A REPORT ON AIR POLLUTION MODELLING

submitted in partial fulfilment of the requirements
for the degree of

(Master Of Science)

in

Mathematics

by

Rohit Mittel

(Admission No.: I15MA031)

under the supervision of

Dr. Twinkle R. Singh

Assistant Professor of Mathematics

DEPARTMENT OF APPLIED MATHEMATICS & HUMANITIES
SARDAR VALLABHBHAI NATIONAL INSTITUTE OF TECHNOLOGY,
SURAT-395007, GUJARAT, INDIA.

2020

Department of Applied Mathematics & Humanities
Sardar Vallabhbhai National Institute of Technology,
(An Institute of National Importance, NIT Act 2007)
Surat-395007, Gujarat, INDIA.

Approval Sheet

The dissertation entitled **"Air Pollution Modelling: Analysis and Prediction of Concentration of Pollutants using Mathematical Models"** by **Rohit Mittel (I15MA031)** is hereby approved for the degree of Master of Science in Mathematics.

Dr. R. K. Meher
(Internal Examiner)

Dr. Sushil Kumar
(Internal Examiner)

Dr. Twinkle R. Singh
(Supervisor)

Dr. Kalpana Meheria
(Chairperson)

Date:
Place: Surat

Department of Applied Mathematics & Humanities
Sardar Vallabhbhai National Institute of Technology,
(An Institute of National Importance, NIT Act 2007)
Surat-395007, Gujarat, INDIA.

Declaration

I hereby declare that the dissertation entitled **"Air Pollution Modelling: Analysis and Prediction of Concentration of Pollutants using Mathematical Models"** is a genuine record of research work carried out by me and no part of this thesis has been submitted to any university or institution for the reward of any degree or diploma.

Rohit Mittel

(I15MA031)

Applied Mathematics and Humanities Department,

Sardar Vallabhbhai National Institute of Technology,

Surat - 395007.

Date:
Place: Surat

Department of Applied Mathematics & Humanities
Sardar Vallabhbhai National Institute of Technology,
(An Institute of National Importance, NIT Act 2007)
Surat-395007, Gujarat, INDIA.

Date:__________

Certificate

This is to certify that the dissertation entitled ”**Air Pollution Modelling: Analysis and Prediction of Concentration of Pollutants using Mathematical Models”** is a bonafide record of research work done by **Rohit Mittel (I15MA031)** under my supervision.

Supervisor

Dr. Twinkle R. Singh

Assistant Professor of Mathematics,

Applied Mathematics and Humanities Department,
Sardar Vallabhbhai National Institute of Technology,Surat.

ACKNOWLEDGEMENTS

I would like to express my deep and sincere thanks to **Dr. Twinkle R. Singh** for her continuous support of my project work. She motivated me, provided valuable guidance and useful suggestions on how to proceed and improve my thesis. I could not have imagined having a better advisor and mentor for my dissertation study.

I am thankful to the Applied Mathematical and Humanities Department and my Institute for providing me the opportunities to work with freedom. I shall always be grateful to Dr. Sushil Kumar, Head of Department of Applied Mathematics & Humanities.

I would like to thank my parents and brother for their constant support and encouragement. I am thankful to all my friends who helped directly and indirectly. They were the source of learning as well as a good company for me throughout the project work.

Rohit Mittel
Surat.

PREFACE

Rohit Mittel

Surat.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Today the world is facing an unprecedented pandemic situation. Many countries have been put under lock down to prevent the rapid spread of the corona virus among the general population, resulting in closure of offices, educational institutions, factories and several other institutions where public gatherings are possible. Governments have no idea on how to deal with this situation and everyone is trying to do their best as per their capabilities and resources available to them, so that they can come out of this situation with the least amount of damage to their economic structure. Governments are using mathematical models to make strategies on dealing with this situation by analysing various scenarios that may arise due to the pandemic and measures taken to control it. Various studies are being conducted which involves the corona virus, the effects of lock down on public health etc. One of the studies conducted also indicated that people living in heavily polluted cities are more prone to be infected by the corona virus than those living in cleaner environment.

A major impact of this lock down has been that a significant decrease in air pollution has been noticed in various cities of many countries which has led to people breathing in fresher air every morning. This has resulted in people becoming aware about the various health problems caused due to the polluted air we breathe.

This project aims to study some of the mathematical models that are being employed in the real world to analyse and predict the air pollution of a region under study. But before we look at the mathematical models we must know what is air pollution, how is it caused, what are its effects on humans as well as the environment and how can it controlled.

1.1 What is Air Pollution?

One can define air pollution in many ways, but to simply put it, when foreign particles having harmful or poisonous effects on the environment contaminate the air is called ***Air pollution***. We can also define, ***Air pollution*** as the presence of chemicals or toxic compounds in the air, at levels such that it reduces the air quality and poses a health risk. These foreign particles, chemicals or toxic compounds are called air pollutants.

Air pollution has increased massively in recent decades. One of the major factors that has led to the increase in air pollution is the rise in population of the world and by 2030 estimates suggest that it will be 8.5 billion. The increasing population is bound to lead to a significant rise in:

- Industrial activities as demands for various commodities will increase and thus leading to increase in industrial pollution,
- in the number of vehicles on the road which in turn will lead to higher emission of harmful particulates into the atmosphere.
- other anthropogenic activities.

The past decade has seen a rise in the efforts taken by governments and other regulatory agencies throughout the world (including India) to reduce air pollution and improve air quality. But still we are facing air quality issues such as photochemical smog and visibility degradation etc.

1.1.1 Pollutants

The foreign particles or substances that are introduced into the atmosphere are called ***pollutants***. These are introduced in the area by emissions from vehicles, industries and various anthropogenic activities. Pollutants can be solid, liquid or gases. The common substances emitted in the atmosphere are Particulate Matter (PM_{10} and $PM_{2.5}$), Carbon Monoxide(CO), Nitrogen Oxides(NO and NO_2) and Ozone. Inhalation of these particulates affects the normal development of lungs and leads to respiratory problems such as asthma. Many studies on effects of air pollution on humans have pointed out that many people have died prematurely every year after inhaling particulate matters present in the atmosphere. Pollution not only affects the humans but also the natural processes of the environment like occurrence of acid rain, Global warming effect, depletion of the ozone layer etc.

Table 1.1 gives a list of some common pollutants, their sources and effects of these pollutants on environment.

1.2 Ways to Reduce Air Pollution

One of the best methods to reduce air pollution is to create and increase awareness among people about causes and harmful effects of air pollution, and about measures that can be taken to prevent air pollution. One of the biggest contributors to air pollution is the huge amount of vehicles on the road that emit these harmful gases. If people take steps like carpooling they can help reduce pollution from vehicles as the number of vehicles on road will decrease. Also pollution from industries also can be curbed by using various filtering devices in the chimneys of the factories which will help in reducing the amount of harmful gases being released in the atmosphere.

Past decade has seen a great rise in development of technologies and how it has become vital in our day to day life. Thus, incorporating available technologies and developing new technologies for creating awareness or developing new devices to help reduce pollution is necessary. But before we can use technology to our advantage we must know where it must be applied and how can it help us. This is where mathematical modelling comes in handy and helps us in understanding what is needed to curb air pollution.

