Estimation of PM$_{2.5}$ Emissions and Source Apportionment Using Receptor and Dispersion Models

Swetha Priya Darshini Thammadi, Sateesh Kumar Pisini, Sanjay Kumar Shukla

Abstract—Source apportionment using Dispersion model depends primarily on the quality of Emission Inventory. In the present study, a CMB receptor model has been used to identify the sources of PM$_{10}$, while the AERMOD dispersion model has been used to account for missing sources of PM$_{2.5}$ in the Emission Inventory. A statistical approach has been developed to quantify the missing sources not considered in the Emission Inventory. The inventory of each grid was improved by adjusting emissions based on road lengths and deficit in measured and modelled concentrations. The results showed that in CMB analyses, fugitive sources - soil and road dust - contribute significantly to ambient PM$_{2.5}$ pollution. As a result, AERMOD significantly underestimated the ambient air concentration at most locations. The revised Emission Inventory showed a significant improvement in AERMOD performance which is evident through statistical tests.

Keywords—CMB, GIS, AERMOD, PM$_{2.5}$, fugitive, emission inventory.

I. INTRODUCTION

PARTICULATE MATTER (PM) levels are very high in Indian cities PM$_{2.5}$: 25–200 μg m$^{-3}$ [27]. A review of air quality trends in India suggests that the levels of PM exceed both 24-h and annual standards at most locations of the Indian National Air Quality Monitoring Program [11]. PM levels in large cities in India could be 5-10 times higher than those in the European cities [26]. These high PM levels may severely impact public health [25] and there are evidences of respiratory health problems which could be related to high pollution levels [27], [22]. Recent studies in India have focused on the chemical characterization of PM$_{10}$ [29], [13], [16]. Source apportionment of PM$_{10}$ has been done using Receptor and Dispersion models [31]. But not much is known about source apportionment of PM$_{2.5}$. Although PM$_{2.5}$ is a subset of PM$_{10}$, its sources, characteristics, health effects and behaviour in the atmosphere could be very different [10], [20], [34].

There is a need to plan and execute PM$_{2.5}$ controls, as PM$_{2.5}$ is a key pollutant with negative impacts on human health [23]. PM$_{2.5}$ can more readily penetrate into the lungs and are therefore likely to have short- and long-term effects such as increased respiratory symptoms, premature death and disease, decreased lung functions. To arrive at PM$_{2.5}$ control strategies, one needs information about the sources and their relative contributions to the ambient air PM levels. The main constraints for effective dispersion modelling are incomplete EI and availability of quality meteorological data. EI is a structured collection of information about emissions of pollutants [28] in a specified area and permits allocation of emitted pollutants to the originating sources. Therefore, it is evident from the above discussion that both receptor and dispersion models have their own benefits and limitations. Dispersion models use meteorology and Emission Inventory (EI) to trace the dispersal trail of a pollutant and thus evaluate its effect at the receptor. Studies have shown greater interest in the past on receptor models, since information on meteorological data and EI (essential for dispersion models) is not required. The receptor model shows the contribution of each source to ambient air pollution level at the receptor location [6]. However, using a receptor model requires chemical characterization (organic and inorganic composition) of PM and is not simple.

Regardless of preference for a receptor or dispersion model, EI is essential for post processing of modelling results to develop air pollution control strategies/action plans. Preparation of EI, in fact, is the first step in planning control of air pollution. The methods and procedures for developing EI for regular point, line and area sources are well established [3]. But identification and quantification of fugitive/ non-point emission sources is challenging. Emission factors for such sources are location and process specific and cannot be applied universally. Some of the important non-point sources include dusts from road [4] soil, pot holes, etc. and these sources should be included in the EI.

Dust emissions from paved roads can be estimated using silt load on the road surface and the average weight of vehicles travelling on the road [33]. This approach is not applicable for partially paved and unpaved roads as silt load may dramatically vary within a short distance. Indian roads in suburban towns are generally poorly paved and hence silt-load based approach cannot be adopted. In India, for these sources, EI is non-existent or incomplete [3]. As seen, preparation of EI is challenging, especially to capture fugitive sources in an urban mix, there is a need to explore new ways to assess fugitive component of emissions.

