TECHNICAL PAPER

ISSN 1047-3289 *J. Air & Waste Manage. Assoc.* **56**:876–888

Effects of Instrument Precision and Spatial Variability on the Assessment of the Temporal Variation of Ambient Air Pollution in Atlanta, Georgia

Katherine S. Wade, James A. Mulholland, Amit Marmur, and Armistead G. Russell
School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA

Ben Hartsell and Eric Edgerton
Atmospheric Research & Analysis, Inc., Cary, NC

Mitch Klein, Lance Waller, Jennifer L. Peel, and Paige E. Tolbert
Rollins School of Public Health, Emory University, Atlanta, GA

ABSTRACT

Data from the U.S. Environmental Protection Agency Air Quality System, the Southeastern Aerosol Research and Characterization database, and the Assessment of Spatial Aerosol Composition in Atlanta database for 1999 through 2002 have been used to characterize error associated with instrument precision and spatial variability on the assessment of the temporal variation of ambient air pollution in Atlanta, GA. These data are being used in time series epidemiologic studies in which associations of acute respiratory and cardiovascular health outcomes and daily ambient air pollutant levels are assessed. Modified semivariograms are used to quantify the effects of instrument precision and spatial variability on the assessment of daily metrics of ambient gaseous pollutants (SO_2, CO, NO_x, and O_3) and fine particulate matter ([$PM_{2.5}$] $PM_{2.5}$ mass, sulfate, nitrate, ammonium, elemental carbon [EC], and organic carbon [OC]). Variation because of instrument imprecision represented 7–40% of the temporal variation in the daily pollutant measures and was largest for the $PM_{2.5}$ EC and OC. Spatial variability was greatest for primary pollutants (SO_2, CO, NO_x, and EC). Population-weighted variation in daily ambient air pollutant levels because of both instrument imprecision and spatial variability ranged from 20% of the temporal variation for O_3 to 70% of the temporal variation for SO_2 and EC. Wind rose plots, corrected for diurnal and seasonal pattern effects, are used to demonstrate the impacts of local sources on monitoring station data. The results presented are being used to quantify the impacts of instrument precision and spatial variability on the assessment of health effects of ambient air pollution in Atlanta and are relevant to the interpretation of results from time series health studies that use data from fixed monitors.

IMPLICATIONS

A quantitative understanding of errors associated with assessment of the temporal variation in ambient air pollution because of instrument imprecision and spatial variability is needed to improve the assessment and interpretation of health risk from time series studies that use data from ambient air pollutant monitors. This work demonstrates that these errors can be substantial and vary widely between pollutants. Primary pollutants are less spatially uniform than secondary pollutants, indicating greater uncertainty in exposure estimates based on ambient concentrations of primary pollutants than secondary pollutants and greater potential for attenuation of health risk results.

INTRODUCTION

In epidemiologic time series studies in which the short-term health effects of ambient air pollution are assessed, instrument error and the spatial variability of air pollution can impact the assessment. A number of studies have addressed the limitations of using central monitoring station data as exposure estimates.[1–5] In population-based epidemiologic studies, however, whereas differences between personal exposures and central station values can be large, the daily mean of personal exposures is likely to be better correlated with a central station value than an individual exposure level.[6] Introduction of uncertainty in the exposure measurement tends to reduce the ability of epidemiologic studies to assess the health effects of air pollution, decreasing the strength of association estimate (bias to null). The attenuation varies across pollutants, because instrument error, which includes instrument bias and instrument precision, and spatial variation differ across pollutants. In population-based time series studies of acute health effects and ambient air pollution, risk ratios are typically estimated per standard increment (either the standard deviation or the interquartile range) of the temporal distribution of ambient pollutant measures to facilitate comparison across multiple air pollutants. Meteorological variables and day-of-week, seasonal, and long-term trends are generally controlled for in the health models. Therefore, a quantitative understanding of the error in estimates of the day-to-day variation of air pollutant levels is needed to better compare results for multiple pollutants.

Results of the Aerosol Research Inhalation Epidemiology Study (ARIES; ref. 7) of emergency department visits

for cardiovascular and respiratory diseases in relation to ambient air pollution in Atlanta, GA, from 1993 to 2000 have been published recently.[8,9] For cardiovascular disease, positive associations were observed with nitrogen dioxide (NO_2), CO, fine particulate matter ($PM_{2.5}$), and $PM_{2.5}$ components organic carbon (OC) and elemental carbon (EC). Risk ratios ranged from 1.02 to 1.03 per standard deviation increase for these pollutant measures. For respiratory disease, positive associations were observed with ozone (O_3), NO_2, CO, and coarse particulate matter (PM_{10}). Central monitoring station data were used in this assessment, including data obtained since August 1998, from the Southeastern Aerosol Research and Characterization Study (SEARCH) site near downtown Atlanta at Jefferson Street (JS).[10] In this paper, the authors address the issue of two types of error in the estimates of the day-to-day variation of air pollutant levels in Atlanta: error associated with instrument imprecision and error associated with spatial variability. Instrument imprecision includes that for a single type of instrument and analytical protocol, as well as that stemming from the use of different instruments or analytical procedures for a single analyte. In the context of the time series health study, instrument accuracy is not a source of error if the error results in a systematic bias but not an error in the assessment of temporal variation. That is, a systematic bias may affect the magnitude of the standard deviation of temporal variation but will not affect the estimated risk ratio or confidence interval per standard deviation. However, differences between instruments and analytical procedures that affect the precision of assessment of temporal variation are sources of error. Factors that affect the spatial heterogeneity of air pollution include the distribution of emission sources, as well as meteorological phenomena, topological features, and pollutant volatility and reactivity.

Instrument precision can be quantified using data from colocated instruments. For continuous and semicontinuous measurements, instrument error typically results from calibration drift, flow rate changes, and changes in atmospheric conditions, such as relative humidity. These errors are minimized through adherence to a stringent quality control protocol. For filter-based measurements, such as PM mass, ions, OC, and EC, instrument error also results from sample handling artifacts and laboratory analysis.

A number of studies have examined the spatial variability of air pollutants in urban areas. When data from a sufficient number of monitoring stations are available, spatial interpolation techniques can be used to provide a spatially continuous representation. In previous work, the authors used a universal kriging procedure for estimating daily O_3 concentrations for each zip code in the 20-county Atlanta metropolitan statistical area (MSA).[11] However, the area of representativeness for a given monitor for a given pollutant can lead to uncertainty in a study.[5–7,11,12] Several studies have used Pearson correlation coefficients to characterize spatial variability in ambient air pollutant concentrations.[13–17] Other studies have used analysis of variance to characterize variability in air pollutant concentrations between specific monitoring sites.[18–20] Analysis of variance provides a measure of variation between monitoring sites relative to variation over time at one monitoring site and is particularly useful when several sites with different average concentrations are considered. Alternatively, the semivariogram, which uses the covariance instead of the correlation, is a well-established tool for conveying information about the spatial variability of environmental pollutants.[21–25] Diem[26] has argued against the use of semivariogram analysis for air pollution, because there are often not enough monitoring stations available to provide a suitable number of points for estimation. In particular, the Atlanta MSA would need data from 100 monitoring stations to create a stable estimate of the semivariogram for O_3. Obviously, this is an unrealistic expectation. Diem[26] suggests that spatial modeling can be used to surmount this obstacle. In this study, however, the authors wish to provide a metric that can be used within existing monitoring networks to assess the spatial accuracy of these networks when applied to continuous surfaces, for example, the epidemiologic study area. In such instances, the exact point of exposure is often not known, and, instead, a central or averaged value is used. The semivariograms developed in this study can be used to qualify the accuracy of this central value.

Although ambient air quality monitoring stations are sited to minimize local source effects in most instances, impacts of specific point and roadway sources are observed, particularly for primary pollutants. For example, Duncan et al.[27] showed the contribution of power plant plume fumigation events on nitrogen oxides ($[NO_x] = NO + NO_2$) and sulfur dioxide (SO_2) concentrations in Atlanta. Kirby et al.[28] showed the contribution of roadways to NO_x. Secondary air pollutants, such as O_3, and a significant fraction of $PM_{2.5}$ mass exhibit high spatial autocorrelation. From analyses of the spatial variability of $PM_{2.5}$ in several urban areas in the southeast, Pinto et al.[29] found high correlations between site pairs and spatial uniformity in concentration fields. In a study of spatial aerosol composition in Atlanta in 1999, Butler et al.[30] concluded that $PM_{2.5}$ mass and major constituents were relatively spatially homogeneous.

In this paper, the authors address the following two questions. First, for each ambient air pollutant of interest in time series studies of the short-term health effects of air pollution in Atlanta, what is the error in the estimate of ambient air pollutant concentration because of instrument imprecision and spatial variability relative to the temporal variation in this ambient air pollutant? Second, to what extent do local and specific point sources impact measurements at each of the ambient air quality monitoring stations in Atlanta?