Pollutants	**Source of Production**	**Effect on Humans and Environment**
Oxides of Carbon (CO_X): Carbon dioxide (CO_2) Carbon monoxide (CO)	Combustion of coal Oil and other fuels for energy production manufacturing and transport Biomass burning	CO_2 has a major role in green house effect, produces weak carbonic acid adding to acid rains, CO affects human health by binding to haemoglobin which may result in asphyxia.
Oxides of Sulphur (CO_X): Sulphur dioxide (SO_2) Sulphur trioxide (SO_3) Sulphate (SO_4)	Combustion of sulphur containing fuel e.g., coal, petroleum extraction and refining; paper manufacturing; municipal incinerating; ore smelting for metal extraction	SO_2 has maximum deleterious effects as it damage to human and other animal lungs , is important precursor to acid rain; adverse effects include corrosion of paint, metals and injury or death to animals and plants.
Oxides of Nitrogen (NO_X): Nitrogen oxide (CO) Nitrogen dioxide (CO_2) Nitrous oxide (N_2O) Nitrate (NO_3)	Burning of fuels; biomass burning; by product in the manufacturing of fertilizers	Form the secondary pollutants: peroxy acetyl nitrate (PAN) and nitric acid (HNO3); suppression of plant growth and tissue damage; cause irritation to eyes, viral infections like influenza; nitrate form in atmosphere impairs the visibility whereas in soil promotes the plant growth.
Suspended Particulate Matter (SPM-solid particles) Dust, soil, sulphate Salts, heavy metal salts, Fire particles of carbon (soot), silica, asbestos, Liquid sprays, mist etc.	Fuel combustion; building constructions; mining; thermal power stations; stone crushing; industrial processes; forest fires; refuse incineration.	Have chronic effects on respiratory system; deposition on the surface of green leaves thus interfering with absorption of C02 and release of 02; blocking of sunlight; particles size that range between 0.1 to 10um, cause greatest lung damage.
Photochemical oxidants Ozone (O_3)	Photochemical reactions in the atmosphere that involve sunlight, oxides of nitrogen and hydrocarbons	Produce haze; irritation to eyes, nose and throat; respiratory problems; blocking of sunlight

Table 1.1: Sources and Effect on Humans and Environment of some common air pollutants

CHAPTER 2

NEED FOR AIR POLLUTION MODELLING

For decades, both researchers and administration around the world have been challenged by the dynamic and heterogeneous nature of the natural environment to develop efficient air pollution management methods. Reducing air pollution impacts and improving air quality, requires constant efforts in order:

- to build extensive inventories of pollutant emissions,
- to determine the source, substance and dispersion rates of these emissions,
- to develop computer-based numerical models based on mass conservation flows,
- To evaluate concentration levels and air pollution intensity at each location in a particular metropolitan area.

The modelling of air pollution thus plays a major role in shaping various policies and strategies on air pollution control.

Air pollution/quality modelling is a term used for computational methods that explain the relationships between emissions, meteorological factors (like wind speed, temperature), ambient concentrations, deposition and other variables. Measuring air pollution provides significant, quantitative information on environmental concentration and deposition but can only explain air quality at particular locations and times, without providing clear guidance on the causes of the problem of air quality. Instead, air pollution modelling will include a more comprehensive deterministic explanation of the issue of air quality, providing an overview of factors and causes (emission sources, meteorological processes, and physical and chemical changes), as well as some advice on mitigation steps.

Modelling can be used to run scenarios, to test theories and understand environmental impact under different emission rates, weather and development scenarios. There are lots of different methods and techniques, but the goal is always the same – make an assessment of pollutant impact over a given area using an existing set of data. To do this we make some assumptions, we use some rules, and we add some data. Air pollution modelling is central in regulatory, academic, and forensic applications.

Air pollution modelling helps in answering the following questions:

- What are the relative contributions to concentrations of air pollutants from various emission sources?
- What emission reductions are needed for outdoor concentrations to meet air quality standards?
- Where should a planned source of emissions be sited?
- What will be the change in concentrations of pollutants if the emissions of precursor air pollutants are reduced by a certain percentage?
- What will be the future state of air quality under certain emission reduction scenarios?
- Which meteorological factors have the most effect on the air quality of a certain region?

Before selecting a suitable model to predict and analyse air pollution we must answer the following questions:

1. What will the model be used for?
2. What input data are available?
3. What meteorological and topographical situation is to be modelled?
4. What computer equipment is available?
5. What funds can be assigned to the modelling project?

Answering these question will help us in identifying which is the best model that we can use to give us the best possible output.

CHAPTER 3

PARAMETERS OF AIR POLLUTION MODELLING

A mathematical model developed for a real world problem requires complete knowledge of factors or parameters that affect the output of the model. Therefore, before we look at the models that are used to analyse air pollution we must familiarise ourselves with the various parameters that influence the spread or dispersion of pollutants in the atmosphere. Parameters affecting air pollution can be broadly classified into three categories:

1. Meteorological Condition: Wind speed and direction, Humidity, Rainfall, Atmospheric temperature, Solar Radiation etc.
2. Emission parameter: Type of source, emission rate, etc.
3. Topographical Factors: Rural or Urban area, terrain elevation(Mountain and Valleys), etc.

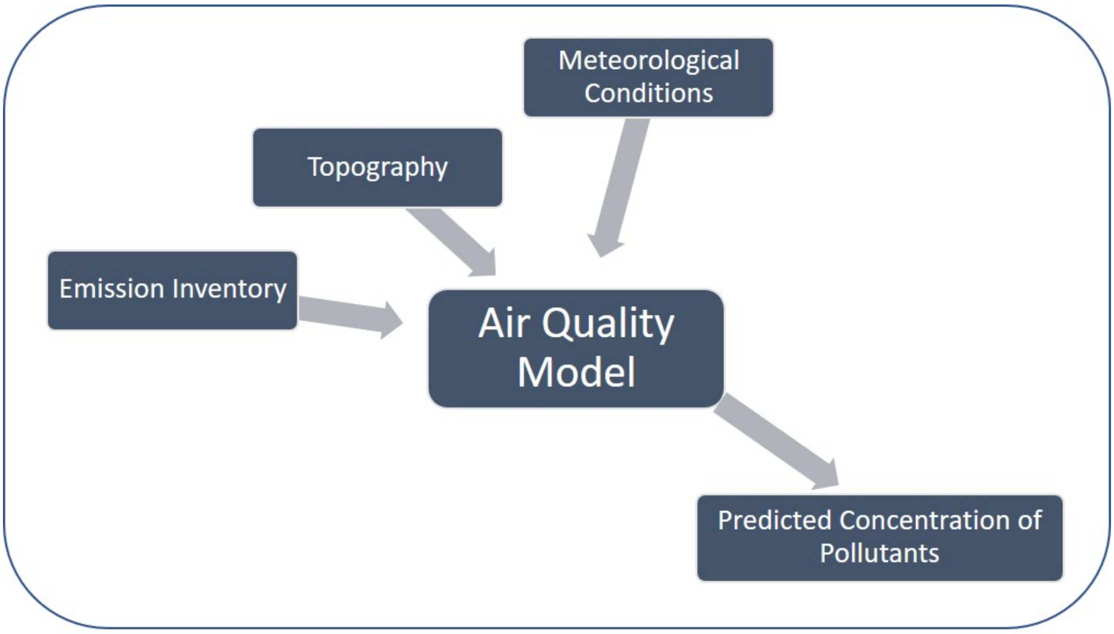


Figure 3.1: Structure of Air Pollution Model

3.1 Meteorological Factors

Wind movements have a significant impact on the fate of air contaminants. It is therefore important to provide a analysis of local weather patterns if we want to examine air pollution in any region. Meteorological data is typically distributed over one-hour time intervals. Study of meteorological data helps:

- to identify source of pollutants
- to predict air pollution events such as inversions and high-pollutant concentration days
- simulate and predict air quality using computer models.

