WORKING PAPERS

EAST-WEST CENTER WORKING PAPERS

EAST-WEST CENTER

The U.S. Congress established the East-West Center in 1960 to foster mutual understanding and cooperation among the governments and peoples of the Asia Pacific region including the United States. Funding for the Center comes from the U.S. government with additional support provided by private agencies, individuals, corporations, and Asian and Pacific governments.

*East-West Center Working Papers* are circulated for comment and to inform interested colleagues about work in progress at the Center.

For more information about the Center or to order publications, contact:

Publication Sales Office
East-West Center
1601 East-West Road
Honolulu, Hawaii 96848-1601

Telephone: 808-944-7145
Facsimile: 808-944-7376
Email: ewcbooks@EastWestCenter.org
Website: www.EastWestCenter.org

EAST-WEST CENTER WORKING PAPERS

EAST-WEST CENTER

Environmental Change, Vulnerability, and Governance Series

No. 53, December 2002

# Determining Spatial Patterns in Delhi's Ambient Air Quality Data Using Cluster Analysis

**Sumeet Saksena, Veena Joshi, and R.S. Patil**

**Sumeet Saksena is a Fellow** in the Research Program at the East-West Center.

**Veena Joshi is Head of the Energy Program** at the Swiss Development Cooperation in New Delhi.

**R.S. Patil is a Professor** at the Centre for Environmental Science and Engineering, Indian Institute of Technology in Mumbai.

*East-West Center Working Papers: Environment Series* is an unreviewed and unedited prepublication series reporting on research in progress. The views expressed are those of the authors and not necessarily those of the Center. Please direct orders and requests to the East-West Center's Publication Sales Office. The price for Working Papers is $3.00 each plus postage. To destinations within the U.S. and its territories: for surface mail, add $3.00 for the first paper plus $0.75 for each additional title or copy sent in the same shipment; for airmail add $4.00 for the first paper plus $1.25 for each additional title or copy in the same shipment. To destinations elsewhere: for surface mail, add $6.00 for the first paper plus $1.50 for each additional title or copy sent in the same shipment; for airmail, add $7.50 for the first paper plus $3.50 for each additional title or copy in the same shipment.

**Determining spatial patterns in Delhi=s ambient air quality data using cluster analysis**

S. Saksena[1], V. Joshi[2] and R.S. Patil[3]

[1] Author to whom correspondence should be addressed. East West Center, Honolulu, 1601 East West Road, Honolulu, USA; fax: 808-944-7298; email: saksenas@eastwestcenter.org

[2] Swiss Development Cooperation, Chanakyapuri, P.O. Box 392, New Delhi - 110 021, India

[3] Centre for Environmental Sciences and Engineering, Indian Institute of Technology, Powai, Mumbai - 400 076, India

## Summary

The purpose of this study was to study the spatial patterns of ambient air quality Delhi in the absence of extensive datasets needed for space-time modeling. A spatial classification was attempted on the basis of ambient air quality data of nine years (1998 is latest year for which published data was available) for three criteria pollutants - nitrogen dioxide, sulphur dioxide, and suspended particulate matter. A hierarchical agglomerative algorithm using the average linkage between groups method and the Euclidean distance metric was used. Cluster analysis indicated that till 1998, by and large, two distinct classes existed. The results of cluster analysis prompted an investigation of systematic biases in the monitored data. No statistically significant differences in the mean concentration of all pollutants were observed between stations belonging to different land-use types (residential and industrial). This fact would be useful, if and when the authorities consider modifying the network or expanding it in Delhi. The results also support the recommendation that Delhi have a uniform standard across all areas. This study has provided a methodology for Indian researchers, practitioners, and regulatory authorities to do an exploratory

study of spatial patterns of air pollution and data quality issues in Indian cities using the National Ambient Air Quality Monitoring System data.

Keyword index: spatial patterns, exposure classification, network analysis, criteria pollutants, quality control

**Cluster analysis of Delhi=s ambient air quality data**

**Introduction and aim**

Air pollution monitoring in India began in 1978, when National Environmental Engineering Research Institute (NEERI) established a network in 10 cities including Delhi. In each city, at least three sampling sites representing residential, commercial and industrial zones were selected (NEERI[1]). Many years later the Central Pollution Control Board (CPCB) also began a monitoring programme. The main function of the CPCB under the Air (Prevention & Control of Pollution) Act 1981 is to improve the quality of air and to prevent, control and abate air pollution in the country. In order to assess the baseline situation, measure trends and evaluate the impact of interventions, it is essential to collect air quality data. Therefore, a National Ambient Air Quality Monitoring (NAAQM) programme was initiated in 1984 with 7 stations at Agra and Annapara. By the end of March 1993, 290 monitoring stations covering 92 cities/towns were operational. There has been no expansion of the network since then. The monitoring stations have been located in different areas, viz., Residential/Rural (R), Industrial (I), and Sensitive (S) to cover the spatial variation in different cities. The last category includes monuments, hill stations, and sanctuaries. The ambient air quality standards that were notified in 1994 are also specific to this locational classification.

Currently, only the criteria pollutants, i.e. Sulphur dioxide ($SO_2$), Nitrogen dioxide ($NO_2$) and Suspended Particulate Matter (SPM) are regularly measured. As per the programme, monitoring for $SO_2$ and $NO_2$ should be conducted at 4-hourly intervals for 24 hours and for SPM at 8-hourly intervals for 24 hours on a bi-weekly basis (CPCB[2]). $SO_2$ is measured by the Improved West & Gaeke method; $NO_2$ by the Jaicob and Hocheiser modified method; and SPM by the Gravimmetric method. (Earlier, SPM was collected continuously on a 24h basis, and gases on a 8 hourly batch basis, sequentially for 24 hours.). In addition, meteorological parameters like wind speed, wind direction, temperature and relative humidity are also measured at each station. Advanced statistical tests have been used to analyze the data (NEERI[3]) B for example, Daniel=s test based on Spearman's rank correlation coefficients for studying yearly variations.

Cluster analysis is the art of finding groups in data (Kaufman and Rousseeuw[4]). Its theoretical basis does not require making assumptions about the mutual independence of samples. It has been used to design and analyse air quality monitoring networks; and for source identification. Sabaton[5] studied the 24-hourly concentrations of $SO_2$ measured at thirty stations in metropolitan Paris, using cluster analysis, aiming at reducing network density with a minimum loss of information. The Indian Standards Institution (now known as Bureau of Indian Standards) has also recommended the use of cluster analysis for network design (ISI[6]). Their recommendations are in turn based on recommendations of Munn[7] . Ibarra *et al.* [8] used cluster analysis to optimize an air quality network in Spain.

