We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

5,200 Open access books available
128,000 International authors and editors
150M Downloads

154 Countries delivered to
TOP 1% Our authors are among the
most cited scientists
12.2% Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Mathematical Modeling of Air Pollutants: An Application to Indian Urban City

P. Goyal and Anikender Kumar
Centre for Atmospheric Sciences, Indian Institute of Technology Delhi
India

1. Introduction

Continuous development and increase of population in the urban areas, a series of problems related to environment such as deforestation, release of toxic materials, solid waste disposals, air pollution and many more, have attracted attention much greater than ever before. The problem of air pollution in cities has become so severe that there is a need for timely information about changes in the pollution level. The air pollution dispersion is a complex problem. It covers the pollutant transport and diffusion in the atmosphere. The pollutant dispersion in the atmosphere depends on pollutant features, meteorological, emission and terrain conditions. Physical and mathematical models are developed to describe the air pollution dispersion. Physical models are small scale representations of the atmospheric flow carried out in wind tunnels. Mathematical models are divided in to statistical and deterministic models. Statistical models are based on analysis of past monitoring air quality data. Deterministic models are based on a mathematical description of physical and chemical processes taking place in the atmosphere. These models are based on mathematical equations, express conservation laws of mass, momentum and energy. Both the models are discussed in this chapter.

Statistical models are also divided into linear and non-linear models. Several studies based on the statistical models have been carried out in different regions to identify local meteorological conditions, most strongly associated with air pollutants concentration to forecast the air quality (McCollister & Willson, 1975; Aron & Aron, 1978; Lin, 1982; Aron, 1984; Katsoulis, 1988; Robeson & Steyn, 1990). Many of the previous studies (Sanchez et. al., 1990; Mantis et al., 1992; Milionis & Davies, 1994) analyzed the meteorological conditions associated with high pollutant concentration. These studies usually produced qualitative or semi quantitative results and shed light on the relation between the meteorological conditions and pollutant concentrations. Shi & Harrison, 1997 developed a linear regression model for the prediction of NO and NO in London. A linear regression model was used by Cogliani, 2001 for air pollution forecast in cities by an air pollution index highly correlated with meteorological variables. Since the relation between air pollutants and meteorological variables is not linear, some non-linear models i.e., Neural Network can also be used to forecast the pollutant concentrations (Bozner et al., 1993; Comrie, 1997).

The deterministic models are divided in to Eulerian, Lagrangian and Gaussian models. In this chapter we discuss only the Eulerian analytical models. The atmospheric diffusion
equation (Seinfeld, 1986) has long been used to describe the dispersion of airborne pollutants in a turbulent atmosphere. The use of analytical solutions of this equation was the first and remains the convenient way for modeling the air pollution problems (Demuth, 1978). Air dispersion models based on analytical solutions possess several advantages over numerical models as all the influencing parameters are explicitly expressed in a mathematically closed form. Analytical models are also useful in examining the accuracy and performance of numerical models. In practice, most of the estimates of dispersion are based on the Gaussian plume model, which assumes the constant wind speed and turbulent eddies with height. Hinrichsen (1986) compared a non-Gaussian model, in which wind speed and turbulence, are not constant with height and observed that non-Gaussian model agreed better than Gaussian model with the observed data.

Several efforts have also been made for the development of non-Gaussian models of point and line sources, since observational studies show that the wind speed and eddy diffusivity vary with vertical height above the ground (Stull, 1988). Analytical solutions of the advection diffusion equation, with wind speed and vertical eddy diffusivity both as power function of vertical height, bounded by Atmospheric Boundary Layer (ABL) are well known for point and line sources (Seinfeld, 1986; Lin & Hildemann, 1996). Taylor’s (1921) analysis and statistical theory suggest that the eddy diffusivity depends on the downwind distance from the source (Arya, 1995). The advection diffusion equation has also solved analytically with wind speed as function of height and eddy diffusivity as a function of downwind distance from the source (Sharan & Modani, 2006). Thus in general, the eddy diffusivity should be a function of both vertical as well as downwind distance (Mooney & Wilson, 1993). Recently (Sharan & Kumar, 2009) formulate the advection diffusion equation considering the wind speed as a function of vertical height and eddy diffusivity as a function of both vertical height and downwind distance applicable only for point source release in reflecting boundary condition. However, Dirichlet (total absorption), Neumann (total reflection) and mixed boundary conditions are also appropriate for calculating the actual ground-level concentration of air pollutants. In addition to these, the few studies have been made for analytical solution of the advection diffusion equation for area sources. Park & Baik (2008) have solved the advection diffusion equation analytically for finite area source with wind speed and vertical eddy diffusivity as power function of vertical height in unbounded region.

The objective of this chapter is to formulate and use the statistical (linear and non-linear) and Eulerian analytical models for prediction/forecast of air pollutants released from point, line and area sources. The analytical models are developed by using four different set of boundary conditions. The model with reflecting boundary condition is used for urban city Delhi, the capital of India and is validated by the observed values of concentration of Respirable Suspended Particulate Matter (RSPM).

2. Mathematical models

2.1 Statistical models

The main role of statistical models is to analyze past monitored air quality data. They are divided into linear and non-linear models. Linear Models as Multiple Linear Regression (MLR) can be used to make a linear empirical relationship between air pollutants and meteorological variables. The methodology of MLR is explained briefly in next section.
2.1.1 Multiple Linear Regression (MLR) models
A forecast/prediction of air pollutants can be made through regression equation in which unknown variable can be expressed as a function of certain number of known variables. There is one dependent variable to be predicted in relation to the two or more independent variables. The general form of MLR can be expressed as

\[ Y = b_1 + b_2 X_2 + \ldots + b_k X_k + e \]  

where Y is dependent variable, X_2, X_3, \ldots, X_k are independent variables, b_1, b_2, \ldots, b_k are linear regression parameters and e is an estimated error term, which is obtained from independent random sampling from the normal distribution with mean zero and constant variance. The purpose of regression modeling is to estimate the b_1, b_2, \ldots, b_k which can be made using minimum square error technique.

The (Eq. (1)) can also be written as

\[ Y = X b + e \]  

where Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, X = \begin{bmatrix} 1 & X_{21} & X_{31} & \ldots & X_{k1} \\ 1 & X_{22} & X_{32} & \ldots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & X_{3n} & \ldots & X_{kn} \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} \text{ and } e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} 

Here Y is an n x 1, X is an n x k, b is a k x 1 and e is an n x 1 matrix. The solution of above equation can be obtained as \( b = (X'X)^{-1} (X'Y) \) using minimum square error technique. Further the F-test has been performed to determine whether a relationship exists between the dependent variables and regressors. The t-test is also performed in order to determine the potential value of each of the regressors’ variables in the regression model. The resulting model can be used to predict future values.

If the relationship between air pollutants and meteorological variables is not linear, non-linear models like Artificial Neural Network can be used for treating the non-linear relationship. The model’s characteristics are described in next section.

2.1.2 Artificial Neural Network
The Artificial Neural Network (ANN) represents an alternative methodology to conventional statistical modeling because of their computational efficiency and generalization ability. ANN models are mathematical models inspired by the biological neurons. The use of ANN as mentioned in the literature is an effective alternative to more traditional statistical techniques for forecasting time series. ANN can be trained to approximate virtually any smooth, measurable and highly nonlinear functions between input and output and requires no prior knowledge to the nature of this relationship (Gardner & Dorling, 1998) and can also be trained to generalize, when presented with new and unseen data. ANNs are made up of interconnected processing elements called neurons or nodes that are arranged in the layers. These layers include an input layer, one or more hidden layers and an output layer which are connected to each neuron of the next layer by the weights. The number of hidden layer is selected based upon the problem complexity. The number of neuron in the input and output layer is problem specific. The information
transfer is allowed only to the next consecutive layer. Each node of the hidden layer receives incoming signals from the nodes of the input layer. Each input value is weighted and based on its relative importance before entering the hidden layer. The total input signal \( NET \) is calculated as

\[
NET = \sum_i W_i X_i
\]  

(3)

The total incoming signal is then passed through a non-linear transfer function \( F \) to produce the outgoing signal \( F(NET) \) of the node.

\[
F(NET) = \frac{1}{1 + e^{-NET}}
\]  

(4)

The output signal of a hidden node is finally passed to the nodes of the next layer (hidden or output), where a similar procedure takes place. There are several transfer functions available such as pure linear, hyperbolic tangent, sigmoid etc. Transfer function plays a key role in training process of neural network because the ANN produce different results sensitive to its transfer function (Wassermann, 1989). The process of optimizing the connection weights is known as training or learning of ANN. This is equivalent to the parameter estimation phase in the conventional statistical models. Iterative techniques are used to get the best values for connection weights by minimizing the performance function i.e. error between model output and the provided target values. The trained network is then used for the forecasting/prediction purpose. ANN has the capability to recognize the patterns in the time series data presented to it and is thus useful in many types of pattern recognition problems.