Reference [19] studied the quality of EI for fine particulates (PM$_{2.5}$) in London using the source contributions calculated from a PMF (Positive Matrix Factorization) receptor model and source apportionment using the AIRQUIS dispersion model. The study by [7] demonstrated an optimized approach

Swetha Priya Darshini Thammadi is with the Indian Institute of Technology Kanpur, India (e-mail: swetha.iitk@gmail.com).
for coupling dispersion and receptor models using a Genetic Algorithm (GA). Reference [24] apportioned PM$_{10}$ sources which were not well-defined in EI using a receptor model, and then augmented these sources in dispersion model predictions. Reference [21] used Ozone Isopleth Plotting Package-Research (OZIPPR) which is a trajectory box model and integrated into the GA to represent ozone transport and chemistry in order to develop ozone control strategies. The study by [18] used a dispersion model to predict concentrations and applied a multiple regression model to estimate emission rates in traffic areas.

An attempt has been made in this paper to apportion PM$_{2.5}$ by developing a mathematical approach for estimating the missing emissions of PM$_{2.5}$. The developed approach uses synergistically the two models: the CMB [9] receptor model and the AERMOD [2] dispersion model to update the EI of missing fugitive sources. The objective of this research is to first identify the missing sources contributing to ambient PM$_{2.5}$ using receptor modelling (as it does not require EI) and then improve the performance of dispersion modelling after updating the EI of missing sources identified by the receptor modelling. This approach of identifying sources using a receptor model, revising EI and dispersion modelling has been demonstrated for a town, Baddi – Nalagarh (30.9412° N latitude, 76.78° E longitude) in the State of Himachal Pradesh, India.

II. METHODOLOGY

Fig. 1 summarizes the stepwise methodology of this study. The AERMOD dispersion modelling has been used and meteorological data (mixing height, wind direction, wind speed, etc.) were obtained from CDAC (Centre for Development of Advanced Computing), Pune and IMD (Indian Meteorological Department). The meteorological data were processed in RAMMET and AERMET which are the processing tools of AERMOD. Stack information was collected from HPSPCB (Himachal Pradesh State Pollution Control Board) for preparation of EI. Source specific EI was developed based on Emission factors given by CPCB [12]. This EI was used as input for AERMOD. It is to be noted that, soil and road dust have not been accounted in the EI of PM$_{2.5}$ as the inventory for fugitive sources was missing. AERMOD was run for all point and area sources separately. Model evaluation was taken up and validated by comparing experimental and model-computed concentrations using statistical tools.

EPA-CMB version 8.2 receptor model was used for identifying the missing sources in EI. The CMB model requires specified ambient data and source profiles to identify and quantify source contributions. The receptor concentrations with appropriate uncertainty estimates and source profile abundances, serve as input data to CMB model. For CMB modelling, CPCB [12] and USEPA [9] reported source profiles were used. The output consists of the amount contributed by each source type to ambient PM$_{2.5}$ concentration. The identified major sources from CMB model can be compared with sources present in EI to get an insight on missing sources. Effort has been made to include missing sources in the existing EI and to improve the model performance of dispersion model using the revised EI.

A. Study Area

The Baddi-Nalagarh (BN) area (Fig. 2) is the most industrialized region in Solan district. The types of industries include - pharmaceutical, textile, chemical, iron, cement, rubber, steel, spinning mills etc. In most offices and institutions, diesel generators were used at the time of power failure. The road condition in the town was quite bad as roads were poorly maintained, broken, partially paved surfaces. It is observed that movement of vehicles causes non-exhaust road dust emission which is a significant amount. The major traffic flow is on national highway 22 which leads to a major tourist destination, Kullu. The BN area is considered for air quality sampling and later used for receptor and dispersion modelling.

Land use of sampling sites: S2, S7, S9, S10, S11, S13 (Industrial); S1, S12 (Industrial & Commercial); S4, S5, S8 (Commercial) and S3, S6 (Residential)

B. GIS Based Emission Inventory

Various maps (Wards, Roads, etc.) of BN were collected from different agencies (e.g.: Census of India, BN Industrial map, Baddi Barotiwala Nalagarh Development Agency, etc.) and digitized using ArcGIS 9.2. World geodetic system (WGS) 1984 (UTM Zone 44 N) map projection system was used for geo referencing the maps. The latitude and longitude of different points including the sampling sites and major landmarks were geo referenced. These geo-referenced maps were first digitized for city boundaries, landmark locations, and road network. All the digitized features were superimposed on 2 km x 2 km grids (total 434 grids). Road lengths in each grid for minor roads (number of vehicles less than 10,000 per day) and major roads (number of vehicles more than 10,000 per day) were calculated from the digitized maps using the ArcGIS tool, ArcMap.