The most common application of cluster analysis for aerometric data relates to source identification. Lin, Lee and Chang[9] investigated the characteristics of acid precipitation collected in Taipei by performing cluster analysis. They were able to identify three potential sources of water soluble ions in rainwater. Miranda *et al.*[10] studied the concentration of elements at two sites in Mexico City using cluster analysis and identified sources. They used the Pearson distance metric and the Wald algorithm for this purpose. A similar analysis was done in Chile (Romo-Groger *et al.*[11]) but using the unweighted pair-group average linkage method. While these studies used the hierarchical agglomerative method, the study by Sanchez Gomez and Ramos Martin[12] used the non-hierarchical K-means method to identify sources in Valladoid, Spain. They concluded that it enables the identification of sources in a similar manner to R-techniques (such as factor analysis). The same method was used by Saucy *et al.*[13] to identify sources of aerosols in the Norwegian Arctic. In addition to considering only pollutant data many studies have also incorporated wind data for cluster analysis (Dorling, Davies, Pierce[14]; Sanchez *et al*.[15]; Ruijgrok and Romer[16]).

Cluster analysis of time series data has also been done to detect temporal patterns (Bohm *et al*.[17]; Lavecchia *et al.*[18]) of ozone. The relationship between emissions and air quality has been studied by Rua, Gimeno, and Hernandez[19,20] using cluster analysis.

Cluster analysis has also been used to group sampling sites with similar characteristics. Wongphatarakul, Friedlander and Pinto[21] studied $PM_{2.5}$ chemical databases from seven sites around the world. Cluster analysis was on of three methods used for this purpose. They observed consistency in results using three different linkage methods, thereby increasing the confidence in cluster analysis as an appropriate tool for such applications. They also concluded that clustering methods provide an initial estimate of source profiles for poorly characterized sampling sites. If a poorly characterized site merges with a group of sampling sites of known source type or types, the poorly characterized site is likely to have similar chemical characteristics or source types.

In Delhi, there are 9 stations, shown in Table 1; 6 operated by CPCB and 3 by the NEERI. Land-use patterns have changed significantly since CPCB adopted the land-use based classification system for its monitoring stations. More importantly, heavy traffic - an ubiquitous phenomenon - would contribute to air pollution even in non-industrial areas. Therefore, it is not safe to assume that ambient levels in industrial areas would necessarily always be higher than in commercial areas, with residential areas being the least polluted.

The objective of our study was to determine if Delhi's sampling sites can be grouped together in any manner. Cluster analysis was used for this in the absence of adequate date for sophisticated space-time modeling. Also, for a city of its size, the sampling network is not dense enough for space-time modeling. This study has provided a simple methodology for Indian researchers, practitioners, and regulatory authorities to do an exploratory study of spatial patterns and/or data quality issues of air pollution in Indian cities using the National Ambient Air Quality Monitoring System data in the absence of the huge data sets needed for more sophisticated space-time modeling. The results have serious implications in terms of improving the monitoring network, sampling strategies and data quality. The results also have a regulatory significance in terms of moving towards a uniform ambient air quality standard for cities such as Delhi.

**Description of study region**

Delhi is the fourth most polluted mega-city of the world vis-a-vis air quality. The city is spread over 1483 $km^2$ (47% urban, 53% rural). The Delhi Metropolitan Area that included satellite towns covers an area of 3182 $km^2$. More than three quarters of the emissions of air pollutants are caused by vehicles. The total number of vehicles in 1995 was 2.4 million (22% cars and 67% two-wheelers). This is more than a nine-fold increase over the total number of vehicles in 1971. Delhi has three big thermal power stations - all coal based. While two of these are located within the city, the third is on the outskirts.

Delhi is not as industrialized as Bombay or Calcutta. Factories are mainly located in west, north-west or east Delhi. There are certain non-polluting industrial estates in south Delhi. Industrial air pollution in Delhi is mainly due to the very large number of small-scale industries. Delhi has many satellite towns which are industrialized to a large extent. Major infrastructural units and large scale industries are located in this region. Delhi and its satellite towns form the National Capital Region - a large area source of pollution. It is estimated that there are about 65,000 industries in Delhi, of which 173 are large scale industries.

## Methodology

*Preparing for the data analysis*

We used the monthly mean concentration of three criteria pollutants B SO2, NO2, and SPM. The data were obtained from published sources (CPCB[2, 22-28]). Missing values were substituted. If, for example, for a particular month the value was missing, then it was substituted by taking the average of the preceding and succeeding months. This was done to preserve the seasonal patterns (as opposed to the effect of the procedure of substituting by the annual average.). In some cases where no value was reported for many consecutive months, the data corresponding to these months for all nine stations were dropped for that year=s analysis. Thus, for 1990 we have not considered the September and October data; for 1991 and 1992 all data were available; for 1993 we did not consider the June, July, August and October data (because NEERI did not report value for these months for its 3 stations); and for 1994 July, August, and September data were not included (again because of non-reporting by NEERI). In addition, data pertaining to the Netaji Nagar (NEERI) station was entirely excluded from analysis for the year 1994 because of missing data for four months. The 1995 data had the least number of missing values.

*Cluster analysis*

Cluster analysis involves splitting a data set into a number of groups of observations which are distinct in terms of typical group values of the variables. The aim is to maximize between-group variance and to minimize within-group variance. Cluster analysis is a classification technique where any number of variables may be used to classify members of the sample.

Classification was done separately for each pollutant, viz., SPM, $SO_2$ and $NO_2$, though classification based on combinations of two or three pollutants at a time is also possible. Using monthly means for the nine stations meant that we worked on a 9 x 12 matrix for each pollutant for each year. We used a hierarchial agglomeration algorithm for clustering. Two distance measures (Euclidean and Squared Euclidean) and four methods for combining clusters ( average linkage between groups, single linkage, complete linkage, and centroid method) were initially used. In addition to using the data as it is, we transformed them prior to analysis in three ways: z-scores of data, log-transformed data, and z-scores of the log-transformed data.

The most robust and consistent results were obtained with log-transformed data, the Euclidean distance and between pairs average linkage method. It is this combination of methods that most often lead to the formation of distinct groups. There is a justification for transforming data because log-normality of aerometric data has been established by many studies (Ott[29]; WHO[30]). Kaufman and Rousseeuw[4] recommend that when using ratio-scale variables (such as concentrations) the data be log-transformed before analysis. There was no need to further transform the data because each classification criterion (parameter/variable) in this case is measured in the same units and is basically the same quantity measured in different months.

The change in the distance coefficient between adjacent steps in the hierarchical algorithm was the basis of deciding the number of distinct clusters that existed in the data. A software package - SPSS for Windows version 9.0 - was used for the analysis.

**Results and discussion**

With SPM as the criterion, it was observed that two distinct clusters existed in the years 1996 – 1998 (Figure 1). Najafgarh and Town Hall were members of a common cluster in all these years. Netaji Nagar was member in 1998, but was the nearest neighboring cluster in 1996 & 1997. No clustering effect was observed in 1995.

