Review on Artificial Neural Network Techniques used for Air Quality Index and Air Pollutants Concentration Prediction

[1]Anvesh Parashar, [2]Abhilash Sonker,
[1]M.tech Student, [2]Assistant Professor MITS,
[1]Dept. of CSE & IT,
[1]MITS Gwalior, India, Gwalior, India.

Abstract: Air pollution epidemiology has received major concerns from various Organizations and others groups because of its Various Health Hazard and Effects on General Public. Various Agencies has established Standards of Air Quality Index that indicate Air Quality and Concentration of Air Pollutants that indicate the risk factor involved in inhaling polluted air. The Decreased Cost of Remote Sensing Devices and Mobile Sensor Help us to monitor the data related to Air Quality but Processing and Prediction of Future Air Quality is also a major Challenge to Build Early Warning System and Various Provision to improve the air quality and minimize the effects on general public.

In this paper we reviewed papers based on the machine learning models particularly neural networks with an aspect of data mining. Mainly the focus is on predicting future trends from the historical data collected using various air quality assessment equipment. The time Series data of air quality is a nonlinear regression problem between predictors and predictand. Here the most suitable choice of algorithm for this problem are algorithms that are able to identify the non-linear structure in the data and is able to predict the future trends efficiently like Support vector machine, Multivariate Adaptive Regression Splines, Neural Networks etc. in which neural network recently emerged as the most popular and efficient method for this task. In this paper different algorithms are used with neural network for complementing or compared, to find the best suitable models for a different number of performances metrices.

Index Terms - Data mining, Machine Learning, Artificial Neural Network, Air Pollution, Forecasting.

1. INTRODUCTION

The modelling of atmospheric pollutant concentrations typically involves the development of a functional relationship between concentrations and the other controlling meteorological parameters. Thus, Air pollutants concentration is dependent on various parameters such as

- Meteorological conditions such as wind speed and direction, the amount of atmospheric, the ambient air temperature, the height to the bottom of any inversion aloft that may be present, cloud cover and solar radiation, Source term (the concentration or quantity of toxins in emission or accidental release source terms) and temperature of the material
- Emissions or release parameters such as source location and height, type of source (i.e., fire, pool or vent stack) and exit velocity, exit temperature and mass flow rate or release rate.
- The location, height and width of any obstructions (such as buildings or other structures) in the path of the emitted gaseous plume, surface roughness or the use of a more generic parameter "rural" or "city" terrain.

To support this a number of studies have shown the existence of high correlation between air pollution and meteorological variables [1]

One approach is to use deterministic models, involving fluid dynamic and chemical transformation to model this relationship, while statistical models [2] use field measurements of emission, meteorological parameters and concentrations to develop a linear or non–linear function between the concentration and these predictor variables. Deterministic models are limited by their requirement for detailed knowledge of source parameters, the topographical structures in the immediate surroundings and the detailed meteorology. These data are not always available in practice. Purely statistical models, when adequately trained, may provide good predictions using routinely available data. However, they are limited by their inability to provide insight into dispersion mechanisms and hence may not be used in "what–if scenario analysis" with respect to changes in emission rates and meteorological conditions without performing a procedure such as a sensitivity analysis.
The Historic Data of ambient Air Quality is in the realm of big data and Extraction of useful data from the data need the application of Various Data mining Techniques and For Classifying/Clustering and Prediction of Future Air Quality Need Using Machine Learning Techniques like Support Vector Machines, K-Means, and Artificial Neural Networks (ANN)

Detailed information about emission rates and its source is limited to a few cities. For the aforementioned reasons, more research is needed in developing air quality models that can capture adequately the variability in observation using limited knowledge of the values of parameters. The ability of the ANN technique to capture the nonlinear behavior of complex processes makes it a suitable tool for developing such models. Neural Networks (NN) have applications for both air pollutant time series modelling and air pollutant concentrations forecasting. The ANN models gave better results than other models, especially for the ANN models that were built with non-linear input and variance in the data with promising results. Specifically, multilayer perceptron ANN models have been used widely in atmospheric sciences in recent years for prediction, function approximation and pattern classification [3]. For Example, A special ANN based on Multi-Layer Perceptron Model for the forecasting of daily average air quality index (AQI) was presented by Jiang et al. (2004) [4].Hence this Paper is More Focused on Review of ANN Technique in the domain of Air Pollution Epidemic related to the important issue of different air pollutant concentration time series approximation and forecasting using NN