2.2 Deterministic models

Deterministic models are based on a mathematical description of physical and chemical processes taking place in the atmosphere. These models are divided into different categories on the basis of source characteristics as point, line and area sources or on the basis of topography of the region as flat or complex terrain. These models can also be classified on the basis of size of the field they are describing:

- Short distance (distance from source less than 30-50 km)
- Mesoscale models (concentration fields of the order of hundreds of kms)
- Continental or planetary circulation models

Finally, models can also be classified on the basis of the time resolution of the concentration produced:

- Episodic models (temporal resolution of less than an hour)
- Short-time models (temporal resolutions greater than or equal to an hour and less than or equal to 24h)
- Climatologically models (with resolution greater than 24h, generally seasonal or annual) (Tirabassi, 2010).

These models are divided into three categories according to different approaches as Eulerian, Lagrangian and Gaussian. Eulerian approach is based on a fixed spatial-temporal grid. The basic equation used in Eulerian air pollution dispersion models is derived from the equation of the pollutant molecular diffusion:
\[
\frac{\partial C}{\partial t} = - \mathbf{U} \cdot \nabla C + D \nabla^2 C + S
\]  
(5)

where: \( C \) - concentration of pollutant in the atmospheric air, 
\( \mathbf{U} \) - wind speed vector of the components \( u, v, w \), 
\( D \) - Molecular diffusion coefficient, 
\( S \) - Represents the sources and sinks of pollutant in the atmosphere i.e. its emission, removal from the atmosphere by dry and wet deposition, chemical reactions, 
\( \nabla \) - Gradient operator, 
\( \nabla^2 \) - Laplacian.

Eq. (5) has been modified for the turbulent flow in the atmosphere. Modifications include the averaging procedure and closure procedure. The wind speed is expressed as the sum of two components, mean and turbulent: \( \mathbf{U} = \bar{\mathbf{U}} + \mathbf{U}' \) and the same can be made for \( C \) as \( C = \bar{C} + C' \) in (Eq. (5)) and hypothesizing a wind with divergence nil:

\[
\frac{\partial \bar{C}}{\partial t} = - \bar{\mathbf{U}} \cdot \nabla \bar{C} + D \nabla^2 \bar{C} + \bar{S}
\]  
(6)

where \( \bar{\mathbf{U}} \) is turbulent concentration flux.

The simplest closure method of (Eq. (6)) is a local first order closure in which \( K \)-theory is used. Assuming that turbulent concentration flux is proportional to the gradient of the average concentration, the following relation is obtained as:

\[
\bar{\mathbf{U}} = -K \nabla \bar{C}
\]  
(7)

where \( K \) (3x3) is turbulent diffusion coefficient.

When \( K \) tensor is diagonal, molecular diffusion is negligible and \( C(x, y, z, t) \) represents the concentration of a non-reactive pollutant (\( \bar{S} = S \)). Thus, the (Eq. (6)) can be written as

\[
\frac{\partial C}{\partial t} = - \bar{\mathbf{U}} \cdot \nabla \bar{C} + \nabla \cdot K \nabla \bar{C} + S
\]  
(8)

Eq. (8) can be solved for \( C \) analytically or numerically if input data for \( \mathbf{U} \), \( K \) and \( S \) are provided with initial and boundary conditions. Exact solution can be obtained by analytical methods, while numerical methods give the only approximate solutions. In this section we will discuss only the Eulerian analytical models.

2.2.1 Eulerian analytical solutions of the advection diffusion equation for point source in different atmospheric boundary conditions

Analytical solutions of the above equation are of fundamental importance in understanding and describing physical phenomena. The deterministic models for the dispersion of pollutants in atmosphere, based on the advection diffusion equation and \( K \)-theory as (Eq. (8)), can be written as:

\[
\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} + v \frac{\partial C}{\partial y} + w \frac{\partial C}{\partial z} = \frac{\partial}{\partial x} \left( K_x \frac{\partial C}{\partial x} \right) + \frac{\partial}{\partial y} \left( K_y \frac{\partial C}{\partial y} \right) + \frac{\partial}{\partial z} \left( K_z \frac{\partial C}{\partial z} \right) + S ,
\]  
(9)

www.intechopen.com
where $K_x$, $K_y$, and $K_z$ are the eddy diffusivities along $x$, $y$ and $z$ directions respectively. The following assumptions are made in solution of Eq. (9):

a. Steady-state condition is considered (i.e., $\frac{\partial C}{\partial t} = 0$).

b. The vertical velocity component ($w$) is neglected in comparison to horizontal velocity components ($u$ and $v$).

c. $x$-axis is oriented in the direction of mean wind (i.e., $u=U$, $v=0$).

d. Downwind diffusion is neglected in comparison to transport due to mean wind

$$\text{i.e.,} \left| U \left( \frac{\partial C}{\partial x} \right) \right| \gg \left| \left( \frac{\partial}{\partial x} \right) \left( K_x \left( \frac{\partial C}{\partial x} \right) \right) \right|. $$

Applications of these assumptions in (Eq. (9)), leads the steady state advection-diffusion equation for dispersion of a non reactive contaminate released from continuous source as (Seinfeld, 1986):

$$U \frac{\partial C}{\partial x} = \frac{\partial}{\partial y} \left( K_y \frac{\partial C}{\partial y} \right) + \frac{\partial}{\partial z} \left( K_z \frac{\partial C}{\partial z} \right),$$

where $x$, $y$, and $z$ are coordinates in the along-wind, cross wind and vertical directions respectively. $C$ is the mean concentration of pollutants and $U$ is the mean wind speed in downwind direction. $K_y$ and $K_z$ are eddy diffusivities of pollutants in the crosswind and vertical directions respectively.

Eq.(10) is solved with the following boundary conditions, in which, $h$ is the top of the inversion/mixed layer:

Dirchlet Boundary Condition (total absorption)

$$C(x,y,z)=0 \ \text{at } z = 0$$

$$C(x,y,z)=0 \ \text{at } z = h \quad (11. a)$$

Neumann Boundary Condition (total reflection)

$$K_y \frac{\partial C(x,y,z)}{\partial z} = 0 \ \text{at } z = 0$$

$$K_z \frac{\partial C(x,y,z)}{\partial z} = 0 \ \text{at } z = h \quad (11. b)$$

Mixed (type-I) Boundary Condition

$$K_y \frac{\partial C(x,y,z)}{\partial z} = 0 \ \text{at } z = 0$$

$$C(x,y,z)=0 \ \text{at } z = h \quad (11. c)$$

Mixed (type-II) Boundary Condition

$$C(x,y,z)=0 \ \text{at } z = 0$$
The pollutant decays in cross wind direction:

\[ C(x, y, z) \rightarrow 0 \quad \text{as} \quad y \rightarrow \pm \infty \]  
(12)

The pollutant is released from an elevated point source of strength \( Q_p \) located at the point \((0, y_s, z_s)\),

\[ U \cdot C(0, y, z) = Q_p \delta(y - y_s) \delta(z - z_s), \]  
(13)

where \( \delta \) is the Dirac delta function.

The transport of contaminant emitted from a source primarily depends on the wind speed \( U \). The formulations of the commonly used dispersion models in air quality studies assume wind speed to be constant. However, it is well known that wind speed increases with height in the lower part of the atmospheric boundary layer (Arya, 1999). The height dependent wind speed can be expressed as

\[ U(z) = \hat{a} z \]  
(14)

where \( U(z_r) \) is the wind speed at reference height \( z_r \) and \( \hat{a} \) depends on atmospheric stability.