It is evident from Figures 2 that with respect to $SO_2$ two clusters existed in all years. Netaji Nagar was always in the distinct cluster – on its own in 1996 and 1998 and with Siri Fort in 1997. However, in the initial years (1990 -1993) three stations (Town Hall, Najafgarh, and Netaji Nagar - all operated by NEERI) were clustered together. In the later years Town Hall (an area characterized by high traffic density) emerged as a distinct cluster in 1994 and in 1995 it was clustered with Najafgarh.

Using $NO_2$ as a clustering criteria we observed that in all years two distinct clusters were formed (Figure 3). The three NEERI stations were clustered together in 1998 and 1997. Before 1997 Town Hall formed a cluster of its own.

The frequent presence of the Netaji Nagar and Town Hall stations (classified as residential areas) in the same cluster as the Najafgarh station (classified as an industrial area) forms the basis of a hypothesis that a systematic bias exists between the stations operated by NEERI and CPCB. Figures 4 to 6 show the frequency distribution of the pollutant levels over the years. It is very clear that median value (50th percentile) for all pollutants measured by CPCB monotonously increased till 1994 and then on remained fairly constant. In the case of the NEERI stations the yearly fluctuations in the median value do not indicate any trend or stability. These patterns for both sets of stations were observed even when an analysis was done separately for industrial and residential areas. We are unable to speculate about the source of the systematic bias.

A non-parametric test (Mann-Whitney U test) indicated that for SPM a significant difference in the means of the two sets of stations (NEERI and CPCB) existed in all years except during 1995 -1997. On the basis of $SO_2$ a significant difference existed in all years except 1992, and on the basis of $NO_2$ a significant difference existed in all years except 1992 and 1997.

We have plotted the yearly variation of the coefficient of variation (CV), also known as Relative Standard Deviation (RSD) in Figures 7-9. This procedure is similar in principle to plotting control charts for ascertaining the quality of experimental data (USEPA[31]). The control chart provides a tool for distinguishing the pattern of random variation from the systematic variation (Lodge[32]). Certainly seasonal effects and variations in source emissions would change the CV, but all monitoring station are expected to be equally affected by these changes. We observed that in the CPCB stations the CV decreased over the years for all pollutants, while in the NEERI stations the trend in CV was erratic. This pattern was observed even when the data were disaggregated by land-use type.

We repeated the cluster analysis for 1998 using only data from the six CPCB stations.

For SPM, we observed that Janak Puri and Siri Fort – the two residential areas - were clustered together, and their combined annual mean was higher than that of the combined mean of the four industrial area stations. We speculate that this may be due to their being downwind of sources of SPM. For $SO_2$, two clusters were observed, with Shahzada Bag forming a group of its own. For $NO_2$ also two clusters were formed, but Ashok Vihar formed a group of its own.

With respect to SPM no significant difference in the means was observed (using the Mann Whitney test) between stations classified as residential and industrial in the period 1990 - 1998. For $SO_2$ there was no difference till 1996. However, for $NO_2$ there did exist a significant difference in the means between these land-use types in 1990, 1991, 1995 and 1996.

Results of both cluster analysis and non-parametric tests suggest that in Delhi there is generally no difference in pollutant levels across land-use types. We speculate that this may be due to the fact that industrial activity has gradually moved out of the city owing to government directives and the build up of commercial areas near and within residential areas (thereby increasing vehicular emissions in residential areas). This fact would be useful, if and when the authorities consider modifying the network or expanding it in Delhi. It is recommended that similar analysis be done using PM10 data when it becomes available. Owing to possibly large background levels of SPM spatial patterns may not get easily revealed. The results also provide some support for the recommendation that for Indian cities like Delhi there must be a single ambient air quality standard, rather than standards classified based on land-use. In the future researchers may also wish to do a cluster analysis based on daily data (not yet available in the public domain in India) and to do this seasonally. It is very likely that spatial patterns of peak levels are quite different from the spatial patterns of grossly averaged levels.

The purpose of this study was to study spatial patterns of air pollution in Delhi over a ten year period using three criteria pollutants. Cluster analysis indicated that till 1998, by and large, two distinct classes existed. The results provided an evidence of a systematic difference between measurements made by the two agencies involved in monitoring air quality. This may be because of a difference in operating and calibration protocols used by different agencies. This is more apparent in the case of $SO_2$ and $NO_2$ measurements, than in the case of SPM measurements (a comparatively simpler pollutant to measure).

This study has provided a methodology for Indian researchers, practitioners, and regulatory authorities to do an exploratory study of spatial patterns and/or data quality issues of air pollution in Indian cities using the National Ambient Air Quality Monitoring System data in the absence of the huge data sets needed for more sophisticated space-time modeling.

**References**


1 NEERI, *Air quality in selected cities in India 1978-79,* National Environmental Engineering Research Institute, 1980, Nagpur.

2 CPCB, **National ambient air quality statistics of India , 1991,** NAAQMS/4/1992-93 Central Pollution Control Board, 1992, New Delhi.

3 NEERI, *Air quality in selected cities in India, 1980-81,* National Environmental Engineering Research Institute, 1983, Nagpur.

4 Kaufman L and Rousseeuw P J, **Finding groups in data,** 1990,Wiley Interscience, New York.

5 Sabaton, C, `Etude d'optimisation d'un reseau de surveillance de la pollution atmospherique doms la region Parissienne', As quoted in: **Atmospheric Pollution**, Benarie M M (ed), 1976, Amsterdam, Elsevier. pp. 159-171.

6 ISI, *Methods for measurement of air quality: Part 14 Guidelines for Planning the sampling of atmosphere and location of monitoring stations. IS: 5182 (Part 14) – 1985,* Indian Standards Institute, 1985, New Delhi.

7 Munn R E, **The design of air quality monitoring networks,** MacMillan Publishers Ltd, 1981, London.

8 Ibarra G, Alonso L, Navazo M, et al.,Application of cluster analysis in optimization process of surface meteorological and air quality networks, Proceedings of the 2nd International Conference on Air Pollution, 1994, Part 1, Barcelona, Spain, Wessex Institute of Technology.

9 Lin N H, Lee H M and Chang M B, Evaluation of the characteristics of acid precipitation in Taipei, Taiwan using cluster analysis, *Water, Air and Soil Pollution,* 1999*,* **113**(1): 241-260

10 Miranda J, Cahill T A, Morales J, et al., Determination of elemental concentrations in atmospheric aerosols in Mexico City using proton induced X-ray emission, proton elastic scattering, and laser absorption, *Atmospheric Environment*, 1994, **28**(14): 2299-2306.

11 Romo-Groger C M, Morales J R, Dinator M I, Llona F, Heavy metals in the atmosphere coming from a copper smelter in Chile, *Atmospheric Environment*, 1994, **28**(4):705-711.

12 Sanchez Gomez M L and Ramos Martin M C, Application of cluster analysis to identify sources of airborne particles, *Atmospheric Environment*, 1987, **21(**7): 1521-1527.

