MODELLING OF ROADSIDE AIR QUALITY USING ARTIFICAL NEURAL NETWORK

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FORTHE AWARD OF THE DEGREE OF

MASTER OF TECHNOLOGY

IN

ENVIRONMENTAL ENGINEERING

by

ANJANA REGHU

(ROLL NO : 2K13/ENE/17)

Under the esteemed guidance of

Dr. S.K SINGH

Dr. ANURADHA SHUKLA

&

Dr.ANUBHA MANDAL

DEPARTMENT OF ENVIRONMENTAL ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

DELHI – 110042

2015

Department of Environmental Engineering
Delhi Technological University
Shahabad, Delhi – 110042

CERTIFICATE

This is to certify that the dissertation / project report entitled "***Modelling of roadside air quality using artificial neural network***" done by Anjana Reghu *(2K13/ENE/17)* is an authentic work carried out by her at the Central Road Research Institute,Delhi under my guidance. The matter embodied in this project has not been submitted earlier for the award of any degree or diploma in this or any other university or institute to the best of my knowledge and belief.

Professor S.K.Singh (Supervisor)
Department of Environmental Engg.
Delhi Technological University,
Delhi -110042

Dr.Anuradha Shukla(Supervisor)
Chief Scientist,Environmental Division
CSIR-Central Road Research Institute
Delhi –110025

Dr.Anubha Mandal(Co- supervisor)
Department of Environmental Engg.
Delhi Technological University
Delhi -110042

Date:
Place:

ACKNOWLEDGEMENT

This thesis is an account of work in the field of roadside air quality modelling using Artificial Neural Network.I would like to express my deepest and sincere gratitude to all the people who have helped in successful completion of this dissertation.

It gives me immense pleasure to take this opportunity to thank my guide ***Prof. S. K. Singh*** for his invaluable guidance, encouragement, patient reviews and giving me a wonderful opportunity to work in CSIR-CRRI on my dissertation .

It would like to specially thank my mentor ***Dr.Anubha Mandal*** for the valuable advice,inputs and support she gave to improve my research work.

I extend my profound gratitude to ***Dr. Anuradha Shukla*** my co-supervisor in CSIR-CRRI,Delhi for her valuable guidance,discussions and providing me with the excellent atmosphere to carry out my work research . I am truly indebted to ***Dr.Ravi Sekar*** for taking out valuable time from his busy schedule to teach and improve my knowledge in statistical aspects of modelling. This dissertation would not have been possible without his patient guidance and support.I thank **Mr.S.K.Peshin** ,Indian Meteorological Department,Delhi for providing me data required in the study. I am also thankful to all other teachers, friends who directly or indirectly helped me in completion of my project successfully.

I dedicate my dissertation to my family .This thesis would not have been possible without their moral support and blessings.Their encouragement and love cheered me up during weary times of the work.

(Anjana Reghu)

DECLARATION

I hereby declare that the thesis entitled **"(Modelling of roadside air quality using artificial neural network)"** submitted by me, for the award of the degree of Master of technology in Environmental Engineering from Delhi Technological University, Delhi is record of bonafied work carried out by me under the supervision of Dr.S.K. Singh (Professor, Department of Environmental Engineering, Delhi Technological University).

I further declare that the work reported in the thesis has not been submitted and will not be submitted either in part or in full, for the award of any degree or diploma in this or in any other institution or university.

Signature of the candidate

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Over the past decade rapid urbanisation and development has led to steady increase in number of vehicles on road in Delhi, which has led to alarming increase in level of air pollution. Air pollution in Delhi has become a serious environmental problem in recent years. The concentration of pollutants like particulate matter (PM_{10},$PM_{2.5}$), nitrogen oxides (NOx) and carbon monoxide (CO) in ambient air have continuously exceeded the threshold limits specially in areas near arterial roads. The high concentration of toxic pollutants in ambient air are a silent and lethal killer .The poor air quality causes serious health ailments such as respiratory diseases, increase in risk of developing cancer, heart diseases and other serious ailments . This leads to tremendous loss of financial resources in form of medical expenses for treatment of affected people. According to data out of the total pollutants discharged in air in Delhi every day, vehicular emissions from transportation sector is the major contributor in the total air pollutant load .Air quality monitoring and prediction systems are the need of the hour for controlling the pollutant concentration to improve urban air quality of Delhi .

The main objective of this study is to develop non-parametric Artificial Neural Network models to predict roadside air quality in Delhi for pollutants like PM_{10} and $PM_{2.5}$,NOx and CO .Air quality prediction is usually carried out by soft computing techniques, fuzzy logic and generic algorithms . Artificial Neural Network (ANN), a soft computing technique has been steadily gaining popularity as an optimised air quality prediction tool among researchers over the past few years. The motivation for using Artificial Neural Network for modelling in this study stems from the capability and efficiency of ANN in computation of highly complex non-linear dynamic systems with large dimensional data. The proposed ANN models considered meteorological parameters like wind speed, gust wind, relative humidity, pressure, temperature and traffic characteristics as inputs and concentration of various pollutants as outputs. Different models were developed for short term prediction and their performance was evaluated on the basis of statistical parameters like MAPE, MAE, RMSE and coefficient of determination (R^2). The weights established after training process were used to find out the relative influence of different vehicles on the pollutant concentrations.ANN modelling can be used as a pre-warning mechanism for air pollution episodes and help the policy makers in formulating suitable mitigation measures to curb pollution arising from vehicular exhaust emissions .

Keywords: Artificial neural network, criteria pollutants, Multilayer perceptron , back propagation learning

Chapter 1
INTRODUCTION

1.1 BACKGROUND

The air pollution from vehicular exhaust emissions has grown at an alarming rate due to increasing number of vehicles from the unplanned and rapid urbanisation in Delhi over the past few years . Delhi was found to be the most polluted city in the world by the report released by WHO in 2014. Delhi has the maximum number of vehicles as compared to other metropolitan cities in India. The number of registered vehicles in Delhi and the surrounding areas of Noida, Greater Noida, Faridabad, Gurgaon, Ghaziabad, Bahadurgarh and Sonipat reached 8 million in March 2015, with nearly 1400 new vehicles being added to this figure every day. According to the study conducted in Delhi, out of the total 3,000 metric tonnes of pollutants discharged in air every day, vehicular emissions from transportation sector contribute to nearly 72 % of total air pollutant load (**Goyal et al. , 2006)** which was around 43 % in 1981 and only 23 % in 1970-71. **(CPCB, 2008)**.

Delhi does not have any major industrial units, though it has a large number of small scale industries .The emissions from these industries does not pose much of a problem in comparison to vehicular induced pollution in the city **.(Suthirto, 2002).**

A detailed analysis of monitored ambient air quality data near traffic intersections presented the following findings.

• CO and NO_X concentrations were reported to be high near traffic intersections during peak traffic hours in morning and evening.

• Particulate matter exceeded the permissible limits continuously ,which was partly due to resuspension of dust from vehicular movement and from the emissions of combustion engines of vehicles**.(Goyal et al. ,2003).**

The air pollution from vehicular exhausts near major highways, arterial roads has become a serious problem in Delhi. A number of studies have shown that pollutant exposure near arterial roadways, intersections are always greater than in other locations in the city. Extensive studies have linked vehicular pollution with the rise in heart attack, lung cancer, respiratory, cardiovascular and pulmonary diseases among the people residing near arterial roads with heavy traffic flow.

In addition to the contribution of rising number of vehicles, air pollution problem in Delhi has been aggravated by the movement of diesel trucks and bus from nearby states of Uttar Pradesh, Haryana and Punjab. Delhi is a major transit point in north India for freight and passenger traffic. Delhi is connected to five major national highways: **NH-1, NH-2, NH-8 , NH-10 and NH-24**. The old, poorly maintained fleet of transport buses and diesel trucks plying from neighbouring states are the main source of high pollutant concentration along the major highways and arterial roads.

1.2 POLLUTANT EMISSIONS FROM VEHICLES

The main pollutants emitted from the vehicles are carbon monoxide, sulphur dioxide, nitrogen dioxide, particulate matter,hydrocarbons, lead and benzene. Air pollutants emitted from vehicles are classified into primary and secondary pollutants. Primary pollutant is directly emitted into atmosphere, while secondary pollutants are formed as a result of chemical reactions between primary pollutants in the atmosphere. The following are the main primary pollutants released from exhausts of vehicles:

1.2.1 Particulate matter ($PM_{2.5}$ and PM_{10}) - Particulate matter (PM) is a mixture of solid particles and liquid droplets found in the air. Particulate matter can be of varied sizes and composition. The main two types of particulate matter are PM_{10} and $PM_{2.5}$ .sParticles less than 10 micrometers (µm) in diameter ,found near roads and dusty industries are known as PM_{10}.$PM_{2.5}$ also known as Respiratory Suspended Particulate Matter (RSPM) are particles less than 2.5µm in diameter, emitted from natural sources like forest fires, industrial combustion sources or formed by reaction of gases in air .The deposition rate of particulate matter depends on prevailing weather conditions like wind speed, temperature, humidity etc. Particulate matter is measured and expressed as mass of particles in cubic meter of air (µg/m3).

The air quality monitoring and analysis has reported that the particulate matter from vehicles is due to following reasons:

1. Resuspension of road dust from vehicular movement.
2. The exhaust emissions of diesel, CNG powered vehicles containing fine particulate matter.

PM can be a primary pollutant when emitted by natural sources into atmosphere or a secondary pollutant when transformed by gaseous precursors hydrocarbons, nitrogen oxides and sulphur dioxides (**Aneja et al. ,2001).** Nitrogen oxides and sulphur dioxides are known to use particulate matter as a surrogate to carry and deposit themselves.Suspended particulate matter in ambient air is very dangerous due to their ability to penetrate deep into respiratory organs, lungs and blood streams causing permanent DNA mutations, heart attacks, lung and respiratory diseases.

The air quality data from monitoring stations maintained by Central Pollution Control Board have shown concentrations of particulate matter in ambient air continuously exceeding the standard safe limit from the past couple of years in Delhi .The average $PM_{2.5}$ and PM_{10} level of Delhi in 2015 was at a whopping 153 ug/m^3 and 286 ug/m^3 respectively which is nearly fifteen times as compared to WHO guidelines. The WHO's safety limit for humans is 10 ug/m3 for $PM_{2.5}$ and 20 ug/m3 for PM_{10} annually.

1.2.2 Nitrogen oxides (NOx) – In combustion engine nitrogen oxides are formed partly by nitrogen compounds in fuel but mostly by direct combination of nitrogen and oxygen atoms in air under high pressure and temperature conditions. Road transport contributes to nearly 40-70% of NOx emissions worldwide. The measured concentrations of NOx near signalised intersections and areas with traffic congestion in Delhi shows predominant effect of vehicles on the NOx levels in Delhi. Nitrogen oxides cause lung irritation and weaken the immune system of body against respiratory infections like pneumonia and influenza. Nitrogen oxides are main precursors of ground level ozone formation which is a health hazard themselves. NOx emissions are also major sources of acid rain that affects both terrestrial and aquatic ecosystems. Airborne nitrates and nitrogen dioxides contribute to pollutant haze, which severely impairs visibility

1.2.3 Carbon monoxide (CO) - Carbon Monoxide is an odourless, colourless and poisonous gas . It is a product of partial oxidation of carbon due to incomplete combustion of fuel rather than full oxidation to carbon dioxide (CO-2) in combustion engines. The air supply in engine is restricted during start up in vehicles which are not tuned well and incomplete combustion occurs releasing carbon monoxide instead of carbon dioxide. Carbon monoxide exposure interferes with blood’s ability to carry oxygen to brain. This causes slowing down of reflexes, affects thinking skills, causes headaches, fatigue and can even prove fatal when at high concentrations. People with chronic illnesses, newborns and foetuses are more susceptible to the effects of carbon monoxide exposure. Hourly carbon monoxide (CO) level in Delhi has crossed 4,000 microgram per cubic metre in many places way above the safe level for residential areas. The concentrations of CO from vehicular emissions showed an increase of

92% in 1996 over the concentrations in 1989, a consequence of increase in vehicular population. **(CPCB,1998).**

The emissions from vehicles depend upon the type of vehicle and the fuel used for combustion. The emissions from different type of vehicles and different fuel type are discussed in the next section.

1.3 FUEL WISE EMISSIONS FROM VEHICLES

1.3.1 Emissions from petrol vehicles

Petroleum is a mixture of large number of different types of hydrocarbons. The most commonly found are alkanes, cycloalkanes, aromatic hydrocarbons and complicated chemicals like asphaltenes. Petroleum also contains many trace elements like carbon, nitrogen, hydrogen, oxygen and sulphur and few trace metals. Petrol vehicles form major share of vehicles on road in Delhi. Two wheelers and cars, usually have petrol combustion engines. The petrol engine works on the principle of spark ignition for fuel ignition. The majority of exhaust emission consists of nitrogen, water and carbon dioxide. Harmful emissions form only a small fraction around 1.1 % of the total emissions from a modern petrol engine. The main pollutants emitted from petrol vehicles are carbon monoxide, hydrocarbons and nitrogen oxides in decreasing order of their contribution.

The introduction of catalytic converters which oxidise pollutants such as CO to less harmful gases such as CO-2 have dramatically reduced emissions from petrol cars. The catalyst cars have lower CO, NO and HC emissions than cars without catalyst. But the CO-2 emissions from catalyst cars are higher due to increase in oxidation of carbon monoxide to carbon dioxide. As a result, a catalyst car uses more fuel and becomes less efficient. Despite these improvements, catalyst petrol cars still produce more carbon monoxide and hydrocarbons than diesel cars although emissions of nitrogen oxides and particulates are much lower than diesel cars. The particulate emissions from petrol cars are very low and almost negligible and is not routinely measured.

1.3.2 Emissions from diesel vehicles

Diesel is derived from fractional distillation of petroleum and is composed of around 75% of saturated hydrocarbons (mostly paraffins and cycloparaffins) and 25% of aromatic hydrocarbons like alkyl benzenes and napthalenes.

In diesel engines fuel ignition takes place as a result of compression of the injected fuel and the inlet air. This generates high temperatures for diesel fuel to ignite when it is injected inside combustion chamber (**Bosch,2005**). Air is primarily composed of nitrogen and oxygen. The nitrogen in air reacts with the oxygen in combustion chamber when temperatures above 1600 degree are attained and NOx is formed. Heat is released with the burning of fuel and carbon monoxide, hydrocarbons are emitted from the unburned diesel and incomplete combustion **(Ibrahim Aslan Res¸itog˘lu et al. ,2015).**

The exhaust emissions from diesel vehicles are primarily composed of carbon dioxide, water and unused portion of engine charge air. Pollutants from diesel vehicles exhausts form only 0.2% of total exhaust emissions and consists of carbon monoxide, nitrogen oxides unburned hydrocarbons and particulate matter .These pollutants mostly originate from various non-ideal processes during combustion such as incomplete combustion of diesel, reactions between diesel components under high temperature and pressure, combustion of lubricating oil, oil additives and non-hydrocarbon components of diesel such as sulphur compounds and fuel additives. The pollutant type and concentration in exhaust gases depends on air-fuel ratio, air fuel concentration, combustion temperature, ignition timing, turbulence in chamber etc.

NOx constitutes highest proportion, nearly 67 % of exhaust gas from diesel vehicles. Nitrogen oxides constitute nitrogen oxides (NO) and nitrogen dioxides (NO_2).NO forms nearly 85–95 % of NO_x. The amount of NOx produced is a function of oxygen concentrations, residence time and maximum temperature in combustion cylinder . Most of the NOx formed is during the early combustion process, when the ignition temperature is the highest as a result of movement of piston to the top of the stroke. The amount of NO_x formed increases by threshold for every 100 degree rise in temperature (**Bosch 2005**).

Particulate matter formation is caused by agglomeration of fine particles of partially burnt fuel, lube oil, ash content of fuel oil, water and sulphates (**Demers and Walters 1999,Maricq 2007**). Particulate matter emissions from diesel engines are considerably higher (nearly six to ten times) than from petrol engines. The formation of particulate matter from combustion process are dependent on factors like fuel quality(sulphur and ash content),combustion temperature and lubrication oil quality (**Burtscher,2005**).The approximate composition of pollutant emissions from diesel vehicles is given in **Figure.1.1**

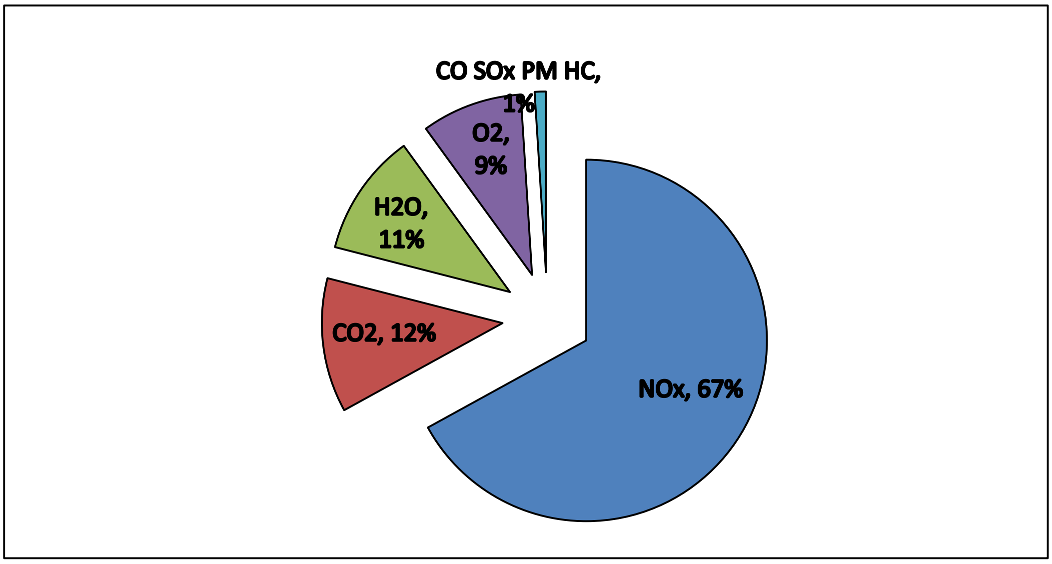


Figure: 1.1: Composition of exhaust gases from diesel vehicles **(Khair and Majewski 2006)**

Diesel engines are lean combustion engines so the emissions of pollutants like CO, CO-2 , HC per km on an average from diesel cars is very less in comparison with petrol cars. Diesel vehicles have no lead emissions. (**U.S E.P.A**)

Diesel vehicles are considered to be the largest contributors to urban air pollution .Diesel engines have greater fuel economy, efficiency, durability and lower operating cost than petrol vehicles. These attractive features make diesel engines more economical than petrol engines in vehicles specially in buses and heavy duty vehicles like trucks in commercial transport.

In 2010, petrol was deregulated by government but diesel was subsidized. This led to a big price gap between petrol and diesel encouraging around 60% of new buyers to purchase diesel cars. Even now more than 40% of new car sales are diesel due to lower price of diesel and higher mileage of diesel cars. Currently, a diesel car emits 5-10 times more particulate matter ($PM_{2.5}$) and nearly 2-3 times more NO_x than a petrol car.

In addition to growing number of diesel vehicles, diesel trucks and HCV coming from neighbouring states are the major contibutor to the increasing particulate matter and NO_x concentrations in Delhi .The pollution check on these vehicles are not conducted regularly and they violate the emission norms blatantly. The heavy duty industry in India is way behind the global emissions standards. In Europe the Euro VI emission standard is currently in force for HCV while India still follows Bharat stage four emission norms which has high CO, NO_x and HC emission limit in comparison to European standards. India plans to move to Euro 5 only by 2020 which is far too late.

Recent studies have reported diesel vehicles have higher emissions of nitrogen oxides and particulate matter as compared to petrol vehicles and are responsible for severe health and environmental problems (**R.Prasad and V.R Bella, 2010**). Diesel particulate matter is considered to be very harmful, because it is composed of very small fine particles that can be easily inhaled. The typical composition of Diesel $PM_{2.5}$ in **Figure: 1.2.**

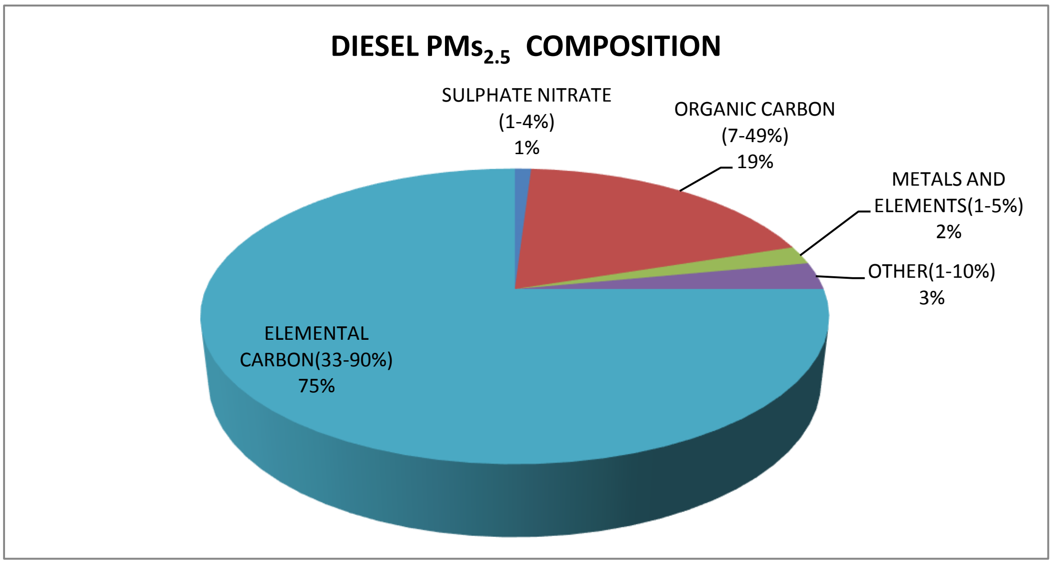


Figure 1.2 :Diesel $PM_{2.5}$ chemical composition**(GmBH report)**

Diesel particulate matter is recognized as a toxic form of fine particulate matter and potential human carcinogen by many international agencies like the World Health Organization, United States Environmental Protection Agency. The exposure to diesel exhaust gases are harmful and causes severe respiratory problems, lung damage and cancer in humans **(Lloyd and Cackette, 2001).**

1.3.3 Emissions from (CNG) vehicles

CNG or compresses natural gas is mainly composed of methane but it may contain ethane, propane and heavy hydrocarbons .It may also have small quantities of nitrogen, oxygen, carbon dioxide, sulphur compounds and water. Natural gas vehicles produces nearly 25 % lesser harmful emissions of nitrogen oxides, particulate matter, volatile hydrocarbons, toxic and carcinogenic pollutants than conventional petroleum and diesel vehicles (**U.S E.P.A)**. However, the total hydrocarbons (THC) emissions from CNG vehicles are relatively high due to presence of methane, a major component of natural gas. Though the harmful emissions from CNG vehicles is less, the emission of green house gas methane is a major disadvantage as methane traps nearly 30 times more heat than carbon dioxide .

A recent study conducted on a small delivery van fitted with three way catalyst and a switch to change modes from CNG and petrol, reported that the emissions of CO, NO_x and non-methane hydrocarbons (NMHC) and were 76%,83% and 88% respectively lower with CNG engine than with petrol **(U.S E.P.A).**

In 2002 ,the Delhi government directed conversion of all commercial buses, taxis, passenger vehicles and three-wheelers to move to compressed natural gas (CNG). The comparison of annual average concentration of CO, PAHs, and SO_2 before and after the implementation of CNG, showed approximately 50 % reduction of the pollutants in ambient air of Delhi except NO_x concentrations, which showed an increase of 10 to 20 % in the subsequent years.. **(K.Ravindra,2006)**

The increasing trend for NO_x concentration maybe related to the steady increase in total number of vehicles every year in Delhi. CNG has high flash-point of (540 °C) in comparison to diesel (232–282 °C). At such high temperatures, more nitrogen from the air gets compressed and reacts with oxygen in the combustion chamber of CNG vehicles, subsequently producing more NO_x **(P. Saxena et al. ,2012).** CNG vehicles emit nearly three times more NO_x than petrol vehicles**.**The concentration of particulate matter in Delhi has crossed permissible limits for both SPM and RSPM after the implementation of CNG. The increasing trend in particulate matter and NO_x may be due to the poor technology of CNG three wheelers, poorly maintained CNG fleet, poor quality of piston rings and improper maintenance of air filters (**Narain and Krupnick,2007).** The air quality of Delhi has not improved substantially after implementation of CNG vehicles .

1.3.4 Emissions from LPG vehicles

LPG is a mixture of petroleum and natural gas which exists in liquid state at normal pressure and temperature. It has simpler composition and lower density than petrol. The main constituents of LPG are propane with butane and isobutane **(Chang et al. , 2001).** LPG-fuelled engines have better performance, efficiency than petrol-fuelled engines and lower exhaust emissions. The particulate emission factor was reported to be nearly 70% lesser with LPG in all operational modes as compared to unleaded petrol **(Z.D. Ristovski et al. ,2004).** The number of LPG vehicles in Delhi is very less and the effect of LPG vehicles exhaust emission is not substantial to be taken into account.

A study was conducted on 27 locations in Delhi on vehicular emissions of criteria pollutants namely particulate matter, nitrogen oxides and carbon monoxide by (**Goyal et al. , 2013**) .The study reported the emissions of PM, NO_x and CO from different fuel and type of vehicles in Delhi based on the data of 2008-2009 from CPCB monitoring stations.

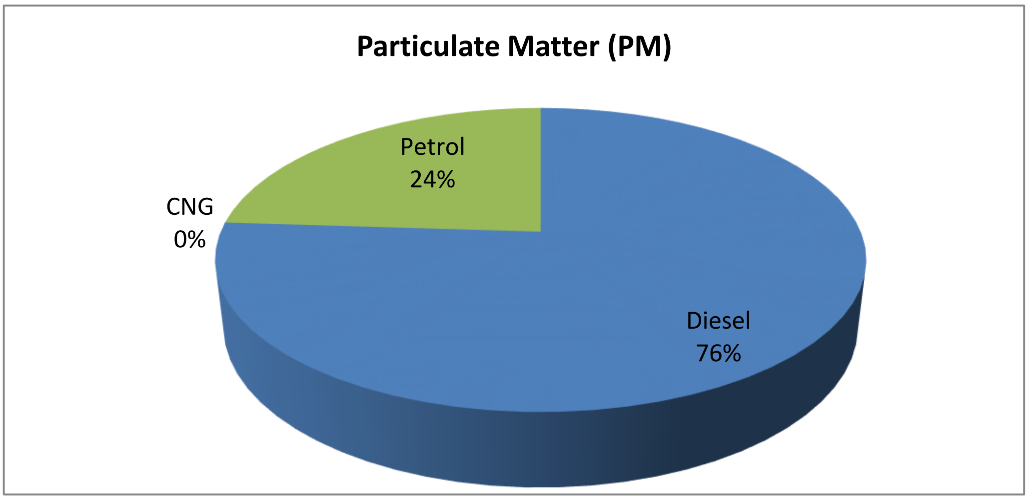


Figure 1.3: Fuel wise emission of Particulate matter**(Goyal et al. , 2013)**

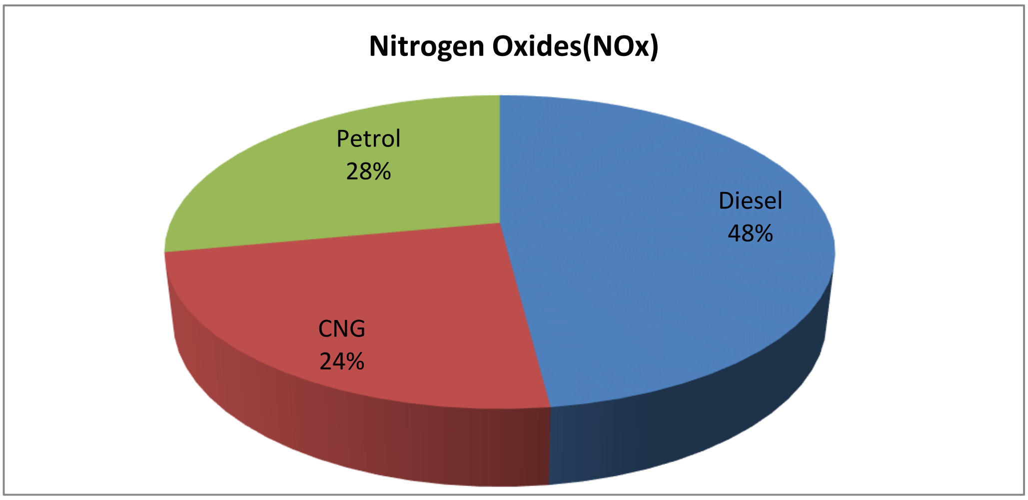


Figure 1.4: Fuel wise emission of Nitrogen Oxides**(Goyal et al. , 2013)**

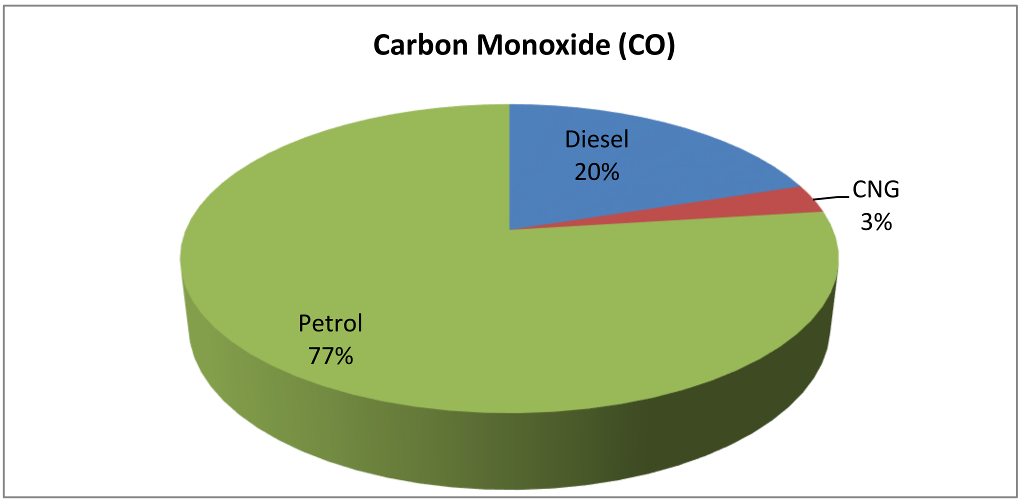


Figure 1.5: Fuel wise emission of Carbon Monoxide **(Goyal et al. ,2013)**

The fuel wise emission analysis revealed that CO emission was mainly due to petrol vehicles while diesel vehicles were the main contributor of NOx and particulate matter emissions. On basis of the results of the study petrol and diesel were found to be major source of CO, NOx and particulate matter in Delhi. (**Goyal et al. , 2013**)

1.4 VEHICLE TYPE WISE EMISSIONS

The vehicle wise emission analysis reported by (**Goyal et al. , 2013)** for Delhi revealed that CO emissions were mainly from two wheelers while the main contributor of NO_x and particulate matter emissions were passenger cars and HCV respectively. The emissions of different pollutants vary according to the different operating conditions of vehicles e.g During idling and decelerating CO emissions were found to be higher than during cruising, while the NO_x and particulate matter emissions were found to be lower during decelerating and idling and than in cruising .

