International Journal for Environmental Rehabilitation and Conservation
Volume II No. 1 2011 [30 – 64] [ISSN 0975 - 6272]

Artificial neural network applications in air quality monitoring and management

Prachi[1], Kumar Nishant[2] and Matta, Gagan[3]

Received: January 11, 2011 | **Accepted**: April 19, 2011 | **Online**: July 20, 2011

With the loss of biodiversity on an unprecedented scale, and addition of pollutants with the potential of altering climates and poisoning environments on a global scale, the pressure to understand and manage the natural environment are far greater now than could ever have been conceived even 50 years ago (Lek and Guegan, 1999).

Air pollutants exert a wide range of impacts on biological, physical, and economic systems. Sulfur dioxide (SO_2); nitrogen oxide (NO_x); nitric oxide (NO), nitrogen dioxide (NO_2); carbon monoxide (CO); Ozone (O_3); respirable suspended particulates (RSPs); etc. are some of the major airborne pollutants contributing to the quality of living in urban areas, especially in densely populated and industrialized areas (Boznar and Mlakar, 2002). Indeed, many epidemiological studies have consistently shown an association between particulate air pollution and cardiovascular and respiratory diseases. The decrease in respiratory efficiency and impaired capability to transport oxygen through the blood caused by a high concentration of air pollutants may be hazardous to certain sensitive groups in the population i.e. those having pre-existing respiratory and coronary artery diseases such as children, asthmatics and elderly people. Consequently, it has become a vital task to accurately keep track of the variation of ambient air pollution levels in urban areas and to take controlling measures (Barai *et al.*, 2000; Pasero and Mesin, 2006).

Keywords: *Artificial Neural Network* | *Environmental Informatics* | *Air Pollution monitoring* | *Pollution modeling* | *Modeling* |

For Correspondence:

[1] Rotary Environmental Lab., Haridwar
[2] Faculty of Engineering, Gurukula Kangri University, Haridwar, India
[3] Department of Zoology and Environmental Science, Gurukula Kangri University, Haridwar, India.

Air pollution control is needed to prevent the situation from becoming worse in the long run. Forecasting of air quality is needed in order to take preventive and evasive actions during episodes of airborne pollution. In this way, by influencing people's daily habits or by placing restrictions on traffic and industry it should be possible to avoid excessive medication, reduce the need for hospital treatment and even prevent premature deaths (Kolehmainen *et al*., 2001).

Studies have shown that pollutants are usually found entrapped into the planetary boundary layer (PBL), which is the lowest part of the atmosphere and has behavior directly influenced by its contact with the ground. In this layer, physical quantities such as flow velocity, temperature, moisture and pollutants display rapid fluctuations (turbulence) and vertical mixing is strong. Conclusively, numerous studies have shown the existence of high correlation between air pollution and meteorological variables (Cogliani, 2001; Pasero and Mesin, 2006).

Other than the meteorological variables, the layout of the city, the existence of green and un-built areas, the geometry, the architectural morphology and the thermal properties of the buildings, the vehicular traffic, the stationary thermal systems, all define the urban environment as a multi-dimensional, multivariable system. Thus, air pollution is a problem that cannot be treated independently of the urban web. Concerning photochemical pollutants, it should be noted that their dynamic nature, accompanied by the strong non-linearities in the underlying physical and chemical mechanisms involved in their creation, chemical transformation and transportation-diffusion, is always among the major challenges for the development of any modeling – forecasting method and tool (Athanasiadis *et al*., 2006).

Admittedly, the main environmental problem that need efficient software tool is the prediction problem. More concrete, it can mean meteorological prediction, air/soil/water pollution prediction, flood prediction and so on. In the last two decades several methods based on artificial intelligence were proposed by taken into account that they can offer more informed methods that use domain specific knowledge and provide solutions faster than the traditional methods, those based on a mathematical formalism (Oprea and Matei).

Numerous researches and studies have been carried out on the subject of nature and dynamic behavior of pollutants, emission, propagation and effects of pollutants. Predicting future dispersion of air pollution is of immense importance since it can provide an effective decision making tool by giving advance warning of excessive pollution beyond the threshold. Also, it enables early air quality control to mitigate the adverse impacts. On the other hand, high dynamism and nonlinear behavior of air pollutant data makes prediction difficult or inaccurate. High capabilities of artificial neural networks e.g. flexible structure and the use of dynamic learning algorithm promote the application of these intelligent systems in this domain (Abbaspour *et al*., 2005).

Natural phenomena are mostly a time series with some degree of randomness. Pollutants in the atmosphere may disperse or concentrate during varied time periods. Previous studies (Giorgio and Piero, 1996) have indicated that the data of ambient air quality are stochastic time series, thereby making it possible to make a short-term forecast on the basis of historical data. Though models may be imperfect, they are the best tool for use in all aspect of air quality planning where prediction is a major component such as for emission control (Melas *et al*., 2000), accidental release of pollutant, land-use planning, traffic planning (Hadjiiski and Hopke, 2000), planning of measurement programs (Rao and Rao, 2000), analyses of measurements/ trends and episode forecasting (Melas *et al*., 2000; Barai, *et al*. 2000).

The modeling and forecasting of environmental parameters involves a variety of approaches. One approach is to use the atmospheric diffusion model to predict future pollutant concentrations. A second is to devise statistical models that attempt to determine the underlying relationship between a set of input variables (original data) and the targets (Shi and Harrison, 1997). Statistical methods, either the time-series methods, which do not use meteorological inputs, or regression and similar methods, which are based on multivariate linear relationship between meteorological conditions and air pollution concentrations, are commonly used (Barai *et al*., 2000).

In other words, though the statistical methods do provide reasonable results, these are essentially incapable of capturing complexity and non-linearity of pollution-weather relationships. To overcome this demerit of statistical methods, Artificial neural networks (ANN), the third approach developed in recent years has become the focus of much attention, largely because they can handle the non-linearity and have been used to model pollutant concentrations with promising results (Boznar *et al*., 1993; Comrie, 1997; Gardner and Dorling, 1996, 1998; Hadjiiski and Hopke, 2000).

Artificial Neural Networks (ANNs)

The concept of artificial neural networks was established in 1943 (McCulloch and Pitts, 1943). Perceptron, the first practical artificial neural network was presented in 1958 (Rosenblatt, 1958). So, their history goes back more than 50 years, but due to the availability of modern computers from the 1980's they have grown to be a competitive tool that has been applied widely since the mid 1990's (Jef *et al*., 2005). One of the reasons for their success is their capability to make regressive approximations of non-linear functions in high dimensional spaces, something that is missing in classical statistics. The flexibility of neural networks (NNs) led to their use in all possible scientific branches. In the last two decades, ANNs have been already explored in various fields like chemical research (Kvasnicka, 1990; Wythoff et al., 1990; Smits et al., 1992), physics research (Dekruger and Hunt, 1994), molecular biology, ecology and environmental sciences and demonstrated remarkable success.

ANNs imitate the learning process of the animal brain (Lippmann, 1987) and can process problems involving very nonlinear and complex data even if the data are imprecise and noisy. ANNs can identify and learn correlated patterns between input data sets and corresponding target values. After training, ANNs can be used to predict the output of new independent input data. This is regarded as an intelligent, cost-effective approach and has received much attention in environmental engineering. ANNs are also known as "universal function approximators" because of their capacity to approximate virtually any continuous nonlinear function with arbitrary accuracy. Thus, ANNs have the ability to solve many complex problems in which a priori knowledge is incomplete or unavailable (Fabbian and Dear, 2007).

It is true that as compared to traditional statistical techniques, a neural network (NN) excels by its flexibility. The main drawback is that a NN which is trained by data from a given measuring location can only forecast for that specific location and it cannot give insight into the physics behind the data: a NN merely learns from examples and it is not suited to generalize to other situations (Jef *et al*., 2005). Thus, there are certain questions on ANN which still remain unsolved and continue to challenge researchers; e.g., the curse of dimensionality, local minima, overfitting, *etc*. (Lek and Guegan, 1999; Barai *et al*., 2000; Boznar and Mlakar, 2002; Lu *et al*., 2004).

Apart from some drawbacks, computing with neural networks is one of the fastest growing fields in the history of artificial intelligence (Comrie, 1997; Gardner and Dorling, 1998; Tecer, 2007). NNs have applications for both air pollutant time series modeling and air pollutant concentrations forecasting. A systematic flow of published papers on such applications starts in the early 1990s, boosted by the constantly improving performance and the decreasing cost of powerful computers in one hand and by the fact that software packages, commercial or open source, which required a minimum programming effort by non-specialists in NN algorithm were becoming widely available.

NNs, are developing constantly. In the initiating phase of NNs, multilayer perceptron seemed to be an effective alternative to more traditional statistical techniques. But later on, many researches were performed using other types and architectures of NNs for the same purpose of air pollution modeling. The present review focuses on the literature related to the important issue of different air pollutant concentration time series approximation and forecasting using NN that appeared until this date. The aim is to provide concise information regarding the NN types and architectures used for this purpose, their results and a critical evaluation between different NN models to some crucial points.

General overview of Artificial Neural Network (ANN)

The experiment with an idea to find a simple non-linear model for a real neuron led to the development of artificial Neural Network (ANN) theory by McCulloh and Pitts in 1943. McCulloh and Pitts tried to model the bio-systems using nets of simple logical operations. This innovation fetched a great interest from various researchers and scientists all over the world. Thus several ANN-based models in different fields were discovered. The technology though had lost its momentum in the late 1969 till 1986 when the back-propagation of error was discovered. ANN-based models have been successfully implemented in a number of disciplines ranging from: Medical, automotive, defense, electronics, aerospace, entertainment, financial and so on. Based on ANN methodologies, Crawford, 2000 reported the acute appendicitis analysis. ANN is a part and parcel of intelligent based systems, designed distinctively to improve the performance of conventional computing techniques. The biggest drawback associated with the so called conventional methods is the inability to learn and identify patterns in dynamic systems. Thus the need to eliminate this shortcoming through learning is proven essential.

ANN models are computer programs that are designed to emulate human knowledge processing, speech, prediction, classification and control. ANN is a cellular information processing system designed and developed on the basis of the perceived notion of the human brain and its neural system. The human brain has 100 billion biological neurons with about 100 000 connections per neuron. Thus, neural networks are basically a collection of interconnected neurons, each one with several inputs and one output. The inputs' weight (importance) and numbers can vary. The output is a function of all inputs. The figure below shows how biological neurons interact with each other.

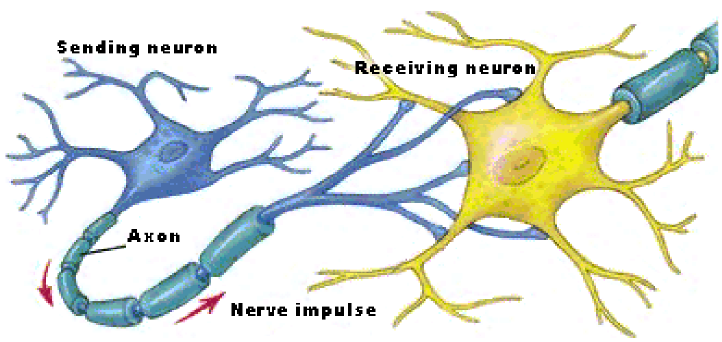

Figure 1: A biological neuron

Above figure shows that a blue neuron is sending an impulse to the yellow neuron. The yellow neuron may receive other impulses (varying in strength) from other neurons, but is only sending one signal (a function of all the inputs). Humans and highly trained animals use the same configuration and summing up to extremely complex networks. Figure 2A shows that how a biological system can be transferred into a computer system via a sketchy representation of artificial neuron.