Let us now look at some meteorological factors and how do they affect the air pollution.

3.1.1 Wind speed and direction

Wind disperses contaminants in various directions, depending on its speed and direction. If the air is still, the accumulation of contaminants in the area can build up because they are not dispersed. When powerful winds are blowing, contaminants are spread rapidly, resulting in lower concentrations of pollutants.

Mathematically, we can describe wind as a velocity vector with direction and speed, the vertical component of which is relatively small than the horizontal component. Continuous release of contaminants at the point of release is diluted by wind speed. In the plume, concentrations are inversely proportional to the wind speed. The wind speed is measured in the meter / second(m / s) unit. Wind speed is determined by means of an instrument called anemometer. The sound anemometer works on the theory that wind speed influences the time it takes for the sound to pass from one point to another. It takes less time for sound to travel with the wind than for sound to travel through the wind. Sonic anemometers can determine both the wind speed and the direction by simultaneously measuring sound wave frequencies in two different directions.

3.1.2 Temperature

Temperature and sunlight (solar radiation) play an significant role in the chemical reactions that occur in the atmosphere to create photochemical smog from other contaminants. Favourable conditions can result in increased smog concentrations.

The most common way to calculate temperature is to use a material with a resistance that varies with temperature, such as platinum wire. These changes are measured by a sensor and translated into a temperature reading.

3.1.3 Solar Radiation

It is important to track solar radiation for use in the simulation of photochemical smog events, as the strength of sunlight has an significant impact on the rate of chemical reactions that cause smog. The cloudiness of the sky, the time of day and the geographical position all influence the strength of the sunlight.

3.1.4 Humidity

Like temperature and solar radiation, water vapour plays a significant part in many thermal and photochemical reactions in the atmosphere. Since water molecules are small and highly polar, they can bind to a large number of substances. If they are attached to particles suspended in the air, they can dramatically increase the amount of light dispersed by the particles (measuring visibility). When water molecules are bound to corrosive gases, such as sulphur dioxide, the gas can dissolve in the water and create an acid solution that can harm health and property.

Water vapour content of air is defined as a percentage of the saturation vapour pressure of water at a given temperature. This is known as the relative humidity. The amount of water vapour in the atmosphere varies greatly, depending on the geographical location, the position of the water sources, the wind direction and the temperature of the air. Relative humidity is usually higher during the summer, when temperatures and rainfall are also at their peak.

3.1.5 Rainfall

Rain has a 'scavenging' effect as it washes particulate matter from the atmosphere and dissolves gaseous contaminants. Particle removal increases visibility. For regions where heavy rainfall is common, air quality is usually higher. When the rain dissolves gaseous contaminants, such as sulphur dioxide, acid rain can form, resulting in potential material or vegetation damage.

3.2 Emission Factors

We can categorize emission sources into four main categories:

1. Mobile sources: like cars, buses on roads, trains and air planes
2. Stationary sources: like industries and factories, power plants and oil refineries
3. Natural sources: like volcanoes, wildfires, sand storms etc.
4. Area sources: like agricultural areas, cities and wood burning fireplaces etc.

Out of these sources mobile sources and stationary sources are the major contributors of air pollution. The presence or absence of any kind of source in an area will affect the air quality of the area.

Rate of emission also impact the concentrations of a pollutant in a region. If the pollutant is being released in the atmosphere at very high rate then its concentration at that rate will be high at the point of release compared to what it will be in the surrounding areas. For example, if there are many industries in a region then the air quality of that region will be very low as compared to the area where there are no or less number of industries. Similarly, high traffic will result in high air pollution in that area than where there is low traffic.

3.3 Topographical Factors

Topographical features such as mountains, valleys and urban areas all have an impact on the air quality of a region. This is because meteorological phenomenons described above are affected as the topography changes. Let us look at how topographical features affect air pollution.

In mountainous regions and valleys, wind speed and direction can vary considerably from one location to another. Also setting up industries and factories, that are one of the major causes of air pollution, is not an easy task in the mountainous regions and hence as there is less emission of pollutants there is less pollution. In addition to the topographical features of the region, the atmospheric and surface thermal characteristics influence the movements of the air, especially at low wind speeds. A region experiencing high rainfall, like north-eastern states of India or an area with huge forest cover will have low concentrations of air pollutants and air quality in such regions will be good as rain will cleanse the air of the pollutants.

For urban settings, wind speed and direction in a street canyon can be very difficult. Regional wind levels can be higher or lower than would otherwise have been the case in the absence of heat emissions from buildings. It can be easily concluded that the flow of air around urban buildings would be entirely different from the flow of air in rural areas. The dispersion of air contaminants is therefore a dynamic phenomenon and the concentrations in the street canyon vary in different wind directions over the urban buildings and the various shapes of the buildings.

CHAPTER 4

MATHEMATICAL MODELS FOR AIR POLLUTION

4.1 Classification of Air Pollution Models

In this chapter, some of the mathematical models that have been developed in order to analyse and predict the air pollution of a certain area are discussed.

Since the realization that air pollution is a growing problem for the world, researchers worldwide have developed a variety of Air Quality Prediction Models, ranging from easy to sophisticated approaches, for forecasting pollutant concentrations. There are several ways to classify the variety of available models according to their specific characteristics. The most important attributes are:

1. Basic model structure:
    - deterministic
    - non-deterministic
2. Frame of reference:
    - Eulerian
    - Lagrangian
3. Time resolution:
    - steady state
    - time dependent
4. Methods of model equations resolution
    - analytical
    - numerical
5. Dimensionality of computational domain:
    - one dimensional,
    - two dimensional,
    - three dimensional
    - multi dimensional
6. Source – receptor relationship:
    - source oriented (point, area, line, volume)

    - receptor oriented (street canyon, intersection model etc.)

7. Time horizon: hour, day, month, year
8. Size: local, regional, national, global
9. Pollutant of concern (SO_2, NO_x, CO, SPM, photochemical oxidant etc.)

