A DESCRIPTIVE ANALYSIS OF TEMPORAL PATTERNS OF AIR POLLUTION IN ATLANTA , GA
AND
AN ASSESSMENT OF MEASUREMENT ERROR IN AIR POLLUTION MONITORING NETWORKS IN ATLANTA, GA

A Thesis
Presented to
The Academic Faculty

By

Katherine Wade

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science in Environmental Engineering

Georgia Institute of Technology

December, 2005

A DESCRIPTIVE ANALYSIS OF TEMPORAL PATTERNS OF AIR POLLUTION IN ATLANTA , GA
AND
AN ASSESSMENT OF MEASUREMENT ERROR IN AIR POLLUTION MONITORING NETWORKS IN ATLANTA, GA

Dr. James Mulholland
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Armistead Russell
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Paige Tolbert
Rollins School of Public Health
Emory University

Date Approved: August 25, 2005

I would like to thank Dr. James Mulholland for all of his guidance and encouragement. I would also like to thank Dr. Ted Russell for his technical help and insight and Dr. Paige Tolbert for supplying public health expertise. Thank you to all three for taking the time to read and comment on this thesis.

I would also like to thank the other graduate students and researchers who have contributed to this work, including Amit Marmur, Jaemeen Baek, and the SOPHIA research team at Emory University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This research is intended to serve as an in-depth analysis of air pollution patterns and monitoring networks in the Atlanta area. A ten year database of carbon monoxide (CO), sulfur dioxide (SO_2), nitrogen oxides (NO_x), ozone (O_3), and particulate matter fine and course ($PM_{2.5}$ and PM_{10}) measurements at 17 monitoring stations across the Atlanta area was developed for use in this research.

Annual, seasonal, weekly, and diurnal profiles of air pollutants are analyzed and described. Several factors are identified that impact these profiles, including changes in emissions, meteorology, and photochemistry. Most sites exhibited decreasing annual average concentrations during the study period, with the exception of O_3 and NO_x, both of which initially increased and then decreased. CO, NO_x, and SO_2 all have the lowest concentrations in the summer months, while O_3 and $PM_{2.5}$ are highest in the summer months. CO, NO_x, and SO_2 are also slightly lower on the weekends. CO and NO_x have peak daily concentrations at rush hour, while O_3 and SO_2 peak in the afternoon hours.

Instrument error was evaluated through audit and calibration data and collocated data. Collocated data is assumed to be a more accurate representation of instrument error; the percent error calculated using collocated data is much higher than that calculated using audit data. Percent errors were similar for all pollutants using audit and calibration data (2-4%) and were similar for all concentration ranges. Percent errors using collocated data were several times larger.

Semivariogram plots are developed to quantify spatial variation of air pollutants. These plots can be interpreted to give the fraction of temporal variation in a pollutant that is actually due to spatial variation. As expected, primary pollutants have higher spatial variation than secondary pollutants. Population weighted averages of the semivariogram function are developed to give a level of uncertainty for a pollutant across the study area.

Pollution rose plots are developed to qualitatively examine local sources that are impacting the monitoring sites used in this research. As expected, primary pollutants have more noticeable local impacts. Point sources are easily identified in SO_2 plots, as are mobile sources in CO and NO_x plots. Pollution roses are also corrected for time of day and season to eliminate false sources.

Chapter 1

INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

Air pollution in urban areas is often monitored through a network of spatially separated samplers. Most large urban areas with air quality problems have government established and regulated monitoring networks for the criteria pollutants: ozone(O_3), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), carbon monoxide (CO), and particulate matter course and fine (PM_{10} and $PM_{2.5}$). Because of the expense of establishing a monitoring network and the availability of data from existing networks, the data collected primarily for regulatory use is often also used for various research objectives, including estimating exposure to air pollution in epidemiologic time-series studies of the health effects of air pollution. In this type of study, the study population spans a region that contains one or more pollutant monitors. The measurement of a pollutant by a monitor is assumed to be the actual pollutant concentration in the atmosphere, and the concentration of the pollutant around the monitor is assumed to be homogenous. Neither of these assumptions is necessarily true. The measurements a sampler gives can be affected by several factors, including statistical sampling error (only discrete portions of air are monitored and they cannot overlap for different monitors), monitor interferences, and reproducibility of methods (APCA, 1995). The combination of these effects on the value a monitor reports is referred to as instrument error. Instrument error can be thought of as

the fact that two identical monitors in the same place may not always report the same concentration of a given pollutant.

In addition to instrument error, many factors influence the observed concentration levels of a pollutant throughout a study area, including proximity to sources, local topography and meteorology. It is difficult to determine if one station is "representative" of an entire region. In many cases, only information from one central monitoring station is used in a study, which increases the amount of error introduced into the epidemiologic study (Goldstein and Landovitz, 1977; Lipfert and Wyzga, 1997; Zeger et al., 2000; McNair et al., 1996). However, in large population-based studies, it has been shown that the average population mean exposure is better correlated with a central monitor value than individual exposures (Schwartz et al., 1996). This does not eliminate the error associated with using a central monitoring station, however, and this error is important as it can reduce the power of an epidemiologic study to determine a relationship between a pollutant exposure level and a health outcome (Pinto et al., 2004; Ito et al., 1995). In one study, Ito et al. investigated the difference between using a single site measurement and a multisite average to estimate total mortality versus PM_{10}. They found considerable variation in PM_{10} concentrations across the study areas (Chicago, Ill and Los Angeles, Ca). The authors concluded that some individual sites were as predictive as the multisite average, but not every site. Thus, the selection of which site(s) to use can greatly affect the outcome of a health effects study. Different air pollutants will have dissimilar levels of uncertainty based on differing spatial variability and instrument error.

The combination of instrument error and spatial variability of air pollution is referred to as measurement error in this thesis. In general, measurement error any

difference between the value of a pollutant used in a study and the actual concentration of the pollutant that the study population is exposed to (Carrothers, 2000; Evans, 2000). In air pollution health effects studies, the health response is often very small and therefore measurement error can be an important source of uncertainty. The focus of this research is to describe and quantify the measurement error of criteria air pollutant measurements used in epidemiologic health effects studies in Atlanta, GA and to qualitatively investigate local sources contributing to the variability. The statistical methods developed can be applied to any urban area.

1.2 Literature Review

1.2.1 Instrument Error

There are several causes of instrument error, including interferences, blank levels, reproducibility of methods, and sampling statistics (APCA, 1995). If a phenomenon interferes with only one sampler, that sampler will not give an accurate measurement. Because the samplers are not watched continuously, there is often no way of knowing if any interference has occurred. The amount of "baseline" pollutant a sampler will register when relatively pure air is sampled is referred to as the blank level. For the same model samplers, this level is assumed to be constant, although in reality it can vary slightly. If each time a sampler is run the method is not exactly the same as for other samplers and sampling events, handler error can be introduced. This aspect of instrument error is minimized by the use of automated, continuous samplers and becomes more important in

remote analysis samplers, such as $PM_{2.5}$. Finally, the amount of air used by any given sampler is statistically a very small fraction of the total amount of air in a sampling area. If the pocket of air sampled is compromised, for instance through a faulty intake valve, the measurement is less likely to be representative of the entire area (APCA, 1995).

In the past, instrument error has largely been dealt with on a purely qualitative level in most studies. It is assumed to be a very small percent of overall concentrations and it is assumed not to vary much among different pollutants. However, in federal monitoring networks a quality assurance system has been established to minimize instrument error (US EPA, 2004). Among the specifications of the Code of Federal Regulations (CFR), network design, siting criteria, quality control/ quality assurance protocol, methodology and data validation are all explicitly outlined to reduce error, including instrument error.

When assessing instrument error, the quality assurance protocol established by the CFR is most important. Performance audits are specified for each air pollutant to examine the accuracy of the monitors and calibrations and collocated monitors are specified to examine the precision of the monitors. Data from these events can then be used to assess the overall instrument error. Several studies directed at the health effects have used quality assurance data to show that instrument error can sometimes be larger than the actual trends in the pollutant and therefore the measurements will not provide meaningful relationships when used in a health effects model (Lipfert and Wyzga, 1997; Carrothers and Evans, 2000; Stanley, 1985). This is particularly relevant in remote analysis samplers as the additional handler error can raise instrument error substantially (APCA, 1995).

1.2.2 Spatial Variability

Several methods of assessing spatial variability of air pollutants have been used. Some research has simply made comparisons of absolute values of pollutants at different locations (Chow et al., 2001; Pinto et al., 2004; Burton et al., 1988). This approach does give a qualitative idea of differences in concentrations over an area, but it gives no information regarding the changes in concentration throughout the area. Other research has looked at the use of different statistics that measure spatial variability, such as the coefficient of divergence (COD) (Pinto et al., 2004). The COD provides a relative measure of homogeneity across a study area, but also does not provide quantitative information regarding spatial variability. Other approaches to assessing spatial variability are described below.

1.2.2.1 Correlation Coefficients

Spatial variability of various air pollutants across urban areas has been studied extensively (Hansen et al., 2003; Mulholland et al., 1998; Tiles and Zimmerman, 1998). In particular, the issue of how representative a monitor is of the surrounding area is of interest. Several studies use correlation coefficients as a measure of variability (Monn et al., 1997; Buzorious et al., 1999; Morowska et al., 2002; Lin et al., 2001; Rao et al., 1995). The correlation coefficient is a simple statistic that measures the interdependence or covariance of two variables, ranging from 0 (no correlation) to +/- 1 (perfect correlation). In an air quality monitoring network, correlation coefficients can be calculated between sites that are measuring the same pollutant simultaneously. A

correlation coefficient of close to 1 between two or more sites indicates less spatial variation in a given pollutant.

For example, Morowska, et al. used Pearson correlation coefficients to characterize spatial variability of O_3, NO_x, and PM_{10} is Brisbane, Australia. In this study, sites were used that were specifically sited to have few local source impacts. Three stations in the network were chosen for the study and correlation coefficients were calculated for the same pollutant at the same point in time at the different stations, each of which consisted of at least 100 observations. The authors further analyzed the data three ways: first, considering all of the data, then considering only the low concentration levels, and finally considering only the high concentration levels. In this way, it is possible to infer if the extreme values of a pollutant are local sources or regional. The authors found that O_3 concentrations had the least variability, followed by NO_x and then PM_{10}. The authors also suggested reducing the number of O_3 monitors in the area, but not the NO_x and PM_{10} monitors.

Correlation coefficients can also be used in the development of interpolation models. For instance, Duc et al. (2000) developed "correlation fields", which are used to determine how far away from a known observation unknown concentrations can be derived. The correlation field can also be used to form regions within a study area by which to group monitoring stations. Other studies have looked at correlation coefficients at various distances from a given source, most often a road (Vardoulakis et al., 2003).

The correlation coefficient gives limited, qualitative information regarding spatial variability. The general assumption is that less correlation implies greater variability. A lack of correlation among monitoring sites indicates that no one site is representative of

the entire study area. The correlation coefficient does not, however, give information on uniformity of concentration throughout a region or how quickly correlation is lost between sites or on the spatial distribution of a given pollutant.

1.2.2.2 Coefficient of Variation

Another metric used to examine spatial variability is the coefficient of variation, which is the standard deviation of a pollutant divided by the mean. This value is useful if one assumes the standard deviation of a set of observations is proportional to the mean. This is useful in air pollution data analysis if two monitoring sites have very different mean concentrations.

Monn et al. (1997) studied PM_{10} and NO_2 near roadsides in Zurich to examine small-scale concentration gradients using a spatial coefficient of variation. They found the larges concentration gradient of PM_{10} was between the site closest to the street and one 15 meters away. Beyond 15 meters, no significant gradients were observed. NO_2 showed a similar, but not as dramatic gradient. PM_{10} showed no significant seasonal or meteorogical patterns and had a spatial coefficient of variation of around 11-13%. NO_2, however, showed a strong seasonal dependence. For the entire study period, the spatial coefficient of variation as 15.4%. For the winter season the value dropped to 2.7% and for the summer it rose to 33.5%. The authors suggested using less NO2 monitors during the winter months.

Hoek et al. (2002) also used variance to examine spatial variability in air pollution. They conducted an analysis of variance between $PM_{2.5}$ monitoring sites in three different urban areas and found that the differences between sites were significantly

larger than the variance within one site. The focus of this study was traffic related air pollution, but the authors found significant differences when using sites that were farther away from roads as well. Several other studies have used the same approach assess spatial variation in other pollutants (Lebret et al., 2000; Roosli et al., 2001).

1.2.2.3 Semivariograms

The methods described above to estimate the spatial variability of air are limited by the idea that data are only available at spatially fixed locations. It can be interpreted that a large variance indicates greater spatial variability but this generality does not have much other useful information, particularly about locations between monitoring sites. A common method of describing spatial data in geostatistics that addresses the uniformity of data over a study area is the semivariogram (see Figure 1.1). The semivariogram is based on the concept of spatial autocorrelation; that is, observations made closer together will be more similar than those made further away. The semivariogram is a function of the spatial distance between two observations, but it is not dependent on the actual locations at which the observations were made. The semivariogram function is often plotted against distance between observations in order to give information about the continuity and variation of observations and has several common attributes. There is usually a distance at which the value of the semivariogram no longer changes. This distance is the range and the value at which the semivariogram remains constant is the sill. The semivariogram is expected to increase from 0 to the sill. At distances outside of the range, observations are assumed to be completely uncorrelated. The shape of the semivariogram near the origin is often of the most interest. A smooth curve implies the

variable being observed is continuous and differentiable. Often one observes a discontinuity at the origin, which is known as the nugget effect. The nugget effect can be interpreted as instrument error usually. The semivariogram is considered a more reliable method of estimating variability than the covariance function in geostatistics. The estimation of the semivariogram is based on the variogram, which is defined as follows:

$$2\gamma(h) = Var(Z(s+h) - Z(s)) = E[(Z(s+h) - Z(S))^2] - [E(Z(s+h) - Z(s))]^2$$

Where h is the distance between observation locations, s_i and s_j and Z is the observation at each location.

Because the expected value of $Z(s_i) = \mu$ the second term is 0. The semivariogram is half of the variogram, defined as follows:

$$\gamma(h) = \frac{1}{2N}\sum_{N} (Z(s_i) - Z(s_j))^2$$

This equation is known as the classic semivariogram estimator, where N is the number of distinct pairs separated by a distance h.

The semivariogram is also theoretically equivalent to the difference in the covariances at distance 0 and at distance h.

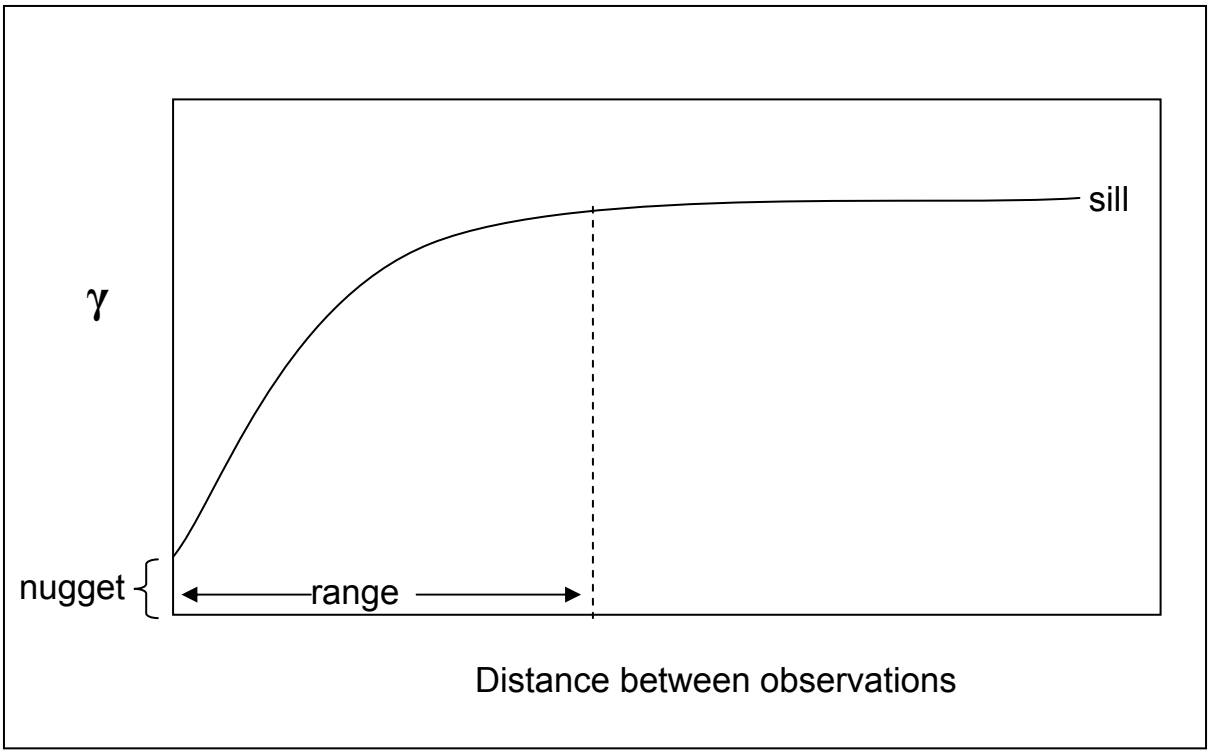


Figure 1.1: Graph of the semivariogram function.

The semivariogram can be visually interpreted to provide qualitative information about pollutant measurements. The nugget effect described above corresponds to instrument error. The initial slope of the curve indicates local source effects. The range of the curve is the distance at which two observations are not expected to have any correlation, similar to the radius of influence described above.

The semivariogram has been used previously for environmental pollutants (Grondona and Cressie, 1991; Casado et al., 1994). A common method of displaying spatial variability is through a contour map, which involves interpolating values of the pollutant of interest between observations. Kriging is a popular geostatistical method of interopolating data which requires the development of a semivariogram. The

semivariogram is used to define the spatial relationship between data points. The concept behind kriging is to take a weighted average of known concentrations to determine an unknown concentration at a different location. This method has been applied successfully to air pollution data in several cases (Iaco et al., 2002; Ferreira et al., 2000). Iaco et al. (2002) developed space-time semivariograms for use in kriging air pollution data in Milan, Italy. They were able to use the resulting contour graphs to distinguish patterns in air pollutants attributed to traffic. The authors showed that the interpolation model based on the semivariograms was able to correctly predict air pollution related to traffic emissions with r-values above .6 for all days tested. Ferreira et al.(2000) used similar methods to study air pollution in Lisbon, also focusing on traffic emissions, and had similar success. However, in neither case was the semivariogram directly interpreted to give meaningful information on the spatial variability of the pollutant of interest. Because of the focus of this research, the use of semivariograms in interpolation techniques will not be addressed.

One variation of the semivariogram that has been used previously in air pollution studies is the cumulative semivariogram (CSV)(Sen, 1989). The CSV is a graph of successive semivariograms:

$$\gamma(h) = \frac{1}{2}\sum_{i=1}^{h}(Z(s_i) - Z(s_j))^2$$

The CSV is capable of showing even small dependencies in pairs of data that may not be evident in a classical semivariogram. At large distances, the CSV is essentially a

straight line, which agrees with the idea of a sill in the classical semivariogram. The area of the CSV closest to the origin is a curve and the distance for which the CSV curves is equivalent to the range in the classical semivariogram. The slope of the CSV at any distance has been shown to be equal to the value of the classical semivariogram at that distance (Sen, 1992).

The cumulative semivariogram has been used in several air pollution studies (Sen 1989 and 1992, Anh et.al., 1997). Sen, who originally developed the CSV, used it as a method of qualitatively evaluating air pollution in Istanbul. Pollutants considered were smoke, PM and SO_2. Each CSV developed for SO_2 showed initial curvature followed by a straight line at large distances. This can be interpreted to mean that at small distances there is greater "dependency" or correlation between observations and at larger distances the observations are independent. The CSV's developed for PM did not show any curvature, indicative of the regional concentration of PM being fairly uniform. Sen also shows that the slope of the long-distance straight line is a function of the standard deviation of the pollutant concentration. As in classical semivariograms, intercepts not at the origin indicate discontinuities in the data.

Anh et al. (1997) also used the cumulative semivariogram to describe spatial variability of air pollution. In this case, the pollutant of interest was ozone and the region was Sydney, Australia. The authors developed a CSV model based on the K-theory of atmospheric diffusion and use it to determine a "radius of influence". They conclude that for ozone, sub-regions of 17km exist in which concentrations are well correlated. The authors also conclude that primary pollutants would have much smaller sub-regions.

Sen (1998) has also applied the cumulative semivariogram to kriging techniques, similar to the classical semivariogram. The CSV is used to determine the weighting of various observations when determining an unknown concentration at a different location. The weightings are inversely proportional to the distance between sites.

There is some speculation that the use of the semivariogram to characterize spatial variation in air pollutants is inappropriate because of a lack of well distributed monitors (Diem, 2003). Specifically, Diem argues that to create a stable semivariogram of ozone concentrations in the Atlanta MSA, 100 monitors would be needed. Obviously this is unreasonable. Diem also suggests that spatial modeling can be used to overcome this obstacle, but that is not the focus of this research.

1.2.3 Sources of variability

Once the issue of spatial variability is addressed and quantified, it is useful to determine the causes of this variability. One cause, which can be quantified relatively easily through the use of the semivariograms, is instrument error. This is error that is based on the fact that different samplers are used at different monitoring sites and no two samplers can be exactly the same. Co-located data, in which two samplers are run simultaneously at the same site, along with audit data, in which an independent person runs standard pollutant levels through a monitor, are used to quantify the amount of instrument error.

A second source of error is due to the impact of local sources. Depending on the size of a source of a given pollutant, it may impact only one or very few monitors within the area. There are already well known sources of air pollutants that can impact sites

differently. Several studies have examined the spatial variability in air pollution due to proximity to roadways (Lebret et al., 2000). For the most part, these studies have been conducted by comparing observations at sites located at varying distances from major roadways. In the Atlanta area, mobile sources are the largest contributor to CO and NO_x concentrations in ambient air. There are several coal fired power plants in the Atlanta area and these are believed to be the most significant sources of SO_2. CO, NO_x and SO_2 are all primary pollutants, meaning they are directly emitted. Secondary pollutants, such as ozone and particulate matter, are formed in the atmosphere through chemical reactions and therefore do not have any "sources". In this research, one would not expect to see as much spatial variability in secondary pollutants because of the lack of sources.

Currently, source apportionment modeling is used to identify sources of $PM_{2.5}$ through analysis of the component fractions (Marmur et al., 2005; Zheng et al., 2002). However, this approach requires detailed measurements that are often not available and is computationally intensive.

One method of qualitatively examining local sources of air pollution is to develop plots of pollution concentrations versus wind direction. If large concentrations are indicated from one direction, that area should be inspected for sources of the pollutant. For example, a large SO_2 spike on a pollution rose plot in a given direction could confirm that SO_2 is coming from a power plant in that direction. Aneja et al. (1999) use wind roses to examine peak ozone concentrations in North Carolina in order to support the hypothesis of regional transport of ozone into the study areas. Ma et al. (2004) conduct similar analyses to determine if long-range transport from China is a source of elevated $PM_{2.5}$ concentrations in Japan. Wind roses have also been used to locate emissions on a

smaller scale. Navazo (2003) used wind roses to indicate traffic and industrial sources of various air pollutants in Bilbao, Spain.

Finally, regional sources and meteorology can also affect the spatial variability of air pollution across a region. If a source of an air pollutant is large enough, its impact on concentrations in a given area could be observed at all monitoring locations. Regional sources could be, for example, very large power plants that have a noticeable affect on SO_2 concentrations in a region of interest. Regional meteorology can also impact local concentrations of an air pollutant. If wind measurements in a region show a strong overall directionality, this could result in a transfer of air pollutants. In this way, another region that has severely elevated levels of a pollutant could serve as a "source" of that pollutant to another region.

1.3 Atlanta Area Epidemiologic Studies

A number of studies have attempted to establish a relationship between ambient air pollution concentrations and cardiovascular health or respiratory health (Schwartz et al., 1996). Recently, analyses have been refined to investigate health outcomes related to specific pollutants in urban environments. In the Atlanta area, the Aerosol Research Inhalation Epidemiology Study (ARIES) has examined emergency department visits for cardiovascular and respiratory diseases relative to ambient pollutant concentrations of particulate matter, ozone, nitrogen dioxide, sulfur dioxide, and carbon monoxide (Metzger et al., 2004; Peel et al., 2005). These studies used air pollution data collected from 1993-2000, including central monitoring site data. The monitoring site used in both studies was the Jefferson Street site, with data collected from 1998-2000.

Peel et al. (2005) found that standard deviation increases in NO_2, O_3, CO and PM_{10} were associated with increases in emergency room visits due to upper respiratory infection, a 2 $\mu g/m^3$ increase in $PM_{2.5}$ was associated with increased pneumonia visits, and standard deviation increases in NO_2 and CO were associated with chronic obstructive pulmonary disease visits. Further, associations remained beyond 3 days for several outcomes and over a week for asthma.

Metzger et al. (2004) examined the same data for associations between cardiovascular conditions. Cardiovascular disease related emergency room visits were found to be associated with NO_2, CO, $PM_{2.5}$ and several $PM_{2.5}$ components. Risk ratios ranged from 1.04 to 1.02, with the strongest associations seen on the day of elevated concentration.

In both of these studies, one concentration of a given pollutant was used to estimate exposure for the entire study population. Several assumptions are inherent when using data from only one site. If it is assumed that this site is representative of the entire study area, then it is also assumed that there are no sources locally affecting air quality only in the immediate vicinity of the monitor and not throughout the study area. It is also assumed that the concentration of a pollutant is homogenous throughout the study area and fluctuations are uniform throughout the study area. For some pollutants, especially secondary pollutants that are not affected by local sources, these assumptions may be true. However, for primary pollutants it is not evident that the concentration observed at one monitoring station is representative of an entire study area. If this is the case, substantial uncertainty due to errors in the exposure estimate and thus the outcome of a

health effects study is introduced. In this research, this error is a major component of measurement error, as described previously.

The suitability of one monitoring site to estimate exposure to a pollutant over a study area should be carefully evaluated for health effects studies. If a central monitoring site is to be used, the size of the area it is representative of should be first established. Different pollutants, with different atmospheric lifetimes, are expected to be representative of different size areas. The amount of error introduced by using spatially-limited data should also be evaluated.

The intent of this research is to develop a metric that can be used in epidemiologic studies to assess the error introduced into a study by using only one, or a limited number or, monitoring site(s). This error is termed "measurement error" and is separated into two components: instrument error and local source impacts.

Chapter 2

OBJECTIVES

The overarching focus of this research is to provide a metric that can be used in epidemiologic studies to quantify the contribution of measurement error, the combination of instrument error and local source impacts, to overall exposure error. This is done through the accomplishment of several secondary objectives.

The first objective of this research was to develop a comprehensive database of air pollution and meteorology data in the Atlanta area. The time frame for this data base was 1993-2003, for use in an epidemiologic study of the same time frame. This database was first used to give descriptive information about air quality in the region. Various temporal averages, yearly, monthly, weekly, and daily, as well as correlation coefficients, were computed for descriptive information. This database is also the origin of the more detailed statistical analysis performed in this research.

The next objective was to quantify instrument error, which is the result of differences in individual monitors, and local source impacts. These are the two components of measurement error. Several questions about measurement error were addressed. First, instrument error was quantified for each pollutant to determine if it is a large component of measurement error and if it is similar in magnitude across pollutants. Next, local source impacts were considered. The impact of local sources was addressed for each pollutant. Similar to instrument error, it was important to establish differences in the magnitude of local source impact across pollutants.

Once the contribution of local source impacts to measurement error was calculated, the final objective of this research was to assess the individual sources impacting the monitoring sites used. This is done to highlight the importance of various sources to different pollutants as well as to illustrate the representativeness of the different stations. Individual monitoring stations were studied to determine if one station has fewer local source impacts and is therefore representative of a larger area than other sites.

Chapter 3

METHODS

This research involved the assembly of a database of air pollution measurements for the metropolitan Atlanta area. The initial database timeframe was 1993-2004, and has now been expanded to include 1968-1992. Data from two monitoring networks, described below, was assembled into a comprehensive database for use by Georgia Tech and Emory researchers. All criteria pollutant data was used, which includes carbon monoxide, oxides of nitrogen, sulfur dioxide, ozone, and particulate matter. Meteorological data was also collected.

3.1 Monitoring Networks

3.1.1 Georgia Department of Natural Resources

The Georgia Department of Natural Resources Environmental Protection Division (EPD) manages a large network of air quality monitoring stations in the metro Atlanta area. The majority of the data used came from this network of monitors. Data was retrieved through the Air Quality System (AQS) database, maintained by the U.S. Environmental Protection Agency (EPA). Sampling methodology is specified in the Code of Federal Regulations (CFR) and all methods described below are the federal reference method or equivalent.

3.1.1.1 Site descriptions

Fifteen EPD managed monitoring sites were used in this study (see Figure 3.1). Most sites only sample for a few pollutants, as described below.

Roswell Road is designed to be a microscale measurement site. The purpose of this monitoring at this location is to monitor peak pollution levels from traffic. A microscale site is only intended to be representative of several meters (up to 100 m) in area. Roswell Road is surrounded by several large roadways, including 3 major highways. The Roswell Road site was established in 1994.

The South Dekalb site is considered to measure an urban scale level of air pollution. These sites are intended to measure pollution coming out of Atlanta as an area source. Located just southeast of downtown Atlanta, the site is immediately surrounded by trees and the area beyond that is mostly residential. South Dekalb is expected to see the worst conditions as it is just downwind of Atlanta. Interstate 285 is just over 1 km to the north of the site.

Most of the sites used in this research are neighborhood scale. This siting is used to estimate pollutant exposure in large populations and is considered representative of an area from 0.5 to 4 km, including Dekalb Tech, Tucker, Georgia Tech, and Confederate Avenue. The Dekalb Tech site is located on a community college campus in a well populated area. Samplers are situated at the southern end of a large campus parking lot. This site was established in 1981. The Tucker site is located in the parking lot of an abandoned building in a mainly commercial area. One large roadway, Lawrenceville Highway, is located approximately 1000 m to the north. This site was established in 1995. The Georgia Tech site is in an urban area and is situated to give average

measurements at the city center. The surrounding area is a mostly a college campus, with some commercial areas. Interstate 75/85 is located directly to the east of the site. The Georgia Tech site was originally established in 1973 but was moved to its present location in 1993. The Confederate Avenue site is in an abandoned parking lot on a National Guard Reservation. The surround area is mostly residential. Interstate 20 is located just over 1 km to the north of the site and I-75/85 is located several kilometers to the west of the site. This site was established in 1991. The site at Fire Station 8 was established in 1973 in a suburban, commercial area. There are several large roads in the vicinity, but no major interstates. E. Rivers School is a suburban site established in 1971. The surrounding area is largely residential. The Fulton County Health Department site is located in the heart of downtown Atlanta. It was the first air quality monitoring site in Georgia, established in 1957. The surrounding area is commercial. Similar to South Dekalb, this site is expected to have the highest levels of air pollution due to its downtown location. Doraville Health Center is located to the north of downtown Atlanta. The surrounding area is mixed commercial, residential and industrial. One large roadway, Buford Highway, is located to the southeast of the site and a major interstate, I-285, is located to the northeast. The Doraville Health Center site was established in 1970. The site in Douglasville is approximately 30 km west of downtown Atlanta. Interstate 20 is several kilometers south of the site and a rail road track runs less than 1 km northwest of the site. The Douglasville site was established in 1995.

Four rural/agricultural sites were also used in this research: Conyers, Yorkville, and Stilesboro. The purpose of these sites is to measure background concentrations. The Conyers site was established in 1976 on a monastery about 40 km east of downtown

Atlanta. The surrounding land is forest and pasture. The Stilesboro site is located on agricultural land about 80 km northwest of downtown Atlanta. Bowen Power Plant is located 3 km directly to the north. This site was established in 1982. Yorkville is a rural site established in 1996 and located on a private farm about 55 km northwest of Atlanta. The site is on a pasture and the surrounding area is covered mainly with hardwoods. The Griffin site is located on University of Georgia Extension Services land. The site is located approximately 60 km south of downtown Atlanta. The Griffin site was established in 1990.

