Journal of Exposure Analysis and Environmental Epidemiology (2001) 11, 21–32
© 2001 Nature Publishing Group All rights reserved 1053-4245/01/$17.00
www.nature.com/jea

Monitor-to-monitor temporal correlation of air pollution and weather variables in the North-Central U.S.

KAZUHIKO ITO, GEORGE D. THURSTON, ARTHUR NÁDAS AND MORTON LIPPMANN

Nelson Institute of Environmental Medicine, New York University School of Medicine, 57 Old Forge Road, Tuxedo, New York 10987

Numerous time series studies have reported associations between daily ambient concentrations of air pollution and morbidity or mortality. Recent personal exposure studies have also reported relatively high longitudinal correlation between personal exposures to particulate matter (PM) and home outdoor PM concentrations, lending support to the health effects reported in time series studies. However, the question remains as to how well the temporal fluctuations in the air pollution levels observed at an outdoor monitor represent the temporal fluctuations in the population exposures to pollution of outdoor origins in a city, and how such representativeness affects the size and significance of risk estimates. Also, such spatio-temporal correlations would vary from pollutant to pollutant, likely influencing their relative significance of statistical associations with health outcomes. In this study, we characterized the extent of monitor-to-monitor correlation over time among multiple monitoring sites for PM less than 10 μm (PM_{10}), gaseous criteria pollutants, and several weather variables in seven central and eastern contiguous states (IL, IN, MI, OH, PA, WI, and WV) during the study period of 1988–1990. After removing seasonal trends, the monitor-to-monitor temporal correlation among the air pollution/weather variables within 100-mile separation distance in these areas could be generally ranked into three groups: (1) temperature, dew point, relative humidity ($r>0.9$); (2) O_3, PM_{10}, NO_2 (r: 0.8–0.6); and (3) CO, SO_2 ($r<0.5$). Using the subsets for separation distance less than 100 miles, regression analyses of these monitor-to-monitor correlation coefficients were also conducted with explanatory variables including separation distance, qualitative (land use, location setting, and monitoring objectives) and quantitative (large and small variance) site characteristics, and region indicators for Air Quality Control Region (AQCR). The separation distance was a significant predictor of monitor-to-monitor correlation decline especially for PM_{10} and NO_2 ($\sim$0.2 drop over 30 miles). Site characteristic variables were, in some cases, significant predictors of monitor-to-monitor correlation, but the magnitude of their impacts was not substantial. Regional differences, as examined by AQCR, were in some cases (e.g., in Metropolitan Philadelphia) substantial. In these areas, the pollutants that had generally poor monitor-to-monitor correlation in the overall seven states data (i.e., for SO_2 and CO) showed higher monitor-to-monitor correlations, comparable with PM_{10} and O_3, within the AQCR. These results are useful in interpreting some of the past time series epidemiological results. The differences in monitor-to-monitor correlations found across pollutants in this work (i.e., $r\sim0.8$ *vs.* $r\sim0.4$) are sufficiently large that they could be a factor in the different pollutant significance levels reported in the epidemiologic literature. It is recommended that future epidemiological studies collect and incorporate information on spatial variability among air pollutants in the analysis and interpretation of their results. *Journal of Exposure Analysis and Environmental Epidemiology* (2001) **11,** 21–32.

Keywords: *air pollution, measurement error, ozone, particulate matter, spatial correlation, weather.*

Introduction

Numerous time series studies have reported associations between daily ambient concentrations of air pollution and morbidity or mortality (U.S. EPA, 1996a). The results from these observational studies have had a major influence on the process of establishing the National Ambient Air Quality Standards (NAAQS) because these studies produce quantitative estimates of the impact of ambient air pollution on the entire population in a given city. In these studies, the daily fluctuations of mortality or morbidity counts in the entire city or metropolitan statistical area (MSA) are regressed on the daily fluctuations of air pollution and weather data measured at a single air monitor or, at best, at a few monitors. However, many of the air monitoring stations used in these studies were mostly routine regulatory-based monitors, and their locations have not been chosen specifically for the support of epidemiological studies. Thus, the question remains as to how well the temporal fluctuations in the air pollution levels observed at a given monitor represent the temporal fluctuations in the population exposures to pollution of outdoor origins in the city, and how such representativeness affects the size and significance of risk estimates.

From the viewpoint of the total human exposures, the ambient concentrations of an air pollutant measured at a monitor may appear to only weakly contribute to the total exposures for an individual to that pollutant, especially when there are indoor sources for the pollutant in question. This is because an individual's exposure to air pollution of outdoor origin depends on such factors as indoor penetration fraction of the pollutant, air exchange, and the fraction

1. Address all correspondence to: Dr. Kazuhiko Ito, Nelson Institute of Environmental Medicine, New York University School of Medicine, 57 Old Forge Road, Tuxedo, NY 10987. Tel.: +1-914-731-3540. Fax: +1-914-351-5472. E-mail: kaz@env.med.nyu.edu
Received 15 February 2000; accepted 6 October 2000.

of the time that the individual spends outdoor. These factors can vary tremendously from person to person, so that if the relationship between personal exposures and ambient concentrations of an air pollutant is compared *across persons* (i.e., cross-sectionally) at a given point in time, a poor correlation may be observed. For example, in the PTEAM study (Ozkaynak et al., 1996), regressing personal PM_{10} exposures of 178 subjects on the ambient PM_{10} concentrations measured in the backyard of each subject resulted in a poor R^2 value (0.14). Note that this correlation is essentially cross-sectional, in that the variation in personal PM_{10} exposures *versus* ambient PM_{10} levels in the backyard across persons contributed to the overall correlation (though it was not purely cross-sectional, as the data were collected over time, up to four subjects per day). The scatter in such cross-sectional relationships depicts the person-to-person difference in the relationship between ambient concentrations and personal exposures.

When the relationship between the personal exposures and ambient concentrations is examined *over time* (i.e., longitudinally) for each individual, a higher correlation may be observed. For example, Wallace (1996), as part of a review of indoor particle exposure, re-calculated cross-sectional and longitudinal regressions using data from the THEES study of Lioy et al. (1990) and showed that higher R^2 values were observed for longitudinal regressions than for cross-sectional regressions. The regressions of personal PM_{10} exposures on outdoor PM_{10} concentrations across the 14 subjects for each day of the 14 study days (i.e., within day cross-sectional regression) resulted in low R^2 values (median=0.06, ranging from 0.00 to 0.39). However, the regression over time of personal PM_{10} exposures on outdoor PM_{10} concentrations over 14 days for each of the 14 subjects (i.e., the longitudinal regression) resulted in higher R^2 values (median=0.46, ranging from 0.02 to 0.82). Wallace discussed that the higher longitudinal correlation observed in this and other data sets was supportive of the time series health effects analyses. More recent studies that investigated the person-to-monitor temporal correlation for particulate matter (PM) (e.g., Tamura et al., 1996; Janssen et al., 1998; 1999) also reported good temporal correlation coefficients (e.g., $r \sim 0.6-0.8$) between personal monitors and fixed outdoor monitors. Moreover, it has been shown by Mage and Buckley (1995) that the average of the personal exposures over many subjects correlates very well with central site monitors ($R^2=0.91$). Thus, from the viewpoint of time series analyses, the generally good mean population-to-monitor longitudinal correlations from the recent exposure studies have provided supportive evidence for the link between ambient air pollution concentrations and short-term health effects.