Similarly, an artificial network is made up of simple interconnected processing elements called *neurons*. The neurons are arranged in a layered structure to complete a network competent of executing parallel and distributed computations. Figure 3 shows a taxonomy tree of NN which clearly depicts the various NN types depending on their architectures. The Architecture of a simple ANN i.e. feedforward neural network is shown in Figure 2B. The attraction of ANN-based models comes with the network's ability to learn, recognize data patterns, and adapt to a changing environment like the human brain. This adaptive characteristic is often called "the human-like reasoning".

The architecture illustrated in Figure 2B, presents a three layered feed-forward network. ANN has a remarkable capability to develop sense from convoluted or imprecise data, extract patterns and detect trends that are too complex often only noticeable by either humans or other computer techniques. A supervised neural network can be trained as a *"guru"* in a given problem space. In broad terms, ANN-based models offer a variety of benefits namely: adaptive learning, self organization, real time operation, fault tolerance via redundant information coding. Thus neural network processes information in the similar way the human brain does.

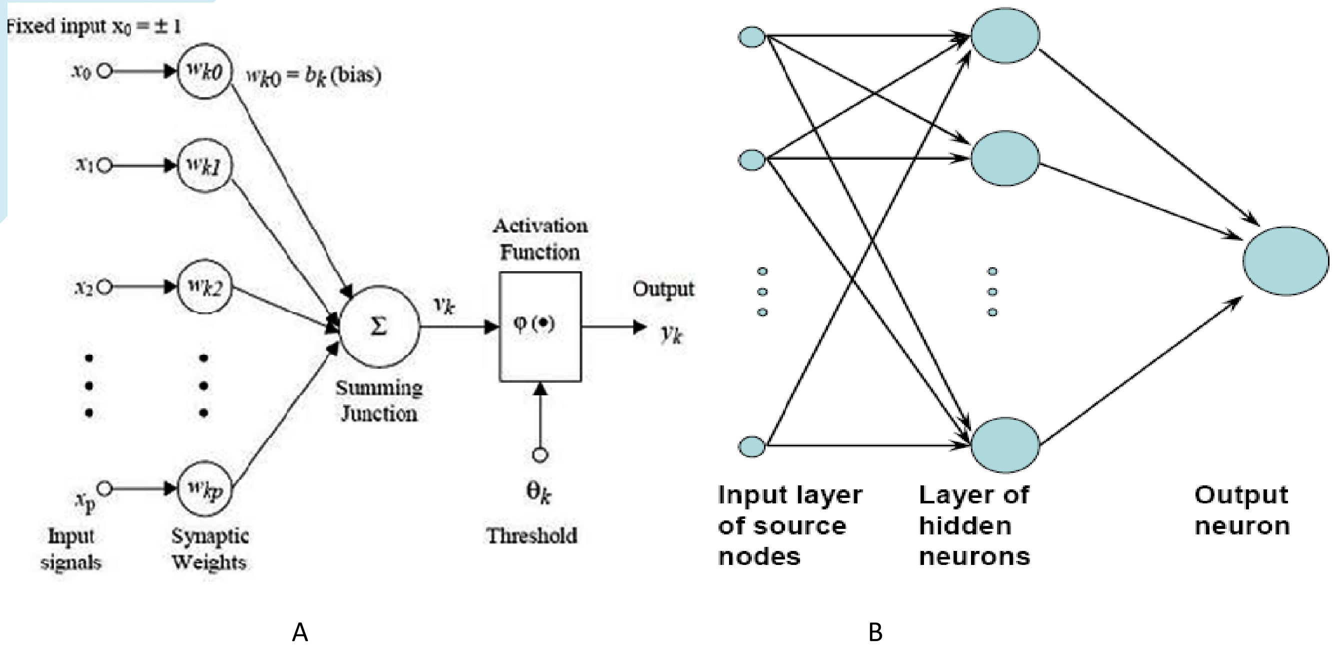

Figure 2: A) Sketchy representation of an artificial neuron. B) Example of a feedforward neural network, with a single hidden layer and a single output neuron (Pasero and Mesin, 2006).

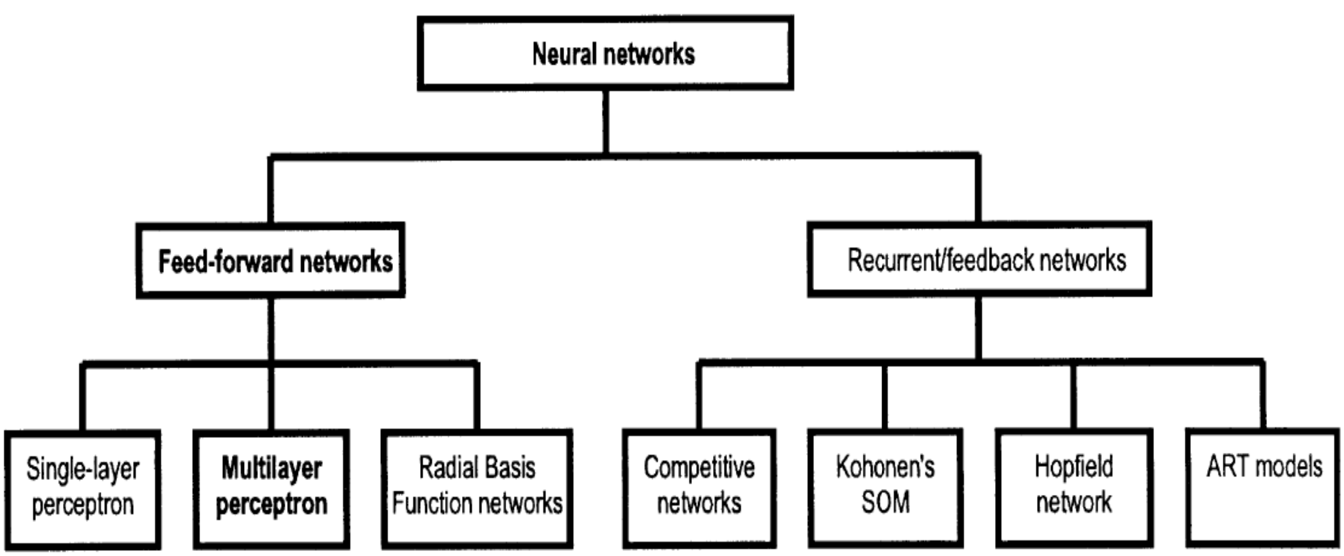

Figure 3: Taxonomy of neural network architectures (Gardner and Dorling, 1998).

Neural Network Topologies

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.

The inner structure of the processing element (neuron) 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.

1. Feedforward 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 backpropagation algorithm, i.e. backpropagation 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. The conceptual basis of the backpropagation algorithm was first presented in by Webos (1974), then independently reinvented by Parker (1982), and presented to a wide readership by Rumelhart *et al.*, 1986 (Lek and Guegan, 1999).

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 behavior 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) Network

A radial basis function network is an artificial neural network 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 output of the network 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 Feedforward MLP network.

Neural Network Models and its forecasting or prediction applications in air pollution –

Boznar *et al*., 1993 presented the method of neural network for short-term air pollution prediction. Their study was on the prediction of SO_2 emissions around the Slovenian thermal power plant (TPP) at Sostanj.

Air dispersion models and Cyclo-stationary Auto Regressive (CSAR) predictors have already been applied to the area but these models failed pertaining to the complexity of the terrain. As input to the model, a historical set of significant meteorological and ecological data was used. The input parameters included the following: the wind data, air temperature, SO_2 concentrations (actual and historical), relative humidity, solar radiation, emissions from the TPP and time of day. The training set of patterns was taken from the beginning of February and the test set of patterns from the later days of the same month and the method was tested on the data collected by the Environmental Information System (EIS) of the Sostanj Thermal Power Plant (TPP). The output feature was the SO_2 concentration measured after half an hour. Some predictions with good results were made for the stations at Zavodnje, Veliki vrh and Velenje. For the station Veliki vrh, wind direction showed a major role in pollution. According to the results the network anticipated higher concentrations, but not high enough. The reason might be the changeable wind direction at the site. For the Zavodnje station, unique terrain that surrounds it found to be significant, especially because of the morning thermal inversions. The results of the prediction on the testing set of patterns have shown that the peaks were slightly shifted. These patterns belong to conditions with the wind blowing from the power plant to Zavodnje when there is no thermal inversion. So, the situation was the same as the one appearing at Veliki vrh. These results show clearly that it is easy for a neural network to recognize inversion conditions, but the prediction of changeable wind direction still needs improvement. For Velenje station, the prediction found good because the peak SO_2 concentration was recognized completely. Thus, the results obtained at all above stations were the best possible. They show, that with a better set of training patterns, prediction of SO_2 concentrations with high accuracy would be possible.

This study set a great landmark in the field of air pollution prediction because of its inclination towards a new and better approach of neural network. Along with this, some questions also raised up like when to stop the learning process to get the best results with the unknown patterns, and how to set the back propagation algorithm parameters to get the best possible convergence of the learning process. The author also mentioned about the preparation of on-line implementation of the method in the software of the central unit of the EIS. Their planning was to implement an automatic algorithm for selection of new training patterns which could enable the automatic updating and adapting of the neural network to the newly measured data.

Yi and Prybutok, 1996 describe a multilayer perceptron that predicts surface ozone concentrations in an industrialised area of North America. The model takes nine input variables to predict the maximum daily surface ozone concentration. These variables include the morning ozone concentration, the maximum daily temperature, CO_2, NO, NO_2 and NO_x levels, and also wind speed and direction. Results from the multilayer perceptron were shown to be better than those obtained from regression analysis (using the same input data). The authors also suggest that the multilayer perceptron outperforms an ARIMA time-series modelling approach, however such comparisons between techniques must be made with care. For example, to fairly compare an ARIMA time-series model with a multilayer perceptron model, requires that both models are constructed using the same data.