2. Artificial Neural Networks

Artificial neural networks (ANN) are inspired by the biological neural structure. Artificial neural networks are a form of learning algorithm with Extensive pattern recognition and machine learning ability. Foundation of ANN comes from mathematical attempts at replicating information processing in biological systems. However, in modern applications they differ significantly from their biological inspiration. The neural network itself is not an algorithm, but rather a framework for many different machine learning algorithms to work together and process complex data inputs. Such systems "learn" to perform tasks by considering examples, generally without being programmed with any task-specific rules. They do this without any prior knowledge. Instead, they automatically generate unique characteristics from the learning mechanism as they process.

With modern memory and processing power, there is a vast potential for complex artificial neural network architectures such as convolutional neural networks and recurrent neural network that have seen recent success in deep learning. The ANN inner structure of the processing element (nodes) in each network is interconnected differently, and the configuration set-up is often referred to as network topology. The behavior of the network relies greatly on the network topology.

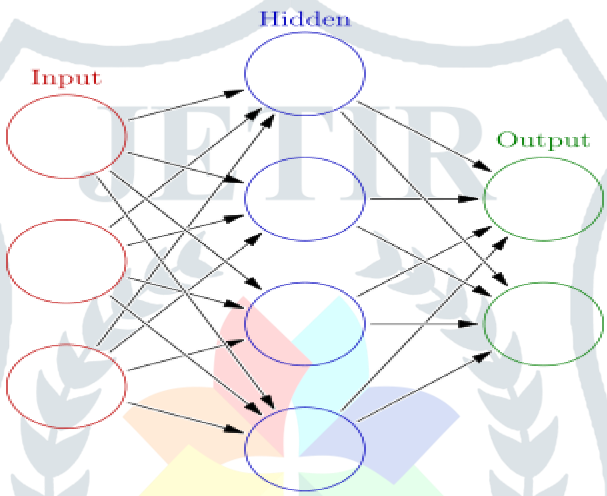

Figure. 1 Basic Example of ANN with Feedforward Topology

2.1 Neural Network Topolgoies

The most commonly used neural networks types are: radial basis function network (RBFN), multi-layered perceptron networks (MLP), and recurrent neural networks (RNNs). The distinction in different network topologies can perhaps be attributed to the arrangement of neurons and the connection patterns of the layers.

1. Feed forward Network MLP feed-forward network is referred to as a directed cyclic graph in which the connections are unidirectional and no loops are introduced in the network, thus each neuron is linked only to neurons in the next layer. This implies no backward links either. Multi-layer feed-forward neural networks trained by back propagation algorithm, i.e. back propagation network (BPN). The BPN is one of the easiest networks to understand. Its learning and update procedure is based on a relatively simple concept: if the network gives the wrong answer, then the weights are corrected, so the error is lessened so future responses of the network are more likely to be correct. Other learning algorithm such as Levenberg Marquardt back propagation also use similar mechanism for updating weights.

2. Recurrent Neural Network unlike MLP network, recurrent structure introduces cycles or loops and backward links in the network. Feedback networks are exceptionally dominant and can get extremely convoluted. The behaviour of these types of networks is known to be changing continuously until they reach an equilibrium point. This implies the state of the network remains at the equilibrium point until the input changes and a new equilibrium needs to be found. Feedback architectures are also referred to as interactive or recurrent, although the latter term is often used to denote feedback connections in single-layer organizations.

3. Radial Basis Function (RBF) that uses radial basis functions as activation functions. RBF emerged as a variant of ANN in late 80's and it is commonly used as a pattern recognition technique, function approximation, time series prediction, and control. Architecture of a radial basis function network involves an input vector which is used as input to all radial basis functions, each with different parameters. The Network output is a linear combination of the outputs from radial basis functions. A linear transfer function is used in the output layer and a nonlinear transfer function (normally the Gaussian) for the hidden layer. The radial basis function network is probably the second widely used type of Artificial Neural Network in contrast to the standard Feed forward MLP network.