However, the person-to-monitor longitudinal correlation is only one aspect of the longitudinal correlations that support the link between ambient air pollution and its effects on the whole population in a given city. Even if the person-to-monitor temporal correlation were high between a person and a nearby monitor, if the temporal correlation between a pair of monitors separated by some distance within a MSA were poor, then the use of a single or a few monitors to represent the MSA would be questionable. Since the distance covered in a MSA can range from about 20 to 100 miles, characterization of the extent of monitor-to-monitor longitudinal correlation within that scale of distance is as important as characterization of person-to-monitor temporal correlation. It is this monitor-to-monitor temporal correlation that this study aims to characterize.

The question of the "adequacy" of the location of a monitor has been investigated in the context of designing an efficient network in the earlier studies (e.g., Clifton et al., 1959; Stalker et al., 1962) and more recently by Hwang and Chan (1997), or in the context of spatial representativeness of air monitors by Kotchmar et al. (1987) and Chan and Hwang (1996). However, the monitor-to-monitor correlations evaluated in these studies were of cross-sectional type, and therefore were not relevant to time series studies. There have been a few studies that examined monitor-to-monitor temporal correlations when multiple monitors were available. For example, when many monitors ($\sim$40 stations) were operating in the 1970s in New York City, Goldstein and Landovitz (1977a,b) investigated a single site's representativeness and, based on the poor monitor-to-monitor correlations of daily fluctuations ($r \sim 0.5-0.6$ for SO_2 and $r \sim 0.3$ and 0.4 for smokeshade), concluded that individual stations "did not adequately represent day-to-day variation of air pollution in the surrounding area." Burnett et al. (1994) described monitor-to-monitor temporal correlations for O_3 and sulfate as a function of separation distance in Ontario, Canada. On the average, the monitor-to-monitor temporal correlation between two monitors separated by 60 miles was about 0.7 for both O_3 and sulfate. Description of monitor-to-monitor temporal correlation in a city in an epidemiological investigation is rather rare, however. This is likely due to the scarcity of monitors available within a city or MSA. To increase the statistical power to characterize monitor-to-monitor temporal correlation in this study, we examine monitor-to-monitor temporal correlation of air pollution and weather variables using multiple monitoring stations in a larger geographic area that contains several major cities.

Another important aim of our study is to compare the monitor-to-monitor temporal correlation across multiple air pollutants (PM_{10}, O_3, SO_2, NO_2, and CO) and weather variables (temperature, dew point, and relative humidity) that are commonly included in time series regression models as covariates. In past time series studies, researchers often interpreted the relative significance of each of the covariates as possible relative importance of causal potency. A potential complication in such an interpretation is that, if

relative error differs across pollutants, then we could have incorrectly ranked the pollutants. For example, if one pollutant had poor monitor-to-monitor temporal correlation than another pollutant (with no such information available to the study because there was only a single monitor, as often is the case), then the pollutant with the poor spatial (monitor-to-monitor) representativeness would be less significant in the health outcome regression, even if the underlying causal potency were the same for the two pollutants. We expect the monitor-to-monitor temporal correlation across the pollutants to vary due to their differences in factors such as physico-chemical characteristics and their source distribution patterns. The person-to-monitor temporal correlation would also be different among the pollution and weather variables due to the differences in personal and indoor/outdoor factors mentioned previously, but determining such difference requires actual personal measurements, which was outside the scope of this study. Thus, in this study, we investigated only monitor-to-monitor temporal correlation, not person-to-monitor correlation.

The specific objectives of this investigation are: (1) to estimate "average" monitor-to-monitor correlation for each of the pollution and weather variables that are typically included in the time series studies of air pollution health effects; and (2) to quantitatively describe, for each pollutant, the monitor-to-monitor correlation as a function of separation distance and qualitative site characterization (e.g., the location setting of a site). Knowing the relative extents of monitor-to-monitor correlation among the explanatory variables is important, as they are likely to influence their corresponding relative significance of associations with health outcomes. We analyzed data from seven Central and Eastern U.S. contiguous states to provide further understanding as to what factors cause a monitor to be poorly correlated with other monitors.

Methods

Data

In order to examine monitor-to-monitor correlation in a large well-monitored area, we chose seven contiguous well-populated Central and Eastern U.S. states: Illinois (IL), Indiana (IN), Michigan (MI), Ohio (OH), Pennsylvania (PA), Wisconsin (WI), and West Virginia (WV). These states cover 312,968 square miles and contain over 56 million people. The study area also includes many major cities, such as Chicago, Cincinnati, Cleveland, Columbus, Detroit, Indianapolis, Milwaukee, Philadelphia, and Pittsburgh.

An environmental dataset providing comparable data for all the criteria pollutants was developed first. Air pollution data for PM_{10}, SO_2, O_3, NO_2, and CO were retrieved from U.S. EPA's Aerometric Information Retrieval System (AIRS) for study period 1988–1990 for these states. The AIRS working file format AMP355 (24-h average for PM_{10} and 1-h average for gaseous pollutants were available) was used. From the hourly gaseous pollutant data, daily average values were computed and used in the subsequent analyses. In the initial retrieval, the numbers of monitoring stations for these states were 287, 295, 241, 80, and 108 for PM_{10}, SO_2, O_3, NO_2, and CO, respectively. However, many of these sites collected data for a short period of time and contained large numbers of missing data. Since most of the PM_{10} data were collected on an every-sixth-day sampling schedule (gaseous pollutants and weather variables were collected every day) at most sites, the data analyses were to be conducted for the PM_{10} samplings days only (total of 183 possible days during the 3-year period) in order to use comparable sample sizes. Therefore, we eliminated those sites that had less than 80% of the PM_{10} every-sixth-day sampling schedules (183 days$\times$0.8=146 days). Since O_3 was often collected only during the warm season (i.e., April–October) in some states, the criterion of the minimum number of days required for O_3 was reduced to 105 days (98% of the PM_{10} sampling days between April and October during the 3-year period). In addition, because we were interested in monitor-to-monitor correlation in mainly densely populated areas, we also eliminated those sites whose characteristics (see below) were "Agricultural," "Forest," or "Rural". This elimination process resulted in 97, 137, 99, 26, and 71 stations for PM_{10}, SO_2, O_3, NO_2, and CO, respectively. We chose weather variables that are often analyzed in time series mortality/morbidity studies: temperature, dew point, and relative humidity. The daily average values of these variables were retrieved from Earth Info (Boulder, CO) database, which compiled National Climatic Data Center's Surface Hourly Observations. There were 45 weather stations in these states. Figure 1 shows the resulting distribution of pollution and weather monitoring sites in these seven states.