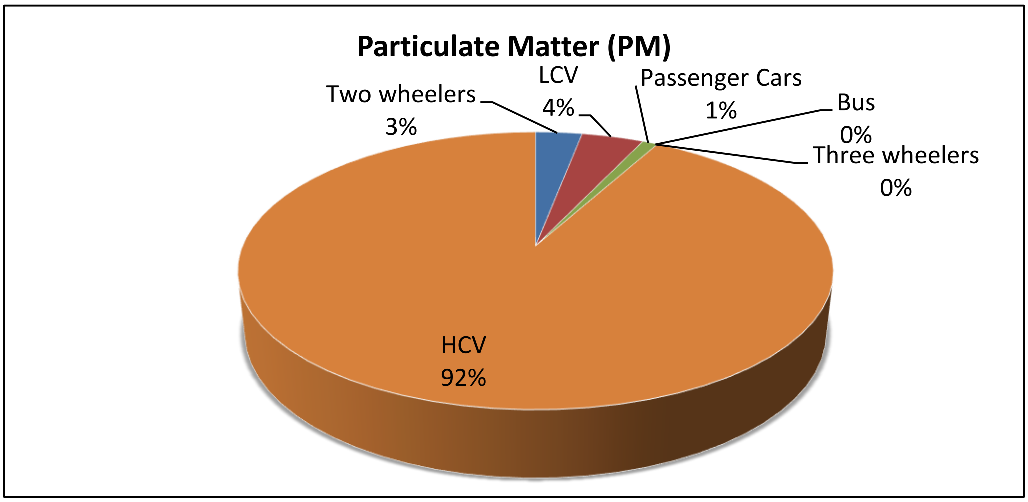


Figure 1.6: Vehicle wise emission of Particulate matter **(Goyal et al. ,2013)**

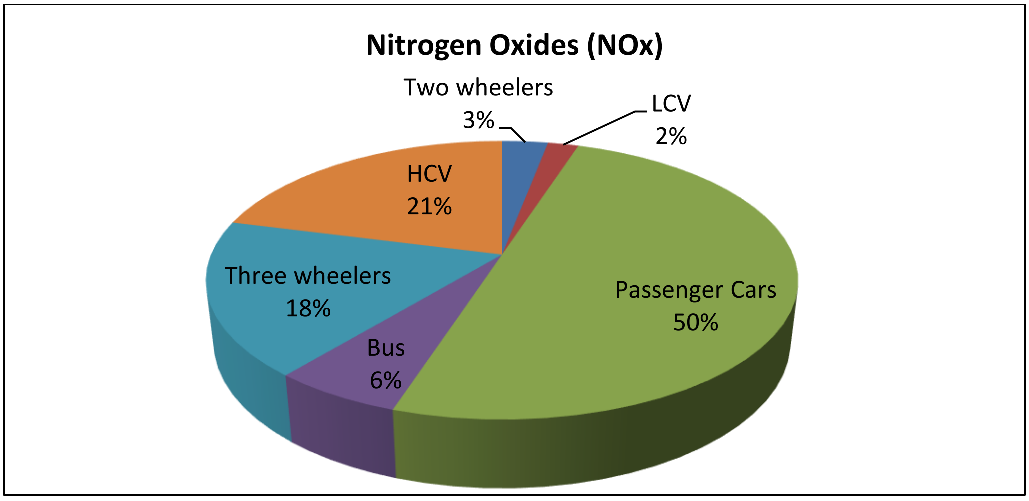


Figure 1.7: Vehicle wise emission of Nitrogen oxides **(Goyal et al. ,2013)**

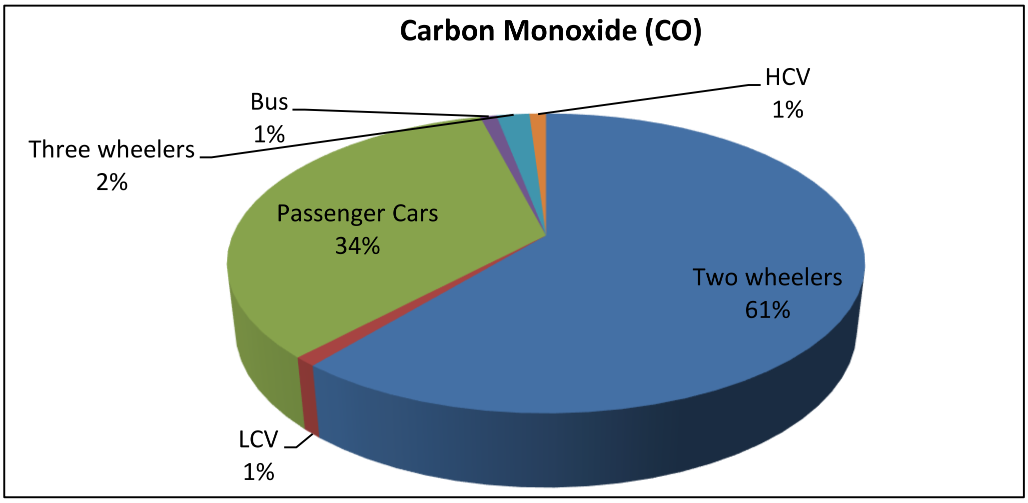


Figure 1.8: Vehicle wise emission of Carbon monoxide **(Goyal et al. ,2013)**

1.5 NEED FOR THE PRESENT STUDY

Air pollutant concentration in ambient air is an important indicator of air quality of a region. The contribution of vehicular exhaust emissions to pollutant concentration in ambient air in areas located near arterial roads, intersections is very high. There is a need for developing robust air monitoring and prediction systems for implementation of proper control measures for ensuring good health of residents .It is not feasible to have air in all areas around traffic intersections and roads with high traffic volume.Air quality models can be used for forecasting using meteorological and traffic parameters at such critical locations. These models can also help in filling in the missing and erroneous data at existing monitoring stations.

Traffic on arterial roads along with the meteorological conditions add complexity in air prediction models .Artificial neural network models have been successfully able to explain the phenomenon of pollutant dispersion in urban areas in studies conducted around the world. ANN consists of adaptive connected processing units with the capability of identifying non–linear relationships between the different parameters from incomplete and noisy data sets. Neural networks can be trained to extract trends and correlations from the data sets to develop appropriate air pollutant prediction models. The ability to capture non-linear relationships and phenomenon is highly useful for developing deterministic models for air pollutant forecasting. ANN can also help in identifying the relative contributions of vehicles on air quality and help agencies formulate remedial measures either by improving the technology of combustion or curbing the movement of particular type of vehicles.

The monitoring station at Central Road Research Institute in Delhi which is located along the busy NH-2 has continuously reported high concentration of $PM_{2.5}$,PM_{10},NO_x and CO exceeding the prescribed permissible limits by CPCB . The pollutant concentration was found to be very high during the peak hours (when the traffic volume and congestion is maximum) and around midnight when the truck movement is maximum. The purpose of this project is to develop ANN models for forecasting of criteria pollutants $PM_{2.5}$,PM_{10}, NOx and CO for the particular location.

1.6 OBJECTIVE OF THE STUDY

The objectives of the present study are given s below.

1. Selection of ensemble of input parameters which influence the complex nature of air pollution from vehicular exhaust emissions
2. To develop and optimise ANN models for $PM_{2.5}$, PM_{10}, NO_x and CO concentrations from hourly data to achieve the most accurate predictions.
3. To evaluate the relative importance of input parameters through process of partitioning of weights obtained during the training process.
4. To evaluate the performance of ANN models developed for the study area.

Chapter 2
LITERATURE REVIEW

This chapter gives an overview of previous studies and research described in publications where Artificial Neural Networks (ANN) has been used for air quality modelling.

Suhasini V. Kottur1, Dr. S. S. Mantha(2015) presented an integrated model using artificial neural network and Kriging to forecast the level of air pollutants at different locations in Mumbai and Navi Mumbai using data from meteorological department and Pollution Control Board. The ANN model was implemented and tested using MATLAB for ANN. The developed models could be used to predict the values of pollutants at monitored and unmonitored locations. Kriging was able to predict pollutant values for unknown locations the high correlation coefficient (R) value showed that desired value of fit was obtained between predicted and observed values.

.

Madhavi Anushka Elangasinghe,Naresh Singhal,Kim N. Dirks,Jennifer A.Salmond(2014) developed a ANN forecasting model for NO_2 at a study area located near a major highway in Auckland, NewZealand. The model consisted of eight predictor variables: Wind direction, wind speed,temperature,relative humidity, solar radiation, month of year, day of week and hour of the day. In the study three optimization techniques were explored: Genetic algorithm, forward selection and backward elimination. Of the three, genetic algorithm technique gave results with smallest absolute mean error. A simplified ANN model was developed after successive removal of predictor variable with less influence of NO_2 concentration. On comparison with linear regression model based on same input parameters ANN model performed significantly better.

Navneeta lal Benjamin, Sarita Sharma, Umesh Pendharker, JKShrivastava (2014) developed two feed –foward neural network for NOx prediction for Mahakal Mandir area located in Ujjain City. The daily ambient air velocity, relative humidity, ambient air temperature, rainfall for the study area was taken in the input data set for period of four years from 2009-2013. The best prediction performance was observed for the model with structure 4-7-1 with Mean Square Error (MSE) as 0.0023 and minimum percentage error as 0.332.

Alka Srivastava and Ashok K.Sharma (2014) did a study on modeling of ambient air pollutants using ANN for an industrial area of Ujjain city air pollutants. Air pollutants like Sox, NOx, SPM and RSPM were considered for the study. The ANN network was built with the inputs as meteorological data and outputs as concentration of various pollutants .The monthly meteorological data like Temperature, Humidity, wind pressure and rainfall from year 2006 - 2012 were used and the pollutant concentration was collected from the State Pollution Control Board. The study estimated the Mean Square Error (MSE) from the model was found satisfactory being in the range of 0.01-0.03. The models developed in the study could be easily implemented to deliver real time prediction, unlike other modeling techniques.

Sudhir Nigam, Rashmi Nigam,Sangeeta Kapoor (2013) attempted to forecast CO based on historical data using (ANN). The eight hourly average CO emission data of eleven years (1996-2006), day time (14.00hrs-22.00hrs) from ITO square in Delhi was used for modelling and simulation study. Model performance was evaluated using Coefficient of Determination (CD), Index of Agreement (IA) and Fractional Bias (FB).The performance of ANN model developed was found to be excellent in capturing the trend of CO dispersion and forecasting the mean levels of CO. The general trend in forecasts was over prediction except at few peaks of observed high CO levels.

Mohammad Arihami,Nima Khamali,Mohammad Mahdi Rajabi (2013) carried out a study on the ability of ANN models to predict criteria pollutants carbon monoxide (CO), oxides of nitrogen (NO_x), nitrogen dioxide (NO_2), nitrogen monoxide (NO), ozone (O_3), and particulate matter (PM_{10}) at Fatemi a densely populated area in Tehran surrounded by streets with heavy traffic . The capacity of ANN models to predict the pollutants influenced by primary emission sources (mainly CO) and secondary photochemical reaction (mainly O_3).A methodology for computing probability of exceeding thresholds of air quality were developed by combining ANNs and Monte Carlo Simulations based on Latin Hypercube Sampling for developing reliable ANN models. High correlations with R^2 more than 0.82 were reported between observed and predicted hourly pollutant concentrations of CO, NO_2, NOx,NO and PM_{10}. However the predicted values by ANN model for O_3 were less accurate than the models developed for other pollutants.

Mouhammad Alkasassbeh,Alaa F.Sheta, Hossam Faris, Hamza Turabeih (2013) did a case study in Salt,Jordan to develop a non-parametric ANN model to predict PM_{10} and TSP. The data collected by eight monitoring stations around Al-Fuhais cement plant was used in the study.

.Artificial neural Network based on Autoregressive with eXternal (ANNARX) was used to develop two high performance models which considered meteorological parameters as inputs.
The model performance was evaluated by measuring the MSE (Mean Square Error), Mean Magnitude of Relative Error (MMRE), Euclidian distance(ED) and the Manhattan distance (MD).The developed model was reported to be good with respect to error evaluation criterion adopted.

Sonaje N P,Mane S J,Kote A S (2013) described the development of ANN model for forecasting of respirable suspended particulate matter (RSPM) for an urban area in Pune,Maharashtra. The six year (September 2004-December 2010) daily data of RSPM used for developing a three layer MLP neural network. The number of epochs was set to 1000 for training. The neural network with architecture 4-2-1 (four inputs, two hidden layers and one neuron as output) was reported to be the best model with RMSE and MAE value as 0.902, 605.454(μg/m3)^2 and 17.45 μg/m3 respectively .

Azman Azid,Hafizan Juahir,Mohd Talib Latif,Sharifuddin Mohd Zain, Mohamad Romizan Osman (2013) conducted a study which described application of ANN to predict air pollutant index (API) for seven air monitoring stations in southern region of Peninsular Malaysia from the seven year database (2005-2011). Feed forward method of ANN was used to predict API using 12 air quality parameters. The models using all the rotated principal component scores was found to be be more efficient and effective**.**

Maitha H. Al Shamisi, Ali H. Assi and Hassan A. N. Hejase (2011) used MATLAB tools to predict monthly average global solar radiation for Al Ain city - UAE. The weather data between 1995 and 2004 were used for training the network, while data from 2005 and 2007 were used for testing. Eleven models with different input combinations were modelled with MLP and RBF ANN techniques. The obtained results reported the superiority of RBF technique over the MLP technique in with deterministic coefficients nearly 90 %and low MBE, MAPE and RMSE values. ANN models showed good performance even if one or more input parameters were unavailable

Ming Cai,Yafeng Yin,Min Xie (2009) used ANN for prediction of hourly air pollutant concentration near an arterial in Guangzhou,China.The pollutants considered in the study were carbon monoxide,particulate matter,nitrogen dioxide and ozone which were measured at the sites using monitoring equipments. Factors taken into consideration were classified into four categories :Meteorological, geographical, traffic characteristics and background concentration .

The hourly averages of the influential factors were taken as input variables and measured concentrations of pollutants as output. Models for each pollutant were developed, trained, validated and tested based on back-propagation technique of ANN. The models were reported to produce accurate prediction of average hourly pollutant concentrations of pollutants 10 hours in advance. The ANN models were found to outperform California line source dispersion model and multiple linear regression models.

Lovro Hrust, Zvjezdana Bencetic Klaic, Josip Kri_zan, Oleg Antonic, Predrag Hercog(2009) developed ANN models for forecasting four air pollutants ((NO_2, O_3, CO and PM_{10}) for a urban residential area located in city of Zagreb,Croatia. Multi layer perceptron(MLP) neural network was for the forecasting model. The data set comprised of 15 minute averages of pollutants were used for model building. The index of agreement for the data set including model building ranged from 0.91-0.97 for the pollutants taken in consideration

Luis A. Dı´az-Robles, Juan C. Ortega, Joshua S. Fu, Gregory D. Reed, Judith C. Chow, John G. Watson, Juan A. Moncada-Herrera(2008) applied a hybrid of ARIMA and ANN to develop a model to forecast air pollutants for Temuco,Chile .The wood burning, industrial and vehicular emissions is a major source of pollution in Chile during cold winters. The hourly and daily time series of PM_{10} and meteorological data during the period of 2000–2006 at the Las Encinas monitoring station in Temuco was used .Normally distributed variables using log transformation was used .Wind direction was found to be insignificant variable, as the sources of PM were more local rather than transported from other regions. The hybrid model was created using the Enterprise Miner tool of the SAS 9.1 software to predict the maximum 24-hr PM_{10} moving average concentrations at Temuco The experimental results of the study reported that the hybrid model was able to capture 100% and 80% of alert and pre-emergency episodes, respectively. The hybrid models were successful in exploiting the unique capabilities of ARIMAX and ANN in linear and non-linear modeling. The hybrid methodology was able to process the air quality prediction not only one month or a season, but also the whole year. To run the hybrid model, the SAS statistical software and meteorological and air quality observed data for the previous day

.

L. H. Tecer (2007) created a three-layered neural network trained by back –propagation algorithm to predict Sulphur dioxide(SO_2) and Particulate Matter (PM)concentration from data of two monitoring stations in Zongulak city, a coastal mining hub in Turkey.One station (Bahçelievler) was situated in an locality with hospital, residential areas and social centres. The second monitoring station was situated in area close to main traffic arterial road close to business and schools in Zonguldak district. The daily average of SO_2 and PM from two monitoring stations were taken as pollutant parameters and meteorological variables as input parameters for a a period of one year from January –December 2002.The daily data was divided into three sets . The odd days for the one year data set were used for training comprising of 175 data sets. The even days was used for testing comprising of 150 data sets. And 16 days were used for validation of the data. The performance evaluation was done by computing the determination coefficient(R^2) and correlation coefficients between the observed and predicted values. The model for SO2 concentration at Bahçelievler station reported R^2 value for training and testing data to be 0.829 and 0.668 respectively. While for PM concentrations R^2 value was found to be 0.820 and 0.808 respectively. The correlation values show significance at 0.01 levels between the observed and predicted values. The P-value from ANOVA analysis was reported to be less than 0.01, which confirms a statistically significant relationship between the variables at 99% confidence level.

S.M.Shiva Nagendra,Mukesh Khare(2006) carried out a study to investigate the use of ANN technique for modelling the complex vehicular exhaust emission for two air quality control regions in Delhi . The first site (AQCR1) was a traffic intersection and the second site (AQCR2) was an arterial road in city of Delhi. The procedure of modelling nitrogen dioxide NO2 dispersion phenomenon using neural networks was explained. The model consisted of 6 traffic characteristic variables and 10 meteorological variables. The meteorological data included 24 hour average observations of temperature, pressure, humidity, wind speed, wind direction, sunshine hours, visibility, cloud cover, mixing height. The traffic characteristics were classified into four groups, : Four wheelers gasoline powered and four wheeler diesel powered, Two wheelers, three wheelers, and source strength of CO and NO_2 which were estimated using emission factors developed by Indian Institute of Petroleum. The results show that the ANN-based NO_2 models with both meteorological and traffic characteristic inputs performed satisfactorily at both the study sites with d = 0.76, for Site 1(AQCR1) and *d* = 0.59, for Site 2 (AQCR2). The study reported marginal decrease in model performance, when only meteorological inputs had been used .The models performed poorly when only traffic characteristics were used as inputs.

Harri Niskaa,Teri Hiltunena, Ari Karppinenb, Juhani Ruuskanena, Mikko Kolehmainena(2004) tested genetic algorithm for designing Multilayer perceptron model for forecasting NO_2 24 hours in advance for Helsinki region. The input parameters considered were traffic and meteorological parameters like wind direction, wind speed, temperature, solar evaluation, friction velocity which determine the atmospheric dispersion conditions The performance assessment was done in context of general performance (IA) and exceedance performance. In all evaluated models, performance were of the order 0.11 (IA=0.89). It was observed that even small amount of hidden neurons was sufficient in the case of two hidden layers .It was reported that evolution of MLP inputs and architecture did not improve the ability of the MLP to forecast high concentrations significantly, which could be due to large under-representation training data

Dahe Jiang, Yang Zhanga, Xiang Hua, Yun Zenga, Jianguo Tanb, Demin Shaob(2003) described the development of ANN model for air pollution index (API) forecasting system for the city of Shanghai in China.Multi layer perceptron (MLP) network was used to develop model which with meteorological data as main input and would predict the next day average API values as output. The pollutants taken into consideration were PM_{10}, SO_2 and NO_2.Though the initial MLP model did not work well and had very low correlation coefficients, new model was built on modified training algorithm based on structure optimization. The new model performed reasonably well with observed API values and is now successfully being used for API forecasting in Shanghai. The authors emphasised on having sufficient number of representative data for training,atleast one year data for API forecasting. Simple structure of MLP models with optimum number of neurons worked better .The authors suggested instant training over batch training and that validation of model should be carried out simultaneously with training.

M.W. Gardner, S.R. Dorling (1999) applied Multi layer perceptron (MLP) neural networks to train and develop model for prediction of hourly NO_X and NO_2 concentration in Central London from hourly meteorological data. The hourly data of two nearby monitoring station from year 1990 -1991 was used to form one combined data set. Unsupervised and supervised training for models were done. The supervised learning of MLP showed better results. The results of study showed that the neural network performs well in comparison with regression models. The study illustrated that Multi layer perceptron networks could resolve complex, non linear patterns of source emissions without supervision. The models developed showed high correlation coefficient upto 0.96 -0.98 between the observed and predicted values and performed better than regression models developed (Shi and Harrison 1997) for the same location earlier.

Dorzdowicz et al. (1997) developed a line source neural network model for estimating hourly mean concentrations of CO in the urban area of Rosario, Italy. Eleven inputs, viz., vehicular flux in terms of vehicles per hour of cars, taxis, median vehicles,trucks and buses, wind speed and direction, solar radiation, humidity, pressure, rain intensity and temperature were used for developing three ANN-based models. The first, with 11 input variables, the second, with seven (excluding humidity, pressure, rain intensity and temperature) and the third with six input variables (excluding solar radiation, humidity, pressure, rain intensity and temperature). These models were validated for each type of network using approximately a set of 100 patterns. The results showed that model predictions were comparable.

CHAPTER 3
AN OVERVIEW OF ARTIFICIAL NEURAL NETWORK

3.1 GENERAL

Artificial neural network is a computational technique that is inspired by the structure, processing method and learning ability of human biological neural network. The history of artificial neural networks began in 1943 with the pioneering work of Warren McCulloch and Walter Pitts, who proposed a simple artificial model of the neuron. The modified and improved version of that model is what artificial neural networks are based on today. The back propagation concept and the updating learning of synaptic weights concept proposed by Rumelhart and Hebbs were path breaking works in the field of Artificial Neural Networks
Over the past few years ANN has grown in its applications in fields of science and engineering. They are proven and powerful method of mathematical modelling in research and practical applications. Neural networks are well suited for modelling non-linear relationships and pattern recognition from a given set of historical data. The review of ANN application in atmospheric sciences was pioneered by Gardner and Dorling . They concluded that artificial neural network give better results than linear regression methods in air modelling.

3.2 INTRODUCTION TO ANN

3.2.1 BASIC FEATURES OF ANN

The human brain consists of 10-500 billion neurons which are the fundamental units of human biological nervous system. A neuron is made up of a cell body, an axon and dendrites as shown in Figure 3.1 .

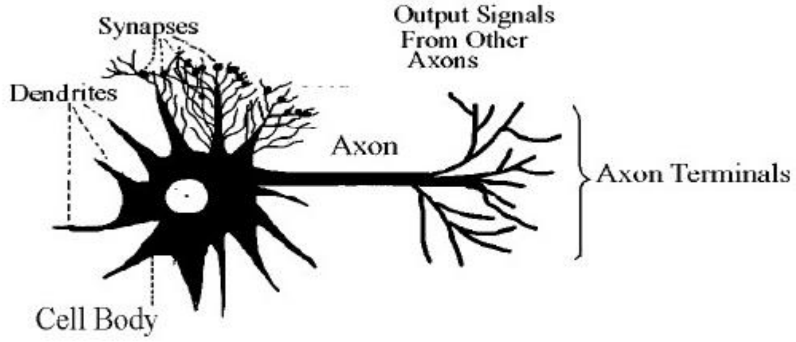


Figure 3.1 : A typical biological neuron

Each neuron is further connected to 100-10000 other neurons and the connections between neurons are called the synapses. A neuron emits an output to the other connected neurons when subjected to a stimulus. The signals are received and processed by other neurons through the path called dendrites. Neurons follow the process of binary numbers. If signal is strong it generates output signals which are transmitted through axon to synapses. As these signal pass through network, different levels of activation are created in neurons. The amount of signals transmitted depends upon the strength of synapses at junction of neurons. During the learning processes of brain the synaptic strengths are modified. These synaptic weights act as memory unit for each inter-connection. The identification and recognition depends on the level of activation of neurons.

The ANN is an assembly of interconnected processing units or artificial neurons whose functionality is loosely based on the human biological neural network. The processing ability of the network is stored as the synaptic weights or inter connections strength obtained by the process of learning from training patterns of data sets. The **Figure 3.2** shows the working of an Artificial Neuron.

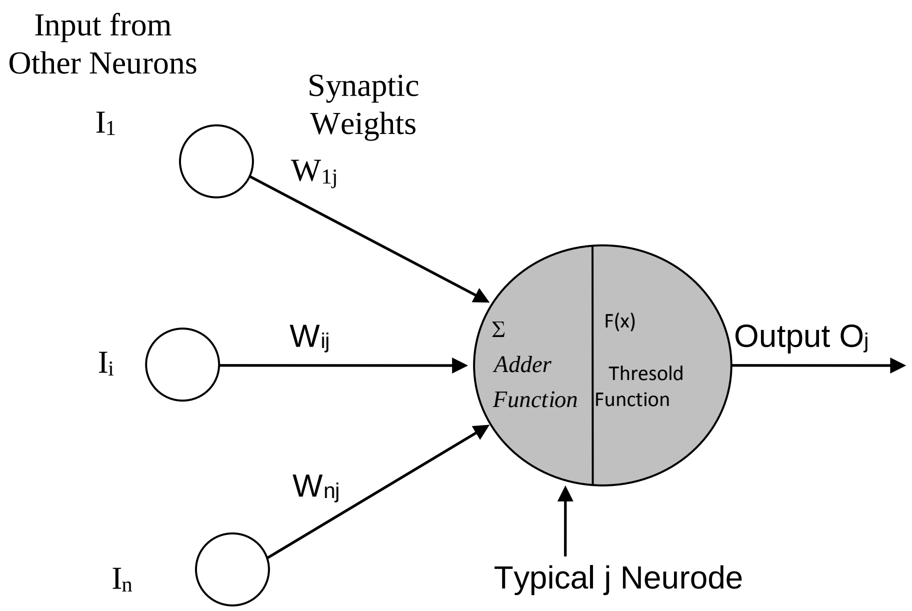


Figure 3.2: Working of an Artificial Neuron (Faghri and Hua, 1993)

MATHEMATICALLY

If $I_1, I_2 \ldots I_n$ are input values of a problem and corresponding synaptic weight values are W_{lj}, $\ldots W_{nj}$, net_j is the summation (over all the incoming neurons) of the product of the incoming neuron's activation and the synaptic weight of the connection at the typical j^{th} neurode expressed as $\sum I_i W_{ij}$. A threshold value is incorporated into output. Threshold value θ_1 is incorporated into the output. Thus the resultant becomes

$$net_j = \sum_{i=1}^{n} I_i W_{ij} + \theta_i \qquad \textbf{Eqn. (3.1)}$$

Where n is the number of input neurons, X is the vector of input neurons W is the vector of the synaptic weights and θ is node threshold, usually taken as the negative weight from the bias unit (a unit whose output is unity)

$$\text{OUTPUT} = f\,(\text{net}_j) \qquad \textbf{Eqn. (3.2)}$$

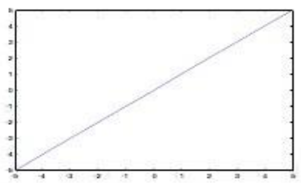
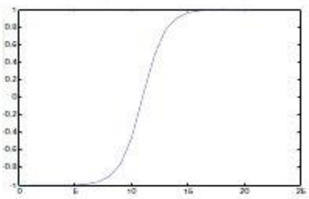
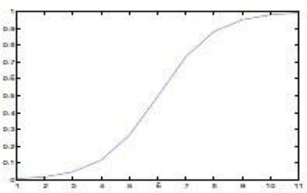
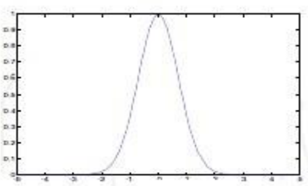
Where f (net) = Activation function

3.2.2 TRANSFER FUNCTION

Transfer function is also called as threshold function, activation function or squashing function used to map the input pattern of neuron to the specified output range. The most commonly used threshold functions are listed in **Table 3.1**. In this study, sigmoid function has been used. The sigmoid function has many advantages some of them are

- It is semi linear
- It has biological basis
- It has greatest slope for inputs equal to zero.