Comrie, 1997 made a direct comparison of the NN models and multiple regression models for weather based ozone (O_3) forecasting in order to identify the best general modeling approach. The ozone data from eight cities around the United States (Seattle, Pittsburg, Chicago, Atlanta, Charlotte, Boston, Tucson and Phoenix) was used. The data used were the daily maximum 1-hour concentrations, for the months of May-September, over the five-year period 1991-1995. U.S. Environmental Protection Agency's (EPA) Aerometric Information and Retrieval System (AIRS) database was used for this data. The variables selected for this study were those most closely related to ozone behavior: daily maximum temperature (TMAX), average daily dew point temperature (DPTP), average daily wind speed (AWND), and daily total sunshine (TSUN). The neural network modeling was performed using the NevProp software package. Feedforward neural network architecture was used along with the back-propagation algorithm to minimize the error. In the choice of a non-linear transfer function, sigmoid worked well. Results showed no dramatic improvements in the performance of neural networks. Model comparison statistics clearly proved that neural networks are somewhat, but not overwhelmingly, better than multiple regression models for weather-based ozone forecasting. It is interesting to note that the relative improvement between unlagged and lagged models is greater for regression models than neural networks (i.e., neural networks using lagged data do not have to rely as much on persistence information in those data). A general observation is that the inclusion of lagged ozone data (the ozone maximum from the previous day) improves both kinds of model. Still, the paper suggests the scope of improvement in individual regression and neural models. Likely strategies suggested were using additional variables (e.g., extra weather elements, synoptic patterns, traffic flow or day-of-the-week information) or changing the nature of the model (e.g., incorporating a priori curvilinear transformations of input data for regressions, or changing the number of hidden nodes and layers in neural networks).

Gardner and Dorling (1998), presented a general introduction and discussion of recent applications of the multilayer perceptron (MLP), one type of ANN, in the atmospheric sciences. Author's main focus was on the architecture of MLP, its applications along with a critical evaluation of back-propagation algorithm. After their exhaustive study, Gordon and Dorling accepted the fact that neural networks are difficult to implement and interpret due to some problems. First, the non-availability of any rule to decide the neural architecture, the number of layers and nodes in those layers. Second, overfitting the training data which results in poor generalization performance. Third, failure of back-propagation algorithm in the case of less number of nodes when it can not be able to converge to a minimum during training. Fourth, curse of dimensionality which affects the speed of backpropagation algorithm. Strategies to reduce the complexity of NN like feature selection and pattern selection were suggested in the paper. So, properly trained multilayer perceptrons still shows the potential to represent relationships, often with surprising accuracy that is not fully understood by the traditional theory.

In their follow up work, Gardner and Dorling, 1999 applied Multilayer perceptron (MLP) neural networks to model hourly NO_x and NO_2 pollutant concentrations in Central London from basic hourly meteorological data. Results have shown that the models perform well when compared to previous attempts by Shi and Harrison (1997) to model the same pollutants using regression based models. The data used were selected that were close to those used in their original study of Shi and Harrison (1997). Hourly NO_x and NO_2 data were obtained from the Department of the Environment, Transport and the Regions (DETR) automatic monitoring network between 1990 and 1991 for two monitoring sites in Central London. Hourly meteorological data were included the Low cloud amount (LOW), Base of lowest cloud (BASE), Visibility (VIS), Dry bulb temperature (DRY), Vapor pressure (VP), Wind speed (WS). The models were trained on data from 1990. Data from 1991 were used as both the validation and test data sets. In the study, Neural networks were trained with and without the emission factor (diurnal nature of NO_x emissions which is a predictor of both NO_x and NO_2 concentrations). MLP 1 and MLP2 models were generated without any explicit information concerning the likely emissions of NO_x or other pollutants. MLP2 outperformed an autoregressive model developed by Shi and Harrison (1997) which included the previous hour lagged NO_x concentrations as an input variable. It is apparent from the results that the lack of any information concerning diurnal emissions does not seriously hinder the MLP neural network model. This would be unlikely to occur in other traditional models where such a complex interaction between predictor variables is not permitted. Two MLP models (MLP3 and MLP4) and two linear regression models (LR3 and LR4) were trained without time of day inputs and using a NO_x emissions factor as an additional input. The performance of both models (MLP3 and MLP4) were extremely similar to the performance of the models with no emissions factor (MLP1 and MLP2). This leads to the conclusion that the emissions factor could be replaced by time of day inputs when using MLP neural networks without any detrimental effects in this work. MLP models are able to make efficient use of proxy data when the optimum predictor variables are unavailable. It was shown when the addition of an hourly sunshine input to the MLP models did not significantly improve the models performance. Thus the lack of intuitively important variables does not necessarily restrict the likelihood of developing a reasonable model given the availability of a number of suitable proxies. This is less likely to be the case when developing regression models where interactions between input variables are not permitted. Tests using previous lagged concentrations illustrate that while the best predictions are made by models with additional 1 h lagged pollutant concentration inputs, the models with 24 h lagged pollutant inputs manage to predict pollution episodes reasonably well. Such models could be implemented as air quality forecast models to provide 24 h forecasts of pollutant levels.

Spelmann, 1999 applied NN model to predict ozone concentrations for London, Harwell and Birmingham cities. In this study, correlation coefficients were found 0.77, 0.72 and 0.53 respectively. The result clearly depicts that neural networks are best as compared to conventional methods.

Barai, *et al*., 2000 compared the performance of neural network models: Recurrent Network Model (RNM), Change Point detection Model with RNM (CPDM), Sequential Network Construction Model (SNCM), and Self Organizing Feature Maps (SOFM) for air quality forecasting. The investigation was carried out for long-term as well as short-term air quality data set. Two cases studies were performed. Case Study 1 demonstrated an example of an annual average emission (long-term) data prediction using the above neural network models for a very limited dataset. The data was taken for 115 counties of California State in United States of America and collected from US EPA website (www.epa.gov). This data is annual average data for 15 years from 1985 to 1999 for seven parameters namely VOC (volatile organic carbon), NO_X (oxides of nitrogen), CO (carbon monoxide), SO_2 (sulfur dioxide), PM10 (particulate matter with size less than 10 microns), $PM_{2.5}$ (particulate matter with size less than 2.5 microns) and NH_3 (ammonia). In the Case Study 2, the data for three parameters namely RPMA (Respiratory Particulate Matter Average), SO_2 (sulfur dioxide) and NO_2 (nitrogen dioxide) was collected for Delhi State at nine locations. These data were daily average concentrations for last two years from 3/7/2000 to 20/8/2001. This data set has been collected from Tata Energy Research Institute web site "www.teri.in". However, only the data for Ashram Chowk was used for carrying out simulation studies. The dataset size had 110 patterns. The models studied for this case, in general showed the prediction of air quality with modest accuracy. However, among all the models implemented, Self-organizing Feature Map (SOFM) based model has performed extremely well in comparison to other models. The performance of various models for different air quality parameters was estimated in terms of mean percentage error (PE). The author concluded that models in general have performed reasonably well even with the limited historical data and the improvement in the performance of models studied could be expected in case more available data.

Tadashi Kondo *et al*., 2000 dealt with the problem of environmental planning and impact assessment by estimating the spatial distribution of each air pollution source at a large area. For the purpose, author implemented the GMDH (Group Method of Data Handling)-type neural network with a feedback loop in order to identify the large-spatial air pollution patterns and compared it with other identification methods. The GMDH-type neural network is a multilayered neural network architecture which is automatically organized by using the heuristic self-organization method and the structural parameters such as the useful input variables, the number of layers, the number of neurons in a hidden layer and the optimum architectures of the neurons in a hidden layer, are automatically determined in this algorithm. The given method was used to accurately estimate the source-receptor matrix. The source-receptor matrix presents a relationship between the multiple air pollution sources and the air pollution concentration at the multiple monitoring stations. Then the air pollution concentration patterns at a large area are identified by using estimated source-receptor matrix. The comparison shows that GMDH-type neural networks are easy to apply for the identification problem of large-spatial air pollution patterns because the optimum neural network architecture is automatically organized. It is also more accurate than other identification methods.

Boznar *et al.*, 2001, in their series of work from Boznar et al. (1993) presented their work on the implementation of feature determination and pattern selection methods in the case of SO_2 concentration prediction for half an hour advance around Slovenia's largest thermal power plant at Šoštanj (ŠTPP). The model was based on a multilayer perceptron neural network with two hidden layers and a sigmoid transfer function in the hidden layer and output neurons. Training was done with backpropagation algorithm. Input features (data) used were meteorological measurements (ground level wind speed and direction, temperature, relative humidity) and ambient concentrations at some of the measuring station for the current half hour interval and for the previous intervals. There was only one output feature – the ambient concentration at a selected measuring station for the following half hour interval (Mlakar, 1972). The solutions of the problems regarding proper training of neural network models suggested by Gardner and Dorling (1998) were finally experimented by Boznar *et al*. (2001, 2002). The important methods used for feature determination were preprocessing, heuristic determination, feature extraction and feature selection. The most important methods for pattern selection were meteorological knowledge-based cluster determination and Kohonen neural network based cluster determination. These feature determination and pattern selection strategies provided the guidelines for finding the most relevant information in the data base available for model construction. The given paper clearly shows the improvement of the model depicted by correlation coefficient, normalized mean square error or fractional bias. Thus, the problems of feature determination and pattern selection are the essential ones that should be solved in order to obtain a good neural network-based air pollution prediction model. The data used for neural network training (learning) should be free of patterns and features with low or no information relevant to pollution prediction. The model will predict the output values of the unknown (during the learning process) patterns well enough in the case of above rules satisfied.

Werner *et al*., 2001 tried to overcome the major demerit of neural network models i.e. its 'black-box' nature. The 'very black-box' nature makes difficult the assessment of the relevance of the different input variables within the model. The neural networks lack the error calculation that some statistical methods can do. Except for the generalization error, i.e. the RSME over all test inputs, neural networks have not much to offer for measuring the quality of the net prediction. Author's new approach was to develop a measure of how sure the network is about its answer. This paper presents an idea of using a two-segmented network, where the first segment works as an input-oriented, (mostly trained by unsupervised methods) classification device, whereas the second segment produces the output based on the classification given by the first segment. In the study, radial basis function self organized map (RBFSOM) networks, which are RBF networks using SOM as hidden layer and the SOM-training to obtain the appropriate centers for the Gaussian functions was used. For each input, the two parameters MA 'maximum activity in the hidden layer' and RA 'maximum relative activity of distant neurons' provide a measure for the quality of the net output. Hence the RBFSOM networks, calculating variable bandwidths and the parameters MA, RA, provide a neural network type that allows the user some insight, how sure the network is about its answer and how reliable the answer can be. Thus, the new network approach found suitable for creating models that are capable to estimate the accuracy of their response even in the situation where only few data for training are available.

Kolehmainen *et al*., 2001 investigated the forecasting capability of the following five methods: regression using periodic components at the year, week and day levels, the MLP and SOM methods applied to the original time series and the MLP and SOM methods applied to the residual of the periodic components. An hourly time series of NO_2 concentrations were used as an example and basic meteorological variables from the city of Stockholm were collected from years 1994-1998. The meteorological variables used were temperature, wind speed, wind direction and solar radiation, in addition to which the hour of the day and month of the year were recorded. The software used for calculation of the periodic components and for neural processing using the MLP was Matlab version 5.4. The neural processing using SOM and the pre-processing of the data were carried out with the neural data analysis (NDA) software package (http://erin.math.jyu."). For data validation, statistical indicators of error, namely the root mean square error (RMSE), its systematic (RMSES) and unsystematic (RMSEU) components, the coefficient of determination (R^2), index of agreement (d), proportion of systematic error (PSE) and bias, were calculated for each of the five methods. Histograms of error residuals for the year 1998 were also plotted and shapes of the error histograms for the periodic components and for the SOM were least satisfactory, while the best profile was achieved by MLP algorithm. However, all the methods gave fairly good estimates but the combination of a linear and neural method did not yield any advantage over the direct application of neural networks.