EXPERIMENTAL WORK

Ambient Air Quality Monitoring Station Data

Ambient air quality data from U.S. Environmental Protection Agency (EPA) Air Quality System ([AQS] formerly known as Aerometric Information Retrieval System), the SEARCH database, and the Assessment of Spatial Aerosol Composition in Atlanta (ASACA) are used to characterize instrument precision and spatial variability of air pollution in the 20-county Atlanta MSA for the 4-yr period 1999–2002; Atlanta monitoring site locations are shown in Figure 1. The JS site is of particular focus here, because

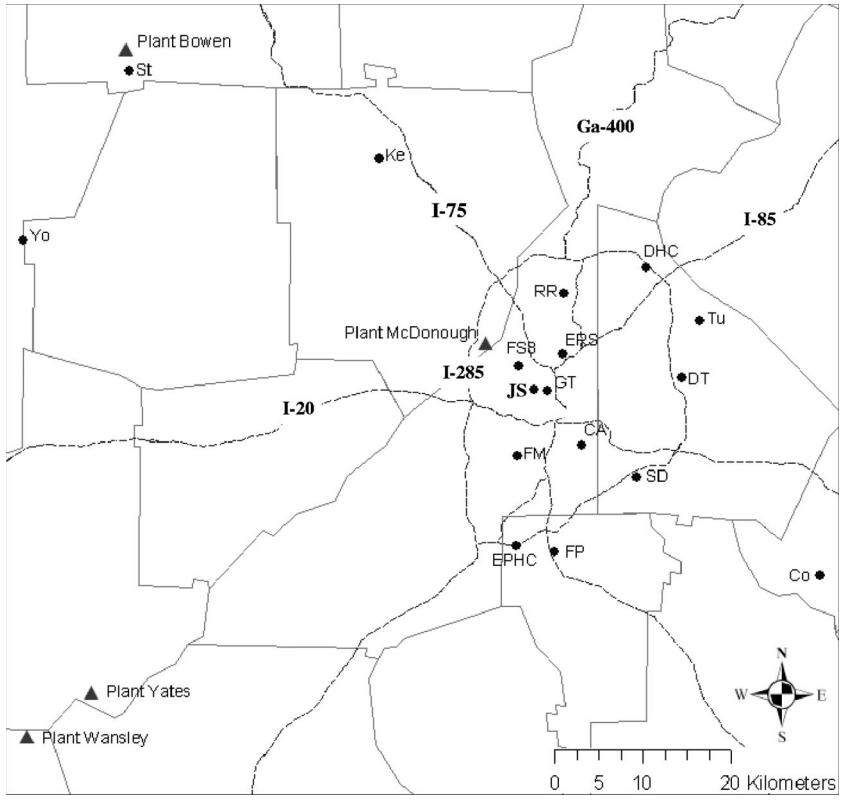

Figure 1. Locations of 20-county Atlanta metropolitan area ambient air quality monitoring sites (●) and coal-fired power plants (▲). Total area shown: 100 × 100 km. County boundaries (—) and interstate highways (----) are also shown. GT = Georgia Tech; ERS = East River School; FS8 = Fire Station 8; FM = Fort McPherson; CA = Confederate Avenue; RR = Roswell Road; DHC = Doraville Health Center; Tu = Tucker; DT = Dekalb Tech; SD = South Dekalb; FP = Forest Park; EPHC = East Point Health Center; Co = Conyers; Ke = Kennesaw; Yo = Yorkville; and St = Stilesboro.

it was the location of the Atlanta Supersite, and Atlanta-based epidemiologic studies rely heavily on data from that location. Pollutants used in this study, for which daily measures were available at multiple sites in the Atlanta MSA, are four pollutant gases (SO_2, CO, NO_x, and O_3), $PM_{2.5}$ total mass, and five $PM_{2.5}$ components (sulfate ion [SO_4^{2-}], nitrate ion [NO_3^-], ammonium ion [NH_4^+], EC, and OC). The completeness of the dataset was very high (>90% for most pollutants). The data for each pollutant at each monitoring site closely approximate a log-normal distribution. Values of the geometric mean, geometric standard deviation, and correlation coefficients with log-transformed data from the JS monitoring station near downtown Atlanta are listed in Tables 1 and 2. The daily metrics listed are the a priori variables used in the health studies: 1-hr maximum for SO_2, CO, and NO_x; 8-hr maximum for O_3; and 24-hr average for $PM_{2.5}$ total mass and major component masses.

SO_2 data from five monitoring sites in the Atlanta MSA are available. SO_2 levels are low in Atlanta, on average between 3 and 6 ppb. Levels at JS were higher than the two AQS sites located near downtown Atlanta, Georgia Tech and Confederate Avenue, possibly because of differences in analytical methods of drying the gas stream and calibration procedures. Unlike the AQS sites, calibration gases are added to the ambient airstream at JS. Thus, the SO_2 measurement at JS accounts for losses in sampling lines. CO data from four monitoring stations are available. Highest CO levels are observed at Roswell Road, which is located near high traffic density roads. NO_x are measured at six sites in the Atlanta MSA. Within the perimeter highway of Atlanta, average NO_x levels exceed 40 ppb. NO_x levels are appreciably lower at the rural Conyers and Yorkville sites. O_3 data from five stations were used. During the period 1999 through 2002, AQS sites only reported data from April through October. O_3 levels are spatially very uniform and highly correlated across sites.

To assess instrument precision, NO_x and O_3 data from AQS and SEARCH monitors at the Yorkville site were used. To assess CO and SO_2 instrument error, AQS audit data were used. Audit tests are performed quarterly by an independent contractor, with standards at three different concentrations run through each sampler.

$PM_{2.5}$ mass data from eight AQS stations, two SEARCH sites, and three ASACA sites were used. AQS stations at East River School, South Dekalb, and Doraville Health Center monitor daily $PM_{2.5}$ mass by the filter-based Federal Reference Method (FRM). AQS stations at Fire Station 8, East Point Health Center, Forest Park, Kennesaw, and Yorkville report $PM_{2.5}$ mass by FRM every 3 days. At the ASACA sites at South Dekalb, Tucker, and Fort McPherson, $PM_{2.5}$ mass is measured continuously by tapered element oscillating microbalance (TEOM). At the SEARCH sites at JS and Yorkville, $PM_{2.5}$ mass is measured both by FRM and TEOM. $PM_{2.5}$ mass instrument precision includes differences in the method of measurement. Data from collocated samplers at East River School, Doraville

Table 1. Average values and correlation coefficients of ambient air pollutants, 1999–2002.

Pollutant Measure	Station[a]	Distance (km)	Geometric Mean	Geometric Standard Deviation	R
SO_2	JS–SEARCH	0	12.4 ppb	2.1	
1-hr max	GT	1.5	8.0 ppb	2.4	0.78
	CA	8.3	6.5 ppb	2.3	0.71
	St	58.6	5.9 ppb	2.9	0.36
	Yo–SEARCH	60.6	6.5 ppb	2.6	0.20
CO	JS–SEARCH	0	0.96 ppm	2.2	
1-hr max	RR	11.5	1.43 ppm	1.6	0.65
	DT	16.8	1.15 ppm	1.8	0.76
	Yo–SEARCH	60.6	0.26 ppm	1.4	0.24
NO_x	JS–SEARCH	0	97.5 ppb	2.1	
1-hr max	GT	1.5	89.1 ppb	2.3	0.86
	SD	15.3	113.6 ppb	2.0	0.74
	Tu	20.4	51.2 ppb	2.0	0.79
	Co	36.8	35.0 ppb	3.4	0.50
	Yo–SEARCH	60.6	9.1 ppb	2.3	0.26
O_3	JS–SEARCH	0	50.0 ppb	1.5	
8-hr max	CA	8.3	50.6 ppb	1.5	0.98
April to October	SD	15.3	46.6 ppb	1.5	0.96
	Co	36.8	49.8 ppb	1.5	0.91
	Yo–SEARCH	60.6	57.4 ppb	1.4	0.85
$PM_{2.5}$ mass	JS–SEARCH	0	15.6 $\mu g/m^3$	1.6	
24-h average	ERS	5.2	16.1 $\mu g/m^3$	1.6	0.87
	SD	15.3	15.6 $\mu g/m^3$	1.7	0.84
	DHC	18.9	16.5 $\mu g/m^3$	1.6	0.88
	FS8–3rd day	3.2	17.8 $\mu g/m^3$	1.7	0.62
	EPHC–3rd day	17.9	16.0 $\mu g/m^3$	1.8	0.65
	FP–3rd day	18.7	16.2 $\mu g/m^3$	1.7	0.87
	Ke–3rd day	31.6	15.9 $\mu g/m^3$	1.7	0.77
	Yo–3rd day	60.6	13.6 $\mu g/m^3$	1.8	0.72
	JS–TEOM	0	18.1 $\mu g/m^3$	1.6	0.91
	FM–TEOM	7.8	17.4 $\mu g/m^3$	1.6	0.83
	SD–TEOM	15.3	16.6 $\mu g/m^3$	1.5	0.91
	Tu–TEOM	20.4	17.6 $\mu g/m^3$	1.7	0.77
	Yo–SEARCH	60.6	11.7 $\mu g/m^3$	1.7	0.85

[a]Station abbreviations and locations shown in Figure 1.

Table 2. Average values and correlation coefficients of $PM_{2.5}$ components, March 1999–August 2000.

