing the ANN modelling. As a consequence, this study tries to fill the gap using a district of Chennai city as its case study. With this aim, our study opts for a best-fitted ANN model from different tested ANN models for forecasting $PM_{2.5}$ pollutants levels in the Manali area of Chennai city.

Data Collection

The diffusion and movement of the particulate matter like $PM_{2.5}$ in the environment are influenced by certain metrological factors such as wind direction, wind speed, temperature, relative humidity, atmospheric pressure, precipitation and solar radiation. These factors have been considered to have a major impact on the maximum PM_{10} concentration for the following day as studied earlier. The $PM_{2.5}$ data and metrological parameters data collected for the Manali district of Chennai City have been used as input parameters.

The daily 24-h mean value of $PM_{2.5}$ concentrations and the metrological data of the Manali have been collected from the Central Pollution Control Board of India and Indian Metrological Department respectively for the period of first of July, 2018 to 30 of Sept, 2019. A data set comprising 457 entries were used for the ANN model. Out of which 70% of the data was used for training, 15% for validation and the remaining 15% for testing purposes. The input layer consists of past data on $PM_{2.5}$, temperature, atmospheric pressure, wind speed, relative humidity and visibility.

Material andMethods

An artificial Neural network is a computing method influenced by biological neural networks.[29] One of the commonly used ANN methods is the Multilayer Perception Neural Network (MLPNN). MLPNN structure usually comprises an input layer with single or multiple hidden layers and an output layer. All three layers consist of a fundamental component known as a neuron or a node. The nodes are interrelated and a weight component indicating the link between the two nodes characterizes the synapses. Each node receives the input values, operates them and then transfers the values to the succeeding layer. Moreover, each node is a neuron utilizing the nonlinear activation function excluding the input nodes. The process is progressed by utilizing weight and its transfer function for obtaining an output value.[18]

Figure 1: The star depicts the location of Manali on the map of India

The output of the MLPNN is represented in the matrix form $y = f(w(f'(w'm + b')) + b)$ where w is the weight matrix that links the hidden layer to the output layer, w' is the weight matrix that links the input layer to the hidden layer and f acts as a transfer function for activating the hidden layer to the output layer, f' acts as a transfer function that links the input layer to the hidden layer, b' define the bias for the hidden layer, b defines the bias for the output layer, m and y refer to the input vector and output vector respectively of the MLPNN.

MLPNNs are competent in operating on input information by utilizing a common method of back-propagation (BP) technique to evaluate synaptic weights.[30] BP is a supervised learning technique based upon the Gradient Descent method for minimising the error in the network by lowering the error curve gradient. The procedure needs a training pattern $(m_1, t_1), (m_2, t_2).......(m_s, t_s)$ made of S ordered pairs of N and M dimensional vectors having N number of input and M number of output. The learning process is initiated by varying the values of synaptic weights of the MLPNN. The process of learning mathematically deals with the reduction of error function as $E = \sum_{i=1}^{s} \|y_i - t_i\|^2$ where y denotes the output vector and t specifies the target.

After the output is contrasted to the target, the error is propagating backwards through the network to create an arrangement in the synaptic weights and biases of the network. This leads to minimising the global difference between the output of the network and the target .[9] Once the minimum difference is achieved, the process of learning ends. This process serves as an entire cycle in which all the information is passed through the network.

Subsequently, in the MLPNNs network, the other determinants are: the number of layers, the transfer function between the various layers, normalization of the data as per the selected transfer function, input and output parameters, training algorithm and methods employed to define several neurons in the hidden layers.

The present study employs the Levenberg-Marquardt (LM) training algorithm for regulating the MLPNN weights and a total of 1000 epochs were applied. LM uses iterations to determine the local minimum of multivariate functions written as a sum of the square for several non-linear and real-valued functions. It is a standard method of curve fitting for the non-linear least square problem by minimising the sum of the square of the error function.