1. Activity Levels

The broad classifications of sources in this area are: (a) industry point sources (stack height ≥ 20 m) (b) line sources (vehicles) (c) industry area sources (stack height < 20 m) and (d) domestic sources (LPG, wood burning, kerosene). The details of industries (e.g. stack details, fuel uses, production etc.) in each grid were collected from HPSPCB. To determine the fraction of vehicle technology classes in service on city streets, video cameras were set up at four locations along the road side and traffic movement was recorded from 08:00 - 11:00 am and 5:00 - 8:00 pm. The traffic volume during lean hours has been extrapolated from the traffic data that was already collected. Parking lot survey procedure given in [28] was adopted in BN region to determine the fractions of various types of vehicles (e.g.: 2 wheelers, 3 wheelers, 4 wheelers, trucks, buses, year of manufacturing, etc.). ArcMap was used to calculate the road length for both major and minor roads in each grid. Based on the road length and number of vehicles on road, total vehicle kilometre travel (VKT) for each vehicle category was estimated in each grid.
For domestic sources, the fuel consumptions of wood, kerosene, and LPG were considered. The population per panchayat was taken from [8] was known and persons per household were estimated. Activity data (kg of fuel/person/day) for each fuel category multiplied by population in each panchayat provided the consumption of wood, kerosene and LPG. The category-wise domestic fuel consumption was assigned to each grid. Emission factors [12] for domestic fuels, vehicles, point sources and industrial productions were used to estimate PM$_{2.5}$ emissions from each grid.

C. Air Quality Monitoring and Chemical Analysis

Air-quality monitoring was carried out at 13 sampling sites (S1 to S13) (Fig. 2) to assess the status of air quality and validation of the model. The period of sampling was May 14, 2012 - June 20, 2012. The WINS Impactor based PM$_{2.5}$ samplers (ECOTECH MODEL AAS 271) were used to collect the particles on Teflon and quartz filter papers. All initial and final weighing of filter papers (Whatman PTFE (Teflon) Membrane, 46.2 mm with support ring) was carried out on Mettler-Toledo MX-5 USA balance having sensitivity of 0.00001 g in a humidity-controlled room. Filters were conditioned in desiccators for 24 hours before and after sampling. The USEPA weighing protocol [32] was followed in this study.

For measuring OC (Organic Carbon) and EC (Elemental Carbon) from the particles collected on quartz filter paper, DRI Model 2001A OC/EC analyzer was used. The analyzer operation is based on preferential oxidation of OC and EC at different temperatures. Its working is based on the fact that OC can be volatilized from the sample in a non-oxidizing helium (He) atmosphere, while EC must be combusted by an oxidizer.

PM$_{2.5}$ data included Metals – Na, Mg, Al, Cu, Si, K, Ca, Fe, Ti, Cr, V, Ni, Mn, Zn, Pb using X-ray fluorescence technique (RIGAKU ZSX Primus II series; Japan). For analyses of ions, filters were extracted using ultra-pure Milli-Q water following the reference method of water soluble inorganic ions (Compendium Method IO-4.2, EPA/625/R-96/010a 1999). Chemical analyses of water soluble inorganic ions were carried out using Ion Chromatography (Model 882 Metrohm, Switzerland).

![Fig. 1 Overall methodology to identify and estimate missing emissions](Image)
III. RESULTS AND DISCUSSION

A. EI and AERMOD Dispersion Modelling

The overall EI of PM$_{2.5}$ for vehicles, domestic fuels, and industries (as an area source and point source) is presented in Fig. 3.

![Fig. 3 Emission Inventory for PM$_{2.5}$](image)

It is to be noted that, road and soil dusts have not been accounted in the EI of PM$_{2.5}$. As seen from the results of CMB (Fig. 4), road and soil dusts contribute significantly to PM$_{2.5}$ at almost all sites. Therefore, the EI is incomplete. An attempt has been made to improve the EI by taking into account the contribution from road and soil dusts by combining the results of CMB8.2 and performance of AERMOD with existing EI (without Road and Soil dust). The EI and meteorology were used to model PM$_{2.5}$ levels using AERMOD.

![Fig. 4 CMB8.2 results showing contribution of each source for PM$_{2.5}$ at all the sites](image)
Thirteen sites were considered for AERMOD modelling (Fig. 2). Each site had a measured sample concentration for two days. Measured and AERMOD computed air quality results are shown in Fig. 5.