$PM_{2.5}$ Component	Station[a]	Distance (km)	Geometric Mean ($\mu g/m^3$)	Geometric Standard Deviation	R
Sulfate	JS–SEARCH	0	4.6	1.9	
	FM–ASACA	7.8	4.3	2.1	0.82
	SD–ASACA	15.3	4.4	2.0	0.84
	Tu–ASACA	20.4	4.3	2.1	0.92
	Yo–SEARCH	60.6	4.5	2.1	0.95
Nitrate	JS–SEARCH	0	0.8	2.1	
	FM–ASACA	7.8	0.7	2.3	0.63
	SD–ASACA	15.3	0.3	2.4	0.64
	Tu–ASACA	20.4	0.8	2.6	0.70
	Yo–SEARCH	60.6	0.6	2.0	0.55
Ammonium	JS–SEARCH	0	1.8	1.9	
	FM–ASACA	7.8	1.6	2.1	0.86
	SD–ASACA	15.3	1.5	2.2	0.82
	Tu–ASACA	20.4	1.7	2.1	0.81
	Yo–SEARCH	60.6	2.4	1.9	0.82
EC	JS–SEARCH	0	1.4	1.9	
	FM–ASACA	7.8	1.2	1.8	0.47
	SD–ASACA	15.3	1.5	1.9	0.47
	Tu–ASACA	20.4	1.2	1.8	0.29
	Yo–SEARCH	60.6	0.7	1.7	0.57
OC	JS–SEARCH	0	4.0	1.7	
	FM–ASACA	7.8	3.8	2.6	0.52
	SD–ASACA	15.3	4.2	2.4	0.50
	Tu–ASACA	20.4	3.8	2.5	0.47
	Yo–SEARCH	60.6	3.2	1.7	0.71

[a]Station abbreviations and locations shown in Figure 1.

Health Center, and Forest Park was used to evaluate instrument precision.

$PM_{2.5}$ major ion (SO_4^{-2}, NO_3^-, and NH_4^+) and carbon fraction (EC and OC) data from March 1999 through August 2000 at five locations in the Atlanta MSA were used (Table 2). Filter-based particle composition monitors (PCMs) were used that included three channels to collect 24-hr integrated samples for analysis of major ions, trace metals, OC, and EC in the $PM_{2.5}$ size range. Ion chromatography was used to quantify water-soluble ionic species, except for ammonium, which was measured using automated colorimetry. EC and OC collected on quartz filters were measured by thermal optical transmittance (TOT) in the ASACA network (South Dekalb, Tucker, and Fort McPherson) and by thermal optical reflectance (TOR) in the SEARCH network (JS and Yorkville). Comparison of these two techniques indicates that whereas the total carbon measurements are in good agreement, EC is lower, and OC is higher for TOT values.[31] At the JS site, multiple measurements of $PM_{2.5}$ components using different methods were made. These data were used to assess instrument precision. For the major ion components, simultaneous PCM and FRM measurements were used. For EC and OC, TOR measurements performed by Atmospheric Research & Analysis, Inc. (ARA), were compared with TOR measurements performed by the Desert Research Institute (DRI).

Spatial Statistical Analysis

Using methods from the field of geostatistics, the authors developed an approach for describing the spatial autocorrelation of air pollutant data. Air pollutant concentration data (C) typically have a lognormal distribution. Here, they take the log transform of the data (Z) and normalize based on the mean (μ_z) and standard deviation (σ_z) of the log-transformed data. That is:

$$Z = \ln C \tag{1}$$

$$Z' = \frac{Z - \mu_z}{\sigma_z} \tag{2}$$

The distribution of normalized data (Z') has a mean of 0 and a standard deviation of 1. Given the strong correlation found between the TOT and TOR techniques for measuring EC and OC,[32] this normalization eliminates most of the bias between these measures, as well as potential bias because of differences in the SEARCH and AQS SO_2 and NO_x monitors.

For a pair of monitoring sites at locations *i* and *j*, the correlation coefficient *R* is defined as follows:

$$R(h) = \frac{1}{n}\sum_{k=1}^{n}\frac{(Z_{i,k} - \mu i)(Z_{i,k} - \mu j)}{\sigma_i \sigma_j} = \frac{1}{n}\sum_{k=1}^{n}(Z'_{i,k} \times Z'_{j,k}) \quad (3)$$

Here, *n* is the number of daily observations, and *h* is the distance between *i* and *j*. Spatial autocorrelation is an attribute of spatial data based on the fact that observations closer together tend to be more alike than observations farther apart. If the spatial process is isotropic, then the correlation coefficient is a function of distance alone (independent of direction). The correlogram plot, a graph of *R* versus the distance *h*, provides a measure of spatial autocorrelation as a function of distance. For small distances, *R* is close to 1; as distance increases, *R* decreases, approaching 0 as the spatial correlation approaches 0.

Another measure of spatial autocorrelation is the semivariogram.[33] The semivariogram (γ) is one half of the variogram:

$$\gamma(h) = \frac{1}{2} Var(Z'_i - Z'_j) \quad (4)$$

It can be shown that the semivariogram for the normalized log-transformed data (Z') is related to the correlation coefficient as follows:

$$\gamma(h) = 1 - \frac{n}{n-1} \times R(h) \cong 1 - R(h) \quad (5)$$

Thus, the graph of the semivariogram versus distance is an inverted form of the correlogram, starting near 0 at short distances if the correlation is close to 1, and increasing with distance as the correlation decreases with distance. In the geostatistical field, the shape of the semivariogram is described by the nugget, which refers to a nonzero semivariogram near the origin, the range, which refers to the separation distance at which the semivariogram levels off, and the sill, which refers to the value of the semivariogram at which it levels off.

In this paper, the authors propose a measure of spatial autocorrelation that can most easily be interpreted in terms of potential impact of measurement error on the epidemiologic study findings. In this population-based epidemiologic study of the short-term health effects of air pollution in Atlanta, instrument precision and spatial variation represent uncertainty in the ambient pollutant measure used as the exposure surrogate, and temporal variation in daily air pollutant levels is a key determinant of the power of time series studies to detect epidemiologic associations. Therefore, the authors are interested in the ratio of the standard deviation of the differences between two sites to the standard deviation of the averages of two sites. For low values of this ratio, the impact of exposure variable uncertainty on health risk assessment is expected to be low. For high values, the impact is likely to be a bias to the null in the point estimate of health risk.

The daily average of the normalized variables at two monitoring sites is $(Z'_i + Z'_j)/2$, and the difference of the measurement at each site to this average is given by $(Z'_i - Z'_j)/2$. The ratio of the standard deviation of the differences to the standard deviation of the averages is the desired metric. It can be shown that these standard deviations are given as follows:

$$\sigma_{average}(h) = \sqrt{1 - \frac{1}{2}\gamma(h)} \quad (6)$$

$$\sigma_{difference}(h) = \sqrt{\frac{1}{2}\gamma(h)} \quad (7)$$

Thus, for large *n*,

$$\gamma'(h) = \frac{\sigma_{difference}(h)}{\sigma_{average}(h)} \cong \sqrt{\frac{1 - R(h)}{1 + R(h)}} \quad (8)$$

Because a plot of this ratio versus distance between monitoring stations has the same general shape as a semivariogram plot, the authors refer to it as a modified semivariogram, $\gamma'(h)$. Its value is 0 for a correlation coefficient of 1 and 1 for a correlation coefficient of 0. This modified semivariogram can be interpreted as the amount of ambient estimate error, because of instrument precision and spatial variability, relative to the amount of temporal variation in the air pollutant data.

The modified semivariogram can be modeled using an exponential function as follows.

$$\gamma'(h) = c_o + c_e\{1 - \exp(-h/a_e)\} \quad (9)$$

Here, c_o is the nugget effect, c_e is the partial sill ($c_o + c_e$) is the sill, and the effective range is $3a_e$.[34] The nugget represents a measure of instrument precision, with a value of 0 for colocated measurements that are exactly correlated and increasing with increasing imprecision. The partial sill represents a measure of spatial variability in the limit of monitors separated by a large distance, and the effective range is the distance at which 95% of the partial sill is reached. Zero values for the nugget and sill mean that the log-transformed, normalized measurements are identical, and, therefore, there is no error because of instrument imprecision and spatial variability over the study area. Values between 0 and 1 correspond with error because of instrument imprecision and spatial variability between 0 and 100% of the temporal variation of the pollutant modeled. Values >1, which correspond with anticorrelated measurements, indicate that error because of instrument imprecision and spatial variability data is greater than the temporal variation of the pollutant of interest.

To assess the impact of point sources and local road sources on the ambient air pollutant monitors, wind rose plots are used. Hourly measurements of pollutants and wind direction were used for the 4-yr period 1999–2002. Wind data bins were 12°. More than 35,000 data points were possible; data completeness was ~90%. Because both wind direction and pollutant concentrations exhibit