Table 3.1 Transfer function in MATLAB **(Maitha.H. et al. ,2014)**

Function Name	Graphical Illustration	Mathematical form
Linear		$f(x) = x$
Hyperbolic Tangent Sigmoid		$f(x) = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$
Logistic Sigmoid		$f(x) = \dfrac{1}{1+e^{-x}}$
Gaussian RBF		$\varphi_j(x) = \exp\left(-\dfrac{1}{2\sigma_j^2}\lVert x - x_j \rVert^2\right)$

3.2.3 TYPES OF ARTIFICIAL NEURAL NETWORK

Artificial Neural Networks can be classified into two categories based on the connection pattern viz.Feed forward neural network and Recurrent or Feedback neural network.

i)Feed forward network

In feed forward neural network, the activation is fed forward from input through hidden layers to output as shown in **Figure 3.3**. In feed forward network graphs have no loops. Feed forward networks are static and they produce only one set of output values from a given input. The responses of these networks are independent of input from the previous network state. The network can be a single layer with input and out or multilayer with input, hidden and output layer . The inputs are connected to other neurons of the hidden layer by their corresponding weights. The output from the last layer is considered as output of the network. The feed forward method involves selection of best number of neurons by experimentation. This is done by adding neurons and checking the network performance till it starts deteriorating. Neural network undergoes training and learning with the data before it is actually applied. The weights of the connecting neurons are adjusted to achieve acceptable value of error**(K.Anitha Kumari et al. , 2013).**

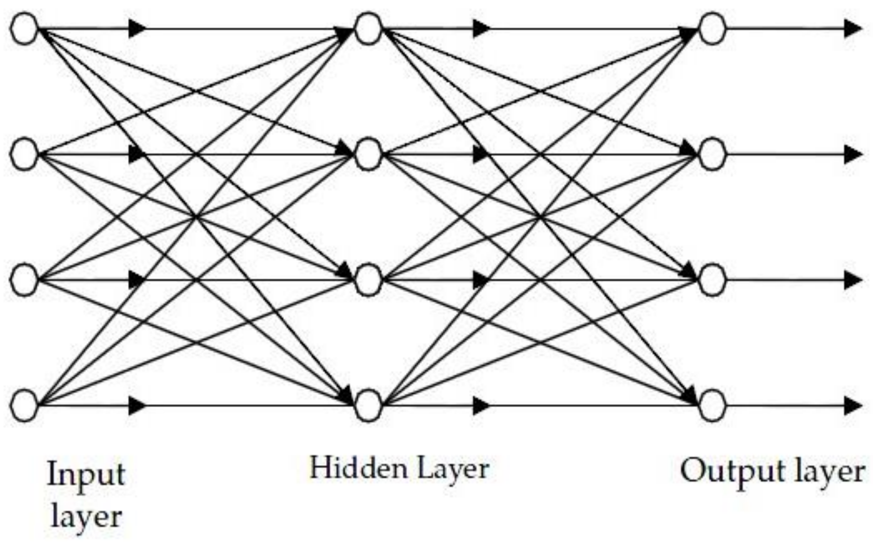


Figure 3.3: A multi layer feed forward network

ii) Recurrent network

Recurrent network or feedback neural network has one or more hidden layers with at least one feedback loop as shown in **Figure.3.4**. The feedback can be a self feedback that is the output of a neuron is fed back in the network to its own input. There are different types of training algorithms in recurrent network but its performance has not been so satisfactory. The practical and theoretical difficulties are the main obstacles in its practical applications.

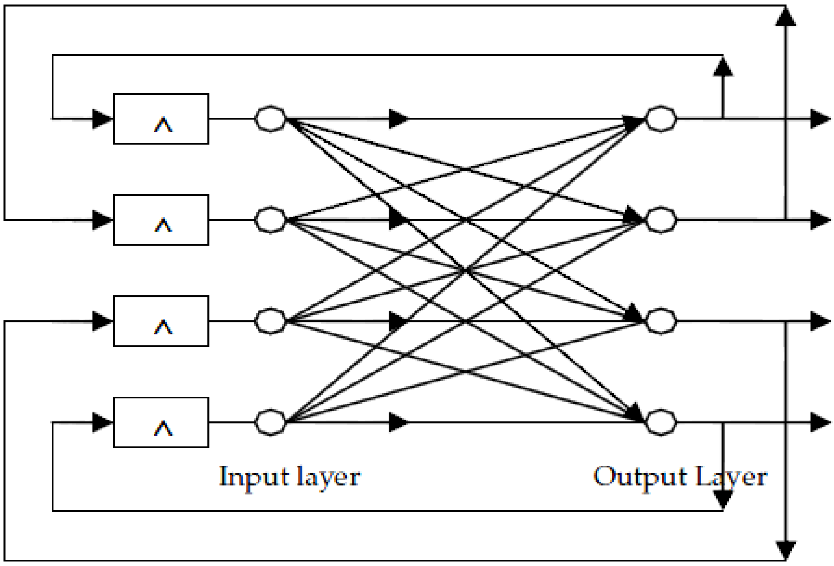


Figure3.4: Recurrent network

For different types of problems, different type of architectures of ANN can be utilised. The **Figure 3.5** gives a overview of the different type of feed forward and recurrent network structures.

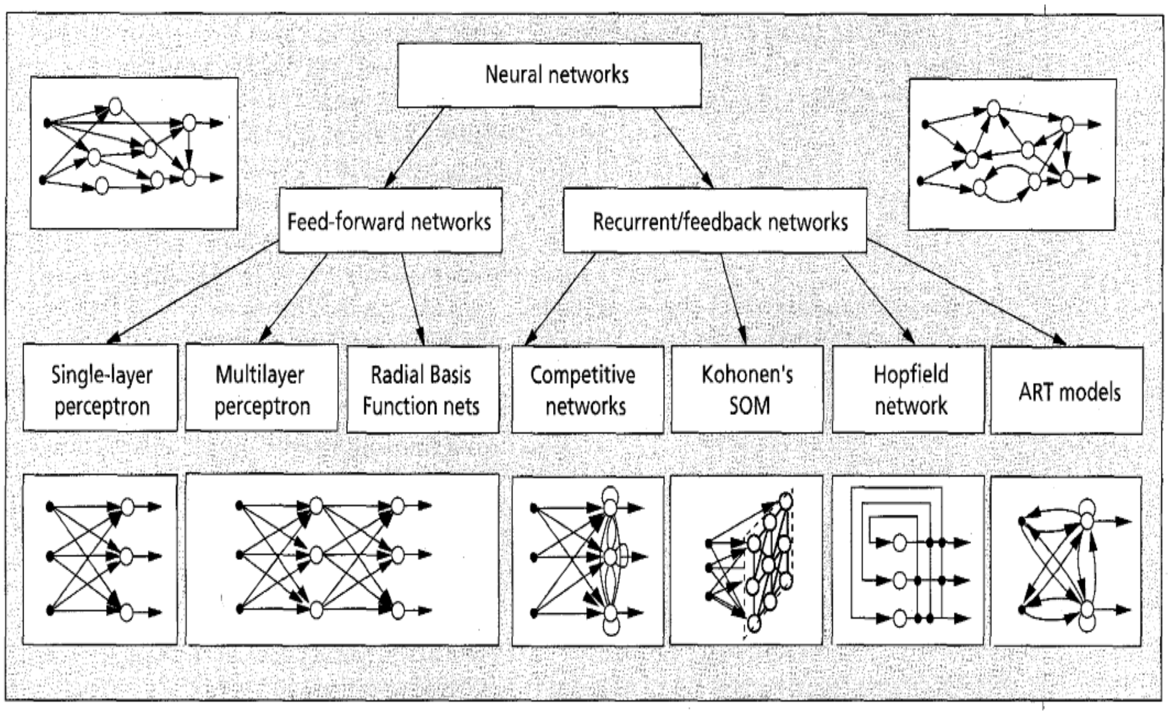


Figure 3.5: Different types of feed forward and recurrent network architectures. (**A.K.Jain,Mao, 1996**)

3.2.4 LEARNING IN NETWORK

The main phase in artificial neural network is learning. The programming for adjusting the network weights is called the training and the training effect is called the learning. The learning can be done either by being given weights processes from set of training data or by adjusting the weights automatically according to some criterion. The network usually learns the connection weight from available training pattern in the data set. Performance is improved by updating the weight in the network iteratively. The learning of ANN generally classified as

Supervised Learning: Also called as learning with teacher, ANN is trained from a set of input output data. The weights are then adjusted to minimize errors in the outputs. Separate set of input-output data set, called the testing data, is provided to test the effectiveness of training.

Unsupervised Learning: Unsupervised learning also known as self organized learning network, trains on the set of input data. The network learns similarities of system and doesn't have error correction or adjustments done according to actual output data (Teacher).

3.3 MULTI-LAYER PERCEPTRON NETWORK USING BACK PROPAGATION LEARNING METHOD

3.3.1 GENERAL

Multi-layer perceptron is a system of layered inter connected neurons which have three arranged layers : Input Layer : Hidden layer (one or more than one), the output layer . The nodes in the three layers are interconnected to nodes in neighbouring layer. Multi layer perceptron is a feed forward type of neural network. The output of the nodes are scaled by connecting weights and fed forward as input to nodes of the next layer in network with forwards direction of information processing. The following equation for MLP was given by MaithaH.Al.Shamisi et al (2011).

$$y_j = f\left(\sum_{i=1}^{n} w_{ji}\, x_i\right) \qquad \textbf{Eqn. (3.3)}$$

Where n is number of neurons (i=1,2,..n).in hidden layer .Each neuron j sums up input signals x_i after scaling the synaptic weights with the strength of respective connections wji from the input layer and computation of the output y_j as the function f of the sum.

MLP is a supervised learning method where it learns through training of data with series of input and associated output.The error signal computed during training is used to determine the degree to which the synaptic weights in network should be adjusted to reduce the overall error to desired level. Back propagation is the mostly employed supervised learning method used for training of MLP. The main logic of the back propagation algorithm is to minimize error difference between the actual input value and the output from the model.

The back propagation algorithm consists of two phases namely forward and backward pass. The forward pass computes the output of network by propagating the input data through hidden layer in the network. The network output is compared with the desired output to compute the error using backward pass and connection weights are then modified to reduce the error in target during the backward pass. The steps involved in back propagation algorithm are briefly discussed below (**Ravi Sekhar, 1999 and 2009**)

3.3.2 BACK PROPAGATION ALGORITHM

Step 1 Network Topology: Network topology includes selection of input nodes which is equivalent to number of explanatory variables considered in modelling, selection of hidden nodes varies from model to model and selection of output nodes which is equivalent to choice of available mode. Figure 3.6 gives the structure of multi layer feed forward back propagation model for modelling relation between meteorological, traffic variables and pollutant concentration.

Step 2 Forward Pass: Compute the activation at hidden and output layer. Computation of error in each pattern.

a) Computing the hidden layer neuron activation (OH_{pj})

$OH_{pj} = F(net_{pj})$ $net_{pj} = \sum W_{ji} I_{pi}$ **Eqn.(3.4)**

Where F( ) is Sigmoid transfer function, $F(a) = 1/1+ e^{-a}$

W_{ji}= Weights from input i to hidden j node, I_{pi} =Value of input node i for the pattern p

b) Compute utput layer neuron activation (OO_{pk})

$OO_{pk} = F(net_{pk})$ $net_{pk} = \sum W_{kj}OH_{pj}$ **Eqn.(3.5)**

W_{kj}= Weights from the hidden node j to the output node k

c) Compute the error for each of the pattern at each output node (E_{pk})

$E_{pk} = 1/2 \sum (T_{pk}-OO_{pk})^2$ **Eqn.(3.6)**

Where T_{pk}=Target value of the output node k for the pattern p

OO_{pk}= Actual value of the output node k for the pattern p

Step 3 Backward Pass: computation of error signal and adjustments of weights between Output to Hidden and Hidden to Input Layer.

a) Computation of the error signal (δ_{pk}) at output layer and adjustment of weights between Output to Hidden Nodes

$\delta_{pk} = (T_{pk}- OO_{pk}) OO_{pk} (1-OO_{pk})$ **Eqn.(3.7)**

$W_{kj}(New) = W_{kj}(old) + \eta \delta_{pk}OH_{pk}+ \alpha[W_{kj}(old)- W_{kj}(old-1)$ **Eqn.(3.8)**

b) Computation of the error signal (δ_{pj}) at the hidden layer and adjustment of synaptic weights between Hidden and input Nodes.

$\delta_{pj} = OH_{pj} [1- OH_{pj}] \sum \delta_{pk} W_{kj}$

Eqn.(5.13)

$W_{ji}(New) = W_{ji}(old) + \eta \delta_{pj} I_{pi} +\alpha[W_{ji}(old)- W_{ji}(old-1)]$ **Eqn.(3.9)**

Step 4 Minimisation of Error: All the above steps are repeated with new training patterns. Best network is decided based on the minimum average MSE obtained during training of neural network.

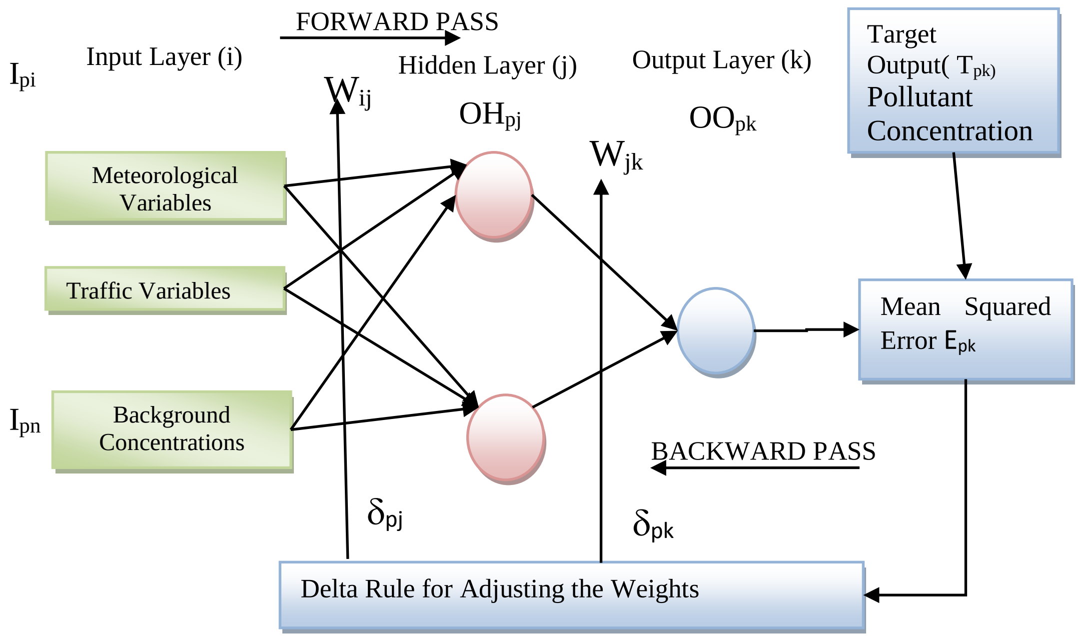


Figure 3.6 : Structure of Multilayer feed forward back propagation ANN for modelling

MLP can model non-linear interactions between input variables only if the transfer function is non –linear. Hence sigmoid transfer function is used for air modelling. To train the multilayer feed forward back propagation network, MATLAB neural network tool box(**MATLAB 2012b**) has been used.

3.3.3 STRENGTH AND LIMITATION OF BACKPROPAGATION LEARNING METHOD

The main strength of back propagation learning method is that it maps the input to output patterns by training, without any prior knowledge. It is simple to implement and can be applied to any network. The major limitations of back propagation method is given below.

- The training of the network may require large number of iterations.
- It may have many examples for correct pattern for a particular data set.
- It can get stuck in local minima to give sub optimal solutions.
- Network paralysis occurs when the adjusted weights are having very large values during training. Large weights forces most of the units to operate at extreme values, in regions where derivative of activation function is very small.

- Multilayer neural network requires large samples to map the input patterns and the network weights need to be adjusted before the network is able to yield an optimal solution**.(Hegazy et al . 1994)**

3.4 BENEFITS AND LIMITATIONS OF ARTIFICIAL NEURAL NETWORK OVER OTHER MODELLING TECHNIQUES

Artificial Neural networks is a promising tool for modelling, simulation and forecasting . In order to utilise ANN it is necessary to understand its potential and limitations. ANN can provide a alternative or complement the traditional modelling techniques. The benefits and limitations of ANN are listed below.

BENEFITS

- **Effective in non-linear environment**: Neural networks can map highly non –linear relationships between input and output variables. Its has been successfully applied in situations where the input and output variables have highly complex non-linear relationship. ANN are very flexible for modelling in variable environment
- **No prior knowledge**: It is a non –parametric approach of modelling which does not need any prior knowledge of functional relationship between the input and output datasets.
- **Better performance**: The performance of artificial neural networks has been tested and proven to be better than other statistical modelling techniques like multivariate regression models by studies conducted by researchers all over the world. (**A.K.Jain &Mao,1996**).

.

LIMITATIONS

- **Training Time**: Neural network may sometimes require long training time which makes it impractical .A minimum of 100 iterations is required for simple problems. However with improvement in computational abilities of computers training time is not a problem now.

  Large Training data: ANN is best suited for models with large training data for optimum results. Also the number of data sets required for training of network to produce good results is not predefined. No guarantee of results. Though most training techniques tune the network, they do not guarantee of operating properly in different conditions. The training can bias a network making it inaccurate in some working areas in addition to getting trapped in local minima..

- **Selection of model parameters**: The inputs parameters need to be selected carefully to give proper output mapping. The number of layers and neurons for best network needs to be experimented to get best performance model. As there is no predefined method of selection trial error and experimentation need to be done before actual building of the neural network. (**N.Sharma et al. ,2005)**

Chapter 4
METHODOLOGY

To achieve the stated objectives presented in Chapter 1, the development of ANN models for the study area was done in a planned and systematic manner. The study methodology is presented in the form of flow chart as shown in **Figure 4.1**.

4.1 LITERATURE REVIEW

As seen from the methodology presented in Figure 4.1, the study was initiated by literature review on work done in air pollution forecasting using ANN. It helped in getting a good understanding of the air modelling done using Artificial Neural Networks in India and abroad.

4.2 DATA COLLECTION AND EXTRACTION

After having the basic background of the research topic from the literature review, the meteorological data was collected from Indian Meteorological Department, Delhi and Traffic data was extracted from the analysis of survey done by CRRI,Delhi details of which are given in Chapter 5 .The collected data sets had a lot of missing and erroneous data .The datasets were pre-processed before building the models. In due course of developing models data had to be pruned to build model with good performance.

4.3 SELECTION OF INPUT PARAMETERS

The most important step in developing ANN models is to select the input variables for the models that have the most significant impact on performance of the model (Faraway and Chatfield 1998). The selection of parameters can be done in three ways

- Selection based on a previous studies of the characteristics of potential input variables
- Training neural networks with different combinations of input variables and selecting parameters that give best performance of network.
- Using linear dependency techniques like correlation

While developing a air pollution model based on vehicular exhaust emissions two aspects need to be considered the most: conditions of pollutant dispersion and sources of pollution. Meteorology is a very dominant factor which influences dispersion, distribution and concentration various air pollutants. Wind speed, wind direction, atmospheric pressure, temperature, relative humidity, rainfall and solar radiation are closely associated with transformation and dispersion of pollutants. These meteorological parameters can be easily and routinely measured at all meteorological stations. Therefore, these meteorological parameters were selected as input factors. As the study period the days with rainfall were very less to be taken as influential factors, hence it was omitted from input parameters. The wind direction data also had a lot of missing data and was not suitable for training of networks so it was also excluded. As discussed in Chapter 1 vehicular exhaust emissions are based on the traffic characteristics like traffic volume, vehicle type and fuel type of vehicles. Traffic data is divided into seven types: two wheelers, three wheelers, four wheelers diesel vehicles, four wheelers petrol vehicles, four wheelers CNG vehicles, diesel trucks and buses and CNG buses. Only weekdays were considered as traffic data of weekends was not available. Based on these considerations and correlation values the different combinations input variables were selected to give best model performance. The correlation of PM_{10}, $PM_{2.5}$ and NOx and CO were strong with each other. These concentrations have been considered as background concentrations and taken as input parameter in separate models.

4.4 BUILDING THE NETWORK

The number of hidden layer, neurons, training function, and transfer function of each layer, partitioning of data and performance function need to be defined suitably before development of models. After reviewing different ANN structures and their suitability for air modelling in Indian conditions, multilayer perceptron with back propagation learning algorithm was selected was modelling. The optimum number of neurons for each model was determined by doing sensitivity analysis on Neuro Solutions version 6 software. The graph of average of minimum MSE is used to select number of neurons to give highest network performance. MLP can model non-linear interactions between input variables only if the transfer function is non –linear. After performing trial and error using different activation functions sigmoid transfer function was selected for modelling in this study as it gave the best results. To train and develop the multilayer feed forward back propagation network, MATLAB neural network tool box(**MATLAB 2012b**) was used.

4.5 TRAINING THE NETWORK

The weights are adjusted during training process in order to make the actual values close to the predicted outputs. In this study three month hourly data from January 2013 –March 2013 was used for training. The available data was divided into three data sets randomly in MATLAB.

1. Training set to adjust the weights of the neural network.
2. Testing set to measure the ability of model to generalize at regular intervals during training
3. Validation set for performance evaluation of the trained ANN model in prediction of pollutant concentrations.

. The training, testing and validation sets were randomly partitioned as 75 %, 15% and 15% of the datasets. The sigmoid function was used as the transfer function .

4.6 PERFORMANCE EVALUATION OF THE MODELS

For evaluating the accuracy of developed models statistical parameters were computed. The statistical parameters considerd were: Mean absolute percentage error, Root mean square error (RMSE) mean absolute error (MAE) and coefficient of determination (R^2).The equations for calculating these statistical parameters are given below:

i) MAPE (Mean Absolute Percentage Error)

MAPE expresses error as a percentage. It gives the mean of absolute error in percentage of the predicted values in comparison with observed values.

$$\boldsymbol{MAPE} = \frac{1}{n}\sum_{i=0}^{n}\frac{|Ai-Pi|}{|Ai|} \times 100 \qquad \textbf{(Eqn.4.1)}$$

Ai is the observed value and *Pi* is the predicted value and n is the number of observations

ii) MAE (Mean Absolute Error)

MAE is the measure of absolute error in the forecast of individual values with the actual values. It is suitable for comparing the absolute error between different models.

$$\boldsymbol{MAE} = \frac{1}{n}\sum_{i=0}^{n}|Ai - Pi| \qquad \textbf{(Eqn.4.2)}$$

Ai is the observed value, *Pi* is the predicted value and n is the number of observations.

iii) RMSE (Root Mean Square Error)

RMSE value signifies the standard deviation of the differences between actual and predicted values of the data set. These individual differences between observed values and predicted values are called residuals . The RMSE aggregates the magnitude of the errors in predictions into a single measure of predictive power. RMSE is a good statistical measure of accuracy for comparing forecasting errors of different models for a particular variable.

$$\boldsymbol{RMSE} = \sqrt{\sum_{i=0}^{n} \frac{(Pi - Ai)}{n}} \qquad \textbf{(Eqn.4.3)}$$

Ai is the observed value; *Pi* is the predicted value and n number of observations.

iv)Coefficient of Determination (R²)

Correlation of determination is a value which indicates goodness of data fit in a statistical model. It is the square of correlation coefficient R.It is defined as the square of covariance of the variables divided by the product of the standard deviations of the sample.

$$\boldsymbol{R^2} = \left(\frac{\sum(P - \bar{P})(A - \bar{A})}{\sqrt{\sum(P - \bar{P})^2 \sum(A - \bar{A})^2}}\right)^{\mathbf{2}} \qquad \textbf{(Eqn.4.4)}$$

The performance of the models were analysed based on the computed MAPE, MAE, RMSE and coefficient of determination values. In addition to this the mean and standard deviation of observed concentration and predicted concentrations of different models were also compared to analyse over and under prediction by the ANN networks.

4.7 INTERPRETATION OF WEIGHTS

ANN lacks method for the interpretation of weights of input variables. This section emphasizes on the innovative method for using neural network for casual analysis of input parameters using partitioning of weights algorithm.

Partitioning of weight algorithm

The method of partitioning of connection weights to find out the relative significance of inputs from the weights of input to hidden layer and hidden to output layer was proposed by Garson in 1991 .It involves partitioning of the weights of connection of each neuron into components associated with the input neuron. The computation algorithm of partitioning of weights for neural network used in this study is described below.

i) The absolute value of connection weights of hidden to output layer is multiplied to the absolute value of input to hidden layer connection weight for each hidden neuron.

This is computed for each of the input variables in order to obtain the product matrix Pjki .

ii) Each hidden neurode,P_{jki} is divided by the sum for all of the input variables to get the matrix Q_{jki}. Example: For a hidden neuron 1. Q11= P_{111}/(P_{111}+ P_{112}+….).

iii) The sum of the product S_1 for each input neuron is formed from previously computed $Q_{jki.}$

$$Si = \sum_{j=1}^{NHN} \sum_{k=1}^{NON} Qjki$$

iv) S_1 is then divided by sum of all the input variables which expressed as percentages to give the relative influence of all output parameters attributed to the given input parameter.Example,for an input neuron 1,the relative importance is equal to (S_1 x 100)/(S_1+ $S_{2+\ldots}$ $S_{i)}$

The results computed from this algorithm are discussed in Chapter 6. Results and Discussions.

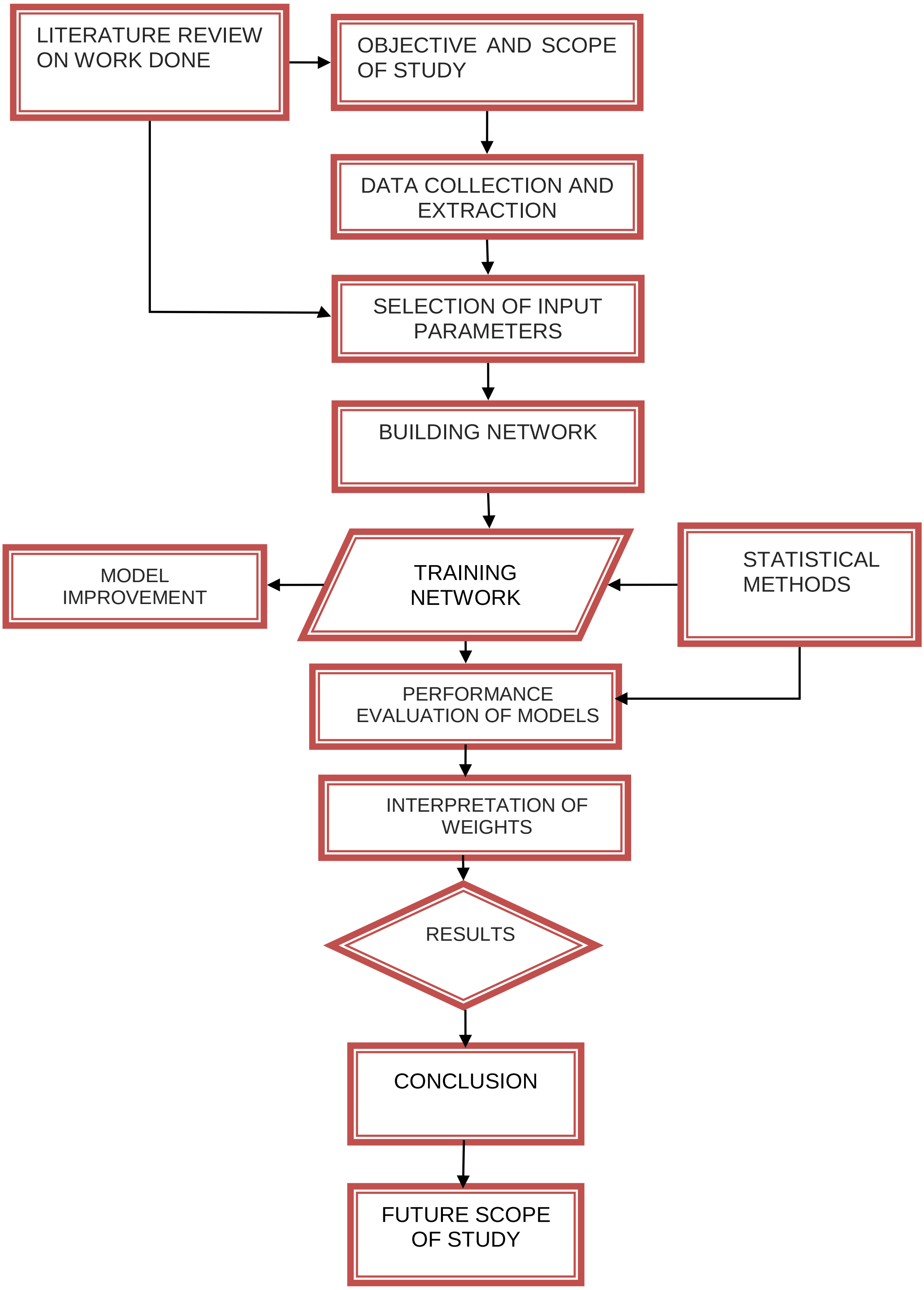


Figure 4.1: The flowchart of study methodology

Chapter 5
STUDY AREA

5.1 SITE DESCRIPTION

The site selected for the study is Central Road Research Institute, located at latitude 28°33' 3.7" N and longitude 16°30.1'E in South Delhi. The study area lies prominently along one of the busiest and important national highway (NH2) that connects the states of Delhi, Uttar Pradesh Haryana,, Bihar, Jharkhand, and West Bengal. There are many residential colonies near CRRI which also lie along the NH2.The Figure 5.1 below shows the map of study area.