The monitoring sites were then characterized. The AIRS database contains site characteristic data elements associated with each air pollution monitor. These include: *Land Use* (Residential, Commercial, Industrial, Agricultural, Forest, Desert, or Mobile); *Location Setting* (Urban, Suburban, or Rural); and *Monitoring Objective* (Maximum Concentration, Population Exposure, Background, or Source). The AIRS database also identified each site's Air Quality Control Region (AQCR). The AQCRs are EPA-designated regional boundaries that were established based on jurisdictional boundaries, urban–industrial concentrations, and other factors such as atmospheric areas for the purpose of providing adequate implementation of air quality standards (U.S. EPA, 1972). AQCR can be intra- or interstate. There were 17 "major" AQCRs that contained at least seven monitors for PM_{10} or SO_2 data in this dataset.

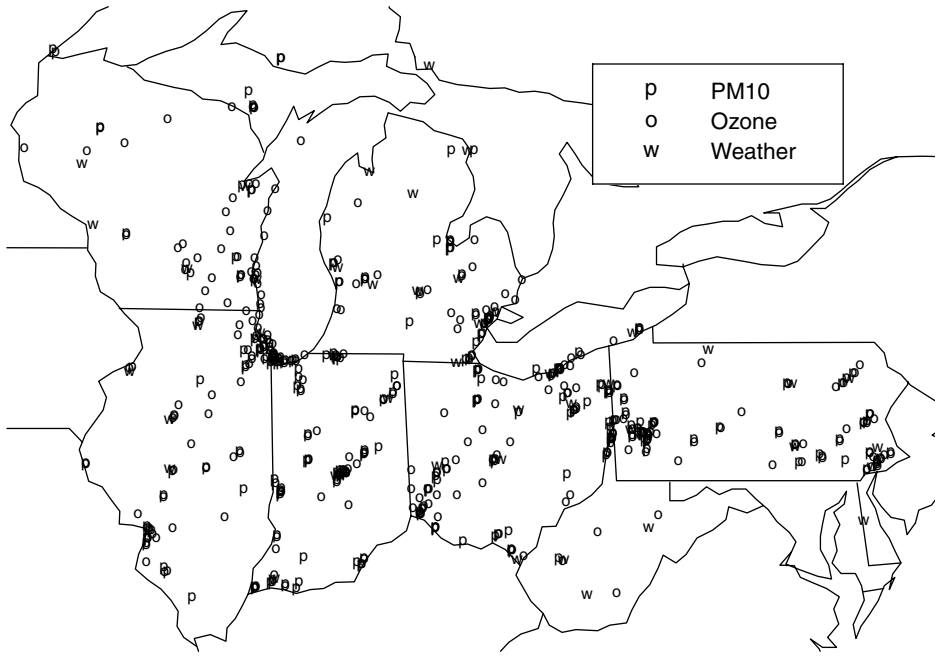

Figure 1. Distribution of air pollution and weather monitoring stations in 7 states.

These characteristics were used in data analysis as group indices or indicator variables. The longitude and latitude information was used to compute the separation distance between each pair of monitoring sites. These data provided a means of evaluating factors affecting the intersite correlations.

Statistical Analysis

The data were examined for temporal correlation after controlling for long-term trends. The air pollution or weather time series from each monitor was detrended using smoothing splines of time, with period corresponding to approximately 1 month and longer, in order to eliminate the influence of seasonal cycles (such influence is especially strong for temperature and O_3). Correlation over time for each pair of monitors was then computed, resulting in $n*(n-1)/2$ correlation coefficients when n sites were available. All the computations were restricted to the every-sixth-day PM_{10} sampling scheduled days. We then characterized the resulting monitor-to-monitor temporal correlation as a function of: (1) separation distance; (2) associated EPA qualitative site characterization (land use, location setting, and monitoring objective); (3) AQCR; (4) variance (the lowest and highest 10 percentiles).

Locally weighted regression, or LOWESS (Cleveland, 1979), was used to smooth the resulting correlation coefficients to graphically visualize the relationships across monitors. We first smoothed the monitor-to-monitor correlation over the separation distance between each pair of monitors (spanning up to approximately 1000 miles). We used 20% of data span for the smoothing window. Because we are particularly interested in characterizing monitor-to-monitor correlation within the geographic scale of a MSA or city (the scale used to aggregate health outcomes in time series studies), we repeated the procedure using correlation with separation distance less than 100 miles. We used a 40% data span for smoothing in the less-than-100-mile data subsets. In addition, we conducted smoothing of monitor-to-monitor by site characteristic strata.

To quantitatively summarize factors that explain monitor-to-monitor correlation, we next conducted regression analyses of the correlation. Note that the regression analyses were conducted for pollution variables only, as the weather stations did not collect individual site characteristic descriptors that are equivalent to pollution stations. Despite the limits of correlation coefficient (-1 to 1), the resulting correlation coefficients were generally normally distributed, except for PM_{10} and NO_2, for which there were some outlying low correlation coefficients. However, preliminary regression analyses indicated that site characteristics and indicator variables for high or low variance could explain these outlying values. Since our preliminary analysis indicated non-linear declines of correlation as a function of distance, with much of the drop in correlation within 30 miles, we also considered a pair of piecewise variables (i.e., the hockey-like shape,

except that the both pieces can have slopes) with the inflection point at 30 miles. Thus, using the data subset for separation distance less than 100 miles, we conducted regression analyses of the monitor-to-monitor correlation (dependent variable) with explanatory variables including the separation distance, site characteristics (indicator variables), a high or low variance (indicator variables for the highest and lowest 10 percentiles), and indicator variables for the 17 major AQCRs.

After the data elimination process described above, the remaining site characteristics categories included: "Industrial," "Commercial," and "Residential" for *Land Use*; "Urban and center city" and "Suburban" for *Location Setting*; and "Population Exposure" and "Maximum Concentration" for *Monitoring Objective*. Thus, each correlation coefficient (for a pair of sites) computed could be classified as either "both," "one only," or "neither" with regard to each site characteristics. Non-singular linear combinations of these indicator variables could be used, but instead, only the indicator variables for "both" (sites having that character) were included because of the ease of interpretation. Prior to regression analyses, we also explored the relationships between the monitor-to-monitor correlation and the explanatory variable candidates to see if any transformation was necessary. Forward stepwise regressions were used to select variables that significantly affected monitor-to-monitor correlation. All the statistical computations were conducted using a statistical package, Splus (StatSci, 1998).

Results

Figure 2 displays the distribution of PM_{10} levels as a function of site characteristics before data restriction (that eliminated "Agricultural" and "Rural" sites). It can be seen that "Agricultural" (AGR) and "Rural" (RUR) sites had lower PM_{10} levels. While the sites with land use = "Industrial" (IND) or monitoring objective = "Maximum Concentration" (MAXC) had larger outlying values, the median levels were comparable between the sites with these characteristics. Similar site characteristic–pollutant concentration patterns were seen in other pollutants (results not shown).

