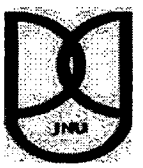
A MODELING STUDY OF AIR POLLUTANTS IN AN AMBIENT URBAN ENVIRONMENT USING ARTIFICIAL NEURAL NETWORK WITH PCA/RBF APPROACH

Thesis submitted to Jawaharlal Nehru University in partial fulfillment of the requirements for the award of the degree of

DOCTOR OF PHILOSOPHY

KUNDAN KUMAR



SCHOOL OF ENVIRONMENTAL SCIENCES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI-110067 (INDIA)
July, 2008

जवाहरलाल नेहरू विश्वविद्यालय
Jawaharlal Nehru University
SCHOOL OF ENVIRONMENTAL SCIENCES
New Delhi- 110 067, INDIA

CERTIFICATE

This is to certify that the research work embodied in this thesis entitled "***A Modeling Study of Air Pollutants in an Ambient Urban Environment using Artificial Neural Network with PCA/RBF Approach***", has bee carried out at the School of Environmental Sciences, Jawaharlal Nehru University, New Delhi, for the partial fulfillment of the degree of Doctor of Philosophy. This work is original and has not been submitted in part or in full, for any other degree or diploma of any other University.

Kundan
KUNDAN KUMAR
(CANDIDATE)

27/10/09

PROF. K. G. SAXENA
(DEAN)

Prof. K. G. SAXENA
Dean
School of Environmental Sciences
Jawaharlal Nehru University
New Delhi - 110067

PROF. V. K. JAIN
(SUPERVISOR)

Gram : JAYENU TEL : 26741557, 26742575, 26742676
SES : 26741538 (Direct) 26704302. 4303. 4304 Fax : 26741502

Acknowledgement

I wish to express my heartfelt gratitude to my adorable supervisor Prof. V.K.Jain for encouraging me to work in this emerging and challenging field of research (neural network modeling), constructive guidance with critical approach and encouragement without which this work would not have been possible.

I am thankful to Prof, Jain ., former Dean, School of Environmental Sciences, and Prof. K.G. Saxena , Dean, School of Environmental Sciences for extending necessary facilities required for carrying out my research work. My sincere thanks to all my teachers especially Dr. Krishan Kumar, Dr. Dimri for their encouragement and blessings.

It is one of the most difficult tasks to acknowledge all those who have lent their helping hand in my research work. My thanks are due to Dr. Shrivastava who always helped me a lot whenever I sought help. I shall remain indebted to Gajender jee for always keeping everything intact in the lab. I take this opportunity to thank my lab colleagues Bishoi, Homdutt, Ujjawal, Gautam, Deepak, Sandeep, Amit, Pradeep, Nasir, Rakesh, Sanyogita and Shushmita.

I shall remain indebted to my friends Sandeep, Akhilananda, Sunil, Mahesh and for their timely help and encouragement.

I am heartily thankful to various experts working in modeling field. Thanks are also due to Ramovatar, Rakesh, Nirmal, Sandeep, Uddhav, Aditya for their wishes and support.

I find fortunate to get such a caring and supporting family whose faith in me kept me in right direction always.

I owe my thanks to my all teachers of the School who taught during the course of study and showed how to proceed for new research.
My heartfelt thanks go to my lab colleague who created conducive atmosphere while working in lab.

I am also thankful to all those friends working in different area to help me in getting inputs while discussions.

I take this opportunity to thank our non-teaching staff community of our School. Mr. Sachdeva, Mr. Rawat, Mr. Amrik, Mrs. Bhatia and all others.

I would like to thank our Technical Officer Dr. P.D. Gaekawad, Mr. B.D. Sharma, Mr. Bhardwaj for always providing necessary support.

Thanks are also due to CPCB, Delhi for providing the required data for this research work. My special thanks are due to BARC for providing software BIKAS and library facility.

Lastly words are not adequate to express my love and gratitude for my parents and my brother whose moral and spiritual support has always stood me in good stead. There are many others whose help was of much importance in completion of this thesis.

Finally I would like to thank the almighty for keeping me on right track throughout against all odds.

Kundan

Kundan Kumar

Dedicated to...

My beloved brother

CONTENTS

List of figures

List of tables

CHAPTER – 1

Chapter 1 – Introduction

Air quality remains a key element in the reduction of risks to health from environmental pollution. Unlike other pollution, it is difficult to escape easily from the adverse impact of air pollution as a person directly breaths the air. Due to heavy industrialization across the globe, air pollution is one of the major environmental hazardous. This is different in the sense that not only the country responsible for polluting the air suffers but also the neighboring countries suffer although that country has no role in polluting the air. The air pollution has adverse impact not only on human being but also on vegetation and non-living things. So much effort is needed to control and reduce the quantum of air pollution.

Several Studies have been carried out across the world regarding monitoring of air pollutants and its adverse impact on living as well as on nonliving. Thereafter different criteria were formulated by the international agencies working in this area. The major pollutants of air are Suspended particulate matter (SPM), Sulphur dioxide (SO_2), Ozone (O_3), Oxides of Nitrogen (NOx), and Carbon monoxide (CO) in urban environments around the world. These pollutants are termed as 'criteria air pollutants or Classical air pollutants. These pollutants are produced by vehicles, industry, burning of solid wastes and natural dusts etc. Air pollution is identified primarily as an urban issue as the density of population, industry and infrastructure such as vehicular transport, buildings etc prevent the pollutants from being dispersed. So presence of one or more contaminants in the atmosphere in such quantity for such duration that is injurious to living being both human as well as plant is termed as Air pollution. Ambient air pollution may be caused either by nature or by anthropogenic activities

Ambient air pollution may be either natural or human generated. The volcanic eruptions of Mt St Helens is an example of natural pollution. On 18 May 1980, the (volcanic) eruption of Mt St Helens in southwest Washington State disrupted the lives of thousands and changed more than 200 square miles of rich forest into a grey, lifeless landscape. The Natural pollutants have the potential to pose a serious health hazard when they are

generated in significant quantities near human habitants. However, with few exceptions, natural pollution has not been much of an environmental concern. On the other hand, anthropogenic air pollution is viewed as a serious problem because the pollutants produced poses a great health hazard.

Recent studies suggest that even at the lower levels of airborne pollution experienced today compared with 1950s there are associations with premature mortality and chronic illness.(Anderson et al., 1987-92). As per WHO report, the global annual death from air pollution are estimated at more than 2.7 million, with 0.9 million in the cities and 1.8 million in rural settings. The United States Environmental Protection Agency in a published report indicated that diesel emissions are responsible for high cancer risks (US-EPA, 2002). Studies show that current levels of air pollution in the cities of many developed and developing countries are associated with increased rates of mortality and morbidity have heightened concern that air pollution continues to pose a threat to public health. (Bascom et al., 1996), Dockery and Pope, (1994). The evidence suggests that small airborne particles are a toxic component of urban air pollution.

So far as harmful effects due to SPM is concerned it posses direct effect as it is inhaled by human beings. SPM may contain fungal spores and pollen, which induce several allergic responses. Due to their acidic nature, sulfate and nitrate ions can even damage respiratory tract. Exposure of human being to high levels of SO_2 has been related to increase in hospital admission for chronic bronchitis (Ciccone et al., 1995), to low birth weights (Rogers et al., 2000) and to reduction in birth rates (Dejmek et al., 2000). The World Health Organization (1979) has determined that the safety limits for SO_2 concentrations are between 100 $\mu g/m^3$ (33 ppb) and 150 $\mu g/m^3$ (50 ppb) for 24 h averages.

From time series analysis of 2 years of daily visits to primary health care clinics in Santiago, Chile, Ostro et al., (1999a) found significant association between PM10 and medical visits for lower respiratory symptoms in children and for upper respiratory symptoms in the older persons. A study conducted recently in Bangkok, Thailand indicated a statistically significant association between PM10 and mortality. A 10 $\mu g/m^3$

change in daily PM10 is associated with 1–2 percent increase in natural mortality, a 1–2 percent increase in cardiovascular mortality, and 3–6 percent increase in respiratory mortality (Ostro et al., 1999b). Particulate matter (PM) has been widely studied due to its potential health impact and a lot of effort was put on for its control. It indicates that finer PM has the strongest health effects (Schwartzet et al., 1996; Borja-Aburto et al., 1998).Dockery et. al., (1993) found that various health effects of PM, from less serious to very serious ones, are associated with its specific chemical and physical components.
Air pollution is causing a serious threat to public health in most of the urban environments in the developing countries. In industrial environment, the major sources of air pollution have been fossil fuel combustion and the manufacture and use of chemicals. Pollution from household emissions, industry and power generation has increased in the past few years, e.g. nitrogen oxides, though other pollutants such as smoke and sulphur dioxide have decreased. The contribution of vehicle emissions is becoming increasingly important.

Indoor air pollutant concentrations including nitrogen dioxide, carbon monoxide, formaldehyde and particulates are often higher than those outdoors, particularly at work There are relatively few studies of the potential adverse health effects of indoor exposure

Stevens (1997) carried out a continuous ambient automobile pollution monitoring programme in the Johannesburg region in and showed that automobile pollution levels which give rise to photochemical smog have slowly increased. Non-methane hydrocarbons in the city centre exceed the U.S. Environmental Protection Agency (USEPA) standard for more than 50 percent of the time. Downwind of the city centre, ozone hourly averages have exceeded the USEPA standard on 8 days during the 1984/85 season.

In Mexico City, such economic damages due to air pollution are estimated at $1.5 billion per year. In Jakarta, 14,000 deaths, about 2 percent of annual deaths, in the cities could be avoided every year if particulate could be kept at the level recommended by the WHO.
Bobak and Feachem (1999) studied air pollution and its impact in Czech Republic and Poland. Overall mortality, post-neonatal mortality and lung cancer due to air pollution

may cause up to 3 percent of all-cause, all-age mortality in the Czech Republic, which is some 9 percent of the mortality gap between this country and Western Europe. The United Kingdom National Air Quality Strategy published a detailed report regarding Air pollution, its causes and impact.

A description of the major contaminants of air is given below:

Suspended particulate matter (SPM) is a complex mixture of organic substances. PM_{10} and $PM_{2.5}$ are potentially harmful. Theses are directly emitted from sources such as non-nuclear power stations, motor vehicles, cement factories and open-cast coal mines. Airborne spores and pollen grains are natural sources of SPM in urban environment. Diesel powered vehicles, are the major source of particulates. A major portion of urban PM_{10} particulates are due to motor vehicles. Domestic coal burning is an important source in rural areas.

Sulphur dioxide (SO_2) is mainly generated by thermal power plants, by transportation activities, combustion of fossil fuels, industrial combustion etc. This is precursor of acid rain. Acid rain causes huge damage to vegetation, material as well as ecology in whole.

Ground level ozone (O_3) is a highly reactive oxidizing agent formed indirectly by the action of sunlight on nitrogen dioxide (NO_2). Peak pollution of Ozone occurs in summer when temperature is high.

Oxides of nitrogen mainly contain nitric oxide (NO) and nitrogen dioxide (NO_2). It is formed by spontaneous conversion of NO to NO_2. Transportation activities and power plants are the main source of this pollutant.

Carbon monoxide (CO) is mainly emitted by Transport vehicles. So in cities this is a concerned subject.

Air pollution has severe adverse effect not only on human being but also on ecological environment of a region (Okita et al., 1996; Pandey et al., 2002, 2004; Srivastava, 2004;

Srivastava et al., 2005). Sulfate aerosols, through the processes of scavenging and acid rain, lead to serious degradation of terrestrial and aquatic ecosystems.

The presence of Volatile Organic Carbons leads to photochemical oxidation causing increased smog episodes, ground level ozone concentrations and they are harmful to the ecosystem (Atkinson, 2000; Derwent, 1995; Kuran and Sojak, 1996; Dewulf and Langenhove, 1997). Many VOCs at higher concentrations have been reported to be toxic, carcinogenic or mutagenic (Edgerton et al., 1989; Duce et al., 1983; Sweet and Vermette, 1992; Kostianen, 1995; Mukund et al., 1996). Many studies have been carried out across the globe to assess the concentrations of hazardous air pollutants in the urban atmospheric environment (Singh et al., 1983; Hunt et al., 1998; Staten Island Report, 1997).

Air pollution is one of the major problems faced by Delhi. The vehicular pollution in Delhi has grown from 64 percent to 72 percent in the last decade and petroleum and diesel consumption grew by 400 percent and 300 percent respectively in the last two decades. The total vehicular pollution load in tones per day in Delhi is as high as 1046.30 compared to as low as 226.25 in Chennai and 293.71 in Kolkata, the study conducted by TERI said. Even Mumbai, which stands second, had recorded only 659.30 TPD vehicular pollution load, way below the national capital.

Our country accounts for nearly one third of the total 150 million asthma afflicted people worldwide. It ranks highest in terms of premature deaths due to outdoor and indoor air pollution. About 30 percent of Delhi's population suffers from respiratory disorders due to air pollution, and the incidence of respiratory diseases in the city is as high as 12 times the national average. Common disease symptoms are eye irritation, throat infection, respiratory discomfort, skin ailments, impaired hearing, chest disease, excessive carboxy haemoglobin and annoyance with noise (http://www.delhitrafficpolice.nic.in/ art8.htm).

In terms of SPM, Delhi is the fourth most polluted city in the world (http://www.gisdevelopment.net/application/natural_hazards/overview/nho0019.htm).

Due to rising population levels, their haphazard distribution and growth and the consequent rise in the levels of infrastructure needed to support them are the main reasons behind this pollution rise. Haphazard industrial development along with transportation activities are equally responsible (Srivastava, 2004). Transportation emissions are estimated to account for 60–70% of total air pollution) (CPCB, 2004. Diesel generating sets and vehicles are major sources which not only generate significant amount of air pollution, but also lead to highly uncomfortable levels of noise pollution. (Pandey et. al., 2001)

Pandey et al., (2005) studied the health risks of NO_2, SPM and SO_2 in Delhi. It was found that health risks due to air pollution in Delhi are highest for children. For all age categories, health risks due to SO_2 are the lowest. Hence, health risk due to SO_2 has been taken as the reference with respect to which HR values due to SPM and NO_2 have been compared. Taking into account all the age categories and their occupancy in different zones, average HR values for NO_2 and SPM turn out to be respectively 22.11 and 16.13 times more than that for SO_2. These pollutants are responsible for the increase in bronchial irritation and asthma. These can also cause irritation in the eyes, besides respiratory diseases.

Patade et al., (1984) observed as part of 3 year study of health morbidity in relation to air pollution around Bombay. There were no significant differences in the trends of each area over 3 years except in the `urban medium' area where with a decline in the pollution, the morbidity decreased. The `urban high' area showed the highest SO_2 and S.P.M. levels. The levels of NO_2 were highest in 1978 but in later 2 years these were similar to the `urban medium' area. The latter site had lower $SO_2$ but similar S.P.M. levels. The `urban low' area showed low SO_2 moderate NO_2 and high S.P.M. levels. The rural area had only high S.P.M. levels.

Among major air pollutants, Volatile organic compounds play a critical role in atmospheric processes. The toxic oxidants created by them are harmful to ecosystem, human health and urban environment. The variability of pollutants is an important factor

in determining human exposure to these chemicals. Very little work has been carried out on Volatile organic compounds especially in India's context. Srivastava et al., (2006) carried out sampling at 15 locations in Mumbai. These locations are in residential, industrial, commercial, traffic intersections and petrol refueling stations. It was concluded that the level of Volatile organic compounds are high as per guide lines set up by WHO.

Knowledge of ambient levels of Volatile organic compounds is necessary to evolve a proper strategy to control tropospheric ozone build up and maintain healthy air quality. Some work has been carried out in Indian context to know the levels of benzene, toluene, xylene and ethyl benzene at few locations (Chattopadhyay et al., 1997; Mohan Rao et al., 1997; Srivastava et al., 2000; Srivastava et al., 2004; Sahu et al., 2004). Srivastava, (2004) studied the source apportionment of ambient Volatile organic compounds in Mumbai

Mondal et al., (1967) measured ground-level concentration of NO_x at 19 important traffic intersection points within the city of Calcutta. Maximum average concentration of 222 μg m^{-3} was observed during winter and a minimum average concentration of 55 μg m^{-3} was observed during peak monsoon.

Naja et al., (2000) measured the surface ozone, CO, CH_4 and oxides of nitrogen at a high altitude site Mt Abu (24.6°N, 72.7°E, 1680 m asl), India for the period 1993–2000. Seasonal variation in ozone shows a pronounced maximum (monthly average about 46 ppbv) in the late autumn and winter. A correlation study between ozone and CO indicates possibility of incomplete photochemical processes over Asia. Annual average mixing ratios of CO, CH_4 and oxides of nitrogen were observed to be about 131.4±35.8 ppbv, 1.63±0.04 ppmv and 1.5±1.4 ppbv, respectively.

Negi et al., (2003) conducted a study of aerosols over four cities in India over a 1year period. It was concluded that the aerosol compositions show a seasonal variation. The levels of pollutant elements were higher during the summer and lower during the winter.

Pandey et al., (1967) reported the diurnal patterns in the concentrations of ozone (O_3), nitrogen dioxide (NO_2), sulphur dioxide (SO_2) and total suspended particulate matter (TSP) in the urban atmosphere of Varanasi city in India during 1989. Ambient concentrations of NO_2 and SO_2 were maximum during winter but ozone and TSP concentrations were highest during summer. The measured maximum concentrations (2-h average) were 150 and 231 $\mu g\ m^{-3}$ (0.078 and 0.086 ppm) for NO_2 and SO_2, respectively, for the winter season. Ozone and TSP concentrations reached a maximum of 160 (0.08 ppm) and 733 $\mu g\ m^{-3}$, respectively, in the summer. NO_2 and SO_2 concentrations peaked in the morning and evening. Peak concentrations of O_3 occurred in the afternoon, generally between noon and 4 p.m.

Singh et al., (1967) carried out measurements of ozone (O_3) and oxides of nitrogen (NO_x) at six sites in Delhi during the winter months of 1993. Concentrations of ozone and NO_x varied between 34 and 126 ppbv and 32 and 272 ppbv, respectively. The paper reported some unusual winter diurnal ozone trends.

Varshney and Aggarwal (1991) measured ozone in the urban environment of Delhi at four different sites during 1989–1990. The amount of ozone in the ambient air varied from 9.4 to 128.31 ppbv exhibiting wide temporal and seasonal variation. The ozone concentration invariably peaked at noontime and remained high during early summer and spring periods. On many occasions 1-h ozone concentration was more than 113 ppbv, which represents the maximum 1-h limit of ozone in ambient air as prescribed by the USEPA.

The air quality standards are developed on the dose-effect/ dose-response relationships. The standards are an integral part of air quality management, which are required to set long-term as well as short-term goals for air quality improvement and formulation of strategies and implementation of various programs for achieving these goals.

CPCB has notified the National Ambient Air Quality Standards (NAAQS) for various Pollutants. These standards are reproduced in Table 1.1

Table 1.1: *National Ambient Air Quality Standards setup by CPCB, Delhi*

		Industrial Areas	Residential, Rural & other Areas	Sensitive Areas	
SulphurDioxide (SO_2)	Annual Average*	80 µg/m3	60 µg/m3	15 µg/m3	- Improved West and Geake Method - Ultraviolet Fluorescence
	24 hours**	120 µg/m3	80 µg/m3	30 µg/m3	
Oxides of Nitrogen as (NO_2)	Annual Average*	80 µg/m3	60 µg/m3	15 µg/m3	- Jacob & Hochheiser Modified (Na-Arsenite) Method
	24 hours**	120 µg/m3	80 µg/m3	30 µg/m3	- Gas Phase Chemiluminescence
Suspended Particulate Matter (SPM)	Annual Average*	360 µg/m3	140 µg/m3	70 µg/m3	- High Volume Sampling, (Average flow rate not less than 1.1 m3/minute).
	24 hours**	500 µg/m3	200 µg/m3	100 µg/m3	
RespirableParticulate Matter (RPM) (size less than 10 microns)	Annual Average*	120 µg/m3	60 µg/m3	50 µg/m3	- Respirable particulate matter sampler
	24 hours**	150 µg/m3	100 µg/m3	75 µg/m3	
Lead (Pb)	Annual Average*	1.0 µg/m3	0.75 µg/m3	0.50 µg/m3	- ASS Method after sampling using EPM 2000 or equivalent Filter paper
	24 hours**	1.5 µg/m3	1.00 µg/m3	0.75 µg/m3	.
Ammonia1	Annual Average*	0.1 mg/ m3	0.1 mg/ m3	0.1 mg/m3	.
	24 hours**	0.4 mg/ m3	0.4 mg/m3	0.4 mg/m3	.
CarbonMonoxide (CO)	8 hours**	5.0 mg/m3	2.0 mg/m3	1.0 mg/ m3	- Non Dispersive Infra Red (NDIR)
	1 hour	10.0 mg/m3	4.0 mg/m3	2.0 mg/m3	

[1.1] Modeling Studies of various Air Pollutants

As discussed earlier the adverse consequences of air pollution, it becomes inevitable that some approaches needed to avoid or at least minimize the harmful impact on living being as well as on non-living due to air pollution. Air quality management and protection require knowledge of the state of the environment encompassing both cognitive and interpretative elements. Efficient air pollution management must dispose of interpretative tools which are able to extrapolate in space and time the values measured.

So in order that the general public can be forewarned about the air quality in an urban area, it becomes desirable to predict pollutant concentration in advance to enable the public to avoid exposure to polluted environment.

Air pollution models have not so far been able to forecast satisfactory ground level concentrations as the influence of important variable is not perfectly described. The deterministic models can not provide an adequate correlation between predictions and observed data paired in time and space.

The literature provides several approaches to the modeling of atmospheric pollution time series data. The auto regressive integrated moving average (ARIMA) model has been used successfully in time series forecasting in many applications. (Ho and Xie, 1998).In this model a linear correlation structure is assumed among the various variables. This model fails miserably as it can not capture non-linear pattern. In particular, the parametric identification methods based on stochastic models of the ARMAX (Auto Regressive Moving Average with exogenous inputs) type are the most popular. The general structure of ARMAX models is of the following type:

$$A(q)\,y(t) = B(q)\,u(t) + C(q)\,e(t)$$

Where $y\ (t)$ are the outputs, $u\ (t)$ the inputs, $e\ (t)$ is a white noise and $A\ (q)$, $B\ (q)$, $C\ (q)$ are the parameter vectors. An exhaustive description of ARMAX model is given by Ljung et al., (1987). However on account of the complexity of the phenomenon involved, prediction based on ARMAX models at times may be unreliable.

According to Van der Wal and Janssen (2000), 45% of the variance of PM10 concentrations may be due to changes in wind direction, temperature and duration of precipitations. Kukkonen et al., (2001) have been able to predict long-term average PM10 concentrations on the basis of a linear relation between PM10 and NOx concentrations, but this model is not accurate for predicting short-term (hourly) average concentrations. This model requires an independent forecast of NOx concentrations.

Modeling atmospheric pollution phenomena is a complex task on account of the influence of a number of external variables (e.g. temperature, humidity, wind speed and direction, etc.), and the presence of time-variance and non-linearity effects which make the traditional identification techniques not particularly appropriate.

Across world-wide experts are introducing new advanced modeling techniques based on the results of recent research on the meteorological parameters. These modeling techniques contain algorithms for calculating the main factors that determine air pollution diffusion in terms of the fundamental parameters of the atmospheric system.

Tiziano Tirabassi (1995) described various types of model developed at the Istituto per lo Studio dei Fenomeni Fisici e Chimici della Bassa e Alta Atmosfera (FISBAT-CNR).Most of the estimates of dispersion of air pollution from continuous point sources are based on the Gaussian approach. Due to the presence of the ground, turbulence is usually not homogeneous in the vertical direction.

Tirabassi developed several model codes employing a non-Gaussian analytical solution of the advection-diffusion equation at FISBAT-CNR, Bologna. : KAPPAG model uses an analytical solution based on the vertical profiles of wind and diffusion coefficients .This is capable of handling multiple sources and multiple receptors in simulating time-varying conditions. KAPPAG-LT was the climatological version of KAPPAG .CISP model elaborates a method for estimating maximum ground level concentrations from a single point VIM model estimates maximum ground-level concentrations in an area with many

emission sources. MAOC provided satisfactorily result for the pollutant concentrations in complex orography.

Karim and Ohno (2000) studied the state of air pollution in Nagoya. In particular, suspended particulate matter (SPM), nitrogen oxides (NOx) emitted by automobiles, and photochemical oxidants especially ozone (O_3), had become major pollution hazardous. They developed an empirical model for evaluating urban SPM concentration in problem areas. The model presented found to be suitable for predicting long-term average SPM concentrations and could be utilized for analyzing the effects of various traffic emission reduction strategies. This model could be utilized for new urban development as an alternative.

Several time series studies have investigated the association between daily adverse health outcomes and daily ambient air pollution concentrations (Chock et al., 2000; Cifuentes et al., 2000; Goldberg et al., 2003; Kelsall et al., 1997, 2000; Kwon et al., 2001; Moolgavkar, 2000; Ostro et al., 1999; Roemer and van Wijnen, 2001; Smith et al., 2000a, b; Stieb et al., 2002; Styer et al., 1995). These studies typically fit a Poisson log–linear model to concurrent time series of daily mortality or morbidity, ambient air pollution and meteorological covariates. Then the fitted models are used to quantify the adverse health effects of ambient air pollution.

Due to the high correlation between ambient air pollutants, the results from studies that focus on a single pollutant can be difficult to interpret in practice (Vedal et al., 2003) (Dominici and Burnett, 2003)

Deterministic models, stochastic models, and artificial intelligent derived approaches have been used for modeling processes. Deterministic models reproduce physical processes, but satisfactory results could not be obtained using them in most of the cases. It is difficult to use deterministic models in areas with high industrial concentration and tormented orography because of the uncertainty inherent to the emission rate of pollutants and the intrinsic structure of mathematical representation (advection– diffusion equation).

The stochastic models aim to forecast future values of a time series starting from current and past values but they are not adequate for concentration forecasting tasks when thermal inversion height is taking place.