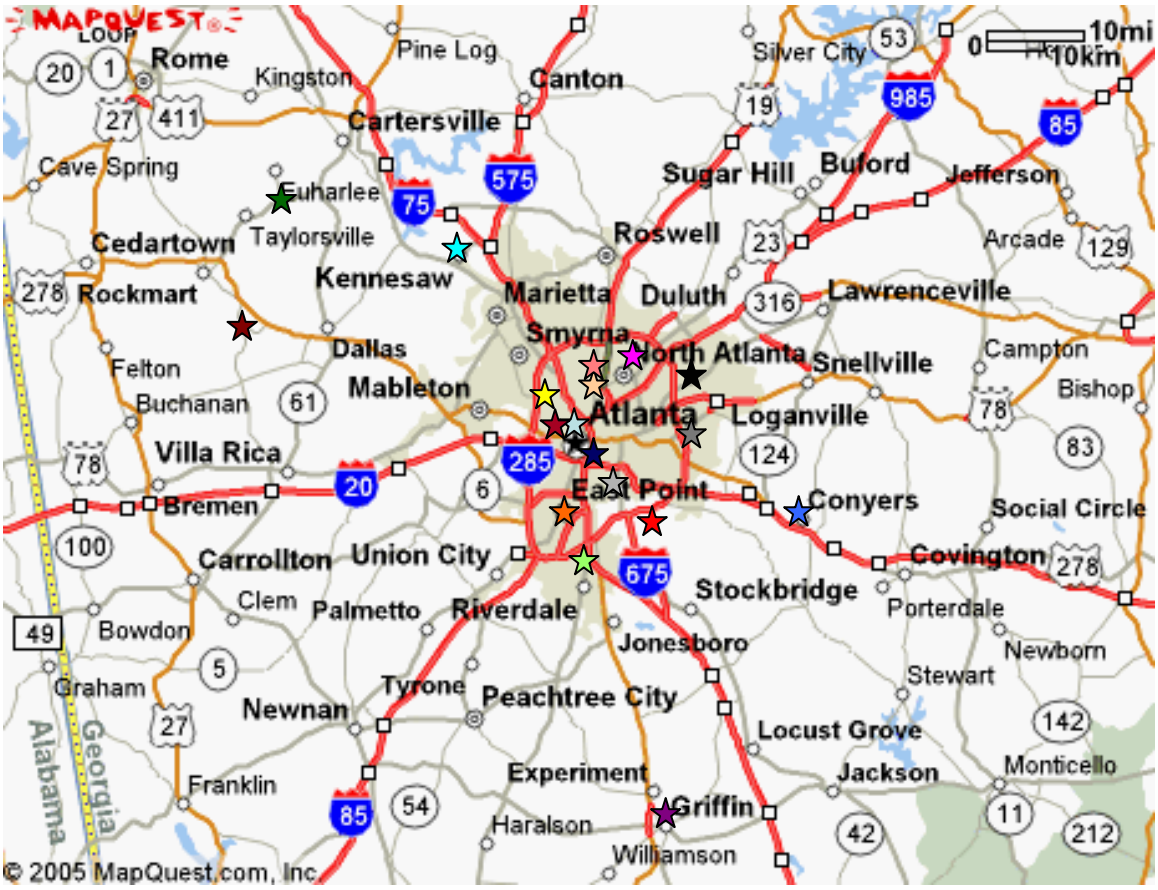


Figure 3.1: Monitoring stations.

3.1.1.2 Measurement Techniques

As specified by the Clean Air Act (CAA), the Atlanta area has two CO monitors. These are located at the Roswell Road monitoring site and the Dekalb Tech monitoring site (see Figure 3.2).

CO is measured continuously at both sites using non-dispersive infrared photometry and gas filter correlation methods. The monitors are just above ground level (3-9 feet above the ground) and intended to sample air at the level most people are inhaling it.

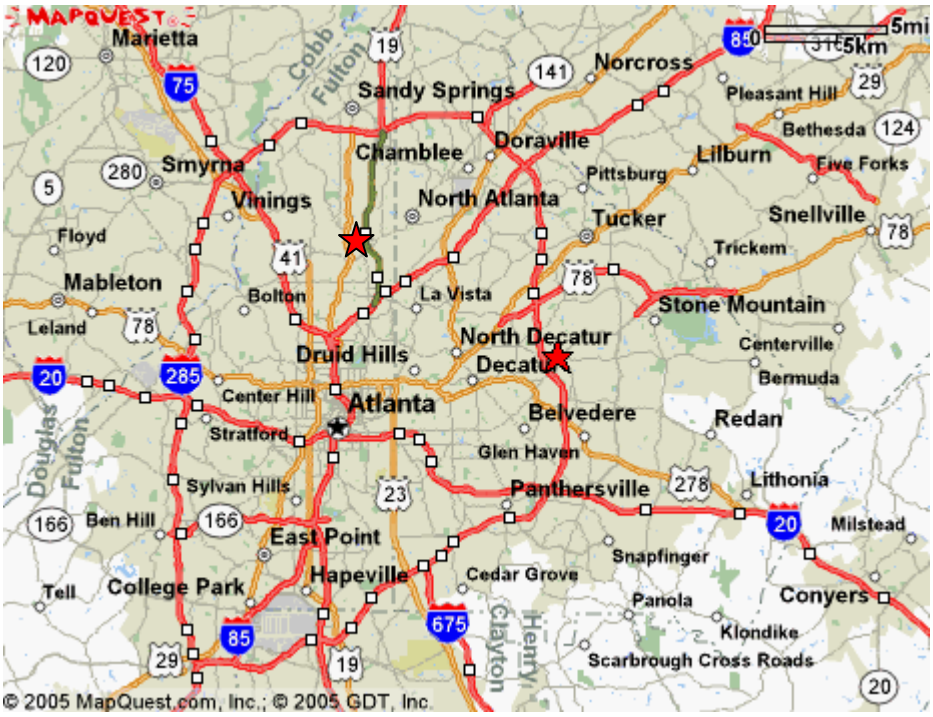


Figure 3.2: CO monitoring stations.

EPD manages five NO_2 and NO_x monitors in metro Atlanta (see Figure 3.3). There are two urban sites, Georgia Tech and South Dekalb, one suburban, Tucker, and two rural or background sites, Conyers and Yorkville.

NO_x and NO_2 are measured continuously using ozone phase chemiluminecence. As with CO, monitors are sited at ground level to sample air from the breathing zone.

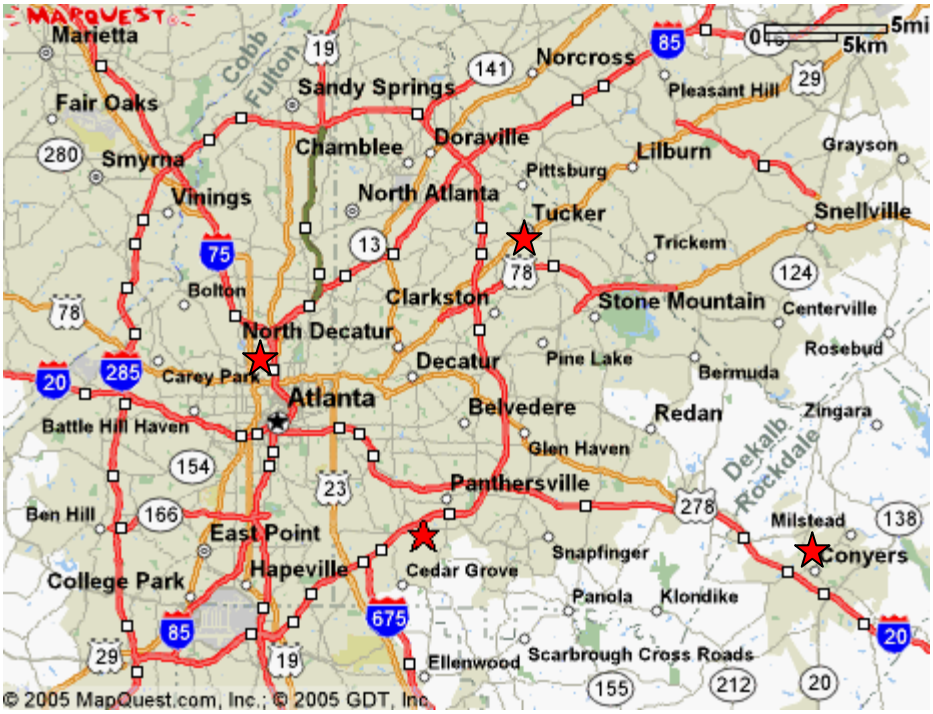


Figure 3.3: NO_2/NO_x monitoring stations.

There are 13 SO_2 monitoring stations in Georgia, three of which are located in the area of interest for this research (see figure 3.4). The sites used were Stilesboro, located

northwest of downtown Atlanta, and Confederate Avenue and Georgia Tech, located in downtown Atlanta.

SO_2 is measured continuously using a potassium tetrachloromercurate bubbler. The detection limit for SO_2 measured in this way is 0.01ppm. There is no height specification for SO_2 monitors.

Figure 3.4: SO_2 monitoring stations; a: urban, b: rural.

Ozone monitoring occurs at 22 sites in Georgia from March through October. Three metro Atlanta sites were used in this research: Confederate Avenue, South Dekalb, and Conyers (see Figure 3.5). Confederate Avenue is a downtown site, South Dekalb is suburban, and Conyers is rural.

Ozone is monitored continuously using the U.V. photometric method. Samplers are placed at ground level to sample at the breathing zone.

Figure 3.5: O_3 monitoring stations.

PM_{10} is measured at 26 sites throughout Georgia. Seven of these sites are within in the Atlanta metro area and are included in this research. These include Georgia Tech, Fire Station 8, E. Rivers School, Fulton County Health Department, Doraville Health Center, Griffin, and Douglas County (see Figure 3.6).

There are two methods of sampling PM_{10} in Georgia. The first is by filter sampling and gravimetric analysis. Twenty-four hour samples are collected on microquartz filters every six days. This is the method used by all of the sites except Georgia Tech. At Georgia Tech, a continuous tapered element oscillating microbalance (TEOM) sampler is used to report continuous concentrations of PM_{10}.

$PM_{2.5}$ is measured at the same sites as PM_{10}, although the $PM_{2.5}$ monitoring network was not established until 1999. The sites used in this research are Yorkville, East Point Health Center, Fire Station 8, E Rivers School, Doraville Health Center, South Dekalb, Kennesaw, and Forest Park (see Figure 3.7). All of these sites are considered urban except Yorkville, which is a rural site, and Kennesaw, which is suburban.

$PM_{2.5}$ samples are collected on Teflon filters and undergo gravimetric analysis. Daily, 24 hour samples are collected at E Rivers School, Doraville Health Center, and South Dekalb. Samples are collected every third day at all other sites.

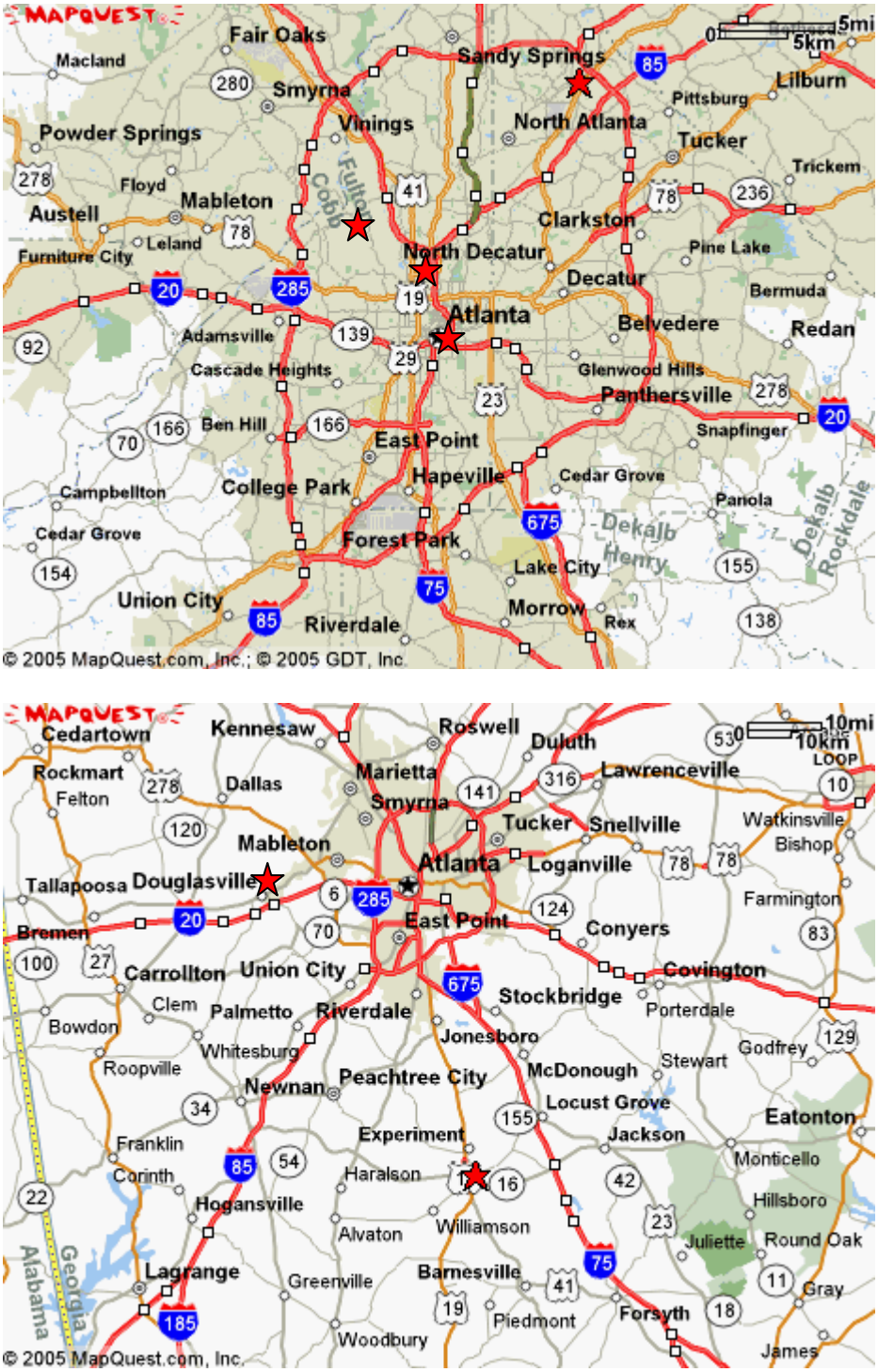


Figure 3.6: PM_{10} monitoring stations; a: urban b: suburban/rural.

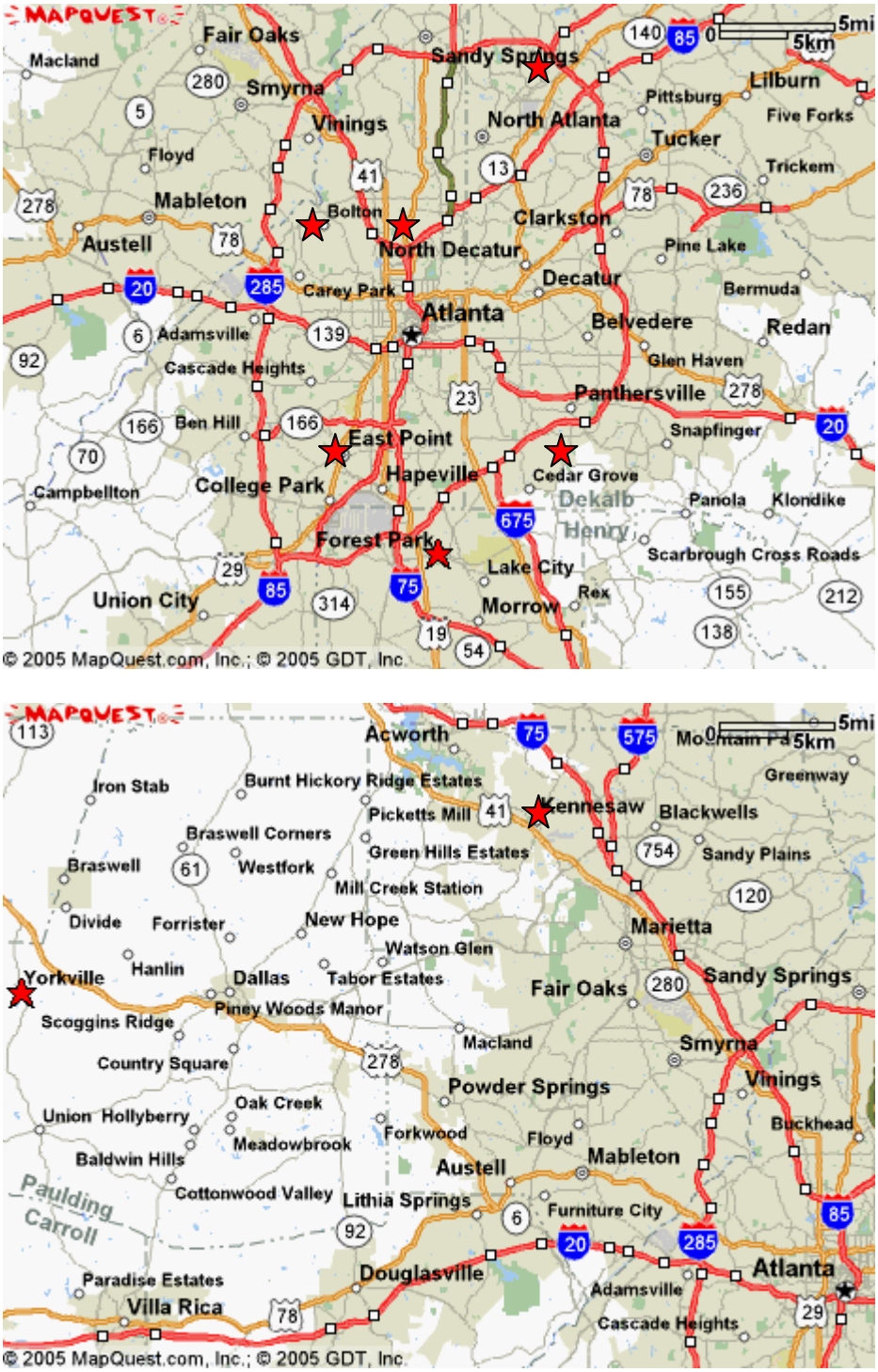


Figure 3.7: $PM_{2.5}$ monitoring stations; a: urban, b: suburban/rural.

Wind speed and direction data was collected for use in the development of pollution rose plots. Whenever possible, meteorological data was used from the same site as that at which pollutant measurements were made. If there was no meteorological data available, the next nearest or most representative site was used. Wind speed and direction data was used from the following sites: Conyers, South Dekalb, Tucker, Yorkville, and Confederate Avenue.

3.1.2 Southeastern Research and Characterization Study

The Southeastern Research and Characterization (SEARCH) study is a collaborative effort between the Electric Power Research Institute and Southern Company to characterize air pollution in the southeast. SEARCH is unique in its use of "supersites"- air quality monitoring sites that contain a large array of pollutant and meteorological measurements. There are eight SEARCH monitoring sites two each in Mississippi, Alabama, Florida, and Georgia. The two Georgia sites are Yorkville and Jefferson Street. Yorkville is also the location of a DNR managed monitoring site as described above. Jefferson Street is on Georgia Power Company property about 4.2 km northwest of the intersection of I-75/85 and I-20. The surrounding area is mixed industrial and residential use. A bus maintenance facility is located 250 m south of the site. All SEARCH stations were established in 1998, although several were located at existing O_3 monitoring stations, including Yorkville.

Measurements used from the Search sites include NO_2, NO_x, SO_2, CO, O_3, PM_{10}, $PM_{2.5}$ and $PM_{2.5}$ components. Meteorological data, including wind speed and direction, was used as well.

At SEARCH sites, all gaseous data is collected every minute. For this research, hourly averages of minute data were used. For most pollutants, the same methods described above were used at SEARCH sites. Ozone is sampled through UV absorption. The minimum detection limit (MDL) for O_3 is 1 part per billion (ppb). NO is sampled via chemiluminescence (CL), with a MDL of 0.5 ppb. NO_2 is measured through photolysis and CL with a 0.1 ppb MDL. This method is believed to be a "truer" means of measuring NO species in the atmosphere and is expected to report lower concentrations that the method used by EPD. CO is sampled using NDIR and has a 5 ppb MDL. SO_2 is not measured via the same method as is used at EPD maintained sites. Instead, SO_2 is measured through UV pulse fluorescence with a 0.2 ppb MDL. This is believed to be a more accurate measurement.

Particle measurements are collected several different ways. Daily, 24 hour filter/gravimetry samples are collected for both $PM_{2.5}$ and PM_{10}. PM2.5 filter samples are also collected using a particle composition monitor (PCM), which provides speciation information (Hansen et al.). They are both also measured using a TEOM sampler.

3.2 Database Design

For this research, a database of air pollutant and meteorological measurements in the Atlanta area was developed. The data used was collected from several public databases, Air Quality System, the SouthEastern Research and Characterization Study

public archive and the National Climatic Data Center archives, and compiled into a hierarchical format.

The Air Quality System (AQS) is a public database maintained by the US Environmental Protection Agency (EPA). Hourly measurements of all gaseous pollutants and $PM_{2.5}$ were requested for the time frame 1993-2003. Daily measurements of $PM_{2.5}$ and PM_{10} were also requested.

The SouthEastern Research and Characterization (SEARCH) study public archive is available online. Hourly measurements were downloaded for all gaseous pollutants, $PM_{2.5}$ and PM_{10} measurements as well as meteorological data.

Meteorological data was requested from the National Climatic Data Center (NCDC) for the Hartsfield Airport monitoring station.

Once the data was collected, AQS data was sorted by pollutant and year and SEARCH data was sorted by site and year into Excel spreadsheet files. Interpolation was then performed if 8 hours of data or less were missing. After interpolation, if a day had 24 observations, the average of these observations was taken. These averages were then transferred into a separate file, the daily file. Each pollutant had a daily file with 24 hour averages for each day from 1993 to 2003 for each site. If there were not 24 observations for a given day at a given site, no value was recorded. Finally, a master file containing 24 hour averages for the entire study period for all pollutants was compiled.

For $PM_{2.5}$ and PM_{10} sites that only recorded 24 hour filter samples, this observation was used in the daily and master files with no alteration.

The entire database is stored in an on-line server to allow access for all researchers on this project.

Chapter 4

DESCRIPTIVE ANALYSIS OF TEMPORAL TRENDS IN ATLANTA AIR QUALITY

The Atlanta Metropolitan Statistical Area is an urban area that has received attention recently for its dramatic growth and poor air quality. The area contains several major interstates with heavy use, large power plants, and various types of industrial plants, all of which contribute to the air pollution problem. Atlanta has been designated by EPA as being out of attainment for both ozone regulatory standards and particulate matter ($PM_{2.5}$) regulatory standards. The Atlanta area was chosen for this research because of the history of elevated pollution levels and because there is a well established air monitoring network for the region.

4.1 Dataset overview

This data used for these analyses consists of hourly measures carbon monoxide, nitrogen oxides, sulfur oxides, ozone, and particulate matter for the years 1993-2003 collected from 17 monitoring sites. If a day was missing less than 8 hours of measurements, these hours were linearly interpolated and used in the analyses. Data completeness was high (average 90% completeness) and around 2% of data was interpolated (see Table 4.1).

Table 4.1: Summary of dataset.

Site	Pollutant	Freq	Network	Time Frame	Completeness-Raw Data	Interpolated
Dekalb Tech	CO	Hourly	EPD	01/1993-06/2003	90.19%	1.74%
Roswell Road	CO	Hourly	EPD	08/ 1994-12/ 2003	97.83%	1.25%
Jefferson Street	CO	Hourly	SEARCH	08/ 1998-12/2002	89.84%	4.02%
Yorkville	CO	Hourly	SEARCH	01/ 1993-12/2002	85.51%	1.92%
Georgia Tech	NO_x/NO_2	Hourly	EPD	01/ 1993-12/2003	89.87%	3.07%
South Dekalb	NO_x/NO_2	Hourly	EPD	01/ 1993-12/2003	87.23%	1.57%
Tucker	NO_x/NO_2	Hourly	EPD	04/1995-12/2003	85.42%	1.94%
Conyers	NO_x/NO_2	Hourly	EPD	04/ 1994-12/2003	94.32%	2.05%
Yorkville	NO_x/NO_2	Hourly	EPD	01/ 1996-12/2003	88.69%	2.13%
Jefferson Street	NO_x/NO_2	Hourly	SEARCH	08/1998-12/2002	84.20%	2.38%
Georgia Tech	SO_2	Hourly	EPD	01/ 1993-12/2003	92.12%	1.26%
Stilesboro	SO_2	Hourly	EPD	01/ 1993-12/2003	96.19%	1.22%
Confederate Avenue	SO_2	Hourly	EPD	01/ 1993-12/2003	95.03%	1.49%
Jefferson Street	SO_2	Hourly	SEARCH	08/ 1998-12/2002	90.15%	3.96%
Yorkville	SO_2	Hourly	SEARCH	01/ 1993-12/2002	88.81%	1.52%
South Dekalb	O_3*	Hourly	EPD	01/ 1993-12/2003	93.90%	1.48%
Confederate Avenue	O_3*	Hourly	EPD	01/ 1993-12/2003	97.59%	1.70%
Conyers	O_3*	Hourly	EPD	01/ 1993-12/2003	97.47%	1.51%
Jefferson Street	O_3	Hourly	SEARCH	08/1998-12/2002	94.20%	2.32%
Yorkville	O_3	Hourly	SEARCH	01/ 1993-12/2002	88.97%	1.97%
Doraville Health Center	$PM_{2.5}$	Daily	EPD	01/ 1999-12/2003	85.93%	
E Rivers School	$PM_{2.5}$	Daily	EPD	01/ 1999-12/2003	84.07%	
South Dekalb	$PM_{2.5}$	Daily	EPD	01/ 1999-12/2003	83.68%	
Fire Station 8	$PM_{2.5}$	3rd day	EPD	01/ 1999-12/2003	89.84%	
East Point Health Center	$PM_{2.5}$	3rd day	EPD	01/ 1999-12/2001	83.61%	
Forest Park	$PM_{2.5}$	3rd day	EPD	01/ 1999-12/2003	88.36%	
Kennesaw	$PM_{2.5}$	3rd day	EPD	01/ 1999-12/2003	88.20%	
Yorkville	$PM_{2.5}$	3rd day	EPD	01/ 1999-12/2003	88.69%	
Jefferson Street	$PM_{2.5}$	Daily	SEARCH	01/ 1998-12/2002	78.64%	
Jefferson Street	$PM_{2.5}$	Hourly	SEARCH	08/ 1998-12/2002	84.78%	
Yorkville	PM_{10}	6th day	EPD	05/ 1998-12/2002	76.70%	
Fire Station 8	PM_{10}	6th day	EPD	01/ 1993-12/2003	91.80%	
Doraville Health Center	PM_{10}	6th day	EPD	01/ 1993-12/2003	88.77%	
Griffin	PM_{10}	6th day	EPD	01/ 1993-12/2003	86.74%	
Fulton Co Health Department	PM_{10}	6th day	EPD	01/ 1993-12/2003	88.97%	
Douglasville	PM_{10}	6th day	EPD	01/ 2000-12/2003	94.26%	
E Rivers School	PM_{10}	6th day	EPD	01/ 2000-12/2003	77.87%	
Georgia Tech	PM_{10}	Hourly	EPD	01/ 1998-12/2003	97.52%	

4.2 Annual Trends

For several pollutants, CO, SO_2, O_3, PM_{10}, and $PM_{2.5}$, annual average concentrations have been decreasing over the past few years. This is particularly evident for CO, where 1 hour maximum, 8 hour maximum, and 24 hour average concentrations have appreciably decreased at all monitoring sites except Yorkville (see Figure 4.1), which is considered a background site as it regularly measures concentrations much lower than the more urban stations. Nitrogen oxides and ozone exhibit a slight increase in concentrations from 1994-1998 and then a decrease from 1998-2003 (see Figure 4.3 and 4.4). The similarity in profiles is not unusual as NOx is an ozone precursor. Although NOx monitoring sites have a large range in concentrations, O3 monitoring sites are relatively uniform throughout the years. PM_{10} and $PM_{2.5}$ are also relatively uniform and exhibit similar magnitudes of decreasing concentration (see Figure 4.5).

Mobile sources are large contributors to CO, NO_x, and PM in Atlanta. Increases in both population and vehicle miles traveled in Atlanta have led to increased mobile emissions over the course of this study period. For the metropolitan Atlanta area, population increased from 2,557,800 in 1990 to 3,429,379 in 2000 (US Census, 1990 & 2000). The Atlanta Regional Commission estimates a further increase to 3,669,300 in 2003. Most of this growth has been in the northern area of the region. At the same time, vehicle miles traveled have increased from 26 miles/person/day to 32 miles/person/day (ARC, 2005).

In 1996, the Georgia Environmental Protection Division (EPD), through the establishment of the Clean Air Force, implemented a mandatory, annual emissions test for all cars older than 3 years. This test checks tailpipe emissions levels, catalytic

converters, and fuel cap seals. This program specifically targets reductions in ozone precursors and PM. Although it is not possible to estimate exact levels of reduction, it is estimated that this program has lowered mobile NO_x and PM emissions (Georgia's Clean Air Force).

Sulfur dioxide concentrations decreased most noticeable between 1993 and 1995. Levels between 1995 and 2000 remained steady or slightly decreased at most sites. After 2000, another slight decrease in concentrations is seen. A noticeable exception is Jefferson Street, which exhibited a large spike in SO2 concentrations, both 1 hour maximum and 24 hour average, in 2000. This is not seen at any other monitoring sites. (See Figure 4.2) Average monthly concentrations were slightly higher in 2000 than in 2001 and 2002 for most of the year and two to three times as high from October through December.

Title IV of the Clean Air Act Amendments of 1990 specifically addresses SO_2 and NO_x from fossil-fuel fired power plants. Phase one of this program called for 263 power plants to reduce SO_2 emissions through a cap-and-trade program. All of the power plants in or near the metropolitan Atlanta area were included in the initial listing; however several plant substitutions (i.e. decreased emissions at power plants not in the Atlanta area were considered in the overall reductions) were made in Georgia to help achieve the lower emissions. Overall SO_2 emissions for the plants were decreased by over 50% by 1995. Phase two of this program included almost all power plants in Georgia. This phase requires all power plants to emit no more than 1.2 pounds of SO_2 per million Btu heat input by 2000. Title IV also calls for a 2 million ton reduction in

NO_x from the power plants identified, however no formal cap-and-trade program or emissions monitoring program was used.

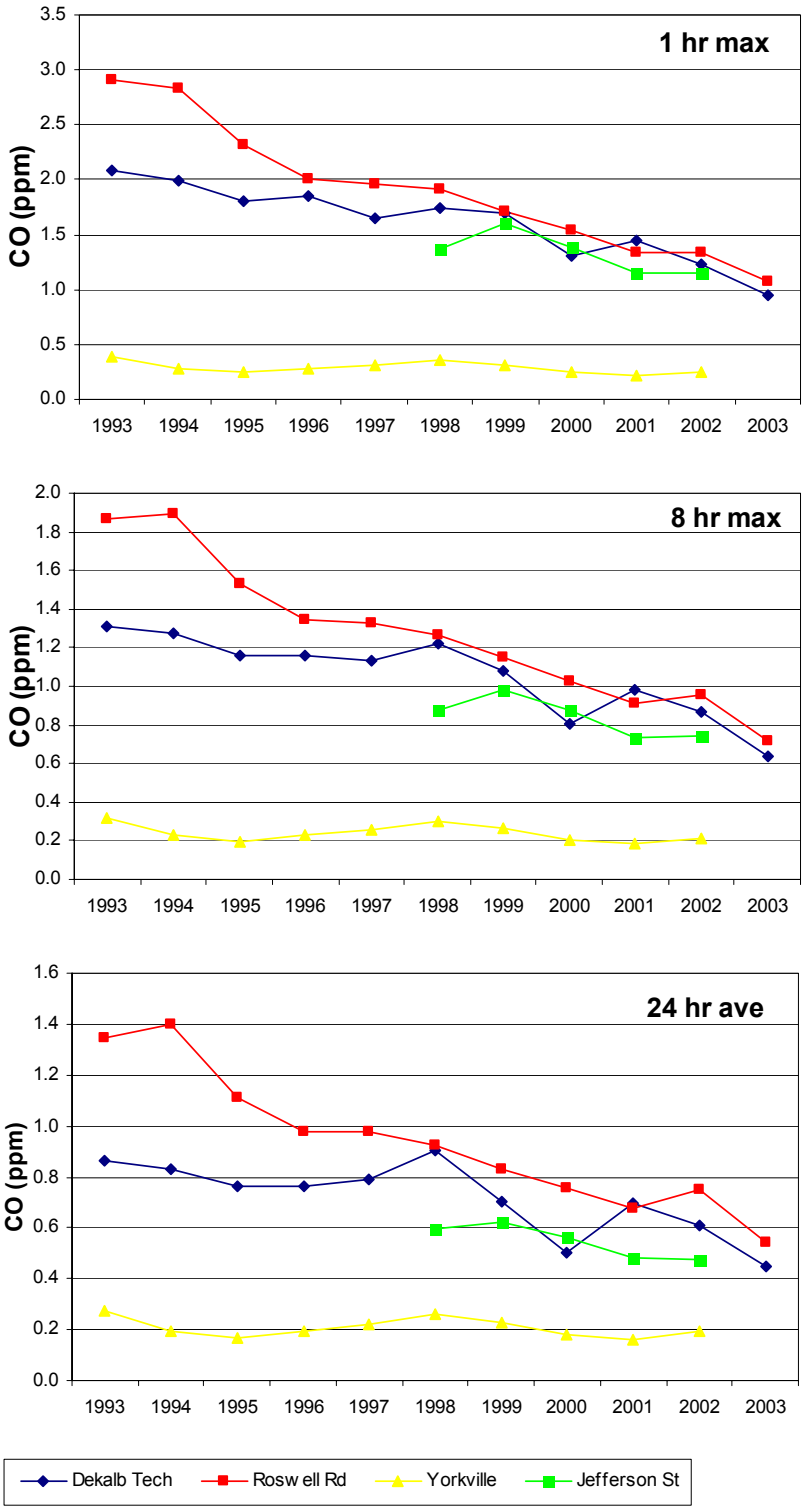


Figure 4.1: CO annual averages.