Figure 3 shows the monitor-to-monitor correlation coefficients calculated (after seasonal detrending) as a function of separation distance, with LOWESS-smoothed correlation superimposed. It is obvious that the correlation coefficients are not only generally higher for the weather variables, especially temperature and dew point, but also tighter in their spread. Generally poorer correlation for SO_2 and CO is also indicated. Among the weather variables, relative humidity's spatial correlation declined more steeply as a function of distance than that for temperature and dew point.

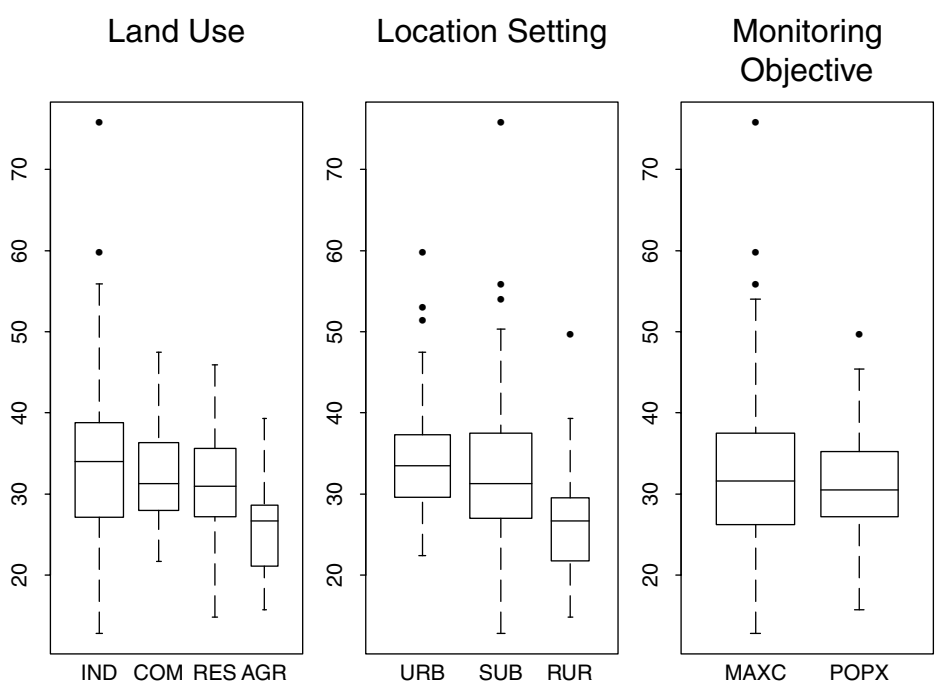

Figure 2. Box-plots of PM_{10} levels (y-axis in μg/m 3) by site characteristics. IND = industrial; COM = commercial; RES = residential; AGR = agricultural; URB = urban; SUB = suburban; RUR = rural; MAXC = maximum concentration; POPX = population exposure.

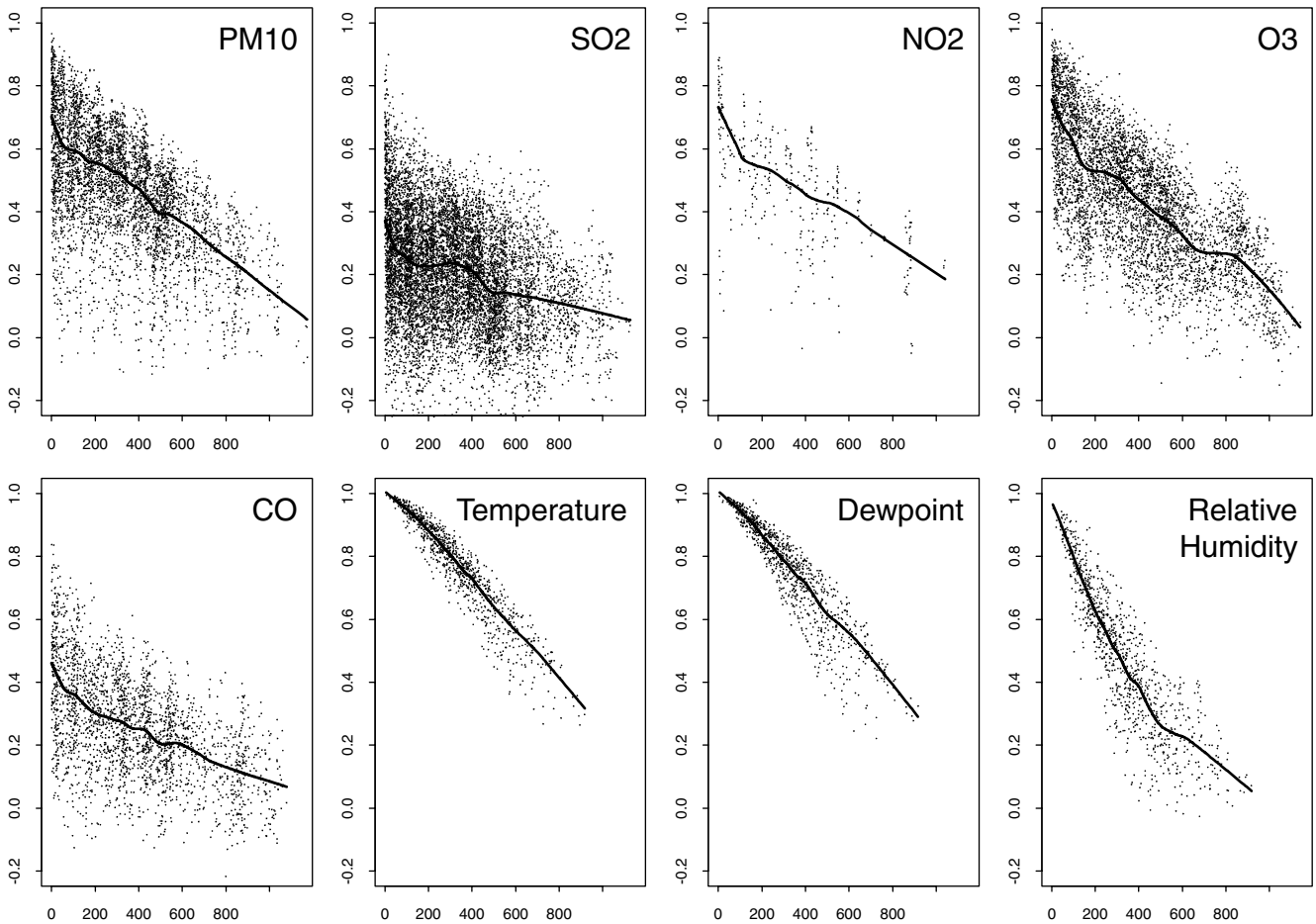

Figure 3. Monitor-to-monitor temporal correlation (y-axis) vs. separation distance (x-axis in miles) for pollution and weather variables.