McCollister and Wilson (1967) worked on to develop two related time series models to forecast concentrations of various air pollutants. The developed model had been tested on carbon monoxide and oxidant data for the Los Angeles basin. Predicted result for 1972 of the daily instantaneous maxima for total oxidant made using only past pollutant concentration data were found to be more accurate than those made by the Los Angeles APCD using meteorological input as well as pollutant concentrations.

Robeson and Steyn (1967) developed three statistical models that estimate daily maximum ozone (O_3) concentrations in the lower Fraser Valley of British Columbia (BC) are specified using measured concentrations from two monitoring stations during the time period 1978–1985. The three models are (1) a univariate deterministic/stochastic model, (2) a univariate autoregressive integrated moving average (ARIMA) model, and (3) a bivariate temperature and persistence based regression model. The bivariate model is superior to both univariate models and ARIMA model.

Ziomas et al., (2002) developed analytical models relating maximum pollutant concentrations in urban areas with meteorological and other variables. The analysis was based on measurements from Greater Athens Area and is of $N0_2$. The meteorological variables, used in analytical modeling for forecasting pollution concentrations, cover the most important atmospheric processes favoring pollution episodes. The evaluation of the developed forecasting models showed that their degree of success is promising.

Saldiva et al., (1986) developed a time series model for examining the relationship between air pollution and respiratory mortality in children under 5 years of age in the metropolitan area of São Paulo, Brazil. Daily records of mortality (excluding neonatal mortality) for the period May 1990 to April 1991 were collected along with daily records of relative humidity, temperature, SO_2, CO, particulates (PM10), O_3, and NO_x

concentrations. Multiple regression methods demonstrated a significant association between mortality due to respiratory diseases and the NO_x levels.

Chameides and Walker (1976) presented a time-dependent photochemical model of tropospheric ozone. Ozone as an active chemical constituent with a local abundance that is determined by photochemical processes rather than by transport processes. The local abundance of odd nitrogen is shown to play a major role in determining the intensity of photochemical production and destruction of ozone, a large local abundance of NO_x (= $NO + NO_2$) leading to large diurnal ozone variations. The photochemical model is shown to reproduce the diurnal variations of ozone and nitrogen dioxide observed near the ground at Research Triangle Park in North Carolina.

Ku and Rao (1994) developed a three-dimensional, grid-based numerical air pollution model for the estimation of air pollutant concentrations in an urban area. The modeling system incorporates the combined influences of heavy transport, turbulent diffusion, chemical transformation, source emissions and surface removal of air contaminants. Recent developments in plume rise and plume penetration processes, objective wind field analysis procedures and numerical solution techniques incorporated into the model are well described.

Computation through neural networks is one of the recently growing areas of artificial intelligence. Neural networks are promising due to their ability to learn highly nonlinear relationships. Two kinds of neural networks were used in the paper titled "The application of neural techniques to the modeling of time series of atmospheric pollution"—multilayer perceptrons and fuzzy neural networks (FNNs). Multilayer perceptrons were used in this study (Narendra et al., 1991).

The concentration levels forecast by a supervised neural net model had taken account of the influence of the variables of the system, such as source emission factors, conditions, local topology, reactions rate, by using an appropriate training on the available experimental data (Faussett, 994).

Gonzalez-Manteiga et al., (1993) have used a semi parametric model to predict SO_2 concentrations 30 min in advance in the neighborhood of a power station in Spain. The inputs to their model are present values and values measured 5 min back of SO_2 concentrations and they report a forecast error of 40%. Boznar et al., (1993) have used a three-layer neural network to predict also atmospheric SO_2 concentrations 30 min in advance in a highly polluted industrialized area of Slovenia. The neural network has 37 inputs, which includes past values of SO_2 concentrations measured at different stations in the study area and meteorological variables as temperature, wind speed and wind direction. Their predictions seem qualitatively good but only graphical representations and not statistics are given. Andretta et al., (2000) have also used a three-layer neural network for forecasting SO_2 ground level concentrations 1 h in advance in the industrial area of Ravenna, Italy. Their network has 10 inputs including past values of sulfur dioxide concentrations, wind speed, wind direction and solar radiation. The output here is binary: 1, if concentration is higher than 40 $\mu g/m^3$ and 0 if it is lower than this value. They report a 16% mean error for 1 h forecasting.

Perez and Trier (2001) have used a neural network to predict atmospheric NO_2 concentrations near a street with heavy traffic in Santiago, Chile. In this case, inputs are NO concentrations on previous hours plus forecasted temperature at the time of the intended prediction. Results show 5 h in advance prediction error of the order of 25% and a strong dependence on the temperature value.

Tirabassi and Rizza (1995) applied the neural filter to a puff model (SPM: Skewed Puff Model). The model utilizes approximate solutions proposed for the dispersion of a cloud of passive contaminants released from an instantaneous near the ground.

Gardner and Dorling (1998) have described the usefulness of a multilayer neural network (a nonlinear model) for applications in atmospheric sciences. Perez et al., (2000) have shown that a three-layer neural network is a useful tool to predict hourly averages of PM2.5 concentrations in the atmosphere of downtown Santiago, Chile, several hours in

advance when hourly concentrations of the previous day and forecasted temperature, relative humidity and wind speed are used as input. Predictions generated with this network are more accurate than those produced with a linear regression that has the same inputs. Silva et al., (2001) have used a MARS algorithm (which is nonlinear) to forecast PM10 concentrations in Santiago, Chile, and they report an accuracy that appears greater than what can be obtained with multilayer neural networks and linear regressions.

Patricio Perez et al., (2002) reported the possibility to predict the maximum of 24 Moving Average of PM10 concentrations on the coming day using multilayer neural networks based on data obtained in a station located at a fixed point to the west of downtown Santiago. The selected station was one of the eight stations that form the network designed to monitor air quality in the city. Inputs to the neural network were hourly averages of PM10 concentrations up to 6 p.m. of the present day and forecasted values of meteorological variables for the coming day. Assuming that maximum and minimum temperature, average wind speed and minimum relative humidity can be forecasted with precision. The neural network algorithm predicts whether the next day the level 240 $\mu g/m^3$ will be exceeded with an accuracy that appears to be slightly greater than that obtained with a linear model that has the same inputs.

Artificial neural networks (ANNs), now a days have become a useful tool for modeling environmental systems. There are various applications of ANNs in many fields. The model based on this technique has been successfully applied to simulate the export of nutrients from river basins (Clalr and Ehrman, (1996)), to forecast salinity (Desileta et. al., 1992) and ozone levels (Comrie, 1997), to predict air pollution (Boznar et al., 1993) and the functional characteristics of ecosystems (Paruelo and Tomasel, 1997), and to model algal growth, and transport in rivers (Whitehead, 1997). Many environmental modelers are trying to modeling with ANNs on datasets for which the use of more conventional techniques (e.g., regression) has been unsuccessful. However, it's a tedious task to use ANNs.

They are generally treated as black box models that are able to capture underlying relationships when presented with input and output data. On many occasion, little consideration is given to potential input data and the mechanism of its working.

The main reason of failure of these models is that they failed to capture non-linear relationships among various parameters. Also it is very difficult to model some complex situations.

An artificial neural network (ANN) is able to treat uncertain and incomplete information like the human brain. The highly non-linear processes of pollutant concentration and its dynamics are only partially known. Only complex computer modeling and simulation can provide a reliable prediction. ANNs seems to be a good approach for modeling pollutant concentration as it has got the capability of pattern recognition, adaptability and hence function approximation. It can also tolerate certain degree of noise level.

ANNs are used for static problems where outputs depend only on current inputs. The memory of conventional ANNs, called long-term memory, is in network connection weights (or any other trainable parameters) that are updated during the training phase. Once an ANN has been successfully trained, the long-term memory remains fixed for the whole testing phase. Using the taxonomy proposed in the spazio temporal connectionist networks (STCNs) also include a short-term memory that allows these networks to deal better with time series. Conventional connectionist networks compute the activation values of all nodes at time t based only on inputs at time t. By contrast, in STCNs the activations of some nodes at time t are computed based on activations at time t-1, or earlier. These activations serve as a short-term memory. Long-term and short-term neural networks have been widely used for forecasting tasks.

In many applications, neural networks have become an alternative method to traditional approaches like deterministic and stochastic methods (McCollister and Wilson KR, 1974). Recent studies show that neural based models could be used for developing air pollution forecasting models (Shiva Nagendra and Khare, 2005). Viotti et al., (2002) used a MLP (multi layer perceptron) to forecast short and middle long-term concentrations

levels for various pollutants like O_3, NOx, NO2, and CO. etc. In Belgian urban environment, Hooyberghs et al., (2005) developed the MLP neural network to forecast the daily average PM10 concentrations one day in advance. Zhang and San used a wavelet neural network to model hourly NOx and NO2 concentrations. Some works were carried out to develop forecasting model for weather and temperature.

Tasadduq et al., (2002) used an MLP neural network trained with the Back Propagation algorithm and a batch learning scheme for the prediction of hourly mean values of ambient temperature 24 hour ahead in Jeddah, Saudi Arabia. Khotanzad et al., (1996) developed a three layer perceptron trained with the Back Propagation algorithm forecasting of a daily temperature for up to seven days in advance. Maqsood et al., (2002) testified a neural network based Hopfield model, MLP and Elman networks for hourly weather forecasting in a Canadian region. Model performance was compared with a standard MLP neural network, with an Elman recurrent network and with a radial basis function neural network. In this case, the MLP and Elman networks have been found as the best forecasters.

ANN was also successfully applied in forecasting meteorological parameters. Luk et al., (2000) developed three alternative ANN models for rainfall forecasting: 1. A multi-layer feed-forward neural network, 2. An Elman neural network and 3. A time delay neural network (TDNN).

The study shows that the three approaches have comparable performance. Kostela et al., (1996) compared the performance of the MLP, FIR and Elman neural networks in four different time series analysis: the performance of the Elman neural networks was better or similar to the other neural networks. It was found that the efficiency of the learning algorithm is a more important factor than the neural model used. For developing ANN model many criteria are involved such as time taken to train the model. This in other way depends on the type of algorithm used for training so that global minima on the error surface get reached in the minimum possible time. Annan and Bishop compared the speed of convergence of six popular training algorithms over two real problems.

Ganguly et al., (1995) used ANN for estimation of petroleum and product properties. The algorithm used in this case was back propagation.

It is clear from the above literature review that there is no universality of approach in forecasting models used for predicting concentration of air pollutants. Although simple linear statistical models have been used, it is the neural net based model, which has proved to be relatively more accurate in prediction of concentration of air pollutant.

The problems associated with neural net approach however pertain to

(i) The choice of number of input variables,
(ii) Efficient training of the neural network and
(iii) Minimization of prediction error.

In order to make more reliable forecast an efficient and well trained neural network requires more judicious choice of input variables by restricting the number of input variables. Thus a neural network based on various approaches has been applied. The present work has been carried out with the following objectives.

[1.2] Objectives

1. To develop a forecasting model for air pollutants NO_2 and NO in Delhi using Artificial Neural Network based on feed forward back propagation approach.
2. To develop a forecasting model for air pollutants NO_2 and NO in Delhi using Artificial Neural Network based on Radial basis function network.
3. To develop a forecasting model for air pollutants NO_2 and NO in Delhi using Artificial Neural Network based on RBF/Principal Component Analysis.
4. Performance evaluation of forecasting models based on the above mentioned approaches.

The thesis is presented in the following manner. The chapter 1 begins with introductory details about modeling study of air pollution. Chapter 2 describes the study area and the

sources of data. Chapter 3 contains theoretical background of Artificial Neural Network and modeling using feed forward back propagation approaches. Chapter 4 describes the modeling using Artificial Neural Network based on Radial Basis Function. Chapter 5 provides work related to Radial Basis Function and Principal component analysis. Chapter 6 presents a comparison of the performance of the developed models followed by conclusion.

CHAPTER – 2

Chapter 2 – Study Area and Sources of data

[2.1] Study Area

Delhi, the capital city of India, is one of the 10 most polluted cities of the world and the third most populated city in India with 13.8 million inhabitants spread over 1483 km^2 (Aneja et al (2001).According to the census report of Government of National Capital Territory of Delhi 1999, the population of Delhi increased from 6.2 million in 1981 to 9.4 million in 1991. Besides, there are 126 000 industrial (small- and medium-scale units) in Delhi (Government of National Capital Territory of Delhi 1999) out of which 98 000 units are categorized as non-conforming by the Master Plan of Delhi (Ministry of Environment and Forests 1997). The domestic sector too contributes greatly to air pollution. Invariably, the transport, domestic, and industrial sectors are believed to be the main contributors towards the rise of ambient air pollution levels. The transportation network in Delhi is predominantly road based with 1284 km of road per 100 km^2. Its urban area has quadrupled from 182 km2 in 1970s to more than 750 km2 in 1999 with the number of industries from 26,000 in 1971 to 137,000 in 1997 (Nils Olof Nylund and Allex (2000). Air pollution in the city is largely due to vehicles. Delhi now has more vehicles than Mumbai, Chennai and Calcutta put together. Delhi had six lakhs vehicles in 1982 that became 20 lakhs in 1992, accounting for about 60% of total pollutants being generated in Delhi. Motor vehicles registration rose by 51 times between 1961 and 1991, whereas, the population increased 4 times over the same period. The vehicles are expected to touch 26 millions in 2011

Vehicles pollution account for about 60% of total air pollutants. Among the industries, thermal power plants are the most prominent contributors to air pollution. (Goyal and Sidhartha, 2003).

Climate and natural sources also contribute to pollution in Delhi. The region has a semi-arid climate with hot summers, heavy rainfalls in the monsoon months, and very cold winters. The summer months have dust storms, while winters have inversions for longer

durations. During only the monsoon season the pollution levels due to frequent washout of pollutants by rain is minimal.

The Central Pollution Control Board (CPCB) has established a national Ambient Air Quality Monitoring (NAMP) network, comprising 295 stations in 92 cities, under the Air (Prevention and Control of Pollution) Act, 1981 to collect, compile and disseminate information on air quality data.

Our study pertains to two sites in Delhi, namely

(i) **Siri Fort**, a residential site

(ii) **ITO**, a heavy traffic zone (sensitive).

The map of study area is shown in the Figure 2.1. The above mentioned sites are shown in the adjoining map of Delhi. The data at the above mentioned sites were collected on hourly average basis. Air pollutants (Nitric oxide (NO), Sulphur dioxide (SO_2), Ozone (O_3), Nitrogen dioxide (NO_2) and carbon monoxides (CO)) were collected at Siri Fort Site for winter (November to February) of 2006 to 2008. Similarly for ITO site Air pollutants (Nitric oxide (NO), Sulphur dioxide (SO_2), Ozone (O_3), Nitrogen dioxide (NO_2), SPM (2.5) and carbon monoxides (CO)) were collected for the same period.

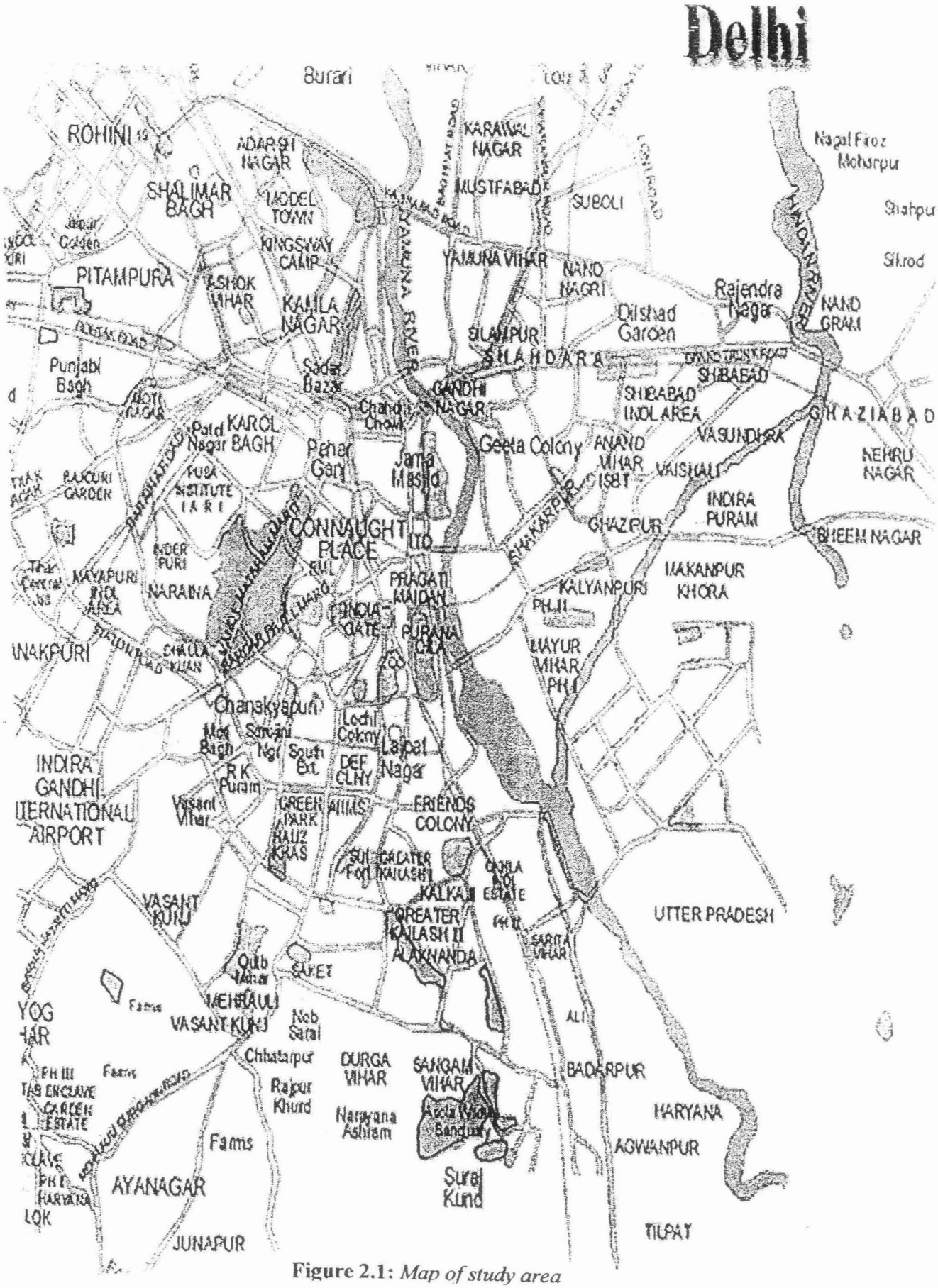
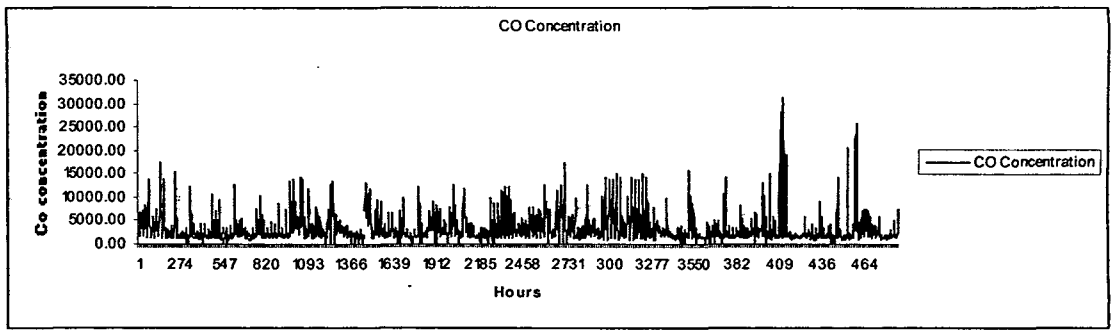
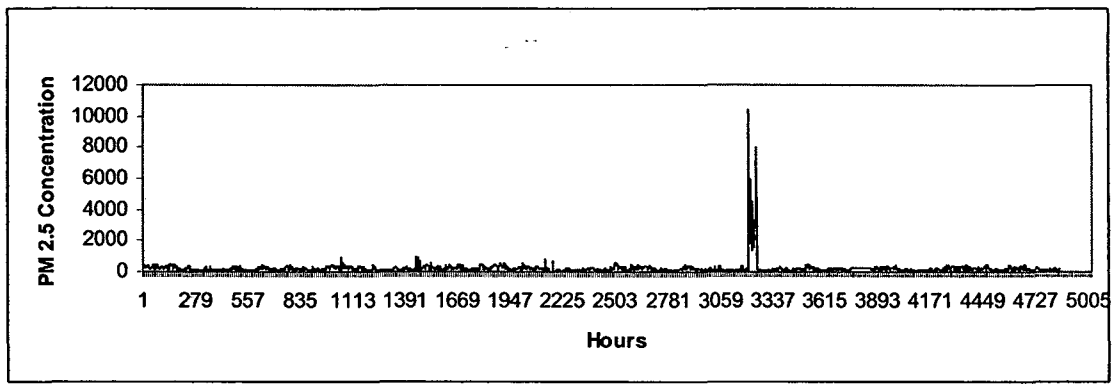
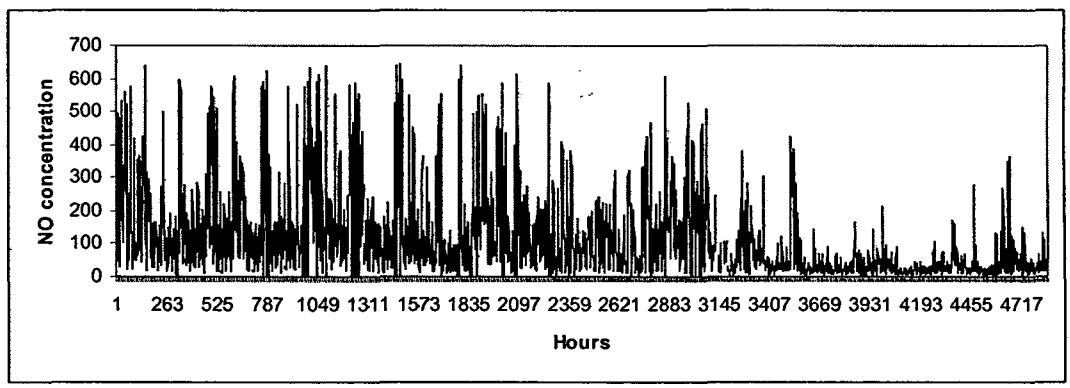


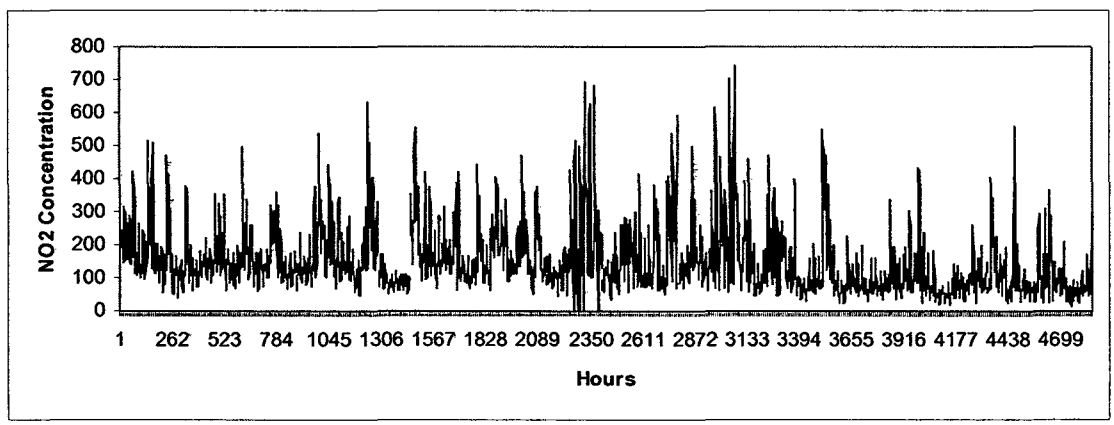
Figure 2.1: *Map of study area*

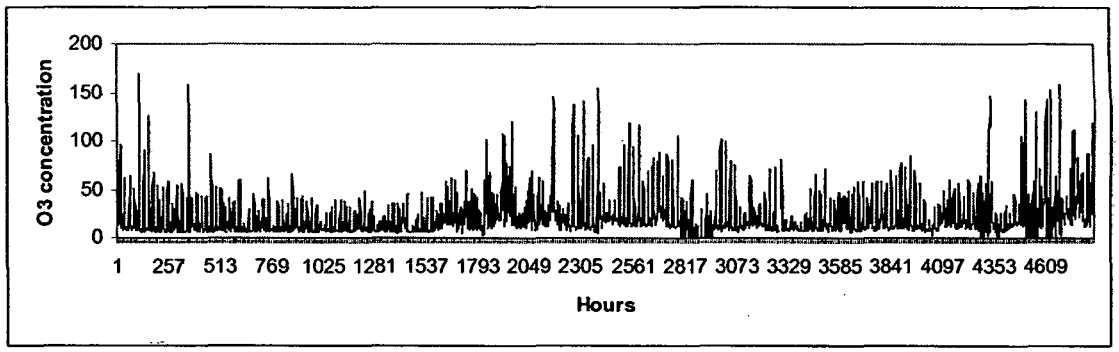
Figure 2.2 - Hourly time series of the data of various Air pollutants at ITO site. Concentrations of the pollutants are in (μgm^{-3})