Further, transfer functions like log-sigmoid, tangent sigmoid and purelin functions to stimulate the connection between the neurons of various layers for forecasting the model probability as output have been used. The formula for Linear (Purelin) transfer function is $f(x) = \frac{1}{1 + e^{-mx}}$ and hyperbolic tangent sigmoid transfer function (Tansig) is $f(x) = \frac{2}{1 + e^{-mx}} - 1$

where x defines the input value of the transfer function, m defines the slope of the parameter and $f(x)$ is the output value in the given formulas. The complete input dataset is normalized to acquire a uniform effect of each input value present in the ANN model. The normalization equation can be defined as $MI_{ij} = \frac{I_{(i,j)} - \min(j)}{\max(j) - \min(j)}$ where I denotes the input value, MI denotes the standardized value, i represents the number of patterns and j signifies the measured value of.[13]

The presence of several hidden layers may lead to overfits and the model may not comply with the new inputs. Therefore a unique hidden layer network was employed to evaluate the variables of the network. Between the input and the hidden layers, a unique hidden layer and a hyperbolic tangent sigmoid transfer function were employed. A linear transfer function was employed between the hidden and output layers. MATLAB R2017b Neural Network Toolbox has been used to calculate ANN.

Transfer functions are used to determine a layer's output from its net input. The mathematical functions LEARNGD and LEARNGDM have been employed to act as learning functions for ing the weight and biases of a network. LEARNGD is the gradient descent weight and bias learning function and LEARNGDM is the gradient descent with momentum weight and bias learning function.

Estimation of the forecasting efficiency

The efficiency of the multiples developed ANN models proposed by Willmott et al[36,37] has been statistically evaluated by employing the mean square error as $RMSE = \frac{1}{n}\sum_{i=1}^{n}(Y_{fi} - Y_{ai})^2$ and mean absolute error as

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_{fi} - Y_{ai}|.$$

Coefficient of determination is:

$$R^2 = 1 - \left(\frac{\sum_{i=1}^{n}(Y_{fi} - Y_{ai})^2}{\sum_{i=1}^{n}(Y_{fi} - \bar{Y})^2}\right)$$

Index of the agreement is:

$$d = 1 - \frac{\sum_{i=1}^{n} (Y_{fi} - Y_{ai})^2}{\sum_{i=1}^{n} \left(| Y_{fi} - \bar{Y} | + | Y_{ai} - \bar{Y} | \right)^2} \quad 0 \le d \le 1$$

where n denotes the number of data, Y_{fi} is the forecasted value, Y_{ai} is the actual value from the data i and $\bar{Y}$ is the average of the actual values. R^2 is a quantity used for illustrating how better the data conform to a statistical model. The R^2 value lies between 0 and 1. The more the value of R^2 lies close to 1, the more the model is regarded as authentic. RMSE is employed to determine the rate of error present in a regression model as well as to indicate the standard deviation of the model forecasting error.

The least possible value of RMSE is considered a decent value for the accuracy of the model. MAE is employed to estimate how close the predicted value is to the actual values. The index d determines the limit to which signs and the magnitude of the forecasted values about the $\bar{Y}$ are associated with the forecasted deviations about $\bar{Y}$ and assess the variation not only in Y_{fi} and Y_{ai} but also in the proportionality of Y_{fi} and Y_{ai}.[27] The value of the index d ranges from 0.0 to 1.0. The former value implies no agreement while the latter defines the ideal agreement. The values of R^2, RMSE, MAE and index of agreement (d) define the ability of the model to explain the real character of the system.

