Research highlights

- ▶ Integrated statistical approach used to evaluate compliance of air quality standards
- ▶ Approach uses statistical distribution model and exceedence probability of standards
- ▶ Approach allowed source reduction estimates for non-compliance criteria pollutants
- ▶ Unlike other criteria pollutants, PM₁₀ and SPM require source emission reductions
- ▶ Use of probability density function needed for realistic estimates of exceedences

Citation details:

Sharma, P., Sharma, P., Jain, S., Kumar, P., 2013. An integrated statistical approach for evaluating the exceedence of criteria pollutants in the ambient air of megacity Delhi. Atmospheric Environment 70, 7-17. <http://dx.doi.org/10.1016/j.atmosenv.2013.01.004>

Citation details:

Sharma, P., Sharma, P., Jain, S., Kumar, P., 2013. An integrated statistical approach for evaluating the exceedence of criteria pollutants in the ambient air of megacity Delhi. *Atmospheric Environment* 70, 7-17.
Web link: <http://dx.doi.org/10.1016/j.atmosenv.2013.01.004>

An integrated statistical approach for evaluating the exceedence of criteria pollutants in the ambient air of megacity Delhi

Pragati Sharma^a, Prateek Sharma^b, Suresh Jain^b and Prashant Kumar^{c, d, *}

^a*Department of Applied Sciences, Bharati Vidyapeeth's College of Engineering, Paschim Vihar, New Delhi-110063, India*

^b*Department of Natural Resources, TERI University, Delhi, 10, Institutional Area, Vasant Kunj, New Delhi – 110070, India*

^c*Department of Civil and Environmental Engineering, Faculty of Engineering and Physical Sciences (FEPS), University of Surrey, Guildford GU2 7XH, United Kingdom*

^d*Environmental Flow (EnFlo) Research Centre, FEPS, University of Surrey, Guildford GU2 7XH, United Kingdom*

ABSTRACT

Like many countries, the Central Pollution Control Board (CPCB), Delhi, in India evaluates exceedences of air pollution levels against the National Ambient Air Quality Standards (NAAQS). One of the mandatory requirements for NAAQS compliance is that the probability of non-exceedence should be at least 0.98, meaning that the formulated framework of NAAQS is essentially statistical. The current practice for assessing the compliance criterion is based on simple computation of the count of number of exceedences in a given year, without giving any consideration to the distribution function followed by different pollutants in the ambient air. This becomes even more important for monitoring stations where continuous monitoring is not done for all 365 days, but assessment is based on a minimum sample of 104 readings recorded in a year. The proper method for evaluating the compliance is the foreknowledge of the population distribution and computation of non-exceedence (or exceedence) probability of NAAQS from the *probability density function (pdf)*. The study proposes an integrated and scientifically robust methodology that is generic in nature and could well be used for assessing the air quality compliance criteria laid out by the NAAQS for India, besides suggesting percent reduction in source emissions to those pollutants that exceed the NAAQS. The usefulness of proposed methodology is exhibited by a case study conducted on four criteria air pollutants – sulphur dioxide (SO₂), nitrogen dioxide (NO₂), suspended particulate matter (SPM), and particulate matter less 10 micron in size (PM₁₀) – monitored in the ambient air of megacity Delhi at six monitoring stations. The collected data at all these sites underwent statistical analysis for the: (i) identification and estimation of the

* Corresponding author: Civil Engineering (C5), Tel. +44 1483 682762; fax: +44 1483 682135; Email addresses: P.Kumar@surrey.ac.uk, Prashant.Kumar@cantab.net (Prashant Kumar)

best-fit distributions, (ii) computation of probability of exceedence of the NAAQS for the non-complying pollutants, (iii) determination of return period of NAAQS violation, and (iv) estimation of percentage source emission reduction to meet the NAAQS criteria for the non-complying pollutants using the statistical rollback theory. It was concluded that the knowledge of *pdf* is a basic and essential requirement for realistically evaluating the compliance of NAAQS.

Key words: Probability density function; Distributional models; Ambient air quality data; Air quality standards; Model identification and estimation; Megacity Delhi

1. Introduction

Ambient air quality standards are employed in many countries to safeguard public health and welfare (Heal et al., 2012). These standards are embodied in the air pollution management legislation of many countries, including India (Table S.1). In some countries such as Australia and the United Kingdom, these are used as guidelines. In addition to local national standards, the World Health Organisation (WHO) has formulated long-term air quality goals, which are generally adopted worldwide. The National Ambient Air Quality Standards (NAAQS) are used to determine the status of ambient air quality in India. The criteria for judging the compliance of NAAQS is: “24-hourly/8-hourly values should be met 98% of time in a year; however, 2% of time it may exceed but not on two consecutive days”. This clearly indicates that the basic framework of NAAQS has a statistical rather than a deterministic form. Thus, knowledge and availability of the *pdf* is essentially built in the compliance criterion. Figure 1 shows a typical *pdf* for a region meeting the compliance criteria laid out by the NAAQS for PM₁₀. The *pdf* can be used to generate the probability of exceedence of NAAQS concentration. Moreover, the *pdf* provides a higher order of description of concentration data for assessing the air quality compliance criteria.

The pollutant concentrations are routinely monitored at representative fixed receptor locations to assess the ambient air quality. Air pollution concentrations are inherently random variables because of their dependence on fluctuations of a variety of source emissions, meteorological and other spatio-temporal variables. When sets of random samples of historical ambient air quality data are available, various statistical characteristics can be determined and assigned to the pollutant concentrations. If certain assumptions are made, this statistical information can be incorporated into distribution functions and thus be used in an organised and efficient manner thereafter. To study the stochastic variability of

concentrations of different air pollutants, the understanding of probability *distribution* functions (PDFs) is necessary, primarily to evaluate the compliance of standards and formulate policies related to control of air pollution (Delbari et al., 2009; Lin et al., 2009; Mijić et al., 2009; Lin et al., 2010; Mohamed et al., 2011; Chu et al., 2012). Using order statistics theory, Georgopoulos and Seinfeld (1982) have shown how various forms of air quality standards can be formulated as properties of the distribution of air pollutant concentration. The output desired from models intended for air quality management is thus, notionally, the probability distribution of air pollutant concentrations. When the probability distribution is correctly chosen, the specific distribution can be used to predict the exceedence of ambient air quality standard (Lu, 2002). Based on the prediction of exceedences, informed decisions can be taken to prioritise the management strategies for different pollutants for a given region.

