Evaluation and optimization of an urban PM$_{2.5}$ monitoring network

JEM
www.rsc.org/jem

Dainius Martuzevicius,†[a] Junxiang Luo,[b] Tiina Reponen,[a] Rakesh Shukla,[b] Anna L. Kelley,[c] Harry St. Clair[c] and Sergey A. Grinshpun*[a]

[a] Center for Health-Related Aerosol Studies, University of Cincinnati, Cincinnati, OH, USA.
   E-mail: Sergey.Grinshpun@uc.edu; Fax: +1-513-558-2263; Tel: +1-513-558-0504
[b] Center for Biostatistical Services, University of Cincinnati, Cincinnati, OH, USA
[c] Hamilton County Department of Environmental Services, Cincinnati, OH, USA

The objective of this study was to evaluate the PM$_{2.5}$ monitoring network established in the Greater Cincinnati and Northern Kentucky metropolitan area for measuring the 24 h integrated PM$_{2.5}$ concentration, as well as—at selected sites—hourly PM$_{2.5}$ concentration and 24 h integrated PM$_{2.5}$ speciation. The data collected during three years at 13 measurement sites were analyzed for spatial and temporal variations. It was found that both daily and hourly concentrations of PM$_{2.5}$ have low spatial variation due to a regional influence of secondary ammonium sulfate. In contrast, the trace element concentrations had high spatial variation. Seasonal variation accounted for most of the total temporal variation (60%), while yearly, monthly, weekly and daily variations were lower. The variance components and cluster analyses were applied to optimize the number of sites for measuring the 24 h PM$_{2.5}$ concentration. It was found that the 13-site network may be optimized by reducing the number of sites to 8, which would result in a relative precision reduction of 9% and a relative cost reduction of 36%. At the same time, the data suggest that the spatial resolution of speciation monitors and real-time PM$_{2.5}$ mass monitors should be increased to better represent spatial and temporal variations of the markers of local air pollution sources.

1. Introduction

In 1997, the US EPA revised the national ambient air quality standards for particulate matter based on the available scientific evidence that linked exposures to ambient particulate matter with adverse health and welfare effects. In the new National Ambient Air Quality Standards (NAAQS) for Particulate Matter,[1–3] PM$_{2.5}$ was introduced as a parameter for characterizing the pollution represented by the fine fraction of ambient aerosol. PM$_{2.5}$ (particulate matter smaller than 2.5 μm in aerodynamic diameter) is presently of a primary monitoring concern in the USA. The US national PM$_{2.5}$ monitoring program, developed and operated following the Code of Federal Regulations (CFR),[1-3] includes numerous PM$_{2.5}$ monitoring networks of monitoring sites designated as State and Local Air Monitoring Stations (SLAMS), National Air Monitoring Stations (NAMS) and Special Purpose Monitoring Stations (SPMS). The Interagency Monitoring of Protected Visual Environments (IMPROVE) is another collaborative monitoring program aiming to establish present visibility levels and trends and to identify sources of anthropogenic impairment. The SLAMS network is the widest and has the largest number of samplers. In 1997, the Federal Reference Method (FRM) sampler network was designed to include nearly 1400 sites. In 2000 the network size was reduced to 1050 FRM sites.

The principles of PM$_{2.5}$ monitoring network design and site selection for the most representative exposure assessment were discussed in the EPA guidance document[4] and by Chow et al.[5] Due to the small size of particles, the properties of PM$_{2.5}$ are different from PM$_{10}$. The major portion of PM$_{2.5}$ mass is comprised of combustion products and secondary aerosols, while PM$_{10}$ mass includes that of PM$_{2.5}$ and, in addition, a mass of particles generated mechanically at nearby sources. The fine PM$_{2.5}$ fraction is often more homogeneously distributed over a space compared to the coarse aerosol fraction, as has been noted in other studies conducted in the north-east part of the USA[6,7] and south-east Canada.[8]

The temporal scales (sampling period, duration and frequency) of PM$_{2.5}$ measurement are important for the assessment of human exposure to outdoor air pollution. In many urban areas, PM$_{2.5}$ has been measured for more than 4 years, giving an estimate of seasonal and annual averages. The measurement duration is typically 24 h. This time interval is sufficient to collect a deposit for accurate gravimetric analysis. However, the 24 h integrated measurement does not allow the investigators to identify the aerosol concentration peaks that have short term duration, but may represent considerable hazard. Since the long-term averages require less frequent sampling, the PM$_{2.5}$ monitoring is conducted mostly using 24 h measurements performed once in 3 d. In some sites, every-day and every-sixth-day measurements are used. The patterns of temporal variations of the PM$_{2.5}$ concentrations have been described by several researchers.[8–11] Seasonal variations were mainly attributed to fluctuations in meteorological factors, while diurnal variations mainly depend on the fluctuations in the intensity of local pollution sources, e.g., traffic.

The review and optimization of regional and national monitoring networks were addressed in several publications.[12–14] The detailed evaluation of urban-scale monitoring networks have recently been done by researchers in Brisbane (Australia)[15] and Santiago (Chile).[16] Although these papers reviewed the PM$_{10}$ and gaseous pollutant networks, they did not address the PM$_{2.5}$ monitoring because the monitoring networks in those cities did not include routine PM$_{2.5}$ measurements.

In our recent study,[17] performed in the Greater Cincinnati area, we have indicated that the PM$_{2.5}$ concentration does not actually depend on the distance from federal highways. This

† On leave from the Department for Environmental Engineering, Kaunas University of Technology, Kaunas, Lithuania.

study has been performed using Harvard-type impactors (not an FRM instrument) at self-established monitoring sites during short-term measurement cycles. The results indicated low spatial variability, and suggested a possible optimization of a continuous monitoring network.

Since the $PM_{2.5}$ standard was introduced in 1997, many urban areas, including the Greater Cincinnati area, have established extensive $PM_{2.5}$ monitoring networks. The long-term monitoring database collected by these networks allows characterization of the $PM_{2.5}$ spatial and temporal variations. This knowledge is crucial for the validation of strategies used for urban $PM_{2.5}$ monitoring. The present paper describes the results of a 3-year $PM_{2.5}$ monitoring campaign that involved an extensive network of multiple monitoring sites operated in the Greater Cincinnati and Northern Kentucky areas.

## 2. Methods

2.1 $PM_{2.5}$ monitoring network

The Greater Cincinnati and Northern Kentucky $PM_{2.5}$ monitoring network consists of 13 measurement sites that host altogether 22 samplers of three different types. These sites are situated in two neighbouring states of the US, Ohio and Kentucky, divided by the Ohio River. A total of 18 samplers are positioned in Ohio and operated by the Hamilton County Department of Environmental Services (HCDOES). Four samplers are located in Northern Kentucky and operated by the Kentucky Department of Environmental Protection (KY DEP). The location of monitoring sites and the $PM_{2.5}$ monitoring equipment operating in each site are presented in Fig. 1. The following sites are situated in Ohio: Lower Price Hill (LPH), Taft (TAF), Norwood (NOR), Winton (WIN), St. Bernard (STB), Carthage (CAR), Oak (OAK), Sacred Heart

School (SHS), Verity (VER), Wild Wood School (WWS), Middletown (MID). The Northern Kentucky sites include Covington (COV) and Fort Thomas (FTH). The following equipment was utilized:

- Filter-based Federal Reference Method (FRM) samplers: Thermo Andersen RAAS2.5-300, RAAS2.5-100 (Thermo Electron Corp., Franklin, MA, USA) and Partisol-Plus Model 2025 Sequential Air Sampler (Rupprecht & Patashnick Co., Inc., East Greenbush, NY, USA);
- Speciation samplers: SASS (Met One Instruments Inc., Grants Pass, OR, USA);
- Real-time particulate concentration monitors: Tapered Element Oscillating Microbalance (TEOM Series 1400a, Rupprecht & Patashnick Co.) and Beta-Attenuation Mass Monitor (BAM 1020, Met One Instruments Inc.).

The $PM_{2.5}$ monitoring on the Ohio side started in January 1999 with three sites and three more sites were added by the end of March 1999. Two $PM_{2.5}$ sites were established in Northern Kentucky in the beginning of the year 2000. Three more $PM_{2.5}$ sites were added in Ohio in October of 2000 making the total number equal to 11. In 2001, the chemical speciation monitoring (using SASS samplers) and an hourly concentration monitoring (using TEOM) were initiated. Another continuous $PM_{2.5}$ monitor (BAM) was added in Middletown (MID) in the northern part of the area in 2003. All monitors in both Ohio and Northern Kentucky are included in the SLAMS network.