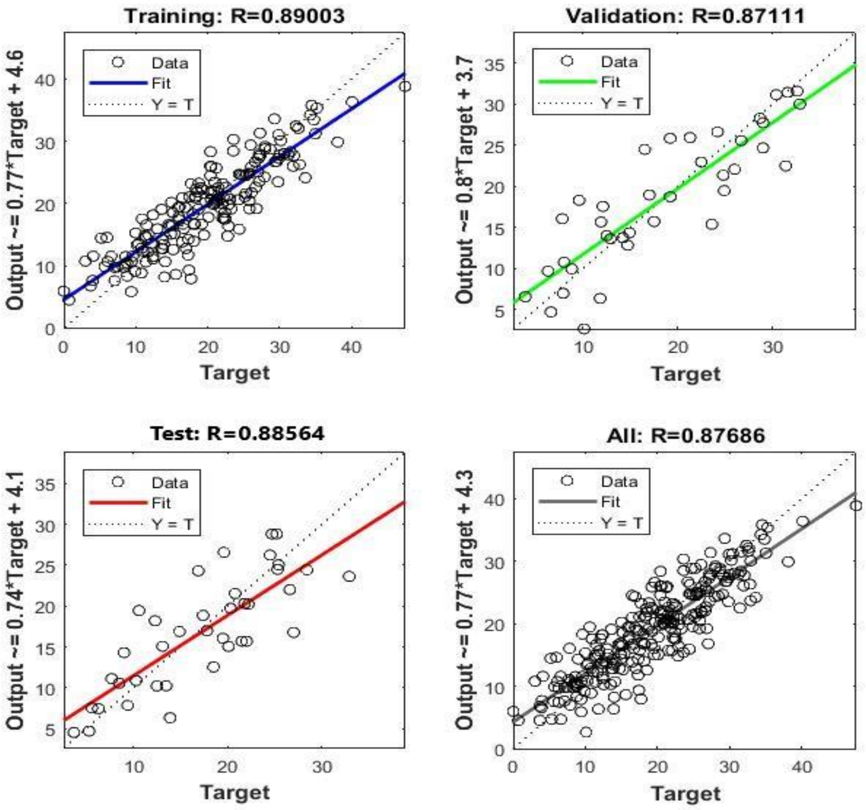

Figure 2: Regression graphs of training, testing, validation and complete of all

Table 1
The comparative performance of various ANNs model

Model	Adaption learning function	Transfer function	Number of hidden neurons	RMSE	MAE	R^2	Index of agreement (d)
ANN_1	Learngd	Purelin-Purelin	10	21.432	11.458	0.727	0.754
		Log-sigmoid		19.672	13.587	0.793	0.684
ANN_2	Learngd	Tansig-Purelin	10	20.543	10.755	0.734	0.763
		Log-sigmoid		18.359	14.563	0.764	0.693
ANN_3	Learngd	Tansig-Tansig	10	19.763	12.567	0.686	0.688
		Log-sigmoid		15.983	11.334	0.834	0.763

ANN_4	Learngdm	Purelin-Purelin		21.534	10.773	0. 754	0.734
		Log-sigmoid	10	18.556	15.335	0.845	0.794
ANN_5	Learngdm	Tansig-Purelin		19.445	10.453	0.734	0.835
		Log-sigmoid	10	20.554	16.332	0.795	0.823
ANN_6	Learngdm	Tansig-Tansig		14.789	8.739	0.881	0.873
		Log-sigmoid	12	23.776	10.781	0.783	0.794
ANN_7	Learngd	Purelin-Purelin		20.443	9.654	0.694	0.773
		Log-sigmoid	12	19.554	16.972	0.776	0.762
ANN_8	Learngd	Tansig-Purelin		22.453	11.564	0. 745	0.763
		Log-sigmoid	12	18.554	14.209	0.674	0.843
ANN_9	Learngd	Tansig-Tansig		18.343	10.786	0.734	0.698
		Log-sigmoid	12	19.556	13.342	0.809	0.796
ANN_{10}	Learngdm	Purelin-Purelin		18.224	11.778	0.679	0.793
		Log-sigmoid	12	18.347	17.345	0.745	0.832
ANN_{11}	Learngdm	Tansig-Purelin		18.542	10.785	0.608	0.823
		Log-sigmoid	14	19.897	16.653	0.689	0.739
ANN_{12}	Learngdm	Tansig-Tansig		20.543	13.675	0.787	0.787
		Log-sigmoid	14	21.674	19.453	0.794	0.833
ANN_{13}	Learngd	Purelin-Purelin		19.653	17.559	0.864	0.846
		Log-sigmoid	14	20.557	17.342	0.678	0.794
ANN_{14}	Learngd	Tansig-Purelin		16.780	14.786	0.796	0.830
		Log-sigmoid	14	19.564	17.452	0. 745	0.843
ANN_{15}	Learngd	Tansig-Tansig		18.562	11.335	0. 837	0.793
		Log-sigmoid	14	23.654	16.632	0.779	0.795
ANN_{16}	Learngdm	Purelin-Purelin		17.764	13.564	0. 745	0.732
		Log-sigmoid	16	19.675	14.543	0.792	0.805
ANN_{17}	Learngdm	Tansig-Purelin		23.564	15.602	0.809	0.769
		Log-sigmoid	16	17.124	19.872	0.835	0.796
ANN_{18}	Learngdm	Tansig-Tansig		20.689	13.267	0.789	0.723
		Log-sigmoid	16	23.302	17.375	0.808	0.693
ANN_{19}	Learngd	Purelin-Purelin		21.673	14.118	0.773	0.675
		Log-sigmoid	16	20.263	17.673	0.789	0.784
ANN_{20}	Learngd	Tansig-Purelin		19.443	13.334	0.754	0.854
		Log-sigmoid	16	21.342	18.004	0.745	0.854
ANN_{21}	Learngd	Tansig-Tansig		17.453	15.564	0.822	0.832
		Log-sigmoid	18	23.785	16.653	0.859	0.813
ANN_{22}	Learngdm	Purelin-Purelin		22.564	12.786	0.839	0. 793
		Log-sigmoid	18	22.702	18.367	0.798	0.784
ANN_{23}	Learngdm	Tansig-Purelin		24.786	13.453	0.689	0.823
		Log-sigmoid	18	19.372	19.334	0.754	0.810
ANN_{24}	Learngdm	Tansig-Tansig		19.856	14.335	0.773	0.830
		Log-sigmoid	18	22.121	17.954	0.794	0.778
ANN_{25}	Learngd	Purelin-Purelin		20.654	12.402	0.804	0.798
		Log-sigmoid	18	17.453	19.302	0.763	0.773
ANN_{26}	Learngd	Tansig-Purelin		18.231	13.891	0.763	0.694
		Log-sigmoid	20	18.245	18.783	0.834	0.754
ANN_{27}	Learngd	Tansig-Tansig		21.432	12.453	0.783	0.783
		Log-sigmoid	20	20.556	17.372	0.853	0.765
ANN_{28}	Learngdm	Purelin-Purelin		20.543	13.564	0.745	0.752
		Log-sigmoid	20	17.334	18.653	0.823	0.834
ANN_{29}	Learngdm	Tansig-Purelin		19.653	12.261	0. 693	0.778
		Log-sigmoid	20	22.344	19.543	0.774	0.784
ANN_{30}	Learngdm	Tansig-Tansig		18.675	14.994	0.789	0.834
		Log-sigmoid	20	21.453	18.453	0.742	0.784