In formulation of dispersion models, \( K_z \) is parameterized as a function of height \( z \) only (Lin & Hildemann, 1996; Park & Baik, 2008). However, based on the Taylor’s analysis and statistical theory, it is revealed that the eddy diffusivity depends on the downwind distance \( x \) (Arya, 1995). Thus, the eddy diffusivity can be a function of \( x \) and \( z \) both (Mooney & Wilson, 1993; Sharan & Kumar, 2009). The modified form of \( K_z(x, z) \) is given as:

\[ K_z(x, z) = K'_z(z)f(x), \]  
(15)

where \( K'_z(z) \) is the form of eddy diffusivity depending on \( z \) and \( f(x) \) is assumed to be function of \( x \). \( K'_z(z) \) is parameterized as a power law profile in \( z \):

\[ K'_z(z) = b z^\beta, \quad b = K'_z(z_r) z_r^{-\beta}, \]  
(16)

where \( K'_z(z_r) \) is the value of \( K'_z(z) \) at height \( z = z_r \) and \( \beta \) depends on atmospheric stability.

Using Taylor’s hypothesis, the lateral eddy diffusivity can be represented by (Huang, 1979; Brown et al., 1997)

\[ K_y(x,z) = \frac{1}{2} U(x) \frac{\sigma_y^2(x)}{dx}, \]  
(17)

where \( \sigma_y \) is the standard deviation of concentration distribution in the crosswind direction.

Based on the analysis of Prairie Grass and some other historical tracer experiments of atmospheric dispersion, Irwin et al., 2007 concluded that the ground level crosswind concentration profile of dispersing plume on average is well characterized as having a Gaussian shape, which is well predicted by all atmospheric transport and diffusion models, regardless of their sophistication. Thus, by assuming the Gaussian concentration
distribution in crosswind direction (Huang, 1979; Irwin et al., 2007), the steady state concentration of a pollutant released from point source in a three dimensional domain can be described as

\[
C(x, y, z) = C(x, z) \frac{\exp\left(-y^2/2\sigma_y^2\right)}{\sqrt{2\pi\sigma_y}}, \tag{18}
\]

where \(C(x, z)\) is the crosswind integrated concentration.

The mathematical formulation of \(C(x,z)\) is obtained by substituting the wind profile (Eq. (14)), diffusivity profile (Eq. (15)) and boundary conditions (Eqs. (11-13)) in the advection-diffusion equation (Eq. (10)) as:

\[
\frac{1}{f(x)} \frac{\partial C}{\partial x} = \frac{b}{a} \frac{\partial}{\partial z} \left( \frac{g^0}{2} \frac{\partial C}{\partial z} \right), \tag{19}
\]

provided \(f(x) \neq 0, \forall x \in (0, \infty)\).

By using the separation of variables technique, the solution of the (Eq. (19)) is assumed in the form:

\[
C(x,z)=X(x) Z(z), \tag{20}
\]

which transforms the (Eq. (19)) into two following ordinary differential equations by taking -\(\lambda^2\) as the separation constant:

\[
\frac{d}{dx} \left(f(x) \lambda^2 X \right) = 0 \tag{21}
\]

\[
\frac{d}{dz} \left(x^2 \frac{dZ}{dz} \right) + \lambda^2 \left( \frac{a}{b} \right) Z = 0 \tag{22}
\]

Eq. (21) has the solution:

\[
X(x) = A \exp\left(-\lambda^2 \int_0^x f(s) ds \right), \tag{23}
\]

where \(A\) is an arbitrary constant.

The solution of Eq. (22) is obtained in different boundary conditions as follows:

i. \(Z = 0\) at \(z = 0, h\)

\[
Z = 0 \tag{24}
\]

represents a Sturm-Liouville eigen value problem.

The solution of (Eq. (22)) with boundary condition (Eq. (24)) is zero for \(\lambda = 0\).

For a non-zero value of \(\lambda\), the transformation of the variables

\[
t = Z \left( \frac{a-b+2}{2} \right) \tag{25}
\]

and
\[ Z(z) = z^{\frac{1-\beta}{2}} G(t) \]  

(26)

in (Eq. (22)) leads to:

\[ t^2 \frac{d^2G}{dt^2} + \frac{dG}{dt} + (k^2t^2 - \mu^2)G = 0, \]

(27)

where

\[ k = \frac{\sqrt{a/b\lambda}}{(\alpha - \beta + 2)/2} \quad \text{and} \quad \mu = \frac{(1-\beta)}{(\alpha - \beta + 2)} \]

(28)

Eq. (27) is the Bessel’s equation, whose solution can be given as:

\[ G(t) = B_1 J_\mu(kt) + B_2 J_{-\mu}(kt), \]

(29)

where \( J_\mu \) and \( J_{-\mu} \) are the Bessel’s functions of first kind of order \( \mu \) and \(-\mu\) respectively. From Eqs. (25), (26) and (29), we have

\[ Z(z) = z^{\frac{1-\beta}{2}} \left\{ B_1 J_\mu \left( k z^{\frac{\alpha - \beta + 2}{2}} \right) + B_2 J_{-\mu} \left( k z^{\frac{\alpha - \beta + 2}{2}} \right) \right\} \]

(30)

Application of the boundary condition (Eq. (24)) at \( z = 0 \) in (Eq. (30)) yields \( B_2 = 0 \) and the condition at \( z = h \) (Eq. (24)) gives rise:

\[ J_\mu \left( k \frac{h}{2} \right) = 0 \]

(31)

The corresponding eigen functions are:

\[ Z_n(z) = z^{\frac{1-\beta}{2}} \left\{ J_\mu \left( k_n z^{\frac{\alpha - \beta + 2}{2}} \right) \right\}, \quad n = 1, 2, 3 \]

(32)

The general solution of (Eq. (19)) is obtained by using (Eqs. (23), (31) and (32)) as:

\[ C(x, z) = z^{\frac{1-\beta}{2}} \sum_{n=1}^{\infty} A_n J_\mu \left( k_n z^{(\alpha - \beta + 2)/2} \right) \exp \left\{ -b(\alpha - \beta + 2)^2 k_n^2 \int f(s) ds \right\}, \]

(33)

where \( A_1, A_2 \ldots A_n \) are the unknown coefficients. The (Eq. (33)) represents the concentration distribution \( C \) through the Fourier-Bessel series (Abramowitz & Stegun, 1972) corresponding to a set of eigen functions \( Z_n \).

Estimation of the coefficients \( A_n \)'s for crosswind integrated concentration:

The source at \( x = 0 \), Eq. (13) gives:

\[ a z^\alpha \left[ \sum_{n=1}^{\infty} A_n J_\mu \left( k_n z^{(\alpha - \beta + 2)/2} \right) \right] = Q_p \delta(z - z_s) \]

(34)
The coefficients $A_n$’s are estimated using the orthogonal property of eigen functions (Abramowitz & Stegun, 1972).

Multiplying (Eq. (34)) by $z^{(1-\beta)/2} J_\mu(k_m z^{(\alpha-\beta+2)/2})$ for $m \geq 1$, and integrating it with respect to $z$ from $0$ to $h$, we get:

$$A_n = Q_p \frac{\alpha - \beta + 2}{\alpha h^{\alpha+\beta+2} z^{(1-\beta)/2}} \frac{1}{J_{\mu+1}(k_m h^{(\alpha-\beta+2)/2})} \int_0^h J_\mu(k_m z^{(\alpha-\beta+2)/2}) dz$$

(35)

Substituting the expression for $A_n$’s for $n \geq 1$ in (Eq. (33)), the final solution is obtained as:

$$C(x,z) = Q_p \left[ \frac{\alpha - \beta + 2}{\alpha h^{\alpha+\beta+2} z^{(1-\beta)/2}} \sum_{n=1}^{\infty} J_\mu \left( \frac{Y_n(z/h)^{(\alpha-\beta+2)/2}}{J_{\mu+1}(Y_n)} \right) \int_0^h J_\mu(Y_n(z/h)^{(\alpha-\beta+2)/2}) \right] \times \exp \left( \frac{-b(\alpha - \beta + 2)^2 Y_n^2}{4 ah^{\alpha+\beta+2} f(s)ds} \right)$$