13 Saucy D A, Anderson J R, Buseck P R, Cluster analysis applied to atmospheric aerosol samples from the Norwegian Arctic, *Atmospheric Environment,* 1987, **21(**7): 1649-1657

14 Dorling S R, Davies T D, Pierce C E, Cluster analysis: a technique for estimating the synoptic meteorological controls on air and precipitation chemistry - method and applications, *Atmospheric Environment,* 1992, **26A**(14):2575-2581.

15 Sanchez M L, Pascual D, Ramos C, Perez, I, Forecasting particulate pollutant concentrations in a city from meteorological variables and regional weather patterns. *Atmospheric Environment*, 1990, **26A**(14):1509-1519.

16 Ruijgrok W and Romer F G, Aspects of wet, acidifying deposition in Arnhem: source regions, correlations, and trends, *Atmospheric Environment*, 1993, **27A**(5):637-653.

17 Bohm M, McCune B and Vandetta, T, Diurnal curves of tropospheric ozone in the western United States, *Atmospheric Environment,* 1991, **25A**(8):1570-1590.

18 Lavecchia C, Angelino E, Bedogoni M, *et al.*, Ozone patterns in the aerological basin of Milan (Italy), *Environmental Software,* 1996, **11**(1-3):73-80

19 Rua A, Gimeno L, and Hernandez E, Relationship between air pollutants emission patterns and rainwater acidity, *Toxicological and Environmental Chemistry,* 1997, **59**(1-4):199-207.

20 Rua A, Gimeno L, and Hernandez E, Relationship between air pollutants emission patterns and concentrations, *Toxicological and Environmental Chemistry,* 1997, **59**(1-4): 189-197.

21 Wongphatarakul V, Friedlander S K, and Pinto J P, A comparative study of $PM_{2.5}$ ambient aerosol chemical databases, *Environmental Science and Technology,* 1998, **32**: 3926-3934.

22 CPCB, **National ambient air quality statistics of India, 1990,** NAAQMS/3/1991-92 Central Pollution Control Board, 11991, New Delhi.

23 CPCB, **National ambient air quality statistics of India , 1992,** NAAQMS/6/1994-95 Central Pollution Control Board, 1994, New Delhi.

24 CPCB, **Ambient air quality** B **status and statistics (1993 & 1994),** NAAQMS/7/1995-96, Central Pollution Control Board, 1995, New Delhi.

25 CPCB, **Ambient air quality** B **status and statistics: 1995,** NAAQMS/8/1996-97, Central Pollution Control Board, 1997, New Delhi.

26 CPCB, **Ambient air quality status: 1996,** NAAQMS/10/1998-99, Central Pollution Control Board, 1998, New Delhi. 115pp.

27 CPCB, **Ambient air quality monitoring status: 1997,** NAAQMS/11/1999-2000, Central Pollution Control Board, 1999, New Delhi. 130pp.

28 CPCB, Air quality status and trends in India, NAAQMS/14/1999-2000, Central Pollution Control Board, 2000, New Delhi. 163pp.

29 Ott W R, **Environmental statistics and data analysis,** Lewis Publishers, 1995,Boca Raton.

30 WHO, *Analysing and interpreting air monitoring data,* WHO Offset Publication No. 51 World Health Organization, 1980, Geneva.

31 USEPA, *Preparation aids for the development of category III quality assurance project plans,* USEPA/600/8-91/005, United States Environmental Protection Agency, 1991, Washington, DC

32 Lodge J P (ed), **Methods of air sampling and analysis,** Lewis Publishers, Inc.1988, Chelsea.

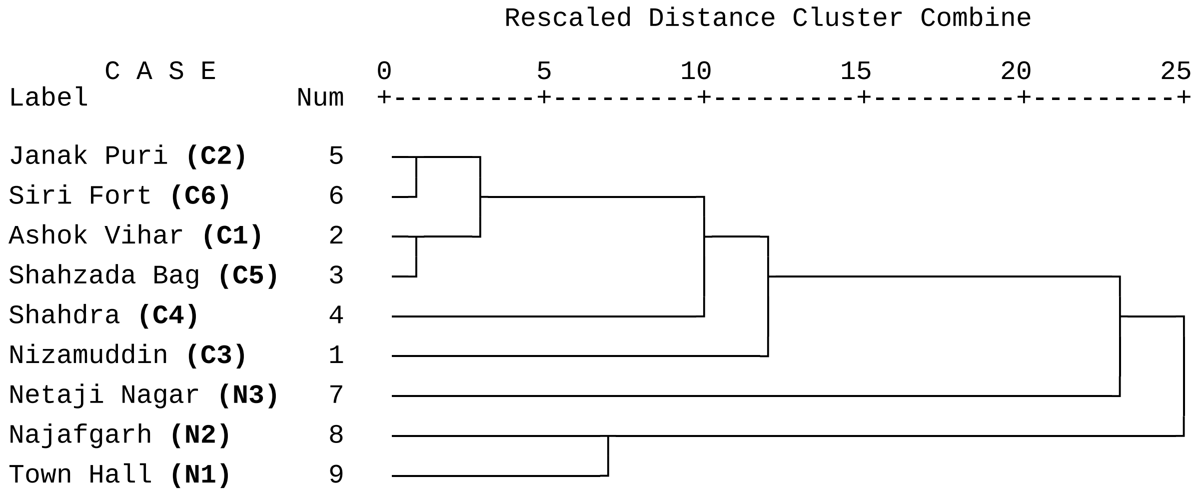
**Table 1.** Location of ambient air quality monitoring stations in Delhi

| Station Code | Location | Operated by | Zone | Land-use type |
|---|---|---|---|---|
| C1 | Ashok Vihar | CPCB | West | Industrial |
| C2 | Janak Puri | CPCB | West | Residential |
| C3 | Nizamuddin | CPCB | South | Industrial |
| C4 | Shahdra | CPCB | East | Industrial |
| C5 | Shahzada Bag | CPCB | East | Industrial |
| C6 | Siri Fort | CPCB | South | Residential |
| N1 | Town Hall | NEERI | Central | Residential |
| N2 | Najafgarh Road | NEERI | West | Industrial |
| N3 | Netaji Nagar | NEERI | South | Residential |

