UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

Risk and Exposure Assessment to Support the Review of the NO2 Primary National Ambient Air Quality Standard: Draft Technical Support Document (TSD)

EPA-452/P-08-002
April 2008

Risk and Exposure Assessment to Support the Review of the NO2 Primary National Ambient Air Quality Standard: Draft Technical Support Document (TSD)

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
Research Triangle Park, North Carolina

Disclaimer

This document has been prepared by staff from the Ambient Standards Group, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, and was written with support from technical documents from ICF International (through Contract No. EP-D-06-115). Any opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily reflect the views of the EPA or ICF International. This is the first draft document submitted to support the NO_2 Risk and Exposure Assessment for review and comment from the Clean Air Scientific Advisory Committee (CASAC) and the general public. Any questions concerning this draft document should be addressed to Dr. Stephen E. Graham, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, C504-06, Research Triangle Park, North Carolina 27711 (email: graham.stephen@epa.gov).

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Acronyms/Abbreviations

AADT	Annual average daily traffic
A/C	Air conditioning
AER	Air exchange rate
AERMOD	American Meteorological Society (AMS)/EPA Regulatory Model
AHS	American Housing Survey
APEX	EPA's Air Pollutants Exposure model, version 4
ANOVA	One-way analysis of variance
AQS	EPA's Air Quality System
BRFSS	Behavioral Risk Factor Surveillance System
CAMD	EPA's Clean Air Markets Division
CASAC	Clean Air Scientific Advisory Committee
CDC	Centers for Disease Control
CHAD	EPA's Consolidated Human Activity Database
CMSA	Consolidated metropolitan statistical area
CO	Carbon monoxide
COV	Coefficient of Variation
CTPP	Census Transportation Planning Package
DVRPC	Delaware Valley Regional Planning Council
EPA	United States Environmental Protection Agency
EOC	Exposure of Concern
GM	Geometric mean
GSD	Geometric standard deviation
hr	Hour
ID	Identification
ISA	Integrated Science Assessment
ISH	Integrated Surface Hourly Database
km	Kilometer
L95	Lower limit of the 95th confidence interval
m	Meter
ME	Microenvironment
max	Maximum
med	Median
min	Minimum
MSA	Metropolitan statistical area
NAAQS	National Ambient Air Quality Standards
NAICS	North American Industrial Classification System
NCEA	National Center for Environmental Assessment
NEI	National Emissions Inventory
NEM	NAAQS Exposure Model
NCDC	National Climatic Data Center
NHAPS	National Human Activity Pattern Study
NHIS	National Health Interview Survey
NO_2	Nitrogen dioxide
NO_x	Oxides of nitrogen
NWS	National Weather Service

O_3	Ozone
OAQPS	Office of Air Quality Planning and Standards
ORD	Office of Research and Development
ORIS	Office of Regulatory Information Systems identification code
POC	Parameter occurrence code
ppb	Parts per billion
PEN	Penetration factor
ppm	Parts per million
PRB	Policy-Relevant Background
PROX	Proximity factor
PVMRM	Plume Volume Molar Ratio Method
RECS	Residential Energy Consumption Survey
SAS	Statistical Analysis Software
SIC	Standard Industrial Code
SD	Standard deviation
se	Standard error
TDM	Travel Demand Modeling
tpy	Tons per year
TRIM	EPA's Total Risk Integrated Methodology
U95	Upper limit of the 95th confidence interval
US DOT	United States Department of Transportation
US EPA	United States Environmental Protection Agency
USGS	United States Geological Survey
VMT	Vehicle miles traveled

1 Introduction

The U.S. Environmental Protection Agency (EPA) is presently conducting a review of the national ambient air quality standards (NAAQS) for nitrogen dioxide (NO_2). Sections 108 and 109 of the Clean Air Act (The Act) govern the establishment and periodic review of the NAAQS. These standards are established for pollutants that may reasonably be anticipated to endanger public health and welfare, and whose presence in the ambient air results from numerous or diverse mobile or stationary sources. The NAAQS are to be based on air quality criteria, which are to accurately reflect the latest scientific knowledge useful in indicating the kind and extent of identifiable effects on public health or welfare that may be expected from the presence of the pollutant in ambient air. The EPA Administrator is to promulgate and periodically review, at five-year intervals, primary (health-based) and secondary (welfare-based) NAAQS for such pollutants. Based on periodic reviews of the air quality criteria and standards, the Administrator is to make revisions in the criteria and standards and promulgate any new standards as may be appropriate. The Act also requires that an independent scientific review committee advise the Administrator as part of this NAAQS review process, a function now performed by the Clean Air Scientific Advisory Committee (CASAC).

This report document, in detail, the methodology and input data used in the risk and exposure assessment for NO_2 conducted in support of the current review of the NO_2 NAAQS. Specifically, this report includes the following:

- Description of the areas assessed and populations considered
- Summary of the air quality modeling methodology and associated input data
- Description of the inhalation exposure model and associated input data
- Evaluation of estimated NO_2 exposures
- Assessment of the quality and limitations of the input data for supporting the goals of the NO_2 NAAQS exposure analysis.

2 Air Quality Characterization

2.1 Air Quality Data Screen

2.1.1 Introduction

The current NO_2 standard of 53 ppb annual arithmetic average was set in 1971 and has been retained since by subsequent reviews (i.e., 1985, 1995). Minor revisions to the standard made in 1985 included an explicit rounding convention, stated annual averages would be determined on a calendar year basis, and indicated an explicit 75% completeness requirement for monitoring (60 FR 52874). Each of these components of the standard were considered in characterizing the air quality monitoring data, beginning first with the selection of valid data.

2.1.2 Approach

NO_2 air quality data from years 1995 through 2006 and associated documentation were downloaded from EPA's Air Quality System (US EPA, 2007a; 2007b). As of the date of the analyses performed, hourly measurements for year 2006 were only available for January 1 through October 31, 2006. A *site* was defined by the state, county, site code, and parameter occurrence code (POC), which gives a 10-digit monitor ID code. The POC identifies collocated measurements at the same monitoring location, so that each measuring instrument is treated as a different site. Typically there was only one POC at a given monitoring location.

As required by the NO_2 NAAQS, a valid year of monitoring data is needed to calculate the annual average concentration. A valid year at a monitoring site is comprised of 75% of valid days in a year, with at least 18 hourly measurements for a valid day (thus at least 274 or 275 valid days depending on presence of a leap year, a minimum of 4,932 or 4,950 hours). This served as a screening criterion for data to be used for analysis.

Site-years of data are the total numbers of years the collective monitors in a location were in operation. For example, from years 1995-2006, the Boston CMSA had 27 total monitors in operation, some of which did not contain sufficient numbers of monitoring values, while others contained upwards of 11 years (Table 1). Thus in summing the number of operating years, this particular location contained a total of 105 site-years of data across the monitoring period.

In all of the subsequent analyses, where hourly values were missing they were treated as such. Reported values of zero (0) concentration were also retained as is. For certain illustrations, values of zero were substituted with 0.5 ppb, derived from one-half the lowest recorded 1-hour concentration (1 ppb).

Table 1. Example of monitors IDs and years of operation using the Boston CMSA.

	Year of monitoring (1995-2006)												Totals	
Monitor ID	**95**	**96**	**97**	**98**	**99**	**00**	**01**	**02**	**03**	**04**	**05**	**06**	**Complete**	**Incomplete**
2303130021	i	c	c	c	i	c	c	c	c	i	i		7	4
2500510021								i					0	1
2500510051		i	c	c	i	i	i						2	4
2500900051								i					0	1
2500920061	c	c	c	c	i	i	c	c	c	c	c	c	10	2
2500940041	c	c	c	c	i	i	c	i	i	i	i	i	5	7
2500950051										i	c	c	2	1
2502100091	c												1	0
2502130031								i	i	i	i	i	0	5
2502500021	c	c	c	c	c	c	c	c	i	c	c	c	11	1
2502500211	c	c	c	c	c	c	c	c					8	0
2502500351	c												1	0
2502500361	c												1	0
2502500401	c	c	c	c	c	c	c	c	c	c	c	i	11	1
2502500411					i	i	c	i	i	i	i	i	1	7
2502500421						i	c	c	c	c	c	c	6	1
2502510031	c	c	c	c	c								5	0
2502700201	c	c	c	c	c	c	c	c	i				8	1
2502700231										c	c	c	3	0
3301100161	c	c	c	c	i								4	1
3301100191					i	c	i						1	2
3301100201							i	c	c	c	c	c	5	1
3301110111										i	i	i	0	3
3301500091	c	c	c	c	c	i	i						5	2
3301500131				i	c	c	c	c	i				4	2
3301500141									i	c	c	c	3	1
3301500151							i	c	i				1	2
Complete	**12**	**10**	**11**	**11**	**7**	**7**	**10**	**10**	**5**	**7**	**8**	**7**	**105**	
Incomplete	**1**	**1**	**0**	**1**	**7**	**6**	**5**	**5**	**8**	**6**	**5**	**5**		**50**

Notes:
c = met criteria for valid year of monitoring data.
i = did not met criteria for valid year of monitoring data.

2.1.3 Results

Of a total of 5,243 site-years of data in the entire NO_2 1-hour concentration database, 1,039 site-years did not meet the above criterion and were excluded from any further analyses. In addition, since shorter term average concentrations are of interest, the remaining site-years of data were further screened for 75% completeness on hourly measures in a year (i.e., containing a minimum of 6,570 or 6,588, depending on presence of a leap year). Twenty-seven additional site-years were excluded, resulting in 4,177 complete site-years in the analytical database. Table 2 provides a summary of the site-years included in the analysis, relative to those excluded, by location and by two site-year groupings.[1] Location selection is defined in the Section 1.2.

---

[1] 14 of 18 named locations and the 2 grouped locations contained enough data to be considered valid for year 2006.

Table 2. Counts of complete site-years of NO_2 monitoring data.

	Number of Site-Years					
	Complete		**Incomplete**		**% Complete**	
Location	**1995-2000**	**2001-2006**	**1995-2000**	**2001-2006**	**1995-2000**	**2001-2006**
Boston	58	47	16	34	78%	58%
Chicago	47	36	20	22	70%	62%
Cleveland	11	11	2	2	85%	85%
Denver	26	10	10	4	72%	71%
Detroit	12	12	4	1	75%	92%
Los Angeles	193	177	16	19	92%	90%
Miami	24	20	1	4	96%	83%
New York	93	81	12	24	89%	77%
Philadelphia	46	39	6	8	88%	83%
Washington	69	66	21	18	77%	79%
Atlanta	24	29	5	1	83%	97%
Colorado Springs	26	0	4	4	87%	0%
El Paso	14	30	11	0	56%	100%
Jacksonville	6	4	0	2	100%	67%
Las Vegas	16	35	4	9	80%	80%
Phoenix	22	27	8	25	73%	52%
Provo	6	6	0	0	100%	100%
St. Louis	56	43	3	9	95%	83%
Other CMSA	1135	1177	249	235	82%	83%
Not MSA	200	243	112	141	64%	63%
Total	4177		1066		80%	

2.2 Selection of Locations

2.2.1 Introduction

The next step in this analysis was to identify similarities and differences in air quality among locations for the purpose of either aggregating or segregating data using a combination of descriptive statistics and health based criteria. *Location* in this context would include a geographic area that encompasses more than a single air quality monitor (e.g., particular city, consolidated metropolitan statistical area or CMSA).

2.2.2 Approach

Criteria were established for selecting sites with high annual means and/or frequent exceedances of potential health effect benchmarks. Selected locations were those that had a maximum annual mean NO_2 level at a particular monitor greater than or equal to 25.7 ppb, which represents the 90th percentile across all locations and site-years, and/or had at least one reported 1-hour NO_2 level greater than or equal to 200 ppb, the lowest level of the potential health effect benchmarks. A *location* in this context would include a geographic area that encompasses more than a single air quality monitor (e.g., particular city, metropolitan statistical area (MSA), or consolidated metropolitan statistical area or CMSA). First, all monitors were identified as either belonging to a CMSA, a MSA, or neither. Then, locations of interest were identified through statistical analysis of the ambient NO_2 air quality data for each site within a location.

2.2.3 Results

Fifteen locations met both selection criteria, that is, having at least one site-year annual mean above 25.72 ppb and at least one exceedance of 200 ppb. Upon further analysis of the more recent ambient data (2001-2006), four additional locations were observed to have met at least one of the criteria (either high annual mean and/or at least one exceedance of 200 ppb). New Haven, CT, while meeting the earlier criteria, did not have any recent exceedances of 200 ppb and contained one of the lowest maximum concentration-to-mean ratios, therefore was not separated out as a specific location. Thus, 14 locations were retained from the initial selection and 4 locations selected from a second screening to provide additional geographical representation. In addition to these 18 specific locations, the remaining sites were grouped into two broad location groupings. The *Other CMSA* location contains all the other sites that are in MSAs or CMSAs but are not in any of the 18 specified locations. The *Not MSA* location contains all the sites that are not in an MSA or CMSA. The selected locations are summarized in Table 3.

The final database for analysis included air quality data from a total of 205 monitors within the named locations, 331 monitors in the Other CMSA group, and 92 monitors in the Not MSA group. Again, the monitors that were retained contained the criteria for estimating a valid annual average concentration described above.

Table 3. Locations selected for NO_2 Air Quality Characterization, associated abbreviations, and values of selection criteria.

Location				Maximum # of Exceedances of 200 ppb	Maximum Annual Mean (ppb)
Type[1]	**Code**	**Description**	**Abbreviation**		
CMSA*	1122	Boston-Worcester-Lawrence, MA-NH-ME-CT	Boston	1	31.1
CMSA	1602	Chicago-Gary-Kenosha, IL-IN-WI	Chicago	0	33.6
CMSA*	1692	Cleveland-Akron, OH	Cleveland	1	28.1
CMSA*	2082	Denver-Boulder-Greeley, CO	Denver	2	36.8
CMSA*	2162	Detroit-Ann Arbor-Flint, MI	Detroit	12	25.9
CMSA*	4472	Los Angeles-Riverside-Orange County, CA	Los Angeles	5	50.6
CMSA	4992	Miami-Fort Lauderdale, FL	Miami	3	16.8
CMSA*	5602	New York-Northern New Jersey-Long Island, NY-NJ-CT-PA	New York	3	42.2
CMSA*	6162	Philadelphia-Wilmington-Atlantic City, PA-NJ-DE-MD	Philadelphia	3	34.00
CMSA*	8872	Washington-Baltimore, DC-MD-VA-WV	Washington DC	2	27.2
MSA*	0520	Atlanta, GA	Atlanta	1	26.6
MSA*	1720	Colorado Springs, CO	Colorado Springs	69	34.8
MSA*	2320	El Paso, TX	El Paso	2	35.1
MSA	3600	Jacksonville, FL	Jacksonville	2	15.9
MSA*	4120	Las Vegas, NV-AZ	Las Vegas	11	27.1
MSA*	6200	Phoenix-Mesa, AZ	Phoenix	37	40.5
MSA	6520	Provo-Orem, UT	Provo	0	28.9
MSA*	7040	St, Louis, MO-IL	St. Louis	8	27.2
MSA/CMSA	-	Other MSA/CMSA	Other CMSA	10	31.9
-	-	Other Not MSA	Not MSA	2	19.7

[1] CMSA is consolidated metropolitan statistical area; MSA is metropolitan statistical area according to the 1999 Office of Management and Budget definitions (January 28, 2002 revision).

* Indicates locations that satisfied both the annual average and exceedance criteria.

2.3 Ambient Monitor Characterization

2.3.1 Introduction

Siting of monitors is of particular importance, recognizing that proximity of local sources could influence on measured NO_2 concentrations. As part of the risk and exposure scope and methods document (US EPA, 2007c), both mobile and stationary sources (in particular power generating utilities using fossil fuels) were indicated as significant contributors to nitrogen oxides (NO_x) emissions in the U.S. Analyses were performed to determine the distance of all location-specific monitors to these source categories. In addition, emissions of NO_x from stationary sources within close proximity of the location-specific monitoring sites were estimated.

2.3.2 Approach

Major road distances to each monitor were calculated using GIS. Distances of monitoring sites to stationary sources and those source's emissions were estimated using data within the 2002 National Emissions Inventory (NEI; US EPA, 2007d). The NEI database reports emissions of NO_x in tons per year (tpy) for 131,657 unique emission sources at various points of release. The release locations were all taken from the latitude longitude values within the NEI. First, all NO_x emissions were summed for identical latitude and longitude entries while retaining source codes for the emissions (e.g., Standard Industrial Code (SIC), or North American Industrial Classification System (NAICS)). Therefore, any facility containing similar emission processes were summed at the stack location, resulting in 40,855 observations. These data were then screened for sources with emissions greater than 5 tpy, yielding 18,798 unique NO_x emission sources. Locations of these stationary source emissions were compared with ambient monitoring locations using the following formula:

$$d = \arccos(\sin\left(lat_1\right) \times \sin(lat_2) + \cos(lat_1) \times \cos(lat_2) \times \cos(lon_2 - lon_1)) \times r$$

where

d = distance (kilometers)
lat_1 = latitude of a monitor (radians)
lat_2 = latitude of source emission (radians)
lon_1 = longitude of monitor (radians)
lon_2 = longitude of source emission (radians)
r = approximate radius of the earth (or 6,371 km)

Location data for monitors and sources provided in the AQS and NEI data bases were given in units of degrees therefore, these were first converted to radians by dividing by 180/π. For each monitor, source emissions with estimated distances within 10 km were retained.

2.3.3 Results

The distribution of the nearest distance of the ambient monitors to major roads for each of the named locations is summarized in Table 4.[2] Physical attributes of individual monitors (e.g., latitude/longitude, probe height) including the distance of the nearest major road is provided in the Appendix A. On average, most monitors are placed at a distance of 50 meters or greater from a major road, however in locations with a large monitoring network such as Boston, Chicago, or New York CMSA, there may be one or two monitors placed within close proximity (<10 meters) of a road.

Table 4. Distribution of the distance (m) of ambient monitors to the nearest major road in selected locations.

		Distance (m) of monitor to nearest major road						
Location	**n**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**
Atlanta	4	488	283	134	134	505	809	809
Boston	21	101	93	7	7	70	337	337
Chicago	12	158	212	2	2	93	738	738
Cleveland	4	114	90	2	2	134	187	187
Colorado Springs	6	196	103	79	79	180	386	386
Denver	7	166	260	18	18	65	748	748
Detroit	3	382	39	339	339	393	415	415
El Paso	7	282	266	33	33	128	718	718
Jacksonville	1	144						
Las Vegas	10	244	286	1	1	181	914	914
Los Angeles	43	155	150	1	2	89	522	570
Miami	4	57	45	15	15	55	103	103
New York	26	145	130	6	6	119	508	508
Philadelphia	10	247	199	45	45	167	630	630
Phoenix	7	190	177	7	7	141	433	433
Provo	1	353						
St Louis	13	126	123	5	5	97	421	421
Washington DC	16	129	104	14	14	83	338	338

[1] n is the number of monitors operating in a particular location between 1995 and 2006. The min, 2.5, med, 97.5, and max represent the minimum, 2.5th, median, 97.5th, and maximum percentiles of the distribution for the distance in meters (m) to the nearest major road. Monitors > 1km from road are not included.

Table 5 contains a summary of the distance of stationary source emissions to monitors within each named location. There were a number of sources emitting >5 tpy of NO_x and located within a 10 km radius for many of the monitors. On average though, most monitors are placed at greater distances from stationary source emissions than roads with most sources at a distance of greater than 5 km. Most of the stationary source emissions of NO_x within a 10 km radius of monitors were less than 50 tpy (Table 6). Details regarding individual monitors are provided in Appendix A.

---

[2] Distances between monitors and major roads were first determined using a Tele-Atlas roads database in a GIS application. For road-monitor pairs that showed particularly close distances, the values were fine-tuned using GoogleEarth® to estimate the distance to road edge.

Table 5. Distribution of the distance (m) of ambient monitors to stationary sources with NO_x emissions >5 tons per year (tpy) and within a 10 kilometers (km)[1] radius.

		Distance of monitor to NO_x emission source (m)						
Location	**n**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**
Atlanta	9	6522	3164	656	656	7327	9847	9847
Boston	595	5333	2603	142	761	5363	9733	9988
Chicago	394	6586	2657	411	770	7277	9834	9994
Cleveland	19	7092	2439	956	956	7278	9884	9884
Colorado Springs	66	6109	2632	782	1034	6340	9847	9933
Denver	140	5655	2593	910	1029	5904	9862	9979
Detroit	87	6889	2254	321	1963	7549	9974	9997
El Paso	126	5694	3185	119	1384	6085	9945	9991
Jacksonville	20	5125	2962	708	708	5720	9558	9558
Las Vegas	18	6700	2184	3837	3837	7237	9950	9950
Los Angeles	523	6003	2435	140	1483	6165	9801	9991
Miami	11	6184	3151	1323	1323	7611	9117	9117
New York	736	6101	2555	103	1383	6467	9818	9983
Philadelphia	382	5837	2474	231	1299	5689	9754	9982
Phoenix	59	6298	2279	833	1312	6355	9803	9890
Provo	7	6558	3664	1214	1214	8178	9433	9433
St Louis	253	6799	2337	396	1989	7120	9863	9990
Washington DC	160	6173	2425	288	704	6254	9777	9973

[1] n is the number of sources emitting >5 tons per year (tpy) of NO_x within a 10 kilometer (km) radius of a monitor in a particular location. The min, 2.5, med, 97.5, and max represent the minimum, 2.5th, median, 97.5th, and maximum percentiles of the distribution for the distance in meters (m) to the source emission.

Table 6. Distribution of NO_x emissions from stationary sources within 10 kilometers (km) of monitoring site, where emissions were >5 tons per year[1].

		Emissions (tpy) of NO_x from sources within 10 km of monitor						
Location	**n**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**
Atlanta	9	709	1621	22	22	35	4895	4895
Boston	595	128	344	5	5	10	1155	3794
Chicago	394	204	919	5	5	10	2204	8985
Cleveland	19	702	612	126	126	284	1476	1476
Colorado Springs	66	387	1091	5	5	19	4205	4205
Denver	140	252	1286	5	5	15	5404	9483
Detroit	87	251	637	5	6	24	2398	3762
El Paso	126	117	286	5	5	31	912	1679
Jacksonville	20	201	407	5	5	31	1642	1642
Las Vegas	18	483	636	18	18	84	1665	1665
Los Angeles	523	70	310	5	5	12	577	4256
Miami	11	24	16	8	8	22	51	51
New York	736	284	1024	5	6	31	3676	9022
Philadelphia	382	154	408	5	5	29	1304	4968
Phoenix	59	85	234	5	5	14	1049	1049
Provo	7	60	38	7	7	83	102	102
St Louis	253	167	1032	5	5	16	848	14231
Washington DC	160	320	1254	6	6	34	6009	10756

[1] n is the number of sources emitting >5 tons per year of NO_x within a 10 kilometer radius of a monitor in a particular location. The min, 2.5, med, 97.5, and max represent the minimum, 2.5th, median, 97.5th, and maximum percentiles of the distribution for the source emission in tons per year (tpy).

2.4 Spatial and Temporal Air Quality Analyses

2.4.1 Introduction

An analysis of the air quality was performed to determine spatial and temporal trends, considering locations, monitoring sites within locations, and time-averaging of ambient NO_2 concentrations collected from 1995 through 2006. The purpose is to present relevant information on the air quality as it relates to both the current form of the standard (annual average concentration) and the exposure concentration and duration associated with adverse health effects (1-hour).

2.4.2 Approach

To evaluate variability in NO_2 concentrations, temporal and spatial distributions of summary statistics were computed in addition to use of statistical tests to compare distributions between years and/or monitors and/or locations. For a given location, the variability within that location is defined by the distribution of the annual summary statistics across years and monitors and by the distribution of the hourly concentrations across hours and monitors. The summary statistics were compiled into tables and used to construct figures for visual comparison and for statistical analysis.

Boxplots were constructed to display the distribution across sites and years (or hours for the hourly concentrations) for a single location. The box extends from the 25th to the 75th percentile, with the median shown as the line inside the box. The whiskers extend from the box to the 5th and 95th percentiles. The extreme values in the upper and lower tails beyond the 5th and 95th percentiles are not shown to allow for similar scaling along the y-axis for the plotted independent variables. The mean is plotted as a dot; typically it would appear inside the box, however it will fall outside the box if the distribution is highly skewed. All concentrations are shown in parts per billion (ppb).

Q-Q plots also display the distribution in the calculated air quality metrics across sites and years (or hours for the hourly concentrations) for a single location. The Q-Q plot is used to compare the observed cumulative distribution to a standard statistical distribution. In this case the observed distributions are compared with a log-normal distribution, so that the vertical scale is logarithmic. The horizontal scale is the quantile of a standard normal distribution, so that if there are N observed values, then the k^{th} highest value is plotted against the quantile probit(p), where probit is the inverse of the standard normal distribution function, and p is the plotting point. The plotting points were chosen as $p = (k-3/8)/(N+1/4)$ for the annual statistics and $p = k/(N+1)$ for the hourly concentrations. If the distribution were exactly log-normal, then the curve would be a straight line. The median value is the y-value when the normal quantile equals zero. The slope of the line is related to the standard deviation of the logarithms, so that the higher the slope, the higher the coefficient of variation (standard deviation divided by the mean for the raw data, before taking logarithms).

In addition to the tabular and graphical comparisons of the summary statistics, the distributions of each variable were compared using various statistical tests. An F-Statistic comparison compares the mean values between locations using a one way analysis of variance

(ANOVA). This test assumed that for each location, the site-year or site-hour variables are normally distributed, with a mean that may vary with the location and a constant variance (i.e., the same for each location). Statistical significance was assigned for p-values less than or equal to 0.05. The Kruskal-Wallis Statistics are non-parametric tests that are extensions of the more familiar Wilcoxon tests to two or more groups. The analysis is valid if the difference between the variable and the location median has the same distribution for each location. If so, this procedure tests whether the location medians are equal. The test is also consistent under weaker assumptions against more general alternatives. The Mood Statistic comparisons are non-parametric tests that compare the scale statistics for two or more groups. The scale statistic measures variation about the central value, which is a non-parametric generalization of the standard deviation. This test assumes that all the groups have the same median. Specifically, suppose there is a total of N values, summing across all the locations to be compared. These N values are ranked from 1 to N, and the j^{th} highest value is given a score of $\{j - (N+1)/2\}^2$. The Mood statistic uses a one-way ANOVA statistic to compare the mean scores for each location. Thus the Mood statistic compares the variability between the different locations assuming that the medians are equal.

2.4.3 Spatial Results

A summary of the important spatial trends in NO_2 concentrations is reported in this section. Detailed air quality results (i.e., by year and within-location) are presented in Appendices B and C, each containing both tabular and graphic summaries of the spatial and temporal concentration distributions.

A broad view of the spatial differences in NO_2 monitoring concentrations across locations is presented in Figures 1 and 2. In general there is variability in NO_2 concentrations between the 20 locations. For example, in Los Angeles, the mean of annual means is approximately 24.3 ppb over the period of analysis, while considering the Not MSA grouping, the mean annual mean was about 7.0 ppb. Phoenix contained the highest mean annual mean of 27.3 ppb. Variability in the annual average concentrations was also present within locations, the magnitude of which varied by location. On average, the coefficient of variation in the annual mean concentrations was about 35%, however locations such as Jacksonville or Provo had COVs as low as 6% while locations such as Las Vegas and Not MSA contained COVs above 60%. Reasons for differing variability arise from the size of the monitoring network in a location, level of the annual mean concentration, underlying influence of temporal variability within particular locations, among others.

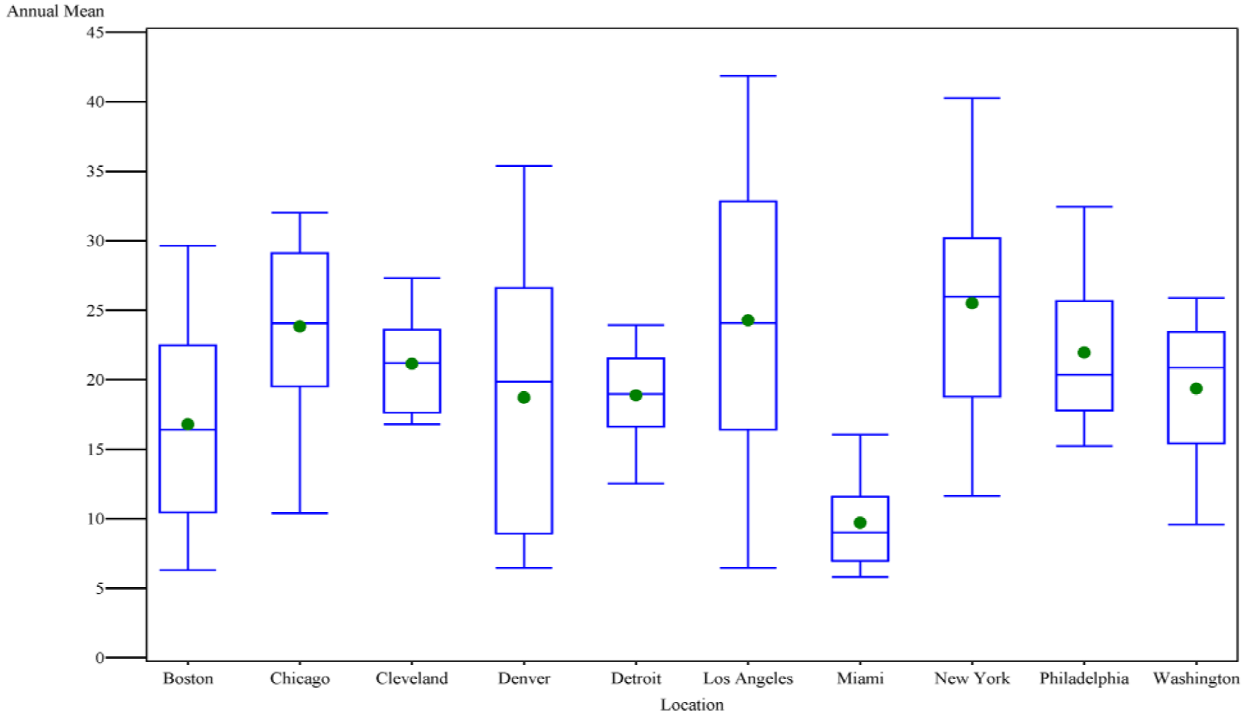


Figure 1. Spatial distributions of annual mean NO_2 ambient monitoring concentrations for selected CMSA locations, years 1995-2006.

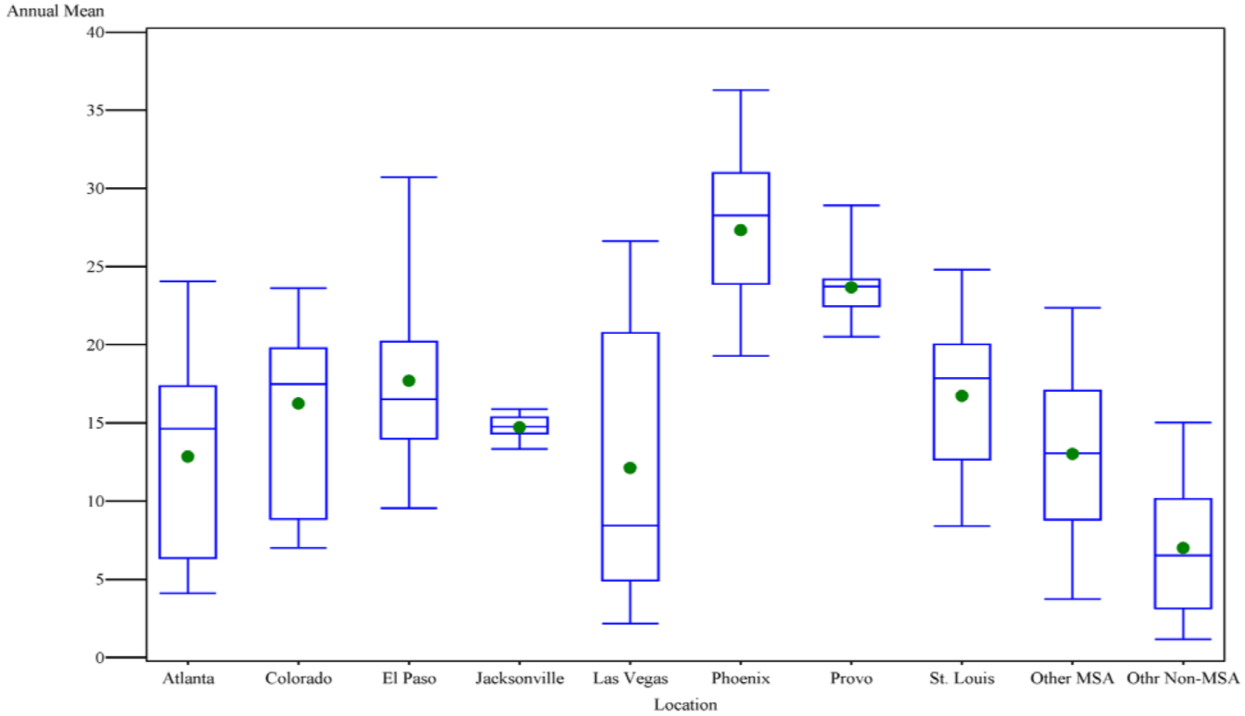


Figure 2. Spatial distributions of annual mean NO_2 ambient monitoring concentrations for selected MSA and grouped locations, years 1995-2006.

Spatial differences in hourly concentrations were of course consistent with that observed for the annual mean concentrations, and as expected there were differences in the COVs across locations, ranging from about 60 to 120%. However, in comparing the 90 percent intervals (from the 5th to the 95th percentiles) of hourly concentrations across locations, the ranges are somewhat similar (for example see Figure 3 for the CMSA locations). This means that the intervals for the annual mean differ more than that of the hourly concentrations between locations likely due to the influence of high 1-hour NO_2 concentrations for certain locations.

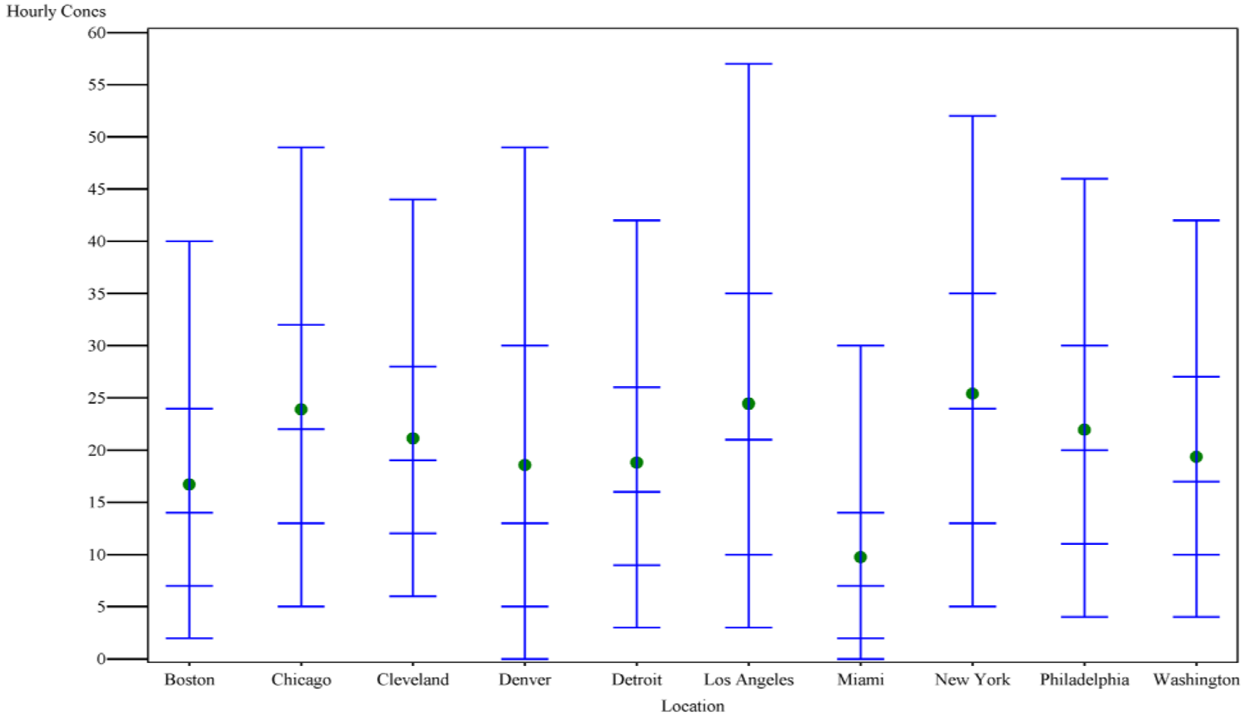
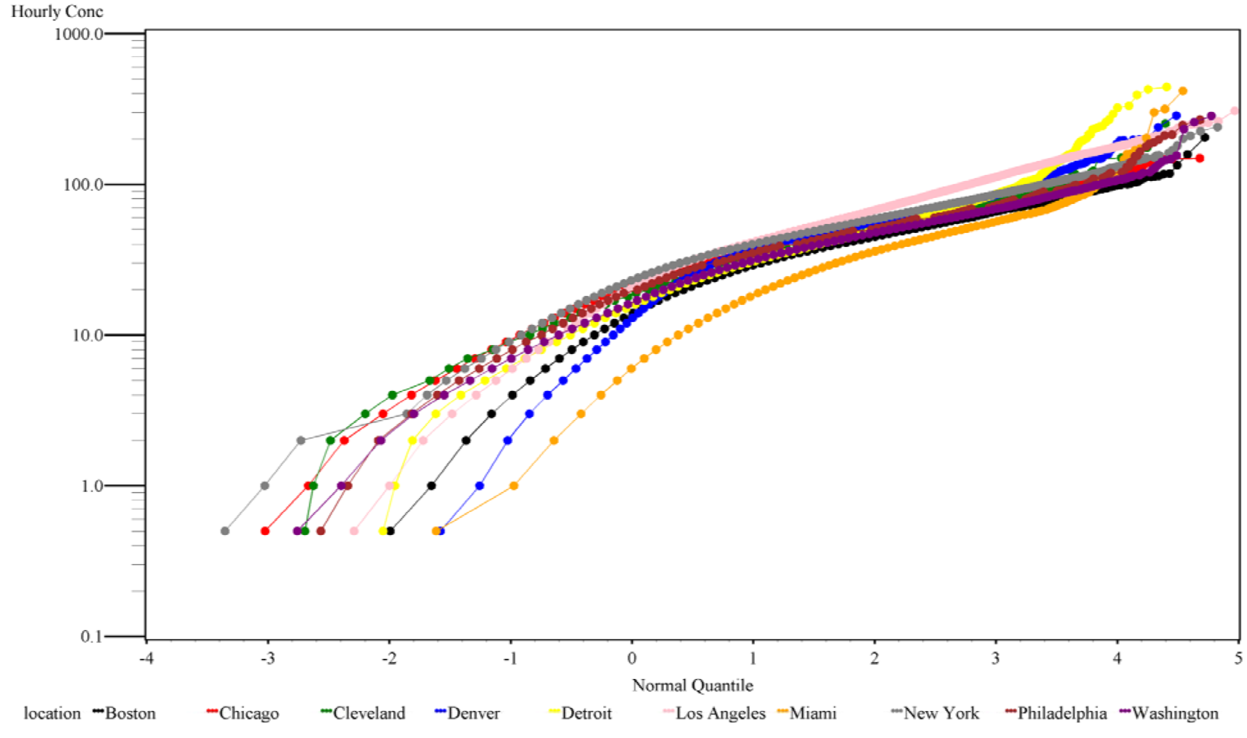


Figure 3. Spatial distributions[3] of hourly NO_2 ambient monitoring concentrations for selected CMSA locations, years 1995-2006.

This presence of extreme NO_2 concentrations is best illustrated in Figure 4 using a Q-Q plot that captures the full concentration distribution for each CMSA location. The Q-Q plots are generally curved rather than straight, such that the distributions do not appear to be log-normal. However, the annual mean and hourly concentration curves do tend to be approximately straight and parallel for values above the median (normal quantile = 0) through the 3rd quantile, suggesting that these upper tails of the distributions are approximately log-normal with approximately the same coefficients of variation. Beyond the 3rd quantile though, each distribution similarly and distinctly curves upwards, indicating a number of uncharacteristic NO_2 concentrations at each location when compared with the rest of their respective concentration distributions.

---

[3] The boxplots for hourly concentrations were created using a different procedure than for the annual statistics, because of the large number of hourly values and the inability of the graphing procedure to allow frequency weights. Therefore, the appropriate weighted percentiles and means were calculated and plotted as shown, but the vertical lines composing the sides of the box were omitted.



Zero values were replaced by 0.5

Figure 4. Spatial distributions of hourly NO_2 ambient concentration for selected CMSA locations, years 1995-2006.

Distributions of each variable (annual means and hourly concentrations) were compared between the different locations using statistical tests. The results in Table 7 show statistically significant differences between locations for both variables and all three summary statistics (means, medians, and scales). This supports the previous observation that the distributions for the different locations are dissimilar.

Table 7. Statistical test results for spatial comparisons of all location parameter distributions.

Concentration Parameter	Means Comparison		Central Values Comparison		Scales Comparison	
	F Statistic	p-value	Kruskal-Wallis	p-value	Mood	p-value
Annual Mean	148	<0.0001	1519	<0.0001	729	<0.0001
Hourly	330272	<0.0001	5414056	<0.0001	1354075	<0.0001

The spatial distributions of NO_2 concentrations within locations were also evaluated. As an example, Figure 5 illustrates the distribution of the annual mean NO_2 concentration at 10 monitoring sites within Philadelphia. The mean annual means vary from a minimum of 14.8 ppb (site 1000310071) to a maximum of 30.5 ppb (site 4210100471). The range of within-site variability can be attributed to the number of monitoring years available coupled with the observed trends in temporal variability across the monitoring period (discussed below in Section 2.4.4).

Distributions of each variable (annual means and hourly NO_2 concentrations) within locations (i.e., site distributions) were compared using statistical tests. The results in Table 8 indicate statistically significant differences within locations for both variables and the central tendency statistics (means and medians), while scales were statistically significant for 38 out of 40 possible tests. This supports the previous observation that the distributions for the different locations are dissimilar.

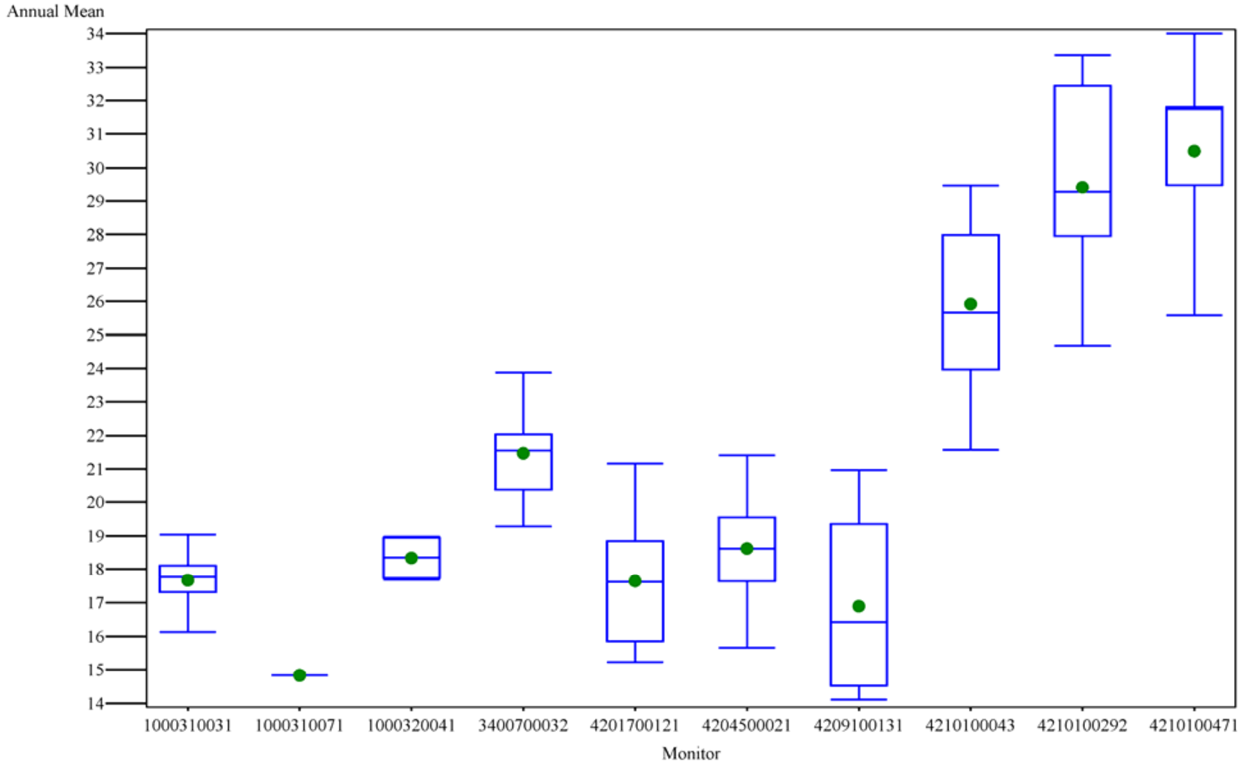


Figure 5. Spatial distribution of annual average NO_2 concentrations among 10 monitoring sites in Philadelphia CMSA, years 1995-2006.

Table 8. Statistical test results for spatial comparisons of within location parameter distributions.

Concentration Parameter	**Location**	**Means Comparison F Statistic**	**p-value**	**Central Values Comparison Kruskal-Wallis**	**p-value**	**Scales Comparison Mood**	**p-value**
Annual Mean	Boston	47.3	<0.001	96.5	<0.001	79.9	<0.001
	Chicago	123	<0.001	76.7	<0.001	68.5	<0.001
	Cleveland	12.1	<0.001	15.4	0.002	7.5	0.058
	Denver	85.3	<0.001	32.0	<0.001	23.0	0.001
	Detroit	13.2	<0.001	13.1	0.001	7.8	0.020
	Los Angeles	49.0	<0.001	325	<0.001	240	<0.001
	Miami	111	<0.001	36.2	<0.001	29.9	<0.001
	New York	106	<0.001	163	<0.001	151	<0.001
	Philadelphia	48.9	<0.001	68.8	<0.001	33.0	<0.001
	Washington DC	48.6	<0.001	104	<0.001	71.2	<0.001
	Atlanta	119	<0.001	45.2	<0.001	28.6	<0.001
	Colorado Springs	8.7	<0.001	18.8	0.009	8.7	0.273
	El Paso	36.0	<0.001	31.6	<0.001	35.3	<0.001
	Las Vegas	137	<0.001	45.4	<0.001	35.2	<0.001

Concentration Parameter	Location	Means Comparison		Central Values Comparison		Scales Comparison	
		F Statistic	p-value	Kruskal-Wallis	p-value	Mood	p-value
	Phoenix	20.4	<0.001	32.2	<0.001	23.6	0.001
	St. Louis	51.5	<0.001	82.1	<0.001	69.0	<0.001
	Other CMSA	82.5	<0.001	2152	<0.001	1934	<0.001
	Not MSA	76.9	<0.001	424	<0.001	372	<0.001
Hourly	Boston	17884	<0.001	312994	<0.001	59896	<0.001
	Chicago	11611	<0.001	142034	<0.001	37224	<0.001
	Cleveland	4191	<0.001	14102	<0.001	1985	<0.001
	Denver	25130	<0.001	104800	<0.001	2864	<0.001
	Detroit	4125	<0.001	10442	<0.001	424	<0.001
	Los Angeles	27288	<0.001	1050310	<0.001	269190	<0.001
	Miami	10669	<0.001	68580	<0.001	43090	<0.001
	New York	20052	<0.001	404234	<0.001	91104	<0.001
	Philadelphia	13759	<0.001	112129	<0.001	4903	<0.001
	Washington	14262	<0.001	223040	<0.001	30974	<0.001
	Atlanta	35917	<0.001	137022	<0.001	17330	<0.001
	Colorado Springs	5541	<0.001	48252	<0.001	3921	<0.001
	El Paso	10503	<0.001	57694	<0.001	18334	<0.001
	Las Vegas	22567	<0.001	136455	<0.001	28972	<0.001
	Phoenix	5626	<0.001	35645	<0.001	6747	<0.001
	St. Louis	14807	<0.001	178180	<0.001	47842	<0.001
	Other CMSA	19557	<0.001	6306431	<0.001	2164452	<0.001
	Not MSA	17630	<0.001	1580139	<0.001	491390	<0.001

2.4.4 Temporal Results

A broad view of the trend of NO_2 monitoring concentrations over time is presented in Figure 6. The annual mean concentrations were calculated for each monitor site within each year to create a distribution of annual mean concentrations for each year. The distribution of annual mean concentrations generally decreases with each increasing year. On average, mean annual mean NO_2 concentrations consistently decrease from a high of 17.5 ppb in 1995 to the most recent mean of 12.3 ppb. Also notable is the consistent pattern in the decreasing concentrations across each years distribution, the shape of each curve is similar indicating that while concentrations have declined, the variability within each year is similar from year to year. The variability within a given year is representing spatial differences in annual average concentrations across the 20 locations.