The type and the location for each monitor were selected according to the US EPA guidelines.[4] The US EPA set the number of sites per state according to previously collected PM monitoring data, pollutant emission rates and meteorological data. Based on the network design guidance, the state and local air pollution control agencies further refined the network. Population exposure and maximum concentrations are the
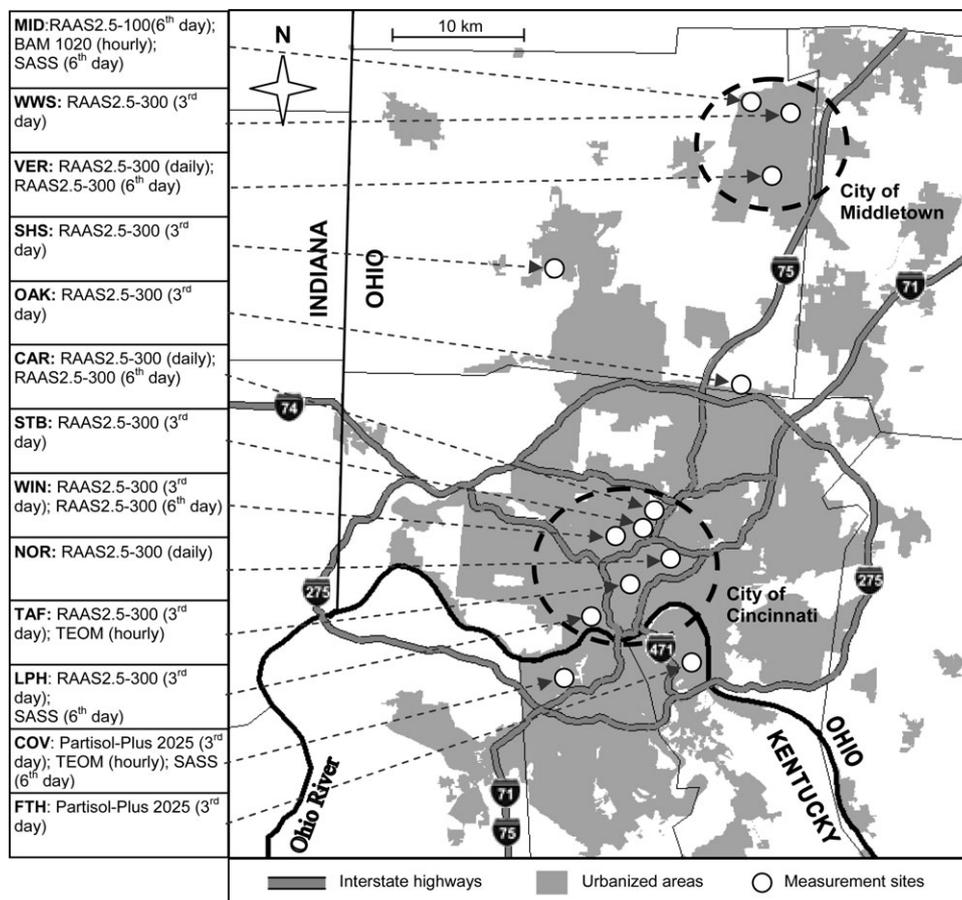
Fig. 1 Map of the Greater Cincinnati $PM_{2.5}$ monitoring network. For each site, its three-letter code, the monitor model and the measurement frequency are indicated.

main factors for selecting locations for monitors. The majority of the PM monitors in south-western Ohio follow a corridor containing highly populated areas with considerable exposure to industrial pollution sources.

The population of 14 counties of the Greater Cincinnati area is 2 143 689,[18] with the Hamilton County being the most populated (831 739 people). It contains six $PM_{2.5}$ monitoring sites. All of these six sites are concentrated in the City of Cincinnati, with a population of 323 885. Butler County (population of 339 828) has five monitoring sites; three of those are situated in the City of Middletown with a population of 51 605.

2.2 $PM_{2.5}$ sampling, analysis and quality control

$PM_{2.5}$ sampling was performed utilizing the FRM samplers[19] in accordance with the EPA regulations[1,3,20] and manufacturer's operating instructions. The 46.2 mm Teflon filters, pore size 2 μm (Whatman Inc., Clifton, NJ, USA) were used to collect airborne particles. All filters were passed through quality assurance checks before shipping to the individual laboratories for gravimetric analysis.[1] Filters were weighed following EPA regulations,[1,20] Standard Operating Procedures and good laboratory practices.

The $PM_{2.5}$ samples were collected every day (sites NOR, CAR, VER), every third day (WWS, SHS, OAK, STB, WIN, TAF, LPH, COV and FTH), and every sixth day (MID). The sampling frequency for the individual sites was determined based on several considerations. Those sites expected to record concentrations nearest to the annual and 24 h standards were operated according to a more frequent schedule. Initial selection for the measurement frequency was based on data generated through the IMPROVE network, as well as the data on pollution sources inventory in the region. It was anticipated that the Greater Cincinnati area would be in non-attainment of the NAAQS for $PM_{2.5}$ or very close to the annual standard of 15 μg m^{-3}. Daily sampling was included in areas of higher population as well as the land use, generating a specific type of pollution, upwind of the site. For the latter it was also important to determine the annual wind direction to help gauge the impact to the citizens downwind of the site(s). Every third day sampling was assigned to the remaining sites while every sixth day sampling was assigned to the collocated units. Three sites of the network (WIN, CAR and VER) had collocated samplers in accordance with CFR.[3] The three sites were selected to represent the areas of denser population with high predicted concentrations due to local pollution sources.

Although the initial network design for chemical speciation was well defined by EPA,[3] there are no samplers for speciation that are defined as an FRM sampler. The instrumentation utilized for speciation sampling is a modification of the existing FRM with various filter media and more channels for concurrent sampling. The different filter media allow the use of various analysis methods to determine the composition of the $PM_{2.5}$ fraction. The following filter media were utilized: Teflon for gravimetric ($PM_{2.5}$ mass) and X-ray fluorescence (elements); nylon for ion chromatography (soluble ions) and quartz for Thermal Optical Transmittance analysis (elemental and organic carbon). The analyses of the filter media for the chemical speciation program were performed by the Research Triangle Institute (Research Triangle Park, NC, USA) through a national contract. In contrast to the $PM_{2.5}$ mass measurements, the speciation samples were collected every sixth day only.

Both TEOM and BAM were used for continuous $PM_{2.5}$ sampling. Although there is no FRM for continuous sampling, these instruments provide "close to a real-time" measurement and have generated the data used for determining the Air Quality Index. The comparisons of TEOM and BAM to FRM samplers have been performed by several investigators.[21–23] The study performed in Austria[22,23] revealed that the two techniques (TEOM and BAM) represented the $PM_{2.5}$ concentrations rather adequately as long as the same heating conditions of the sample air flow were maintained. It should be acknowledged that in our tests, no heating was applied for the BAM inlet, while the temperature of the TEOM sample flow was maintained at 50 °C. Both monitors utilized PM_{10} inlet heads and sharp-cut cyclones for separating the $PM_{2.5}$ fraction.

While required for the FRM and speciation units, monthly flow checks and quarterly flow audits were performed on all samplers. The monthly flow checks were done by the operators of each unit using flow devices, which were referenced annually to the National Institute of Standards and Technology (NIST, Gaithersburg, MD, USA) traceable standards. The quarterly flow audits were performed by the Ohio Environmental Protection Agency, Division of Air Pollution Control for the HCDOES and by the Quality Assurance Section of the Kentucky Division for Air Quality for the KY DEP.

2.3 Data analysis techniques

The statistical analysis of the collected data was performed using Microsoft Excel 2003 (Microsoft Corp., Redmond, WA, USA) and SAS 8.02 (SAS Institute Inc., Cary, NC, USA) software. In order to quantitatively characterize the spatial variation of $PM_{2.5}$ concentration, the coefficient of variation ($CV_{spatial}$) was calculated from every third day measurement for each cycle. It represented the variation among 12 simultaneously operated monitoring stations based on daily measurements, thus giving us a distribution of spatial CVs over a large number of days ($n = 330$). The data from the MID station, which was operated on an every sixth day schedule, was excluded from calculations. Since the CV value depends on the number of data points, it can not be used as a single factor to describe the variation. In addition, simple linear regression analysis was performed to determine the coefficient of determination (r^2) and slope (b_1). The r^2 values were significant ($\alpha = 95\%$) in all cases for $PM_{2.5}$ 24 h and hourly concentrations due to a large number of data points ($n \geq 270$ for 24 h measurements and $n \geq 2335$ for hourly measurements). For speciation data, r^2 values greater than 0.08 were considered as significant ($\alpha = 95\%$) due to a lower number of measurements ($n = 49$).