Results and Discussion

Multiple pairs of the transfer function for the hidden layer and output layer with varied adaption functions have been evaluated by changing the number of neurons in the hidden layer to attain the optimal performance of the ANN model. Various combinations of tangent sigmoid and linear (purelin) transfer functions have been evaluated to obtain the best set to produce better results. The training efficiency results of the various ANNs network with transfer functions have been represented in table 1. The results obtained in table 1 show that ANN_6 network attained the better statistical values than the other ANNs network evaluated in terms of the least value RMSE and MAE and the highest value R^2 and d. Based on the obtained optimum model, the simulation of the network has been performed and depicted in figure 2.

Figure 2 depicts that the training of the ANN_6 model is quite efficient as the value of the R has been found 0.88564 for testing and 0.87111 for validation. ANN_6 network includes 5 inputs: parameters temperature, atmospheric pressure, wind speed, relative humidity and visibility with 12 neurons in the hidden layer and one output variable- $PM_{2.5}$ concentration. The best-fits MLPNN model forecasting has been achieved with the learning function LEARNGDM and with the pair of Tansig-Tansig transfer functions producing a 5-12-1 MLPNN network. The RMSE, MAE, R^2 and d values for the best-fit ANN_6 model are 14.789, 8.739, 0.881 and 0.873 respectively attained in the present study.

The results attained in this study are quite convincing when compared to previous studies. Robeson and Steyn[28] obtained the RMSE values for forecasting the maximum of ozone as 19.5 and 15.5 for the two locations they studied. Xie[38] attains the RMSE values as 15.84, 16.43 and 20.55 with three different methods of ANNs for forecasting the $PM_{2.5}$ concentration. The value of d attained by Sharma et al[31] for forecasting the ambient air quality of Delhi city is 86.93%.

Moreover, the comparisons of the daily (24-h mean value) concentrations of $PM_{2.5}$ value have been done with the daily admissible concentration of $PM_{2.5}$ as prescribed by the Central Pollution Control Board of India for the Manali area of Chennai.

The figure shows the $PM_{2.5}$ concentrations for the period of 01/01/2018 to 01/09/2019. The fluctuation signifies daily (24-h mean value) records of $PM_{2.5}$ concentrations and the horizontal line defines the daily admissible limit (60 μg/m³). The figure clearly shows that there are hardly a few days in which the actual $PM_{2.5}$ concentrations is less than the admissible limit of $PM_{2.5}$. This wide rise in the level of $PM_{2.5}$ from the permissible limit leads to forecasting mandatory.