However, the most important and popular way of classifying air pollution models is based on the basic model structure and the approach used for the closure of the turbulent diffusion equation which is widely used in urban air pollution modelling. Based on model structure air pollution model can be divided as:

- Deterministic or Dispersion Models
- Non-deterministic Models
  - Statistical Models
  - Physical Models

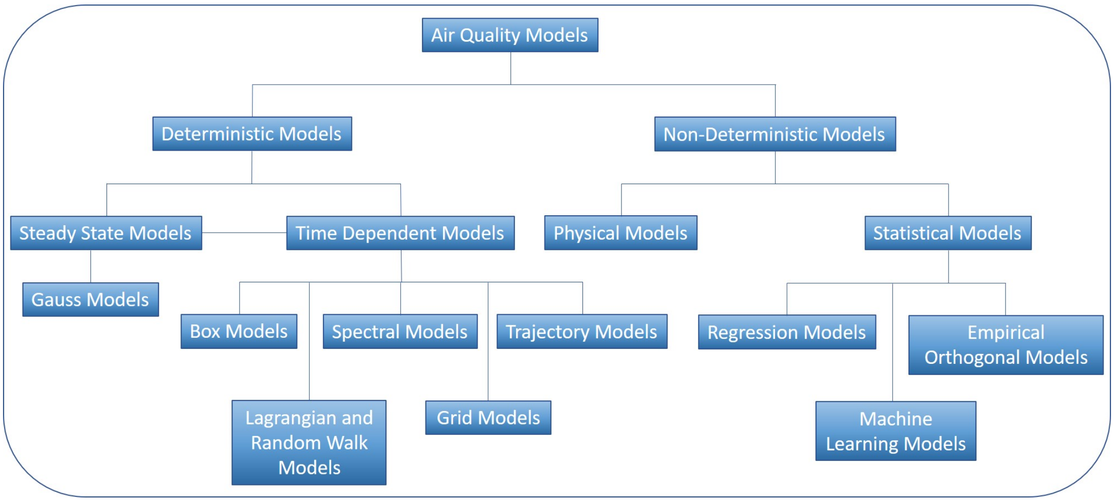


Figure 4.1: Classification of Air Pollution Models

4.2 Air Pollution Models

4.2.1 Deterministic or Dispersion Model

Deterministic mathematical models(DMMs) or Dispersion models measure concentrations of pollutants from emissions inventories and meteorological variables based on solutions of different equations that describe the respective physical processes. In other words, a dispersion model is a collection differential equations formed by comparing the rate of change in the concentration of contaminants to average wind and turbulent diffusion, which, in turn, is derived from the theory of mass conservation. It simulates the process of atmospheric dispersion, which involves combining the contaminants with the existing air. The pollutants are carried by the wind and their concentration is reduced by atmospheric turbulence. The dispersion model estimates the concentration rates of the pollutants at any point in space and at any time, depending on the availability of meteorological data.

The dispersion model aims to resolve the following issues:

- Evaluation of compliance with air quality guidelines and/or requirements for concentrations of pollutants from proposed installations(including known concentrations from other sources);
- Evaluation of the effects of the emissions of a plant that is to undergo a process change;
- Determination of appropriate stack heights;
- Evaluation of the contribution of individual plants to the total concentrations;
- Designing networks for ambient air monitoring;
- Evaluating the effects and efficacy of regulation and mitigation approaches (e.g. the influence of pollution standards);
- forecasting pollution episodes;
- Evaluation of risks and preparation for the control of rare occurrences, such as accidental leakage of hazardous substances;
- estimating concentrations of contaminants that are too difficult or costly to measure;
- Estimate the effect of geophysical factors on dispersion (e.g. elevation of terrain, presence of water bodies and land use);

- Tracking the source of accidental releases of hazardous substances.

A dispersion model can replace monitoring and save cost and time.

Not every dispersion model can tackle the many processes that occur in the atmosphere during air pollution transport. Usually, the model considers some of these phenomena and leaves others out. Selecting a specific model depends on the task that needs to be addressed.

Box Models

Box models are based on the theory of conservation of mass. The receptor is considered as a box into which contaminants are released and undergo chemical and physical processes. Input to the model is simple meteorology.

However, following inputs of the initial conditions a box model simulates the formation of pollutants within the box without providing any information on the local concentrations of the pollutants. Box models are not suitable to model the particle concentrations within a local environment, as it does not provide any information on the local concentrations, where concentrations and particle dynamics are highly influenced by local changes to the wind field and emissions.

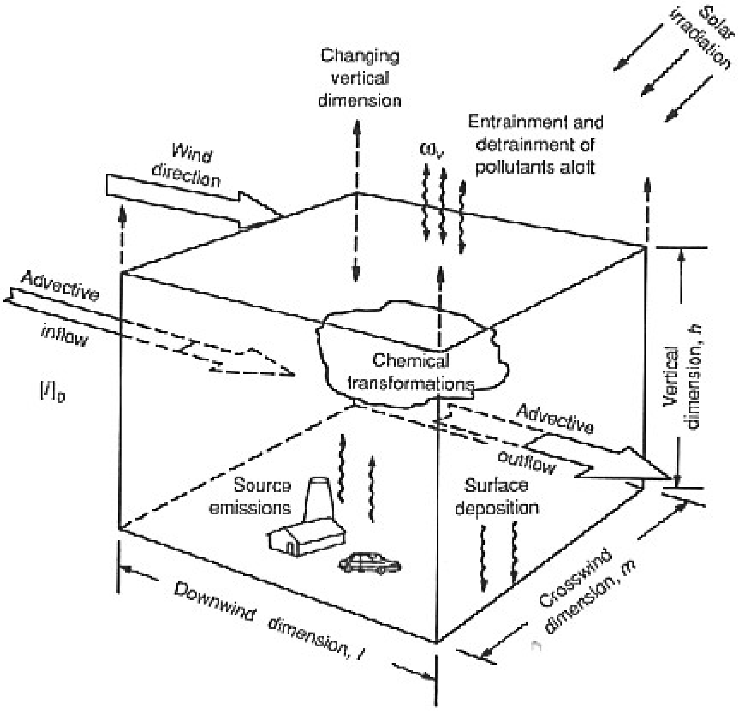


Figure 4.2: Graphical Representation of The Box Model

To estimate the pollutant concentration in a city, we consider air pollution in a rectangular box under the following major simplifying assumptions:

- The city is a rectangle with dimensions W (downwind) and L (crosswind), both in units [m]. The box is defined by $W \cdot L \cdot H[m^3]$ where $H[m]$ is called the mixing height;
- The air pollution emission rate of the city is $Q[g/s]$, which is independent of space and time (continuous emission). Q is related to emission rate per unit area, $q[g/s.m^2]$ by:
$$Q = q \cdot (W \cdot L) \qquad [g/s] \tag{4.1}$$
- The mass of pollutant emitted from the source remains in the atmosphere. No pollutant leaves or enters through the top of the box, nor through the sides that are parallel to the wind direction. No deposition, including gravitational settling or turbulent impaction occurs. No material is removed through chemical transformation (destruction rate equals zero).
- Atmospheric turbulence produces complete and spatial uniform mixing of pollutants within the box.
- The turbulence is strong enough in the upwind direction that the pollutant concentrations [mass/volume] from releases of the sources within the box and those due to the pollutant masses entering the box from upwind side are spatially uniform in the box.
- The wind blows with an average (constant) velocity $u[m/s]$.