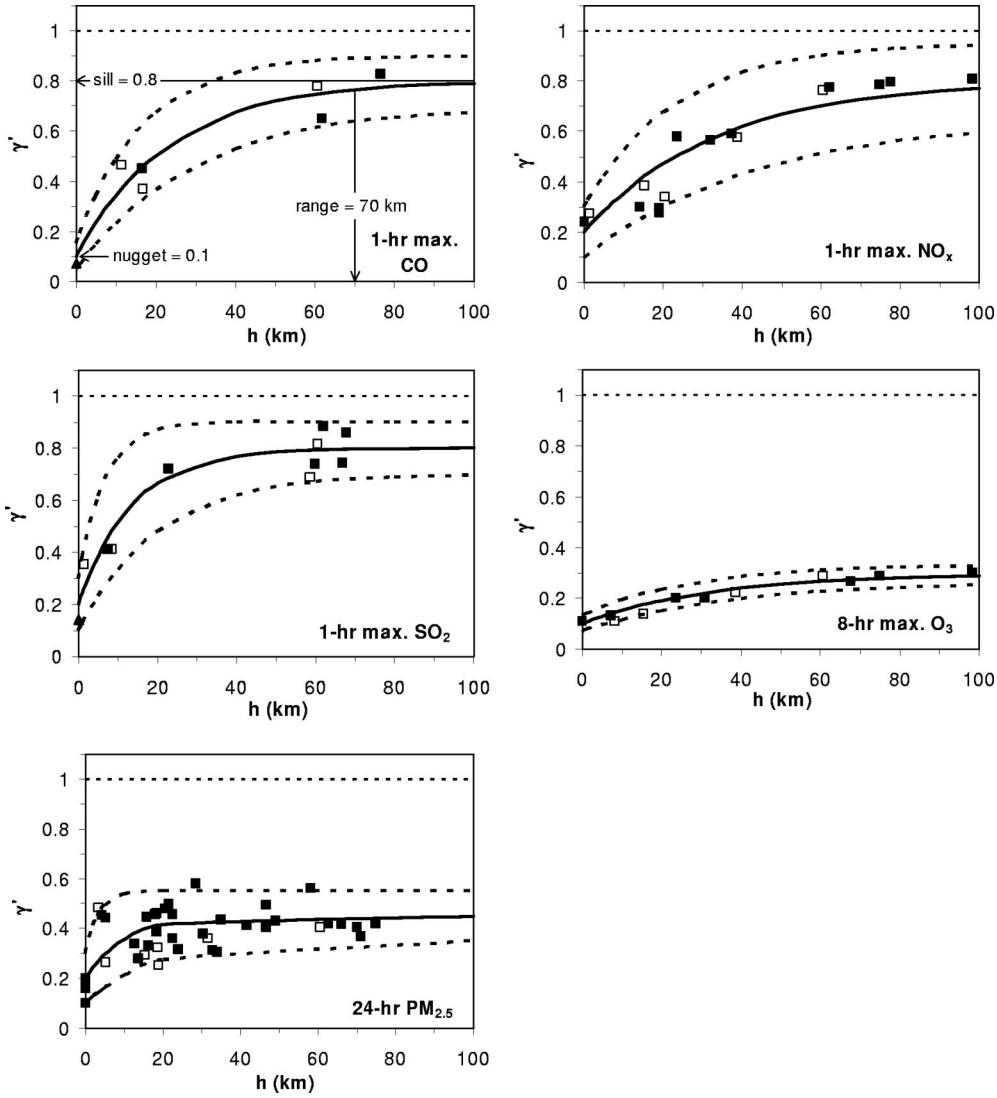

Figure 2. Modified semivariograms for ambient air pollutant measures in Atlanta, 1999–2002. ○, points that include JS data. ▲, use of audit data. Exponential model results are shown by solid curve (best fit) and dashed curves (upper and lower bound estimates); model parameters (nugget, range, sill) are shown on the CO semivariogram and listed in Table 3.

diurnal and seasonal patterns, effects of these patterns were estimated and removed. To estimate possible effects of diurnal and seasonal patterns in wind direction and pollutant levels on the wind rose plots, wind rose plots were constructed using the actual wind direction data and average monthly and diurnal (twelve 2-hr averages) concentrations. Deviation of these wind rose plots from a perfect circle indicate effects of confounding of wind direction and pollutant level because of seasonal and diurnal trends. Plots corrected for diurnal and seasonal patterns were generated by subtracting the seasonal and diurnal deviations from the original wind rose plot using raw data. With seasonal and diurnal effects removed, high concentrations in the wind rose plots suggest the direction of sources that impact the station.

RESULTS AND DISCUSSION

Instrument Precision and Spatial Variation Effects

Modified semivariograms for 1-hr maximum CO, 1-hr maximum NO_x, 1-hr maximum SO_2, 8-hr maximum O_3, and 24-hr average $PM_{2.5}$ mass are shown in Figure 2. Exponential model best-fit values and upper and lower bounds of the nugget (c_o), sill ($c_o + c_e$), and effective range ($3a_e$) are estimated for each of the semivariogram plots (Table 3).

The nugget or instrument imprecision estimate as assessed using colocated instrument data ranged from 10% of the temporal variation for CO and O_3 to 20% for SO_2, NO_x, and $PM_{2.5}$ mass. The greater instrument imprecision

Table 3. Parameter estimates for semivariogram functions of gaseous pollutants and $PM_{2.5}$ mass.

Pollutant	Nugget	Effective Range	Sill	Population Weighted γ′
1-hr max SO_2	0.2 ± 0.1	40 ± 20	0.8 ± 0.1	0.7 ± 0.1
1-hr max CO	0.1 ± 0.05	70 ± 20	0.8 ± 0.1	0.6 ± 0.1
1-hr max NO_x	0.2 ± 0.1	100 ± 30	0.8 ± 0.1	0.6 ± 0.1
8-hr max O_3	0.1 ± 0.03	100 ± 20	0.3 ± 0.03	0.2 ± 0.03
24-hr $PM_{2.5}$ mass	0.2 ± 0.1	30 ± 20	0.45 ± 0.1	0.4 ± 0.1

for SO_2, NO_x, and $PM_{2.5}$ mass is likely because of differences in measurement techniques among the AQS, SEARCH, and ASACA networks for these pollutants.

As expected, the primary pollutants (CO, NO_x, and SO_2) exhibited much greater spatial variability than the predominantly secondary pollutants (O_3 and $PM_{2.5}$ mass). That is, estimates of the sills of the primary pollutant semivariograms are 0.8, indicating that, at large distances, the uncertainty in the primary pollutant variable because of instrument precision and spatial variability is ~80% of the temporal variation in that pollutant. Of these three primary pollutant gases, the effective range in the semivariogram was smallest for SO_2 (40 km). As will be shown later, this is likely because of point-source plume events that impact SO_2 measurements at the ambient monitoring stations. The larger effective ranges for CO (70 km), NO_x (100 km), and O_3 (100 km) are likely because of the predominance of dispersed mobile sources for CO and NO_x and the regional atmospheric formation of O_3. Sills for O_3 and $PM_{2.5}$, on the other hand, are estimated to be 0.3 and 0.45, respectively. O_3 is a secondary pollutant, formed by atmospheric reactions. Source apportionment modeling indicates that 60–70% of $PM_{2.5}$ mass at JS is secondary.[35,36] Because concentrations of these pollutants are driven more by regional meteorology and chemistry than by specific sources, spatial variation for the secondary pollutants accounts for a smaller fraction of temporal variation than for the primary pollutants. Although values of O_3 were lower at the South Dekalb monitoring site because of titration from local NO sources, as described later, the O_3 correlation with other monitoring sites is strong.

Modified semivariograms for 24-hr integrated measurements of $PM_{2.5}$ major ions (SO_4^{-2}, NO_3^-, and NH_4^+)