A number of studies have fitted different types of PDFs to the air pollutant concentration data. These distributions are lognormal (Mage and Ott, 1984; Kao and Friedlander, 1995; Burkhardt et al., 1998), Weibull (Georgopoulos and Seinfeld, 1982), pseudo-lognormal (Mage, 1984) and type V Pearson distribution (Morel et al., 1999). Numerous right-skewed distributions have also been used to fit atmospheric pollutant concentrations with varying results when compared to the lognormal *pdfs* (Graedel et al., 1977; Kalpasanov and Kurchatova, 1976). There have also been few studies on the concentration of distribution of aerosols (Savoie and Prospero, 1977; Savoie et al., 1987) and even fewer studies of urban aerosol distribution (Morel et al., 1999). Jia et al. (2008) have characterised the distributions of personal exposures to volatile organic compounds (VOCs) and provide national-level and representative statistics regarding VOC exposures, that can be used as important input for risk assessment and probabilistic analysis. A description of the general attributes of statistical distribution models is provided by Jakeman et al. (1988). Application of statistical distributions for generic applications have been comprehensively reviewed by Pollack (1975), Bencala and Seinfeld (1976), Georgopoulos and Seinfeld (1982), and Jakeman and Taylor (1989). Bai et al. (1992) provide a very good account of application of statistical distribution modelling in air quality management.

The information generated by the statistical distribution models has also been used to compute the percentage source emission reductions to meet the NAAQS for the non-complying pollutants. This has been carried out by applying the statistical theory of rollback (STR) model. For instance, Ott (1995) provides the theoretical details of the STR. Larsen (1961) originally proposed the rollback approach and showed the procedure to apply STR

model to air quality data. Larsen (1961, 1971) gives empirical deterministic form of STR model. Georgopoulos and Seinfeld (1982) suggest a stochastic version of the STR model. Their formulation is consistent with the STR developed by Ott (1995). The STR, when combined with statistical distribution models, form a powerful technique for predicting the effect of regulatory decisions on concentrations in the environment. The theory provides prediction of concentrations distribution for the post– and pre–emission control states (Lonati et al., 2011). The STR has been applied to compute percentage source reduction for various pollutants such as sulphur dioxide, SO₂ (Ott, 1995), nitrogen oxide, NO₂ (Georgopoulos and Seinfeld, 1982; Chaloupka et al., 2008), carbon monoxide, CO (Mage and Ott, 1984; Ott, 1995) and particulate matter $\leq 10 \mu\text{m}$ size, PM₁₀ (Lu, 2002). The combined technique can also be used for predicting standards and the return period for the highest level concentrations subsequent to emission reductions (Mijić et al., 2009).

The CPCB collects ambient air quality data through its network of fixed monitoring stations under the National Air Monitoring Programme (NAMP). At present there are 566 stations spread over the entire country, out of which 523 are manually run and remaining 43 are continuously run (automatic) in order to assess the air quality status. Only 28 stations are operational out of these 43 automatic run stations. The data on manually run stations is not collected on continuous basis; a minimum of 104 measurements are taken in a year (twice a week 24-hourly at uniform interval). The CPCB evaluates the NAAQS based on simple computation of the count of total number of exceedences in a given year; no considerations are given to the statistical distribution that produces such exceedences. This may give some proximity to the true air quality status for areas covered by continuously run automatic stations, but would be a gross oversimplification of the compliance criterion where entire data set for the whole year is not collected. Such a generalised approach can only be made when the information and knowledge of the complete statistical distribution of pollutant concentration is available. Moreover, in order to carry out a more detailed comprehensive source apportionment study, it is important to statistically link the statistical distribution of pollutant concentration with desired source emission to meet the NAAQS. This study aims to fill this research gap by proposing a scientifically robust, integrated statistical approach, for assessing the air quality exceedence criteria laid out by the NAAQS in India, or similar compliance criteria elsewhere. In order to demonstrate the proposed approach data for four criteria pollutants – SO₂, NO₂, SPM and PM₁₀ – monitored at six NAMP stations in the megacity Delhi are statistically analysed for: (i) identifying and estimating the best-fit distributions, (ii) computing probability of exceedence of the NAAQS for the non-complying

pollutants, (iii) determining return period of NAAQS violation, and (iv) estimating percentage source emission reduction to meet the NAAQS criteria for the non-complying pollutants using the STR theory.

2. Materials and methods

2.1 History and description of sampling sites

In India, the Air (Prevention and Control of Pollution) Act, 1981 (CPCB, 2010) establishes the national policy and provides the legislative support for the prevention and control of air pollution. The CPCB is the national apex body for the assessment, monitoring and control of pollution, under the Ministry of Environment and Forest, Government of India. The executive responsibilities for enforcement of various environment related Acts, including the Air Act, 1981 are carried out by the Board. The CPCB initiated the National Ambient Air Quality Monitoring (NAAQM) programme in the year 1984-85; this programme was later renamed as “NAMP”, as described in George et al. (2012). Within the purview of NAMP, six NAAQM stations were established in various parts of megacity Delhi (see Figure 2). The underlining ideas were to: (i) assess the current status of air quality in Delhi, (ii) its seasonal variation with increased industrialisation and urbanisation, and (iii) to understand any possible increase in various air pollution generation activities. Table 1 provides the detailed description of sampling locations, types of sampling areas, height of monitoring stations, and pollutants measured. Further, details on methods of measurements and instruments used to measure various pollutants are provided in Supplementary Information (SI) Table S.1. Figure 2 shows the geographical location of all the monitoring stations in Delhi. Further details of the site can be found elsewhere (e.g. Pandey et al., 2005; Garg, 2011; George et al., 2012).

2.2 Data acquisition

The monitoring of criteria pollutants (SO₂, NO₂, SPM, and PM₁₀) is carried out by CPCB for 24 hours, once in three working days. This includes 4-hourly sampling for gaseous pollutants and 8-hourly sampling for particulate matter. The CPCB follows strict quality control protocol and duly consider accuracy and precision of analytical air quality data. As described in Table S.1, the NAAQS state the measurement methods of various pollutants that are followed by the CPCB (2003) for measuring ambient air quality.