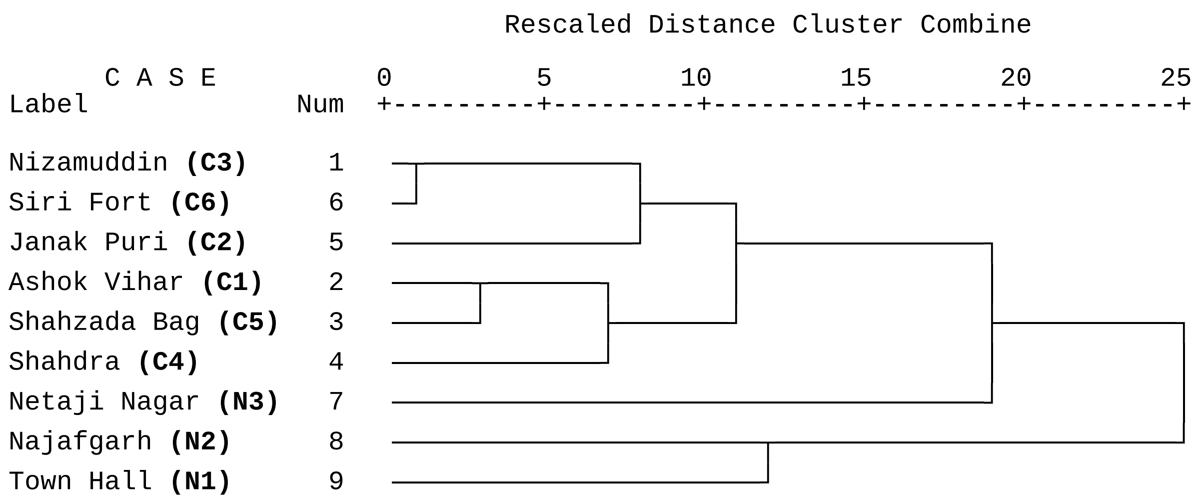
Note: Till 1991 Town Hall was classified under the commercial land-use type.

Figure 1 Dendogram of SPM data

a) 1996




b) 1997



c) 1998

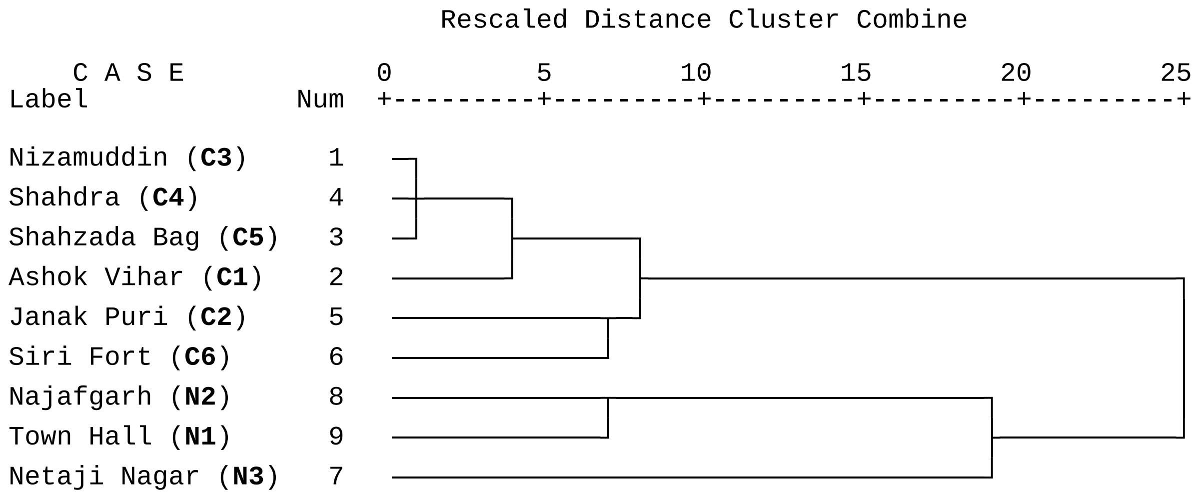
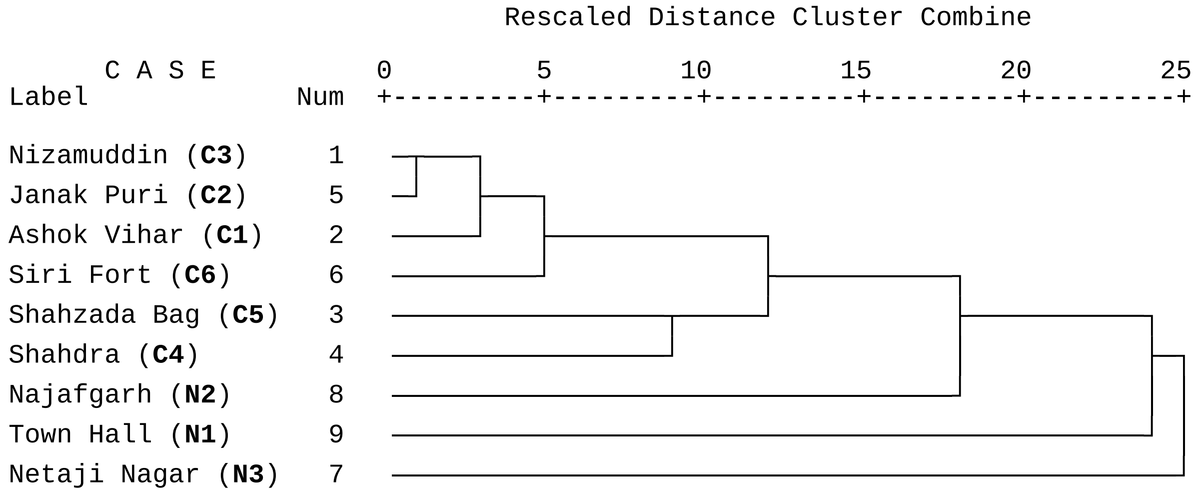
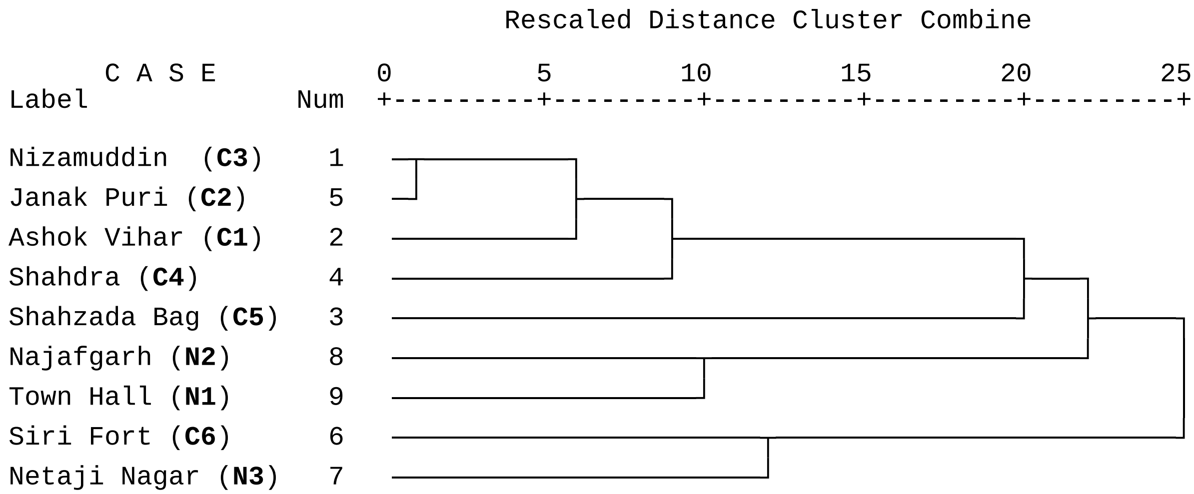