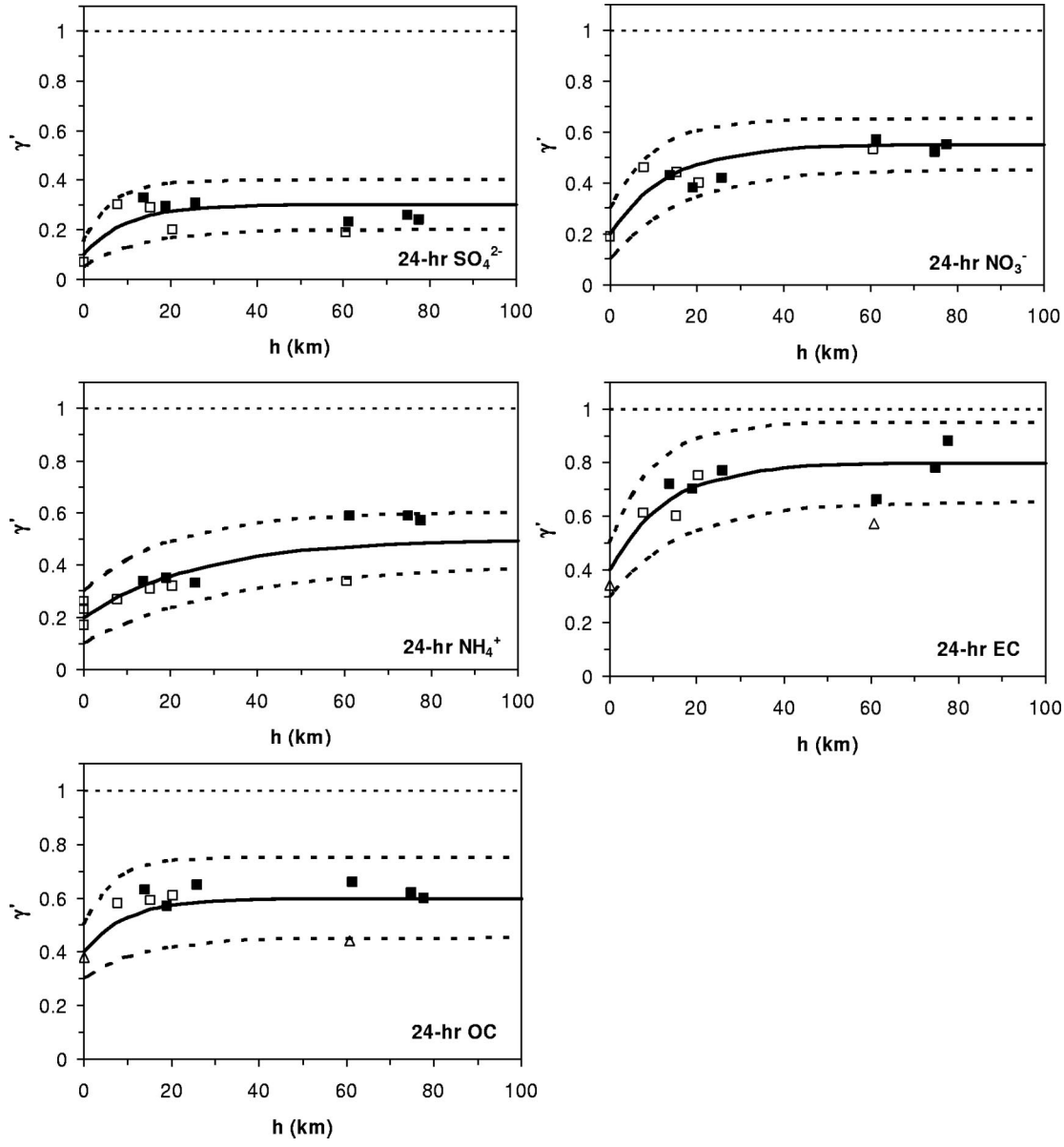

Figure 3. Modified semivariograms for $PM_{2.5}$ components in Atlanta, March 1999 to August 2000. Hollow symbols, points that include JS data. ▲, TOR method used for both EC and OC measurements. Exponential model results are shown by solid curve (best fit) and dashed curves (upper and lower bound estimates); model parameters (nugget, range, sill) are listed in Table 4.

Table 4. Parameter estimates for semivariogram functions of $PM_{2.5}$ components.

$PM_{2.5}$ Component (24-hr)	Nugget	Effective Range	Sill	Population Weighted γ'
SO_4^{2-}	0.1 ± 0.05	30 ± 10	0.3 ± 0.1	0.25 ± 0.1
NO_3^-	0.2 ± 0.1	40 ± 10	0.55 ± 0.1	0.5 ± 0.1
NH_4^+	0.2 ± 0.1	80 ± 20	0.5 ± 0.1	0.4 ± 0.1
EC	0.4 ± 0.1	40 ± 10	0.8 ± 0.15	0.7 ± 0.15
OC	0.4 ± 0.1	30 ± 10	0.6 ± 0.15	0.55 ± 0.15

and carbon fractions (EC and OC), shown in Figure 3 and Table 4, are consistent with the findings cited above. The semivariogram nuggets are higher for the EC and OC measurements, likely because of differences in the TOR analysis protocol used by ARA and DRI. The semivariogram sills for the predominantly secondary components of $PM_{2.5}$ (SO_4^{-2}, NO_3^-, NH_4^+, and, to some extent, OC) are lower than the sill for the primary $PM_{2.5}$ component (EC). Source apportionment modeling suggests that ~50% of OC mass at JS is secondary organic aerosol, although there is high uncertainty associated with this estimate.[36] The EC and OC semivariogram plots also suggest that error in the ambient estimate results from differences between the TOR method (SEARCH data) and the TOT method (ASACA data). EC and OC semivariograms from the SEARCH site pair (JS and Yorkville, hollow triangles at 60 km in the EC and OC plots in Figure 3) are lower than from other pairs. Comparison of SEARCH data, obtained using the TOR method, and more recent data from EPA Speciation Trends Network, using the TOT method, suggest that EC and OC measurement precision is improving.

The semivariograms shown in Figures 2 and 3 indicate that no one station appears to be an outlier, suggesting that local and regional influences play similar roles at each site. Data from the JS monitoring station, used for the points represented by hollow symbols in Figures 2 and 3, were used for the primary analyses in the epidemiologic studies because of the wide range of measures available at this site and its central location. These data appear to be as representative as data from any of the other monitoring stations in the Atlanta area for the pollutants shown.

On the other hand, the semivariogram analysis suggests that error in the ambient air pollutant measures used in the health studies is likely to reduce the power of health risk assessment, particularly for primary pollutants. As a reference point for comparing ambient estimate error, the authors computed population-weighted semivariograms for the metropolitan Atlanta area; these are listed in the last columns of Tables 3 and 4. Fifty-seven percent of the patients in the emergency department database for the 20-county area (and a similar proportion of the 20-county population) live within a 30-km radius of downtown. For primary pollutant variables (SO_2, CO, NO_x, and EC), instrument precision and spatial variability result in a population-weighted uncertainty of ~60–70% of the temporal variation. For O_3, on the other hand, uncertainty in the ambient concentration is only ~20% of the temporal variation. Work is ongoing to incorporate these ambient air pollution error estimates into the health risk models.

Point and Local Source Impacts

To assess the impact of point sources and local road sources on the ambient air pollutant monitors, wind rose plots (both uncorrected and corrected for seasonal and diurnal effects) were constructed for those pollutants for which hourly data are available (Figures 4–8). Hourly measurements of pollutants and wind direction were used for the 4-yr period 1999–2002. Two sites did not have wind direction data available for this time period, Georgia Tech and Roswell Road. JS wind direction data were used for the Georgia Tech analysis because of the proximity of these sites, and Confederate Avenue data were used for the Roswell Road analysis because of the similarity in

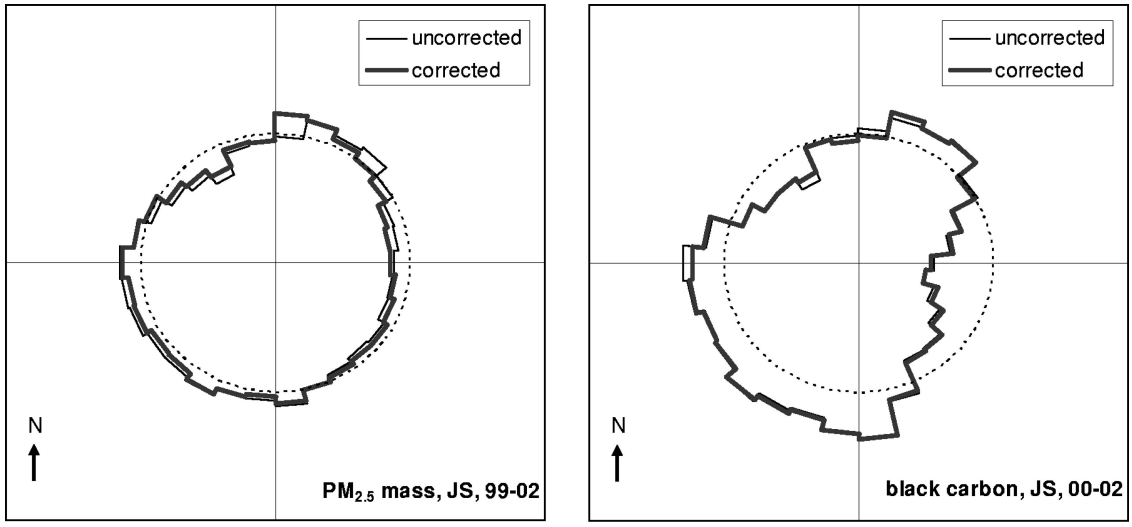

Figure 4. Wind rose plots for $PM_{2.5}$ mass and the BC component of $PM_{2.5}$ measured hourly at JS monitoring station, 1999–2002. Seasonal and diurnal effects are removed from corrected curves. Dashed circle, average value. Full scale is 32 $\mu g/m^3$ for $PM_{2.5}$ mass and 2.5 $\mu g/m^3$ for BC.