It was observed that the model under-predicted PM 2.5 levels at seven locations in the range of 30% - 78%. The measured and AERMOD modelled values for PM2.5 have been plotted in Fig. 6. The slope and intercepts of the best fit line are statistically significant at the 5% level of significance. The intercept indicates that even if emissions are zero, there exists significant measured concentration of PM 2.5 at about 15.45 µg/m3. The slope of the best fit line also suggests under prediction by the model. The R² indicating the model fit, although significant at the 5% level of significance, can still be stated as modest.

It can be inferred from the model results shown in Fig. 6 that EI is incomplete and there is a need to identify missing sources and judiciously distribute the missing sources to improve the model performance and update the EI for better decision making.

B. Identification of Missing Sources: Application of CMB 8.2

The CMB8.2 model was run at 13 sites (Fig. 2) to obtain the contribution of various sources to ambient PM2.5 levels. The acceptance criteria for model results are (i) R² (model fitting) be greater than or equal to 0.8 and (ii) percentage of measured mass reconciled by the model should be in the range 0.8 - 1.2. Receptor modelling results were analysed in terms of R² (model fitting) and percentage of measured mass reconciled by the model. The model could apportion the measured mass in the range 80-120%, which was acceptable as per [15].

It is observed from Fig. 4 that soil and road dust is a prominent source (greater than 30%) of PM2.5 in the study area in seven sites. It is to be noted while major contribution to PM2.5 is from road and soil dust for these sites but the existing EI (Fig. 3) does not include these sources which are fugitive in nature. It can be argued that incomplete EI without soil and road dust emission has resulted in under prediction of PM2.5 levels by the dispersion model. Thus, there is a need to improve the EI and model performance for developing proper control strategies. It is to be noted that fugitive soil and road dust were not considered in the EI. It can be argued that incomplete EI without soil and road dust emission has resulted in under-prediction of PM2.5 levels by the dispersion model, and therefore, there is a need to improve the EI and model performance for developing proper control strategies.

Soil and Road Dust Emission and Revised EI

The following approach has been adopted for accounting the soil and road dust emission which may vary from one grid to another. The road length (paved, partially paved and unpaved) in each grid is taken as an indicator of road and soil dust emission. For estimating this emission, Road factor in the ith grid (Rfi) is defined as:

\[ R_{fi} = \frac{R_i}{\sum R_i} \]  

where \( R_i \) = sum of all road lengths in ith grid; and \( \sum R_i \) = sum of all road lengths in all grids.

The deviation in measured and modelled concentration in ith grid \( \Delta C_i \) is taken as an indicator of missing soil and road dust emission which should relate to \( R_{fi} \).

The deviation in measured and modelled concentration in ith grid \( \Delta C_i \) = (\( C_{measured} - C_{modelled} \)) is taken as an indicator of missing soil and road dust emission which should relate to \( R_{fi} \). For the sites where measurements of PM2.5 were available, first \( \Delta C_i \) was estimated (using average of measured and modelled concentration at each sampling site) and a relationship between \( \Delta C_i \) and \( R_{fi} \) (road factor of the grid where the sampling site is located) is developed (Fig. 7). Two sites have same road factor value and hence there seems to be 12 values in Fig. 7.

The linear relation between \( \Delta C_i \) and \( R_{fi} \) indicates that \( \Delta C_i \).
depends on \( R_f \) as:

\[
\Delta C_i = 228.11 R_f i + 10.457
\]  

(2)

As per the first principle, pollutant concentration, \( C \) is proportional to emission rate, \( Q \) (kg/d). Therefore, the missing emission rate, \( \Delta Q_i \) (in \( i^{th} \) grid) can be estimated as:

\[
\Delta Q_i = \frac{Q_i}{C_i} \Delta C_i
\]  

(3)

where \( C_i \) is the modelled concentration.

By substituting \( \Delta C_i \) from (2) in (3);

\[
\Delta Q_i = \frac{Q_i}{C_i} (228.11 R_f i + 10.457)
\]  

(4)

Equation (4) has been used for estimating \( \Delta Q_i \) (having the same units as \( Q_i \)) in all the grids. Thus the EI has been revised to include road and soil dust and the results are presented in Fig. 8.