Conclusion

The present study has attempted to attain the best-fit ANN model by testing multiple models based on statistical values to forecast the one-day $PM_{2.5}$ concentrations in the Chennai city. The results attained for forecasting are almost error-free. The level of forecasting accuracy attained using ANN modelling is quite satisfactory and requires further advancement to attain absolute accuracy in forecasting

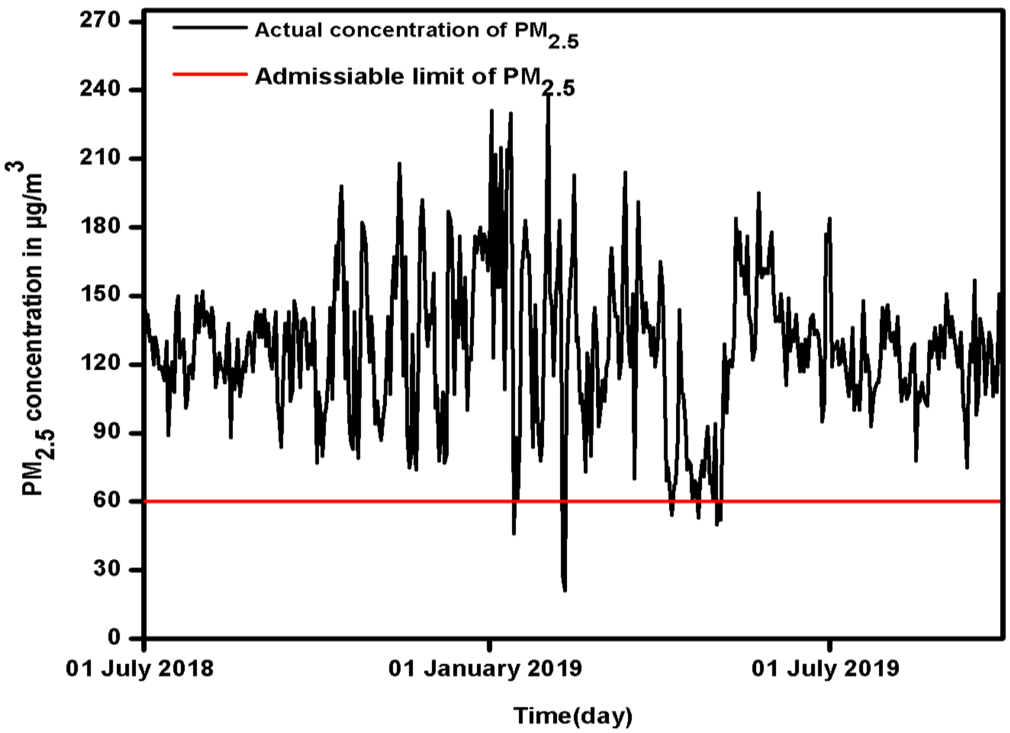

Figure 3: The comparison between the actual $PM_{2.5}$ level and the admissible $PM_{2.5}$ level

The actual daily concentration of $PM_{2.5}$ on comparing with a permissible limit of National Ambient Air Quality Standards of India along with forecasting accuracy of ANN modelling can prove to be quite beneficial in framing an appropriate strategy to manage the highly deteriorating level of $PM_{2.5}$ level in Chennai City.

References

1. Ahmad A., Othman M.R. and Latif M.T., Early study of Surfactants in Indoor Dust and Their Connection With Street Dust, *Int. J. Environ. Res.*, **3(3)**, 403-410 **(2009)**

2. Altiner M. and Kuvvetli Y., Prediction of Grinding Behavior of Low-Grade Coal Based on Its Moisture Loss by Neural Networks, *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, **39(1-2)**, 1250-1257 **(2017)**

3. Biancofiore F., Busilacchio M., Verdecchia M., Tomassetti B., Aruffo E., Bianco S., Di Tommaso S., Colangeli C., Rosatelli G. and Di Carlo P., Recursive Neural Network Model for Analysis and Forecast of PM_{10} and $PM_{2.5}$, *Atmospheric Pollution Research*, **8(4)**, 652–659 **(2017)**