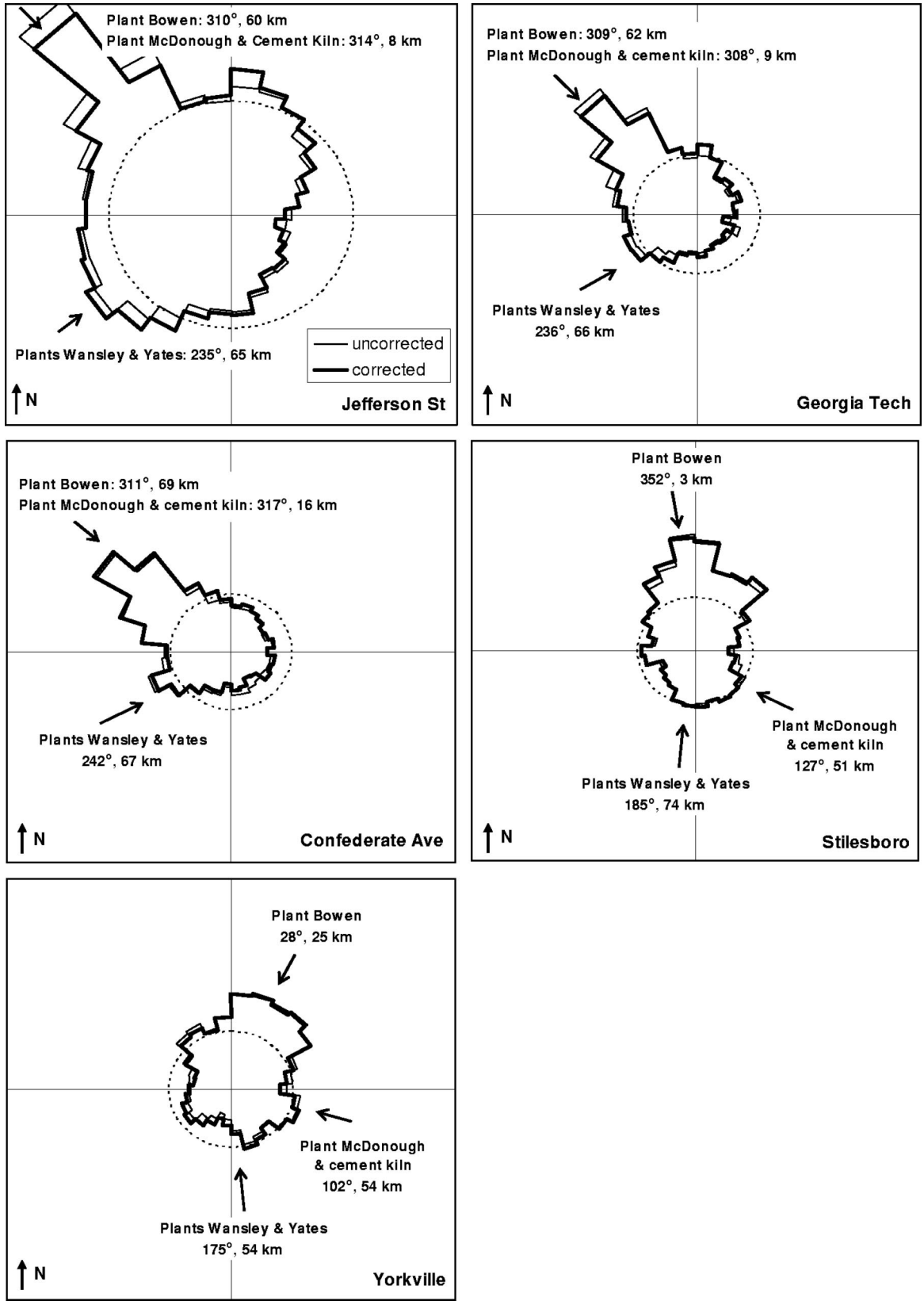

Figure 5. SO_2 wind rose plots, 1999–2002. Full scale is 10 ppb. Dashed circle, average value. Distances and directions of coal-fired power plants and a coal-fired cement facility from monitoring stations are shown.

topography at these sites. With seasonal and diurnal effects removed, high concentrations in the wind rose plots suggest the direction of sources that impact the station. The degree to which sources impact a monitoring site is also an indicator of the degree to which the pollutant fields are isotropic. That is, greater source impact results in

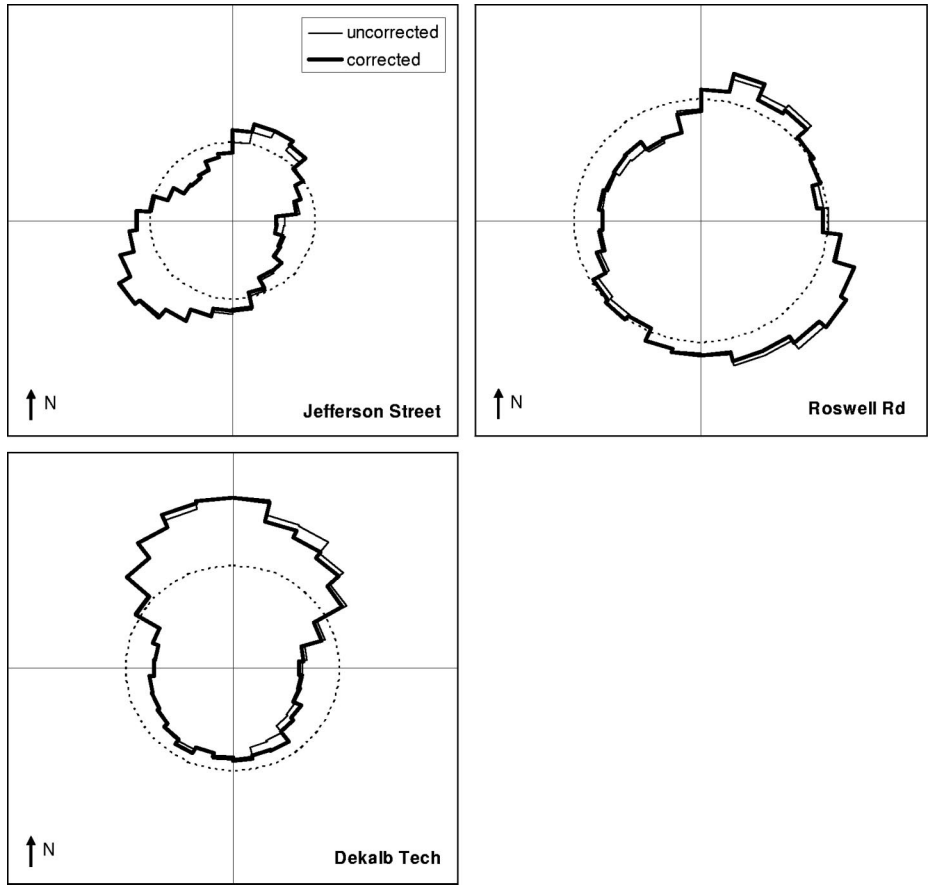

Figure 6. CO wind rose plots for 1999–2002. Full scale is 1.5 ppm.

more scatter in the semivariogram plots because of anisotropic pollutant fields.

In Figure 4, wind rose plots at the JS monitoring station are shown for $PM_{2.5}$ mass and the black carbon (BC) component of $PM_{2.5}$. The latter was measured by aethalometer and should correlate approximately with EC.[37,38] Plots of SO_2, CO, NO_x, and O_3 are shown in Figures 5–8, respectively.

The plots of CO, NO_x, and BC at JS are similar. Possible sources of these pollutants are roadways to the south and west, a trucking facility and major rail yard to the north, and a bus maintenance facility to the south. The peak in these pollutants to the northeast is consistent with the alignment of two major highways (I-85 and Ga-400; Figure 1).

As expected, the plots of O_3 and $PM_{2.5}$ mass do not show strong effects of wind direction. O_3 is a secondary pollutant, and, as stated previously, 60–70% of $PM_{2.5}$ mass is estimated to be secondary.[35,36] Secondary pollutants are less affected by emission sources than primary pollutants. The authors do note, however, that O_3 minima are observed in directions where NO_x peaks occur. This is likely because of O_3 inhibition by radical scavenging and the titration of O_3 by NO, indicating NO_x inhibition. O_3 tends to peak on weekends at JS when NO_x levels are lower, consistent with NO_x inhibition.

The SO_2 wind rose plot at JS has a large peak when winds come from the northwest and smaller peaks when winds come from the southwest and north. The peak corresponding with the winds from the north may be because of the trucking facility. Analysis of the peak SO_2 concentrations in the 0–12° category indicates that the peaks occur most frequently in the morning when activity at the trucking facility is greatest. The northwest and southwest peaks typically occur in late afternoon. These peaks are likely because of vertical mixing effects on coal combustion plumes (Figure 5), as evidenced by the alignment of SO_2 peaks at each monitoring site with the direction of coal-fired power plants and a coal-fired cement facility. Power plant emissions are injected above the atmospheric mixed layer at night and can be downwardly mixed after the breakup of the nocturnal inversion. Plant Bowen, located 60 km northwest of Atlanta, is the largest of the coal-fired power plants; Plant McDonough, located much closer to downtown Atlanta and also to the northwest, is the smallest of these sources. There is also a coal-fired cement kiln near Plant McDonough.

Observed CO and NO_x peaks in these pollutants are consistent with the directions of major roadways (Figures 6 and 7). The data show the extent to which the primary pollutant monitoring data are impacted by local source effects. In the case of NO_x at South Dekalb, the large peak to the north is likely because of close proximity of two major highways (I-285 and I-20) in that direction (Figure 1).

Finally, O_3 wind rose plots using data from four monitoring stations are shown in Figure 8. These plots demonstrate minimal local source impacts on O_3, as expected. As already mentioned, NO_x inhibition likely accounts for