The temporal variation of the $PM_{2.5}$ concentration was estimated using the variance components analysis. This type of analysis allows partitioning of the total variation in a measurement data set in terms of various components (factors).[24,25] Temporal components of the total variance in the $PM_{2.5}$ were split into yearly, seasonal, monthly, weekly and daily effects. The calculations of the variance components were performed using restricted maximum likelihood method in the random-effect ANOVA model available from the MIXED procedure of the SAS Software.[26] The data were fitted to the following nested model:

$$\log(X_{ijklmn}) = \mu_y + Y_j + P_{k(j)} + M_{l(k)} + W_{m(jkl)} + D_{n(jklm)} \quad (1)$$

where $\log(X_{ijklmn})$ represents the $PM_{2.5}$ concentration, adjusted to a log-normal distribution; μ_y is the overall long-term mean value of the $PM_{2.5}$ concentration determined for a specific site; Y_j is the random yearly effect, ($j = 1...J$); $P_{k(j)}$ is the random seasonal effect (nested in a year), ($k = 1...K$); $M_{l(k)}$ is the random monthly effect (nested in season), ($l = 1...L$); $W_{m(jkl)}$ is the random weekly effect (nested in month), ($m = 1...M$); and $D_{n(jklm)}$ is the random daily effect (nested in week), ($k = 1...K$), confounded with the effect of error.

It is assumed that Y_j, $P_{k(j)}$, $M_{l(k)}$, $W_{m(jkl)}$ and $D_{n(jklm)}$ are normally distributed with means 0 and variances σ_Y^2, σ_P^2, σ_M^2, σ_W^2 and σ_D^2, respectively, and that the random effects were considered mutually independent.

Hence, for monitoring data collected in i stations during a certain time period, the variance of the mean exposure $Var(\overline{\log(X)})$ is the sum of variances of all components of the

nested model. Since the number of measurement data is not infinite, the Finite Population Correction was applied for computing the variance of the overall mean value as follows:

$$\mathrm{Var}(\overline{\log(X)}) = \left(\frac{1}{y} - \frac{1}{Y}\right)\hat{\sigma}_Y^2 + \frac{1}{y}\left(\frac{1}{p} - \frac{1}{P}\right)\hat{\sigma}_P^2$$
$$+ \frac{1}{yp}\left(\frac{1}{m} - \frac{1}{M}\right)\hat{\sigma}_M^2 + \frac{1}{ypm}\left(\frac{1}{w} - \frac{1}{W}\right)\hat{\sigma}_W^2$$
$$+ \frac{1}{ypmw}\left(\frac{1}{d} - \frac{1}{D}\right)\hat{\sigma}_D^2$$

(2)

where $\hat{\sigma}_{(\cdot)}^2$ is the estimate of the respective variance component; y, p, m, w and d are, respectively, numbers of years during which a specific data sample was collected, seasons in a year, months in a season, weeks in a month, and days in a week—all related to a specific data sample; Y, P, M, W and D are the corresponding number of years, seasons, months, weeks and days in the entire monitoring data (statistical population), *i.e.*, the sampling period. In this analysis, we have included the 3-year $PM_{2.5}$ concentration data, obtained from the sites operating according to an every third day schedule. Thus, $Y = 3$, $P = 4$, $M = 3$, $W = 4$, $D = 2.5$. The value $D = 2.5$ represents the number of every third day measurements during one week.

LeMasters *et al.*[24] have successfully used the variance components analysis for balancing cost and precision in an occupational exposure study. The nested model was appended with the "site" effect in order to asses the contribution of spatial variability to the total variability of the measurement data. The model equation can be transformed as follows:

$$\log(X_{ijklmn}) = \mu_y + S_i + Y_j + P_{k(j)} + M_{l(k)}$$
$$+ W_{m(jkl)} + D_{n(jklm)}$$

(3)

where S_i represents a random effect of site (incorporating the time dimension), $(S_i = 1\ldots S)$.

The variance is calculated accordingly:

$$\mathrm{Var}(\overline{\log(X)}) = \frac{\hat{\sigma}_S^2}{s} + \left(\frac{1}{y} - \frac{1}{Y}\right)\hat{\sigma}_Y^2 + \frac{1}{y}\left(\frac{1}{p} - \frac{1}{P}\right)\hat{\sigma}_P^2$$
$$+ \frac{1}{yp}\left(\frac{1}{m} - \frac{1}{M}\right)\hat{\sigma}_M^2 + \frac{1}{ypm}\left(\frac{1}{w} - \frac{1}{W}\right)\hat{\sigma}_W^2$$
$$+ \frac{1}{ypmw}\left(\frac{1}{d} - \frac{1}{D}\right)\hat{\sigma}_D^2 + \frac{1}{sypmwd}\hat{\sigma}_\varepsilon^2$$

(4)

If some of the monitoring sites are removed during the process of optimization, precision of the overall mean is expected to decrease. To estimate this drop of precision, the number of sites (s) can be altered when computing the variance of the overall mean of the $PM_{2.5}$ concentration for the alternative

(abbreviated) network. The precision of collected measurement data can be expressed as an inverse of a standard error (relative to the variance). The "precision" of the measurement data should not be confused with the "accuracy" of the measurement procedure. It is assumed that the cost is directly proportional to the number of observations. To compare the abbreviated network that can be proposed for the optimization to the existing full-size one, two indices were calculated: the Relative Precision (*RP*) and Relative Cost (*RC*):

$$RP = \frac{SE_{\mathrm{e.n.}}}{SE_{\mathrm{a.n.}}} \times 100\%$$

(5)

$$RC = \frac{N_{\mathrm{obs.\ a.n.}}}{N_{\mathrm{obs.\ e.n.}}} \times 100\%$$

(6)

where $SE_{\mathrm{e.n.}}$ and $SE_{\mathrm{a.n.}}$ represent standard error of measurement data in existing and abbreviated networks, respectively; $N_{\mathrm{obs.}}$ represents number of observations (samples).

In addition to the variance components analysis, the variable cluster analysis (VARCLUS procedure in SAS, disjoint clustering[26]), was used for the network optimization. This type of analysis allows grouping several sites into one cluster according to the similarities between the data recorded in certain sites. The disjoint clustering groups variables based on a correlation matrix. The clusters are chosen to maximize the variation accounted for by the first principal component. The cluster analysis returns a dendritic tree (dendogram), which shows how sites can be grouped.

Table 1 summarizes the information on the PM data collected during the monitoring campaign. Only the data collected in all sites during the same time span were utilized in the analysis. Thus, we analyzed the daily average concentration data measured during almost three years, which adds up to 330 data points of every-third day measurements. For hourly measurements, the data from 3 samplers collected during 10 months of the year 2003 were analyzed; for the chemical speciation, 1-year data from the three stations taken at 6 day intervals were available for the analysis.

3. Results and discussion

3.1 Spatial variation of $PM_{2.5}$ concentration

The spatial variation and the relationship of the $PM_{2.5}$ concentration measured with 24 h filter-based samplers among 12 monitoring sites operating every third day was quantitatively characterized by the coefficient of variation and parameters of simple linear regression analysis described in Section 2.3. The 13th site (MID), which recorded the $PM_{2.5}$ concentration every sixth day was not used in this analysis.

The site-specific average concentration during the analyzed period across the 12 stations ranged from 15.2 ± 8.1 μg m^{-3} (FTH) to 18.1 ± 8.9 μg m^{-3} (CAR). The median CV_{spatial}

Table 1 $PM_{2.5}$ monitoring summary

Data type	Monitor	Collected by	Monitoring period	Number of sites	Number of samples per monitor (*n*)
$PM_{2.5}$ 24 h	RAAS2.5	HCDOES	October 3, 2000–July 31, 2003	10	330[a]
	Partisol 2025	KY DEP		2	
$PM_{2.5}$ 1 h	TEOM 1400, BAM 1020	HCDOES	February 27, 2003–December 31, 2003	2	6072 2335[b]
	TEOM 1400	KY DEP		1	6072
$PM_{2.5}$ speciation 24 h	SASS	HCDOES	August 6, 2002–August 13, 2003	2	52[c]
	SASS	KY DEP		1	

[a] Every third day measurements. [b] The data collected from February 27 through June 5, 2003 represent the instrument validation and have not been included in the EPA Air Quality System (AQS). [c] Every sixth day measurements.

appeared to be relatively low (11.2%). This confirms the hypothesis about a uniform pattern of the $PM_{2.5}$ concentration across the region. The $CV_{spatial}$ value is in good agreement with the result of our previous study[17] ($CV_{spatial} = 7.6\%$), which was performed using a different set of sites and different PM sampling instruments (Harvard-type impactors). The highest spatial variation of 31.6% among the sites occurred on October 30, 2000, when the lowest concentration of 6.9 µg m^{-3} was recorded at the WWS site, and the highest concentration of 18.8 at the NOR site. Meteorological data, recorded during this day indicate low night temperatures (+3 °C), low wind speeds (indicators of thermal inversions) and hazy conditions during daytime. The measurements conducted at the sites, situated in the industrialized areas (NOR, CAR) identified higher concentration levels. The lowest spatial variation of 3.0% was found on June 22, 2002, with the $PM_{2.5}$ concentration ranging from 37.7 µg m^{-3} (FTH) to 41.6 µg m^{-3} (CAR). On this day the haze and fog conditions were also recorded, together with relatively low wind speeds. In contrast to the day when high spatial variability occurred, the daily temperature reached 32 °C. The relatively uniform concentrations may have occurred due to secondary aerosol formation across the region under these meteorological conditions.