(36)

in which $Y_n$ is given as:

$$J_\mu(Y_n) = 0.$$  

(37)

ii. The solution of Eq. (22) for crosswind integrated concentration by substituting the wind profile (Eq. (14)) and diffusivity profile (Eq. (15)) in (Eq. (10)) for Neumann boundary condition (Eq. (11.b)) is obtained from earlier study by Sharan & Kumar (2009) as:

$$C(x,z) = Q_p \left[ \frac{\alpha + 1}{\alpha h^{\alpha+\beta+2}} + \frac{\alpha - \beta + 2}{\alpha h^{\alpha+\beta+2} z^{(1-\beta)/2}} \sum_{n=1}^{\infty} J_\mu \left( \frac{Y_n(z/h)^{(\alpha-\beta+2)/2}}{J_{\mu+1}(Y_n)} \right) \right] \times \exp \left( \frac{-b(\alpha - \beta + 2)^2 Y_n^2}{4 ah^{\alpha+\beta+2} f(s)ds} \right)$$

(38)

in which $Y_n$ is given as:

$$J_{\mu+1}(Y_n) = 0.$$  

(39)

iii. The (Eq. (22)), along with the following boundary conditions corresponding to (Eq. (11.c)):

$$b z^{(\beta)/2} \frac{dZ}{dz} = 0 \text{ at } z = 0 \quad \text{and} \quad Z = 0 \text{ at } z = h$$

(40)

The solution of (Eq. (22)) with boundary condition (Eq. (40)) is zero for $\lambda = 0$.

For a non-zero value of $\lambda$, the boundary condition (Eq. (40)) at $z = 0$ in (Eq. (30)) yields $B_1 = 0$ and the condition at $z = h$ (Eq. (40)) gives rise:
The corresponding eigen functions are:

\[ Z_n(z) = z^{\frac{1-\beta}{2}} J_{-\mu} \left( k_n z^{(\alpha-\beta+2)/2} \right), \quad n = 1, 2, 3, \quad \beta > 1 \]  

The general solution of (Eq. (19)) is obtained by using (Eqs. (23), (41) and (42)) as:

\[ C(x,z) = z^{\frac{1-\beta}{2}} \sum_{n=1}^{\infty} A_n J_{-\mu} \left( k_n z^{(\alpha-\beta+2)/2} \right) \exp \left\{ -\frac{b(\alpha-\beta+2)^2 k_n^2}{4a} \int_0^h f(s) \, ds \right\}, \quad \beta > 1 \]  

where \( A_1, A_2, \ldots \) are the unknown coefficients.

The source condition at \( x = 0 \) (Eq. (13)), gives:

\[ az^{\frac{1-\beta}{2}} \sum_{n=1}^{\infty} A_n J_{-\mu} \left( k_n z^{(\alpha-\beta+2)/2} \right) = Q_p \delta(z - z_a) \]  

The coefficients \( A_n \)'s are estimated using the orthogonal property of eigen functions (Abramowitz & Stegun, 1972).

Multiplying (Eq. (44)) by \( z^{(1-\beta)/2} J_{-\mu} \left( k_m z^{(\alpha-\beta+2)/2} \right) \), \( m \geq 1 \), and integrating it with respect to \( z \) from 0 to \( h \), we get:

\[ A_n = Q_p \frac{\alpha-\beta+2}{ah^{\alpha-\beta+2}} \frac{1-\beta}{2} \left( \frac{z_a}{k_n} \right)^{\frac{(\alpha-\beta)+2}{2}} J_{-\mu} \left( k_n z_a^{(\alpha-\beta+2)/2} \right) J_{-\mu+1} \left( k_n z_a^{(\alpha-\beta+2)/2} \right) \]  

Substituting the expression for \( A_n \)'s, \( n \geq 1 \) into (Eq. (45)), the final solution of an elevated point source at crosswind is obtained as:

\[ C(x,z) = Q_p \frac{\alpha-\beta+2}{ah^{\alpha-\beta+2}} \left( \frac{z}{k_n} \right)^{\frac{(\alpha-\beta)+2}{2}} \sum_{n=1}^{\infty} J_{-\mu} \left( Y_n \right) Y_n(z/h)^{\frac{(\alpha-\beta)+2}{2}} J_{-\mu} \left( Y_n (z/h)^{\frac{(\alpha-\beta)+2}{2}} \right) \]  

\[ \exp \left\{ -\frac{b(\alpha-\beta+2)^2}{4a h^{\alpha-\beta+2}} \int_0^h f(s) \, ds \right\}, \quad \beta > 1 \]

in which \( Y_n \) is given as:

\[ J_{-\mu}(Y_n) = 0. \]
iv. The (Eq. (22)), along with the following boundary conditions corresponding to (Eq. (11.d)):

\[
Z = 0 \text{ at } z = 0 \text{ and } b z^\beta \frac{dZ}{dz} = 0 \text{ at } z = h \tag{48}
\]

The solution of (Eq. (22)) with (Eq. (48)) is zero for \( \lambda = 0 \).

For a non-zero value of \( \lambda \), the boundary condition (Eq. (48)) at \( z = 0 \) in (Eq. (30)) yields \( B_2 = 0 \) and the condition at \( z = h \) (Eq. (48)) gives rise:

\[ J_{\mu - 1} \left( \frac{a - \beta + 2}{2} \right) \Big|_{h} = 0 \tag{49} \]

The corresponding eigen functions are:

\[ Z_n(z) = z^\frac{\alpha - \beta}{2} J_{\mu} \left( k_n z^{(\alpha - \beta + 2)/2} \right), \quad n = 1, 2, 3, \tag{50} \]

The general solution of (Eq. (19)) is obtained by using (Eqs. (22), (49) and (50)) as:

\[
C(x,z) = z^\frac{1-\beta}{2} \sum_{n=1}^{\infty} A_n J_{\mu} \left( k_n z^{(\alpha - \beta + 2)/2} \right) \exp \left\{ - \frac{b(a - \beta + 2)^2 k_n^2}{4a} \int_0^h f(s) \, ds \right\}, \tag{51} \]

where \( A_1, A_2, \ldots \) are the unknown coefficients. The Eq. (51) represents the concentration distribution \( C \) as the Fourier-Bessel series (Abramowitz & Stegun, 1972) corresponding to a set of eigen functions \( Z_n \).

Estimation of the coefficients \( A_n \)'s for crosswind integrated concentration

The source condition at \( x = 0 \) (Eq. (13)), gives:

\[
\alpha z^a \left[ z^\frac{1-\beta}{2} \sum_{n=1}^{\infty} A_n J_{\mu} \left( k_n z^{(\alpha - \beta + 2)/2} \right) \right] = Q_p \delta(z - z_a) \tag{52} \]

The coefficients \( A_n \)'s are estimated using the orthogonal property of eigen functions (Abramowitz & Stegun, 1972).

Multiplying (Eq. (52)) by \( z^{(1-\beta)/2} J_{\mu} \left( k_n z^{(\alpha - \beta + 2)/2} \right) \), \( m \geq 1 \), and integrating it with respect to \( z \) from 0 to \( h \), we get:

\[
A_n = Q_p \frac{a - \beta + 2}{a h^{\alpha - \beta + 2}} \left( \frac{1-\beta}{2} \right) J_{\mu} \left( k_n z_a^{(\alpha - \beta + 2)/2} \right) \tag{53} \]

Substituting the expression for \( A_n \)'s, \( n \geq 1 \) into (Eq. (51)), the final solution of an elevated point source at crosswind is obtained as:
Mathematical Modeling of Air Pollutants: An Application to Indian Urban City

\[ C(x,z) = Q_p \left[ \frac{\alpha - \beta + 2}{ah^{\alpha - \beta + 2}} \right] (z \zeta_a)^{1-\beta/2} \sum_{n=1}^{\infty} \frac{Y_n \left( \frac{z}{\zeta_a} \zeta_a \right)^{\left(\alpha - \beta + 2\right)/2}}{J^2_p \left( Y_n \right)} \times \exp \left[ - \frac{b (\alpha - \beta + 2)^2 Y_n^2}{4ah^{\alpha - \beta + 2}} \right] \int_{0}^{\infty} f(s) ds \]  

(54)

where \( Y_n \) is given as:

\[ J_{\mu-1} (Y_n) = 0. \]  

(55)