These assumptions lead to a steady state situation and the accumulation rate is zero. All the terms can then be easily quantified and calculated. If $\chi(t)[g/m3]$ denotes the pollutant concentration in the box as a function of time t and χ_{in} the (constant) concentration in the incoming air mass, the:

Flow-in rate = $L \cdot H \cdot u \cdot \chi_{in}$

Flow-out rate = $L \cdot H \cdot u \cdot \chi(t)$

Emission rate = Q

Destruction rate = 0

Accumulation rate= $W \cdot L \cdot H \cdot \dfrac{d\chi(t)}{dt}$

Then the differential equation emerges:

$$W \cdot L \cdot H \cdot \frac{d\chi(t)}{dt} = Q + L \cdot H \cdot u \cdot \chi_{in} - L \cdot H \cdot u \cdot \chi(t) \tag{4.2}$$

which has the solution:

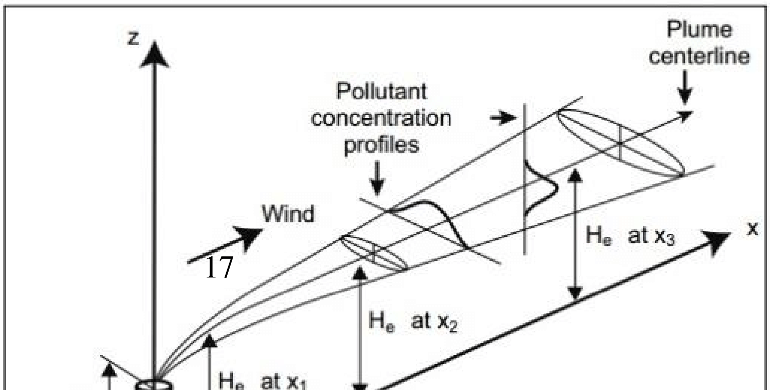
$$\chi(t) = Q/(L \cdot H \cdot u) \cdot (1 - exp(-u \cdot t/W)) \tag{4.3}$$

For longer times t the concentration approaches a steady state ($\chi(t) = Q/(L \cdot H \cdot u)$) which corresponds to zero accumulation rate.

There are several drawbacks of box models. Firstly, some of the assumptions are unrealistic (e.g. wind speed independent of height or uniformity of air pollutant concentrations throughout the box). Secondly, the model does not distinguish a source configuration of a large numbers of small sources emitting pollutants at low elevation (cars, houses, small industry, and open burning) from that of a small numbers of large sources emitting larger amounts per source at higher elevation (power plants, smelters, and cement plants). Both types of sources are simply added to estimate a value for the emission rate per unit area (q). Of two sources with the same emission rate, the higher elevated one leads to lower ground level concentrations in reality. As there is no way to deal with this drawback, box models are unlikely to give reliable estimates, except perhaps under very special circumstances.

Gaussian dispersion models

The Gaussian Plume Model is one of the earliest models still widely used to calculate the



maximum ground level impact of plumes and the distance of maximum impact from the source. These models are extensively used to assess the impacts of existing and proposed sources of air pollution on local and urban air quality. An advantage of Gaussian modeling systems is that they can treat a large number of emission sources, dispersion situations, and a receptor grid network, which is sufficiently dense spatially (of the order of tens of meters). Figure 4.2 shows a buoyant Gaussian air pollutant dispersion plume. The width of the plume is determined by σ_y and σ_z, which are defined by stability classes. The Gaussian model is based on the following assumptions:

- continuous emissions [mass/time unit, usually g/s];
- conservation of mass;
- steady-state meteorological conditions for the travel time of pollutant from source to receptor;
- concentration profiles in the crosswind direction and in the vertical direction (both perpendicular to the path of transport) are represented by Gaussian or normal distributions
- that atmospheric stability and all other meteorological parameters are uniform and constant throughout the layer into which the pollutants are discharged, and in particular that wind speed and direction are uniform and constant in the domain;
- that turbulent diffusion is a random activity and therefore the dilution of the

pollutant can be described in both horizontal and vertical directions by the Gaussian or normal distribution;

- that the pollutant is released at a height above the ground that is given by the physical stack height and the rise of the plume due to its momentum and buoyancy (together forming the effective stack height);
- that the degree of dilution is inversely proportional to the wind speed;
- that pollutant material reaching the ground level is reflected back into the atmosphere;
- that the pollutant is conservative, i.e., not undergoing any chemical reactions, transformation or decay.

A Gaussian model is the solution of the basic equations for transport and diffusion in the atmosphere assuming stationarity in time and complete homogeneity in space. A Gaussian dispersion model is normally used for considering a point source such as a factory smoke stack. It attempts to compute the downwind concentration resulting from the point source. The origin of the coordinate system is placed at the base of the stack, with the x axis aligned in the downwind direction. The contaminated gas stream, which is normally called ”plume”, is shown rising from the smokestack and then leveling off to travel in the x direction and spreading in the y and z directions as it travels.

The plume normally rises to a considerable height above the stack because it is emitted at a temperature higher than that of ambient air and with a vertical velocity component. For Gaussian plume calculation, the plume is assumed to start from a point with coordinates (0,0,H), where H is called the effective stack height and is the sum of physical stack height (h) and the plume rise (dh).

It should be kept in mind that the Gaussian plume approach tries to calculate only the average values without making any statement about instantaneous values. The results obtained by Gaussian plume calculations should be considered only as averages over periods of at least 10 minutes, and preferably one-half to one hour. The Gaussian plume model so far allows one to estimate the concentration at a receptor point due to a single emission source for a specific meteorology. The spatial dynamics of pollution dispersion is described by the following type of equation in a Gaussian model:

$$\chi(x,y,z) = \frac{Q}{2\cdot\pi\cdot u\cdot\sigma_y\cdot\sigma_z}\cdot\left[exp\left\{\frac{y^2}{2\cdot\sigma_y^2}\right\}\cdot\left\{exp\left(\frac{(H-z)^2}{2\cdot\sigma_z^2}\right)+exp\left(\frac{(H+z)^2}{2\cdot\sigma_z^2}\right)\right\}\right] \tag{4.4}$$

In this equation $\sigma_y = a\cdot x^p$ is the standard deviation of the concentration distribution in the crosswind direction, in [m] at the downwind distance x; and $\sigma_z = b\cdot x^q$ is the standard deviation of the concentration distribution in the vertical direction, in [m], at the downwind distance x, a, b, p and q are constants depending on the stability of the atmosphere.

Where

$\chi(x,y,z)$: pollutant concentration at point (x,y,z);

u: wind speed (in the x ”downwind” direction, m/s)

σ: represents the standard deviation of the concentration in the x and y direction, i.e., in the wind direction and cross-wind, in meters;

Q: is the emission strength (g/s)

H: is the effective stack height, see below.

From the above equation, the concentration at any point (x,y,z) in the model domain, from a constant emission rate source, in steady state can be calculated.

Gaussian plume models are also applied to estimate multi-source urban concentrations. The procedure is to estimate the concentration at various locations for each of the point, area and line sources in the city for each meteorological condition and then sum up over all sources, all wind directions, all wind speeds, and all stability classes, weighted by the frequency of their occurrence.

4.2.2 Statistical Models

Statistical models are techniques based fundamentally on statistical analysis of calculated atmospheric concentrations. The statistical models derive the concentrations of pollutants from the meteorological and traffic parameters using statistical methods after an effective statistical relationship has been empirically established from the measured concentrations. Such models are not deterministic in the sense that they do not create or simulate a cause-effect, physical relationship between emissions and atmospheric concentrations. Regression, multiple regression and time-series techniques are some of the main approaches for statistical modelling.

In general, the statistical models are best applicable for short-term forecasts of concentrations on a local scale. Their simplicity forms an alternative to deterministic models.

Multiple Linear Regression (MLR)

Multiple regression is an extension of simple linear regression. It is used when we want to predict the value of a variable based on the value of two or more other variables. The variable we want to predict is called the dependent variable (or sometimes, the outcome, target or criterion variable). The variables we are using to predict the value of the dependent variable are called the independent variables (or sometimes, the predictor, explanatory or regressor variables).

For example, you could use multiple regression to understand whether exam performance can be predicted based on revision time, test anxiety, lecture attendance and gender. Alternately, you could use multiple regression to understand whether daily cigarette consumption can be predicted based on smoking duration, age when started smoking, smoker type, income and gender.