Vaziri, 2001 compared artificial neural network (ANN) and regression modeling to model motor vehicle air pollution emission attributes for Tehran (Iran). The emission data consisted of 2000 records extracted from the Tehran Air Pollution Database. The relevant emission information (consisted of variables namely, emission density of CO, hydrocarbons, oxygen and carbon dioxide), vehicle information and meteorological information were also used. The testing data showed an average RMSE reduction of 90%, when four ANN's predictions were compared. In the study, developed ANNs were found superior to the developed regression models as is clear from the RMSE. Thus, the regression modeling is ineffective in modeling the relationships among air pollutant, meteorological and motor vehicle characteristics but ANN seem to be a very useful tool.

Perez *et al.*, 2001 have shown that a three-layer neural network may be a useful tool to predict PM2.5 concentrations in the atmosphere of downtown Santiago, Chile several hours in advance when hourly concentrations of the previous day were used as input. The paper also stressed on the fact that improvement of predictions is possible by using another neural network for noise reduction on the original series which resulted in significant prediction error. In the same year, Perez, 2001 predicted the SO_2 concentrations in the atmosphere of Santiago, Chile and then showed the comparison of predictions using persistence, linear regressions and feed forward neural networks.

The neural network used here had 37 inputs, among which were the past values of SO_2 concentrations measured at different stations in the study area plus meteorological variables as temperature, wind speed and wind direction. The data corresponded to hourly averages for the period that goes from 18 May 1995–30 September 1995, and from 18 May 1996–30 September 1996 which include the winter seasons of years 1995 and 1996. From the results shown, this paper concluded that knowledge of the sequence of past values of sulfur dioxide concentrations is important in order to estimate its future values. The relevance of these past values appears to be greater than in the case of NO and NO_2 predictions (Perez and Trier, 2001). Meteorological conditions at the time of the intended prediction have also an important effect. Being able to forecast the weather seems of greater importance as compared with the case of PM2.5 prediction (Perez et al., 2001). It is remarkable that after training a three-layer neural network with 1995 data, author was able to generate reasonably good predictions for 1996 data. Thus, the results clearly indicated the superiority of NN as compared to others.

Lu *et al.*, 2002 developed an improved neural network model which combines both the principal component analysis (PCA) technique and the radial basis function (RBF) network to analyze and predict the concentration variations of six pollutants: Respirable suspended particle (RSP) , SO_2, NO_x, NO, NO_2 and CO measured hourly during 1999 at the Causeway Bay roadside gaseous monitoring station. The concentration of these six pollutants were used as original input variables and the corresponding RSP concentration level was used as the output of the models. In the study, PCA was used to reduce and orthogonalize the original variables. Such orthogonalization make a neural network more easily trained due to filtering of the noise that exists in the data set. The orthogonalized variables treated are then used as input vectors in a RBF neural network model to forecast the pollutant levels, e.g., the RSP level in the downtown area of Hong Kong until the network training error achieves the given error. This improved method was evaluated based on hourly time series RSP concentrations collected at the Causeway Bay roadside gaseous monitoring station in Hong Kong during 1999. The simulation results showed the effectiveness of the model. The proposed algorithm involving the combination of RBF and PCA technique proved to be much more effective than the traditional RBF network and also acted as a better alternative to multilayer feed-forward neural networks because of simpler network architecture and faster learning speed without any compromise with the generalization capability of the network. PCA technique removes the curse of dimensionality faces by MLP. In this research, an adaptive method was also used to decide the number of hidden nodes based on a trial-and-error procedure. Although it may give more nodes than actually needed, such an adaptive method is an effective potential RBF approach in pollutant modeling. In this study, although the performance of the PCA/ RBF network for long-term prediction is not studied here, the better properties expressed in the given case study provide sufficient evidence to conclude that it is worth pursuing further research using the model proposed. Although this work ignored the impact of climatic conditions, satisfactory results are still obtained using the improved PCA/RBF method. It can be foreseen that, if meteorological parameters are considered, better prediction results can be achieved.

Viotti *et al.*, 2002 used ANN to forecast short and middle long-term concentration levels for benzene, NOx, CO and ozone. The results shown a good accord with the monitored data and allowed the use of ANN as the forecasting model on a 24–48 h basis requiring only the meteorological conditions and the traffic level.

Abdul-Wahab and Al-Alawi, 2002 applied NN to predict ozone concentrations as a function of meteorological conditions and various air quality parameters. The results of their study indicate that the ANN is a promising method for air pollution modeling.

Chelani *et al.*, 2002 predicted SO_2 concentration by using an ANN, and the predicted values were compared with the measured concentrations at three sites in Delhi. A multivariate regression model was also used for comparison with the results obtained by using the neural network model. The study results indicated that the neural network was able to give better predictions with less residual mean square error than those given by multivariate regression models.

Morabito *et al.*, 2002 reported a study of estimation and short time prediction of atmospheric pollutants using a multi-resolution dynamic forecasting system (MDFS), a combined approach of both neural networks (NN) and wavelet analysis concepts. Particularly, the prediction of the hydrocarbon (HC) in the air was carried out and an environmental database that refers to the Southern Italy area of the Messina Strait was used. The available database clearly showed the characteristics of regularity (cyclicality) of the atmospheric parameters: for example, the temperature attains the maximum values during mid-day, while during night and early morning it reaches its minimum values. The same can be said for the atmospheric pressure, whose regularity is strictly related to the period of the year the observations have been taken. But given system is not able to generalize in other periods of the year due to the effect that weather has on the presence and stabilization of the pollutants in the air. That means, a set of non-linear models need to be designed in different background weather conditions. For evaluation of the performance of the combined approach, standard deviation was used as a reference value. The proposed MDFS proved useful in predicting experimental time series at an improved accuracy level with respect to other known approaches. The use of a most complex system that utilizes the different time series derived from the multi-resolution decomposition permits us to achieve better resolution in prediction and finally, the best accuracy results. The MDFS is also most amenable to implementation in parallel processing.

Dorling *et al.*, 2003 worked on APPETISE project to quantitatively inter-compare the performance of Deterministic (DET) and Statistical Air Quality Models. For the same, a case study of Helsinki was used for the prediction of NO_x, NO_2 and PM10 concentration. The experimental data of Helsinki from 1996-1999 was used and included the concentration data of PM10, NO_x and NO_2, meteorological data and traffic flow data. Output was recorded as the concentration time series of PM10 and NO_2 at the stations of Toolo and Vallila of Helsinki. Index of agreement as a statistical evaluation of model performance showed that NN model performance is better for NO_2, compared with that for PM10. The comparison of NN model and Deterministic model says that NN are computationally more effective but DET models can be more easily extended to other locations and time periods. A tool named as JANN (Java Artificial Neural Network) was developed for air pollution modelling using Multi-layer Perceptron Neural Networks.

Gariazzo & Tirabassi, 2003 tested the combination of two methodologies i.e. ANN and dispersion modeling for air pollutant prediction. Results have shown better performance (correlation factor, r=0.90) as compared to earlier studies.

Lu *et al.*, 2004 again applied the combined approach of PCA/RBF to forecast RSP, NO_x and NO_2 concentrations in an hourly time series in Mong Kok urban area, Hong Kong. This case study also validated the superiority of the PCA/RBF network over the simple RBF network as found in the earlier work.

Zhang *et al.*, 2004 developed an idea of using adaptive neural network (AWNN) as a tool for prediction of air pollution abatement scenarios, and focused on avoiding overfit during training of AWNN. The wavelet concept has already been applied by Morabito *et al.*, 2002. The simplicity of the structure and the algorithm of AWNN, improves its efficiency of on-line learning. Author obtained the data in the form of hourly NO_x and NO_2 from the monitoring site of the Bureau of the Environment of Heilong Jiang in Jiamu Si City in 1995 and 1998. The hourly metrological data were obtained for the same period from the Bureau of the Weather of Heilongjiang. The meteorological variables used in this work were similar to that used by Gardner and Dorling, 1999. About 39000 data were taken to train or test AWNN. Instead of the emission factors, the network is given two additional time of day inputs consisting of the sine and cosine of the time of day normalized between 0 and 24 h. The samples were divided into two sets—training set (data of 1995) and testing set (data of 1998). The MAE-mean absolute error was used as performance statistics calculated over the whole year. Results have shown that the accuracy of AWNN is low in the beginning of learning, but the errors gradually decrease after about seven days learning. The predicted line derives from the actual line between about 10-15 January because the windy weather is very different from that of early month, which forces AWNN to learn again to suit the new weather condition; in the same time, this also indicates AWNN can predict well again after several days learning when it meets a new condition. Thus, the WNN can work similar to other static NN when the condition of prediction is stable. Results have also shown that the new algorithm ensures good generalization of AWNN during online learning when the condition of prediction is unstable, but the classic algorithm cannot ensure the generalization during the same course.

Ahmad *et al.*, 2004 utilized the combination of ANN and genetic algorithm (GA) to find the optimal operating conditions in boiler emission processes from the palm oil mill. In the study, the procedure adopted was as follows. First GA writes the selected input parameters that are written in the text file. The text file is then read by the ANN and received as a new input parameter. Then, ANN will predict the output value. The output generated from this prediction is compared with the pollutant limit. If the value exceeds the limit, GA generates new input parameters from the GA operator, i.e. mutation and crossover. These steps are repeated until the optimal input values of fuel are found. This was an iterative process at the end of which the GA arrives at the optimum set of fuel parameters which produce emission within acceptable limit. Commercial software of MATLAB Version 5.3 and Neural Network Toolbox from Mathworks Inc. (MATLAB User's Guides, 1998) were used to generate neural network Modeling.

By considering the equipment provided and the operation time for palm oil mill, the data collected were 120 sets for the first mill and 65 sets for the second mill in 15 min interval time for several days. The data taken were then divided into two major parts: simulation (training and validation) and testing according to the rule set by Environmental Protective Agency of United States of America (EPAUSA). The tool developed from this study can be utilized to predict and control any pollution. Thus, GA was employed to find the optimal operating conditions so that the overlimit release of emission is reduced to the allowable limit. This paper has shown a new dimension of regulating air pollution directly from industries.

Niska *et al.*, 2004 worked on the problem of neural network architecture selection which is a time consuming task. The paper presented a study where a parallel genetic algorithm (GA) was employed for selecting the inputs and designing the high-level architecture of a multi-layer perceptron model for forecasting hourly concentrations of nitrogen dioxide at a busy urban traffic station in Helsinki. In this context, the evolutionary and genetic algorithms (GA) have proven to be powerful techniques due to their ability to solve linear and non-linear problems by exploring all regions of the state space and exploiting promising areas through genetic operations. The main drawbacks related to the using of GAs for optimizing NNs have been high computational requirement and complex search space.