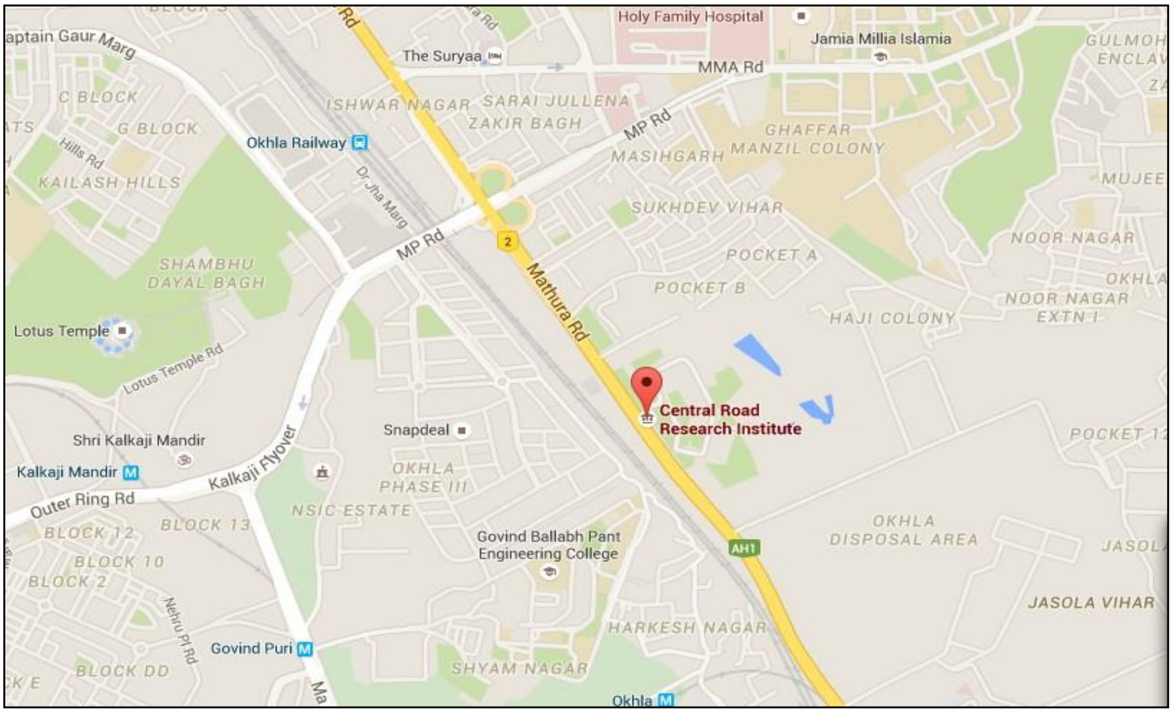


Figure 5.1 : The map of study area :Central Road Research Institute Delhi –Mathura Road NH-2

The NH 2 which lies in the front of CRRI connects Delhi to Agra, Faridabad other towns in Uttar Pradesh road and has one of the highest traffic densities in the region with a lot of interstate and local city traffic. Agra is one of the top tourist destinations of the country. Faridabad is an important industrial township of Haryana and as it is close to the national capital it is part of the national capital region. The region has a number of residential townships. This stretch is the only direct link from Faridabad to Delhi and has a very high traffic density with over sixty thousand vehicles plying daily. This area was suitable for studying the effect of vehicular emissions on air quality and developing air pollutant model based on vehicular exhaust emissions.

The study area is also in close proximity with Okhla. Industrial Area and Timarpur municipal waste incinerator plant. The Okhla industrial area located at 2.5 km from the monitoring station mainly has pharmaceutical, machinery and plastic manufacturing units, packaging industries,exporters,call centers,MNC's,offices,banks and other social complexes.Timarpur municipal waste incinerator plant is located around 1 km on the backside of CRRI complex.

5.2 MONITORING STATION

The current study is based on data collected at single monitoring station in CRRI.The monitoring station has been set up and maintained by Indian Meteorological Department, Lodhi Road.

Figure 5.2 : Monitoring station inside CRRI,Delhi premises

The real time monitoring of pollutants levels of $PM_{2.5}$, PM_{10}, NOx, SO_2, CO and ozone is displayed on monitor near the station. The details of monitoring instruments and parameters analysed is given in detail in next section.

5.3 MONITORING INSTRUMENTS USED FOR MEASURING POLLUTANTS

i) $PM_{2.5}$ AND PM_{10}

Equipment: Thermo Fisher FH62C14

Method of measurement: Gravimetric, TOEM, Beta attenuation

Principle: Radiometric particulate mass monitor provides real-time measurements. The FH62C14 measures mass concentration of PM_{10}, $PM_{2.5}$ in ambient air in real-time. The FH62C14 monitor displays time-averaged measurements of integral beta attenuation mass sensor.

ii) NO_X

Equipment : ThermoFisher Model 42i Trace Level

Method of measurement: Modified Jocob and Hochheise (SodiumArsenite) Chemiluminescence

Principle : Ozone (O_3) and nitric oxide (NO) react to emit characteristic luminescence with an intensity which is linearly proportional to NO concentration. Infrared light emissions are emitted from decaying of electronically excited NO_2 molecules to the lower energy states. The ozone required for the chemiluminescent reaction is generated by the ozonator. Ozone reacts with NO in the sample in reaction chamber to produce excited NO_2 molecules. A photomultiplier tube (PMT) detects the luminescence produced during this reaction. The NOx and NO concentrations are computed in the NOx and NO modes and stored in memory. The difference between the NOx and NO concentrations is used to compute the NO_2 concentration.

iii)CO

Equipment: Model 48i Gas Filter Correlation CO Analyzer

Method of measurement: Non Dispersive Infrared (NDIR) spectroscopy

Principle : Carbon monoxide absorbs the infrared radiation at a wavelength of 4.6 microns. The equipment uses internally stored calibration curve to linearize the instrument to measure and give output over any range upto concentration of 10,000 ppm.

5.4 DATA COLLECTION AND EXTRACTION

In the current study the hourly meteorological variables: Wind speed, gust wind, relative humidity, atmospheric pressure temperature, and solar radiation and hourly pollutant concentrations of $PM_{2.5}$, PM_{10}, NO_X and CO for the period of January 2013 to March 2013 were considered. The meteorological data and hourly concentrations of pollutants was provided by Indian Meteorological Department, Lodhi Road. The hourly traffic data of study area was provided by Central Research Institute, Delhi. A total of 2090 data sets were compiled but a lot of outliers, missing and erroneous data had to be removed before using in ANN for modelling .

The statistical properties of the variables in model is given in **Table 5.1** below.

Table 5.1 : Statistical properties of input and output variables in ANN model

PARAMETER	MEAN	MEDIAN	MINIMUM	MAXIMUM
WIND SPEED ($m\ s^{-1}$)	0.27	0.00	0.00	4.50
GUST WIND($m\ s^{-1}$)	5.61	0.00	0.00	58.80
TEMPERATURE(℃)	18.15	17.80	4.20	34.10
HUMIDITY (%)	56.29	60.00	14.00	88.00
SOLAR RADIATION($W\ m^{-2}$)	76.14	0	0.00	736.00
PRESSURE	989.76	990	980.00	998.00
PM_{10} ($\mu g/m^3$)	274.89	237.06	6.62	934.02
$PM_{2.5}$ ($\mu g/m^3$)	159.36	133.62	4.26	604.65
NO_X ($\mu g/m^3$)	312.69	271.57	7.77	964.30
CO ($\mu g/m^3$)	3131.98	2656.50	11.45	13018.10

(Source: Based on data from Indian Meteorological Department, Delhi)

The figures below give the variation of pollutant concentrations with the National Ambient Air Quality Standards (NAAQS) during the study period.

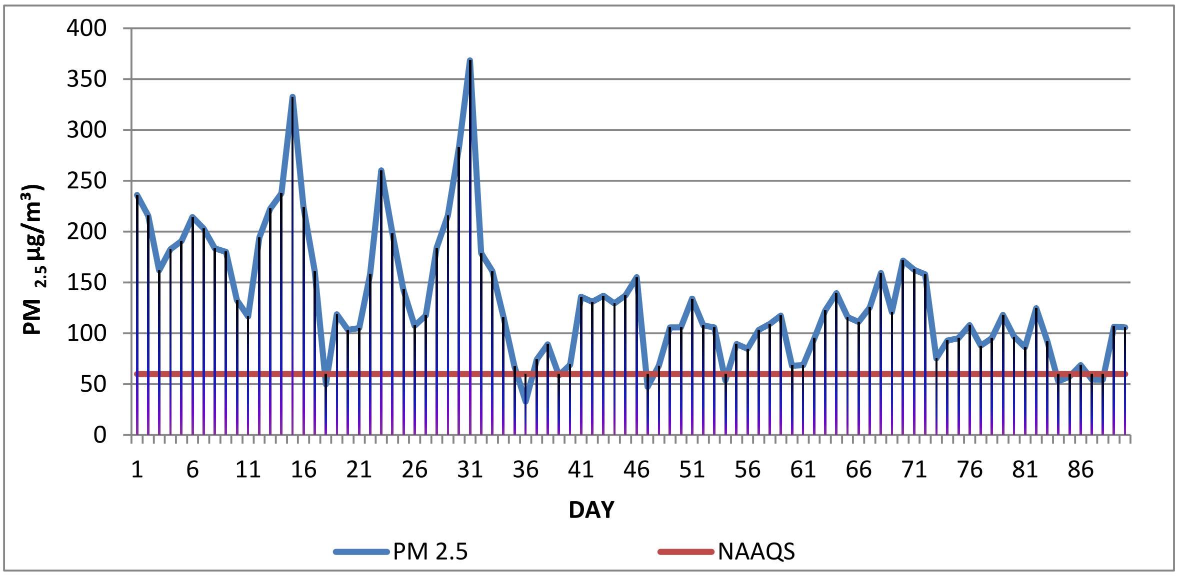


Figure 5.3: Variation of daily average concentration of $PM_{2.5}$ with 24 hour NAAQS

(Source: Based on data from Indian Meteorological Department, Delhi)

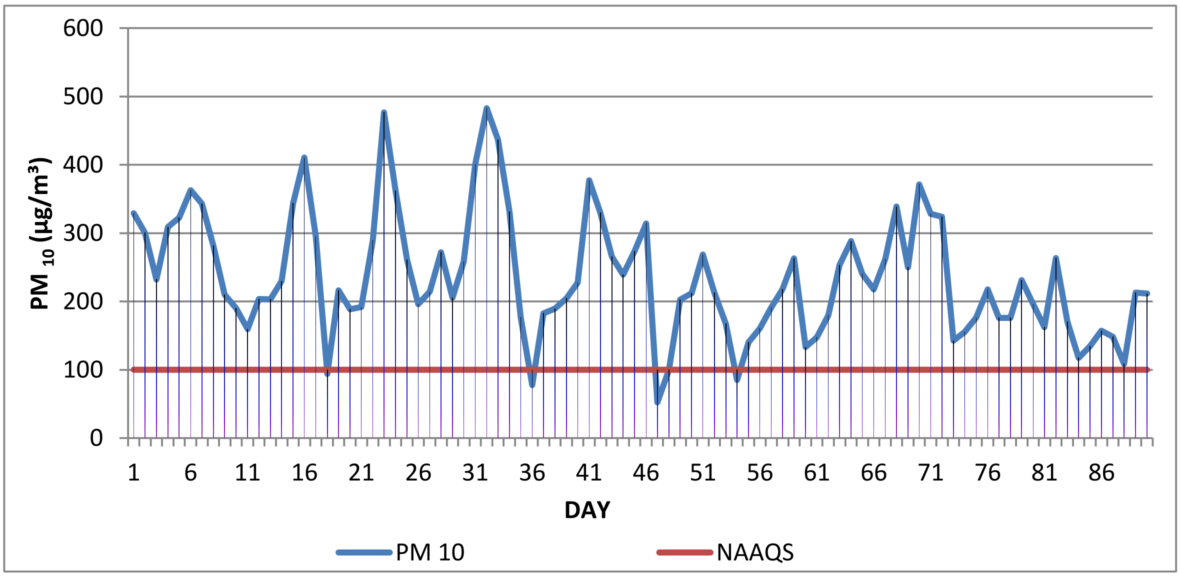


Figure 5.4 : Variation of daily average concentration of PM_{10} with 24 hour NAAQS

(Source: Based on data from Indian Meteorological Department, Delhi)

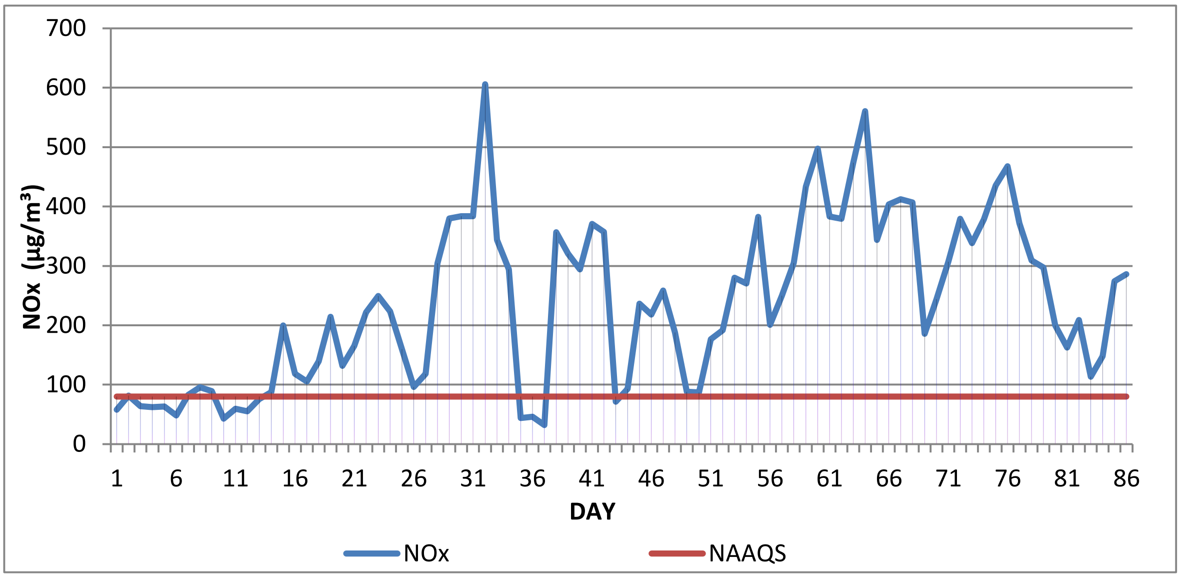


Figure 5.5: Variation of daily average concentration of NOx with 24 hour NAAQS

(Source: Based on data from Indian Meteorological Department, Delhi)

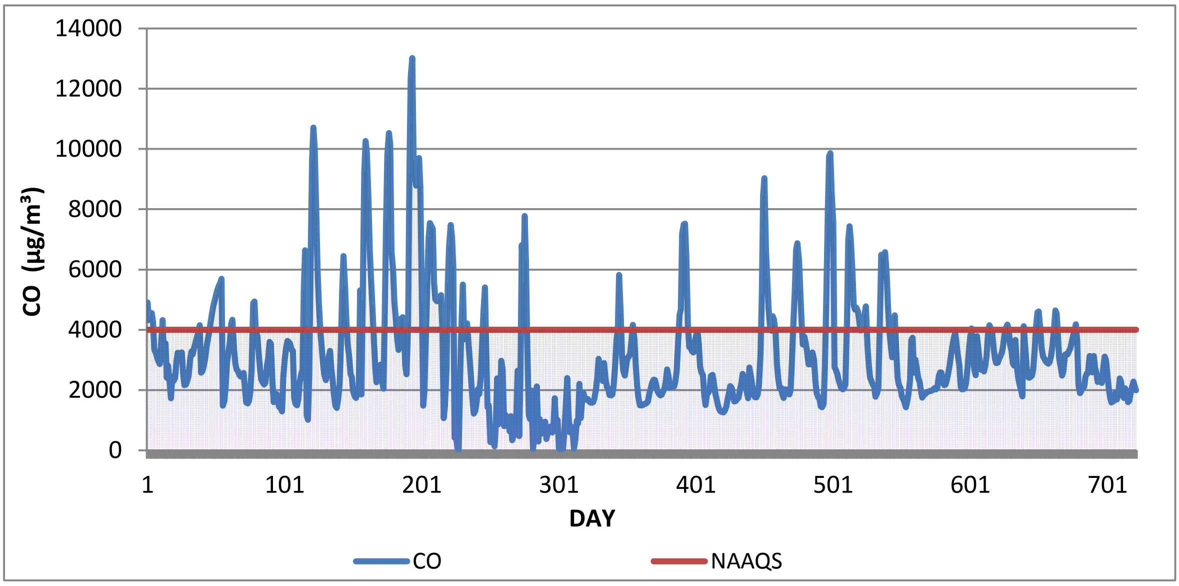


Figure 5.6: Variation of hourly concentration of CO variation with NAAQS standard

(Source: Based on data from Indian Meteorological Department, Delhi)

The concentrations of $PM_{2.5}$ and PM_{10} exceeded permissible limits continuously in the study period. The days in which 24 hour average concentration exceeded the acceptable limit for $PM_{2.5}$ and PM_{10} was found to be 94.4 and 90 percentile respectively. The study period also saw high concentration of NOx in ambient air with nearly 85.5 percentile of days exceeding the permissible standard of 80 µg/m³. The CO concentration was high but with comparatively lesser incidences of exceeding the permissible limit than the other pollutants.

Chapter 6
RESULTS AND DISCUSSIONS

6.1 GENERAL

Previous studies on air modelling of vehicular exhaust emissions done using ANN reported that best model performance is achieved by using all meteorological parameters (Wind speed, gust wind speed, atmospheric pressure, temperature, relative humidity and solar radiation) and traffic characteristics (Two wheelers, Three wheelers, four wheelers petrol, four wheelers diesel car etc.). The models developed in this study use both meteorological and traffic data as input. The correlation of the different input variables with pollutants during study period is given below.

Table 6.1: Correlation matrix of pollutants with meteorological parameters

POLLUTANTS	WIND SPEED	GUST WIND	TEMP.	RELATIVE HUMIDITY	PRESSURE	SOLAR RADIATION
$PM_{2.5}$	-0.32	-0.28	-0.52	0.50	0.40	-0.21
PM_{10}	-0.38	-0.32	-0.46	0.52	0.35	-0.28
NOx	-0.25	-0.24	-0.11	0.21	0.03	-0.17
CO	-0.27	-0.27	-0.25	0.28	0.19	-0.18

Table 6.1 shows the correlation of meteorological parameters on the pollutants. The high concentration of pollutants during study period were due to low wind speed,temperature,solar radiation and high relative humidity and pressure as seen from the correlation matrix. Temperature, relative humidity and pressure shows a strong influence on the particulate matter concentration in the study area than on NOx and CO concentrations.

The negative values of correlation of wind speed, gust wind, temperature and solar radiation suggest negative influence of these parameters on pollutants.The negative values of correlation of wind speed can be explained as speed increases, the concentrations of pollutants decreases due to dilution and dispersion. The temperature during study period was low, resulting in high concentration of pollutants during study period due to temperature inversion ,which explains the negative correlation value. Increase in temperature leads to turbulence and mixing of different layers causing dilution of pollutant concentration. As the study period was during the winter and early spring months the solar radiation was comparitively less . This created stable air condition with practically no vertical movement of air causing stagnation of the accumulated pollutants in ambient air.

Table 6.2: Correlation matrix of pollutants with traffic parameters

POLLUTANT	TWO WHEELERS	THREE WHEELERS	PETROL CARS	DIESEL CARS	CNG CARS	CNG BUSES	DIESEL.TRUCKS
$PM_{2.5}$	0.70	0.63	0.73	0.76	0.72	0.39	0.56
PM_{10}	0.72	0.70	0.75	0.77	0.75	0.38	0.48
NOx	0.26	0.21	0.34	0.31	0.29	0.11	0.16
CO	0.71	0.60	0.76	0.76	0.74	0.47	0.28

Table: 6.2 shows the correlation of traffic parameters on the pollutants. The correlation matrix shows that $PM_{2.5}$, PM_{10} and CO have strong influence with vehicles suggesting main source of these pollutants may be vehicular emissions from the highway in study area.. The low NOx correlation values suggests that NOx concentrations in study area may not be from the vehicular sources.

In addition to the meteorological and traffic variables it was observed that PM_{10} ,$PM_{2.5}$ and NOx ,CO showed strong correlation between them ,which suggests that emission of these pollutants may be from similar sources.

i)CORRELATION BETWEEN PARTICULATE MATTER

The Figure 6.1 shows the scatter plot of $PM_{2.5}$ concentrations against PM_{10} concentrations during the study period

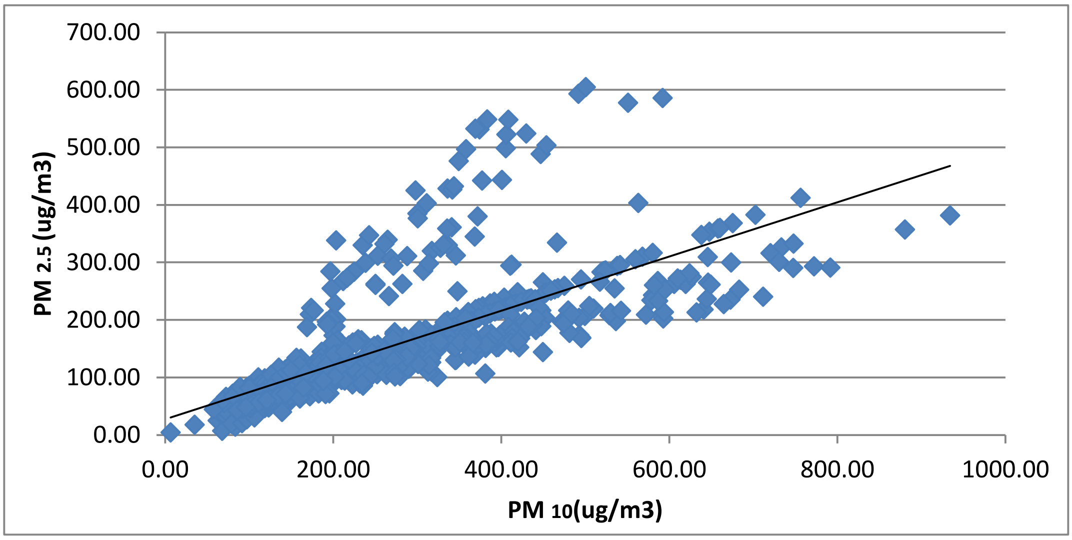


Figure 6.1: The scatter plot of PM $_{2.5}$ concentration against PM_{10} concentration

$$\mathbf{PM}_{2.5} = 0.4715*\mathbf{PM}_{10} + 27.309 \qquad \mathbf{R}^2 = 0.5284 \;, \qquad \mathbf{R}=0.72 \qquad \textbf{(Eqn.6.1)}$$

The regression equation 6.1 was computed from the data of the monitoring station. The equation reveals that $PM_{2.5}$ concentrations increases with increasing PM_{10} concentrations. The correlation and regression analysis indicate strong linear relationship between $PM_{2.5}$ and PM_{10}. This suggests that $PM_{2.5}$ and PM_{10} can be predicted as a function of PM_{10} and PM $_{2.5}$ respectively.A separate ANN model was developed for $PM_{2.5}$ and PM_{10} using PM_{10} and $PM_{2.5}$ concentrations respectively as inputs in addition to meteorological and traffic input variables .

ii)CORRELATION BETWEEN NOx and CO

The **Figure 6.2** shows the scatter plot of NOx concentrations against CO concentrations during the study period.

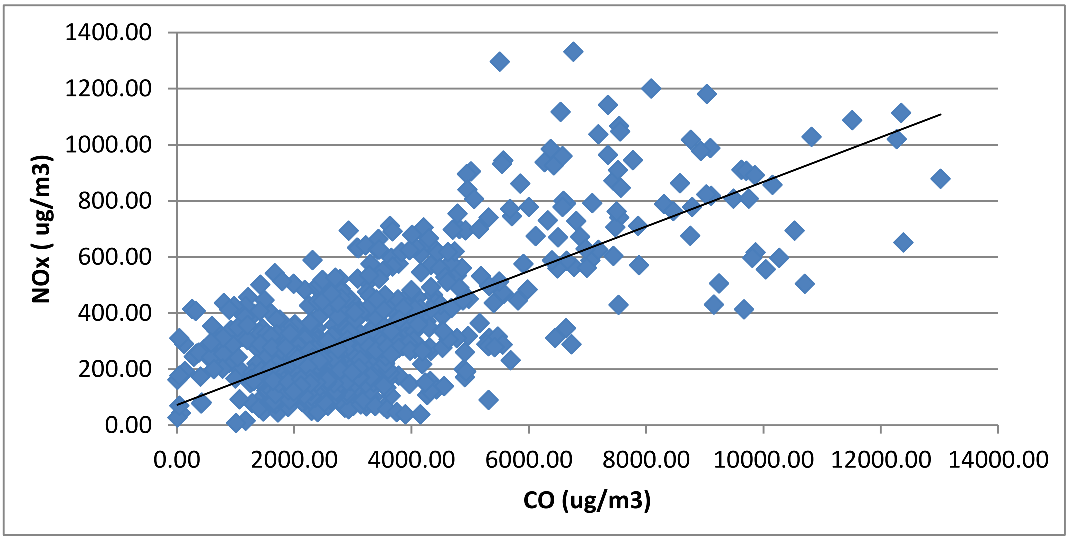


Figure 6.2: The scatter plot of NO_x concentration against CO concentration

$\mathbf{NO_x} = 0.0796^*\mathbf{CO} + 71.575 \qquad R^2 = 0.5164 \qquad R=0.671$ **(Eqn.6.2)**

The results of correlation and regression analysis indicate a moderately linear relationship between NO_x and CO .A separate ANN model was developed for prediction of NOx and CO as a function of CO and NO_x respectively in addition to meteorological and traffic input variables.

6.2 ANN MODELS FOR POLLUTANTS

Multilayer perceptron back propagation neural network were developed for each pollutant. The fixing of hidden nodes were tried on trial and error basis which was exhaustive and yielded poor results. The optimal number of hidden neurons were found using the average minimum MSE graph by running the data on Neuro solutions 6.It was seen that the average MSE decreased by increasing the number of hidden neurons upto a certain limit. Less number of neurons causes inadequate mapping of input and output data while more number of neurons leads to over fitting of the model. On the basis results of sensitivity analysis, optimum number of neurons were determined and used in MATLAB to develop network with best performance for the pollutants with given dataset. The performance evaluation of each model has been discussed in detail in this section .The values of momentum rate α and learning rate η were kept at 0.01 and 0.9 respectively. The relative influence of input parameters on each pollutant in model has also been discussed .

6.3 MODELS FOR $PM_{2.5}$

For $PM_{2.5}$ two ANN based models were developed RSPM-1 and RSPM-2. A total of 739 complete data sets were extracted from raw data and were randomly separated for training, testing and validation.