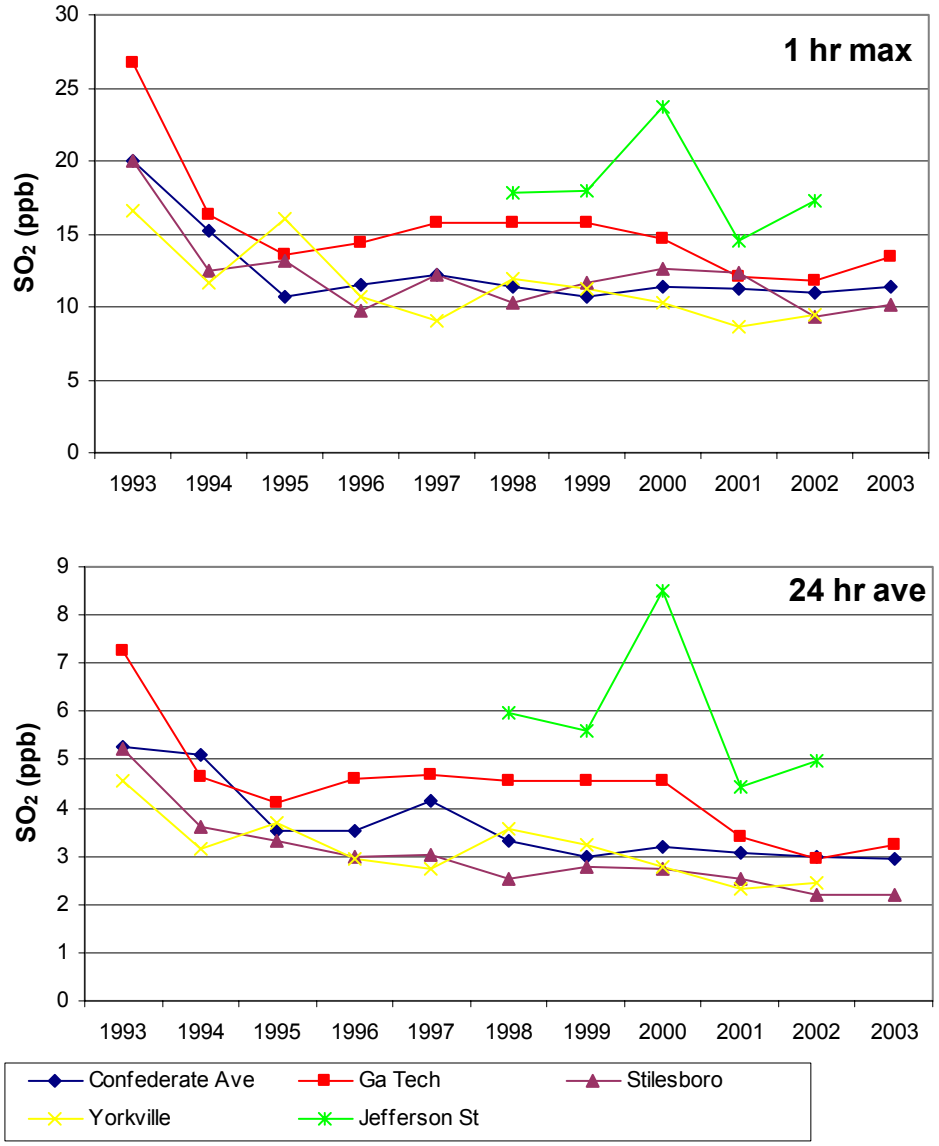


Figure 4.2: SO_2 annual averages.

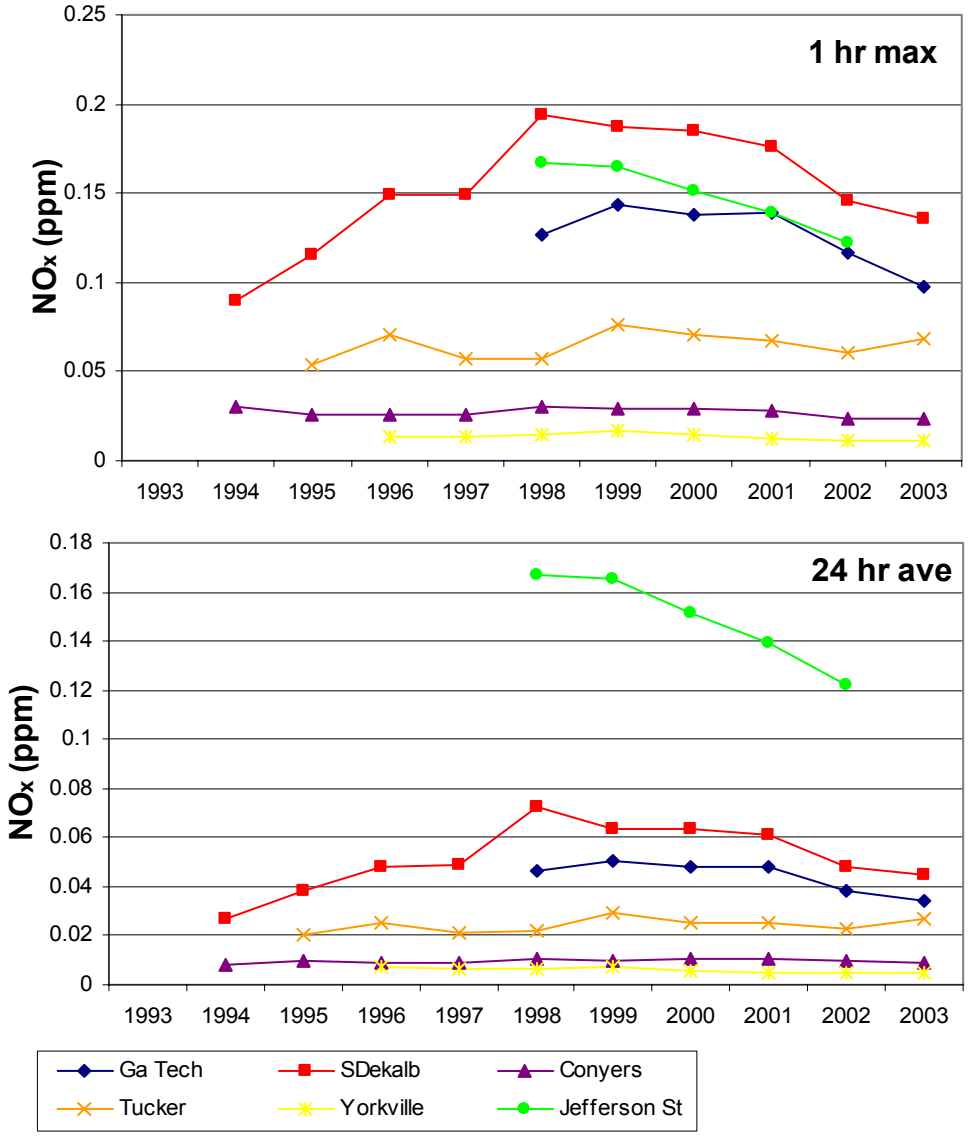


Figure 4.3: NO_x annual averages.

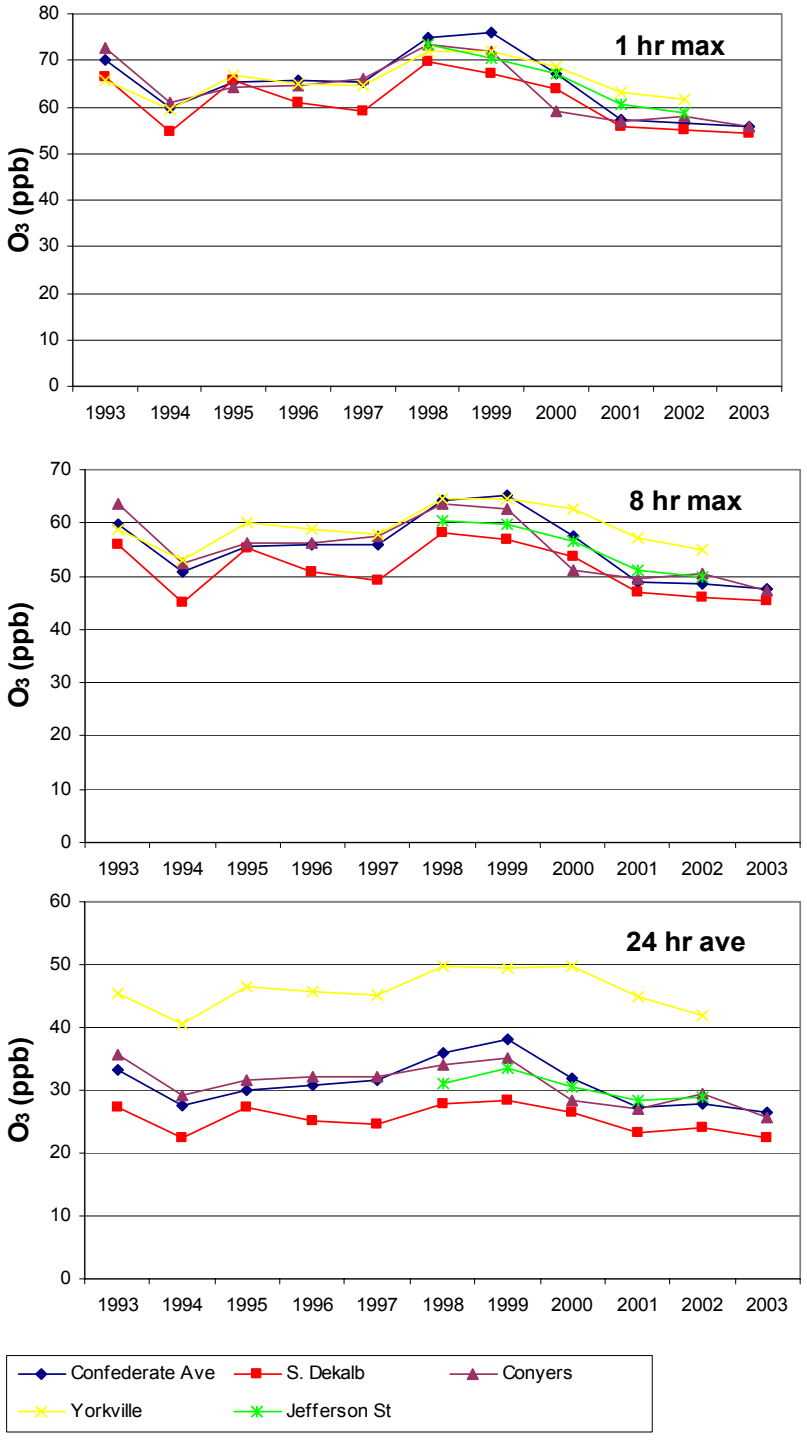
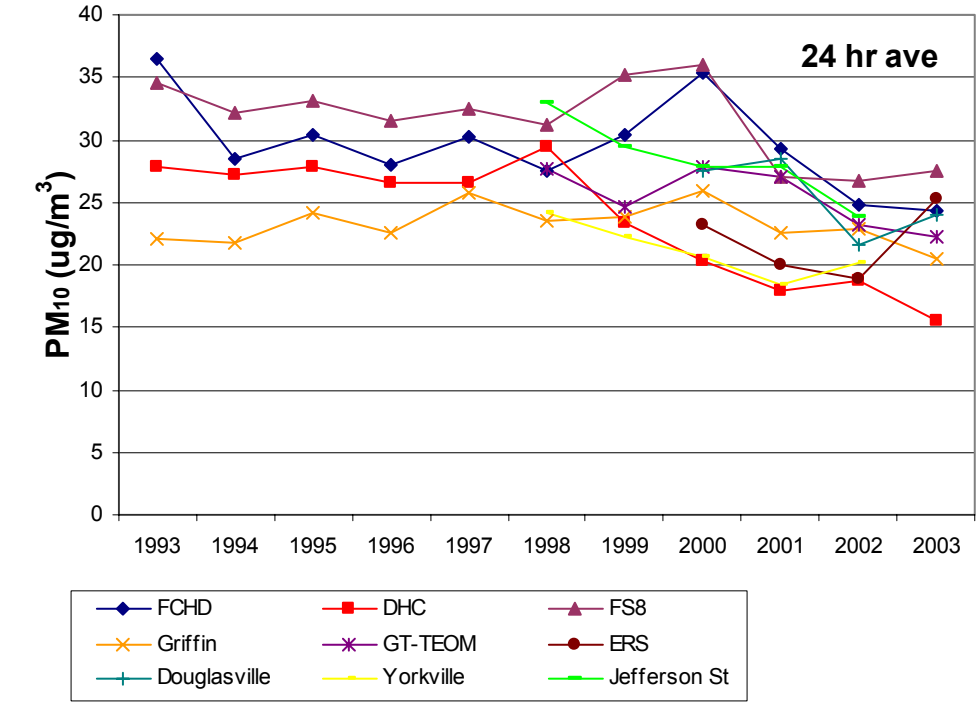


Figure 4.4: O_3 annual averages.



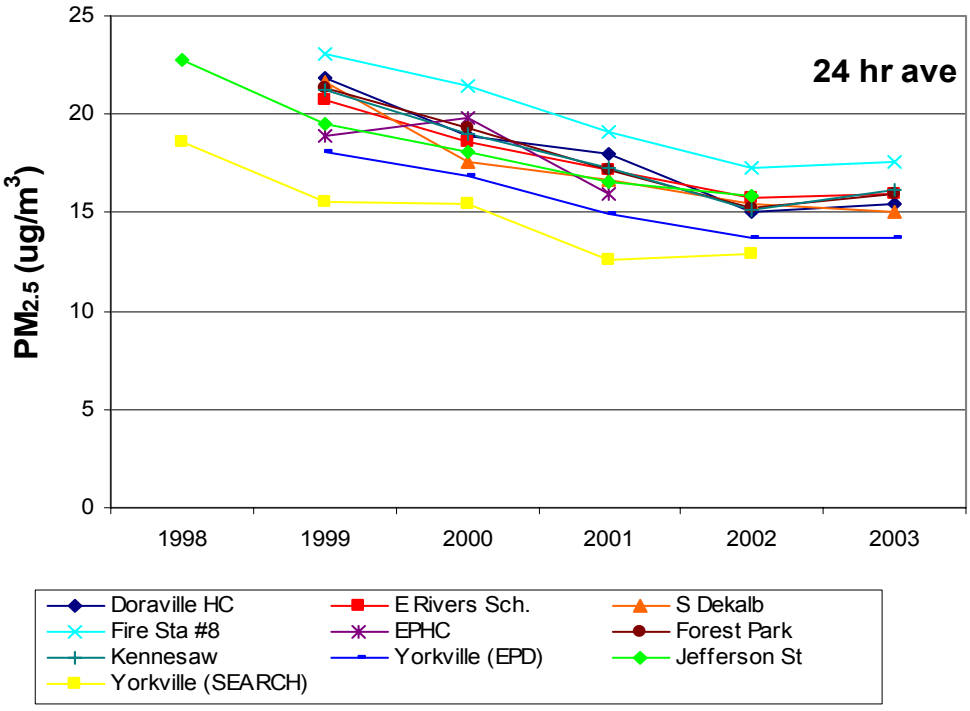


Figure 4.5: PM_{10} and $PM_{2.5}$ annual averages.

4.3 Monthly Profiles

Many air pollutants vary by season due to the effects of meteorology and photochemistry. The Atlanta area receives a large quantity of solar radiation between May and August, increasing photochemistry. Temperatures are lowest, and vertical mixing most limited, during the winter months, November through February.

Carbon monoxide's seasonal profile is based largely on temperature. In the cooler winter months, there is less vertical mixing in the atmosphere. However, emissions of CO do not change. The emissions that are well mixed vertically in the summer are trapped at the lower levels of the troposphere in the winter, leading to higher concentrations (see Figure 4.6).

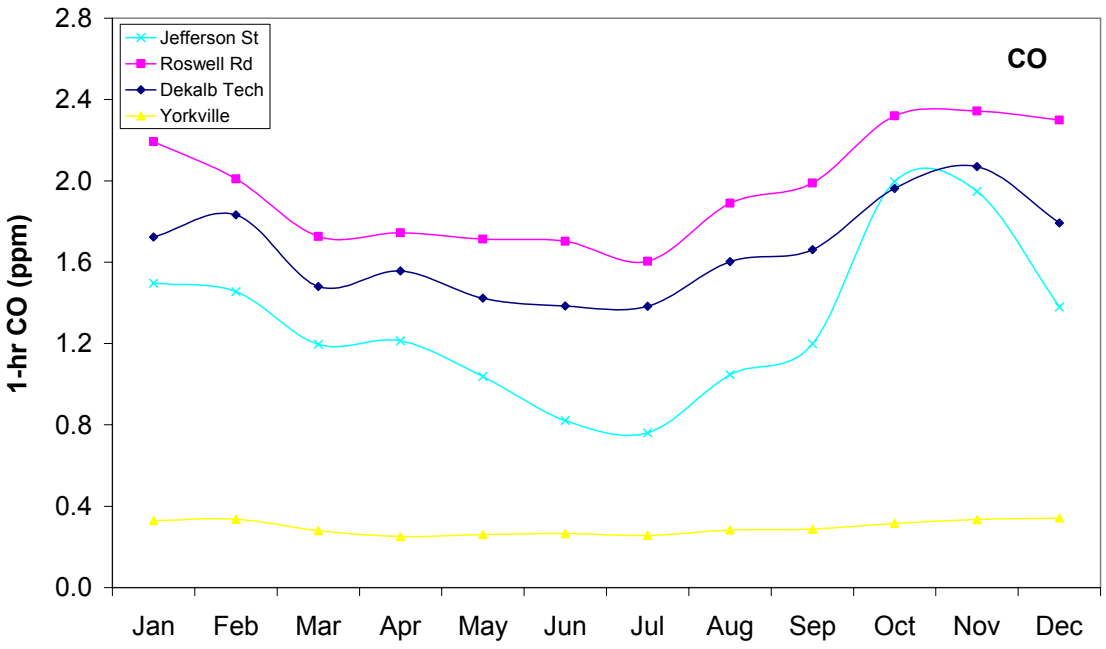


Figure 4.6: Seasonal profiles of CO. Data from 1993-2003 except Jefferson Street (1998-2002) and Yorkville (1996-2002).

Nitrogen oxides exhibit a similar seasonal pattern as CO (see Figure 4.7). This is in part because of the vertical mixing discussed above. There is also less photochemistry in the winter due to less solar radiation. Photochemical reactions to form ozone lower NO_x concentrations in the summer, as do reactions leading to the formation of peroxy acetic nitric anhydride (PAN), which is a reservoir for NO species. The lack of these reactions in the winter leads to elevated NO_x concentrations.

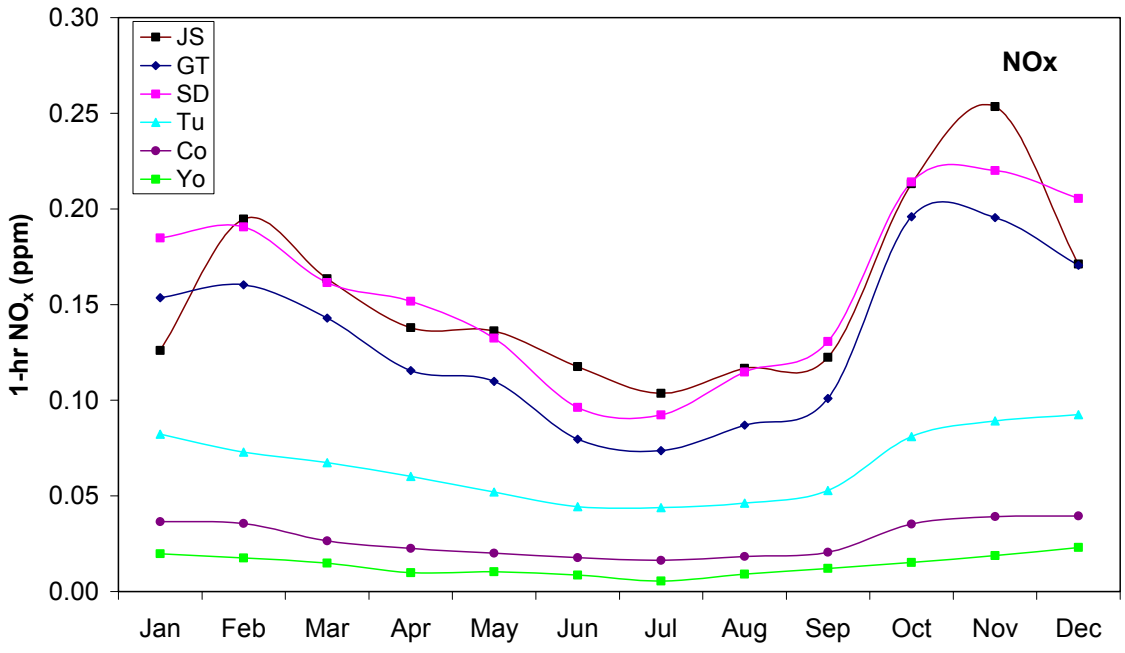


Figure 4.7: Seasonal profiles of NO_x. Data from 1993-2003 except Jefferson Street (1998-2002) and Yorkville (1996-2002).

The variation in SO_2 by season is largely due to changes in emissions. The major sources of SO_2 are coal-fired power plants, and electricity demand peaks in summer due to air conditioning. In the peak summer months, oil and natural gas are used to supplement coal fuel in the power plants affecting the Atlanta area. This compensates for the increase in power generation for air conditioning. Nonetheless, there appears to be a July peak (see Figure 4.8).

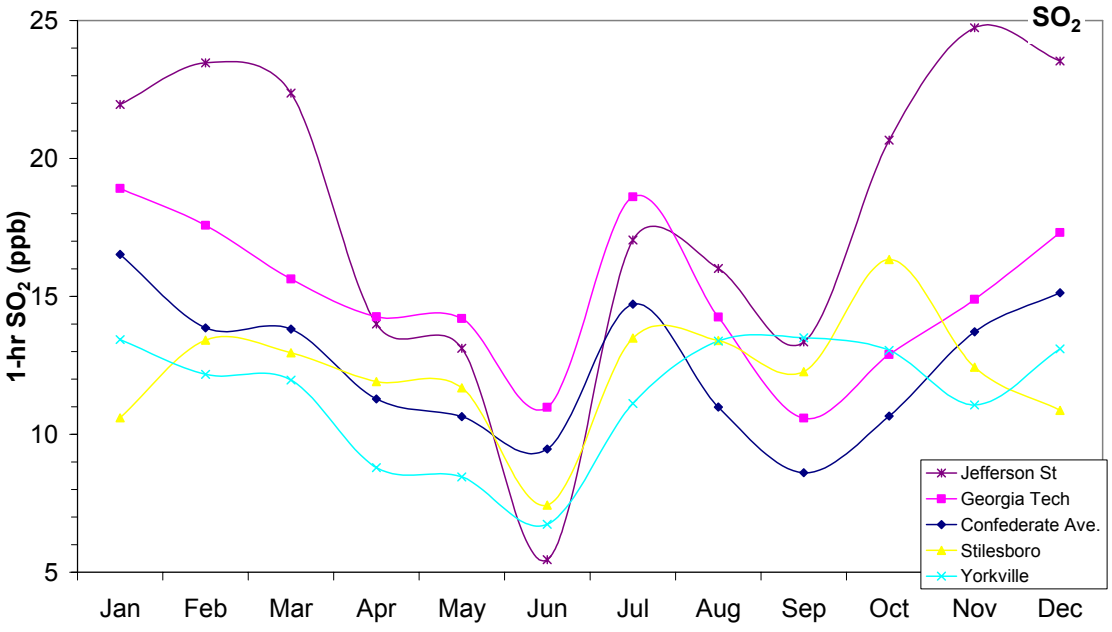


Figure 4.8: Seasonal profiles of SO2. Data from 1993-2003 except Jefferson Street (1998-2002) and Yorkville (1996-2002).

Because there are an abundance of ozone precursors (reactive organic compounds (ROGs) and NO_x) in the Atlanta area during all months, seasonal ozone patterns are driven by photochemistry. ROGs have a large biogenic source in the Atlanta area and do

not limit O_3 formation. The summer months have the largest amount on incoming solar radiation, as well as high levels of NO_x and ROGs, and therefore the largest amount of ozone production (see Figure 4.9).

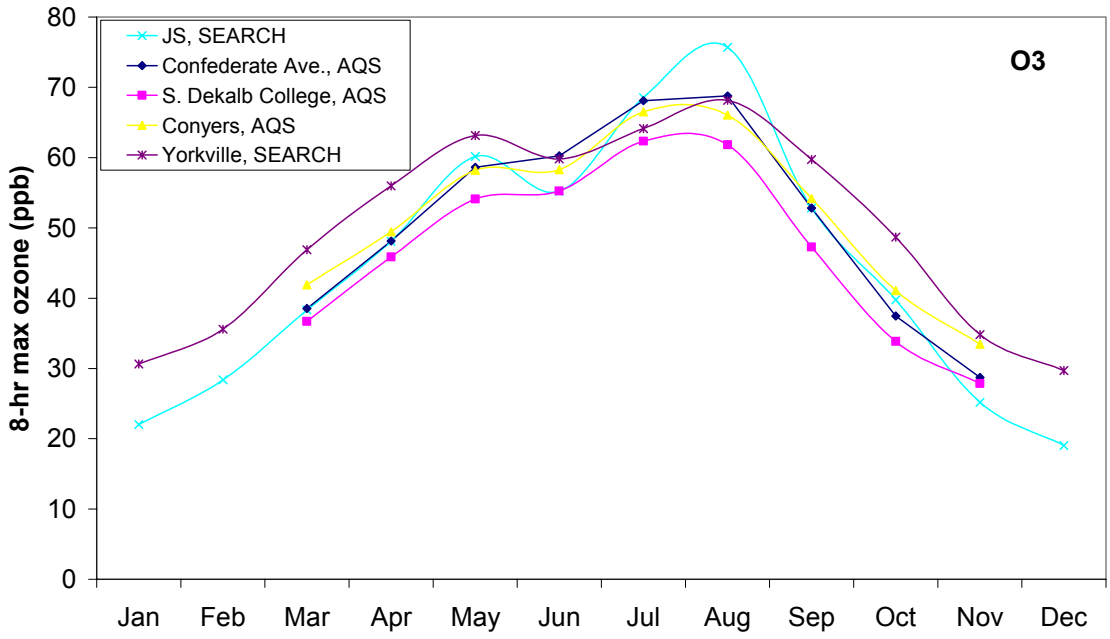


Figure 4.9: Seasonal profiles of ozone. Data from 1993-2003 except Jefferson Street (1998-2002) and Yorkville (1996-2002).

Particulate matter displays two peak concentrations annually- first in July/August and then in November (see Figure 4.10). The July/August peak is largely due to increased solar radiation leading to increased formation of secondary components of $PM_{2.5}$.

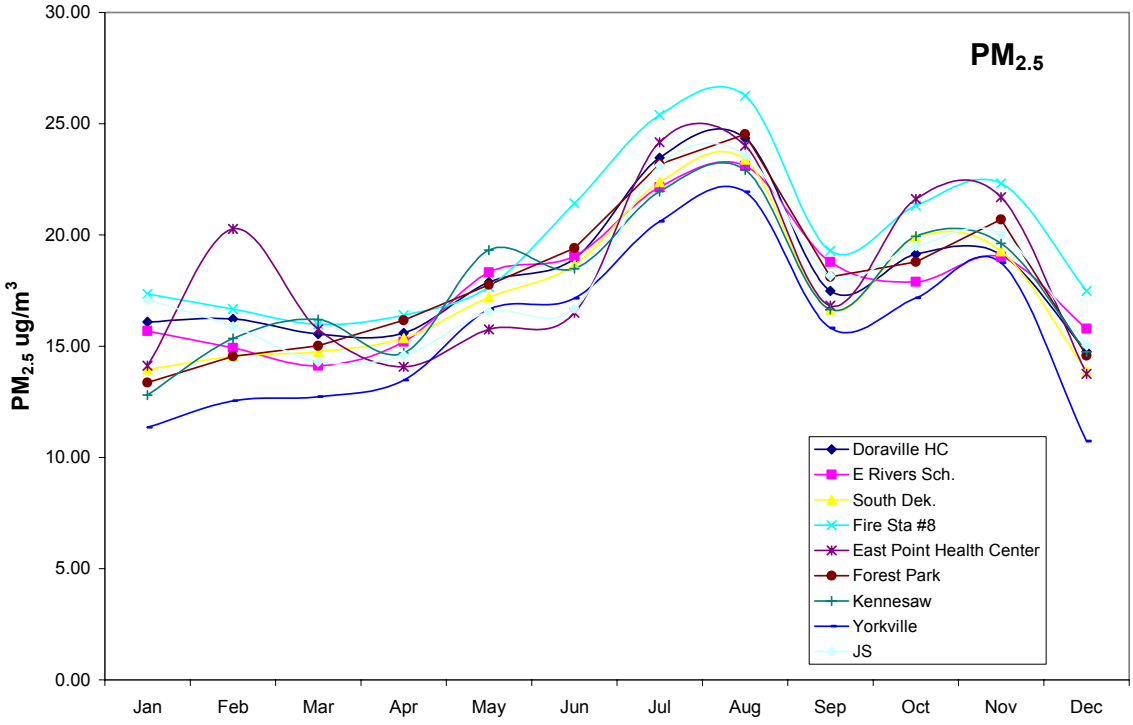


Figure 4.10: Seasonal profiles of $PM_{2.5}$. Data from 1998-2003 except Jefferson Street (1998-2002).

4.4 Weekly Profiles

Air pollution concentrations also exhibit weekly patterns. These patterns are dependent only on patterns in emissions as meteorology and photochemistry are not affected by day of week. Several major sources of air pollution, however, do exhibit weekly emissions profiles.

As mentioned above, carbon monoxide is mainly emitted through mobile sources. There is not a strong weekly profile in mobile emissions, but there is a slight tendency

towards less miles traveled on the weekends. This is evident in the urban CO monitoring sites, particularly the Roswell Road site (see Figure 4.11). This site is situated to monitor emissions from mobile sources on several major interstates (GA 400 and I-75/85) that are used largely for commuting during the week. It is expected that on the weekends there is less traffic and therefore fewer emissions. Dekalb Tech and Jefferson Street exhibit only a slight decrease in concentrations on the weekend.

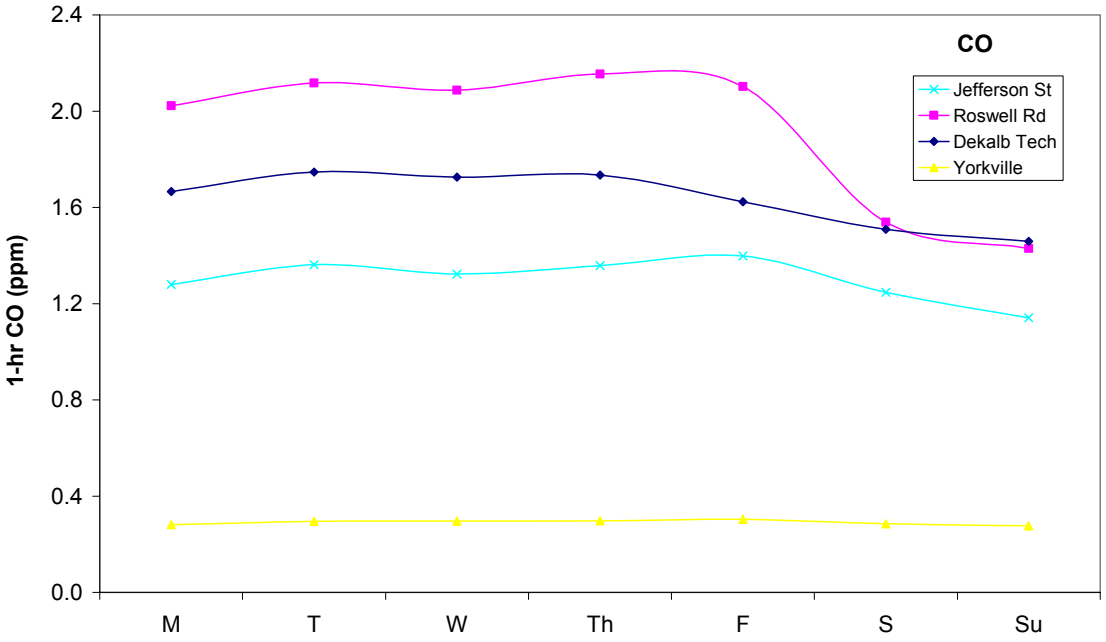


Figure 4.11: Day of week profiles of CO. Data from 1993-2003 except Jefferson Street (1998-2002) and Yorkville (1996-2002).

Nitrogen oxides are also largely emitted by mobile sources and so reflect the same weekly profile as CO. Weekly patterns are more dramatic for NO_x than CO, with all of sites except the most rural (Yorkville) displaying a decrease in concentrations on the weeknd (see Figure 4.12). Photochemical reactions of NO_x to ozone are not affected by

day of week and exaggerate the decrease in NO_x concentrations. At the most urban sites, Jefferson Street, Georgia Tech, and South Dekalb, there is also a one day build-up in concentrations from Monday to Tuesdays.