Input data consisted: the concentration data comprising of hourly concentrations of NO_2, NO_x, O_3, PM10, SO_2 and CO monitored at the urban air quality monitoring station in Toolo (in Helsinki central) and meteorological data.

The test data set was extracted from the APPETISE (Air pollution Episodes: Modelling Tools for Improved Smog Management, http://www.uea.ac.uk/env/appetise/) database. The data set comprised the concentrations of airborne pollutants and meteorological soundings and observations, monitored in Helsinki metropolitan area during the years 1996–1999. Out of which, the data from the years 1996-1998 was used as training data and data from the year 1999 as model validation data. The fitness of model was assessed from the observed and predicted values by calculating the index of agreement (IA), which is a dimensionless measure limited to the range of 0–1 and thus, allows the comparison of different models. The results showed that GA is an applicable technique in this domain; it is capable of searching feasible high-level architectures and particularly reducing the need of computational efforts by eliminating irrelevant inputs. In the case of air quality forecasting this can also imply smaller costs due to the smaller amount of measurements required. No clear connection between architectural issues and performance was found, so it was suggested to use somewhat simpler architecture instead of complex one in order to minimize the risk of noise over-fitting.

Sahin *et al.*, 2004 applied multi-layer perceptron NN model to predict daily CO concentrations using metereological variables as predictors for the European part of Istanbul, Turkey.

Nunnari *et al.*, 2004 did inter-comparison between several statistical techniques for modeling SO_2 concentration at a point by using neural networks, fuzzy logic, generalized additive techniques, and other recently proposed statistical approaches. They found that artificial neural network-based models and neuro-fuzzy models were the most promising.

Abbaspour *et al.*, 2005 presented a new structure based on neural networks e.g. cerebellar model (CMAC model). This model can be used as the black box approach which shows superior ability to memorize and predict unknown time series data. The CMAC model was modified into TD-CMAC to predict future behavior of air pollutant density (carbon monoxide) by using the current and past values of observed and collected data. The simulations were conducted for 1-hour-ahead prediction and for 24-hour-ahead prediction (one-day-ahead) of air pollutants time series to compare the prediction capability of TD-CMAC, conventional CMAC (TW-CMAC) and a multi-layer perceptron (MLP) neural network. Time series data used in this evaluation were taken from CO values recorded and averaged at Villa station in Tehran, Iran from October 3rd 2001 to March 14th 2002 at one-hour intervals (3912 samples). In these simulations, the models, learning algorithms, output mapping and inputs were the same, but the outputs were different. The efficiency evaluation revealed that under equal conditions TD-CMAC has better prediction capabilities than the TW-CMAC and MLP neural network. Still the short-term prediction provides better result than the long-term prediction. Features like usage of less memory with lower hardware implementation costs, simplicity and ability to conduct online implementation, all makes TD-CMAC model the better option than MLP and TW-CMAC.

Jef *et al.*, 2005 examined the feasibility of a neural network short term forecasting model for ambient PM10 concentrations in Belgium. In this research, the dataset consisted of half-hourly PM10 concentration from the ten measuring stations of Belgium for the period 1997-2001. The two input parameters of Model 1 were the forecasted boundary layer height and the PM10 measurements of the morning of day0. This model got a reasonable accuracy and was tested in online operational mode (IRCELCELINE). By extending the model with other input parameters: cloud cover, day of week and wind direction, the increase in the performance was noticed with more forecast accuracy because these parameters contained explanatory value for the PM10 phenomenon in Belgium, complementary to the two inputs of Model 1. The author clearly stated that gain in additional accuracy from the use of temperature or the wind speed was not marginal. Thus the features carrying useful information for the prediction of PM10 concentrations were selected from the dataset. This brings the author to the conclusion that day to day fluctuations of PM10 concentrations in Belgian urban areas are to a large extent driven by meteorological conditions and to a lesser extent by changes in anthropogenic sources.

Sharma *et al.*, 2005 gave a review of ANN techniques in vehicular pollution modeling.

Nagendra and Khare, 2005 discussed the development and performance evaluation of ANN-based vehicular exhaust emission (VEE) models for predicting 8-h average CO concentrations at two Air Quality Control Regions (AQCRs), one representing a traffic intersection (AQCR1) and other, an arterial road (AQCR2), in the city of Delhi, India. The eight hourly CO concentration data have been collected from Central Pollution Control Board (CPCB), New Delhi for a period of 3 years from January 1997 to December 1999, for both the AQCRs. The meteorological data including 8-h average observations of cloud cover, pressure, mixing height, sunshine hours, visibility, temperature, wind speed, wind direction and humidity have been collected from Indian Meteorological Department, New Delhi. The eight hourly average traffic characteristics data have been collected from Central Road Research Institute (CRRI), New Delhi for the respective AQCRs. The vehicles have been classified into four groups i.e. two wheelers, three wheelers, four wheelers gasoline powered and four wheeler diesel powered, for which emission factors, developed by Indian Institute of Petroleum, have been used for estimating CO and NO_2 source strengths. The performance of all the developed models was evaluated on the basis of index of agreement (d) and other statistical parameters. The forecast performance of the developed models, with meteorological and traffic characteristics (d=0.78 for AQCR1 and d=0.69 for AQCR2) and with only meteorological inputs (d=0.77 for AQCR1 and d=0.67 for AQCR2), were comparable with the measured data. The results show that ANN-based CO models, with both meteorological and traffic characteristic variables and with only meteorological variables show best performance on the test data set at both the AQCRs. The study also shows that elimination of traffic characteristic variables from the model inputs causes negligible effect on model performance. However, the models developed with only traffic characteristic inputs showed poor performance on the test data set at both the AQCRs and reflected their inability to take into account the 'lag-effect'.

Cigizoglu *et al.*, 2005, studied the air pollution parameter estimation in Istanbul based on two ANN methods, radial basis function (RBF) and feed forward back propagation (FFBP), and comparison was done with multi-linear regression (MLR) method. In the study the SO2, NO and CO measurements in the Asian part of Istanbul, Ümraniye, for November 1999-April 2003 period were used. One measurement was taken during a day. The data was not continuous and had gaps. The paper is divided in two parts. Firstly, FFBP and RBF methods were employed to estimate SO2 using the past and present NO measurements. Secondly, the study was repeated again to estimate SO2 but this time using the past and present CO measurements. Both methods provided significantly superior estimations compared with the conventional MLR method. The performance evaluation criteria of FFBP for the testing period was slightly better than RBF. However, multiple FFBP simulations were required until obtaining satisfactory performance criteria and this total duration was longer than the unique RBF application. The presented study has shown that both ANN methods could be successfully employed to infill the gaps in the air pollution record.

Elminir *et al.*, 2006 applied ANN model to estimate air pollutant concentrations (PM10, CO and NO_2) in Egypt. The work was based upon measurements of different indicators performed during the period 2000 to 2002 at Abbassyia station, Egypt. The input variables used in the model were ambient air temperature, relative humidity, wind speed and wind direction. The output represented PM10, CO, and NO_2 concentrations. The generalization of the model was tested by correlation coefficient. Based on the results obtained, annual correlation coefficient was found to be 0.997, 0.969 and 0.978 for PM10, CO and NO_2 respectively. These results show that approximately 98% of the variation in the dependent variables (output parameters) can be explained by the independent variables (input parameters) selected. Then the author has used the mean bias error (MBE) to describe how much the ANN model underestimates or overestimates the situation. MBE results indicate that, ANN model always tend to over-estimate air pollutant concentrations, but remain in a domain of errors for which this model can be applied with good accuracy. Thus the overall results demonstrated that, ANN model can estimate air pollutant concentrations in urban area, for the given data set with an accuracy of approximately 96%. Elminir *et al.,* 2006 also showed their concern for the inherent limitations of ANN. The main limitation is the extension of model in terms of time period and location; this always requires training with locally measured data. According to the author, the ANN models cannot therefore be recommended for analyzing various air pollution abatement scenarios for future years.

Pasero *et al*., 2006 evaluated the performance of MLP and Support Vector Machines (SVM) for the accurate prediction of the time evolution of air pollutant. The main aim of this research was the medium-term forecasting of the air pollutants mean and maximum values by means of meteorological actual and forecasted data. Author focused on the selection of features and the modeling and processing techniques based on the theory of ANN, using MLP and SVM. The study provided an example of application based on data measured every hour by a station located in the urban area of the city of Goteborg, Sweden. The database considered was based on meteorological and air pollutant information sampled for the time period January 2004 to October 2005. MLP performance, both for the samples under the threshold and for the samples above the threshold, increased when the number of input features increased. Results were obtained with 5115 samples of days under the threshold and 61 samples of days above the threshold. It can be noted that the probability to have a false alarm is really low (0.82%) while the capability to forecast when the concentrations are above the threshold is about 80%. In the case using SVM method, the probability to have a false alarm was higher than in the MLP (0.96%) but the capability to forecast when the concentrations were above the threshold was nearly 90%. Thus, the use of Multi Layer Perceptrons and Support Vector Machines were proposed as an efficient strategy to perform an accurate prediction of the time evolution of air pollutant concentration.

Athanasiadis *et al,* 2006 studied the application of classification data mining algorithms for the development of an operationally efficient OF module and the comparison of their performance with statistical analysis methods. Algorithms used were iBK, Kstar, Nnge (Nearest Neighbor With Generalization), rule-based classifiers, as Conjunctive Rules, OneR, Decision Tables, decision trees (C4.5, ADTrees), along with Bayesian Classifiers (NaiveBayes), and Neural networks (Voted Perceptron). All these algorithms are implemented in WEKA (open source software) and was used for the data mining experiments in this work. In this study, the overall forecasting performance of all algorithms was found within accepted range but the most efficient algorithm for the experiments conducted is C4.5(J48), which is a decision tree learning algorithm. Author managed to conclude that classification methods should be considered as appropriate for operational air quality forecasting applications. In addition to this, results clearly indicate that classification methods could be employed together with statistical methods and other "fast" data analysis and prediction algorithms, for the creation of operational air quality forecasting modules that may effectively support operational air quality management on a day-to-day basis, in line with contemporary EU environmental legislation.

Bianchini *et al.*, 2006 proposed a novel approach of using cyclostationary neural network (CNN) architecture to model and predict hourly NO_2 concentration being independent from exogenous data. With the supposition that the relevant meteorological influence is already present in the time series, it uses only the time series of NO and NO_2 concentrations for prediction and no meteorological data was taken into account. CNN is composed of a number of MLP blocks equal to the estimated cyclostationary period in the analyzed phenomenon. Cyclostationary process is a subclass of non-stationary processes which vary periodically with time. The experiment was done on the data gathered from ARPA (Regional agency for environmental protection), Lombardia (Northern Italy). The task consisted in modeling the NO_2 time series, based on the past concentrations of NO and NO_2. In this case, it was evident that a strong correlation exists between the past NO data and the current value of the NO_2, with a daily periodicity. This means that the NO_2 pollution at time t + 1 depends on the NO sampled at t − 24, t − 48, etc. Therefore, the considered process has a cyclostationary period T = 24. Consequently, a CNN model composed by 24 MLP blocks was used to face the prediction task. Completely independent of exogenous data, such as weather condition (i.e. pressure, wind, humidity, etc.) and geographic information, focus was only be on the NO_2 estimation and the resulting model expected to be much more robust against noise and prediction errors. Preliminary experiments were very promising and have shown a significant improvement in performance, together with a low computational cost and faster speed for the CNN learning phase with respect to standard statistical tools. CNN performed better near the peak hours.