6.3.1: MODEL RSPM-1

i)Topology of network

The model RSPM-1 considered meteorological and traffic parameters as input and concentration of $PM_{2.5}$ as output. The **Figure 6.3** shows graph of the average of minimum MSE for each trial with increment of hidden nodes in the hidden layer. Five hidden neurons were considered initially and incremented by one hidden neuron upto a maximum of 40 hidden neurons. Based on results of the sensitivity analysis results 34 neurons were considered optimum for mapping between input and output variables in present model. The **Figure 6.4** shows the network topology for RSPM-1 model.

ii) Network Training, testing and validation

Out of the 739 data sets used for development of model data sets were randomly divided into 75%, 15% and 15% for training, testing and validation of network. The network was stopped after 500 iterations with one complete pass through the set of input and target during training of the network. The synaptic weights were stored for network validation. The results of training, testing and validation of the network in MATLAB 2012b is given in **Figure 6.5.** The variation of predicted and observed concentrations of $PM_{2.5}$ by RSPM-1 model is given in **Figure 6.6**

•

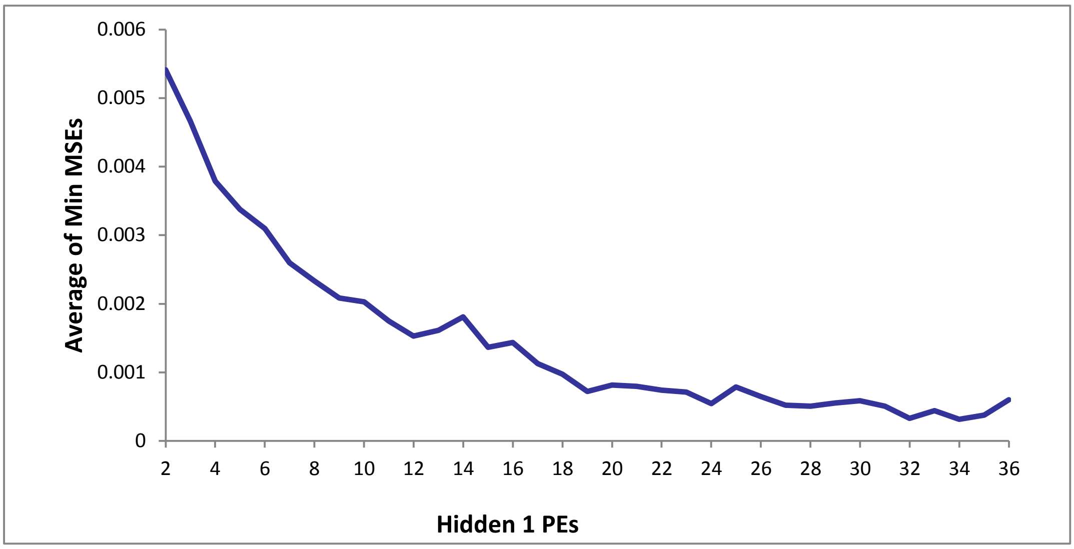


Figure 6.3: Estimated optimum number hidden nodes for RSPM-1 Model

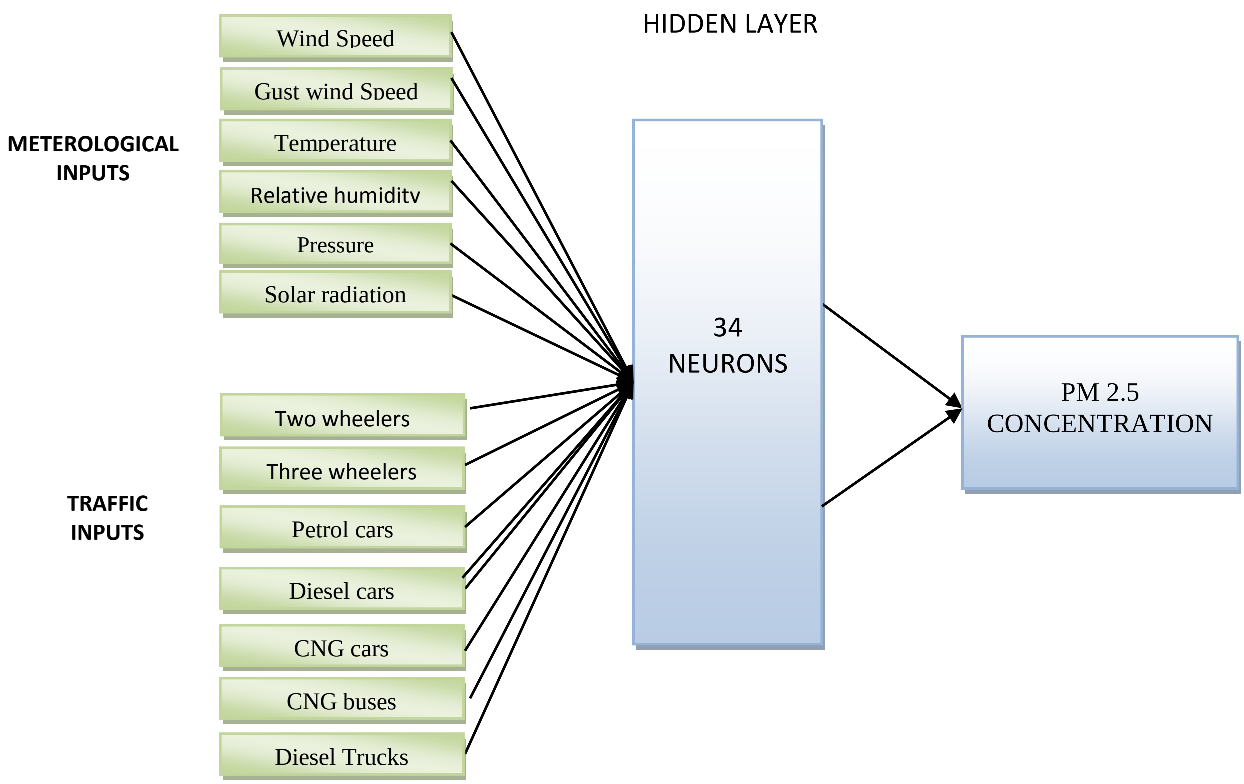


Figure 6.4: MLP back propagation network topology for RSPM-1 Model

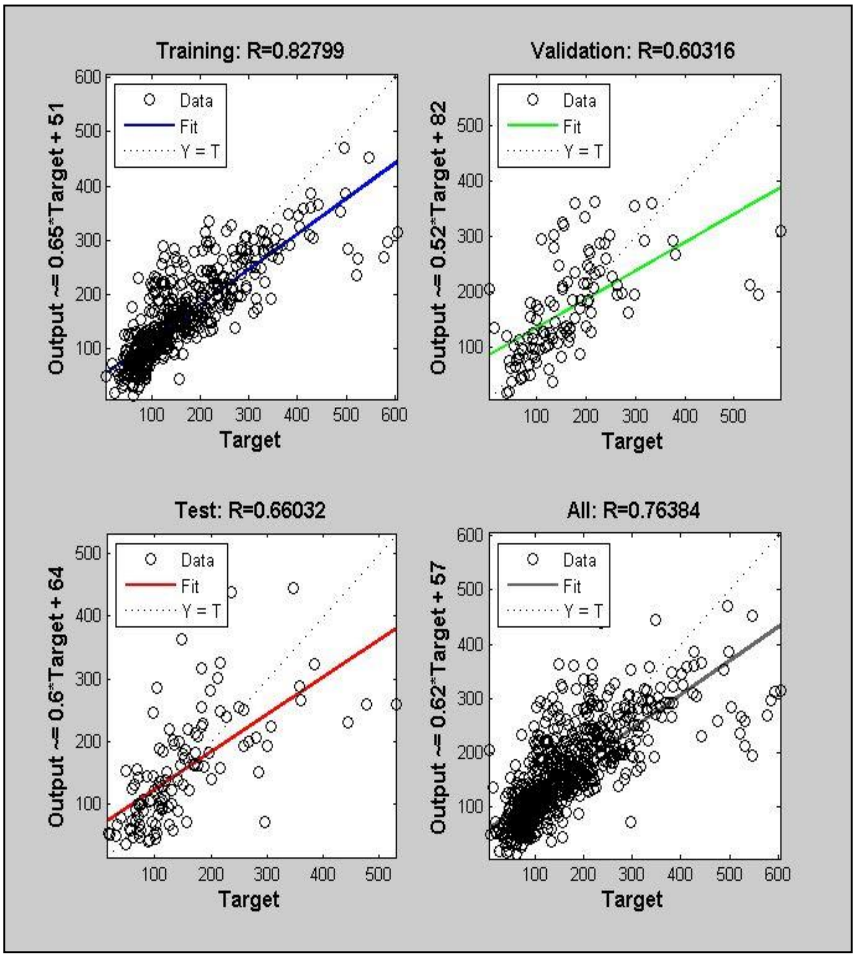


Figure 6.5: The training, testing, validation in MATLAB for RSPM-1 Model

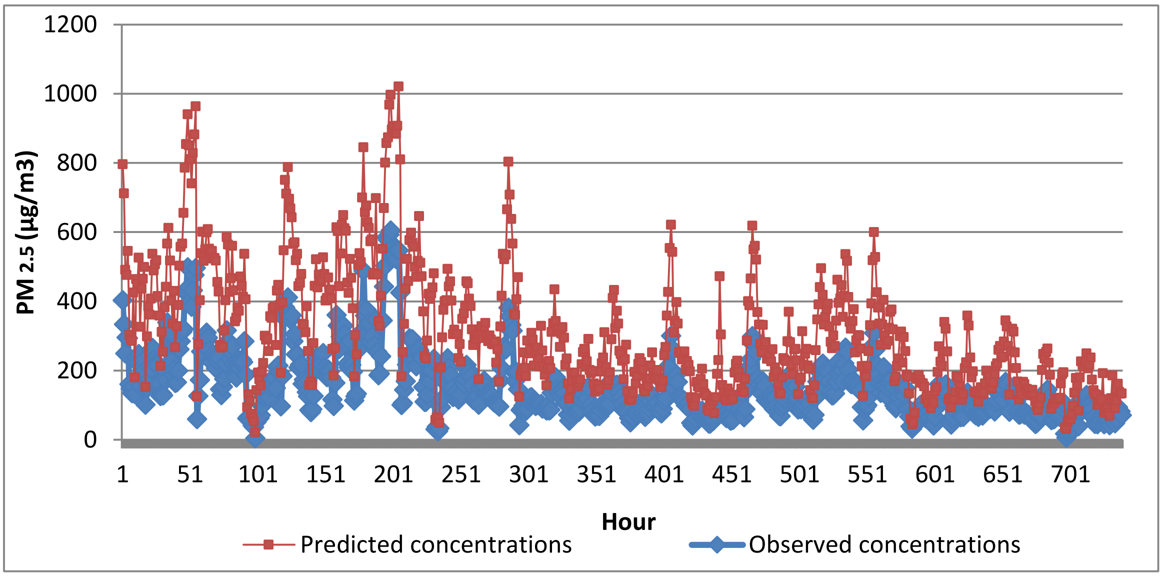


Figure 6.6: The variation in predicted and measured values by RSPM-1 Model

6.3.2: MODEL RSPM-2

i)Topology of network

The RSPM-2 model considered meteorological, traffic parameters and background concentrations (PM_{10}) as input and concentration of $PM_{2.5}$ as output .The **Figure 6.7** shows graph of the average of minimum MSE with increment of neurons in hidden layer. Based on results of sensitivity analysis results 35 hidden nodes were considered optimum for mapping between input and output variables in this model. The **Figure 6.8** shows the network topology for RSPM-2 model

ii) Network Training, testing and validation

Out of the 739 data sets used for development of model data sets were randomly divided into 75%, 15% and 15% for training, testing and validation of network. The network was stopped after 400 iterations and synaptic weights were stored for network validation and to determine the relative influence of input variables on the pollutants. The results of training, testing and validation of the network in MATLAB 2012b is given in **Figure 6.9.**The variation of predicted and measured value of $PM_{2.5}$ concentration by RSPM-2 model is given in **Figure 6.10.**

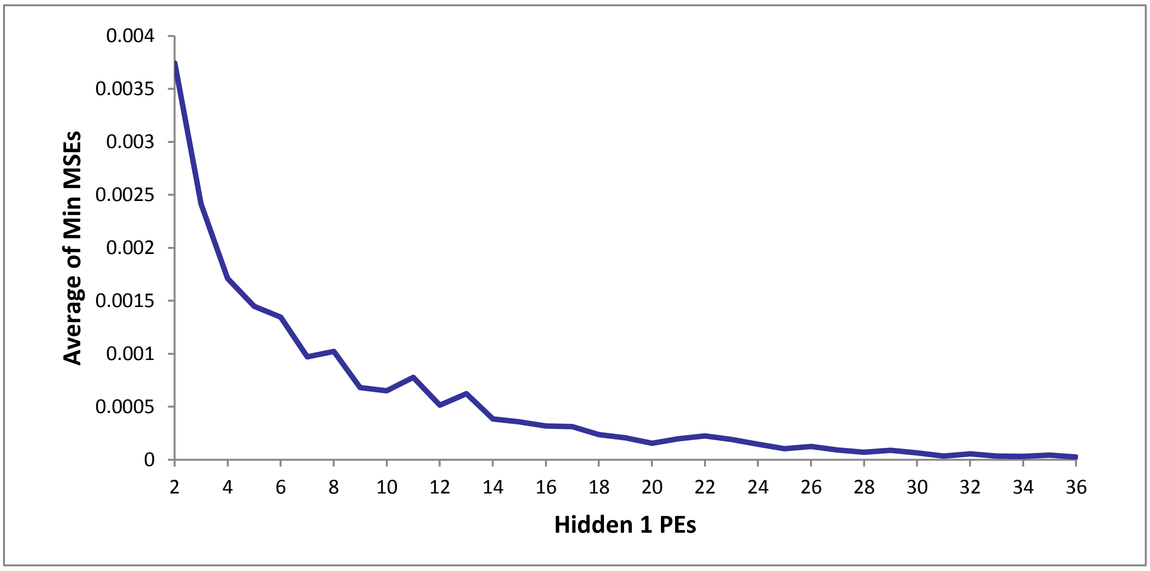


Figure 6.7: Estimated optimum number hidden nodes for RSPM-2 Model

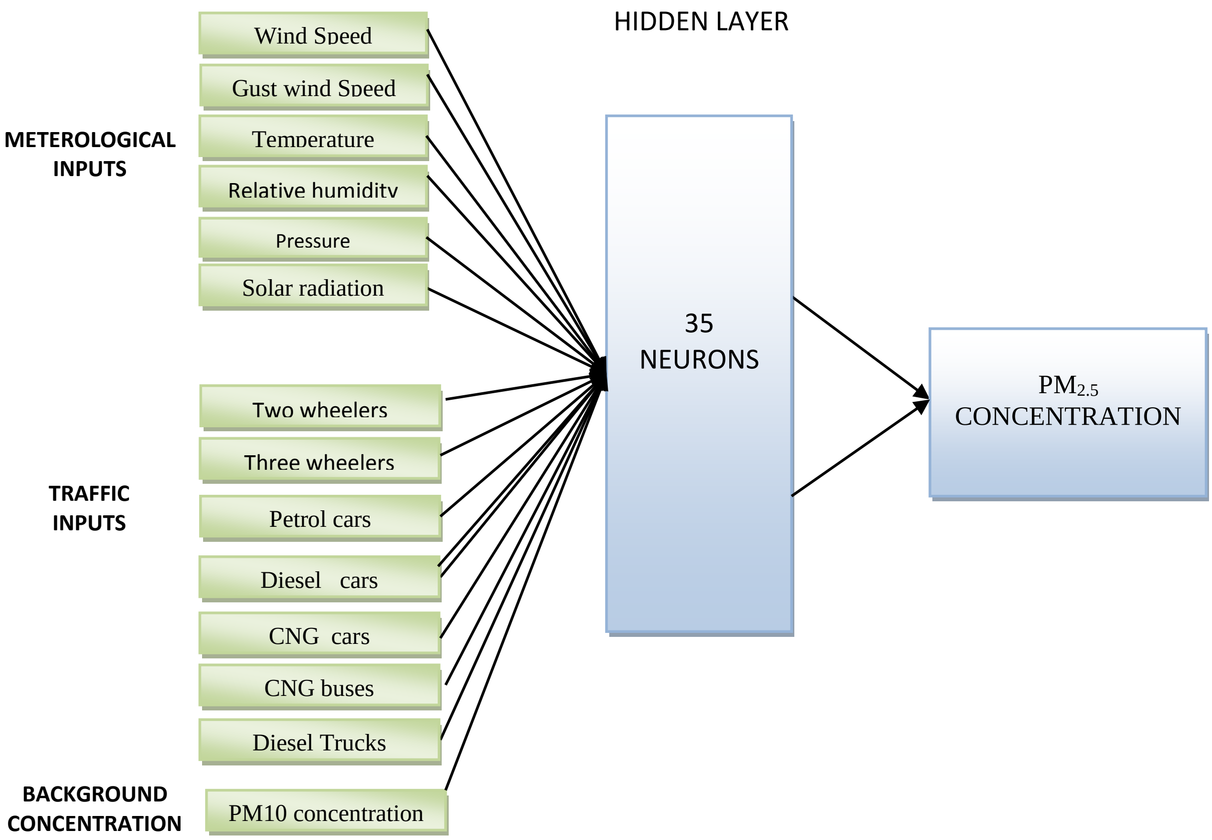


Figure 6.8: MLP back propagation network topology for RSPM-2 Model

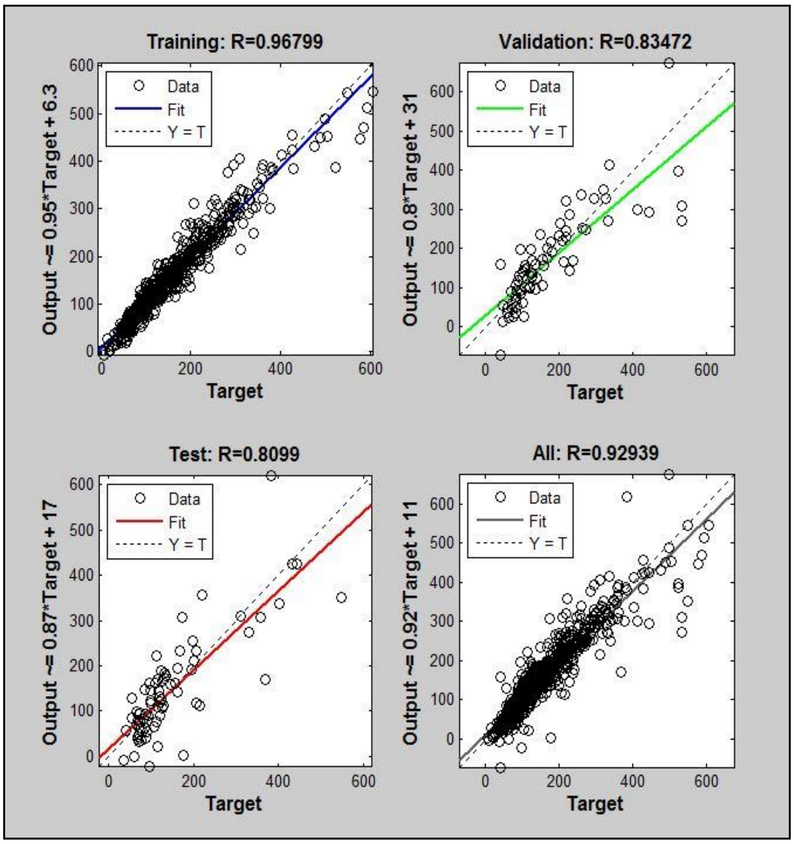


Figure 6.9: The training, testing, validation in MATLAB for RSPM-2 Model

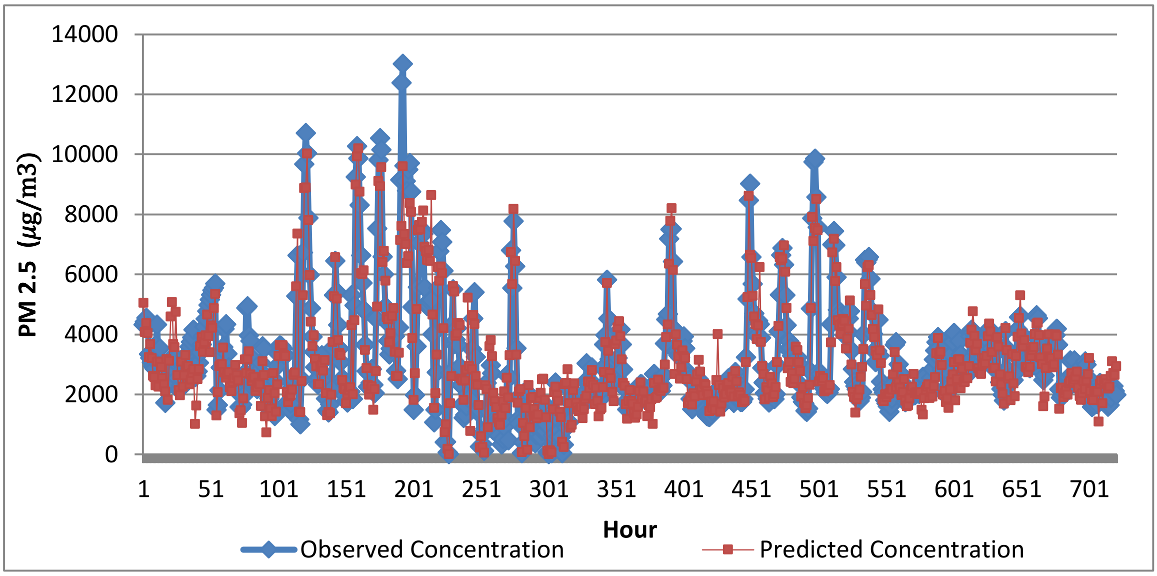


Figure 6.10: The variation between predicted and measured values for RSPM-2 Model

6.3.3 COMPARISON BETWEEN $PM_{2.5}$ MODELS

The table 6.3 lists the performance statistics of model RSPM-1 and RSPM-2 .$\bar{O}$ and $\bar{P}$ are the mean of the observed and predicted concentrations . The σo and σp denotes standard deviation of observed and predicted values. The MAPE, MAE, RMSE and R^2 values are given in detail in Chapter 4 are also listed in the table.

Table 6.3 Performance statistics of models RSPM-1 and RSPM-2 for $PM_{2.5}$.

MODEL	STATISTICAL PARAMETERS							
	$\bar{O}$	$\bar{P}$	σo	σp	MAPE	MAE	RMSE	R^2
	(µg/m3)	(µg/m3)	(µg/m3)	(µg/m3)	%	(µg/m3)	(µg/m3)	Testing
RSPM-1	154.69	156.82	99.15	89.85	31.11	42.69	64.27	0.43
RSPM-2	154.69	153.79	99.15	97.66	21.23	28.39	45.18	0.64

The mean of the predicted values is higher than that of observed concentrations for model RSPM-1. This explains the tendency of over prediction of the model as observed in Figure 6.6.The MAPE,MAE,RSME value of the RSPM-1 model is higher than RSPM-2 model. This shows that RSPM-2 model has better prediction capacity than RSPM-1. The R^2value of model RSPM-2 is 0.64 for testing which suggests good prediction capacity of the model over model RSPM-1.

6.3.4 RELATIVE INFLUENCE OF INPUT PARAMETERS ON $PM_{2.5}$ CONCENTRATION

The relative influence of the input parameters on PM $_{2.5}$ was evaluated by examining the weights obtained after training process between input-hidden layer and hidden-output layer. The method to determine the relative importance of various input parameters was proposed by Garson (1991) which is explained in Chapter 4 section 7. The same algorithm has been used in the study to assess the relative influence of input variables on PM $_{2.5}$ concentration.

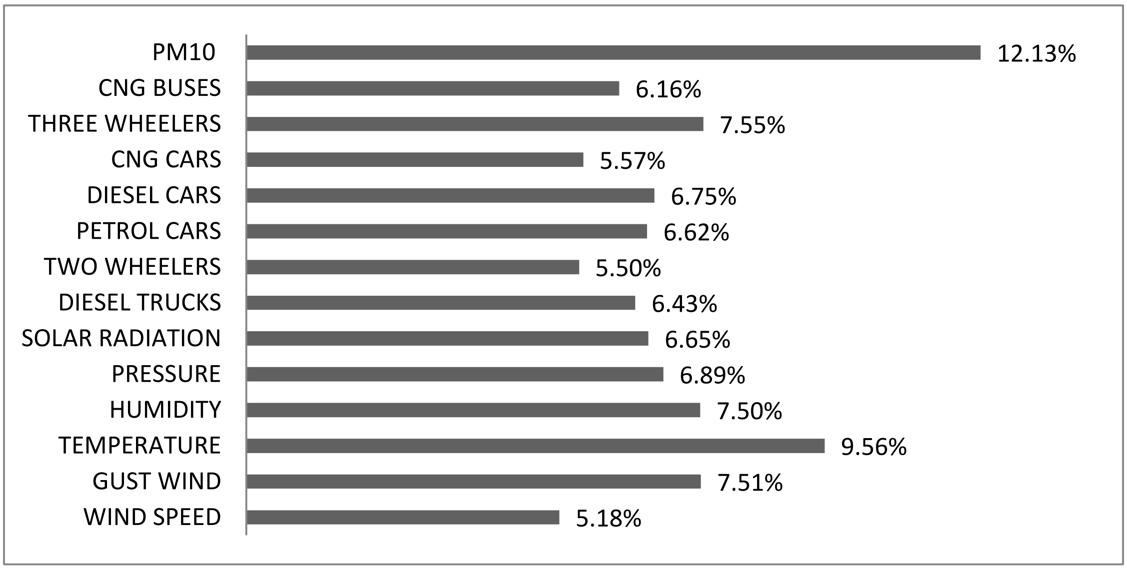


Figure 6.11: The relative influence of input parameters on $PM_{2.5}$ concentration

The **Figure 6.11** shows that the meteorological parameters: temperature, gust wind, humidity, and three wheelers, diesel cars influences the PM $_{2.5}$ concentration in the study area the most. It is also seen that on including PM_{10} concentration as input apart from meteorological and traffic parameters , the model performance of PM $_{2.5}$ improves considerably as observed from **Table 6.3**. This explains the high influence of PM_{10} as input parameter in **Figure 6.11**.

6.4 MODELS FOR PM_{10}

For PM_{10} two ANN based models were developed PM-1 and PM-2. A total of 739 complete data sets were extracted from raw data and were randomly separated for training, testing and validation.

6.4.1: MODEL PM-1

i)Topology of network

The PM-1 model considered meteorological, traffic parameters as input and concentration of PM_{10} as output .The Figure **6.12** shows graph of the average of minimum MSE with increment of neurons in hidden layer. Initially, seven hidden neurons were considered and incremented by one neuron upto a maximum of 35 neurons. Based on results of sensitivity analysis results 32 neurons were considered optimum for mapping between input and output variables in this model. The **Figure 6.13** shows the network topology for PM-1 model

ii) Network Training, testing and validation

Out of the 739 data sets used for development of model data sets were randomly divided into 75%, 15% and 15% for training, testing and validation of network. The network was stopped after 600 iterations and synaptic weights were stored for network validation. The results of training, testing and validation of the network in MATLAB 2012b is given in **Figure 6.14.**The variation of predicted and measured value of PM_{10} concentration by PM-1 model is given in **Figure 6.15.**