3. REVIEW OF LITERATURE

Gardner and Dorling [5] presented recent applications of the multilayer perceptron (MLP). Author's main focus was on the architecture of MLP, its applications along with a critical evaluation of back-propagation algorithm. In this implementation neural networks are found difficult to implement and interpret due to some problems. First, the no prior rules are available to decide the neural architecture, the number of layers and nodes in those layers. Second, over fitting the training data resulting in poor generalization. Third, ANN Applications in Air Quality Monitoring and Management back-propagation algorithm in the case of less number of nodes when it cannot be able to converge to a minimum during training. Fourth, curse of dimensionality which affects the speed of back propagation algorithm. To reduce the complexity of NN like feature selection and pattern selection were suggested in the paper. So, properly trained MLP still shows the potential to represent relationships, often with surprising accuracy.

Kolehmainen et al. [6] compared two principally different Neural Network Methods: Self-organizing maps (SOM) and multilayer perceptron (MLP). SOM is a form of competitive learning in which a neural network learns the structure of data, here a variation of SOM is used. MLP learns complex relationship between input and output variables, Here MLP with Levenberg-Marquardt Back-Propagation Algorithm is used. Here periodic components are removed and evaluated with respect to NN. Hourly time series on NO_2 and basic meteorological variables of Stockholm were collected between 1994 to1998. values for forecasting were calculated in three ways: using periodic components alone, applying NN to residual values after removing the periodic components and directly applying NN to the original data. The Best result in found in directly applying NN to the original data. A Combination of the periodic regression method and NN does not give any advantages over a direct application of Neural Algorithms.

Kukkonen et al. [7] used Neural Network, a linear statistical model (LIN) and deterministic modelling system (DET) for the prediction of urban NO_2 and PM_{10} concentrations. The model considers the hourly concentration time series of NO2 and PM10 of Helsinki from 1996-1999 by two stations. The model utilized both stationary and mobile emission sources for its data and pre-processed meteorological data as input data. The emission source from various mobile sources such as road traffic, harbours and marine traffic, aviation etc. and stationary sources such as power plants, other point sources and residential heating. Modelling Systems are used for Data such as Combined application of road network dispersion model (CAR-FMI) for mobile sources dispersion and for the urban dispersion modelling system (UDM-FMI) for evaluating dispersion from stationery sources. Both these models are called Multiple source Gaussian Urban Dispersion Models. NN used for this research is mostly based on Feed-Forward Back-Propagation MLP based on previous similar studies [5][6].

Five neural Network models are used in which it is assumed that observed data by the system is corrupted by a Gaussian noise process with constant variance(homoscedastic) and with variance varies with time(heteroscedastic).so NN-HoG(homoscedastic),NN-HeG(Neural network with Heteroscedastic Gaussian noise),NN-2HeG(two components),NN-3HeG(three Components) and NN-Lapl(Noise term following Laplacian Distribution) and compared with LIN and DET. The missing concentration data have been replaced using the hybrid method, i.e., a combination of linear interpolation and SOM .Here relative performance of NN models with regard to LINs/DETs are analysed rigorously and the non-linear NN models performed slightly better in terms of the model performance values and with proper training with appropriate data then this model requires less effort and provide better utilization than other models such as DETs and LINs. But Here it is Indicated that when NN is trained using localized data then the NN cannot be utilized efficiently for air pollution abatement scenarios for future years

Niska et al. [8] presented research on coarse grained Genetic algorithm within Neural Network structure of Multi-layer Perceptron for next day concentrations of nitrogen dioxide at a busy urban traffic station in Helsinki with observational data of 4 years 1996-1999.The concentration data comprised of NO_2,NO_x,O_3,PM_{10}, SO_2 and CO with pre-processed meteorological data. Missing values are imputed using hybrid methods such as linear interpolation and SOM. the problem is a non-linear regression problem so MLP is used for predicting NO_2 concentration of next day(T+24h). This MLP Structure is used with Coarse Grained GA to employ parallelism in this evolutionary method. The input parameters optimized using GA where different input parameters are encoded as population of GA and used to tune input parameter of Neural Network but the tuning of control parameter of this evolutionary algorithm is largely an empirical task as one could apply different fitness function to different inputs and form a different result thus this approach is very noisy in nature.