Boger, 2006 explained the modeling of NO_x emission from the production plants in Iran utilizing ANN model. For the training of ANN, the Guterman-Boger (GB) training algorithm set was used because this algorithm can easily train large scale ANN models, as it starts from non-random initial connection weights, obtained by the assumption that the inputs and outputs of the training data set are linearly related. Initially, the ANN model was trained with the database at 5 minutes interval collected in the January-July 2005 period to predict the NO_x concentration. Subsequent analysis of a trained ANN model has shown that causal index algorithm for knowledge extraction was found to be very useful in relating each input change influence on the relative magnitude and sign changes of each output. Results of this analysis showed that gas absorption temperature is one of the important operating parameter affecting the NO_x emission and then an ANN model based on the daily averages was trained, thus eliminating the diurnal temperature change effect. Overall, the 5 minute NO_x concentration at the stack modeling results have given a mean average error of 0.6% between the actual measurements and the ANN model, with a standard deviation of 6.7%. The daily average ANN model was trained to give the total NO_x emission and gave a mean model error of 0.006 Kg/Hr NO_x, with a standard deviation of 0.61 Kg/Hr. Thus, the ANN models trained from daily plant feature averages proved more informative than the 5 minute data, although it may be the results of the importance of the diurnal temperature changes in this plant.

Ozcan *et al.*, 2006 proposed a study of estimation of methane values of Istanbul (Turkey) landfill area using multi-layered ANN approach for the measured data using 4 parameters (CH_4, Co_2, CO, temperature) as inputs during 1 year (July 2002-April 2003). In the present model, NN was trained and tested using Matlab 6.0. Statistical performance indices: 1) Mean bias error (Bias); 2) Mean absolute error (MAE); 3) Root mean square error (RMSE); 4) Correlation coefficient (r); and 5) index of agreement were calculated: Bias, 3.04; MAE, 7.98; RMSE, 10.95; R, 0.81; and d, 0.88. Extraordinary performance of ANN was found indicating it as a promising technique for parameter estimation of landfill areas.

Agirre-Basurko *et al.*, 2006 presented the ANN model to forecast O_3 and NO_2

concentrations in the Bilbao area, Spain. The primary goal of the work was to build an accurate statistical model to forecast O_3 and NO_2 levels k hours ahead (k = 1, 4, 5, 6, 7, 8) in the Bilbao area. Meteorological variables, air pollution variables and traffic variables (the number of vehicles NV, the occupation percentage OP, and the variable KHZ (NV/OP), which gives an idea of the velocity) were used to develop the model. The data used in this work were hourly current data and historical data from the air pollution network and the traffic network of Bilbao during years 1993–1994. The study was limited to four stations in Bilbao, namely Elorrieta, Txurdinaga, Mazarredo and Deusto. Two multilayer perceptron-based models (MLP1 and MLP2) and one multiple linear regression model (MLR) were developed. After introducing the appropriate inputs, the outputs of the models were the forecasted O_3 or NO_2 levels at time t+k (k = 1,., 8).

The MLP neural networks used in this work were trained using the scaled conjugate gradient (SCG) algorithm. At the same time, in order to avoid overtraining, the early stopping technique was applied: the first 85% of the data from 1993 were the training set, the last 15% of the data from 1993 formed the validation set and the data from 1994 were chosen to be the test set. A joint study of the values obtained from the statistics of the Model Validation Kit showed that MLP models performed better than the MLP based model in 75% of cases (in six forecasts up to eight) in the Bilbao area. Moreover, for k = 1, 4, 5, 6, 7, 8 h ahead the MLP2 model provided the most accurate forecasts of O_3 and NO_2 at time t+k in the studied area. Nevertheless, for k = 2, 3 the most accurate forecasts of O_3 and NO_2 at time t+k were provided by the LR model (Agirre, 2003). The cause of this uncertainty was unknown that time but suggested to get resolved by choosing new dataset for MLP.

Lira *et al*., 2007 investigated the forecasting capability of linear models (such as ARX, ARMAX, output-error and Box-Jenkins), and neural networks. Input data included 24-h PM10 concentration of the present day and meteorological variables from the city of Uberlandia (Brazil) during the period of 2003-2005 and the output foreseen by the models obtained as 24-h PM10 concentration, with horizon of prediction of up to three days ahead. PM10 concentration data was collected by School of Chemical Engineering of the Federal University of Uberlândia (UFU). The meteorological data used in study were obtained in the climatic station of the Institute of Geography of the UFU. The software Matlab was used for NN and linear models were adjusted using the System Identification toolbox. For the neural network model (MLP), the learning algorithm used was Levenberg-Marquardt back-propagation (Neural Network toolbox). The transfer functions selected for the layers were hyperbolic tangent for the hidden layer and linear for the output layer. For the evaluation of the model's performance, three statistical measures were selected, namely the root mean square error (RMSE), the coefficient of determination (R^2) and the index of agreement (d). Results have shown that fairly good estimates can be achieved by all of the models, but Box-Jenkins model presented best fit and predictability.

Iliadis *et al.*, 2007 developed a reliable ANN model for estimating the tropospheric ozone concentrations at the site of Lykovryssi, Athens. Input variables included both the pollution parameters (CO, NO, NO2 and PM10) and meteorological parameters (mean air temperature, total solar radiation, mean pressure at sea level, the relative humidity, mean wind speed, NW_SE direction wind component and SW-NE direction wind component). Thus, in total, 4232 data records were used; 74% of them were used for the training procedure, while 26% of them were used for the testing procedure. Missing values were excluded. Input Contribution of input parameters was done during the testing phase by the Neuralworks II Plus software and the results have clearly shown that the CO and NO play the most important role in the formulation of Ozone. The Modular ANN gave the R^2 = 0.8827 and the RMS Error=0.1164 in the training phase, hereas in the testing process the R^2 = 0.742 and the RMS Error=0.156. Thus, the paper closed its discussion with the conclusion that ANN has proven to be a powerful tool offering a very reliable approach towards ozone concentration estimation. It can be applied reliably towards the design of environmental protection policy and management.

Tecer, 2007 proposed the application of ANN to predict the concentrations of SO_2 and PM at two different stations in Zonguldak city (Turkey). The 24-hour SO_2 and PM concentration averages were used as pollutant parameters for the period January-December 2002. The input for the model included the meteorological variables provided by the Governmental Meteorology Office and SO_2 and PM concentrations pertaining to the day the meteorological data was taken. In this study, a three-layer recurrent network was used. Results have shown that the predicted pollutant concentrations were concordant with the observed concentrations on a majority of the days. The best results were obtained in the training data set. At the Bahçelievler station, the determination coefficient (R2) between observed and predicted values of SO2 concentrations for the training data was 0.829 and for testing data was found to be 0.668. For PM concentrations, R2 was 0.820 for the training set, and 0.808 for the testing set, respectively. The determination coefficient indicates that the fitted model explains the percentage of the variability between observed values and the neural network model predictions. ANOVA analysis was performed to check the fitting between observed values and ANN prediction. The P-value in the ANOVA analysis was found to be less than 0.01, which indicated a statistically significant relationship between the variables at the 99% confidence level. Author concluded that knowledge of the sequence of past values of air pollutant concentrations is of considerable significance in order to predict its future values. Meteorological conditions at the intended time of prediction also have an important effect on air pollution modeling. It was also shown that the neural network model provides a good agreement with measured values of air pollutants.

Caselli *et al.*, 2008 presented the prediction of PM10 concentration utilizing feed-forward neural network and its comparison with multivariate regression model. Data used were the meteorological data and data of PM10 from San Nicola monitoring station during the period January 2005 to March 2006. The prediction of PM10 concentration

12h and 36h before gave an RMSE of 26% and 30% respectively and the correlation coefficient between the observed and predicted values was 0.7 which showed a good accuracy by ANN. Comparison of results obtained by neural network with those of multivariate regression for one and two days forecasting was done and found in concordance with previous researches till date. Neural network outweighed multivariate regression modeling.

Abdul-Wahab *et al.,* 2008 developed ANN models for the prediction of ground-level SO_2 in the Sultanate of Oman. This study indicated the potential of the neural network approach for capturing the non-linear interactions between SO_2 levels and meteorological variables and for the identification of the relative importance of these variables. In this work, two ANN models were generated with different objectives. The first model (Model I) was used to predict SO_2 levels at certain receptors from the Mina Al-Fahal refinery in Oman. Second model (Model II) was used for the prediction of first three maximum SO_2 concentrations and their corresponding locations with respect to the refinery. The Feed Forward network using the Back-Propagation (BP) algorithm was used to develop these two models. The training process was performed using the NeuroShell simulator. The R^2 value for the training set was 0.9693. The statistical analysis of these results indicated that the R^2 value for the testing set was 0.9666, the mean squared error was 8.902 and the mean absolute error was 1.173. Thus, the model prediction was in good agreement with the actual results.

The models were also used to determine meteorological conditions that most affect SO_2 concentrations. To find the percent contribution of each of the input variables with respect to the output variables, the partitioning method of the connection weights proposed by Garson was used. The method involves partitioning the hidden-output connection weights of each hidden neuron into components associated with each input neuron. Some of these methods are (1) Partial Derivatives or 'PaD'; (2) the "weights method or the partitioning method used in the current work; (3) the "perturb" method; (4) the "Profile" method; (5) the "Classical Stepwise"; (6) the "improved stepwise a"; (7) the "improved stepwise". Using this method, for Model I, it was found that the highest contributing variable that affects the SO_2 concentration level is the wind direction; while in Model II, wind direction, stability and wind speed are the highest contributors. It was found that the neural network model has given superior predictions consistently and the results produced were encouraging.

Popentiu *et al.,* 2009 applied the artificial neural network approach in order to monitor the urban pollution level from "Moldocim" cement plant, in Tasca (Romania). Their work stressed on the fact that the study of air pollution should be done with regard to the weather patterns of the local area because the pollutant concentration is strongly influenced by the movements and characteristics of the wind into which they are emitted. For this purpose, a professional computer aided modeling and pollution control tool (referred to as *PoLogCem Software*- Pollution Logistic Cement) was used. Results have clearly shown the considerable impact of wind speed and wind direction on pollution distribution.

Postolache *et al.*, 2009 proposed a novel strategy to assess the indoor air quality for the first time. This work presented a Wi-Fi indoor-outdoor air quality monitoring network along with the combination of capabilities of tin oxide sensors using an advanced sensor data processing based on multilayer perceptron neural networks for an accurate measurement of air quality. Neural network algorithm was implemented in JavaScript for the calculation of several air quality values.