Figure 4 shows the smoothed spatial correlation using the cases with separation distance of less than or equal to 100 miles (to focus on the scale of a city or MSA typically used in time series health effects studies). In this scale, the decline in correlation over distance is not apparent for temperature and dew point. It can be seen that the smoothed correlation for PM_{10}, O_3, and NO_2 essentially overlapped. CO and SO_2 showed poorer monitor-to-monitor correlation, but the correlation coefficients for CO were somewhat higher than for SO_2. Also, the decline in correlation as a function of distance for pollution variables appears to be steeper within a 30-mile distance. Thus, without taking into account the site characteristics and AQCR difference, the correlation can be ranked into three distinct groups: (1) temperature, dew point, relative humidity ($r>0.9$); (2) O_3, PM_{10}, NO_2 (r: 0.8–0.6); and (3) CO, SO_2 ($r<0.5$). Among the second group, O_3 appears to generally have a slightly higher correlation than PM_{10} or NO_2.

Note that, as could be seen in Figure 3, the distribution of monitor-to-monitor correlation for PM_{10} appears to be somewhat skewed. That is, there are a relatively small number of very low correlation coefficients (i.e., outliers). To identify what factor(s) contributes to the low correlation, the distribution of correlation was also examined by site characteristics. Figure 5 shows boxplots of the monitor-to-monitor correlation for PM_{10} by a 10-mile increment (for

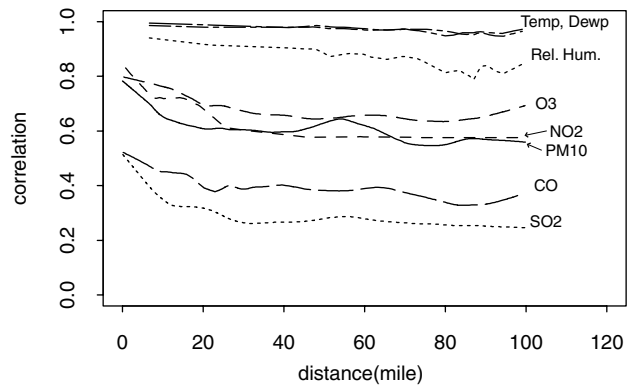

Figure 4. Smoothed monitor-to-monitor correlation (distance $\leq$ 100 miles).

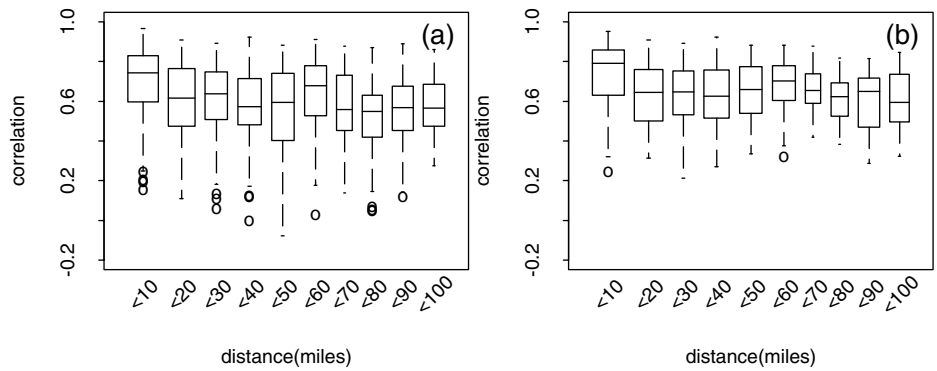

Figure 5. Monitor-to-monitor correlation (distance $\leq$ 100) by distance and with (a) and without (b) industrial sites.

distance <100 miles) with and without the "Industrial" sites. It is clear that most of the very low correlation coefficients are those from the "Industrial" sites. However, the medians are not appreciably different between the correlation with industrial sites and the correlation without industrial sites, suggesting that the "Industrial" sites do not alone explain these low correlation coefficients.

Smoothing was also conducted for data subsets stratified by site characteristics. Figure 6 shows the resulting smoothed correlation for PM_{10} and SO_2. It can be seen that, for land use, the smoothed correlation for "Industrial" sites was lower than that for "Residential" or "Commercial". The smoothed correlation by location setting did not show appreciable difference for both PM_{10} and SO_2. Monitoring objectives ("Population exposure" and "Maximum concentrations") did not separate smoothed correlation for PM_{10}. However, for SO_2, the smoothed correlation was somewhat smaller for "Maximum concentration" than

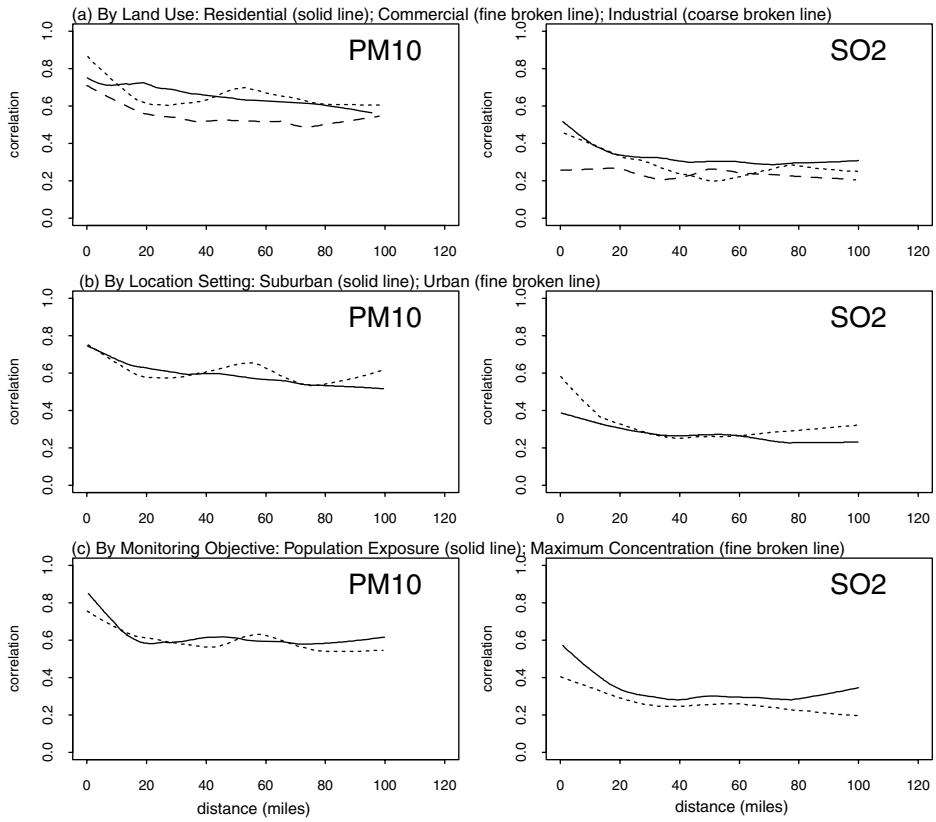

Figure 6. Smoothed monitor-to-monitor correlation by site characteristics for PM_{10} and SO_2.