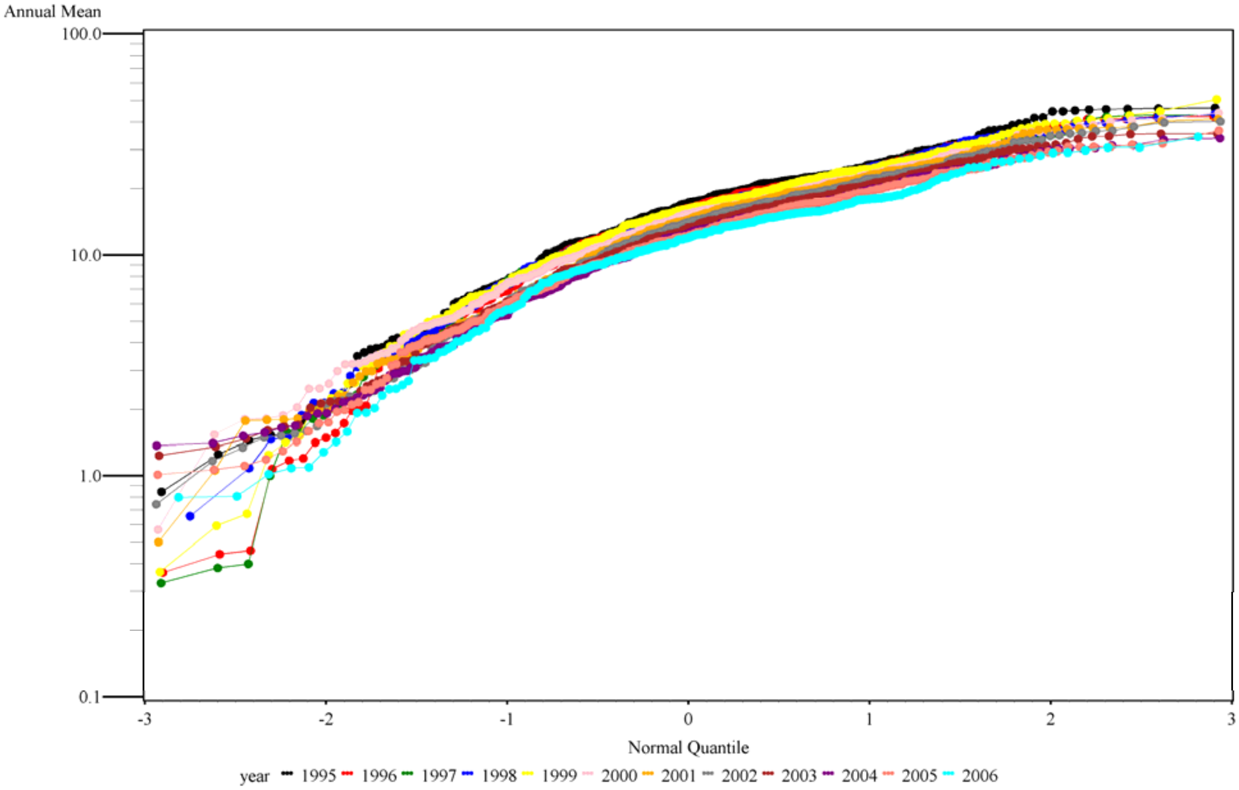


Figure 6. Temporal distributions of annual mean NO_2 concentrations for all monitors, years 1995-2006.

In general, temporal trends within a location were also consistent with the trends observed in all monitors, particularly where the location's monitoring network was comprised of several monitors. For example, Figure 7 illustrates the temporal distributions of annual average NO_2 concentration in the Philadelphia CMSA, each comprising between 4 and 8 monitors in operation per year. Clearly NO_2 concentrations are decreased with increasing calendar year of monitoring with the lowest NO_2 concentrations in the more recent years of monitoring. The pattern of variability in NO_2 concentration within a year at this location is also similar when comparing across years based on similarities in the shape of each years respective curve.

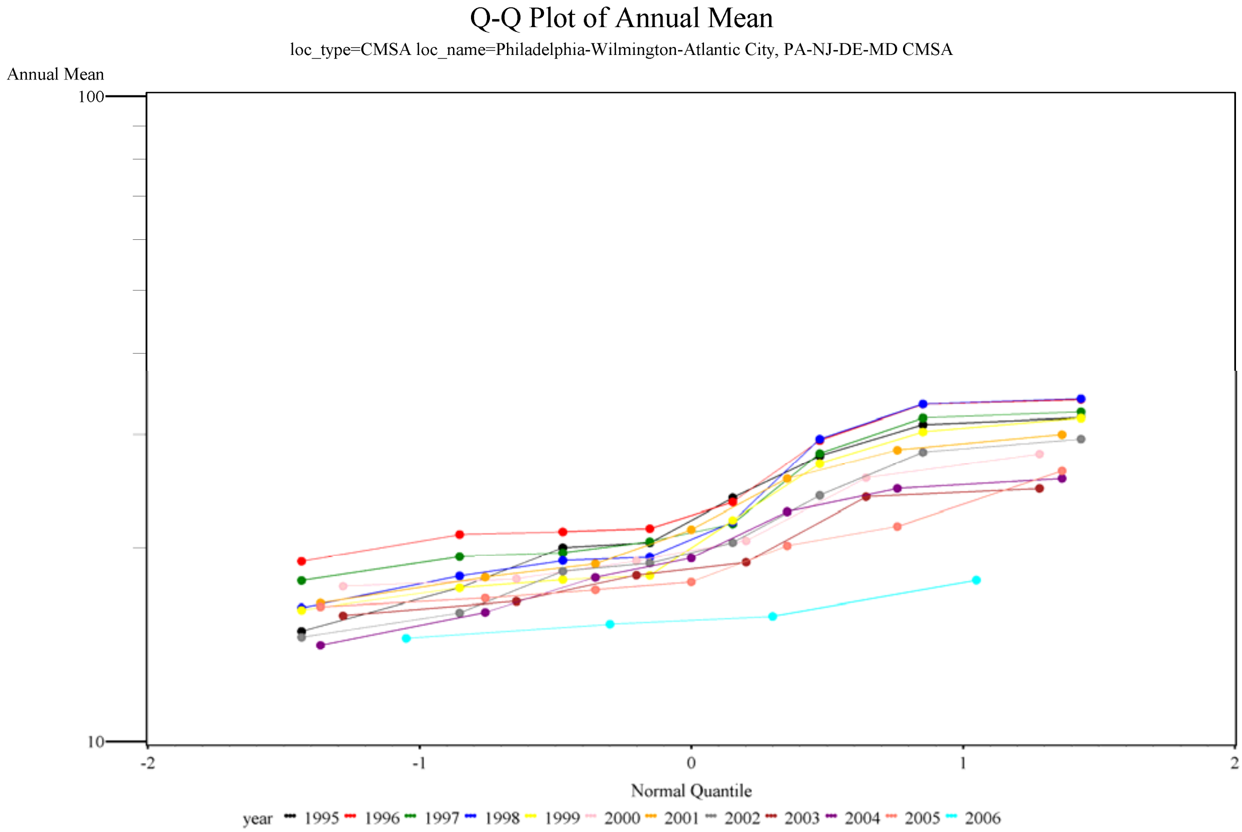
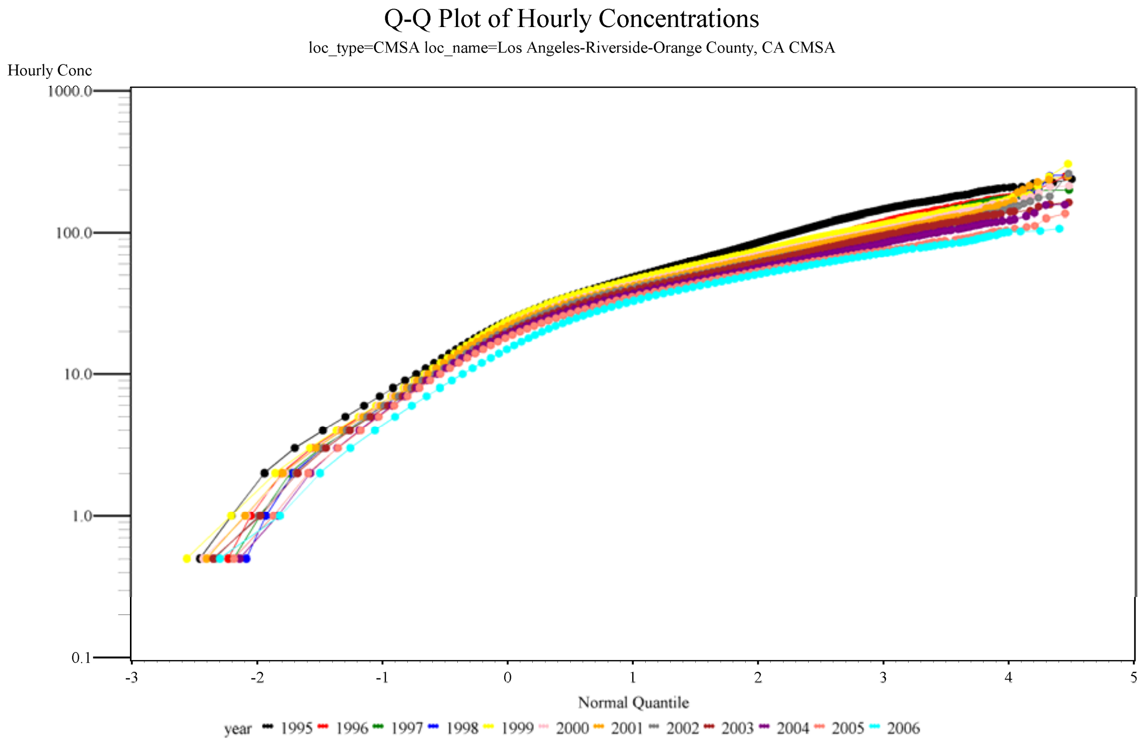


Figure 7. Temporal distributions of annual mean NO_2 concentrations for the Philadelphia CMSA, years 1995-2006.

In general, temporal trends within a location considering the hourly concentration data were consistent with the above, particularly where the monitoring network was comprised of several monitors. For example, Figure 8 illustrates the temporal distribution for hourly NO_2 concentration in the Los Angeles CMSA, comprising between 26 and 36 monitors in operation per year. NO_2 concentrations are decreased with increasing calendar year of monitoring with the distribution of hourly concentrations lowest in the more recent years of monitoring. The pattern of variability in NO_2 concentration within a year at this location is also similar when comparing across years based on similarities in the shape of each years respective curve.



Zero values were replaced by 0.5

Figure 8. Temporal distribution of hourly NO_2 concentrations in the Los Angeles CMSA, years 1995-2006.

These temporal trends were confirmed by statistical comparison tests. The means and medians of the annual means and hourly concentrations compared across the different years were statistically significant (Table 9). The Mood test shows that for the annual means, the scales were also significantly different. Note, however, that the Mood test derivation assumes that the medians of the annual means are the same for each year, whereas the plots and the Kruskall-Wallis test result implies that the medians are not the same. As noted before, Figure 8 indicates that the Q-Q curves for different years have similar slopes but different intercepts, which implies that the annual means for different years have different mean values but similar coefficients of variation. In fact the coefficients of variation of the annual means are nearly identical for different years, ranging from 52 % to 55 %.

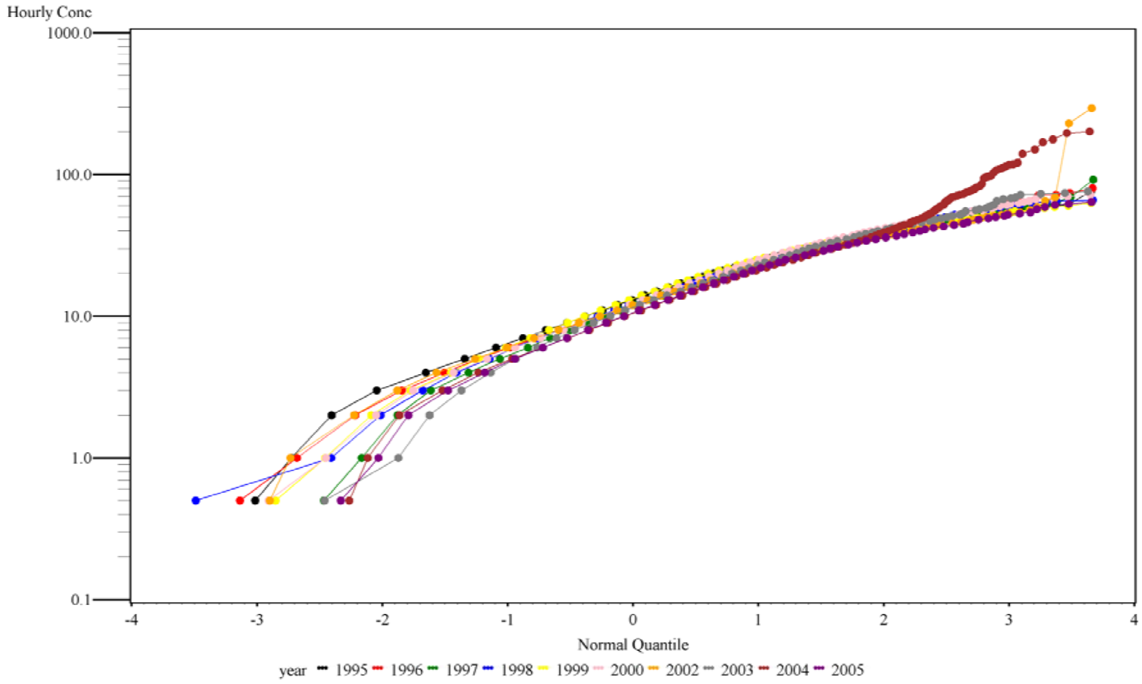
Table 9. Statistical test results for temporal comparisons of all location parameter distributions.

Concentration Parameter	Means Comparison		Central Values Comparison		Scales Comparison	
	F Statistic	p-value	Kruskal-Wallis	p-value	Mood	p-value
Annual Mean	15.0	<0.0001	146	<0.0001	32.5	0.0006
Hourly	47432	<0.0001	494826	<0.0001	24238	<0.0001

There were some exceptions to this temporal trend, particularly when considering the distribution of hourly concentrations and where a given location had only few monitors per year. Using Jacksonville as an example, Figure 9 illustrates the same temporal trend in NO_2 concentrations as was observed above for much of the distribution, however distinctions are noted at the upper tails of the distribution for two years of data, 2002 and 2004. For Jacksonville, each years' hourly concentration distribution was based on only a single monitor.

Where few monitors exist in a given location, atypical variability in one or a few monitors from year to year can greatly influence the distribution of short-term concentrations, particularly at the upper percentiles.

The same follows for assignment of statistical significance to temporal trends within locations. While annual average concentrations are observed to have declined over time within a location, the number of sites were typically few thus limiting the power of the statistical tests. Only Los Angeles, El Paso, Phoenix, and Other CMSA were significant ($p<0.05$) for the central tendency tests, while only Los Angeles and Other CMSA were significant ($p<0.05$) for scale (data not shown). All hourly concentrations comparison tests for years within each location were significant for all three test statistics (mean, median, scale).



Zero values were replaced by 0.5

Figure 9. Temporal distributions of hourly NO_2 concentrations in the Jacksonville MSA, years 1995-2006, one monitor.

One final temporal trend worthy of mention is that associated with the Not MSA grouped location. There is very little difference in annual average concentrations across the 1995-2006 monitoring period. While percentage-wise the reduction in concentration is about 25%, on a concentration basis, this amounts to a maximum of about 2 ppb reduced over the 11 year period (Figure 10). In considering, the past 5 years, there was even less of a reduction in annual average concentration with about a 0.5 ppb difference between 2001 and 2006. This could indicate that many of these monitoring sites are receiving relatively less impact from local sources of NO_2 (e.g., emissions from major roads and stationary sources) compared with the other locations. Therefore, the areas that these monitors represent may also be less likely to see significant benefit by changes in source emissions and/or NO_2 standard levels compared with the named CMSA/MSA locations.

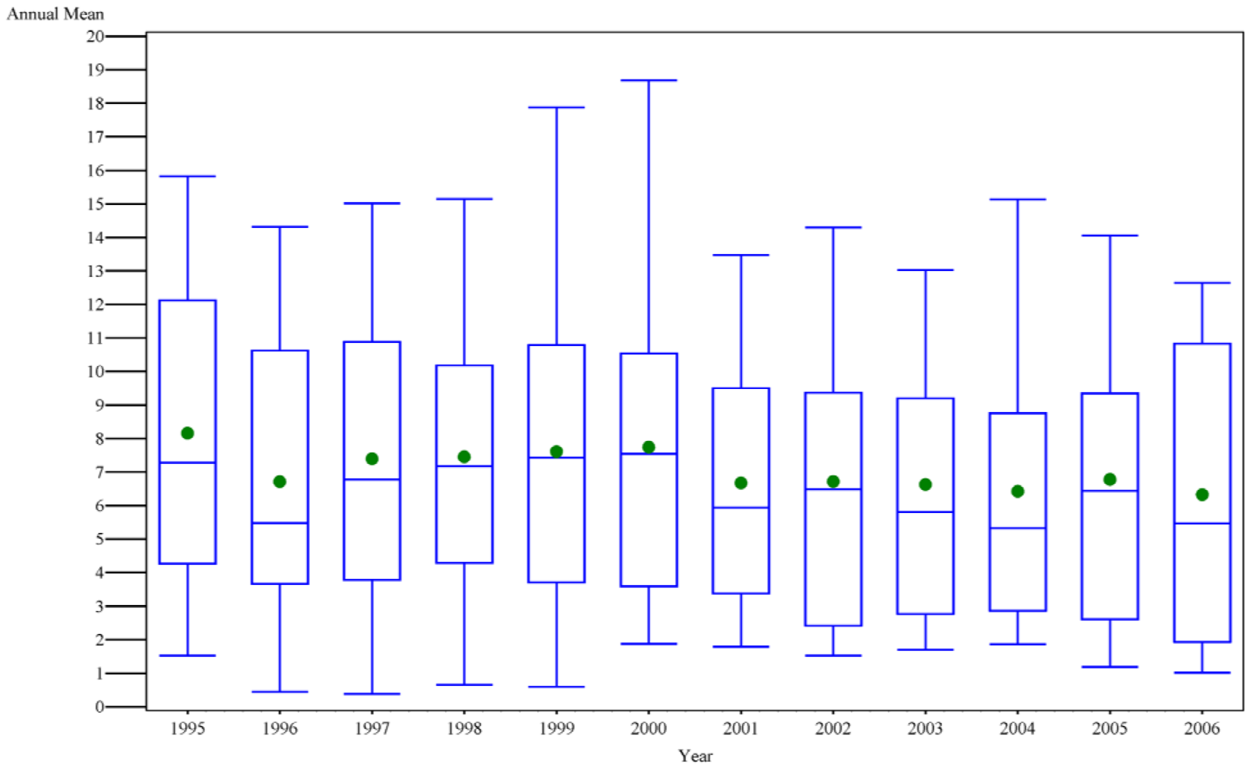


Figure 10. Temporal distribution of annual average NO_2 concentrations in the Not MSA group location, years 1995-2006.

2.5 Air Quality Simulation

2.5.1 Introduction

Every location across the U.S. meets the current NO_2 annual standard (US EPA, 2007e). Even considering air quality data as far back as 1995, no location/monitoring site exceeded the current standard. Therefore, simulation of air quality data was required to evaluate just meeting the current standard or standards that are more stringent.

In developing a simulation approach to adjust air quality to meet a particular standard level, policy-relevant background (PRB) levels in the U.S. were first considered. Policy-relevant background is defined as the distribution of NO_2 concentrations that would be observed in the U.S. in the absence of anthropogenic (man-made) emissions of NO_2 precursors in the U.S., Canada, and Mexico. Estimates of PRB have been reported in the draft ISA (Section 1.5.5) and the Annex (AX2.9), and for most of the continental U.S. the PRB is estimated to be less than 300 parts per trillion (ppt). In the Northeastern U.S. where present-day NO_2 concentrations are highest, this amounts to a contribution of about 1% percent of the total observed ambient NO_2 concentration (AX2.9). This low contribution of PRB to NO_2 concentrations provides support for a proportional method to adjust air quality, i.e., an equal adjustment of air quality values across the entire air quality distribution to just meet a target value.

Next, the variability in NO_2 concentrations was evaluated to determine whether a proportional approach would be reasonable if applied broadly across all years of data. Since the adjustment factor to meet the current standard would likely increase with increasing year, it was of interest to determine the trend in both the hourly concentrations and variability by year. Figure 11 presents a summary of the annual average and hourly mean concentrations, as well as the coefficient of variation (COV, standard deviation as a percent of the mean) for each respective mean. Sample size for the annual average concentrations was about 350 per year, while hourly concentrations numbered about 3 million per year.

As expected, there was no observed difference in the mean concentrations when comparing each concentration metric within a year. The mean of the annual averages of all monitors is nearly identical to the mean of the hourly concentrations. However, statistically significant decreases in concentration are evident from year-to-year ($p<0.0001$), with concentrations decreasing by about 30% across the monitoring period. Contrary to this, there is no apparent trend in the COV for the annual average concentrations across the 12 years of data, generally centered about 53%. The COV of the hourly concentrations is larger than the annual COV as expected, however it increases with increasing year. The hourly COV ranges from a low of 84% in 1998 to a high of 92% in 2006, amounting to a relative percent difference of only 10% across the entire monitoring period. A non-parametric Mann-Whitney U-test indicates that there is a significant difference in the COVs when comparing each year-group ($p=0.004$). This may result in a small upward bias in the number of estimated exceedances of short-term (1-hour) potential health benchmark levels if using a proportional roll-up on the more recent monitoring data relative to that estimated by rolling up the historic data to just meet the current standard. While the trend of increasing COV is apparent across the entire monitoring period, based on the limited difference in COV from year-to-year for both the annual and hourly concentration data within each year-group (each is <4%), it is concluded that a proportional method could be broadly applied.

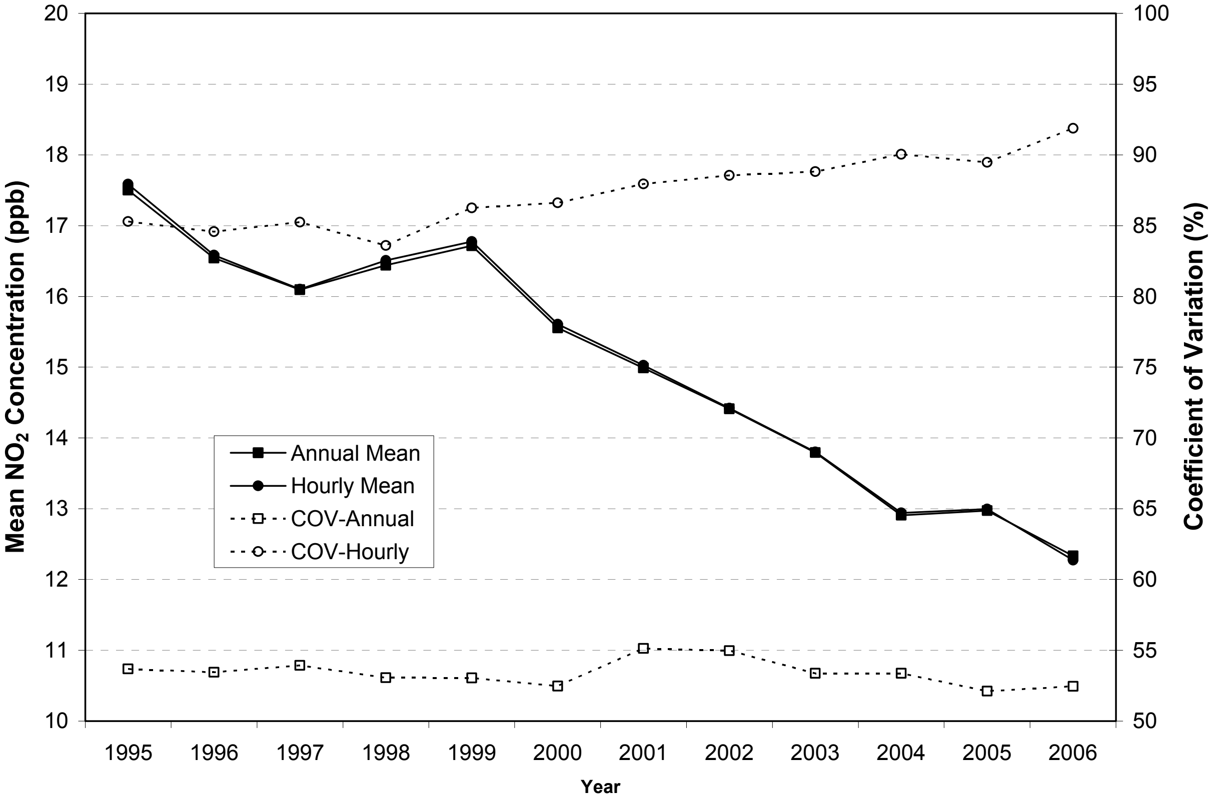


Figure 11. Trends in hourly and annual average NO_2 ambient monitoring concentrations and their associated coefficients of variation (COV) for all monitors, years 1995-2006.

2.5.2 Approach

For the air quality characterization, data were first separated into two groups, an historic set of monitoring data (1995-2000) and one containing the most recent air quality (2001-2006). This grouping would further reduce any potential influential monitoring data affecting the variability in hourly concentrations that may exist in one year to the next within a location. Typically, ambient concentrations are not adjusted higher to simulate just meeting alternative standards, therefore older historical data may be of use in better representing scenarios that are at or near the current NO_2 standard. To date, the following air quality scenarios have been considered:

- "*as is*" representing the historical and recent ambient monitoring hourly concentration data as reported by US EPA's Air Quality System (AQS).
- "*simulated*" concentrations to just meet the current NO_2 NAAQS (53 ppb annual average).

Based on the form of the standard and observed trends in ambient monitoring, such as the retention of similar hourly and annual COVs over time while annual average concentrations significantly decrease over the same time period, NO_2 concentrations were proportionally modified at each location using the maximum annual average concentration that occurred in each year. To just meet the current standard adjustment factors F for each location (i) and year (j) were derived by the following

$$F_{ij} = 53 / C_{\max,ij}$$

where,

F_{ij} = Adjustment factor (unitless)
$C_{max,ij}$ = Maximum annual average NO_2 concentration at a monitor in a location i (ppb)

Values for each air quality adjustment factor used for each location are given in Tables 10 and 11. It should be noted that a different monitor could have been used for each year to estimate F, the selection dependent only on whether the monitor contained the highest annual concentration for that year in the particular location. For each location and calendar year, all the hourly concentrations were multiplied by the same constant value F to make the highest annual mean equal to 53 ppb for that location and year. For example, for Boston in 1995, the maximum annual mean was 30.5 ppb, giving an adjustment factor of $F = 53/30.5 = 1.74$. All hourly concentrations in Boston in 1995 were multiplied by 1.74. Then, using the adjusted hourly concentrations, the distributions of the annual means and annual number of exceedances are computed in the same manner as the as-is scenario.[4]

---

[4] Because of the large database, we did not implement this procedure exactly as stated. For the annual means we computed and applied the adjustment factors directly to each annual mean. For the hourly concentrations we used the frequency distributions of the rounded hourly values, so that, in effect, we applied the adjustment factors to the hourly values after rounding them to the nearest integer. This has a negligible impact on the calculated number of exceedances.

Table 10. Maximum annual average NO$_2$ concentrations and air quality adjustment factors (*F*) to just meet the current standard, historic monitoring data.

Location	Metric	1995	1996	1997	1998	1999	2000
Boston	Max Annual Mean	30.5	31.0	30.4	30.7	29.7	29.0
	F	1.74	1.71	1.74	1.73	1.79	1.83
Chicago	Max Annual Mean	32.2	32.0	33.6	32.2	31.5	32.0
	F	1.64	1.66	1.58	1.64	1.68	1.66
Cleveland	Max Annual Mean	27.3	25.9	28.1	27.3	24.5	23.1
	F	1.94	2.04	1.89	1.94	2.16	2.30
Denver	Max Annual Mean	34.8	33.1	33.9	35.3	19.4	14.9
	F	1.52	1.60	1.56	1.50	2.73	3.55
Detroit	Max Annual Mean	21.6	21.5	25.9	22.9	18.0	23.9
	F	2.45	2.47	2.05	2.31	2.94	2.22
Los Angeles	Max Annual Mean	46.2	42.3	43.2	43.4	50.6	43.9
	F	1.15	1.25	1.23	1.22	1.05	1.21
Miami	Max Annual Mean	14.7	16.0	16.6	15.2	16.8	15.7
	F	3.60	3.30	3.19	3.49	3.15	3.37
New York	Max Annual Mean	41.7	42.2	41.1	41.9	41.5	40.6
	F	1.27	1.26	1.29	1.26	1.28	1.31
Philadelphia	Max Annual Mean	31.8	33.9	32.4	34.0	31.7	27.9
	F	1.67	1.56	1.63	1.56	1.67	1.90
Washington DC	Max Annual Mean	26.2	26.9	25.9	27.2	25.4	23.5
	F	2.02	1.97	2.05	1.95	2.09	2.26
Atlanta	Max Annual Mean	18.8	26.6	25.2	24.1	23.8	22.9
	F	2.81	1.99	2.10	2.20	2.22	2.31
Colorado Springs	Max Annual Mean	23.2	23.6	19.8	20.5	19.3	34.8
	F	2.28	2.24	2.68	2.59	2.75	1.52
El Paso	Max Annual Mean	23.3	35.1	33.6	30.7	27.7	24.3
	F	2.27	1.51	1.58	1.72	1.91	2.18
Jacksonville	Max Annual Mean	15.8	14.9	14.4	15.0	15.9	15.4
	F	3.36	3.55	3.69	3.52	3.34	3.45
Las Vegas	Max Annual Mean	27.1	26.7		25.3	26.6	25.1
	F	1.96	1.99		2.09	1.99	2.12
Phoenix	Max Annual Mean	32.6	31.6	32.0	35.0	40.5	36.3
	F	1.63	1.68	1.66	1.52	1.31	1.46
Provo	Max Annual Mean	22.6	24.3	23.3	23.9	24.1	23.6
	F	2.35	2.18	2.27	2.22	2.20	2.25
St. Louis	Max Annual Mean	26.2	24.8	24.8	25.8	27.2	26.3
	F	2.02	2.14	2.14	2.05	1.95	2.02
Other CMSA	Max Annual Mean	31.9	30.3	29.4	31.0	29.3	26.5
	F	1.66	1.75	1.80	1.71	1.81	2.00
Not MSA	Max Annual Mean	19.1	14.5	19.7	18.8	19.7	18.7
	F	2.78	3.66	2.69	2.82	2.69	2.83

Table 11. Maximum annual average NO_2 concentrations and air quality adjustment factors (*F*) to just meet the current standard, recent monitoring data.

Location	Metric	2001	2002	2003	2004	2005	2006
Boston	Max Annual Mean	29.7	25.3	22.5	25.0	23.4	22.5
	F	1.79	2.10	2.36	2.12	2.26	2.35
Chicago	Max Annual Mean	31.9	32.4	30.9	29.3	29.6	30.6
	F	1.66	1.63	1.72	1.81	1.79	1.73
Cleveland	Max Annual Mean	23.6	22.3	21.7	22.2	21.5	18.2
	F	2.25	2.38	2.45	2.38	2.46	2.91
Denver	Max Annual Mean	36.8	35.4	21.4	27.2	27.6	29.1
	F	1.44	1.50	2.47	1.95	1.92	1.82
Detroit	Max Annual Mean	23.2	21.4	22.0	18.9	19.6	15.9
	F	2.29	2.47	2.41	2.80	2.71	3.34
Los Angeles	Max Annual Mean	41.2	40.2	35.3	33.7	30.9	29.7
	F	1.29	1.32	1.50	1.57	1.72	1.78
Miami	Max Annual Mean	15.8	14.3	12.9	13.0	13.5	
	F	3.35	3.71	4.12	4.08	3.92	
New York	Max Annual Mean	40.3	39.7	32.0	30.5	36.5	34.2
	F	1.32	1.33	1.65	1.74	1.45	1.55
Philadelphia	Max Annual Mean	29.9	29.5	24.7	25.6	26.3	17.8
	F	1.77	1.80	2.15	2.07	2.02	2.98
Washington DC	Max Annual Mean	24.3	24.8	26.0	24.0	24.1	19.6
	F	2.18	2.14	2.04	2.20	2.20	2.70
Atlanta	Max Annual Mean	23.3	19.4	16.4	17.0	17.4	17.9
	F	2.27	2.73	3.23	3.12	3.05	2.96
Colorado Springs	Max Annual Mean						
	F						
El Paso	Max Annual Mean	21.7	21.4	19.9	18.0	17.3	18.0
	F	2.45	2.48	2.66	2.94	3.06	2.94
Jacksonville	Max Annual Mean		14.6	14.3	13.7	13.3	
	F		3.62	3.70	3.88	3.97	
Las Vegas	Max Annual Mean	22.5	22.3	21.4	19.7	19.9	
	F	2.35	2.38	2.48	2.69	2.67	
Phoenix	Max Annual Mean	37.1	34.7	34.3	31.4	31.5	30.6
	F	1.43	1.53	1.54	1.69	1.68	1.73
Provo	Max Annual Mean	24.1	24.8	21.8	22.3	20.5	28.9
	F	2.20	2.14	2.43	2.37	2.58	1.83
St. Louis	Max Annual Mean	24.7	22.9	20.3	22.3	16.8	15.0
	F	2.15	2.32	2.60	2.37	3.15	3.52
Other CMSA	Max Annual Mean	26.5	27.4	26.4	25.3	24.0	18.5
	F	2.00	1.93	2.01	2.09	2.21	2.87
Not MSA	Max Annual Mean	16.5	16.4	15.5	15.8	17.1	15.6
	F	3.21	3.23	3.42	3.36	3.11	3.39

2.6 Method for Estimating On-Road Concentrations

2.6.1 Introduction

As an additional step in the air quality characterization, the potential impact of motor vehicles on the surrogate exposure metrics was evaluated. Several studies have shown that concentrations of NO_2 are at elevated levels when compared to ambient concentrations measured at a distance from the roadway (e.g., Rodes and Holland, 1981; Gilbert et al., 2003; Cape et al., 2004; Pleijel et al., 2004; Singer et al., 2004). On average, concentrations on or near a roadway are from 1.5 to 2 times greater than ambient concentrations (US EPA, 2007f), but on occasion, as high as 7 times greater (Bell and Ashenden, 1997; Bignal et al., 2007). A strong relationship between measured on-road NO_2 concentrations and those with increasing distance from the road has been reported under a variety of conditions (e.g., variable traffic counts, different seasons, wind direction) and can be described (e.g., Cape et al., 2004) with an exponential decay equation of the form

$$C_x = C_b + C_v e^{-kx} \qquad \text{eq (1)}$$

where,

C_x = NO_2 concentration at a given distance (x) from a roadway (ppb)
C_b = NO_2 concentration (ppb) at a distance from a roadway, not directly influenced by road or non-road source emissions
C_v = NO_2 concentration contribution from vehicles on a roadway (ppb)
k = Rate constant describing NO_2 combined formation/decay with perpendicular distance from roadway (meters^{-1})
x = Distance from roadway (meters)

As a function of reported concentration measures and the derived relationship, much of the decline in NO_2 concentrations with distance from the road has been shown to occur within the first few meters (approximately 90% within 10 meter distance), returning to near ambient levels between 200 to 500 meters (Rodes and Holland, 1981; Bell and Ashenden, 1997; Gilbert et al., 2003; Pleijel et al., 2004). At a distance of 0 meters, referred to here as *on-road*, the equation reduces to the sum of the non-source influenced NO_2 concentration and the concentration contribution expected from vehicle emissions on the roadway using

$$C_r = C_a (1 + m) \qquad \text{eq (2)}$$

where,

C_r = 1-hour on-road NO_2 concentration (ppb)
C_a = 1-hour ambient monitoring NO_2 concentration (ppb) either *as is* or modified to just meet the current standard
m = Modification factor derived from estimates of C_v/C_b (from eq (1))

and assuming that $C_a = C_b$.[5]

2.6.2 Derivation of On-Road Factors

A literature review was conducted to identify published studies containing NO_2 concentrations both on-roads and with various distances from roadways. Principal criteria for inclusion in this analysis were that either tabular, graphical, or equations were provided in the paper that related distances from roadways and associated NO_2 concentrations. Eleven papers were identified using these criteria, spanning several countries, various time periods, roadway locations, seasons, and wind direction (Table 12). The final data set contained 501 data points, encompassing multiple NO_2 measurements from a total of 56 individual roads.

Table 12. Reviewed studies containing NO_2 concentrations at a distance from roadways.

First Author	Year	Country/State	Season	Type	Wind Direction
Bell	1987	Wales	Summer, winter	Rural	Up, down
Bignal	2004	England	Summer, fall	Urban	Combined
Cape	2002	Scotland	Annual	Urban	Combined
Gilbert	2001	Quebec	Summer	Urban	Down, up, combined
Maruo	2001	Japan	Summer	Urban	Combined
Monn	1995	Switzerland	Summer, Winter	Urban	Combined
Nitta	1982	Japan	Not reported	Urban	Combined
Pleijel	1994	Sweden	Summer	Rural	Combined
Rodes	1978	California	Summer	Urban	Down
Roorda-Knape	1995	Holland	Summer	Urban	Combined
Singer	2001	California	Spring through fall	Urban	Up, Down

Although there were, on occasion, several roads within a particular study, data for factors thought to influence on-road concentrations were very limited or were not distinct for all studies. The relationship noted in eq (1) was solved using the data collected from the above reviewed literature and employing the SAS procedure ***proc nlin***, generally as follows,

```
proc nlin data=no2 maxiter=1000 noprint NOITPRINT;
   parms    Cb=0 to 80 by 1
            Cv= 0 to 80 by 1
            k= 0 to 1 by .025;
   model Cr=Cb + Cv*exp(-k*x);
   by author road season wind;
   output out=outdata parms=Cb Cv k;
run;
```

The procedure was run for all individual roads identified within each study location. Results of this analysis were screened for data that yielded no unique solutions (lack of model convergence) or irrational parameters. Criteria for censoring data included the following, as well as the number of individual roads censored:

- Model did not converge (n=5)

---

[5] Note that C_a differs from C_b since C_a may include the influence of on-road as well as non-road sources. However, it is expected that for most monitors the influence of on-road emissions is minimal so that $C_a \cong C_b$.

- $k<0$ (n=1)
- $k>1$ (n=2)
- Both $k=0$ and $C_v=0$ (n=1)
- Extremely large C_v (>8,000 ppb; n=2)
- $C_b<0$ (n=1)

Data were evaluated for trends using available influential factors and considering the number of samples available for potential groupings. In general, the measurements reported in the summer and resultant parameter estimates were observed as distinct from the measures and parameter estimates from other seasons. The data were then grouped accordingly into two seasonal groups, *summer* and *not summer*, containing 23 and 21 samples, respectively. These two groups were also censored for any unusual parameter estimates. Resulting criteria for censoring the grouped data included the following:

- Extreme value of *k* compared with others in group (n=1)
- Extreme values of estimated *m* due to combined low estimated C_b relative to high estimated C_v (n=2)

Two approaches were considered for estimating *m* from the C_v and C_b pairs in each season. The first approach was to regress C_b on C_v (either with or without an intercept) and use the fitted slope to estimate *m*. Ignoring meteorological effects, Equation 1 implies that C_v results solely from on-road emission sources and that C_b results solely from non-road emission sources. Since these two source types are likely to have quite different diurnal profiles, we expect the hourly C_v and C_b values to be approximately independent.[6] Regressing C_b against C_v would imply that there is some correlation between the values, which would be inconsistent with the conceptual model underlying Equation 1. Further, if C_b were regressed against C_v using an intercept, the physical meaning of the intercept would be unclear.

An empirical method was selected for the approach to estimate *m* based on the two seasonal sets of ratios of C_v/C_b. The resulting distribution for each group is presented in Figure 12. Neither group could be assigned to a particular distribution (e.g., normal, lognormal, exponential, gamma). Means from the two seasons were tested for significant difference using a Student's t (p=0.026), while the season distributions were compared using a Kolmogorov-Smirnov test (p= 0.196). It was decided to retain the groups as separate to allow for some apportioning of variability resulting from an apparent seasonal influence, even though the statistical test results were mixed.

---

[6] Although the fact that C_v and C_b are subject to the same meteorology introduces some correlation, because meteorology tends to vary on a longer time scale than hourly, it is likely to have less influence than the emissions on the correlation between hourly concentrations.

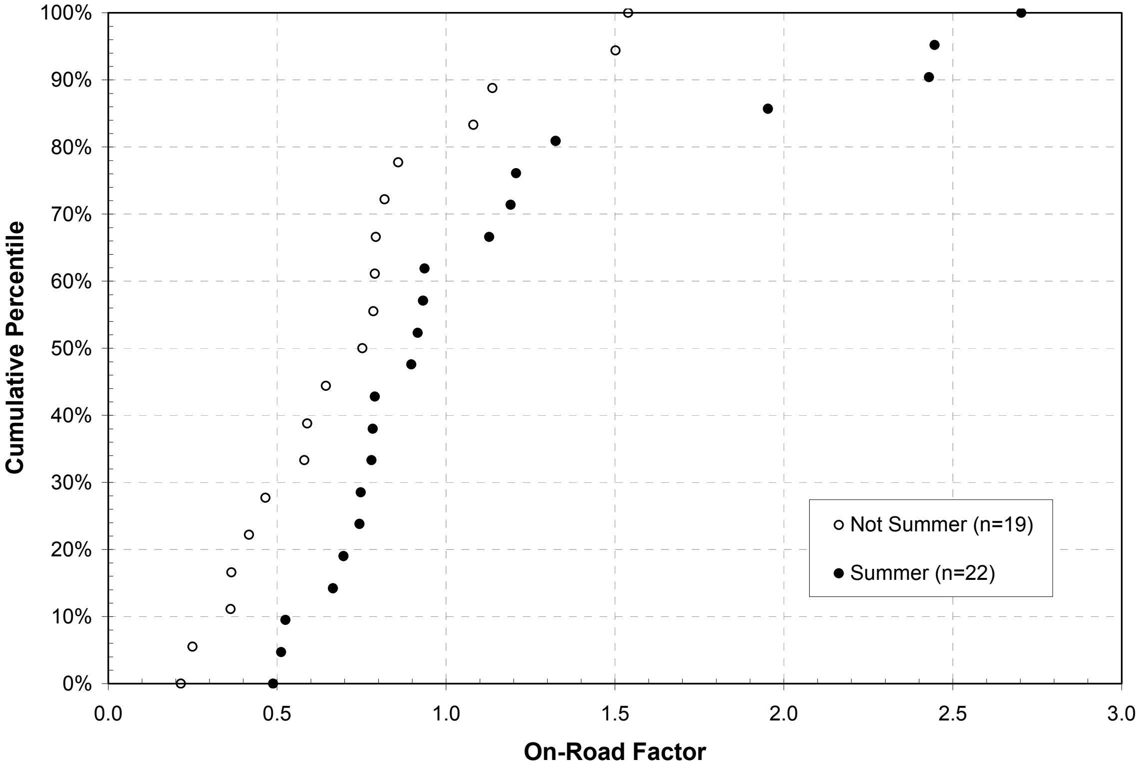


Figure 12. Distribution of on-road factors (C_v/C_b or *m*) for two season groups.

2.6.3 Application of On-Road Factors

The purpose of this particular analysis was to estimate on-road concentrations using eq (2) above along with the required inputs, namely, the hourly ambient monitoring concentrations and derived on-road factors. The derived on-road factors for the two season groups could not be assigned a particular statistical distribution (e.g., normal, lognormal, gamma) with confidence. Therefore, an empirical approach was selected to still allow for some seasonal variability in the on-road concentration estimates. Summer months were first defined as *June*, *July*, *August*, while the remaining months were not summer. Although there may be distinctions among what may be designated as a summer month across the U.S., the reviewed data are not robust to allow for such an application.

Each monitor site was then randomly assigned two on-road factors selected from the derived empirical distribution for a given year, one for summer months and one for the other months, using the appropriate distribution. Because the influence of on-road and non-road sources is likely different in each location and at each monitor, it would be expected that the empirical relationship between the two values C_v and C_b to vary from place to place. If source category emissions data for each study location were available to derive an eq (1) regression, that could have been used to match each of the study locations here, or, perhaps, each of the monitoring sites, to a similar eq (1) study area for assigning an appropriate ratio. However, since this information was not available, an empirical approach was used to randomly match the literature-derived ratios to the NO_2 site-seasons.

A particular *summer* on-road factor has a 1/22 chance of selection, while a specific *not summer* value has a 1/19 probability of selection, based on respective sample sizes. This random assignment was repeated for all site-years of data. Hourly NO_2 concentrations were estimated for each site-year of data in a location using eq (2) and the randomly assigned on-road factors. Finally, the process was simulated 100 times for each site-year of hourly data. For example, the Boston CMSA location had 210 random selections from the on-road distributions applied independently to the total site-years of data (105). Following 100 simulations, a total of 10,500 site-years of data were generated using this procedure (along with 21,000 randomly assigned on-road values selected from the appropriate empirical distribution).

Simulated on-road NO_2 concentrations were used to generate concentration distributions for the annual average concentrations and distributions for the number of exceedances of short-term potential health effect benchmark levels. Means and median values are reported to represent the central tendency of each parameter estimate. Since there were multiple simulations performed at each location using all available site-years of data, results for the upper percentiles were expanded to the 95th, 98th and 99th percentiles of the distribution, rather than estimate a 95% interval as was done above for the non-road scenarios. It is more appropriate to apply the parameter estimates outside the central tendencies to particular sites, areas within locations, or for certain conditions. Minimum values for the annual mean and annual number of exceedances were also estimated. One approach would have been to use the minimum values across the 100 simulations. However, that approach may not give the lowest possible value, because it is unlikely that in 100 simulations for a site-year there is a simulation where both seasonal adjustment factors are chosen to be the lowest values of $1 + m$. To obtain the lowest value, two simulations were conducted for each site-year. The Summer seasonal adjustment factor was set to the lowest possible value (1.49) and the Not-Summer seasonal adjustment factor was the lowest possible value (1.22). The annual means and exceedances for those two separate simulations were used to compute the minimum values for each distribution.

In addition and as part of the air quality characterization, the approach described in Section 2.7 below was used to estimate the number of short-term concentrations above selected levels that might occur on roadways using the estimated hourly C_r values, associated with air quality as is. For evaluating just meeting the current annual standard the approach described in Section 2.5 was used before estimating on-road concentrations.

2.6.4 Interpretation of Estimated On-Road Concentrations

The simulated on-road concentrations, simply put, are estimates. The algorithm is not designed to estimate concentrations on a particular roadway, all roads, or to estimate on-road exposures in a location. The algorithm assumes that the monitor is measuring the concentrations that would be observed at a distance (>200m or so) of a particular road (could be any road type). It then follows that the monitors within a location are linked proportionally to the distribution of roads (and types) in a location. This is likely not the case, particularly in locations with few monitoring sites, therefore available monitors will likely be either over- or under-representative of some roadway types.

The simulation is designed to estimate the potential concentrations associated with potential on-road exposures, developing central tendencies and bounds to be interpreted qualitatively with the expected emissions that would occur on-roads within a location. That is, the higher-traveled roadways would be better represented by on-road concentration estimates at the upper tails of the distribution, while other roads with less traffic density would be better represented at the lower tails of the distribution. Additional consideration should be given to where few monitor sites were available in a location, or even where monitor sites are more densely distributed within a particular area of a location, before interpreting estimated concentrations.

2.7 Estimation of Potential Health Effect Benchmark Exceedances

2.7.1 Introduction

A principal goal of the exposure assessment was to develop a model that estimates the frequency of high short-term exposures, considering just meeting the current standard and any alternative standards under consideration. Since the current standard is on an annual average basis, the relationship between that NO_2 concentration and short-term NO_2 concentrations needed to be evaluated. As part of the prior review, McCurdy (1994) used a non-linear regression (i.e., exponential) to describe the relationship between annual average concentrations and occurrence of short-term peak concentrations at two locations (i.e., one for Los Angeles and one for all other locations combined). At the time of the McCurdy (1994) analysis, there were at least a few monitors with reported annual average concentrations at or above the current standard for the Los Angeles analysis. The non-linear model was applied to estimate number of exceedances given selected annual average concentrations, and reasonably estimated the average number of exceedances at selected annual average concentration levels.

The same type of regression model was explored as a first step in this analysis as well as evaluating the feasibility of other models (i.e., a logistic regression, and another assuming a possion distribution) using air quality monitoring data from 1995-2006 (see Appendix D). Each of these models were developed for each location and applied to estimate the number of potential health effect benchmark exceedances at various annual average concentration levels. Following the construction of the models, a few issues with the approach became evident. Because of the limited number of exceedances above 200 ppb in most locations, the best models could only be developed at concentrations lower than this level. Second, some of the locations yielded inadequate models (e.g., non-convergence) that led to a regrouping of the original 20 locations identified above. Third, the predictability of the developed relationships using the varied regression approaches was questionable. Consistently, predictions above the observed maximum annual average concentrations were orders of magnitude higher than the maximum observed number of exceedances. The same occurred with the McCurdy (1994) analysis, though to a lesser degree, since there were at least a few site-years with concentrations at the annual standard. However, upper bound estimations in that 1994 analysis needed to be stunted once predictions were made for concentrations outside of the range of the measured data. Confidence intervals in each both the McCurdy (1994) analyses and those generated as part of this analysis were extremely large.

It is due to these issues surrounding the applicability of these statistical models to the current ambient monitoring data given the reduced annual average concentrations that a new approach was developed and applied. An empirical model was employed to avoid the difficulties in extrapolating outside the range of the data, combined with the concentration roll-up procedure described in Section 2.5 to estimate short-term concentration exceedances that might occur at concentrations just meeting the current standard.