This study uses the 24-hourly average data for SO₂, NO₂, SPM and PM₁₀, collected from January 2003 to December 2006 at the six monitoring stations highlighted in Figure 2. The sampling has been done using respirable dust sampler (Model APM 460) using IS 5182 Part 2, IS 5182 Part 6, IS 5182 Part 23 methods for SO₂, NO₂ and PM₁₀, respectively

(Kandlikar, 2007; CPCB, 2009; CPCB, 2011). This data set had less than 5% values missing (i.e. less than 5 of any 100 raw data readings were unavailable), which were estimated by linear interpolation between prior and subsequent known points in the data sets. All the data sets were well above detection limit and hence there were no issues related to method detection limits. It may be noted that the SPM has now been revoked from the regulatory limits since 2009 and the reason this is appearing in our estimates is that the data analysed in our article is for periods when SPM standards were in operation.

2.3 Methodology

2.3.1 Modelling the statistical distribution of pollutant concentration

Statistical distribution modelling of pollutants involves a two-step procedure. In the first step, most suitable distribution model from a range of candidate models is identified using the goodness-of-fit tests. This is done by testing the null hypothesis (H_0) that pollution data comes from a particular distribution type against the alternative hypothesis (H_1) that the data comes from some other distribution form. In the present study, the goodness-of-fit was evaluated using Anderson and Darling (A-D), Kolmogorov-Smirnov (K-S) and Chi-square (χ^2) tests, and by visually examining the probability plots (Kottegoda and Rosso, 2008). The A-D test was given preference over other tests since it is suitable for fitting distributions both in the middle and higher percentile ranges. Thus, A-D test served the primary criterion for selection and other tests helped to confirm or improve the selection. The mathematical details of A-D test are provided in the SI Section S.1. In the second step, the parameters of the identified distribution are estimated by various methods to get unbiased model parameters. There are several methods of parameter estimation reported in literature, viz. method of moments (MoM), method of maximum likelihood estimator (MLE) and method of least square estimator (LSE) (Kottegoda and Rosso, 2008). Among these, MLE and MoM are the most preferred methods. Mage and Ott (1984) carried out Monte-Carlo experiments and concluded that the MLE provides the best estimates of the parameters than MoM. Moreover, the MLE solutions for model parameters are statistically efficient and provide estimates that have minimum variance (Ayyub and McCuen, 2011). In the present study, MLE has been used for estimating model parameters and has been described in SI Section S.2.

The following parametric forms of the distributions have been used in the present study to fit and analyse the air quality data: normal, lognormal, three-parameter lognormal, exponential, two-parameter exponential, Weibull, three-parameter Weibull, gamma, three-

parameter gamma, logistic, loglogistic and three-parameter loglogistic distribution. The statistical characteristics of these distributions have been provided in Table 2.

2.3.2 Evaluation of compliance of NAAQS and prediction of exceedences

One of the primary objectives of the present study is the application of statistical distributions of air pollutant concentrations to assess the degree of compliance with NAAQS. These standards define acceptable probability with which such concentrations can be exceeded, i.e. determination of probability of exceedance of NAAQS. The probability that a particular concentration level 'x' will be exceeded in a single observation is given by the complementary cumulative distribution function (*cdf*):

$$\bar{F}(x) = \Pr\{C > x\} = 1 - F(x)$$

where

$$F(x) = \Pr(C \leq x) \quad (1)$$

Where $\bar{F}(x)$ and $F(x)$ are complementary *cdf* and *cdf*, respectively; Pr, C and x represent the probability, pollutant concentration as a random variable and pollutant concentration value, respectively. Thus, the larger the concentration level x, the smaller is $\bar{F}(x)$. The theory and derivation of relationship between *cdf*, number of exceedences and return period has been provided in section SI S.3.

The *cdfs*, as well as certain statistical properties of these random variables, can be determined by applying the methods and results of *order statistics*. Georgopoulos and Seinfeld (1982) provide the theory behind the development of models for predicting the probability of exceedance of NAAQS and the return period for a given concentration (C_{NAAQS}). The theory is based on order statistics and assumes the concentration levels measured in successive non-overlapping periods are independent of one another; and the random variables are identically distributed (Jakeman et al., 1988; Georgopoulos and Seinfeld, 1982). It may be noted that air pollutant concentration levels have high autocorrelation and do violate the independence assumption. The independence of air quality data does pose some constraint on the application of statistical distribution models (SDMs) when the pollutant data is highly autocorrelated and non-stationary. However, autocorrelation is a problem for shorter time averages than the longer ones (Jakeman et al., 1988). Hadley and Toumi (2003) also observed that though the hourly data may be strongly auto-correlated, daily data are more likely to be statistically independent. The derivation of relevant relations between pdf, cdf and return period has been provided in the SI Section S.3.

The probability of exceeding the NAAQS, $\Pr[X > C_{\text{NAAQS}}]$ can be predicted by integrating the *pdf* and deriving the *cdf* and then the complementary *cdf* from it, using:

$$\Pr[C \leq c_{\text{NAAQS}}] = F_C(x) = \int_{-\infty}^{+x} f_C(z) dz. \quad (2)$$

Thus, the probability that C exceeds the NAAQS or the probability of exceedance is given by the complementary function of $F_C(x)$ as:

$$\Pr[C \geq c_{\text{NAAQS}}] = \bar{F}_C(x) = 1 - F_C(x) = 1 - \int_{-\infty}^{+x} f_C(z) dz. \quad (3)$$

The number of violations in a year for 24 hourly averages of SO_2 , NO_2 , SPM and PM_{10} can be estimated by multiplying the exceedance probability derived from Eq. (3) with the total number of 365 days in a year.

2.3.3 Modelling the percentage source emission reduction - the rollback model

The reduction in emission source strength ' R ' required to meet an air quality goal, C_{NAAQS} is calculated by the simple equation of STR model (de Nevers et al., 1977).

$$R = \frac{\kappa c_p - c_s}{\kappa c_p - c_b} \quad (4)$$

Where κ is the growth factor for future emission sources (i.e., the ratio of future source strength to present source strength in the absence of controls), c_b , c_s and c_p are background concentration, NAAQS prescribed concentrations, and present highest concentration observed corresponding to c_s , respectively. For example, if c_s is a value of the 8-hourly average concentration of CO not to be exceeded more than 2% of time in a year and then c_p would be the highest 8-hourly average CO concentration observed in the region in the present year. Following assumptions are made for using the above model.

- The spatial distribution of emission sources remains unchanged with time.
- Growth factor for all sources is common.
- The average meteorological conditions remain the same.
- Non-reactive pollutant species.