In an urban area, where the daily $PM_{2.5}$ concentrations are rather uniformly distributed, the CV may serve as a good indicator of potential data outliers or local pollution peak levels. Those could be detected solely based on the change in the overall $CV_{spatial}$ without screening the data from each monitor separately.

The simple linear regression analysis confirmed the uniform distribution across the metropolitan area. The data collected in all the sites demonstrated a relatively low scattering around the regression line. The results from the regression analysis are shown in Fig. 2. Fig. 2 (a) and (b) represent two extreme cases in terms of data scattering. Fig. 2(a) shows the correlation between $PM_{2.5}$ concentration in STB and CAR stations, which were located approximately 8 km from one another. The correlation between these stations is very high ($r^2 = 0.96$) and the linear regression slope is close to unity (0.97). Lower correlations were observed for the stations located far away from each other. The greatest distance was between COV/FTH, located in Northern Kentucky, and VER/WWS, located in Middletown, OH (Fig. 1). The distance between these locations is approximately 45 km. However, the lowest correlation was noticed between the WWS and the LPH sites [$r^2 = 0.84$, Fig. 2(b)]. The latter is situated in the river valley near
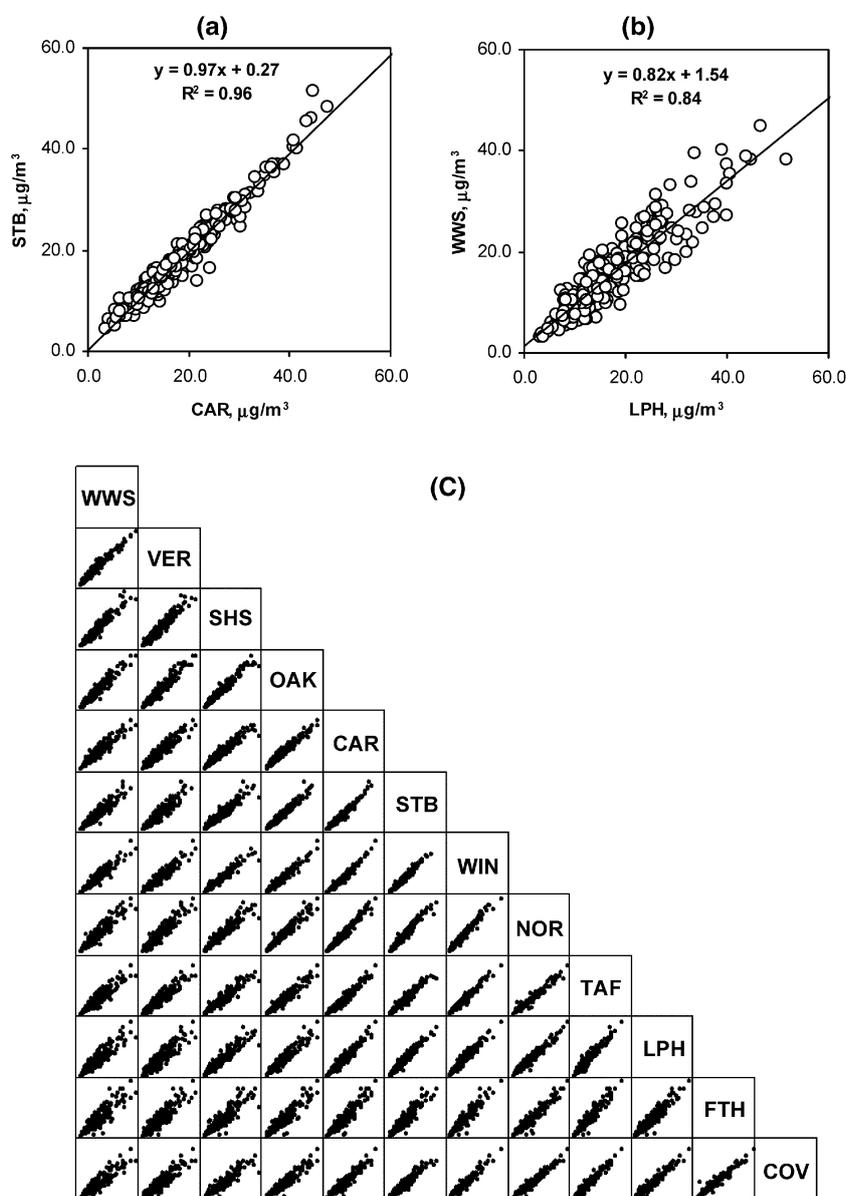
Fig. 2 Spatial variation of the daily averaged $PM_{2.5}$ concentration: (a) the example representing the lowest variation; (b) the example representing highest variation and (c) regression for all sites.

downtown Cincinnati. This finding can be explained by the influence of local topography on the local meteorological conditions and, consequently, on the $PM_{2.5}$ concentrations. Due to a lower wind speed and more frequent thermal inversions, the daily fluctuation pattern of the $PM_{2.5}$ concentration was slightly different in the river valley, compared to those outside the valley, while the overall r^2 value was still relatively high. However, the slope value $b_1 = 0.82$ suggests that concentrations in LPH were constantly higher than in WWS.

The entire set of the scatter plot graphs describing spatial relationship between sites are presented in Fig. 2(c). While scattering was low and the r^2 value was relatively high in all pair-wise comparisons, the slope shows that the $PM_{2.5}$ was higher in the sites located in the inner city (*e.g.*, LPH, STB, CAR) as compared to those located outside the city limit, (*e.g.*, OAK and SHS).

The low spatial variability implies that the most significant part of the $PM_{2.5}$ originates from regional rather than from local sources. The low spatial variation is likely to occur due to high levels of background ammonium sulfate and organic carbon concentration in the particles, as described in Section 3.2. Ammonium sulfate is chemically stable and can therefore be transported over large distances. The major SO_2 emission sources (coal-burning power plants) impacting the region are distributed along the Ohio river. They are far away from the Greater Cincinnati area so that there is a sufficient time for the SO_2 to SO_4^{2-} conversion process, and the pollutants are well dispersed before impacting the sites. Therefore, the sulfate concentrations are not too different across the monitoring location. Organic carbon compounds are also products of combustion processes. A significant amount of OC is transported regionally, while it also receives a considerable contribution from local combustion sources. This low spatial variability of $PM_{2.5}$ concentrations supports the results from other studies conducted in the north-east part of the USA and south-east Canada.[6–8]

The hourly $PM_{2.5}$ was measured at three sites—COV and TAF (by TEOM) and MID (by BAM). As seen from Fig. 3, the COV and TAF sites, which were relatively close to each other, demonstrate good correlation with respect to the $PM_{2.5}$ hourly concentration ($r^2 = 0.85$, $b_1 = 0.94$). The r^2 value is close to that determined for the 24 h measurement ($r^2 = 0.95$). Both sites (COV and TAF) were close to the downtown Cincinnati area. At the same time, they were located on the hills outside of the river valley, which made them accessible for the regionally transported air parcels. Although no hourly particle composition data are available, regional secondary ammonium sulfate and organic matter are most likely to be the key contributors to the hourly $PM_{2.5}$ mass. The correlation of the hourly data between these stations and the MID station was substantially lower: $r^2 = 0.43$, $b_1 = 0.99$ for TAF *vs.* MID and $r^2 = 0.41$, $b_1 = 0.93$ for COV *vs.* MID (compared to $r^2 = 0.87$ and $r^2 = 0.83$ for the 24 h values, respectively). The lower r^2 may be due to

different air masses or sources impacting the MID site compared to COV and TAF. MID is approximately 40 km from COV and TAF (Fig. 1), and, therefore, meteorological conditions at the two downtown sites are often different from those at the MID site. The time needed for the transportation of an air parcel through the region affects the hourly variation in $PM_{2.5}$ concentration between two remote areas. The data analysis on the temporal variation showed that high hourly concentrations (peaks) may not necessarily be identified by the 24 h concentration measurement, as discussed in Section 3.3. Thus, the 24 h concentration data show slightly better spatial correlation than the hourly concentration data.