### 2.2.2 Point Source Models

The steady state three dimensional solution of a point source can be obtained from (Eq. (18)) and \( f(x) \), which is expressed as a linear function of downwind distance, i.e., \( f(x) = \gamma Ux \) (Sharan & Modani, 2006), in different boundary conditions:

i. The solution in Dirichlet Boundary condition is given as

\[
C(x,y,z) = \frac{Q_p}{\sqrt{2\pi} \sigma} \left[ \frac{\alpha - \beta + 2}{ah^{\alpha - \beta + 2}} \right] (z \zeta_a)^{1-\beta/2} \sum_{n=1}^{\infty} \frac{Y_n \left( \frac{z}{\zeta_a} \zeta_a \right)^{\left(\alpha - \beta + 2\right)/2}}{J^2_p \left( Y_n \right)} \times \exp \left[ - \frac{b (\alpha - \beta + 2)^2 Y_n^2}{8ah^{\alpha - \beta + 2}} \right] \times \exp \left[ - \frac{(y - y_a)^2}{2\sigma_y^2} \right],
\]

where \( \mu = (1 - \beta) / (\alpha - \beta + 2) \), \( J_\mu \) is the Bessel function of order \( \mu \) and \( Y_n \)'s are obtained from the equation:

\[ J_{\mu} (Y_n) = 0 \]  

(57)

ii. The solution in Neumann Boundary condition is given as

\[
C(x,y,z) = \frac{Q_p}{\sqrt{2\pi} \sigma} \left[ \frac{\alpha + 1 - \alpha + 2}{ah^{\alpha + 1}} \right] \sum_{n=1}^{\infty} \frac{J_{-\mu} \left( \frac{z}{\zeta_a} \zeta_a \right)^{\left(\alpha + 1 - \alpha + 2\right)/2}}{J^2_p \left( Y_n \right)} \times \exp \left[ - \frac{b (\alpha + 1 - \alpha + 2)^2 Y_n^2}{8ah^{\alpha + 1}} \right] \times \exp \left[ - \frac{(y - y_a)^2}{2\sigma_y^2} \right],
\]

where \( \mu = (1 - \beta) / (\alpha - \beta + 2) \), \( J_{-\mu} \) is the Bessel function of order is \( -\mu \) and \( Y_n \)'s are obtained from the equation:

\[ J_{-\mu} (Y_n) = 0 \]  

(59)

iii. The solution in Mixed (type-I) Boundary condition is given as
\[
C(x,y,z) = \frac{Q_p}{\sqrt{2\pi\nu_y^2}} \left[\frac{\alpha - \beta + 2}{ah^{\alpha-\beta+2}} (zz_x)^{1-\beta/2} \sum_{n=1}^{\infty} J_\mu(\nu_y/h)^{(\alpha-\beta+2)/2} J_{\mu-1}(\nu_y/h)^{(\alpha-\beta+2)/2}\right]
\times \exp\left(-\frac{b y U (x-x_s)^2 (\alpha - \beta + 2) y_n^2}{8ah^{\alpha-\beta+2}}\right) \times \exp\left(\frac{(y - y_s)^2}{2\sigma_y^2}\right),
\]
(60)

where \( \mu = (1-\beta)/(\alpha - \beta + 2) \), \( J_\mu \) is the Bessel function of order \( \mu \) and \( y_n \)'s are obtained from the equation:

\[
J_{\mu-1}(\nu_y) = 0
\]
(61)

iv. The solution in Mixed (type-II) Boundary condition is given as

\[
C(x,y,z) = \frac{Q_p}{\sqrt{2\pi\nu_y^2}} \left[\frac{\alpha - \beta + 2}{ah^{\alpha-\beta+2}} (zz_x)^{1-\beta/2} \sum_{n=1}^{\infty} J_\mu(\nu_y/h)^{(\alpha-\beta+2)/2} J_{\mu-1}(\nu_y/h)^{(\alpha-\beta+2)/2}\right]
\times \exp\left(-\frac{b y U (x-x_s)^2 (\alpha - \beta + 2) y_n^2}{8ah^{\alpha-\beta+2}}\right) \times \exp\left(\frac{(y - y_s)^2}{2\sigma_y^2}\right),
\]
(62)

where \( \mu = (1-\beta)/(\alpha - \beta + 2) \), \( J_\mu \) is the Bessel function of order \( \mu \) and \( y_n \)'s are obtained from the equation:

\[
J_{\mu-1}(\nu_y) = 0
\]
(63)

### 2.2.3 Correspondence between present solutions to the earlier solutions

In present solution, wind speed is a function of \( z \) and vertical eddy diffusivity is considered a function of \( x \) and \( z \) as \( K_z(x, z) = bx^2 f(x) \). If \( f(x) \) is equal to 1.0, the solutions of Eqs. (56), (58), (60) and (62) are become same as the solutions obtained by Lin & Hildemann (1996). When wind speed is power law function of \( z \) and vertical eddy diffusivity is function of \( x \) only, i.e. \( K_z(x, z) = f(x) = K(x) \) (for \( \beta = 0 \) and \( b = 1.0 \)), the expression for the two dimensional concentration i.e., Eq. (38) is same as solution by Sharan & Modani (2006).

### 2.2.4 Line source models

A line source can be considered as a superposition of point sources. The solution for finite line source can be obtained by integrating point source solution from \( y_1 \) to \( y_2 \) with unit source strength \( Q \) with the same \( y_n \)'s as in point source in different boundary conditions:

i. The solution in Dirichlet Boundary condition is given as

\[
C(x,y,z) = \frac{Q}{2} \left[\frac{\alpha - \beta + 2}{ah^{\alpha-\beta+2}} (zz_x)^{1-\beta/2} \sum_{n=1}^{\infty} J_\mu(\nu_y/h)^{(\alpha-\beta+2)/2} J_{\mu-1}(\nu_y/h)^{(\alpha-\beta+2)/2}\right]
\times \exp\left(-\frac{b y U (x-x_s)^2 (\alpha - \beta + 2) y_n^2}{8ah^{\alpha-\beta+2}}\right) \times \left[\text{erf}\left(\frac{y - y_1}{\sqrt{2}\sigma_y}\right) - \text{erf}\left(\frac{y - y_2}{\sqrt{2}\sigma_y}\right)\right]
\]
(64)
ii. The solution in Neumann Boundary condition is given as

\[
C(x, y, z) = \frac{Q}{2} \left[ \frac{\alpha + 1}{ah^{a+1}} + \frac{\alpha - \beta + 2}{ah^{a-\beta+2}} \sum_{n=1}^{\infty} J_{\mu}[Y_n(z/h)]^{(a-\beta+2)/2} J_{\mu}[Y_n(z_s/h)]^{(a-\beta+2)/2} \right] \cdot \exp\left( -\frac{byU(x-x_s)^2}{8ah^{a-\beta+2}} \right) \cdot \left[ erf\left( \frac{y - y_1}{\sqrt{2}a_y} \right) - erf\left( \frac{y - y_2}{\sqrt{2}a_y} \right) \right]
\]

iii. The solution in Mixed (type-I) Boundary condition is given as

\[
C(x, y, z) = \frac{Q}{2} \left[ \frac{\alpha - \beta + 2}{ah^{a-\beta+2}} \sum_{n=1}^{\infty} J_{\mu}[Y_n(z/h)]^{(a-\beta+2)/2} J_{\mu}[Y_n(z_s/h)]^{(a-\beta+2)/2} \right] \cdot \exp\left( -\frac{byU(x-x_s)^2}{8ah^{a-\beta+2}} \right) \cdot \left[ erf\left( \frac{y - y_1}{\sqrt{2}a_y} \right) - erf\left( \frac{y - y_2}{\sqrt{2}a_y} \right) \right]
\]

iv. The solution in Mixed (type-II) Boundary condition is given as

\[
C(x, y, z) = \frac{Q}{2} \left[ \frac{\alpha - \beta + 2}{ah^{a-\beta+2}} \sum_{n=1}^{\infty} J_{\mu}[Y_n(z/h)]^{(a-\beta+2)/2} J_{\mu}[Y_n(z_s/h)]^{(a-\beta+2)/2} \right] \cdot \exp\left( -\frac{byU(x-x_s)^2}{8ah^{a-\beta+2}} \right) \cdot \left[ erf\left( \frac{y - y_1}{\sqrt{2}a_y} \right) - erf\left( \frac{y - y_2}{\sqrt{2}a_y} \right) \right]
\]

where \(\text{erf}\) is the error function defined by

\[
\text{erf}(a) = \frac{2}{\sqrt{\pi}} \int_{0}^{a} e^{-t^2} dt
\]

The solution of an infinite line source of source strength \(Q_s\) can be obtained by integrating point source along the crosswind direction \(y_s = -\infty\) to \(\infty\), which is different than the finite line source formulation, since \(\int_{-\infty}^{\infty} \exp\left( -\frac{(y - y_s)^2}{2\sigma_y^2} \right) dy_s = 1\).