Multiple regression also allows you to determine the overall fit (variance explained) of the model and the relative contribution of each of the predictors to the total variance explained. For example, you might want to know how much of the variation in exam performance can be explained by revision time, test anxiety, lecture attendance and gender ”as a whole”, but also the ”relative contribution” of each independent variable in explaining the variance.

Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of multiple linear regression (MLR) is to model the linear relationship between the explanatory (independent) variables and response (dependent) variable.

In essence, multiple regression is the extension of ordinary least-squares (OLS) regression that involves more than one explanatory variable.

The Formula for Multiple Linear Regression Is:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \ldots + \beta_n x_{in} + \epsilon \tag{4.5}$$

where, for $i = n$ observations:

y_i = dependent variable

x_{i1} = explanatory variable

β_0 = y-intercept (constant term)

β_1 = slope coefficients for each explanatory variable

ϵ = the model's error term(also knows as residuals)

Explaining Multiple Linear Regression

A simple linear regression is a function that allows an analyst or statistician to make predictions about one variable based on the information that is known about another variable. Linear regression can only be used when one has two continuous variables — an independent variable and a dependent variable. The independent variable is the parameter that is used to calculate the dependent variable or outcome. A multiple regression model extends to several explanatory variables.

The multiple regression model is based on the following assumptions:

- There is a linear relationship between the dependent variables and the independent variables.
- The independent variables are not too highly correlated with each other.
- y_i observations are selected independently and randomly from the population.

- Residuals should be normally distributed with a mean of 0 and variance σ.

The coefficient of determination (R-squared) is a statistical metric that is used to measure how much of the variation in outcome can be explained by the variation in the independent variables. R^2 always increases as more predictors are added to the MLR model even though the predictors may not be related to the outcome variable.

R^2 by itself can't thus be used to identify which predictors should be included in a model and which should be excluded. R^2 can only be between 0 and 1, where 0 indicates that the outcome cannot be predicted by any of the independent variables and 1 indicates that the outcome can be predicted without error from the independent variables. When interpreting the results of a multiple regression, beta coefficients are valid while holding all other variables constant (”all else equal”). The output from a multiple regression can be displayed horizontally as an equation, or vertically in table form.

Time Series Techniques

In the last decade, the time series analysis techniques have been widely used to describe the dispersion of air pollutants on a local scale. Especially the ARIMA and ARIMAX stochastic models have been adopted in many applications. In an ARIMA (Auto Regressive Integrated Moving Average) model, the concentrations at a certain instant are expressed as linear combinations of previous concentration values and random terms (noise), which are specified in a statistical sense (i.e., are properly described in terms of a random process). Thus, in ARIMA models the physical causes of phenomena (meteorological variables and emission rates of the sources) are not distinguished in the input. Such models represent a ”black-box” approach. All possible uncertainties of the model are taken into account by a ”noise” variable with assigned statistical properties.

4.2.3 Physical or Fluid Models

Physical models have a high predictive capacity for air quality. Such models use scaling methods to transform the measured concentrations into ambient concentrations of contaminants. In a physical model, a real-world procedure is replicated, using a

Model Name		**Limitations of the Model**
Deterministic Models	Gaussian Models Numerical Models	It has high cost in terms of time and storage data, Requires large amount of input data, Gaussian based plume models perform poorly when wind speeds are less than 1 m/s.
Statistical Models	Time Series Models Regression Models	Require long historical data sets and lack of physical interpretation, Regression modelling often underperforms when used to model non-linear systems, Statistical models are site specific.
Physical Models	Wind tunnel Simulation	Construction and operational costs are high, Simulation of real time air pollution is expensive

Table 4.1: Limitations Of Various Air Pollution Models

smaller scale physical experiment in a laboratory, simulating the essential features of the original processes being studied. Typical experimental instruments, such as wind tunnels or water tunnels are employed, in which the atmospheric flows, for which boundary layer wind tunnels (wind tunnel modelling) are used. This method of physical simulation conducted in the wind tunnel, in which atmospheric currents were modelled using air as a fluid medium, is often referred to as fluid simulation by numerous researchers.

CHAPTER 5

CASE STUDY

5.1 Data Collection

First of all the daily Air quality index (AQI) as a comprehensive assessment of air quality concentration of criteria pollutants namely Respirable Suspended Particulate Matter (RSPM), Sulphur dioxide (SO2), Nitrogen dioxide (NO2) and Suspended Particulate Matter (SPM) has been calculated at ITO (a busiest traffic intersection) for a period of seven years 2000-2006), monitored continuously by Central Pollution Control Board (CPCB). A method of US nvironmental Protection Agency (USEPA) has been used for estimating the AQI, in which the sub-index and breakpoint pollutant concentrations depend on Indian National Ambient Air Quality Standard (NAAQS). There are primarily two steps involved in formulating an AQI: first the formation of sub-indices of each pollutant, second the aggregation (breakpoints) of sub indices. The Breakpoint concentration of each pollutant, used in calculation of AQI, is based on Indian NAAQS and results of epidemiological studies indicating the risk of adverse health effects of specific pollutants. It has been noticed that different breakpoint concentrations and different air quality standards has been reported in literature (Environmental Protection Agency, 1999). In India, to reflect the status of air quality and its effects on human health, the range of index values has been designated as “Good (0-100)”, “Moderate (101-200)”, “Poor (201-300)”, “Very Poor (301-400)” and “Severe (401-500)” (Nagendra et al. (2007)) as shown in Table below.

5.1.1 Air Quality Index

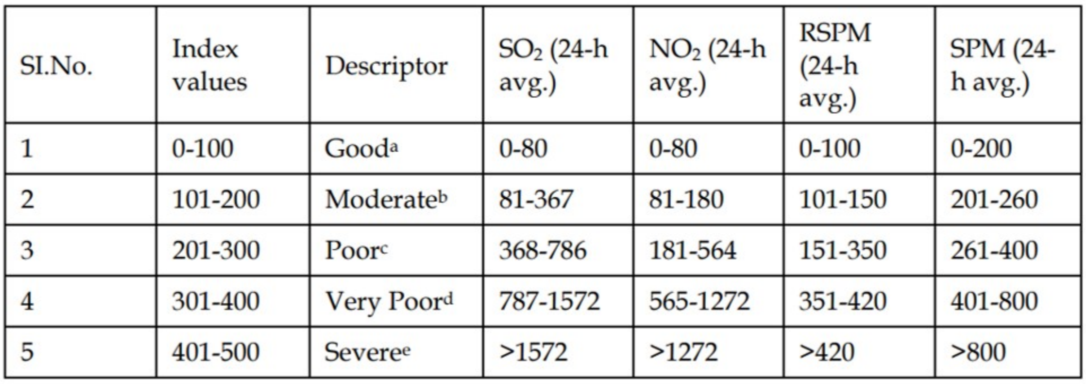
Air Quality Index (AQI) is an index that provides the public with the level of pollution associated with its health effects. The AQI focuses on the various health effects that people might experience based on the level and hours of exposure to the pollutant concentration. The AQI values are different from country to country based on the air quality standard of the country. The higher the AQI level greater is the risk of health related problems. The following table shows different classifications of AQI levels:

The AQI is calculated from the pollutant concentration data using the following formula:

$$I_p = \frac{I_{Hi} - I_{L0}}{BP_{Hi} - BP_{L0}}(C_p - BP_{L0}) + I_{L0} \tag{5.1}$$

I_p → Index of the Pollutant
I_{Hi} → AQI corresponding to BP_{Hi}
I_{L0} → AQI corresponding to BP_{L0}
C_p → Concentration of Pollutant P
BP_{Hi} → Breakpoint that is higher than or equal to C_p
BP_{L0} → Breakpoint that is lesser than or equal to C_p

Breakpoint is the concentration point which seperates each AQI classification level.