4. Blanes V.V., Cantuaria M.L. and Nadimi E.S., A Novel Approach for Exposure Assessment in Air Pollution Epidemiological Studies Using Neuro-Fuzzy Inference Systems: Comparison of Exposure Estimates and Exposure-Health Associations, *Environmental Research*, **154**, 196–203 **(2017)**

5. Ceylan Z. and Bulkan S., Forecasting PM_{10} levels using ANN and MLR: A case study for Sakarya City, *Global NEST Journal*, **10(10)**, 20-30 **(2018)**

6. Chattopadhyay S. and Chattopadhyay G., Modeling and prediction of monthly total ozone concentrations by use of an artificial neural network based on principal component analysis, *Pure Appl Geophys.*, **169(10)**, 1891–1908 **(2012)**

7. Chiou P., Tang W., Lin C.J., Chu H.W. and Ho T.C., Comparison of Atmospheric Aerosols between Two Sites over Golden Triangle of Texas, *Int. J. Environ. Res.*, **3(2)**, 253-270 **(2009)**

8. Díaz-Robles L., Ortega J., Fu J., Reed G., Chow J., Watson J. and Moncada H.J., A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: the case of Temuco, Chile, *Atmos Environ.*, **42**, 8331–8340 **(2008)**

9. Emamgholizadeh S., Kashi H., Marofpoor I. and Zalaghi E., Prediction of Water Quality Parameters of Karoon River (Iran) by Artificial Intelligence-Based Models, *International Journal of Environmental Science and Technology*, **11(3)**, 645–656 **(2014)**

10. Fang D. and Wang J., A Novel Application of Artificial Neural Network for Wind Speed Estimation, *International Journal of Sustainable Energy*, **36(5)**, 415–429 **(2017)**

11. Gardner M.W. and Dorling S.R., Artificial Neural Networks (the Multilayer Perceptron)-a Review of Applications in the Atmospheric Sciences, *Atmospheric Environment*, **32(14–15)**, 2627–2636 **(1998)**

12. Goswami S.H., Road Traffic Noise: A Case Study of Balasore Town, Orissa, India, *Int. J. Environ. Res*., **3(2)**, 309-316 **(2009)**

13. Keskin M.E. and Terzi Ö., Artificial Neural Network Models of Daily Pan Evaporation, *Journal of Hydrologic Engineering*, **11(1)**, 65-70 **(2006)**

14. Kukkonen J., Partanen L., Karppinen A., Ruuskanen J., Junninen H., Kolehmainen M., Niska H., Dorling S., Chatterton T., Foxall R. and Cawley G., Extensive evaluation of neural network models for the prediction of NO_2 and PM_{10} concentrations, compared with adeterministic modelling system and measurements in central Helsinki, *Atmospheric Environment*, **37**, 4539-4550 **(2003)**

15. Liu J.C. and Peng R.D., Health effect of mixtures of ozone, nitrogendioxide and fine articulates in 85 US counties, *Air Qual Atmos Health*, **11**, 311–324 **(2018)**

16. Luna A., Paredes M., Oliveira G. and Correa S., Prediction of ozone concentration in tropospheric levels using artificial neural networks and support vector machine at Rio de Janeiro, Brazil, *Atmos Environ*., **98**, 98–104 **(2014)**

17. Mckendry I., Evaluation of artificial neural networks for fine particulate pollution (PM10 and PM2.5) forecasting, *J Air Waste Manage Assoc*., **52**, 1096-1101 **(2002)**

18. Moazami S., Roohollah N., Bahman J.A., Bijan Y., Sadegh P. and Salman S., Reliable Prediction of Carbon Monoxide Using Developed Support Vector Machine, *Atmospheric Pollution Research*, **7(3)**, 412–418 **(2016)**

19. Ordieres J., Vergara E., Capuz R. and Salazar R., Neural network prediction model for fine particulate matter ($PM_{2.5}$) on the US-Mexico border in El Paso (Texas) and Ciudad Juarez (Chihuahua), *Environ Model Softw*., **20**, 547–559 **(2005)**

20. Özgür C. and Erdi T., Prediction of density and kinematic viscosity of biodiesel by Artificial Neural Networks, *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, **39(10)**, 985–991 **(2017)**

21. Ozturk Z.C., Kunt F. and Dursun S., Application of artificial neural network in environmental engineering problems, In 5th International Conference of Ecosystems ICE2015, Tirana, Albania, 368-375 **(2015)**