The first three assumptions refer to stationarity of the data generation process. The fourth assumption also indirectly refers to stationarity of a process as a “non-stationarity is likely to be strongest in data sets involving reactive pollutants (Georgopoulos and Seinfeld, 1982; Jakeman et al., 1988). Many modifications to Eq. (4) have appeared in the literature, representing efforts to relax some of the above assumptions (Larsen, 1969, Horie and Overton, 1974, de Nevers and Morris, 1975). Georgopoulos and Seinfeld (1982) and Mage

and Ott (1984) have modified the Eq. (4) on the arguments that the correct form of the STR model should be in the following form (Lu, 2002):

$$R = \frac{E\{c_p\} - E\{c\}_s}{E\{c_p\} - c_b} \quad (5)$$

where, $E\{c\}_s$ is the expected concentration of a distribution such that the extreme value for which corresponds to c_s . $E\{c_p\}$ is expected concentration of present distribution. Thus, if c_s is the SPM daily average ground concentration not to be exceeded more than 2% of time in a year ($\text{Pr}[\text{SPM} > c_s] = 2/100 = 0.02$), then the $E\{c\}_s$ is the expected daily SPM average concentration of a distribution whose probability of concentration exceeding $200 \mu\text{g m}^{-3}$ (for residential areas) equals 0.02. $E\{c_p\}$ would be the expected SPM daily concentration of the present distribution. Eq. (5) reveals that the concentration of the pollutant is proportional to emission source strength and the c_b in this equation can be neglected. Once the distribution type is decided, the values of $E\{c_p\}$ and $E\{c\}_s$ can be calculated from the present distribution characteristics of pollutant concentration. Then the emission source reduction, R , can be estimated.

3. Results and discussions

3.1 Model identification and parameter estimation

The *pdfs* for all the four pollutants monitored at the six NAAQM stations were identified and the respective parameters estimated using the methodology described under the Sections 2.3.1 and 2.3.2. Figures 2 to 5 show the best fit probability plots for the 24-hourly average SO_2 , NO_2 , SPM, and PM_{10} at the six NAAQM stations, respectively. Table 3 provides a comprehensive summary of various best-fit pdf models applied to chosen pollutants. These results demonstrate that theoretical distributions fit adequately to most of the data according to the goodness-of-fit criteria. However, few outlying values were observed in the uppermost percentile region for SO_2 at Ashok Vihar, Shahadra, Shahzada Bagh and Sirifort stations (Figures 3a, d, e and f), NO_2 at Nizamuddin (Figure 4c) and SPM at Shahzada Bagh (Figure 5e) despite moderate to good fit. The SO_2 data showed maximum number of outlier high values, especially during winter months, resulting in deviation in the fitted distribution (in the upper percentile region). The SO_2 , NO_2 and SPM data at all the above-mentioned stations was visually inspected and the models were again fitted after removing the outliers. There was only a modest improvement in the model fit, and hence the originally fitted distributions were retained.

The deviations seen in Figures 3-6 from the theoretical distribution can be ascribed to multiple circumstances such as values reported below detection limit, missing values, high outlying observations, anomalies and measurement errors (Jia et al., 2008; Le et al., 2007; Ott, 1995). However, these circumstances could be tackled one by one, leaving the high outlying observations during winter season as a dominant cause of deviations. For instance, the data considered here was well above the detection limit. The missing values, though a few, were replaced by the estimated values as described in Section 2.2. A visual examination of the data suggested that high outlying values were mostly observed during the winter season when the inversion conditions prevail, resulting in elevated pollution levels.

Table 3 shows that right-skewed distributions fit the best to the four pollutants at all the six monitoring stations, with the exception of NO₂ at Ashok Vihar station, which follows a normal distribution. The distributions fitting the data are lognormal, gamma, Weibull, loglogistic and logistic distributions. Ott (1995) observed that the right-skewed distributions tend to resemble each other, and it is very difficult to fit a single model to different data set under all conditions. Although he also showed through the “Theory of Successive Random Dilutions” that ambient air quality data follows lognormal distribution. Different studies have observed that several other right-skewed distributions also equally fit or provide a better fit than the lognormal distribution (e.g. Kalpasanov and Kurchatova, 1976; Graedel et al., 1977; Morel et al., 1999; Lu, 2002). However, these studies have empirically fitted theoretical distributions to the existing data sets, rather than explain the underlying stochastic process. In the present study an attempt has been made to empirically fit the best distribution to the data set.

3.2 Assessment of NAAQS compliance

As elaborated under Section 1, the criterion of judging the compliance of NAAQS is: “24-hourly/8-hourly values should be met 98% of time in a year; however, 2% of time it may exceed but not on two consecutive days”. To evaluate the compliance of NAAQS and prediction of exceedance, together with corresponding return period, a detailed analysis was carried out for each pollutant measured at all the six monitoring stations. Table 4 gives a compact summary of the probability of exceedence, comparison of number of violations of the NAAQS in the following year for observed and predicted values, compliance of NAAQS and return period for various NAAQM stations for different pollutants. A closer inspection of Table 4 reveals that the NAAQS criteria for SO₂ and NO₂ are being completely met throughout Delhi and no violation of NAAQS was observed at any of the six monitoring

stations. However, a number of exceedences over the NAAQS for SPM and PM₁₀ were found at all the stations.

The identified *pdfs* for all the pollutants were plotted for assessing the compliance to standards. However, only the *pdfs* of NAAQS complying and non-complying pollutants, i.e. SO₂ (Sirifort station) and PM₁₀ (Nizamuddin station) have been shown in Figure 6 for brevity reasons. The vertical reference line in Figure 7 corresponds to the NAAQS. This reference line ought to separate the area of *pdf* into two parts – the left area equal to 0.98 and the right area 0.02 for a complying pollutant, which represent the probability of non-exceedence and exceedence, respectively. The air quality criterion is being completely met for SO₂ (Figure 7a), but a gross violation is observed for PM₁₀ (Figure 7b) since the reference line is too much towards the left side. The above discussions suggest that the *pdf* for each pollutant along with reference line corresponding to the NAAQS makes it a very useful graphical tool to assess the compliance/non-compliance of air quality criterion.

3.3 Estimation of source emission reduction

Once the non-compliance of pollutants is known, the next step is to suggest percentage reduction in source emissions (*R*) to meet any standards (i.e. NAAQS in our case). This reduction was calculated using the STR model described in Section 2.3.3. Table 5 gives the summary of various factors required to estimate *R*, using the STR model, along with the estimated values of *R* for chosen pollutants. As can be seen from Table 5, SO₂ and NO₂ do not require any source reduction in the regions covered by the six NAAQM stations. However, SPM and PM₁₀ require a source emission reduction at all the six stations in order to meet the compliance criteria. The source emission reduction required varies from 30.86 to 89.10% for SPM and 55.88 to 85.71% for PM₁₀ for different monitoring stations.