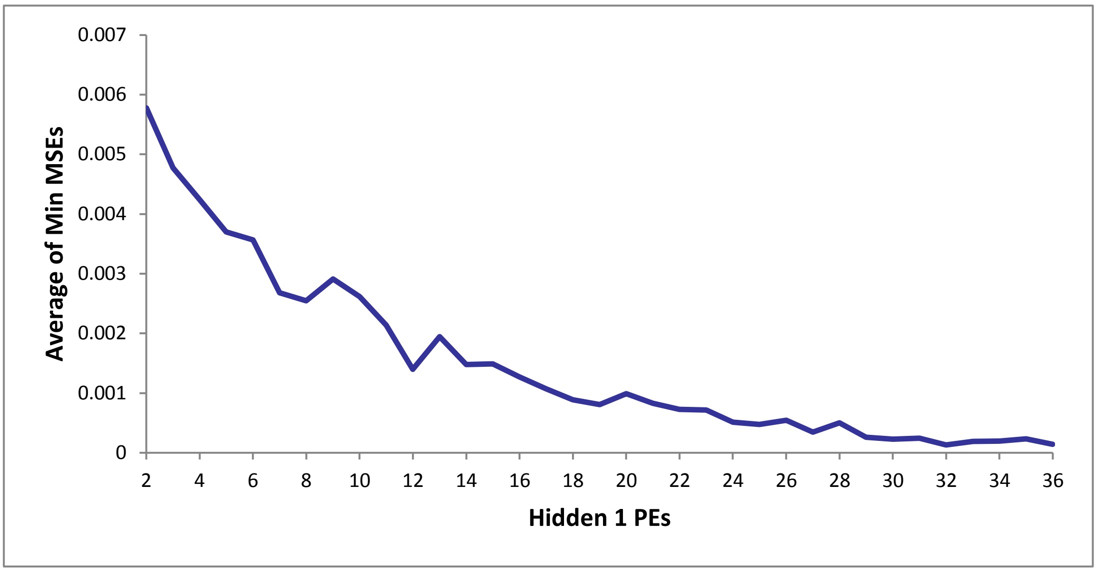


Figure 6.12: Estimated optimum number hidden nodes for PM-1 Model

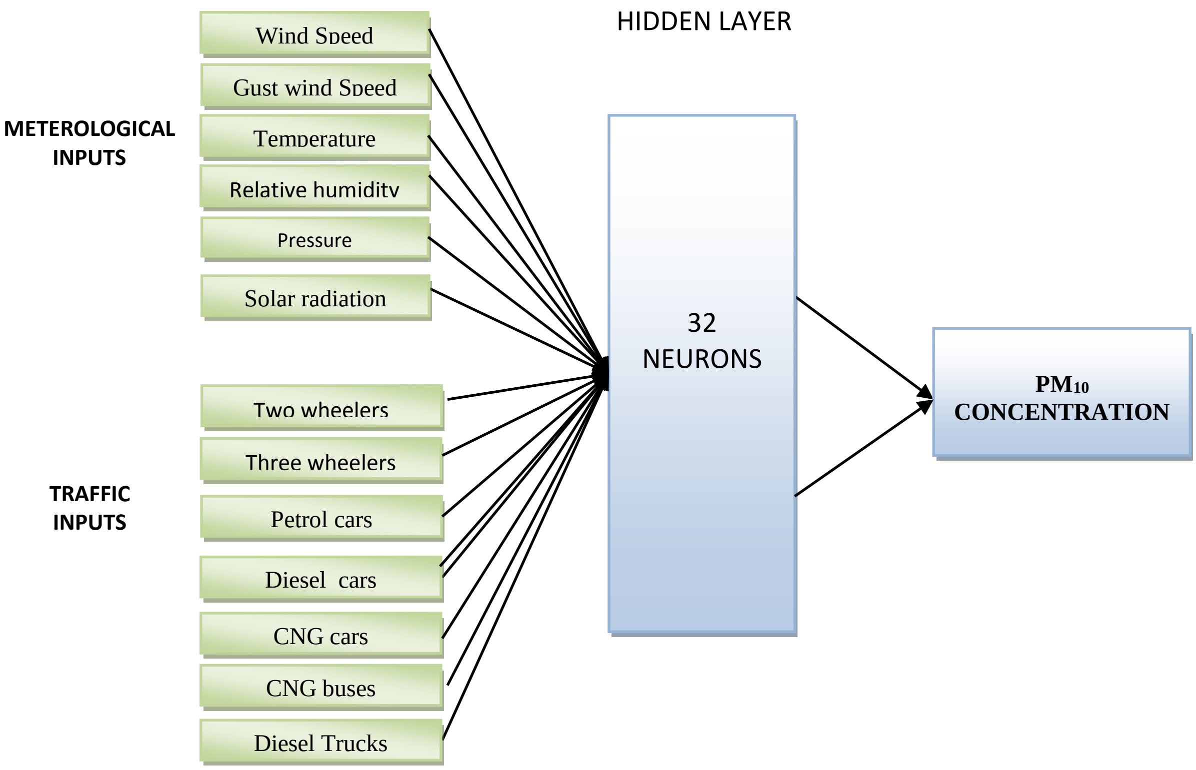


Figure 6.13 : MLP back propagation network topology for PM-1 Model

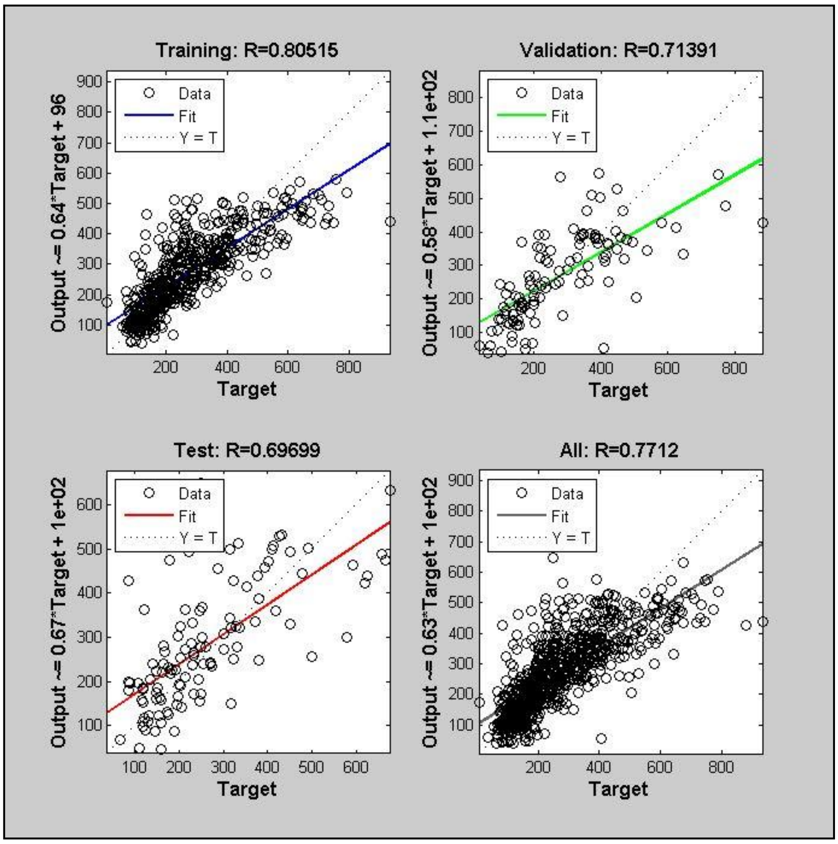


Figure 6.14: The training, testing, validation in MATLAB for PM-1 Model

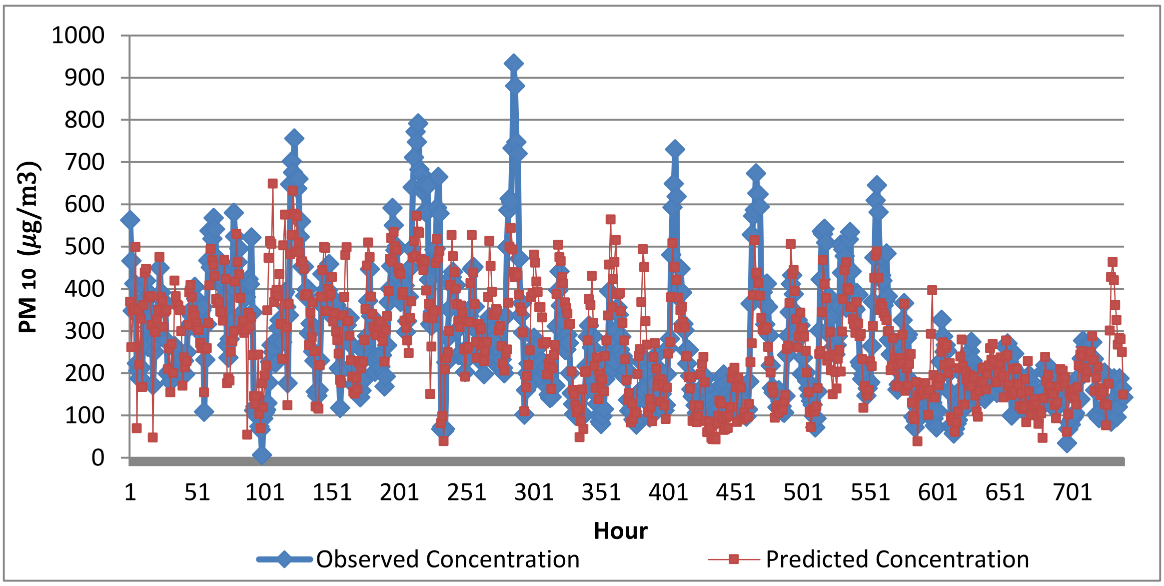


Figure 6.15: The variation in predicted and measured values by PM-1 Model

6.4.2: MODEL PM-2

i)Topology of network

The second model PM-2 considered meteorological, traffic parameters and background concentrations ($PM_{2.5}$) as input and concentration of PM_{10} as output. The Figure **6.16** shows results of sensitivity analysis. Five hidden neurons were considered initially and incremented by one hidden neuron to a maximum of 40 hidden neurons. 36 hidden nodes were considered optimum for mapping between input and output variables in this model. The **Figure 6.18** shows the network topology for PM-2 model

ii) Network Training, testing and validation

Out of the 739 data sets used for development of model data sets were randomly divided into 75%, 15% and 15% for training, testing and validation of network. The network was stopped after approximately 600 iterations and synaptic weights were stored for network validation and to determine the relative influence of input parameters. The results of training, testing and validation of the network in MATLAB 2012b is given in **Figure 6.19.** The variation of predicted and measured value of PM_{10} concentration by PM-2 model is given in **Figure 6.20**

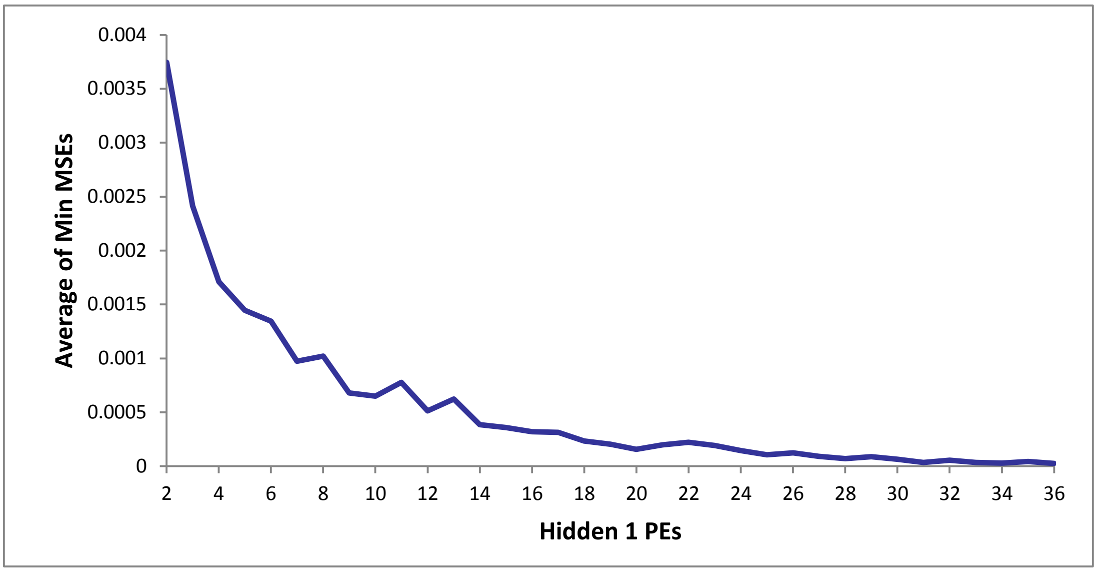


Figure 6.16: Estimated optimum number hidden nodes for PM-2 Model

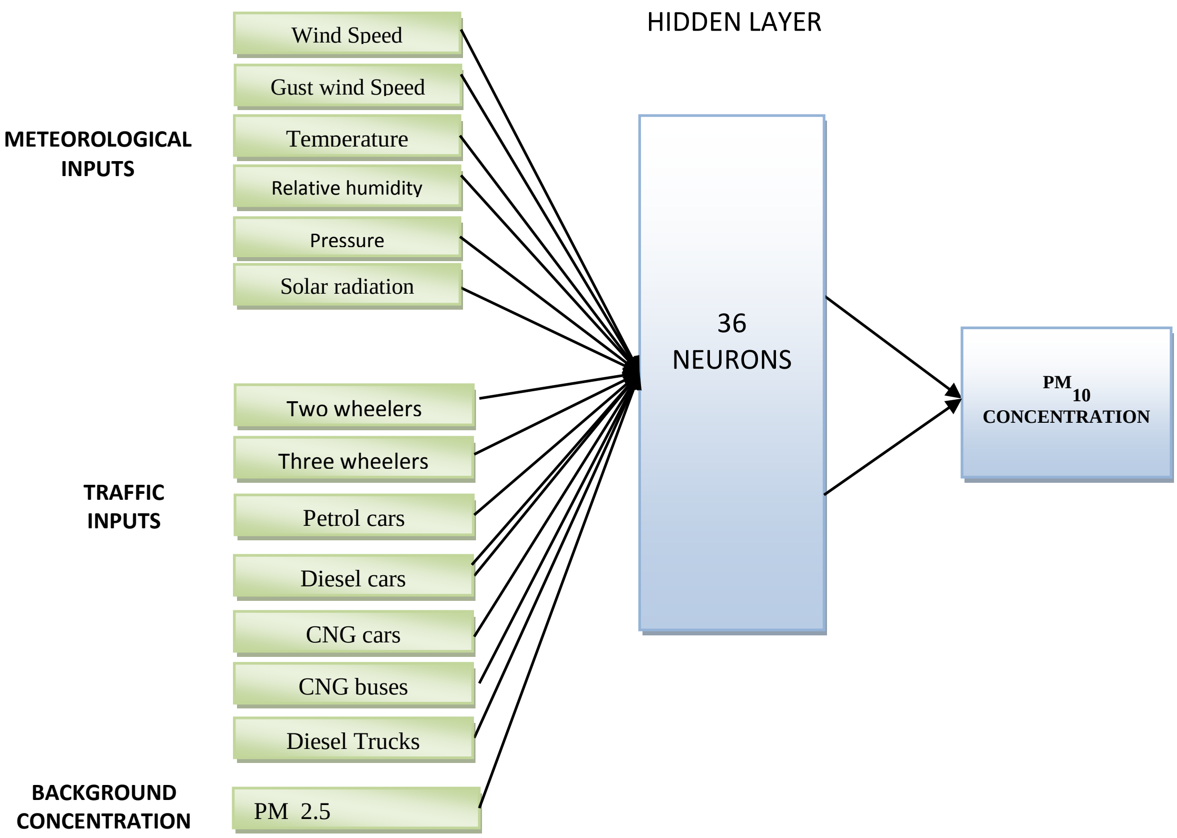


Figure 6.17: MLP back propagation network topology for PM- 2 Model

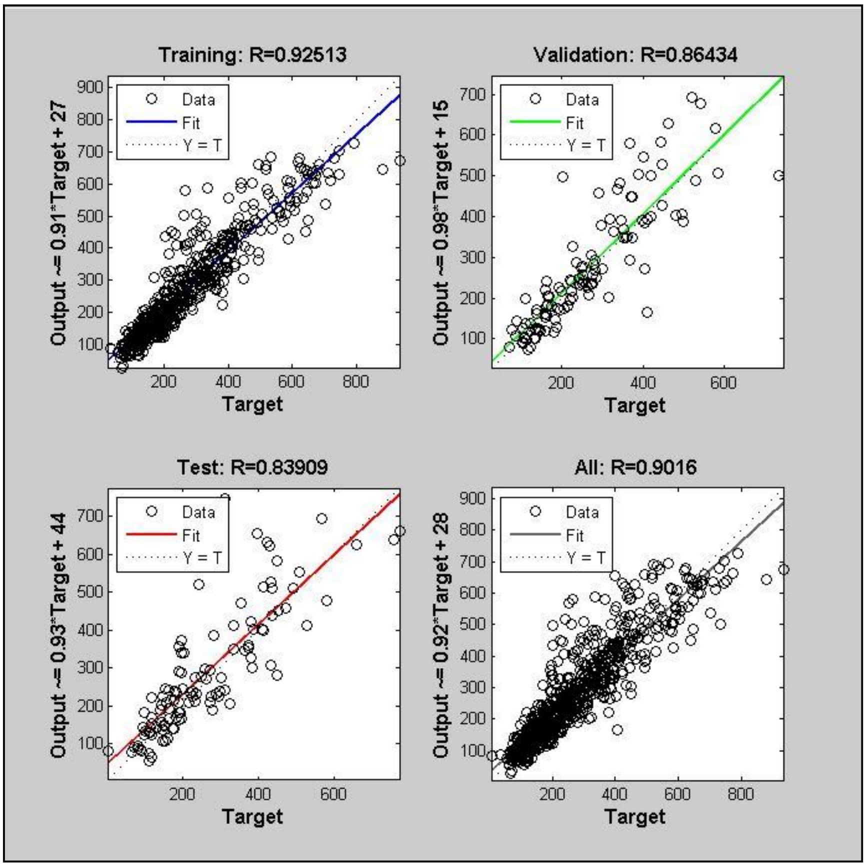


Figure 6.18: The training, testing, validation in MATLAB for PM-2 Model

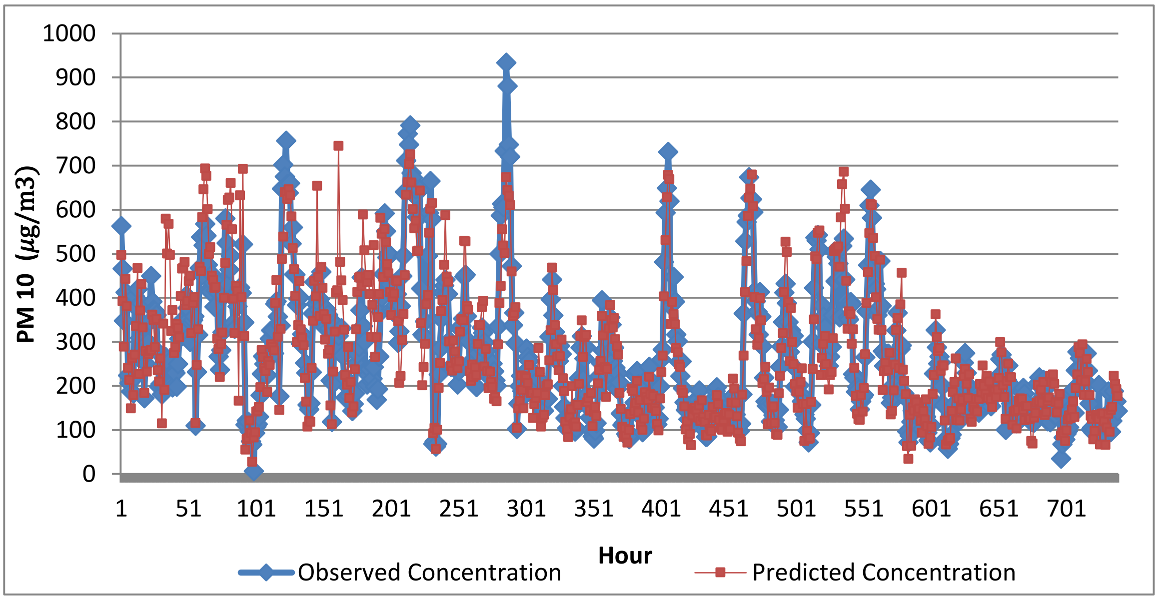


Figure 6.19: The variation between predicted and measured values for PM-2 Model

6.4.3 COMPARISON BETWEEN PM_{10} MODELS

The **Table 6.4** lists the performance statistics of model PM-1and PM-2 .$\bar{O}$ and $\bar{P}$ are mean of the observed and predicted concentrations . The σo and σp denotes the standard deviation of the observed and predicted concentrations. The MAPE, MAE, RMSE and R^2 values of the predicted values from the models are given in the **Table.6.4**

Table 6.4. Performance statistics of models PM-1 and PM-2 for PM_{10}.

MODEL	STATISTICS							
	$\bar{O}$	$\bar{P}$	σo	σp	MAPE	MAE	RMSE	R^2
	(µg/m3)	(µg/m3)	(µg/m3)	(µg/m3)	%	(µg/m3)	(µg/m3)	Testing
PM-1	270.18	271.51	152.88	125.57	33.31	69.78	97.59	0.47
PM-2	270.18	276.51	152.88	155.65	20.16	45.94	68.73	0.70

The MAPE, MAE, RSME value of the PM-1 model is very high than PM-2 model. This shows that PM-1model has poor prediction capacity than PM-2.There is a large difference in the standard deviation of PM-1with standard deviation of observed data. This explains the inability of model to produce actual variations of PM_{10} concentrations. The mean and standard deviation of predicted values for model PM-2 is higher than mean of observed values. This indicates that model PM-2 has tendency of over prediction than model PM-1. The R^2 value of PM-2 model is 0.70 for testing which suggests that the model has good forecasting capacity than PM-1model.

6.4.4 RELATIVE INFLUENCE OF INPUT PARAMETERS ON PM_{10} CONCENTRATION

The relative importance of input parameters on PM_{10} concentration in the study area is given in **Figure 6.20** which was computed by examining the synaptic weights obtained after the training process.

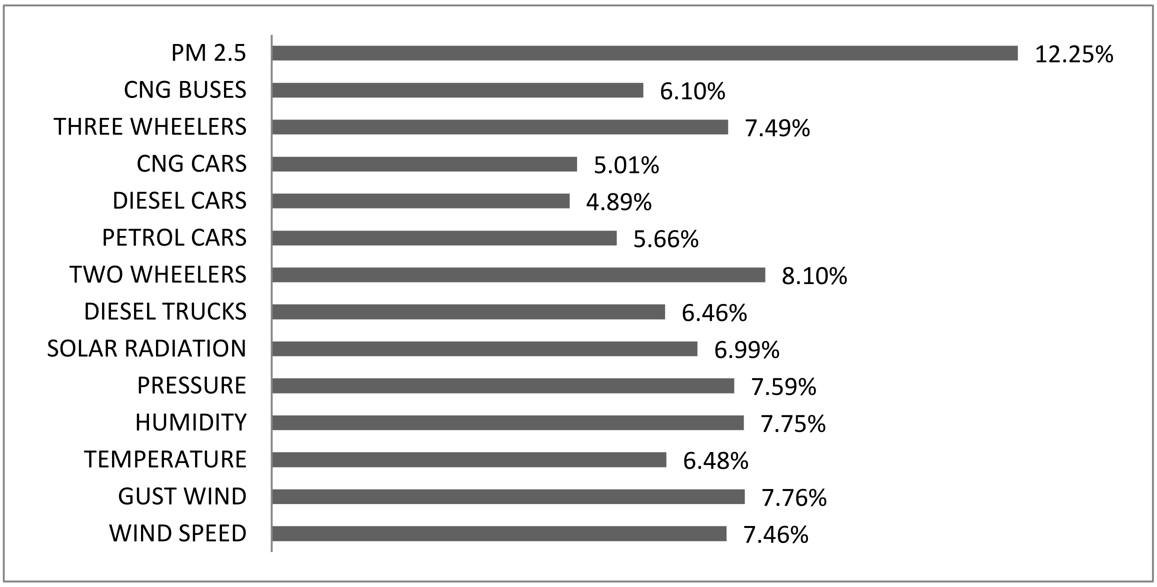


Figure 6.20 : The relative influence of parameters on PM10 concentration

The **Figure 6.20** shows that meteorological parameters: gust wind, pressure and humidity, and three wheelers, two wheelers influences the PM_{10} concentration most in the study area. It is also seen that on including $PM_{2.5}$ concentration as input improves the model performance of PM_{10} considerably as observed from **Table 6.4**. This explains the high influence of $PM_{2.5}$ as input parameter in **Figure 6.20**.

6.5 MODELS FOR NOx

For NOx two ANN based models were developed NO-1 and NO-2. A total of 722 complete data sets were extracted from raw data and were randomly separated for training, testing and validation.

6.5.1: MODEL NO-1

i)Topology of network

The NO-1 model considered meteorological and traffic parameters as input and concentration of NO as output. **Figure 6.21** shows the average of minimum MSE of each trial with increment of hidden nodes in the hidden layer. Two hidden neurons were considered initially and incremented by one hidden neuron to a maximum of 36 hidden neurons. Based on the sensitivity analysis results given in **Figure 6.21**, 34 hidden nodes were considered optimum for mapping between input and output variables in present model. The **Figure 6.22** shows the network topology for NO-1 model

ii) Network Training, testing and validation

Out of the 722 data sets used for development of model data sets were randomly divided into 75%, 15% and 15% for training, testing and validation of network. The network was stopped after approximately 300 iterations and synaptic weights were stored for network validation. The results of training, testing and validation of the network in MATLAB 2012b is given in **Figure 6.23.** The variation of predicted and measured value of NOx concentration by NO-1 model is given in **Figure 6.24.**

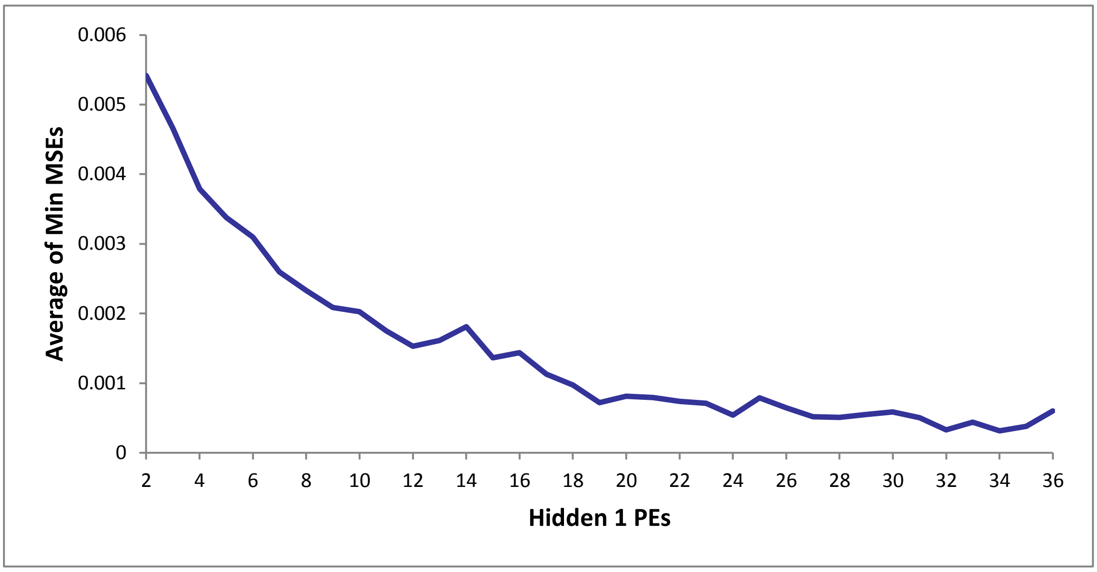


Figure 6.21: Estimated optimum number hidden nodes for NO-1 Model

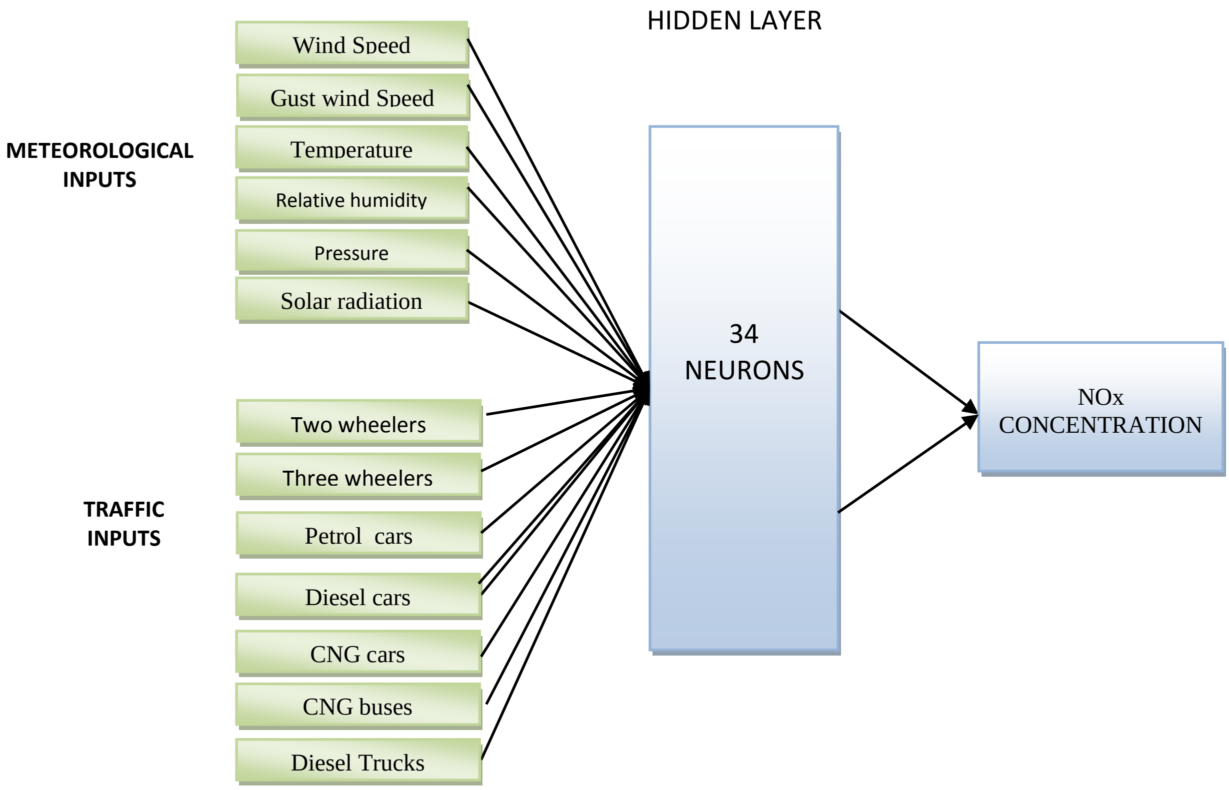


Figure 6.22: MLP back propagation network topology for NO-1 Model

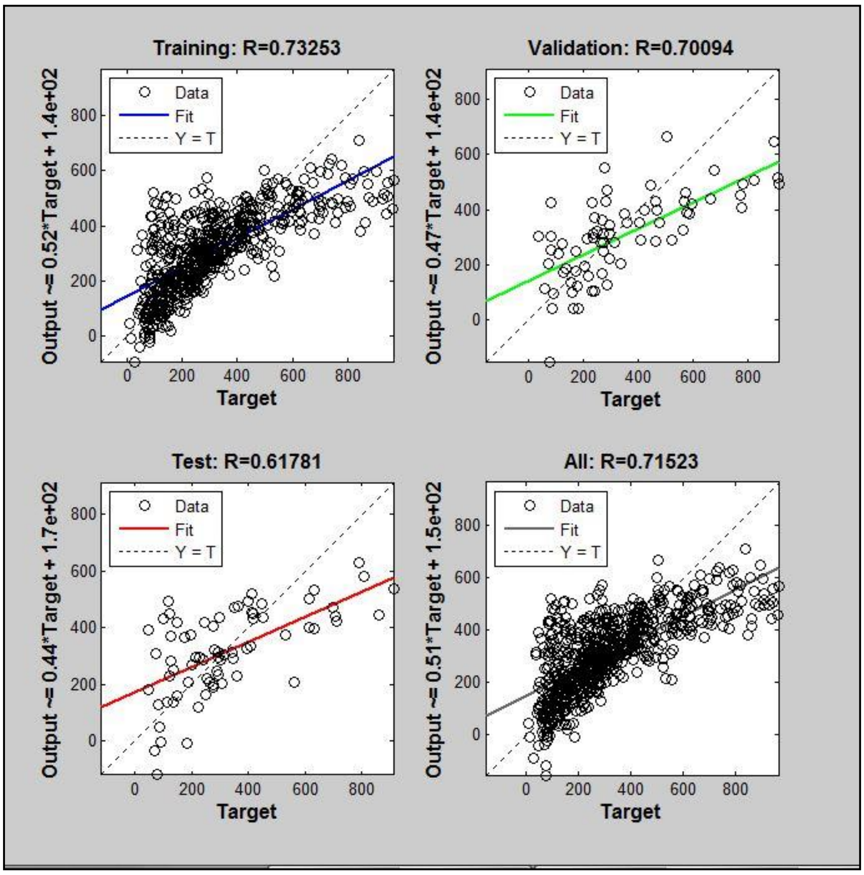


Figure 6.23: The training, testing, validation in MATLAB for NO-1 Model

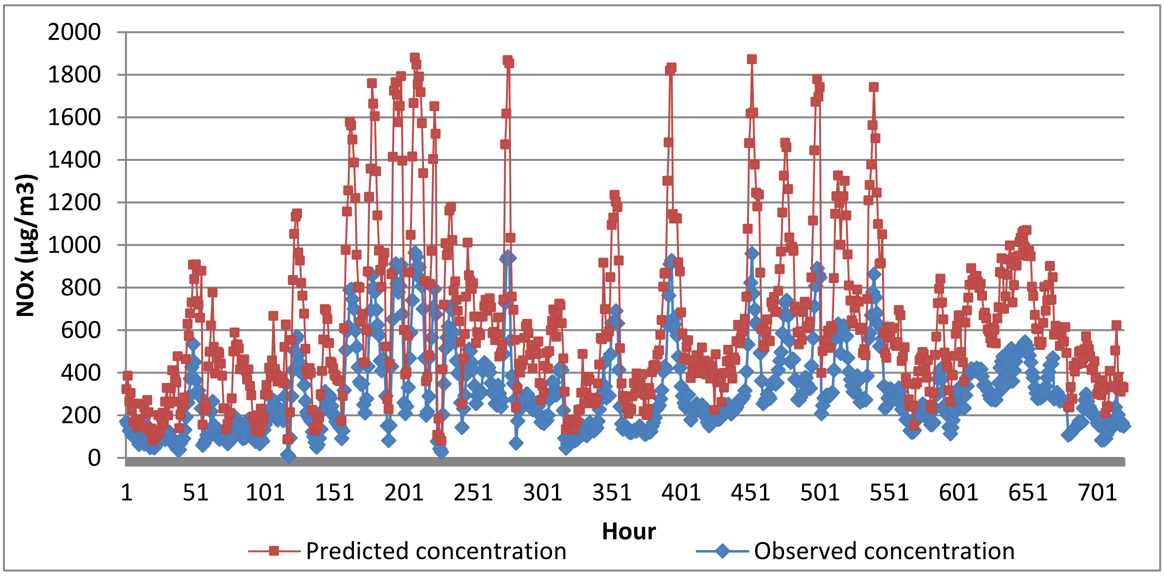


Figure 6.24: The variation in predicted and measured values by NO-1 Model

6.5.2: MODEL NO-2

i)Topology of network

The second model NO-2 considered meteorological, traffic parameters and background concentrations (CO) as input and concentration of NO as output .The Figure **6.25** shows results of the sensitivity analysis. Seven hidden neurons were considered initially and incremented by one hidden neuron to a maximum of 40 hidden neurons. 36 hidden nodes were considered optimal for mapping between inputs, output variables in this model. The **Figure 6.26** shows the network topology for NO-2 model

ii) Network Training, testing and validation

Out of the 722 data sets used for development of model data sets were randomly divided into 75%, 15% and 15% for training, testing and validation of network. The network was stopped after 400 iterations.The results of training, testing and validation of the network in MATLAB 2012b is given in **Figure 6.27.**The variation of predicted and measured value of NO concentration by NO-2 model is given in **Figure 6.28.**