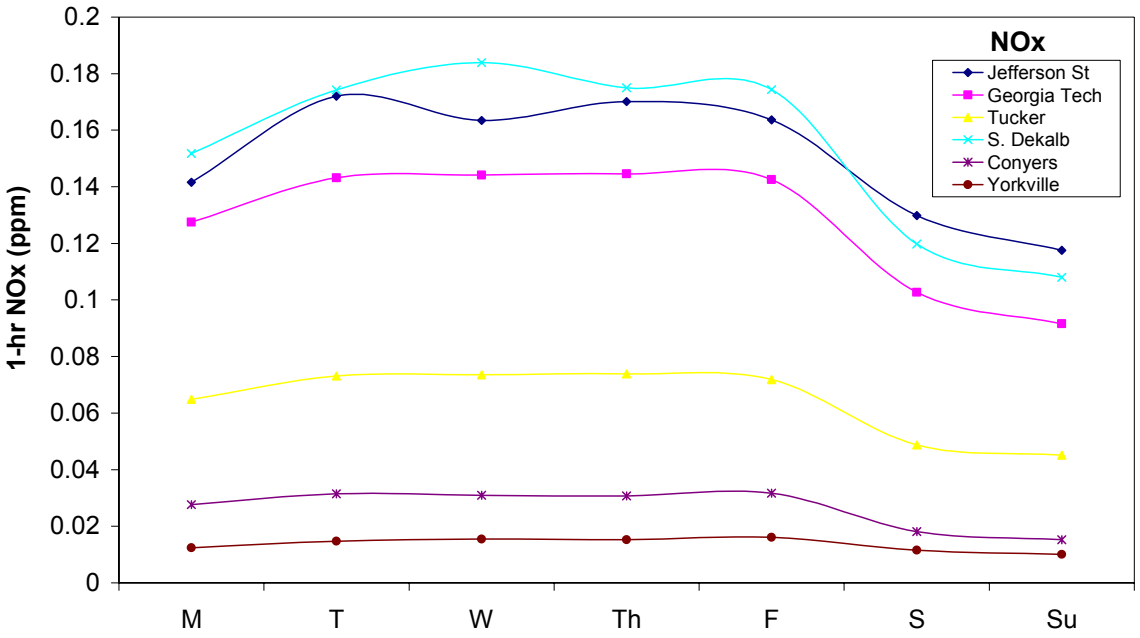


Figure 4.12: Day of week profiles of NO_x.

In the urban sites, Jefferson Street, Georgia Tech, and Confederate Avenue, there is a slight downward trend in SO_2 concentrations towards the end of the week (see Figure 4.13). This is possible due to less diesel traffic on the weekends.

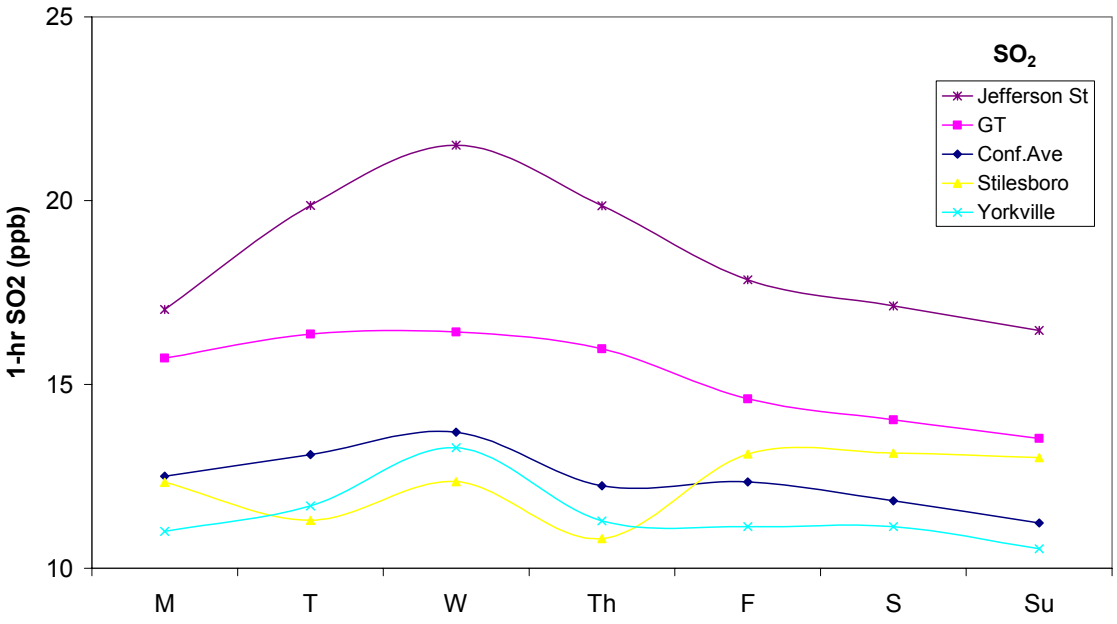


Figure 4.13: Day of week profiles of SO_2. Data from 1993-2003 except Jefferson Street (1998-2002) and Yorkville (1996-2002).

Because ozone is a secondary pollutant and not directly emitted into the atmosphere, it has little variation by day of week. What is seen at urban sites is a gradual build up of ozone concentrations throughout the week due to the increase in ozone precursors (see Figure 4.14). This is most dramatic at Jefferson Street and does not occur at all at the rural sites, Conyers and Yorkville. The Jefferson Street and Yorkville sites include winter data (November-February) which could affect the patterns observed.

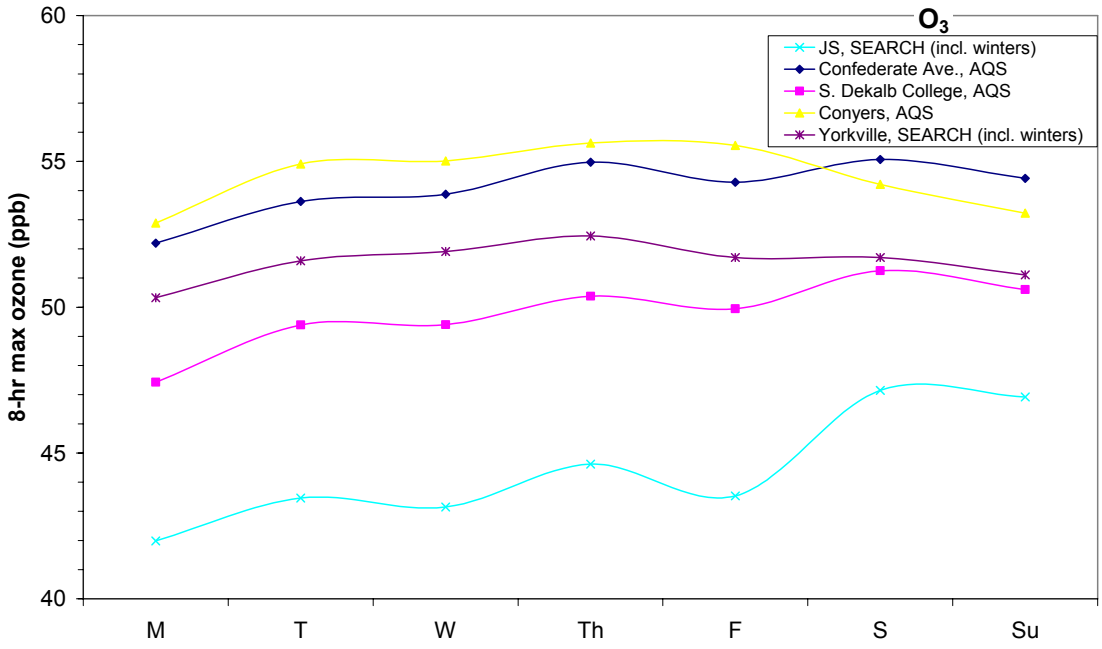


Figure 4.14: Day of week profiles of O_3. Data from 1993-2003 (March-October) except Jefferson Street (1998-2002, all months) and Yorkville (1996-2002, all months).

Particulate matter is a combination of primary emissions and secondary formation. There are several components of $PM_{2.5}$, each with unique emissions sources. This leads to little day of week variation in $PM_{2.5}$ concentrations (see Figure 4.15). In Atlanta, mobile sources, both gasoline and diesel, are a significant fraction of $PM_{2.5}$, possible explaining the slight drop in concentrations at most sites on the weekends.

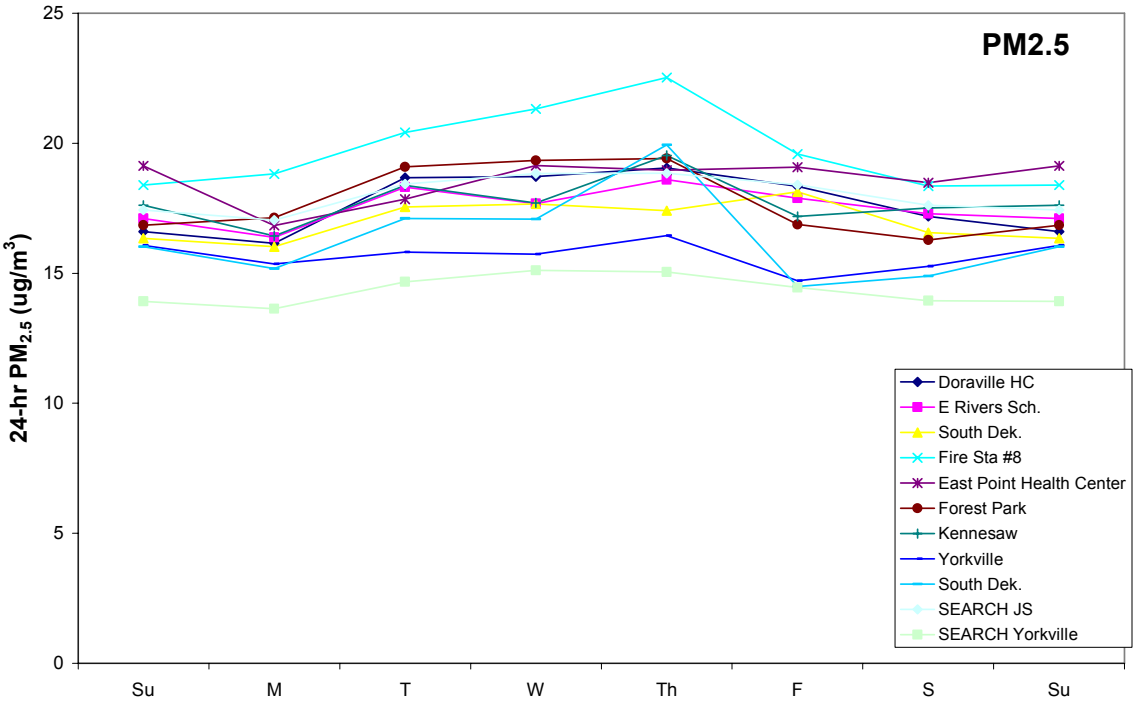


Figure 4.15: Day of week profiles of $PM_{2.5}$. Data from 1998-2003 except Jefferson Street (1998-2002) and SEARCH Yorkville (1998-2002).

4.5 Diurnal Profiles

Photochemistry, meteorology, and emissions of air pollutants all have distinctive diurnal patterns. As a result, air pollution concentrations conform to unique diurnal profiles. Solar radiation peaks between 10 am and 2 pm and two major rush hours are observed between 6 am and 8 am and between 5 pm and 7 pm.

Carbon Monoxide (CO) is a primary pollutant, i.e. it is directly emitted into the atmosphere. It is a relatively inert gas in the atmosphere. In Atlanta, CO concentrations

are mainly impacted by mobile emissions. Peak CO concentrations corresponding with rush hour events (hours 6-9 and 17-20) are evident when one examines graphs of the diurnal profiles of CO (see Figure 4.16). For urban sites near large roadways, Jefferson Street, Roswell Road and Dekalb Tech, the peak concentration corresponding to morning rush hour is most prominent. In the evenings and overnight, CO concentrations remain elevated due to a stable boundary layer that does not allow much vertical mixing. At the background site, Yorkville, there are only a few small roadways nearby and so no diurnal pattern in observed.

It is also evident when examining the yearly diurnal profiles that although overall CO concentrations are decreasing in the Atlanta area, the diurnal profiles are not changing.

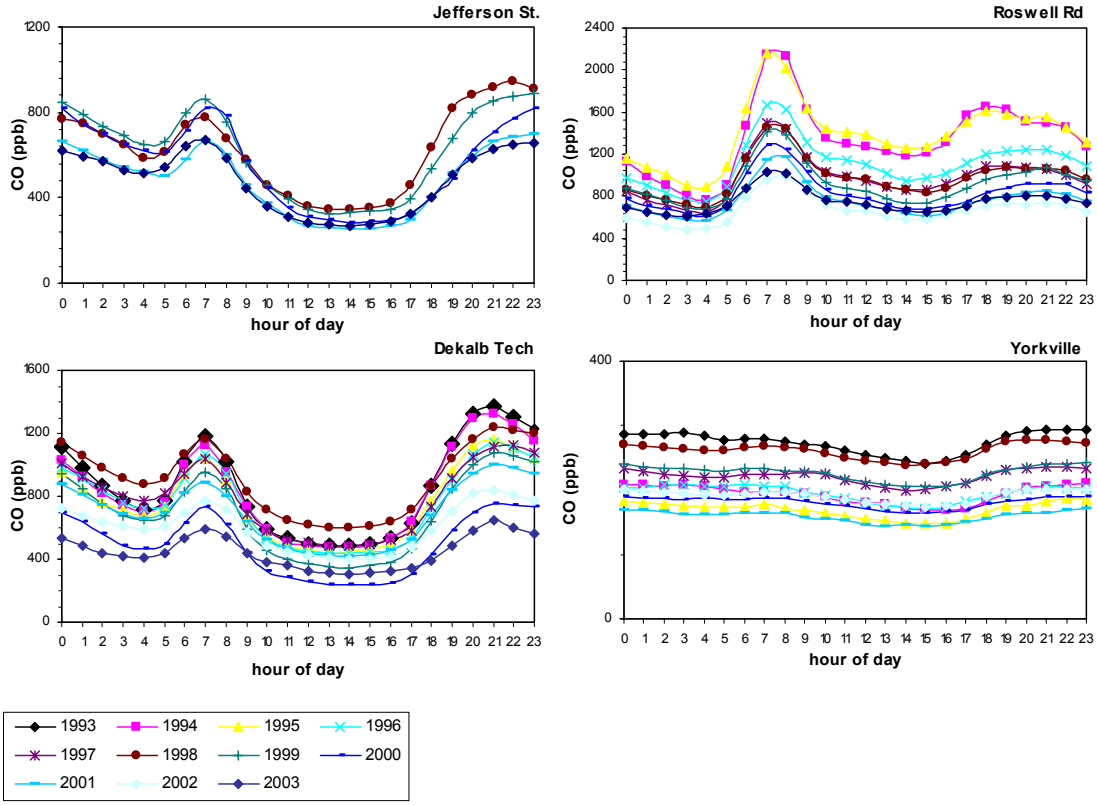


Figure 4.16: Diurnal profiles of CO.

Similar to CO, nitrogen oxides (NO_x) are also emitted by mobile sources. Diurnal NO_x profiles show the same morning rush hour peak as CO (see Figure 4.17). However, NO_x is also affected by photochemistry. NO_x is a precursor to ozone, which is produced at times of high photochemical energy, in the middle of the afternoon. The afternoon decrease in NO_x concentrations (hours 12-15) is much more dramatic than that observed in the CO profiles. This is because NO_x emissions are decreasing at the same time as an increase in photochemistry and therefore scavenging of NO_x to form ozone. Urban sites show the strongest diurnal NO_x concentrations while rural sites exhibit relatively constant levels throughout the day.

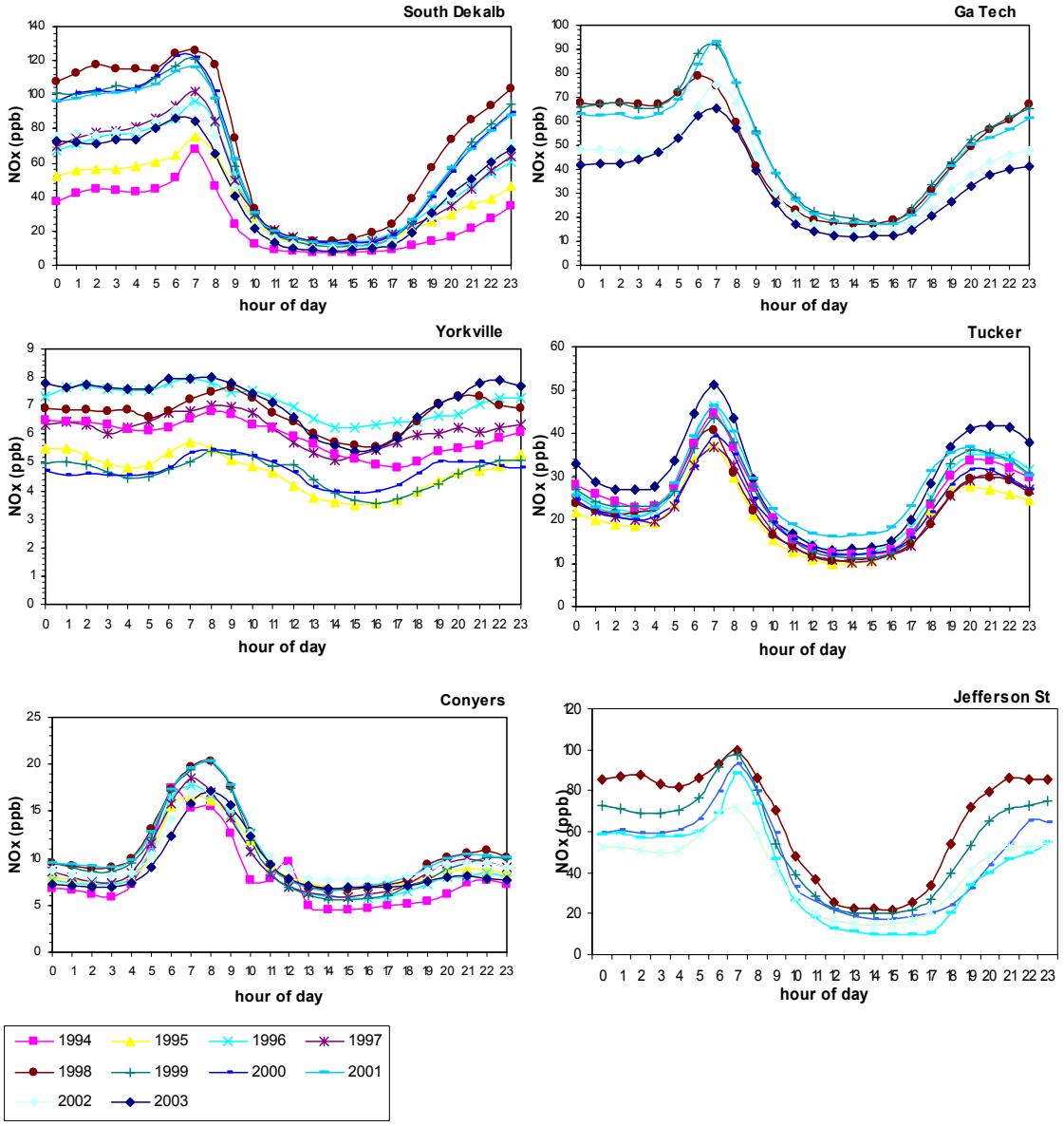


Figure 4.17: Diurnal profiles of NO_x.

The largest source of sulfur dioxide in the Atlanta area is power plant emissions. Other sources include diesel vehicles and industrial emissions. None of these sources emits with a strong diurnal pattern. This is reflected in the graphs of the diurnal profiles of SO_2 (see Figure 4.18). The only site that has an obvious diurnal profile is Stilesboro. This is probably due to power plant fumigation episodes. Power plants generally emit much higher in the atmosphere than other pollution sources. Because of the boundary layer, these emissions often stay high in the atmosphere until there is sufficient warming to begin mixing. When this happens, the power plant emissions are mixed downwards. This warming usually begins late morning (about hour 9). Stilesboro is the closest monitoring site to Plant Bowen, a large, coal-fired power plant, and so it exhibits the largest diurnal profile in SO_2 emissions. Confederate Avenue, Georgia Tech, and Yorkville have weaker diurnal profiles and no pattern is evident at Jefferson Street.

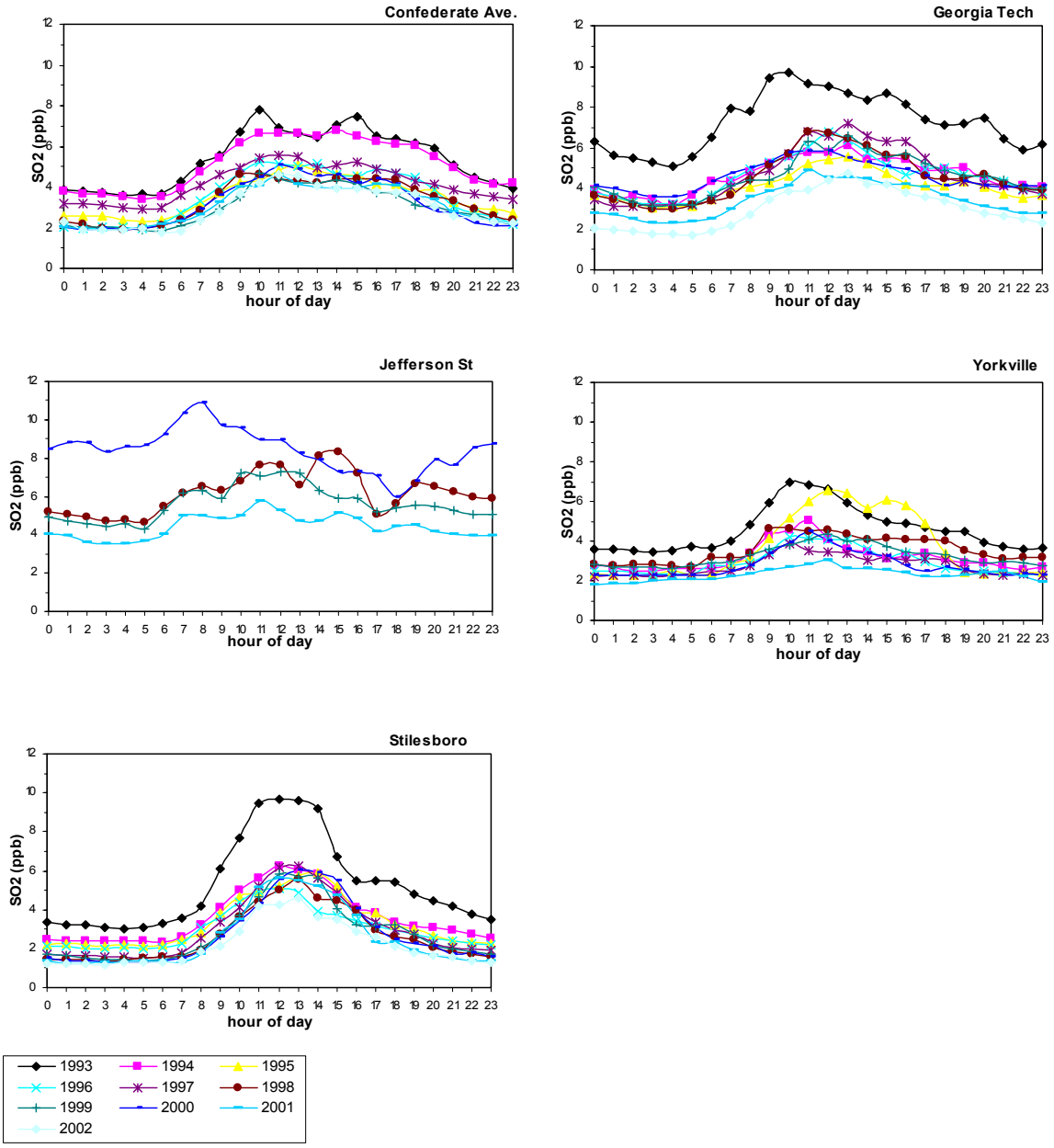


Figure 4.18: Diurnal profiles of SO_2.

Unlike CO, NO_x, and SO_2, ozone (O_3) is formed in the atmosphere through photochemical reactions. Because of this, ozone concentrations are highest during the late afternoon, when there is the most solar radiation to drive atmospheric photochemistry. There is only one, large peak in O_3 concentrations in the afternoon (hours 13-16) that is observed at all sites (see Figure 4.19). At the rural site, Yorkville, this peak is much smaller than at the urban sites.

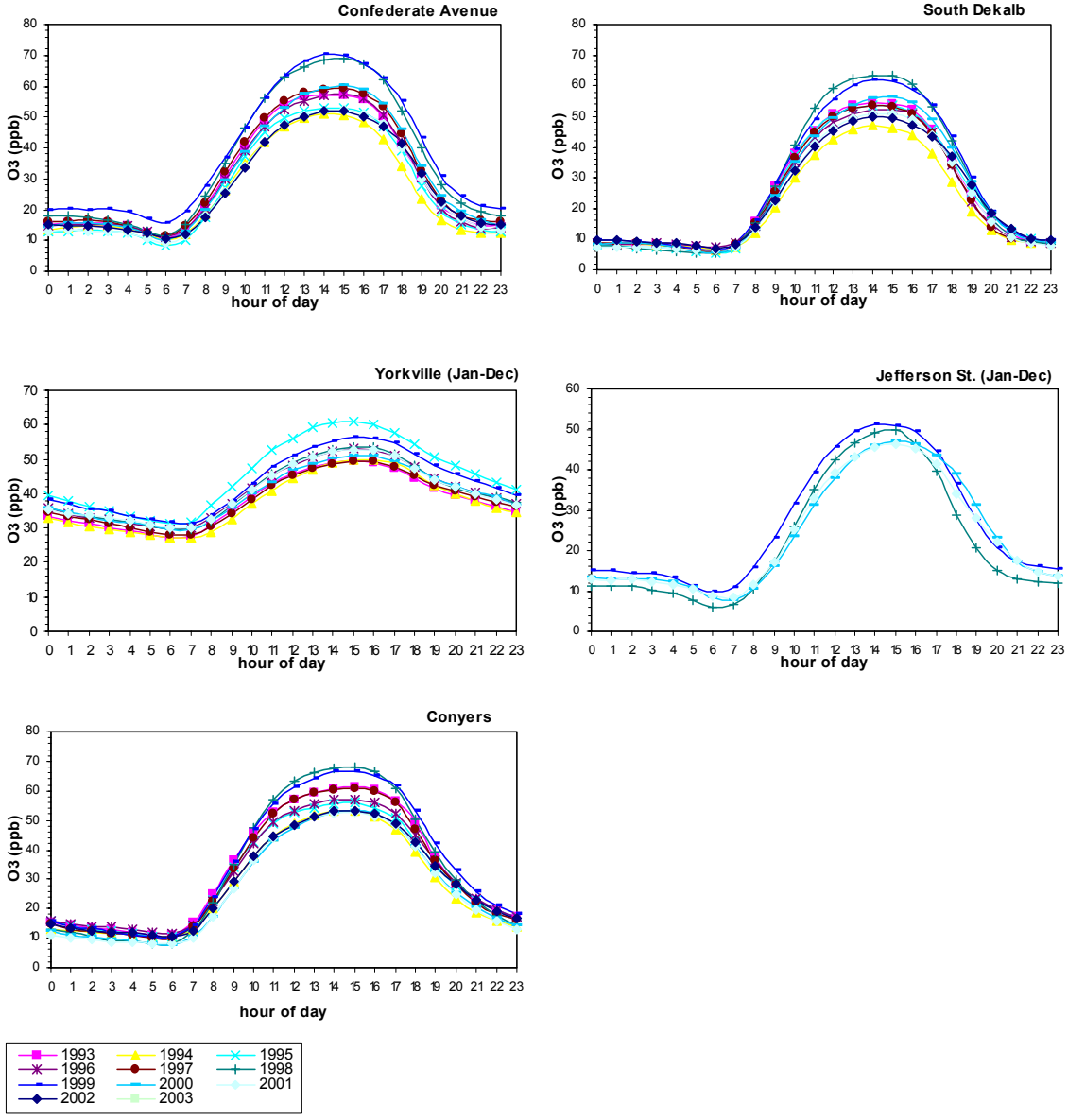


Figure 4.19: Diurnal profiles of O_3.

Hourly $PM_{2.5}$ data was only available for Jefferson Street for this study period. The diurnal profiles observed are very similar to those seen for CO and NO_x at this site, implying that mobile sources are contributing largely to the diurnal variability(see Figure 4.20). This site is near a diesel bus facility which could be impacting the $PM_{2.5}$ levels at this site and not others.

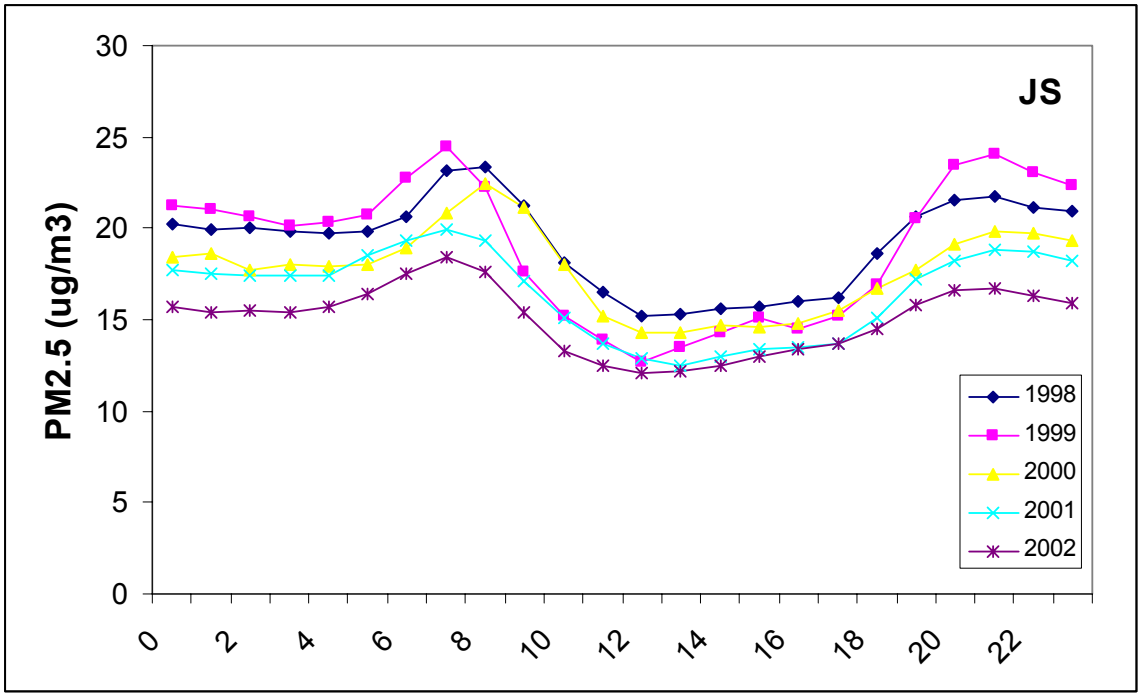


Figure 4.20: Diurnal profiles of $PM_{2.5}$.

4.6 Conclusions

There are several factors that affect temporal patterns in air pollutants: variations in emissions, meteorology, and photochemistry. Each pollutant has a unique temporal profile, and it is important to consider this when performing data analysis on air pollutants. Diurnal profiles are the most distinct; that is, all sites have similar diurnal profiles for the same pollutant. Weekly and seasonal profiles are also evident for most pollutants. In the sites considered in this research, temporal profiles were similar at like sites, i.e. urban sites shared similar temporal profiles. This can give an indication of source categories that are affecting monitoring sites.

Chapter 5

INSTRUMENT ERROR

5.1 Introduction

In a health effects study, one measure of a given pollutant may be used in an epidemiologic model to analyze the health outcome. The measurement given by the sampler is assumed to be the actual concentration of the pollutant and this concentration is often assumed to be constant throughout the study area. One criticism of this type of study is that when the health outcome response is very small, exposure error can have a large impact. Exposure error is the difference in the pollutant concentration measured and used in a model and the actual pollutant concentration the population is exposed to. The effect of exposure error is a bias to the null, or a lack of power in the model to determine the true relationship (Lipfert and Wyzga, 1997). There are many sources of exposure error, including instrument error, spatial variability, temporal variability, and variability in human behavior (Carrothers and Evans, 2000). Instrument error and spatial variability are collectively referred to here as measurement error. This section of research addresses instrument error.

This research uses data from 13 different monitoring stations in the Atlanta area. Several different networks are used; however the majority of the data is from federally mandated ambient air monitoring stations. The data collected by these stations is mandated by the Clean Air Act (CAA) and the Code of Federal Regulations, and has

multiple uses: to establish attainment/nonattainment designations for the National Ambient Air Quality Standards (NAAQS), to evaluate the effectiveness of control programs, to assess the impact of air pollution on public health, to track the progress of State Implementation Plans (SIPs), to aid in the development of control strategies, to verify emissions inventories and dispersion modeling, and to monitor developing temporal trends (Sleva et al.). Determining the impact of air pollution on public health is only one objective, and is often not the primary objective of the agency managing air pollution monitoring networks. This implies that air sampling monitors should not be assumed to be located so as to minimize exposure error. The quantification of exposure error, and therefore measurement error, is important in understanding the total uncertainty in a health effects study.