5. Conclusions

The aim of this study was to present a methodology based on an integrated statistical approach which could be used for the compliance assessment of air quality standards and suggesting percent reductions in emissions of various pollutants. For demonstrating the usefulness of the proposed methodology, ambient concentrations data of the four criteria pollutants (i.e. SO₂, NO₂, SPM, PM₁₀) was collected at six NAMP stations in megacity Delhi. These data were analysed statistically to firstly identify the *pdfs* for all pollutants at these monitoring stations using the goodness-of-fit criteria. The respective parameters were then estimated using the MLE method. The results revealed that the *pdf* for air quality data are asymmetrical and positively skewed. The distribution form was found to vary with location

and pollutant type and thus, the data could not be represented by a single distribution function. Further, the identified *pdfs* were used to compute the exceedence probability of NAAQS. The results suggested that the NAAQS criterion for SO₂ and NO₂ is being completely met throughout Delhi and no violation of NAAQS was observed at any of the monitoring stations. The SPM and PM₁₀ concentrations exceeded the NAAQS at all the stations (Ashok Vihar, Janak Puri, Nizamuddin, Shahzada Bagh, Shahadra, Sirifort). Finally, STR model was used to estimate the percentage reductions in source emissions in order to meet the NAAQS for the non-complying pollutants. The results of emission reduction were found to depend on the distribution type and the tail properties of the distribution. Information on the type of distribution is therefore essential for predicting the emission reduction of various pollutants. It may be noted that the rollback model gives the overall source emission reduction required for a region to attain compliance status. However, it does not apportion the contribution from different sources, for which a more detailed source apportionment study using dispersion modelling is required.

The evaluation of ambient air quality is primarily based on the monitoring of pollutant concentrations and their compliance with corresponding standards, such as NAAQS in our case. Inherent in the compliance criteria is the knowledge and availability of the *pdf*. However, the current practice followed by the CPCB Delhi in India is to simply count and add the number of violations of the NAAQS in a given year, without giving any consideration to the population distribution of data followed by the pollutant. This is of specific relevance for those stations where continuous monitoring is not done and limited data, spread over different seasons, is collected. At present, 92% of the 566 monitoring stations run by CPCB are manual, where complete data set for the entire year is not collected and thus information about the entire population is not available. For accurate assessment of the compliance, identification of *pdf* for each station for each pollutant is an essential requirement and would thus form the desired scientific foundation for evaluating the NAAQS criteria. It may be noted that the *pdf* can be used for estimating the exceedence probability of NAAQS in a year, which is a compliance requirement, and does not provide a means to estimate such exceedences on multiple time intervals. For such situations the time-series models are more appropriate and can explicitly predict the air quality at different time intervals. The proposed methodology can be adopted for assessing the air quality exceedence criteria laid out by the NAAQS in India, or similar compliance criteria elsewhere.

6. Acknowledgements

The authors are grateful to the Central Pollution Control Board, Delhi, for providing the necessary data. Prashant Kumar thanks the UKIERI grant which enabled him to interact with the co-authors and write this joint article.

7. References

- Ayyub, B.M. and McCuen, R.H., 2011. Probability, Statistics, & Reliability for Engineers and Scientists, CRC Press, Boca Raton.
- Bai, J., Jakeman, A.J., McAleer, M., 1992. Estimation and discrimination of alternative air pollution models. *Ecological Modelling* 64, 89–124.
- Bencala, K.E., Seinfeld, J.H., 1976. On frequency distributions of air pollutant concentrations. *Atmospheric Environment* 10, 941–950.
- Burkhardt, J., Sutton, M.A., Milford, C., Storeton-West, R.L., Fowler, D., 1998. Ammonia concentrations at a site in Southern Scotland from 2 yr of continuous measurements. *Atmospheric Environment* 32, 325–331.
- Chaloulakou, A., Mavroidis, I., Gavriil, I., 2008. Compliance with the annual NO₂ air quality standard in Athens. Required NO_x levels and expected health implications. *Atmospheric Environment* 42, 454–465.
- Chu, H.J., Yu, H.L., Kuo, Y.M., 2012. Identifying spatial mixture distributions of PM_{2.5} and PM₁₀ in Taiwan during and after a dust storm. *Atmospheric Environment* 54, 728–733.
- CPCB, 2003. Guidelines for Ambient Air Quality Monitoring. National Ambient Air Quality Monitoring Series: NAAQMS/.../2003-04.
- CPCB, 2009. National ambient air Quality status 2008, Central pollution control board Ministry of environment & forests, National ambient air quality monitoring Series: NAAQMS//2009-10, New Delhi.
- Available at: http://cpcb.nic.in/upload/NewItems/NewItem_147_report-2008.pdf
- CPCB, 2010. Pollution Control Acts, Rules and Notifications issued thereunder. Pollution Control Law Series: PCL/02/2010, CPCB, Ministry of Environment & Forests, Govt. of India, Parivesh Bhawan, East Arjun Nagar, Delhi- 110032
- Available at: http://www.cpcb.nic.in/NewItem_19_PollutionControlLaw.pdf
- CPCB, 2011. Guidelines for the Measurement of Ambient Air Pollutants Volume-I; Guidelines for Manual Sampling & Analyses, Central Pollution Control Board, Ministry of Environment & Forests, Govt. of India, Parivesh Bhawan, East Arjun Nagar, Delhi- 110032
- Available at: <http://www.cpcb.nic.in/NAAQSMannualVolumeI.pdf>
- de Nevers, N., Morris, J.R., 1975. Rollback Modeling: Basic and Modified. *Journal of the Air Pollution Control Association* 25, 943–947.
- de Nevers, N., Neligan, R.E., Slater, H.H., 1977. Air Pollution. Third Edition, Volume. 6, Stern, A.C., Academic Press. New York.
- Delbari, M., Afrasiab, P., Loiskandl, W., 2009. Using sequential Gaussian simulation to assess the field-scale spatial uncertainty of soil water content. *Catena* 79, 163–169.
- Garg, A., 2011. Pro-equity effects of ancillary benefits of climate change policies: a case study of human health impacts of outdoor air pollution in New Delhi. *World Development* 39, 1002–1025.