Although the unheated BAM sample is expected to give a higher concentration of $PM_{2.5}$ than the heated TEOM sample, the slopes of the regression lines appeared to be close to unity (0.99 for TAF *vs.* MID and 0.93 for COV *vs.* MID). Visually, the scatter plots seem to indicate that higher concentration levels were observed in MID, but this is not revealed by the quantitative information obtained using the least squares method.

3.2 Spatial distribution of the $PM_{2.5}$ chemical compounds

As indicated above, three monitoring sites, COV, LPH and MID, were equipped with the speciation monitors (Fig. 1). The COV and LPH sites were relatively close to each other (approx. 2.5 km), while the MID site was located approx. 47 km to the north. Since the number of speciation sites is so small, the *CV* was not used to characterize the variation for specific compounds. Instead, the linear regression graphs were plotted. The spatial relationship between sites was analysed for the following ions and elements: SO_4^2, NO_3^-, NH_4^+ (markers of combustion processes and secondary aerosol formation); organic carbon (OC) and elemental carbon (EC) (combustion and traffic emissions); Si, Ca, (crustal matter); Fe, V, Zn, Mn and Ni (local industrial processes and traffic).[27] For most elements, the concentrations were significantly greater than the limit of detection (LOD), except for nickel and vanadium (for which some measurement data were comparable to LOD). Thus, some of the variations described below may result from instrumental errors.

The regression analysis of the $PM_{2.5}$ constituents revealed different patterns of spatial distribution for different elements. Some of the constituents were of regional nature, and had low spatial variation, while others represented the influence of local sources and their concentration varied substantially from one site to another. The scatter-plots between three stations for two elements, sulfur and zinc, are presented in Fig. 4. These elements represented two extremes in terms of the spatial variation. Sulfur (proportional to sulfate) is a marker of coal-powered plant emissions, which are transported regionally and subsequently form secondary aerosol particles. Zinc represents local sources, mostly as a component of emissions
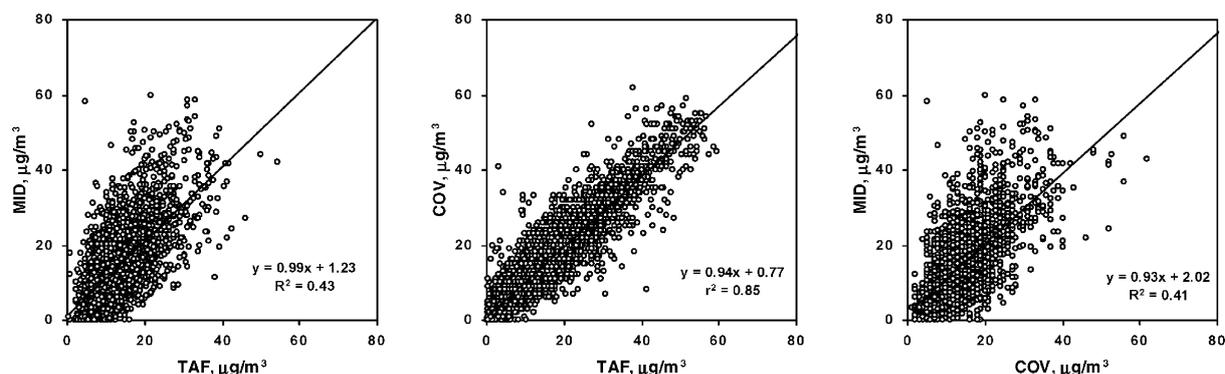

Fig. 3 Spatial variation of the hourly $PM_{2.5}$ concentration among three sites, monitoring hourly $PM_{2.5}$ concentration.
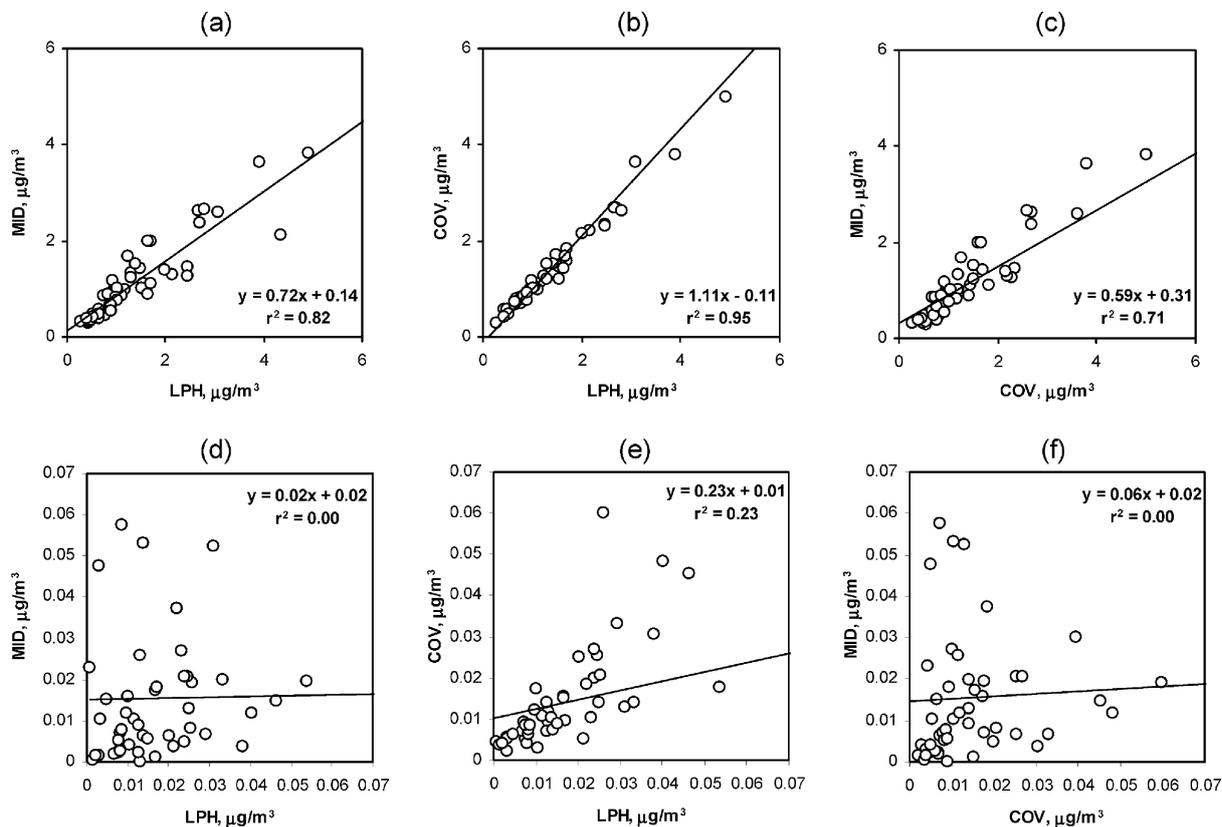
Fig. 4 Spatial variation (regression analysis) of sulfur (a, b and c) and zinc (d, e and f) among three speciation sites.

from metal-working industry and a trace element in traffic emissions. Both elements were detected by the XRF analysis.

Fig. 4 (a), (b) and (c) shows good correlation between the concentrations of sulfur obtained in these sites. The highest spatial correlation is observed between sites located relatively close to each other—LPH and COV [Fig. 4(b) with $r^2 = 0.95$]. The slope of the regression line ($b_1 = 1.11 \pm 0.04$) implies that sulfur concentrations measured in the COV site are slightly higher than those in the LPH site. This could be explained by the influence of the topography: LPH site is positioned on the floor of the river valley, while the COV site is situated on the top of the valley hill, and thus is more susceptive to the regional air parcels containing secondary sulfate. More distant sites demonstrate lower correlation. In general, sulfur (or sulfate) follows spatial patterns similar to those of the $PM_{2.5}$ concentration. As sulfate is clearly seen as one of the major constituents of the $PM_{2.5}$ fraction, this finding explains the low variation of $PM_{2.5}$ concentration in the area, described in Section 3.1.