SI.No.	Index values	Descriptor	SO_2 (24-h avg.)	NO_2 (24-h avg.)	RSPM (24-h avg.)	SPM (24-h avg.)
1	0-100	Good[a]	0-80	0-80	0-100	0-200
2	101-200	Moderate[b]	81-367	81-180	101-150	201-260
3	201-300	Poor[c]	368-786	181-564	151-350	261-400
4	301-400	Very Poor[d]	787-1572	565-1272	351-420	401-800
5	401-500	Severe[e]	>1572	>1272	>420	>800

Figure 5.1: Propose sub-index and breakpoint pollutant concentration for Indian-AQI.

The overall AQI is now determined on the basis of the AQI for above pollutant ‘p’ and highest among them is declared as the overall AQI for that day.

The above estimated daily AQI along with meteorological variables like daily maximum temperature (tmax), minimum temperature (tmin), daily temperature range (difference between daily maximum and minimum temperature, trange), average temperature (tavg), wind speed (wsp), wind direction index (wdi), relative humidity (rh), vapor pressure (vp), station level pressure (slp), rainfall (rf), sunshine hours (ssh), cloud cover (cc), visibility (v) and radiation (rd), monitored at Safdarjung airport by Indian Meteorological Department (IMD), Delhi, have been used in statistical models to forecast daily AQI one day in advance. This study has been carried out for four different seasons namely summer (March, April, May), monsoon (June, July, August), post monsoon (September, October, November) and winter (December, January, February). The location ITO has been chosen in the present study due to various reasons: (i.) a busiest traffic intersection, (ii.) air pollutants concentration of RSPM, SO2, NO2 and SPM is monitored continuously and (iii.) the meteorological station Safdarjung airport is within 10 km radius. These models include previous day's AQI and meteorological variables as input and yield daily forecasting of AQI. The input and output is normalized between -1 to +1 using the minimum and maximum of the time series before any preprocessing. Forecast of daily AQI in Delhi has been obtained by Multiple Linear Regression Model.

5.2 AQI by Multiple Linear Regression Model

The following MLR equations for different seasons are resulted through training of AQI and meteorological data of 2000-2005 using statistical packages in R and python programming languages for summer, monsoon, post monsoon and winter respectively:

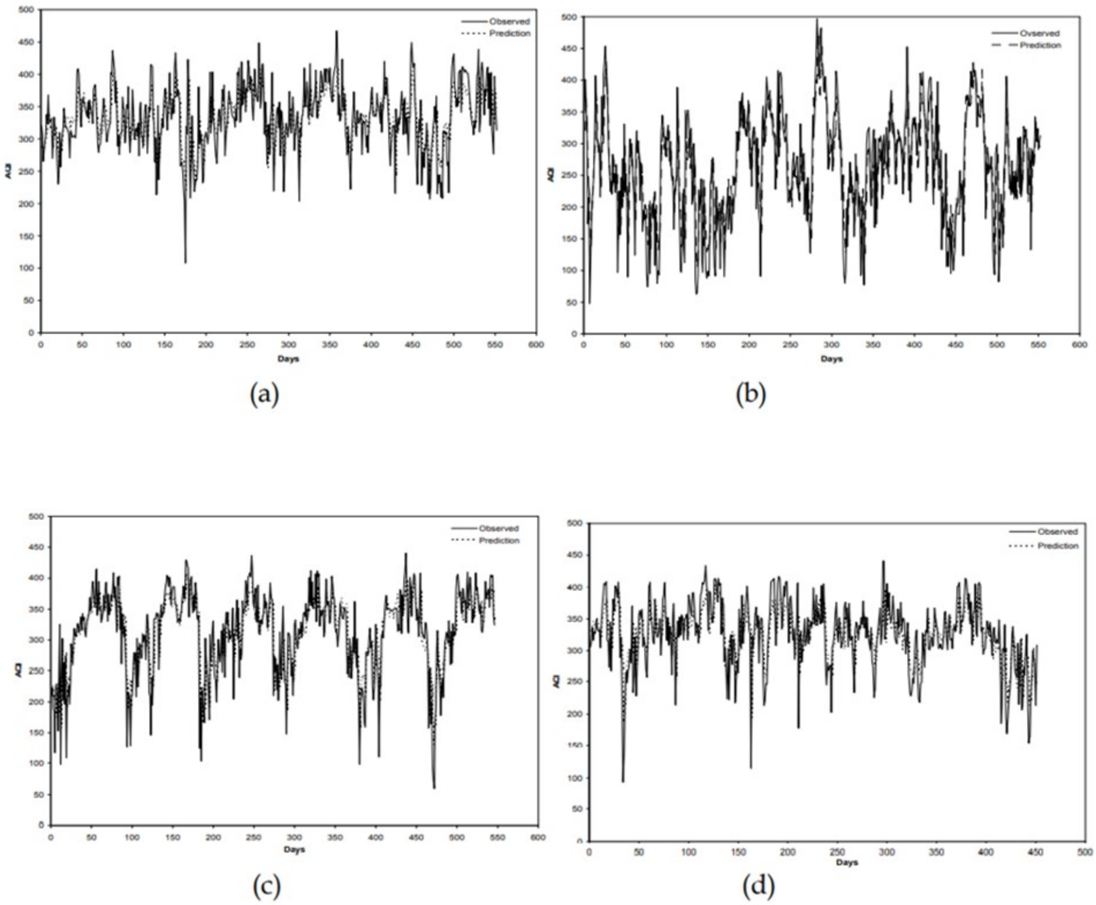
$$[AQI] = 0.0478 + 0.504[AQI_{d-1}] - 0.079[rh] + 0.126[t_{max}] - 0.068[cc] \tag{5.2}$$

$$[AQI] = 0.181 + 0.599[AQI_{d-1}] - 0.282[rh] - 0.128[v] - 0.155[t_{min}] \tag{5.3}$$

$$
\begin{gathered}
[AQI] = -0.324 + 0.537[AQI_{d-1}] + 0.573x[slp]- \\
0.112[vp] + 0.070[ssh]–0.135[v] + 0.066[t_{max}]
\end{gathered}
\tag{5.4}
$$

$$
\begin{gathered}
[AQI] = 0.171 + 0.503[AQI_{d-1}] - 0.191[v] - 0.115[cc]- \\
0.169[wsp] - 0.157[rh] + 0.151[rf]
\end{gathered}
\tag{5.5}
$$

The previous day's AQI is the common variable in all four equations. The above equations have also been used to forecast the daily AQI of 2000-2005 for all four seasons, which have been compared with the observed AQI, used as trained data of the corresponding seasons during 2000-2005 and are shown graphically below.



(a) (b)

(c) (d)

Figure 5.2: Comparison of observed and MLR model predicted values of daily AQI in (a) Summer, (b) Monsoon, (c) Post Monsoon and (d) Winter seasons during the years 2000-2005.