Arjun Akkala *et al.*, 2010 for the first time demonstrated the use of ANN approach to model Radon concentrations in Ohio, US. Along with ANN, a new technique named as knowledge based neural networks (KBNNs) was used to improve the reliability of the estimations from ANN. In the KBNNs used in this work, knowledge in the form of radon concentration data and uranium concentration data is embedded into the neural network structure. In the ANN approach, the fact that radon is formed from uranium by the decay chain has been considered. And the neural networks employed in this work were 3-layer multi-layer perceptrons. The databases of Radon concentration were requested from laboratories, university researchers, and others to compile a unified database consisting of 145,996 measurements for 1492 zip code areas in Ohio. It was observed that the houses and other structures built above uranium-bearing rocks or sediments may have higher indoor radon levels. This fact is corroborated by the uranium concentration map of Ohio with the radon concentration map. The observation represented knowledge, which could be used for further improving ANN model accuracies. Three approaches were applied for modeling radon and uranium concentrations using KBNNs, namely Prior Knowledge Input (PKI), Source Difference Model (SDM), and Space-Mapped Neural Network (SMNN).

Author compared the performance of these techniques to conventional interpolation techniques and the MLP3 approach. The SMNN has given the best performance in terms of *MAE*, *Fa2*, *RMSE*, and *FB*; whereas the PKI method gave the best performance in term of *NMSE*. Hence, the SMNN method was found to be more reliable in the scenario of predicting missing radon concentrations in Ohio. It could also be seen that all the knowledge based approaches performed more efficiently as compared to rest of the techniques such as RBF and Kriging and also simple MLP3 approach.

Selvaraj *et al.*, 2010 developed a method for short term prediction of ozone concentration using the neural network technique. The study area Chenbagaramanputhur, a rural place in Kanyakumari district was used. The parameters used to develop the model were: mean surface ozone concentration as response, nitrogen dioxide concentration as predictor, mean temperature as predictor and prevailing % relative humidity as predictor. Results shown the mean square error of the data during testing: 0.15375 ppb and the accuracy of testing data: 99.85 %. The above results validated the proposed model. Hence it is concluded that the given model can be used for predicting surface ozone concentration with nitrogen dioxide, temperature, % relative humidity as predictors.

In a conference Boznar *et al.*, presented their research work to construct a forecasting model of ozone concentration values of the following day. The intention behind this project was to inform citizens that a day with ozone alarm value is coming. Authors concentrated their research on the problem of maximal hourly value of ozone concentration that would appear in the following day. The tools, multilayer perceptron neural network and fuzzy logic both were used and all were implemented in the Matlab software package. The selected input features were: air temperature, global solar radiation, NO, NO_2, NO_x, CO, O_3, 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 (all three taken from the available measured database). 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 results. So, the author concluded that both the models are good enough to use the model for informing citizens about possibilities of high and alarm concentrations. Nunnari *et al.*, 2004 also tested the use of fuzzy logic and NN for SO_2 prediction and came up with the same results.

Kapegeridis *et al.*, presented their work on the development of a modular neural network (MNN) based on RBF networks for the prediction of 24h average PM10 concentrations for the next day at the city center of Kozani. The developed system was converted to a computer code (a DLL application extension) that was integrated with the Integrated Air Quality Management System (IAQMS) in August 2007. Since then, it has been used to forecast the 24h average PM10 concentration in Kozani. The daily forecasts become available through the Laboratory's website (http://www.airlab.edu.gr). The MNN was developed to receive several inputs from air pollution and environment monitoring stations at the center of Kozani and was trained by the data during the period 1 June 2004 and 31 May 2005 (12 months). The prediction capacity of the presented system has been demonstrated using observed and predicted concentrations collected since it became operational as part of the IAQMS. The study shows a mean absolute error: 13.7% on the validation set and 26.3% during application.

Conclusion

The above studies and discussion provide us an understanding of artificial neural network (ANN) technique in air pollution modelling. ANN models are computer programs that are designed to emulate human knowledge processing, speech, prediction, classification and control. ANN is a cellular information processing system designed and developed on the basis of perceived notion of the human brain and its neural system. This concept of artificial neural networks was established in 1943 (McCulloch and Pitts, 1943). Later, in 1958, the first practical artificial neural network was presented: *the perceptron* (Rosenblatt, 1958). In the last two decades, ANNs have been already explored in various fields like chemical research (Kvasnicka, 1990; Wythoff et al., 1990; Smits et al., 1992), physics research (Dekruger and Hunt, 1994), molecular biology, ecology and environmental sciences and demonstrated remarkable success.

In environmental studies, air pollution is an extremely significant issue that should be focused on all around the globe as it affects human health, ecosystems and the environment, destroys aesthetical conditions, gives harm to matters, buildings, and in turn economy. It is defined as a condition, in which the substances that result from both natural and anthropogenic activities are present in the air at concentrations sufficiently high above their normal ambient levels producing considerable impacts on humans, animals, vegetation, or materials.

Literature survey shows 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, particularly ANNs. All the meteorological variables and factors have a non-linear relationship with air quality, which can be accurately captured by non-linear models such as ANNs and Support Vector Machines.

The facts and the quality of results provided by ANNs make them more attractive to apply than other models. The advantages of these models are that they do not require very exhaustive information about air pollutants, reaction mechanisms, meteorological parameters or traffic flow and that they have the ability of allowing nonlinear relationships between very different predictor variables. Due to the growing development of computer-aided analysis, its easy accessibility to all researchers has also facilitated the application of ANNs in air pollution modeling.

Boznar *et al*. (1993) presented the earliest paper based on the use of neural networks for the prediction of sulphur dioxide (SO_2). Since then several authors have developed different neural network-based models to forecast air pollutant concentrations.

Easiest in implementation, multilayer perceptron (MLP) is the most used ANN method. Almost in all the findings where the comparison between ANN and regression method comes, the former takes the superiority level, mainly because of its nature to deal with non-linear relationships between the predictors, the more flexibility and the ability to make efficient use of proxy data when the optimum predictor variables are unavailable and the better predictions than those given by other methods e.g. multivariate regression models, statistical linear models, etc. But, the main advantages of neural networks (NN) are that air pollution concentration can be predicted with time series and basic meteorological variables. This enables the models to be easily constructed. MLPs perform very well with proxy data and are successful with both secondary (ozone) and primary pollutants. Secondary pollutant like Ozone, its formation is a complex, non-linear process that neural networks are able to capture more accurately without many of the usual limiting assumptions of other statistical methods. In the case, a pollutant shows a serial correlation i.e. present day concentration of a pollutant depends on previous day's concentration, regression model fails because it assumes that observations are statistically independent events. Although the inclusion of lagged data in regression modeling has shown an improvement in the accuracy of regression predictions, still ANN serves a much better alternative. The performance of MLP found superior to conventional statistical and stochastic methods in continuous flow series forecasting.

Instead of all the rewards, there are limitations with implementation of ANN. Some shortfalls faced when using multilayer perceptron (MLP/FFBP) can be listed as the dependency on different random weight assignments in the beginning of each training, their tendency to trap by the local minima, network overfitting, and curse of dimensionality and black box nature of NN. Performance of a network also depends on the choice of network architecture, the number of layers and nodes in those layers. There are no rules to help in this process and so, the number of input and output nodes is determined by the complexity of problem at hand. Different number of a particular layer neurons need to be tried in a particular study and the number that gives a minimum mean square error (MSE) are chosen for the final. This process increases the total duration of NN functioning.

The reliability factor that ANNs are gaining in the field of prediction, also draws more concentration for its advancement. Research shows the testing of different combinations of ANN and other algorithms which can include Genetic algorithm (GA) and Principal component analysis (PCA). GA technique is for the search of feasible high-level architectures of neural network and PCA works in the reduction of the dimensionality of input data and its orthogonalization. PCA and RBF combination executes prediction task with much accuracy than the conventional models.

Feature determination and pattern selection strategies considers as the most important step in the model construction because the selection of proper information with which to train the neural network is crucial for model's efficiency. Feature determination strategies seek for the most important and significant independent variables that describe the studied process. Pattern selection strategies construct the group of training samples in such a way that the model can learn the non-linear function behavior over the whole domain. Both strategies in the ideal case should ensure that the whole domain of the studied function is uniformly and densely enough covered with the training samples. Kohonen neural network based and the expert knowledge based pattern selection strategies have shown significant improvement in comparison to the model trained with unselected patterns. Approaches like Self organizing map (SOM) and two-segmented network are able to remove black-box nature of neural networks.

The different possibilities of using ANN variants in air pollution modeling includes Radial basis function (RBF), Change point detection model with RNM (CPDM), Sequential network construction model (SNCM), Self organizing feature maps (SOFM), Group methods of data handling (GMDH), TD-CMAC (an extension to the conventional Cerebellar Model Arithmetic Computer), AWNN etc. Among the above ANN methods, RBF, SOFM and TD-CMAC came up as the best tools. RBF networks possess the property of best approximation and shows a clear improvement in the learning efficiency. A new algorithm, adaptive wavelet neural network (AWNN) in which wavelet neural network (WNN) is a kind of RBF possesses many advantages than the general networks, such as faster convergence, avoiding local minimum, easy decision and adaptation of structure, so acts as an efficient tool for the prediction of air pollutants concentration along with the ability of avoiding overfit during the training.

Consequently, neural network technique is regarded as a reliable and cost-effective method for the task of prediction. In an operational setting for any specific place, neural network models could be improved through experimentation and fine-tuning. Likely strategies include using additional variables (e.g., extra weather elements, synoptic patterns, traffic flow or day-of-the-week information) or changing the nature of the model (e.g., changing the number of hidden nodes and layers in neural networks). Use of Fuzzy modeling independently and with ANN is also showing brighter side of soft-computing application in air pollution modeling.

References

Abbaspour, M., Rahmani, A.M. and Teshnehlab, M. (2005). Carbon monoxide prediction using novel intelligent network. International Journal of Environmental Science & Technology, 1, pp. 257-264.

Adame, J.M.B., Januchs, M.G.C., Andina, D. and Corona, A.V. (2009). Monitoring strategy to assessment the air pollution level in Salamanca (México).

Ahmad, A.L., Azid, I.A., Yusof, A.R., Seetharamu, K.N. (2004). Emission control in palm oil mills using artificial neural network and genetic algorithm. Computers and Chemical Engineering, 28, pp. 2709–2715.

Athanasiadis, I.N., Karatzas, K.D. and Mitkas, P.A. (2006). Classification techniques for air quality forecasting. In the fifth workshop on binding environmental sciences and artificial intelligence, 17th European conference on artificial intelligence, pp. 41-47.

Barai, S.V., Dikshit, A.K., Sharma, S.: Neural Network Models for Air Quality Prediction: A Comparative Study.

Basurko, E.A., Berastegi, G.I. and Madariaga, I. (2006). Regression and multilayer perceptron-based models to forecast hourly O3 and NO2 levels in the Bilbao area. Environmental Modelling & Software, 21, pp. 430–446.

Bianchini, M., Iorio, E.D., Maggini, M., Mocenni, C., Pucci, A. (2006). A Cyclostationary Neural Network Model for the Prediction of the NO2 Concentration. In the proceedings of European Symposium on Artificial Neural Networks, pp. 67-72.