Table 1. Stepwise regression coefficients (t-ratio) of monitor-to-monitor correlation on site characteristics (case with separation distance $\leq$100 miles only).

	Number of observations	Intercept	Drop in 30 miles	Land Use			Location setting		Monitor objective		Variance indicator[a]		AQCR indicator[b]	
				IND	COM	RES	URB	SUB	POPX	MAXC	High	Low	Highest	Lowest
PM_{10}	969	0.760	−0.171	−0.069	0.037	0.048					−0.085	−0.085	NE Pennsylvania	Chicago
		(45.6)	(−9.33)	(−4.19)	(1.83)	(2.40)					(−5.93)	(5.93)	0.212 (2.08)	−0.069 (−3.44)
O_3	865	0.711	−0.084			0.034		−0.029	−0.032	0.034	0.049	0.049	Cleveland	Chicago
		(34.8)	(−4.10)			(2.91)		(−2.69)	(−2.35)	(1.41)	(3.77)	(3.77)	0.240 (6.05)	0.085 (5.11)
NO_2	70	0.719	−0.204				0.114							
		(19.7)	(−3.75)				(2.37)							
CO	442	0.472	−0.066		−0.034	0.066							NE Pennsylvania	SW Pennsylvania
		(22.1)	(−2.69)		(−1.80)	(2.86)							0.287 (5.92)	0.139 (2.81)
SO_2	1895	0.321	−0.060	−0.039		0.037	0.024	−0.018	0.031	−0.028			Philadelphia	St. Louis
		(19.7)	(−3.47)	(−2.96)		(3.08)	(2.07)	(−1.99)	(2.19)	(−2.58)			0.444 (6.05)	−0.176 (−3.62)

Note: (a) Indicator variables for variables with the highest 10th percentile and the lowest 10th percentile variance; (b) the AQCR indicator variables with the highest coefficient and the lowest coefficients are listed.

Abbreviation keys: IND, Industrial; COM, Commercial; RES, Residential; URB, Urban; SUB, Suburban; POPX, Population Exposure; MAXC, Maximum Concentration.

for "Population exposure". For other pollutants, the site characteristics did not influence smoothed correlation, except for NO_2 and CO (results not shown). For NO_2, the sites with location setting = "Urban" showed higher correlation than those with "Suburban". For CO, the sites with land use = "Residential" showed higher correlation than the "Commercial" sites. Thus, these site characteristics do appear to influence the monitor-to-monitor correlation, but the magnitude of their impacts was not substantial.

Table 1 shows results from the forward stepwise regressions. Note that the t-ratios should not be directly compared across the pollutants because their available numbers of observations are different. The R^2 values for these models were relatively small (0.15, 0.21, 0.23, 0.31, and 0.17 for PM_{10}, O_3, NO_2, CO, and SO_2, respectively). For all the pollutants, the intercept was the most significant predictor. The intercepts showed the essentially the same ranking/grouping as those for the correlation smoothed over the distance discussed above: PM_{10}, O_3, and NO_2 showing similar intercepts ($\sim$0.7–0.75), followed by CO (0.47), then SO_2 (0.32). The piecewise slope variable within the 30-mile distance was selected in the stepwise regression for all the pollutants, while neither the linear distance term nor the piecewise slope variable beyond 30 mile was selected for any pollutant. The computed decay in correlation as a function of distance at the 30-mile separation point was substantial for PM_{10} and NO_2 (-0.17 and -0.20, respectively), but more modest for O_3 (-0.08), CO (-0.07), and SO_2 (-0.06).

The coefficients for the specific site characteristics (Land Use, Location Setting, and Monitor Objective) reflect the magnitude of a gain or loss in correlation (from the intercept) when a pair of sites has a given site characteristic (e.g., Land Use = "Industrial"). While sometimes significant and mostly consistent with the expectation (i.e., "Industrial" sites reduce correlation), their magnitude of impacts (mostly $<|0.1|$) was not substantial. These results are generally consistent with the results from smoothing by site characteristics (Figure 6). The large variance of temporal fluctuations (indicator for the highest 10 percentiles in variance) for PM_{10} reduced correlation, while the small variance (the lowest 10 percentiles in variance) for CO reduced monitor-to-monitor correlation. Thus, the large variance may be indicative of strong local source impacts that contribute to "error" in representing larger region-wide fluctuations of pollution, whereas the small variance may be indicative of the lack of "signal" that contributes to monitor-to-monitor correlation.

As can also be seen in Table 1, the magnitudes of some of the regression coefficients for certain AQCRs were substantial. Of all the pollutants, the largest positive regression coefficient for the AQCR was that for the SO_2 sites in Metropolitan Philadelphia (0.444), although the number of sites within a 100-mile distance in this AQCR was small (four sites producing six correlation coefficients, or 0.3% of the total number of correlation coefficients for SO_2). In fact, the median monitor-to-monitor correlation coefficient of the four SO_2 sites in Metropolitan Philadelphia was 0.795, which is close to the intercept plus the coefficient ($0.321+0.444=0.765$). The second largest regression coefficient for the AQCR was the CO coefficient (0.287) for Northeast Pennsylvania — Upper Delaware Valley. There were five CO sites in this area (10 correlation coefficients), with a median monitor-to-monitor correlation of 0.690. The largest negative regression coefficient for the AQCR was that for the SO_2 sites in Metropolitan St. Louis (-0.176), making the already generally low monitor-to-monitor correlation for SO_2 even poorer. An examination of the six SO_2 sites (15 monitor-to-monitor correlation coefficients) in Metropolitan St. Louis, in fact, showed negative correlation coefficients (range: -0.326 to 0.609, with a median of 0.053). Thus, the regional difference in monitor-to-monitor correlation, as modeled with the AQCR indicator variables, can be substantial, but the number of such AQCRs, as well as the number of sites associated with these AQCRs, was relatively small in this analysis.

Discussion

Without taking into account the site characteristics and regional differences, the monitor-to-monitor temporal correlation among the air pollution/weather variables within a 100-mile separation distance in the combined study areas could be generally ranked into three groups: (1) temperature, dew point, relative humidity ($r>0.9$); (2) O_3, PM_{10}, NO_2 (r: 0.8–0.6); and (3) CO, SO_2 ($r<0.5$), with SO_2 generally showing smaller correlation ($\sim$0.3) than that for CO ($\sim$0.4). Within the 100-mile separation distance, the monitor-to-monitor correlation for the pollution variables declined as a function of separation distance, with much of the drop observed within the first 30 miles.