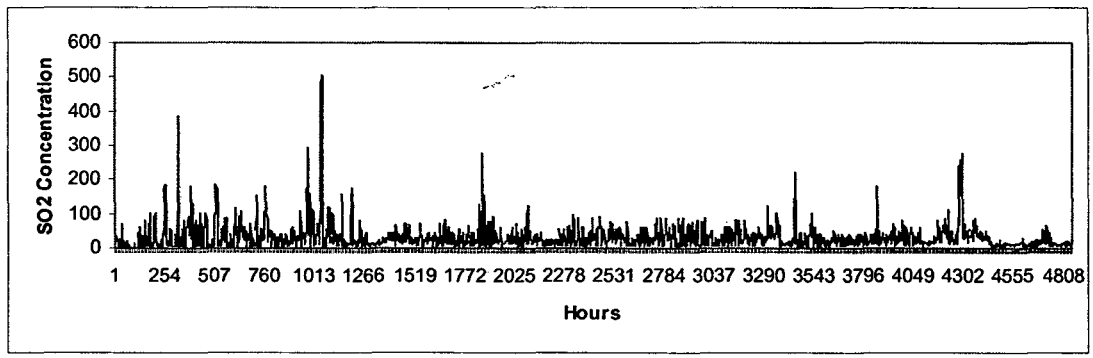
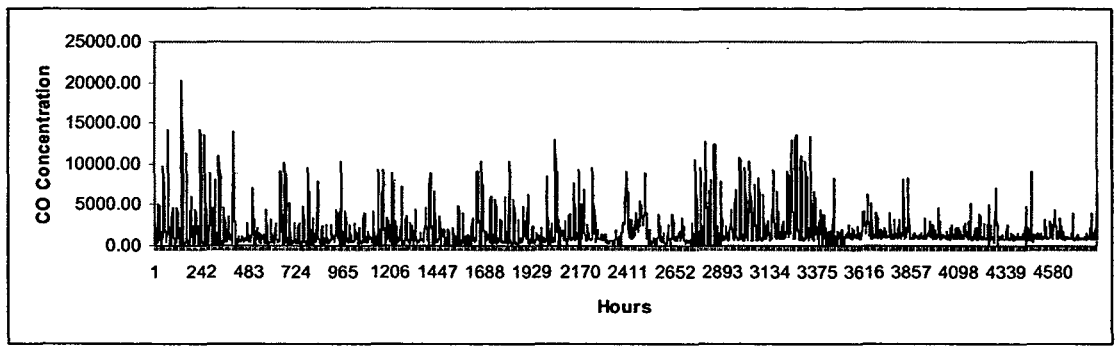
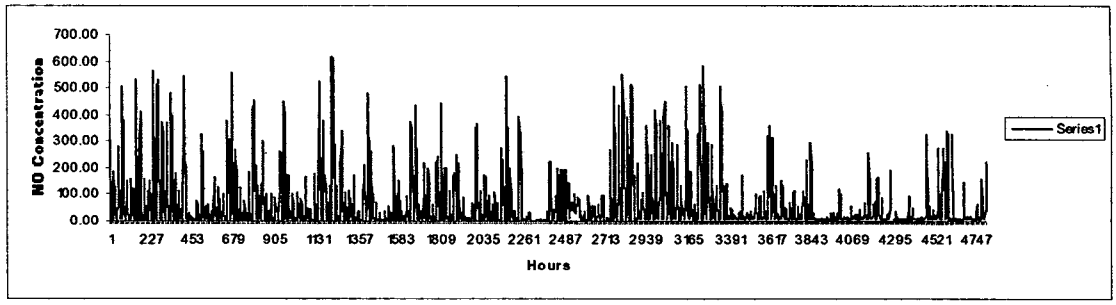
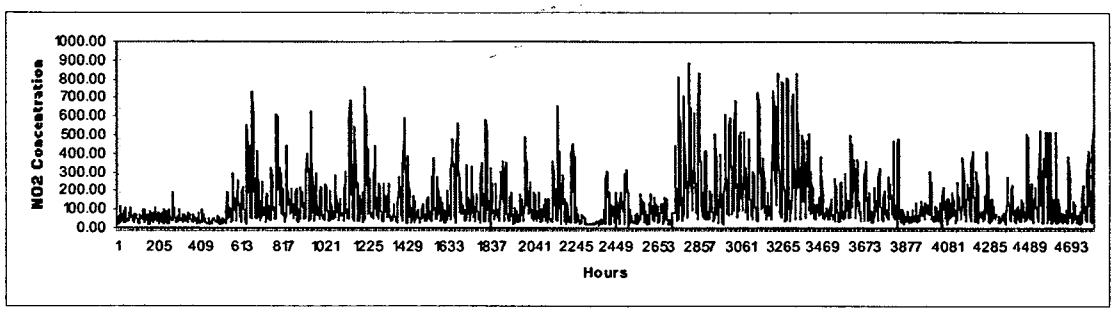
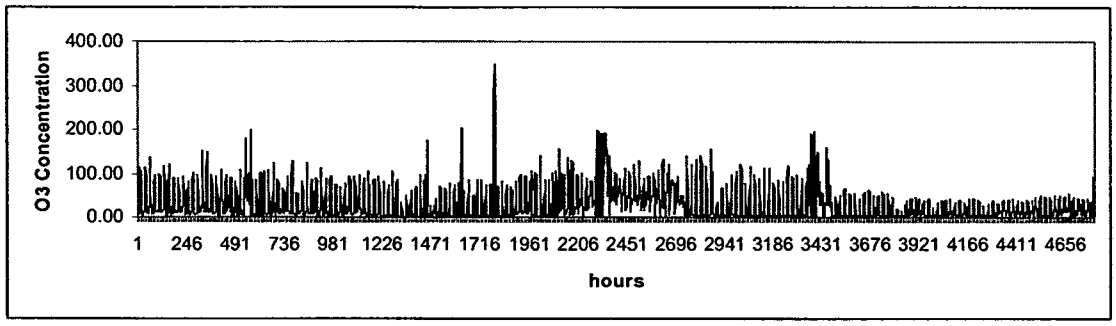


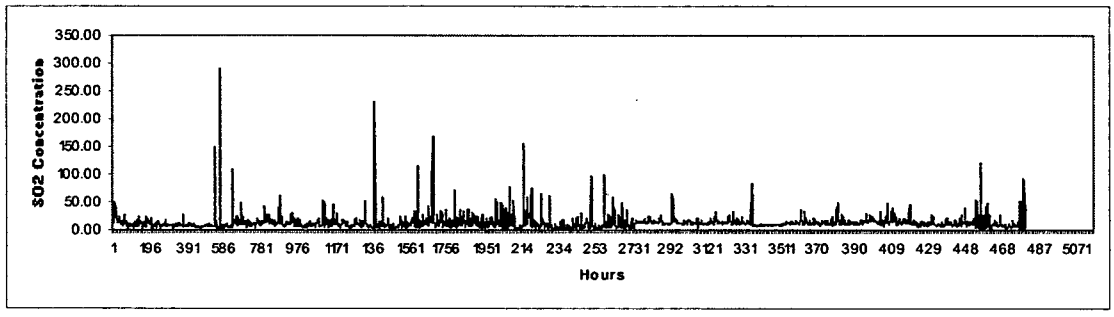
Figure 2.3 Hourly time series of the data of various Air pollutants at Siri Fort site. Concentrations of the pollutants are in (μgm^{-3})




NO Concentration
700.00
600.00
500.00
400.00
300.00
200.00
100.00
0.00
1 227 453 679 905 1131 1357 1583 1809 2035 2261 2487 2713 2939 3165 3391 3617 3843 4069 4295 4521 4747
Hours
Series1


NO2 Concentration
1000.00
900.00
800.00
700.00
600.00
500.00
400.00
300.00
200.00
100.00
0.00
1 205 409 613 817 1021 1225 1429 1633 1837 2041 2245 2449 2653 2857 3061 3265 3469 3673 3877 4081 4285 4489 4693
Hours


O3 Concentration
400.00
300.00
200.00
100.00
0.00
1 246 491 736 981 1226 1471 1716 1961 2206 2451 2696 2941 3186 3431 3676 3921 4166 4411 4656
hours


SO2 Concentration
350.00
300.00
250.00
200.00
150.00
100.00
50.00
0.00
1 196 391 586 781 976 1171 136 1561 1756 1951 214 234 253 2731 292 3121 331 3511 370 390 409 429 448 468 487 5071
Hours

CHAPTER –3

Chapter 3: Artificial Neural Network with Feedforward Back Propagation Approach

[3.1] Introduction

The question arises why there is need to use neural networks. There is need of techniques to be developed so that meaningful work can be done on complicated or imprecise data. Here meaningful work means that some pattern or trend can be obtained that are too complex to be logical either by human being or digital computers. Human brain performs work of pattern recognition, perception and other works in daily life many times faster than the fastest computer system. This led to development of the neural network based on human brain structure. After training a neural network becomes expert and can provide prediction on unseen situation. Other benefits are its capability of adaptive learning, self-organization, real time operation, fault tolerance, etc.

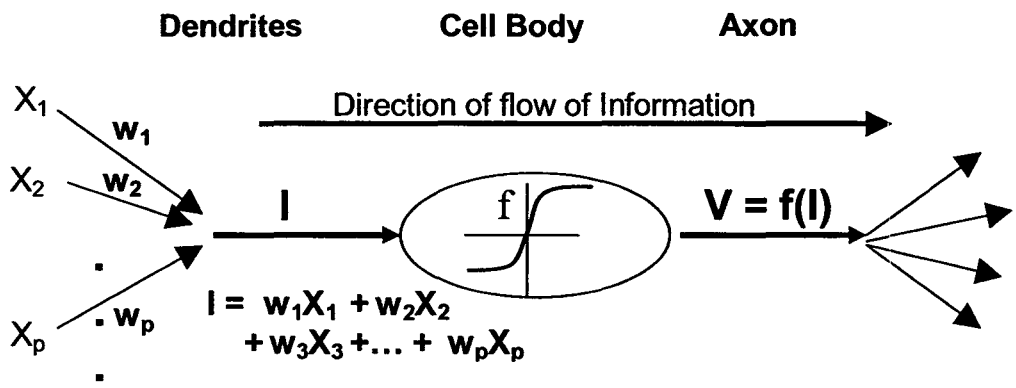
[3.2] Theoretical Background

A *neural network* is a massively parallel distributed processor that has a natural capacity for storing experiential knowledge and making it available for use whenever required. Its structure and functioning are like that of the brain. (Haykin,1998). First Knowledge is acquired by the network through a learning process. Connection weights are used to store the knowledge. Learning is a process by which the free parameters of a neural network are adapted through a continuing process of stimulation by the environment in which the network is to be developed. The type of learning is determined by the manner in which the parameter changes take place.

[3.2.1] Models of a Neuron

This is the simple model of a neuron. This shows how functioning resembles with that of human brain. $X_1, X_2, \ldots X_p$ are the inputs (source),

An Artificial Neuron



- Receives Inputs X_1 X_2 ... X_p from other neurons or environment
- Inputs fed-in through connections with 'weights'
- Total Input = Weighted sum of inputs from all sources
- Transfer function (Activation function) converts the input to output
- Output goes to other neurons or environment

Single layer perceptron – This is the most simple network structure in which activation function directly converts input to output.

Multi layer perceptron – The basic multi layer perceptron network is constructed by putting the units in layers so that each unit in a layer take as input only the outputs of units in the previous layer. If there are such two layers then it is said to be two layer network. The input signal proceeds through the network in a forward direction on a layer by layer basis. These are multi layer perceptrons (MLPs). A three layer network will have two hidden layers. This depends on the number of hidden layers used in the network.

[3.2.2] Types of network connections

Neurons are connected by a network of input to another neuron. These paths are unidirectional. A neuron may receive input from many neurons but produce a single

output. The neuron in a layer many communicate with each other or they may not have any connection.

[3.2.3] Learning algorithm

Learning algorithm is defined as a prescribed set of well-defined rules for the solution of a learning problem. Learning algorithm is developed as per requirement of the problem. There are various learning algorithm in use. Some of them are listed below

(1) First order error back propagation algorithm.

(2) Clustering algorithm.

(3) Global optimization techniques.

(4) Univariate random optimization.

[3.2.4] Activation function

The activation functions form a basic part of the neuron model. It limits the output of a neuron within specified range. This way the desired output was obtained with minimized error. The normalized amplitude range of the output of the neuron is in the interval [0, 1] for sigmoid function. There are various types of activation function. Some of them are listed below as –

1. **Unipolar sigmoid transfer function** – It is the most popular transfer function among neural network users. It is also known as "LOGISTIC" function. Large positive values asymptotically approach 1. While large negative values are squashed to 0. The sigmoid function is given by:

$$T(f) = \frac{1}{1 + \exp(-\lambda * f)}$$

The sigmoid limits the output in the interval [0, 1].

2. **Bipolar sigmoid transfer function** – The bipolar sigmoid transfer function is given by:

$$T(f) = -1 + \frac{2}{1 + \exp(-\lambda * f)}$$

3. **Hyperbolic tangent transfer function** – The hyperbolic tangent transfer function is given by

$$T(f) = \frac{(\exp(f) - \exp(-f))}{(\exp(f) + \exp(-f))}$$

It has a similar sigmoid shape to the logistic function, but the values are spread through the interval [-1, 1].

4. **Gaussian transfer function** – The Gaussian transfer function is given by

$$T(f) = \exp(-f^2)$$

The Gaussian is symmetric around the origin, where it takes its maximum value of 1, falling off as the distance of '*f*' from the origin.

5. **Linear transfer function** – The linear transfer function is simply

$$T(f) = f$$

This function passed on the value returned by the input function as is. This function is normally only be used in the output layer of the network.

6. **Bipolar logarithmic transfer function** – the bipolar logarithmic transfer function is given by

$$T(f) = \begin{cases} -\ln(1-f), & \text{if } f < 0 \\ \ln(1+f); & \text{if } f > 0 \\ 0; & \text{else} \end{cases}$$

[3.2.5] Learning process

Learning is a process by which the free parameters of a neural network are adopted through a process of stimulation by the environment. This takes place in the following sequence. First, neural network gets stimulated by the environment, then undergoes changes as a result of this simulation and then responds in a new way to the environment. There are three types of learning process.

1. Supervised learning
2. Reinforcement learning
3. Self organized or unsupervised learning

In supervised learning both the inputs and the outputs are adjusted as per error signal. While in unsupervised learning, the inputs are supplied to the network but outputs are not provided.

[3.2.6] Rate of learning

The back propagation algorithm uses method of steepest descent while training, the rate of learning η, plays a very important role. If we keep learning rate η to be very small then in the network changes in the weight will be smaller and hence trajectory of the weight space will be smooth. But this will reduce the speed of faster convergence. At the same time if we increase the value of learning parameter then process becomes faster but at the same time network becomes oscillatory. So optimum value of learning rate has to be provided to the network.

[3.2.7] Stopping criteria

Once training was initiated we keep on observing graphs between Room Mean Square Error (RMSE) and number of iteration. When this error gets minimized then training stops. This way we keep on repeating the procedure for various samples of training data. The model so obtained was applied for validation.

[3.2.8] Learning strategies

There are various learning strategies associated with neural network. These are namely,

1. Pattern association
2. Pattern recognition
3. Function approximation
4. Control
5. Filtering and Beamforming

[3.3] Development of the Model

As discussed, ANN basically consists of inputs source (data) transfer function and outputs. The important thing is how to transform the available data and then the number of inputs to be selected for training purpose. The various parameters for training the model has to be made. This is the most difficult part of the network. Many authors have worked in this area.

Preparation of Data – The entire data set of winter season for year 2006 to 2008 was randomly divided into training and testing set. 70% of data was taken for training, 20% for testing and 10% for predication.

(i) Normalization – Further input data was normalized between 0 and 1. This was done using the formula

$$x_i = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Maximum Normalization value was taken as 0.9 and minimum normalization value was taken as 0.1

Model inputs - From the available literature studies, all the available data set (variables) were taken as input. All available variables (Nitric oxide (NO), Sulphur dioxide (SO_2), Ozone (O_3), Nitrogen dioxide (NO_2) and Carbon monoxides (CO)) were considered for Siri Fort site while Air Pollutants (Nitric oxide (NO), Sulphur dioxide (SO_2), Ozone (O_3), Nitrogen dioxide (NO_2), SPM (2.5) and carbon monoxides (CO) were taken for ITO site. as potential input variable. The model that performed best was retained. The maximum variation in input value was applied so that the models retain maximum information about the input source.

Network structure – ANN structure comprises of hidden layer modes and the number of modes in each of these layers. It determines the number of parameters needed for a model performance. If sufficient number of parameters were not found then it will be difficult in getting the convergence. If we increase the parameter too much then the network will take longer time to converge. In many cases even a single layer is sufficient to approximate a function. Unnecessarily taking many layers may complicate the problem and the desired result may not be obtained.

Trial and error method was used to meet the convergence criteria. In our case two hidden layer provided satisfactory result.

Activation function – In this Feed forward back propagation model sigmoidal function was used, from Input to first layer – sigmoidal function, first layer to 2nd layer – sigmoidal function and second layer to output – sigmoidal function.

Parameters estimation (Training) – The various steps followed is described below.

Step 1: The connection weights were assigned random values.

Step 2: Momentum constant was selected in the range 0.1 to 0.8.

Step 3: Two hidden layer were used. Trial and error method was followed.

Step 4: Learning rate constant were selected in the range 0.1 to 0.7.

Step 5: A training sample was presented to the network. After presenting all the parameters to the network, train the network command executed. Training continues till we achieve the given criteria. After sufficient number of iterations, output was produced.

Step 6: Error function was calculated as

$$E = \frac{1}{2}\sum (O_i - P_i)^2$$

where E is the error function.

O_i is the desired output,

P_i is the predicted output.

Step 7: The connection weights (w) are adjusted using the gradient descent rule of optimization

$$\Delta w(t) = \sum_{s=1}^{\varepsilon} -\eta \frac{\partial E}{\partial w} + \mu \Delta w(t-1),$$

where s is the training sample resented to network, η is the learning rate, and μ is the momentum value.

All the steps from 2 to 7 were repeated until specified criteria are met.

Once training was initiated we keep on observing graphs between Root Mean Square Error (RMSE) and number of iteration. When this error gets minimized then training

stops. This way we keep on repeating the procedure for various samples of training data. The model so obtained was applied for validation.

Model validation – After completion of training process, the model performance needs to be estimated using test data. A part of data has been kept separated while training from master data file. Now this data was presented to the model and after getting the specified error goal the best model was chosen for our purpose for prediction.

Optimal structure - The architecture corresponding to the minimal RMSE on the validation set is chosen as the optimal structure.

Performance indices – The model performance was evaluated by the following statistical criteria, namely, the determination coefficient R^2, the root mean square error RMSE , Fractional Bias (FB) and the index of agreement (IA) . Their formulas are mentioned below.

$$R = \sqrt{\frac{\sum_{i=1}^{n}(O_i - \overline{O})^2 - \sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2}}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|P_i - O_i|$$

$$IA = 1 - \frac{\sum_{i=1}^{n}|P_i - O_i|^2}{\sum_{i=1}^{n}(|P_i - \overline{O}| + |O_i - \overline{O}|)^2}$$

$$FB = \frac{\overline{P} - \overline{O}}{0.5(\overline{P} + \overline{O})}$$

P stands for predicted value and O stands for observed value and $\overline{P}$ and $\overline{O}$ stands for corresponding mean values

[3.4] Representation of model performance parameters

Figure 3.1: *RMSE during training the network of NO at ITO*

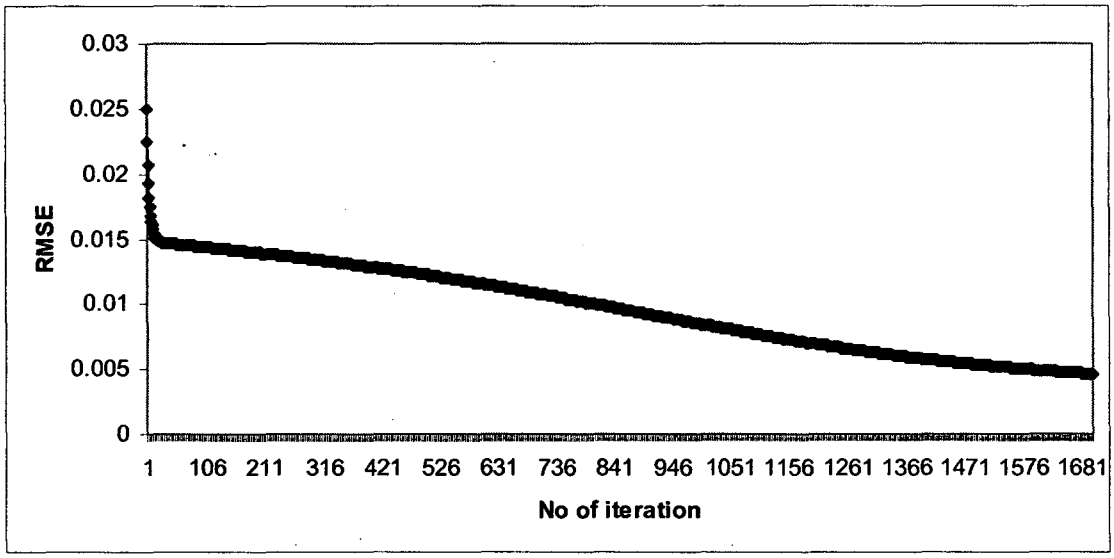


Figure 3.2: *RMSE during training the network of NO_2 at ITO*

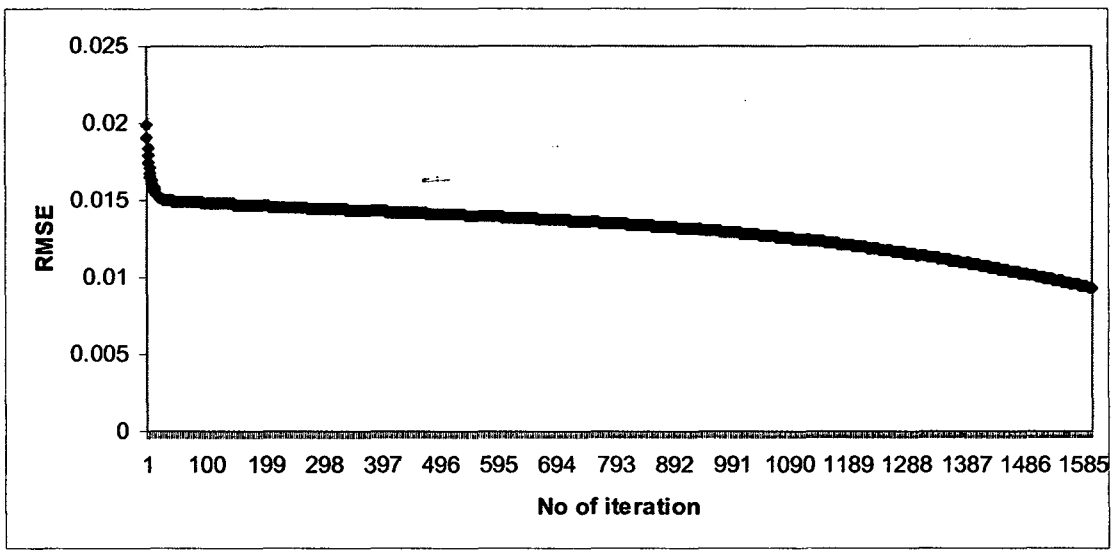


Figure 3.3: *RMSE during training the network of NO at Siri Fort*

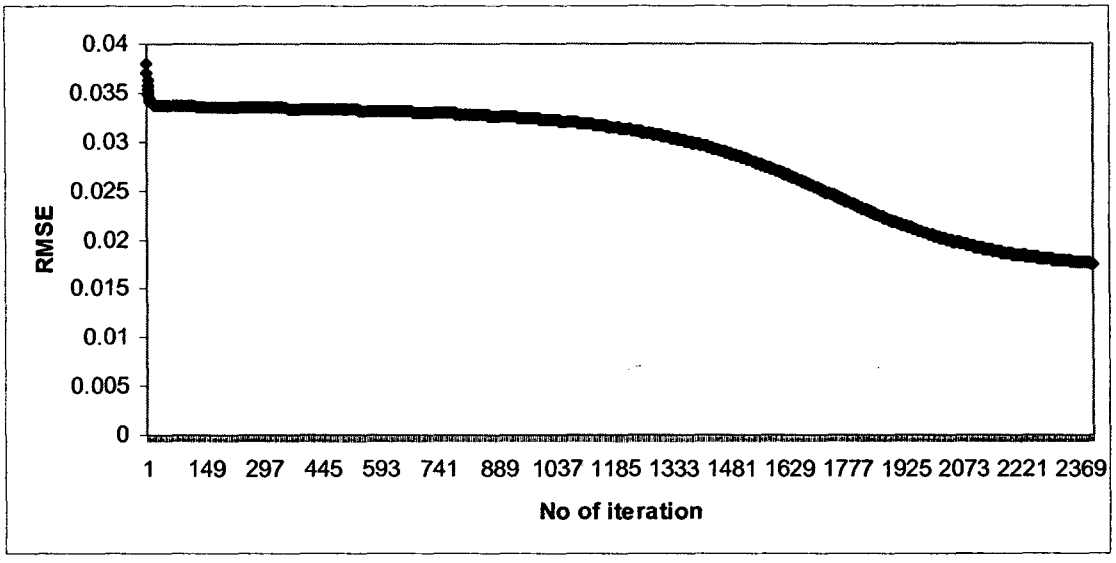


Figure 3.4: *RMSE during training the network of NO_2 at Siri Fort*

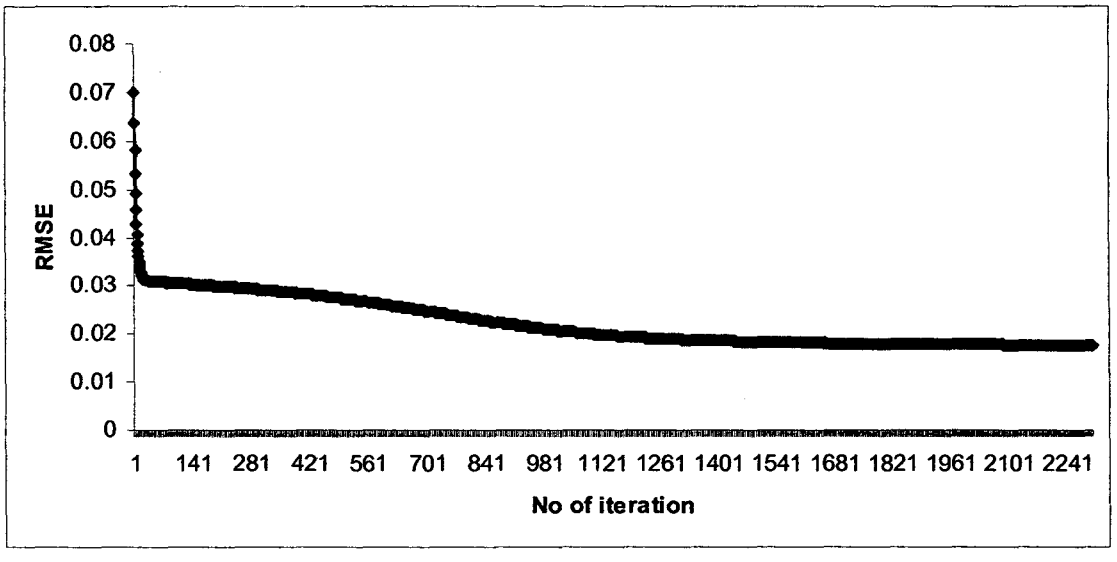


Table 3.1: *Network architecture of feedforward propagation network at Siri Fort site*