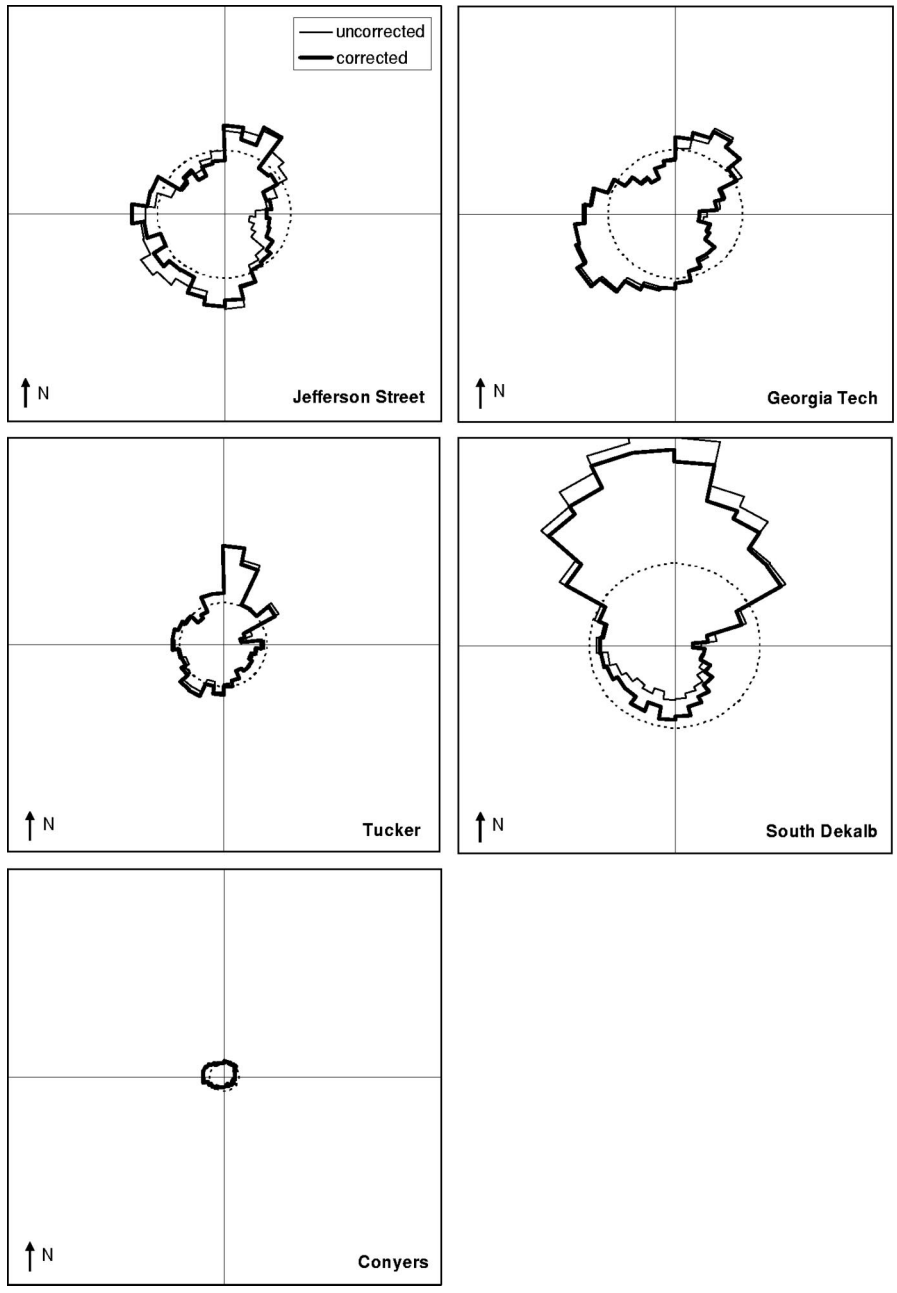

Figure 7. NO_x wind rose plots for 1999–2002. Full scale is 150 ppb.

the shape of the O_3 wind rose plot at the JS station near downtown Atlanta. Given the large NO_x peak associated with winds from I-285 north of South Dekalb, one might have expected more NO_x inhibition in the O_3 wind rose plot at SD. One explanation is that the time-of-day correction removed evidence of source impacts as well. Note that the uncorrected wind rose plot does show lower O_3 levels when winds are from the north. The diurnal pattern of wind direction at South Dekalb is quite strong, with 80% of the winds from the north at night and 80% of the winds from the south during the day. The authors tested their correction by using JS O_3 data with the South Dekalb wind direction data and found the correction to be reliable. A more plausible explanation, therefore, is that the more suburban South Dekalb site is less NO_x inhibited than the more urban JS site because of greater biogenic volatile organic compound emissions at South Dekalb.

CONCLUSIONS

Effects of instrument precision and spatial variability on the characterization of ambient air pollution in Atlanta have been assessed. For quantifying error in the ambient air pollution variables used in large population time series health analyses, a modified semivariogram is proposed. Population-weighted uncertainty in primary pollutant levels because of instrument imprecision and spatial variability was found to be 60–70% of the temporal variation. As expected, secondary pollutants are much more spatially homogeneous and correlated than primary pollutants, with a population-weighted uncertainty of only

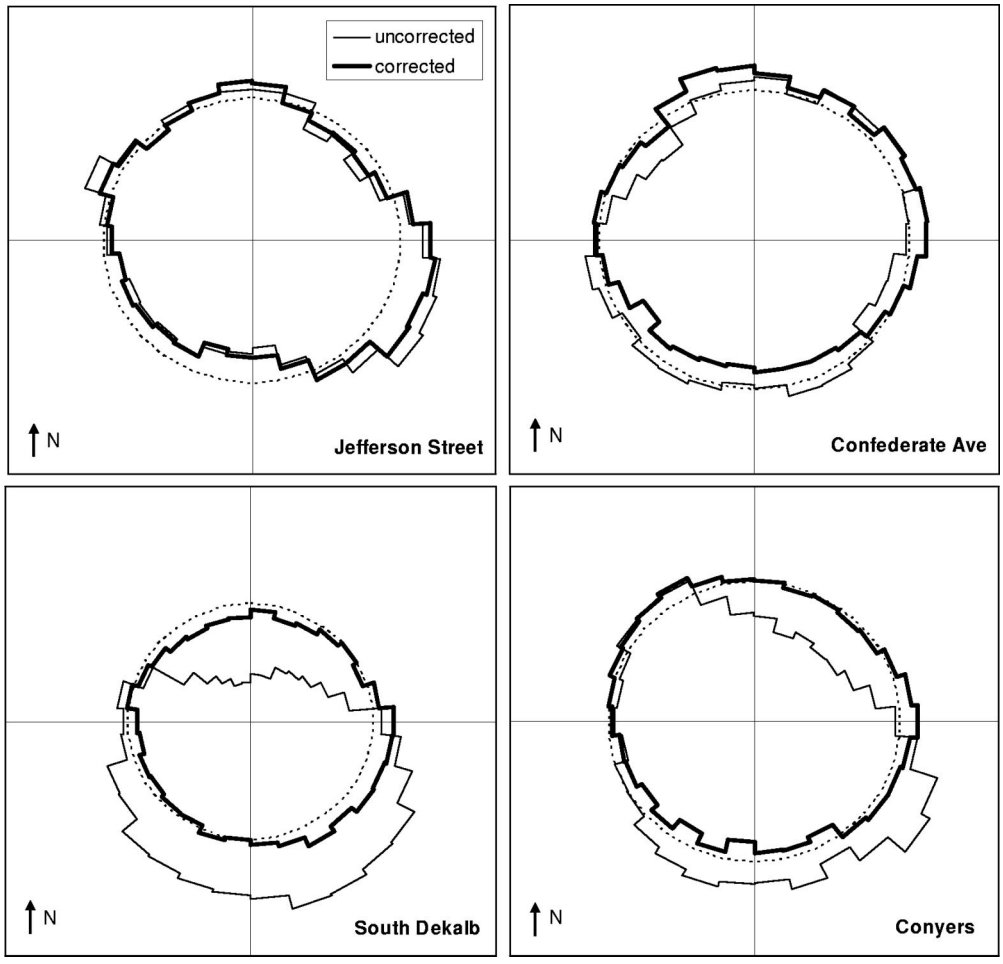

Figure 8. O_3 wind rose plots for April through October, 1999–2002. Full scale is 50 ppb.

20% for O_3. Error because of instrument imprecision was greatest for air pollutants measured by different analytical techniques, such as the elemental and OC fractions of $PM_{2.5}$. Local source impacts for primary air pollutants were identified at several Atlanta monitoring stations using wind rose plots with adjustments to remove diurnal and seasonal pattern effects. The use of spatial average values, obtained using interpolation and population-weighting methods, rather than data from one central monitoring station, can dampen these local impacts on exposure measures used in population-based epidemiologic studies of the health effects of air pollution.

Exposure variable error estimates are needed to interpret findings from epidemiologic studies of ambient air pollution and health. Work is ongoing to incorporate the ambient estimate errors presented in this paper into the health risk models.

ACKNOWLEDGMENTS

This work was supported by grants from the U.S. Environmental Protection Agency (R82921301-0), National Institute of Environmental Health Sciences (R01ES11294 and R01ES11199), and Electric Power Research Institute (EP-P4353/C2124). This research used air quality data from a monitoring station operated by Aerosol Research Inhalation Epidemiology Study and managed by Ron Wyzga and Alan Hansen of Electric Power Research Institute.