2.7.2 Approach

An empirical approach was selected to estimate exceedances at each location. A total of four air quality scenarios were evaluated using the empirical model for each of two distinct ambient monitoring periods, resulting in a total of eight separate analyses. The available NO_2 air quality were divided into two groups; one contained data from years 1995-2000, representing an historical data set; the other contained the monitoring years 2001-2006, representing recent ambient monitoring. Each of these monitoring year-groups were evaluated considering the NO_2 concentrations as they were reported and representing the conditions at that time (termed in this assessment "as is"). This served as the first air quality scenario. The second scenario considered the ambient NO_2 concentrations simulated to just meeting the current standard of 0.053 ppm annual average. The 3rd and 4th scenarios followed in similar fashion, however these scenarios used the ambient monitoring data to estimate NO_2 concentrations that might occur on roadways to generate on-road concentrations for as is air quality and for ambient concentrations just meeting the current standard. Again, each of these four scenarios was evaluated using both the historical and recent data air quality data sets.

Since all of the NO_2 ambient monitoring sites are represented by this analysis, the generated results are considered a broad characterization of national air quality and human exposures that might be associated with these concentrations. The output of this air quality characterization was used to estimate the number of times per year specific locations experience levels of NO_2 that could cause adverse health effects in susceptible individuals. Each location that was evaluated contained one to several monitors operating for a few to several years, generating a number of site-years of data. The number of site-years in a location were used to generate a distribution of two exposure and risk characterization metrics; the annual average concentrations and the numbers of exceedances that did (observed data) or could occur (simulated data) in a year for that location. The mean and median values were reported to represent the central tendency of each metric for the four scenarios in each air quality year-group, while the minimum value served to represent the lower bound. Since there were either multiple site-years or numerous simulations performed at each location using all available site-years of data, results for the upper percentiles included the 95th, 98th and 99th percentiles of the distribution.

2.7.3 Results

2.7.3.1 Air Quality Monitoring Data As Is

As mentioned previously, air quality data were separated into two groups, one representing historic data (1995-2000) and the other more recent data (2001-2006). Detailed statistics regarding concentration distributions for particular locations and specific monitoring years are provided in Appendices B and C. All of the results in Tables 13-15 are based on air quality data as reported by the AQS.

Table 13 provides descriptive statistics for ambient NO_2 concentrations and the site-years available for each location and air quality grouping. For example, in Boston, there were 58 complete site-years during 1995-2000, for which the annual mean concentrations ranged from about 5 ppb to 31 ppb with a mean annual mean of about 18 ppb. Los Angeles, New York, Phoenix and Denver (recent data only), had higher annual average concentrations at the mean and upper percentiles, considering both the recent and historic air quality data, compared with other locations. Annual average NO_2 concentrations have decreased on average by 14% at most of the locations when comparing the historic to the recent year groups, although the mean annual average concentrations increased by about 67% at the Denver location using the more recent data.

The number of short-term concentration exceedances follows in Tables 14 and 15 given potential health effect benchmark levels of 200, 250, and 300 ppb. For example, the numbers of exceedances of 200 ppb ranged from 0 to 1 with a mean estimated number of exceedances of 0 for Boston (Table 14). During the years 2001-2006, annual average ambient NO_2 concentrations ranged from 14 to 23 ppb in Detroit considering 122 site-years of data (Table 14). On average there was one exceedance of 200 ppb in Detroit across the total time period, however was as high as 12 given a particular year and site (Table 15).

In general, the number of exceedances of the selected benchmark levels was low when considering either air quality year-group and at any location. The average number of exceedances of the lowest potential health effect benchmark level across each location was primarily one or less, with very few locations deviating from this estimate. Where locations had a larger mean estimate, it was largely driven be a single site-year of data that contained a number of concentration exceedances. For example, the Colorado Springs mean estimate of exceedances was 3 for the entire area (a total of 8 monitors in operation at some time over the 6 year period of 1995-2000), however there was one-site year that contained 69 concentrations above 200 ppb (Table 14). That particular monitor (ID 0804160181) does not appear to have any unusual attributes; the closest major road is beyond a distance of 160 meters, the closest stationary source emitting > 5 tpy over 4 km away, and most sources within 10 km are emitting on average 430 tpy (Appendix A). However, one particular source is noted as driving the estimated mean emissions upwards. A power generating utility (NAICS code 221112) located at a 7.2 km distance contained an emission estimate of 4205 tpy, while 9 of the 11 sources located within 10 km are under 100 tpy (data not provided). It is not known at this time whether this particular facility is influencing the observed concentration exceedances at this specific monitoring site.

The same can be stated for the Phoenix location across the same time period, whereas a single year from one monitor (ID 0401330031) was responsible for all observed exceedances of 200 ppb (Table 14). While located closer to the roadway (at 78 m) than the Colorado Springs monitor, 9 of 10 stationary sources located within 10 km of this monitor emitted less than 60 tpy (one was at 272 tpy), none of which were located within 5 km. It is not known if observed exceedances of 200 ppb at this monitor are a result of proximity of major roads or stationary sources. Detroit contained the largest number of exceedances of 200 ppb (a maximum of 12) when considering the air quality data from years 2001-2006 (Table 15). Again, all of those exceedances occurred at one monitor (ID 2616300192) during one year (2002). Twelve sources of NO_x emissions are located within 2.6 to 5 km of this monitor, contributing between 6 and 27 tpy. The number of exceedances of higher potential health effect benchmark levels (i.e., 250 and 300 ppb) were of course less than those observed for 200 ppb, most of which were zero, with maximum numbers isolated the same aforementioned cities.

Table 13. Monitoring site-years and annual average NO_2 concentrations for two monitoring periods, historic and recent air quality data (as is).

Location	1995-2000							2001-2006						
	Site-Years	Annual Mean (ppb) [1]						Site-Years	Annual Mean (ppb) [1]					
		mean	min	med	p95	p98	p99		mean	min	med	p95	p98	p99
Boston	58	18	5	19	31	31	31	47	15	5	13	25	30	30
Chicago	47	24	9	24	32	34	34	36	24	16	23	32	32	32
Cleveland	11	23	17	23	28	28	28	11	19	14	19	24	24	24
Denver	26	16	6	9	35	35	35	10	26	18	27	37	37	37
Detroit	12	19	12	19	26	26	26	12	19	14	19	23	23	23
Los Angeles	193	26	4	26	45	46	46	177	22	4	22	36	37	40
Miami	24	10	6	9	17	17	17	20	9	6	8	15	16	16
New York	93	27	11	27	41	42	42	81	23	10	24	36	40	40
Philadelphia	46	23	15	21	33	34	34	39	20	14	19	29	30	30
Washington	69	20	9	22	26	27	27	66	18	7	19	25	26	26
Atlanta	24	14	5	15	25	27	27	29	12	3	14	19	23	23
Colorado Springs[2]	26	16	7	17	24	35	35	-	-	-	-	-	-	-
El Paso	14	23	14	23	35	35	35	30	15	8	16	21	22	22
Jacksonville	6	15	14	15	16	16	16	4	14	13	14	15	15	15
Las Vegas	16	14	3	8	27	27	27	35	11	1	9	22	23	23
Phoenix	22	30	24	30	36	40	40	27	25	11	24	35	37	37
Provo	6	24	23	24	24	24	24	6	24	21	23	29	29	29
St. Louis	56	18	5	19	26	26	27	43	15	8	15	22	25	25
Other CMSA	1135	14	1	14	24	26	28	1177	12	1	12	20	22	24
Not MSA	200	8	0	7	16	19	19	243	7	1	6	14	16	16

[1] The mean is the sum of the annual means for each monitor in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the annual mean.

[2] Colorado Springs monitoring data were collected as part of short-term study completed in September 2001, therefore there are no 2001-2006 data.

Table 14. Number of exceedances of short-term (1-hour) potential health effect benchmark levels in a year, 1995-2000 historic NO_2 air quality (as is).

	Exceedances of 200 ppb 1-hour [1]						**Exceedances of 250 ppb 1-hour [1]**						**Exceedances of 300 ppb 1-hour [1]**					
Location	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**
Boston	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Chicago	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cleveland	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0
Denver	0	0	0	1	2	2	0	0	0	0	1	1	0	0	0	0	0	0
Detroit	0	0	0	3	3	3	0	0	0	1	1	1	0	0	0	1	1	1
Los Angeles	0	0	0	1	2	4	0	0	0	0	0	2	0	0	0	0	0	0
Miami	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
New York	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0
Philadelphia	0	0	0	0	3	3	0	0	0	0	0	0	0	0	0	0	0	0
Washington	0	0	0	0	1	2	0	0	0	0	1	1	0	0	0	0	0	0
Atlanta	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Colorado Springs	3	0	0	3	69	69	1	0	0	0	23	23	0	0	0	0	4	4
El Paso	0	0	0	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0
Jacksonville	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Las Vegas	1	0	0	11	11	11	0	0	0	3	3	3	0	0	0	3	3	3
Phoenix	2	0	0	0	37	37	0	0	0	0	3	3	0	0	0	0	0	0
Provo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
St. Louis	0	0	0	0	0	8	0	0	0	0	0	4	0	0	0	0	0	0
Other CMSA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Not MSA	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0

[1] The mean number of exceedances represents the number of exceedances occurring at all monitors in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, and p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the number of exceedances in any one year within the monitoring period.

Table 15. Number of exceedances of short-term (1-hour) potential health effect benchmark levels in a year, 2001-2006 recent NO_2 air quality (as is).

	Exceedances of 200 ppb 1-hour						Exceedances of 250 ppb 1-hour						Exceedances of 300 ppb 1-hour					
Location	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**
Boston	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chicago	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cleveland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Denver	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Detroit	1	0	0	12	12	12	1	0	0	8	8	8	0	0	0	5	5	5
Los Angeles	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0
Miami	0	0	0	2	3	3	0	0	0	2	3	3	0	0	0	2	3	3
New York	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Philadelphia	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0
Washington	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Atlanta	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
El Paso	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Jacksonville	1	0	1	2	2	2	0	0	0	1	1	1	0	0	0	0	0	0
Las Vegas	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Phoenix	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Provo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
St. Louis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Other CMSA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Not MSA	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0

[1] The mean number of exceedances represents the number of exceedances occurring at all monitors in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, and p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the number of exceedances in any one year within the monitoring period.

2.7.3.2 Simulated Air Quality Data to Just Meet The Current Standard

Descriptive statistics for the ambient NO_2 concentrations simulated to just meet the current standard is presented in Table 16. Note both year groups (1995-2000 and 2001-2006) contain maximum concentrations of 53 ppb for the annual average concentration, a direct consequence of the concentration roll-up procedure. On average, the mean simulated annual average NO_2 concentrations are about 2.1-2.4 higher than the respective year group as is concentrations, however the actual range in this factor could be as low as 1.2 or as high as 3.8 depending on year group and particular location selected. This is a function of the location and year specific factors used to simulate just meeting the current standard (Tables 10 and 11) that are also of similar range.

As expected, the number of estimated short-term concentration exceedances is greater when considering the current standard and considering all potential health effect benchmark levels (Tables 17 and 18). For example, depending on location and year-group, the number of exceedances of 200 ppb on average could be a low as 0 or as high as 88 considering air quality simulated to meet the current standard, although about 75% the location-year groups were estimated as containing a mean of less than 10 exceedances in a year. Median estimates of exceedances of 200 ppb were typically well below that of the mean, 85% were either 1 or less, indicating an upward bias in the mean influenced by the number of exceedances at the upper ends of the distribution.

The same three cities noted in the *as is* evaluation above contained the highest mean and maximum number of exceedances when considering the simulated historic air quality data (i.e., Colorado Springs, Detroit, Phoenix). The same reasoning applies here, primarily the influence of concentration exceedances at a single monitor. Miami and Jacksonville are also indicated as having a relatively higher estimate of mean number of exceedances than the other locations, however this is driven mainly by the small network size (n=1 for Jacksonville, n=5 for Miami). Having a limited number of monitors in a given location could bias the mean estimate in either direction (high or low), most notable here where there were an unusual number of peak concentrations in a given year. In addition, Miami contained some of the lowest annual average concentrations (Table 13), yielding the highest air quality simulation factors across all years of data (Tables 10-11). That coupled with a high COV (~130%) for hourly concentrations at a two of the monitors in Miami (IDs 1201180021, 1208600271) clearly played a significant role in the higher estimated number of exceedances. Denver also contained a high COV (~110%) for the earlier air quality period (1995-2000), likely associated with the higher estimate of maximum exceedances at this location (141) following the concentration roll-up compared with only 2 observed exceedances when considering the air quality as is. Both the mean and maximum estimate of exceedances for Provo (ID 4904900021) during 2001-2006 were also likely influenced by the small network size (n=1) in this location and one particular year (2006) that contained a number of concentrations above 150 ppb prior to the concentration roll-up.

Table 16. Estimated annual average NO_2 concentrations for two monitoring periods, historic and recent air quality data adjusted to just meet the current standard (0.053 ppm annual average).

	1995-2000							2001-2006						
	Site-Years	Annual Mean (ppb) [1]						Site-Years	Annual Mean (ppb) [1]					
Location		**mean**	**min**	**med**	**p95**	**p98**	**p99**		**mean**	**min**	**med**	**p95**	**p98**	**p99**
Boston	58	32	10	33	53	53	53	47	32	11	28	53	53	53
Chicago	47	39	15	40	53	53	53	36	41	27	39	53	53	53
Cleveland	11	47	37	53	53	53	53	11	48	41	53	53	53	53
Denver	26	29	10	29	53	53	53	10	47	33	53	53	53	53
Detroit	12	45	26	51	53	53	53	12	49	42	50	53	53	53
Los Angeles	193	31	4	32	52	53	53	177	33	5	33	53	53	53
Miami	24	34	19	31	53	53	53	20	35	19	32	53	53	53
New York	93	35	14	35	53	53	53	81	35	15	35	53	53	53
Philadelphia	46	39	25	35	53	53	53	39	41	26	40	53	53	53
Washington	69	42	20	45	53	53	53	66	40	19	44	53	53	53
Atlanta	24	32	11	31	53	53	53	29	34	9	40	53	53	53
Colorado Springs[2]	26	38	14	45	53	53	53	-	-	-	-	-	-	-
El Paso	14	43	30	40	53	53	53	30	42	24	43	53	53	53
Jacksonville	6	53	53	53	53	53	53	4	53	53	53	53	53	53
Las Vegas	16	29	7	17	53	53	53	35	28	4	21	53	53	53
Phoenix	22	45	36	44	53	53	53	27	40	19	40	53	53	53
Provo	6	53	53	53	53	53	53	6	53	53	53	53	53	53
St. Louis	56	37	11	39	53	53	53	43	38	19	38	53	53	53
Other CMSA	1135	26	1	26	43	48	50	1177	25	1	26	43	48	51
Not MSA	200	22	1	20	51	53	53	243	22	3	20	46	53	53

[1] The mean is the sum of the annual means for each monitor in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the annual mean.

[2] Colorado Springs monitoring data were collected as part of short-term study completed in September 2001, therefore there are no 2001-2006 data.

Table 17. Estimated number of exceedances of short-term (1-hour) potential health effect benchmark levels in a year, 1995-2000 NO$_2$ air quality adjusted to just meet the current standard (0.053 ppm annual average).

	Exceedances of 200 ppb 1-hour						**Exceedances of 250 ppb 1-hour**						**Exceedances of 300 ppb 1-hour**					
Location	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**
Boston	0	0	0	1	1	2	0	0	0	0	1	1	0	0	0	0	0	1
Chicago	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Cleveland	3	0	0	24	24	24	1	0	0	10	10	10	0	0	0	3	3	3
Denver	8	0	0	19	141	141	2	0	0	5	28	28	1	0	0	4	9	9
Detroit	13	0	13	25	25	25	4	0	2	15	15	15	2	0	1	10	10	10
Los Angeles	1	0	0	5	8	9	0	0	0	0	2	2	0	0	0	0	0	2
Miami	10	0	8	27	34	34	2	0	0	6	15	15	1	0	0	2	8	8
New York	0	0	0	1	2	3	0	0	0	0	1	3	0	0	0	0	0	1
Philadelphia	0	0	0	1	12	12	0	0	0	0	9	9	0	0	0	0	5	5
Washington	1	0	0	4	9	17	0	0	0	1	3	3	0	0	0	1	2	2
Atlanta	4	0	0	19	21	21	0	0	0	2	3	3	0	0	0	1	1	1
Colorado Springs	30	0	0	180	241	241	15	0	0	123	135	135	8	0	0	72	83	83
El Paso	4	0	1	14	14	14	1	0	0	6	6	6	0	0	0	2	2	2
Jacksonville	12	2	15	20	20	20	2	0	1	7	7	7	0	0	0	1	1	1
Las Vegas	3	0	0	28	28	28	1	0	0	13	13	13	1	0	0	11	11	11
Phoenix	12	0	0	57	198	198	4	0	0	4	92	92	1	0	0	0	31	31
Provo	1	0	0	5	5	5	0	0	0	0	0	0	0	0	0	0	0	0
St. Louis	0	0	0	1	1	15	0	0	0	0	0	14	0	0	0	0	0	13
Other CMSA	0	0	0	1	3	6	0	0	0	0	1	1	0	0	0	0	0	1
Not MSA	4	0	0	18	53	87	1	0	0	4	15	42	1	0	0	1	8	21

[1] The mean number of exceedances represents the number of exceedances occurring at all monitors in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, and p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the number of exceedances in any one year within the monitoring period.

Table 18. Estimated number of exceedances of short-term (1-hour) potential health effect benchmark levels in a year, 2001-2006 NO_2 air quality adjusted to just meet the current standard (0.053 ppm annual average).

	Exceedances of 200 ppb 1-hour						**Exceedances of 250 ppb 1-hour**						**Exceedances of 300 ppb 1-hour**					
Location	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**
Boston	0	0	0	1	5	5	0	0	0	0	1	1	0	0	0	0	0	0
Chicago	1	0	0	2	15	15	0	0	0	0	0	0	0	0	0	0	0	0
Cleveland	1	0	1	4	4	4	0	0	0	1	1	1	0	0	0	1	1	1
Denver	2	0	1	7	7	7	0	0	0	2	2	2	0	0	0	1	1	1
Detroit	8	0	1	45	45	45	4	0	0	34	34	34	3	0	0	28	28	28
Los Angeles	0	0	0	1	5	6	0	0	0	0	0	1	0	0	0	0	0	1
Miami	17	0	11	66	69	69	3	0	0	18	23	23	1	0	0	11	19	19
New York	0	0	0	1	2	5	0	0	0	0	1	1	0	0	0	0	0	0
Philadelphia	1	0	0	2	25	25	0	0	0	1	7	7	0	0	0	0	1	1
Washington	0	0	0	2	5	6	0	0	0	1	1	2	0	0	0	0	1	1
Atlanta	8	0	0	48	56	56	1	0	0	9	10	10	0	0	0	2	5	5
El Paso	7	0	6	24	27	27	1	0	0	3	6	6	0	0	0	0	1	1
Jacksonville	31	7	22	72	72	72	15	1	7	46	46	46	7	0	1	25	25	25
Las Vegas	1	0	0	3	12	12	0	0	0	0	2	2	0	0	0	0	0	0
Phoenix	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Provo	88	0	0	526	526	526	34	0	0	205	205	205	0	0	0	1	1	1
St. Louis	0	0	0	2	5	5	0	0	0	1	1	1	0	0	0	0	1	1
Other CMSA	0	0	0	1	3	5	0	0	0	0	1	2	0	0	0	0	0	1
Not MSA	3	0	0	17	44	57	1	0	0	4	14	20	1	0	0	2	8	9

[1] The mean number of exceedances represents the number of exceedances occurring at all monitors in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, and p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the number of exceedances in any one year within the monitoring period.

2.7.3.3 Simulated On-Road Concentrations, Air Quality Data As Is

Descriptive statistics for estimated on-road NO_2 concentrations are presented in Table 19. These on-road concentrations were generated by using the simulation procedure described in Section 2.6 as applied to air quality data as is. On average, the simulated on-road annual average concentrations are approximately a factor of 1.8 higher compared with their respective ambient concentrations (see Table 13). This factor is consistent with the range of 1.5 to 2 reported in the ISA (US EPA, 2007f) for studies that compared on-road to ambient NO_2 concentrations. Los Angeles, New York, Phoenix and Denver (recent data only for this location), were locations estimated to contain higher on-road concentrations at the mean and upper percentiles considering both the recent and historic air quality data compared with other locations. This is a direct result of these locations already containing the highest *as is* concentrations prior to the on-road simulation.

As a point of reference, the median of the simulated concentration estimates for Los Angeles were compared with NO_2 measurements provided by Westerdahl et al. (2005) for arterial roads and freeways in the same general location during spring 2003. Although the averaging time is not the same,[7] comparison of the medians could be considered appropriate. Median on-road concentrations from Westerdahl et al. (2005) ranged from 31 to 55 ppb and compare well with the median of 40 ppb estimated here for years 2001-2006.

When considering the number of exceedances of 200 ppb estimated to occur on-road, most locations, on average, would have had less than 10 in a year. As observed with the ambient NO_2 concentrations, the median frequency of exceedances in most locations were estimated to be typically 1 or less per year, considering both the historic and recent air quality data (Tables 20 and 21). However, the number of exceedances at each location were consistently less when considering the recent air quality compared with the historic air quality. There were a few exceptions to these generalities, such as the high number of estimated on-road exceedances of 200 ppb for the Colorado Springs and Provo locations. Again, these were the result of these locations having few monitoring sites and a number influential NO_2 concentrations at the upper percentiles of the distribution in one or a few site-years. When considering the two largest groups (all of the other CMSA/MSA and Not CMSA), it is estimated that, on average, about 1 or less exceedances per year of 200 ppb could occur. The 95 percent interval indicates as many as 14 exceedances at a particular site within that large grouping for a given year considering the historic data, while only as many as 4 when considering the more recent data.

There were similarities in the estimated distributions for Chicago, Los Angeles, and New York. Each of these locations are large CMSA, contain several monitoring sites, and have an abundance of roads and associated vehicles.[8] Based on the calculations here, each of these locations was estimated to have on average, about 10 exceedances of 200 ppb per year on-roads. Assuming that the on-road exceedances distribution is proportionally representing the distribution of roadways within each location, about one-half of the roads in these areas would

---

[7] Table 13 here considers the median of the annual average while Westerdahl et al. (2005) reported median concentrations averaged over 2 to 4 hours. In general, there are no differences for the mean annual averages versus the mean hourly averages (see Appendix B), the main difference in these two metrics is in the variability (and hence the various percentiles of the distribution outside the central tendency).

[8] Of the named locations, Chicago, Los Angeles, and New York contain the highest daily vehicle miles traveled (Federal Highway Administration (FHWA, 2005)).

not have any concentrations in excess of 200 ppb. This is because the median value for exceedances of 200 ppb in most locations is zero. However, Tables 20 and 21 indicate that there is also a possibility of tens to just over a hundred exceedances in a year on certain roads/sites.

Table 19. Estimated annual average on-road concentrations for two monitoring periods, historic and recent ambient air quality (as is).

	1995-2000							2001-2006						
	Site-Years	Annual Mean (ppb)[1]						Site-Years	Annual Mean (ppb)[1]					
Location		mean	min	med	p95	p98	p99		mean	min	med	p95	p98	p99
Boston	5800	33	7	33	59	67	71	4700	27	7	25	51	57	60
Chicago	4700	44	11	44	68	75	79	3600	43	20	42	66	72	76
Cleveland	1100	42	22	41	61	65	67	1100	36	18	35	51	54	58
Denver	2600	29	8	19	67	78	81	1000	48	23	46	74	83	87
Detroit	1200	35	15	34	52	57	59	1200	34	18	34	47	52	54
Los Angeles	19300	48	5	47	87	97	104	17700	41	5	40	71	80	85
Miami	2400	19	7	17	33	38	39	2000	17	7	15	30	33	36
New York	9300	50	14	49	81	91	96	8100	43	12	41	70	79	85
Philadelphia	4600	43	19	40	68	76	80	3900	37	18	34	57	63	68
Washington	6900	37	12	38	56	61	64	6600	33	9	33	52	57	61
Atlanta	2400	26	6	25	49	57	60	2900	21	4	23	40	43	47
Colorado Springs	2600	30	9	30	51	64	73	-	-	-	-	-	-	-
El Paso	1400	42	17	40	67	75	82	3000	27	10	27	42	45	48
Jacksonville	600	28	18	27	37	39	41	400	25	17	25	34	36	37
Las Vegas	1600	26	4	16	56	62	63	3500	20	2	15	45	50	53
Phoenix	2200	54	30	52	76	83	88	2700	45	14	43	70	79	84
Provo	600	43	29	42	58	62	64	600	43	26	41	61	69	70
St. Louis	5600	33	7	33	51	58	61	4300	27	10	27	44	49	52
Other CMSA	113500	26	1	25	47	53	57	117700	21	1	21	39	45	48
Not MSA	20000	14	0	12	31	35	39	24300	12	1	11	27	31	33

[1] The mean is the sum of the annual means for each monitor in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, and p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the annual mean.

Table 20. Estimated number of exceedances of short-term (1-hour) potential health effect benchmark levels in a year on-roads, 1995-2000 historic NO_2 air quality (as is).

	Exceedances of 200 ppb 1-hour						Exceedances of 250 ppb 1-hour						Exceedances of 300 ppb 1-hour					
Location	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**
Boston	3	0	0	14	37	54	1	0	0	2	10	15	0	0	0	0	1	3
Chicago	12	0	0	79	142	183	2	0	0	15	31	53	0	0	0	2	6	10
Cleveland	10	0	0	74	108	129	2	0	0	12	30	49	1	0	0	1	10	17
Denver	7	0	0	41	94	102	2	0	0	9	17	33	1	0	0	4	6	7
Detroit	10	0	2	48	72	86	4	0	1	21	34	35	2	0	0	14	21	26
Los Angeles	45	0	4	236	417	550	13	0	0	71	146	211	4	0	0	21	48	78
Miami	0	0	0	4	6	8	0	0	0	1	4	6	0	0	0	0	3	4
New York	20	0	1	109	230	384	5	0	0	28	65	129	1	0	0	5	14	31
Philadelphia	5	0	0	31	60	84	1	0	0	4	11	15	0	0	0	1	4	7
Washington	4	0	0	23	43	58	0	0	0	3	7	11	0	0	0	1	2	2
Atlanta	4	0	0	31	57	87	1	0	0	3	11	21	0	0	0	1	1	2
Colorado Springs	20	0	0	170	264	320	11	0	0	106	181	216	6	0	0	47	119	159
El Paso	7	0	2	33	58	76	2	0	0	9	19	30	1	0	0	5	7	11
Jacksonville	0	0	0	1	2	4	0	0	0	0	1	1	0	0	0	0	0	0
Las Vegas	6	0	0	37	66	97	1	0	0	11	15	19	1	0	0	6	11	11
Phoenix	36	0	3	256	319	390	14	0	0	107	200	280	7	0	0	26	103	181
Provo	2	0	0	9	33	34	0	0	0	0	1	4	0	0	0	0	0	0
St. Louis	2	0	0	14	25	35	0	0	0	1	8	12	0	0	0	0	4	10
Other CMSA	1	0	0	6	18	32	0	0	0	1	3	6	0	0	0	0	1	2
Not MSA	1	0	0	2	7	14	0	0	0	1	2	4	0	0	0	0	1	2

[1] The mean number of exceedances represents the number of exceedances occurring at all monitors in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, and p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the number of exceedances in any one year within the monitoring period.

Table 21. Estimated number of exceedances of short-term (1-hour) potential health effect benchmark levels in a year on-roads, 2001-2006 historic NO_2 air quality (as is).

	Exceedances of 200 ppb 1-hour						**Exceedances of 250 ppb 1-hour**						**Exceedances of 300 ppb 1-hour**					
Location	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**
Boston	1	0	0	2	8	17	0	0	0	0	1	4	0	0	0	0	0	0
Chicago	10	0	0	50	142	188	2	0	0	11	29	44	0	0	0	1	6	8
Cleveland	3	0	0	21	36	42	1	0	0	4	7	9	0	0	0	1	3	3
Denver	8	0	1	39	69	82	2	0	0	8	15	20	0	0	0	1	7	7
Detroit	5	0	0	29	44	45	2	0	0	16	22	28	1	0	0	13	14	21
Los Angeles	11	0	0	70	131	183	2	0	0	13	29	48	0	0	0	2	7	13
Miami	0	0	0	3	7	13	0	0	0	2	5	5	0	0	0	2	4	5
New York	9	0	0	48	90	143	2	0	0	8	19	25	0	0	0	1	3	6
Philadelphia	1	0	0	6	14	29	0	0	0	1	1	2	0	0	0	0	1	1
Washington	1	0	0	6	14	21	0	0	0	0	1	2	0	0	0	0	0	0
Atlanta	1	0	0	8	16	25	0	0	0	1	3	6	0	0	0	0	1	2
El Paso	1	0	0	6	9	15	0	0	0	1	1	2	0	0	0	0	0	0
Jacksonville	3	0	1	15	23	24	2	0	0	8	15	15	1	0	0	5	8	8
Las Vegas	1	0	0	6	15	23	0	0	0	0	1	3	0	0	0	0	0	0
Phoenix	3	0	0	21	44	61	0	0	0	2	5	7	0	0	0	0	0	0
Provo	70	0	0	547	662	662	33	0	0	234	606	612	13	0	0	3	423	435
St. Louis	1	0	0	2	7	14	0	0	0	0	1	2	0	0	0	0	0	1
Other CMSA	0	0	0	1	5	10	0	0	0	0	1	1	0	0	0	0	0	0
Not MSA	0	0	0	1	4	8	0	0	0	0	2	3	0	0	0	0	1	2

[1] The mean number of exceedances represents the number of exceedances occurring at all monitors in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, and p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the number of exceedances in any one year within the monitoring period.

2.7.3.4 Simulated On-Road Concentrations, Simulated Air Quality Data To Just Meet The Current Standard

Descriptive statistics for estimated on-road NO_2 concentrations with just meeting the current standard are presented in Table 22. These on-road concentrations were generated by using the simulation procedure described in Section 2.6 and applied to simulated air quality data to just meet the current standard using the approach described in Section 2.5. On average, the simulated on-road annual average concentrations are also about 1.8 time higher than the ambient concentrations rolled-up to just meet the current standard (see Table 16), similar to what was observed for this relationship considering the air quality (as is).

The mean number of estimated exceedances of 200 ppb ranges from tens to several hundreds (Tables 23 and 24), sharply increased from the previous on-road estimates using the air quality (*as is*). Some of the highest exceedance estimates occurred in the locations described previously as being influenced by a few concentrations at the upper percentiles of their distributions in a small number of years and/or monitoring sites (e.g., Miami, Colorado Springs, Provo). Compared to the means, median estimated exceedances of 200 ppb are lower, on average by about 60%, indicating the presence of highly influential data at the upper percentiles of the distribution at each location. This is evident when considering the 95th – 99th percentiles, where several hundred to around two thousand exceedances of 200 ppb were estimated. However, the estimated number of exceedances is lower for locations containing more site-years of data than for the locations with the fewest site-years. This trend is consistent with those described earlier, whereas estimates of exceedances in the simulated data for the large urban areas are stabilized by greater sample size (both the number of monitors and 1-hour values). The median number of exceedances of 200 ppb at the locations containing a larger monitoring network (i.e. at least 40 site-years per year-group) was estimated to be between 10 and 100 per year. Upper bounds for the locations with the greatest number of monitoring sites approach around 1,000 estimated on-road exceedances per year upon just meeting the current standard.

It should be noted that the estimated on-road concentrations and number exceedances for many of the locations were higher for the 2001-2006 rolled-up data when compared with the 1995-2000 rolled-up data. To obtain generally comparable results across the two time periods, the assumption for the concentration roll-up was that a similar level of variability be maintained from year-to-year (or year-group to year-group). As described in section 2.5 of the draft TSD, a slight increase in hourly COV occurred from 1995-2006 (~10% for all locations). The effect may have finally emerged in this combined simulation by generating a greater number of concentrations above the potential health effect benchmarks that may have previously been just below the threshold in the earlier on-road simulations considering the *as is* ambient concentrations.

Table 22. Estimated annual average on-road concentrations for two monitoring periods, air quality data adjusted to just meet the current standard (0.053 ppm annual average).

	1995-2000							**2001-2006**						
	Site-	**Annual Mean (ppb)[1]**						**Site-**	**Annual Mean (ppb)[1]**					
Location	**Years**	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**Years**	**mean**	**min**	**med**	**p95**	**p98**	**p99**
Boston	5800	58	13	57	103	117	125	4700	58	14	53	105	120	126
Chicago	4700	72	18	72	112	123	130	3600	74	35	72	113	124	130
Cleveland	1100	86	47	84	123	128	136	1100	88	53	86	123	130	146
Denver	2600	53	12	49	112	124	129	1000	85	42	85	124	130	141
Detroit	1200	81	33	83	124	129	133	1200	90	54	87	123	129	134
Los Angeles	19300	56	6	55	102	114	122	17700	61	7	60	105	116	123
Miami	2400	62	24	56	111	124	128	2000	63	25	57	112	126	129
New York	9300	64	18	62	104	117	123	8100	63	18	61	103	119	125
Philadelphia	4600	71	31	67	111	123	128	3900	74	33	71	111	125	128
Washington	6900	77	26	77	116	124	130	6600	73	23	74	114	124	128
Atlanta	2400	57	14	55	111	126	129	2900	61	12	66	111	126	129
Colorado Springs	2600	69	18	73	118	127	131	-	-	-	-	-	-	-
El Paso	1400	77	38	74	122	129	138	3000	75	30	74	112	124	128
Jacksonville	600	96	67	95	128	131	144	400	96	67	94	129	139	145
Las Vegas	1600	53	8	34	113	125	130	3500	50	5	36	112	124	129
Phoenix	2200	82	46	78	115	127	129	2700	72	24	71	110	125	127
Provo	600	96	67	95	129	139	144	600	95	67	93	128	131	138
St. Louis	5600	68	14	68	106	118	124	4300	69	25	67	106	118	126
Other CMSA	113500	46	1	46	84	95	103	117700	46	1	45	84	95	102
Not MSA	20000	39	1	35	90	104	115	24300	39	3	35	89	101	109

[1] The mean is the sum of the annual means for each monitor in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, and p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the annual mean.

Table 23. Estimated number of exceedances of short-term (1-hour) potential health effect benchmark levels in a year on-roads, 1995-2000 historic NO_2 air quality adjusted to just meet the current standard (0.053 ppm annual average).

	Exceedances of 200 ppb 1-hour[1]						Exceedances of 250 ppb 1-hour[1]						Exceedances of 300 ppb 1-hour[1]					
Location	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**
Boston	78	0	13	411	677	790	23	0	1	131	257	334	8	0	0	43	106	131
Chicago	172	0	61	727	1001	1170	59	0	7	303	512	643	22	0	0	137	230	322
Cleveland	321	1	195	1045	1221	1439	124	0	38	566	663	761	51	0	5	304	380	392
Denver	214	0	23	1261	1921	2215	97	0	5	511	1142	1574	45	0	1	228	582	908
Detroit	405	2	284	1227	1439	1589	175	2	97	576	776	872	80	0	40	317	424	482
Los Angeles	100	0	18	489	791	927	33	0	2	173	318	432	12	0	0	62	127	184
Miami	363	1	260	1045	1334	1427	162	0	93	579	737	791	72	0	32	316	396	430
New York	77	0	11	412	693	930	23	0	1	127	258	420	8	0	0	40	91	171
Philadelphia	114	0	27	570	797	942	32	0	4	181	308	364	9	0	0	52	104	138
Washington	219	0	101	852	1070	1185	73	0	18	351	457	525	27	0	2	158	220	270
Atlanta	251	0	42	1094	1472	1640	106	0	7	535	843	947	45	0	1	277	435	514
Colorado Springs	304	0	77	1320	1756	1879	120	0	11	565	769	930	60	0	1	294	371	416
El Paso	178	0	82	692	951	1105	57	0	24	215	347	447	21	0	8	78	162	200
Jacksonville	610	40	549	1426	1515	1801	263	2	195	773	839	1002	114	0	66	407	443	470
Las Vegas	238	0	26	1107	1674	1882	89	0	5	574	688	860	36	0	1	280	369	422
Phoenix	250	0	105	953	1326	1435	83	0	17	379	466	563	33	0	3	181	296	364
Provo	443	1	230	1643	1871	2058	135	0	32	543	697	817	43	0	2	208	303	339
St. Louis	148	0	48	620	871	966	46	0	6	259	356	432	16	0	0	99	163	200
Other CMSA	52	0	6	268	444	592	15	0	0	84	156	231	5	0	0	25	57	90
Not MSA	95	0	7	549	928	1203	39	0	1	221	438	635	17	0	0	91	198	318

[1] The mean number of exceedances represents the number of exceedances occurring at all monitors in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, and p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the number of exceedances in any one year within the monitoring period.

Table 24. Estimated number of exceedances of short-term (1-hour) potential health effect benchmark levels in a year on-roads, 2001-2006 recent NO_2 air quality adjusted to just meet the current standard (0.053 ppm annual average).

	Exceedances of 200 ppb 1-hour[1]						**Exceedances of 250 ppb 1-hour[1]**						**Exceedances of 300 ppb 1-hour[1]**					
0**Location**	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**	**mean**	**min**	**med**	**p95**	**p98**	**p99**
Boston	87	0	12	458	753	990	23	0	1	137	263	330	7	0	0	38	93	132
Chicago	176	0	61	805	1022	1139	59	0	7	335	560	620	23	0	0	128	295	354
Cleveland	387	14	268	1117	1322	1735	149	0	65	573	676	846	62	0	15	326	407	428
Denver	277	0	113	964	1233	1560	87	0	22	337	430	557	28	0	5	125	203	283
Detroit	440	17	309	1214	1444	1628	166	0	90	513	689	744	67	0	25	265	322	385
Los Angeles	106	0	23	533	788	893	31	0	2	186	290	363	10	0	0	59	115	150
Miami	406	3	306	1173	1345	1416	193	0	113	669	855	923	88	0	35	367	542	588
New York	84	0	14	458	709	872	25	0	1	149	295	413	8	0	0	49	110	177
Philadelphia	174	0	60	726	973	1184	51	0	7	239	383	521	16	0	1	77	153	227
Washington	208	0	83	874	1171	1310	63	0	10	327	426	558	21	0	1	127	181	224
Atlanta	335	0	135	1293	1647	1755	143	0	21	687	973	1093	61	0	4	339	510	656
El Paso	389	4	257	1251	1604	1737	144	0	66	530	858	971	54	0	20	221	350	441
Jacksonville	607	56	542	1385	1642	1743	273	5	202	789	924	1027	125	1	74	436	490	557
Las Vegas	278	0	43	1319	1929	2196	101	0	6	680	828	1045	42	0	0	354	502	565
Phoenix	149	0	19	758	1172	1352	33	0	1	203	303	370	7	0	0	48	70	95
Provo	516	1	345	1664	1966	2115	228	0	72	729	818	847	134	0	5	643	693	694
St. Louis	182	0	69	762	1100	1216	59	0	8	302	468	576	20	0	1	127	211	260
Other CMSA	64	0	6	333	569	740	19	0	0	105	207	300	6	0	0	31	72	120
Not MSA	101	0	7	569	874	1095	39	0	1	232	419	569	16	0	0	95	184	264

[1] The mean number of exceedances represents the number of exceedances occurring at all monitors in a particular location divided by the number of site-years across the monitoring period. The min, med, p95, p98, and p99 represent the minimum, median, 95th, 98th, and 99th percentiles of the distribution for the number of exceedances in any one year within the monitoring period.

2.8 Variability and Uncertainty

This uncertainty analysis first identifies the sources of the assessment that do or do not contribute to uncertainty, and provide a rationale for why this is the case. A qualitative evaluation follows for the types and components of uncertainty, resulting in a matrix describing, for each source of uncertainty, both the direction and magnitude of influence has on exposure estimates. The bias direction indicates how the source of uncertainty is judged to influence estimated concentrations, either the concentrations are likely “over-“ or “under-estimated”. In the instance where two types or components of uncertainty result in offsetting direction of influence, the uncertainty was judged as “both”. The magnitude indicates an estimated size of influence the uncertainty has on estimated concentrations. “Minimal” uncertainty was noted where quantitative evidence indicates the influence is either conditional and/or limited to few components in type. A characterization of “moderate” was assigned where multiple components of uncertainty existed within a given type and act in similar direction, however the presence of all at once may be dependent on certain conditions. “Major” uncertainty was used where multiple components of uncertainty exist within a given type, the components have few limiting conditions, and the components consistently act in similar bias direction. “Unknown” was assigned where there was no evidence reviewed to judge the uncertainty associated with the source. Table 25 provides a summary of the sources of uncertainty identified in the air quality characterization and the judged bias and magnitude of each.

2.8.1 Air Quality Data

One basic assumption is that the AQS NO_2 air quality data used are quality assured already. Reported concentrations contain only valid measures, since values with quality limitations are either removed or flagged. There is likely no selective bias in retention of data that is not of reasonable quality, it is assumed that selection of high concentration poor quality data would be just as likely as low concentration data of poor quality. Given the numbers of measurements used for this analysis, it is likely that even if a few low quality data are present in the data set, they would not have any significant effect on the results presented here. Therefore, the air quality data and database used likely contributes minimally to uncertainty. Temporally, the data are hourly measurements and appropriately account for variability in concentrations that are commonly observed for NO_2 and by definition are representative of an entire year. In addition, having more than one monitor does account for some of the spatial variability in a particular location. However, the degree of representativeness of the monitoring data used in this analysis can be evaluated from several perspectives, one of which is how well the temporal and spatial variability are represented. In particular, missing hourly measurements at a monitor may introduce bias (if different periods within a year or different years have different numbers of measured values) and increase the uncertainty. Furthermore, the spatial representativeness will be poor if the monitoring network is not dense enough to resolve the spatial variability (causing increased uncertainty) or if the monitors are not evenly distributed (causing a bias). Additional uncertainty regarding temporal and spatial representation by the monitors is expanded below.

2.8.2 Measurement Technique for Ambient NO_2

One source of uncertainty for NO_2 air quality data is due to interference with other oxidized nitrogen compounds. The ISA points out positive interference, commonly from HNO_3, of up to 50%, particularly during the afternoon hours, resulting in overestimation of concentrations.

Also, negative vertical gradients exist for monitors (2.5 times higher at 4 meter vs. 15 meter vertical siting (draft ISA, section 2.5.3.3), thus monitors positioned on rooftops may underestimate exposures. Only 7 of the 177[9] monitors in the named locations contained monitoring heights of 15 meters or greater, with nearly 60% at 4 meters or less height, and 80% at 5 meters or less in height. Not accounting for this potential vertical gradient in NO_2 concentrations may generate underestimates of exceedances for some site-years, however the overall impact of inferences made for the locations included in this assessment is likely minimal since most monitors sited at less than 4-5 meters in vertical height.

2.8.3 Temporal Representation

Data are valid hourly measures and are of similar temporal scale as the potential health effect benchmark concentrations. There are frequent missing values within a given valid year which contribute to the uncertainty as well as introducing a possible bias if some seasons, day types (e.g., weekday/weekend), or time of the day (e.g., night or day) are not equally represented. Since a 75 percent daily and hourly completeness rule was applied, some of these uncertainties and biases were reduced in these analyses. Data were not interpolated in the analysis. Similarly, there may be bias and uncertainty if the years monitored vary significantly between locations. Although monitoring locations within a region do change over time, the NO_2 network has been reasonably stable over the 1995-2006 period, particularly at locations with larger monitoring networks, so the impact to uncertainty is expected to be minimal regarding both bias direction and magnitude. It should also be noted that use of the older data in some of the analyses here carries the assumption that the sources present at that time are the same as current sources, adding uncertainty to results if this is not the case. Separating the data into two 5 year groups (historic and recent) before analysis reduces the potential impact from changes in national- or location-specific source influences and is judged to have a minimal magnitude.

2.8.4 Spatial Representation

Relative to the physical area, there are only a small number of monitors in each location. Since most locations have sparse siting, the monitoring data are assumed to be spatially representative of the locations analyzed here. This includes areas between the ambient monitors that may or may not be influenced by similar local sources of NO_2. For these reasons the uncertainty and bias due to the spatial network may be moderate, although the monitoring network design should have addressed these issues within the available resources and other monitoring constraints. This air quality characterization used all monitors meeting the 75 percent completeness criteria, without taking into account the monitoring objectives or land use for the monitors. Thus, there will be some lack of spatial representation and likely moderate uncertainty due to the inclusion/exclusion of some monitors that are very near local sources (including mobile sources).

2.8.5 Air Quality Adjustment Procedure

The primary uncertainty of the empirical method used to estimate exceedances under the current-standard scenario is due to the uncertainty of the true relationship between the annual mean concentrations and the number of exceedances. The empirical method assumes that if the annual means change then all the hourly concentrations will change proportionately. However,

---

[9] 28 monitors did not have height reported (therefore, 177 + 28 = 205 total number of monitors in named locations)

different sources have different temporal emission profiles, so that applied changes to the annual mean concentrations at monitors may not correspond well to all parts of the concentration distribution equally. Similarly, emissions changes that affect the concentrations at the site with the highest annual mean concentration will not necessarily impact lower concentration sites proportionately. This could result in overestimations in the number of exceedances at lower concentration sites within a location, however it is likely to be minimal given that the highest concentrations typically were measured at the monitoring sites with the highest annual average concentrations within the location (draft TSD, Appendix C). This minimal bias would apply to areas that contain several monitors, such as Boston, New York, or Los Angeles. Universal application of the proportional simulation approach at each of the locations was done for consistency and was designed to preserve the inherent variability in the concentration profile. A few locations were noted that may have an exceptional number of exceedances as a result of the air quality adjustment approach, particularly those locations with few monitoring sites that contained very low annual average concentrations and/or atypical variability in hourly concentrations. These locations (e.g., Miami, Jacksonville, Provo) could contain moderate overestimations at the upper tails of the concentration distribution, leading to bias in number of estimated exceedances at both the upper percentiles and the mean for the scenarios using the air quality simulated to just meet the current standard.

2.8.6 On-Road Concentration Simulation

On-road and ambient monitoring NO_2 concentrations have been shown to be correlated significantly on a temporal basis (e.g., Cape et al., 2004) and motor vehicles are a significant emission source of NO_x, providing support for estimating on-road concentrations using ambient monitoring data. The relationship used in this analysis to estimate on-road NO2 concentrations was derived from data collected in measurement studies containing mostly long-term averaging times, typically 14-days or greater in duration (e.g., Roorda-Knape, 1998; Pleijel et al., 2004; Cape et al, 2004), although one study was conducted over a one-hour time averaging period (Rodes and Holland, 1981). This is considered appropriate in this analysis to estimate on-road hourly concentrations from hourly ambient measures, assuming a direct relationship exists between the short-term peaks to time-averaged concentrations (e.g., hourly on-road NO_2 concentrations are correlated with 24-hour averages). While this should not impact the overall contribution relationship between vehicles and ambient concentrations on roads, the decay constant k will differ for shorter averaging times. The on-road concentration estimation also assumes that concentration changes that occur on-road and at the monitor are simultaneous (i.e., within the hour time period of estimation). Since time-activity patterns of individuals are not considered in this analysis, there is no bias in the number of estimated exceedances. The long-term data used to develop the model were likely collected over variable meteorological conditions (e.g., shifting wind direction) and other influential attributes (e.g., rate of transformation of NO to NO_2 during the daytime versus nighttime hours) than would be observed across shorter time periods. This could result in either over- or under-estimations of concentrations, depending on the time of day. The variability in NO_2 concentration within an hour is also not considered in this analysis, that is, the on-road concentration at a given site will likely vary during the 1-hour time period. If considering personal exposures to individuals within vehicles that are traveling on a road, it is likely that their exposure concentrations would also vary due to differing roadway concentrations. This could also result in either over- or under-estimations of concentrations, depending on the duration of travel and type of road traveled on.

On-road concentrations were not modified in this analysis to account for in-vehicle penetration and decay. This indicates that in-vehicle concentrations would be overestimated if using the on-road concentrations as a surrogate, given that reactive pollutants (e.g., $PM_{2.5}$) tend to have a lower indoor/outdoor (I/O) concentration ratio (Rodes et al., 1998). Chan and Chung (2003) report mean (I/O) ratios of NO_2 for a few roadways and driving conditions in Hong Kong. On highways and urban streets, the value is centered about 0.6 to 1.0, indicating decay of NO_2 as it enters the vehicle.

At locations where traffic counts are very low (e.g., on the order of hundreds/day) the on-road contribution has been shown to be negligible (Bell and Ashenden, 1997; Cape et al., 2004), therefore any rural areas just meeting the standard with minimal traffic volumes would likely have resulted in small overestimations of NO2 concentrations using eq (2). For any monitor that is sited in close proximity of the roadway (14 monitors were sited at <10 m from a major road), on-road concentrations may have been overestimated using eq (2), since the assumption is that the ambient concentration is equivalent to the non-source impacted concentration. In some locations (i.e., Boston, Chicago, Denver, Los Angeles, Miami, St. Louis, and Washington DC), at least half of the monitors used in this analysis are sited < 100 m from a major road (see Table 5, section 2.3.3), a distance noted by some researchers a possibly receiving notable impact from vehicle emissions (e.g., Beckerman et al., 2008). In addition, NO_x is primarily emitted as NO (e.g., Heeb et al., 2008; Shorter et al., 2005), with substantial secondary formation due predominantly to $NO + O_3 \rightarrow NO_2 + O_2$. Numerous studies have demonstrated the O_3 reduction that occurs near major roads, reflecting the transfer of odd oxygen to NO to form NO_2, a process that can impact NO_2 concentrations both on- and downwind of the road. Some studies report NO_2 concentrations increasing just downwind of roadways and that are inversely correlated with O_3 (e.g., Beckerman et al., 2008), suggesting that peak concentration of NO_2 may not always occur on the road, but at a distance downwind. Uncertainty regarding where the peak concentration occurs (on-road or at a distance from the road) in combination with the form of the exponential model used to estimate the on-road concentrations (the highest concentration occurs at zero distance from road) could also lead to overestimation. However, the interpretation of the estimate is what may be most uncertain, that is whether the exceedances are occurring on the road or nearby.