Model Parameters	NO	NO_2
No. of inputs	4	4
Processing time	375 s	345 s
No. of layers	2	2
No. of nodes in hidden layer 1	9	8
No. of nodes in hidden layer 2	5	5
Error tolerance	10^{-6}	10^{-6}
Momentum constant	0.8	0.8
No. of iteration	1700	1600
Learning rate	0.7	0.6

Table 3.2: *Network architecture of feedforward propagation network at ITO site*

Model Parameters	NO	NO_2
No. of inputs	5	5
Processing time	480 s	455 s
No. of layers	2	2
No. of nodes in hidden layer 1	10	9
No. of nodes in hidden layer 2	6	6
Error tolerance	10^{-6}	10^{-6}
Momentum constant	0.8	0.8
No. of iteration	2400	2300
Learning rate	0.5	0.5

Table 3.3: *Forecasting result of air pollutants NO and NO_2 at ITO site*

No. of Days	NO (% of error)	NO_2 (% of error)
1.	6.61	15.40
2.	12.78	20.31
3.	15.45	19.09
4.	-17.25	-11.36
5.	-14.2	31.57
6.	12.07	29.27
7.	17.01	-30.03
8.	-5.24	18.12
9.	-3.55	19.96
10.	-3.56	-49.66
11.	21.85	32.62
12.	25	83.86
13.	1.42	-58.28
14.	18.15	74.91
15.	-12.34	51.57
16.	16.45	-41.33
17.	-24.85	58.74
18.	15.33	23.32
19.	4.5	-37.48
20.	6.41	35.31
21.	-6.22	11.84
22.	-1.72	54.13
23.	-4.32	-24.4
24.	3.83	28.90
25.	18.8	-31.41
26.	6.34	11.45
27.	-17.42	16.73
28.	-33.46	7.18
29.	29	16.62
30.	22.2	11.13
31.	-11.3	2.56
32.	6.49	17.13
33.	-19.2	52.20
34.	20.12	-66.98
35.	-21	45.5
36.	19.95	-21.95
37.	15.6	-43.55
38.	-11.42	66.08
39.	-6.9	34.89
40.	10.41	20.18
41.	17.7	-27.03
42.	-13.56	4.10
43.	-29.9	17.85
44.	38.48	12.25
45.	21.9	25.73
46.	5.61	22.20
47.	-28.74	6.24
48.	6.61	20.80

Table 3.4: *Forecasting result of air pollutants NO and NO_2 at Siri Fort site*

No. of Days	NO (% of error)	NO_2 (% of error)
1.	31.28	34.30
2.	21.7	41.18
3.	-31.97	17.68
4.	28.66	-32.2
5.	32.67	15.46
6.	-15.07	29.57
7.	42.34	-31.58
8.	14.76	16.72
9.	-10.7	-33.84
10.	38.66	16.95
11.	24.35	28.79
12.	-31.55	19.61
13.	14.1	-60.89
14.	23.93	15.76
15.	-24.06	10.89
16.	11.56	-14.53
17.	52.41	11.34
18.	-20.81	20.11
19.	19.42	44.25
20.	21.18	19.45
21.	44.82	-39.82
22.	-12.02	33.29
23.	29.05	34.94
24.	-24.83	26.47
25.	51.31	-13.93
26.	-46.76	16.64
27.	22.83	56.83
28.	6.69	54.19
29.	34.55	27.85
30.	-53.69	-30.12
31.	18.06	25.63
32.	25.03	46.37
33.	-19.5	62.45
34.	25.51	-53.39
35.	-14.04	12.06
36.	24.62	22.63
37.	17.69	-46.9
38.	23.99	13.63
39.	-12.88	23.12
40.	23.8	-14.43
41.	16.72	28.93
42.	61.74	30.76
43.	-29.98	-56.2
44.	18.74	10.91
45.	26.29	10.18
46.	-42.98	31.5
47.	31.28	-31.95
48.	21.7	43.38

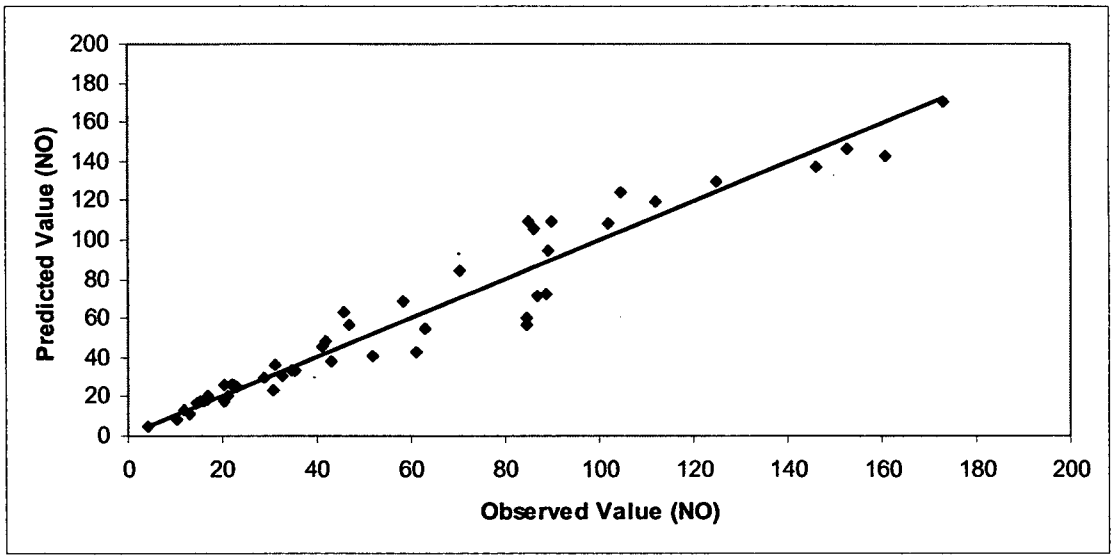


Figure 3.5: *Scatter diagram of predicted versus observed value of NO at ITO from 1st to 2nd Feb. 2008. Value of* $R^2 = 0.84$.

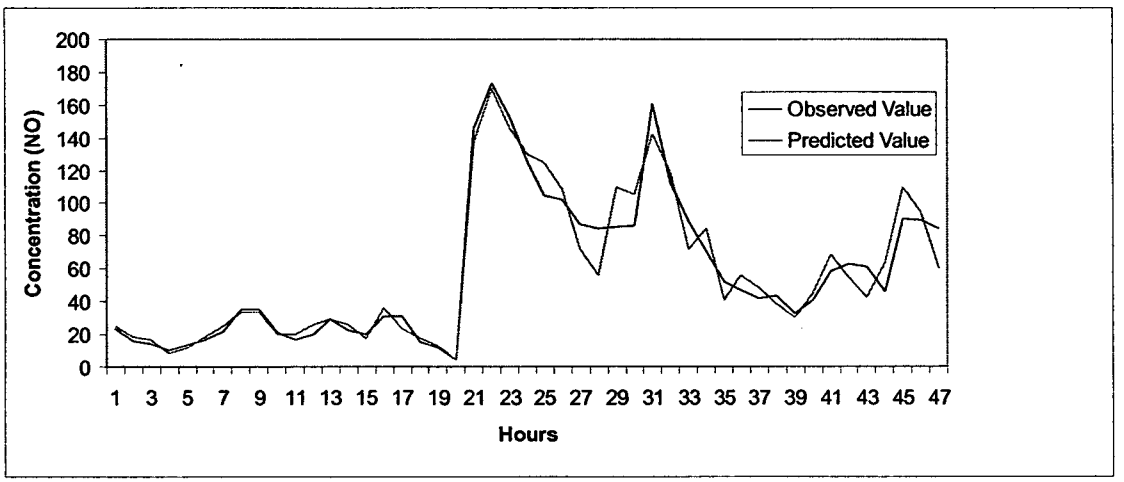


Figure 3.6: *Hourly time series of the observed and predicted concentration of NO* (μgm^{-3}) *at ITO from 1st to 2nd Feb. 2008.*

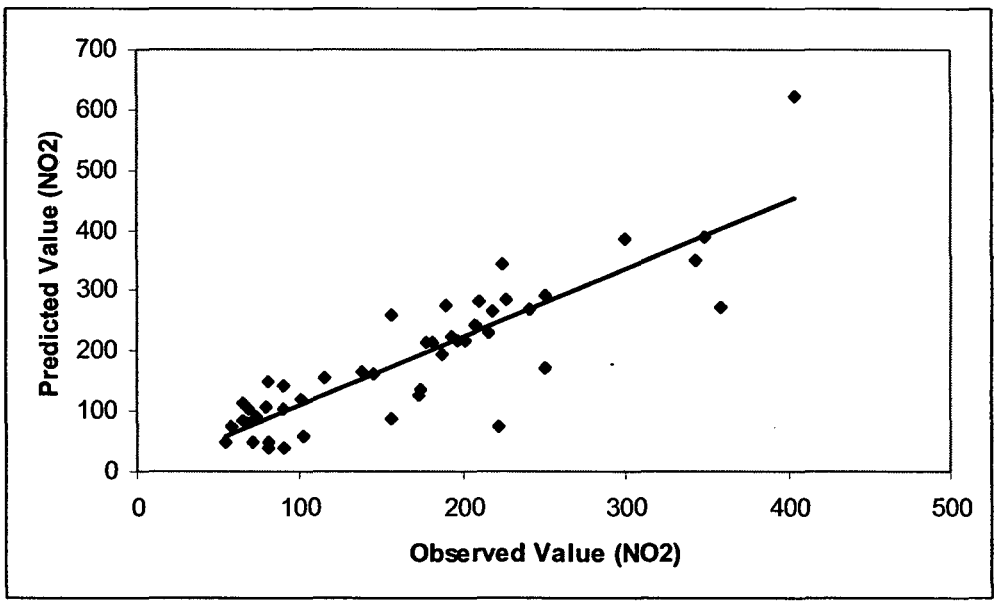


Figure 3.7: *Scatter diagram of predicted versus observed value of NO_2 at ITO from 1st to 2nd Feb. 2008. Value of R^2= 0.80 .*

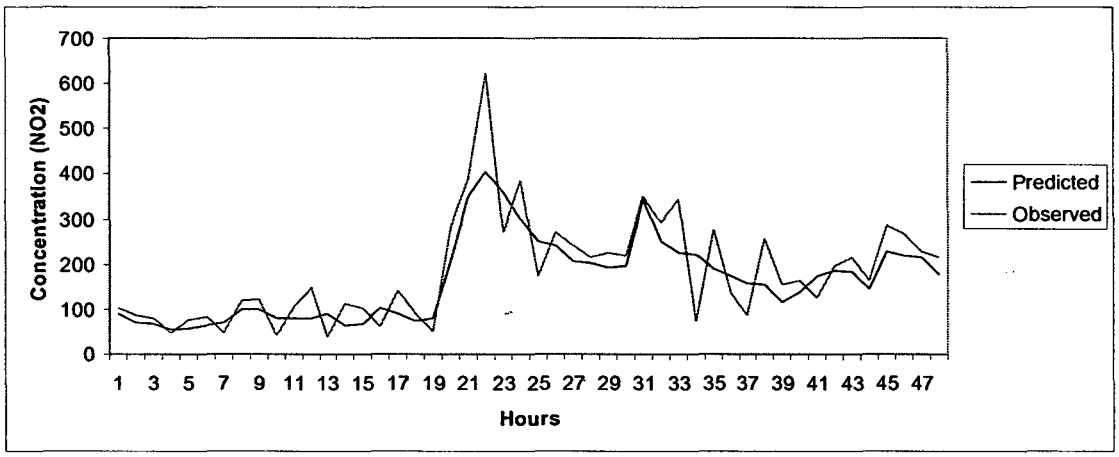


Figure 3.8: *Hourly time series of the observed and predicted concentration of NO_2 (μgm^{-3}) at ITO from 1st to 2nd Feb. 2008.*

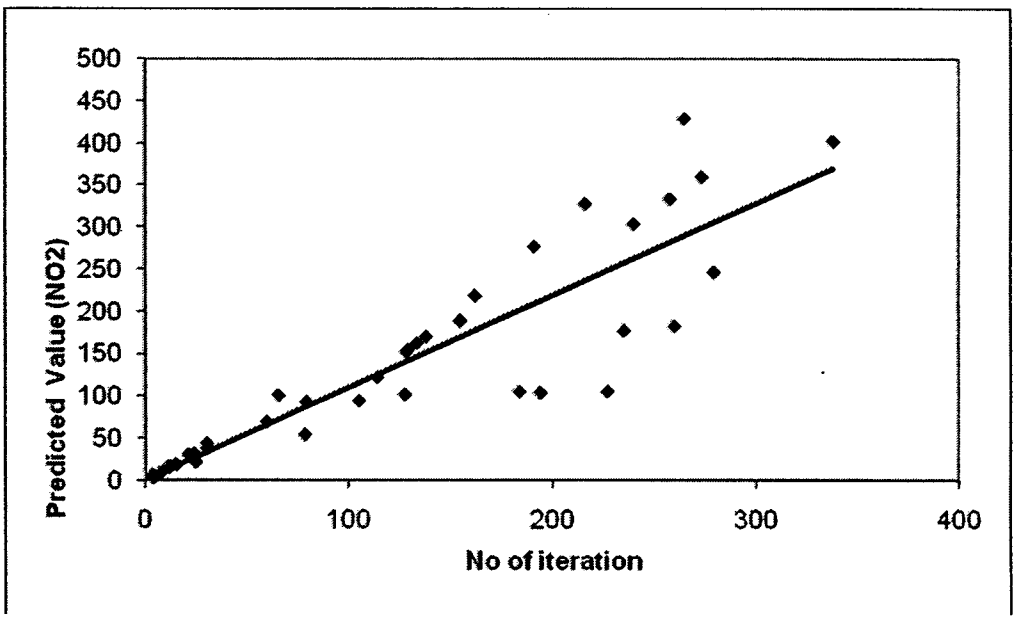


Figure 3.9: *Scatter diagram of predicted versus observed value of NO at Siri Fort from 19th to 20th Feb. 2008. Value of* R^2=0.79 .

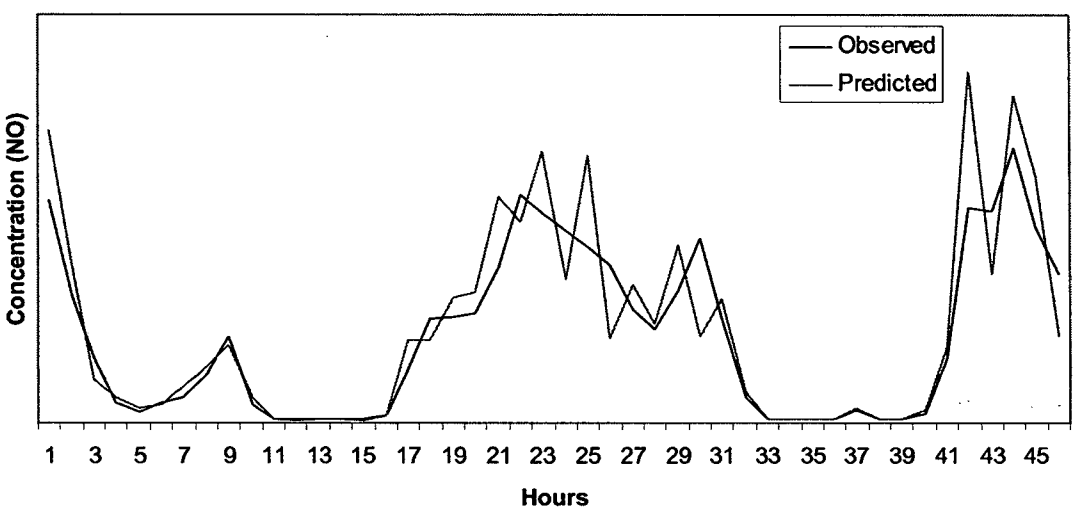


Figure 3.10: *Hourly time series of the observed and predicted concentration of NO* (μgm^{-3}) *at Siri Fort from 19th to 20th Feb. 2008.*

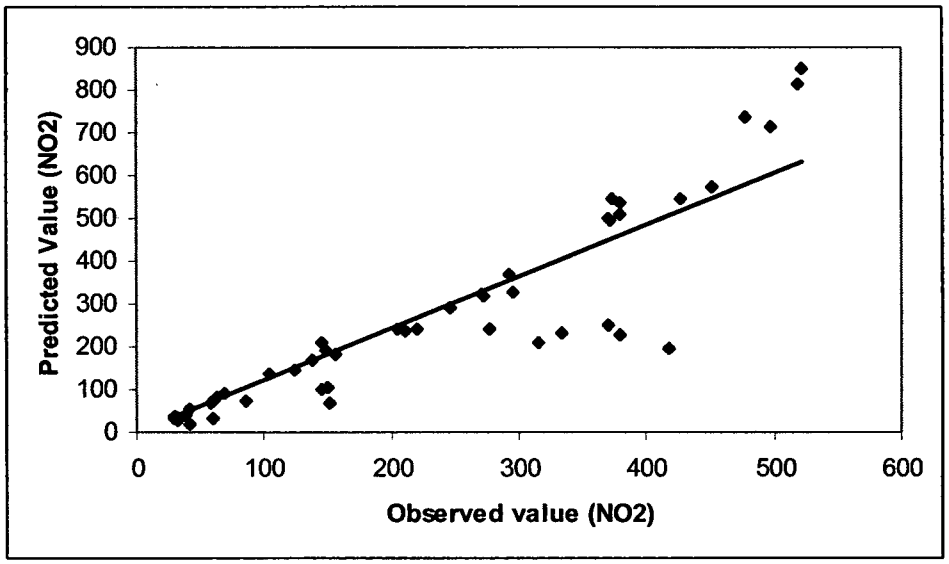


Figure 3.11: *Scatter diagram of predicted versus observed value of NO_2 at Siri Fort from 19th to 20th Feb. 2008. Value of R^2= 0.74*

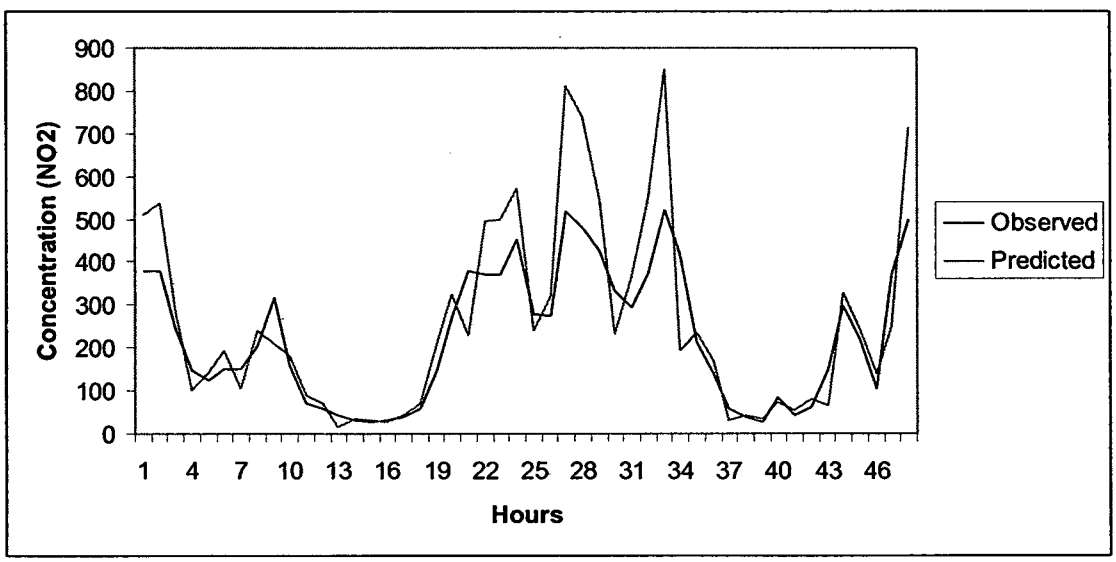


Figure 3.12: *Hourly time series of the observed and predicted concentration of NO_2 (μgm^{-3}) at Siri Fort from 19th to 20th Feb. 2008.*

CHAPTER – 4

Chapter 4 – Artificial Neural Network Modeling with Radial Basis Function Approach

[4.1] Introduction

In this chapter, we deal with another type of neural network known as Radial basis function (RBF) network which has important universal approximation properties. The need for use of this function arises as most MLP networks having back propagation algorithm face problem of local minima convergence. This undermines the performance of the model. To overcome this difficulty, the radial basis function network was developed and successfully used for many applications. (Roadknight et al., 1997) This network provides an effective method for training of an MLP network. For function approximation this radial basis function network is best suited. Its learning speed is very fast as compared to simple BP algorithm.

RBF network uses memory-based learning for their design. Specifically, learning is viewed as a curve-fitting problem in high-dimensional space (Broomhead and Lowe, 1989). Learning process takes place by finding a surface in a multidimensional space that provides a best fit to the training data and then generalization (i.e., response of the network to input data not seen before) is performed by the use of this multidimensional surface to interpolate the test data.

[4.2} Theoretical Description of the Radial basis function network

Radial – basis functions were first introduced in the solution of the real multivariate interpolation problem. The R*adial basis functions (RBF) network* consists of three layers.

1. The first layer is that of input layer. The input layer is made up of source nodes.
2. The second layer is the hidden layer in the network. This applies a Gaussian function from the input space to the hidden space.

3. The function is linear between hidden layer and output layer. We discuss here the Generalized Radial Basis Function network. (Haykin, 1998)

The transfer function used here is given as $F(x) = e^{\frac{(x-\mu)^2}{2\sigma^2}}$

[4.2.1] Generalized Radial-Basis Function Networks

To overcome computational difficulties due to weight space, the complexity of the network need to be reduced. This can be done by applying some approximation to the regularized solution.

using *Galerkin's method.* the approximated solution F (x) is expended on a finite basis, as shown by (Poggio and Girosi, 1990a)

$$F(x) = \sum_{i=1}^{m_1} w_i \varphi_i(x) \tag{4.1}$$

where $\{\varphi_i (x \mid i = 1, 2... m_1\}$ is a new set of basis functions.

where $m_1 \leq n$, and ω_i constitute a new set of weights.

$$\varphi_i(x) = G(\| x - t_i \|), \qquad i = 1,2,...,m_1 \tag{4.2}$$

where the set of centers $\{t_i \mid i = 1, 2..., m_1\}$ is to be determined. This particular choice of basis functions is the only one that guarantees that in the case of $m_1 = N$, and

$$t_i = x_i, \qquad i = 1,2,...,N$$

from eq. (4.1) and (4.2), we get,

$$F(x) = \sum_{i=1}^{m_1} w_i G(x, t_i)$$

$$= \sum_{i=1}^{m_1} w_i G(\| x - t_i \|)$$

now to determine the new set of weights $\{\omega_i \mid i = 1, 2, ..., m_1\}$ so as to minimize the new cost functional $\xi(F)$ defined by

$$\xi(F) = \sum_{i=1}^{N} \left(d_i - \sum_{j=1}^{m_1} w_j G(\| x_i - t_j \|) \right)^2 + \lambda \| DF \|^2 \tag{4.3}$$

The first term on the right-hand side of Eq. (4.3) can be written as the squared Euclidean norm $\| d - Gw \|^2$, where

$$d = [d_1, d_2 ..., d_N]^T$$

$$G = \begin{bmatrix} G(x_1,t) & G(x_1,t_2) & \cdots & G(x_1,t_{m_1}) \\ G(x_2,t_1) & G(x_2,t_2) & \cdots & G(x_2,t_{m_1}) \\ \vdots & \vdots & & \vdots \\ G(x_N,t_1) & G(x_N,t_2) & \cdots & G(x_N,t_{m_1}) \end{bmatrix}$$

$$w = [w_1, w_2, ..., w_{m_1}]^T$$

The second term on the right-hand side of Eq. (4.3) as

$$\begin{aligned} \| DF \|^2 &= (DF, DF)_H \\ &= \left[\sum_{i=1}^{m_1} w_i G(x,t_i), \tilde{D}D \sum_{i=1}^{m_1} w_i G(x;t_i) \right]_H \\ &= \left[\sum_{i=1}^{m_1} w_i G(x,t_i), \sum_{i=1}^{m_1} w_i \delta t_i \right]_H \\ &= \sum_{j=1}^{m_1} \sum_{i=1}^{m_1} w_i w_i G(t_j,t_i \\ &= w^T G_0 w \end{aligned}$$

The matrix G_o is a symmetric m_1- by $- m_1$ matrix, defined by

$$G_0 = \begin{bmatrix} G(x_1,t_1) & G(x_1,t_2) & \cdots & G(t_1,t_{m_1}) \\ G(x_2,t_1) & G(x_2,t_2) & \cdots & G(t_2,t_{m_1}) \\ \vdots & \vdots & & \vdots \\ G(t_{m_1},t_1) & G(t_{m_1},t_2) & \cdots & G(t_{m_1},t_{m_1}) \end{bmatrix}$$

minimization Eq.(4.3) with respect to the weight vector **w**

we get, $(G^T G + \lambda G_0)w = G^T d$

As λ approaches zero, the weight vector w converges to the pseudoinverse (minimum-norm) solution to the over determined least-squares data-fitting problem for $m_1 < N$, (Broomhead and Lowe, 1988)

$$w = G^+ d, \qquad \lambda = 0$$

where G^+ is the pseudoinverse of matrix G; i.e,

$$G^+ = (G^T G)^{-1} G^T$$

$$\begin{aligned} \| x \|_C^2 &= (Cx)^T (Cx) \\ &= x^T C^T C x \end{aligned}$$

where C is an m_o-by-m_o *norm weighting matrix*, and m_o is the dimension of the input vector x.