REFERENCES

1. Goldstein, I.F.; Landovitz, L. Analysis of Air Pollution Patterns in New York City—2. Can One Aerometric Station Represent the Area Surrounding It?; *Atmos. Environ.* **1977**, *11,* 53-57.
2. Lipfert, F.W.; Wyzga, R.E. Air Pollution and Mortality: The Implications of Uncertainties in Regression Modeling and Exposure Measurement; *J. Air & Waste Manage. Assoc.* **1997**, *47,* 517-523.
3. Zeger, S.L.; Thomas, D.; Dominici, F.; Samet, J.M.; Schwartz, J.; Dockery, D.; Cohen, A. Exposure Measurement Error in Time-Series Studies of Air Pollution: Concepts and Consequences; *Environ. Health Perspect.* **2000**, *108,* 419-426.
4. McNair, L.A.; Harley, R.A.; Russell, A.G. Spatial Inhomogeneity in Pollutant Concentrations, and Their Implications for Air Quality Model Evaluation; *Atmos. Environ.* **1996**, *30,* 4291-4301.
5. Ito, K., Kinney, Ito, K.; Kinney, P.L.; Thurston, G.D. Variations in PM10 Concentrations within Two Metropolitan Areas and Their Implications for Health Effects Analyses. *Inhal. Toxicol.* **1995**, *7,* 735-745.
6. Schwartz, J.; Dockery, D.W.; Neas, L.M. Is Daily Mortality Associated Specifically with Fine Particles?; *J. Air & Waste Manage. Assoc.* **1996**, *46,* 927-939.
7. Van Loy, M.; Bahadori, T.; Wyzga, R.; Hartsell, B.; Edgerton, E. The Aerosol Research and Inhalation Epidemiology Study (ARIES): PM2.5 Mass and Aerosol Component Concentrations and Sampler Intercomparisons; *J. Air Waste Manage. Assoc.* **2000**, *50,* 1446-1458.
8. Metzger, K.B.; Tolbert, P.E.; Klein, M.; Peel, J.L.; Flanders, W.D.; Todd, K.; Mulholland, J.A.; Ryan, P.B.; Frumkin, H. Ambient Air Pollution and Cardiovascular Emergency Department Visits; *Epidemiol.* **2004**, *15,* 46-56.
9. Peel, J.L.; Tolbert, P.E.; Klein, M.; Metzger, K.B.; Flanders, W.D.; Todd, K.; Mulholland, J.A.; Ryan, P.B.; Frumkin, H. Ambient Air Pollution and Respiratory Emergency Department Visits; *Epidemiol.* **2005**, *16,* 184-174.
10. Hansen, D.A.; Edgerton, E.S.; Hartsell, B.E.; Jansen, J.J.; Kandasamy, N.; Hidy, G.M.; Blanchard, C.L. The Southeastern Aerosol Research and Characterization Study: Part 1—Overview; *J. Air & Waste Manage. Assoc.* **2003**, *53,* 1460-1471.

11. Mulholland, J.A.; Butler, A.J.; Wilkinson, J.; Russell, A.G.; Tolbert, P.E. Temporal and Spatial Distributions of Ozone in Atlanta: Regulatory and Epidemiologic Implications; *J. Air & Waste Manage. Assoc.* **1998**, *48,* 418-426.
12. Tiles, S.; Zimmerman, J. Investigation on the Spatial Scales of the Variability in Measured Near-Ground Ozone Mixing Ratios; *Geophys. Res. Lett.* **1998**, *25,* 3827-3830.
13. Monn, C.; Carabias, V.; Junker, M.; Waeber, R.; Karrer, M.; Wanner, H.U. Small-Scale Spatial Variability of Particulate Matter < 10 μm (PM_{10}) and Nitrogen Dioxide; *Atmos. Environ.* **1997**, *31*, 2243-2247.
14. Buzorius, G.; Hameri, K.; Pekkanen, J.; Kulmala, M. Spatial Variation of Aerosol Number Concentration in Helsinki City; *Atmos. Environ.* **1999**, *33,* 553-565.
15. Morawska, L.; Vishvakarman, D.; Mengersen, K.; Thomas, S. Spatial Variation of Airborne Pollutant Concentrations in Brisbane, Australia and Its Potential Impact on Population Exposure Assessment; *Atmos. Environ.* **2002**, *36,* 3545-3555.
16. Lin, T.; Young, L.; Wang, C. Spatial Variations of Ground Level Ozone Concentrations in Areas of Different Scales; *Atmos. Environ.* **2001**, *35,* 5799-5807.
17. Rao, S.T.; Zalewsky, E.; Zurbenko, I.G. Determining Temporal and Spatial Variations in Ozone Air Quality; *J. Air & Waste Manage. Assoc.* **1995**, *45,* 57-61.
18. Hoek, G.; Meliefste, K.; Cyrys, J.; Lewne, M.; Bellander, T.; Brauer, M.; Fischer, P.; Gehring, U.; Heinrich, J.; van Vliet, P.; Brunekreef, B. Spatial Variability of Fine Particle Concentrations in Three European Areas; *Atmos. Environ.* **2002**, *36,* 4077-4088.
19. Lebret, E.; Briggs, D.; Reeuwijk, H.; Fischer, P.; Smallbone, K.; Harssema, H.; Kriz, B.; Gorynski, P.; Elliott, P. Small Area Variations in Ambient NO 2 Concentrations in Four European Areas; *Atmos. Environ.* **2000**, *34,* 177-185.
20. Roosli, M.; Theis, G.; Kunzli, N.; Staehelin, J.; Mathys, P.; Oglesby, L.; Camenzind, M.; Braun-Fahrlander, Ch. Temporal and Spatial Variation of the Chemical Composition of PM_{10} at Urban and Rural Sites in the Basel Area, Switzerland; *Atmos. Environ.*, **2001**, *35,* 3701-3713.
21. Grondona, M.O.; Cressie, N. Using Spatial Considerations in the Analysis of Experiments. *Technometrics* **1991**, *33,* 381-392.
22. Casado, L.S.; Rouhani, S.; Cardelino, C.A.; Ferrier, A.J. Geostatistical Analysis and Visualization of Hourly Ozone Data; *Atmos. Environ.* **1994**, *28,* 2105-2118.
23. Sen, Z. Regional Air Pollution Assessment by Cumulative Semivariogram Technique; *Atmos. Environ.* **1995**, *29,* 543-548.
24. Sen, Z. An Application of a Regional Air Pollution Estimation Model over Istanbul Urban Area; *Atmos. Environ.* **1998**, *32,* 3425-3433.
25. Anh, V. Duc, H.; Shannon, I. Spatial Variability of Sydney Air Quality by Cumulative Semivariogram; *Atmos. Environ.* **1997**, *31,* 4073-4080.
26. Diem, J.E. A Critical Examination of Ozone Mapping from a Spatial-Scale Perspective; *Environ. Pollut.* **2003**, *125,* 369-383.
27. Duncan, B.N.; Stelson, A.W.; Kiang, C.S. Estimated Contribution of Power Plants to Ambient Nitrogen Oxides Measured in Atlanta, Georgia in August 1992; *Atmos. Environ.* **1995**, *29,* 3043-3054.
28. Kirby, C.; Greig, A.; Drye, T. Temporal and Spatial Variations in Nitrogen Dioxide Concentrations across an Urban Landscape: Cambridge, UK; *Environ. Monit. Assess.* **1998**, *52,* 65-82.
29. Pinto, J.P.; Lefohn, A.S.; Shadwick, D.S. Spatial Variability of $PM_{2.5}$ in Urban Areas in the United States; *J. Air & Waste Manage. Assoc.* **2004**, *54,* 440-449.
30. Butler, A.J.; Andrew, M.S.; Russell, A.G. Daily Sampling of $PM_{2.5}$ in Atlanta: Results of the First Year of the Assessement of Spatial Aerosol Composition in Atlanta Study; *J. Geophys. Res.* **2003**, *108,* 8415-8426.
31. Chow, J.C.; Watson, J.G.; Crow, D.; Lowenthal, D.H.; Merrifield, T. Comparison of IMPROVE and NIOSH Carbon Measurements; *Aerosol Sci. Technol.* **2001**, *34,* 23-34.
32. Solomon, P.; Baumann, K.; Edgerton, E.; Tanner, R.; Eatough, D.; Modey, W.; Maring, H.; Savoie, D.; Natarajan, S.; Meyer, M.B.; Norris, G. Comparison of Integrated Samplers for Mass and Composition during the 1999 Atlanta Supersites Project; *J. Geophys. Res.* **2003**, *108,* SOS11-1-26.
33. Cressie, N. *Statistics for Spatial Data*. Wiley and Sons: New York, NY, 1993.
34. Waller, L.A.; Gotway, C.A. *Applied Spatial Statistics for Public Health Data;* Chapter 8: Spatial Exposure Data; Wiley and Sons: New York, NY, 2004.
35. Kim, E.; Hopke, P.K.; Edgerton, E.S. Improving Source Identification of Atlanta Aerosol Using Temperature Resolved Carbon Fractions in Positive Matrix Factorization; *Atmos. Environ.* **2004**, *38,* 3349-3362.
36. Marmur, A.; Unal, A.; Mulholland, J.A.; Russell, A.G. Optimization-Based Source Apportionment of $PM_{2.5}$ Incorporating Gas-to-Particle Rations; *Environ. Sci. Technol.* **2005**, *39,* 3245-3254.
37. Wilson, W.E.; Chow, J.C.; Claiborn, C.; Fusheng, W.; Engelbrecht, J.; Watson, J.G. Monitoring of Particulate Matter Outdoors; *Chemosphere* **2002**, *49,* 1009-1043.
38. Jeong, C.; Hopke, P.K.; Kim, E.; Lee, D. The Comparison between Thermal-Optical Transmittance Elemental Carbon and Aethelometer Black Carbon Measured at Multiple Monitoring Sites; *Atmos. Environ.* **2004**, *38,* 5193-5204.

About the Authors

Katherine Wade was a graduate research assistant at Georgia Institute of Technology when this work was done and is currently at Sonoma Technology. James Mulholland and Armistead Russell are professors in the School of Civil and Environmental Engineering at the Georgia Institute of Technology, and Amit Marmur is a graduate research assistant there and on staff at the Environmental Protection Division of the Georgia Department of Natural Resources. Lance Waller and Paige Tolbert are professors and Mitch Klein a research assistant professor at the Rollins School of Public Health at Emory University. Jennifer Peel was at Emory University as SOPHIA study director and is currently assistant professor at Colorado State University. Address correspondence to: Jim Mulholland, Environmental Engineering, Georgia Institute of Technology, 311 Ferst Dr., Atlanta, GA 30332-0512; phone: +1-404-894-1695; fax: +1-404-894-8266; e-mail: james.mulholland@ce.gatech.edu.