This research focuses on two ways to assess instrument error: accuracy and precision quality assurance data from the Environmental Protection Agency (EPA) and collocated sampler data.

5.2 Accuracy and Precision Data

Any environmental program that is run by or on behalf of the EPA is required to have a quality assurance program (Sleva et al.). Specifically, several aspects of air monitoring are established in the Code of Federal Regulations (CFR), part 40, including network design, siting criteria, methodology and data validation, and an adequate quality assurance program. The quality assurance program is outlined through data quality objectives (DQOs) that establish objectives for precision, accuracy, representativeness,

comparability, and comprehensiveness (Sleva, et al.). DQOs are often used by decision makers as the level of uncertainty acceptable in decision making and have been developed for the determination of NAAQS compliance. Each state is required to develop a Quality Assurance Program Plan (QAPP) which is then certified by the EPA.

5.2.1 Automated Samplers

All of the gaseous pollutant data used in this research are from automated samplers. These are samplers that run continuously and collect and analyze samples without user interference. Accuracy and precision data are collected for these samplers through multipoint audits and calibrations.

Multipoint audits are required once per year for automated samplers to test accuracy. During the audit, standards of three known concentrations are injected into the sampler during normal sampling operation. The standards must be traceable to the National Institute of Standards and Technology (NIST), a NIST-traceable Reference Material (NTRM) or a NIST-certified Gas Manufacturer's Internal Standards (GMIS). The three concentrations are usually a factor of ten different and are intended to test the entire range of the instrument (see Table 5.1). The output is compared to the known value. Errors of greater than +/- 20% require the instrument to be re-audited after adjustments have been made.

Table 5.1: Ranges of standards used for accuracy audits, in ppm.

	SO_2 and O_3	NO_2 and NO_x	CO
1	0.03-0.08	0.03-0.08	3-8
2	0.15-0.20	0.15-0.20	15-20
3	0.35-0.45	0.35-0.45	35-45
4	0.80-0.90	------	80-90

To test precision in automated samplers, calibration data are used. Calibrations are required biweekly. During the calibration, standards of a known value within the range normally monitored by the sampler are injected into the sampler. For SO_2, NO_2/NO_x, and O_3 the range is from 0.08 – 0.1 ppm and for CO the range is 8-10 ppm. Similar to the accuracy audits, standards must be traceable to the NIST, NTRM or GMIS. The output is then compared to the known value. If the error exceeds +/- 15%, the instrument is re-calibrated after adjustments have been made.

5.2.2 Remote Analysis Samplers

Remote analysis samplers are samplers that collect an air sample at one point, for a given amount of time, which is then transferred to a lab for analysis. Particulate matter is sampled in this way. Because there is human handling of the sample and transfer between locations, instrument error is expected to be higher for samples collected in this manner. Accuracy and precision data for these samplers are collected through flow check audits and collocated sampler data.

Flow check audits must be performed on $PM_{2.5}$ samplers once per year. A certified flow rate transfer standard is compared to the instrument's output. The recorded value must be +/- 10% of the standard value or the sampler should be re-audited.

Precision of $PM_{2.5}$ samplers is verified through the use of collocated samplers. Within a $PM_{2.5}$ monitoring network, 15% of samplers must have a collocated sample. Special consideration is given to areas that are likely to exceed the National Ambient Air Quality Standard, such as Atlanta. In this case, 80% of the collocated samplers must be in areas where the concentration has exceeded 90% of the NAAQS 24-hour average standard. In each area that has experienced a violation, there should be at least one collocated sampler. The collocated samplers should be between two and four meters apart and must use the same collection method. Calibration, sampling and analysis of collocated samplers should occur at the same time. One sampler must be designated the "primary" sampler for the site. Collocated data are reviewed annually to determine any systematic problems. In the Atlanta area, Doraville Health Center, E Rivers School, and Forest Park monitoring stations have collocated samplers. Limited data were available for this study period.

5.2.3 EPD/SEARCH Collocated Samplers

Both the Georgia Environmental Protection Division (EPD) and the Southeastern Research and Characterization Study (SEARCH) maintain monitoring stations at the Yorkville site. The SEARCH station includes all of the gaseous and PM measurements considered in this research. The EPD station includes ozone, NO_2/NO_x, and particulate matter. The monitoring methods are identical in both networks for these pollutants. The

stations are located approximately 4 meters apart. This is sufficient to consider the continuous gaseous pollutants monitored by each network at Yorkville to be collocated.

5.3 Results

Instrument error can be caused by several factors, including statistical sampling error (only discrete portions of air are monitored and they cannot overlap for different monitors), monitor interferences, and reproducibility of methods (APCA, 1995). Instrument error can be thought of as the fact that two identical monitors in the same place may not always report the same concentration of a given pollutant. The clearest way to assess instrument error for the purposes of this research is to use collocated data. When that is not possible, audit and calibration data are used.

5.3.1 Audit and Calibration Data for Continuous Monitoring Samplers

Audit and calibration data was used to calculate percent error, which represents sampler accuracy, for CO, SO_2, O_3 and NO_x. Audit and calibration data was collected from EPD for all sites used in this research for the time frame of interest (1993-2003). The percent error was calculated as the absolute value of the difference between the observed value and the standard value divided by the standard value. This value ranged between 2 and 5 percent with standard deviations between 2 and 4 percent. (See Table 5.2)

Each set of audit and calibration data was composed of data from three distinct concentration ranges, as described above. (See Figure 5.1) To verify that the percent

error is similar throughout the whole range of the sampler, each concentration range was analyzed individually. Percent errors were similar at all concentration levels. (See Table 5.3)

Table 5.2: Percent error using audit data

Pollutant	% Error	St Dev	R	Regression
CO	3.88%	3.06%	0.99	Obs = - 0.448 + 1.04 Std
NO_x	4.47%	4.03%	0.98	Obs = 0.00119 + 0.980 Std
O_3	2.15%	2.65%	1.0	Obs = 0.00261 + 0.973 Std
SO_2	3.11%	2.76%	0.99	Obs = - 0.00315 + 1.02 Std

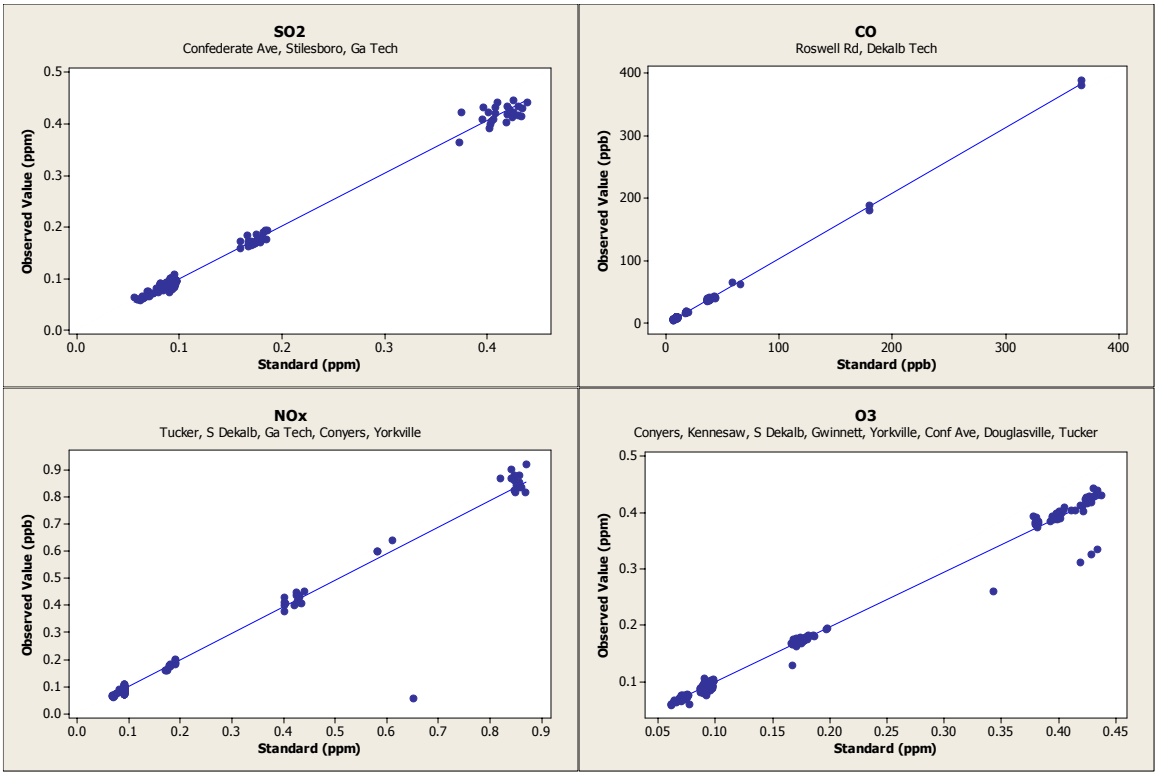


Figure 5.1: Scatterplots of audit/calibration data.

Table 5.3: Percent errors for different concentration ranges.

Range	CO	SO_2	O_3	NO_x
1	3.93%	3.10%	2.23%	4.65%
2	3.23%	3.17%	1.88%	3.52%
3	3.28%	3.46%	2.83%	3.49%

5.3.2 Collocated Data

For ozone, NO_2/NO_x, and $PM_{2.5}$ data from collocated samplers were used to assess instrument precision. The correlation coefficient between collocated samplers was calculated for each pollutant, as well as the percent error and standard deviation of the percent error, as above. The R-values were high for all pollutants (above 0.9) (see Figure 5.2); however, for pollutants with both audit and calibration data and collocated data, the percent error was much higher for the collocated data than for the audit and calibration data. (See Table 5.3 for values.) This suggests that the accuracy of the instrument, as measured by the audit and calibration data, is only a small part of the total instrument error, as measured by the collocated samplers.

Table 5.4: Percent error using collocated monitors.

Pollutant	Site	Years	% Error	St Dev	R	Regression
$PM_{2.5}$	Jefferson St (FRM and PCM)	98-02	12.13%	11.66%	0.96	PCM = - 0.397 + 0.954 FRM
	Jefferson St (FRM and TEOM)	98-02	20.12%	19.92%	0.91	TEOM = - 1.20 + 0.969 FRM
PM_{10}	FS8, DHC, ERS	95-03	5.83%	5.49%	0.98	Prim = 1.73 + 0.915 Col
NO_x	Yorkville	02	26.38%	26.42%	0.95	SEARCH = 0.002 + 0.949 AQS
O_3	Yorkville	99-02	8.30%	9.46%	0.98	SEARCH = 4.62 + 992 AQS

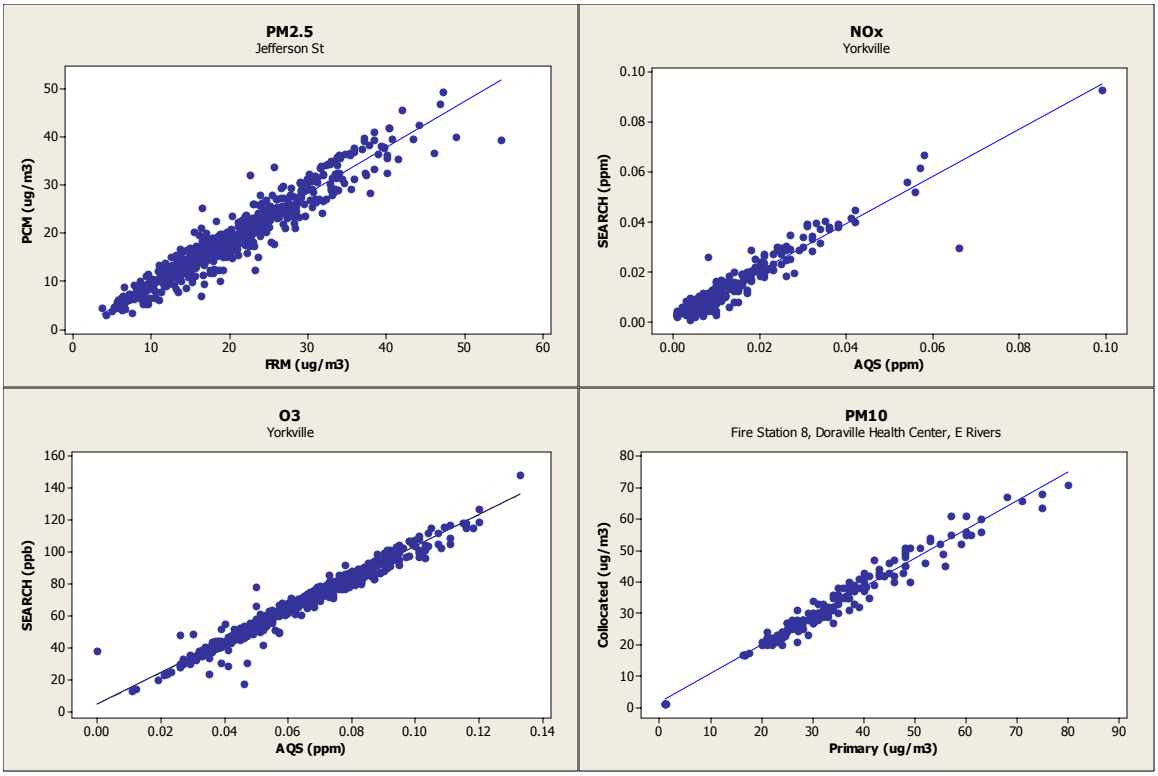


Figure 5.2: Scatterplots of collocated data.

5.3.3 Using Audit and Calibration Data and Collocated Data to Assess Instrument Error

Whenever possible, collocated instrument data was used to assess instrument. This was done for $PM_{2.5}$, PM_{10}, O_3, and NO_x. For SO_2 and CO, audit and calibration data was used. These two different types of data give different information about instrument error. The error derived from collocated data is assumed to be a more accurate description of instrument error. Because the collocated samplers are running simultaneously and very close together, any discrepancy between measurements can be described as errors due to instrument functioning and not due to local source impacts. Audit and calibration data gives information about the accuracy of an instrument, which is only one component of instrument error. Comparing percent error calculated from audit and calibration data with that calculated from collocated data for PM, O_3, and NO_x shows that using audit and calibration data gives a smaller percent error. This should be taken into consideration when working with audit and calibration data.

In this research, γ is a semivariogram function developed to describe the fraction of the temporal variation in a pollutant that is actually due to the spatial variation of that pollutant. The function is defined as:

$$\gamma(h) = c_o + c_e \{1 - exp(-h/a_e)\} \qquad \textbf{Eq 4}$$

Where h is the distance between two sites, c_o is the instrument error, c_e is the partial "sill" of the function and a_e is the "range" of the function divided by 3. (See section 7 for a detailed discussion of this function.) Air quality data fits this curve when the R-value between two stations is converted to γ as follows:

$$\gamma' = \sqrt{\frac{1-R}{1+R}} \qquad \textbf{Eq 2}$$

When h is zero, i.e. there is no distance between measurements, the only term in the function is c_o. To empirically fit c_o PM, O_3, and NO_x, the R-value between collocated samplers is used in equation 2. For SO_2 and CO, the percent error calculated from audit and calibration data is multiplied by the average value of the pollutant concentration during the study period and divided by the standard deviation of the concentration. These values are reported in Table 5.5.

Table 5.5: Instrument error.

Pollutant	Instrument error
SO_2	0.14
CO	0.07
NO_x	0.24
O_3	0.11
$PM_{2.5}$	0.10
PM_{10}	0.11

5.4 Conclusions

Measurement error, an important component of exposure error in an epidemiologic study, includes instrument error and spatial variability. While spatial variability is often assumed to be much larger than instrument error, it is still important to quantify instrument error to understand overall uncertainty in a study using ambient air pollutant data. Two methods of quantifying instrument error were performed in this analysis: using audit and calibration data to calculate the percent variation of a sampler and using collocated data to calculate the correlation between samplers at the same site. Both calculations give an indication of how much of the total variability observed in air pollutants is attributable to variations in the samplers used. As expected, error between collocated samplers was much larger than the error derived using audit and calibration data, and it is expected to be a more accurate measurement of instrument error.

Chapter 6

SPATIAL VARIATION

6.1 Introduction

Spatial variation of air pollutants is an important part of measurement error. In an epidemiologic study, measurement error is the difference in what an air pollution sampler measures and what a population is actually exposed to. Spatial variation refers to the fact that air pollution monitors are only measuring concentrations at one, or a few, discrete point(s) in space. Often it is assumed that this concentration is the same throughout the study area or that there is a constant concentration gradient throughout the study area. Neither of these assumptions is true, however. The distribution of local sources, for primary pollutants, can greatly increase the spatial variability of a given pollutant. At the same time, the lack of sources of secondary pollutants and the reliance on atmospheric chemistry for their formation can lead to a decrease in spatial variability of a given pollutant. Every pollutant should therefore be considered to have a unique pattern of spatial variability. In order to assess the overall uncertainty of an epidemiologic health – affects study, it is important to quantify spatial variability of the pollutants used across the study area.

6.1.1 Quantifying Spatial Variability of Air Pollutants

There are several common methods of examining variability of air pollutants across a study area. The simplest and most common is through the use of the Pearson correlation coefficient (Monn et al., 1997; Buzorious et al., 1999; Morowska et al., 2002; Lin et al., 2001; Rao et al., 1995). This statistic is a measure of the covariance between two variables. In an air monitoring network, the correlation coefficient would be calculated between two samplers at different locations monitoring the same pollutant simultaneously. A correlation of 1 is assumed to imply less spatial variability of that pollutant.

Other methods of evaluating the spatial variability of air pollutants include the coefficient of variation (Monn et al., 1997; Hoek et al., 2002; Lebret et al., 2000; Roosli et al., 2001), which is the standard deviation between sites divided by the mean, and simple comparisons of absolute concentrations of a pollutant at different sites (Chow et al., 2001; Pinto et al., 2004; Burton et al., 1988). While these methods all give a general indication of agreement between sites, they do not give any information as to how the agreement changes with distance.

This research utilizes a variation on a common method of describing spatial data in geostatistics, the semivariogram (Waller and Gotway, 2004). The semivariogram addresses the uniformity of data over a study area. It is based on the concept of spatial autocorrelation; that is, observations made closer together will be more similar than those made further away. The semivariogram is a function of the spatial distance between two observations, but it is not dependent on the actual locations at which the observations were made. The semivariogram function is often plotted against distance between

observations in order to give information about the continuity and variation of observations and has several common attributes.

6.2 Methods

6.2.1 Semivariogram derivation

Specialized semivariograms were developed to measure the amount of temporal variability that is explained by spatial variability. The derivation of the semivariogram formula used in this research follows.

$$\mu_x \equiv \frac{\sum x}{n} \qquad \mu_y \equiv \frac{\sum y}{n}$$

$$sd_x \equiv \sqrt{\frac{n\sum x^2 - \left(\sum x\right)^2}{n(n-1)}} \qquad sd_y \equiv \sqrt{\frac{n\sum y^2 - \left(\sum y\right)^2}{n(n-1)}}$$

$$R \equiv \frac{\sum (x-\mu_x)(y-\mu_y)}{n(sd_x sd_y)}$$

Where x and y are logarithms of the daily (i.e.24 hour average) observations at two stations during the study period. The actual observations are assumed to have lognormal distributions. Probability density function plots of the data support this assumption, as will be shown.

Values x and y are then normalized:

$$x' \equiv \frac{x-\mu_x}{sd_x} \qquad y' \equiv \frac{y-\mu_y}{sd_y}$$

The mean of the normalized values as well as the sum of the normalized values are both 0. The standard deviation of the normalized values is 1. The R value can be rewritten using the normalized values:

$$R = \frac{\sum (x'y')}{n}$$

A new term, the daily spatial average value, is defined:

$$z' \equiv \frac{x'+y'}{2}$$

note:

$$x', y' = z' \pm \frac{x'-y'}{2} = z' \pm \Delta z'$$

Therefore, the spatial variation in z' is $\Delta z' = \frac{x'-y'}{2}$.

The standard deviation of the spatial variation in z' is then calculated.

$$sd_{\Delta z'} = sd\left(\frac{x'-y'}{2}\right) = \frac{sd(x'-y')}{2} = \frac{1}{2}\sqrt{\frac{n\sum (x'-y')^2 - \left(\sum (x'-y')\right)^2}{n(n-1)}}$$

This can be rearranged and simplified as follows.

$$sd_{\Delta z'} = \frac{1}{2}\sqrt{\frac{n\sum x'^2 + n\sum y'^2 - 2n\sum (x'y') - \left(\sum x'\right)^2 - \left(\sum y'\right)^2 + 2\sum x'\sum y'}{n(n-1)}}$$

$$sd_{\Delta x'} = \frac{1}{2}\sqrt{\frac{n\sum x'^2 - \left(\sum x'\right)^2}{n(n-1)} + \frac{n\sum y'^2 - \left(\sum y'\right)^2}{n(n-1)} + \frac{2\sum x'\sum y'}{n(n-1)} - \frac{2n\sum (x'y')}{n(n-1)}}$$

$$sd_{\Delta x'} = \frac{1}{2}\sqrt{sd_{x'} + sd_{y'} + \frac{2\sum x' \sum y'}{n(n-1)} - \frac{2n^2 R}{n(n-1)}}$$

$$sd_{\Delta z'} = \frac{1}{2}\sqrt{1+1+0-\frac{2nR}{n-1}}$$

If n is very large, $\frac{n}{n-1} \approx 1$, therefore

$$sd_{\Delta z'} = \frac{1}{2}\sqrt{2-2R} = \sqrt{\frac{1-R}{2}} \qquad \textbf{Eq 1}$$

Next, the ratio of the standard deviation in the spatial variation ($sd_{\Delta z'}$) to the standard deviation in the value ($sd_{z'}$) is calculated.

$$\mu_{z'} = 0$$
$$sd_{z'} = \sqrt{\frac{1+R}{2}} > 1 \qquad \textbf{Eq 2}$$

The derivation of $\sqrt{\frac{1+R}{2}}$ is similar to the above derivation of $\sqrt{\frac{1-R}{2}}$ with the substitution of $\frac{x'+y'}{2}$ for z' (versus $\frac{x'-y'}{2}$).

This gives the following equation:

$$\frac{sd_{\Delta z'}}{sd_{z'}} = \sqrt{\frac{1-R}{1+R}} \qquad \textbf{Eq 3}$$

In this equation, $sd_{\Delta z'}$ is referred to as the spatial standard deviation and $sd_{z'}$ is the temporal standard deviation.

It can also be shown that the ratio of the spatial to temporal standard deviation is the same for z as it is for z'.

6.2.2 Application to Atlanta Air Quality Data

Air quality data was collected from the Georgia Environmental Protection Division's (EPD) ambient monitoring network, the Southeastern Research and Characterization Study (SEARCH), and the ASACA network. The time frame used was from 1998-2002 in order to include as many sites as possible (SEARCH data is available only during those years). The distribution of the data was examined and it was determined that all pollutants have a lognormal distribution (see Table 6.1, Table 6.2 for distribution parameters and Figure 6.1 for distribution fit). The lognormal of all of the data was calculated and used for the semivariogram analysis. The intercept, or "nugget effect", was calculated using the instrument error. For pollutants for which collocated data was available, PM, NO_x, and O_3, the R-value of the collocated samplers was calculated and used in equation 3 to determine the intercept. For pollutants without any collocated samplers, audit and calibration data was used. The average percent difference between the standard concentrations used in the audits and calibrations and the measured concentration was calculated. This value was then multiplied by the average concentration of the pollutant for the given time frame and divided by the standard deviation to give the intercept.

Once the intercept is established for a pollutant, equation 3 is then calculated for each pair of sites, i.e. for each distance. The line is then fitted to the following model:

$$\gamma(h) = c_o + c_e \{1 - \exp(-h/a_e)\} \qquad \textbf{Eq 4}$$

Here, c_o is the calculated nugget effect, c_e is the partial sill ($c_o + c_e$ is the sill), and the effective range (typically defined as the distance at which the autocorrelation is 0.05) is $3a_e$ (Waller and Gotway, 2004).

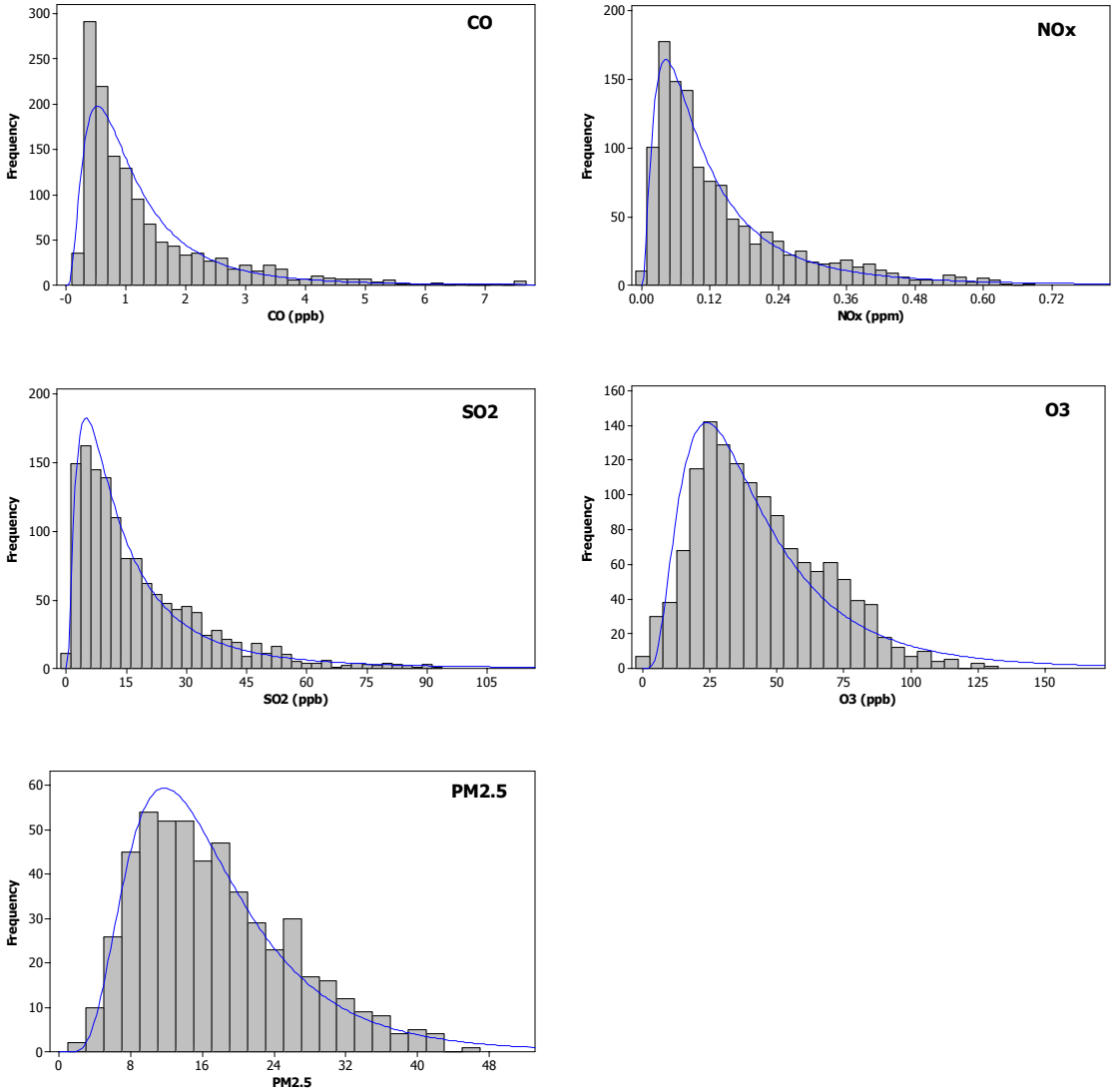


Figure 6.1: Distribution fitting for Jefferson Street data. The blue line is a standard lognormal distribution given the mean and standard deviation of the data.

Table 6.1: Lognormal distribution parameters, gaseous pollutants and $PM_{2.5}$ mass.

pollutant measure	station	median value	geometric standard deviation
SO_2	JS – SEARCH	12.84 / 2.5 ppb	2.61 / 2.47
1-hr max /	GT	8 / 2.75	3.0 / 2.13
24-hr avg	CA	7 / 2.19	3.0 / 2.06
	St	6 / 1.73	3.11 / 1.89
	Yo	6.42 / 2.05	2.42 / 2.32
CO	JS – SEARCH	0.87 / 0.39 ppm	2.17 / 1.77
1-hr max /	RR	1.5 / 0.74	1.71 / 1.62
24-hr avg	DT	1.2 / 0.55	1.92 / 1.80
	Yo	0.24 / 0.18	1.41 / 1.32
NO_x	JS – SEARCH	94.7 / 31.6 ppb	2.48 / 2.36
1-hr max /	GT	86 / 30.2	2.50 / 2.34
24-hr avg	SD	136 / 38.5	2.84 / 2.94
	Tu	51 / 19.8	2.17 / 2.04
	Co	29 / 10.8	3.70 / 3.41
	Yo	9 / 3.9	2.57 / 2.29
O_3	JS – SEARCH	48.3 / 39.8 / 22.2 ppb	1.8 / 1.9 / 2.0
1-hr max /	CA	60.0 / 51.6 / 29.5	1.6 / 1.6 / 1.6
8-hr max /	SD	57.0 / 48.3 / 24.7	1.6 / 1.7 / 1.6
24-hr avg,	Co	57.0 / 49.9 / 28.0	1.5 / 1.5 / 1.5
Apr-Oct	Yo	53.2 / 48.4 / 38.5	1.5 / 1.5 / 1.5
$PM_{2.5}$ mass	JS – SEARCH	16.0 $\mu g/m^3$	1.6
24-hr avg	ERS	16.5	1.6
	SD	16.0	1.7
	DHC	16.8	1.6
	FS8 – 3rd day	18.1	1.7
	EPHC – 3rd day	16.7	1.8
	FP – 3rd day	16.1	1.7
	Ke – 3rd day	16.5	1.7
	Yo – 3rd day	13.5	1.8
	JS – TEOM	17.9	1.6
	FM – TEOM	17.5	1.6
	SD – TEOM	16.6	1.5
	Tu – TEOM	18.1	1.7
	Yo – SEARCH	11.9	1.7

Table 6.2: Lognormal distribution parameters, $PM_{2.5}$ components.

$PM_{2.5}$ component	station	median	geometric standard deviation
sulfate	JS – SEARCH	4.6 $\mu g/m^3$	1.9
	FM – ASACA	4.3	2.1
	SD – ASACA	4.4	2.0
	Tu – ASACA	4.3	2.1
	Yo – SEARCH	4.5	2.1
nitrate	JS – SEARCH	0.8 $\mu g/m^3$	2.1
	FM – ASACA	0.7	2.3
	SD – ASACA	0.3	2.4
	Tu – ASACA	0.8	2.6
	Yo – SEARCH	0.6	2.0
ammonium	JS – SEARCH	1.8 $\mu g/m^3$	1.9
	FM – ASACA	1.6	2.1
	SD – ASACA	1.5	2.2
	Tu – ASACA	1.7	2.1
	Yo – SEARCH	2.4	1.9
elemental carbon	JS – SEARCH	1.4 $\mu g/m^3$	1.9
	FM – ASACA	1.2	1.8
	SD – ASACA	1.5	1.9
	Tu – ASACA	1.2	1.8
	Yo – SEARCH	0.7	1.7
organic carbon	JS – SEARCH	4.0 $\mu g/m^3$	1.7
	FM – ASACA	3.8	2.6
	SD – ASACA	4.2	2.4
	Tu – ASACA	3.8	2.5
	Yo – SEARCH	3.2	1.7

Table 6.3: Parameter estimates for semivariogram functions of gaseous pollutants.

pollutant	nugget	effective range	sill
SO_2	0.14	45	0.84
CO	0.07	60	0.77
NO_x	0.24	105	0.84
O3	0.11	120	0.31
$PM_{2.5}$-mass	0.10	30	0.45

Table 6.4: Parameter estimates for semivariogram functions of $PM_{2.5}$ components.

$PM_{2.5}$ component	nugget	effective range	sill
SO_4^{2-}	0.07	45	0.37
NO_3^-	0.19	36	0.49
NH_4^+	0.24	75	0.49
EC	0.34	45	0.84
OC	0.38	30	0.63

6.3 Results

6.3.1 Semivariograms

Semivariograms were developed for CO (1 hour maximum), SO_2 (1 hour maximum), NO_x (1 hour maximum), ozone (8 hour maximum) and $PM_{2.5}$ (24 hour average) (See Figure 6.2 and Tables 6.5-6.9). These sampling intervals were chosen based on current EPA regulations. The model given in equation 4 was then fitted to the points. The nugget, or c_0 was calculated based on an assessment of the instrument error. The sill and range were fitted to the data by hand. See Table 6.3 for fitted parameter values. The sills of the primary pollutants, CO, SO_2, and NO_x, have sills ranging from 0.77 to 0.84. The sills of the secondary pollutants, ozone and $PM_{2.5,}$ are much lower, 0.31 and 0.45 respectively. The higher sills are indicative of greater spatial variability of the primary pollutants, which was expected. At large distances, 40 to 80 km, the spatial variation and instrument error of the primary pollutants are around 80% of the temporal variation. For secondary pollutants, the spatial variation and instrument error are around 40% of the temporal variation at similar distances. This is expected, as concentrations of ozone and $PM_{2.5}$ are influenced by meteorology and atmospheric chemistry as opposed to local sources.