This model was executed for 150 generation with elitism and 10 independent runs for minimising risk of poor convergence. Parameter used were 25 hidden nodes, learning algorithm of scaled conjugated gradient back-propagation, the performance function of RMSE, hyperbolic sigmoid tangent for the hidden layers and linear for the output layer. The model is evaluated and Index of Agreement (IA), A dimensionless Entity is used to compare different models. The paper concluded that evolving input and architecture of the model does not improve the model significantly but a more robust and reliable system is obtained.it is found that GA as a technique for improving model's computational cost by eliminating irrelevant input but here no significant improvement is seen and application of simpler architecture is suggested to minimising the risk of noise over-fitting. training instances). Therefore, further research is needed on issues such as boosting, where the frequency of high concentration values is increased, enhancing the error term by using some regularization technique and recurrent neural networks, where the temporal patterns are better considered.

Challoner et al. [9] researched to indoor exposure of Air Pollutants as most of the research is based on modelling outdoor environment yet quality of indoor air is an essential determinant of a person well beings as a person spends average 90% of time indoor. Predictions are made using two modelling techniques, the Personal-exposure Activity Location Model (PALM), to predict outdoor air quality at a particular building, and Artificial Neural Networks, to model the indoor/outdoor relationship of the building. Initially, ANN models were developed to determine the dynamic relationship between the measured outdoor and indoor air quality of several monitored buildings. The Personal-exposure Activity Location Model (PALM) model was then used to predict the outdoor air quality at any particular building in the city of Dublin and thus provide an input into the ANN models to predict indoor air quality. The three commercial building used are in close proximity of each other two of which are mechanical ventilated and one is naturally ventilated. Feed-forward ANN is used which take previously collected time series data (indoor concentration and outdoor meteorological data), Levenberg-Marquardt algorithm (LMA) is used to model procedure equation. This algorithm is used for numerical solution to the problem of minimising a function. LMA interpolates between Gaussian-newton algorithm and gradient descent which is a first order optimisation algorithm.

$$(J^T.J.+\lambda diag(J^T.J).)\delta = J^T.[y - f(\beta)] \qquad (1)$$

Where:
J -Local gradient of f with respect to β
y -Independent and dependent variables
δ -Increment
β -Parameter

Matlab toolbox called "Neural Network Time series Tool" using a non-linear auto-regression with external input networks (NARX) modelling technique was used to calculate relation between indoor and outdoor concentrations of PM2.5 and NO2, and meteorological data. The NARX network is a two-layer feed forward time delay neural network (TDNN) which uses a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. In order to train the system, the feedback loops between the output and input (which are usually closed) were opened. pre-set time lag of two-time steps, between input variables and target reactions was initially selected. The meteorological variables chosen were; time of day, barometer level pressure (hPa), sea level pressure(hPa), temperature (C), relative humidity (%), wind speed (knots), wind direction (knots), Pasqual atmospheric stability class, global solar radiation and outdoor pollutant concentrations. The PALM-GIS model was used to predict the outdoor pollution levels at the locations of the test sites. The PALM-GIS model uses custom Python scripts to integrate various air dispersion models (such as the Operational Street Pollution Model, the General Finite Line Source Model and Gaussian Dispersion models) with a Geographic Information Systems (ArcGIS) platform. Both models ANN and PALM-GIS are tested for mechanically and naturally ventilated buildings. And it shown a promising result in predicting indoor NO2 concentration over a range of input data but prediction of PM2.5 shown poor performance a longer training and improved models considering various factor effecting it is need.

Rahman et al. [12] used Neural network for prediction of state of atmospheric air of industrial city Sterlitamak and developed two models: Temporal (for short term prediction of air pollutants for next days) and Spatial model (for prediction of air pollutants in any spot of the city. The short-term prediction model on basis of the feedforward neural network. Developing a neural network model for short-term prediction of content of air pollutants, specific for the city with developed chemical and petrochemical industry, such as: dust, ammonia, hydrogen sulphide, phenol, vinyl chloride, nitrogen dioxide. The offered model provides forecast with an advance up to several days depending on the meteorological characteristics of the following days. An important issue of the development was the selection of input data for the neural network model: concentration of the air pollutants in previous time periods, concentration of the air pollutants at the moment of application of the model and the meteorological parameters for the next day direction and speed of wind, air temperature, atmosphere pressure, unfavourable meteorological conditions mode, which is due to the high closeness of connection between them. the authors. In addition, a research for determination of the optimal number of the neurons in the hidden layers was carried out. However, the variation of the number of the neurons in the hidden layers did not lead to a significant change in the quality of the neural network model. As for activation functions of the neurons, there are used hyperbolic tangent in the neurons of the first layer and linear function in the neurons of the second layer.