When the separation distance, site characteristics, and regional difference were modeled in regression analyses for the 100 miles or less separation distance subset, the intercepts were the most significant predictors and their values were similar to the correlation smoothed over distance. The separation distance was a significant predictor of monitor-to-monitor correlation decline, but more so for PM_{10} and NO_2 than for other pollutants. The results indicate that it makes difference whether two monitors are located 2 or 30 miles away. While the qualitative site characteristics were, in some cases, significant predictors of monitor-to-monitor correlation, the magnitude of their impacts was generally not substantial. In some cases, the regional difference, as modeled with the AQCR indicator variables, could be substantial, to the extent that the ranking of the

pollutants' monitor-to-monitor correlation was different from that of the overall data (e.g., SO_2 in Philadelphia and CO in Northeast Pennsylvania). Thus, while there is a general ranking of monitor-to-monitor correlation among pollutants, the regional difference and monitor separation distance can make exceptions. It should also be pointed out that the relatively low R^2 values of these regression models ($\sim$0.2–0.3) indicate that predicting a monitor-to-monitor correlation coefficient for any particular pair of sites based on the explanatory variables included was difficult.

There are some possible explanations for the ranking of pollution monitor-to-monitor correlations observed in this study. SO_2 and CO are both primary pollutants and are likely highly influenced by local sources, whereas O_3 and NO_2 are largely secondary pollutants, and therefore likely more uniformly distributed within the scale of a city or MSA. PM_{10}, in this region of U.S. in the summer months, contains a large fraction (from 50% to 80% in the summer) of sulfate, which is a secondary pollutant derived from SO_2. Wilson and Suh (1997) examined monitor-to-monitor correlation of PM_{10}, $PM_{2.5}$, and $PM_{10-2.5}$ in Philadelphia and St. Louis, and found that monitor-to-monitor correlation coefficients for $PM_{2.5}$ were high ($r\sim0.9$), but low for $PM_{10-2.5}$ ($r\sim0.4$), indicating that fine particles have smaller errors in representing community-wide exposures. Thus, fine particles are expected to show even better monitor-to-monitor correlation than PM_{10}. Also, the situation may be quite different in the Western U.S., where sources of PM are different from those in the Northeast or Central U.S. We are currently extending our study to consider larger areas that include West Coast states.

Interpretation of the monitor-to-monitor temporal correlation in the context of exposure characterization error and its effects on the health risk estimates is an important issue for time series studies. Recently, Zeger et al. (2000) laid out a conceptual framework for assessing measurement error effects in air pollution mortality studies. In their model, the pollution measurement difference between personal exposure and the average of multiple monitor measurements was decomposed into three components: (1) the error due to having aggregate rather than individual exposure data; (2) the difference between the average personal exposure and the true ambient pollutant level; and (3) the difference between the true and the measured ambient concentration. They then described how each component of error would affect the risk estimate (mortality regression coefficient) by substituting the decomposed terms in an equation that described the relationship between individuals' effects and aggregate level. The first component is considered to be of Berksonian error, in that it does not bias the risk estimate. The second component is not of the Berksonian type and is likely to bias the risk estimate. The third component is considered largely of the Berksonian type if the average of available monitors is an unbiased estimate of the true spatially averaged ambient level. The introduction of this conceptual framework is a major advancement, from which empirical and simulation analyses may be developed. However, it should be noted that Zeger et al.'s model assumed, as mentioned in their paper, that a high degree of spatial homogeneity in ambient levels existed, and, for the "measured ambient concentration," the average of multiple monitor measurements was assumed to be available. Zeger et al. also mentioned that the third component represented the instrumental error, which would tend to be cancelled out by spatial averaging across multiple unbiased ambient monitors in a region. However, we point out that many of the past time series studies relied on data from a single monitor (e.g., the time series analysis of the Harvard Six Cities study). For PM, the currently available data in most cities and MSAs were collected every sixth day (which did not provide sufficient sample size for mid-sized cities) and therefore, the mortality analyses often relied on only one or a few, at most, monitors that collected data on daily basis. We also argue that based on our results, we cannot generally assume a high degree of spatial homogeneity, and the extent of such homogeneity varies from pollutant to pollutant. Thus, Zeger et al.'s model, while excellent in terms of framework, did not directly address the possible impact of error due to monitor-to-monitor variability on risk estimates that, as our data suggest, could be important.

Does the knowledge of monitor-to-monitor temporal correlation of a pollutant allow correction or adjustment of the health outcome regression coefficients in a time series study? The answer, unfortunately, is no, or at least not with the monitor-to-monitor correlation alone because: (1) the extent of the error due to person-to-monitor needs to be considered; (2) the classical additive error may not be applicable; and (3) the effects of "error" in multiple regression model requires consideration of intercorrelation among covariates. What we can infer, however, from these results, is that for the pollutants that have poor monitor-to-monitor correlation, such as SO_2 and CO, we have, on the average, a worse chance of getting accurate population exposure estimates using a single or a few monitoring sites, compared to other pollution and weather variables.

The added monitor-to-monitor error would have two effects on epidemiologic risk estimates: (1) it would bias the pollutant relative risk estimates downward; and (2) it would increase the confidence band around the pollution relative risk estimate. For example, in a simplified linear additive error scheme, a monitor-to-monitor correlation is the attenuation factor (also called a reliability ratio; Fuller, 1987) that reduces the risk coefficient by that number. Thus, a pollutant that is measured at a monitor with an underlying monitor-to-monitor correlation of 0.8 could have twice the excess risk of a pollutant with an underlying monitor-to-monitor correlation of 0.4, if they had the same underlying effects. In addition, the pollutant significance would be even

more affected, as the standard error of risk coefficient would increase as well. In time series studies in which the reported pollution effects are often marginally significant, such an extent of difference in monitor-to-monitor correlation could clearly make a large difference in whether a pollutant is deemed a "significant" or "non-significant" predictor of health outcomes.

Despite the limitation of the use and interpretation of these monitor-to-monitor temporal correlation, the ranking of these monitor-to-monitor correlation among variables appears to be, to some extent, consistent with these predictors' relative significance in recent time series mortality studies reviewed (U.S. EPA, 1996b), in which temperature was often the most significant predictor of mortality; PM indices and O_3 were often the most significant predictors among air pollution indices; and SO_2 and CO often did not significantly explain mortality variation. There are also exceptions in which SO_2 or CO was a significant predictor of daily measure of health outcomes. For example, in two recent mortality–air pollution time series studies (Moolgavkar and Luebeck, 1996; Kelsall et al., 1997) for Philadelphia data (recall that, in our results, SO_2's monitor-to-monitor correlation was as good as that for PM_{10}), both Total Suspended Particles (TSPs) and SO_2 were significant predictors of mortality in single pollutant models, and both of their coefficients were reduced when they were included in the model simultaneously. This suggests that the intercorrelations of pollutants, as well as the relative extent of error, can play roles in the relative importance of these pollutants in health outcome regression models.