The element Zn shows high spatial variation as demonstrated by the scatter plots presented in Fig. 4 (d), (e) and (f). The close sites, LPH and COV, display somewhat better, although still low, correlation between them ($r^2 = 0.23$). The concentration of zinc depends on the intensity of local sources and the local meteorological conditions. Both the downtown Cincinnati (LPH) and Middletown (MID) are highly industrialized areas, housing metal-working and other industrial facilities.

The regression analysis was also performed on the data obtained for selected elements and ions. Table 2 presents the coefficient of determination r^2 and the slope b_1. A slope value smaller than one indicates that concentration is higher in the site which is assigned as an independent variable X in the linear regression equation $Y = b_0 + b_1X$. For example, organic carbon (OC) concentrations are constantly higher in the LPH site, located in downtown Cincinnati close to major highways compared to Middletown ($b_1 = 0.67$) or Covington ($b_1 = 0.65$).

Obviously, for the elements with low values of r^2, the coefficient b_1 cannot be used as an indicator, because the regression line does not reflect the linear relationship between the stations.

As seen from Table 2, ammonium and OC demonstrate high spatial correlation between the three stations. Spatial distribution of ammonium is similar to that of sulfate and $PM_{2.5}$. Ammonium in the $PM_{2.5}$ is present in the form of ammonium sulfate $(NH_4)_2SO_4$ or ammonium bisulfate NH_4HSO_4 (depending on the particle acidity[28]), as well as ammonium nitrate. Other products of coal combustion, such as OC and elemental carbon (EC), are also transported regionally. However, since they have numerous local sources, e.g., small industrial facilities and motor vehicle traffic, the spatial correlation between sites is slightly lower for these compounds, compared to sulfate and ammonium: $r^2 = 0.74$ for LPH vs. MID with respect to organic carbon and $r^2 = 0.47$ for LPH vs. MID with respect to elemental carbon. For both OC and EC, the slope indicates that monitoring stations located in the city of Cincinnati (COV and LPH) constantly record higher concentrations of these pollutants, compared to those measured in the city of Middletown. For example, LPH vs. MID, $b_1 = 0.67$ for OC and $b_1 = 0.32$ for EC. The crustal element Si also showed a high spatial correlation. There is evidence that particulate Si is subject to long-range transport.[29] Concentrations may depend on the efficiency of particle re-suspension from the soil, which is affected by wind speed and traffic volume. Higher concentrations of Si also occur in the sites within Cincinnati city limits, as indicated by slope values ($b_1 = 0.74$–78) due to re-suspension of crustal matter associated with the street traffic and possibly due to local activities resulting in fugitive dust generation.

The trace elements (Fe, Pb, Mn, Ni, V and Zn) demonstrated no spatial correlation, or it was very low. It was concluded that the greatest contribution of these elements to the $PM_{2.5}$ comes from local pollution sources, making the regional influence practically undistinguishable.

Table 2 Spatial variation of selected elements and ions among three speciation monitoring sites. The coefficient b_1 represents slope of the linear regression line equation $Y = b_0 + b_1X$; r^2 represents the coefficient of determination

|	$X = $ LPH, $Y = $ MID		$X = $ LPH, $Y = $ COV		$X = $ COV, $Y = $ MID	
	b_1	r^2	b_1	r^2	b_1	r^2
$PM_{2.5}$	0.67	0.77	0.90	0.87	0.67	0.71
NH_4^+	0.70	0.80	0.90	0.90	0.68	0.69
NO_3^-	0.36	0.11	0.21	0.00	−0.01	0.02
SO_4^{2-}	0.70	0.79	1.09	0.96	0.58	0.69
EC	0.32	0.47	0.58	0.70	0.53	0.59
OC	0.67	0.74	0.65	0.68	0.79	0.66
Ca	0.49	0.33	0.51	0.56	0.84	0.45
Fe	0.57	0.19	0.68	0.31	0.23	0.05
Pb	0.04	0.00	0.01	0.00	−0.03	0.00
Mn	0.38	0.06	0.60	0.28	0.18	0.02
Ni	−0.08	0.01	−0.52	0.01	−0.02	0.02
Si	0.76	0.74	0.98	0.85	0.78	0.89
V	−0.06	0.00	0.43	0.20	−0.11	0.01
Zn	0.02	0.00	0.22	0.23	0.06	0.00

3.3 Mass closure. Temporal variations of the $PM_{2.5}$ concentration and chemical composition

The $PM_{2.5}$ mass balance was calculated utilizing conventional correction factors for assessing oxygen and hydrocarbon mass in oxides and hydrocarbons. A conversion factor of 1.6 was used for the estimation of organic matter (OM) from organic carbon.[30] Oxides of the most abundant trace elements (As, Cr, Cu, K, Mg, Ni, P, Pb, Se, Sr, V and Zn) were represented by a TEO (trace element oxides) group. Average monthly concentrations of $PM_{2.5}$ and its constituents for the three stations during the one-year period are presented in Fig. 5.

The mass balance calculations showed that the composition of the particulate matter is rather similar in the three speciation stations (COV, LPH and MID), suggesting the influence of regional background aerosol. The background $PM_{2.5}$ concentration in the Greater Cincinnati area is considerably affected by high concentrations of organic matter and secondary ammonium sulfate, as already mentioned in Section 3.2. The annual average organic matter contribution to the total $PM_{2.5}$ mass was 34% in all three monitoring sites, while the sulfate concentration varied as follows: 26% in COV, 21% in LPH, and 19% in MID. The city sites demonstrated a higher contribution from elemental carbon (3.8% for COV, 3.9% for LPH *vs.* 2.6% for MID). The increased concentrations of EC observed in the city sites may be explained by the influence of interstate highways I-71 and I-75, carrying more than 300 000 vehicles per day, of which approx. 10% are heavy-duty diesel vehicles. However, the trace element oxides (TEO) concentration and its relative contribution to the total mass were higher in the MID site (2.4%), as compared to the city sites (1%). As pointed out in Section 3.1, the Middletown area contains numerous metal-working industrial facilities, which result in the elevated concentrations of trace metals.

The seasonal variation of $PM_{2.5}$ concentrations in all three sites followed the same patterns. The concentration tended to increase during both winter and summer. This pattern can be explained by the influence of regional and local meteorology. The seasonal fluctuation of the $PM_{2.5}$ concentration is mainly associated with changing concentrations of sulfate and nitrate. While sulfate concentrations increased in summer, nitrate concentrations peaked in winter (Fig. 5). Increased formation of secondary sulfate during summer is driven by the increased solar radiation and photochemical activity, as well as by higher ambient temperature and humidity. These factors, along with an increased emission rate from local industrial sources, might have caused a significant increase of $PM_{2.5}$ and elemental concentrations in the MID site during several days at the end

of July, 2003. Abnormally high concentrations of some trace metals (2.3 µg m^{-3} for Sn and 5.2 µg m^{-3} for As), which were recorded during those days, caused an increase in the average relative contribution of trace elements to the mass balance. At
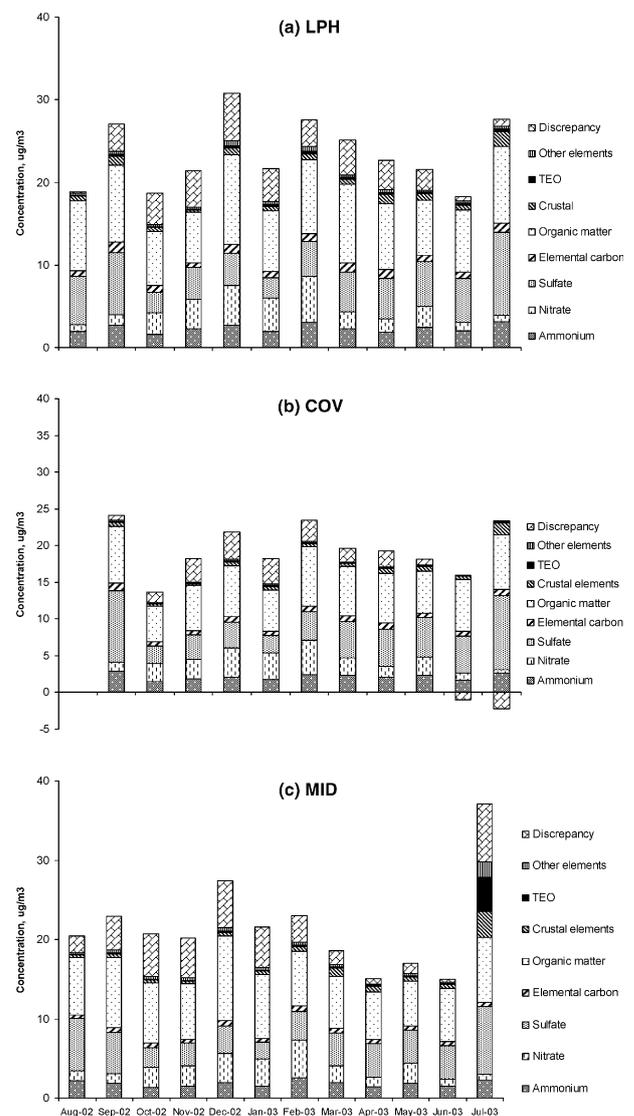
**Fig. 5** $PM_{2.5}$ monthly average concentration and composition in three speciation sites.