Different learning Algorithm in a Feed-Forward Network are used such as Broyden-Fletcher-Goldfarb-Shanno algorithm, Conjugate Gradient Back-propagation algorithm with Powell-Beale restarts, Gradient descent algorithm, Gradient descent algorithm with adaptive learning rate etc. here Conjugate Gradient Back-propagation algorithm with Powell-Beale restarts showed best Adequacy in prediction of ammonia with 73% assessment of adequacy. Other model is developed for spatial prediction model is based on Elman recurrent neural network with learning algorithms such as Gradient descent algorithm, Gradient descent with adaptive learning, Gradient descent with momentum, Gradient descent with momentum and adaptive learning, Bayesian regularization. The best adequacy of the spatial prediction model is achieved in the Elman recurrent neural network with the learning algorithm based on gradient descent with momentum and adaptive learning with 86.7% assessment of adequacy. Both models are implemented on Matlab software. The paper concluded with the note that the neural network models should be trained periodically with additional daily experimental data to improve its performance further and also some other complementing techniques should also be used with NN.

Azid et al. [13] used a combination of Principal Component Analysis (PCA) and ANN to predict Air Quality Index (AQI) of Malaysia obtained from Malaysian Department of Environment (DOE) for data of Seven Year from January 1, 2005, to December 31, 2011. The air quality variables used in this study are CO, O_3, PM_{10},SO_2, NO_2, CH_4 , non-methane hydrocarbon (NmHC), and total hydrocarbon (THC) in the form of hourly reading. A total of 202,080 data points (8 variables×25,260 data set) were utilized in this analysis. The total number of missing data in the data points was very small (~3%) compared to the overall data. In order to facilitate the data analysis, the nearest neighbour method was applied. The dimension of a huge data set can be trimmed down by using PCA, which is considered as one of the most prevalent and useful statistical methods for uncovering the potential structure of a set of variables. This method is used for explaining the variance of a large set of interrelated variables by transforming them into a new, smaller set of uncorrelated (independent) variables, namely principal components. PCs are orthogonal and uncorrelated to each other and have linear combinations of the original variables. Bartlett's test of sphericity was performed at the beginning of the PCA in order to examine the correlation of the variables used in the PCA and The Kaiser-Meyer-Olkin (KMO) test was carried out in order to measure the sampling adequacy.

Two models were developed and compared, namely MLP-FF-ANN model A (as a reference) and MLP-FFANN model B. The eight variables were used as input layers in model A. After the varimax rotation result of factor scores obtained in PCA was used as input layers in model B. The two VFs generated by rotated PCA indicate that only five (CH_4, NmHC, THC, O_3, and PM_{10}) out of eight parameters were significant and responsible for air quality variations in the study area. Based on the Malaysian Ministry of Transport data, it is believed that motor vehicles are one of the major factors that contribute to the formation of these pollutants. Thus, this study indicated that for the future and effective management of the Malaysian air quality, the effort of controlling point and non-point pollution sources should be prioritized. R^2 and RSME are used as performance metrics for the models. The results of factor scores after varimax rotation were used as input variables in Model B for AQI prediction with the R^2 value of 0.618, which is better than the R^2 value (0.615) in model A. The reduction of the amount of input data in model B enables faster predictive outcome without changing significantly the predictive power as compared to model A. It means that the factor scores after varimax rotation are more efficient and effective due to reduction of predictor variables without losing important information. Additionally, the combination of factor scores after varimax rotation and ANN has been proven as useful tools in air pollution modelling. Moreover, this technique can be a simple alternative model to provide reliable estimates of AQI by using only limited information.

Russo A et al. [14] applied recent methods in stochastic data analysis for discovering a set of few stochastic variables that represent the relevant information on a multivariate stochastic system, used as input for artificial neural network models for air quality forecast. A method for deriving Eigen variable of a system coupled with stochastic variables, which were then used as input for training ANN models, this reduce the input data by a huge factor The reduced set of variables including these derived variables is therefore proposed as an optimal variable set for training neural network models in forecasting geophysical and weather properties and the predictive power of the system is maintained even with reduced information the reduction of the amount of input data optimizes the ANN models by enabling faster prediction without changing significantly the predictive power. This is because the variable incorporates temporal correlations between independent and spatially separated monitoring stations. Finally brief a discuss other possible applications of such optimized neural network models is made.