There are several simulation studies that examined joint effects of correlation and error (Zidek et al., 1996; Marcus and Chapman, 1998; Carrothers and Evans, 2000; Zeger et al., 2000). These studies generally suggest that transfer of effects from one (causal) pollutant to another (non-causal) pollutant is possible under some conditions (i.e., strong correlation between the covariates and large error in the causal predictor). Since intercorrelations among pollution/weather variables, as well as monitor-to-monitor temporal correlation, are expected to vary from region to region, an examination of this issue in real data requires analysis of data from various locations. We did not investigate the intercorrelation among pollutants in this study, but plan to investigate the issue in our ongoing study, in which a larger geographic area will be covered.

It should be re-emphasized that the monitor-to-monitor variability is only one aspect of the exposure characterization error. For example, the monitor-to-monitor temporal correlation of temperature may be very high, but in a city where residential air conditioning or heating is prevalent, the ambient air temperature may not be a good index of population exposure to temperature in that city. Another type of error, monitoring site-to-personal difference, is also likely different among the weather/pollution variables. In addition, the averaging of concentrations over multiple sites would tend to reduce spatial variability as a problem. An overall assessment of the relative contribution of these exposure characterization errors, which will vary from study to study, will be possible when data for all of these components of the interpollutant errors are available.

Acknowledgments

This analysis was conducted as part of EPA-funded projects STAR grant (R82-5271) and EPA PM Center Grant (R82-7351). The manuscript has not been reviewed by EPA and does not reflect EPA's policies. This research is also part of a Center Program supported by the National Institute of Environmental Health Sciences (ES00260).

References

Burnett R.T., Dales R.E., Raizenne M.E., Krewski D., Summers P.W., Roberts G.R., Raad-Young M., Dann T., and Brook, J. Effects of low ambient levels of ozone and sulfates on the frequency of respiratory admissions to Ontario hospitals. *Environ. Res.* 1994: 65: 172–194.

Carrothers T.J., and Evans J.S. Assessing the impact of differential measurement error on estimates of fine particle mortality. *J. Air Waste Manage. Assoc.* 2000: 50: 65–72.

Chan C.C., and Hwang J.S. Site representativeness of urban air monitoring stations. *J. Air Waste Manage. Assoc.* 1996: 46: 755–760.

Cleveland W.S. Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* 1979: 74: 829–836.

Clifton M., Kerridge D., Moulds W., Pemberton J., and Donaghue J.K. The reliability of air pollution measurements in relation to the siting of instruments. *Int. J. Air Pollut.* 1959: 2: 188–196.

Fuller W.A. Measurement Error Models. John Wiley and Sons, New York, 1987.

Goldstein I.F., and Landovitz L. Analysis of air pollution patterns in New York City: I. Can one station represent the large metropolitan area? *Atmos. Environ.* 1977a: 11: 47–52.

Goldstein I.F., and Landovitz L. Analysis of air pollution patterns in New York City: II. Can one aerometric station represent the area surrounding it? *Atmos. Environ*. 1977b: 11: 53–57.

Hwang J.S., and Chan C.C. Redundant measurements of urban air monitoring networks in air quality reporting. *J. Air Waste Manage. Assoc.* 1997: 47: 614–619.

Janssen N.A., Hoek G., Brunekreef B., Harssema H., Mensink I., and Zuidhof A. Personal sampling of particles in adults: relation among personal, indoor, and outdoor air concentrations. *Am. J. Epidemiol.* 1998: 147: 537–547.

Janssen N.A., Hoek G., Brunekreef B., and Harssema H., Brunkreef B. Personal exposures to fine particles in children correlates closely with ambient fine particles. *Arch. Environ. Health* 1999: 54: 95–101.

Kelsall J.E., Samet J.M., Zeger S.L., and Xu J. Air pollution and mortality in Philadelphia, 1974–1988. *Am. J. Epidemiol.* 1997: 146: 750–762.

Kotchmar D.J., McMullen T., and Hasselblad V. Adequacy of a single monitoring site for defining mean outdoor concentrations of fine particles in a demarcated residential community. *JAPCA* 1987: 37: 377–381.

Lioy P.J., Waldman J.M., Buckley T., Butler J., and Pietarinen C. The personal, indoor, and outdoor concentration of PM_{10} measured in an

industrial community during the winter. *Atmos. Environ. Part B* 1990: 24: 57–66.

Mage D.T., and Buckley T.J. The relationship between personal exposures and ambient concentrations of particulate matter. Presented at the 88th Annual Meeting of the Air and Waste Management Association. Air and Waste Management Association, San Antonio, TX, Pittsburgh, PA, June, 1995, paper no. 95-MP18.01.

Marcus A., and Chapman R. Estimating the health effects of fine particles from epidemiological studies: how serious are problems of measurement error correlation, and confounding? Vol. II. $PM_{2.5}$: a fine particle standard. In: Chow J., and Koutrakis P. (Eds.), The Proceedings of an International Specialty Conference, January 28–30, 1998.

Moolgavkar S.H., and Luebeck E.G. A critical review of the evidence on particulate air pollution and mortality. *Epidemiology* 1996: 7: 420–428.

Ozkaynak H., Xue J., Spengler J., Wallace L., Pellizzari E., and Jenkins, P. Personal exposure to airborne particles and metals: results from the particle TEAM study in Riverside, California. *J. Expos. Anal. Environ. Epidemiol.* 1996: 6: 57–78.

Stalker W.W., Dickerson R.C., and Kramer G.D. Sampling station and time requirements for urban air pollution surveys, Part IV. *J. Air Pollut. Control Assoc.* 1962: 12: 361–375.

StatSci, a Division of MathSoft, Inc., Seattle, WA, 1998.

Tamura K., Ando M., Sagai M., and Matsumoto, Y. Estimation of levels of personal exposure to suspended particulate matter and nitrogen dioxide in Tokyo. *Environ. Sci. (Tokyo)* 1996: 4: 37–51.

U.S. EPA. National Office for Environmental Assessment. Office of Research and Development, Air Quality Criteria for Particulate Matter. Research Triangle Park, NC, 1996a, EPA/600/P-95/001cF.

U.S. EPA. Office of Air Quality Planning and Standards. Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information OAQPS Staff Paper. Research Triangle Park, NC, 1996, EPA-452\R-96-013.

U.S. EPA. Office of Air Programs Federal Air Quality Control Regions. Office of Air Programs Publications No. AP-102, Raleigh, NC, 1972.

Wallace L. Indoor particles: a review. *J. Air Waste Manage. Assoc.* 1996: 46: 98–126.

Wilson W.E., and Suh H.H. Fine particle and coarse particles: concentration relationships relevant to epidemiological studies. *J. Air Waste Manage. Assoc.* 1997: 47: 1238–1249.

Zeger S.L., Thomas D., Dominici F., Samet J.M., Schwartz J., Dockery D., and Cohen A. Exposure measurement error in time series studies of air pollution: concepts and consequences. *Environ. Health Perspect.* 2000: 108: 419–426.

Zidek J., Wong H., Le N.D., and Burnett R. Causality, measurement error and multicollinearity in epidemiology. *Environmetrics* 1996: 7: 441–451.