Semivariograms were also developed for $PM_{2.5}$ components (see Figure 6.3). Similar patterns were observed in primary versus secondary components. Fitted parameter values are reported in Table 6.4. Elemental carbon (EC) is the only purely primary component and it has a much higher sill than the other components, which are secondary. This is consistent with the findings above.

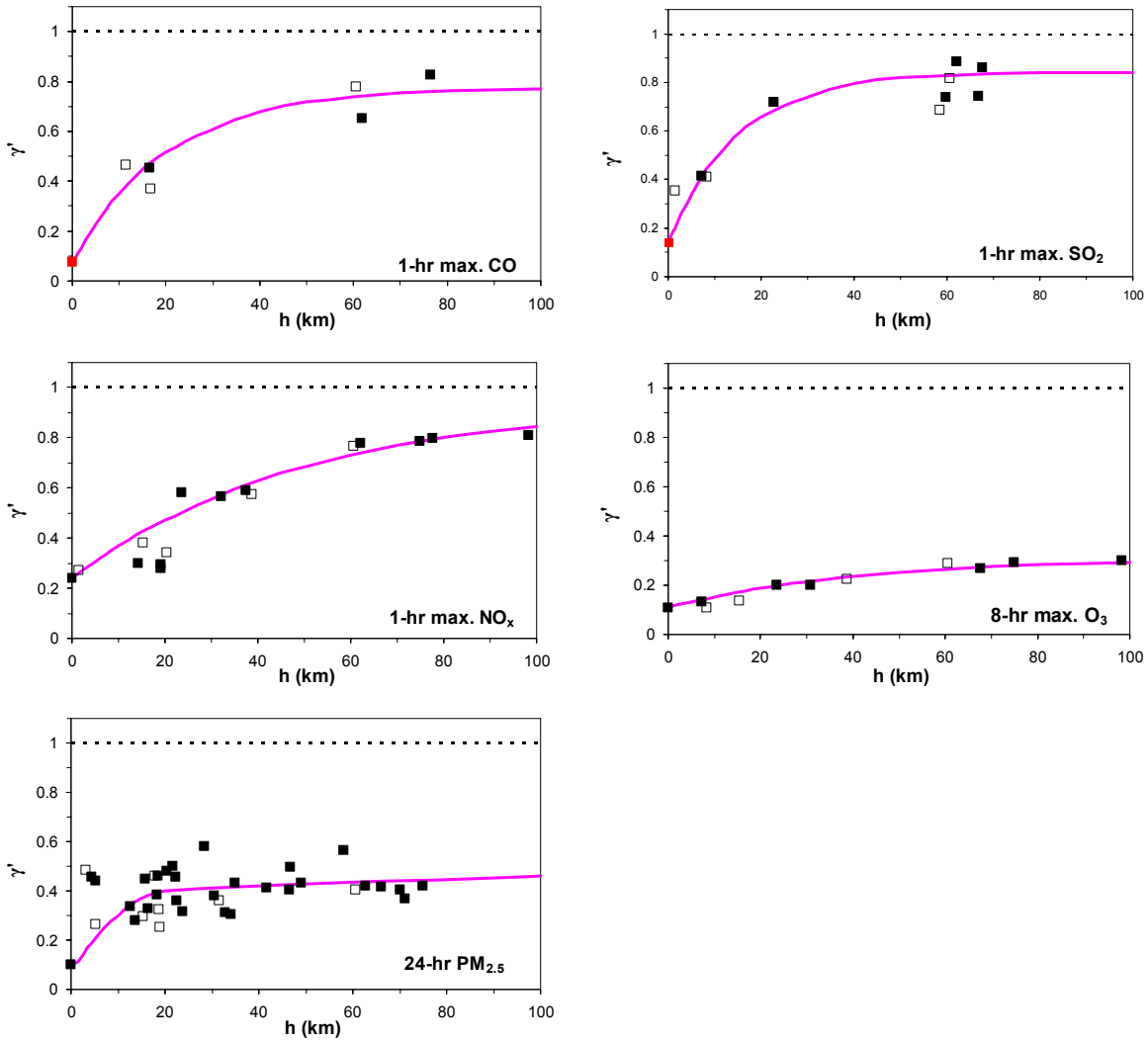


Figure 6.2: Normalized semivariograms for ambient air pollutant measures in Atlanta, 1999-2002. Hollow boxes represent JS data. Red boxes represents audit data.

Table 6.5: Semivariogram fit for CO.

		D (km)	R	γ'
DT	RR	16.497	0.661412	0.451437
	Yo	76.568	0.187724	0.826978
RR	Yo	61.888	0.40506	0.650712
DT	JS	16.841	0.760948	0.368445
RR	JS	11.504	0.646342	0.46348
Yo	JS	60.575	0.244126	0.779458

Table 6.6: Semivariogram fit for NO_x.

		D (km)	R	γ'
GT	SD	14.154	0.835454	0.299414
	Co	37.436	0.482177	0.591072
	Tu	19.085	0.8563	0.27823
	Yo	62.051	0.24828	0.776019
SD	Co	23.588	0.497105	0.579579
	Tu	19.092	0.838879	0.296005
	Yo	74.857	0.235861	0.786323
Co	Tu	32.031	0.515874	0.565129
	Yo	98.283	0.209572	0.80838
Tu	Yo	77.578	0.222265	0.797688
GT	JS	1.504	0.86239	0.271825
SD	JS	15.356	0.744851	0.3824
Co	JS	38.754	0.503241	0.574856
Tu	JS	20.407	0.792831	0.339932
Yo	JS	60.575	0.262126	0.76461

Table 6.7: Semivariogram fit for SO_2.

		D (km)	R-1 hr	γ'
CA	GT	7.299	0.708195	0.413312
	St	66.783	0.286747	0.744518
	Yo	67.722	0.149892	0.859822
GT	St	59.864	0.292499	0.739858
	Yo	62.05116	0.12187	0.884725
St	Yo	22.79262	0.31688	0.720237
CA	JS	8.31	0.712181	0.410001
GT	JS	1.504	0.779276	0.352211
St	JS	58.598	0.358802	0.686939
Yo	JS	60.575	0.198799	0.817519

Table 6.8: Semivariogram fit for O_3.

		D (km)	R(1hr)	γ'
Yo	Yo	0		0.10941
CA	SD	7.244	0.96850675	0.131743
	Co	30.83	0.91122688	0.199435
	Yo	67.722	0.85231151	0.266742
SD	Co	23.588	0.91456729	0.201417
	Yo	74.857	0.82901746	0.290417
Co	Yo	98.283	0.80191462	0.300542
CA	JS	8.31	0.97158686	0.107833
SD	JS	15.356	0.95715775	0.13718
Co	JS	38.754	0.88257809	0.22214
Yo	JS	60.575	0.85416343	0.287365

Table 6.9: Semivariogram fit for $PM_{2.5}$.

		D (km)	R	γ'
DHC	ERS	13.748	0.855424	0.279143
	SD	23.857	0.81995	0.314533
	FS8	18.362	0.741054	0.385655
	EPHC	34.938	0.683506	0.433586
	FP	34.059	0.829334	0.305441
	Ke	32.846	0.821537	0.313008
	Yo	71.071	0.763612	0.366109
	JS-FRM	18.922	0.880017	0.252627
	Yo-FRM	71.071	0.840605	0.294278
ERS	SD	16.327	0.803779	0.329823
	FS8	5.165	0.674696	0.440734
	EPHC	22.406	0.65613	0.45567
	FP	22.559	0.771842	0.358844
	Ke	30.495	0.747065	0.380496
	Yo	62.762	0.69879	0.42108
	JS-FRM	5.19	0.869258	0.264468
	Yo-FRM	62.762	0.81384	0.320364
SD	FS8	18.473	0.648874	0.461464
	EPHC	15.746	0.666595	0.447271
	FP	12.699	0.795017	0.337929
	Ke	46.592	0.720088	0.403399
	Yo	74.857	0.698071	0.421672
	JS-FRM	15.356	0.840392	0.294491
	Yo-FRM	74.857	0.800076	0.333263
FS8	EPHC	20.486	0.627379	0.478508
	FP	21.624	0.601903	0.498513
	Ke	28.399	0.495498	0.580816
	Yo	58.184	0.519874	0.562048
	JS-FRM	3.238	0.622693	0.482203
	Yo-FRM	58.184	0.659883	0.452664
EPHC	FP	4.361	0.657872	0.454275
	Ke	46.746	0.60507	0.496036
	Yo	66.041	0.703241	0.417411
	JS-FRM	17.86	0.652214	0.4588
	Yo-FRM	66.041	0.72168	0.402065
FP	Ke	49.036	0.686171	0.431416
	Yo	70.11	0.719642	0.403773
	JS-FRM	18.65	0.809665	0.32431
	Yo-FRM	70.11	0.743872	0.383241
Ke	Yo	41.668	0.708312	0.413214
	JS-FRM	31.636	0.768728	0.361602
	Yo-FRM	41.668	0.79135	0.341286
Yo	JS-FRM	60.575	0.717728	0.405374
	Yo-FRM	0	0.857436	0.277044
JS-FRM	Yo-FRM	60.575	0.851903	0.28279

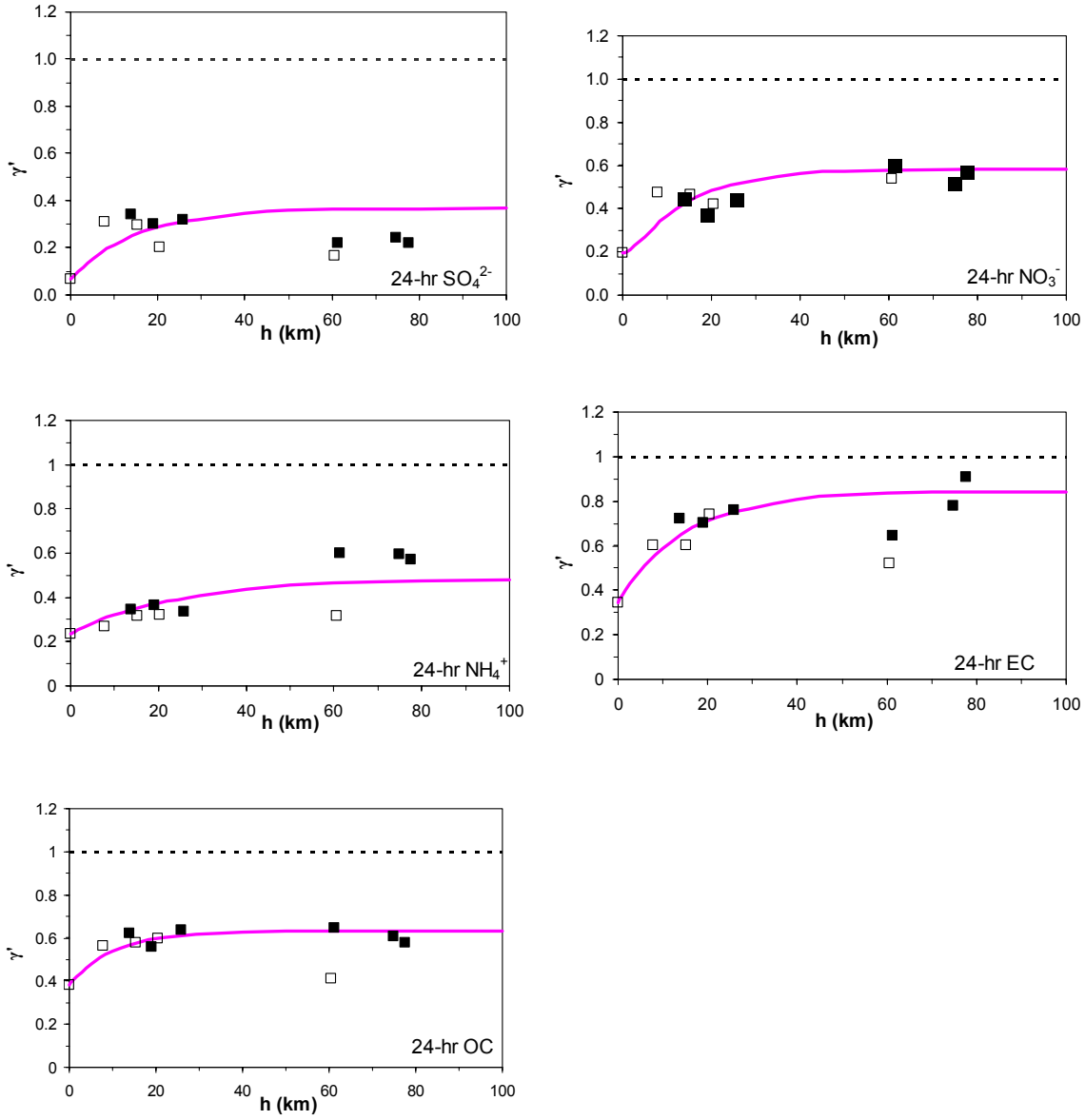


Figure 6.3: Normalized semivariograms for $PM_{2.5}$ components in Atlanta, March 1999 – August 2000. Hollow boxes represent JS data.

6.3.2 Population Weighted Averages

After the semivariogram function was fitted to pollutant data from the Atlanta area, the semivariogram values were weighted by population to determine one value for use over the entire metropolitan Atlanta area. The area was divided into nine approximately circular regions of varying distances from the city center (See Figure 6.4). The population for each of these rings was calculated using 2000 Census Data at a zip-code level. The percent of the total population for the area was calculated for each ring and then multiplied by the corresponding semivariogram value for each pollutant (See Table 6.10). This value is a population-weighted average value of the semivariogram which can be used for the entire study area.

Table 6.10: Population weighted averages; a: SO_2, NO_x, and CO; b: O_3, and $PM_{2.5}$.

Distance	Pop	% of Pop	SO_2	SO_2-weighted	NO_x	NO_x-weighted	CO	CO-weighted
1	21043	0.42%	0.188	0.001	0.258	0.001	0.109	0.000
5	358994	6.75%	0.341	0.023	0.321	0.022	0.229	0.015
10	700485	6.82%	0.483	0.033	0.390	0.027	0.350	0.024
20	1667701	19.31%	0.658	0.127	0.502	0.097	0.517	0.100
30	2864304	23.89%	0.748	0.179	0.587	0.140	0.618	0.148
40	3921781	21.12%	0.794	0.168	0.650	0.137	0.680	0.144
50	4275918	7.07%	0.818	0.058	0.697	0.049	0.717	0.051
60	4720572	8.88%	0.830	0.074	0.733	0.065	0.740	0.066
70	5007896	5.74%	0.836	0.048	0.760	0.044	0.753	0.043
average				**0.686**		**0.559**		**0.575**

Distance	Pop	% of Pop	O_3	O_3-weighted	$PM_{2.5}$	$PM_{2.5}$-weighted
1	21043	0.42%	0.114	0.000	0.109	0.000
5	358994	6.75%	0.133	0.009	0.203	0.014
10	700485	6.82%	0.154	0.010	0.302	0.021
20	1667701	19.31%	0.188	0.036	0.402	0.078
30	2864304	23.89%	0.215	0.051	0.439	0.105
40	3921781	21.12%	0.236	0.050	0.453	0.096
50	4275918	7.07%	0.252	0.018	0.458	0.032
60	4720572	8.88%	0.265	0.024	0.460	0.041
70	5007896	5.74%	0.275	0.016	0.461	0.026
average				**0.205**		**0.398**

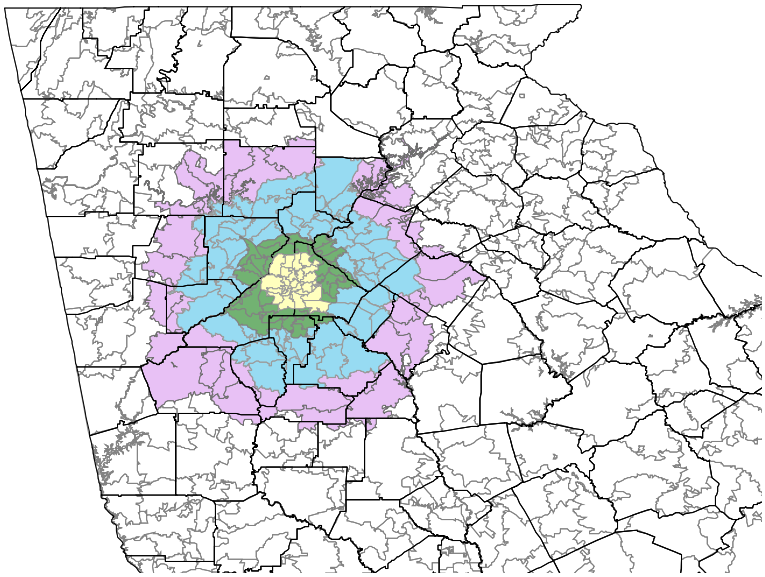


Figure 6.4: Area within 60 km (purple), 40 km (blue), 20 km (green), and 10 km (yellow) of Atlanta's city center.

6.4 Conclusions

The results of the semivariogram analysis show significant spatial variability for primary pollutants and indicate that this spatial variability is a nontrivial component of exposure error. The semivariograms also show noticeable differences between pollutants. In the Atlanta area, approximately 90% of the population lives within a 30km radius of downtown. If one is considering SO_2 in a health effects study in Atlanta, it must be noted that at 30km, almost 80% of the temporal variation (which is what is used to establish a relationship between pollutant and health effect) is actually spatial variation. Using only one monitoring site for the study area would not be appropriate for SO_2. However, for ozone, only 30% of the temporal variation at 30km is actually spatial variation and one monitoring site may be sufficient for the needs of the study.

For points that include Jefferson Street data, the boxes are hollow to distinguish them. This is because Jefferson Street data is often chosen for epidemiologic studies due to the central location of the site and the wide array of measurements there. The semivariograms show that Jefferson Street fits the semivariogram model as well as any other station. The lack of outliers indicates no one station can be termed as being more representative than another station.

These results suggest that, particularly for primary pollutants, data from multiple sites should be used in epidemiologic health effects studies to minimize measurement error. Interpolation methods using population weighting over an entire study area would likely give a more accurate description of the population's exposure to a given pollutant.

Chapter 7

LOCAL SOURCE IMPACTS

7.1 Introduction

In an air quality monitoring network, it is important to examine individual sites for impacts from local sources. If data from a site is being used to estimate population exposure to the pollutant, monitoring sites should be situated so as to minimize the impact from local sources, which could bias the exposure estimate and lead to error in a health effects study.

The main source categories of the pollutants used in this research are well known for the Atlanta area. CO, NO_x and SO_2 are all primary pollutants, meaning they are directly emitted. In the Atlanta area, mobile sources are the largest contributor to CO and NO_x concentrations in ambient air. There are several coal fired power plants in the Atlanta area and these are believed to be the most significant sources of SO_2. Secondary pollutants, such as ozone and particulate matter, are formed in the atmosphere through chemical reactions and therefore do not have any sources. In this research, one would not expect to see any "local sources" of secondary pollutants.

The relative location of these sources to monitoring sites should be observed for descriptive purposes. One method of qualitatively examining local sources of air pollution is to develop plots of pollution concentrations versus wind direction. If large concentrations are indicated from one direction, that area should be inspected for sources of the pollutant. For example, a large SO_2 spike on a pollution rose plot in a

given direction could confirm that SO_2 is coming from a power plant in that direction. Looking at an individual site's plot will give information on local sources while comparing plots for the same pollutant across monitoring sites can give information about regional pollution sources.

7.2 Emissions Inventories

As part of the US Environmental Protection Agency's National Emissions Inventory (NEI), the state of Georgia Environmental Protection Division (EPD) compiles an emissions inventory for all criteria pollutant and volatile organic compound (VOC) point source emissions throughout the state on a regular basis. The 1999 Georgia emissions inventory was used for this research. Point sources within the Atlanta region were identified and are summarized in Table 7.1 and mobile sources in Table 7.2. Highway vehicle emissions are responsible for over 68% of CO emissions and 53% of NO_x emissions every year, but only about 2% of PM_{10} and SO_2 and 4% of $PM_{2.5}$. Off-highway mobile emissions are responsible for an additional 20% of CO and 13% of NOx, as well as 1% of PM_{10}, 2.5% of $PM_{2.5}$ and 1.44% of SO_2.

Power plants in the Atlanta region, including Plants Bowen, Wansley, Yates and McDonough, have the largest emissions of all the pollutants considered here. Transcontinental Gas Pipeline also has large emissions of CO and NO_x, however this facility is to the south of the city, not close to any CO monitors and over 10 miles south of the nearest NO_x monitor (South Dekalb). Lafarge Building Materials (formerly Blue Circle Cement) also has large emissions of NO_x, as well as SO_2. This facility is directly

northwest of the Jefferson Street monitoring station and could have an impact on measurements at several urban monitoring sites.

Table 7.1: Point Source emissions inventory for the Atlanta area, 1999, TPY (GA EPD, 2000).

Facility Name	Map	Longitude	Latitude	CO	NO_x	PM_{10}	PM_{25}	SO_2
BJ Sanitary Landfill & Recyclihg Center		-84.236	33.932	81.900	31.900			
Boral Bricks, Inc; Atlanta Plant	A	-84.501	33.821	207.400	59.000	55.573		191.820
Caraustar Mill Group, Inc.	B	-84.649	33.818	16.426	364.107	53.443	30.853	711.628
Cellofoam North America, Inc.		-84.083	33.783		1.450			
CITGO Doraville Terminal		-84.275	33.910		4.122			
ConocoPhillips Co. - Doraville Products Terminal		-84.273	33.915	5.081	2.028			0.005
Dart Container Corporation of Georgia		-84.118	33.724	10.990	23.840			
Delta Air Lines, Inc. Atlanta Station		-84.442	33.644	0.395	0.762			
Delta Air Lines, Inc. General Office Facilities		-84.431	33.653	3.289	6.536			
Delta Air Lines, Inc. TOC	C	-84.418	33.644	46.259	101.598			
Earthgrains Baking Companies, Inc. - Atlanta Bakery		-84.275	33.692	3.450	4.352	0.530	0.078	0.027
Emory University		-84.324	33.795	34.370	35.050	3.180	0.000	0.840
Engineered Fabrics Corporation		-85.037	34.002	2.730	3.874	0.185	0.062	0.042
Ford Motor Company Atlanta Assembly		-84.402	33.656		17.851			
General Motors Assembly Plant		-84.281	33.906	26.870	35.500			
GEORGIA INSTITUTE OF TECHNOLOGY		-84.395	33.773	18.686	22.078	1.245	0.418	0.155
Georgia Power Company, Bowen Steam-Electric Generating Plant	1	-84.923	34.125	2125.241	36897.367	5214.511	2256.840	160534.656
Georgia Power Company, McDonough Steam-Electric Generating Plant	2	-84.475	33.825	321.983	4886.804	332.497	145.907	27993.310
Georgia Power Company, Wansley Steam-Electric Generating Plant	3	-85.038	33.407	1750.524	20492.708	1135.163	523.715	73551.250
Georgia Power Company, Yates Steam-Electric Generating Plant	4	-84.905	33.463	566.086	8987.153	647.035	283.360	41546.200
Lafarge Building Materials	D	-84.470	33.823	36.220	1252.890	32.620	27.480	1617.000
Live Oak Landfill & Recycling Center	E	-84.339	33.671	185.200	34.000			
Lockheed Martin Aeronautics Company - Marietta		-84.529	33.926	38.960	53.180	3.500	3.260	24.910
Magellan Terminals Holdings, L.P.		-84.253	33.919	9.321	3.728			
Marathon Ashland Petroleum LLC		-84.418	33.781	8.610	3.450			

METAL COATERS OF GEORGIA		-84.538	33.979	2.128	4.820	0.144	0.144	0.015
Owens Corning - Atlanta		-84.546	33.763	41.597	3.838	69.296	44.986	67.583
Owens Corning - Fairburn Plant		-84.617	33.539	27.630	602.060	2.510		0.190
Owens-Brockway Glass Container Inc. - Atlanta, GA plant	F	-84.421	33.671	45.512	710.451	239.449	239.449	284.821
Pan Glo Atlanta/Div. Of Russell T. Bundy Assoc. Inc.		-84.388	33.664		0.484			
PPG Architectural Finishes, Inc.		-84.434	33.689	1.200	1.400	0.109	0.109	0.001
Printpack Inc.		-84.960	33.739	2.250	2.990	0.210		0.020
Rexam Beverage Can Company		-84.390	33.630	4.138	4.971	0.195		0.028
RM Clayton WRC	G	-84.456	33.822	196.450	41.817			4.130
Southwire Company		-85.071	33.565	220.947	20.985	38.157	18.518	5.091
Stevens Graphics		-84.411	33.738	0.510	0.600	0.050		0.010
The Sherwin-Williams Company		-84.339	33.569	1.498	1.846			0.011
Transcontinental Gas Pipe Line - Station 120	H	-84.255	33.569	700.980	2347.440			
Tuscarora Incorporated		-84.048	33.675	2.088	1.813	0.095	0.051	0.579
Weyerhaeuser Impak Center		-84.180	33.858	0.762	0.907			0.005
WinCup - Stone Mountain		-84.194	33.828	0.198	34.057			2.922
Woodbridge Corporation		-84.119	33.738	1.430	1.720			0.008
Young Refining Corp.		-84.731	33.767	16.221	21.340	1.052	0.351	12.217

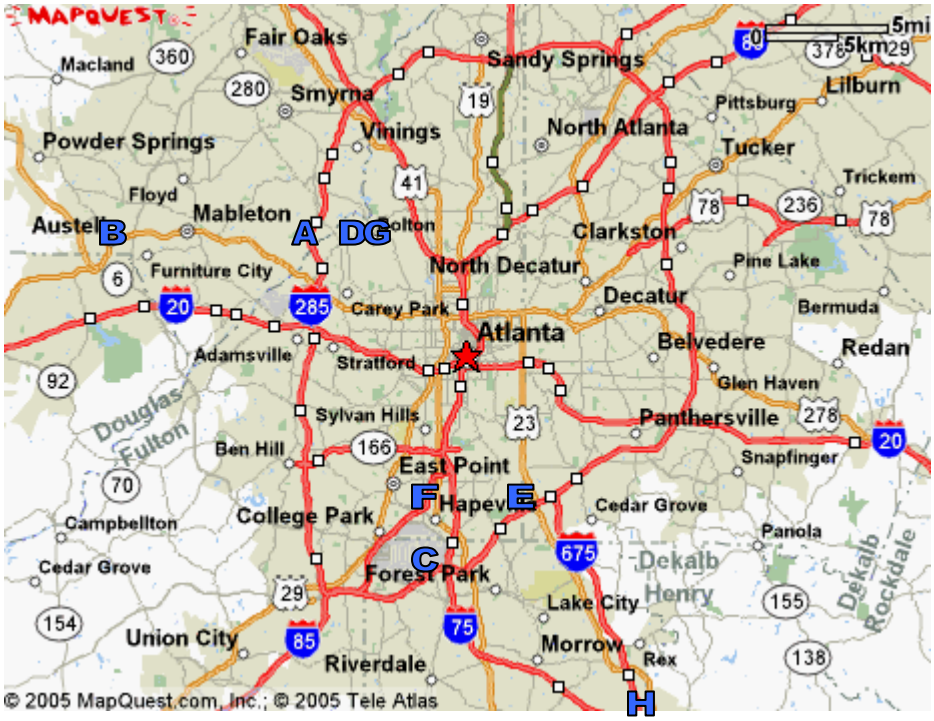


Figure 7.1: Selected point sources.

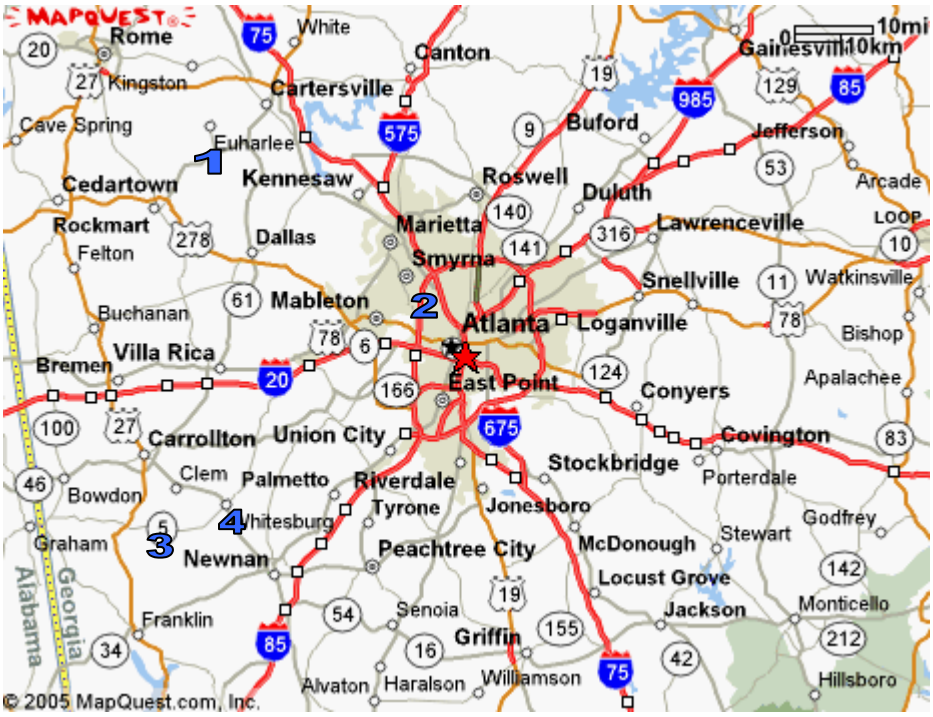


Figure 7.2: Power plants in the Atlanta area.

Table 7.2: Mobile emissions inventory for the 20-county Atlanta MSA, by County, 1999, TPY (US EPA, 2000).

COUNTY		CO	NO_x	PM_{10}	$PM_{2.5}$	SO_2
Barrow	Highway Vehicles	13285.14	1755.52	53.49	41.77	66.02
	Off-Highway	1694.97	334.35	18.29	16.08	23.5
Bartow	Highway Vehicles	29792.82	4102.58	123.24	96.98	149.71
	Off-Highway	4490.5	1196.31	58.28	52.92	80.88
Carroll	Highway Vehicles	32344.49	4302.14	129.11	101.04	159.33
	Off-Highway	5240.2	593.8	36.91	33.77	43.95
Cherokee	Highway Vehicles	40970.41	5352.13	158.12	123.24	198.35
	Off-Highway	15342.93	605.35	74.83	68.77	61.8
Clayton	Highway Vehicles	80427.72	8273.56	225.93	167.32	328.05
	Off-Highway	16350.75	7200.45	92	84.5	114.86
Cobb	Highway Vehicles	174327	18732.51	543.06	402.13	789.42
	Off-Highway	67162.23	4146.36	416.57	382.44	458.46
Coweta	Highway Vehicles	34956.3	4815.13	145.22	114.3	176.38
	Off-Highway	5514.12	722.56	58.76	53.65	79.46
DeKalb	Highway Vehicles	216140.2	23026.69	664.74	491.11	972.3
	Off-Highway	66115.8	2848.2	297.26	272.16	282.6
Douglas	Highway Vehicles	30835.27	3553.47	101.74	77.51	136.83
	Off-Highway	3071.74	313.84	25.35	23.24	35.12
Fayette	Highway Vehicles	27294.16	3439.31	104.32	81.05	131.81
	Off-Highway	7563.19	545.34	49.76	45.38	54
Forsyth	Highway Vehicles	26266.41	3873.36	125.21	99.72	145.9
	Off-Highway	13330.42	628.53	85.1	78.34	70.69
Fulton	Highway Vehicles	260908.5	28205.96	819.65	608.15	1186.34
	Off-Highway	76688.11	8247.2	694.15	631.33	1166.24
Gwinnett	Highway Vehicles	151188.9	17374.14	514.73	387.61	715.46
	Off-Highway	76010.62	4016.82	401.04	368.43	397.98
Henry	Highway Vehicles	38576.04	5378.27	164.81	129.9	198.49
	Off-Highway	6697.31	820.15	82.09	75.09	109.37
Newton	Highway Vehicles	23830.08	3277.02	98.7	77.62	119.86
	Off-Highway	3481.14	430.01	37.6	34.41	43.67
Paulding	Highway Vehicles	21630.31	3189.76	103.05	82.07	120.16
	Off-Highway	2599.96	315.77	33.83	30.89	51.59
Pickens	Highway Vehicles	4809.33	715.33	22.73	18.08	26.5
	Off-Highway	804.4	77.13	7.1	6.5	7.98
Rockdale	Highway Vehicles	21060.62	2683.43	81.88	63.61	103.07
	Off-Highway	5206.41	481.55	38.89	35.74	46.5
Spalding	Highway Vehicles	18455.19	2325.1	69.81	54.18	88.32
	Off-Highway	3145.34	322.29	25.26	23.05	28.37
Walton	Highway Vehicles	15466.59	1960.26	58.73	45.68	74.38
	Off-Highway	3018.86	307.76	24.49	22.5	28.33

7.3 Methods

7.3.1 Development of Pollution Roses

To assess the impact of local sources on individual monitoring sites, pollution roses were developed using hourly pollutant concentrations and simultaneous wind direction data for each site and each pollutant. The observations were divided into 30 different wind direction bins of 12 degrees each. The average concentration for each bin was then plotted versus the wind direction on a 360° scatter plot.