Rescaled Distance Cluster Combine
C A S E
Label
Num
0
5
10
15
20
25
Nizamuddin (**C3**) 1
Shahdra (**C4**) 4
Shahzada Bag (**C5**) 3
Ashok Vihar (**C1**) 2
Janak Puri (**C2**) 5
Siri Fort (**C6**) 6
Najafgarh (**N2**) 8
Town Hall (**N1**) 9
Netaji Nagar (**N3**) 7

Figure 2 Dendogram of $SO_2$ data

a) 1996




b) 1997



c) 1998

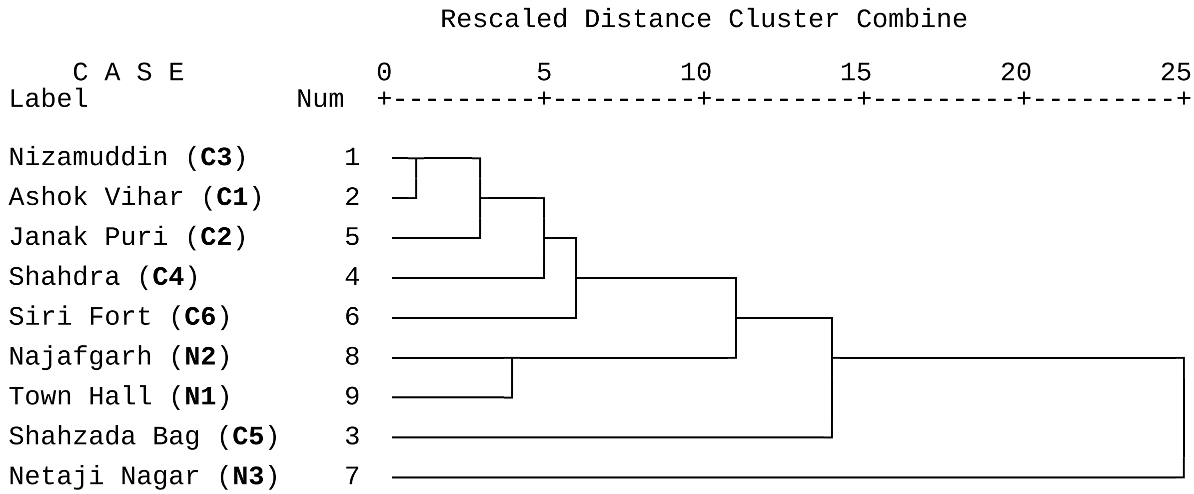
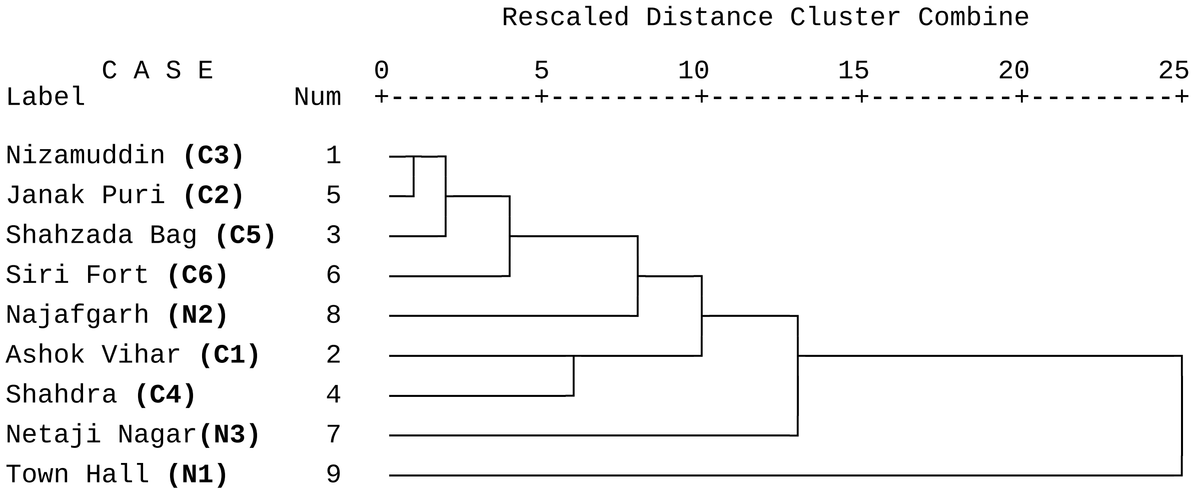
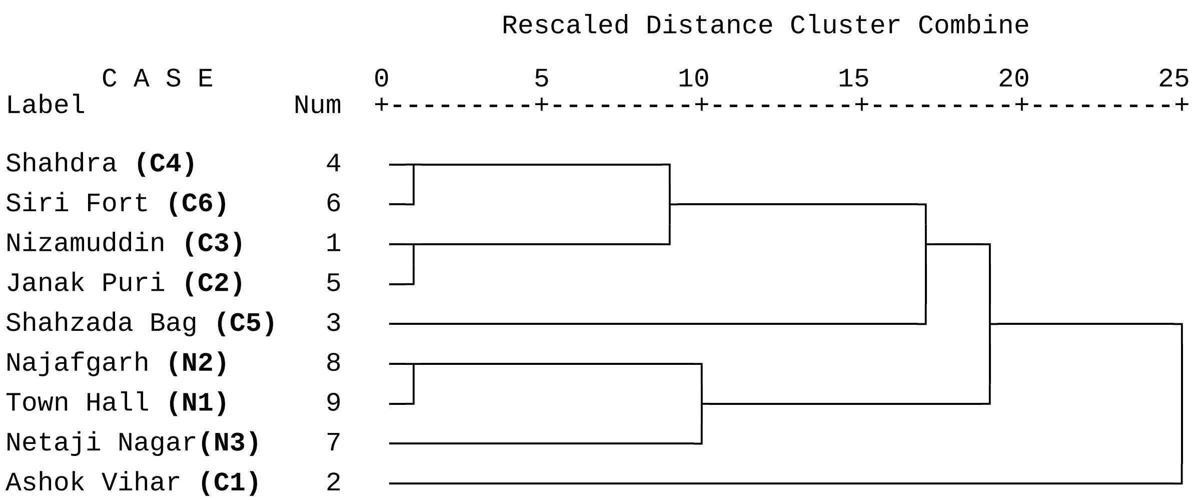

Rescaled Distance Cluster Combine
C A S E
Label
Num
0
5
10
15
20
25
Nizamuddin (**C3**) 1
Ashok Vihar (**C1**) 2
Janak Puri (**C2**) 5
Shahdra (**C4**) 4
Siri Fort (**C6**) 6
Najafgarh (**N2**) 8
Town Hall (**N1**) 9
Shahzada Bag (**C5**) 3
Netaji Nagar (**N3**) 7