The function may be written as

$$F(x) = \sum_{i=1}^{m_1} e_i G(\| x - t_i \|_C)$$

A Gaussian radial-based function $G(\| x - t_i \|_C)$ at centered at t_i and having weight matrix C can be written as

$$G(\| x-t_i \|_C = \exp[-x-t_i)^T C^T C(x-t_i)]$$
$$= \exp\left[-\frac{1}{2}(x-t_i)^T \sum^{-1}(x-t_i)\right]$$

[4.2.2] Universal Approximation Theorem

Let $G: \Re^{m_0} \to \Re$ be an integrable bounded function such that G is continuous and

$$\int_{\Re^{m_0}} G(x)dx \neq 0$$

and

$$F(x) = \sum_{i=1}^{m_1} w_i G\left(\frac{x-t_i}{\sigma}\right)$$

Where $\sigma > 0, w_i \in \Re$ and $t_i \in \Re^{m_0}$ for $i = 1,2...,m_1$.

[4.2.3] Normalized RBF network

By applying *spherical symmetry* of the kernel *K(x)*, (Krzyzak et al., 1996)

$$K\left(\frac{x-x_i}{h}\right) = K\left(\frac{\| x-x_i \|}{h}\right) \qquad \text{for all } i$$

where $\| . \|$ denotes the Euclidean norm of the enclosed vector. So the *normalized radial basis function*

$$\psi_N(x.x_i) = \frac{K\left(\frac{\| x-x_i \|}{h}\right)}{\sum_{j=1}^{N} K\left(\frac{\| x-x_j \|}{h}\right)}, \qquad i = 1,2,...,N$$

$$\sum_{i=1}^{N} \psi_N(x,x_i) = 1 \qquad \text{for all } x$$

The input-output mapping function of the normalized RBF network takes the following form:

$$F(x) = \frac{\sum_{i=1}^{N} w_i \exp\left(-\frac{\| x - x_i \|^2}{2\sigma^2}\right)}{\sum_{j=1}^{N} \exp\left(-\frac{\| x - x_j \|^2}{2\sigma^2}\right)}$$

[4.2.4] Learning methods

There are various techniques for learning of Radial Basis function network. Some of the approaches are mentioned below.

1. **Random Selection of Fixed Centers** – This is the simplest approach for the learning process. In this method, locations of centers are chosen randomly from the training data (Broomhead and Lowe 1998). In this case, the normalized radial basis function centered at t_i may be defined as

$$G(\| x - t_i \|^2 = \exp\left(-\frac{m_1}{d_{\max}^2} \| x - t_i \|^2\right), \qquad i = 1,2,...,m_1$$

where m_1 is the number of centers and d_{max} is the maximum distance between the chosen centers. In effect, the standard deviation (i.e., width) of all the Gaussian radial-basis functions is fixed at

$$\sigma = \frac{d_{\max}}{\sqrt{2m_1}}$$

2. **Self-organized Selection of Centers** – In the first case, some limitation came to light. For optimum performance large training set was required. So another method of hybrid learning was developed (Moody and Darken 1989). For optimum performance large training set was required needed which consists of the following two processes.

(i) Appropriate locations for centers were estimated in the hidden layer of radial basis function.

(ii) Then supervised learning method was applied by estimating all linear weights of the output layer.

3. **Supervised selection of centers** – This approach was developed by (Poggio and Girosi, 1990). This is a supervised approach in which all the parameters of the radial basis function network undergo a supervised learning process. We fixed up the error goal. Accordingly other parameters adjusted.

   So error goal is defined as

$$E = \frac{1}{2}\sum_{j=1}^{N} e_j^2$$

   where N is size of training data e_i is the error of the output

$$e_j = d_j - F(x_j)$$
$$= d_j - \sum_{i=1}^{M} w_i(\| x_j - t_i \|)$$

[4.3] Development of the Model

As discussed, ANN basically consists of inputs source (data) transfer function and outputs. The important thing is how to transform the available data and then the number of inputs to be selected for training purpose. In this Radial basis function network the transfer function used is a Gaussian Function between input layer and hidden layer. The linear function is used between hidden layer and output layer. The various parameters for training the model has to be made as we have done in the feedforward back propagation network.

Preparation of Data

(ii) The entire data set of winter season for year 2006 to 2008 was randomly divided into training and testing set. 70% of data was taken for training, 20% for testing and 10% for predication.

(iii) Normalization – Further input data was normalized between 0 and 1. This was done using the formula

$$x_i = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Maximum Normalization value was taken as 0.9 and minimum normalization value was taken as 0.1

Model inputs – From the available literature studies, all the available data set (variables) were taken as input. All available variables (Nitric oxide (NO), Sulphur dioxide (SO_2), Ozone (O_3), Nitrogen dioxide (NO_2) and Carbon monoxides (CO)) were considered for Siri Fort site while Air Pollutants (Nitric oxide (NO), Sulphur dioxide (SO_2), Ozone (O_3), Nitrogen dioxide (NO_2), SPM (2.5) and carbon monoxides (CO) were taken for ITO site. as potential input variable. The model that performed best was retained. The maximum variation in input value was applied so that the models retain maximum information about the input source.

Network structure – ANN structure comprises of hidden layer modes and the number of modes in each of these layers. It determines the number of parameters needed for a model performance. If sufficient number of parameters were not found then it will be difficult in getting the convergence. If we increase the parameter too much then the network will take longer time to converge. In many cases even a single layer is sufficient to approximate a function. Unnecessarily taking many layers may complicate the problem and the desired result may not be obtained.

Trial and error method was used to meet the convergence criteria. In our case two hidden layer provided satisfactory result.

Once the model inputs were fed into the network, the remaining procedural steps for training the model were the same as mentioned in section (3.3) in the earlier chapter 4.

Activation function – In this Radial Basis Function network Gaussian function was used. From Input to hidden layer and linear function was used between hidden layer and output layer.

After initiating the training we keep on observing graphs between Root Mean Square Error (RMSE) and number of iteration. When this error gets minimized then training stops. This way we keep on repeating the procedure for various samples of training data. The model so obtained was applied for validation.

Model validation – After completion of training process, the model performance needs to be estimated using test data. A part of data has been kept separated while training from master data file. Now this data was presented to the model and after getting the specified error goal the best model was chosen for our purpose for prediction.

Optimal structure - The architecture corresponding to the minimal RMSE on the validation set is chosen as the optimal structure.

Performance indices – The model performance was evaluated by the following statistical criteria, namely, the determination coefficient R^2, the root mean square error RMSE , Fractional Bias and the index of agreement.

[4.4] Representation of model performance parameters

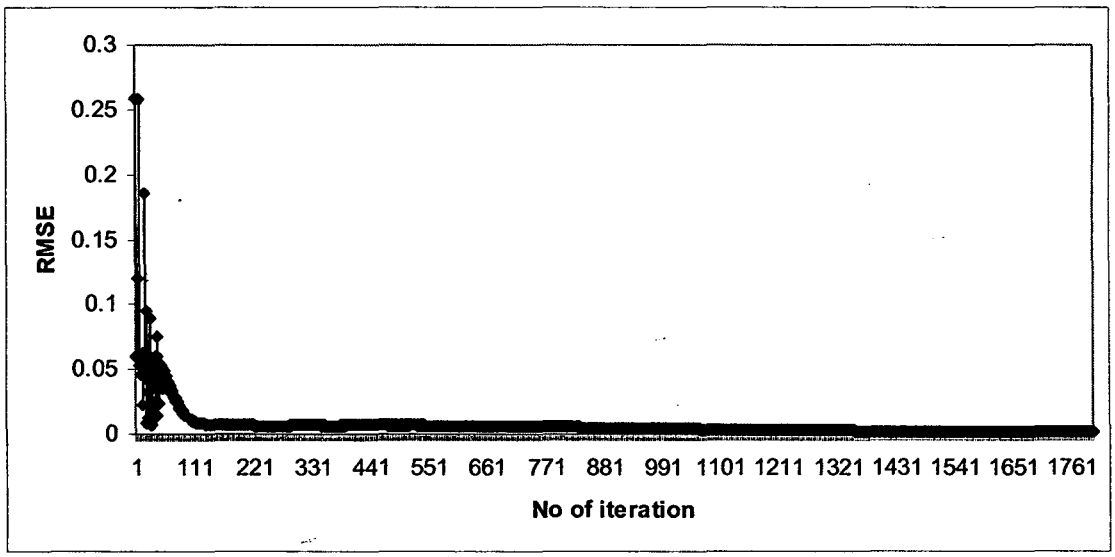


Figure 4.1 *RMSE during training the network of NO at ITO*

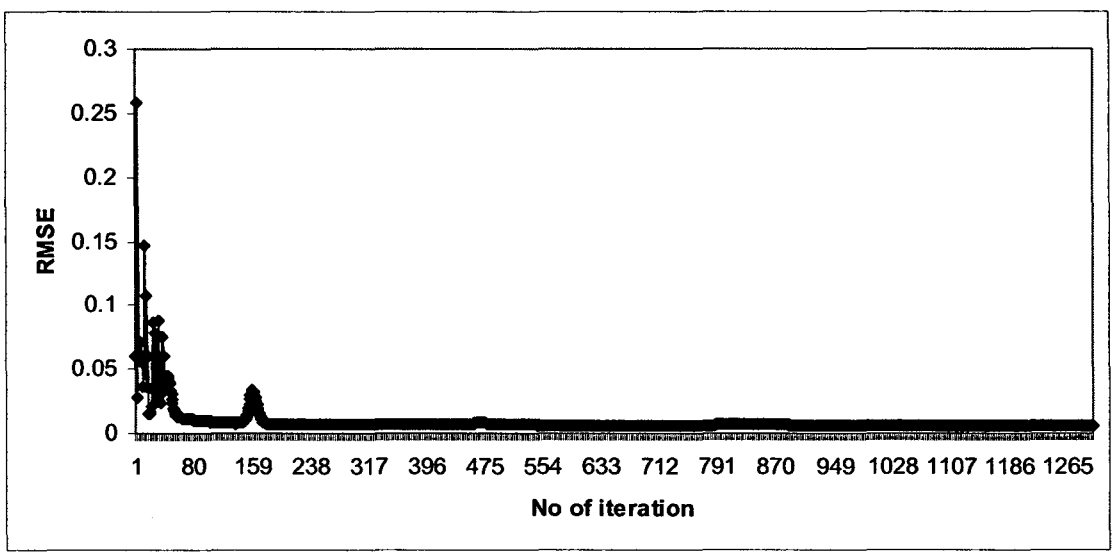


Figure 4.2 *RMSE during training the network of NO_2 at ITO*

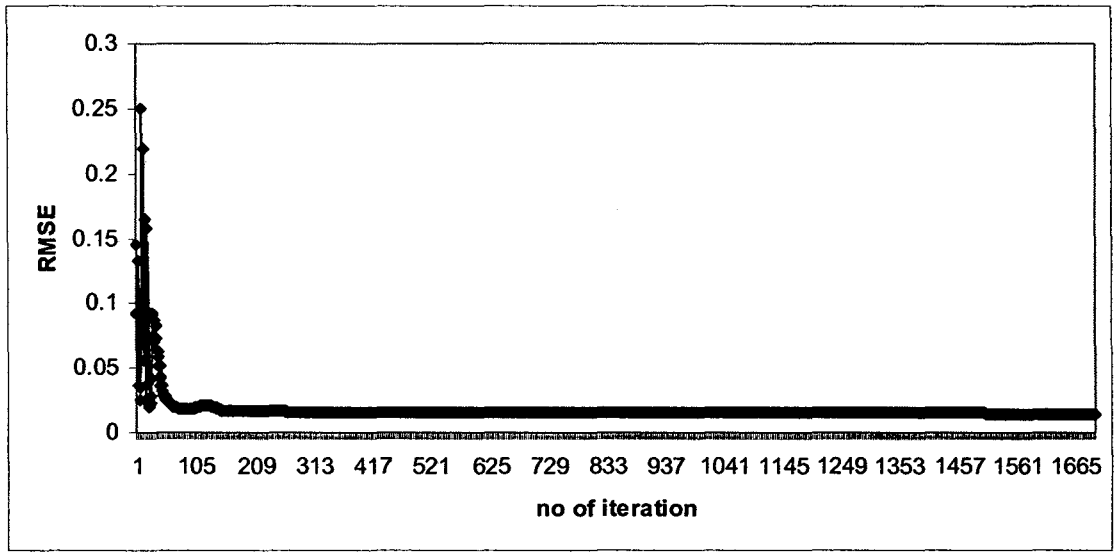


Figure 4.3 *RMSE during training the network of NO at Siri Fort*

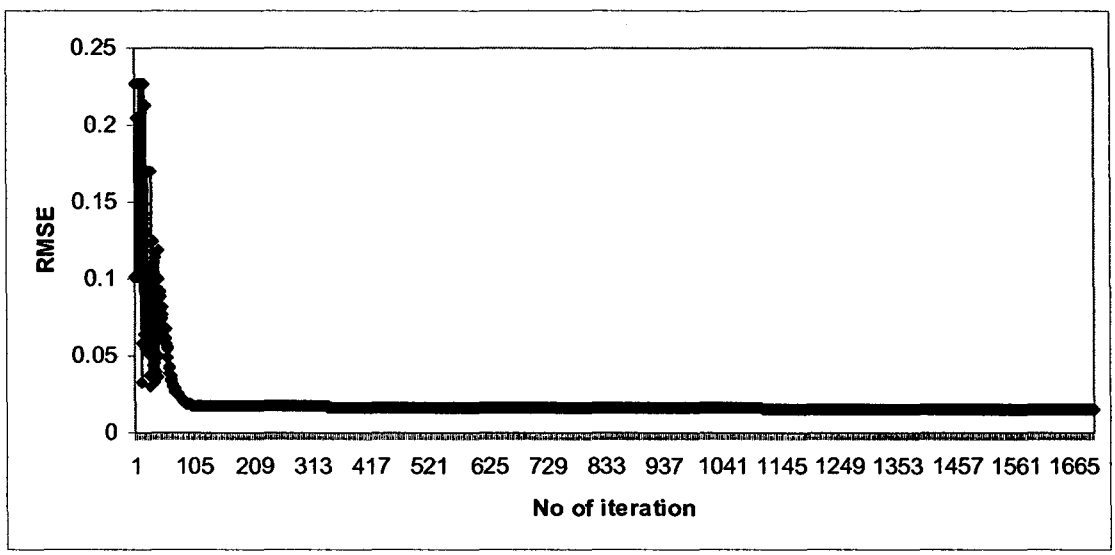


Figure 4.4 *RMSE during training the network of NO_2 at Siri Fort*

Table 4.1: *Network architecture of RBF network at ITO site*

Model Parameters	NO	NO_2
No. of inputs	5	5
Processing time	380 s	325 s
No. of layers	1	1
No. of nodes in hidden layer 1	12	10
No. of nodes in hidden layer 2		
Error tolerance	10^{-6}	10^{-6}
Momentum constant	0.7	0.7
No. of iteration	1800	1300
Learning rate	0.7	0.7

Table 4.2: *Network architecture of RBF network at Siri Fort site*

Model Parameters	NO	NO_2
No. of inputs	4	4
Processing time	390	405
No. of layers	1	1
No. of nodes in hidden layer 1	12	14
No. of nodes in hidden layer 2		
Error tolerance	10^{-6}	10^{-6}
Momentum constant	0.6	0.6
No. of iteration	1800	1900
Learning rate	0.7	0.7

Table 4.3: *Forecasting result of air pollutants NO and NO_2 at ITO*

No. of Days	NO (% of error)	NO_2 (% of error)
1.	26.21	24.62
2.	12.78	14.71
3.	15.45	34.28
4.	-17.25	16.35
5.	-14.2	8.30
6.	12.07	26.32
7.	17.01	17.89
8.	-25.24	11.21
9.	-33.45	12.33
10.	-13.56	20.42
11.	21.85	11.77
12.	25	13.91
13.	11.42	5.67
14.	18.15	30.75
15.	-12.34	9.17
16.	36.45	8.29
17.	-24.85	21.45
18.	35.33	26.84
19.	14.55	7.49
20.	26.41	24.85
21.	-26.24	35.24
22.	-22.72	23.78
23.	-14.32	19.06
24.	33.83	24.56
25.	18.8	16.11
26.	16.34	8.98
27.	-17.42	25.51
28.	-33.46	13.14
29.	29	17.89
30.	22.2	28.82
31.	-11.3	8.46
32.	16.49	29.74
33.	-19.2	14.27
34.	20.12	17.45
35.	-21	6.66
36.	19.95	17.97
37.	15.6	8.98
38.	-11.42	35.72
39.	-26.9	18.1
40.	10.41	28.32
41.	27.89	17.04
42.	-13.56	33.05
43.	-29.9	61.58
44.	38.48	28.70
45.	21.9	27.10
46.	15.61	16.5
47.	-28.74	23.36
48.	26.21	13.08

Table 4.4: *Forecasting result of air pollutants NO and NO_2 Siri Fort*

No. of Days	NO (% of error)	NO_2 (% of error)
1.	28.56	9.67
2.	37.03	-25.35
3.	-24.08	19.23
4.	21.99	27.64
5.	16.28	-19.51
6.	28.84	26.72
7.	-15.31	71.31
8.	28.63	-38.27
9.	13.74	30.36
10.	31.47	18.23
11.	-24.57	11.37
12.	39.67	-55.83
13.	14.31	27.83
14.	-44.47	10.59
15.	17.5	15.23
16.	42.3	-18.78
17.	-27.84	32.76
18.	41.69	26.65
19.	28.87	-33.85
20.	-24.68	10.30
21.	24.72	-23.14
22.	42.58	-20.12
23.	53.37	-23.67
24.	-13.57	31.64
25.	21.61	-14.5
26.	14.35	8.06
27.	25.69	55.75
28.	-13.67	45.69
29.	31.29	-19.83
30.	21.61	17.29
31.	13.76	15.71
32.	-24.52	40.06
33.	31	-58.75
34.	14.5	42.29
35.	18.65	19.31
36.	-21.72	26.05
37.	19.36	-33.5
38.	57.29	8.08
39.	62.35	70.23
40.	-56.53	-52.31
41.	28.65	51.78
42.	85.14	17.75
43.	-24.76	-28.12
44.	30.82	18.10
45.	27.1	34.08
46.	-45.47	-42.55
47.	17.97	17.33
48.	29.23	38.98

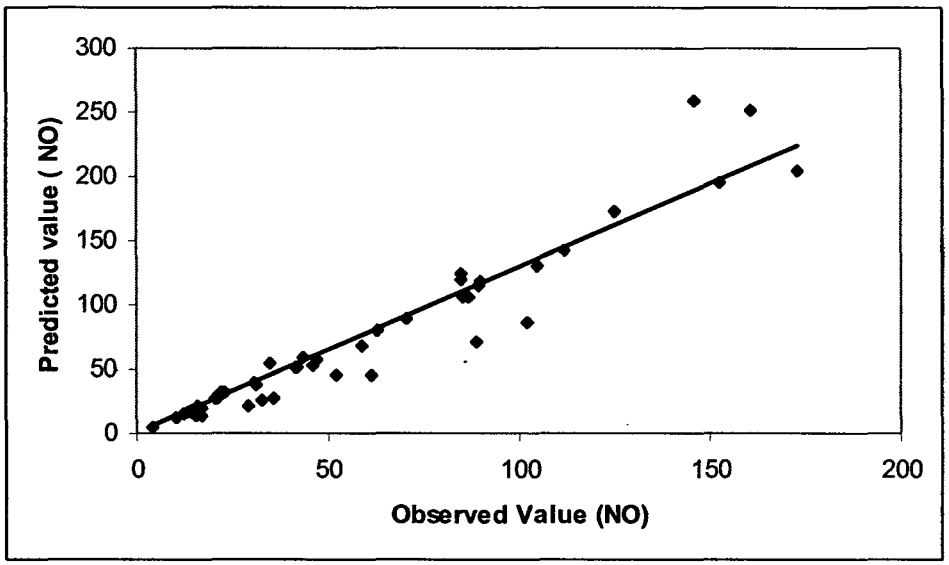


Figure 4.5: *Scatter diagram of predicted versus observed value of NO at ITO from 1st to 2nd Feb. 2008. Value of* $R^2 = 0.89$.

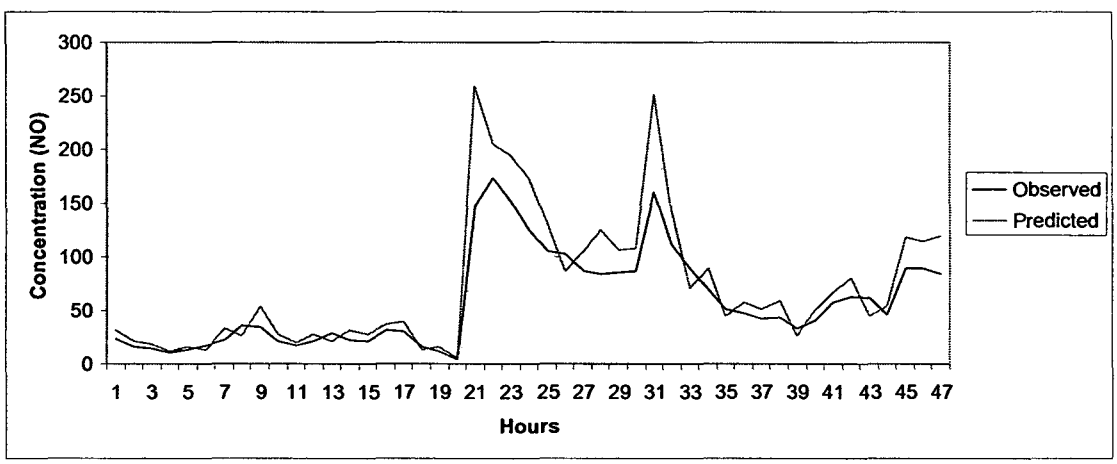


Figure 4.6: *Hourly time series of the observed and predicted concentration of NO* (μgm^{-3}) *at ITO from 1st to 2nd Feb. 2008.*

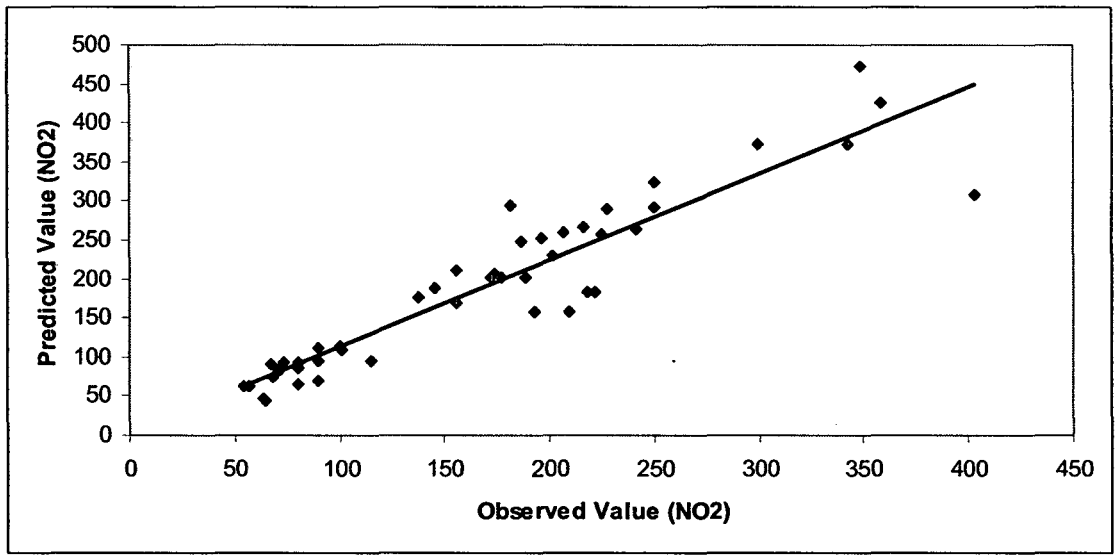


Figure 4.7: *Scatter diagram of predicted versus observed value of NO_2 at ITO from1st to 2nd Feb. 2008. Value of R^2= 0.88.*

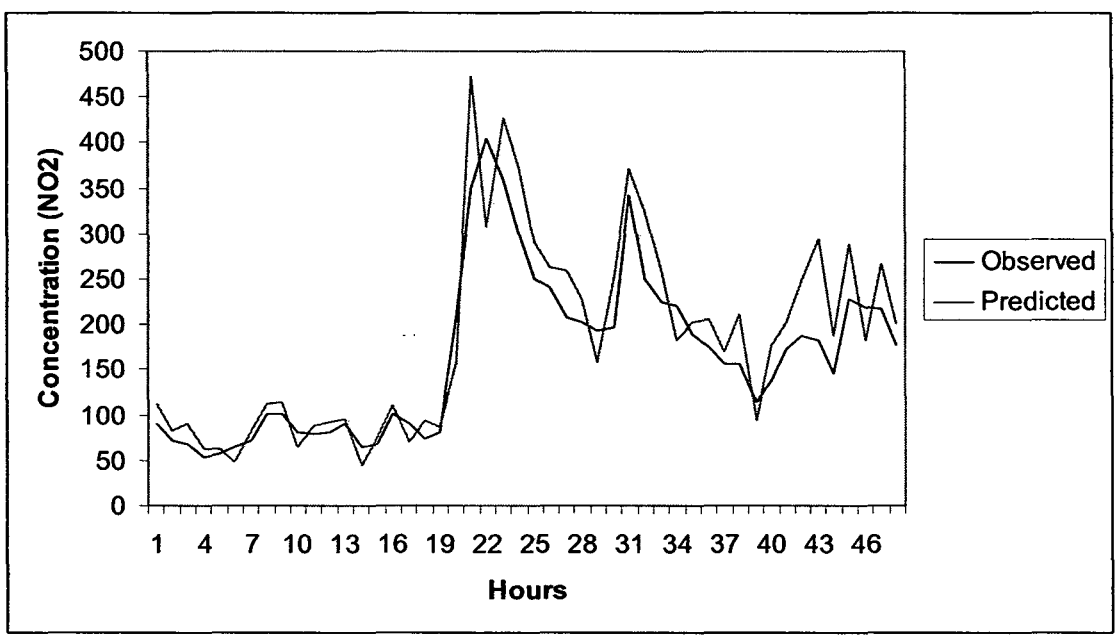


Figure 4.8: *Hourly time series of the observed and predicted concentration of NO_2 (μgm^{-3}) at ITO from 1st to 2nd Feb. 2008.*

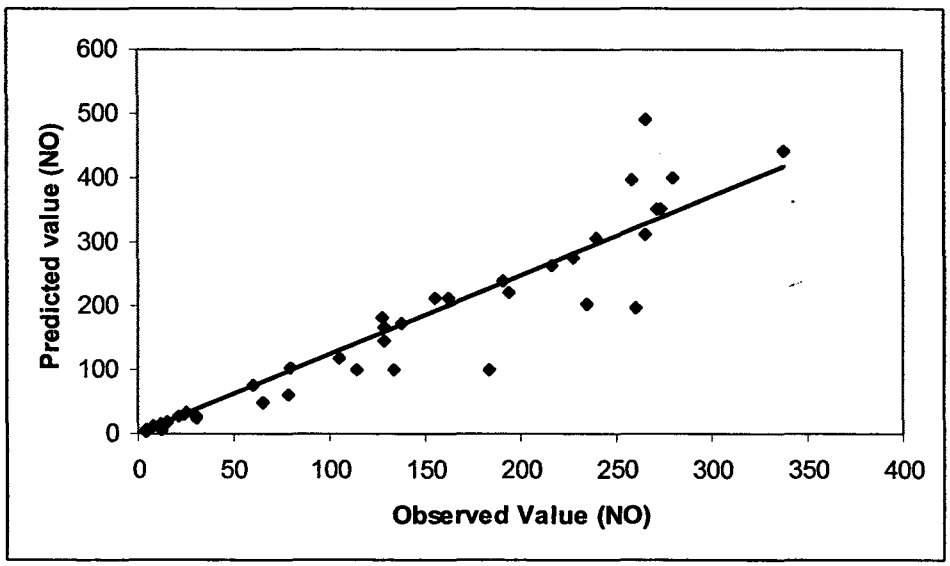


Figure 4.9: *Scatter diagram of predicted versus observed value of NO at Siri Fort from 19th to 20th Feb. 2008. Value of* $R^2 = 0.87$. .