Another source of uncertainty is the extent to which the near-road study locations represent the locations studied in these analyses. The on-road and near-road data were collected in a few locations, most of them outside of the United States. The source mixes (i.e., the vehicle fleet) in study locations may not be representative of the U.S. fleet. Without detailed information characterizing the emissions patterns for the on-road study areas, there was no attempt to match the air quality characterization locations to specific on-road study areas, which might have improved the precision of the estimates. However, since concentration ratios were selected randomly from all the near-road studies and applied to each monitor individually, and since we estimated overall minimum and upper bounds using multiple simulations, the analysis provides a reasonable lower and upper bound estimate of the uncertainty.

2.8.7 Health Benchmark

The choice of potential health effect benchmarks, and the use of those benchmarks to assess risks, can introduce uncertainty into the risk assessment. For example, the potential health effect benchmarks used were based on studies where volunteers were exposed to NO_2 for varying lengths of time. Typically, the NO_2 exposure durations were between 30 minutes and 2 hours. This introduces some uncertainty into the characterization of risk, which compared the potential health effect benchmarks to estimates of exposure over a 1-hour time period. Use of a 1-hour averaging time could over- or under-estimate risks. In addition, the human exposure studies evaluated airways responsiveness in mild asthmatics. For ethical reasons, more severely affected asthmatics and asthmatic children were not included in these studies. Severe asthmatics and/or asthmatic children may be more susceptible than mildly asthmatic adults to the effects of NO_2 exposure. Therefore, the potential health effect benchmarks based on these studies could underestimate risks in populations with greater susceptibility.

Table 25. Summary of qualitative uncertainty analysis for the air quality characterization.

Source	Type	Bias Direction	Magnitude
Air Quality Data	Database quality	both	minimal
Ambient Measurement	Interference	over	moderate
	Vertical siting	under	minimal
Temporal Representation	Scale	none	none
	Missing data	both	minimal
	Years monitored	both	minimal
	Source changes	over	minimal
Spatial Representation	Scale	both	moderate
	Monitor objectives	both	moderate
Air Quality Adjustment	Temporal scale	over	moderate
	Spatial scale	over	moderate
On-Road Simulation	Temporal scale	both	minimal
	Decay	over	minimal
	Spatial scale	over	moderate
	Model used	over	minimal
	Non US studies used	unknown	unknown
Health Benchmarks	Averaging time	unknown	moderate
	Susceptibility	under	moderate

Notes:
Bias Direction: indicates the direction the source of uncertainty is judged to influence either the concentration or risk estimates.
Magnitude: indicates the estimated size of influence.
minimal – influence is either conditional and/or limited to few components in type
moderate – multiple components of uncertainty existed within a given type and act in similar direction, however the presence of all at once may be dependent on certain conditions.
major – multiple components of uncertainty exist within a given type, the components have few limiting conditions, and the components consistently act in similar bias direction.

2.9 References

Bell S and Ashenden TW. (1997). Spatial and temporal variation in nitrogen dioxide pollution adjacent to rural roads. *Water Air Soil Pollut*. 95:87-98.

Beckerman B, Jerrett M, Brook JR, Verma DK, Arain MA, Finkelstein MM. (2008). Correlation of nitrogen dioxide with other traffic pollutants near a major expressway. *Atmos Environ*. 42:275-290.

Bignal KL, Ashmore MR, Headley AD, Stewart K, Weigert K. (2007). Ecological impacts of air pollution from road transport on local vegetation. *Applied Geochemistry*. 22:1265-1271.

Cape JN, Tang YS, van Dijk N, Love L, Sutton MA, Palmer SCF. (2004). Concentrations of ammonia and nitrogen dioxide at roadside verges, and their contribution to nitrogen deposition. *Environ Pollut*. 132:469-478.

Chan AT and Chung MW. (2003). Indoor-outdoor air quality relationships in vehicle: effect of driving environment and ventilation modes. *Atmos Environ*. 37: 3795-3808.

FHWA. (2005). Highway Statistics 2005, Urbanized Areas - 2005, Miles and Daily Vehicle-Miles of Travel (Table HM-71). Available at: http://www.fhwa.dot.gov/policy/ohim/hs05/htm/hm71.htm.

Gilbert NL, Woodhouse S, Stieb DM, Brook JR. (2003). Ambient nitrogen dioxide and distance from a major highway. *Sci Total Environ*. 312:43-46.

Heeb NV, Saxer CJ, Forss A-M, Brühlmann S. (2008). Trends of NO-, NO_2-, and NH_3-emissions from gasoline-fueled Euro-3- to Euro-4-passenger cars. *Atmos Environ*. 42(10):2543-2554.

Maruo YY, Ogawa S, Ichino T, Murao N, Uchiyama M. (2003). Measurement of local variations in atmospheric nitrogen dioxide levels in Sapporo, Japan using a new method with high spatial and high temporal resolution. *Atmos Environ*. 37:1065-1074.

McCurdy TR. (1994). Analysis of high 1 hour NO2 values and associated annual averages using 1988-1992 data. Report to the Office of Air Quality Planning and Standards, Durham NC.

Monn Ch, Carabias V, Junker M, Waeber R, Karrer M, Wanner HU. (1997). Small-scale spatial variability of particulate matter <10 μm (PM_{10}) and nitrogen dioxide. *Atmos Environ*. 31(15)2243-2247.

Nitta H, Sato T, Nakai S, Maeda K, Aoki S, Ono M. (1993). Respiratory health associated with exposure to automobile exhaust. I. Results of cross-section studies in 1979, 1982, 1983. *Arch Environ Health*. 48(1):53-58.

Pleijel H, Karlsson GP, Gerdin EB. (2004). On the logarithmic relationship between NO2 concentration and the distance from a highroad. *Sci Total Environ*. 332:261-264.

Rodes C, Sheldon L, Whitaker D, Clayton A, Fitzgerald K, Flanagan J, DiGenova F, Hering S, Frazier C. (1998). Measuring Concentrations of Selected Air Pollutants Inside California Vehicles. California Environmental Protection Agency, Air Resources Board. Final Report, December 1998.

Rodes CE and Holland DM. (1981). Variations of NO, NO_2 and O_3 concentrations downwind of a Los Angeles freeway. *Atmos Environ*. 15:243-250.

Roorda-Knape MC, Janssen NAH, De Hartog JJ, Van Vliet PHN, Harssema H, Brunekreef B. (1998). Air Pollution from traffic in city districts near major roadways. *Atmos Environ*. 32(11)1921-1930.

Shorter JH, Herndon S, Zahniser MS, Nelson DD, Wormhoudt J, Demerjian KL, Kolb CE. (2005). Real-Time measurements of nitrogen oxide emissions from in-use New York City transit buses using a chase vehicle. *Environ Sci Technol*. 39:7991-8000.

Singer BC, Hodgson AT, Hotchi T, Kim JJ (2004). Passive measurement of nitrogen oxides to assess traffic-related pollutant exposure for the East Bay Children's Respiratory Health Study. *Atmos Environ*. 38:393-403.

US EPA. (2007a). US EPA Air Quality System (AQS). Download Detailed AQS Data. Available at: http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm.

US EPA. (2007b). Field Guide to Air Quality Data (v1.0.0). February 28, 2007. Available at: http://www.epa.gov/ttn/airs/aqsdatamart/documentation/index.htm**.**

US EPA. (2007c). Nitrogen Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment. September 2007. Office of Air Quality Planning and Standards. Available at: http://www.epa.gov/ttn/naaqs/standards/nox/s_nox_cr_pd.html.

US EPA. (2007d)**.** ALLNEICAP Annual 11302007 file posted at: http://www.epa.gov/ttn/chief/net/2002inventory.html#inventorydata.

US EPA. (2007e). Air Trends. Nitrogen Dioxide. http://www.epa.gov/airtrends/nitrogen.html.

Westerdahl D, Fruin S, Sax T, Fine PM, Sioutas C. (2005). Mobile platform measurements of ultrafine particles and associated pollutant concentrations on freeways and residential streets in Los Angeles. *Atmos Environ*. 39:3597-3610.

US EPA. (2007f). Integrated Science Assessment for Oxides of Nitrogen – Health Criteria (First External Review Draft) and Annexes (August 2007). Research Triangle Park, NC: National Center for Environmental Assessment. Available at: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=181712.

3 Exposure Assessment and Health Risk Characterization

3.1 Introduction

This section documents the methodology and input data used in the inhalation exposure assessment for NO_2 conducted in support of the current review of the NO_2 primary NAAQS. Two important components of the analysis include the approach for estimating temporally and spatially variable NO_2 concentrations and simulating contact of humans with these pollutant concentrations. Both air quality and exposure modeling approaches have been used here to generate estimates of 1-hour NO_2 exposures within selected urban areas of the U.S. Details on the approaches used are provided below and include the following:

- Description of the areas assessed and populations considered
- Summary of the air quality modeling methodology and associated input data
- Description of the inhalation exposure model and associated input data
- Evaluation of estimated NO_2 exposures using modeling methodology
- Assessment of the quality and limitations of the input data for supporting the goals of the NO_2 NAAQS exposure analysis.

The selected modeling approach was both time and labor intensive. To date, only the exposure and risk results for the Philadelphia case-study are complete and are presented in this draft document. Location-specific input data for Philadelphia and the other selected case-study areas are presented where collected (mainly meteorological data) to provide information on the relative variability of the input data to be used.

3.1.1 Selection of Study Areas

The selection of areas to include in the exposure analysis takes into consideration the location of field and epidemiology studies, the availability of ambient monitoring and other input data, the desire to represent a range of geographic areas, population demographics, general climatology, and results of the ambient air quality characterization.

Locations of interest were initially identified through a similar statistical analysis of the ambient NO_2 air quality data described above for each site within a location. Criteria were established for selecting sites with high annual means and/or high numbers of exceedances of potential health effect benchmark concentrations. The analysis considered all data combined, as well as the more recent air quality data (2001-2006) separately.

The 90th percentile served as the point of reference for the annual means, and across all complete site-years for 2001-2006, this value was 23.5 ppb. Seventeen locations contained one or more site-years with an annual average concentration at or above the 90th percentile. When combined with the number of 1-hour NO_2 concentrations at or above 200 ppb, only two locations fit these criteria, Philadelphia and Los Angeles. Considering the short-term criterion alone, Detroit contained the greatest number of exceedances of 200 ppb (numbering 12 for years 2001-2006). Two additional locations were selected by considering geographic/climatologic representation and also their historic ambient concentrations. Atlanta (1 exceedance of 200 ppb and a maximum annual average concentration of 26.6 ppb for years 1995-2006) and Phoenix (maximum annual mean concentration of 37.1 ppb for 2001-2006 and 37 exceedances of 200 ppb for years 1995-2006) were selected to represent the southern and western region of the US from the pool of remaining locations with either exceedances of the 90th percentile annual mean concentration or 200 ppb 1-hour.

To summarize, the following 5 urban areas were selected for a detailed exposure analysis:

- Philadelphia, PA
- Atlanta, GA
- Detroit, MI
- Los Angeles, CA
- Phoenix, AZ

3.1.2 Exposure Periods

The exposure periods modeled were 2001 through 2003 to envelop the most recent year of travel demand modeling (TDM) data available for the respective study locations (i.e., 2002) and to include a 3 years of meteorological data to achieve a degree of stability in the dispersion and exposure model estimates.

3.1.3 Populations Analyzed

A detailed consideration of the population residing in each modeled area was included where the exposure modeling was performed. The assessment includes the general population (All Persons) residing in each modeled area and considered susceptible and vulnerable populations as identified in the ISA. These include population subgroups defined from either an exposure or health perspective. The population subgroups identified by the ISA (US EPA, 2007a) that were included and that can be modeled in the exposure assessment include:

- Children (ages 5-18)
- Asthmatic children (ages 5-18)
- All persons (all ages)
- All Asthmatics (all ages)

In addition to these population subgroups, individuals anticipated to be exposed more frequently to NO_2 were considered, including those commuting on roadways and persons residing near major roadways. To date, this document provides a summary of the subpopulations of interest (all asthmatics and asthmatic children), supplemented with additional exposure and risk results for the total population where appropriate.

3.2 Dispersion Modeling

Air quality data used for input to APEX were generated using AERMOD, a steady-state, Gaussian plume model (EPA, 2004). For each identified case-study location, the following steps were performed

1. **Collect and analyze general input parameters.** Meteorological data, processing methodologies used to derive input meteorological fields (e.g., temperature, wind

speed, precipitation), and information on surface characteristics and land use are needed to help determine pollutant dispersion characteristics, atmospheric stability and mixing heights.

2. **Estimate emissions**. The emission sources modeled included, major stationary emission sources, on-road emissions that occur on major roadways, and fugitive emissions.
3. **Define receptor locations**. Three sets of receptors were identified for the dispersion modeling, including ambient monitoring locations, census block centroids, and links along major roadways.
4. **Estimate concentrations at receptors**. Hourly concentrations were estimated for each year of the simulation (years 2001 through 2003) by combining concentration contributions from each of the emission sources and accounting for sources not modeled.

The AERMOD model predictions were then used as input to the APEX model to estimate population exposure concentrations for Philadelphia County. Hourly NO_2 concentrations were estimated for each of 3 years (2001-2003) at each of the defined receptor locations (census blocks and roadway links) using hourly NO_x emission estimates and dispersion modeling. Relevant input data collected for Philadelphia as well some of the data collected as part of the other selected case-study locations to be evaluated in the second draft risk and exposure assessment are presented below.

3.2.1 Meteorological Inputs

All meteorological data used for the AERMOD dispersion model simulations were processed with the AERMET meteorological preprocessor, version 06341. This section describes the input data and processing methodologies used to derive input meteorological fields for each of the five regions of interest.

3.2.1.1 Data Selection

Raw surface meteorological data for the 2001 to 2003 period were obtained from the Integrated Surface Hourly (ISH) Database,[10] maintained by the National Climatic Data Center (NCDC). The ISH data used for this study consists of typical hourly surface parameters (including air and dew point temperature, atmospheric pressure, wind speed and direction, precipitation amount, and cloud cover) from hourly Automated Surface Observing System (ASOS) stations. No on-site observations were used.

Surface meteorological stations for this analysis were those at the major airports of each of the five cities in the study:

- Atlanta: Atlanta Hartsfield International (KATL)
- Detroit: Detroit Metropolitan (KDTW)
- Los Angeles: Los Angeles International (KLAX)
- Philadelphia: Philadelphia International (KPHL)
- Phoenix: Phoenix Sky Harbor International (KPHX).

---

[10] http://www1.ncdc.noaa.gov/pub/data/techrpts/tr200101/tr2001-01.pdf

The selection of surface meteorological stations for each city minimized the distance from the station to city center, minimized missing data, and maximized land-use representativeness of the station site compared to the city center.

The total number of surface observations per station, and the percentage of those observations accepted by AERMET (i.e., those observations that were both not missing and within the expected ranges of values), are shown by Table 26.

Note that instances of calm winds are not rejected by the AERMET processor, but are later treated as calms in the dispersion analysis. There were 2,538 hours in Atlanta with calm winds (10% of the total hourly records), 1,924 in Detroit (7%), 3,190 in Los Angeles (12%), 1,772 in Philadelphia (7%), and 3,559 in Phoenix (14%) (see Table 27).

Table 26. Number of AERMET raw hourly surface meteorology observations and percent acceptance rate, 2001-2003.[a]

Surface Variable	**Atlanta (KATL) N=26281**	**Detroit (KDTW) N=26271**	**Los Angeles (KLAX) N=26276**	**Philadelphia (KPHL) N=26268**	**Phoenix (KPHX) N=26279**
	% Accepted	**% Accepted**	**% Accepted**	**% Accepted**	**% Accepted**
Precipitation	100	100	100	100	100
Station Pressure	99	99	99	99	99
Cloud Height	99	99	99	99	99
Sky Cover	97	97	97	95	97
Horizontal Visibility	99	99	99	99	100
Temperature	99	99	100	99 *	85 *
Dew Point Temperature	99	99	100	99	99
Relative Humidity	99	99	100	99	99
Wind Direction	94	97	92	97	91
Wind Speed	99	99	100	99	99

Notes:
[a] Percentages are rounded down to the nearest integer. All data obtained from the NCDC ISH database.
[*] The majority of unaccepted records are due to values being out of range.
▨ ≤95% of observations were accepted.

Table 27. Number of calms reported by AERMET by year and location.

	Atlanta	**Detroit**	**Los Angeles**	**Philadelphia**	**Phoenix**
2001	917	547	1051	610	1152
2002	856	619	1019	470	1233
2003	765	758	1120	692	1174
Total	**2538**	**1924**	**3190**	**1772**	**3559**

Mandatory and significant levels of upper-air data were obtained from the NOAA Radiosonde Database.[11] Upper air observations show less spatial variation than do surface observations; thus they are both representative of larger areas and measured with less spatial frequency than are surface observations. The selection of upper-air station locations for each city minimized both the proximity of the station to city center and the amount of missing data in the records. The selected stations are:

- Atlanta: Peachtree City (KFFC)
- Detroit: Detroit/Pontiac (KDTX)
- Los Angeles: Miramar Naval Air Station near San Diego (KNKX)
- Philadelphia: Washington Dulles Airport (KIAD)
- Phoenix: Tucson (KTWC).
-

The total number of upper-air observations per station per height interval, and the percentage of those observations accepted by AERMET, are shown in Table 28.

---

[11] http://raob.fsl.noaa.gov/

Table 28. Number and AERMET acceptance rate of upper-air observations 2001-2003.

Height Level	Variable	Atlanta (KFFC)		Detroit (KDTX)		Los Angeles (KNKX)		Philadelphia (KIAD)		Phoenix (KTWC)	
		n	% Accepted	n	% Accepted	n	% Accepted	n	% Accepted	n	% Accepted
Surface	Pressure	2124	100	2125	100	2166	100	2152	100	2143	99 *
	Height	2124	100	2125	100	2166	100	2152	100	2143	99 *
	Temperature	2124	100	2125	100	2166	100	2152	100	2143	87 *
	DewPoint Temperature	2124	100	2125	100	2166	100	2152	100	2143	99 *
	WindDirection	2124	99	2125	100	2166	99	2152	100	2143	100
	WindSpeed	2124	89 *	2125	98 *	2166	99	2152	85 *	2143	100
0-500m	Pressure	3418	100	4577	100	5775	100	4320	100	3611	100
	Height	3418	100	4577	100	5775	100	4320	100	3611	100
	Temperature	3418	100	4577	100	5775	100	4320	100	3611	97 *
	DewPoint Temperature	3418	99	4577	99 *	5775	99	4320	99	3611	100
	WindDirection	3418	29	4577	64	5775	47	4320	63	3611	63
	WindSpeed	3418	29	4577	64	5775	47	4320	62	3611	62
500-1000m	Pressure	4133	100	3059	100	6058	100	3702	100	2797	100
	Height	4133	100	3059	100	6058	100	3702	100	2797	100
	Temperature	4133	100	3059	100	6058	100	3702	100	2797	100
	DewPointTemperature	4133	99 *	3059	98 *	6058	99	3702	99 *	2797	99 *
	WindDirection	4133	62	3059	50	6058	62	3702	73	2797	88
	WindSpeed	4133	62	3059	50	6058	62	3702	73	2797	88
1000-1500m	Pressure	4336	100	4739	100	4473	100	4204	100	1473	100
	Height	4336	100	4739	100	4473	100	4204	100	1473	100
	Temperature	4336	100	4739	100	4473	100	4204	100	1473	100
	DewPointTemperature	4336	96 *	4739	96 *	4473	98 *	4204	97 *	1473	99 *
	WindDirection	4336	72	4739	67	4473	71	4204	71	1473	54
	WindSpeed	4336	72	4739	67	4473	71	4204	71	1473	54
1500-2000m	Pressure	3203	100	3351	100	2478	100	3354	100	1889	100
	Height	3203	100	3351	100	2478	100	3354	100	1889	100
	Temperature	3203	100	3351	100	2478	100	3354	100	1889	100
	DewPointTemperature	3203	95 *	3351	95 *	2478	96 *	3354	95 *	1889	95 *
	WindDirection	3203	50	3351	46	2478	50	3354	50	1889	54
	WindSpeed	3203	50	3351	46	2478	50	3354	50	1889	54
2000-2500m	Pressure	3171	100	3078	100	2229	100	3246	100	3453	100
	Height	3171	100	3078	100	2229	100	3246	100	3453	100
	Temperature	3171	100	3078	100	2229	100	3246	100	3453	100
	DewPointTemperature	3171	94 *	3078	92 *	2229	94 *	3246	93 *	3453	94 *
	WindDirection	3171	52	3078	50	2229	51	3246	50	3453	82

Height Level	**Variable**	**Atlanta (KFFC)**		**Detroit (KDTX)**		**Los Angeles (KNKX)**		**Philadelphia (KIAD)**		**Phoenix (KTWC)**	
		n	**% Accepted**	**n**	**% Accepted**	**n**	**% Accepted**	**n**	**% Accepted**	**n**	**% Accepted**
	WindSpeed	3171	52	3078	50	2229	51	3246	50	3453	82
2500-3000m	Pressure	4318	100	4257	100	2769	100	3736	100	2213	100
	Height	4318	100	4257	100	2769	100	3736	100	2213	100
	Temperature	4318	100	4257	100	2769	100	3736	100	2213	100
	DewPointTemperature	4318	94 *	4257	90 *	2769	90 *	3736	90 *	2213	90 *
	WindDirection	4318	74	4257	71	2769	73	3736	64	2213	55
	WindSpeed	4318	74	4257	71	2769	73	3736	64	2213	55
3000-3500m	Pressure	2840	100	2932	100	2754	100	3614	100	2344	100
	Height	2840	100	2932	100	2754	100	3614	100	2344	100
	Temperature	2840	100	2932	99	2754	100	3614	100	2344	100
	DewPointTemperature	2840	92 *	2932	88 *	2754	91 *	3614	90 *	2344	88 *
	WindDirection	2840	49	2932	48	2754	69	3614	65	2344	54
	WindSpeed	2840	49	2932	48	2754	69	3614	65	2344	54
3500-4000m	Pressure	2964	100	2775	100	2014	100	2830	100	2423	100
	Height	2964	100	2775	100	2014	100	2830	100	2423	100
	Temperature	2964	100	2775	99	2014	100	2830	100	2423	100
	DewPointTemperature	2964	90 *	2775	84 *	2014	86 *	2830	87 *	2423	85 *
	WindDirection	2964	49	2775	49	2014	53	2830	50	2423	55
	WindSpeed	2964	49	2775	49	2014	53	2830	50	2423	55
>4000 m	Pressure	7895	87 *	7279	77 *	6136	82 *	7619	88 *	7483	58 *
	Height	7895	73 *	7279	70 *	6136	64 *	7619	71 *	7483	71 *
	Temperature	7895	100	7279	98 *	6136	99 *	7619	99 *	7483	99 *
	DewPointTemperature	7895	81 *	7279	74 *	6136	76 *	7619	79 *	7483	69 *
	WindDirection	7895	53	7279	59	6136	59	7619	55	7483	65
	WindSpeed	7895	53	7279	59	6136	59	7619	55	7483	65

Notes:

[a] Percentages are rounded down to the nearest integer. All data obtained from the NCDC ISH database.

* The majority of unaccepted records are due to values being out of range

Shading:

░ ≤95 of observations were accepted.

▒ ≤75 of observations were accepted.

▓ ≤50 of observations were accepted.

3.2.2 Surface Characteristics and Land Use Analysis

In addition to the standard meteorological observations of wind, temperature, and cloud cover, AERMET analyzes three principal variables to help determine atmospheric stability and mixing heights: the Bowen ratio[12], surface albedo[13] as a function of the solar angle, and surface roughness[14].

The January 2008 version of AERSURFACE was used to estimate land-use patterns and calculate the Bowen ratio, surface albedo, and surface roughness as part of the AERMET processing. AERSURFACE uses the US Geological Survey (USGS) National Land Cover Data 1992 archives (NLCD92)[15]. Three to four land-use sectors were manually identified around the surface meteorological stations using this land-use data. These land-use sectors are used to identify the Bowen ratio and surface albedo, which are assumed to represent an area around the station of radius 10 km, and to calculate surface roughness by wind direction.

A monthly temporal resolution was used for the Bowen ratio, albedo, and surface roughness for all five meteorological sites. Because the five sites were located at airports, a lower surface roughness was calculated for the 'Commercial/Industrial/Transportation' land-use type to reflect the dominance of transportation land cover rather than commercial buildings. Los Angeles and Phoenix are arid regions, which increases the calculated albedo and Bowen ratio values and decreases the surface roughness values assigned to the 'Shrubland' and 'Bare Rock/Sand/Clay' land-use types to reflect a more desert-like area. Philadelphia and Detroit each have at least one winter month of continuous snow cover, which tends to increase albedo, decrease Bowen ratio, and decrease surface roughness for most land-use types during the winter months compared to snow-free areas.

Seasons were assigned for each site based on 1971-2000 NCDC 30-year climatic normals and on input from the respective state climatologists. Table 29 provides the seasonal definitions for each city.

Table 29. Seasonal specifications by study location.

Location	Winter (continuous snow)	Winter (no snow)	Spring	Summer	Fall
Atlanta		Dec, Jan, Feb	Mar, Apr, May	Jun, Jul, Aug	Sep, Oct, Nov
Detroit	Dec, Jan, Feb, Mar		Apr, May	Jun, Jul, Aug	Sep, Oct, Nov

---

[12] For any moist surface, the Bowen Ratio is the ratio of heat energy used for sensible heating (conduction and convection) to the heat energy used for latent heating (evaporation of water or sublimation of snow). The Bowen ratio ranges from about 0.1 for the ocean surface to more than 2.0 for deserts. Bowen ratio values tend to decrease with increasing surface moisture for most land-use types.

[13] The ratio of the amount of electromagnetic radiation reflected by the earth's surface to the amount incident upon it. Value varies with surface composition. For example, snow and ice vary from 80% to 85% and bare ground from 10% to 20%.

[14] The presence of buildings, trees, and other irregular land topography that is associated with its efficiency as a momentum sink for turbulent air flow, due to the generation of drag forces and increased vertical wind shear.

[15] http://seamless.usgs.gov/

Los Angeles			Apr, May, Jun	Jul, Aug, Sep	Oct, Nov, Dec, Jan, Feb, Mar
Philadelphia	Dec, Jan, Feb		Mar, Apr, May	Jun, Jul, Aug	Sep, Oct, Nov
Phoenix			Apr, May, Jun	Jul, Aug, Sep	Oct, Nov, Dec, Jan, Feb, Mar

Season definitions provided by the AERSURFACE manual as follows:

Winter (continuous snow):	Winter with continuous snow on ground
Winter (no snow):	Late autumn after frost and harvest, or winter with no snow
Spring:	Transitional spring with partial green coverage or short annuals
Summer:	Midsummer with lush vegetation
Fall:	Autumn with unharvested cropland

Further discussion of the land use and surface analysis, as well as a discussion of the difference in results from employing the new AERSURFACE tool is given in Appendix E.

3.2.3 Meteorological Analysis

The AERMET application location and elevation were taken as the center of each modeled city, estimated using Google Earth version 4.2.0198.2451 (beta). They are as follows:

- Atlanta: 33.755 °N, 84.391 °W, 306 m
- Detroit: 42.332 °N, 83.048 °W, 181 m
- Los Angeles: 34.053 °N, 118.245 °W, 91 m
- Philadelphia: 39.952 °N, 75.164 °W, 12 m
- Phoenix: 33.448 °N, 112.076 °W, 330 m

For each site in this study, the 2001-2003 AERSURFACE processing was run three times – once assuming the entire period was drier than normal, once assuming the entire period was wetter than normal, and once assuming the entire period was of average precipitation accumulation. These precipitation assumptions influence the Bowen ratio, discussed above.

To create meteorological input records that best represent the given city for each of the three years, the resulting surface output files for each site were then pieced together on a month-by-month basis, with selection based on the relative amount of precipitation in each month. Any month where the actual precipitation amount received was at least twice the 1971-2000 NCDC 30-year climatic normal monthly precipitation amount was considered wetter than normal, while any month that received less than half the normal amount of precipitation amount was considered drier than normal; all other months were considered to have average surface moisture conditions.

Table 30 indicates the surface moisture condition for each month-location combination for this study. The final meteorological record includes wet conditions for the Bowen ratio for the month-location combinations shown in green and dry conditions for those in orange. All other region-month combinations used an average Bowen ratio.

Table 30. Monthly precipitation compared to NCDC 30-year climatic normal, 2001-2003.

City	2001											
	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
Atlanta	55.9%	77.6%	169.3%	91.7%	77.1%	185.9%	50.0%	28.3%	53.9%	25.8%	21.4%	58.7%
Detroit	27.4%	125.6%	29.8%	89.2%	106.0%	61.6%	39.4%	82.0%	108.2%	280.7%	75.8%	66.0%
Los Angeles	157.7%	237.0%	52.7%	175.0%	4.9%	0.0%	0.0%	0.0%	0.0%	10.9%	119.9%	60.0%
Philadelphia	74.8%	103.6%	144.2%	43.9%	102.9%	180.1%	29.9%	26.0%	67.1%	30.6%	17.9%	64.6%
Phoenix	214.4%	111.5%	71.7%	429.9%	14.8%	13.1%	68.8%	48.6%	0.0%	3.0%	28.0%	95.9%
City	**2002**											
	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
Atlanta	105.5%	54.7%	102.2%	50.5%	86.0%	77.2%	50.8%	21.6%	157.5%	191.7%	131.4%	137.3%
Detroit	132.7%	73.9%	66.6%	123.9%	104.2%	25.4%	133.7%	36.8%	60.1%	50.7%	83.5%	47.8%
Los Angeles	76.1%	17.7%	28.4%	27.5%	42.7%	4276.6%	15656.2%	1358.3%	829.8%	1730.1%	277.0%	216.6%
Philadelphia	69.9%	17.7%	96.4%	52.7%	89.2%	93.9%	51.0%	59.0%	89.1%	202.7%	94.2%	117.9%
Phoenix	6.2%	0.0%	7.7%	25.2%	0.0%	0.0%	119.7%	0.0%	66.7%	46.8%	57.7%	18.8%
City	**2003**											
	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
Atlanta	40.2%	75.4%	132.1%	95.6%	252.5%	198.5%	100.2%	95.2%	59.5%	47.9%	102.3%	70.9%
Detroit	13.4%	34.1%	59.8%	64.9%	155.0%	59.7%	28.2%	109.0%	124.7%	99.7%	99.2%	86.9%
Los Angeles	0.0%	121.7%	69.4%	78.7%	397.0%	0.0%	0.0%	0.0%	0.0%	111.5%	71.1%	64.9%
Philadelphia	53.2%	165.0%	102.7%	62.0%	108.5%	246.2%	46.5%	86.1%	120.8%	162.8%	92.9%	158.6%
Phoenix	68.8%	413.1%	48.2%	69.3%	0.0%	0.0%	61.6%	45.2%	36.2%	26.9%	90.6%	20.5%

Shading:	
▓	At least twice the normal precipitation level
□	Less than twice the normal precipitation level and greater than half the normal amount
░	Less than or equal to half the normal monthly precipitation amount

3.2.4 On-Road Emissions Preparation

3.2.4.1 Philadelphia County Data Sources

Information on traffic data in the Philadelphia area was obtained from the Delaware Valley Regional Planning Council (DVRPC[16]) via their most recent, baseline travel demand modeling (TDM) simulation – that is, the most recent simulation calibrated to match observed traffic data. DVRPC provided the following files.

- Shapefiles of TDM outputs for the 2002 baseline year for all links in their network.
- Input files for the MOBILE6.2 emissions model that characterize local inputs that differ from national defaults, including fleet registration distribution information.
- Postprocessing codes they employ for analysis of TDM outputs into emission inventory data, to ensure as much consistency as possible between the methodology used for this study and that of DVRPC. These include DVRPC's versions of the local SVMT.DEF, HVMT.DEF, and FVMT.DEF MOBILE6.2 input files describing the vehicle miles traveled (VMT) by speed, hour, and facility, respectively, by county in the Delaware Valley area.
- A lookup table used to translate average annual daily traffic (AADT) generated by the TDM into hourly values.

Although considerable effort was expended to maintain consistency between the DVRPC approach to analysis of TDM data and that employed in this analysis, including several personal communications with agency staff on data interpretation, complete consistency was not possible due to the differing analysis objectives. The DVRPC creates countywide emission inventories. This study created spatially and temporally resolved emission strengths for dispersion modeling.

Emission Sources and Locations

The TDM simulation's shapefile outputs include annual average daily traffic (AADT) volumes and a description of the loaded highway network. The description of the network consists of a series of nodes joining individual model links (i.e., roadway segments) to which the traffic volumes are assigned, and the characteristics of those links, such as endpoint location, number of lanes, link distance, and TDM-defined link daily capacity.[17]

To reduce the scope of the analysis, the full set of links in the DVRPC network was first filtered to include only those roadway types considered *major* (i.e., freeway, parkway, major arterial, ramp), and that had AADT values greater than 15,000 vehicles per day (one direction).

However, the locations of links in the model do not necessarily agree well with the roads they are attempting to represent. While the exact locations of the links may not be mandatory for DVRPC's travel demand modeling, the impacts of on-road emissions on fixed receptors is crucially linked to the distance between the roadways and receptors. Hence, it was necessary to modify the link locations from the TDM to the best known locations of the actual roadways.

---

[16] http://www.dvrpc.org/

[17] The TDM capacity specifications are not the same as those defined by the Highway Capacity Manual (HCM). Following consultation with DVRPC, the HCM definition of capacity was used in later calculations discussed below.

The correction of link locations was done based on the locations of the nodes that define the end points of links with a GIS analysis, as follows.

A procedure was developed to relocate TDM nodes to more realistic locations. The nodes in the TDM represent the endpoints of links in the transportation planning network and are specified in model coordinates. The model coordinate system is a Transverse Mercator projection of the TranPlan Coordinate System with a false easting of 31068.5, false northing of -200000.0, central meridian: -75.00000000, origin latitude of 0.0, scale factor of 99.96, and in units of miles. The procedure moved the node locations to the true road locations and translated to dispersion model coordinates. The Pennsylvania Department of Transportation (PA DOT) road network database[18] was used as the specification of the true road locations. The nodes were moved to coincide with the nearest major road of the corresponding roadway type using a built-in function of ArcGIS. Once the nodes had been placed in the corrected locations, a line was drawn connecting each node pair to represent a link of the adjusted planning network.

To determine hourly traffic on each link, the AADT volumes were converted to hourly values by applying DVRPC's seasonal and hourly scaling factors. To determine hourly traffic on each link, the AADT volumes were converted to hourly values by applying DVRPC's seasonal and hourly scaling factors. The heavy-duty vehicle fraction – which is assumed by DVRPC to be about 6% in all locations and times – was also applied.[19] Another important variable, the number of traffic signals occurring on a given link, was taken from the TDM link-description information.

Several of these parameters are shown in the following set of tables.

- Table 31: hourly scaling factors
- Table 32: seasonal scaling factors
- Table 33: number of signals per roadway mile
- Table 34: statistical summaries of AADT volumes for links included in the study.

---

[18] http://www.pasda.psu.edu/

[19] As shown by Figure 13, NO_x emissions from HDVs tend to be higher than their LDV counterparts by about a factor of 10. However, the HDV fraction is less than 10% of the total VMT in most circumstances, mitigating their influence on composite emission factors, although this mitigating effect is less pronounced at some times than others. For example, nighttimes on freeways tend to show a smaller reduction in HDV volume than in total volume, and thus an increased HDV fraction. This effect is not captured in most TDMs or emission postprocessors and – both to maintain consistency with the local MPO's vehicle characterizations and emissions modeling and due to lack of other relevant data – was also not included here. The net result of this is likely to be slightly underestimated emissions from major freeways during late-night times.

Table 31. Hourly scaling factors (in percents) applied to Philadelphia County AADT volumes.

Road Type	Region	0:00	1:00	2:00	3:00	4:00	5:00	6:00	7:00	8:00	9:00	10:00	11:00
Freeway	CBD	1.23	0.86	0.74	0.84	1.23	2.50	4.87	6.52	6.47	5.75	4.99	5.02
	Fringe	1.23	0.86	0.74	0.84	1.23	2.50	4.87	6.52	6.47	5.75	4.99	5.02
	Urban	1.23	0.86	0.74	0.84	1.23	2.50	4.87	6.52	6.47	5.75	4.99	5.02
	Suburban	0.96	0.64	0.54	0.61	0.90	2.16	5.39	7.33	6.85	5.52	4.90	4.94
	Rural	0.71	0.48	0.38	0.48	0.95	2.54	6.05	7.77	6.79	5.22	4.64	4.78
Arterial	CBD	1.43	0.96	0.61	0.50	0.58	1.17	2.89	5.50	6.87	5.87	5.37	5.17
	Fringe	1.53	0.97	0.62	0.47	0.54	1.10	2.99	5.77	6.53	5.60	5.14	4.86
	Urban	1.13	0.68	0.52	0.45	0.63	1.68	4.26	6.68	6.86	5.47	5.09	5.17
	Suburban	0.70	0.40	0.32	0.33	0.55	1.71	4.51	7.04	6.84	5.37	4.95	5.36
	Rural	0.60	0.36	0.34	0.41	0.77	2.29	5.47	7.37	6.62	5.36	5.09	5.35
Local	CBD	1.11	0.71	0.45	0.37	0.41	0.97	2.39	4.82	6.72	6.50	4.60	4.93
	Fringe	1.00	0.55	0.37	0.21	0.39	0.98	1.98	5.31	5.91	5.78	5.14	5.19
	Urban	1.19	0.74	0.53	0.43	0.54	1.32	3.37	6.54	6.86	5.09	4.65	4.95
	Suburban	0.53	0.29	0.21	0.20	0.37	1.25	3.94	7.51	7.50	5.24	4.66	5.22
	Rural	0.55	0.32	0.25	0.30	0.57	1.89	5.26	7.93	6.84	4.94	4.57	4.89
Ramp	CBD	1.23	0.86	0.74	0.84	1.23	2.50	4.87	6.52	6.47	5.75	4.99	5.02
	Fringe	1.23	0.86	0.74	0.84	1.23	2.50	4.87	6.52	6.47	5.75	4.99	5.02
	Urban	1.23	0.86	0.74	0.84	1.23	2.50	4.87	6.52	6.47	5.75	4.99	5.02
	Suburban	0.96	0.64	0.54	0.61	0.90	2.16	5.39	7.33	6.85	5.52	4.90	4.94
	Rural	0.71	0.48	0.38	0.48	0.95	2.54	6.05	7.77	6.79	5.22	4.64	4.78
Road Type	**Region**	**12:00**	**13:00**	**14:00**	**15:00**	**16:00**	**17:00**	**18:00**	**19:00**	**20:00**	**21:00**	**22:00**	**23:00**
Freeway	CBD	4.97	5.77	6.40	6.60	7.02	6.76	6.27	4.20	3.52	3.06	2.50	1.92
	Fringe	4.97	5.77	6.40	6.60	7.02	6.76	6.27	4.20	3.52	3.06	2.50	1.92
	Urban	4.97	5.77	6.40	6.60	7.02	6.76	6.27	4.20	3.52	3.06	2.50	1.92
	Suburban	5.05	5.19	5.90	6.80	7.58	7.67	6.51	4.27	3.34	2.97	2.32	1.66
	Rural	4.92	5.01	5.75	7.12	7.88	8.18	6.27	4.31	3.45	2.97	2.10	1.27
Arterial	CBD	5.27	5.57	5.95	6.63	7.39	7.81	6.36	4.78	4.05	3.74	3.18	2.36
	Fringe	5.52	5.40	6.08	6.88	7.36	8.08	6.24	4.98	4.21	3.82	3.13	2.19
	Urban	5.42	5.54	6.16	7.04	7.39	7.42	6.08	4.74	3.77	3.31	2.61	1.93
	Suburban	5.75	5.71	6.12	7.05	7.66	7.98	6.42	4.81	3.83	3.13	2.15	1.34
	Rural	5.55	5.50	6.00	7.11	7.82	7.98	6.26	4.48	3.50	2.80	1.88	1.11
Local	CBD	6.26	6.74	6.88	6.78	7.64	8.10	6.57	4.96	3.96	3.02	2.88	2.25
	Fringe	6.31	5.64	6.64	7.32	7.85	9.52	6.25	5.50	5.29	2.87	2.46	1.56
	Urban	5.25	5.40	6.44	7.35	7.80	7.85	6.41	5.02	4.04	3.46	2.79	2.01
	Suburban	5.78	5.57	6.01	7.11	8.20	8.98	6.83	5.02	3.83	2.90	1.82	1.05
	Rural	5.20	5.11	5.89	7.41	8.53	8.93	6.75	4.82	3.64	2.70	1.73	0.99
Ramp	CBD	4.97	5.77	6.40	6.60	7.02	6.76	6.27	4.20	3.52	3.06	2.50	1.92
	Fringe	4.97	5.77	6.40	6.60	7.02	6.76	6.27	4.20	3.52	3.06	2.50	1.92
	Urban	4.97	5.77	6.40	6.60	7.02	6.76	6.27	4.20	3.52	3.06	2.50	1.92
	Suburban	5.05	5.19	5.90	6.80	7.58	7.67	6.51	4.27	3.34	2.97	2.32	1.66
	Rural	4.92	5.01	5.75	7.12	7.88	8.18	6.27	4.31	3.45	2.97	2.10	1.27

Table 32. Seasonal scaling factors applied to Philadelphia County AADT volumes.

Season	Road Type	Factor
Winter	Freeway	0.945
Spring	Freeway	1.006
Summer	Freeway	1.041
Autumn	Freeway	1.009
Winter	Arterial	0.942
Spring	Arterial	1.004
Summer	Arterial	1.041
Autumn	Arterial	1.013
Winter	Local	0.933
Spring	Local	1.012
Summer	Local	1.05
Autumn	Local	1.004
Winter	Ramp	0.944
Spring	Ramp	1.005
Summer	Ramp	1.041
Autumn	Ramp	1.011

Table 33. Signals per mile, by link type, applied to Philadelphia County AADT volumes.

	Region Type				
Functional Class	**CBD**	**Fringe**	**Rural**	**Suburban**	**Urban**
Freeway	0	0	0	0	0
Local	8	6	1.5	3	5
Major Arterial	8	6	1	2	4
Minor Arterial	8	6	1.3	2	4
Parkway	4	2	0.5	1	1.5
Ramp	0	0	0	0	0

Table 34. Statistical summary of AADT volumes (one direction) for Philadelphia County AERMOD simulations.

Statistic	Road Type	CBD	Fringe	Suburban	Urban
Count	Arterial	186	58	210	580
	Freeway	11	10	107	98
	Ramp	0	4	3	1
Minimum AADT	Arterial	15088	15282	15010	15003
	Freeway	15100	18259	15102	15100
	Ramp		16796	15679	16337
Maximum AADT	Arterial	44986	44020	48401	44749
	Freeway	39025	56013	68661	68661
	Ramp		40538	24743	16337
Average AADT	Arterial	21063	21196	20736	22368
	Freeway	25897	40168	33979	31294
	Ramp		24468	18814	16337

Emission Source Strength

On-road mobile emission factors were derived from the MOBILE6.2 emissions model as follows. The DVRPC-provided external data files describing the vehicle miles traveled (VMT) distribution by speed, functional class, and hour, as well as the registration distribution and *Post-1994 Light Duty Gasoline Implementation* for Philadelphia County were all used in the model runs without modification. To further maintain consistency with the recent DVRPC inventory simulations and maximize temporal resolution, the DVRPC's seasonal particulate matter (PM) MOBILE6 input control files were also used. These files include county-specific data describing the vehicle emissions inspection and maintenance (I/M) programs, on-board diagnostics (OBD) start dates, VMT mix, vehicle age distributions, default diesel fractions, and representative minimum and maximum temperatures, humidity, and fuel parameters. The simulations are designed to calculate average running NO_x emission factors.[20]

These input files were modified for the current project to produce running NOx emissions in grams per mile for a specific functional class (Freeway, Arterial, or Ramp) and speed. Iterative MOBILE6.2 simulations were conducted to create tables of average Philadelphia County emission factors resolved by speed (2.5 to 65 mph), functional class, season, and year (2001, 2002, or 2003) for each of the eight combined MOBILE vehicle classes (LDGV, LDGT12, LDGT34, HDGV, LDDV, LDDT, HDDV, and MC)[21]. The resulting tables were then consolidated into speed, functional class, and seasonal values for combined light- and heavy-duty vehicles. Figure 13 shows an example of the calculated emission factors for Autumn, 2001.

---

[20] Basing the present emissions model input files on MPO-provided PM, rather than NO_x input files should not cause confusion. MPO-provided PM files were used because they contain quarterly rather than annual or biannual information. In all cases the output species were modified to produce gaseous emissions. Further, many of the specified input parameters do not affect PM emissions, but were included by the local MPO to best represent local conditions, which were preserved in the present calculations of NO_x emissions. This usage is consistent with the overall approach of preserving local information wherever possible.

[21] HDDV - Heavy-Duty Diesel Vehicle, HDGV - Heavy-Duty Gasoline Vehicle, LDDT - Light-Duty Diesel Truck, LDDV - Light-Duty Diesel Vehicle, LDGT12 - Light-Duty Gasoline Truck with gross vehicle weight rating ≤ 6,000 lbs and a loaded vehicle weight of ≤ 5,750 lbs, LDGT 34 - Light-Duty Gasoline Truck with gross vehicle weight rating between 6,001 - 8,500 and a loaded vehicle weight of ≤ 5,750 lbs, LDGV - Light-Duty Gasoline Vehicle, MC - Motorcycles.

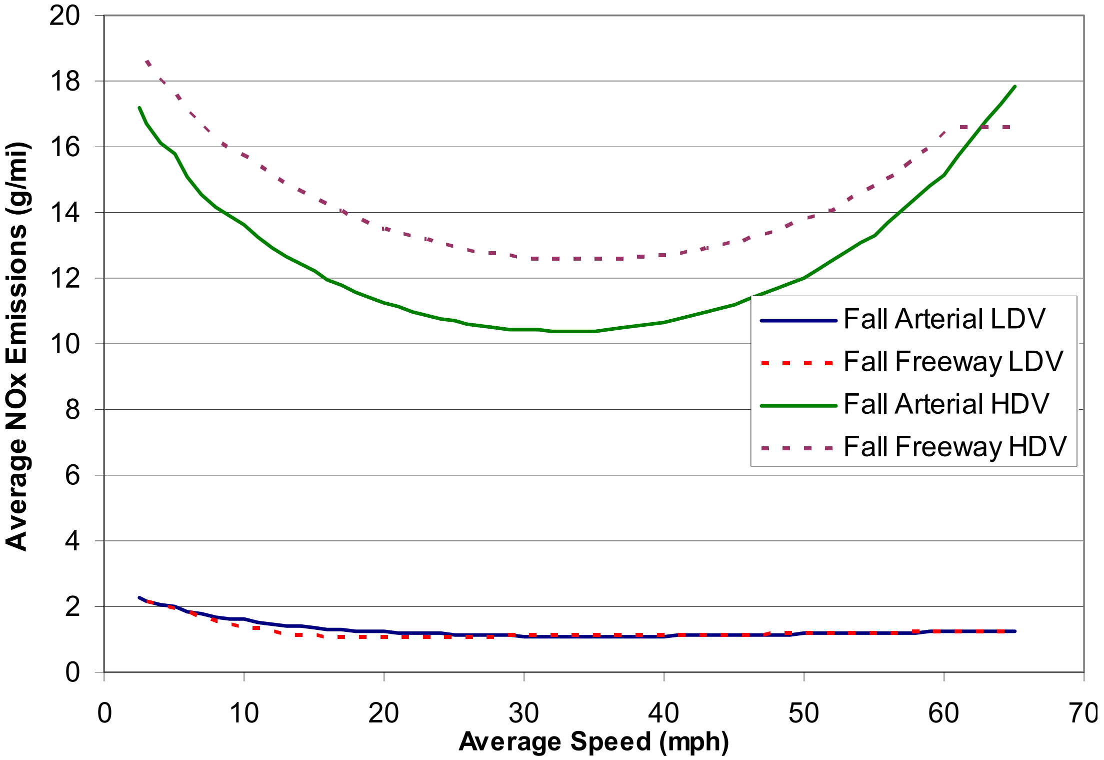


Figure 13. Example of Light- and heavy-duty vehicle NO_x emissions grams/mile (g/mi) for arterial and freeway functional classes, 2001.

To determine the emission strengths for each link for each hour of the year, the Philadelphia County average MOBILE6.2 speed-resolved emissions factor tables were merged with the TDM link data, which had been processed to determine time-resolved speeds. The speed calculations were made as follows.

The spatial-mean speed of each link at each time was calculated following the methodology of the Highway Capacity Manual.[22] Generally, the spatial-mean speed calculation is a function of the time-resolved volume-to-capacity ratio, with capacity the limiting factor. In the case of freeway calculations, this is determined by the HDV fraction, posted speed, and the general hilliness of the terrain, which was assumed to be uniformly flat for this region. The case of arterials without intersections is similar, but also considers urban effects. The case of arterials with intersections further considers the number of signals and length of each link and signalization parameters. It was assumed that all signals are identical, operating with a 120-second cycle and a protected left turn phase. Each link's speed is calculated independently. For example, a series of adjacent arterial links could show very different spatial-mean speeds if one link contains one or more intersections. That is, no up- or down-stream impacts are considered on individual link speeds. Speeds were assumed to be equal for light- and heavy-duty vehicles.

Table 35 shows the resulting average speed for each functional class within each TDM region. Several values are shown as N/A, due to the focus only on major links as discussed above.

---

[22] As defined in Chapter 9 of Recommended Procedure for Long-Range Transporation Planning and Sketch Planning, NCHRP Report 387, National Academy Press, 1997. 151 pp., ISBN No: 0-309-060-58-3.

Table 35. Average calculated speed by link type.