The model performance has been tested using statistical analyses [30] for (i) fractional bias (FB), (ii) normalized mean square error (NMSE), (iii) coefficient of correlation (r) and (iv) index of agreement (d) along with parameters of best fit line (slope and intercept) for existing and revised EI (Table I).

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>STATISTICAL PARAMETERS FOR PM(_{2.5})</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM(_{2.5}) (existing EI)</td>
<td>PM(_{2.5}) (revised EI)</td>
</tr>
<tr>
<td>Slope</td>
<td>0.45</td>
</tr>
<tr>
<td>Intercept</td>
<td>15.45</td>
</tr>
<tr>
<td>FB</td>
<td>0.46</td>
</tr>
<tr>
<td>NMSE</td>
<td>0.44</td>
</tr>
<tr>
<td>r</td>
<td>0.20</td>
</tr>
<tr>
<td>d</td>
<td>0.38</td>
</tr>
</tbody>
</table>

FB is a nonlinear operator that varies between \(-2\) and \(+2\). For best model performance FB has an ideal value of 0 [1]. A negative FB value indicates that the model over estimates, and a positive value suggests that the model under estimates. NMSE is an unbiased dimensionless statistic and is preferred over FB. It measures the random spread of the values around the mean, i.e., it deals with scatter or variance [17]. A value of zero for NMSE is considered to a perfect modelling fit. Smaller values of NMSE indicate better model performance, and acceptable criterion is NMSE \( \leq 0.5 \). The index of agreement (d) is a measure of relative error in estimated model results. It is a dimensionless number and ranges from 0 to 1, where 0 indicates complete disagreement between modelled and observed values and 1 indicates perfect fit. The coefficient of correlation (r) between observed and modelled values is also an independent indicator of model performance. The index of agreement (d) is used because the coefficient of correlation (r) cannot account for additive differences or differences in proportionality [5], [14].

It has been observed that the model performance has improved with revised EI for the above stated statistical tests: (i) values of FB decreased from 0.45 to 0.86 (ii) NMSE levels decreased from 0.44 to 0.08 (iii) ‘r’ increased from 0.20 to 1.07 (iv) estimated values of ‘d’ increased from 0.38 to 0.80. In this study, the value of NMSE decreased from 0.44 to 0.08, and it is reasonable claim that model performance is satisfactory and has improved with revised EI. The overall ‘r’ was found to increase from 0.20 to 1.07 which is statistically significant at 5% level of significance, indicating linear association between observed and modelled concentrations. The value of ‘d’ improved from 0.38 to 0.80 suggesting that the model is adequate for application and decision making. These statistical analyses validate that the model with revised EI can describe physical phenomena well and can be used for further interpretation and application. Fig. 9 presents a clear depiction of measured vs. modelled concentration with existing and revised EI.

### III. CONCLUSIONS

Based on the results and discussion, the following general conclusions can be made:

1. The CMB receptor model and AERMOD dispersion model can be combined to identify the missing sources, revise the EI and improve the performance of AERMOD dispersion model.
2. The CMB analyses showed that fugitive sources, soil and road dusts contribute substantially to ambient pollution.
3. AERMOD significantly underestimated the ambient air

Fig. 8 Measured and AERMOD modelled PM\(_{2.5}\) levels with revised EI

![Fig. 8 Measured and AERMOD modelled PM\(_{2.5}\) levels with revised EI](image)

The measured PM\(_{2.5}\) concentration is compared with the modelled PM\(_{2.5}\) concentration with existing and revised EI.

Fig. 9 PM\(_{2.5}\) measured vs. modelled concentration with revised EI

![Fig. 9 PM\(_{2.5}\) measured vs. modelled concentration with revised EI](image)

The measured PM\(_{2.5}\) concentration is compared with the modelled PM\(_{2.5}\) concentration with existing and revised EI.
concentration because the fugitive sources were not considered in the EI.

4. The existing EI was improved in each emission grid by adjusting emission as per road lengths and deficit in measured and computed concentrations.

5. The revised EI showed a significant improvement in AERMOD performance in terms of statistical tests, such as fractional bias, normalized mean square error, coefficient of correlation and index of agreement.

6. This method can be used for adopting effective and appropriate air pollution control strategies.

ACKNOWLEDGMENTS

We gratefully acknowledge the Himachal Pradesh State Pollution Control Board (HPSPCB) for their grant towards this study as per letter no. HPSPCB/EPS-Study of Solan/11e21727 dated 11-1-2012.

REFERENCES


Quality Research. Volume 8: 54–64.