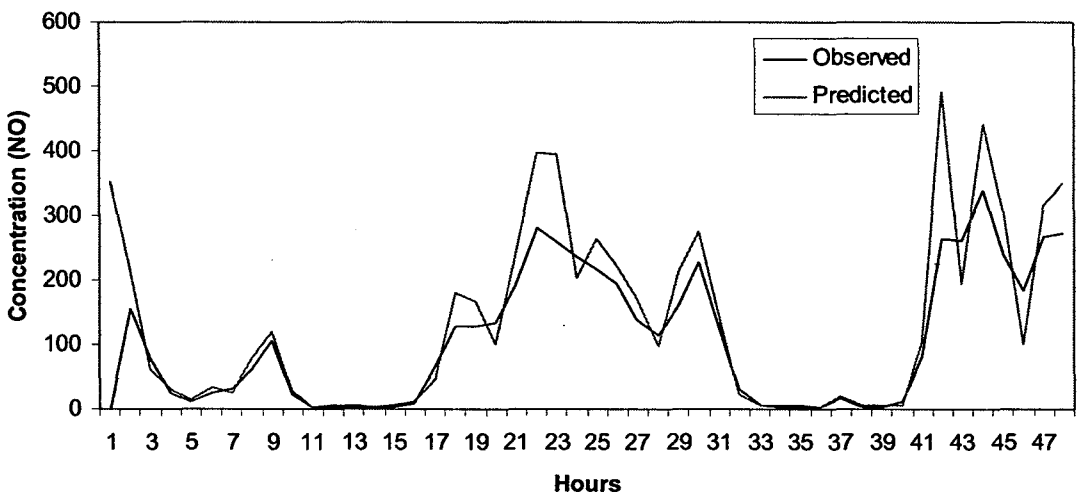


Figure 4.10: *Hourly time series of the observed and predicted concentration of NO* (μgm^{-3}) *at Siri Fort from 19th to 20th Feb. 2008.*

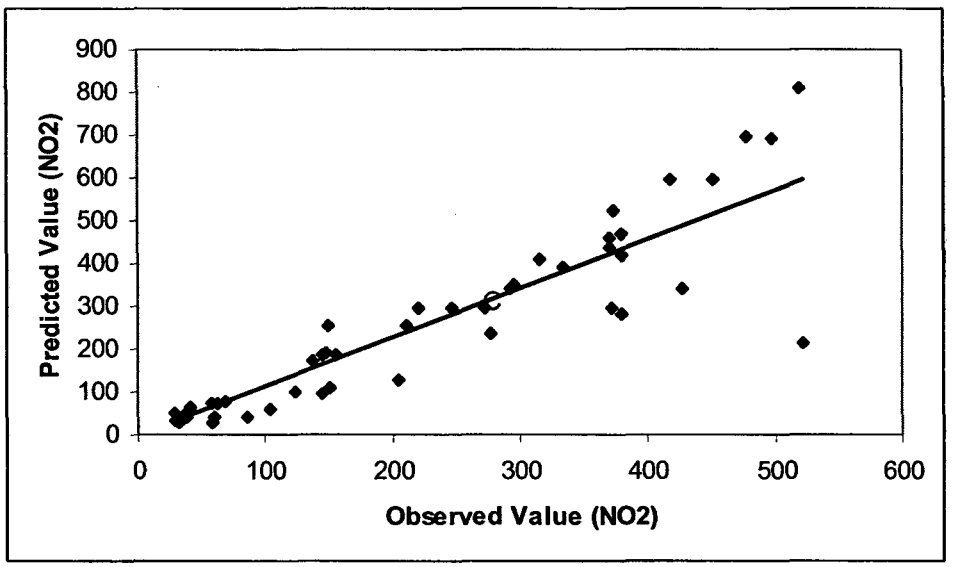


Figure 4.11: *Scatter diagram of predicted versus observed value of NO_2 at Siri Fort from 19th to 20th Feb. 2008. Value of R^2= 0.82 .*

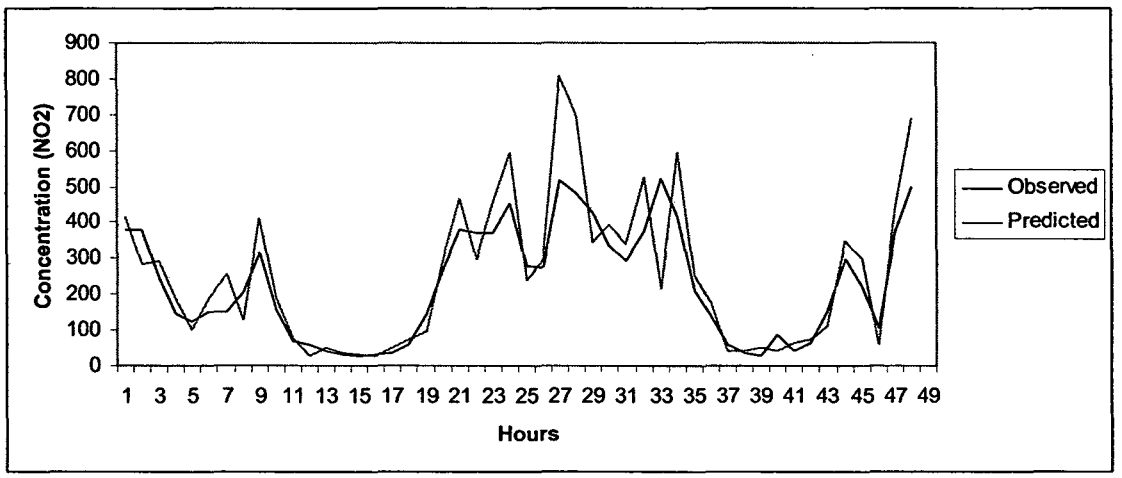


Figure 4.12: *Hourly time series of the observed and predicted concentration of NO_2 (μgm^{-3}) at Siri Fort from 19th to 20th Feb. 2008.*

CHAPTER – 5

Chapter 5 – Artificial Neural Network Modeling with Principal Component Analysis based Radial Basis Function Approach

[5.1] Introduction

This chapter deals with the application of Principal Component Analysis (PCA) in the implementation of Radial Basis Function Neural Network. Several studies on neural networks reported the problem due to large input space (Curse of Dimensions). As most of the operations are matrix based, a large size of matrix creates a lot of complexity. Moreover, we do not have any knowledge regarding percentage of contribution of each input in the behaviour of the system Therefore it becomes necessary to apply PCA technique to our data set to observe any improvement in the efficiency of model performance. PCA is a technique for reducing the dimensions of the input data set.

PCA transform a number of correlated variables into a smaller number of uncorrelated variables called principal components.

In PCA, matrix approach can be used to calculate principal components through a direct computation of eigenvalues and eigenvectors of the covariance matrix (Thurston and Spengler, 1985; Harrison et al., 1996; Kaminski et al., 2000).Due to high dimensionality of input points, RBF network is not much effective as the input points may contain certain unwanted input variables. So using PCA method, the input data points get minimized and orthogonalised. After this operation, then the RBF is applied.

[5.2] Theoretical Background of Principal Component Analysis Method.

Mathematical description of PCA method was described by various authors. (Haykin, 1998). We use Hebbian based learning algorithm. Matrix based approach was applied in formulation of its algorithm for software implementation. We describe here its mathematical representation.

Suppose **X** is an *m*-dimensional random vector.

Let the random vector **X** has zero mean;

$$E[X] = 0$$

Let q denote a unit vector of dimension *m*, onto which the vector **X** is to be projected.

This projection is defined by the product of the vectors **X** and q as shown by

$$A = X^T q = q^T X \tag{5.1}$$

Subject to the constraint

$$\| q \| = (q^T q)^{1/2} = 1$$

The mean value of the projection **A** is zero too,

$$E[A] = q^T E[X] = 0$$

Sot the variance of **A** is same as its mean-square value

$$\begin{aligned} \sigma^2 &= E[A]^2 \\ &= E[(q^T X)(X^T q)] \\ &= q^T E[XX^T] q \\ &= q^T R q \end{aligned}$$

The matrix R is the correlation matrix of size *m*-by-*m* of the random vector **X**, given by $R = E[XX^T]$

As correlation matrix **R** is symmetric, so

$$R^T = R$$

if a and b are any *m*-by-1 vectors, then

$$\begin{aligned} a^T Rb &= b^T Ra \\ \psi(q) &= \sigma^2 \\ &= q^T Rq \end{aligned} \tag{5.2}$$

so $\Psi(q)$ can be considered as a variance probe

[5.2.1] Eigenstructure of principal component analysis

Now to find those unit vectors q along which Ψ (q) has local maxima or minima), subject to a constraint on the Euclidean norm of q. we need to find the correlation matrix. If q is a unit vector such that the variance probe Ψ (q) has an extremal value, then for any small perturbation δq,

$$\psi(q+\delta q) = \psi(q) \tag{5.3}$$

From Eq. (5.2),

$$\begin{aligned} \psi(q+\delta q) &= (q+\delta q)^T R(q+\delta q) \\ &= q^T Rq + 2(\delta q)^T Rq + (\delta q)^T R\delta q \end{aligned}$$

Ignoring the second –order term $(\delta q)^T R\,\delta q$

This may be written as

$$\begin{aligned} \psi(q+\delta q) &= q^T Rq + 2(\delta q)^T Rq \\ &= \psi(q) + 2(\delta q)^T Rq \end{aligned} \tag{5.4}$$

From Eq. (5.3) in (5.4) implies that

$$(\delta q)^T Rq = 0$$

we use only those perturbations for which the Euclidean norm of the perturbed vector q + δq remain equal to unity; that is

$$\| q + \delta q \| = 1$$

$$(q + \delta q)^T (q + \delta q) = 1$$

from Eq. (5.1), we require that to a first order in δq,

$$(\delta q)^T q = 0$$

so perturbation δq must be orthogonal to q,

$$(\delta q)^T Rq - \lambda (\delta q)^T q = 0$$

$$(\delta q)^T (Rq - \lambda q) = 0 \qquad (5.5)$$

from Eq. (5.5) to hold, it is necessary and sufficient to have

$$Rq = \lambda q \qquad (5.6)$$

This is the equation that governs the unit vectors q for which the variance probe $\Psi(q)$ has local maxima or minima values.

Equation (5.6) an eigenvalue problem. Now to calculate Eigen value of the above mentioned equation.

Let the eigenvalues of the m-by-m matrix $\boldsymbol{R}$ be denoted by $\lambda_1, \lambda_2, \ldots\ldots \lambda_m$ and the associated eigenvectors be denoted by $q_1, q_2, \ldots\ldots, q_m$ respectively.

writing the matrix eqn.

$$Rq_j = \lambda_j q_j, \qquad j = 1,2,\ldots,m \qquad (5.7)$$

after solving this we get the eigen value.

Now arranging the corresponding eigenvalues in decreasing order:

$$\lambda_1 > \lambda_2 > \ldots > \lambda_j > \ldots > \lambda_m$$

so that $\lambda_1 = \lambda_{max.}$ Let the associated eigenvectors be used to construct an *m*-by-*m* matrix:

$$Q = [q_1, q_2, ..., q_j, ..., q_m]$$

We may then combine the set of *m* equations represented in (5.6) into a single equation:

$$RQ = Q\Lambda \tag{5.8}$$

where is a diagonal matrix defined by the eigenvalues of matrix R

$$\Lambda = diag[\lambda_1, \lambda_2, ..., \lambda_j, ..., \lambda_m]$$

The matrix **Q** is an orthogonal unitary matrix

The eigenvectors of R satisfy the conditions of Orthonormality:

$$q_i^T = \begin{cases} 1, & j=i \\ 0, & j \neq i \end{cases}$$

so equation is

$$Q^T Q = 1$$

$$Q^T = Q^{-1}$$

Writing Eq. (5.8) in orthogonal similar transformation form:

$$Q^T RQ = \Lambda \tag{5.9}$$

$$q_j^T R q_k = \begin{cases} \lambda_j, & j=j \\ 0, & k \neq j \end{cases}$$

The orthogonal similarity (unitary) transformation of Eq. (5.9) transforms the correlation matrix **R** into a diagonal matrix of eigenvalues.

The correlation matrix R is represented by

$$R = \sum_{i=1}^{m} \lambda_i q_i q_i^T \tag{5.10}$$

the outer product $q_i q_i{}^T$ is of rank 1 for all i .

Equations (5.9) and (5.10) are two equivalent representation of the eigen decomposition of the correlation matrix R.

Principal component analysis and eigen decomposition of matrix R are the same. From Eqs. (5.2) and (5.10) we get that the variance probes and eigenvalues are equal, as

$$\psi(q_i) = \lambda_j, \qquad j = 1,2,...,m$$

The eigenvectors of the correlation matrix R represents the principal directions along which the variance probes $\Psi(q_j)$ have their extrema.

[5.3] Development of the Model

Having preprocessed the data for principal components, the methodology as outlined in section 4.3 is followed

Preparation of Data

(iv) The entire data set of winter season for year 2006 to 2008 was randomly divided into training and testing set. 70% of data was taken for training, 20% for testing and 10% for predication.

(v) Normalization – Further input data was normalized between 0 and 1. This was done using the formula

$$x_i = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

Maximum Normalization value was taken as 0.9 and minimum normalization value was taken as 0.1

Model inputs – At Siri Fort site, the number of input data consists of air pollutants (NO, SO_2, O_3, NO_2, CO). After applying PCA, the dimensions of the input points get reduced from 5 to 3 corresponding to air pollutants only. The model that performed best was retained. The maximum variation in input value was applied so that the models retain maximum information about the input source. For ITO site, the no of input variables among Air pollutant data get reduced to four.

After selecting the model inputs the remaining procedural steps for training the model viz., Network structure, Activation function, Training were the same as mentioned in section (4.3) in the earlier chapter 4.

Stopping criteria was applied to get the acceptable results and then the model was validated with the available data set. This way best model was selected for prediction purpose.

Optimal structure - The architecture corresponding to the minimal RMSE on the validation set is chosen as the optimal structure.

Performance indices – The model performance was evaluated by the following statistical criteria, namely, the determination coefficient R^2, the root mean square error RMSE , Fractional Bias and the index of agreement.

[5.4] Representation of model performance parameters

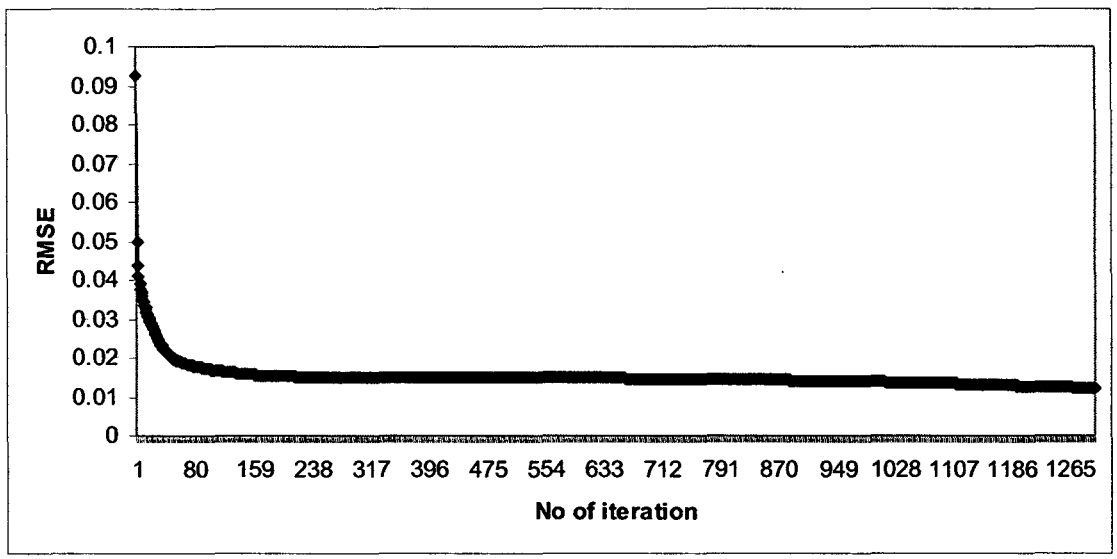


Figure 5.1 *RMSE during training the network of NO at Siri Fort*

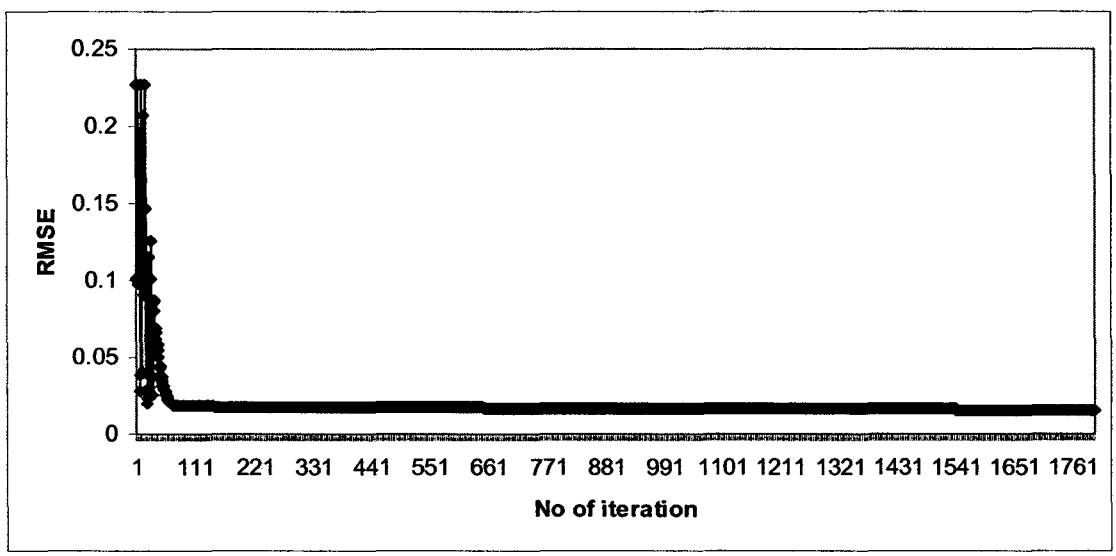


Figure 5.2 *RMSE during training the network of NO_2 at Siri Fort*

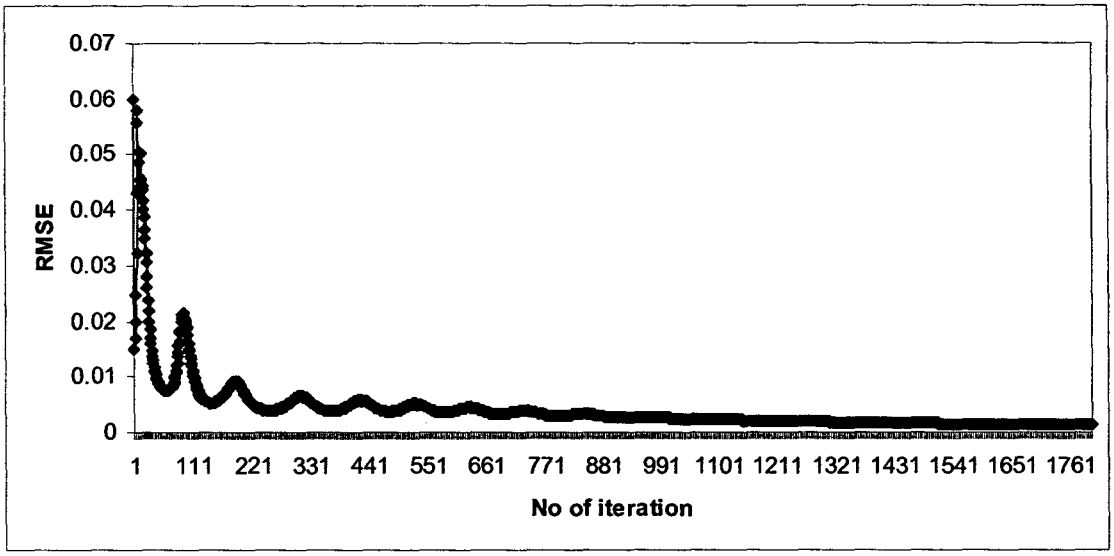


Figure 5.3 *RMSE during training the network of NO at ITO*

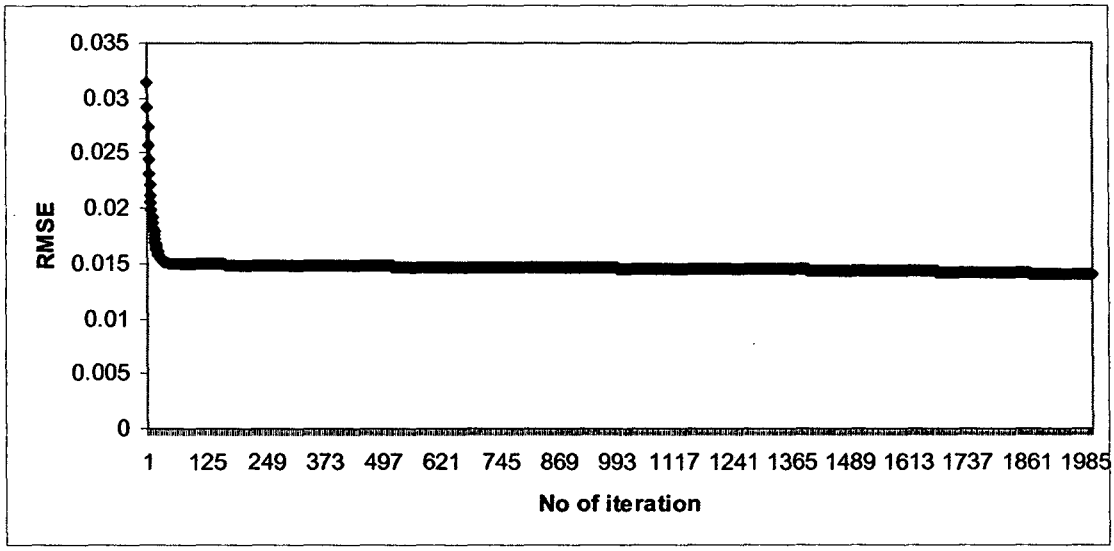


Figure 5.4 *RMSE during training the network of NO at ITO*

Table 5.1: *Network architecture of RBF/PCA network at ITO site*

Model Parameters	NO	NO_2
No. of inputs	5	5
Processing time	390s	330 s
No. of layers	1	1
No. of nodes in hidden layer 1	12	10
No. of nodes in hidden layer 2		
Error tolerance	10^{-12}	10^{-12}
Momentum constant	0.5	0.6
No. of iteration	1800	1300
Learning rate	0.7	0.7