22. Patrico P. and Jorge R., An integrated neural network model for PM_{10} forecasting, *Atmospheric Environment*, **40(16)**, 2845-2851 **(2006)**

23. Perez P., Trier A. and Reyes J., Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago, Chile, *Atmos Environ*, **34**, 1189-1196 **(2000)**

24. Perez P., Combined model for PM10 forecasting in a large city, *Atmos Environ*, **60**, 271-276 **(2012)**

25. Pope C.A., Ezzati M., Cannon J.B., Allen R.T., Jerrett M. and Burnett R.T., Mortality risk and PM 2.5 air pollution in the USA: an analysis of a national prospective cohort, *Air Qual Atmos Health*, **11**, 245-252 **(2018)**

26. Prasad K., Amit K.G. and Pramila G., Development of ANFIS Models for Air Quality Forecasting and Input Optimization for Reducing the Computational Cost and Time, *Atmospheric Environment*, **128**, 246-62 **(2016)**

27. Rao S.T., Sistla G., Petersen W.B., Irwin J.S. and Turner D.B., Evaluation of the performance of RAM with the regional air pollution study database, *Atmospheric Environment*, **19**, 229–245 **(1985)**

28. Robeson S.M. and Steyn D.G., Evaluation and comparison of statistical forecast models for daily maximum ozone concentrations, *Atmos Environ*, **24B**, 303–312 **(1990)**

29. Rozlach Z., Data-driven Modelling in River Channel Evolution Research: Review of Artificial Neural Networks, *J. Int. Environ. Appl. & Sci.*, **10**, 384-398 **(2015)**

30. Rumelhart D.E., Learning Representations by Error Propagation, In Rumelhart D.E., McClelland J.L. and PDP Research Group, Parallel Distributed Processing, **1 (1986)**

31. Sharma P., Chandra A. and Kaushik S.C., Forecasts using Box–Jenkins models for the ambient air quality data of Delhi City, *Environ Monit Assess*., **157(1–4)**, 105–112 **(2009)**

32. Sun W., Zhang H., Palazoglu A., Singh A., Zhang W. and Liu S., Prediction of 24-hour-average $PM_{2.5}$ concentration using a hidden Markov model with different emission distributions in Northern California, *Sci. Total Environ*., **443**, 93-103 **(2013)**

33. Thomas S. and Jacko R., Model for forecasting expressway fine particulate matter and carbon monoxide concentration: application of regression and neural network models, *J. Air Waste Manage Assoc*, **57**, 480–488 **(2007)**

34. Ul-Saufie A., Yahaya A., Ramli N., Rosaida N. and Hamid H., Future daily PM_{10} concentrations prediction by combining regression models and feed forward back propagation models with principle component analysis (PCA), *Atmos Environ.*, **77**, 621–630 **(2013)**

35. Voukantsis D., Karatzas K., Kukkonen J., Räsänen T., Karppinen A. and Kolehmainen M., Intercomparison of air quality data using principal component analysis and forecasting of $PM_{2.5}$ and PM_{10} concentrations using artificial neural networks, in Thessaloniki and Helsink, *Sci Total Environ*, **409**, 1266–1276 **(2011)**

36. Willmott C.J., On the validation of models, *Physical Geography*, **2**, 184-194 **(1981)**

37. Willmott C.J., Ackleson S.G., Davis R.E., Feddema J.J., Klink K.M. and Legates D.R., Statistics for the evaluation and comparison of models, *Journal of Geophysical Research*, **90**, 8995-9005 **(1985)**

38. Xie J., Deep neural network for PM 2.5 pollution forecasting based on manifold learning, In Proceedings of the International Conference on Sensing, Diagnostics, Prognostics and Control, SDPC 2017, 236–240 **(2017)**

39. Xing Y.F., Xu Y.H., Shi M.H. and Lian Y.X., The impact of $PM_{2.5}$on the human respiratory system, *Journal of Thoracic Disease*, **8(1)**, 69-74 **(2016)**

40. Zafra C., Ángel Y. and Torres E., ARIMA analysis of the effect of land surface coverage on PM_{10} concentrations in a high-altitude megacity, *Atmospheric Pollution Research*, **8(4)**, 660–668 **(2017)**.

(Received 17th May 2022, accepted 22nd July 2022)