George, K.V., Patil, D.D., Kumar, P., Alappat, B.J., 2012. Field comparison of cyclonic separator and mass inertial impactor for PM₁₀ monitoring. *Atmospheric Environment* 60, 247-252.

Georgopoulos, P.G. and Seinfeld, J.H., 1982. Statistical distributions of air pollutant concentrations. *Environmental Science Technology* 16, 401A-416A.

Graedel, T. E., 1977. The Wind Boxplot: An improved Wind Rose. *Journal of Applied Meteorology and Climatology* 16, 448-450.

Hadley, A. and Toumi, R., 2003. Assessing changes to the probability distribution of sulphur dioxide in the UK using a lognormal model. *Atmospheric Environment* 37, 1461-1474.

Heal, M.R., Kumar, P., Harrison, R.M., 2012. Particles, air quality, policy and health. *Chemical Society Reviews* 41, 6606-6630.

Horie, Y., Overton, J., 1974. US Environmental Protection Agency.

Jakeman, A.J., Taylor, J.A., 1989. Identification, estimation, and simulation of frequency distributions of pollutant concentrations for air quality management. In: *Encyclopaedia of Environmental Control Technology (Air Pollution Control – Vol-II)* (edited by Cheremisinoff, P.N.), Gulf Publishing Company, Houston, Texas, USA.

Jakeman, A.J., Simpson, R.W., Taylor, J.A., 1988. Modelling distributions of air pollutant concentrations - III. The hybrid deterministic-statistical distribution approach. *Atmospheric Environment* 22, 163-174.

Jia C.R., D'Souza, J., Batterman, S., 2008. Distributions of personal VOC exposures: A population-based analysis. *Environmental International* 34, 922-931.

Kalpasanov, Y., Kurchatova, G., 1976. A study of the statistical distribution of chemical pollutants in air. *Journal of the Air Pollution Control Association* 26, 981-985.

Kandlikar, M., 2007. Air pollution at a hotspot location in Delhi: Detecting trends, seasonal cycles and oscillations. *Atmospheric Environment* 41(28), 5934-5947.

Kao, A.S., Friedlander, S.K., 1995. Frequency distributions of PM₁₀ chemical components and their sources. *Environmental Science and Technology* 29, 19-28.

Kottegoda, N.T., Rosso, R., 2008. *Applied Statistics for Civil and Environmental Engineers*. Blackwell Publishing Ltd, UK.

Larsen, R., 1961. A method for determining source reduction required to meet air quality standards. *Journal of the Air Pollution Control Association* 11, 71-76.

Larsen, R., 1969. A new mathematical model for air pollutant concentration averaging time and frequency. *Journal of the Air Pollution Control Association* 19, 24-30.

Larsen, R., 1971. A mathematical model for relating air quality measurements to air quality standards. Publication No. AP-89, USEPA, Research Triangle Park, NC, USA.

Le, H.Q., Batterman, S., Godwin, C., 2007. Reproducibility and imputation of air toxics data. *Journal of Environment Monitoring* 9, 1358-1372.

Lin, Y.P., Chen, B.Y., Shyu, G.S., Chang, T.K., 2010. Combining a finite mixture distribution model with indicator kriging to delineate and map the spatial patterns of heavy metal pollution in soil. *Environmental Pollution* 158, 235-244.

Lin, Y.P., Chu, H.J., Wang, C.L., Yu, H.H., Wang, Y.C., 2009. Remote Sensing Data with the conditional Latin hypercube sampling and geostatistical approach to delineate landscape changes induced by large chronological physical disturbances, *Sensors* 9, 48-174.

475 Lonati, G., Cernuschi, S., Giugliano, M., 2011. The duration of PM₁₀ concentration in a large
 476 metropolitan area. *Atmospheric Environment* 45, 137–146.
 477 Lu, H.C., 2002. The statistical characters of PM₁₀ concentrations in Taiwan area.
 478 *Atmospheric Environment* 36, 491-502.
 479 Mage, D.T., 1984. Pseudo-lognormal distributions. *Journal of Air Pollution and Control*
 480 *Association* 31, 374-376.
 481 Mage, D.T., Ott, W.R., 1984. An evaluation of the method of frcatiles, moments and
 482 maximum likelihood for estimating parameters when sampling air quality data from a
 483 stationary lognormal distribution. *Atmospheric Environment* 18, 163-171.
 484 Mijic, Z., Tasic, M., Rajšic, S., Novakovic, V., 2009. The statistical characters of PM₁₀ in
 485 Belgrade area. *Atmospheric Research* 92, 420–426
 486 Minitab Inc., 2003. Minitab Statistical Software, Release 14 for Windows. State College,
 487 Pennsylvania.
 488 Mohamed, N., Abdullah, M.M.A., Tan, C., Ramli, N.A., Yahaya, A.S., Fitri, N.F.M.Y., 2011.
 489 Modelling of PM₁₀ concentration for industrialized area in Malaysia: A case study in
 490 Shah Alam. *Physics Procedia* 22, 318–324.
 491 Morel, B., Yen, S., Cifuentes, L., 1999. Statistical distribution for air pollutant applied to the
 492 study of the particulate problem in Santiago. *Atmospheric Environment* 33, 2575-2585.
 493 Ott, W.R., 1995. *Environmental Statistics and Data Analysis*. Lewis Publishers.
 494 Pandey, J.S., Kumar, R., Devotta, S., 2005. Health risks of NO₂, SPM and SO₂ in Delhi
 495 (India). *Atmospheric Environment* 39, 6868–6874.
 496 Pollack, R.I., 1975. *Studies of Pollutant Concentration Frequency Distributions*. USEPA,
 497 Publication No. 650/4–75–004, Research Triangle Park, North Carolina.
 498 Savoie, D. L., Prospero, J. M., 1977. Aerosol Concentration Statistics for the Northern
 499 Tropical Atlantic. *Journal of Geophysical Research* 82, 5954-5964.
 500 Savoie, D.L., Prospero, J.M., Nees, R.T., 1987. Nitrate, non-sea-salt sulfate, and mineral
 501 aerosol over the northwestern Indian Ocean. *Journal of Geophysical Research* 92, 933-
 502 942.

503 **List of figure captions**

504 **Figure 1.** Probability density function for a NAAQS complying pollutant - The vertical
505 reference line corresponds to the NAAQS for 24-hourly average PM_{10} concentrations.