The above Figure reflects that the trained values of AQI are matching well with observed values. The same set of Eqs. (5.1-5.4) have been used for forecasting the daily AQI in all four seasons of the year 2006 which have been shown graphically in Figures below for summer, monsoon, post monsoon and winter seasons respectively. The observed values of AQI of year 2006 for each season have also been plotted in the Fig. 2 in order to validate the forecasted values of AQI.

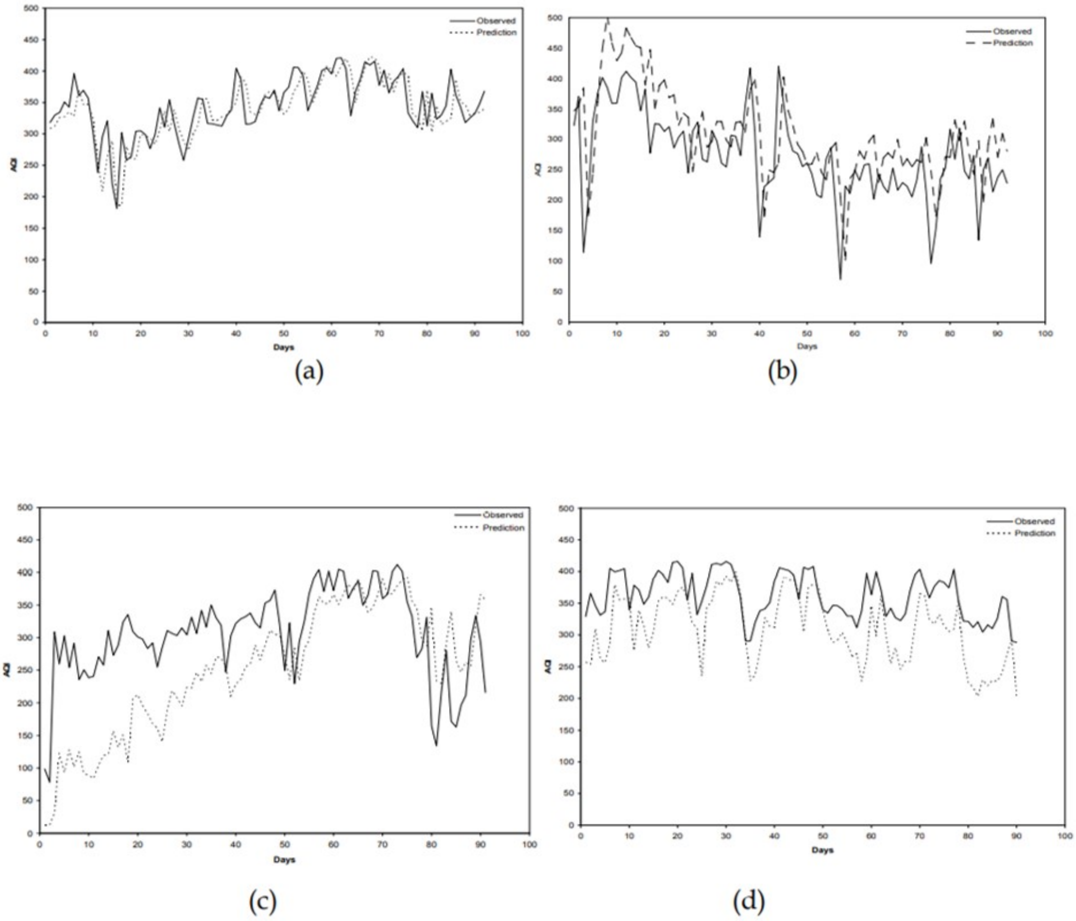
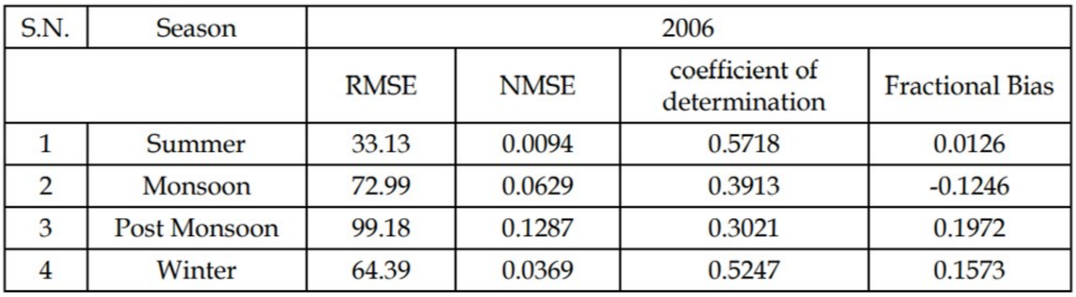


Figure 5.3: Comparison of observed and model's predicted values of daily AQI in (a) Summer, (b) Monsoon, (c) Post Monsoon and (d) Winter seasons during the years 2000-2005.

The quantitative analysis of comparison of forecasted and observed values of AQI has been made through statistical parameters. The NMSE and coefficient of determination (R2) are found as (0.0094, 0.5718) in summer season which are followed by (0.0369, 0.5247) in winter; (0.0629, 0.3913) in monsoon and (0.1287, 0.3021) in post monsoon seasons, showing good performance of the model as the ideal values of NMSE and

R2 are 0 and 1 respectively. The values of R2 in four different seasons reflect that the model's forecasted and observed AQI could be correlated explained by the selected input variables as approximately 57% in summer, 52% in winter, 39% in monsoon and 30% in post monsoon seasons. However, the fractional bias shows the under-prediction in all the seasons except Monsoon.



S.N.	Season	2006			
		RMSE	NMSE	coefficient of determination	Fractional Bias
1	Summer	33.13	0.0094	0.5718	0.0126
2	Monsoon	72.99	0.0629	0.3913	-0.1246
3	Post Monsoon	99.18	0.1287	0.3021	0.1972
4	Winter	64.39	0.0369	0.5247	0.1573

Figure 5.4: Comparison of MLR model predicted and observed values in years 2000-2005 and year 2006.

CHAPTER 6

CONCLUSION

In the present study, the statistical (linear and non-linear) and analytical dispersion models of air pollutants released from point, line and area sources are discussed. Air quality index has been forecasted using MLR and ANN models. Performance of both the modes has been compared and observed that ANN model is doing better than MLR. The analytical models are formulated by considering the wind speed as a power law profile of vertical height above the ground and vertical eddy diffusivity as an explicit function of downwind distance and vertical height in different boundary conditions. A case study of Delhi has been made for predict 24 hourly concentration of RSPM through the application models with Neumann boundary condition in the month of Jan 2008. The input parameters namely emission inventory and meteorological variables are pre-processed as per the requirement of the model. The different types of primary and secondary sources of RSPM due to vehicular, domestic, industries and power plants have been used in emission inventory for year 2008. Some traditional methods for determining the meteorological field are also discussed in this chapter, which are used as an input to these analytical models. The analytical models are evaluated with observed concentration at different locations in Delhi obtained from CPCB and NEERI, which show that the concentration levels obtained from the models are always high in comparison to the observed values and NAAQS. However, the models are performing satisfactory. Although the present models have the limitation as the longitudinal diffusion is neglected in comparison to the advection and are not considering the wind directions at different vertical heights. These models can be used for other Indian urban cities.