Table 5.2: *Network architecture of RBF/PCA network at Siri Fort site*

Model Parameters	NO	NO_2
No. of inputs	4	4
Processing time	325 s	400 s
No. of layers	1	1
No. of nodes in hidden layer 1	10	09
No. of nodes in hidden layer 2		
Error tolerance	10^{-12}	10^{-12}
Momentum constant	0.5	0.6
No. of iteration	1300	1800
Learning rate	0.7	0.7

Table 5.3: *Forecasting result of air pollutants NO and NO_2 at ITO site.*

No. of Days	NO (% of error)	NO_2 (% of error)
1.	38.48	28.57
2.	29.4	15.5
3.	27.11	25.54
4.	15.64	31.15
5.	23.51	33.92
6.	20.77	31.05
7.	48.14	10.28
8.	24.04	23.47
9.	56.06	10.34
10.	30.77	12.72
11.	13.66	11.28
12.	31.87	24.42
13.	28.71	14.67
14.	44.97	38.4
15.	31.66	11.35
16.	20.63	24.59
17.	26.97	15.05
18.	14.98	20.72
19.	31.4	42.95
20.	19.06	28.17
21.	77.95	12.65
22.	18.4	10.39
23.	28.03	24.65
24.	38.61	23.24
25.	23.83	36.9
26.	14.92	28.32
27.	21.94	16.63
28.	47.28	29.04
29.	25.27	7.28
30.	25.17	16.94
31.	56.3	9.78
32.	28.02	32.6
33.	20.31	11.36
34.	28.14	6.75
35.	12.73	17.35
36.	22.55	20.10
37.	24.09	12.68
38.	38.06	19.78
39.	19.40	12.95
40.	24.5	29.16
41.	15.64	17.55
42.	27.32	14.51
43.	26.52	15.58
44.	17.22	26.04
45.	31.11	17.88
46.	28.31	5.11
47.	41	10.75

Table 5.4: *Forecasting result of air pollutants NO and NO_2 at Siri Fort*

No. of Days	NO (% of error)	NO_2 (% of error)
1.	17.36	8.94
2.	30.16	26.36
3.	11.61	13.37
4.	-35.77	-40.27
5.	49.5	25.32
6.	-12.4	-38.9
7.	14.12	19.61
8.	42.87	34.13
9.	-38.65	-27.86
10.	69.50	41.82
11.	-12.54	26.50
12.	24.21	35.58
13.	12.43	24.39
14.	-26.50	-10.42
15.	16.8	30.65
16.	24.12	-59.98
17.	-14	48.55
18.	35.53	30.37
19.	52.72	-58.96
20.	-10.41	12.4
21.	17.98	24.33
22.	13.94	-2.32
23.	42.19	23.8
24.	-37.97	32.02
25.	23.67	33.17
26.	19.84	-21.31
27.	10.18	56.15
28.	-35.94	-46.56
29.	17.94	22.21
30.	53.76	-19.49
31.	58.24	11.42
32.	-10.18	40.41
33.	49.41	58.86
34.	8.49	-43.21
35.	16.51	20.09
36.	13.98	31.58
37.	-26.89	18.23
38.	21.71	-17.38
39.	24.5	13.73
40.	12.49	40.85
41.	40.52	-44.81
42.	-76.21	28.74
43.	49.34	25.57
44.	12.18	19.69
45.	-23.72	-32.96
46.	32.18	17.71
47.	12.43	-15.81
48.	-26.5	37.30

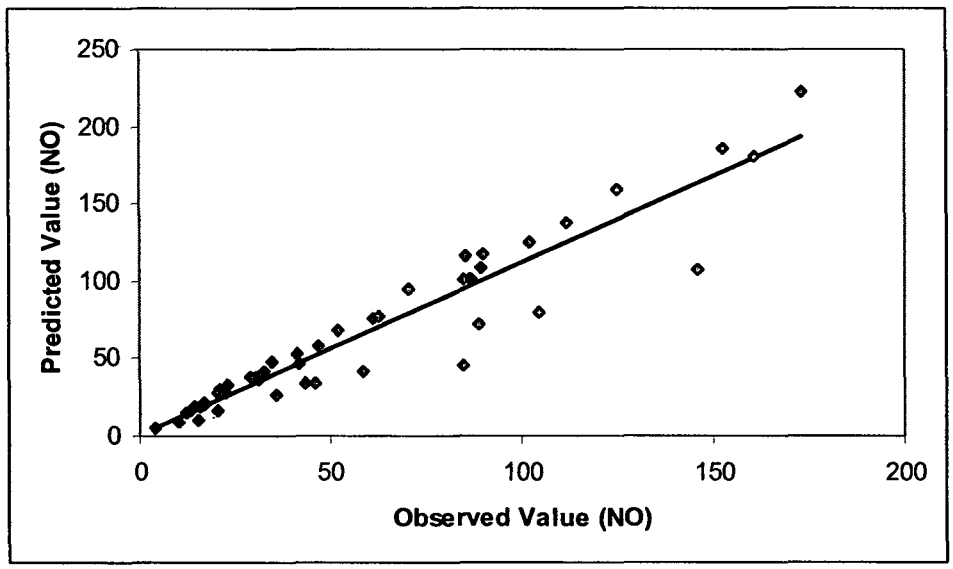


Figure 5.5: *Scatter diagram of predicted versus observed value of NO at ITO from 1st to 2nd Feb. 2008. Value of* $R^2 = 0.85$.

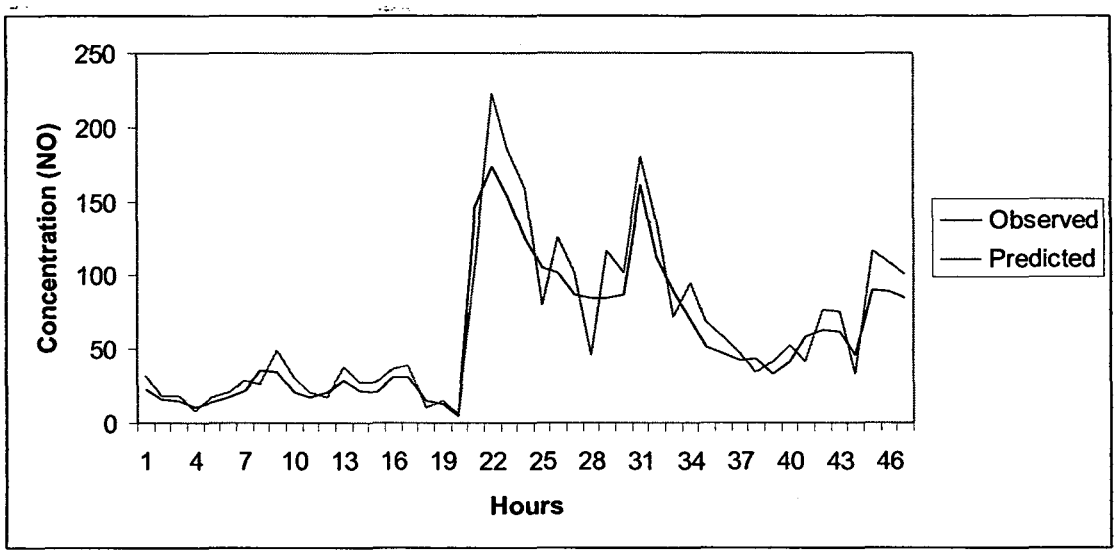


Figure 5.6: *Hourly time series of the observed and predicted concentration of NO* (μgm^{-3}) *at ITO from 1st to 2nd Feb. 2008.*

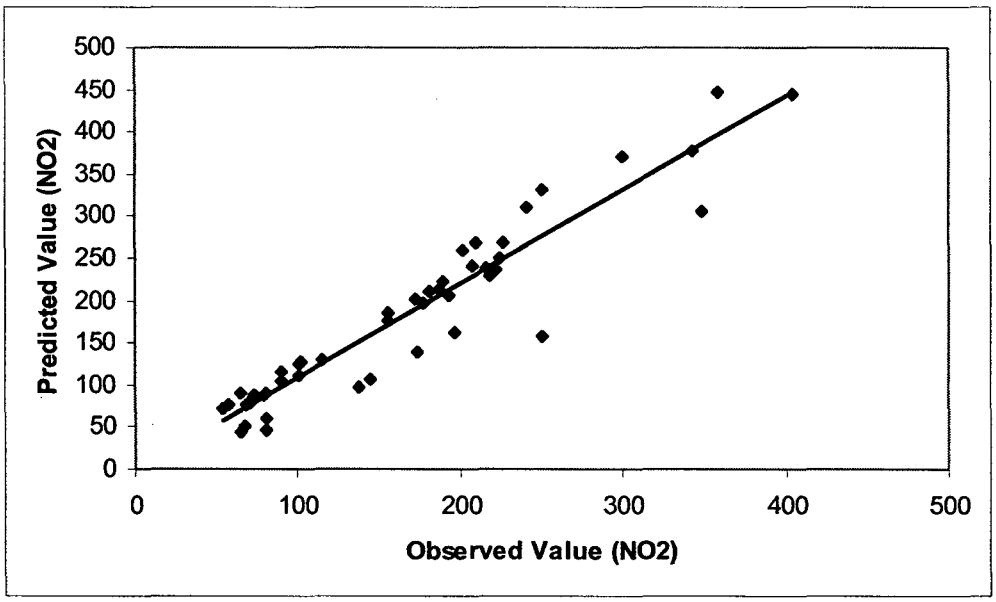


Figure 5.7: *Scatter diagram of predicted versus observed value of NO_2 at ITO from 1st to 2nd Feb. 2008. Value of R^2= 0.86 .*

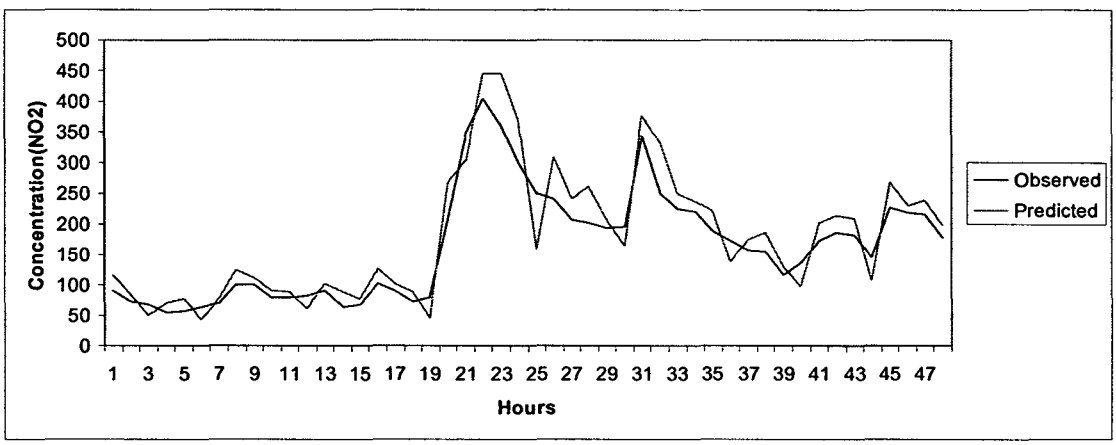


Figure 5.8: *Hourly time series of the observed and predicted concentration of NO_2 (μgm^{-3}) at ITO from 19th to 20th Feb. 2008.*

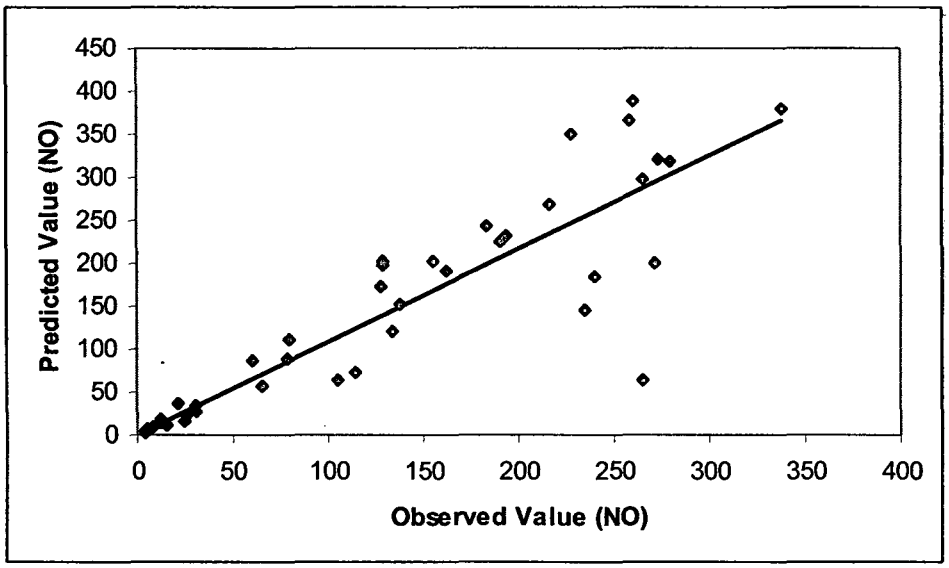


Figure.5.9: *Scatter diagram of predicted versus observed value of NO at Siri Fort from 19th to 20th Feb. 2008. Value of* $R^2 = 0.81$.

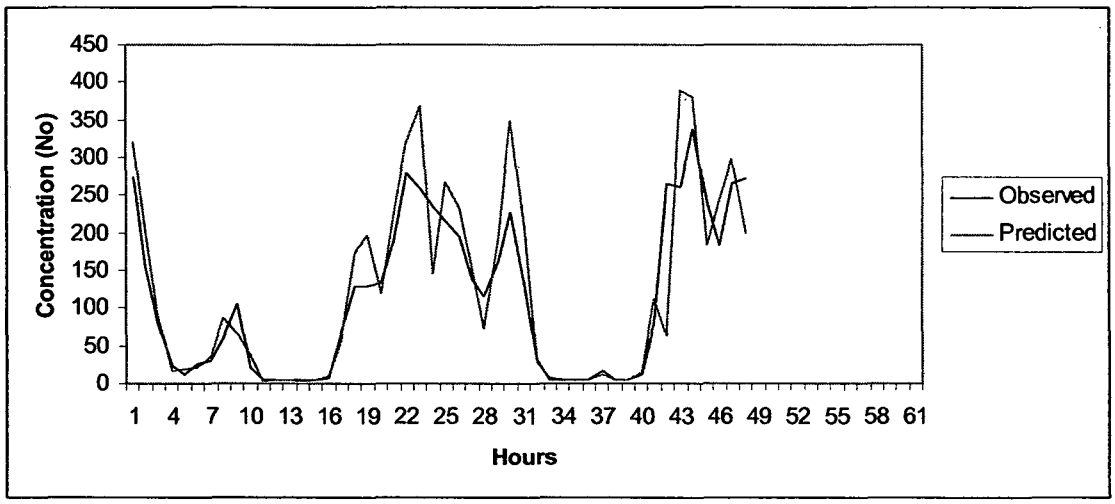


Figure 5.10: *Hourly time series of the observed and predicted concentration of NO* (μgm^{-3}) *at Siri Fort from 19th to 20th Feb. 2008.*

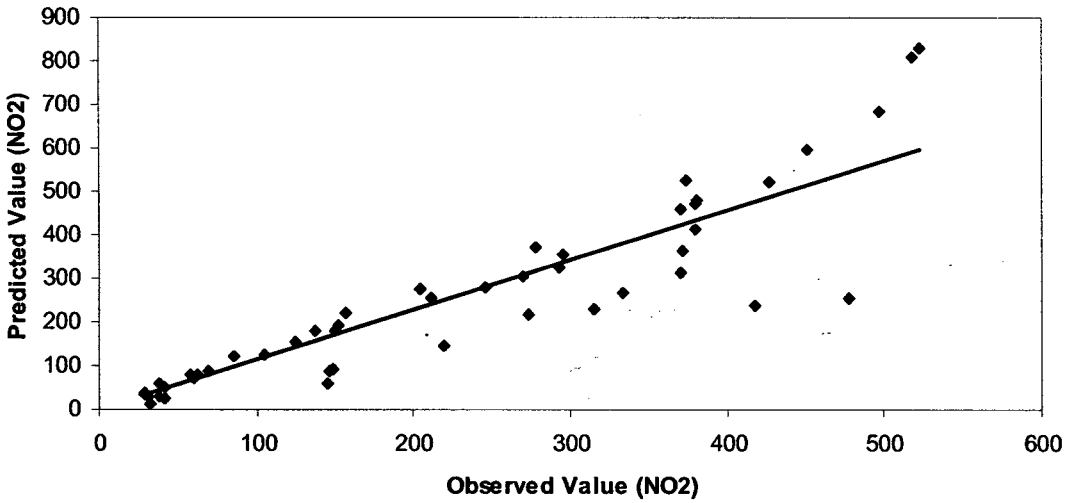


Figure 5.11: *Scatter diagram of predicted versus observed value of NO_2 at Siri Fort from 19th to 20th Feb. 2008. Value of R^2= 0.79*

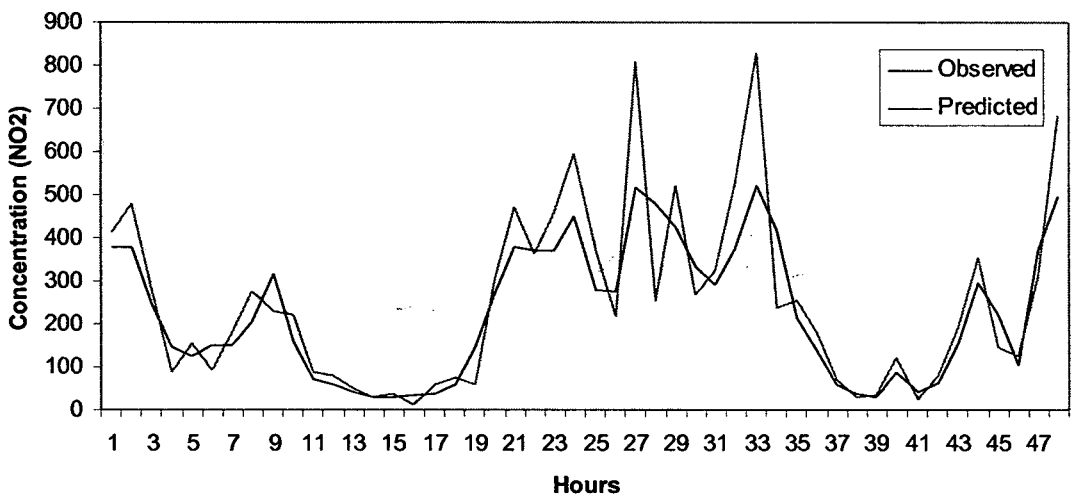


Figure 5.12: *Hourly time series of the observed and predicted concentration of NO_2 (μgm^{-3}) at Siri Fort from 19th to 20th Feb. 2008.*

CHAPTER – 6

Chapter 6 – Performance Evaluation of the Developed Forecasting Models

Performance of the developed models based on three different approaches namely, ANN based on feedforward back propagation, ANN based on RBF and ANN based on RBF in conjunction with PCA, has been evaluated in terms of the network architecture, time taken to train the model and finally the error in forecasting.

Tables 6.1 and 6.2 show the comparison of the models based on the above mentioned three criteria. It can be seen from the Table 6.1 that both for Siri Fort and ITO, the network architecture of RBF is the simplest as indicated by the number of nodes and hidden layers. RBF network is having only one hidden layer while feedforward back propagation network is having two hidden layers. It may be noted that the number of nodes are least in case of RBF network at both the sites. The number of nodes and hidden layers of RBF at ITO are also minimum viz. 12 and 10 respectively. Further it is to be noted that architecture is simple in case of Siri fort site for feedforward network as well as RBF network as compared to that at ITO site.

A comparison of the time taken to train the models as given in Table 6.2 shows that all the models took almost the same time (~6 minutes) for training the network. It may be mentioned here that the time taken to train the model is the measure of the number of iterations needed for the convergence (stabilization) of the RMSE. The table 6.3, 6.4 and 6.5 shows the value of index of agreement (IA), the coefficient of determination (R^2) and fractional bias (FB) for all the models respectively. The results depicted here is quite satisfactory.

An examination of Table 6.5 where RMSE values in respect of various modeling approach are depicted, shows that although the model performance of both feedforward back propagation and RBF approaches is quite satisfactory, but it is the RBF network which yields minimum RMSE values for forecasting. Therefore forecasting accuracy is more in case of RBF network compared to feedforward back propagation network.

The above details have been listed in the adjoining tables given below.

Table 6.1: *A comparison of various modeling approaches in terms of network architecture*

Site	Air Pollutant	Feedforward Network			RBF Network		PCA/RBF Network	
		No. of Hidden Layers	No. of Nodes		No. of Hidden Layers	No. of Nodes	No. of Hidden Layers	No. of Nodes
			1st Layer	2nd Layer				
Siri Fort	NO	2	9	5	1	12	1	10
	NO_2	2	8	5	1	14	1	9
ITO	NO	2	10	6	1	12	1	12
	NO_2	2	9	6	1	10	1	10

Table 6.2: *A comparison of various modeling approaches in terms of time taken to train the model*

Site	Air Pollutant	Feedforward Network	RBF Network	PCA/RBF Network
		CPU Time	CPU Time	CPU Time
Siri Fort	NO	375	390	325
	NO_2	345	405	400
ITO	NO	480	380	390
	NO_2	455	325	330

Table 6.3: *A comparison of various modeling approaches in terms of IA*

Site	Air Pollutant	Feedforward Network	RBF Network	PCA/RBF Network
		IA Value	*IA Value*	*IA Value*
Siri Fort	NO	0.86	0.94	0.92
	NO_2	0.86	0.89	0.88
ITO	NO	0.89	0.94	0.94
	NO_2	0.86	0.94	0.94

Table 6.4: *A comparison of various modeling approaches in terms of* R^2 Value

Site	Air Pollutant	Feedforward Network	RBF Network	PCA/RBF Network
		R^2 Value	R^2 Value	R^2 Value
Siri Fort	NO	0.79	0.87	0.81
	NO_2	0.74	0.82	0.79
ITO	NO	0.84	0.89	0.85
	NO_2	0.80	0.88	0.86

Table 6.5: *A comparison of various modeling approaches in terms of RMSE*

Site	Air Pollutant	Feedforward Network	RBF Network	PCA/RBF Network
		% of Av. RMSE	% of Av. RMSE	% of Av. RMSE
Siri Fort	NO	26.94	24.75	27.92
	NO_2	29.45	26.25	29.45
ITO	NO	26.15	22.94	22.18
	NO_2	28.65	19.82	20.12

Table 6.6: *A comparison of various modeling approaches in terms of Fractional Bias* Value

Site	Air Pollutant	Feedforward Network	RBF Network	PCA/RBF Network
		FB Value	FB Value	FB Value
Siri Fort	NO	0.08	0.18	0.09
	NO_2	0.14	0.11	0.10
ITO	NO	0.06	0.22	0.085
	NO_2	0.10	0.11	0.086

Conclusion

The results of the present study clearly demonstrate that Artificial Neural Network modeling can be used as an important tool for air quality forecasting. Winter season was chosen for our work as this season remains more polluted. All models performed well. The index of agreement, the coefficient of determination, RMSE and fractional bias were computed and the values obtained have been shown in the tables 6.3, 6.4, 6.5 and 6.6 respectively. The results obtained are within accepted limit for developing the model. Among them, the RBF model performed successfully for winter season for both the residential as well as sensitive sites. Therefore, it would be interesting to see whether RBF networks can provide accurate forecasts for other categories of sites (areas) and seasons.

Since we were having hourly average data of air pollutants only, applying PCA did not improve model performance significantly. May be for high dimensional input variables including meteorological parameters it may be of much importance.

It would be worthwhile to test the application of various other algorithms and optimization methods so that more accuracy can be obtained for forecasting. Since ANN works on pattern recognition approach, it sometimes gives unacceptably large error.

Scope of Future Work

The present study describes, unambiguously, the importance of RBF networks for the purpose of forecasting future values of air pollutants, especially NO and NO_2, in ambient atmosphere of Delhi. In the present case only the air pollutants' concentration data were taken as input parameters to forecast NO and NO_2 concentrations. Due to non availability of meteorological data at the monitoring sites, these data were not included. Since the influence of meteorological variables on the concentrations of ambient air pollutants vary with space and time, it would be interesting to study the performance of RBF networks vis-à-vis other ANN models at different locations and during different seasons taking into account the meteorological variables. Similarly, the inclusion of emission factors in the modeling exercise may improve the model performance. Therefore it would be worthwhile including meteorological parameters and emission factors in a future study. Further I would like to examine the RBF networks with different advanced learning algorithms such as genetic algorithm, Bayesian algorithm etc.

Recently, different dynamical and static network architectures are being proposed for forecasting purposes in sequential as well as batch training mode. It would be worthwhile to examine the performance of Radial Basis Function, when integrated with the existing network architectures such as gamma function network, Bi-directional recurrent network, Bi-directional adaptive networks etc. Optimization needs to be done using various algorithms.

REFERENCES

Astrup, P., (1972). Some Physiological and Pathological effects of moderate CO exposure. *Brit Med J*, 4, 447-52.

Atkinson, R., (2000).Atmospheric Chemistry of VOCs and NOx. *Atmospheric Environment*, 34, 2063–2101.

Badran, F. And Thiria, S. (1991). Wild ambiguity removal by the use of neural network techniques. *Journal of Geophysical research,* 96(C 11), 20521-20529.

Bankert, R. L. (1994). Cloud classification of AVHHR imagery in maritime regions using a probabilistic neural network. *Journal of Applied Metrology*, 33, 909-918.

Barlow, H. B., (1989). "Unsupervised learning," *Neural Computation*, 1, 295–311.

Bascom, R., Bromberg, P.A. and Costa, D.L., (1996). Health effects of outdoor air pollution. *Am J Respir Crit Care Med*, 153, 477-498.

Battiti, R. (1992). First- and second- order methods for learning between steepest descent and Newton's method. *Neural Computation*, 4, 141-166.

Becker, S. and Hinton, G.E., (1982). "A self-organizing neural network that discovers surfaces in random-dot stereograms." *Nature (London)*, 355, 161–163.