Figure 3 Dendogram of $NO_2$ data

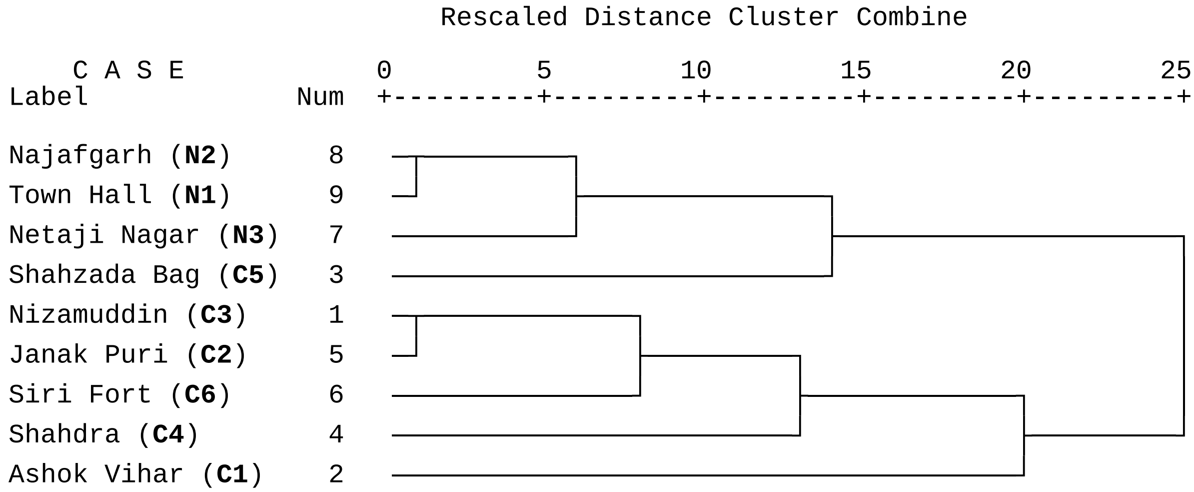
a) 1996




c) 1997



i) 1998


Rescaled Distance Cluster Combine
C A S E
Label
Num
0
5
10
15
20
25
Najafgarh (N2) 8
Town Hall (N1) 9
Netaji Nagar (N3) 7
Shahzada Bag (C5) 3
Nizamuddin (C3) 1
Janak Puri (C2) 5
Siri Fort (C6) 6
Shahdra (C4) 4
Ashok Vihar (C1) 2