Boger, Z. (2006). Artificial Neural Networks Modeling to Reduce Industrial Air Pollution.

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

Boznar, M., Lesjak, M., and 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, Vol. 27B, pp. 221-230.

Božnar, M.Z., and Primož Mlakar, P. (2001). Artificial neural network-based environmental models. Air Pollution Modeling and Its Application XIV, pp. 483-492.

Boznar, M.Z. and Mlakar, P. (2002). Use of neural networks in the field of air pollution modelling. Air Pollution Modeling and Its Application XV, pp. 375-383.

Caselli, M., Trizio, L., Gennaro, G.D. and Ielpo, P. (2008). PM10 Forecasting by a Feed Forward Neural Network and Comparison with Multivariate Linear Regression Model.

Chelani, A.B., Raoi, C.V., Phadke, K.M. and Hasan, M.Z. (2002). Prediction of sulfur dioxide concentration using artificial neural networks. Environmental Modelling & Software, 17, pp. 161-168.

Cigizoglu, H.K., Alp, K. and Kömürcü, M. (2005). Estimation of air pollution parameters using artificial neural networks. Advances in Air Pollution Modeling for Environmental Security, pp. 63–75.

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

Comrie, A.C. (1997). Comparing Neural Networks and Regression Models for Ozone Forecasting. Air & Waste Management Association, 47, pp. 653-663.

Dorling, S., Cawley, G., Kukkonen, J. and Karppinen, A. (2003). Air quality forecasting using neural network approaches. Air quality forecasting meeting, Culham.

Elminir, H.K. and Galil, H.A. (2006). Estimation of air pollutant concentrations from meteorological parameters using artificial neural network. Journal of Electrical Engineering, 57, pp. 105–110.

Fabbian, D. and Dear, R.D. (2007). Application of Artificial Neural Network Forecasts to Predict Fog at Canberra International Airport. Weather and forecasting, 22, pp. 372-381.

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.

Gardner, M.W., Dorling, S.R. (1999): Neural network modelling and prediction of hourly NO_x and NO_2 concentrations in urban air in London. Atmospheric Environment, 33, pp. 709-719.

Giorgio, F. and Piero, M. (1996). Mathematical models for planning and control-ling air quality. Proceedings of IIASA Workshop, 17.

Hadjiiski, L., and Hopke, P. (2000). Application of artificial neural networks to modeling and prediction of ambient ozone concentrations. Journal of Air and Waste Management Association, 50, 894-901.

Hall, T., Brooks, H.E. and Doswell, C.A. (1999). Precipitation forecasting using a neural network. Whether Forecasting, 14, pp. 338–345.

Iliadis, L.S., Spartalis, S.I., Paschalidou, A.K. and Kassomenos, P. (2007). Artificial Neural Network Modelling of the surface Ozone Concentration. International Journal of Computational and Applied Mathematics, 2, pp. 125–138.

Jef, H., Clemens, M., Gerwin, D., Frans, F. and Olivier, B. (2005). A neural network forecast for daily average PM10 concentrations in Belgium. Atmospheric Environment, 39, pp. 3279-3289.

Kapageridis, I., Evagelopoulos, V. and Triantafyllou, A. Prediction of PM10 Concentrations using a Modular Neural Network System and Integration with an Online Air Quality Management System.

Kolehmainen, M., Martikainen, H., Ruuskanen, J. (2001). Neural networks and periodic components used in air quality forecasting. Atmospheric Environment, 35, pp. 815-825. [15]

Kondo, T., and Pandya, A.S. (2000): GMDH-type Neural Networks with a Feedback Loop and their Application to the Identification of Large-spatial Air Pollution Patterns. In the Proceedings of 39th SICE Annual Conference International Session Papers, 112A-4, pp. 1-6.

Lek, S., Guegan, J.F. (1999): Artificial neural networks as a tool in ecological modeling, an introduction. Ecological Modelling, 120, pp. 65–73.

Lippmann, R. P. (1987). An introduction to computing with neural nets. IEEE Acoustics, Speech Signal Process. Magazine, 4, pp. 4–22.

Lira, T.S., Barrozo, M.A.S. and Assis, A.J. (2007). Air quality prediction in Uberlândia, Brazil, using linear models and neural networks. 17th European Symposium on Computer Aided Process Engineering – ESCAPE17, pp 1-6.

Lu, W.Z., Wang, W.J., Fan, H.Y., Leung, A.Y.T, Xu, Z.B., Lo, S.M. and Wong, J.C.K. (2002). Prediction of Pollutant Levels in Causeway Bay Area of Hong Kong Using an Improved Neural Network Model. Journal of Environmental Engineering, 128, pp. 1146–1157.

Lu, W.Z., Wang, W.J., Wang, X.K., Yan, S.H. and Lam, J.C. (2004). Potential assessment of a neural network model with PCA/RBF approach for forecasting pollutant trends in Mong Kok urban air, Hong Kong. Environmental Research, 96, pp. 79–87.

Marzban, C. (2004). The ROC curve and the area under it as a performance measure. Weather Forecasting, 19, pp. 1106–1114.

Melas, D., Kioutsioukis, I., and Ziomas, I.C. (2000). Neural network models for predicting peak photochemical pollutant levels. Journal of Air and Waste Management Association, 50, 495-501.

Morabito, F.C. (2002). Wavelet neural network processing of urban air pollution. pp. 1-6.

Nagendra, S.M.S. and Khare, M. (2005). Modelling urban air quality using artificial neural network. Clean Technology Environment Policy, 7, pp. 116–126.

Nikov, A., Karaca, F., Alagha, O., Kurt, A., Hakkoymaz, H. (2005): AirPolTool: A web-based tool for Istanbul air pollution forecasting and control, to be published in Proc. of 3rd International Symposium on Air Quality Management at Urban.

Niska, H., Hiltunen, T., Karppinen, A., Ruuskanen, J., Mikko Kolehmainen, M. (2004). Evolving the neural network model for forecasting air pollution time series. Engineering Applications of Artificial Intelligence, 17, pp. 159–167.

Nunnari, G., Dorling, S., Schlink, U., Cawley, G., Foxall, R. and Chatterton, T. (2004). Modelling SO_2 concentration at a point with statistical approaches. Environmental Modelling & Software, 19, pp. 887-905.

Oprea, M. and Matei, A. Applying Artificial Neural Networks in Environmental Prediction Systems. Recent advances in automation and information, pp. 110-115.

Osródka, L., Wojtylak, M. and Krajny, E. Forecasting of high-level air pollution in urban-industrial agglomeration by means of numerical weather forecasting.

Ozcan, H.K., Ucan, O.N., Sahin, U., Borat, M. and Bayat, C. (2006). Artificial neural network modeling of methane emissions at Istanbul Kemerburgaz-Odayeri landfill site. Journal of Scientific & Industrial Research, 65, pp. 128-134.

Parker, D.B. (1982). Learning logic. Invention report S81-64, File 1, Office of Technology Licensing, Stanford University.

Pasero, E. and Mesin, L. 2006. Artificial Neural Networks to Forecast Air Pollution. pp. 221-240.

Perez, P. (2001): Prediction of sulfur dioxide concentrations at a site near downtown Santiago, Chile. Atmospheric Environment, 35, pp. 4929-4935.

Perez, P. and Reyes, J. (2001): Prediction of Particulate Air Pollution using Neural Techniques. Neural Computing & Applications, 10, pp. 165–171.

Popentiu, F., Lungu, A. and Creosteanu, R. (2009). Modelling wind speed with neural networks for monitoring the air pollution. Fascicula de Energetică, 15, pp. 280-283.

Postolache, O.A., Pereira, J.M.D. and Girão, P.M.B.S. (2009). Smart Sensors Network for Air Quality Monitoring Applications. IEEE transactions on instrumentation and measurement, 58, pp. 3253-3262.

Rao, M.N., and Rao, H.V. (2000). Air Pollution. Tata McGraw-Hill, New Delhi.

Rummelhart, D. E., and McClelland, J.L. (1986). Parallel Distributed Processing. The MIT Press, 1.

Sahin, U., Ucan, O.N., Soyhan, B. and Bayat, C. (2004). Modeling of CO distribution in Istanbul using artificial neural networks. Fresen Environ Bull, 13, pp. 839-845.

San, Y. and Zhang, Z. (2004). Adaptive wavelet neural network for prediction of hourly NO and NO_2 concentrations. Proceedings of the 2004 Winter Simulation Conference, pp. 1-9.

Selvaraj, R.S., Elampari, K., Gayathri, R. and Jeyakumar, S.J. (2010). A neural network model for short term prediction os surface ozone at tropical city. International Journal of Engineering Science and Technology, 2, pp. 5306-5312.

Sharma, N., Chaudhry, K.K. and Rao, C.V.C. (2005). Vehicular pollution modeling using artificial neural network technique: A review. Journal of Scientific & Industrial Research, 64, pp. 637-647.

Shi, J. P. and R.M. Harrison (1997): Regression modelling of hourly NO_x and NO_2 concentrations in urban air in London. Atmospheric Environment, 31, pp. 4081-4094.

Spellman, G. (1999). An application of artificial neural networks to the prediction of surface ozone concentrations in the United Kingdom, Application Geography, 19, pp. 123-136.

Tecer, L.H. (2007). Prediction of SO_2 and PM Concentrations in a Coastal Mining Area (Zonguldak, Turkey) Using an Artificial Neural Network. Polish Journal of Environmental Studies, 16, pp. 633-638.

Tirabassi, T., Pelliccioni, A. and Garazzo, C. (2003). Coupling of neural network and dispersion models: A novel methodology for air pollution models, Journal of Environmental Pollution, 20, pp. 136-146.

Tresp, V., R. Neuneier, and S. Ahmad, 1995: Efficient methods for dealing with missing data in supervised learning. Advances in Neural Information Processing Systems 7, G. Tesauro, D. S. Touretzky, and T. K. Leen, Eds., The MIT Press, 689–696.

Vaziri, M. (2001): Modeling vehicle air pollution attributes by artificial neural network. Scientia Iranica, 8, pp. 10-15.

Votti, P. and Genova, P. D. (2002). Atmospheric urban pollution: applications of artificial neural network (ANN) to the city of Perugia, Ecological Modeling, 148, pp. 27-46.

Wahab-Abdul, S.A. and Al-Alawi, S. M. (2002). Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks, Environ Model Softw, 17, pp. 219-228.

Wahab, S.A.A. and Alawi, S.M.A. (2008). Prediction of Sulfur Dioxide (SO_2) Concentration Levels from the Mina Al-Fahal Refinery in Oman Using Artificial Neural Networks. American Journal of Environmental Sciences, 4, pp. 473-481.

Webos, P. (1974). Beyond regression: new tools for prediction and analysis in the behavioral sciences. Thesis, Harvard University.

Werner, H., Obach, M. (2001): New neural network types estimating the accuracy of response for ecological modeling. Ecological Modelling, 146, pp. 289–298.

Yi, J. and Prybutok, R. (1996). A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialised urban area. Environmental Pollution, 92, pp. 349-357.