Bellman, R.E. (1961). *Adaptive Control Processes*. Princeton University Press, Princeton, NJ.

Benediktsson, J. A., Swain, P. H. and Ersoy, O. K., (1990). Neural network approaches versus statistical methods in classification of multi source remote sensing data. *IEEE Transaction on Geosciences and Remote Sensing*, 28 (4), 540-552.

Benton, J., fuhrer, J., Gimeno, B. S., Skarby, L. And Sanders, G. (1995). Results from the UN/ ECE ICP – crops indicate the extent of exceedance of the critical levels of ozone in Europe. *Water, Air and Soil Pollution*, 85, 1473-1478.

Bishop, C. M. (1995). *Neural networks for pattern Recognition.* Clarendon Press, Oxford. Box, G.E.P. and Jenkins, G.M. Time series analysis: forecasting and control. *Holden-Day. S. Francisco*: ISBN 0-8162-1104-3.

Boznar, M., Lesjak, M. and lakar, A., (1993). Neural network-based method for short-term predictions of ambient SO_2 concentrations in highly polluted industrial areas of complex terrain. *Atmos. Environ.*, 27B (2), 221-230.

Boznar, M., Lesjak, M. and Mlakar, P. (1993). A neural network- based methods for the short-term predictions of ambient So_2 concentrations in highly polluted industrial areas of complex terrain. *Atmospheric Environmental*, B 27 (2), 221-230.

Breiman, L., (1994).Comment on "Neural network: A review from a statistical perspective" by B. Cheng and D.M. Titterington. *Stat. Sci.,* 9 (l), 38-42.

Broomhead, D. S. and Lowe, D., (1988). "Multivariable functional interpolation and adaptive networks." *Complex Systems*, 2, 321– 355.

Chakraborty, K., Mehrotra, K., Mohan C.K. and Ranka, S., (1992). Forecasting the behavior of multivariate time series using neural networks. *Neural Networks,* 6, 961-970.

Chameides, William, L. and Walker, C.G., (1976). A Time-Dependent Photochemical Model for Ozone nears the Ground. *Journal of Geophysical Research*, 81(3), 413-420.

Chattopadhyay, G., Samanta, G., Chatterjee, S. and Chakroborti, P., (1997). Determination of benzene, toluene and xylene in ambient air of Calcutta for three years during winter. *Environmental Technology,* 18, 211–218.

Cheng, B. and Titterington, D.M., (1994). Neural network A review from a statistical perspective, *St and Sci.* (l), 2-54.

Chock, D.P., Winkler, S.L. and Chen, C., (2000). A study of the association between daily mortality and ambient air pollutant concentrations in Pittsburgh, Pennsylvania. *Journal of the Air & Waste Management Association,* 50, 1481–1500.

Cifuentes, L.A., Vega, J., Kopfer, K. and Lave, L.B., (2000). Effect of the fine fraction of particulate matter versus the coarse mass and other pollutants on daily mortality in Santiago, Chile. *Journal of the Air & Waste Management Association,* 50, 1287–1298.

Clalr, T.A., and Ehrman, J.M., (1996). Variations in discharge and dissolved organic carbon and nitrogen export from terrestrial basins with changes in climate: A neural network approach, *Limnol. Oceanogr*, 41(5), 921-927.

Comon, P., (1994). "Independent component analysis: A new concept?" *Signal Processing*, 36, 287–314.

Comrie, A. C. (1997). Comparing neural networks and regression models for ozone forecasting. *Journal of Air and Waste Management,* 47, 653-663.

Comrie, A.C., (1997). Comparing neural networks and regression models for ozone forecasting, *J. Air Waste Manage. Assoc.*, 47(6), 653-663.

Connor, J.T., Martin, R.D. and Atlas L.E., (1994). Recurrent neural networks and robust time series prediction. *IEEE tins. Neural Network,* 5 (2), 240-254.

CPCB, (2004). National Ambient Air Quality Status 2002.*Central Pollution Control Board,* Ministry of Environment and Forests, India, p. 2.

Darral, N.M., (1994).The Sensitivity of Net Photosynthesis in Several Plant Species to Short-term Fumigation with Sulphur Dioxide. *Environment International,* 20(2), 166-167.

Del, L.B., Golden, Wang, Q. and Kumar, F.T., (1992) Predicting salinity in the Chesapeake Bay using back propagation, *Comput. Oper. Res,* 19(3/4), 227-285.

Derwent, R.G., (1995). Volatile organic compounds in atmosphere. *Issues in Environmental Science and Technology, 4,* 1–15.

DeVries, B. and Principe, J.C., (1992). "The gamma models a new neural model for temporal processing." *Neural Networks*, 4, 565–576.

Dewulf, J. and Langenhove, H. Van, (1997). Analytical techniques for determination of measurement data of 7 chlorinated C1 and C2 hydrocarbons and 6 monocyclic aromatic hydrocarbons in remote air masses. *Atmospheric Environment,* 31, 3291–3307.

Dockery, D.W. and Pope, C.A., (1994). Acute respiratory effects of particulate air pollution. *Annu Rev Public Health*, 15, 107-132.

Dockery, D.W., Pope, C.A., Xu, X., Splenger, J.D., Ware, J.H., Fay, M.E., Ferris Jr., B.G. and Speizer, F.E., (1993). An association between air pollution and mortality in six US cities. *The New England Journal of Medicine,* 329, 1753–1759.

Doshi, V.B., Patade, V.D. and Shetye, S.V., (1984). Relationship of COHb and ischemic heart diseases; Preliminary study and correlation between COHb levels and health morbidity in normal community residing at a busy road junction. *Lung India*, 68-75.

Dubourg, W.R., (1996). Estimating the mortality costs of lead emissions in England and Wales. *Energy Policy*, 24, 621 – 625.

Duce, R.A., Mohnen, V.A., Zimmerman, P.R., Grosjean, D.,Cautreels, W.J., Chatfield, R., Jaenicke, R., Ogren, J.A.,Pellizzari, E.D. and Wallace, G.T., (1983). Organic material in global troposphere. *Reviews of Geophysical Space Physics,* 21, 921–952.

Edgerton, S.A., Holdren, M.W. and Smith, D.L., (1989). Inter urban comparison of ambient volatile organic compounds concentration. *JAPCA,* 39, 729–732.

Elizindo, D., hoogenboom, G. and McClendon, R. W. (1994). Development of a neural network model to predict daily solar radiation. *Agricultural and Forest Meteorology,* 7, 115-132.

Flood, I. and Kartam, M., (1994) Neural networks in civil engineering. I: Principle and understanding, *J. Comp. Civil Eng.*, 8 (2), 131-148.

Ganguly, S., Sadhukhan, J., Deoki, N. and Anderson, J. A., (1995). Artificial Neural Network Based Estimation of Petroleum and Product Properties, *Introduction to Neural Networks* (Cambridge, MA: MIT Press).

Genqay and Liu, T., (1997). Nonlinear niodelling and prediction with feedforward and recurrent networks, *Physica*, 108 (l/2), 119-134.

Goldberg, M.S., Burnett, R.T. and Valois, M.F., (2003). Associations between ambient air pollution and daily mortality among persons with congestive heart failure. *Environmental Research*, 91, 8–20.

Goldsmith, J.R., (1968). In: Stern, A.C. (Ed.), *Effect of Air Pollution on Human Health in Air Pollution*, vol. 1. Academic Press, New York/London.

Goyal, P. and Sidhartha, (2003). Present scenario of air quality in Delhi: a case study of CNG implementation. *Atmospheric Environment*, 37, 5423-5431.

Haasoun, M.H., (1995).*Fundamentals of Artificial Neural Networks,* MIT Press, Cambridge.

Hall, J.V., (1996). Assessing health effects of air pollution. *Atmospheric Environment,* 30, 743–746.

Handbook on air pollution and health. London: HMSO, 1997.

Hannan, J.M. and Bishop, J.M., (1997). A comparison of fast training algorithms over two real problems In: Proceedings of the fifth international conference *ANN*, Cambridge, UK, 1–6.

Hansen, J.V. and Nelson, R.D., (1997). Neural networks and traditional time series methods: A synergistic combination in state economic forecasts, *IEEE mans. Neural Network*, 8 (4), 863-873.

Harrison, R.M., Smith, D.J.T., Luhana, L., (1996). Source apportionment of atmospheric poycyclic aromatic hydrocarbons collected from urban locations in Birmingham, U.K. *Environ. Sci. Technol*, 30, 825-832.

Harrison, R.M., Deacon, A.R., Jones, M.R., (1997). Sources and processes affecting concentrations of PM10 and PM2.5 particulate matter in Birmngham (UK). *Atmos. Environ*, 31, 4103-4117.

Haykin, S., (1999). *Neural Networks: A Comprehensive Foundation*, 2nd ed. (Englewood Cliffs, NJ: Prentice-Hall).

Health effects Institute: (1995). Diesel exhaust. A critical analysis of emissions, exposure and health effects, Special report.

Hertz, J.A., Krogh, A. and Palmer, R.G., (1991). *Introduction to the Theory of Neural Computation,* Addison-Wesley,Redwood City.

Highlights, (2004), *Parivesh,* Central pollution control board, MOEF.

Hoffer, T.E., Farber, R.J. and Ellis, E.C., (1982). Background continental ozone levels in the rural US southwestern desert. *Sci. Total Env,* 23, 17–30.

Hooyberghs, J., Mensink, C., Dumont, G., Fierens, F. and Brasseur O., (2005).A neural network forecast for daily average PM10 concentrations in Belgium. *Atmospheric Environment,* 39(18), 3279–89.

Hunt Jr., W.F., Faoro, R.B. and Freas, W., (1998). Report of the Interim Database for State and local Air Toxic Volatile Chemical Measurements. EPA-450/4-86-012, *US Environmental Protection Agency,* Washington DC.

Ioannis, C., Ziomas, Dimitrios M., Christos, S., Zerefos, Alkiviadis, F. and Hanasios, G., (2002). Forecasting peak pollutant levels from meteorological variables. *Atmospheric Environment*, 36(27), 4323-4335.

Jia-Yeong, K. and Rao T.V., (1994). Numerical simulation of air pollution in urban areas: Model development. *Atmospheric Environment,* 28(1), 149-158.

Kamat, S.R. and Doshi, V.B., (1987).Sequential health effect study in relation to air pollution in Bombay, India. *Eur J Epidemiol*, 3, 265277.

Kamat, S.R., (1999). Will pollution soon be mankind's biggest killer. *Medisite*, 1, 14-19.

Kamat, S.R., Doshi, V.B., Patade, V.D. and Naik, M., (1984). Third year analyses on regularly followed sample of Bombay air pollution study population and correlation with other factors. *Lung India*, 2, 110-130.

Kamat, S.R., Godkhindi, K.D. and Shah, V.N., (1980). Correlation of health morbidity to air pollution levels in Bombay city. Results of prospective study at one year. *J Postgrad Med*, 26, 45-62.

Kamat, S.R., Godkhindi, K.D. and Shah, V.N., (1984). Prospective 3 year study of health morbidity in relation to air pollution in Bombay, India: I: Methodology and early results up to 2 years. *Lung India*, 2, 120.

Kaminski, W., Skrzypski, J., Strumillo, P., (2000). Forecasting of air pollution in urban areas by means of artificial neural networks. In: Sucharov, L.J. (Ed.), *Urban Transport and the Environment for the 21st Century*. WIT Press, Southampton, Boston, pp. 114-124.

Karim, M.and Masud, O., (2000). Takashi Air quality planning and empirical model to evaluate SPM concentrations. *Journal of Environmental Engineering*, 126(12), 1116-1124.

Kelsall, J.E., Samet, J.M., Zeger, S.L. and Xu, J., (1997). Air pollution and mortality in Philadelphia, 1974–1988. *American Journal of Epidemiology*, 146, 750–762.

Khotanzad, A, Davis, M.H. and Abaye, A., (1996). An artificial neural network hourly temperature forecaster with applications in load forecasting. *IEEE Transaction on Power Syste*m, 11(2).

Kleeberger, S.R., Longphre, M. and Tankersley, C.G., (1999). Mechanism of response to 0_3 exposures: the role of mast cells in mice. *Health Effects Inst* Report No 85.

Koskela, T., Lehtokangas, M., Saarinen, J. and Kaski, K., (1996). Time series prediction with multilayer perceptron, FIR and Elman neural networks. In: Proceedings of the world congress on neural networks. *INNS Press*, 491–496.

Kostianen, R., (1995). Volatile organic compounds in the indoor air of normal and sick houses. *Atmospheric Environment*, 29 (6), 693–702.

Kremer, S.C., (2001). Spatio-temporal connectionist networks: a taxonomy and review. *Neural Computation*, 13(2), 249–306.

Kuran, P. and Sojak, L., (1996). Environmental analysis of volatile organic compounds in water and sediments by gas chromatography. *Journal of Chromatography A.*, 733, 119–141.

Lloyd, S.A., (1993) stratospheric ozone depletion. *Lancet*, 342, 1156 – 1158.

Luk, K.C., Ball, J.E. and Sharma, A., (2000). A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting. *Journal of Hydrology*, 227, 56–65.

Maier H.R. and Dandy, D.C., (1996).The Use of artificial neural networks for the prediction of water quality parameters, *Water and Resources, 32* (4), 1013-1022.

Maqsood, I., RiazKhan M. and Abraham, A., (2002). Weather forecasting models using ensembles of neural networks. In Third international conference on intelligent systems design and applications, intelligent systems design and applications, advances in soft computing. *Germany Springer*, 33–42.

Maren, A. Haon C. and Pap, R., (1990). *Handbook of Neural Computing Applications,* Academic Press, San Diego, CA.

Martin, B. and Richard, G.A., (1999). Air pollution and mortality in Central and Eastern Europe, An estimate of the impact. *European Journal of Agronomy*, 10(3), 185-195.

Masters, T., (1993). *Practical Neural Network Recipe* in C++, Academic Press, San Diego.

McCollister, G.M. and Wilson, K.R., (1974). Linear stochastic models for forecasting daily maxima and hourly concentrations of air pollutants. *Atmospheric Environment*, 9, 416–23.

Melia, J., Ellman, R. and Chamberlain, J., (1994). Meeting the health of the nation target for skin cancer: problems with tackling prevention and monitoring trends. *Journal of Public Health Medicine,* 16, 225 – 232.

Melia, J., Ellman, R. and Chamberlain, J., (1994). Meeting the health of the nation target for skin cancer: problems with tackling prevention and monitoring trends. *Journal of Public Health Medicine*, 16, 225 – 232.

Mohan Rao, A.M., Pandit, G.C., Sain, P., Sharma, S., Krishnamoorthy, T.M. and Nambi, K.S.V., (1997). Non methane hydrocarbons in industrial locations of Mumbai. *Atmospheric Environment,* 31 (7), 1077–1085.

Mondal, R., Sen, G.K., Chatterjee, M., Sen, B.K. and Sen, S., (1984). Ground-level concentration of nitrogen oxides (NO_x) at some traffic intersection points in Calcutta. *Atmospheric Environment,* 18(1), 29-39

Moody and Darken, (1989). "Fast learning in networks of locally-tuned processing unites," *Neural Computation*, 1, 281–294.

Moolgavkar, S.H., (2003). Air pollution and daily mortality in two US counties: season-specific analyses and exposure–response relationships. *Inhalation Toxicology*, 15, 877–907.

Mukund, R., Kelly, T.J. and Spicer, C.W., (1996). Source attribution of ambient air toxics and other VOCs in Colombus-Ohio. *Atmospheric Environment,* 30 (20), 3457–3470.

NAAQS, (1998). National Ambient Air Quality Standards Notification, 1998. *Central Pollution Control Board*, India.

Naja, M., Lal, S. and Chand, D., (2000). Diurnal and seasonal variabilities in surface ozone at a high altitude site Mt Abu (24.6°N, 72.7°E, and 1680 m ASL) in India. *Atmospheric Environment,* 34(4), 629-633.

NEERI, (2003–2004). Annual Report. *National Environmental Engineering Research Institute*, Nagpur, India

Negi, B. S., Sadasivan and Mishra, U.C., (2003). Aerosol composition and sources in urban areas in India. *Atmospheric Environment*, 37(30), 4205-4215.

Oja, E., (1982). "A simplified neuron model as a principal component analyzer," *J. Math. Biol.*, 15, 267–273.

Okita, T., Hara, H. and Fukuzaki, N., (1996). Measurements of atmospheric SO_2 and SO_4 and determination of the wet scavenging coefficient of sulfate aerosols for the winter monsoon season over the Sea of Japan. *Atmospheric Environment,* 30, 3733–3739.

Ostro, B.D., Hurley, S. and Lipsett, M.J., (1999). Air pollution and daily mortality in the Coachella Valley, California: a study of PM10 dominated by coarse particles. *Environmental Research,* 81, 231–238.

Pandey, J., Agrawal, M., Khanam, M., Narayan, D. and Rao, D.N., (1967). Air pollutant concentrations in Varanasi, India. *Atmospheric Environment,* 21(6), 1259-1266.

Pandey, J.S., Khan, S. and Khanna, P., (2001). Modeling and quantification of temporal risk gradients (TRG) for traffic zones of Delhi City in India. *Journal of Environmental Systems,* 28 (1), 55–69.

Pandey, J.S., Khan, S., Joseph, V. and Kumar, R., (2002). Aerosol scavenging: model application and sensitivity analysis in the Indian context. *Environmental Monitoring and Assessment,* 74, 105–116.

Pandey, J.S., Kumar, R. and Devotta S., (2005). Health risks of NO_2, SPM and SO_2 in Delhi (India). *Atmospheric Environment,* 39, 6868–6874

Parivesh News letter: June (1999). Central Pollution Control Board, Ministry of Environment.

Park, J. and Sandberg, I.W., (1993). "Approximation and radial-basis function networks," *Neural computation*, 5, 305–316.

Paruelo, J.M. and Tomasel, F., (1997). Prediction of functional characteristics of ecosystem. A comparison of artificial neural networks and regression models. *Ecol. Model.*, 98 (2/3), 173-186.

Patade, V.D., Doshi, V.B., Kamat, S.R., Ubare, A.G., Konde, K.N., Shetye, V.S.V., Shah, V., Gregrat, J.K., Papewar, V.N., Sonaje, A.G., Varsha, Shah., Janki, K., Gregrat, V. and Sanjivani, V., (1984) Feb. Fluctuations of daily air pollution levels and respiratory symptoms by health diary in 4 communities from the prospective health survey, Bombay. *Lung India,* 2(1), 29-43

Poggio,T. and F. Girosi, (1990). "Networks for approximation and learning," *Proc. IEEE*, 78, 1481–1497.

Ripley, B.D., (1993). Statistical aspect.6 of neural networks, In *Networks and Chaos-Statistical and Prorbabilistic Aspects,* (Edited by O.E. Barndorff-Nielsen, Jensen J.L. and Kendall K.S.), 40-123, Chapman and Hell, London.

Roadknigh, C.M., Balls, G.R., Mills, G.E., Palmer-Brown, D., (1997). Modeling complex environmental data. *IEEE Trans. Neural Networks* 8, 852-861.

Robeson, S.M. and Steyn, D.G., (1973) Evaluation and comparison of statistical forecast models for daily maximum ozone concentrations. *Atmospheric Environment,* 7(11), 1033-1061.

Rojas, R., (1996). Neural Networks: A Systematic Introduction, *Springer-Verlag*, Berlin.

Rumelhart, D. E., Hinton, G.E. and Williams, J.R., (1986). "Learning internal representations by error propagation," (Cambridge, MA: *MIT Press*), 1, 8.

Sahu, S.K., Pandit, G.G. and Sadasivan, S., (2004). Precipitation scavenging of polycyclic aromatic hydrocarbons in Mumbai, India. *Science of the Total Environment*, 318 (13), 245–249.

Saldiva, P.H.N., Lichtenfels, A.J.F.C., Paiva, P.S.O., Barone, I.A., Martins, M.A., Massad, E., Pereira, J.C.R., Xavier, V.P., Singer, J.M. and Bohm, G.M., (1986). Association between AirPollution and Mortality Due to Respiratory Diseases in Children in São Paulo, Brazil: A Preliminary Report. *Atmospheric Research,* 20(1), 53-65.

Sandberg, I. W., (1991b). "Approximation theorems for discrete-time systems," *IEEE Trans. Circuits Sys.,* 38(5), 564–566.

Sandberg, I.W. and Xu, L., (1997). "Uniform approximation and gamma networks," *Neural Networks*, 10, 781–784.

Sandberg, I.W., (1991a). "Structure theorems for nonlinear systems," *Multidimensional Sys. Sig. Process,* 2, 267–286.

Schroeder, W.H., Dohson, M., Kane, D.M. and Johnson, N.D. (1987). Toxic trace elements associated with air borne particulate matter: a review. *Journal of Air Pollution Control Association,* 33, 1267–1285.

Shetye, S.V., Doshi, V.B. and Patade, V.D., (1984). Trends in cause of death in 4 zones of Bombay over 1971-79. *Lung India*, 44-49.

Shiva Nagendra S.M. and Khare, M., (2005). Artificial neural network approach for modelling nitrogen dioxide dispersion from vehicular exhaust emissions. *Ecological modelling*, corrected proof available online.

Singh, A., Sarin, S.M., Shanmugam, P., Sharma, N., Attri A.K. and Jain, V.K., (1978). Ozone distribution in the urban environment of Delhi during winter months. *Atmospheric Environment,* 12(11), 2185-2196.

Singh, H.B., Salas, L.J., Sites, R. and Shigeishi, H., (1983). Measurements of Hazardous Organic Chemicals in Ambient Atmosphere. EPA-600/3-83-002, *US Environmental Protection Agency*, Washington DC.

Srivastava, A., (2004). Source apportionment of ambient VOCs in Mumbai City. *Atmospheric Environment,* 38, 6829–6843.

Srivastava, A., (2004). Source apportionment of ambient VOCs in Mumbai City. *Atmospheric Environment,* 38 (39), 6829–6843.

Srivastava, A., Joseph, A.E. and Nair, S., (2004). Ambient levels of benzene in Mumbai City. *International Journal of Environmental Health and Research*, 14 (3), 215–222.

Srivastava, A., Joseph, A.E., Patil, S., More, A., Dixit, R.C. and Prakash, M., (2005). Air toxics in ambient air of Delhi. *Atmospheric Environment*, 39, 59–71.

Srivastava, P.K., Pandit, G.C., Sharma, S. and Mohan, A.M., (2000). Volatile organic compounds in indoor environments in Mumbai, India. *The Science of the Total Environment*, 255, 161–168.

Srivastavaa, Anjali., Joseph, A.E. and Devotta S., (2006). Volatile organic compounds in ambient air of Mumbai, India. *Atmospheric Environment,* 40, 892–903.

Staten Island Report, (1997). Staten Island /New Jersey Urban Air Toxic Assessment Project Report, vol. I Summary. EPA/902/ R-93-001a, *US Environmental Protection Agency Research,* Triangle Park, NC.

Stern, F.B., Halperin, W.E., Hornung, R.W., Ringenberg, V.L. and McCammon, C.S., (1988). Heart disease morbidity among bridge and tunnel officers exposed to CO. *Am J Epidemiol,* 128, 1276-88.

Stevens, C.S., (1987). Ozone formation in the greater Johannesburg region. *Atmospheric Environment,* 2(14), 799-805.

Stevens, C.S., (1997). The $NO/NO_2/O_3$ photostationary state and rate of photolysis of NO_2 in central Johannesburg. *Atmospheric Environment,* 31(20), 3421-3427.

Steyn, D.G., Bottenheim, G.W. and Thomson, R.B., (1987). Overview of tropospheric ozone in the Lower Fraser Valley, and the Pacific '93 field study. *Atmospheric Environment,* 21(3), 523-530.

Sweet, C.W. and Vermette, S.J., (1992). Toxic volatile organic compounds in urban air in Illinois. *Environment Science & Technology,* 26, 165.

Tager, I.B., (1999). Methods Development for epidemiologic investigations of the health effects of prolonged 0_3 exposures. *Health effects Inst* Report No.85.

Tang, Z., AImeide C.de. and Fihwick, P.A., (1991).Time series forecasting using neural networks vs. Box-Jenkins methodology, simulation 57 (S), 303-310.

Tasadduq, I., Shafiqur, R. and Khaled, B., (2002). Application of neural networks for the prediction of hourly mean surface temperatures in Saudi Arabia. *Renewable Energy,* 25, 545–54.

TERI, (2001). Review of past and on-going work on urban air quality in India. Report submitted to the World Bank, December, *Tata Energy Research Institute*, http://www.worldbank.org/sarurbanair.

The United Kingdom National Air Quality Strategy. (1997). London: Department of the Environment.

Thom, S.R., (1997). Ischiropoulos Mechanism of oxidative stress from levels of CO; *Health Effects Inst*. Report No 80.

Thurston, G.D., Spenger, J.D., (1985). A quantitative assessment of source contributions to inhalable particulates in metropolitan, *Boston. Atmos. Environ.* 19,9-25.

Urban Air Act (1996): Urban air quality management strategy in Asia Greater Mumbai report: *World Bank*, USA. US-EPA, (2002). Health assessment document for diesel exhaust, EPA/600/8-90/057F, May.

Van Hulle, M. M., (2000). *Faithful Representations and Topographic Maps: From Distortion-to-Information-Based Self Organization* (New York: Wiley).

Varshney, C.K. and Aggarwal, M., (1991). Ozone pollution in the urban atmosphere of Delhi. *Atmospheric Environment Part B. Urban Atmosphere,* 25(1), 109-112.

Viotti P, Liuti G. and Di Genova, P. (2002). Atmospheric urban pollution: applications of an artificial neural network (ANN) to the city of Perugia. *Ecological Modelling,* 148(1), 27–46.

Werbos, P. J., (1990). "Back propagation through time: What it does and how to do it," *Proc. IEEE*, 78, 1550–1560.

Whitehead, P.G., Howard, A. and Arulmani, C., (1997). Modelling algal growth and transport in rivers-A comparison of time series analysis, dynamic mass balance and neural network techniques. *Hydrobiologia*, 349, 39-46.