Because of seasonal and diurnal patterns in both pollutant concentrations and wind direction, the pollution roses were corrected for season and for time of day. To perform the time of day correction each wind direction bin was further sorted into one of twelve time of day bins, each composed of two hours. A new variable, TOD, was calculated as follows.

$$TOD_j = \sum_{i=1}^{12} \left(B_{ij} A_i \right)$$

Where i is the time of day bin from 1-12 and j is the wind direction bin from 1-30. A and B are further defined as:

$$A_i = \frac{\sum_{j=1}^{30} \left(\bar{x}_{ij} n_{ij} \right)}{\sum_{j=1}^{30} n_{ij}} \qquad B_{ij} = \frac{n_{ij}}{\sum_{i=1}^{12} n_{ij}}$$

Where $\bar{x}_{ij}$ is the average of the ith time of day bin of the jth wind direction bin and n_{ij} is the number of observations in the ith time of day bin of the jth wind direction bin. A_i is the expected concentration for each time of day bin over all wind directions and B_{ij} is the ratio of the number of observations for each time of day bin in each wind direction

bin to the total number of observations for each wind direction bin. Therefore, TOD_j is the expected concentration for each wind direction bin based on the expected concentrations for each time of day bin. The error attributed to the time of day is then calculated.

$$TOD_{err,j} = TOD_j - \frac{\sum_{j=1}^{30} TOD_j}{30}$$

This is the TOD metric for each wind direction bin minus the average TOD metric. This is then subtracted from the original average value for each time of day bin and wind direction bin in order to correct for time of day.

The same procedure is used to find the seasonal correction. In that case, each wind direction bin is divided into twelve seasonal bins, one for each month. The correction factor is derived the same way as TOD_{err} and is then also subtracted from the original average value for each time of day bin within each wind direction bin to correct for season.

7.3.2 Application to Atlanta Air Quality Data

The methods detailed above were applied to hourly air quality data from the Atlanta area. Data was collected from EPA's Air Quality System (AQS) and the Southeastern Research and Characterization study (SEARCH). Pollution roses were developed using the non-normalized concentrations of each pollutant and the wind direction measured simultaneously. Whenever possible, wind direction measurements from the same site as the pollutant measurements were used. For three sites, this was not possible and the most representative measurements available were used. These sites

were Stilesboro, where Tucker data was used, Georgia Tech, where Jefferson Street data was used, and Roswell Road, where Confederate Avenue data was used. Corrected plots were also calculated.

To further examine the possible confounding of diurnal patterns of pollutant concentrations and wind direction on the pollution roses, each dataset was examined by time of day. The data for each graph was divided into four hour bins to examine possible changes in apparent local sources with time of day.

7.4 Results

7.4.1 Pollution rose plots

Pollution rose plots for CO, SO_2, NO_x, and O_3 are shown in figures 7.3-7.6 (black line). Plots for black carbon and $PM_{2.5}$ mass at Jefferson Street are shown in figure 7.7 (black line). Elevated concentrations in a particular direction indicate a potential local source in that direction. For some pollutants, the wind rose plots allowed regional pollutants to be identified as well.

CO, NO_x and black carbon exhibit peaks for all sites in the direction of major roadways. The largest peak of these pollutants at Jefferson Street , to the northeast, aligns with a major interstate (85). The Georgia Tech NO_x pollution rose shows a similar alignment. Smaller peaks at Jefferson Street and Georgia Tech correspond to roadways to the south and west, a trucking facility and major railyard to the north, and a bus maintenance facility to the south. The South Dekalb NO_x pollution rose is heavily

influenced by mobile emissions from interstate 285 directly to the north. CO at Dekalb Tech shows a large peak to the north, which is consistent with the fact that the monitoring site is located at the southern end of a large parking lot. The Roswell Road site does not show distinct peaks as the site is surrounded by large interstates. Point sources for CO and NO_x listed in Table 7.1 are not observed in the wind rose plots as they are most likely overwhelmed by mobile source emissions.

Sulfur dioxide pollution rose plots show large peaks in the direction of Plant Bowen, the largest power plant in Georgia, for all monitoring sites. Peaks are also seen in the directions of Plants Wansley, Yates, and McDonough as well as for the cement facility.

Ozone pollution rose plots are relatively circular for most monitoring sites. Exceptions are Conyers and South Dekalb, which have depressed concentrations when wind direction is from the north. This could be due to the elevated NO_x levels inhibiting ozone formation.

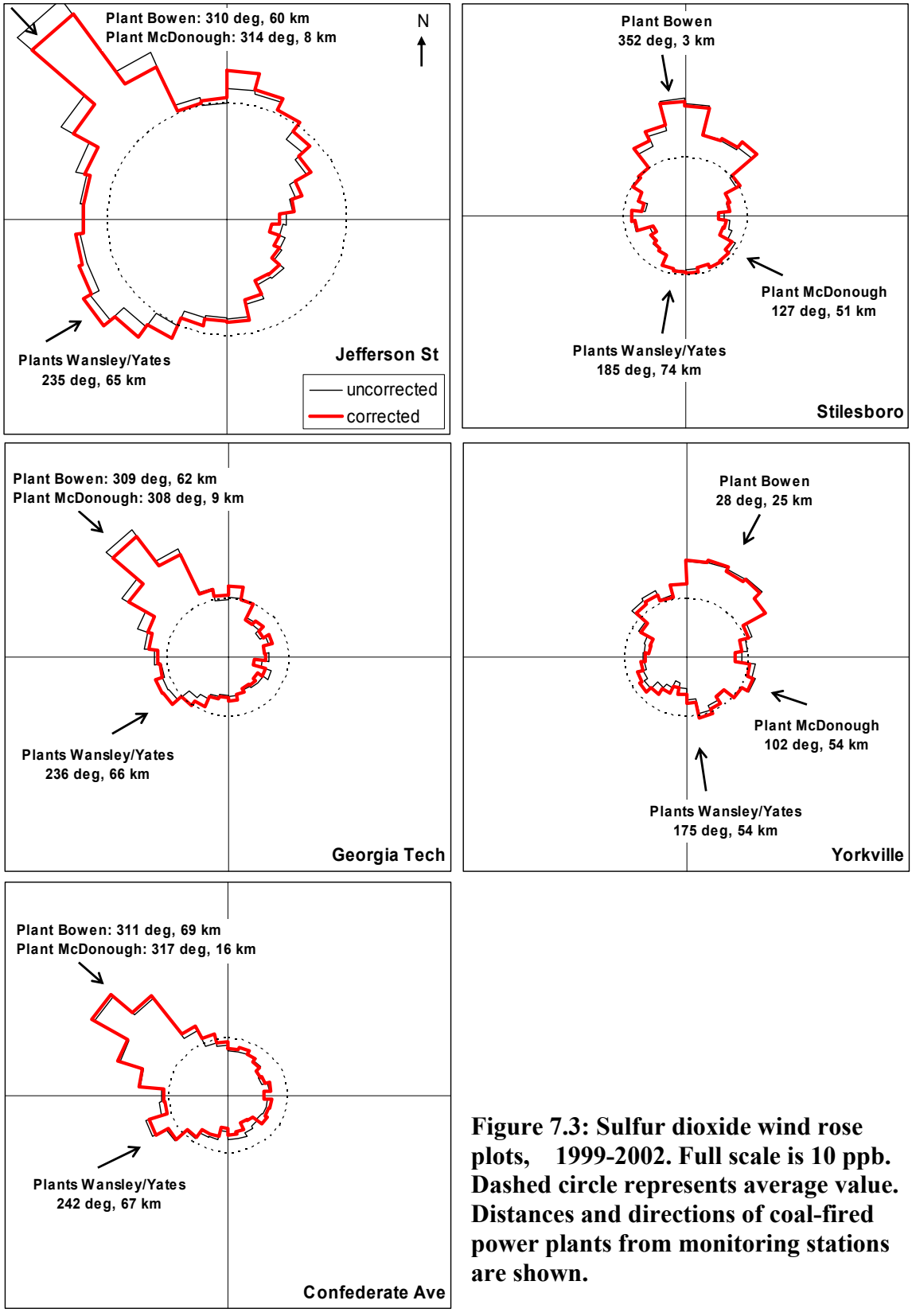


Figure 7.3: Sulfur dioxide wind rose plots, 1999-2002. Full scale is 10 ppb. Dashed circle represents average value. Distances and directions of coal-fired power plants from monitoring stations are shown.

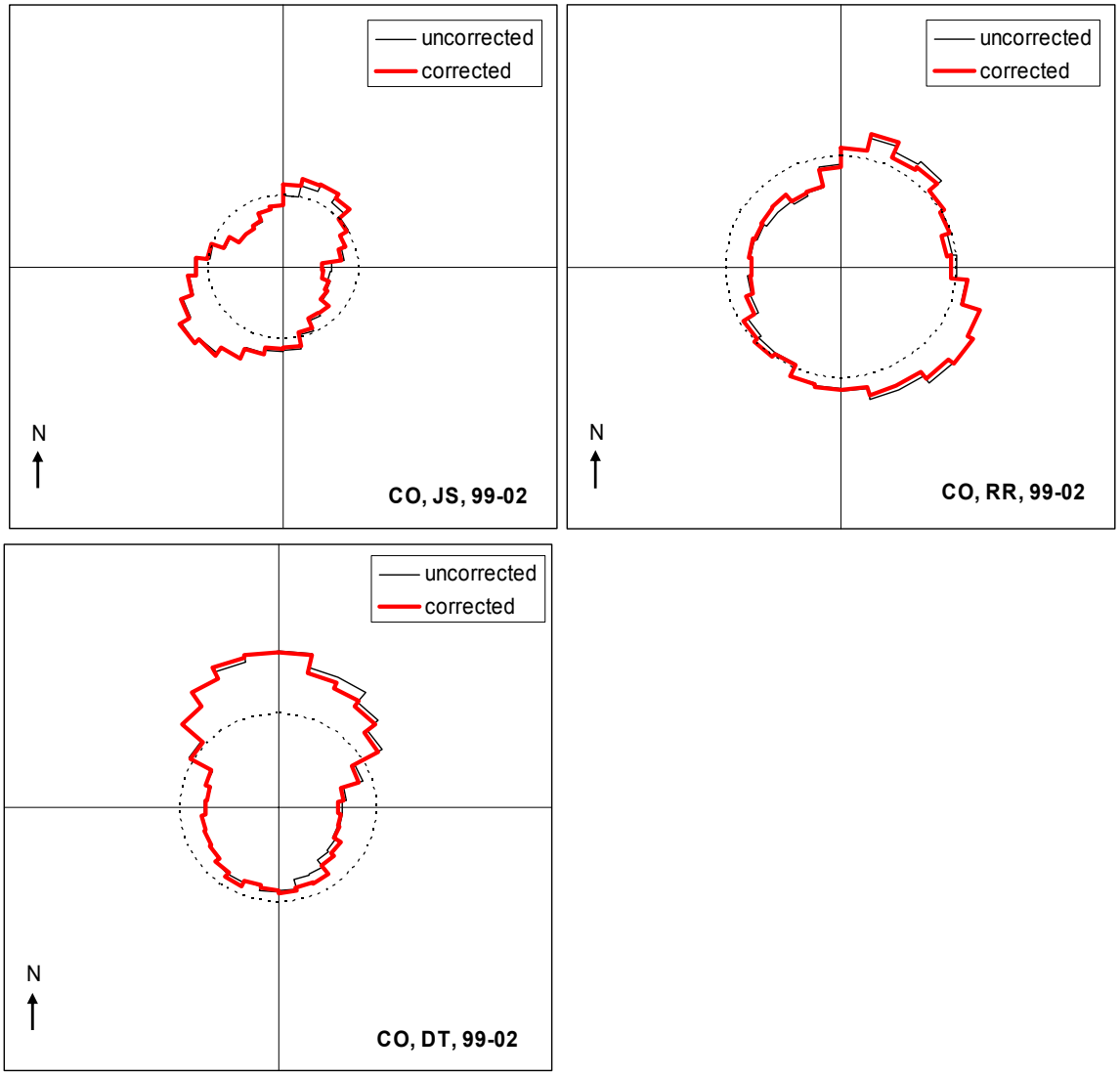


Figure 7.4: CO wind rose plots for 1999-2002. Dashed line is average value. Full scale is 1.5 ppm.

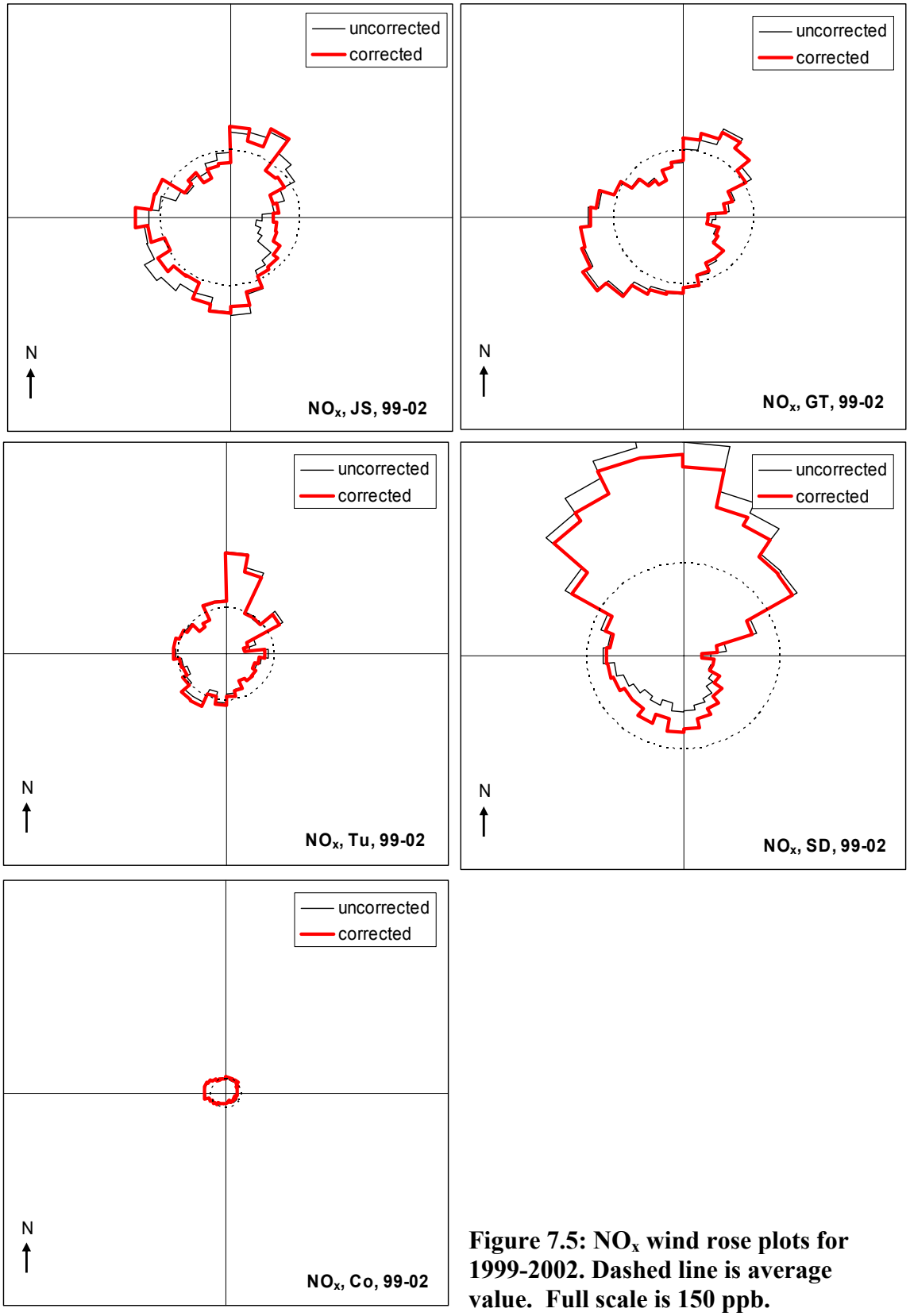


Figure 7.5: NO_x wind rose plots for 1999-2002. Dashed line is average value. Full scale is 150 ppb.

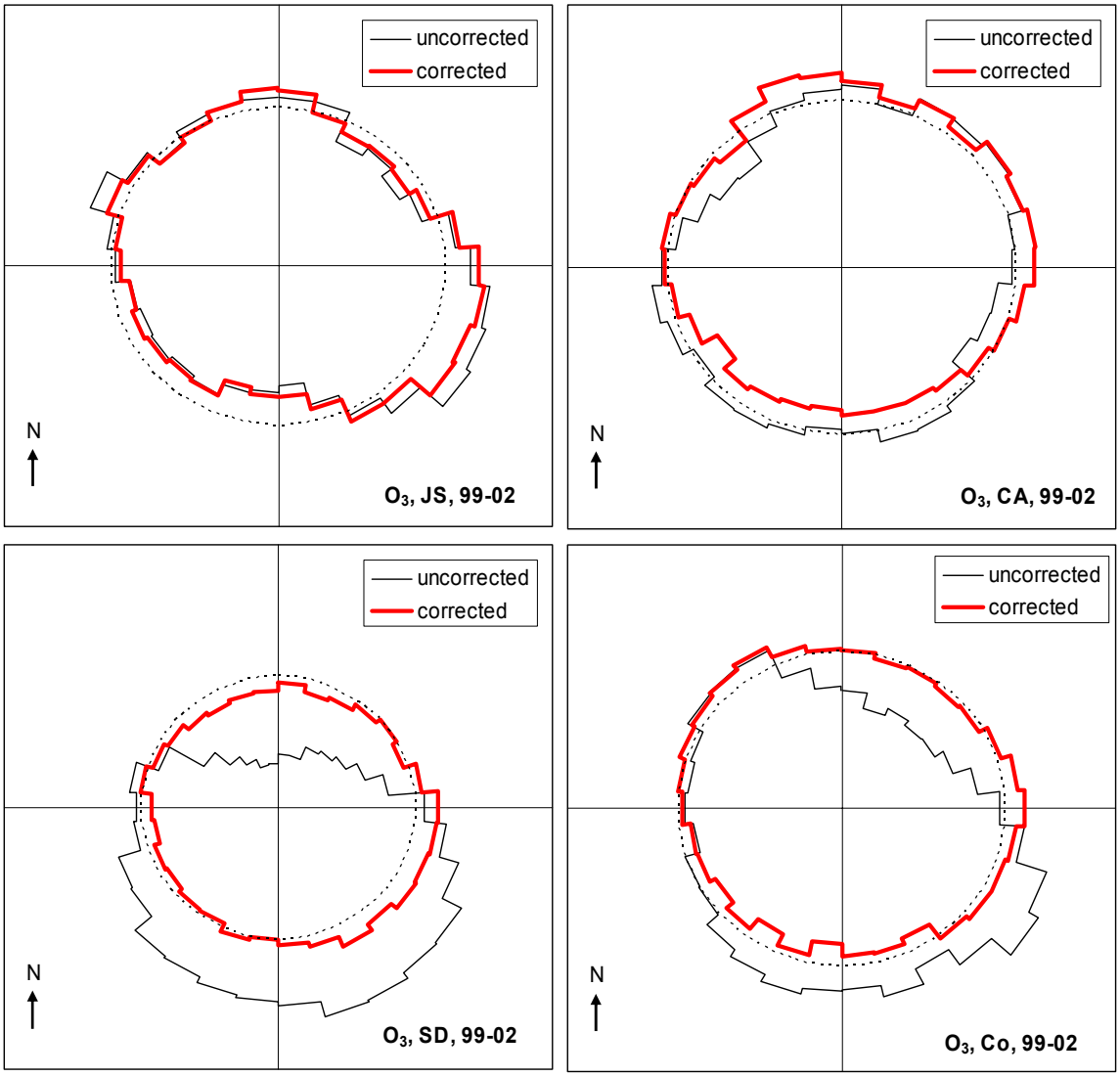


Figure 7.6: Ozone wind rose plots for April through October, 1999-2002. Dashed line is average value. Full scale is 50 ppb.

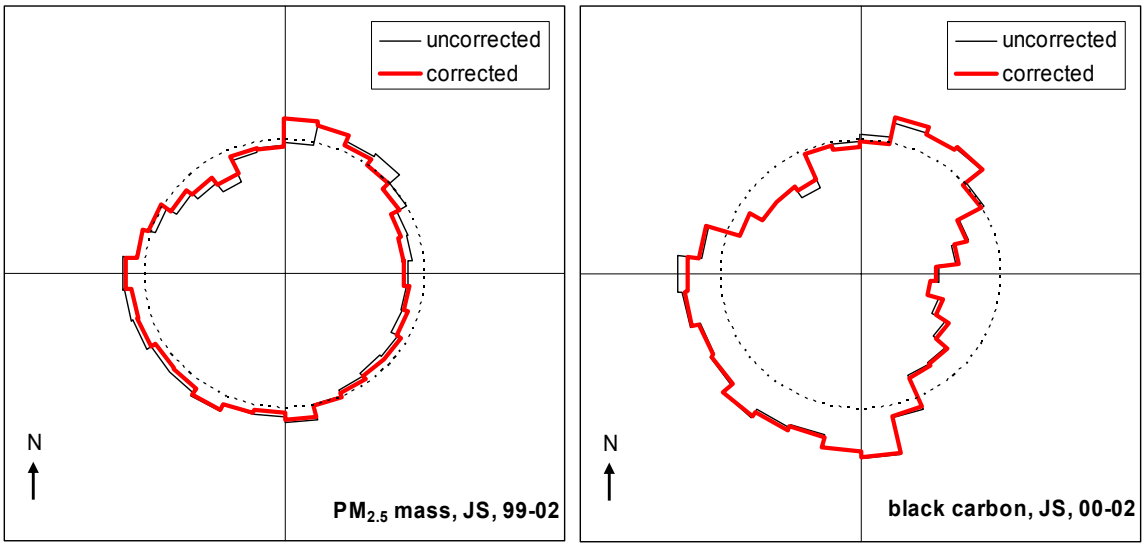


Figure 7.7: Pollution rose plots for $PM_{2.5}$ mass (99-02) and black carbon (00-02) at Jefferson Street.

7.4.2 Time of day and seasonal corrections

Pollution rose plots were also examined by time of day. (See Figures 7.8 – 7.11.) Most pollutants had the same pattern at varying concentrations throughout the day, as expected. South Dekalb and Conyers NO_x pollution rose plots do show an additional source of NO_x during the evening hours, around the evening rush hour that does not appear on the average plots.

Wind direction was then examined by time of day. (See Figures 7.12 – 7.17.) Any correlation between wind direction and time of day, if that time of day had peak pollutant concentrations, could lead to false "sources" appearing on the wind rose plots. For example, at South Dekalb and Conyers, the percentage of wind coming from the south is much higher during the daytime hours (8 am to 3 pm). Ozone peaks between 1 pm and 3 pm, which could lead to an apparent "source" of ozone from the south.

All pollution rose plots were therefore corrected for time of day and seasonal affects. In most cases the corrections were very small, indicating the sources observed are true sources and not due to time of day or seasonal variations in wind direction. However, the corrections did eliminate the apparent shift in ozone concentrations at South Dekalb and Conyers. The corrected wind rose plots for all ozone sites are much more uniform, as expected (See Figures 7.3-7.7, red line).

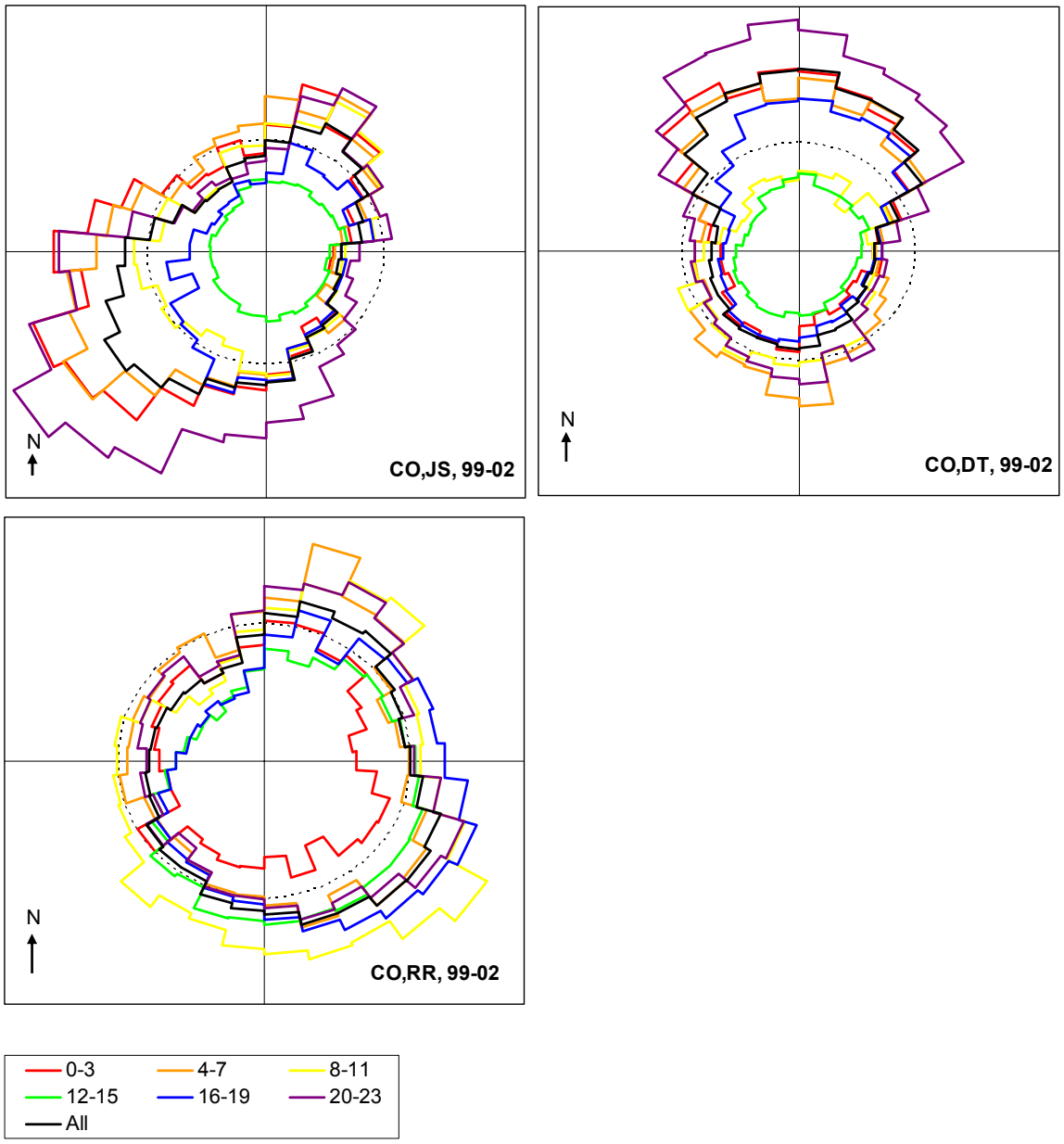


Figure 7.8: CO pollution rose plots by time of day.

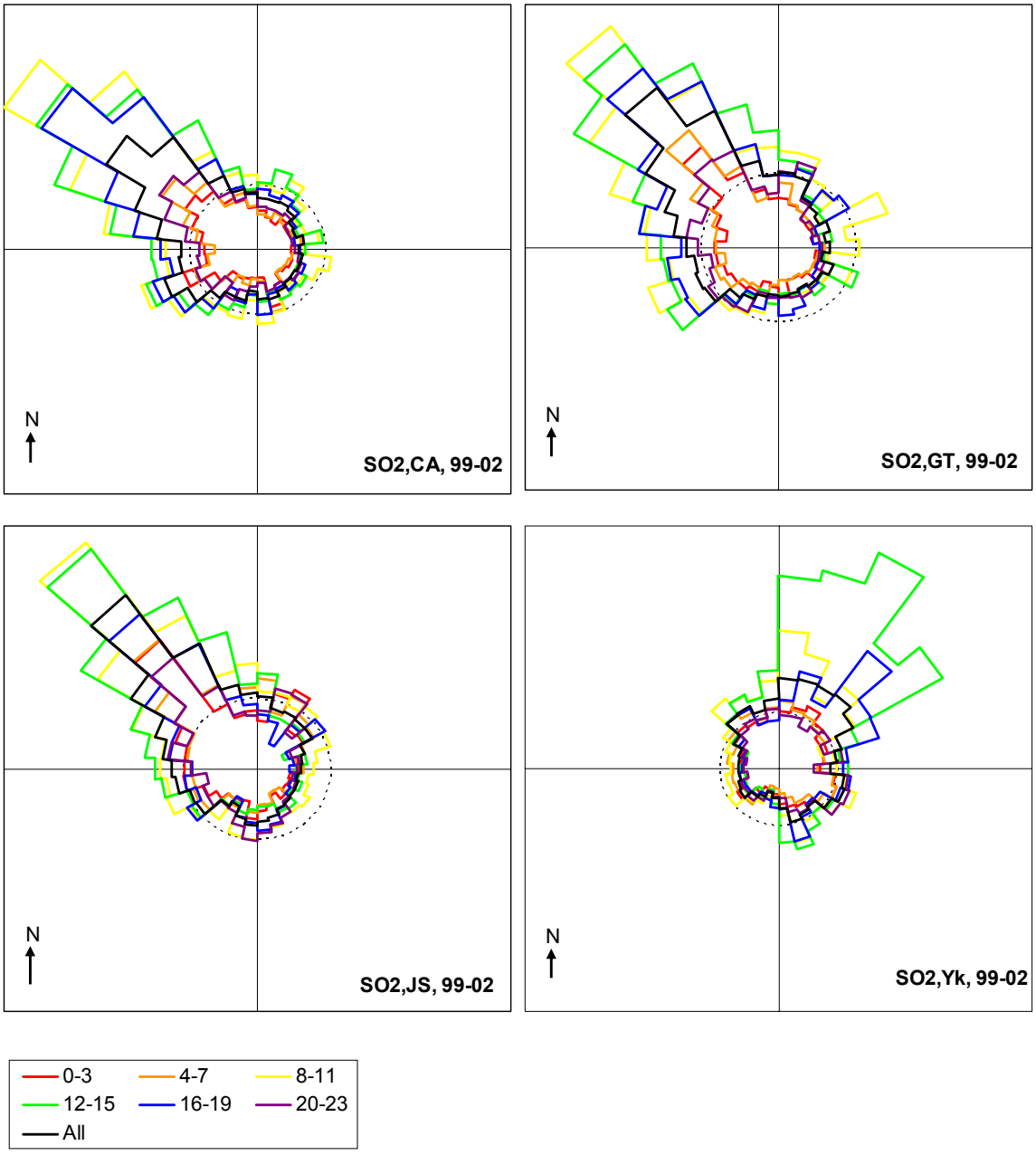


Figure 7.9: SO_2 pollution rose plots by time of day.

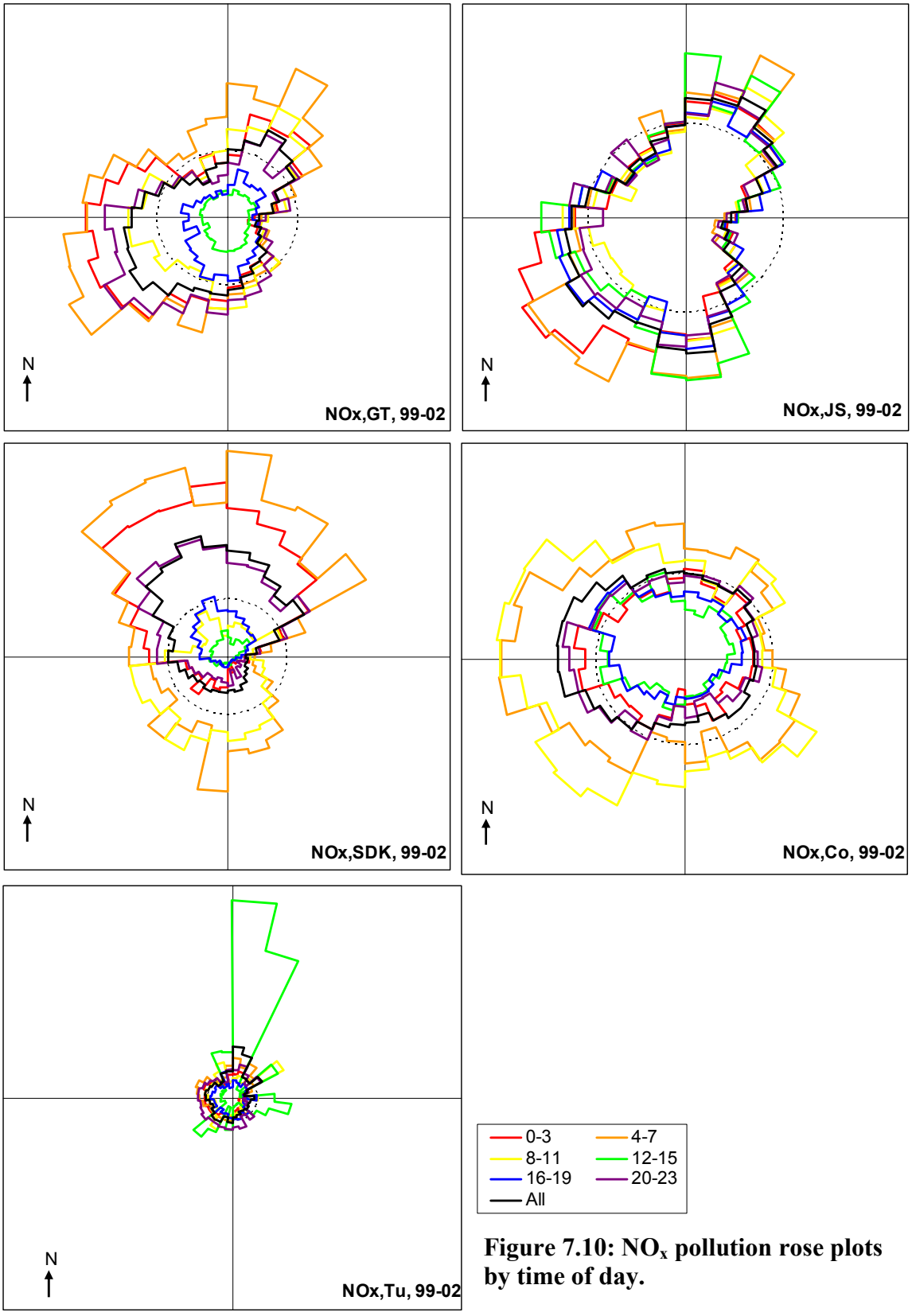


Figure 7.10: NO_x pollution rose plots by time of day.