Boznar et al. [15] in a conference presented research work to construct a forecasting model of ozone concentration values. The objective behind this project was to make early warning system for alarming ozone concentration in upcoming days. Authors concentrated their research on the problem of maximal hourly value of ozone concentration that would appear in the following day. MLPNN and fuzzy logic are used. MPNN was trained using Levenberg – Marquardt method and FL was trained using Takagi Sugeno fuzzy models, Gustafson – Kessel algorithm for their structure configuration and backpropagation training. Both were implemented in the Matlab software package. The selected input features were: air temperature, global solar radiation, NO, NO2, NOx, CO, O3, prognostic vector wind speed for the day of prediction, sinus of prognostic vector wind direction for the day of prediction, prognostic maximal hourly air temperature for the day of prediction. Crucial point was the features and patterns used. Both models tested used the same input features and the same learning patterns set. When the models were constructed, they were tested on the same independent verification set of patterns. Although, the verification set was relatively small, still MPNN and FL have given satisfactory result.

4. Results and Discussion

4.1 Summary of literature review

S. No	References	Objective	Model	Remark
1	Gardner & dorling [5]	Applications of the multilayer perceptron (MLP) in atmospheric sciences	MLP with a critical evaluation of back-propagation algorithm	Various shortcoming found in using ANN like over fitting, complex implementation, curse of dimensionality
2	Kolehmainen et al. [6]	Time series analysis of NO2 concentration with meteorological variables to predict future concentration	Comparison in SOM and MLP	Show fairly good performance and anticipate for unknown parameters but improvement suggested
3	Kukkonen et al. [7]	Prediction of urban NO2 and PM10 concentrations	Neural network, a Linear Statistical model (LIN) and Deterministic modelling system (DET)	NN trained using localized data then the NN cannot be utilized efficiently for air pollution abatement scenarios for future years
4	Niska et al. [8]	Evolving the neural network model for forecasting air pollution time series	Genetic algorithm (GA) with MLP	Using GA reduced computational efforts by eliminating irrelevant inputs
5	Challoner et al. [9]	Predicting the indoor air quality from outdoor monitors of NO_2 & $PM_{2.5}$	Levenberg-Marquardt algorithm, a type of Feed-Forward ANN	Emphasis on meteorological data effecting accuracy of model is discussed & extensive research is needed

6	Rahman et al. [12]	Using NN for prediction of air pollution index in industrial city	FFNN with the learning algorithm based on conjugate gradient back-propagation with Powell-Beale restarts & Elmal recurrent neural network	Specialized algorithm is used for specific prediction with promising results & high accuracy with greater iterations
7	Azid et al. [13]	To predict air pollutants: O_3, CO, NO_2, SO_2 and PM10 in states of Malaysia	Principle component analysis (PCA) with MLP-FFNN	The PCA-MLP showed better predictive ability in the determination of AQI with fewer variables,
8	Russo a et al. [14]	To reduce input variable for training NN without effecting its predicting power using stochastic variable	Stochastic variables as optimal input for NN	Reducing inputs using stochastic variable optimize data and make faster convergence in NN
9	Boznar et al. [15]	To make early warning system for alarming ozone concentration in upcoming days	MLPNN and Fuzzy logic	Comparison of fuzzy logic and MLPNN showed similar result for the same data set

Table 4.1 summarizing Objective of the reviewed papers and techniques used.

In environmental studies, air pollution is a significant issue focused all around the globe as it affects human health, ecosystems and the environment, buildings and in turn economy. It is defined as a condition, in which the substances that result from both natural and artificial activities are present in the air at concentrations significantly high above their normal levels causing considerable impacts on humans, animals, vegetation, or materials.

Literature surveys suggests a lot of methods for the prediction of air quality ranging from numerical, mathematical and statistical methods (e.g., regression) to techniques based on artificial intelligence, In this paper particularly ANNs. All the meteorological variables and factors have a non-linear relationship with air quality, which can be accurately captured by nonlinear models such as ANNs and Support Vector Machines.