	Average Speed (mph)				
	CBD	**Fringe**	**Suburban**	**Urban**	**Rural**
Ramp	N/A	35	35	35	N/A
Arterial	34	31	44	32	N/A
Freeway	51	62	66	62	N/A

The resulting emission factors were then coupled with the TDM-based activity estimates to calculate emissions from each of the 1,268 major roadway links. However, many of the links were two sides of the same roadway segment. To speed model execution time, those links that could be combined into a single emission source were merged together. This was done only for the 628 links (314 pairs) where opposing links were paired in space and exhibited similar activity levels within 20% of each other.

Other Emission Parameters

Each roadway link is characterized as a rectangular area source with the width given by the number of lanes and an assumed universal lane width of 12 ft (3.66 m). The length and orientation of each link is determined as the distance and angle between end nodes from the adjusted TDM locations. In cases where the distance is such that the aspect ratio is greater than 100:1, the links were disaggregated into sequential links, each with a ratio less than that threshold. There were 27 links that exceeded this ratio and were converted to 55 segmented sources. Thus, the total number of area sources included in the dispersion simulations is 982. Table 36 shows the distribution of on-road area source sizes. Note that there are some road segments whose length was zero after GIS adjustment of node location. This is assumed to be compensated by adjacent links whose length will have been expanded by a corresponding amount.

Table 36. On-road area source sizes.

	Segment Width (m)	**Lanes**	**Segment Length (m)**
Minimum	3.7	1.0	0.0
Median	11.0	3.0	220.6
Average	13.7	3.8	300.2
1-σ Deviation	7.7	2.1	259.5
Maximum	43.9	12.0	1340.2

Resulting daily emission estimates were temporally allocated to hour of the day and season using MOBILE6.2 emission factors, coupled with calculated hourly speeds from the postprocessed TDM and allocated into SEASHR emission profiles for the AERMOD dispersion model. That is, 96 emissions factors are attributed to each roadway link to describe the emission strengths for 24 hours of each day of each of four seasons and written to the AERMOD input control file.

The release height of each source was determined as the average of the light- and heavy-duty vehicle fractions, with an assumed light- and heavy-duty emission release heights of 1.0 ft

(0.3048 m) and 13.1 ft (4.0 m), respectively.[23] Because AERMOD only accepts a single release height for each source, the 24-hour average of the composite release heights is used in the modeling. Since surface-based mobile emissions are anticipated to be terrain following, no elevated or complex terrain was included in the modeling. That is, all sources are assumed to lie in a flat plane.

3.2.5 Stationary Sources Emissions Preparation

3.2.5.1 Philadelphia Data Sources

Data for the parameterization of major point sources in Philadelphia comes primarily from two sources: the 2002 National Emissions Inventory (NEI; US EPA, 2007b) and Clean Air Markets Division (CAMD) Unit Level Emissions Database (US EPA, 2007c). These two databases have complimentary information.

The NEI database contains stack locations, emissions release parameters (i.e., height, diameter, exit temperature, exit velocity), and annual emissions for 707 NO_x-emitting stacks (206 of which are considered fugitive release points) in Philadelphia County. The CAMD database, on the other hand, has information on hourly NO_x emission rates for all the units in the US, where the units are the boilers or equivalent, each of which can have multiple stacks. The alignment of facilities between the two databases is not exact, however. Some facilities listed in the NEI, are not included in the CAMD database. Of those facilities that do match, in many cases there is no clear pairing between the individual stacks assigned within the databases.

3.2.5.2 Data Source Alignment

To align the information between the two databases and extract the useful portion of each for dispersion modeling, the following methodology was used.

1. Attention was limited stacks within the NEI data base that (a) lie within Philadelphia County and (b) were part of a facility with total emissions from all stacks exceeding 100 tpy NO_x.
2. Individual stacks that had identical stack physical parameters and were co-located within about 10 m were combined to be simulated as a single stack with their emissions summed.
3. All fugitive releases were removed from the list, to be analyzed as a separate source group.

The resulting 19 distinct, combined stacks from the NEI are shown in Table 37.

The CAMD database was then queried for facilities that matched the facilities identified from the NEI database. Facility matching was done on the facility name, Office of Regulatory Information Systems (ORIS) identification code (when provided) and facility total emissions to ensure a best match between the facilities. Once facilities were paired, individual units and

---

[23] 4.0 m includes plume rise from truck exhaust stacks. See Diesel Particulate Matter Exposure Assessment Study for the Ports of Los Angeles and Long Beach, State of California Air Resources Board, Final Report, April 2006.

stacks in the data bases were paired, based on annual emission totals. Table 38 shows the matching scheme for the seven major facilities in Philadelphia County.[24]

In Table 38, there are sometimes multiple CAMD units that pair with a single NEI combined stack. In these cases the hourly emission rates from the matching CAMD units are summed for each hour. For example, in the case of stack 859 for "Sunoco, Inc – Philadelphia" five CAMD hourly records are summed into a single hourly record. Then each resulting hourly value is scaled by a factor of 1032.8 / 938.9 = 1.10, so that the annual total matches the NEI annual total.

Similarly, there are sometimes multiple combined stacks that pair with single units. In this case the CAMD values are disaggregated according to NEI-defined stack contributions. For example, "Sunoco, Inc – Philadelphia" stack 855's profile is determined by taking the hourly profile from CAMD unit number 52106-150101, and scaling each value by a factor of 26.2 tpy / 48.2 tpy total = 0.54. Then each resulting hourly value is scaled by a factor of 48.2/162.1 = 0.3 so that the sum of the annual totals for the 4 stacks corresponding to unit number 52106-150101 matches the NEI total. For consistency, in each case the 2001 and 2003 hourly emission profiles were determined using the same scaling factors, but applied to the respective CAMD emission profile.

It is clear from Table 38 that most facilities agree well in total annual NO_x emissions between the two databases. However, in the case of the "Sunoco Chemicals (Former Allied Signal)" facility, nearly half of the NEI emissions (without fugitives) do not appear in the CAMD database. The reason for this is unknown and no information was readily available on the relative accuracy of the two databases.

Figure 14 illustrates the discrepancy versus fraction of hours with positive emissions, according to the CAMD data base. The figure suggests that the discrepancies are not primarily the result of facilities with episodic emissions (i.e., "peak load" facilities). Although there is good agreement on facility-wide emissions between the two data bases, there are larger discrepancies between CAMD unit emissions and NEI stack emissions. This is to be expected given the discrepancy in resolution between the two data bases.

---

[24] Note that Jefferson Smurfit does not exist in the CAMD database. The matching here was based on facility types as follows. Smurfit in PA was taken as a packaging/recycling facility, and the stack assumed to be a Cogen facility, based on information in the NEEDS database (http://www.epa.gov/interstateairquality/pdfs/NEEDS-NODA.xls). The best matched cogen plant in Philadelphia County in both the NEEDS and CAMD database is the Gray's Ferry Cogen Partnership (ORIS 54785), which was a reasonable match for Smurfit's total emissions. It was assumed that the hourly emission profile also matches well.

Table 37. Combined stacks parameters for stationary NO_x emission sources in Philadelphia County.

Stack No	NEI Site ID	Facility Name	SIC Code	NAICS Code	ORIS Facility Code	Stack Emiss (tpy)	Stack X (deg)	Stack Y (deg)	Stack Ht (m)	Exit Temp (K)	Stack Diam (m)	Exit Vel (m/s)	Facility Emiss Incl Fugitive (tpy)
817	NEIPA2218	EXELON GENERATION CO - DELAWARE STATION	4911	221112	3160	4.82	-75.1358	39.96769	49	515	4.2	0	297.8
818	NEIPA2218	EXELON GENERATION CO - DELAWARE STATION	4911	221112	3160	287.8	-75.1358	39.96769	64	386	3.7	17	297.8
819	NEI40720	JEFFERSON SMURFIT CORPORATION (U S)	2631	32213		0.148	-75.2391	40.03329	16	477	0.4	19	228.4
820	NEI40720	JEFFERSON SMURFIT CORPORATION (U S)	2631	32213		113.8	-75.2391	40.03329	53	427	2.4	10	228.4
821	NEI40720	JEFFERSON SMURFIT CORPORATION (U S)	2631	32213		114.46	-75.2391	40.03329	53	477	2.4	12	228.4
855	NEI40723	Sunoco Inc. - Philadelphia	2911	32411		26.2	-75.2027	39.92535	24	450	2.1	9	3112.2
856	NEI40723	Sunoco Inc. - Philadelphia	2911	32411		1.3	-75.2003	39.91379	24	644	1.5	22	3112.2
857	NEI40723	Sunoco Inc. - Philadelphia	2911	32411		1.4	-75.203	39.92539	25	511	1.9	10	3112.2
858	NEI40723	Sunoco Inc. - Philadelphia	2911	32411		19.3	-75.2027	39.92535	25	527	1.9	11	3112.2
859	NEI40723	Sunoco Inc. - Philadelphia	2911	32411		1032.8	-75.2124	39.90239	61	489	5.8	11	3112.2
860	NEI7330	SUNOCO CHEMICALS (FORMER ALLIED SIGNAL)	2869	325998		0.033	-75.0715	40.00649	5	476	0.5	7	160.9
861	NEI7330	SUNOCO CHEMICALS	2869	325998		49.1	-75.0715	40.00649	41	422	1.4	22	160.9

Stack No	NEI Site ID	Facility Name	SIC Code	NAICS Code	ORIS Facility Code	Stack Emiss (tpy)	Stack X (deg)	Stack Y (deg)	Stack Ht (m)	Exit Temp (K)	Stack Diam (m)	Exit Vel (m/s)	Facility Emiss Incl Fugitive (tpy)
		(FORMER ALLIED SIGNAL)											
862	NEI7330	SUNOCO CHEMICALS (FORMER ALLIED SIGNAL)	2869	325998		34.6	-75.0715	40.00649	42	422	1.6	17	160.9
863	NEI7330	SUNOCO CHEMICALS (FORMER ALLIED SIGNAL)	2869	325998		77.2	-75.0715	40.00649	42	422	1.6	22	160.9
864	NEIPA101353	TRIGEN - SCHUYLKILL	4961	22		128.6	-75.1873	39.94239	69	450	4.9	6	190.1
865	NEIPA101353	TRIGEN - SCHUYLKILL	4961	22		61.5	-75.1873	39.94239	78	450	7.3	2	190.1
866	NEIPA101356	GRAYS FERRY COGENERATION PARTNERS	4911	22	54785	143.2	-75.1873	39.94239	78	396	5.5	20	233.5
867	NEIPA101356	GRAYS FERRY COGENERATION PARTNERS	4911	22	54785	90.3	-75.1873	39.94239	85	443	3.2	21	233.5
868	NEIPA2222	TRIGEN - EDISON	4961	62		130.5	-75.1569	39.94604	78	589	3.7	9	130.5

Table 38. Matched stacks between the CAMD and NEI database.

NEI Facility Name	NEI *Comb. Stack* Number	NEI *Comb. Stack* Emiss (tpy)	NEI Unit Emiss (tpy)	NEI Facility Emiss (tpy, w/out Fugitive)	CAMD Facility Name	CAMD Units *	CAMD Unit Emiss (tpy) *	CAMD Comb. Unit Totals (tpy)	CAMD Facility Totals (tpy)	Stack δ (%, relative to CAMD value)	Stack δ (tpy)	Facility δ (% relative to CAMD value)	Facility δ (tpy)
Exelon Generation Co - Delaware Station	817	4.8	4.8	**292.6**	Delaware	3160-9	1.542	1.542	**289.3**	213%	3.3	**1%**	**3.3**
	818	287.8	287.8			3160-71	123.8	287.8		0%	0.0		
						3160-81	164						
Sunoco Inc. - Philadelphia	855	26.2	48.2	**1081.0**	Philadelphia Refinery	52106-150101	162.1	162.1	**1101.0**	-70%	-113.9	**-2%**	**-20.3**
	856	1.3											
	857	1.4											
	858	19.3											
	859	1032.8	1032.8			52106-150137	194.2	938.9		10%	93.9		
						52106-150110	162.1						
						52106-150138	194.2						
						52106-150139	194.2						
						52106-150140	194.2						
Sunoco Chemicals (Former Allied Signal)	860	0.0	160.9	**160.9**	Sunoco Chemicals Frankford Plant	880007-52	84.5	84.5	**84.5**	90%	76.4	**90%**	**76.4**
	861	49.1											
	862	34.6											
	863	77.2											
Trigen - Schuylkill	864	128.6	128.6	**190.1**	Trigen Energy -	50607-23	163.1	163.1	**178.7**	-21%	-34.5	**6%**	**11.4**

NEI Facility Name	**NEI *Comb. Stack* Number**	**NEI *Comb. Stack* Emiss (tpy)**	**NEI Unit Emiss (tpy)**	**NEI Facility Emiss (tpy, w/out Fugitive)**	**CAMD Facility Name**	**CAMD Units ***	**CAMD Unit Emiss (tpy) ***	**CAMD Comb. Unit Totals (tpy)**	**CAMD Facility Totals (tpy)**	**Stack δ (%, relative to CAMD value)**	**Stack δ (tpy)**	**Facility δ (% relative to CAMD value)**	**Facility δ (tpy)**
	865	61.5	61.5		Schuykill	50607-24	2.9	15.6		293%	45.9		
						50607-26	12.7						
Grays Ferry Cogeneration Partners	866	143.2	143.2	**233.5**	Grays Ferry Cogen Partnership	54785-2	143.2	143.2	**233.5**	0%	0.0	**0%**	**0.0**
	867	90.3	90.3			54785-25	90.3	90.3		0%	0.0		
Trigen - Edison	868	130.5	130.5	**130.5**	Trigen Energy Corporation-Edison St	880006-1	19.8	111	**111.0**	18%	19.4	**18%**	**19.4**
						880006-2	17.3						
						880006-3	36.1						
						880006-4	37.8						
Jefferson Smurfit Corporation (U S) ***	819	0.1	228.4	**228.4**		54785-2	143.2	233.5	**233.5**	-2%	-5.1	**-2%**	**-5.1**
	820	113.8				54785-25	90.3						
	821	114.5											

Notes:
* In the format "ORIS ID - UNIT ID"
** All CAMD values are for 2002
*** Jefferson Smurfit not in CAMD; will use Grays Ferry as surrogate

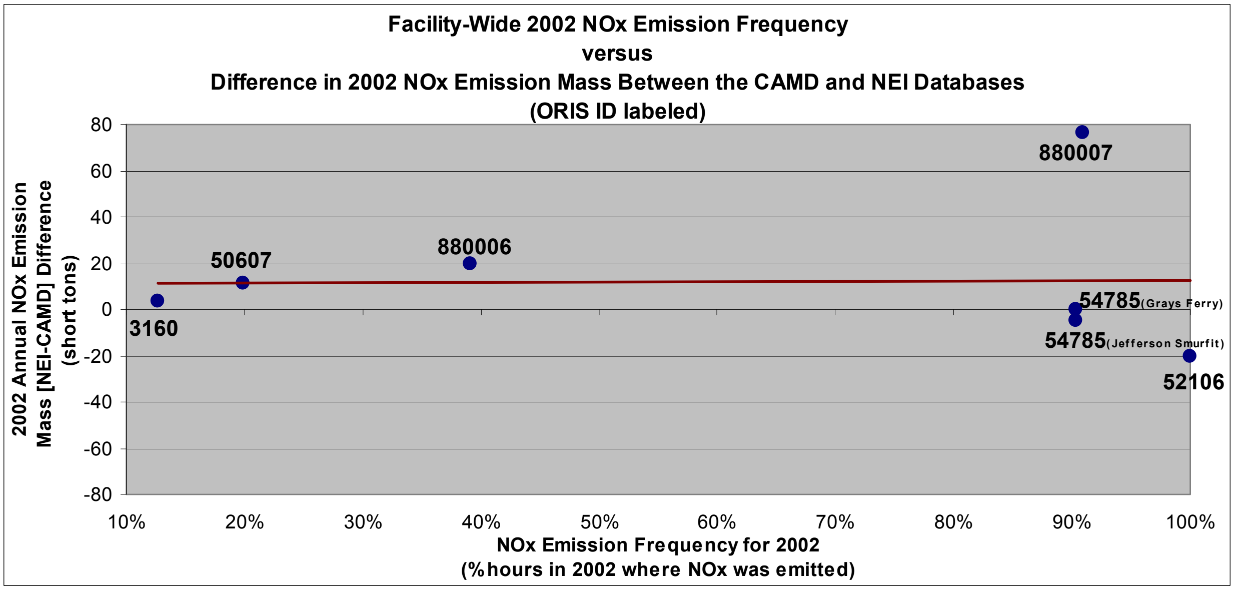


Figure 14. Differences in facility-wide annual NOx emission totals between NEI and CAMD data bases for Philadelphia County 2002.

3.2.6 Fugitive and Airport Emissions Preparation

3.2.6.1 Philadelphia County

Fugitive emission releases in Philadelphia County, as totaled in the NEI database, were modeled as area sources with the profile of these releases determined by the overall facility profile of emissions. In addition, emissions associated with the Philadelphia International Airport were estimated.

Fugitive Releases

Thirty five *combined stacks* were identified during the point source analysis (see previous section) that were associated with facilities considered major emitters, but where the emissions from the stacks are labeled *Fugitive* in the NEI. These stacks have zero stack diameter, zero emission velocity, and exit temperature equal to average ambient conditions (295 K). Thus, we determined it was not appropriate to include these in the point source group simulation.

These 35 stacks occur at only two facilities in the County: Exelon Generation Co – Delaware Station (NEI Site ID: NEIPA2218) and Sunoco Inc. – Philadelphia (NEI Site ID: NEI40723). Consequently, they were grouped by facility. The Sunoco emissions further fall into two distinct categories based on release heights. Thus, to accommodate all these sources most efficiently, we created three area source groups: one for Sunoco emissions at 3.0 m, one for Sunoco emissions greater than 23.0 m, and one for Exelon. The “stacks” within the NEI and their parameters comprising each of these sources are shown in Table 39 along with their groupings and the resulting combined area source parameters.

Table 39. Emission parameters for the three Philadelphia County fugitive NO_x area emission sources.

Grp. No.	NEI Site ID	Facility Name	NEI 2002 Emissions (tpy)	Stack X	Stack Y	Stack Height (m)	Stacks Used for Emission Profile [1]	Scaled Emissions (tpy) [2]		
								2001	2002	2003
1	NEIPA 2218	EXELON GENERATION CO - DELAWARE STATION	0.1	-75.13582	39.96769	5				
			5.1	-75.12528	39.96680	8				
			5.2			**6.5**	**817+818**	**4.8**	**5.2**	**6.4**
2	NEI40 723	Sunoco Inc. - Philadelphia	65.3	-75.21408	39.90811	3				
			350.9	-75.21300	39.90878	3				
			12.7	-75.20972	39.90467	3				
			355.7	-75.20945	39.90778	3				
			31.1	-75.20876	39.90185	3				
			6.2	-75.20845	39.90708	3				
			182.4	-75.20809	39.91580	3				
			1.1	-75.20707	39.90946	3				
			7.5	-75.20651	39.90988	3				
			1.0	-75.20301	39.91362	3				
			2.0	-75.20114	39.91273	3				
			49.4	-75.20090	39.91621	3				
			106.3	-75.20079	39.91615	3				
			188.5	-75.20047	39.91366	3				
			87.8	-75.20043	39.91377	3				
			36.1	-75.20024	39.91406	3				
			9.7	-75.20020	39.91410	3				
			61.2	-75.19995	39.91596	3				
			13.6	-75.19766	39.91696	3				
			17.0	-75.19751	39.91696	3				
			17.2	-75.19735	39.91590	3				
			12.2	-75.19723	39.91597	3				
			12.6	-75.19720	39.91698	3				
			23.7	-75.19713	39.91596	3				
			19.2	-75.19699	39.91599	3				
			10.0	-75.19644	39.91493	3				
			1,680.4			**3.0**	**855+856+857+858+859**	**1,873.8**	**1,681.4**	**2,202.4**
3	NEI40 723	Sunoco Inc. - Philadelphia	79.5	-75.21322	39.90899	23				
			13.1	-75.20833	39.90278	26				
			15.3	-75.20850	39.90246	27				
			2.5	-75.20844	39.90239	27				
			10.2	-75.20838	39.90231	27				
			19.0	-75.20828	39.90237	27				
			211.2	-75.20889	39.90279	30				
			350.8			**26.7**	**855+856+857+858+859**	**391.2**	**351.0**	**459.8**

[1] See Table 37 for stack definitions.

[2] Scaled emissions are determined by summing the scaled, hourly values from the CAMD database, as used in the dispersion modeling.

In the case of the Sunoco emissions, the vertices of the area sources were determined by a convex hull encapsulating all the points. In the case of Excelon, only two points are provided, which is insufficient information to form a closed polygon. Instead, the boundary of the facility was digitized into a 20-sided polygon. Figure 15 shows the locations of these polygons.

Emission profiles for the fugitive releases were determined from the CAMD hourly emission database in a method similar to that for the point sources. We determined scaling factors based on the ratio of the 2002 fugitive releases described by the NEI to the total, non-fugitive point source releases from the same facility. All stacks within that facility were combined on an hourly basis for each year and the fugitive to non-fugitive scaling factor applied, ensuring that the same temporal emission profile was used for fugitives as for other releases from the facility, since the origins of the emissions should be parallel. We created external hourly emissions files for each of the three fugitive area sources with appropriate units (grams per second per square meter).

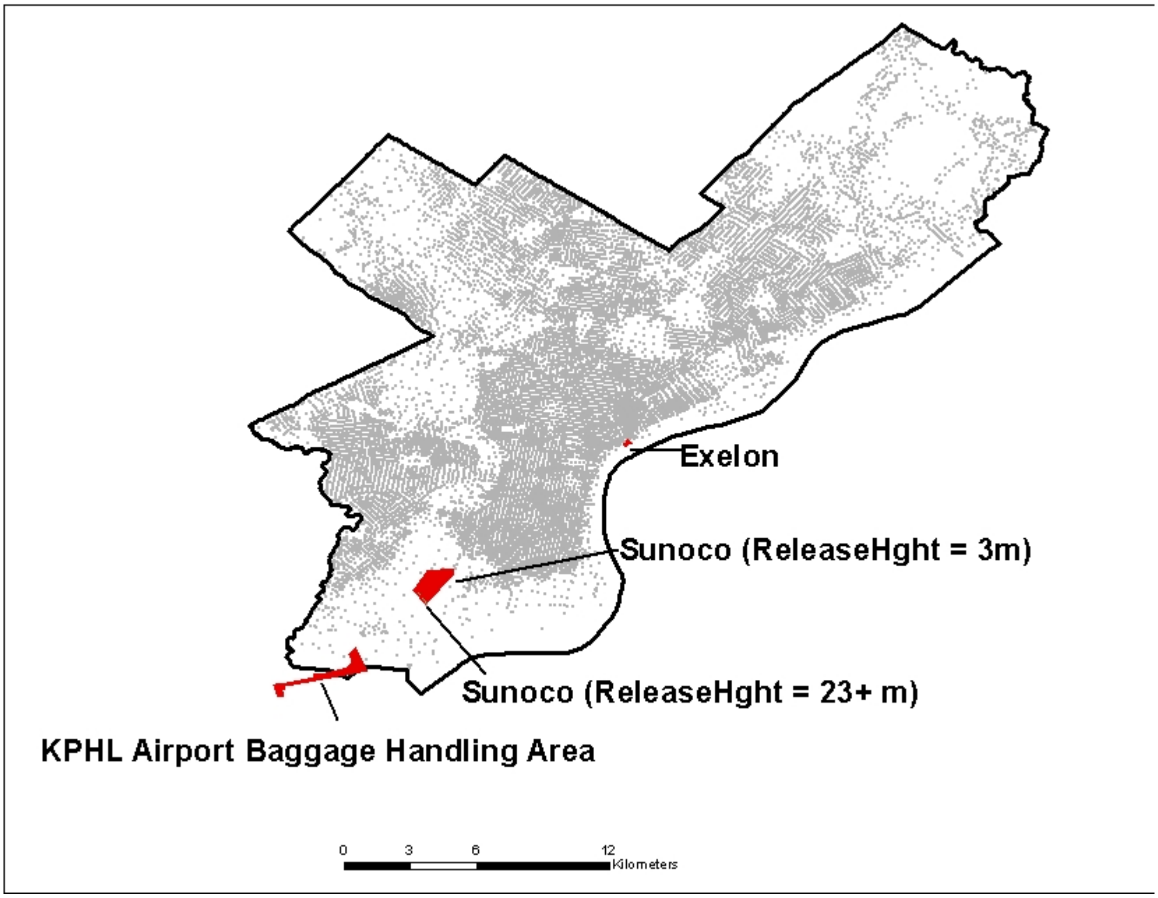


Figure 15. Locations of the four ancillary area sources. Also shown are centroid receptor locations.

Philadelphia International Airport

Another significant source of NOx emissions in Philadelphia County not captured in the earlier simulations is from operation of the Philadelphia International Airport (PHL). PHL is the only major commercial airport in the County and is the largest airport in the Delaware Valley.

The majority of NOx emissions in the NEI[25] database attributable to airports in Philadelphia County are from non-road mobile sources, specifically ground support equipment. There is another airport in the County: Northeast Philadelphia Airport. However, because it serves general aviation, is generally much smaller in operations than PHL, and has little ground support equipment activity – which is associated primarily with commercial aviation – all airport emissions in the County were attributed to PHL. The PHL emissions were taken from the non-road section of the 2002 NEI, and are shown by Table 40.

Table 40. Philadelphia International airport (PHL) NO_x emissions

State and County	SCC	NOx (tpy)	SCC Level 1 Description	SCC Level 3 Description	SCC Level 6 Description	SCC Level 8 Description
Philadelphia, PA	2265008005	4.6	Mobile Sources	Off-highway Vehicle Gasoline, 4-Stroke	Airport Ground Support Equipment	Airport Ground Support Equipment
	2267008005	5.1	Mobile Sources	LPG	Airport Ground Support Equipment	Airport Ground Support Equipment
	2270008005	196.2	Mobile Sources	Off-highway Vehicle Diesel	Airport Ground Support Equipment	Airport Ground Support Equipment
	2275020000	0.01	Mobile Sources	Aircraft	Commercial Aircraft	Total: All Types
	2275050000	2.5	Mobile Sources	Aircraft	General Aviation	Total
PHL Total		**208.4**				

As with the fugitive sources discussed above, the airport emissions are best parameterized as area sources. The boundary of the area source was taken as the region of operation of baggage handling equipment, including the terminal building and the region surrounding the gates. This region was digitized into an 18-sided polygon of size 1,326,000 m^2, and included in the AERMOD input control file.

The activity profile for PHL was taken to have seasonal and hourly variation (SEASHR), based on values from the EMS-HAP model.[26] These factors are disaggregated in the EMS-HAP model database based on source classification codes (SCCs), which were linked to those from the NEI database. The EMS-HAP values provide hourly activity factors by season, day type, and hour; to compress to simple SEASHR modeling, the hourly values from the three individual day types were averaged together. The total emissions for each SCC were then disaggregated into seasonal and hourly components and the resulting components summed to create total PHL emissions for each hour of the four annual seasons. These parameterized emissions were then normalized to the total cargo handling operational area, to produce emission factors in units of grams per second per square meter and included in the AERMOD input file. Figure 15 also shows the location of the PHL area source.

---

[25] http://www.epa.gov/ttn/chief/net/2002inventory.html

[26] EPA 2004, User's Guide for the Emissions Modeling System for Hazardous Air Pollutants (EMS-HAP) Version 3.0, EPA-454/B-03-006.

3.2.7 Receptor Locations

3.2.7.1 Philadelphia County

Three sets of receptors were chosen to represent the locations of interest. First, all NOx monitor locations, shown by Table 41, within the Philadelphia CMSA were included as receptor locations. Although all receptors are assumed to be on a flat plane, they are placed at the standard breathing height of 5.9 ft (1.8 m).

Table 41. Philadelphia CMSA NO_x monitors.

CMSA	Site ID	Latitude	Longitude
Philadelphia-Wilmington-Atlantic City, PA-NJ-DE-MD	100031003	39.7611	-75.4919
	100031007	39.5511	-75.7308
	100032004	39.7394	-75.5581
	340070003	39.923	-75.0976
	420170012	40.1072	-74.8822
	420450002	39.8356	-75.3725
	420910013	40.1122	-75.3092
	421010004	40.0089	-75.0978
	421010029	39.9572	-75.1731
	421010047	39.9447	-75.1661

The second receptor locations were selected to represent the locations of census block centroids near major NO_x sources. GIS analysis was used to determine all block centroids in Philadelphia County that lie within a 0.25 mile (400 m) of the roadway segments and also all block centroids that lie within 6.2 miles (10 km) of any major point source. 12,982 block centroids were selected due to their proximity to major roadways; 16,298 centroids were selected due to their proximity to major sources. The union of these sets produced 16,857 unique block centroid receptor locations, each of which was assigned a height of 5.9 ft (1.8 m). The locations of centroids that met either distance criteria – and were thus included in the modeling – is shown by Figure 16.

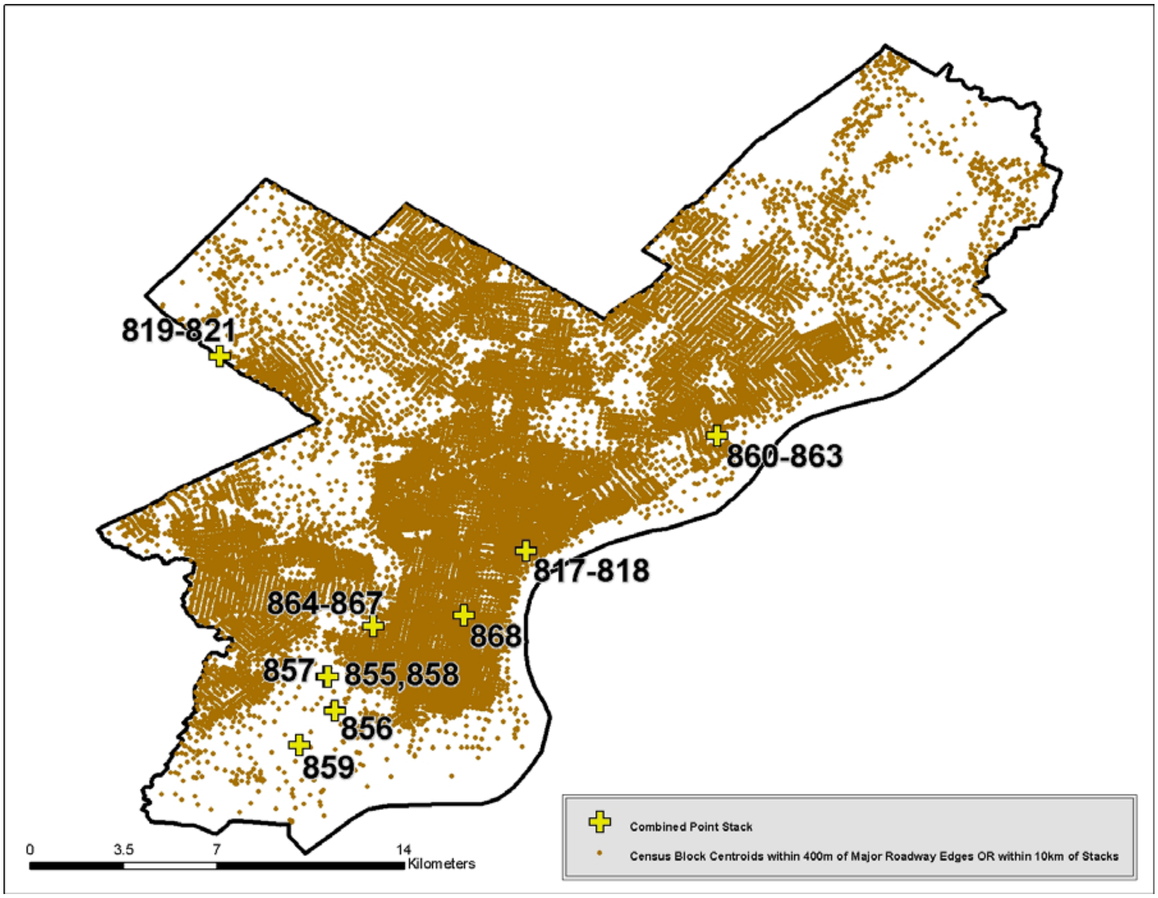


Figure 16. Centroid locations within fixed distances to major point and mobile sources.

The third set of receptors was chosen to represent the on-road microenvironment. For this set, one receptor was placed at the center of each of the 982 sources.

The distance relationship between the road segments and block centroids can be estimated by looking at the distance between the road-centered and the block centroid receptors. Figure 17 shows the histogram of the shortest distance between each centroid receptor and its nearest roadway-centered receptor.

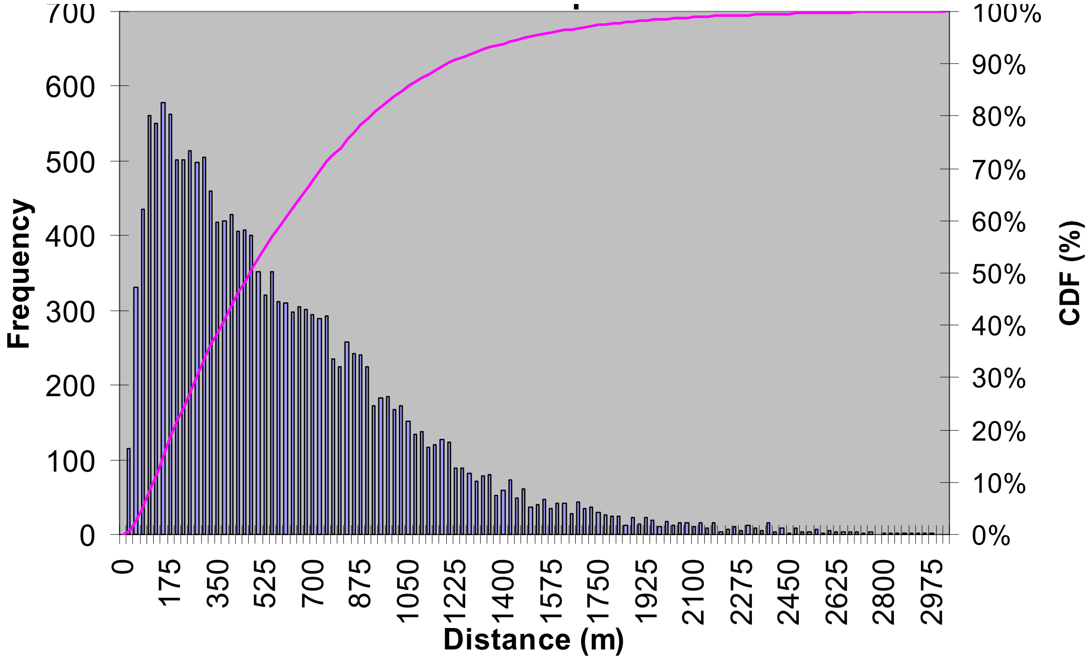


Figure 17. Frequency distribution of distance between each Census receptor and its nearest road-centered receptor.

The centroids selected were those within 10 km of any major point source or 400 m from any receptor edge, so the distances to the nearest major road segment can be significantly greater than 400 m. The mode of the distribution is about 150 m and the median distance to the closest roadway segment center is about 450 m. However, these values represent the distances of the block centroids to road centers instead of road edges, so that they overestimate the actual distances to the zone most influenced by roadway by an average of 14 m and a range of 4 m to 44 m (see Table 36 above).

3.2.8 Other Modeling Specifications

Since each of the case-study locations were MSA/CMSAs, all emission sources were characterized as urban. The AERMOD *toxics* enhancements were also employed to speed calculations from area sources. NO_x chemistry was applied to all sources to determine NO_2 concentrations. For the each of the roadway, fugitive, and airport emission sources, the ozone limiting method (OLM) was used, with plumes considered ungrouped. Because an initial NO_2 fraction of NO_x is anticipated to be about 10% or less (Finlayson-Pitts and Pitts, 2000; Yao et al., 2005), a conservative value of 10% for all sources was selected. For all point source simulations the Plume Volume Molar Ratio Method (PVMRM) was used to estimate the conversion of NO_x to NO_2, with the following settings:

1. Hourly series of O_3 concentrations were taken from EPA's AQS database[27]. The complete national hourly record of monitored O_3 concentrations were filtered for the four monitors within Philadelphia County (stations 421010004, 421010014, 421010024, and 421010136). The hourly records of these stations were then averaged together to provide an average Philadelphia County concentrations of O_3 for each hour of 2001-2003.

---

[27] http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm

2. The equilibrium value for the NO_2:NO_x ratio was taken as 75%, the national average ambient ratio.[28]
3. The initial NO_2 fraction of NO_x is anticipated to be about 10% or less. A default value of 10% was used for all stacks (Finlayson-Pitts and Pitts, 2000).

3.2.9 Air Quality Concentration Estimation

The hourly concentrations estimated from each of the three source categories were combined at each receptor. Then a local concentration, reflecting the concentration contribution from emission sources not included in the simulation, was added to the sum of the concentration contributions from each of these sources at each receptor. The local concentration was estimated from the difference between the model predictions at the local NO_2 monitors and the observed values. It should be noted that this local concentration may also include any model error present in estimating concentration at the local monitoring sites. Table 42 presents a summary of the estimated local concentration added to the AERMOD hourly concentration data.

Table 42. Comparison of ambient monitoring and AERMOD predicted NO_2 concentrations.

Year and Monitor ID	**Annual Average NO_2 concentration (ppb)**			
	Monitor	**AERMOD Initial**	**Difference[1]**	**AERMOD Final[2]**
2001				
4210100043	26	7	18	19
4210100292	28	22	6	33
4210100471	30	20	10	32
mean			11	
2002				
4210100043	24	7	17	18
4210100292	28	21	7	32
4210100471	29	19	10	31
mean			11	
2003				
4210100043	24	7	17	13
4210100292	25	22	3	28
4210100471*	25	26	-1	32
mean			6	

[1] the difference represents concentrations attributed to sources not modeled by AERMOD and model error.
[2] the mean difference between measured and modeled was added uniformly at each receptor hourly concentration to generate the AERMOD final concentrations.
* monitor did not meet completeness criteria used in the air quality characterization.

---

[28] Appendix W to CFR 51, page 466. http://www.epa.gov/scram001/guidance/guide/appw_03.pdf.

3.3 Human Exposure Modeling using APEX

The Air Pollutants Exposure model (APEX) is a personal computer (PC)-based program designed to estimate human exposure to criteria and air toxic pollutants at the local, urban, and consolidated metropolitan levels. APEX, also known as TRIM.Expo, is the human inhalation exposure module of EPA's Total Risk Integrated Methodology (TRIM) model framework (US EPA, 1999), a modeling system with multimedia capabilities for assessing human health and ecological risks from hazardous and criteria air pollutants. It is being developed to support evaluations with a scientifically sound, flexible, and user-friendly methodology. Additional information on the TRIM modeling system, as well as downloads of the APEX Model, user's guide, and other supporting documentation, can be found on EPA's Technology Transfer Network (TTN) at http://www.epa.gov/ttn/fera.

3.3.1 History

APEX was derived from the National Ambient Air Quality Standards (NAAQS) Exposure Model (NEM) series of models, developed to estimate exposure to the criteria pollutants (e.g., carbon monoxide (CO), ozone O_3). In 1979, EPA began by assembling a database of human activity patterns that could be used to estimate exposures to indoor and outdoor pollutants (Roddin et al., 1979). These data were then combined with measured outdoor concentrations in NEM to estimate exposures to CO (Biller et al., 1981; Johnson and Paul, 1983). In 1988, OAQPS began to incorporate probabilistic elements into the NEM methodology and use activity pattern data based on various human activity diary studies to create an early version of probabilistic NEM for O_3 (i.e., pNEM/O_3). In 1991, a probabilistic version of NEM was extended to CO (pNEM/CO) that included a one-compartment mass-balance model to estimate CO concentrations in indoor microenvironments. The application of this model to Denver, Colorado has been documented in Johnson et al. (1992). Additional enhancements to pNEM/O_3 in the early- to mid-1990's allowed for probabilistic exposure assessments in nine urban areas for the general population, outdoor children, and outdoor workers (Johnson et al., 1996a; 1996b; 1996c). Between 1999 and 2001, updated versions of pNEM/CO (versions 2.0 and 2.1) were developed that relied on activity diary data from EPA's Consolidated Human Activities Database (CHAD) and enhanced algorithms for simulating gas stove usage, estimating alveolar ventilation rate (a measure of human respiration), and modeling home-to-work commuting patterns.

The first version of APEX was essentially identical to pNEM/CO (version 2.0) except that it was capable of running on a PC instead of a mainframe. The next version, APEX2, was substantially different, particularly in the use of a personal profile approach (i.e., simulation of individuals) rather than a cohort simulation (i.e., groups of similar persons). APEX3 introduced a number of new features including automatic site selection from national databases, a series of new output tables providing summary exposure and dose statistics, and a thoroughly reorganized method of describing microenvironments and their parameters. Most of the spatial and temporal constraints of pNEM and APEX1 were removed or relaxed by version 3.

The version of APEX used in this exposure assessment is APEX4, described in the APEX User's Guide and the APEX Technical Support Document (US EPA, 2006a; 2006b), referred to as the APEX User's Guide and TSD.

3.3.2 Model Overview

APEX estimates human exposure to criteria and toxic air pollutants at the local, urban, or consolidated metropolitan area levels using a stochastic, microenvironmental approach. The model randomly selects data for a sample of hypothetical individuals from an actual population database and simulates each hypothetical individual's movements through time and space (e.g., at home, in vehicles) to estimate their exposure to a pollutant. APEX simulates commuting, and thus exposures that occur at home and work locations, for individuals who work in different areas than they live.

A **microenvironment** is a three-dimensional space in which human contact with an environmental pollutant takes place and which can be treated as a well-characterized, relatively homogeneous location with respect to pollutant concentrations for a specified time period.

APEX can be conceptualized as a simulated field study that would involve selecting an actual sample of specific individuals who live in (or work and live in) a geographic area and then continuously monitoring their activities and subsequent inhalation exposure to a specific air pollutant during a specific period of time.

The main differences between APEX and an actual field study are that in APEX:

- The sample of individuals is a virtual sample, not actual persons. However, the population of individuals appropriately balanced according to various demographic variables and census data using their relative frequencies, in order to obtain a representative sample (to the extent possible) of the actual people in the study area
- The activity patterns of the sampled individuals (e.g., the specification of indoor and other microenvironments visited and the time spent in each) are assumed by the model to be comparable to individuals with similar demographic characteristics, according to activity data such as diaries compiled in EPA's Consolidated Human Activity Database (or CHAD; US EPA, 2002; McCurdy et al., 2000)
- The pollutant exposure concentrations are estimated by the model using a set of user-input ambient outdoor concentrations (either modeled or measured) and information on the behavior of the pollutant in various microenvironments;
- Variation in ambient air quality levels can be simulated by either adjusting air quality concentrations to just meet alternative ambient standards, or by reducing source emissions and obtaining resulting air quality modeling outputs that reflect these potential emission reductions, and
- The model accounts for the most significant factors contributing to inhalation exposure – the temporal and spatial distribution of people and pollutant concentrations throughout the study area and among microenvironments – while also allowing the flexibility to adjust some of these factors for alternative scenarios and sensitivity analyses.

APEX is designed to simulate human population exposure to criteria and air toxic pollutants at local, urban, and regional scales. The user specifies the geographic area to be modeled and the number of individuals to be simulated to represent this population. APEX then generates a personal profile for each simulated person that specifies various parameter values required by the model. The model next uses diary-derived time/activity data matched to each personal profile to generate an exposure event sequence (also referred to as *activity pattern* or *diary*) for the modeled individual that spans a specified time period, such as one year. Each event in the sequence specifies a start time, exposure duration, geographic location, microenvironment, and activity performed. Probabilistic algorithms are used to estimate the pollutant concentration associated with each exposure event. The estimated pollutant concentrations account for the effects of ambient (outdoor) pollutant concentration, penetration factors, air exchange rates, decay/deposition rates, and proximity to emission sources, depending on the microenvironment, available data, and estimation method selected by the user. Because the modeled individuals represent a random sample of the population of interest, the distribution of modeled individual exposures can be extrapolated to the larger population. The model simulation can be broadly described in five steps that follow:

1. **Characterize the study area**. APEX selects census tracts within a study area – and thus identifies the potentially exposed population – based on user-defined criteria and availability of air quality and meteorological data for the area.
2. **Generate simulated individuals**. APEX stochastically generates a sample of hypothetical individuals based on the census data for the study area and human profile distribution data (such as age-specific employment probabilities).
3. **Construct a sequence of activity events**. APEX constructs an exposure event sequence spanning the period of the simulation for each of the simulated individuals and based on the activity pattern data.
4. **Calculate hourly concentrations in microenvironments**. APEX users define microenvironments that people in the study area would visit by assigning location codes in the activity pattern to the user-specified microenvironments. The model then calculates hourly concentrations of a pollutant in each of these microenvironments for the period of simulation, based on the user-provided microenvironment descriptions and hourly air quality data. Microenvironmental concentrations are calculated for each of the simulated individuals.
5. **Determine exposures**. APEX estimates a concentration for each exposure event based on the microenvironment occupied during the event. These values can be averaged by clock hour to produce a sequence of hourly average exposures spanning the specified exposure period. These hourly values may be further aggregated to produce daily, monthly, and annual average exposure values.

3.3.3 Study Area Descriptions

The APEX study area has traditionally been on the scale of a city or slightly larger metropolitan area, although it is now possible to model larger areas such as combined statistical areas (CSAs). In this analysis the study area is defined by a single or few counties. The demographic data used by the model to create personal profiles is provided at the census block level. For each block the model requires demographic information representing the distribution of age, gender, race, and work status within the study population. Each block has a location specified by latitude and longitude for some representative point (e.g., geographic center). The current release of APEX includes input files that already contain this demographic and location data for all census tracts, block groups, and blocks in the 50 United States, based on the 2000 Census.

Philadelphia County is comprised of 17,315 blocks containing a population of 1,517,550 persons. For this analysis the population studied was limited those residents of Philadelphia

County residing in census blocks that were either within 400 meters of a major roadway or within 10 km of a major emission source (see section 3.2.7 for definition). This was done to maintain balance between the representation of the study area/objectives and the computational load regarding file size and processing time. There were 16,857 such blocks containing a population of 1,475,651.

3.3.3.1 Air Quality Data

Air quality data input to the model were generated by air quality modeling. Principal emission sources included both mobile and stationary sources as well as fugitive emissions. The methodology was described previously in Section 3.2. Briefly, hourly NO_2 concentrations were estimated for each of 3 years (2001-2003) at each of the defined receptor locations using hourly NO_x emission estimates and dispersion modeling.

In APEX, the ambient air quality data are assigned to geographic areas called districts. The districts are used to assign pollutant concentrations to the blocks/tracts and microenvironments being modeled. The ambient air quality data are provided by the user as hourly time series for each district. As with blocks/tracts, each district has a representative location (latitude and longitude). APEX calculates the distance from each block/tract to each district center, and assigns the block/tract to the nearest district, provided the block/tract representative location point (e.g., geographic center) is in the district. Each block/tract can be assigned to only one district. In this assessment the district was synonymous with the receptor modeled in the dispersion modeling (see Sections 3.2).

3.3.3.2 Meteorological Data

Ambient temperatures are input to APEX for different sites (locations). As with districts, APEX calculates the distance from each block to each temperature site and assigns each block to the nearest site. Hourly temperature data are from the National Climatic Data Center Surface Airways Hourly TD-3280 dataset (NCDC Surface Weather Observations). Daily average and 1-hour maxima are computed from these hourly data.

There are two files that are used to provide meteorological data to APEX. One file, the meteorological station location file, contains the locations of meteorological data recordings expressed in latitude and longitude coordinates. This file also contains start and end dates for the data recording periods. The temperature data file contains the data from the locations in the temperature zone location file. This file contains hourly temperature readings for the period being modeled for the meteorological stations in and around the study area. Table 43 lists the meteorological stations used for each modeled area.

Table 43. The meteorological stations used for each study area.

Location	World Meteorological Organization ID (call sign)
Atlanta, GA	72219 (KATL)
Detroit, MI	72537 (KDTW)
Los Angeles, CA	72295 (KLAX)
Philadelphia County, PA	72408 (KPHL)
Phoenix	72278 (KPHX)

3.3.4 Simulated Individuals

APEX stochastically generates a user-specified number of simulated persons to represent the population in the study area. Each simulated person is represented by a personal profile, a summary of personal attributes that define the individual. APEX generates the simulated person or profile by probabilistically selecting values for a set of profile variables (Table 44). The profile variables could include:

- Demographic variables, generated based on the census data;
- Physical variables, generated based on sets of distribution data;
- Other daily varying variables, generated based on literature-derived distribution data that change daily during the simulation period.

APEX first selects demographic and physical attributes for each specified individuals, and then follows the individual over time and calculates his or her time series of exposure.

Table 44. Examples of profile variables in APEX.

Variable Type	Profile Variables	Description
Demographic	Age	Age (years)
	Gender	Male or Female
	Home block	Block in which a simulated person lives
	Work tract	Tract in which a simulated person works
	Employment status	Indicates employment outside home
Physical	Air conditioner	Indicates presence of air conditioning at home
	Gas Stove	Indicates presence of gas stove at home

Due to the large size of the air quality input files, the modeled area was separated into three sections. The number of simulated persons in each model run (3 sections per 3 years) was set to 50,000, yielding a total of 150,000 persons simulated for each year. The parameters controlling the location and size of the simulated area were set to include the county(s) in the selected study area. The settings that allow for replacement of CHAD data that are missing gender, employment or age values were all set to preclude replacing missing data. The width of the age window was set to 20 percent to increase the pool of diaries available for selection. The variable that controls the use of additional ages outside the target age window was set to 0.1 to further enhance variability in diary selection. See the APEX User's Guide for further explanation of these parameters. The total population simulated for Philadelphia County was approximately 1.48 million persons, of which there a total simulated population of 163,000 asthmatics. The model simulated approximately 281,000 children, of which there were about 48,000 asthmatics. Due to random sampling, the actual number of specific subpopulations modeled varied slightly by year.