### 2.2.5 Area source models

The concentration of air pollutants due to a finite area source at \((x, y, z)\) is calculated as a superposition of finite line sources extending from \(x_1\) to \(x_2\) in \(x\) direction is obtained as

\[
C(x, y, z) = \int_{x_1}^{x_2} C(x-x_s, y, z) dx_s,
\]

where \(C(x-x_s, y, z)\) is equivalent to the finite line source as obtained in section 2.2.4. However, the source strength \(Q_s\) is replaced by \(Q_s\) in (Eqs. 64, 65, 66 and 67) with same \(Y_s\)’s as in point source.

For an infinite area source with uniform strength \(Q_s\), the solution is obtained as:

\[
C(x, z) = \int_{x_1}^{x_2} C(x-x_s, z) dx_s,
\]
where \( C(x-x_s, z) \) is equivalent to the crosswind integrated concentration. However, the source term \( Q_p \) is replaced by \( Q_s \) in (Eqs. 36, 38, 46 and 54) with same \( \gamma \)'s as in point source.

Eqs. 68 and 69 are the solutions for finite and infinite area sources and can be solved numerically.

If wind speed and vertical eddy diffusivity are considered as function of \( z \), i.e. \( K_z(x, z) = \beta z^\nu \) with unbounded region (\( h \to \infty \)) and ground level source (\( z \to 0 \)) in (Eq. (68)), the expression obtained for the concentration is same as in Park & Baik, 2008.

3. Case study of Delhi

A case study of Delhi has been attempted through the application of two different categories namely statistical and analytical models to fulfill the objective of the chapter.

### 3.1 Application of statistical models

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 (SO\(_2\)), Nitrogen dioxide (NO\(_2\)) 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 Environmental 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 1.

The formula (EPA, (1999)) used to calculate AQI for four criteria pollutants RSPM, SO\(_2\), NO\(_2\) and SPM from 2000-2006 is given below:

\[
I_P = \left[ \frac{(I_{Hi} - I_{Lo})}{(BP_{Hi} - BP_{Lo})} \right] (C_P - BP_{Lo}) + I_{Lo},
\]

(70)

where \( I_P \) = the AQI for pollutant ‘p’,

\( C_P \) = actual ambient concentration of the pollutant ‘p’,

\( BP_{Hi} \) = the breakpoint in Table 1 that is greater than or equal to \( C_P \),

\( BP_{Lo} \) = the breakpoint in Table 1 that is less than or equal to \( C_P \),

\( I_{Hi} \) = the subindex value corresponding to \( BP_{Hi} \),

\( I_{Lo} \) = the sub index value corresponding to \( BP_{Lo} \).
Table 1. Propose sub-index and breakpoint pollutant concentration for Indian-AQI.

<table>
<thead>
<tr>
<th>SI.No.</th>
<th>Index values</th>
<th>Descriptor</th>
<th>SO$_2$ (24-h avg.)</th>
<th>NO$_2$ (24-h avg.)</th>
<th>RSPM (24-h avg.)</th>
<th>SPM (24-h avg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0-100</td>
<td>Good$^a$</td>
<td>0-80</td>
<td>0-80</td>
<td>0-100</td>
<td>0-200</td>
</tr>
<tr>
<td>2</td>
<td>101-200</td>
<td>Moderate$^b$</td>
<td>81-367</td>
<td>81-180</td>
<td>101-150</td>
<td>201-260</td>
</tr>
<tr>
<td>3</td>
<td>201-300</td>
<td>Poor$^c$</td>
<td>368-786</td>
<td>181-564</td>
<td>151-350</td>
<td>261-400</td>
</tr>
<tr>
<td>4</td>
<td>301-400</td>
<td>Very Poor$^d$</td>
<td>787-1572</td>
<td>565-1272</td>
<td>351-420</td>
<td>401-800</td>
</tr>
<tr>
<td>5</td>
<td>401-500</td>
<td>Severe$^e$</td>
<td>&gt;1572</td>
<td>&gt;1272</td>
<td>&gt;420</td>
<td>&gt;800</td>
</tr>
</tbody>
</table>

All the values of SO$_2$, NO$_2$, RSPM and SPM are in µg/m$^3$.

$^a$. Good: Air quality is acceptable; however, for some pollutants there may be a moderate health concern for a very small number of people.

$^b$. Moderate: Members of sensitive groups may experience health effects.

$^c$. Poor: Members of sensitive groups may experience more serious health effects.

$^d$. Very poor: Triggers health alter, everyone may experience more serious health effects.

$^e$. Severe: Triggers health warnings of emergency conditions.

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 ($t_{\text{max}}$), minimum temperature ($t_{\text{min}}$), daily temperature range (difference between daily maximum and minimum temperature, $t_{\text{range}}$), average temperature ($t_{\text{avg}}$), 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, SO$_2$, NO$_2$ 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 two models MLR and ANN independently. The performance of both the models has been assessed with respect to the statistical parameters.

### 3.1.1 AQI by MLR model

The following MLR equations for different seasons are resulted through training of AQI and meteorological data of 2000-2005 using SPSS software for summer, monsoon, post monsoon and winter respectively:
[\text{AQI}] = 0.0478 + 0.504 \times [\text{AQI}_{d-1}] -0.079 \times [\text{rh}] +0.126 \times [\text{t}_{\text{max}}] - 0.068 \times [\text{cc}] \quad (71)

[\text{AQI}] = 0.181 + 0.599 \times [\text{AQI}_{d-1}] - 0.282 \times [\text{rh}] - 0.128 \times [v] - 0.155 \times [\text{t}_{\text{min}}] \quad (72)

[\text{AQI}] = -0.324 + 0.537 \times [\text{AQI}_{d-1}] + 0.573 \times [\text{slp}] - 0.112 \times [\text{vp}] + 0.070 \times [\text{ssh}] - 0.135 \times [v] + 0.066 \times [\text{t}_{\text{max}}] \quad (73)

[\text{AQI}] = 0.171 + 0.503 \times [\text{AQI}_{d-1}] - 0.191 \times [v] - 0.115 \times [\text{cc}] - 0.169 \times [\text{wsp}] - 0.157 \times [\text{rh}] + 0.151 \times [\text{rf}] \quad (74)

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 Figs.1 (a), (b) (c) and (d).

Fig. 1. 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 Fig. 1 reflects that the trained values of AQI are matching well with observed values (calculated directly from formulation Eq. (70)). The same set of Eqs. (71-74) have been used for forecasting the daily AQI in all four seasons of the year 2006 which have been shown graphically in Figs. 2 (a), (b), (c) and (d) 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.

Fig. 2. 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 year 2006.

The quantitative analysis of comparison of forecasted and observed values of AQI has been made through statistical parameters, which are summarized in Table 2. The NMSE and coefficient of determination ($R^2$) 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 $R^2$ are 0 and 1 respectively. The values of $R^2$ 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.
Table 2. Comparison of MLR model predicted and observed values in years 2000-2005 and year 2006.

<table>
<thead>
<tr>
<th>S.N.</th>
<th>Season</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
</tr>
<tr>
<td>1</td>
<td>Summer</td>
<td>33.13</td>
</tr>
<tr>
<td>2</td>
<td>Monsoon</td>
<td>72.99</td>
</tr>
<tr>
<td>3</td>
<td>Post Monsoon</td>
<td>99.18</td>
</tr>
<tr>
<td>4</td>
<td>Winter</td>
<td>64.39</td>
</tr>
</tbody>
</table>

3.1.2 AQI by ANN model

The ANN model for different seasons is developed through training of AQI and meteorological data of 2000-2005 using MATLAB. The weights of a network are iteratively modified to minimize the total mean squared error between the desired target and actual output values. The above model has also been trained using observed daily AQI (same values as in MLR) of 2000-2005 for all four seasons. The comparison of observed and trained AQI has been shown graphically in Figs.3 (a), (b) (c) and (d).