506 **Figure 2.** Location of ambient air quality monitoring stations in Delhi

507 **Figure 3.** Probability plots for SO_2 at (a) Ashok Vihar, (b) Janak Puri (c) Nizamuddin (d)
508 Shahadra, (e) Shahzada Bagh and (f) Sirifort monitoring stations.

509 **Figure 4.** Probability plots for NO_2 at (a) Ashok Vihar, (b) Janak Puri (c) Nizamuddin (d)
510 Shahadra, (e) Shahzada Bagh and (f) Sirifort monitoring stations.

511
512 **Figure 5.** Probability plots for SPM at (a) Ashok Vihar, (b) Janak Puri (c) Nizamuddin
513 (d) Shahadra, (e) Shahzada Bagh and (f) Sirifort monitoring stations.

514
515 **Figure 6.** Probability plots for PM_{10} at (a) Ashok Vihar, (b) Janak Puri (c) Nizamuddin
516 (d) Shahadra, (e) Shahzada Bagh and (f) Sirifort monitoring stations.

517
518 **Figure 7.** The best fit pdf for (a) SO_2 at Sirifort (complying NAAQS), and (b) PM_{10} at
519 Nizamuddin (non-complying NAAQS) monitoring station; the NAAQS standard is shown
520 with the vertical line.

List of tables

Table 1. Description of the NAAQM stations and type of pollutants measured at each of them. Details of instruments used to measure various pollutants can be found in SI Table S.1.

Monitoring station	Location	Type of area	Monitoring height	Parameters monitored
Ashok Vihar	North-North West	Residential	5 m	SPM, PM ₁₀ , NO ₂ , SO ₂
Shazada Bagh	North-West	Industrial	10 m	SPM, PM ₁₀ , NO ₂ , SO ₂
Nizamuddin	South-East	Mixed use	10 m	SPM, PM ₁₀ , NO ₂ , SO ₂
Janak Puri	South-West	Residential	10 m	SPM, PM ₁₀ , NO ₂ , SO ₂
Shahadra	North-East	Industrial	10m	SPM, PM ₁₀ , NO ₂ , SO ₂
Sirifort	South	Residential	5 m	SPM, PM ₁₀ , CO, NO ₂ , SO ₂

526 **Table 2.** Properties of statistical distributions used for fitting ambient air quality data (source:
527 Minitab Inc., 2003).
528

Statistical distribution	Probability density function	Parameters			
		Location	Scale	Shape	Threshold
Normal	$\frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$	μ	σ	N.A.	N.A.
Lognormal	$\frac{1}{\alpha x \sqrt{2\pi}} \exp\left[-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right]$	μ	σ	N.A.	N.A.
Three-parameter lognormal	$\frac{1}{\sqrt{2\pi}\sigma(x-\lambda)} \exp\left[-\frac{(\ln(x-\lambda)-\mu)^2}{2\sigma^2}\right]$	μ	σ	N.A.	λ
Exponential	$\frac{1}{\theta} \exp\left[-\frac{x}{\theta}\right]$	N.A.	θ	N.A.	N.A.
Two-parameter exponential	$\frac{1}{\theta} \exp\left[-\frac{(x-\lambda)}{\theta}\right]$	N.A.	θ	N.A.	λ
Weibull	$\frac{\beta}{\alpha^\beta} (x)^{\beta-1} \exp\left[-\left(\frac{x}{\alpha}\right)^\beta\right]$	N.A.	β	α	N.A.
Three-parameter Weibull	$\frac{\beta}{\alpha^\beta} (x-\lambda)^{\beta-1} \exp\left[-\left(\frac{x-\lambda}{\alpha}\right)^\beta\right]$	N.A.	β	α	λ
Gamma	$\frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right)$	N.A.	β	α	N.A.
Three-parameter gamma	$\frac{1}{\Gamma(\alpha)\beta^\alpha} (x-\gamma)^{\alpha-1} \exp\left(-\frac{x-\gamma}{\beta}\right)$	N.A.	β	α	λ
Logistic	$\frac{\exp\left(\frac{x-\mu}{\sigma}\right)}{\sigma \left[1 + \exp\left(\frac{x-\mu}{\sigma}\right)\right]^2}$	μ	σ	N.A.	N.A.
Loglogistic	$\frac{\exp\left(\frac{\ln x - \mu}{\sigma}\right)}{\sigma \left[1 + \exp\left(\frac{\ln x - \mu}{\sigma}\right)\right]^2}$	μ	σ	N.A.	N.A.
Three-parameter loglogistic	$\frac{\exp\left(\frac{\ln(x-\lambda) - \mu}{\sigma}\right)}{\sigma \left[1 + \exp\left(\frac{\ln(x-\lambda) - \mu}{\sigma}\right)\right]^2}$	μ	σ	N.A.	λ

529

Table 3. Best fit *pdf* models for the pollutants monitored at the six NAAQM monitoring station (24-hourly average) with estimated parameters and goodness-of-fit test.

Pollutant	Station	<i>pdf</i> model	ML estimates of distribution parameters				Goodness-of-fit test	
			Location	Shape	Scale	Threshold	A-D	p-value
SO ₂	Ashok Vihar	Loglogistic	2.02	N.A.	0.17	N.A.	3.47	<0.005
	Janak Puri	Weibull	N.A.	4.19	11.38	N.A.	3.88	<0.010
	Nizamuddin	Weibull	N.A.	4.30	11.57	N.A.	2.08	<0.010
	Shahadra	Lognormal	2.24	N.A.	0.36	N.A.	1.84	<0.005
	Shahzada Bagh	Lognormal	2.08	N.A.	N.A.	0.35	3.79	<0.005
	Sirifort	Weibull	N.A.	3.66	8.77	N.A.	3.58	<0.010
NO ₂	Ashok Vihar	Normal	40.34	N.A.	12.10	N.A.	0.43	0.3
	Janak Puri	Weibull	6.01	N.A.	49.27	N.A.	0.48	0.24
	Nizamuddin	Logistic	45.30	N.A.	4.91	N.A.	0.68	0.046
	Shahadra	Gamma	N.A.	13.54	2.80	N.A.	0.36	>0.25
	Shahzada Bagh	Lognormal	3.77	N.A.	0.28	N.A.	0.46	0.25
	Sirifort	Loglogistic	3.52	N.A.	0.15	N.A.	0.42	>0.25
SPM	Ashok Vihar	Gamma	N.A.	4.56	72.06	N.A.	0.55	0.18
	Janak Puri	Gamma	N.A.	5.42	61.53	N.A.	0.73	0.06
	Nizamuddin	Lognormal	5.66	N.A.	0.46	N.A.	0.48	0.24
	Shahadra	Lognormal	5.76	N.A.	0.48	N.A.	0.41	0.33
	Shahzada Bagh	Gamma	N.A.	7.53	46.75	N.A.	0.52	0.21
	Sirifort	Gamma	N.A.	4.53	70.82	N.A.	0.27	>0.25
PM ₁₀	Ashok Vihar	Lognormal	4.71	N.A.	0.53	N.A.	0.34	0.50
	Janak Puri	Gamma	N.A.	4.35	32.60	N.A.	0.29	>0.25
	Nizamuddin	Gamma	N.A.	4.68	25.80	N.A.	0.22	>0.25
	Shahadra	Lognormal	4.81	N.A.	0.51	N.A.	0.29	0.62
	Shahzada Bagh	Gamma	N.A.	5.56	26.01	N.A.	0.35	>0.25
	Sirifort	Lognormal	4.64	N.A.	0.56	N.A.	0.35	<0.46