3.3.4.1 Population Demographics

APEX takes population characteristics into account to develop accurate representations of study area demographics. Specifically, population counts by area and employment probability estimates are used to develop representative profiles of hypothetical individuals for the simulation.

APEX is flexible in the resolution of population data provided. As long as the data are available, any resolution can be used (e.g., county, census tract, census block). For this application of the model, census block level data were used. Block-level population counts come from the 2000 Census of Population and Housing Summary File 1 (SF-1). This file contains the 100-percent data, which is the information compiled from the questions asked of all people and about every housing unit.

As part of the population demographics inputs, it is important to integrate working patterns into the assessment. In the 2000 U.S. Census, estimates of employment were developed by census information (US Census Bureau, 2007). The employment statistics are broken down by gender and age group, so that each gender/age group combination is given an employment probability fraction (ranging from 0 to 1) within each census tract. The age groupings used are: 16-19, 20-21, 22-24, 25-29, 30-34, 35-44, 45-54, 55-59, 60-61, 62-64, 65-69, 70-74, and >75. Children under 16 years of age were assumed to be not employed.

Since this analysis was conducted at the census block level, block level employment probabilities were required. It was assumed that the employment probabilities for a census tract apply uniformly to the constituent census blocks.

3.3.4.2 Commuting

In addition to using estimates of employment by tract, APEX also incorporates home-to-work commuting data. Commuting data were originally derived from the 2000 Census and were collected as part of the Census Transportation Planning Package (CTPP) (US DOT, 2007). The data used contain counts of individuals commuting from home to work locations at a number of geographic scales. These data were processed to calculate fractions for each tract-to-tract flow to create the national commuting data distributed with APEX. This database contains commuting data for each of the 50 states and Washington, D.C.

Commuting within the Home Tract

The APEX data set does not differentiate people that work at home from those that commute within their home tract.

Commuting Distance Cutoff

A preliminary data analysis of the home-work counts showed that a graph of log(flows) versus log(distance) had a near-constant slope out to a distance of around 120 kilometers. Beyond that distance, the relationship also had a fairly constant slope but it was flatter, meaning that flows were not as sensitive to distance. A simple interpretation of this result is that up to 120 km, the majority of the flow was due to persons traveling back and forth daily, and the numbers of such persons decrease fairly rapidly with increasing distance. Beyond 120 km, the majority of the flow is made up of persons who stay at the workplace for extended times, in which case the separation distance is not as crucial in determining the flow.

To apply the home-work data to commuting patterns in APEX, a simple rule was chosen. It was assumed that all persons in home-work flows up to 120 km are daily commuters, and no persons in more widely separated flows commute daily. This meant that the list of destinations for each home tract was restricted to only those work tracts that are within 120 km of the home tract. When the same cutoff was performed on the 1990 census data, it resulted in 4.75% of the home-work pairs in the nationwide database being eliminated, representing 1.3% of the workers. The assumption is that this 1.3% of workers do not commute from home to work on a daily basis. It is expected that the cutoff reduced the 2000 data by similar amounts.

Eliminated Records

A number of tract-to-tract pairs were eliminated from the database for various reasons. A fair number of tract-to-tract pairs represented workers who either worked outside of the U.S. (9,631 tract pairs with 107,595 workers) or worked in an unknown location (120,830 tract pairs with 8,940,163 workers). An additional 515 workers in the commuting database whose data were missing from the original files, possibly due to privacy concerns or errors, were also deleted.

Commuting outside the study area

APEX allows for some flexibility in the treatment of persons in the modeled population who commute to destinations outside the study area. By specifying “KeepLeavers = No” in the simulation control parameters file, people who work inside the study area but live outside of it are not modeled, nor are people who live in the study area but work outside of it. By specifying “KeepLeavers = Yes,” these commuters are modeled. This triggers the use of two additional parameters, called LeaverMult and LeaverAdd. While a commuter is at work, if the workplace is outside the study area, then the ambient concentration is assumed to be related to the average concentration over all air districts at the same point in time, and is calculated as:

$$Ambient\ Concentrat\,ion = LeaverMult\ \times avg\,(t) + LeaverAdd \qquad \text{eq (3)}$$

where:

Ambient Concentration = Calculated ambient air concentrations for locations outside of the study area (ppm or ppm)
LeaverMult = Multiplicative factor for city-wide average concentration, applied when working outside study area
avg(t) = Average ambient air concentration over all air districts in study area, for time *t* (ppm or ppm)
LeaverAdd = Additive term applied when working outside study area

All microenvironmental concentrations for locations outside of the study area are determined from this ambient concentration by the same function as applies inside the study area.

Block-level commuting

For census block simulations, APEX requires block-level commuting file. A special software preprocesser was created to generate this files for APEX on the basis of the tract-level commuting data and finely-resolved land use data. The software calculates commuting flows between census blocks for the employed population according to the following equation.

$$Flow_{block} = Flow_{\,tract} \times F_{pop} \times F_{land}$$

where:

$Flow_{block}$ = flow of working population between a home block and a work block.
$Flow_{tract}$ = flow of working population between a home tract and a work tract.
F_{pop} = fraction of home tract's working population residing in the home block.
F_{land} = fraction of work tract's commercial/industrial land area in the work block

Thus, it is assumed that the frequency of commuting to a workplace block within a tract is proportional to the amount of commercial and industrial land in the block.

3.3.4.3 Profile Functions

A *Profile Functions* file contains settings used to generate results for variables related to simulated individuals. While certain settings for individuals are generated automatically by APEX based on other input files, including demographic characteristics, others can be specified using this file. For example, the file may contain settings for determining whether the profiled individual's residence has an air conditioner, a gas stove, etc. As an example, the *Profile Functions* file contains fractions indicating the prevalence of air conditioning in the cities modeled in this assessment (Figure 18). APEX uses these fractions to stochastically generate air conditioning status for each individual. The derivation of particular data used in specific microenvironments is provided below.

```
AC_Home
! Has air conditioning at home
TABLE
INPUT1 PROBABILITY 2     “A/C probabilities”
0.85 0.15
RESULT INTEGER 2         “Yes/No”
1 2
#
```

Figure 18. Example of a profile function file for A/C prevalence.

3.3.4.4 Asthma Prevalence Rates

One of the important population subgroups for the exposure assessment is asthmatic children. Evaluation of the exposure of this group with APEX requires the estimation of children's asthma prevalence rates. The proportion of the population of children characterized as being asthmatic was estimated by statistics on asthma prevalence rates recently used in the NAAQS review for O_3 (US EPA, 2007d; 2007e). Specifically, the analysis generated age and gender specific asthma prevalence rates for children ages 0-17 using data provided in the National Health Interview Survey (NHIS) for 2003 (CDC, 2007). These asthma rates were characterized by geographic regions, namely Midwest, Northeast, South, and West. Adult asthma prevalence rates for Philadelphia County were obtained from the Behavioral Risk Factor Surveillance System (BRFSS) survey information (PA DOH, 2008). The average rates for adult males and females in Philadelphia for 2001-2003 were 7% and 12%, respectively. These rates were assumed to apply to all adults uniformly. Table 45 provides a summary of the prevalence rates used in the exposure analysis by age and gender for each of the four regions as applied to the five study areas in the exposure assessment.

Table 45. Asthma prevalence rates by age and gender for 4 regions.

Region (Study Area)	Age	Females				Males			
		Prevalence	se	L95	U95	Prevalence	se	L95	U95
Midwest (Detroit)	0	0.070	0.036	0.021	0.203	0.031	0.015	0.010	0.090
	1	0.071	0.020	0.037	0.130	0.063	0.018	0.033	0.115
	2	0.073	0.018	0.042	0.124	0.108	0.021	0.070	0.163
	3	0.075	0.019	0.042	0.132	0.158	0.027	0.107	0.228
	4	0.081	0.022	0.044	0.144	0.216	0.037	0.145	0.308
	5	0.095	0.026	0.051	0.171	0.178	0.035	0.113	0.270
	6	0.092	0.029	0.045	0.178	0.128	0.028	0.078	0.204
	7	0.090	0.026	0.047	0.166	0.121	0.026	0.074	0.193
	8	0.086	0.022	0.048	0.149	0.128	0.027	0.079	0.200
	9	0.110	0.027	0.063	0.186	0.147	0.030	0.093	0.226
	10	0.162	0.035	0.098	0.255	0.177	0.030	0.120	0.254
	11	0.196	0.039	0.123	0.298	0.190	0.030	0.131	0.266
	12	0.212	0.040	0.137	0.313	0.195	0.031	0.135	0.272
	13	0.170	0.034	0.107	0.258	0.169	0.028	0.115	0.242
	14	0.140	0.026	0.092	0.209	0.168	0.026	0.117	0.235
	15	0.133	0.023	0.091	0.192	0.180	0.026	0.130	0.243
	16	0.140	0.022	0.098	0.198	0.201	0.030	0.142	0.277
	17	0.165	0.040	0.093	0.275	0.237	0.058	0.132	0.388
Northeast (Philadelphia)	0	0.068	0.066	0.007	0.442	0.048	0.033	0.010	0.200
	1	0.072	0.038	0.021	0.221	0.046	0.018	0.019	0.108
	2	0.075	0.022	0.038	0.145	0.052	0.015	0.027	0.097
	3	0.077	0.020	0.042	0.138	0.068	0.018	0.037	0.120
	4	0.082	0.023	0.043	0.151	0.100	0.023	0.059	0.164
	5	0.116	0.030	0.063	0.205	0.149	0.029	0.094	0.226
	6	0.161	0.037	0.092	0.266	0.207	0.042	0.129	0.316
	7	0.185	0.041	0.108	0.298	0.228	0.045	0.143	0.343
	8	0.171	0.040	0.096	0.284	0.222	0.043	0.142	0.332
	9	0.145	0.035	0.080	0.246	0.212	0.041	0.136	0.316
	10	0.135	0.031	0.078	0.223	0.177	0.037	0.108	0.275
	11	0.141	0.031	0.084	0.227	0.166	0.035	0.102	0.259
	12	0.166	0.034	0.102	0.259	0.183	0.036	0.116	0.276
	13	0.174	0.034	0.109	0.266	0.171	0.031	0.113	0.250
	14	0.151	0.029	0.095	0.232	0.170	0.029	0.115	0.244
	15	0.146	0.028	0.093	0.221	0.182	0.029	0.127	0.254
	16	0.146	0.031	0.088	0.232	0.204	0.032	0.142	0.284
	17	0.157	0.054	0.068	0.322	0.242	0.061	0.133	0.399
	18+	0.070		0.040	0.140	0.120		0.090	0.150
South (Atlanta)	0	0.034	0.013	0.015	0.077	0.041	0.019	0.015	0.110
	1	0.052	0.012	0.031	0.085	0.070	0.016	0.041	0.116
	2	0.071	0.014	0.046	0.109	0.102	0.017	0.070	0.146
	3	0.088	0.017	0.056	0.134	0.129	0.021	0.088	0.184
	4	0.099	0.019	0.064	0.150	0.144	0.024	0.099	0.205
	5	0.119	0.022	0.079	0.175	0.165	0.024	0.118	0.224
	6	0.122	0.023	0.079	0.182	0.164	0.025	0.116	0.226
	7	0.112	0.022	0.072	0.170	0.133	0.023	0.090	0.194

Region (Study Area)	Age	Females				Males			
		Prevalence	se	L95	U95	Prevalence	se	L95	U95
	8	0.093	0.019	0.059	0.144	0.138	0.023	0.095	0.197
	9	0.091	0.018	0.059	0.139	0.168	0.025	0.121	0.230
	10	0.108	0.020	0.071	0.162	0.178	0.025	0.130	0.240
	11	0.132	0.023	0.090	0.191	0.162	0.022	0.119	0.218
	12	0.123	0.020	0.085	0.175	0.145	0.020	0.106	0.195
	13	0.097	0.017	0.065	0.142	0.143	0.019	0.105	0.192
	14	0.095	0.016	0.064	0.137	0.153	0.019	0.116	0.200
	15	0.100	0.016	0.070	0.141	0.151	0.017	0.116	0.194
	16	0.115	0.016	0.084	0.156	0.140	0.018	0.105	0.185
	17	0.145	0.029	0.091	0.223	0.122	0.026	0.075	0.193
West (Los Angeles) (Phoenix)	0	0.013	0.010	0.002	0.067	0.031	0.025	0.004	0.186
	1	0.031	0.013	0.012	0.078	0.046	0.019	0.017	0.116
	2	0.054	0.015	0.029	0.098	0.063	0.014	0.036	0.106
	3	0.074	0.018	0.043	0.127	0.078	0.019	0.044	0.136
	4	0.077	0.021	0.042	0.137	0.091	0.025	0.048	0.168
	5	0.077	0.021	0.042	0.139	0.113	0.029	0.060	0.201
	6	0.073	0.020	0.039	0.131	0.121	0.029	0.068	0.207
	7	0.081	0.020	0.047	0.138	0.127	0.028	0.075	0.208
	8	0.091	0.020	0.055	0.146	0.132	0.027	0.080	0.208
	9	0.102	0.023	0.061	0.167	0.151	0.028	0.096	0.229
	10	0.122	0.027	0.074	0.194	0.164	0.026	0.112	0.233
	11	0.127	0.027	0.079	0.200	0.170	0.026	0.117	0.240
	12	0.131	0.025	0.084	0.197	0.175	0.027	0.120	0.248
	13	0.120	0.024	0.076	0.183	0.162	0.028	0.107	0.237
	14	0.111	0.021	0.071	0.167	0.165	0.026	0.112	0.236
	15	0.112	0.021	0.074	0.166	0.170	0.025	0.120	0.236
	16	0.122	0.022	0.081	0.180	0.179	0.025	0.127	0.246
	17	0.144	0.040	0.076	0.256	0.192	0.043	0.111	0.312

Notes:
se – Standard error
L95 – Lower limit on 95th confidence interval
U95 – Upper limit on 95th confidence interval

3.3.5 Activity Pattern Sequences

Exposure models use human activity pattern data to predict and estimate exposure to pollutants. Different human activities, such as spending time outdoors, indoors, or driving, will have varying pollutant exposure concentrations. To accurately model individuals and their exposure to pollutants, it is critical to understand their daily activities.

The Consolidated Human Activity Database (CHAD) provides data for where people spend time and the activities performed. CHAD was designed to provide a basis for conducting multi-route, multi-media exposure assessments (McCurdy et al., 2000). The data contained within CHAD come from multiple activity pattern surveys with varied structures (Table 46), however the surveys have commonality in containing daily diaries of human activities and personal attributes (e.g., age and gender).

There are four CHAD-related input files used in APEX. Two of these files can be downloaded directly from the CHADNet (http://www.epa.gov/chadnet1), and adjusted to fit into the APEX framework. These are the human activity diaries file and the personal data file, and are discussed below. A third input file contains metabolic information for different activities listed in the diary file, these are not used in this exposure analysis. The fourth input file maps five-digit location codes used in the diary file to APEX microenvironments; this file is discussed in the section describing microenvironmental calculations (Section 3.3.6).

3.3.5.1 Personal Information file

Personal attribute data are contained in the CHAD questionnaire file that is distributed with APEX This file also has information for each day individuals have diaries. The different variables in this file are:

- The study, person, and diary day identifiers
- Day of week
- Gender
- Employment status
- Age in years
- Maximum temperature in degrees Celsius for this diary day
- Mean temperature in degrees Celsius for this diary day
- Occupation code
- Time, in minutes, during this diary day for which no data are included in the database

3.3.5.2 Diary Events file

The human activity diary data are contained in the events file that is distributed with APEX. This file contains the activities for the nearly 23,000 people with intervals ranging from one minute to one hour. An individuals' diary varies in length from one to 15 days. This file contains the following variables:

- The study, person, and diary day identifiers
- Start time of this activity
- Number of minutes for this activity
- Activity code (a record of what the individual was doing)
- Location code (a record of where the individual was)

Table 46. Summary of activity pattern studies used in CHAD.

Study Name	Location	Study time period	Ages	Persons	Person-days	Diary type /study design	Reference
Baltimore	A single building in Baltimore	01/1997-02/1997, 07/1998-08/1998	72-93	26	292	Diary	Williams et al. (2000)
California Adolescents and Adults (CARB)	California	10/1987-09/1988	12-17 18-94	181 1,552	181 1,552	Recall /Random	Robinson et al. (1989); Wiley et al. (1991a)
California Children (CARB)	California	04/1989-02/1990	0-11	1,200	1,200	Recall /Random	Wiley et al. (1991b)
Cincinnati (EPRI)	Cincinnati MSA	03/1985-04/1985, 08/1985	0-86	888	2,587	Diary /Random	Johnson (1989)
Denver (EPA)	Denver MSA	11/1982-02/1983	18-70	432	791	Diary /Random	Johnson (1984); Akland et al. (1985)
Los Angeles: Elementary School Children	Los Angeles	10/1989	10-12	17	51	Diary	Spier et al. (1992)
Los Angeles: High School Adolescents	Los Angeles	09/1990-10/1990	13-17	19	42	Diary	Spier et al. (1992)
National: NHAPS-Air	National	09/1992-10/1994	0-93	4,326	4,326	Recall /Random	Klepeis et al. (1996); Tsang and Klepeis (1996)
National: NHAPS-Water	National	09/1992-10/1994	0-93	4,332	4,332	Recall /Random	Klepeis et al. (1996); Tsang and Klepeis (1996)
Washington, D.C. (EPA)	Wash. DC MSA	11/1982-02/1983	18-98	639	639	Diary /Random	Hartwell et al. (1984); Akland et al. (1985)

3.3.5.3 Construction of Longitudinal Activity Sequences

Typical time-activity pattern data available for inhalation exposure modeling consist of a sequence of location/activity combinations spanning a 24-hour duration, with 1 to 3 diary-days for any single individual. Exposure modeling requires information on activity patterns over longer periods of time, e.g., a full year. For example, even for pollutant health effects with short averaging times (e.g., NO_2 1-hour average concentration) it may be desirable to know the frequency of exceedances of a concentration over a long period of time (e.g., the annual number of exceedances of a 1-hour average NO_2 concentration of 200 ppb for each simulated individual).

Long-term multi-day activity patterns can be estimated from single days by combining the daily records in various ways, and the method used for combining them will influence the variability of the long-term activity patterns across the simulated population. This in turn will influence the ability of the model to accurately represent either long-term average high-end exposures, or the number of individuals exposed multiple times to short-term high-end concentrations.

A common approach for constructing long-term activity patterns from short-term records is to re-select a daily activity pattern from the pool of data for each day, with the implicit assumption that there is no correlation between activities from day to day for the simulated individual. This approach tends to result in long-term activity patterns that are very similar across the simulated population. Thus, the resulting exposure estimates are likely to underestimate the variability across the population, and therefore, underestimate the high-end exposure concentrations or the frequency of exceedances.

A contrasting approach is to select a single activity pattern (or a single pattern for each season and/or weekday-weekend) to represent a simulated individual's activities over the duration of the exposure assessment. This approach has the implicit assumption that an individual's day-to-day activities are perfectly correlated. This approach tends to result in long-term activity patterns that are very different across the simulated population, and therefore may over-estimate the variability across the population.

Cluster-Markov Algorithm

A new algorithm has been developed and incorporated into APEX to represent the day-to-day correlation of activities for individuals. The algorithms first use cluster analysis to divide the daily activity pattern records into groups that are similar, and then select a single daily record from each group. This limited number of daily patterns is then used to construct a long-term sequence for a simulated individual, based on empirically-derived transition probabilities. This approach is intermediate between the assumption of no day-to-day correlation (i.e., re-selection for each time period) and perfect correlation (i.e., selection of a single daily record to represent all days).

The steps in the algorithm are as follows.

1. For each demographic group (age, gender, employment status), temperature range, and day-of-week combination, the associated time-activity records are partitioned into 3 groups using cluster analysis. The clustering criterion is a vector of 5 values: the time spent in each of 5 microenvironment categories (indoors – residence; indoors – other building; outdoors – near road; outdoors – away from road; in vehicle).
2. For each simulated individual, a single time-activity record is randomly selected from each cluster.
3. A Markov process determines the probability of a given time-activity pattern occurring on a given day based on the time-activity pattern of the previous day and cluster-to-cluster transition probabilities. The cluster-to-cluster transition probabilities are estimated from the available multi-day time-activity records. If insufficient multi-day time-activity records are available for a demographic group, season, day-of-week combination, then the cluster-to-cluster transition probabilities are estimated from the frequency of time-activity records in each cluster in the CHAD data base.

Details regarding the Cluster-Markov algorithm and supporting evaluations are provided in Appendix F.

3.3.6 Calculating Microenvironmental Concentrations

Probabilistic algorithms are used to estimate the pollutant concentration associated with each exposure event. The estimated pollutant concentrations account for the effects of ambient (outdoor) pollutant concentration, penetration factor, air exchange rate, decay/deposition rate, and proximity to microenvironments can use the transfer factors method while the others use the mass balance emission sources, depending on the microenvironment, available data, and the estimation method selected by the user.

APEX calculates air concentrations in the various microenvironments visited by the simulated person by using the ambient air data for the relevant blocks, the user-specified estimation method, and input parameters specific to each microenvironment. APEX calculates hourly concentrations in all the microenvironments at each hour of the simulation for each of the simulated individuals using one of two methods: by mass balance or a transfer factors method.

3.3.6.1 Mass Balance Model

The mass balance method simulates an enclosed microenvironment as a well-mixed volume in which the air concentration is spatially uniform at any specific time. The concentration of an air pollutant in such a microenvironment is estimated using the following processes:

- Inflow of air into the microenvironment
- Outflow of air from the microenvironment
- Removal of a pollutant from the microenvironment due to deposition, filtration, and chemical degradation
- Emissions from sources of a pollutant inside the microenvironment.

Table 47 lists the parameters required by the mass balance method to calculate concentrations in a microenvironment. A proximity factor ($f_{proximity}$) is used to account for differences in ambient concentrations between the geographic location represented by the ambient air quality data (e.g., a regional fixed-site monitor or modeled concentration) and the geographic location of the microenvironment (e.g., near a roadway). This factor could take a value either greater than or less than 1. Emission source (ES) represents the emission rate for the emission source and concentration source (CS) is the mean air concentration resulting from the source. $R_{removal}$ is defined as the removal rate of a pollutant from a microenvironment due to deposition, filtration, and chemical reaction. The air exchange rate ($R_{air\ exchange}$) is expressed in air changes per hour.

Table 47. Mass balance model parameters.

Variable	Definition	Units	Value Range
$f_{proximity}$	Proximity factor	unitless	$f_{proximity} \geq 0$
CS	Concentration source	ppb	$CS \geq 0$
$R_{removal}$	Removal rate due to deposition, filtration, and chemical reaction	1/hr	$R_{removal} \geq 0$
$R_{air\ exchange}$	Air exchange rate	1/hr	$R_{air\ exchange} \geq 0$
V	Volume of microenvironment	m^3	$V > 0$

The mass balance equation for a pollutant in a microenvironment is described by:

$$\frac{dC_{\mathrm{ME}}(t)}{dt} = \Delta C_{in} - \Delta C_{out} - \Delta C_{removal} + \Delta C_{source} \qquad \text{eq (4)}$$

where:

$dC_{\mathrm{ME}}(t)$ = Change in concentration in a microenvironment at time t (ppb),

ΔC_{in} = Rate of change in microenvironmental concentration due to influx of air (ppb/hour),

ΔC_{out} = Rate of change in microenvironmental concentration due to outflux of air (ppb/hour),

$\Delta C_{removal}$ = Rate of change in microenvironmental concentration due to removal processes (ppb/hour), and

ΔC_{source} = Rate of change in microenvironmental concentration due to an emission source inside the microenvironment (ppb/hour).

Within the time period of an hour each of the rates of change, ΔC_{in}, ΔC_{out}, $\Delta C_{removal}$, and ΔC_{source}, is assumed to be constant. At each hour time step of the simulation period, APEX estimates the hourly equilibrium, hourly ending, and hourly mean concentrations using a series of equations that account for concentration changes expected to occur due to these physical processes. Details regarding these equations are provided in the APEX User's Guide. APEX reports hourly mean concentration as hourly concentration for a specific hour. The calculation then continues to the next hour by using the end concentration for the previous hour as the initial microenvironmental concentration. A description of the input parameters estimates used for microenvironments using the mass balance approach is provided below.

3.3.6.2 Factors Model

The factors method is simpler than the mass balance method. It does not calculate concentration in a microenvironment from the concentration in the previous hour and it has fewer parameters. Table 48 lists the parameters required by the factors method to calculate concentrations in a microenvironment without emissions sources.

Table 48. Factors model parameters.

Variable	Definition	Units	Value Range
$f_{proximity}$	Proximity factor	unitless	$f_{proximity} \geq 0$
$f_{penetration}$	Penetration factor	unitless	$0 \leq f_{penetration} \leq 1$

The factors method uses the following equation to calculate hourly mean concentration in a microenvironment from the user-provided hourly air quality data:

$$C_{ME}^{hourlymean} = C_{ambient} \; x \; f_{proximity} \; x \; f_{penetration} \qquad \text{eq (5)}$$

where:

$C_{ME}^{hourlymean}$ = Hourly concentration in a microenvironment (ppb)

$C_{ambient}$ = Hourly concentration in ambient environment (ppb)

$f_{proximity}$ = Proximity factor (unitless)

$f_{penetration}$ = Penetration factor (unitless)

The ambient NO_2 concentrations are from the air quality data input file. The proximity factor is a unitless parameter that represents the proximity of the microenvironment to a monitoring station. The penetration factor is a unitless parameter that represents the fraction of pollutant entering a microenvironment from outside the microenvironment via air exchange. The development of the specific proximity and penetration factors used in this analysis are discussed below for each microenvironment using this approach.

3.3.6.3 Microenvironments Modeled

In APEX, microenvironments represent the exposure locations for simulated individuals. For exposures to be estimated accurately, it is important to have realistic microenvironments that match closely to the locations where actual people spend time on a daily basis. As discussed above, the two methods available in APEX for calculating pollutant levels within microenvironments are: 1) factors and 2) mass balance. A list of microenvironments used in this study, the calculation method used, and the parameters used to calculate the microenvironment concentrations can be found in Table 49.

Table 49. List of microenvironments and calculation methods used.

Microenvironment		**Calculation Method**	**Parameter Types used** [1]
No.	**Name**		
1	Indoors – Residence	Mass balance	AER and DE
2	Indoors – Bars and restaurants	Mass balance	AER and DE
3	Indoors – Schools	Mass balance	AER and DE
4	Indoors – Day-care centers	Mass balance	AER and DE
5	Indoors – Office	Mass balance	AER and DE
6	Indoors – Shopping	Mass balance	AER and DE
7	Indoors – Other	Mass balance	AER and DE
8	Outdoors – Near road	Factors	PR
9	Outdoors – Public garage - parking lot	Factors	PR
10	Outdoors – Other	Factors	None
11	In-vehicle – Cars and Trucks	Factors	PE and PR
12	In-vehicle - Mass Transit (bus, subway, train)	Factors	PE and PR
0	Not modeled		

[1] AER=air exchange rate, DE=decay-deposition rate, PR=proximity factor, PE=penetration factor

Each of the microenvironments is designed to simulate an environment in which people spend time during the day. CHAD locations are linked to the different microenvironments in the *Microenvironment Mapping* File (see below). There are many more CHAD locations than microenvironment locations (there are 113 CHAD codes versus 12 microenvironments in this assessment), therefore most of the microenvironments have multiple CHAD locations mapped to them.

3.3.6.4 Microenvironment Descriptions

Microenvironment 1: Indoor-Residence

The Indoors-Residence microenvironment uses several variables that affect NO_2 exposure: whether or not air conditioning is present, the average outdoor temperature, the NO_2 removal rate, and an indoor concentration source. The first two of these variables affect the air exchange rate.

Since the selection of an air exchange rate distribution is conditioned on the presence or absence of an air-conditioner, for each modeled area the air conditioning status of the residential microenvironments is simulated randomly using the probability that a residence has an air conditioner. For this study, location-specific air conditioning prevalence was taken from the American Housing Survey of 2003 (AHS, 2003a; 2003b). Previous analyses (US EPA, 2007d) detail the specification of uncertainty estimates in the form of confidence intervals for the air conditioner prevalence using the following:

$$\text{Standard Error}\,(P) = \sqrt{\frac{3850\,P\,(1-P)}{N}},$$

$$\text{Confidence Interval}\,(P) = P \pm 1.96 \times \text{Standard Error}\,(P)$$

where P is the estimated percentage and N is the estimated total number of housing units. Table 50 contains the values for air conditioning prevalence used for each modeled location.

Table 50. Air conditioning prevalence estimates with 95% confidence intervals.

AHS Survey	Housing Units	A/C Prevalence (%)	se	L95	U95
Atlanta	797,687	97.0	1.2	94.7	99.3
Detroit	1,877,178	81.4	1.8	78.0	84.9
Los Angeles	3,296,819	55.1	1.7	51.7	58.4
Philadelphia	1,943,492	90.6	1.3	88.1	93.2
Phoenix	-	-	-	-	-

Notes:
se – Standard error
L95 – Lower limit on 95th confidence interval
U95 – Upper limit on 95th confidence interval

Air exchange rate data for the indoor residential microenvironment were obtained from US EPA (2007d). Briefly, residential air exchange rate (AER) data were obtained from several studies (Avol et al., 1998; Williams et al., 2003a, 2003b; Meng et al., 2004; Weisel et al., 2004; Chillrud at al, 2004; Kinney et al., 2002; Sax et al., 2004; Wilson et al., 1986, 1996; Colome et al., 1993, 1994; Murray and Burmaster, 1995). Influential characteristics (e.g., temperature, air conditioning), where reported in the study, were also compiled for use in statistical analyses. Descriptive statistics were generated for each location/variable type and evaluated using statistical comparison testing (e.g., ANOVA). Based on the summary statistics and the statistical comparisons, different AER distributions were fit for each combination of A/C type, city, and temperature. In general, lognormal distributions provided the best fit, and are defined by a geometric mean (GM) and standard deviation (GSD). To avoid unusually extreme simulated AER values, bounds of 0.1 and 10 were selected for minimum and maximum AER, respectively.

Fitted distributions were available for one of the cities modeled in this assessment, Los Angeles. For the other 4 of the cities to be modeled, a distribution was selected from one of the other locations thought to have similar characteristics to the city to be modeled, qualitatively considering factors that might influence AERs. These factors include the age composition of housing stock, construction methods, and other meteorological variables not explicitly treated in the analysis, such as humidity and wind speed patterns. The distributions used for these each of the modeled locations are provide in Table 51.

Table 51. Geometric means (GM) and standard deviations (GSD) for air exchange rates by city, A/C type, and temperature range.

Area Modeled	Study City	A/C Type	Temp (ºC)	N	GM	GSD
	Houston	Central or Room A/C	<=20	15	0.4075	2.1135
			20-25	20	0.4675	1.9381
			25-30	65	0.4221	2.2579
			>30	14	0.4989	1.7174
		No A/C	<=10	13	0.6557	1.6794
			10-20	28	0.6254	2.9162
			>20	12	0.9161	2.4512
	Inland California	Central or Room A/C	<=25	226	0.5033	1.9210
			>25	83	0.8299	2.3534
		No A/C	<=10	17	0.5256	3.1920
			10-20	52	0.6649	2.1743
			20-25	13	1.0536	1.7110
			>25	14	0.8271	2.2646
Los Angeles	Los Angeles	Central or Room A/C	<=20	721	0.5894	1.8948
			20-25	273	1.1003	2.3648
			25-30	102	0.8128	2.4151
			>30	12	0.2664	2.7899
		No A/C	<=10	18	0.5427	3.0872
			10-20	390	0.7470	2.0852
			20-25	148	1.3718	2.2828
			>25	25	0.9884	1.9666
Philadelphia and Detroit	New York City	Central or Room A/C	<=10	20	0.7108	2.0184
			10-25	42	1.1392	2.6773
			>25	19	1.2435	2.1768
		No A/C	<=10	48	1.0165	2.1382
			10-20	59	0.7909	2.0417
			>20	32	1.6062	2.1189
Atlanta (No A/C)	Outside California	Central or Room A/C	<=10	179	0.9185	1.8589
			10-20	338	0.5636	1.9396
			20-25	253	0.4676	2.2011
			25-30	219	0.4235	2.0373
			>30	24	0.5667	1.9447
		No A/C	<=10	61	0.9258	2.0836
			10-20	87	0.7333	2.3299
			>20	44	1.3782	2.2757
Atlanta (A/C)	Research Triangle Park, NC	Central or Room A/C	<=10	157	0.9617	1.8094
			10-20	320	0.5624	1.9058
			20-25	196	0.3970	1.8887
			>25	145	0.3803	1.7092

For this analysis, the same NO_2 removal rate distribution was used for all microenvironments that use the mass balance method. This removal rate is based on data provided by Spicer et al. (1993). A total of 6 experiments, under variable source emission characteristics including operation of gas stove, were conducted in an unoccupied test house. A distribution could not be described with the limited data set, therefore a uniform distribution was approximated by the bounds of the 6 values, a minimum of 1.02 and a maximum of 1.45 h^{-1}.

An excerpt from the APEX input file describing the indoor residential microenvironment is provided in Figure 19. The first section of the input file excerpt specifies the air exchange rate distributions for the microenvironment. Average temperature and air conditioning presence, which are city-specific, were coded into air exchange rate *conditional variables, C1* and *C2*, respectively. Average temperatures were separated into five categories (variable *C1*, numbered 1-5): 50 º F, 50-68 º F, 68-77 º F, 77-86 º F, and 86 º F and above. For variable *C2*, air conditioning status can range from *1* to *2* (*1* for having air conditioning, *2* for not having it). The air exchange rate estimates generated previously in the form of lognormal distributions were entered into the appropriate temperature and A/C category for each location for a total of ten distributions (i.e., 5 temperature distributions by 2 air conditioning distributions). In the input file example however, there are actually four AER distributions for homes with an air conditioner and three for those without; the last few distributions for each air conditioning setting were the same due to the available data to populate the field. The parameter estimates for the removal factor (DE) is also shown following the AER data.

```
Micro number     = 1        !     Indoors - residence - AIR EXCHANGE RATES
Parameter Type   = AER
Condition # 1    = AvgTempCat
Condition # 2    = AC_Home
ResampHours      = NO
ResampDays       = YES
ResampWork       = YES
Block DType Season Area C1  C2  C3  Shape      Par1    Par2  Par3 Par4  LTrunc UTrunc
1     1     1      1    1   1   1   Lognormal  0.711   2.018  0    .      0.1    10
1     1     1      1    2   1   1   Lognormal  1.139   2.677  0    .      0.1    10
1     1     1      1    3   1   1   Lognormal  1.139   2.677  0    .      0.1    10
1     1     1      1    4   1   1   Lognormal  1.244   2.177  0    .      0.1    10
1     1     1      1    5   1   1   Lognormal  1.244   2.177  0    .      0.1    10
1     1     1      1    1   2   1   Lognormal  1.016   2.138  0    .      0.1    10
1     1     1      1    2   2   1   Lognormal  0.791   2.042  0    .      0.1    10
1     1     1      1    3   2   1   Lognormal  1.606   2.119  0    .      0.1    10
1     1     1      1    4   2   1   Lognormal  1.606   2.119  0    .      0.1    10
1     1     1      1    5   2   1   Lognormal  1.606   2.119  0    .      0.1    10

Micro number     = 1              !    DECAY RATES
Pollutant = 1
Parameter Type   = DE
ResampHours      = NO
ResampDays       = NO
ResampWork       = YES
Block DType Season Area C1  C2  C3  Shape      Par1  Par2  Par3 Par4  LTrunc UTrunc
1     1     1      1    1   1   1   Uniform    1.02  1.45   .    .     1.02   1.45
```

Figure 19. Example input file from APEX for Indoors-residence microenvironment.

Indoor source contributions. A number of studies, as described in section 2.5.5 of the NO_x ISA, have noted the importance of gas cooking appliances as sources of NO_2 emissions. An indoor emission source term was included in the APEX simulations to estimate exposure to indoor sources of NO_2. Three types of data were used to implement this factor:

- The fraction of households in the Philadelphia MSA that use gas for cooking fuel
- The range of contributions to indoor NO_2 concentrations that occur from cooking with gas
- The diurnal pattern of cooking in households.

The fraction of households in Philadelphia County that use gas cooking fuel (i.e., 55%) was taken from the *US Census Bureau's American Housing Survey for the Philadelphia Metropolitan Area: 2003*.

Data used for estimating the contribution to indoor NO_2 concentrations that occur during cooking with gas fuel were derived from a study sponsored by the California Air Resources Board (CARB, 2001). For this study a test house was set up for continuous measurements of NO_2 indoors and outdoors, among several other parameters, and conducted under several different cooking procedures and stove operating conditions. A uniform distribution of concentration contributions for input to APEX was estimated as follows.

- The concurrent outdoor NO_2 concentration measurement was subtracted from each indoor concentration measurement, to yield net indoor concentrations
- Net indoor concentrations for duplicate cooking tests (same food cooked the same way) were averaged for each indoor room, to yield average net indoor concentrations
- The minimum and maximum average net indoor concentrations for any test in any room were used as the lower and upper bounds of a uniform distribution

This resulted in a minimum average net indoor concentration of 4 ppb and a maximum net average indoor concentration of 188 ppb.

An analysis by Johnson et al (1999) of survey data on gas stove usage collected by Koontz et al (1992) showed an average number of meals prepared each day with a gas stove of 1.4. The diurnal allocation of these cooking events was estimated as follows.

- Food preparation time obtained from CHAD diaries was stratified by hour of the day, and summed for each hour, and summed for total preparation time.
- The fraction of food preparation occurring in each hour of the day was calculated as the total number of minutes for that hour divided by the overall total preparation time. The result was a measure of the probability of food preparation taking place during any hour, given one food preparation event per day.
- Each hourly fraction was multiplied by 1.4, to normalize the expected value of daily food preparation events to 1.4.

The estimated probabilities of cooking by hour of the day are presented in Table 52. For this analysis it was assumed that the probability that food preparation would include stove usage was the same for each hour of the day, so that the diurnal allocation of food preparation events would be the same as the diurnal allocation of gas stove usage. It was also assumed that each cooking event lasts for exactly 1 hour, implying that the average total daily gas stove usage is 1.4 hours.

Table 52. Probability of gas stove cooking by hour of the day.

Hour of Day	Probability of Cooking (%)[1]
0	0
1	0
2	0
3	0
4	0
5	5
6	10
7	10
8	10
9	5
10	5
11	5
12	10
13	5
14	5
15	5
16	15
17	20
18	15
19	10
20	5
21	5
22	0
23	0

[1] Values rounded to the nearest 5%. Data sum to 145% due to rounding and scaling to 1.4 cooking events/day.

Microenvironments 2-7: All other indoor microenvironments

The remaining five indoor microenvironments, which represent Bars and Restaurants, Schools, Day Care Centers, Office, Shopping, and Other environments, are all modeled using the same data and functions (Figure 20). As with the Indoor-Residence microenvironment, these microenvironments use both air exchange rates and removal rates to calculate exposures within the microenvironment. The air exchange rate distribution (GM = 1.109, GSD = 3.015, Min = 0.07, Max = 13.8) was developed based on an indoor air quality study (Persily et al, 2005; see US EPA, 2007d for details in derivation). The decay rate is the same as used in the Indoor-Residence microenvironment discussed previously. The Bars and Restaurants microenvironment included an estimated contribution from indoor sources as was described for the Indoor-Residence, only there was an assumed 100% prevalence rate and the cooking with the gas appliance occurred at any hour of the day.

```
Micro number     = 2       !    Bars & restaurants     - AIR EXCHANGE RATES
Parameter Type    = AER
ResampHours      = NO
ResampDays       = YES
ResampWork       = YES
Block DType Season Area C1  C2  C3   Shape        Par1   Par2  Par3 Par4  LTrunc UTrunc
1        1          1        1    1    1    1   LogNormal  1.109  3.015   0     .      0.07       13.8

Micro number     = 2               !    DECAY RATES
Pollutant = 1
Parameter Type    = DE
ResampHours      = NO
ResampDays       = YES
ResampWork       = YES
Block DType Season Area C1  C2  C3   Shape      Par1   Par2  Par3 Par4  LTrunc UTrunc
1        1         1         1     1     1   1   Uniform     1.02    1.45   .      .      1.02       1.45
```

Figure 20. Example input file from APEX for all Indoors microenvironments, other than Indoors-residence.

Microenvironments 8 and 9: Outdoor microenvironments

Two outdoor microenvironments, the Near Road and Public Garage/Parking Lot, used the factors method to calculate pollutant exposure. Penetration factors are not applicable to outdoor environments (effectively, PEN=1). Proximity factors were developed from the AERMOD concentration predictions, i.e., the block-centroid-to-nearest-roadway concentration ratios. Based on the resulting sets of ratio values, the ratio distributions were stratified by hour of the day into 3 groups as indicated by the “hours-block” specification in the example file in Figure 21. The lower and upper bounds for sampling were specified as the 5th and 95th percentile values, respectively, of each distribution.

```
Micro number     = 8       !    Outdoor near road      PROXIMITY FACTOR
Pollutant = 1
Parameter Type    = PR
Hours - Block      =       1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 1 1
ResampHours      = YES
ResampDays       = YES
ResampWork       = YES
Block DType Season Area C1  C2  C3   Shape      Par1   Par2  Par3 Par4  LTrunc UTrunc  ResampOut
1    1    1     1    1   1   1  LogNormal   1.251  1.478  0.    .    0.86  2.92     Y
2    1    1     1    1   1   1  LogNormal   1.555  1.739  0.    .    0.83  4.50     Y
3    1    1     1    1   1   1  LogNormal   1.397  1.716  0.    .    0.73  4.17     Y
```

Figure 21. Example input file from APEX for outdoor near road microenvironment.

Microenvironment 10: Outdoors-General.

The general outdoor environment concentrations are well represented by the modeled concentrations. Therefore, both the penetration factor and proximity factor for this microenvironment were set to 1.

Microenvironments 11 and 12: In Vehicle- Cars and Trucks, and Mass Transit

Penetration factors were developed from data provided in Chan and Chung (2003). Inside-vehicle and outdoor NO_2 concentrations were measured with for three ventilation conditions, air-recirculation, fresh air intake, and with windows opened. Since major roads were the focus of this assessment, reported indoor/outdoor ratios for highway and urban streets were used here. Mean values range from about 0.6 to just over 1.0, with higher values associated with increased ventilation (i.e., window open). A uniform distribution was selected for the penetration factor for Inside-Cars/Trucks (ranging from 0.6 to 1.0) due to the limited data available to describe a more formal distribution and the lack of data available to reasonably assign potentially influential characteristics such as use of vehicle ventilation systems for each location. Mass transit systems, due to the frequent opening and closing of doors, was assigned a uniform distribution ranging from 0.8 to 1.0 based on the reported mean values for fresh air intake and open windows.

Proximity factors were developed as described above for Microenvironments 8 and 9

3.3.6.5 Microenvironment Mapping

The *Microenvironment Mapping* file matches the APEX Microenvironments to CHAD Location codes. Table 53 gives the mapping used for the APEX simulations.

Table 53. Mapping of CHAD activity locations to APEX microenvironments.

```
CHAD Loc.  Description                          APEX micro
---------  -------------------------------------------------
U          Uncertain of correct code          =   -1  Unknown
X          No data                            =   -1  Unknown
30000      Residence, general                 =    1  Indoors-Residence
30010      Your residence                     =    1  Indoors-Residence
30020      Other residence                    =    1  Indoors-Residence
30100      Residence, indoor                  =    1  Indoors-Residence
30120      Your residence, indoor             =    1  Indoors-Residence
30121      ..., kitchen                       =    1  Indoors-Residence
30122      ..., living room or family room    =    1  Indoors-Residence
30123      ..., dining room                   =    1  Indoors-Residence
30124      ..., bathroom                      =    1  Indoors-Residence
30125      ..., bedroom                       =    1  Indoors-Residence
30126      ..., study or office               =    1  Indoors-Residence
30127      ..., basement                      =    1  Indoors-Residence
30128      ..., utility or laundry room       =    1  Indoors-Residence
30129      ..., other indoor                  =    1  Indoors-Residence
30130      Other residence, indoor            =    1  Indoors-Residence
30131      ..., kitchen                       =    1  Indoors-Residence
30132      ..., living room or family room    =    1  Indoors-Residence
30133      ..., dining room                   =    1  Indoors-Residence
30134      ..., bathroom                      =    1  Indoors-Residence
30135      ..., bedroom                       =    1  Indoors-Residence
30136      ..., study or office               =    1  Indoors-Residence
30137      ..., basement                      =    1  Indoors-Residence
30138      ..., utility or laundry room       =    1  Indoors-Residence
30139      ..., other indoor                  =    1  Indoors-Residence
30200      Residence, outdoor                 =   10  Outdoors-Other
30210      Your residence, outdoor            =   10  Outdoors-Other
30211      ..., pool or spa                   =   10  Outdoors-Other
30219      ..., other outdoor                 =   10  Outdoors-Other
30220      Other residence, outdoor           =   10  Outdoors-Other
```

```
30221      ..., pool or spa                      =   10  Outdoors-Other
30229      ..., other outdoor                    =   10  Outdoors-Other
30300      Residential garage or carport         =    7  Indoors-Other
30310      ..., indoor                           =    7  Indoors-Other
30320      ..., outdoor                          =   10  Outdoors-Other
30330      Your garage or carport                =    1  Indoors-Residence
30331      ..., indoor                           =    1  Indoors-Residence
30332      ..., outdoor                          =   10  Outdoors-Other
30340      Other residential garage or carport   =    1  Indoors-Residence
30341      ..., indoor                           =    1  Indoors-Residence
30342      ..., outdoor                          =   10  Outdoors-Other
30400      Residence, none of the above          =    1  Indoors-Residence
31000      Travel, general                       =   11  In Vehicle-Cars_and_Trucks
31100      Motorized travel                      =   11  In Vehicle-Cars_and_Trucks
31110      Car                                   =   11  In Vehicle-Cars_and_Trucks
31120      Truck                                 =   11  In Vehicle-Cars_and_Trucks
31121      Truck (pickup or van)                 =   11  In Vehicle-Cars_and_Trucks
31122      Truck (not pickup or van)             =   11  In Vehicle-Cars_and_Trucks
31130      Motorcycle or moped                   =    8  Outdoors-Near_Road
31140      Bus                                   =   12  In Vehicle-Mass_Transit
31150      Train or subway                       =   12  In Vehicle-Mass_Transit
31160      Airplane                              =    0  Zero_concentration
31170      Boat                                  =   10  Outdoors-Other
31171      Boat, motorized                       =   10  Outdoors-Other
31172      Boat, other                           =   10  Outdoors-Other
31200      Non-motorized travel                  =   10  Outdoors-Other
31210      Walk                                  =   10  Outdoors-Other
31220      Bicycle or inline skates/skateboard   =   10  Outdoors-Other
31230      In stroller or carried by adult       =   10  Outdoors-Other
31300      Waiting for travel                    =   10  Outdoors-Other
31310      ..., bus or train stop                =    8  Outdoors-Near_Road
31320      ..., indoors                          =    7  Indoors-Other
31900      Travel, other                         =   11  In Vehicle-Cars_and_Trucks
31910      ..., other vehicle                    =   11  In Vehicle-Cars_and_Trucks
32000      Non-residence indoor, general         =    7  Indoors-Other
32100      Office building/ bank/ post office    =    5  Indoors-Office
32200      Industrial/ factory/ warehouse        =    5  Indoors-Office
32300      Grocery store/ convenience store      =    6  Indoors-Shopping
32400      Shopping mall/ non-grocery store      =    6  Indoors-Shopping
32500      Bar/ night club/ bowling alley        =    2  Indoors-Bars_and_Restaurants
32510      Bar or night club                     =    2  Indoors-Bars_and_Restaurants
32520      Bowling alley                         =    2  Indoors-Bars_and_Restaurants
32600      Repair shop                           =    7  Indoors-Other
32610      Auto repair shop/ gas station         =    7  Indoors-Other
32620      Other repair shop                     =    7  Indoors-Other
32700      Indoor gym /health club               =    7  Indoors-Other
32800      Childcare facility                    =    4  Indoors-Day_Care_Centers
32810      ..., house                            =    1  Indoors-Residence
32820      ..., commercial                       =    4  Indoors-Day_Care_Centers
32900      Large public building                 =    7  Indoors-Other
32910      Auditorium/ arena/ concert hall       =    7  Indoors-Other
32920      Library/ courtroom/ museum/ theater   =    7  Indoors-Other
33100      Laundromat                            =    7  Indoors-Other
33200      Hospital/ medical care facility       =    7  Indoors-Other
33300      Barber/ hair dresser/ beauty parlor   =    7  Indoors-Other
33400      Indoors, moving among locations       =    7  Indoors-Other
33500      School                                =    3  Indoors-Schools
33600      Restaurant                            =    2  Indoors-Bars_and_Restaurants
33700      Church                                =    7  Indoors-Other
33800      Hotel/ motel                          =    7  Indoors-Other
33900      Dry cleaners                          =    7  Indoors-Other
34100      Indoor parking garage                 =    7  Indoors-Other
34200      Laboratory                            =    7  Indoors-Other
```

```
34300      Indoor, none of the above          =    7  Indoors-Other
35000      Non-residence outdoor, general     =   10  Outdoors-Other
35100      Sidewalk, street                   =    8  Outdoors-Near_Road
35110      Within 10 yards of street          =    8  Outdoors-Near_Road
35200      Outdoor public parking lot /garage  =    9  Outdoors-Public_Garage-Parking
35210      ..., public garage                 =    9  Outdoors-Public_Garage-Parking
35220      ..., parking lot                   =    9  Outdoors-Public_Garage-Parking
35300      Service station/ gas station       =   10  Outdoors-Other
35400      Construction site                  =   10  Outdoors-Other
35500      Amusement park                     =   10  Outdoors-Other
35600      Playground                         =   10  Outdoors-Other
35610      ..., school grounds                =   10  Outdoors-Other
35620      ..., public or park                =   10  Outdoors-Other
35700      Stadium or amphitheater            =   10  Outdoors-Other
35800      Park/ golf course                  =   10  Outdoors-Other
35810      Park                               =   10  Outdoors-Other
35820      Golf course                        =   10  Outdoors-Other
35900      Pool/ river/ lake                  =   10  Outdoors-Other
36100      Outdoor restaurant/ picnic         =   10  Outdoors-Other
36200      Farm                               =   10  Outdoors-Other
36300      Outdoor, none of the above         =   10  Outdoors-Other
```

3.3.7 Exposure Calculations

APEX calculates exposure as a time series of exposure concentrations that a simulated individual experiences during the simulation period. APEX determines the exposure using hourly ambient air concentrations, calculated concentrations in each microenvironment based on these ambient air concentrations (and indoor sources if present), and the minutes spent in a sequence of microenvironments visited according to the composite diary. The hourly exposure concentration at any clock hour during the simulation period is determined using the following equation:

$$C_i = \frac{\sum_{j=1}^{N} C_{ME(j)}^{hourlymean} \quad t_{(j)}}{T} \qquad \text{eq (6)}$$

where:

C_i = Hourly exposure concentration at clock hour *i* of the simulation period (ppb)

N = Number of events (i.e., microenvironments visited) in clock hour *i* of the simulation period.