Table 3 Contribution of various temporal factors to the total temporal variation in the PM$_{2.5}$ concentration (%)

Component of variance	Monitoring site		
	NOR	VER	CAR
Year	0.00	0.00	0.00
Season	59.9	65.7	65.8
Month	13.7	4.71	5.31
Week	19.5	20.9	20.7
Day	6.96	8.63	8.15

the same time, the concentrations of S and OC increased 2–3 times compared to other days of the same month. Since these concentration levels remained elevated during the month of August, we do not attribute this increase to a possible sample contamination. Most likely, the emissions of a local coal combustion facility, rich in As and Sn, influenced the PM$_{2.5}$ composition. An increase in the nitrate concentration observed in winter is influenced by the higher stability and the decreased volatility of ammonium nitrate, NH$_4$NO$_3$, during cooler months. The discrepancy between the reconstructed mass and the actual mass measured by the gravimetric method might have appeared due to the loss of particle-bound water during filter conditioning.

In order to quantify the temporal variation of the PM$_{2.5}$ concentration, the variance components analysis was performed using a nested model. The temporal factor was split into five sub-factors: year, season, month, week and day, as described in Section 2.3. The data of three FRM sites, measuring the PM$_{2.5}$ concentration each day (NOR, VER and CAR) were analysed. The results, presented in Table 3, indicate that during three years of sampling, the PM$_{2.5}$ concentration variation from year to year was negligible in all sampling sites. The seasonal variation was the greatest fraction (59.9 to 65.8%) of the total variation and weekly variations accounted for 19.5 to 20.7% of the total. Month-to-month and day-to-day variations had lower contributions. All three stations revealed similar temporal variability patterns. Thus, the variance components analysis of temporal variation by site also supported the finding of a low spatial variation, which was obtained using the regression analysis.

In addition to the analysis of temporal variation of 24 h PM$_{2.5}$ concentrations, the hourly measurement data were examined for the maximum concentrations in order to compare the performance of the filter-based FRM monitors and the real-time instruments. Fig. 6 demonstrates increased pollution represented by the fine particulates during the holiday fireworks, as was measured in the TAF site in July 2002. The PM$_{2.5}$ concentration reached 110 µg m^{-3} and remained elevated for 6 h. However, the FRM monitor showed no concentration increase. It is concluded that in the event of increased air pollution, the real-time monitor provides a better
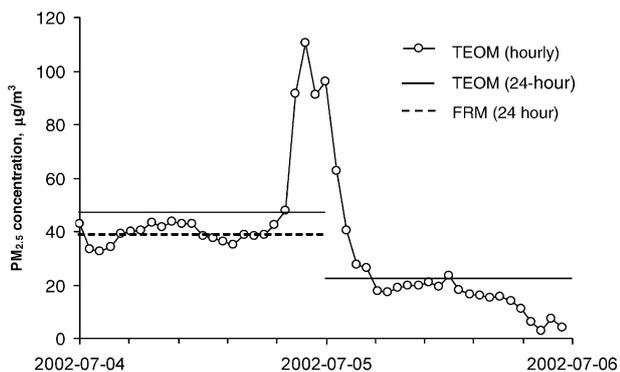
Fig. 6 Increase of hourly PM$_{2.5}$ concentration due to a local pollution event, as measured in the TAF site.

indication of the levels of exposure to ambient particulate matter.

3.4 Optimization of the spatial distribution of PM$_{2.5}$ monitoring sites

The low spatial variation of the daily averaged PM$_{2.5}$ concentrations implies that the number of sites equipped with filter-based FRM monitors may be reduced without considerable loss in the measurement precision of the overall average value. To evaluate the feasibility of the reduction of the number of monitoring sites, variance components and cluster analyses were applied to the 24 h PM$_{2.5}$ data collected every third day during the 3 year period (COV, FTH, LPH, TAF, NOR, WIN, STB, CAR, OAK, SHS, VER and WWS).

Fig. 7(a) shows the effect of decreasing the number of monitoring sites (N) on the relative precision (RP) and relative cost (RC). The results were normalized assuming that the current network configuration represents the highest precision and costs. Obviously, reducing the number of sites should reduce the relative precision. For example, if N decreases from 12 to 9 sites, RP drops by 7% and RC drops by 27%; if N decreases to 6 sites, RP drops by 16% and RC drops by 53%. It must be acknowledged that the cost of measurements and number of sites do not follow a linear relationship in practice. Thus, the estimation of the drop in the relative cost may not be precise.

While this type of analysis may give a good estimate of the precision drop and cost reduction by considering scenarios involving different numbers of sites, it does not imply how sites can be grouped and merged spatially. The optimization of spatial distribution of monitoring sites can be achieved using cluster analysis. The results of the cluster analysis are presented in Fig. 7(b). This type of analysis allows for a spatial grouping of stations according to the similarity in partitioning the total variance. The values of the X axis (proportion of variance explained by the first principal component) are used for computational purposes and do not provide a useful quantitative estimate in terms of selection of clusters. However, the values close to one support findings on low spatial variability determined by other methods. The dendogram indicates that the primary clusters are formed by sites NOR and TAF; CAR, STB and WIN; COV and FTH; VER and WWS. These clusters match the spatial allocation of the measurement sites (Fig. 1). Further (secondary) clustering involves stations which are situated within a relatively close distance. The MID site, which contains a speciation monitor and a PM$_{2.5}$ mass monitor operating every sixth day, could serve as a location for the one every third day measurement in the city of Middletown.

By using the two analyses, the optimization of the number of sites can be effectively addressed. However, neither of the analyses indicates where the "cut" can be made most effectively. The final decision on the number of sites, which is supposed to be made based on the theoretical approach described above, has to be validated with respect to practical considerations (not directly incorporated into the analyses in this paper). The latter include the local pollutions sources surrounding the sites, as well as population density and geographical factors.

The following clusters of sites were proposed to be merged: NOR and TAF; CAR, STB, and WIN; and VER and WWS. The merger is represented by the vertical dashed line in Fig. 7(b). This means that the monitoring network would consist of 8 FRM samplers measuring the PM$_{2.5}$ concentration every third day. The relative precision of the measurement results would decrease by 9%, while the relative costs would decrease by 36% [see Fig. 7(a)].

A further reduction of the number of sites could include the Northern Kentucky stations COV and FTH. However, such a modification deals with administrative issues since the above
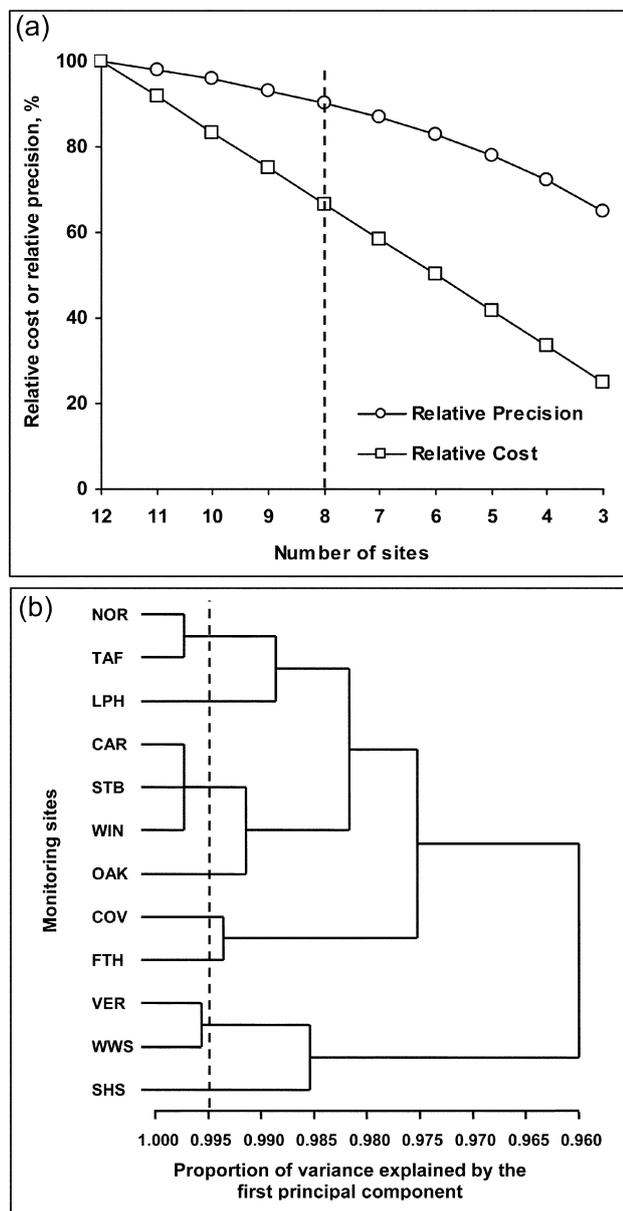
Fig. 7 The decrease in relative precision (*RP*) and relative cost (*RC*) by minimizing the number of monitoring sites compared to the current design as calculated by: (a) variance components analysis and (b) cluster analysis. The dashed line represents the proposed cut in the number of sites.