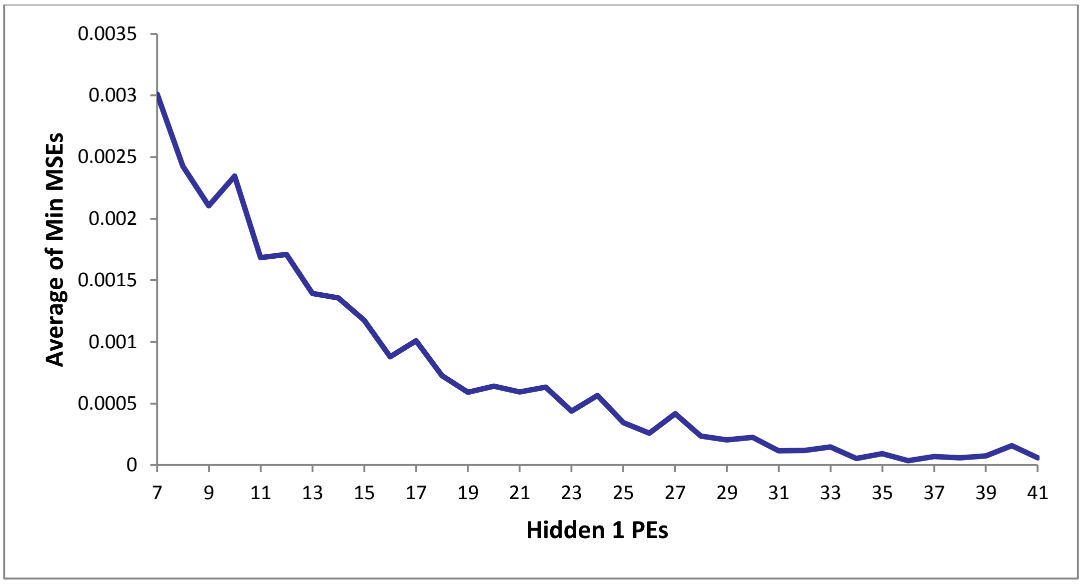


Figure 6.25: Estimated optimum number hidden nodes for NO-2 Model

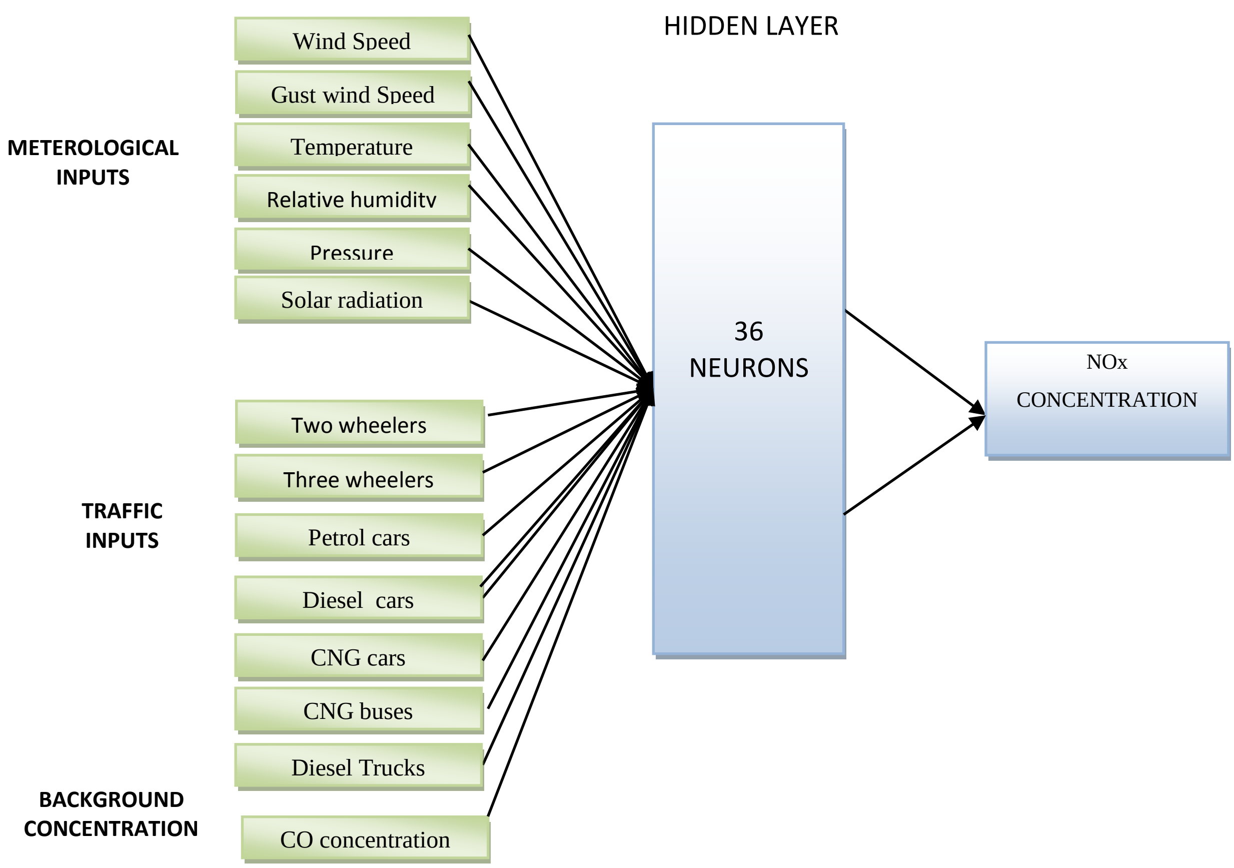


Figure 6.26: MLP back propagation network topology for NO-2 Model

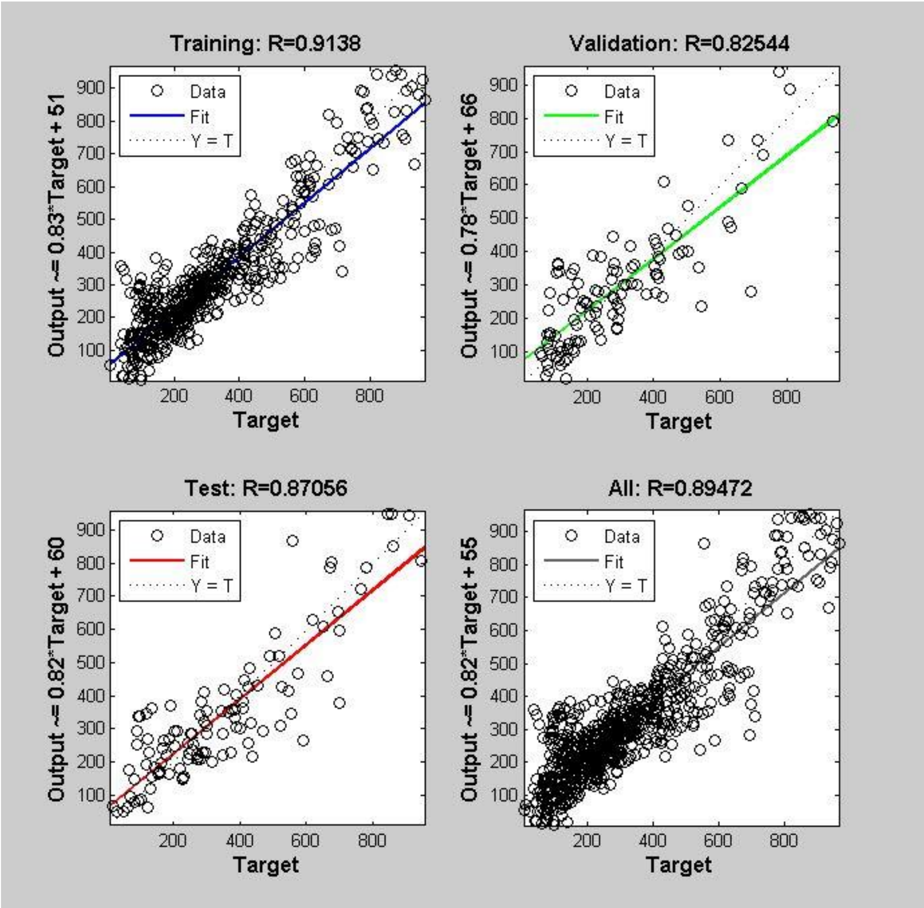


Figure 6.27: The training, testing, validation in MATLAB for NO-2 Model

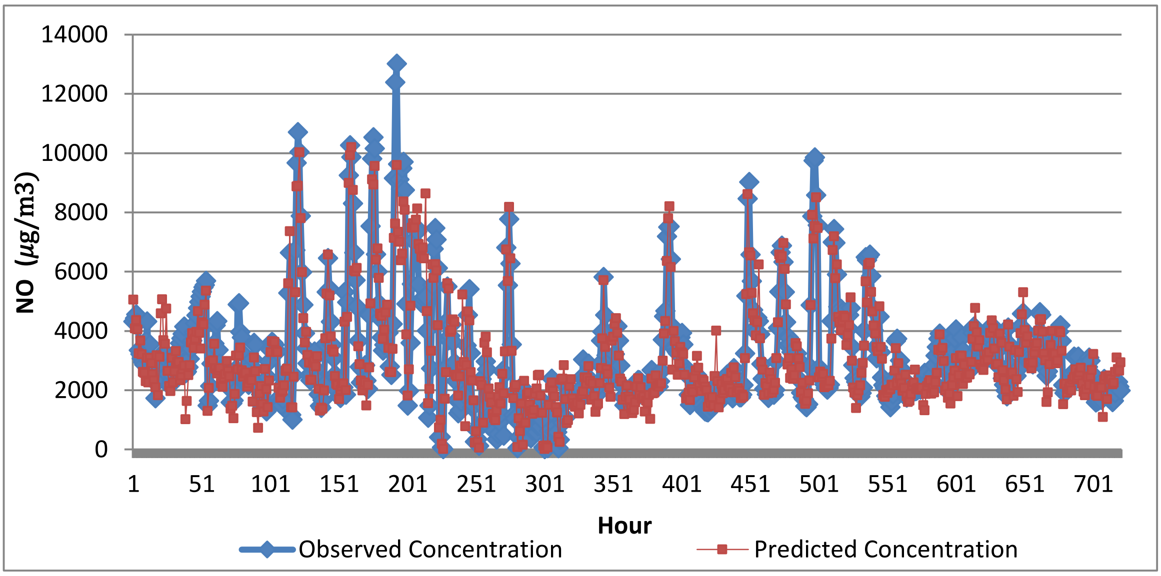


Figure 6.28 : The variation between predicted and measured values for NO-2 Model

6.5.3 COMPARISON BETWEEN NO_x MODELS

The **Table 6.5** lists the performance statistics of model NO-1 and NO-2 .$\bar{O}$ and $\bar{P}$ are mean of the observed and predicted concentrations . The σo and σp denotes standard deviation of the observed and predicted concentrations. The MAPE, MAE, RMSE and R^2 values of the predicted values are also given in the table.

Table 6.5: Performance statistics of models NO-1 and NO-2 for NOx.

MODEL	STATISTICS							
	$\bar{O}$	$\bar{P}$	σo	σp	MAPE	MAE	RMSE	R^2
	(μg/m3)	(μg/m3)	(μg/m3)	(μg/m3)	%	(μg/m3)	(μg/m3)	Testing
NO-1	312.69	318.16	202.35	140.33	49.11	141.45	189.83	0.51
NO-2	312.69	311.89	202.35	186.13	30.97	64.18	90.46	0.70

The MAPE, MAE and RMSE value is very high in NO-1 model indicating poor capability of the model to handle variations in NOx concentration. The mean of NO-1 model is higher than the observed mean indicating the tendency of model to over predict the concentrations which can be seen in **Figure 6.24**. The mean of the model predictions of NO-2 is close to the mean of observed NOx concentrations indicating better forecasting capacity than NO-1 model.

6.5.4 RELATIVE INFLUENCE OF INPUT PARAMETERS ON NOx CONCENTRATION

The relative importance of input parameters on NOx concentration in the study area is given in **Figure 6.29** which was computed by examining the synaptic weights obtained after the training process.

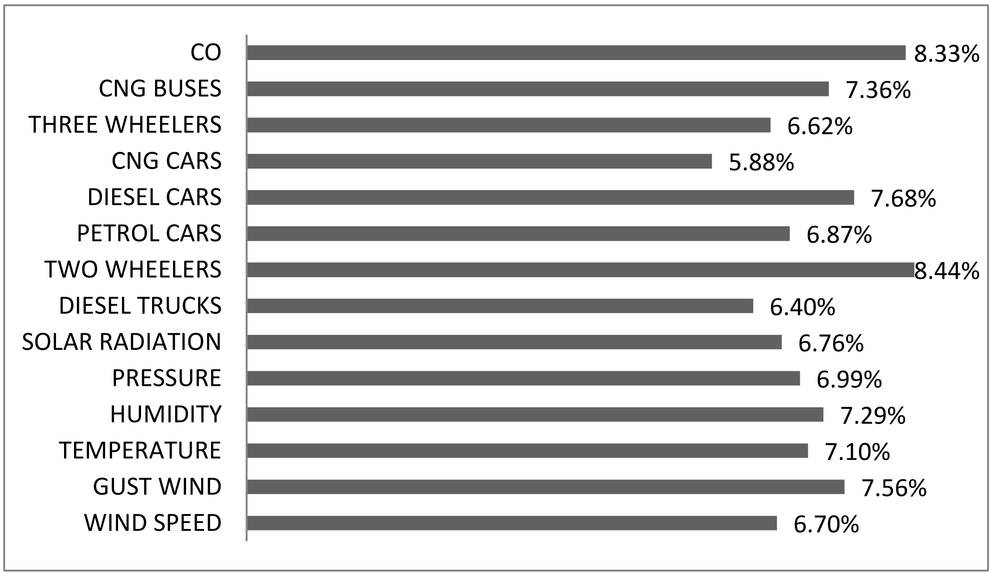


Figure 6.29: The relative influence of input parameters on NOx concentration

The **Figure 6.29** shows that the meteorological parameters: gust wind, temperature, humidity and two wheelers, diesel cars and CNG buses influences the NOx concentration most in the study area. It is also seen that on including CO concentration as input improves the model performance of NOx considerably as observed from **Table 6.5**. This explains the high influence of CO as input parameter in **Figure 6.29**.

6.6 MODELS FOR CO

For CO two ANN based models were developed CO-1 and CO-2. A total of 722 complete data sets were extracted from raw data and were randomly separated for training, testing and validation.

6.6.1: MODEL CO-1

i)Topology of network

The model CO-1 considered meteorological and traffic parameters as input and concentration of CO as output. The **Figure 6.30** shows graph of the average of minimum MSE of each trial with increment of hidden nodes in the hidden layer. Eight hidden neurons were considered initially and incremented by one hidden neuron upto a maximum of 36 hidden neurons. Based on results of the sensitivity analysis results 34 hidden nodes were considered optimum for mapping between input and output variables in present model. The **Figure 6.31** shows the network topology for CO-1 model.

ii) Network Training, testing and validation

Out of the 722 data sets used for development of model data sets were randomly divided into 75%, 15% and 15% for training, testing and validation of network. The network was stopped after 600 iterations with one complete pass through the set of input and target during training of the network. The synaptic weights were stored for network validation. The results of training, testing and validation of the network in MATLAB 2012b is given in **Figure 6.32.** The variation of predicted and observed concentrations of CO by CO-1 model is given in **Figure 6.33.**

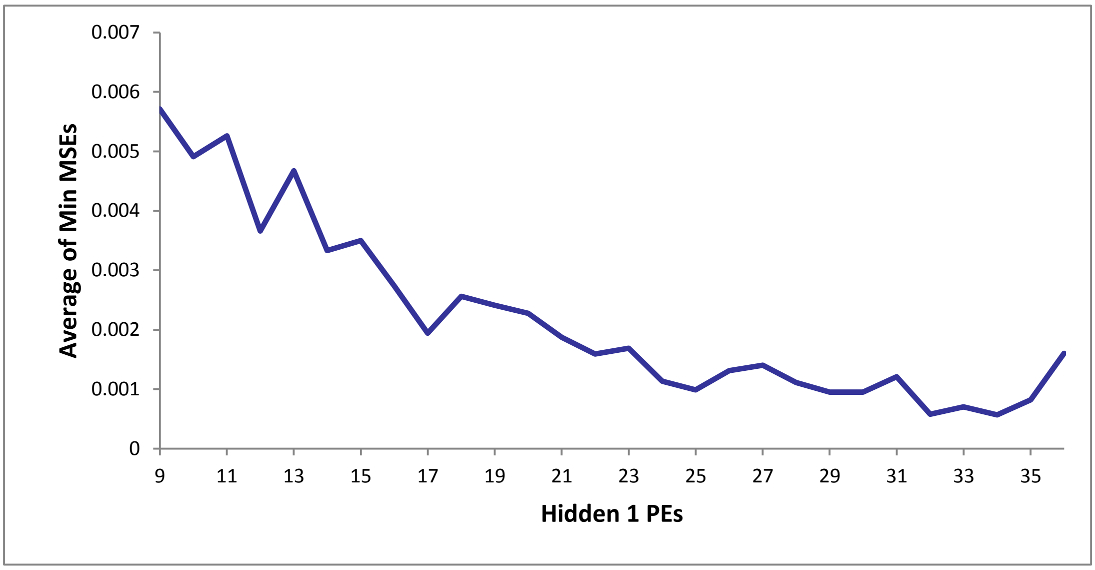


Figure6.30: Estimated optimum number hidden nodes for CO-1 Model

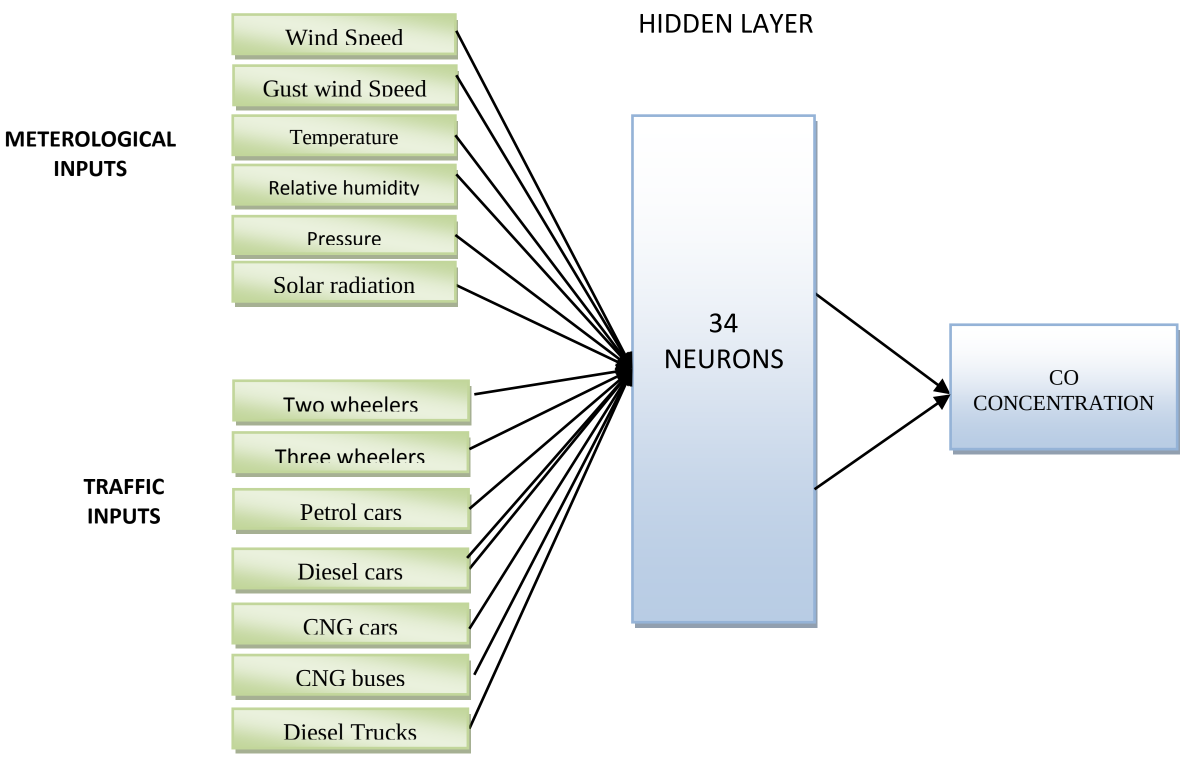


Figure 6.31: MLP back propagation network topology for CO-1 Model

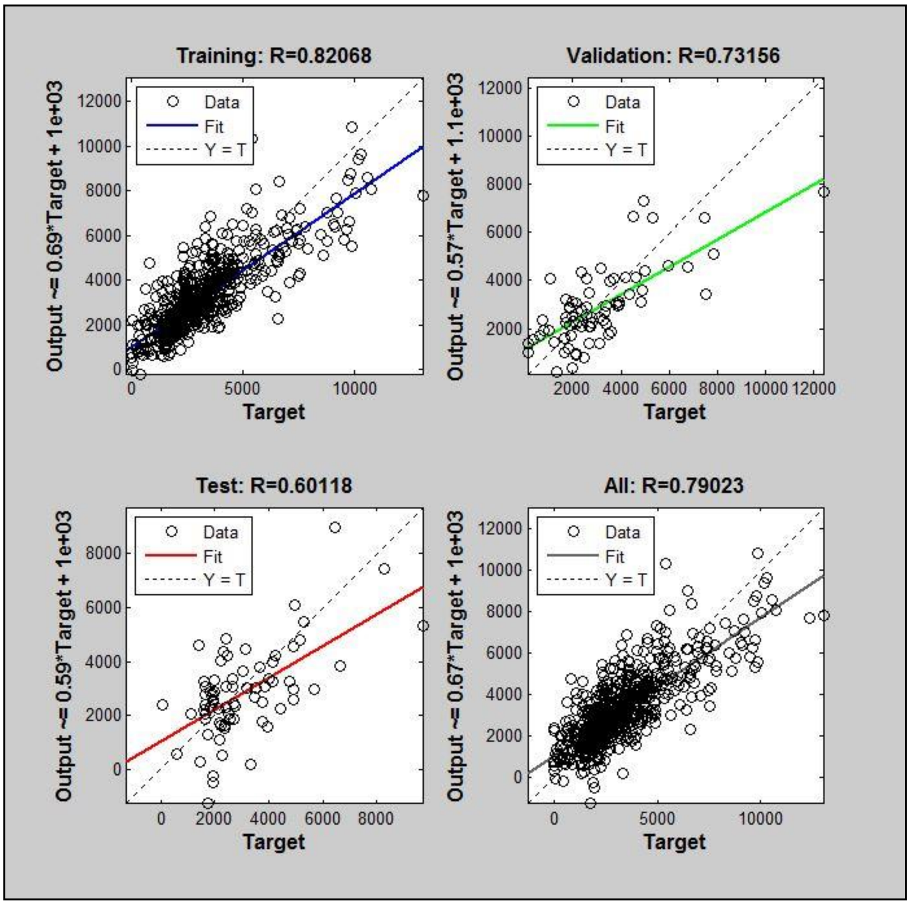


Figure 6.32 : The training,testing,validation in MATLAB for CO-1 Model

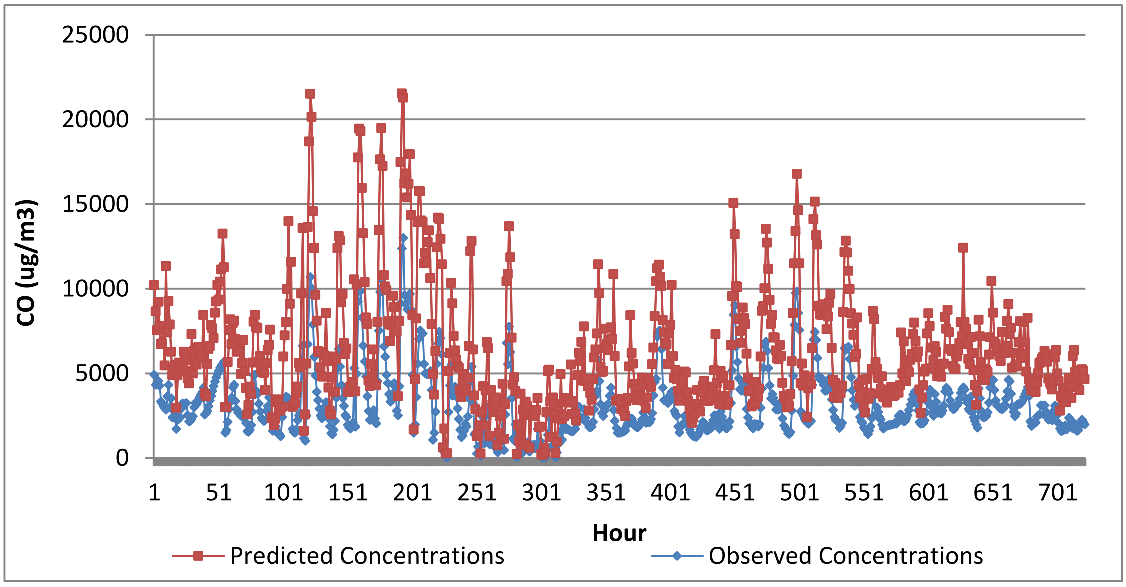


Figure 6.33 : The variation in predicted and measured values by CO-1 Model

6.6.2: MODEL CO-2

i)Topology of network

The CO-2 considered meteorological, traffic parameters and background concentrations (NO_X) as input and concentration of CO as output .The Figure **6.34** shows the average of minimum MSE of each trial with increment of hidden nodes in hidden layer. Five hidden neurons were considered initially and incremented by one hidden neuron upto a maximum of 38 hidden neurons. Based on the sensitivity analysis the results given in **Figure 6.34**, 38 hidden nodes were considered optimal for this network. The **Figure 6.35** shows the network topology for CO-2 model

ii)Network Training, testing and validation

Out of the 722 data sets used for development of model data sets were randomly divided into 75%, 15% and 15% for training, testing and validation of network. The network was stopped after 400 iterations and synaptic weights were stored for network validation and to determine the relative influence of input parameters. The results of training, testing and validation of the network in MATLAB 2012b is given in **Figure 6.36.**The variation of predicted and measured value of CO concentration by CO-2 model is given in **Figure 6.37.**