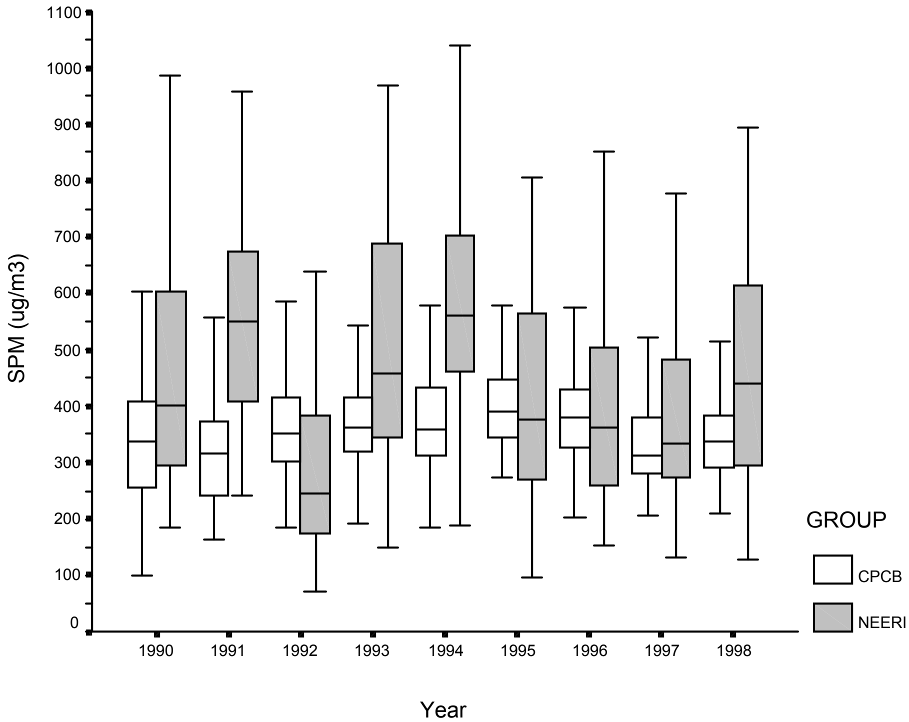



**Figure 4 Yearly variation in SPM concentrations across groups**

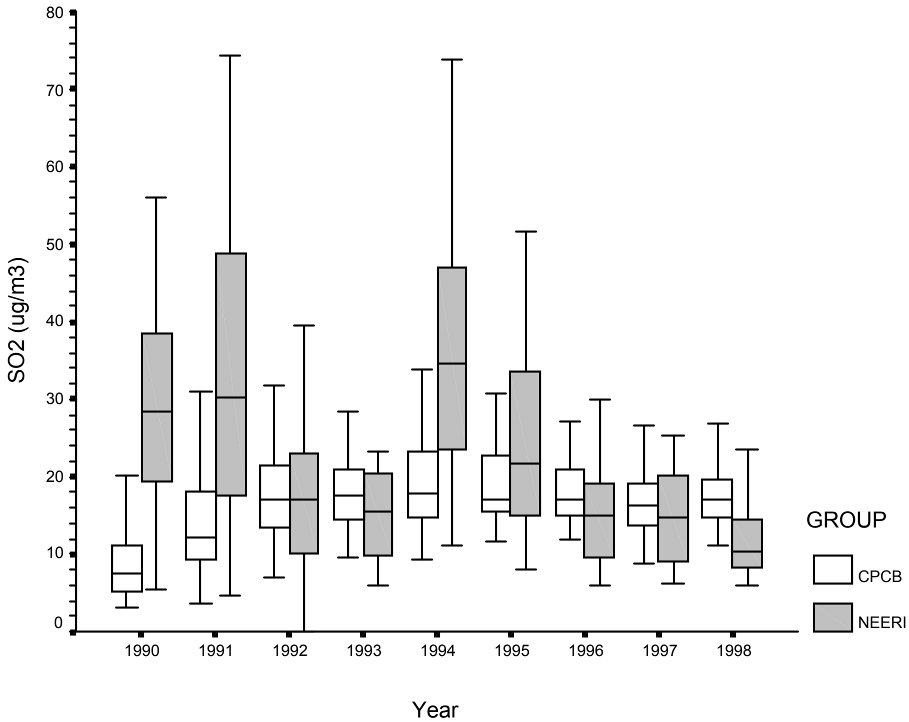



**Figure 5 Yearly variation in SO2 concentrations across groups**

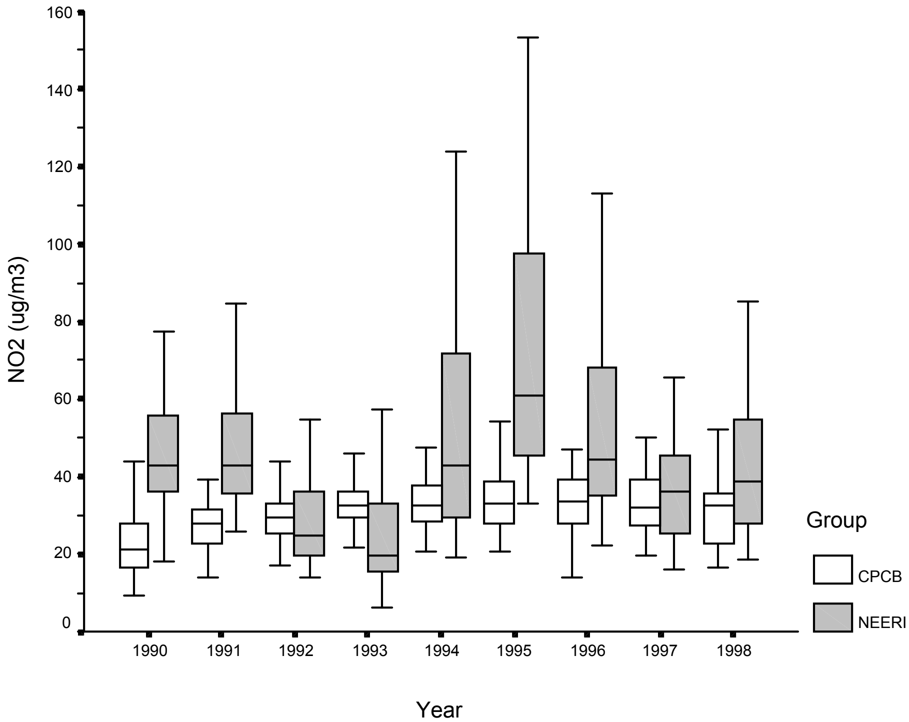



**Figure 6 Yearly variation in NO2 concentrations across groups**

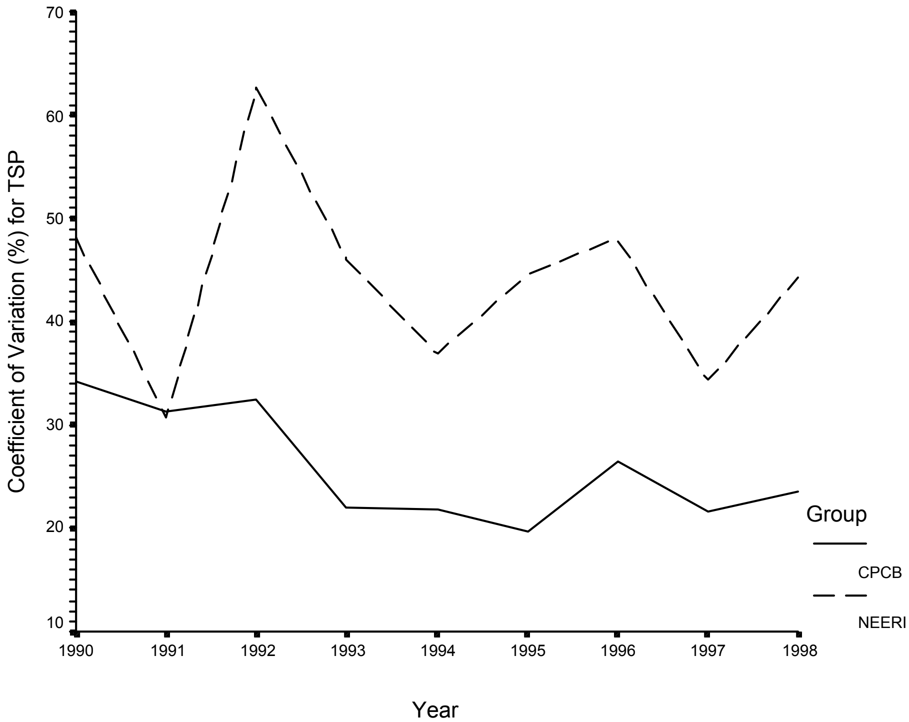



**Figure 7 Yearly variation in coefficient of variation of SPM across groups**


**Figure 8 Yearly variation in coefficient of variation of SO2 across groups**

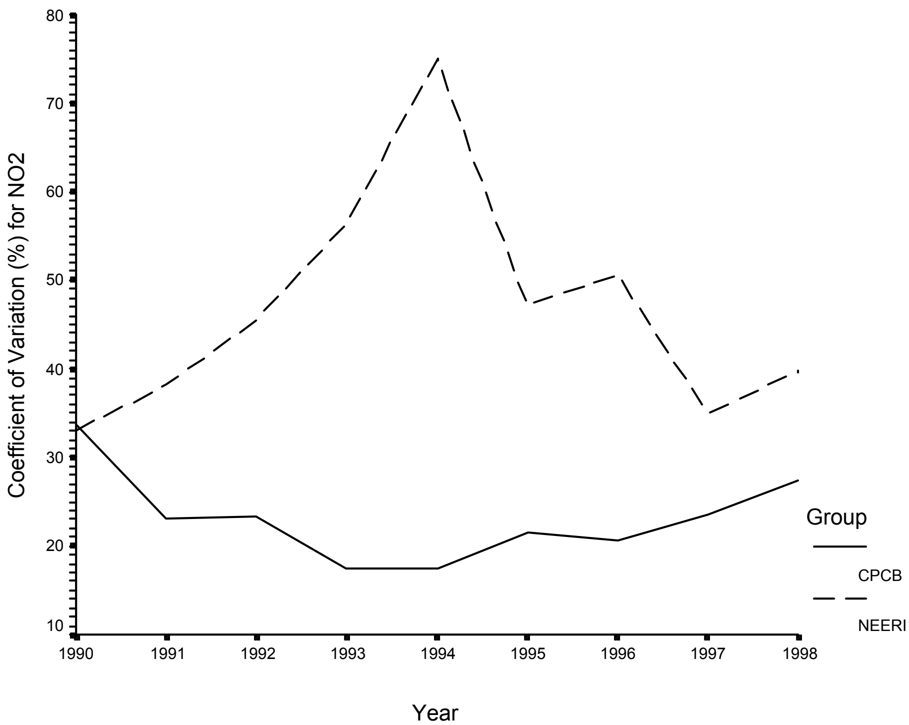



**Figure 9 Yearly variation in coefficient of variation of NO2 across groups**