Fig. 3. Comparison of observed and ANN model predicted values of daily AQI in (a) Summer, (b) Monsoon, (c) Post Monsoon and (d) Winter seasons during the years 2000-2005.
The Fig. 3 reflects that the trained values of AQI are showing the same trend as observed values. The same trained architectures have been used for forecasting the daily AQI in all four seasons of the year 2006, which have been shown graphically in Figs. 4 (a), (b), (c) and (d) 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. 4 in order to validate the forecasted values of AQI.

![Comparison of observed and ANN model forecasted values of daily AQI in (a) Summer, (b) Monsoon, (c) Post Monsoon and (d) Winter seasons during the year 2006.](image)

The comparison of forecasted and observed values of AQI has been made through statistical parameters in Table 3, which revels that the NMSE and coefficient of determination ($R^2$) are found as (0.0118, 0.4976) in summer season which are followed by (0.0277, 0.3986) in winter; (0.0540, 0.3619) in monsoon and (0.0718, 0.3816) in post monsoon seasons, showing good performance of the model as the ideal values of NMSE and $R^2$ are 0 and 1 respectively. The values of $R^2$ in four different seasons reflect that the model’s forecasted and observed AQI could be correlated explained by the selected input variables as approximately 50% in summer, 40% in winter, 38% in post monsoon and 36% in monsoon seasons. However, the fractional bias shows the under-prediction in all the seasons except Monsoon.
Table 3. Comparison of ANN model predicted and observed values in years 2000-2005 and year 2006

<table>
<thead>
<tr>
<th>S.N.</th>
<th>Season</th>
<th>RMSE</th>
<th>NMSE</th>
<th>Coefficient of determination</th>
<th>Fractional Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Summer</td>
<td>36.88</td>
<td>0.0118</td>
<td>0.4976</td>
<td>0.0269</td>
</tr>
<tr>
<td>2</td>
<td>Monsoon</td>
<td>64.27</td>
<td>0.0540</td>
<td>0.3619</td>
<td>-0.0224</td>
</tr>
<tr>
<td>3</td>
<td>Post Monsoon</td>
<td>76.44</td>
<td>0.0718</td>
<td>0.3816</td>
<td>0.1355</td>
</tr>
<tr>
<td>4</td>
<td>Winter</td>
<td>56.52</td>
<td>0.0277</td>
<td>0.3986</td>
<td>0.1330</td>
</tr>
</tbody>
</table>

3.1.3 Comparison of MLR and ANN models
On the basis of above exercise, it can be seen that ANN model is giving better results than MLR model with respect to RMSE and NMSE as shown in Tables 2 and 3, although, the $R^2$ is showing almost same values in both the models throughout all four seasons.

3.2 Application of analytical models
The above discussed analytical model with Neumann boundary condition is applied to simulate the hourly concentration of RSPM in the month of January (representative of winter season) 2008 due to point, line and area sources. The most important parameters of the models are emission inventory and meteorological variables, which are pre-process according to the model’s requirement.

A gridded emission inventory of RSPM has been developed over an area of 26 km x 30 km of Delhi. The total area has been divided into 195 square grids of size 2 km x 2 km. Emission of RSPM has been estimated in each grid due to all anthropogenic sources viz., domestic, industries, power plants and vehicles for the year 2008 using the primary and secondary data. The emission of RSPM from domestic sector has been calculated on the basis of fuel consumption data and the emission factor of the corresponding fuel. The emission estimation from industrial sources has been made by using the data obtained from concerned agencies / department. The emission of RSPM from power plants has been calculated on the basis of information given by Delhi Pollution Control Committee (DPCC). However, estimation of vehicular emission in each grid are made by using the following mathematical formulation, based on total number of registered vehicles, emission factors, vehicle kilometer traveled (VKT) of each type of vehicle, different type of roads:

$$E_i = N_i \times e_i \times VKT_i$$  \hspace{1cm} (75)

where $N_i$ is the number of vehicle of $i^{th}$ category, $E_i$ is the emission of pollutant by vehicle of $i^{th}$ category ($g$/day), $e_i$ is emission factor of $i^{th}$ category vehicle for the pollutant ($g$/km-vehicle). $VKT_i$ is vehicle kilometer travel per day for $i^{th}$ category vehicle (km/day).

Further, the total emission of pollutant has been calculated by summing the emissions from each category of vehicles:

$$E = \sum_i E_i$$  \hspace{1cm} (76)
The emissions of RSPM due to vehicular sources in Delhi for the year 2008 have been apportioned into each grid w. r. t. to the road lengths. The gridded emission inventory shows the spatial distribution of emissions of RSPM due to all types of sources (domestic, industries, power plants and vehicles) in Fig. 5. The emission from domestic sources has been distributed uniformly in all the grids and the emission from industries and traffic intersections have been apportioned according to their locations. The emissions of RSPM due to three power plants namely Indraprastha, Badarpur and Rajghat are superimposed in the grids as per their locations. It has been found that vehicles are the major source, contributing 57% of RSPM among all the estimated sources, which is followed by power plants 29% of total RSPM. While the contributions of domestic and other industrial sources are 8% and 6% respectively. The Fig.5 reflects that grids, in which power plants are located in addition to traffic intersections and industries, have higher values of emissions compared to others. This emission inventory is used as input to the model.

From an atmospheric pollution perspective, the most important season in Delhi is the winter lasting from December to February. This period is dominating by cold, dry air and ground based inversion with low wind conditions (<1 ms⁻¹), which increase the concentration of pollutants (Anfossi et al., 1990). For practical reasons, the January month of 2008 is used in this case study. The hourly meteorological data, measured at Safdarjung Airport from India.
Meteorological Department (IMD) is used as second input file to the models. There are some traditional methods, which are used to determination, the stability classes to produce the input in the model. The some traditional methods based on (i.) the wind velocity, sun radiation intensity and cloud cover; (ii.) measurement of the wind direction fluctuations; (iii.) the vertical temperature gradient; (iv.) the Richardson number can used to determination, the stability classes.

<table>
<thead>
<tr>
<th>Stability Class</th>
<th>Stability class of Pasquill</th>
<th>$\sigma_\theta$ (in degree)</th>
<th>Vertical Temperature Gradient ($^\circ$C/m $10^{-2}$)</th>
<th>Richardson Number at 2 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very unstable</td>
<td>A</td>
<td>25.0</td>
<td>&lt;-1.9</td>
<td>-0.9</td>
</tr>
<tr>
<td>Moderately unstable</td>
<td>B</td>
<td>20.0</td>
<td>-1.9 to -1.7</td>
<td>-0.5</td>
</tr>
<tr>
<td>Slightly unstable</td>
<td>C</td>
<td>15.0</td>
<td>-1.7 to -1.5</td>
<td>-0.15</td>
</tr>
<tr>
<td>Neutral</td>
<td>D</td>
<td>10.0</td>
<td>-1.5 to -0.5</td>
<td>0</td>
</tr>
<tr>
<td>Slightly stable</td>
<td>E</td>
<td>5.0</td>
<td>-0.5 to 1.5</td>
<td>0.4</td>
</tr>
<tr>
<td>Moderately stable</td>
<td>F</td>
<td>2.5</td>
<td>1.5 to 4.0</td>
<td>0.8</td>
</tr>
</tbody>
</table>


Note: $\sigma_\theta$ is the standard deviation of horizontal wind direction.

Table 5. Classification of Atmospheric Stability based on wind direction fluctuations, vertical temperature gradient and the Richardson number.

In this study atmospheric stability is measured on the basis of surface wind speed, cloud cover and solar insolation (strong, moderate, slight). These stability are classified according to Pasquill’s stability classes of A, B, C, D, E and F, which range from extremely unstable to extremely stable as given by Turner (1969). The dispersion parameters have also been calculated based on the stability parameters as discussed in section 2.