$C_{ME(j)}^{hourlymean}$ = Hourly mean concentration in microenvironment *j* (ppm)

$t_{(j)}$ = Time spent in microenvironment *j* (minutes)

T = 60 minutes

From the hourly exposures, APEX calculates time series of 1-hour average exposure concentrations that a simulated individual would experience during the simulation period. APEX then statistically summarizes and tabulates the hourly (or daily, annual average) exposures. In this analysis, the exposure indicator is 1-hr exposures above selected health effect benchmark levels. From this, APEX can calculate two general types of exposure estimates:

counts of the estimated number of people exposed to a specified NO_2 concentration level and the number of times per year that they are so exposed; the latter metric is in terms of person-occurrences or person-days. The former highlights the number of individuals exposed at least *one or more* times per modeling period to the health effect benchmark level of interest. APEX can also report counts of individuals with multiple exposures. This person-occurrences measure estimates the number of times per season that individuals are exposed to the exposure indicator of interest and then accumulates these estimates for the entire population residing in an area.

APEX tabulates and displays the two measures for exposures above levels ranging from 200 to 300 ppb by 50 ppb increments for 1-hour average exposures. These results are tabulated for the population and subpopulations of interest.

To simulate just meeting the current standard, dispersion modeled concentration were not rolled-up as done in the air quality characterization. A proportional approach was used as done in the Air Quality Characterization, but to reduce processing time, the health effect benchmark levels were proportionally reduced by the similar factors described for each specific location and simulated year. Since it is a proportional adjustment, the end effect of adjusting concentrations upwards versus adjusting benchmark levels downward within the model is the same. The difference in the exposure and risk modeling was that the modeled air quality concentrations were used to generate the adjustment factors. Table 54 provides the adjustment factors used and the adjusted potential health effect benchmark concentrations to simulate just meeting the current standard. When modeling indoor sources, the indoor concentration contributions needed to be scaled downward by the same proportions.

Table 54. Adjustment factors and potential health effect benchmark levels used by APEX to simulate just meeting the current standard

Simulated Year (factor)	Potential Health Effect Benchmark Level (ppb)	
	Actual	Adjusted
2001 (1.59)	150	94
	200	126
	250	157
	300	189
2002 (1.63)	150	92
	200	122
	250	153
	300	184
2003 (1.64)	150	91
	200	122
	250	152
	300	183

3.3.8 Exposure Model Output

All of the output files written by APEX are ASCII text files. Table 55 lists each of the output data files written for these simulations and provides descriptions of their content. Additional output files that can produced by APEX are given in Table 5-1 of the APEX User's Guide, and include hourly exposure, ventilation, and energy expenditures, and even detailed event-level information, if desired. The names and locations, as well as the output table levels (e.g., output percentiles, cut-points), for these output files are specified by the user in the simulation control parameters file.

Table 55. Example of APEX output files.

Output File Type	Description
Log	The *Log* file contains the record of the APEX model simulation as it progresses. If the simulation completes successfully, the log file indicates the input files and parameter settings used for the simulation and reports on a number of different factors. If the simulation ends prematurely, the log file contains error messages describing the critical errors that caused the simulation to end.
Profile Summary	The *Profile Summary* file provides a summary of each individual modeled in the simulation.
Microenvironment Summary	The *Microenvironment Summary* file provides a summary of the time and exposure by microenvironment for each individual modeled in the simulation.
Sites	The *Sites* file lists the tracts, districts, and zones in the study area, and identifies the mapping between them.
Output Tables	The *Output Tables* file contains a series of tables summarizing the results of the simulation. The percentiles and cut-off points used in these tables are defined in the simulation control parameters file.

3.4 Exposure Modeling Results

3.4.1 Overview

The results of the exposure and risk characterization are presented here for Philadelphia County. Several scenarios were considered for the exposure assessment, including two averaging time for NO_2 concentrations (annual and 1-hour), inclusion of indoor sources, and for evaluating just meeting the current standard. To date, year 2002 served as the base year for all scenarios, years 2001 and 2003 were only evaluated for a limited number of scenarios. Exposures were simulated for four groups; children and all persons, and the asthmatic population within each of these.

The exposure results summarized below focus on the population group where exposure estimations are of greatest interest, namely asthmatic individuals. Complete results for each of these two population subgroups is provided in Appendix G. However, due to certain limitations in the data summaries output from APEX, some exposure data could only be output for the entire population modeled (i.e., all persons - includes asthmatics and healthy persons of all ages). The summary data for the entire population (e.g., annual average exposure concentrations, time spent in microenvironments at or above a potential health effect benchmark level) can be representative of the asthmatic population since the asthmatic population does not have its microenvironmental concentrations and activities estimated any differently from those of the total population.

3.4.2 Annual Average Exposure Concentrations (as is)

Since the current NO_2 standard is 0.053 ppm annual average, the predicted air quality concentrations, the measured ambient monitoring concentrations, and the estimated exposures were summarized by annual average concentration. The distribution for the AERMOD predicted NO_2 concentrations at each of the 16,857 receptors for years 2001 through 2003 are illustrated in Figure 22. Variable concentrations were estimated by the dispersion model over the three year period (2001-2003). The NO_2 concentration distribution was similar for years 2001 and 2002, with mean annual average concentrations of about 21 ppb and a COV of just over 30%. On average, NO_2 annual average concentrations were lowest during simulated year 2003 (mean annual average concentration was about 16 ppb), largely a result of the comparably lower local concentration added (Table 26). While the mean annual average concentrations were lower than those estimated for 2001 and 2002, a greater number of annual average concentrations were estimated above 53 ppb for year 2003. In addition, year 2003 also contained greater variability in annual average concentrations as indicated by a COV of 53%.

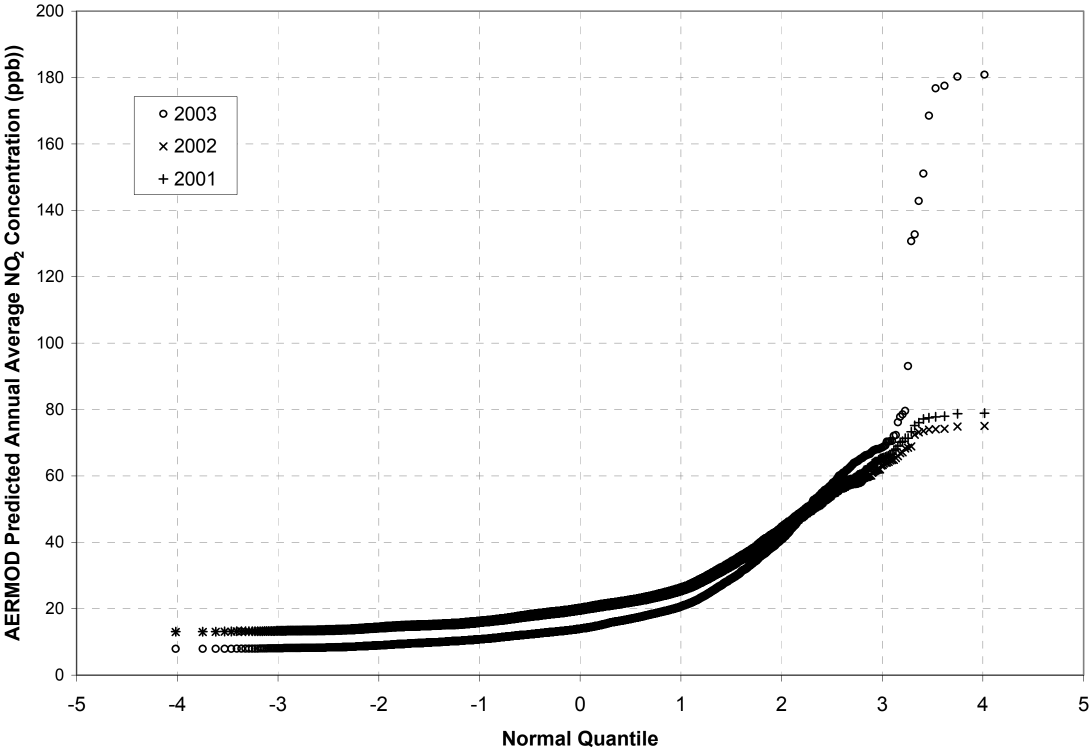


Figure 22 . Distribution of AERMOD predicted annual average NO_2 concentrations at each of the 16,857 receptors in Philadelphia County for years 2001-2003

The hourly concentrations output from AERMOD were input into the exposure model, providing a range of estimated exposures output by APEX. Figure 23 illustrates the annual average exposure concentrations for the entire simulated population (both asthmatics and healthy individual of all ages), for each of the years analyzed and where indoor sources were modeled. While years 2001 and 2002 contained very similar population exposure concentration distributions, the modeled year 2003 contained about 20% lower annual average concentrations. The lower exposure concentrations for year 2003 are similar to what was observed for the predicted air quality (Figure 22), however, all persons were estimated to contain exposures below an annual average concentration of 53 ppb, even considering indoor source concentration contributions. Again, while Figure 23 summarizes the entire population, the data are representative of what would be observed for the population of asthmatics or asthmatic children.

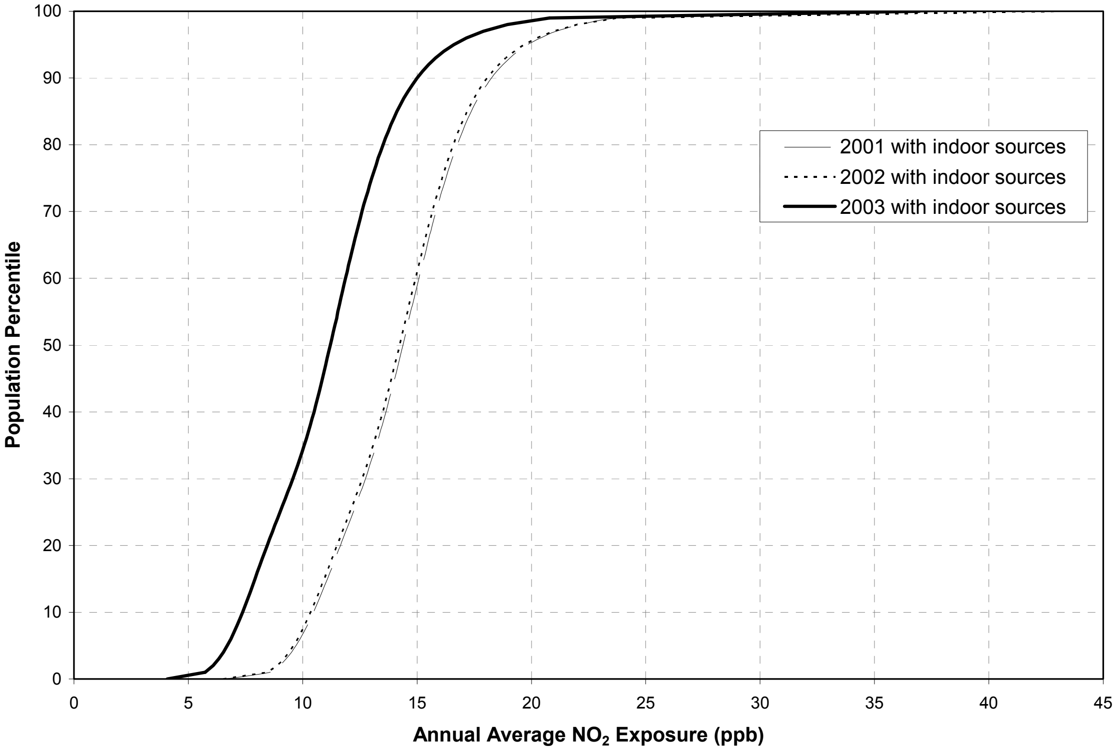


Figure 23. Estimated annual average total NO_2 exposure concentrations for all simulated persons in Philadelphia County, using modeled 2001-2003 air quality (as is), with modeled indoor sources.

The AERMOD predicted air quality and the estimated exposures for year 2002 were compared using their respective annual average NO_2 concentrations (Figure 24). As a point of reference, the annual average concentration for 2002 ambient monitors ranged from 24 ppb to 29 ppb. Many of the AERMOD predicted annual average concentrations were below that of the lowest ambient monitoring concentration of 24 ppb, although a few of the receptors contained concentrations above the highest measured annual average concentration . Estimated exposure concentrations were below that of both the modeled and measured air quality. For example, exposure concentrations were about 5 ppb less than the modeled air quality when the exposure estimation included indoor sources, and about 10 ppb less for when exposures were estimated without indoor sources. In comparing the estimated exposures with and without indoor sources, indoor sources were estimated to contribute between 1 and 5 ppb to the total annual average exposures.

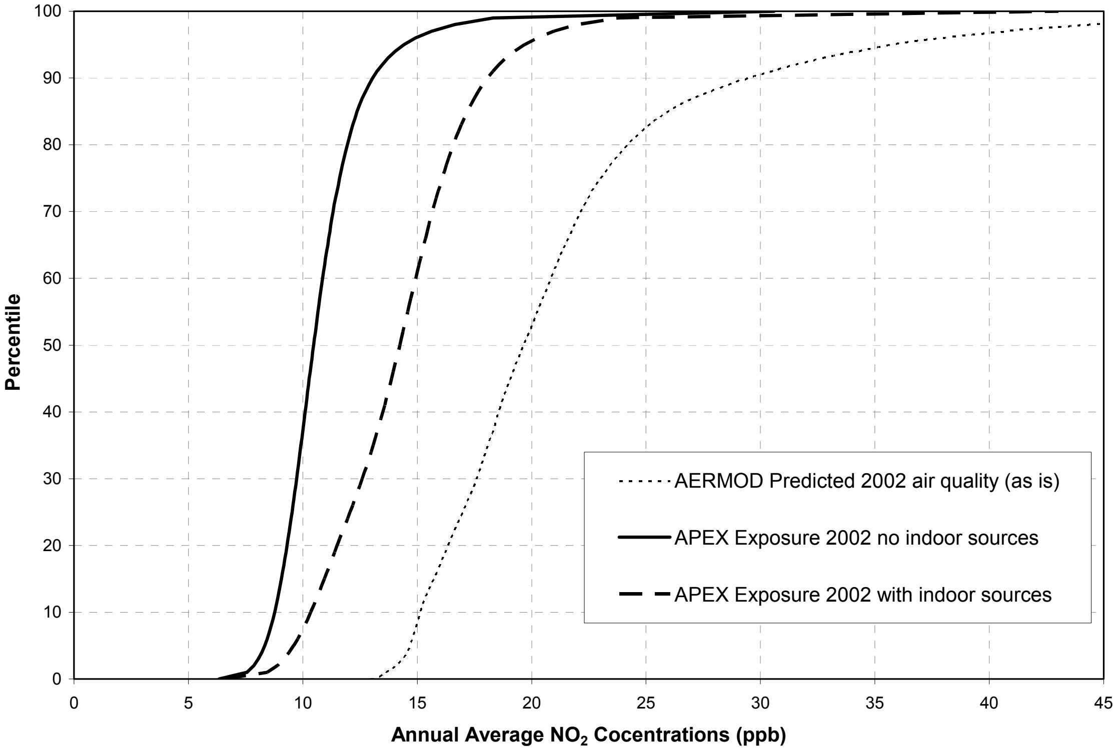


Figure 24. Comparison of AERMOD predicted and ambient monitoring annual average NO_2 concentrations (as is) and APEX exposure concentrations (with and without modeled indoor sources) in Philadelphia County for year 2002.

3.4.3 One-Hour Exposures (as is)

Since there is interest in short-term exposures, a few analyses were performed using the APEX estimated exposure concentrations. As part of the standard analysis, APEX reports the maximum exposure concentration for each simulated individual in the simulated population. This can provide insight into the proportion of the population experiencing any NO_2 exposure concentration level of interest. In addition, exposures are estimated for each of the selected potential health effect benchmark levels (200, 250, and 300 ppb, 1-hour average). An exceedance was recorded when the maximum exposure concentration observed for the individual was above the selected level in a day (therefore, the maximum number of exceedances is 365 for a single person). Estimates of repeated exposures are also recorded, that is where 1-hour exposure concentrations were above a selected level in a day added together across multiple days (therefore, the maximum number of multiple exceedances is also 365). Persons of interest in this exposure analysis are those with particular susceptibility to NO_2 exposure, namely individuals with asthma. The health effect benchmark levels are appropriate for estimating the potential risk of adverse health effects for asthmatics. The majority of the results presented in this section are for the simulated asthmatic population. However, the exposure analysis was performed for the total population to assess numbers of persons exposed to these levels and to provide additional information relevant to the asthmatic population (such as time spent in particular microenvironments), although most of the results for the total population are reported in Chapter 3 of the TSD.

3.4.3.1 Maximum Estimated Exposure Concentrations

A greater variability was observed in maximum exposure concentrations for the 2003 year simulation compared with years 2001 and 2002 (Figure 25). While annual average exposure concentrations for the total population were the lowest of the 3-year simulation, year 2003 contained a greater number of individual maximum exposures at and above the lowest potential health effect benchmark level. When indoor sources are not modeled however, over 90% of the simulated persons do not have an occurrence of a 1-hour exposure above 200 ppb in a year.

3.4.3.2 Number of Estimated Exposures above Selected Levels

When considering the total asthmatic population simulated in Philadelphia County and using current air quality of 2001-2003, nearly 50,000 persons were estimated to be exposed at least one time to a one-hour concentration of 200 ppb in a year (Figure 26). These exposures include both the NO_2 of ambient origin and that contributed by indoor sources. The number of asthmatics exposed to greater concentrations (e.g., 250 or 300 ppb) drops dramatically and is estimated to be somewhere between 1,000 – 15,000 depending on the 1-hour concentration level and the year of air quality data used. Exposures simulated for year 2003 contained the greatest number of asthmatics exposed in a year consistently for all potential health effect benchmark levels, while year 2002 contained the lowest number of asthmatics. Similar trends across the benchmark levels and the simulation years were observed for asthmatic children, albeit with lower numbers of asthmatic children with exposures at or above the potential health effect benchmark levels.

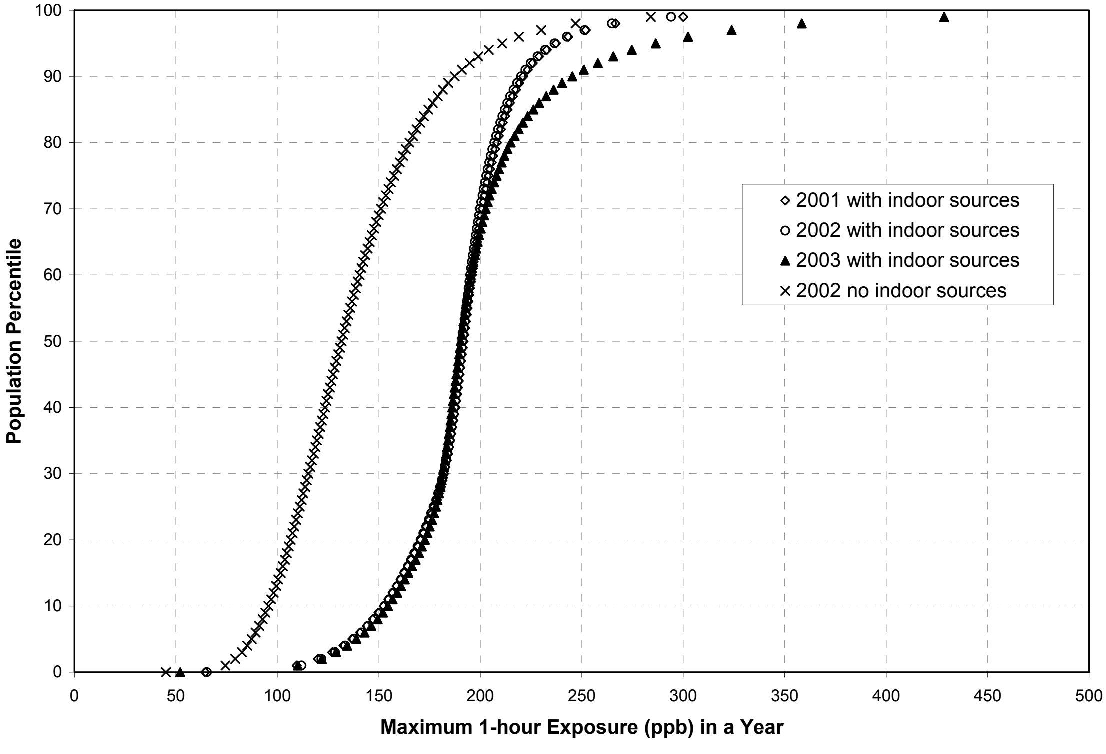


Figure 25. Estimated maximum NO_2 exposure concentration for all simulated persons in Philadelphia County, using modeled 2001-2003 air quality (as is), with and without modeled indoor sources. Values above the 99th percentile are not shown

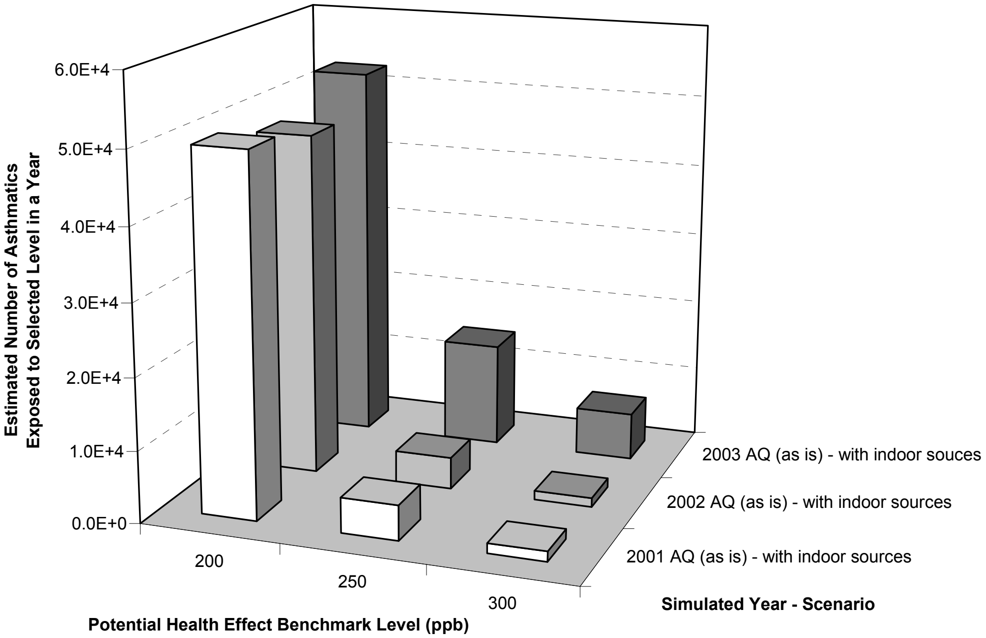


Figure 26. Estimated number of all simulated asthmatics in Philadelphia County with at least one NO_2 exposure at or above the potential health effect benchmark levels, using modeled 2001-2003 air quality (as is), with modeled indoor sources.

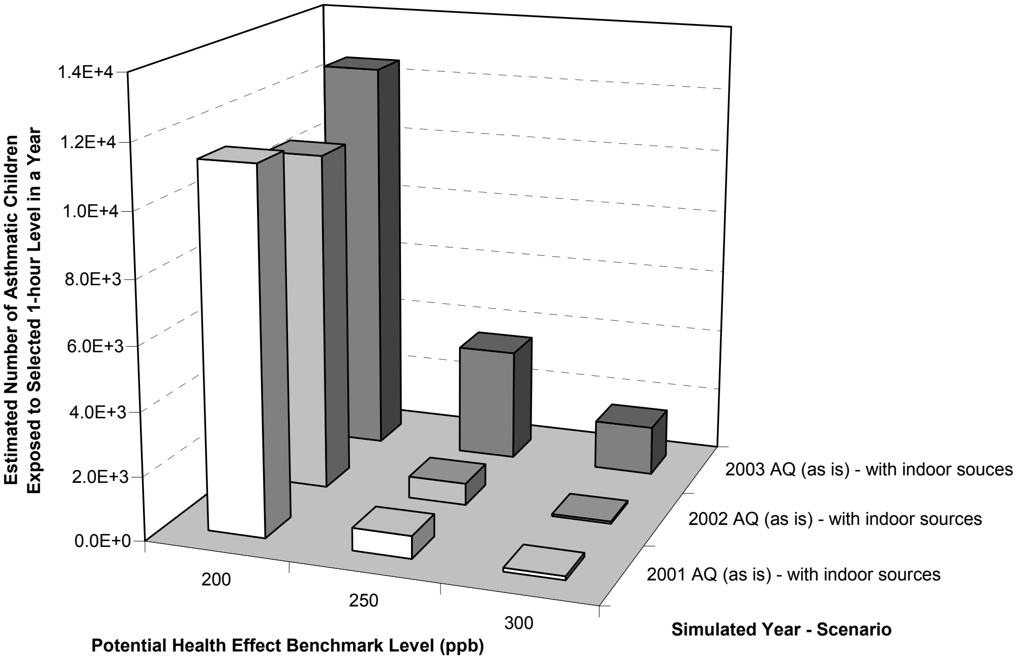


Figure 27. Estimated number of simulated asthmatic children in Philadelphia County with at least one NO_2 exposure at or above the potential health effect benchmark levels, using modeled 2001-2003 air quality (as is), with modeled indoor sources.

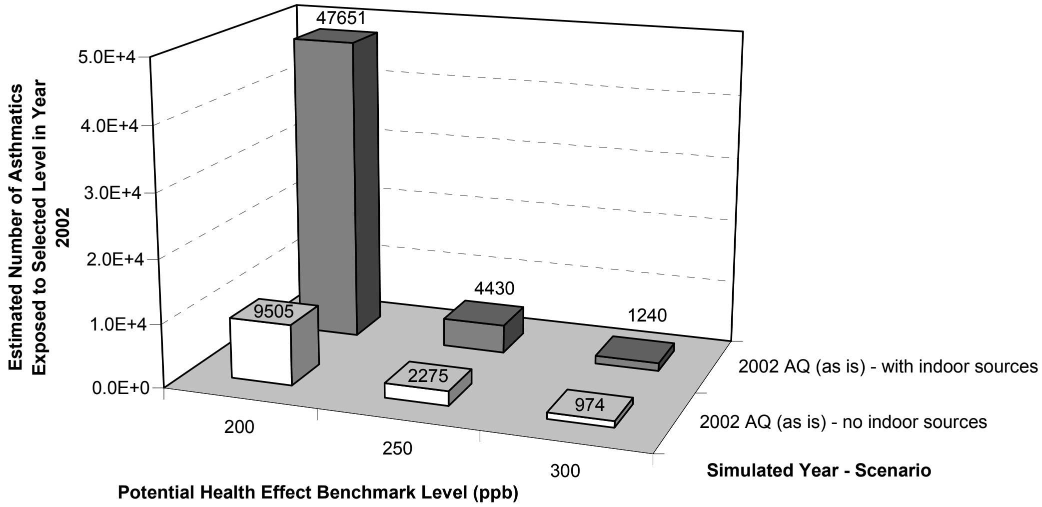


Figure 28. Comparison of the estimated number of all simulated asthmatics in Philadelphia County with at least one NO_2 exposure at or above potential health effect benchmark levels, using modeled 2002 air quality (as is) , with and without modeled indoor sources.

For example, nearly 12,000 were estimated to be exposed to at least a one-hour NO_2 concentration of 200 ppb in a year (Figure 27). Additional exposure estimates were generated using the modeled 2002 air quality (as is) and where the contribution from indoor sources was not included in the exposure concentrations. APEX allows for the same persons to be simulated, i.e., demographics of the population were conserved, as well as using the same individual time-location-activity profiles generated for each person. Figure 28 compares the estimated number of asthmatics experiencing exposures above the potential health effect benchmarks, both with indoor sources and without indoor sources included in the model runs. The number of asthmatics at or above the selected concentrations is reduced by between 50-80%, depending on benchmark level, when not including indoor source (i.e., gas cooking) concentration contributions.

An evaluation of the time spent in the 12 microenvironments was performed to estimate where simulated individuals are exposed to concentrations above the potential health effect benchmark levels. Currently, the output generated by APEX is limited to compiling the microenvironmental time for the total population (includes both asthmatic individuals and healthy persons) and is summarized to the total time spent above the selected potential health effect benchmark levels. As mentioned above, the data still provide a reasonable approximation for each of the population subgroups (e.g., asthmatics or asthmatic children) since their microenvironmental concentrations and activities are not estimated any differently from those of the total population by APEX.

As an example, Figure 29 (a, b, c) summarizes the percent of total time spent in each microenvironment for simulation year 2002 that was associated with estimated exposure concentrations at or above 200, 250, and 300 ppb (results for years 2001 and 2003 were similar). Estimated exposures included the contribution from one major category of indoor sources (i.e., gas cooking). The time spent in the indoor residence and bars/restaurants were the most important for concentrations ≥200 ppb, contributing to approximately 75% of the time persons were exposed (Figure 29a). This is likely a result of the indoor source concentration contribution to each individual's exposure concentrations. The importance of the particular microenvironment however changes with differing potential health effect benchmark levels. This is evident when considering the in-vehicle and outdoor near-road microenvironments, progressing from about 19% of the time exposures were at the lowest potential health effect benchmark level (200 ppb) to a high of 64% of the time exposures were at the highest benchmark level (300 ppb, Figure 29c).

The microenvironments where higher exposure concentrations occur were also evaluated for the exposure estimates generated without indoor source contributions. Figure 30 illustrates that the time spent in the indoor microenvironments contributes little to the estimated exposures above the selected benchmark levels. The contribution of these microenvironments varied only slightly with increasing benchmark concentration, ranging from about 2-5%. Most of the time associated with high exposures was associated with the transportation microenvironments (In-Vehicle or In-Public Transport) or outdoors (Out-Near Road, Out-Parking Lot, Out-Other). The importance of time spent outdoors near roadways exhibited the greatest change in contribution with increased health benchmark level, increasing from around 30 to 44% of time associated with concentrations of 200 and 300 ppb, respectively

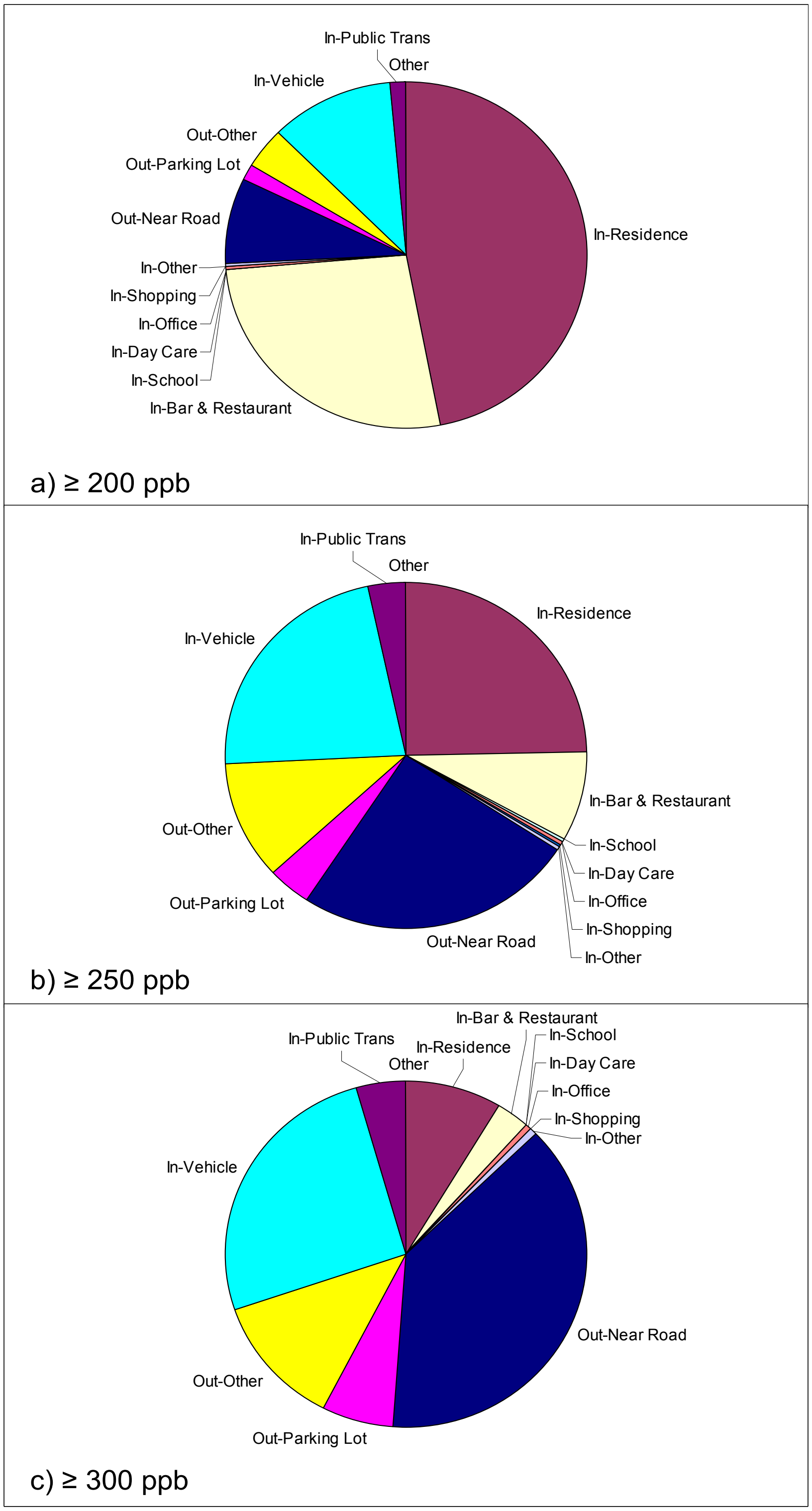


Figure 29. Fraction of time all simulated persons in Philadelphia County spend in the twelve microenvironments associated with the potential NO2 health effect benchmark levels, a) ≥ 200 ppb, b) ≥ 250 ppb, and c) ≥ 300 ppb, year 2002 simulation with indoor sources.

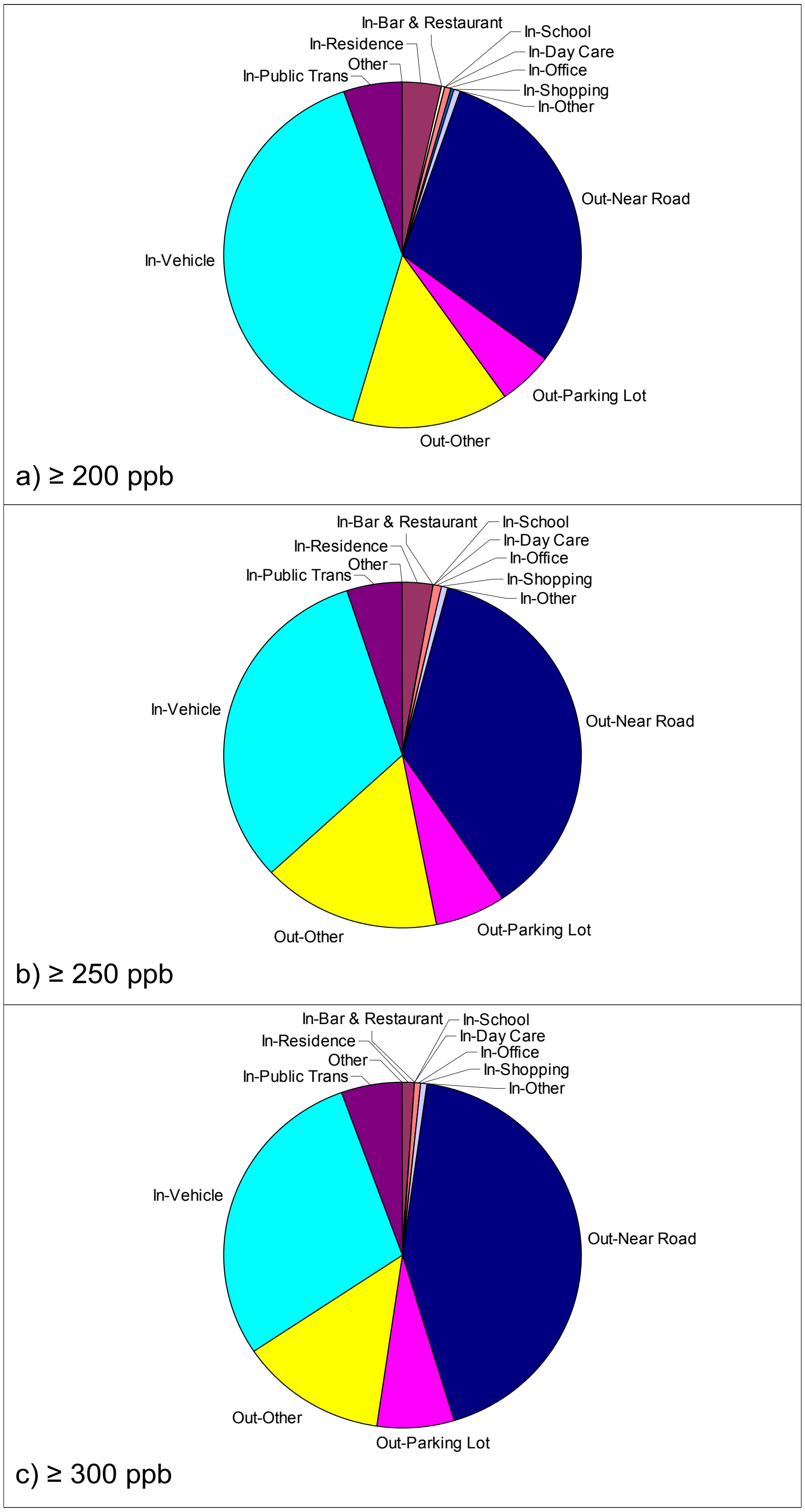


Figure 30. Fraction of time all simulated persons in Philadelphia County spend in the twelve microenvironments associated with the potential NO_2 health effect benchmark levels, a) ≥ 200 ppb, b) ≥ 250 ppb, and c) ≥ 300 ppb, year 2002 simulation without indoor sources

3.4.3.3 Number of Repeated Exposures Above Selected Levels

In the analysis of persons exposed, the results show the number or percent of those with at least one exposure at or above the selected potential health effect benchmark level. Given that the benchmark is for a small averaging time (i.e., one-hour) it may be possible that individuals are exposed to concentrations at or above the potential health effect benchmark levels more than once in a given year. Since APEX simulates the longitudinal diary profile for each individual, the number of times above a selected level is retained for each person. Figure 31 presents such an analysis for the year 2003, the year containing the greatest number of exposure concentrations at or above the selected benchmarks. Estimated exposures include both those resulting from exposures to NO_2 of ambient origin and those resulting from indoor source NO_2 contributions. While a large fraction of individuals experience at least one exposure to 200 ppb or greater over a 1-hour time period in a year (about 32 percent), only around 14 percent were estimated to contain at least 2 exposures. Multiple exposures at or above the selected benchmarks greater than or equal to 3 or more times per year are even less frequent, with around 5 percent or less of asthmatics exposed to 1-hour concentrations greater than or equal to 200 ppb 3 or more times in a year.

Exposure estimates for year 2002 are presented to provide an additional perspective, including a lower bound of repeated exposures for this population subgroup and for exposure estimates generated with and without modeled indoor sources (Figure 32). Most asthmatics exposed to a 200 ppb concentration are exposed once per year and only around 11 percent would experience 2 or more exposures at or above 200 ppb when including indoor source contributions. The percent of asthmatics experiencing multiple exposures a and abovet 250 and 300 ppb is much lower, typically less than 1 percent of all asthmatics are exposed at the higher potential benchmark levels. Also provided in Figure 32 are the percent of asthmatics exposed to selected levels in the absence of indoor sources. Again, without the indoor source contribution, there are reduced occurrences of multiple exposures at all of the potential health effect benchmark levels compared with when indoor sources were modeled.

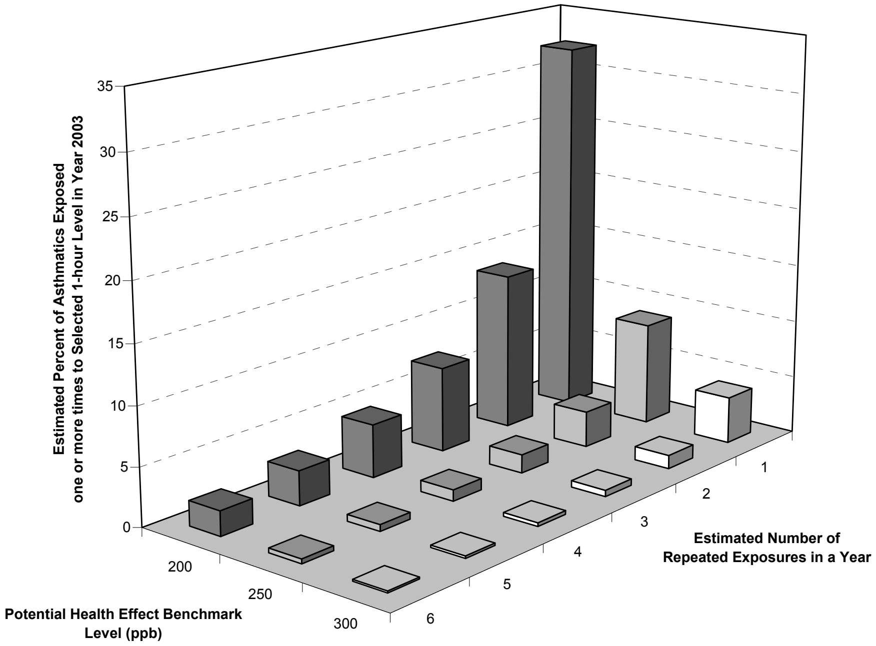


Figure 31. Estimated percent of all asthmatics in Philadelphia County with repeated NO_2 exposures above potential health effect benchmark levels, using 2003 modeled air quality (as is), with modeled indoor sources.

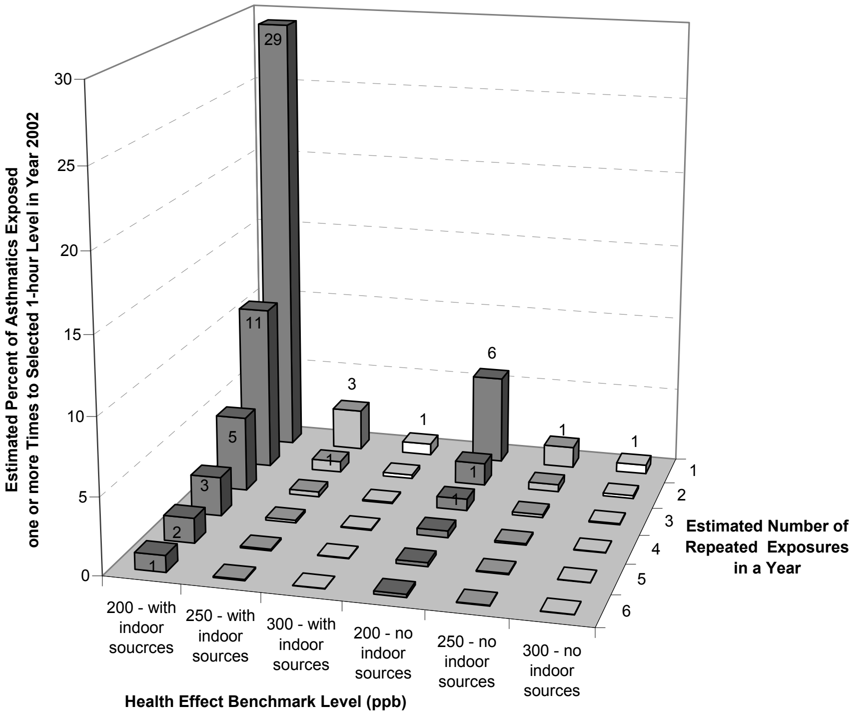


Figure 32. Estimated percent of all asthmatics in Philadelphia County with repeated NO_2 exposures above potential health effect benchmark levels, using modeled 2002 air quality (as is), with and without indoor sources.

3.4.4 One-Hour Exposures Associated with Just Meeting the Current Standard

To simulate just meeting the current NO_2 standard, the potential health effect benchmark level was adjusted in the exposure model, rather than adjusting all of the hourly concentrations for each receptor and year simulated. Similar estimates of short-term exposures (i.e., 1-hour) were generated for the total population and population subgroups of interest (i.e., asthmatics and asthmatic children).

3.4.4.1 Number of Estimated Exposures above Selected Levels

In considering exposures estimated to occur associated with air quality simulated to just meet the current annual average NO_2 standard, the number of persons experiencing concentrations at or above the potential health effect benchmarks increased. To allow for reasonable comparison, the number of persons affected considering each scenario is expressed as the percent of the subpopulation of interest. Figure 33 illustrates the percent of asthmatics estimated to experience at least one exposure at or above the selected potential health effect benchmark concentrations, with just meeting the current standard and including indoor source contributions. While it was estimated that about 30% percent of asthmatics would be exposed to 200 ppb (1-hour average) at least once in a year for as is air quality, it was estimated that around 80 percent of asthmatics would experience at least one concentration above the lowest potential health effect benchmark level in a year representing just meeting the current standard. Again, estimates for asthmatic children exhibited a similar trend, with between 75 to 80 percent exposed to a concentration at or above the lowest potential health effect benchmark level at least once per year for a year just meeting the current standard (data not shown). The percent of all asthmatics experiencing the higher benchmark levels is reduced to between 31 and 45 percent for the 250 ppb, 1-hour benchmark, and between 10 and 24 percent for the 300 ppb, 1-hour benchmark level associated with air quality representing just meeting the current annual average standard.

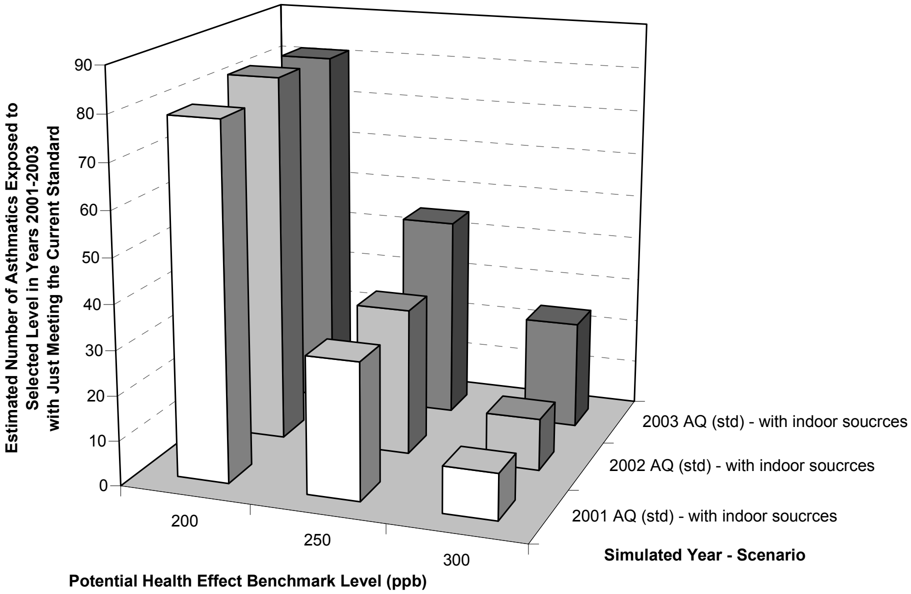


Figure 33. Estimated percent of all asthmatics in Philadelphia with at least one exposure at or above the potential health effect benchmark level, using modeled 2001-2003 air quality just meeting the current standard, with modeled indoor sources

In evaluating the influence of indoor source contribution for the scenario just meeting the current standard, the numbers of individuals exposed at selected levels are reduced without indoor sources, ranging from about 26 percent lower for the 200 ppb level to around 11 percent for the 300 ppb level when compared with exposure estimates that accounted for indoor sources (Figure 34).