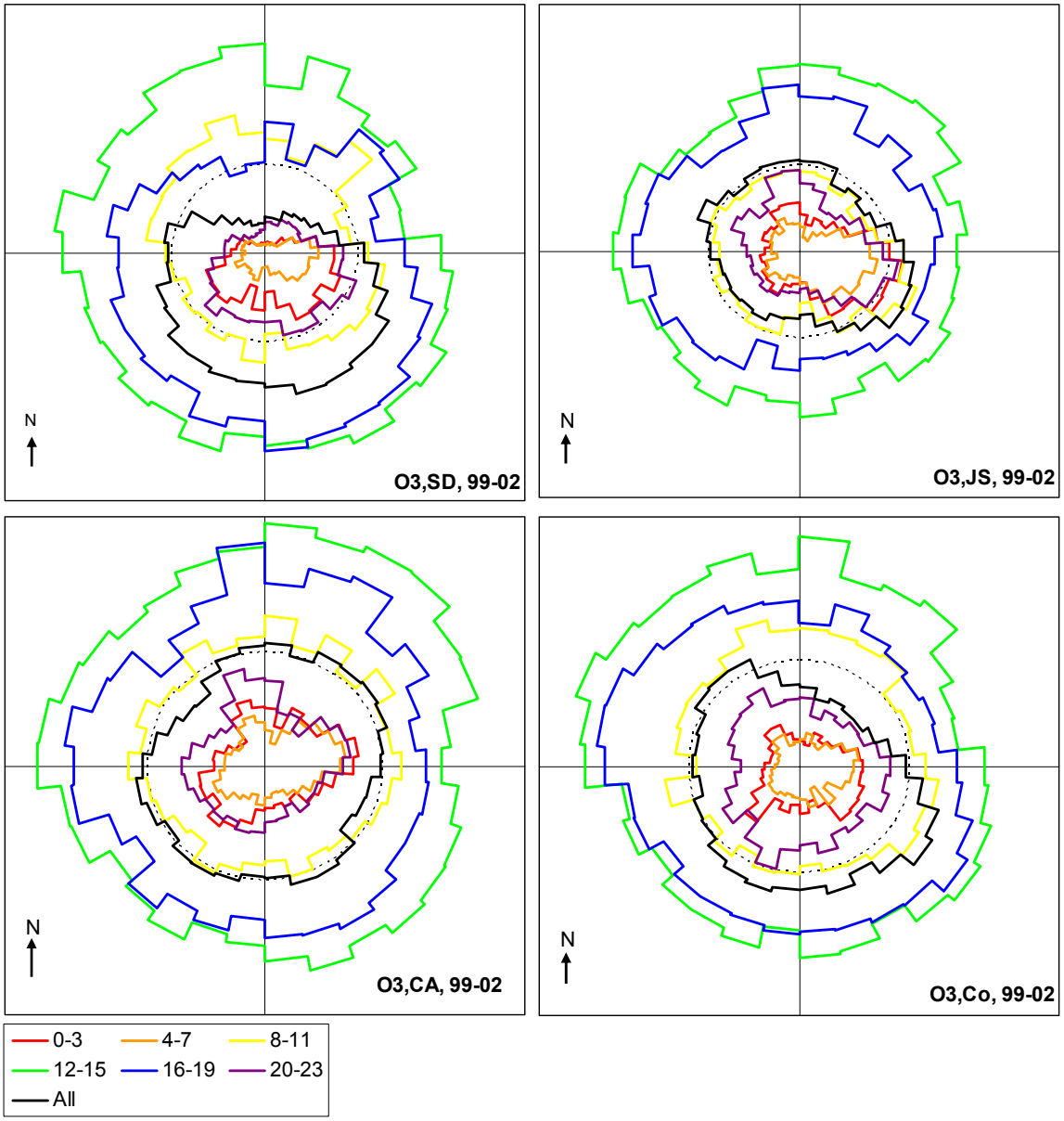


Figure 7.11: Pollution rose plots for ozone by time of day.

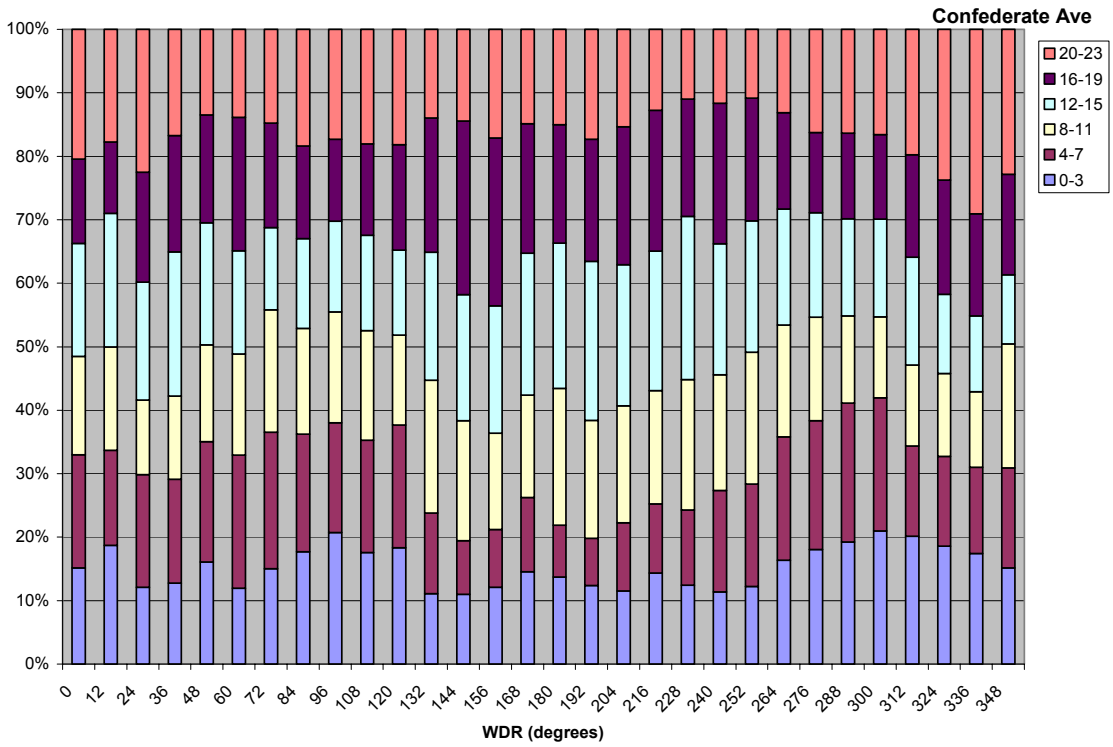


Figure 7.12: Wind direction profile at Confederate Avenue.

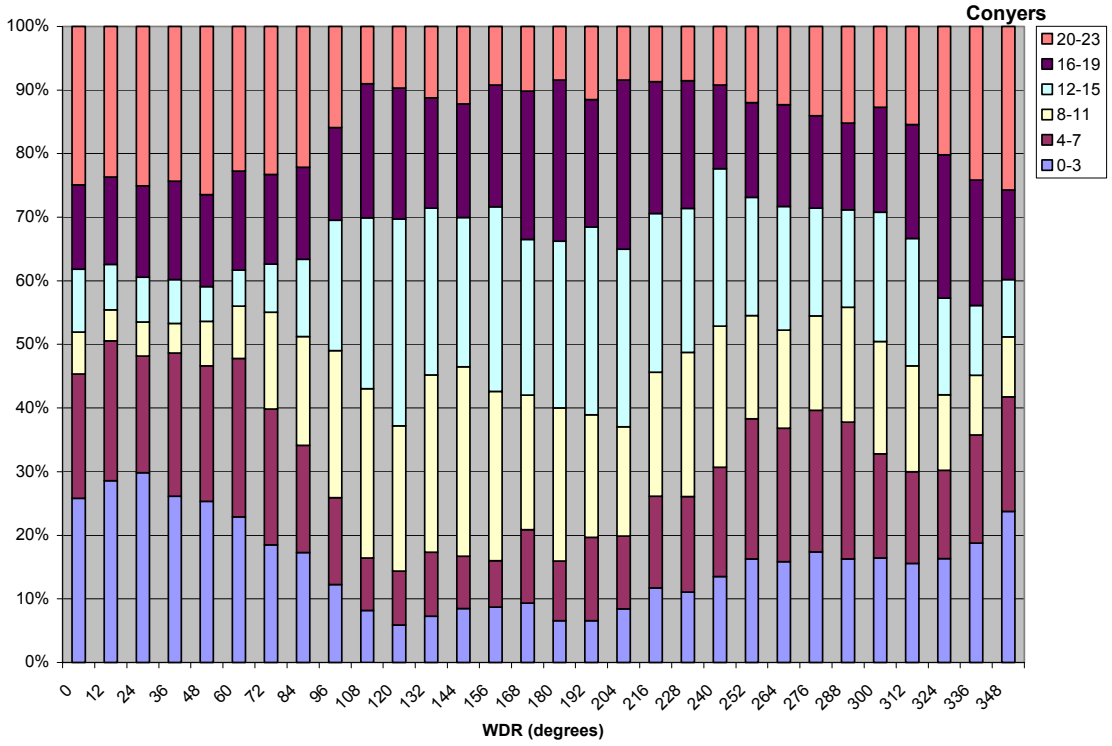


Figure 7.13: Wind direction profile at Conyers.

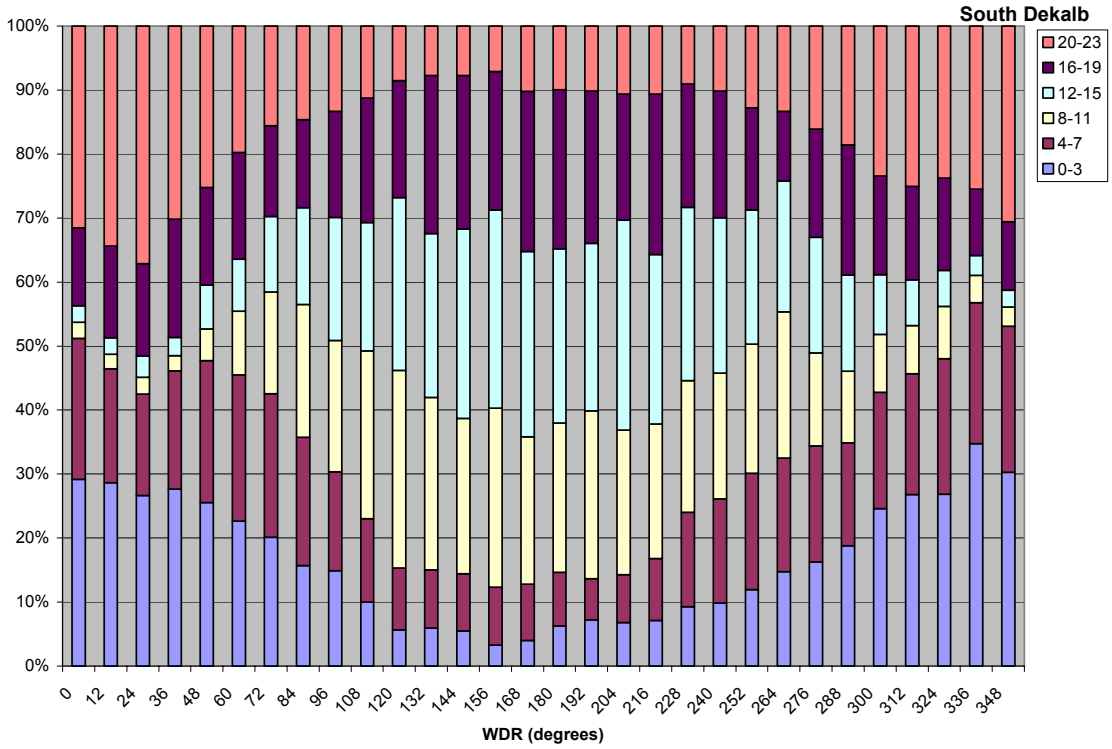


Figure 7.14: Wind direction profile for South Dekalb.

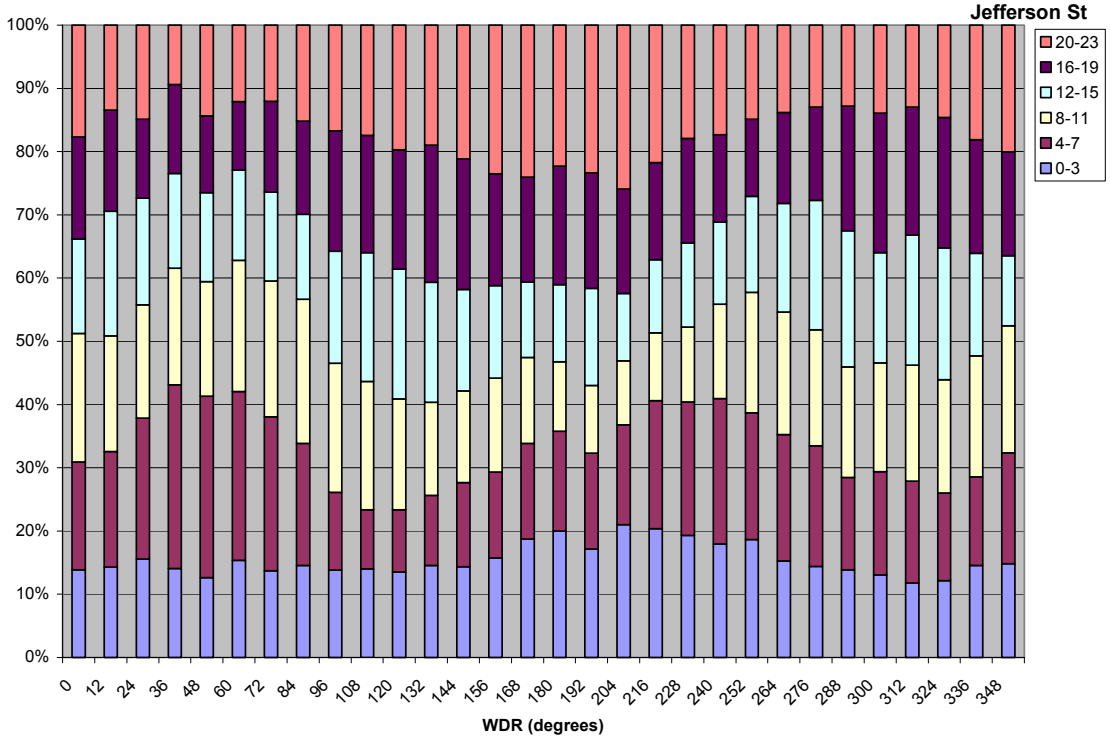


Figure 7.15: Wind direction profile for Jefferson Street.

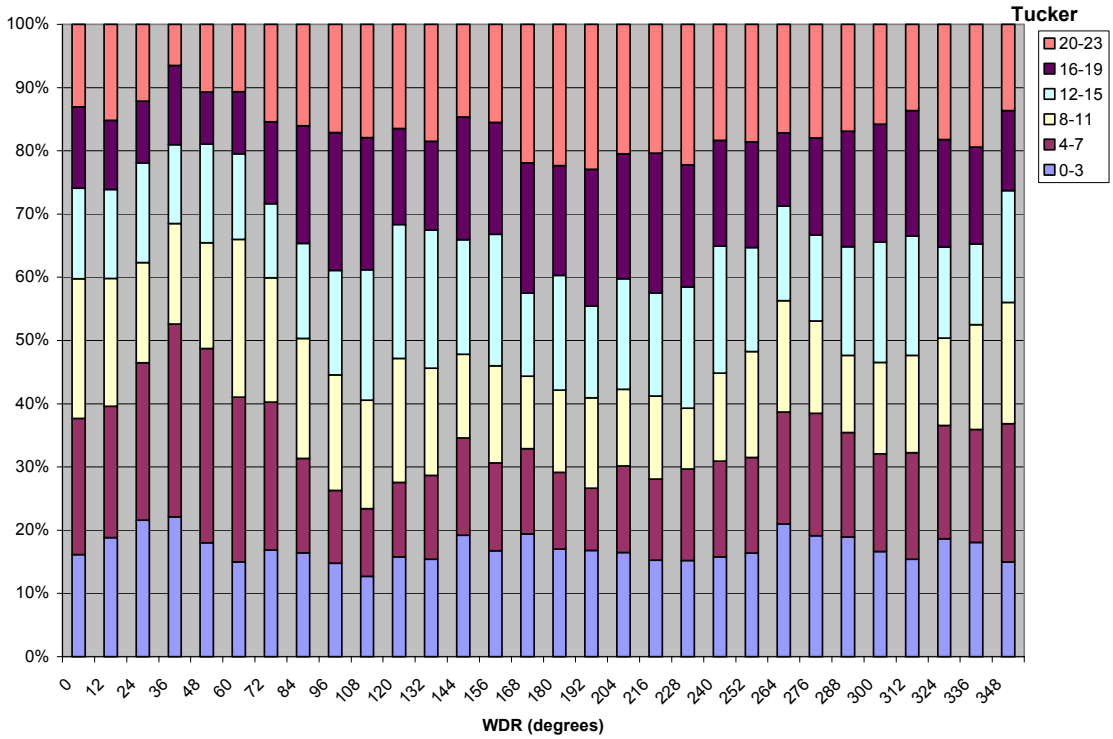


Figure 7.16: Wind direction profile for Tucker.

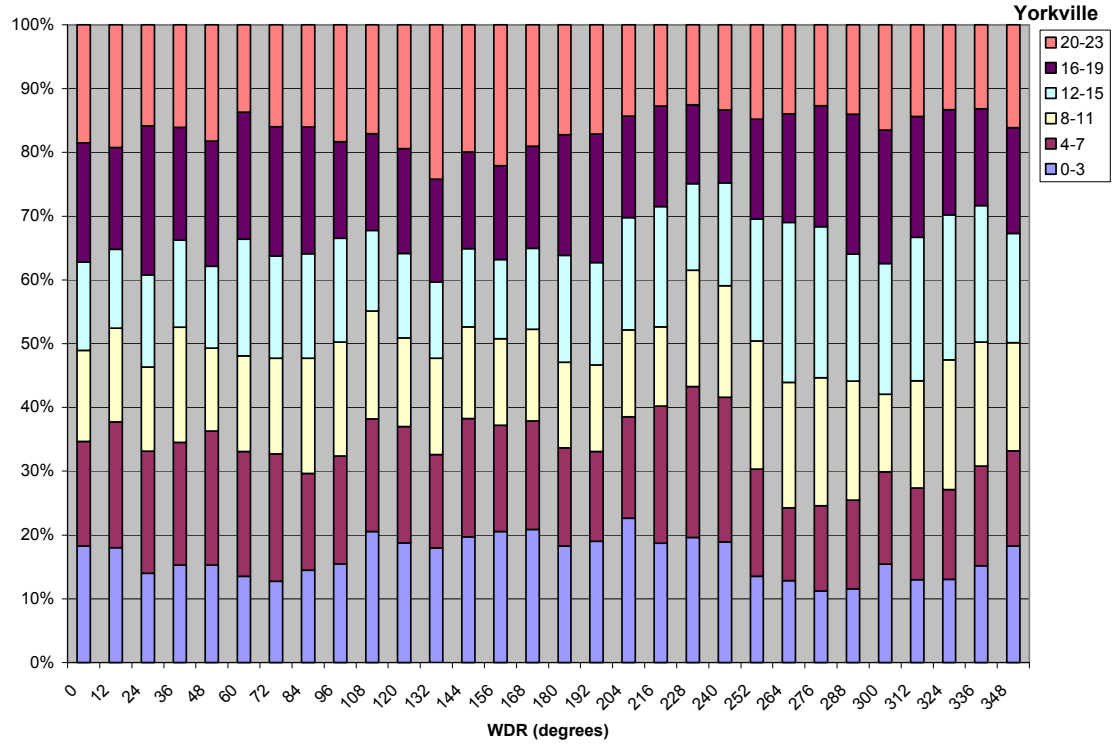


Figure 7.17: Wind direction profile for Yorkville.

7.5 Discussion

7.5.1 Comparison of emission inventory and pollution rose plots

Final pollution rose plots give a good indication of mobile source impacts on a monitoring site. For example, both CO and NO_x plots have peaks in directions of major highways. However, point sources are not as clearly defined for CO and NO_x. The point sources are most likely eclipsed by the much larger mobile contribution to emissions.

However, in the SO_2 pollution rose plots, where there is very little mobile source contribution, point sources are easy to identify. The five largest SO_2 emitters, Plants Bowen, McDonough, Wansley, and Yates and Lafarge Building Materials, are located in peak directions for all SO_2 monitoring sites.

7.5.2 Polllution rose plots and semivariograms

Variations in the windrose plots can help explain the scatter of data points seen in the semivariograms developed in Section 6. Because wind direction is not uniform or constant, local source impacts are anisotropic. Different sites will not register the impact from a local source at the same time due to variations in wind direction.

If a site has many large local impacts, the initial slope of the semivariogram function is expected to be steeper than if a site has no local source impacts. This is true for SO_2, CO, and NO_x, which have more sources apparent on their respective pollution rose plots than $PM_{2.5}$ and O_3 and also have a steeper semivariogram slope.

7.5.3 Source apportionment and $PM_{2.5}$ pollution rose

Another, more detailed, method of determining local sources of $PM_{2.5}$ is source apportionment. This is a much more computationally intensive method that involves analyzing the components of $PM_{2.5}$ to determine potential sources that are contributing to the overall concentration. Source apportionment of $PM_{2.5}$ at the Jefferson Street station has been performed and has determined that while a large fraction of $PM_{2.5}$ at Jefferson Street is secondary, which agrees with the largely circular shape of the pollution rose plot, and diesel emissions are the largest source of primary $PM_{2.5}$ (Zheng, 2002; Marmur, 2005). The Jefferson Street $PM_{2.5}$ pollution rose is slightly elongated to the northeast, as is the black carbon pollution rose plot. Diesel vehicles are one of the primary contributors to black carbon concentrations; therefore, the similarities between the $PM_{2.5}$ and black carbon plots could indicate a diesel source to the northeast of the Jefferson Street site, which would agree with the source apportionment results.

7.6 Conclusions

Local sources influencing pollution concentrations observed at monitoring stations in the Atlanta area were investigated through the construction of pollution rose plots and comparison to the emissions inventory for the Atlanta area. This analysis supported the hypothesis, based on examination of the emissions inventory, that power plants and a cement facility are influencing SO_2 observations. CO and NO_x concentrations appear to be largely influenced by mobile sources. Examining the pollution rose plots by time of day confirms that similar impacts, at varying concentrations, are experienced throughout the day. Correction of the plots for time of

day and season had little affect on the original plots. South Dekalb and Conyers ozone plots were the notable exception as corrections led to an appreciably more uniform pollution rose.

Variations in the windrose plots can help explain the scatter of data points seen in the semivariograms developed in Section 6. Because wind direction is not uniform or constant, local source impacts are anisotropic. Different sites will not register the impact from a local source at the same time due to variations in wind direction.

Chapter 8

CONCLUSIONS

8.1 Conclusions

This research attempted to first characterize temporal trends in air pollution in the Atlanta area on several different temporal scales and then to quantify and qualitatively describe the impacts of instrument error and local sources on the monitoring sites.

Temporal patterns in air pollutants are important to consider when performing analysis on large datasets. Factors affecting these patterns include emissions rates, meteorology and photochemistry. Each pollutant will be affected differently by combinations of these factors and will therefore have unique temporal profiles. Annual, seasonal, weekly, and diurnal profiles were examined for all pollutants. Sites with similar surroundings, i.e. urban or rural, had the most similar profiles.

It is important to quantify instrument error to understand overall uncertainty in a study using ambient air pollutant data. Audit and calibration data and collocated monitor data was used to assess instrument error. Collocated monitors are expected to give a more accurate indication of instrument error. Percent errors were similar for all pollutants using audit and calibration data (2-4%) and were similar for all concentration ranges. Percent errors using collocated data were several times larger.

Semivariograms were developed to assess spatial variation of air pollutants. The semivariograms can be interpreted as the amount of temporal variation of a pollutant that is actually spatial variation. There are noticeable differences in semivariograms among pollutants. For example, at 30km, almost 80% of the temporal variation of SO_2 concentrations is actually spatial variation. However, for ozone, only 30% of the temporal variation at 30km is actually spatial variation. Therefore, using only one monitoring site for the study area would not be appropriate for SO_2 but could give enough information for O_3. These results suggest that, particularly for primary pollutants, data from multiple sites should be used in epidemiologic health effects studies to minimize measurement error.

Local sources influencing pollution concentrations observed at monitoring stations in the Atlanta area were investigated through the construction of pollution rose plots and comparison to the emissions inventory for the Atlanta area. This analysis supported the hypothesis, based on examination of the emissions inventory, that power plants and a cement facility are influencing SO_2 observations. CO and NO_x concentrations appear to be largely influenced by mobile sources, which are most likely eclipsing any effects from local sources. Examining the pollution rose plots by time of day confirms that similar impacts, at varying concentrations, are experienced throughout the day. Correction of the plots for time of day and season had little affect on the original

plots. South Dekalb and Conyers ozone plots were the notable exception as corrections led to an appreciably more uniform pollution rose.

8.2 Suggestions for future work

A longer time frame database (1968-2004) has already been completed and will be used for more detailed, historical temporal analysis of air quality in the Atlanta area. This data can also be used to look for temporal patterns in local source emissions. For example, semivariograms and pollution roses using different time frames can be examined.

The semivariograms developed by this research should be used to quantify exposure error in epidemiologic health-effects studies. The effects of measurement error on the outcomes of these studies are to introduce a bias to the null and to widen the confidence interval of the outcome. Although measurement error cannot be eliminated from the study, epidemiologists can use the population-weighted average for the semivariogram function to determine how much error is attributed to measurement error.

The semivariogram functions can also be used to help establish new monitoring networks. Based on this research, one could determine how many monitors for a given pollutant are necessary to achieve a given level of uncertainty. For example, for the same level of uncertainty across a study area, a much smaller number of ozone monitors would be needed than sulfur dioxide monitors.

REFERENCES

Aneja, V.P.; Oomen, R.G.; Riordan, A.J.; Arya, S.P.; Wayland, R.J.; Murray, G.C. Ozone patterns for three metropolitan statistical areas in North Carolina, USA. *Atmos. Environ.* **1999**, 33, 5081-5093.

Anh, V. Duc, H., Shannon, I. Spatial variability of Sydney air quality by cumulative semivariogram. *Atmos. Environ.* **1997**, *31*, 4073-4080.

Burton, C.S. Ozone air quality models. *JAPCA* **1988**, *38*, 1119-1128.

Buzorius, G.; Hameri, K.; Pekkanen, J.; Kulmala, M. Spatial variation of aerosol number concentration in Helsinki city. *Atmos. Environ.* **1999**, *33*, 553-565.

Carrothers, T.J.; Evans, J.S. Assessing the impact of differential measurement error on estimates of fine particle mortality. *J. Air & Waste Manage. Assoc.* **2000,** *50*, 65-74.

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.

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.

De Iaco, S.; Myers, D.E.; Posa, D. Space-time variograms and a functional form for total air pollution measurements. *Comput. Stat. Data Anal.* **2002**, *41*, 311-328.

Diem, J.E. A critical examination of ozone mapping from a spatial-scale perspective. *Environ. Pollution.* **2003**, *125*, 369-383.

Duc, H.; Shannon, I.; Azzi, M. Spatial distribution characteristics of some air pollutants in Sydney. *Mathematics and Computers in Simulation* **2000**, 54, 1-21.

Ferreira, F.; Tente, H.; Torres, P.; Cardoso, S.; Palma-Oliviera, J.M. Air quality monitoring and management in Lisbon. *Environ Monit Assess* **2000**, *65*, 443-450.

Georgia Environmental Protection Division. 1999 Point source emissions inventory. *Personal correspondence*, **2002.**

Goldstein, I.F. and Landovitz, L. Analysis of air pollution patterns in New York City Em Dash 2. Can one aerometric station represent the area surrounding it? *Atmos. Environ.* **1977**, *11*, 53-57.

Grondona, M.O. and Cressie, N. Using spatial considerations in the analysis of experiments. *Technometrics* **1991***, 33*, 381-392.

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.

Hoek, G. et al. Spatial variability of fine particle concentrations in three European areas. *Atmos. Environ.* **2002**, *36*, 4077-4088.

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.

Kirby, C.; Greig, A.; Drye, T. Temporal and spatial variations in nitrogen dioxide concentrations across an urban landscape: Cambridge, UK. *Environ. Monitoring and Assessment* **1998**, *52*, 65-82.

Krautler, C. Growth patterns in metro Atlanta. *Atlanta Regional Commission*. **2005.**

Lebret, E., Briggs, D., Reeuwijk, H., Fischer, P., Smallbone, K., Harssema, H., Kriz, B., Gorynski, P., Elliott, P. Small area variations in ambient NO2 concentrations in four European areas. *Atmos. Environ.* **2000**, *34*, 177-185.

Lin, T.; Young, L.; Wang, C. Spatial variations of ground level ozone concentrations in areas of different scales. *Atmos. Environ.* **2001**, *35*, 5799-5807.

Lipfert, F.W. and 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.

Ma, C.; Oki, Y.; Tohno, S.; Kasahar, M. Assessment of wintertime atmospheric pollutants in anurban area of Kansai, Japan. *Atmos. Environ.* **2004**, *38*, 2939-2949.

Marmur, A.; Ulpner, A., Mulholland, J.A.; Russell, A.G. Optimization-Based Source Apportionment of PM2.5 Incorporating Gas-to-Particle Ratios. *Environ. Sci. Technol.* **2005,** *39*, 3245-3254.

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.

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. *Epidemiology* **2004**, *15*, 46-56.

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.

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.

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.

Navazo, M.; Durana, N.; Alonso, L.; Garcia, J.A.; Ilardia, J.L.; Gomez, J.C.; Gangoiti, G. Volatile organic compounds in urban and industrial atmospheres: measurement techniques and data analysis*. Intern. J. Environ. Anal. Chem.* **2002**, 83, 199-217.

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. *Epidemiology* **2005**, *16*, 184-174.

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.

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.

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.

Schwartz, J.; Dockery, D.W.; Neas, L.M. Is daily morality associated specifically with fine particles? *J. Air & Waste Manage. Assoc.* **1996**, *46*, 927-939.

Sen, Z. Regional air pollution assessment by cumulative semivariogram technique. *Atmos. Environ.* **1995**, *29*, 543-548.

Sen, Z. An application of a regional air pollution estimation model over Istanbul urban area. *Atmos. Environ.* **1998**, *32*, 3425-3433.

Stanley, T.W. Intent of a performance audit. *Transactions - Quality Assurance in Air Pollution Measurements,* **1985**, 673-680.

Tiles, S.; Zimmerman, J. Investigation on the spatial scales of the variability in measured near-ground ozone mixing ratios. *Geophysical Research Letters*. **1998**, *25*, 3827-3830.

US Census Bureau. http://www.census.gov/ **2000.**

US EPA. Code of Federal Regulations, Title 40. http://www.epa.gov/epahome/cfr40.htm. **2000.**

US EPA. National emissions inventory. http://www.epa.gov/ttn/chief/eiinformation.html. **2002.**

Vardoulakis, S.; Fisher, B.E.; Pericleous, K.; Gonzalez-Flesca, N. Modelling air quality in street canyons: A review. *Atmos. Environ.* **2003**, *37*, 155-182.

Villasenor, R.; Ortiz, E.; Watson, J.; Chow, J. Spatial and temporal variations in ambient $PM_{2.5}$ and PM_{10} in Mexico City. *J. Aerosol Sci.* **2000**, *31*, S901-S902.

Waller, L.A.; Gotway, C.A. *Applied Spatial Statistics for Public Health Data.* Chapter 8: Spatial Exposure Data. Wiley and Sons: New York, 2004.

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

Zheng, M.; Cass, G.R.; Schauer, J.J.; Edgerton, E.S.; Source Apportionment of PM2.5 in the Southeastern United States Using Solvent-Extractable Organic Compounds as Tracers. *Environ. Sci. Tecnol*. **2002**, *36*, 2361-2371.