stations are operated by a different authority (KY DEP). The sites forming secondary clusters could also be merged spatially, but this does not seem practically feasible because they represent different community monitoring zones.

The analyses confirm the finding that local pollution sources do not have a considerable influence on the 24 h $PM_{2.5}$ concentrations. However, the concentrations of the $PM_{2.5}$ constituents and the hourly $PM_{2.5}$ concentrations do not necessarily follow this pattern, as shown in Section 3.1. Three speciation monitors and three real-time mass monitors provide data that better represent the cases of the peak concentrations of $PM_{2.5}$ and its constituents. Thus, while decreasing the number of $PM_{2.5}$ FRM monitors, the spatial resolution of speciation and real-time monitors should be increased.

Conclusions

This study was undertaken to analyze data from three years of urban $PM_{2.5}$ monitoring, with the goal of evaluating and optimizing the spatial distribution of $PM_{2.5}$ monitors. The

Greater Cincinnati and Northern Kentucky $PM_{2.5}$ monitoring network was evaluated by estimating spatial and temporal variations of the $PM_{2.5}$ 24 h integrated and hourly concentrations, as well as the 24 h $PM_{2.5}$ chemical composition. The optimization was performed using variance components and cluster analyses. The results indicate that the 24 h $PM_{2.5}$ concentration across the area was distributed rather uniformly and was mostly influenced by regional transport of ammonium sulfate and organic carbon. In contrast, the trace element concentrations had a high spatial variation. The hourly $PM_{2.5}$ concentration also demonstrated rather low spatial variation, although, higher than that of daily concentrations.

The temporal analysis indicated that the highest variation of $PM_{2.5}$ concentrations occurred from season to season followed by week-to-week; yearly, monthly and daily variations were relatively low. The real-time monitors were capable of detecting peak increases of $PM_{2.5}$ concentration within a day, while these fluctuations in aerosol concentration were missed by the FRM monitors. Thus, although the real-time monitors may have measurement errors due to their measurement principles, the spatial density of these monitors should be increased in order to better represent the spatial distribution of $PM_{2.5}$ concentrations.

The optimization of the spatial distribution of the FRM monitors indicated that several measurement sites can be merged spatially without any sizable loss in precision. Our proposed reduction of the number of FRM sites includes merging 3 clusters of sites, thus reducing the number of stations from 12 to 8. This scenario would decrease the relative precision by only 9%, while the relative cost would drop by 36%. A further reduction in the number of stations is possible, which would only moderately decrease the measurement precision of the $PM_{2.5}$. This can be attributed to the low spatial variation of $PM_{2.5}$ mass concentration across the area. However, the spatial density of speciation monitors, similarly to the real-time $PM_{2.5}$ mass monitors, should be increased to better represent the spatial and temporal variation of local pollution source markers.

References

1   US EPA, 40 CFR, Part 50, Federal Register Vol. 62, No. 138, 1997.
2   US EPA, 40 CFR, Part 53, Federal Register Vol. 62, No. 138, 1997.
3   US EPA, 40 CFR, Part 58, Federal Register Vol. 62, No. 138, 1997.
4   US EPA, *Guidance For Network Design and Optimum Site Exposure For PM$_{2.5}$ and PM$_{10}$*, EPA-454/R-99-022, Research Triangle Park, NC, 1997.
5   J. C. Chow, J. P. Engelbrecht, J. G. Watson, W. E. Wilson, N. H. Frank and T. Zhu, *Chemosphere*, 2002, **49**(9), 961–978.
6   R. M. Burton, H. H. Suh and P. Koutrakis, *Environ. Sci. Technol.*, 1996, **30**, 400–407.
7   A. T. DeGaetano and O. M. Doherty, *Atmos. Environ.*, 2004, **38**, 1547–1558.
8   J. R. Brook, A. H. Wiebe, S. A. Woodhouse, C. V. Audette, T. F. Dann, S. Callaghan, M. Piechowski, E. Dabek-Zlotorzynska and J. F. Dloughy, *Atmos. Environ.*, 1997, **31**, 4223–4236.

9   I. Orlic, X. Wen, T. H. Ng and S. M. Tang, *Nucl. Instrum. Methods*, 1999, **B150**, 457–464.

10 M. A. Pohjola, A. Kousa, J. Kukkonen, J. Härkönen, A. Karppinen, P. Aarnio and T. Koskentalo, *Water Air Soil Pollut.*, 2002, **Focus 2**, 5–6.

11 R. Gehrig and B. Buchmann, *Atmos. Environ.*, 2003, **37**2571–2580.

12 K. L. Demerjian, *Atmos. Environ.*, 2000, **34**, 1861–1884.

13 P. Paatero, P. K. Hopke, J. Hoppenstock and S. I. Eberly, *Environ. Sci. Technol.*, 2003, **37**, 2460–2476.

14 US EPA, *National Air Quality and Emissions Trends Report. 2003 Special Studies Edition*, EPA 454/R-03-005, Research Triangle Park, NC, 1997.

15 L. Morawska, K. Vishakarman, K. Mengersen and S. Thomas, *Atmos. Environ.*, 2002, **36**, 3545–3555.

16 C. Silva and A. Quiroz, *Atmos. Environ.*, 2003, **37**, 2337–2345.

17 D. Martuzevicius, S. A. Grinshpun, T. Reponen, R. Górny, R. Shukla, J. Lockey, S. Hu, R. McDonald, P. Biswas, L. Kliucininkas and G. LeMasters, *Atmos. Environ.*, 2004, **38**, 1091–1105.

18 US Census Bureau, County population datasets, http://www.census.gov.

19 US EPA, 40 CFR, Part 50, Appendix L, Federal Register Vol. 62, No. 138, 1997.

20 US EPA, *Quality Assurance Guidance Document 2.12*, Research Triangle Park, NC, 1998.

21 A. Charron, R. M. Harrison, S. Moorcroft and J. Booker, *Atmos. Environ.*, 2004, **38**, 415–423.

22 H. Hauck, A. Berner, B. Gomiscek, S. Stopper, H. Puxbaum, M. Kundi and O. Preining, *J. Aerosol Sci.*, 2004, **35**, 1135–1149.

23 H. Hauck, A. Berner, T. Frischer, B. Gomiscek, M. Kundi, M. Neuberger, H. Puxbaum and O. Preining, *Atmos. Environ.*, 2004, **38**, 3905–3915.

24 G. K. Lemasters, R. Shukla, Y. D. Li and J. Lockey, *J. Occup. Environ. Med.*, 1996, **38**, 39–45.

25 C. Peretz, P. Goldberg, E. Kahan, S. Grady and A. Goren, *Ann. Occup. Hyg.*, 1997, **41**, 485–500.

26 SAS Institute Inc., SAS OnlineDoc®, Version 8, Cary, NC, 1999.

27 Health Effects Institute, *Emissions from diesel and gasoline engines measured in highwaytunnels*, Boston, MA, 2002.

28 S. L. Rees, A. L. Robison, A. Khlystov, C. Stanier and S. N. Pandis, *Atmos. Environ.*, 2004, **38**, 3305–3318.

29 R. A. VanCuren and T. A. Cahill, *J. Geophys. Res.*, 2002, **107**(D24), 4804, doi: 10.1029/2002JD002204.

30 B. J. Turpin and H. J. Lim, *Aerosol. Sci. Technol.*, 2001, **35**, 602–610.