In the above solutions discussed in section 2, the exponent $\alpha$ has relationship with Pasquill’s Stability Classes (Hanna et al., 1982) and $\beta = 1 - \alpha$ is based on the Schmidt’s conjugate law.
In these formulations \( f(x) \) is expressed as a linear function of downwind distance as \( f(x) = \gamma \bar{U} x \). In this function \( \bar{U} \) is average velocity and \( \gamma \) is turbulence parameter. This turbulence parameter \( \gamma \) is parameterized as the square of turbulent intensity using Taylor statistical theory of diffusion \( \gamma = \left( \frac{\sigma_w}{U} \right)^2 \) (Arya, 1995; 1999). Turbulent intensity can be expressed as \( \left( \frac{\sigma_w}{U} \right) = \tan(\sigma_\phi) \), where \( \sigma_\phi \) is the standard deviation of vertical wind direction in radians. For small \( \sigma_\phi \), the turbulent intensity which is also depends on atmospheric stability is approximated as (Arya, 1999): \( \frac{\sigma_w}{U} = \sigma_\phi \).

<table>
<thead>
<tr>
<th>Pasquill’s Stability Classes</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha ) for Urban</td>
<td>0.15</td>
<td>0.15</td>
<td>0.20</td>
<td>0.25</td>
<td>0.40</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Table 6. Approximate Correspondence between Pasquill’s Stability Classes and \( \alpha \) (Hanna, 1982).

In the above solutions, standard deviation of concentration distribution in crosswind direction is represented by a power of downwind distance \( \sigma_y = R \sigma_x \) (Seinfeld, 1986). These \( R \) and \( r \) are constants as depending on the atmospheric stability.

<table>
<thead>
<tr>
<th>Pasquill’s Stability Classes</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_\phi )</td>
<td>≥ 11.5</td>
<td>10.0-11.5</td>
<td>7.8-10.0</td>
<td>5.0-7.8</td>
<td>2.4-5.0</td>
<td>&lt; 2.4</td>
</tr>
</tbody>
</table>

Table 7. Approximate Correspondence between Pasquill’s Stability Classes and Turbulence parameters \( \sigma_\phi \) (Arya, 1999).

The 24 hourly averaged ground level concentration of RSPM from point, line and area sources for Jan 2008 has been obtained by using the emission inventory and meteorological data as input parameters to the models. The power plants are considered as the point sources. Most of the analytical methods for predicting the concentration from stack need the plume rise height (\( \Delta h \)) of stack (power plants). There are numerous methods for calculating the plume rise height; in this study it has been calculated, given by Carpenter et al., 1970:

<table>
<thead>
<tr>
<th>Source</th>
<th>Coefficient</th>
<th>Stability Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Turner, 1969; Martin, 1976)</td>
<td>R</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>0.443</td>
<td>0.324</td>
</tr>
<tr>
<td></td>
<td>r</td>
<td>0.894</td>
</tr>
</tbody>
</table>

Table 8. Approximate Correspondence between Pasquill’s Stability Classes and Gaussian plume dispersion parameters (Seinfeld, 1986).
\[
\Delta h = \frac{114G(1)^{1/3}}{u},
\]
(77)

where \( I = gV_s d^2(T_s - T_a)/4T_a \ m^3/s^3; G = 1.58 - 41.4 \left( \frac{\Delta \theta}{\Delta z} \right) \) is dimensionless; \( \frac{\Delta \theta}{\Delta z} \) is potential temperature gradient, \(^\circ\)K/m; the constant 114 in this equation has unit of \( m^{2/3} \); \( V_s \) is the stack gas exit velocity, in m/s; \( d \) is the stack exit diameter in m; \( u \) is the speed at the stack exit in m/s; \( T_s \) is the stack gas temperature; \( T_a \) is the environmental temperature and \( g \) is the gravitational acceleration, m/s\(^2\). This method is especially useful because the potential temperature gradient factor is used for adjusting the different stability conditions.

The spatial distribution of RSPM concentration obtained from models (Eqs. (58), (64) and (68)) with background concentration 40 µg/m\(^3\) for Delhi (Kansal et al., 2011) is shown in Fig. 2 in the form of isopleths, which indicates the hot spots of RSPM ranging from 300-800 µg/m\(^3\) at different locations namely I.T.O., Nizamuddin, Badarpur power station, I.P. powers station, Raighat power station, AIIMS, Sirifort and Dhaula Khuan, which are major traffic intersections or power plants. This figure also shows the higher concentration near the Shahzada Bagh, which has maximum number of industries.

Fig. 6. 24 hourly averaged concentration of RSPM (µg/m\(^3\)) due to all types of sources in Jan 2008.
The concentrations of RSPM, cumulative sum of concentrations due to point, line and area sources, is evaluated against observed concentrations obtained from CPCB and NEERI, at different locations, as shown in Table 9.

<table>
<thead>
<tr>
<th>Location</th>
<th>Observed (µg/m³)</th>
<th>Model predicted (µg/m³)</th>
<th>National Ambient Air Quality Standard (µg/m³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitampura</td>
<td>353.00</td>
<td>411.30</td>
<td>100.00</td>
</tr>
<tr>
<td>Sirifort</td>
<td>374.00</td>
<td>449.45</td>
<td>100.00</td>
</tr>
<tr>
<td>Janakpuri</td>
<td>279.00</td>
<td>345.49</td>
<td>100.00</td>
</tr>
<tr>
<td>Shahzada Bagh</td>
<td>478.74</td>
<td>603.01</td>
<td>100.00</td>
</tr>
<tr>
<td>Sarojini Nagar</td>
<td>363.77</td>
<td>401.28</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Table 9. Comparison of 24 hourly averaged predicted and observed concentration of RSPM at different locations in Delhi in Jan 2008.

The table 9 shows that the models predictions are higher than observed as well as NAAQS. However, these values are well within a factor of two of observed values, which satisfies the criteria of Chang and Hanna (1982) for assessing the performance of the model with conclusion that the models are performing satisfactory.

4. 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.

5. Appendix A

The statistical measures, which have been used for statistical evaluation of the performance of models has been given by Chang & Hanna (2004) as follows:
5.1 Coefficient of correlation (R)
Coefficient of correlation (R) is relative measure of the association between the observed and predicted values. It can vary from 0 (which indicates no correlation to +1 (which indicates perfect correlation). A value of R close to 1.0 implies good agreement between the observed and predicted values i.e. good model performance.

\[
R = \frac{(C_o - \bar{C})(C_p - \bar{C}_p)}{\sigma_{C_p} \sigma_{C_o}}
\]

5.2 Coefficient of determination (R^2)
Coefficient of determination (R^2), which is the square of coefficient of correlation, determines the proportion of variance that can be explained by the model.

5.3 Root Mean Square Error (RMSE)
RMSE, is a measure of the differences between values predicted by a model and the observed values and is expressed as follows:

\[
RMSE = \sqrt{\frac{(C_o - C_p)^2}{n}}
\]

5.4 Normalized Mean Square Error (NMSE)
NMSE, as a measure of performance, emphasizes the scatter in the entire data set and is defined as follows:

\[
NMSE = \frac{(C_o - C_p)^2}{\bar{C}_o \cdot \bar{C}_p}
\]

The normalization by \( \bar{C}_o \cdot \bar{C}_p \) ensures that NMSE will not be biased towards models that over predict or under predict. Ideal value for NMSE is zero. Smaller values of NMSE denote better model performance.

5.5 Fractional Bias (FB)
It is a performance measure known as the normalized or fractional bias of the mean concentrations:

\[
FB = \frac{(\bar{C}_o - \bar{C}_p)}{0.5(\bar{C}_o + \bar{C}_p)},
\]

where:
- \( C_p \): model predictions,
- \( C_o \): observations,
- Overbar (\( \bar{C} \)): average over the dataset, and
- \( \sigma_C \): standard deviation over the data set.
6. References


Air pollution has been a major transboundary problem and a matter of global concern for decades. High concentrations of different air pollutants are particularly harmful to large cities residents, where numerous anthropogenic activities strongly influence the quality of air. Although there are many books on the subject, the one in front of you will hopefully fulfill some of the gaps in the area of air quality monitoring and modeling, and be of help to graduate students, professionals and researchers. The book is divided in five sections, dealing with mathematical models and computing techniques used in air pollution monitoring and forecasting; air pollution models and application; measuring methodologies in air pollution monitoring and control; experimental data on urban air pollution in China, Egypt, Northeastern U.S, Brazil and Romania; and finally, the health effects due to exposure to benzene, and on the influence of air pollutants on the acute respiratory diseases in children in Mexico.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:


InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821