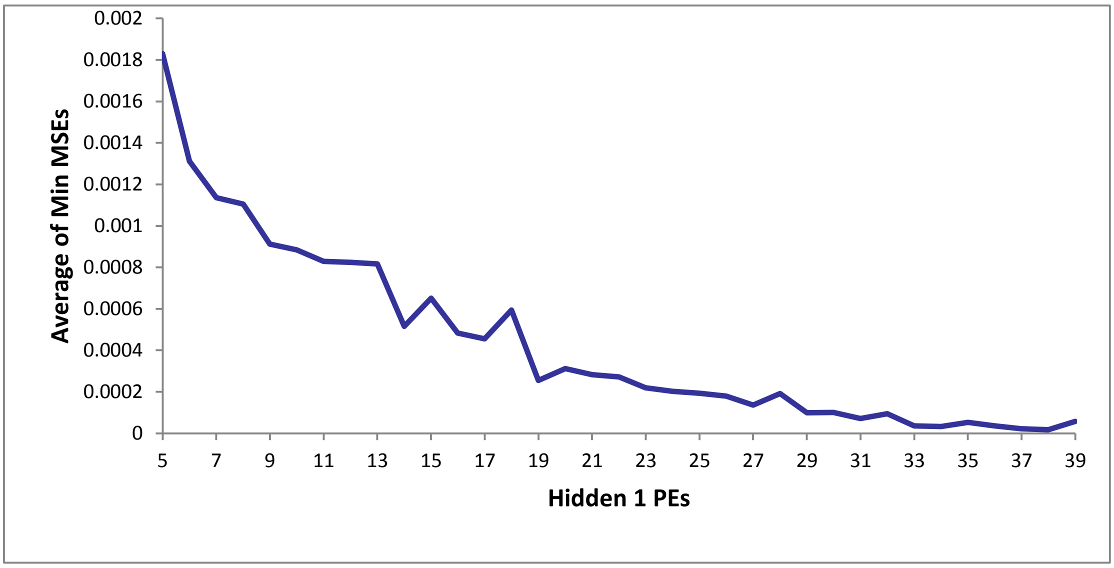


Figure6.34: Estimated optimum number hidden nodes for CO-2 Model

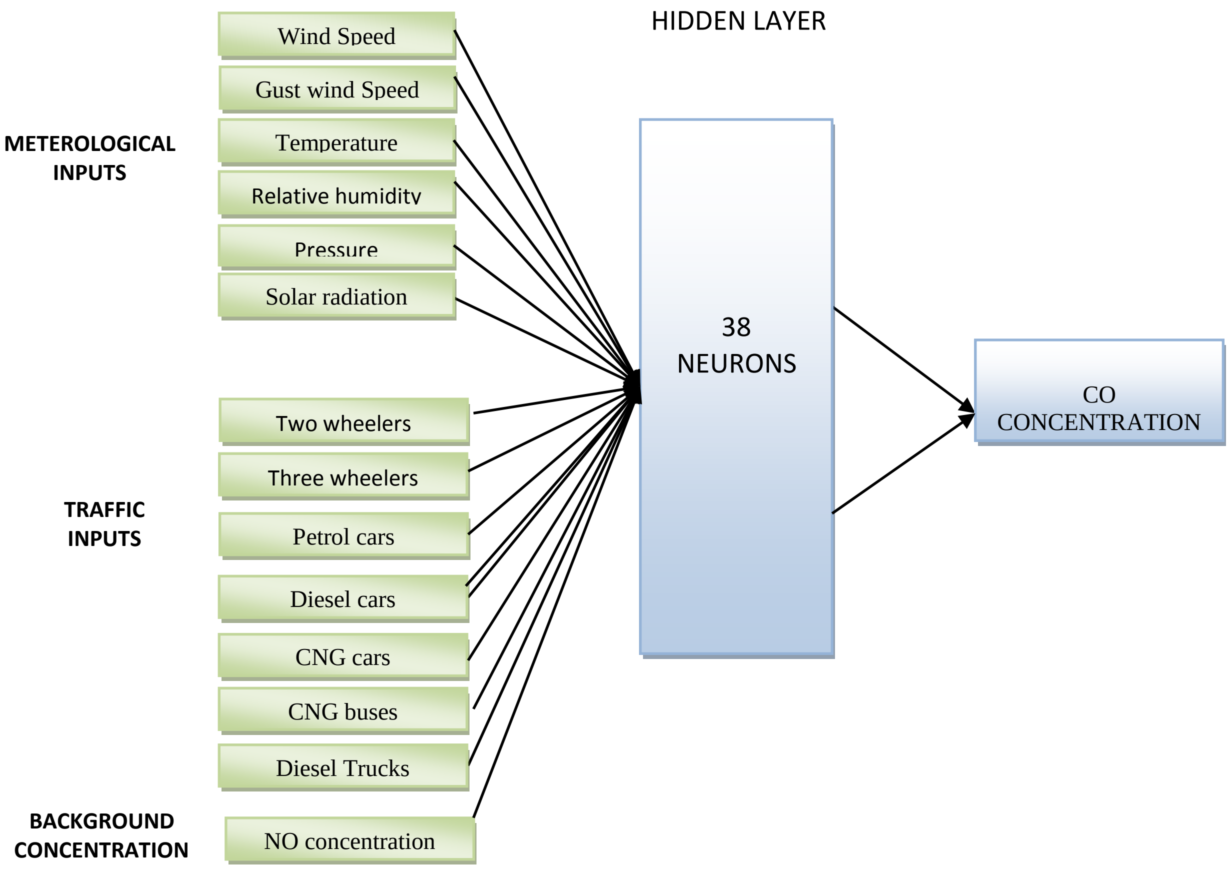


Figure 6.35: MLP back propagation network topology for CO-2 Model

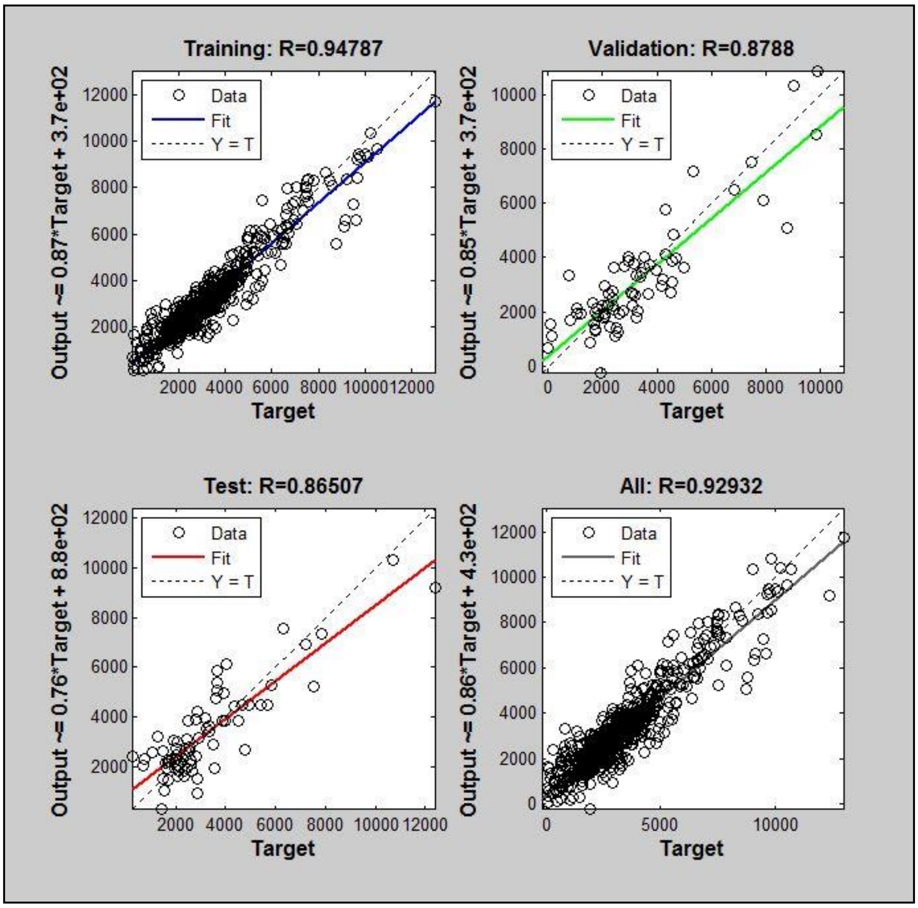


Figure 6.36: The training, testing, validation in MATLAB for CO-2 Model

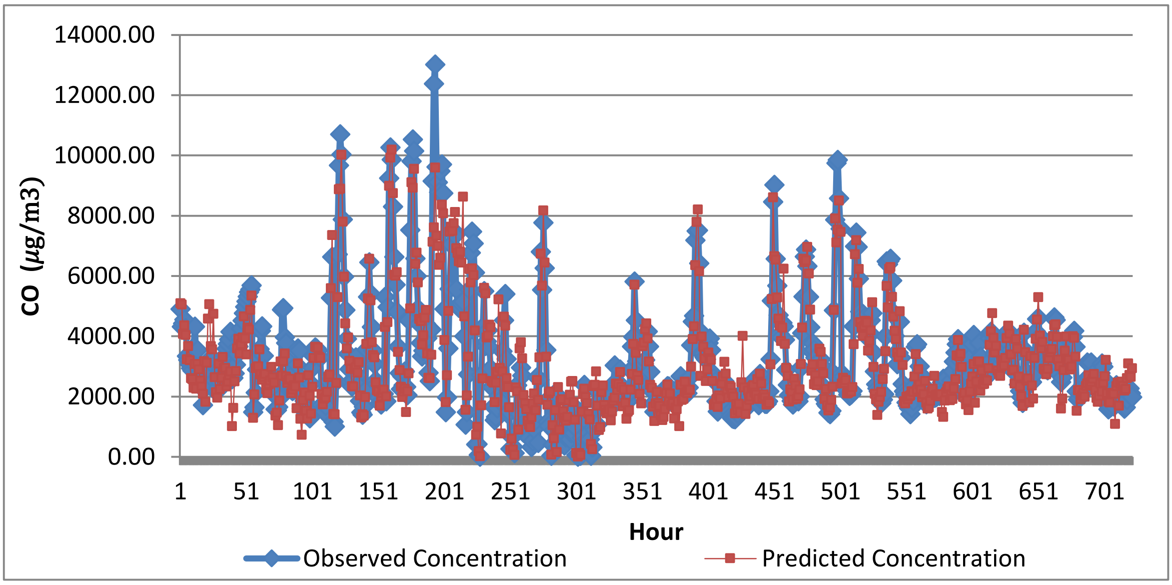


Figure 6.37 : The variation in predicted and observed concentrations of CO-2 model

6.6.3 COMPARISON BETWEEN CO MODELS

The Table 6.6 lists the performance statistics of model CO-1 and CO-2 .$\bar{O}$ and $\bar{P}$ are mean of the observed and predicted concentrations . The σo and σp denotes standard deviation of the observed and predicted concentrations. The MAPE, MAE, RMSE and R^2 values of the predicted values are also given in the table.

Table 6.6 Performance statistics of models CO-1 and CO-2 for CO.

MODEL	STATISTICS							
	$\bar{O}$	$\bar{P}$	σo	σp	MAPE	MAE	RMSE	R^2
	(μg/m3)	(μg/m3)	(μg/m3)	(μg/m3)	%	(μg/m3)	(μg/m3)	Testing
CO-1	3131.98	3008.38	1933.03	1943.56	38.11	572.83	748.92	0.36
CO-2	3131.98	3087.38	1933.03	1857.68	23.82	508.11	736.89	0.73

The difference between the standard deviations between the predicted and observed values is high . This may be the reason of models not being able to produce variations in test data. The mean of both the model is higher than that of observed values indicating tendency of models to over predict. In comparison CO-2 model fares better with lower RMSE,MAPE and MAE values than CO-1model. .

6.6.4 RELATIVE INFLUENCE OF INPUT PARAMETERS ON CO

The relative importance of input parameters on CO concentration in the study area is given in **Figure 6.38** which was computed by examining the synaptic weights obtained after the training process.

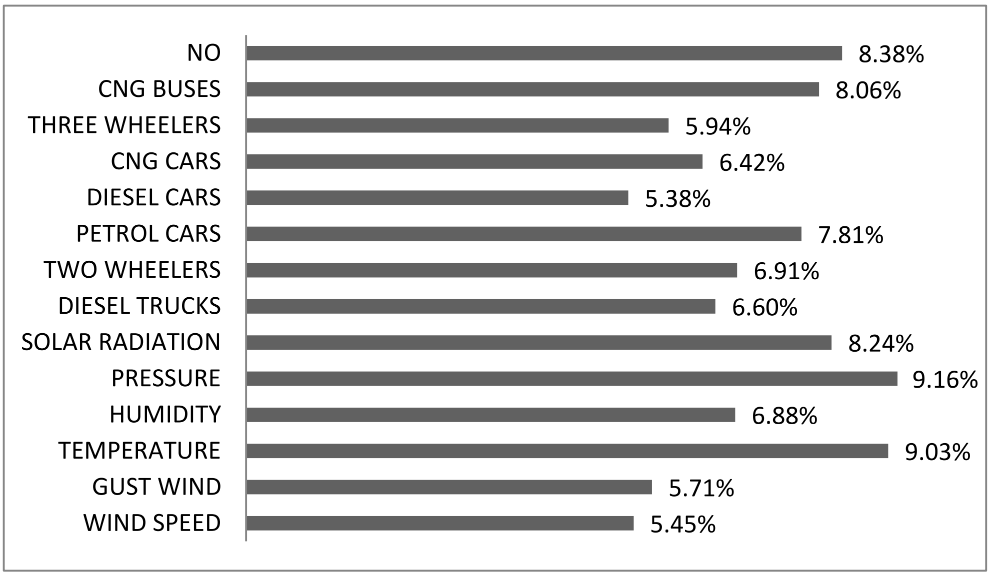


Figure 6.38 : The relative influence of parameters on CO concentration

The **Figure 6.38** shows that temperature, pressure and solar radiation plays an important role in influencing CO concentration .The CNG buses,petrol cars ,two wheelers and CNG buses strongly influence the CO concentrations in decreasing order in the study area.The **Table: 6.6** clearly shows that model performance of CO increases by considering NO_X as input parameter . This explains the influence of NO_X as input parameter as seen in **Figure 6.38**.

Chapter 7
CONCLUSIONS AND RECOMMENDATIONS

7.1 CONCLUSIONS

The following are the conclusions drawn from this study

1. Models RSPM-1, PM-1, NO-1 and CO-1 gives reasonably good predictions, but have tendency to over predict.These models which take meteorological and traffic characteristics as inputs can be effectively used to forecast hourly concentrations of pollution.
2. Models RSPM-2, PM-2, NO-2 and CO-2 which consider meteorological, traffic and background concentrations give better prediction results. These models can be used to find out the missing and erroneous data of monitoring station when other pollutant concentration is known.
3. The performance evaluation of models shows that ANN is suitable for forecasting of particulate matter with good results. The ANN models for NOx and CO gives less accurate predictions with tendency of over prediction.
4. The partitioning of weights of models gives an idea of influence of the traffic and meteorological parameters on the pollutant concentration in the area.

7.2 RECOMMENDATIONS

1. The improved models with larger dataset can be used when the monitoring station is not working properly to give real time predictions and fill missing or erroreous data.The models also eliminate the continuous monitoring of all the pollutants.
2. The relative influence of the traffic on each pollutant can be used to find out the vehicles which affect the pollutant concentration the most and mitigation measures can be formulated to improve the ambient air quality near the area.
3. A user friendly interface can be developed using the trend and pattern of the input parameters on pollutants which would give the air quality details on inputting the meteorological and traffic data for the area.

REFERENCES

1. Alka Srivastava and Ashok K. Sharma **(2014), "**Modeling of Ambient Air Pollutants through Artificial Neural Network in Industrial Area of Ujjain City", IJESRT,2014 ISSN: 2277-9655
2. Aneja, V.P., Agarwal, A., Roelle, P.A., Phillips, S.B., Tong, Q.S., Watkins, N.,Yablonsky,R, **(2001).** Measurements and analysis of criteria pollutants in New Delhi, India. Environment International 27, 35-42.
3. Amirsasha Bnanankhah, Farhad Nejadkoorki **(2012),** "Artificial Neural Networks: A Non-Linear TooL for Air Quality Modeling and Monitoring", International Conference on Applied Life Sciences (ICALS2012) Turkey, September 10-12, 2012
4. AM Aenab, SK Singh, AJ Lafta **(2013),** Critical Assessment of Air Pollution by ANOVA Test and Human Health Effects, Atmospheric Environment, Volume 71, pp 84-91
5. Anil .K Jain,Jianchang Mao**(1996),** "Artificial Neural Networks : A tutorial",IEEE's computer magazine,pp 31-44
6. Anunay A. Gour, S.K. Singh, S.K. Tyagi, Anubha Mandal**(2013),** "Weekday/Weekend Differences in Air Quality Parameters in Delhi, INDIA", International Journal of Research in Engineering & Technology (IJRET), ISSN 2321-8843, Vol. 1, Issue 5, 69-76, October, 2013
7. Azid, I. A. Yusoff, A. R. Seetharamu, K. N**.(2003),** "Application backpropagation neural network in predicting palm oil mill emission",AJSTD., 20: pp. 71-86.
8. A.Vinod Kumar, Rashmi S. Patil and K.S.V. Nambi**(2001),**"Source Apportionment of Suspended Particulate Matter at Two Traffic Junctions in Mumbai India",Atmospheric Environment 35, 4245, 2001
9. Basheer,I.A. Hajmeer, M**.(2000),"**Artificial neural networks: fundamentals, computing, design, and application", Journal Microbiological Methods . 2000. 43: pp. 3-31.
10. Bindhu Lal, Sanjaya Shankar Tripathy **(2012),** " Prediction of dust concentration in open cast coal mine using artificial neural network",Atmospheric Pollution Research 3 (2012) 211-218
11. Bosch **(2005),"** Emissions-control technology for diesel engines". Robert Bosch GmbH, Germany
12. Burtscher H **(2005)** , "Physical characterization of particulate emissions from diesel engines: a review." Aerosol Sci 36:896–932

13. Chan, C. W. Huang, G. H.**(2003)** "Artificial intelligence for management and control of pollution minimization and mitigation processes. Engineering Application of Artificial Intelligence", 2003,16: pp. 75-90.
14. CPCB (Central Pollution Control Board), **2000.** Air Quality Status and Trends in India, NAAQMS/14/1999–2000, New Delhi, 163 pages.

15. CPCB (Central Pollution Control Board), **2008.** Epidemiological Study on Effect of Air Pollution on Human Health (Adults) in Delhi, Environmental Health Series: EHS/1/2008, New Delhi, 317 pages.

16. CPCB. (1998). "Annual Report **(1997-98**)", Central Pollution Control Board, Ministry of Environment and Forest, Government of India, New Delhi
17. Dahe Jianga, Yang Zhanga, Xiang Hua, Yun Zenga, Jianguo Tanb, Demin Shaob**(2004)** "Progress in developing an ANN model for airpollution index forecast", Atmospheric Environment 38 (2004) 7055–7064
18. Demers D, Walters G (**1999**), "Guide to exhaust emission control options." BAeSAME, Bristol
19. E. MARAZIOTIS,L. SAROTIS,C. MARAZIOTI,P. MARAZIOTI**(2008)**, "STATISTICAL ANALYSIS OF INHALABLE (PM10) AND FINE PARTICLES(PM2.5) CONCENTRATIONS IN URBAN REGION OF PATRAS, GREECE", Global NEST Journal, Vol 10, No 2, pp 123-131, 2008
20. F. Nejadkoorki. Advanced Air Pollution.InTech.**2011**.
21. Gour, A.A., Singh, S.K., Tyagi, S.K. and Mandal, A. **(2015)** Variation in Parameters of Ambient Air Quality in National Capital Territory (NCT) of Delhi (India), Atmospheric and Climate Sciences; 5:13-22.
22. Gardner MW, Dorling SR. **(1996), "**Neural network modeling of the influence of local meteorology on surface ozone concentrations", Proceeding of 1st International Conference on Geo Computation, University of Leeds, 1996, 359-370.
23. Gardner MW, Dorling SR. **(1998),** "Artificial neural network (the multi-layer feed-forword neural networks)-a review of application in the atmospheric sciences", Atoms Environ 1998; (30):2627-2637.
24. Goh,C.T.A(1994) " Seismic Liquefaction potential Assessed by Neural Networks",Journal of Geotechnical Engg.ASCE Vol.120 No.9
25. Goyal, P., Sidhartha (**2003**)," Present scenario of air quality in Delhi: a case study of CNG implementation", Atmospheric Environment 37 (2003) 5423-5431

26. Goyal, S.K., Ghatge, S.V., Nema, P., S, M.T**.(2006).** Understanding urban vehicular pollution problem vis–a–vis ambient air quality—case study of a megacity (Delhi, India). Environmental Monitoring and Assessment 119, 557–569.
27. Goyal, Dhirendra Mishra and Anikender Kumar **(2013),"** Vehicular emission inventory of criteria pollutants in Delhi Goyal et al. SpringerPlus 2013, 2:216
28. Harri Niska, Teri Hiltunen, Ari Karppinen, Juhani Ruuskanen, Mikko Kolehmainen**(2004),** " Evolving the neural network model for forecasting air pollution time series", Engineering Applications of Artificial Intelligence 17 (2004) 159–167
29. Hegazy,T.,Fazio,P, and Moselhi,O **(1994),** "Developing practical neural network applications using back-propagation" Journal of Micro computers in civil engineering Vol.pp 145-159
30. Ibrahim Aslan Res¸itog˘lu,Kemal Altinisik,Ali Keskin **(2015)**, "The pollutant emissions from diesel-engine vehicles and exhaust after treatment systems", Clean Techn Environ Policy (2015) 17:15–27
31. K Ravindra, E Wauters, SK Tyagi, S Mor, R Van Grieken **(2006)** , "Assessment of air quality after the implementation of compressed natural gas (CNG) as fuel in public transport in Delhi, India", Environmental monitoring and assessment 115 (1), 405-417
32. Khair MK, Majewski WA **(2006)** "Diesel emissions and their control*",* SAE Inernational. ISBN: 978-0-7680-0674-2
33. Kurt, A. Gulbagci, B. Karaca, F. Alagha, O**(2008),** "An online air pollution forecasting system using neural networks", Environmental International., 34:pp. 592-598.
34. L. H. Tecer **(2007),** "Prediction of SO2 and PM Concentrations in a Coastal Mining Area (Zonguldak, Turkey) Using an Artificial Neural Network", Polish J. of Environ. Stud. Vol. 16, No. 4 (2007), 633-638
35. Lovro Hrust, Zvjezdana Bencetic´ Klaic´, Josip Kri_zan, Oleg Antonic´, Predrag Hercog **(2009),** "Neural network forecasting of air pollutants hourly concentrations using optimised temporal averages of meteorological variables and pollutant concentrations Atmospheric Environment 43 (2009) 5588–5596
36. Luis A. Dı ´ az-Robles, Juan C. Ortega, Joshua S. Fu, Gregory D. Reed, Judith C. Chow, John G. Watson, Juan A. Moncada-Herrera**(2008)**, "A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile", Atmospheric Environment 42 (2008) 8331–8340
37. Lloyd AC, Cackette TA (**2001), "** Diesel engines: environmental impact and control.", Journal Air Waste Managment Assoc 51:809–847

38. Maitha H. Al Shamisi, Ali H. Assi and Hassan A. N. Hejase **(2011),** “Using MATLAB to Develop Artificial Neural Network Models for Predicting Global Solar Radiation in Al Ain City – UAE”, ISBN: 978-953-307-656-0
39. Mohammad Arihami,Nima Khamali,Mohammad Mahdi Rajabi **(2013) “**Predicting hourly air pollutant levels using artificial neural networks coupled with uncertainty analysis by Monte Carlo simulations”, Environ Sci Pollut Res (2013) 20:4777–4789
40. Madhavi Anushka Elangasinghe,Naresh Singhal,Kim N. Dirks,Jennifer A.Salmond**(2014),** “Development of an ANN–based air pollution forecasting system with explicit knowledge through sensitivity analysis”, Atmospheric Pollution Research 5 (2014) 696-708
41. Matter Concentration Using Artificial Neural Networks in an Urban Area”, IRACST – Engineering Science and Technology: An International Journal (ESTIJ) Vol.3, No.4
42. Maricq MM **(2007**), “ Chemical characterization of particulate emissions from diesel engines: a review.” Aerosol Sci 38:1079–1118
43. McCulloch, Warren; Walter Pitts **(1943).** "*A* Logical Calculus of Ideas Immanent in Nervous Activity*"*. Bulletin of Mathematical Biophysics 5 (4): 115–133.
44. Ming Cai,Yafeng Yin,Min Xie **(2009**) Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach Transportation Research Part D 14 (2009) 32–41
45. MODELING”, International Journal of Computer Applications (0975 – 8887) International Conference on Current Trends in Advanced Computing “ICCTAC-2013”
46. MOEF, 1997. White Paper on Air Pollution in Delhi with an Action Plan, Ministry of Environment and Forests, Government of India.
47. Mouhammd Alkasassbeh, Alaa F. Sheta, Hossam Faris and Hamza Turabieh, **(2013),”**Prediction of PM10 and TSP Air Pollution Parameters Using Artificial Neural Network Autoregressive, External Input Models: A Case Study in Salt, Jordan*”,* Middle-East Journal of Scientific Research 14 (7): 999-1009
48. Mohammad Arihami,Nima Khamali,Mohammad Mahdi Rajabi **(2013) “**Predicting hourly air pollutant levels using artificial neural networks coupled with uncertainty analysis by Monte Carlo simulations”, Environ Sci Pollut Res (2013) 20:4777–4789
49. M.W. Gardner*, S.R. Dorling**(1999)**, “ Neural network modelling and prediction of hourly NO*x* and NO2 concentrations in urban air in London”, Atmospheric Environment 33 (1999) 709719
50. NATIONAL AMBIENT AIR QUALITY STATUS (**2008)** ,CPCB annual report

51. NAAQMS/17/2000-2001, 2001. Ambient air quality monitoring status of Delhi Central Pollution Control Board, Delhi

52. Navneeta lal Benjamin, Sarita Sharma, Umesh Pendharker, JK Shrivastava(2014), "Air quality prediction using artificial neuralNetwork**",** IJCS 2014; 2(4): 07-09
53. Narain and Krupnick (**2007), "** The Impact of Delhi's CNG Program on Air Quality"Resources for the Future.
54. Nejadkoorki, F. Baroutian,**(2012**)," Forecasting extreme PM10 concentrations using artificial neural networks. International",Journal Of Environmental Research, 6(1):pp. 277-284.
55. N.Sharma KK chaudhary,CV Chalapati Rao **(2005)**, "Vehicular p[ollution modeling using artificial neural network technique:Review",Journal of Scientific & Industrial research Vol.64,pp. 637-647
56. P Saxena,bharadwaj R,Ghosh C. (2012) " Status of air pollutants after implementation of CNG in Delhi" ,Curr World Environ 2012;7(1):109-115
57. Parida, M., Chandra, S., Ravi Sekhar, C. (2002) "Artificial Neural Network Based Mode Choice Analysis", Highway Research Bulletin, IRC, No.67, Indian Roads Congress, New Delhi
58. R. Prasad, V.R. Bella **(2010) "** Review on Diesel Soot Emission, its Effect and Control." *Bulletin of Chemical Reaction Engineering and Catalysis*, 5(2): 69-86.
59. Rajola, V., Singh S.K. and Kanan,G.K. **(2002)** " Status of air quality in Delhi based on air quality index : criteria for review of existing standards " Journal. Of Environmental Protection, Vol 22, No. 12.
60. Ravi Sekhar Ch.**(1999),** "Mode Choice Analysis Using Neural Networks" Doctoral Dissertation, IIT Roorkee
61. Santosh Kumar Nanda, Debi Prasad Tripathy, S.S. Mahapatra**(2011), "** Application of Legendre Neural Network for Air Quality Prediction",The 5th PSU-UNS International Conference on Engineering and Technology (ICET-2011),
62. Shivaji Bhandarkar **(2013),** "Vehicular Pollution, Their Effect on Human Heatlh and Mitigation Measures",Vehicle Engineering(VE) Volume 1 Issue 2, June 2013
63. Shivraj Sahai, C. Sharma, S. K. Singh and P.K. Gupta **(2010),** "Ambient Carbon monoxide during winter and summer agricultural burning seasons in rural northern India", Asian Journal of Water Environment and Pollution , Vol. 7 ,No. 4. pp 99-104
64. S.M. Shiva Nagendra, Mukesh Khare (**(2002)**, "Line Source Emission Modelling Review", Atoms Environ 36 (13),2083-

65. S.M. Shiva Nagendra, Mukesh Khare (**2006**), “Artificial neural network approach for modelling nitrogen dioxide dispersion from vehicular exhaust emissions*”*,Ecological Modelling 190 (2006) 99–115
66. Sonaje N P,Mane S J,Kote A S **(2013), “**Modelling of Respirable Suspended Particulate
67. Sonia Tyagi, Preeti Teotia, Arun Kumar, Mukesh Chand, Rajpal Tyagi and Hirdesh Kumar**(2015),** “Study of carbon monoxide in ambient air of Delhi”,WORLD JOURNAL OF PHARMACY AND PHARMACEUTICAL SCIENCES Volume 4, Issue 03, 645-654.
68. Sudhir Nigam, Rashmi Nigam,Sangeeta Kapoor **(2013),** “FORECASTING CARBON MONOXIDE CONCENTRATION USING ARTIFICIAL NEURAL NETWORK
69. Suhasini V. Kottur, Dr. S. S. Mantha**(2015),** “An Integrated Model using Artificial Neural Network (ANN) and Kriging for Forecasting Air Pollutants using Meteorological Data”,International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 1, January 2015
70. Suthirto, P., 2002. There's Cancer in the Air. Hindustan Times,New Delhi.
71. **U.S E.P.A ,**The United States Environmental Protection Agency**,http://www.epa.gov/**
72. W. Kaminski J. Skrzypski, E.Jach-Szakiel,Application of Artificial Neural Networks (ANNs) to Predict Air Quality Classes in Big Cities 19th International Conference on Systems Engineering