ANN Models is found to be superior then other models. Majority of the paper reviewed here are based on FF-MLP Models with some other ANN Models and sometimes compared with other machine learning paradigms such as SOM, Fuzzy Logic, PCA, GA etc. Sometimes these techniques are used to complement the Shortcoming or enhancing the performance of the ANN Models. The Overall Performance of the Models was improved by enhancing these models using each iterative application and using different complementing techniques. Sometimes the ANN Shows slight improvement and these studies need to be improved and investigated further.

5. Acknowledgment

We would also like to show our gratitude to the Prof. Abhilash Sonker, Dept. of CSE/IT, MITS for sharing their pearls of wisdom with us during the course of this research, and we thank my friend Prasoon Purwar for their so-called insights. We are immensely grateful to for their comments on an earlier version of the manuscript, although any errors are our own and should not tarnish the reputations of these esteemed persons.

References

[1] Cogliani, E. (2001). Air pollution forecast in cities by an air pollution index highly correlated with meteorological variables, Atm. Env., 35, pp. 2871- 2877.

[2] Shi, J. P. and R.M. Harrison (1997): Regression modelling of hourly NOx and NO2 concentrations in urban air in London. Atmospheric Environment, 31, pp. 4081-4094.

[3] Gardner, M.W.,Dorling, S.R., 1999. Neural network modelling and prediction of hourly NOX & NO2 concentrations in urban air in London. Atmospheric Environment 33, 709–719.

[4] Jiang, D., Zhang, Y., Hu, X., Zeng, Y., Tan, J., Shao, D., 2004. Progress in developing an ANN model for air pollution index forecast. Atmos. Environ. 38, 7055e7064.

[5] Gardner, M. W. and Dorling, S. R. (1998): Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences. Atmospheric Environment, Vol. 32, No. 14/15, pp. 2627-2636.

[6] Kolehmainen M, Martikainen H, Ruuskanen J. Neural networks and periodic components used in air quality forecasting. Atmos Environ. 2001;35(5):815–25.

[7] Kukkonen J, Partanen L, Karppinen A, Ruuskanen J, Junninen H, Kolehmainen M, Niska H, Dorling S, Chatterton T, Foxall R, et al. Extensive evaluation of neural network models for the prediction of no 2 and pm 10 concentrations, compared with a deterministic modelling system and measurements in central helsinki. Atmos Environ. 2003;37(32): 539–550.

[8] Niska H, Hiltunen T, Karppinen A, Ruuskanen J, Kolehmainen M. Evolving the neural network model for forecasting air pollution time series. Eng Appl Artif Intell. 2004:159–167

[9] Challoner A, Pilla F, Gill L. Prediction of indoor air exposure from outdoor air quality using an artificial neural network model for inner city commercial buildings. Int J Environ Res Public health. 2015;12(12): 15233–53.

[10] Levenberg, K. A method for the solution of certain problems in least squares. Q. Appl. Math. 1944, 2, 167–168. 31.

[11] Marquardt, D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. SIAM J. Appl. Math. 1963, 11, 431–441. [CrossRef]

[12] Rahman P A,Panchenko A A,Safarov A M , Using neural networks for prediction of air pollution index in industrial city ,2017 IOP Conf. Ser.: Earth Environ. Sci. 87 042016

[13] Azid, Azman & Juahir, Hafizan & Toriman, Mohd & Kamarudin, Mohd khairul amri & Mohd Saudi, Ahmad Shakir & Noraini Che Hasnam, Che & Azlina Abdul Aziz, Nor & Azaman, Fazureen & Latif, Mohd Talib & Farihan Mohamed Zainuddin, Syahrir & Osman, Romizan & Yamin, Mohammad. (2014). Prediction of the Level of Air Pollution Using Principal Component Analysis and Artificial Neural Network Techniques: a Case Study in Malaysia. Water Air and Soil Pollution. 225. 10.1007/s11270-014-2063-1.

[14] Russo, A., Raischel, F., & Lind, P. G. (2013). Air quality prediction using optimal neural networks with stochastic variables. Atmospheric Environment, 79, 822–830.doi:10.1016/j.atmosenv.2013.07.072

[15] Boznar, M.Z., Mlakar, P. and Grasic, B. Neural networks based ozone forecasting, 2004 ,9th Int. Conf. on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes. Pp. 356-360