Table 4. Probability of exceedence, comparison of number of violations of NAAQS in the following year and return period for various NAAQM monitoring stations for different pollutants.

Station	Probability	Pollutants			
		SO ₂	NO ₂	SPM	PM ₁₀
Ashok Vihar	Pr ($X \leq x_{\text{NAAQS}}$) ¹ (in %)	100	99.95	20.67	42.46
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Predicted)	0	0.05	79.33	57.54
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Observed)	0	0	80.30	42.20
	NAAQS criteria being met (Yes/No)	Yes	Yes	No	No
	Number of days exceeding NAAQS in the following year (predicted)	0	0	56	37
	Number of days exceeding NAAQS in the following year (observed)	0	0	57	27
	Return period (in days)	∞	∞	1.26	2.37
Janak Puri	Pr ($X \leq x_{\text{NAAQS}}$) ¹ (in %)	100	100	18.10	30.01
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Predicted)	0	0	81.90	69.99
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Observed)	0	0	93.70	77
	NAAQS criteria being met (Yes/No)	Yes	Yes	No	No
	Number of days exceeding NAAQS in the following year (predicted)	0	0	52	43
	Number of days exceeding NAAQS in the following year (observed)	0	0	59	47
	Return period (in days)	∞	∞	1.22	1.43
Nizamuddin	Pr ($X \leq x_{\text{NAAQS}}$) ¹ (in %)	100	99.92	21.70	40.49
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Predicted)	0	0.08	78.30	59.51
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Observed)	0	0	81.60	54.70
	NAAQS criteria being met (Yes/No)	Yes	Yes	No	No
	Number of days exceeding NAAQS in the following year (predicted)	0	0	60	38
	Number of days exceeding NAAQS in the following year (observed)	0	0	62	35
	Return period (in days)	∞	∞	1.28	1.68

¹24-hourly/8-hourly values should be met 98% of time in a year.

²However, 2% of time, it may exceed but not on two consecutive days.

542 **Table 4.** (contd.)
543

Station	Probability	Pollutants			
		SO ₂	NO ₂	SPM	PM ₁₀
Shahadra	Pr ($X \leq x_{\text{NAAQS}}$) ¹ (in %)	100	100	83.07	65.25
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Predicted)	0	0	16.93	34.75
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Observed)	0	0	24.70	43.10
	NAAQS criteria being met (Yes/No)	Yes	Yes	No	No
	Number of days exceeding NAAQS in the following year (predicted)	0	0	13	25
	Number of days exceeding NAAQS in the following year (observed)	0	0	19	31
	Return period (in days)	∞	∞	5.9	2.88
Shahzada Bagh	Pr ($X \leq x_{\text{NAAQS}}$) ¹ (in %)	100	99.98	87.32	65.26
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Predicted)	0	0.02	12.68	34.74
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Observed)	0	0	28.10	53.30
	NAAQS criteria being met (Yes/No)	Yes	Yes	No	No
	Number of days exceeding NAAQS in the following year (predicted)	0	0	7	21
	Number of days exceeding NAAQS in the following year (observed)	0	0	16	32
	Return period (in days)	∞	∞	7.90	2.90
Sirifort	Pr ($X \leq x_{\text{NAAQS}}$) ¹ (in %)	100	99.65	22.12	47.74
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Predicted)	0	0.35	77.88	52.26
	Pr ($X \geq x_{\text{NAAQS}}$) ² (in %) (Observed)	0	0	78.40	57.10
	NAAQS criteria being met (Yes/No)	Yes	Yes	No	No
	Number of days exceeding NAAQS in the following year (predicted)	0	0	58	33
	Number of days exceeding NAAQS in the following year (observed)	0	0	58	36
	Return period (in days)	∞	∞	1.28	1.91

544 ¹24-hourly/8-hourly values should be met 98% of time in a year.

545 ²However, 2% of time, it may exceed but not on two consecutive days.

546

Table 5. The emission source reduction required to meet NAAQS for different pollutants at various NAAQM stations

Station/ ↓ Pollutants →	SO ₂	NO ₂	SPM			PM ₁₀		
	R (%)	R (%)	E{C _p } (µg/m ³)	E{C} _s (µg/m ³)	R (%)	E{C _p }	E{C} _s (µg/m ³)	R (%)
Ashok Vihar	0	0	328.80	35.87	89.10	127.86	36.02	71.83
Janak Puri	0	0	320.61	41.92	86.92	120.21	38.12	82.09
Nizamuddin	0	0	333.61	83.98	74.82	141.82	20.27	85.71
Shahzada Bagh	0	0	319.29	80.96	74.46	120.89	22.93	81.03
Shahadra	0	0	353.40	214.79	39.22	139.31	61.46	55.88
Sirifort	0	0	351.90	243.31	30.86	144.60	53.82	62.78

R = The emission source reduction (in %) required to meet NAAQS

E{c_p} = expected concentration of present distribution.

E{c}_s = expected concentration of a distribution such that the extreme value for which corresponds to c_s.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Citation details:

Sharma, P., Sharma, P., Jain, S., Kumar, P., 2013. An integrated statistical approach for evaluating the exceedence of criteria pollutants in the ambient air of megacity Delhi. *Atmospheric Environment* 70, 7-17.

Web link: <http://dx.doi.org/10.1016/j.atmosenv.2013.01.004>