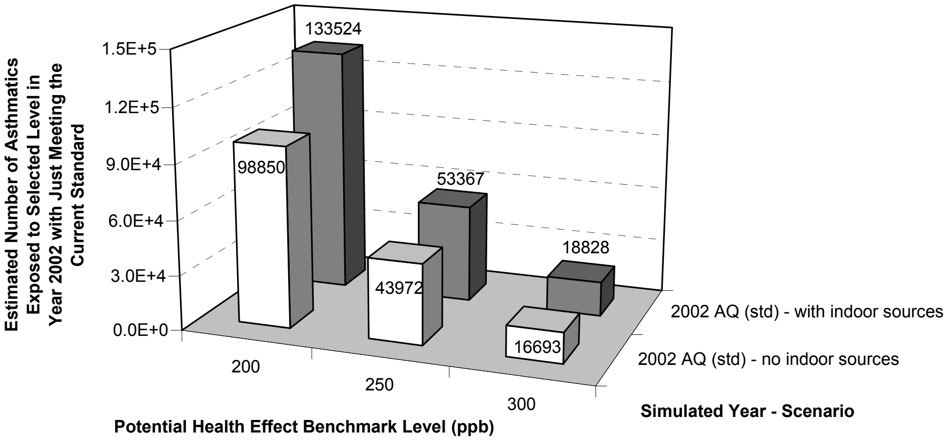


Figure 34. Estimated number of all asthmatics in Philadelphia with at least one exposure at or above the potential health effect benchmark level, using modeled 2002 air quality just meeting the current standard, with and without modeled indoor sources.

3.4.4.2 Number of Repeated Exposures Above Selected Levels

For air quality simulated to just meet the current standard, repeated exposures at the selected potential health effect benchmarks are more frequent than that estimated for the modeled as is air quality. Figure 35 illustrates this using the simulated asthmatic population for year 2002 data as an example. Many asthmatics that are exposed at or above the selected levels are exposed more than one time. Repeated exposures above the potential health effect benchmark levels are reduced however, when not including the contribution from indoor sources. The percent of asthmatics exposed drops with increasing benchmark level, with progressively fewer persons experiencing multiple exposures for each benchmark level.

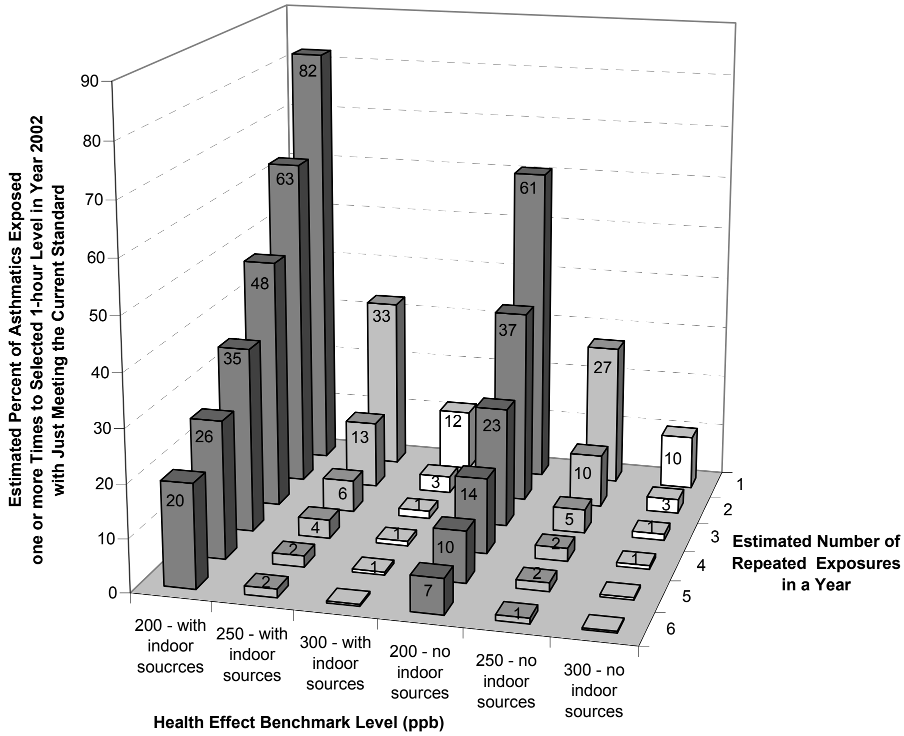


Figure 35. Estimated percent of asthmatics in Philadelphia County with repeated exposures above health effect benchmark levels, using modeled 2002 air quality just meeting the current standard, with and without modeled indoor sources

3.5 Variability and Uncertainty

3.5.1 Introduction

The methods and the model used in this assessment conform to the most contemporary modeling methodologies available. APEX is a powerful and flexible model that allows for the realistic estimation of air pollutant exposure to individuals. Since it is based on human activity diaries and accounts for the most important variables known to affect exposure, it has the ability to effectively approximate actual conditions. In addition, the input data selected were the best available data to generate the exposure results. However, there are constraints and uncertainties with the modeling approach and the input data that limit the realism and accuracy of the model results.

All models have limitations that require the use of assumptions. Limitations of APEX lie primarily in the uncertainties associated with data distributions input to the model. Broad uncertainties and assumptions associated with these model inputs, utilization, and application include the following, with more detailed analysis summarized below and presented previously (see US EPA, 2007d; Langstaff, 2007).

- The CHAD activity data used in APEX are compiled from a number of studies in different areas, and for different seasons and years. Therefore, the combined data set may not constitute a representative sample for a particular study scenario.

- Commuting pattern data were derived from the 2000 U.S. Census. The commuting data address only home-to-work travel. The population not employed outside the home is assumed to always remain in the residential census tract. Furthermore, although several of the APEX microenvironments account for time spent in travel, the travel is assumed to always occur in basically a composite of the home and work block. No other provision is made for the possibility of passing through other blocks during travel.
- APEX creates seasonal or annual sequences of daily activities for a simulated individual by sampling human activity data from more than one subject. Each simulated person essentially becomes a composite of several actual people in the underlying activity data.
- The model currently does not capture certain correlations among human activities that can impact microenvironmental concentrations (for example, cigarette smoking leading to an individual opening a window, which in turn affects the amount of outdoor air penetrating the microenvironment).
- Certain aspects of the personal profiles are held constant, though in reality they change as individuals age. This is only important for simulations with long timeframes, particularly when simulating young children (e.g., over a year or more).

3.5.2 Input Data Evaluation

Modeling results are heavily dependent on the quality of the data that are input to the system. As described above, several studies were reviewed, and data from these studies were used to develop the parameters and factors that were used to build the microenvironments in this assessment. A constraint on this effort is that there are a limited number of NO_2 exposure studies to use for evaluation.

The input data used in this assessment were selected to best simulate actual conditions that affect human exposure. Using well characterized data as inputs to the model lessens the degree of uncertainty in exposure estimates. Still, the limitations and uncertainties of each of the data streams affect the overall quality of the model output. These issues and how they specifically affect each data stream are discussed in this section.

3.5.2.1 Meteorological Data

Meteorological data are taken directly from monitoring stations in the assessment areas. One strength of these data is that it is relatively easy to see significant errors if they appear in the data. Because general climactic conditions are known for each area simulation, it would have been apparent upon review if there were outliers in the dataset. However, there are limitations in the use of these data. Because APEX only uses one temperature value per day, the model does not represent hour-to-hour variations in meteorological conditions throughout the day that may affect both NO_2 formation and exposure estimates within microenvironments.

3.5.2.2 Air Quality Data

Air quality data used in the exposure modeling was determined through use of EPA's recommended regulatory air dispersion model, AERMOD (version 07026 (US EPA, 2004)), with meteorological data discussed above and emissions data based on the EPA's National Emissions Inventory for 2002 (US EPA, 2007b) and the CAMD Emissions Database (US EPA, 2007c) for stationary sources and mobile sources determined from local travel demand modeling and EPA's MOBILE6.2 emission factor model. All of these are high quality data sources. Parameterization of meteorology and emissions in the model were made in as accurate a manner as possible to ensure best representation of air quality for exposure modeling. Further, minor sources not included in the dispersion modeling were captured and any remaining long-term errors in the results corrected through use of local concentrations derived from monitor observations. Thus, the resulting air quality values are free of systematic errors to the best approximation available through application of modeled data.

3.5.2.3 Population and Commuting Data

The population and commuting data are drawn from U.S. Census data from the year 2000. This is a high quality data source for nationwide population data in the U.S. However, the data do have limitations. The Census used random sampling techniques instead of attempting to reach all households in the U.S., as it has in the past. While the sampling techniques are well established and trusted, they introduce some uncertainty to the system. The Census has a quality section (http://www.census.gov/quality/) that discusses these and other issues with Census data.

In addition to these data quality issues, certain simplifying assumptions were made in order to better match reality or to make the data match APEX input specifications. For example, the APEX dataset does not differentiate people that work at home from those that commute within their home tract, and individuals that commute over 120 km a day were assumed to not commute daily. In addition to emphasizing some of the limitations of the input data, these assumptions introduce uncertainty to the results.

Furthermore, the estimation of block-to-block commuter flows relied on the assumption that the frequency of commuting to a workplace block within a tract is proportional to the amount of commercial and industrial land in the block. This assumption introduces additional uncertainty.

3.5.2.4 Activity Pattern Data

It is probable that the CHAD data used in the system is the most subject to limitations and uncertainty of all the data used in the system. Much of the data used to generate the daily diaries are over 20 years old. Table 46 indicates the ages of the CHAD diaries used in this modeling analysis. While the specifics of people's daily activities may not have changed much over the years, it is certainly possible that some differences do exist. In addition, the CHAD data are taken from numerous surveys that were performed for different purposes. Some of these surveys collected only a single diary-day while others went on for several days. Some of the studies were designed to not be representative of the U.S. population, although a large portion of the data are from National surveys. Furthermore, study collection periods occur at different times of the year, possibly resulting in seasonal differences. A few of these limitations are corrected by the approaches used in the exposure modeling (e.g., weighting by US population demographics for a particular location, adjusting for effects of temperature on human activities).

A sensitivity analysis was performed to evaluate the impact of the activity pattern database on APEX model results for O_3 (see Langstaff (2006) and US EPA (2007d)). Briefly, exposure results were generated using APEX with all of the CHAD diaries and compared with results generated from running APEX using only the CHAD diaries from the National Human Activity Pattern Study (NHAPS), a nationally representative study in CHAD. There was very good agreement between the APEX results for the 12 cities evaluated, whether all of CHAD or only the NHAPS component of CHAD is used. The absolute difference in percent of persons above a particular concentration level ranged from -1% to about 4%, indicating that the exposure model results are not being overly influenced by any single study in CHAD. It is likely that similar results would be obtained here for NO_2 exposures, although remains uncertain due to different averaging times (1-hour vs. 8-hour average).

3.5.2.5 Air Exchange Rates

There are several components of uncertainty in the residential air exchange rate distributions used for this analysis. US EPA (2007d) details an analysis of uncertainty due to extrapolation of air exchange rate distributions between-CMSAs and within-CMSA uncertainty due to sampling variation. In addition, the uncertainty associated with estimating daily air exchange rate distributions from air exchange rate measurements with varying averaging times is discussed. The results of those investigations are briefly summarized here.

Extrapolation among cities

Location-specific distributions were assigned in the APEX model, as detailed in the indoors-residential microenvironment. Since specific data for all of the locations targeted in this analysis were not available, data from another location were used based on similar influential characteristics. Such factors include age composition of housing stock, construction methods, and other meteorological variables not explicitly treated in the analysis, such as humidity and wind speed patterns. In order to assess the uncertainty associated with this extrapolation, between-CSA uncertainty was evaluated by examining the variation of the geometric means and standard deviations across cities and studies.

The analysis showed a relatively wide variation across different cities in the air exchange rate geometric mean and standard deviation, stratified by air-conditioning status and temperature range. This implies that the air exchange rate modeling results would be very different if the matching of modeled locations to study locations was changed. For example, the NO_2 exposure estimates may be sensitive to the assumption that the Philadelphia air exchange rate distributions can be represented by the New York City air exchange rate data.

Within CSA uncertainty

There is also variation within studies for the same location (e.g., Los Angeles), but this is much smaller than the variation across CMSAs. This finding tends to support the approach of combining different studies for a CMSA. In addition, within-city uncertainty was assessed by using a bootstrap distribution to estimate the effects of sampling variation on the fitted geometric means and standard deviations for each CMSA. The bootstrap distributions assess the uncertainty due to random sampling variation but do not address uncertainties due to the lack of representativeness of the available study data or the variation in the lengths of the AER monitoring periods.

1,000 bootstrap samples were randomly generated for each AER subset (of size N), producing a set of 1,000 geometric mean and geometric standard deviation pairs. The analysis indicated that the geometric standard deviation uncertainty for a given CSA/air-conditioning-status/temperature-range combination tended to have a range of at most from *fitted GSD-1.0* hr^{-1} to *fitted GSD+1.0* hr^{-1}, but the intervals based on larger AER sample sizes were frequently much narrower. The ranges for the geometric means tended to be approximately from *fitted GM-0.5*

hr^{-1} to *fitted GM+0.5* hr^{-1}, but in some cases were much smaller. Figure 36 illustrates such results for Los Angeles as an example.

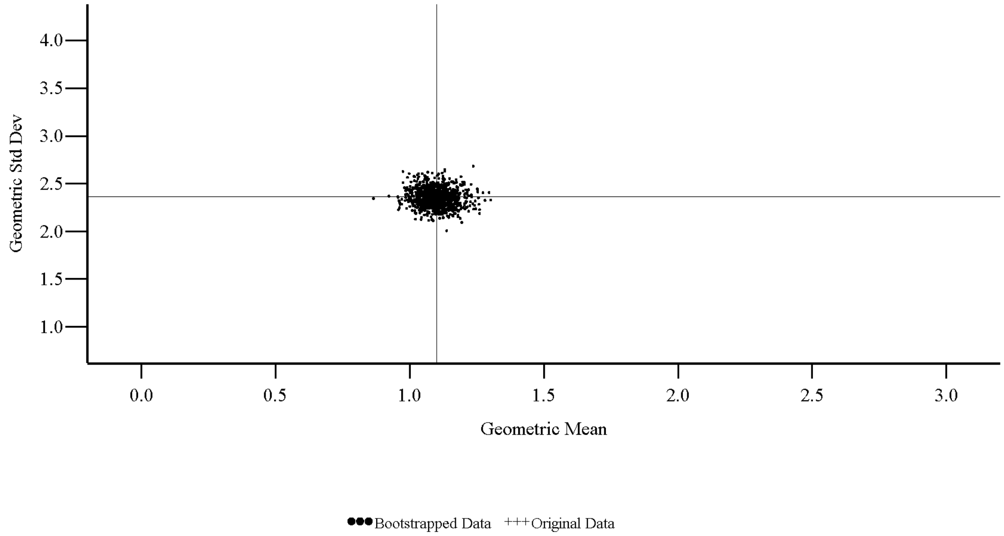


Figure 36. Geometric mean and standard deviation of air exchange rate bootstrapped for Los Angeles residences with A/C, temperature range from 20-25 degrees centigrade (from US EPA, 2007d).

Variation in measurement averaging times

Although the averaging periods for the air exchange rates in the study data varied from one day to seven days, the analyses did not take the measurement duration into account and treated the data as if they were a set of statistically independent daily averages. To investigate the uncertainty of this assumption, correlations between consecutive 24-hour air exchange rates measured at the same house were investigated using data from the Research Triangle Park Panel Study (US EPA, 2007d). The results showed extremely strong correlations, providing support for the simplified approach of treating multi-day averaging periods as if they were 24-hour averages.

3.5.2.6 Air Conditioning Prevalence

Because the selection of an air exchange rate distribution is conditioned on the presence or absence of an air-conditioner, for each modeled area, the air conditioning status of the residential microenvironments was simulated randomly using the probability that a residence has an air conditioner, i.e., the residential air conditioner prevalence rate. For this study we used location-specific data from the American Housing Survey of 2003. US EPA (2007d) details the specification of uncertainty estimates in the form of confidence intervals for the air conditioner prevalence rate, and compares these with prevalence rates and confidence intervals developed from the Energy Information Administration's Residential Energy Consumption Survey (RECS) of 2001 for more aggregate geographic subdivision (e.g., states, multi-state Census divisions and regions).

Air conditioning prevalence rates for the 5 locations from the American Housing Survey (Table 50) ranged from 55% for Los Angeles to 97% for Atlanta. Reported standard errors were relatively small, ranging from less than 1.2% for Atlanta to 1.8% for Detroit. The corresponding 95% confidence intervals are also small and range from approximately 4.6% to 6.9%. The

RECS prevalence estimates and confidence intervals compared with the similar locations using AHS data were mixed. Good agreements between the AHS and RECS confidence intervals was found for Atlanta and Detroit. Poor agreement with the AHS for either the Census Region or Census Division estimates was shown for Los Angeles and Philadelphia, with estimates of those owning A/C lower when considering the RECS data. However, since the AHS survey results are city-specific and were based on a more recent survey, the AHS prevalence estimates were used for the APEX modeling.

Furthermore, some residences use evaporative coolers, also known as “swamp coolers,” for cooling. The estimation of air exchange rate distributions from measurement data used here did not take into account the presence or absence of an evaporative cooler. Based on statistical comparison tests (i.e., F-test, Kruskal-Wallis, Mood) for where information was available to generate AER distributions with and without swamp cooler ownership, it was determined that presence or absence of such data did not alter the statistical air exchange model (US EPA, 2007d).

3.6 References

AHS. (2003a). American Housing Survey for 2003. Available at: http://www.census.gov/hhes/www/housing/ahs/ahs.html.

AHS. (2003b). Source and Accuracy Statement for the 2003 AHS-N Data Chart. Available at: http://www.census.gov/hhes/www/housing/ahs/03dtchrt/source.html.

Akland GG, Hartwell TD, Johnson TR, Whitmore RW. (1985). Measuring human exposure to carbon monoxide in Washington, D. C. and Denver, Colorado during the winter of 1982-83. *Environ Sci Technol*. 19:911-918.

Avol EL, Navidi WC, Colome SD. (1998) Modeling ozone levels in and around southern California homes. *Environ Sci Technol*. 32:463-468.

Biller WF, Feagans TB, Johnson TR, Duggan GM, Paul RA, McCurdy T, Thomas HC. (1981). A general model for estimating exposure associated with alternative NAAQS. Paper No. 81-18.4 in Proceedings of the 74th Annual Meeting of the Air Pollution Control Association, Philadelphia, PA.

CARB. (2001). Indoor air quality: residential cooking exposures. Final report. California Air Resources Board, Sacramento, California. Available at: http://www.arb.ca.gov/research/indoor/cooking/cooking.htm.

CDC. (2007). National Center for Health Statistics. National Health Interview Survey (NHIS) Public Use Data Release (2003). Available at: http://www.cdc.gov/nchs/about/major/nhis/quest_data_related_1997_forward.htm.

Chan AT and Chung MW. (2003). Indoor-outdoor air quality relationships in vehicle: effect of driving environment and ventilation modes. *Atmos Environ*. 37:3795-3808.

Chilrud SN, Epstein D, Ross JM, Sax SN, Pederson D, Spengler JD, Kinney PL. (2004). Elevated airborne exposures of teenagers to manganese, chromium, and iron from steel dust and New York City's subway system. *Environ Sci Technol*. 38:732-737.

Colome SD, Wilson AL, Tian Y. (1993). California Residential Indoor Air Quality Study, Volume 1, Methodology and Descriptive Statistics. Prepared for the Gas Research Institute, Pacific Gas & Electric Co., San Diego Gas & Electric Co., Southern California Gas Co.

Colome SD, Wilson AL, Tian Y. (1994). California Residential Indoor Air Quality Study, Volume 2, Carbon Monoxide and Air Exchange Rate: An Univariate and Multivariate Analysis. Chicago, IL. Prepared for the Gas Research Institute, Pacific Gas & Electric Co., San Diego Gas & Electric Co., Southern California Gas Co. GRI-93/0224.3

Finlayson-Pitts BJ and Pitts JN. (2000). Chemistry of the Upper and Lower Atmosphere. Academic Press, San Diego CA. Page 17.

Hartwell TD, Clayton CA, Ritchie RM, Whitmore RW, Zelon HS, Jones SM, Whitehurst DA. (1984). Study of Carbon Monoxide Exposure of Residents of Washington, DC and Denver, Colorado. Research Triangle Park, NC: U.S. Environmental Protection Agency, Office of Research and Development, Environmental Monitoring Systems Laboratory. EPA-600/4-84-031.

Johnson TR and Paul RA. (1983). The NAAQS Exposure Model (NEM) Applied to Carbon Monoxide. EPA-450/5-83-003. Prepared for the U.S. Environmental Agency by PEDCo Environmental Inc., Durham, N.C. under Contract No. 68-02-3390. U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.

Johnson T. (1984). A Study of Personal Exposure to Carbon Monoxide in Denver, Colorado. Research Triangle Park, NC: U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory. EPA-600/4-84-014.

Johnson T. (1989). Human Activity Patterns in Cincinnati, Ohio. Palo Alto, CA: Electric Power Research Institute. EPRI EN-6204.

Johnson T, Capel J, Olaguer E, Wijnberg L. (1992). Estimation of Ozone Exposures Experienced by Residents of ROMNET Domain Using a Probabilistic Version of NEM. Prepared by IT Air Quality Services for the Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.

Johnson T, Capel J, McCoy M. (1996a). Estimation of Ozone Exposures Experienced by Urban Residents Using a Probabilistic Version of NEM and 1990 Population Data. Prepared by IT Air Quality Services for the Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, September.

Johnson T, Capel J, Mozier J, McCoy M. (1996b). Estimation of Ozone Exposures Experienced by Outdoor Children in Nine Urban Areas Using a Probabilistic Version of NEM. Prepared for the Air Quality Management Division under Contract No. 68-DO-30094, April.

Johnson T, Capel J, McCoy M, Mozier J. (1996c). Estimation of Ozone Exposures Experienced by Outdoor Workers in Nine Urban Areas Using a Probabilistic Version of NEM. Prepared for the Air Quality Management Division under Contract No. 68-DO-30094, April.

Johnson T, Mihlan G, LaPointe J, Fletcher K. (1999). Estimation Of Carbon Monoxide Exposures and Associated Carboxyhemoglobin Levels In Denver Residents Using pNEM/CO (version 2.0). Prepared for the U.S. Environmental Protection Agency under Contract No. 68-D6-0064, March 1999.

Kinney PL, Chillrud SN, Ramstrom S, Ross J, Spengler JD. (2002). Exposures to multiple air toxics in New York City. *Environ Health Perspect*. 110:539-546.

Klepeis NE, Tsang AM, Behar JV. (1996). Analysis of the National Human Activity Pattern Survey (NHAPS) Respondents from a Standpoint of Exposure Assessment. Washington, DC: U.S. Environmental Protection Agency, Office of Research and Development. EPA/600/R-96/074.

Koontz, M. D., L. L. Mehegan, and N. L. Nagda. 1992. Distribution and Use of Cooking Appliances That Can Affect Indoor Air Quality, Report No. GRI-93/0013. Gas Research Institute, Chicago.

Langstaff JE. (2007). OAQPS Staff Memorandum to Ozone NAAQS Review Docket (OAR-2005-0172). Subject: Analysis of Uncertainty in Ozone Population Exposure Modeling. [January 31, 2007]. Available at: http://www.epa.gov/ttn/naaqs/standards/ozone/s_o3_cr_td.html.

McCurdy T, Glen G, Smith L, Lakkadi Y. (2000). The National Exposure Research Laboratory's Consolidated Human Activity Database, *J Exp Anal Environ Epidemiol*. 10: 566-578.

Meng QY, Turpin BJ, Korn L, Weisel CP, Morandi M, Colome S, Zhang JJ, Stock T, Spektor D, Winer A, Zhang L, Lee JH, Giovanetti R, Cui W, Kwon J, Alimokhtari S, Shendell D, Jones J, Farrar C, Maberti S. (2004). Influence of ambient (outdoor) sources on residential indoor and personal PM2.5 concentrations: Analyses of RIOPA data. *J Expos Anal Environ Epidemiol*. 15:17-28.

Murray DM and Burmaster DE. (1995). Residential air exchange rates in the United States: empirical and estimated parametric distributions by season and climatic region. *Risk Analysis*. 15(4):459-465.

PA DOH. (2008). Behavioral Risk Factor Surveillance System. Pennsylvania Department of Health, Bureau of Health Statistics and Research. Available at: http://www.dsf.health.state.pa.us/health/cwp/view.asp?a=175&Q=242623.

Persily A and Gorfain J. (2004). Analysis of ventilation data from the U.S. Environmental Protection Agency Building Assessment Survey and Evaluation (BASE) Study. National Institute of Standards and Technology, NISTIR 7145, December 2004.

Persily A, Gorfain J, Brunner G. (2005). Ventilation design and performance in U.S. office buildings. ASHRAE Journal. April 2005, 30-35.

Robinson JP, Wiley JA, Piazza T, Garrett K, Cirksena K. (1989). Activity Patterns of California Residents and their Implications for Potential Exposure to Pollution. California Air Resources Board, Sacramento, CA. CARB-A6-177-33.

Roddin MF, Ellis HT, Siddiqee WM. (1979). Background Data for Human Activity Patterns, Vols. 1, 2. Draft Final Report. Prepared for Strategies and Air Standards Division, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, N.C.

Sax SN, Bennett DH, Chillrud SN, Kinney PL, Spengler JD. (2004). Differences in source emission rates of volatile organic compounds in inner-city residences of New York City and Los Angeles. *J Expos Anal Environ Epidemiol*. 14(S):95-109.

Spicer CW, Kenny DV, Ward GF, Billick IH (1993). Transformations, lifetimes, and sources of NO_2, HONO, and HNO_3 in indoor environments. *JAWMA*. 43(11):1479-1485.

Spier CE, Little DE, Trim SC, Johnson TR, Linn WS, Hackney JD. (1992). Activity patterns in elementary and high school students exposed to oxidant pollution. *J Expo Anal Environ Epidemiol*. 2:277-293.

Tsang AM and Klepeis NE. (1996). Descriptive Statistics Tables from a Detailed Analysis of the National Human Activity Pattern Survey (NHAPS) Data. U.S. Environmental Protection Agency. EPA/600/R-96/148.

US Census Bureau. (2007). Employment Status: 2000- Supplemental Tables. Available at: http://www.census.gov/population/www/cen2000/phc-t28.html.

US DOT. (2007). Part 3-The Journey To Work files. Bureau of Transportation Statistics (BTS). Available at: http://transtats.bts.gov/.

US EPA. (1999). Total Risk Integrated Methodology. Website: http://www.epa.gov/ttnatw01/urban/trim/trimpg.html.

US EPA. (2002). Consolidated Human Activities Database (CHAD) Users Guide. Database and documentation available at: http://www.epa.gov/chadnet1/.

US EPA. (2004). AERMOD: Description of Model Formulation. Office of Air Quality Planning and Standards. EPA-454/R-03-004. Available at: http://www.epa.gov/scram001/7thconf/aermod/aermod_mfd.pdf.

US EPA. (2006a). Total Risk Integrated Methodology (TRIM) - Air Pollutants Exposure Model Documentation (TRIM.Expo / APEX, Version 4) Volume I: User's Guide. Office of Air Quality Planning and Standards, Research Triangle Park, NC. June 2006. Available at: http://www.epa.gov/ttn/fera/human_apex.html.

US EPA. (2006b). Total Risk Integrated Methodology (TRIM) - Air Pollutants Exposure Model Documentation (TRIM.Expo / APEX, Version 4) Volume II: Technical Support Document.

Office of Air Quality Planning and Standards, Research Triangle Park, NC. June 2006. Available at: http://www.epa.gov/ttn/fera/human_apex.html.

US EPA. (2007a). Integrated Science Assessment for Oxides of Nitrogen – Health Criteria (First External Review Draft) and Annexes (August 2007). Research Triangle Park, NC: National Center for Environmental Assessment. Available at: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=181712.

US EPA. (2007b). 2002 National Emissions Inventory Data & Documentation. Available at: http://www.epa.gov/ttn/chief/net/2002inventory.html.

US EPA. (2007c). Clean Air Markets - Data and Maps. Emissions Prepackaged Data Sets. Available at: http://camddataandmaps.epa.gov/gdm/index.cfm?fuseaction=emissions.wizard.

US EPA. (2007d). Ozone Population Exposure Analysis for Selected Urban Areas (July 2007). Research Triangle Park, NC: Office of Air Quality Planning and Standards. EPA-452/R-07-010. Available at: http://epa.gov/ttn/naaqs/standards/ozone/s_o3_cr_td.html.

US EPA. (2007e). Review of the National Ambient Air Quality Standards for ozone: assessment of scientific and technical information. OAQPS Staff paper (July 2007). Research Triangle Park, NC: Office of Air Quality Planning and Standards. EPA-452/R-07-007a. Available at: http://epa.gov/ttn/naaqs/standards/ozone/s_o3_cr_sp.html.

Weisel CP, Zhang JJ, Turpin BJ, Morandi MT, Colome S, Stock TH, Spektor DM, Korn L, Winer A, Alimokhtari S, Kwon J, Mohan K, Harrington R, Giovanetti R, Cui W, Afshar M, Maberti S, Shendell D. (2004). Relationship of Indoor, Outdoor and Personal Air (RIOPA) study; study design, methods and quality assurance/control results. *J Exp Anal Environ Epidemiol*. 15:123-137.

Wiley JA, Robinson JP, Piazza T, Garrett K, Cirksena K, Cheng Y-T, Martin G. (1991a). Activity Patterns of California Residents: Final Report. California Air Resources Board, Sacramento, CA. ARB/R93/487. Available from: NTIS, Springfield, VA., PB94-108719.

Wiley JA, Robinson JP, Cheng Y-T, Piazza T, Stork L, Pladsen K. (1991b). Study of Children's Activity Patterns: Final Report. California Air Resources Board, Sacramento, CA. ARB-R-93/489.

Williams R, Suggs J, Creason J, Rodes C, Lawless P, Kwok R, Zweidinger R, Sheldon L. (2000). The 1998 Baltimore particulate matter epidemiology-exposure study: Part 2. Personal exposure associated with an elderly population. *J Expo Anal Environ Epidemiol*. 10(6):533-543.

Williams R, Suggs J, Rea A, Leovic K, Vette A, Croghan C, Sheldon L, Rodes C, Thornburg J, Ejire A, Herbst M, Sanders, Jr W. (2003a). The Research Triangle Park particulate matter panel study: PM mass concentration relationships. *Atmos Environ*. 37:5349-5363.

Williams R, Suggs J, Rea A, Sheldon L, Rodes C, Thornburg J. (2003b). The Research Triangle Park particulate patter panel study: modeling ambient source contribution to personal and residential PM mass concentrations. *Atmos Environ*. 37:5365-5378.

Wilson AL, Colome SD, Baker PE, Becker EW. (1986). Residential Indoor Air Quality Characterization Study of Nitrogen Dioxide, Phase I, Final Report. Prepared for Southern California Gas Company, Los Angeles.

Wilson AL, Colome SD, Tian Y, Baker PE, Becker EW, Behrens DW, Billick IH, Garrison CA. (1996). California residential air exchange rates and residence volumes. *J Expos Anal Environ Epidemiol*. 6(3):311-326.

Yao X, Lau NT, Chan CK, Fang M. (2005). The use of tunnel concentration profile data to determine the ratio of NO_2/NO_x directly emitted from vehicles. *Atmos Chem Phys Discuss*.

5:12723–12740. Available at: http://www.atmos-chem-phys-discuss.net/5/12723/2005/acpd-5-12723-2005.pdf.

United States
Environmental Protection
Agency

Office of Air Quality Planning and Standards
Air Quality Strategies and Standards Division
Research Triangle Park, NC

EPA-452/P-08-002
April 2008

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

Risk and Exposure Assessment to Support the Review of the NO2 Primary National Ambient Air Quality Standard: Draft Technical Support Document (TSD)

Appendices

EPA-452/P-08-002
April 2008

Risk and Exposure Assessment to Support the Review of the NO2 Primary National Ambient Air Quality Standard: Draft Technical Support Document (TSD)

Appendices

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
Research Triangle Park, North Carolina

Appendix A. Ambient Monitor Characterization

Appendix A contains details regarding physical attributes of each monitor used within the named locations (i.e., 18 specific locations were defined; it does not include the broadly grouped locations of "Other CMSA" or Not MSA). Each of these monitors met the criteria for containing a valid number of reported concentrations and were used throughout the air quality characterization. Data provided include monitor location and purpose, ground height and elevation above sea level, and distance to the nearest major roadway (Table A-1). In addition, the distances and emissions of stationary sources that emit > 5 tons NO_x per year were calculated for each monitor (Table A-2).

Table A-1. Attributes of location-specific ambient monitors used for air quality characterization and the distance to nearest major roadway.

							Monitor[3]		Roadway[4]	
Location	**ID**	**Lat**	**Long**	**Land Use**	**Location Type[1]**	**Objective[2]**	**Ht (m)**	**Elev (m)**	**Dist (m)**	**Type**
Atlanta	130890002	33.69	-84.29	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	5	308	432	3
Atlanta	130893001	33.85	-84.21	RESIDENTIAL	RURAL	OTHER	5	0	579	2
Atlanta	131210048	33.78	-84.40	COMMERCIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	5	290	134	3
Atlanta	132230003	33.93	-85.05	AGRICULTURAL	RURAL	GENERAL/BACKGROUND	4	417	>1000	-
Atlanta	132470001	33.59	-84.07	AGRICULTURAL	RURAL	POPULATION EXPOSURE	5	219	809	3
Boston	230313002	43.08	-70.75	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	-	40	70	2
Boston	250051005	42.06	-71.15	AGRICULTURAL	RURAL	POPULATION EXPOSURE	4	61	17	3
Boston	250092006	42.47	-70.97	COMMERCIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	5	52	158	3
Boston	250094004	42.79	-70.81	RESIDENTIAL	SUBURBAN	MAX OZONE CONCENTRATION	4	1	15	3
Boston	250095005	42.76	-71.11	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	-	0	337	3
Boston	250210009	42.32	-71.13	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	4	0	144	3
Boston	250250002	42.35	-71.10	COMMERCIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	5	6	7	2
Boston	250250021	42.38	-71.03	RESIDENTIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	4	6	7	3
Boston	250250035	42.33	-71.12	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	-	0	158	3
Boston	250250036	42.33	-71.12	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	-	0	158	3
Boston	250250040	42.35	-71.04	INDUSTRIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	4	0	37	3
Boston	250250041	42.32	-70.97	COMMERCIAL	RURAL	POPULATION EXPOSURE	6	10	>1000	-
Boston	250250042	42.33	-71.08	COMMERCIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	5	6	26	3
Boston	250251003	42.40	-71.03	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	59	228	4
Boston	250270020	42.27	-71.80	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	3	145	44	3
Boston	250270023	42.27	-71.79	COMMERCIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	4	145	49	3
Boston	330110016	42.99	-71.46	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	5	75	168	3
Boston	330110019	43.00	-71.47	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	-	61	70	3
Boston	330110020	43.00	-71.47	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	5	61	70	3
Boston	330150009	43.08	-70.76	COMMERCIAL	SUBURBAN	UNKNOWN	3	3	48	3
Boston	330150013	43.00	-71.20	RESIDENTIAL	RURAL	OTHER	1	0	>1000	-
Boston	330150014	43.08	-70.75	RESIDENTIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	2	4	266	3

Table A-1. Attributes of location-specific ambient monitors used for air quality characterization and the distance to nearest major roadway.

							Monitor[3]		Roadway[4]	
Location	**ID**	**Lat**	**Long**	**Land Use**	**Location Type[1]**	**Objective[2]**	**Ht (m)**	**Elev (m)**	**Dist (m)**	**Type**
Boston	330150015	43.08	-70.76	COMMERCIAL	SUBURBAN	POPULATION EXPOSURE	4	3	38	3
Chicago	170310037	41.98	-87.67	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	-	183	17	3
Chicago	170310063	41.88	-87.63	MOBILE	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	3	181	68	3
Chicago	170310064	41.79	-87.60	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	15	180	346	3
Chicago	170310075	41.96	-87.66	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	15	180	136	3
Chicago	170310076	41.75	-87.71	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	186	2	3
Chicago	170313101	41.97	-87.88	MOBILE	SUBURBAN	HIGHEST CONCENTRATION	3	197	20	2
Chicago	170313103	41.97	-87.88	MOBILE	SUBURBAN	HIGHEST CONCENTRATION	4	195	20	2
Chicago	170314002	41.86	-87.75	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	184	118	3
Chicago	170314201	42.14	-87.80	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	8	198	239	2
Chicago	170314201	42.14	-87.80	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	8	198	239	2
Chicago	170318003	41.63	-87.57	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	179	2	3
Chicago	171971011	41.22	-88.19	AGRICULTURAL	RURAL	GENERAL/BACKGROUND	5	181	>1000	-
Chicago	180890022	41.61	-87.30	INDUSTRIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	5	183	738	1
Chicago	180891016	41.60	-87.33	RESIDENTIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	14	183	187	3
Cleveland	390350043	41.46	-81.58	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	287	187	2
Cleveland	390350060	41.49	-81.68	COMMERCIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	4	206	2	4
Cleveland	390350066	41.46	-81.58	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	5	287	187	2
Cleveland	390350070	41.46	-81.59	RESIDENTIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	4	278	81	3
Colorado Springs	080416001	38.63	-104.72	INDUSTRIAL	RURAL	UNKNOWN	4	1673	>1000	-
Colorado Springs	080416004	38.92	-104.81	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	4	1931	150	1
Colorado Springs	080416005	38.76	-104.76	AGRICULTURAL	URBAN AND CENTER CITY	UNKNOWN	4	1747	79	3
Colorado Springs	080416006	38.92	-105.00	RESIDENTIAL	RURAL	UNKNOWN	4	2313	199	2
Colorado Springs	080416009	38.64	-104.71	INDUSTRIAL	RURAL	UNKNOWN	4	1707	>1000	-
Colorado Springs	080416011	38.85	-104.83	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	3	1832	198	3
Colorado Springs	080416013	38.81	-104.82	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	3	1823	386	4
Colorado Springs	080416018	38.81	-104.75	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	3	1795	163	2
Denver	080013001	39.84	-104.95	AGRICULTURAL	RURAL	POPULATION EXPOSURE	4	1559	748	3
Denver	080050003	39.66	-105.00	COMMERCIAL	SUBURBAN	HIGHEST CONCENTRATION	4	1654	138	2

Table A-1. Attributes of location-specific ambient monitors used for air quality characterization and the distance to nearest major roadway.

							Monitor[3]		Roadway[4]	
Location	**ID**	**Lat**	**Long**	**Land Use**	**Location Type[1]**	**Objective[2]**	**Ht (m)**	**Elev (m)**	**Dist (m)**	**Type**
Denver	080310002	39.75	-104.99	COMMERCIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	-	1589	18	3
Denver	080590006	39.91	-105.19	INDUSTRIAL	RURAL	UNKNOWN	-	1774	65	3
Denver	080590008	39.88	-105.17	INDUSTRIAL	RURAL	GENERAL/BACKGROUND	4	1715	31	3
Denver	080590009	39.86	-105.20	INDUSTRIAL	RURAL	GENERAL/BACKGROUND	4	1848	99	3
Denver	080590010	39.90	-105.24	AGRICULTURAL	RURAL	UNKNOWN	4	1877	63	2
Detroit	260990009	42.73	-82.79	COMMERCIAL	SUBURBAN	UNKNOWN	-	189	415	3
Detroit	261630016	42.36	-83.10	RESIDENTIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	4	191	393	5
Detroit	261630019	42.43	-83.00	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	192	339	3
El Paso	481410027	31.76	-106.49	COMMERCIAL	URBAN AND CENTER CITY	GENERAL/BACKGROUND	5	1140	33	4
El Paso	481410028	31.75	-106.40	RESIDENTIAL	SUBURBAN	SOURCE ORIENTED	5	1126	718	3
El Paso	481410037	31.77	-106.50	COMMERCIAL	URBAN AND CENTER CITY	MAX OZONE CONCENTRATION	4	1143	128	3
El Paso	481410044	31.77	-106.46	COMMERCIAL	URBAN AND CENTER CITY	MAX PRECURSOR EMISSIONS IMPACT	5	1128	38	3
El Paso	481410055	31.75	-106.40	COMMERCIAL	URBAN AND CENTER CITY	UPWIND BACKGROUND	5	0	127	3
El Paso	481410057	31.66	-106.30	RESIDENTIAL	SUBURBAN	GENERAL/BACKGROUND	5	0	450	3
El Paso	481410058	31.89	-106.43	RESIDENTIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	5	0	478	3
Jacksonville	120310032	30.36	-81.64	COMMERCIAL	SUBURBAN	UNKNOWN	3	7	144	1
Las Vegas	320030022	36.39	-114.91	INDUSTRIAL	RURAL	SOURCE ORIENTED	3.5	0	122	2
Las Vegas	320030023	36.81	-114.06	RESIDENTIAL	RURAL	POPULATION EXPOSURE	4	490	303	3
Las Vegas	320030073	36.17	-115.33	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	3.5	0	515	2
Las Vegas	320030078	35.47	-114.92	DESERT	RURAL	REGIONAL TRANSPORT	4	1094	25	3
Las Vegas	320030539	36.14	-115.09	MOBILE	SUBURBAN	POPULATION EXPOSURE	3.5	533	11	3
Las Vegas	320030557	36.16	-115.11	RESIDENTIAL	SUBURBAN	UNKNOWN	3	567	1	3
Las Vegas	320030563	36.18	-115.10	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	570	254	3
Las Vegas	320030601	35.98	-114.84	COMMERCIAL	SUBURBAN	POPULATION EXPOSURE	4	0	52	3
Las Vegas	320031019	35.79	-115.36	DESERT	RURAL	GENERAL/BACKGROUND	4	950	914	3
Las Vegas	320032002	36.19	-115.12	COMMERCIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	3.5	0	240	3
Los Angeles	060370002	34.14	-117.92	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	2	183	329	3
Los Angeles	060370016	34.14	-117.85	RESIDENTIAL	SUBURBAN	UNKNOWN	6	275	300	3
Los Angeles	060370030	34.04	-118.22	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	5	65	50	3
Los Angeles	060370113	34.05	-118.46	MOBILE	URBAN AND CENTER CITY	UNKNOWN	5	91	190	3

Table A-1. Attributes of location-specific ambient monitors used for air quality characterization and the distance to nearest major roadway.

Location	ID	Lat	Long	Land Use	Location Type[1]	Objective[2]	Monitor[3]		Roadway[4]	
							Ht (m)	Elev (m)	Dist (m)	Type
Los Angeles	060370206	33.96	-117.84	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	-	300	>1000	-
Los Angeles	060371002	34.18	-118.32	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	5	168	58	3
Los Angeles	060371103	34.07	-118.23	RESIDENTIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	13	87	55	3
Los Angeles	060371201	34.20	-118.53	COMMERCIAL	SUBURBAN	UNKNOWN	6	226	206	3
Los Angeles	060371301	33.93	-118.21	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	7	27	29	3
Los Angeles	060371601	34.01	-118.06	COMMERCIAL	SUBURBAN	POPULATION EXPOSURE	6	75	78	3
Los Angeles	060371701	34.07	-117.75	COMMERCIAL	SUBURBAN	UNKNOWN	6	270	15	3
Los Angeles	060372005	34.13	-118.13	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	4	250	385	3
Los Angeles	060374002	33.82	-118.19	RESIDENTIAL	SUBURBAN	UNKNOWN	6	6	1	3
Los Angeles	060375001	33.92	-118.37	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	-	21	10	3
Los Angeles	060375005	33.95	-118.43	RESIDENTIAL	SUBURBAN	UPWIND BACKGROUND	4	21	149	3
Los Angeles	060376002	34.39	-118.53	COMMERCIAL	SUBURBAN	POPULATION EXPOSURE	-	375	2	3
Los Angeles	060376012	34.38	-118.53	COMMERCIAL	SUBURBAN	UNKNOWN	-	397	143	3
Los Angeles	060379002	34.69	-118.13	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	5	725	61	3
Los Angeles	060379033	34.67	-118.13	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	3	725	146	3
Los Angeles	060590001	33.83	-117.94	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	5	45	225	3
Los Angeles	060590007	33.83	-117.94	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	10	225	3
Los Angeles	060591003	33.67	-117.93	RESIDENTIAL	SUBURBAN	UNKNOWN	6	0	202	3
Los Angeles	060595001	33.93	-117.95	RESIDENTIAL	SUBURBAN	UNKNOWN	82	82	570	3
Los Angeles	060650012	33.92	-116.86	COMMERCIAL	SUBURBAN	POPULATION EXPOSURE	4	677	432	1
Los Angeles	060655001	33.85	-116.54	RESIDENTIAL	SUBURBAN	UNKNOWN	6	171	75	3
Los Angeles	060658001	34.00	-117.42	RESIDENTIAL	SUBURBAN	UNKNOWN	4	250	133	3
Los Angeles	060659001	33.68	-117.33	RESIDENTIAL	SUBURBAN	UNKNOWN	-	1440	522	4
Los Angeles	060710001	34.90	-117.02	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	8	690	64	3
Los Angeles	060710012	34.43	-117.56	COMMERCIAL	RURAL	UNKNOWN	-	4100	30	3
Los Angeles	060710014	34.51	-117.33	RESIDENTIAL	SUBURBAN	UNKNOWN	4	876	18	3
Los Angeles	060710015	35.78	-117.37	INDUSTRIAL	SUBURBAN	UNKNOWN	-	498	42	3
Los Angeles	060710017	34.14	-116.06	MOBILE	URBAN AND CENTER CITY	UNKNOWN	4	607	64	3
Los Angeles	060710306	34.51	-117.33	RESIDENTIAL	SUBURBAN	UNKNOWN	4	913	38	3
Los Angeles	060711004	34.10	-117.63	RESIDENTIAL	URBAN AND CENTER	UPWIND BACKGROUND	6	369	349	2

Table A-1. Attributes of location-specific ambient monitors used for air quality characterization and the distance to nearest major roadway.

Location	ID	Lat	Long	Land Use	Location Type[1]	Objective[2]	Monitor[3]		Roadway[4]	
							Ht (m)	Elev (m)	Dist (m)	Type
					CITY					
Los Angeles	060712002	34.10	-117.49	INDUSTRIAL	SUBURBAN	UNKNOWN	5	381	81	3
Los Angeles	060711234	35.76	-117.40	DESERT	RURAL	OTHER	1	545	>1000	-
Los Angeles	060714001	34.42	-117.28	RESIDENTIAL	SUBURBAN	UNKNOWN	-	1006	111	3
Los Angeles	060719004	34.11	-117.27	COMMERCIAL	SUBURBAN	HIGHEST CONCENTRATION	5	0	169	3
Los Angeles	061110005	33.20	-117.37	UNKNOWN	UNKNOWN	POPULATION EXPOSURE	1	320	63	3
Los Angeles	061110007	32.71	-117.15	COMMERCIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	5	244	89	3
Los Angeles	061111003	34.45	-119.27	MOBILE	SUBURBAN	UNKNOWN	-	231	18	2
Los Angeles	061111004	34.45	-119.23	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	262	56	3
Los Angeles	061112002	34.28	-118.68	RESIDENTIAL	SUBURBAN	HIGHEST CONCENTRATION	4	314	471	1
Los Angeles	061112003	34.28	-119.31	RESIDENTIAL	SUBURBAN	GENERAL/BACKGROUND	2	3	90	1
Los Angeles	061113001	34.26	-119.14	RESIDENTIAL	RURAL	POPULATION EXPOSURE	4	43	307	3
Miami	120110003	26.28	-80.28	INDUSTRIAL	RURAL	HIGHEST CONCENTRATION	6	3	22	3
Miami	120110031	26.27	-80.30	RESIDENTIAL	SUBURBAN	MAX PRECURSOR EMISSIONS IMPACT	4	3	103	4
Miami	120118002	26.09	-80.11	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	3	>1000	-
Miami	120860027	25.73	-80.16	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	16	2	15	3
Miami	120864002	25.80	-80.21	COMMERCIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	4	5	87	3
New York	090010113	41.18	-73.19	COMMERCIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	4	3	8	3
New York	090019003	41.12	-73.34	FOREST	RURAL	POPULATION EXPOSURE	5	4	508	4
New York	090090027	41.30	-72.90	COMMERCIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	3.67	11	237	1
New York	090091123	41.31	-72.92	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	9	18	14	2
New York	340030001	40.81	-73.99	RESIDENTIAL	SUBURBAN	UNKNOWN	4	61	82	3
New York	340030005	40.90	-74.03	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	3	6	172	5
New York	340130011	40.73	-74.14	INDUSTRIAL	URBAN AND CENTER CITY	UNKNOWN	4	3	232	1
New York	340130016	40.72	-74.15	INDUSTRIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	5	3	6	1
New York	340131003	40.76	-74.20	COMMERCIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	4	48.45	25	3
New York	340170006	40.67	-74.13	COMMERCIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	5	3	266	3
New York	340210005	40.28	-74.74	RESIDENTIAL	SUBURBAN	MAX OZONE CONCENTRATION	4	30	442	1
New York	340230011	40.46	-74.43	AGRICULTURAL	RURAL	POPULATION EXPOSURE	4	21	298	3
New York	340273001	40.79	-74.68	AGRICULTURAL	RURAL	UNKNOWN	5	274	227	3

Table A-1. Attributes of location-specific ambient monitors used for air quality characterization and the distance to nearest major roadway.

							Monitor[3]		Roadway[4]	
Location	**ID**	**Lat**	**Long**	**Land Use**	**Location Type[1]**	**Objective[2]**	**Ht (m)**	**Elev (m)**	**Dist (m)**	**Type**
New York	340390004	40.64	-74.21	INDUSTRIAL	SUBURBAN	HIGHEST CONCENTRATION	4	5.4	37	4
New York	340390008	40.60	-74.44	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	0	99	3
New York	360050080	40.84	-73.92	RESIDENTIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	15	15	122	3
New York	360050083	40.87	-73.88	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	15	24	132	5
New York	360050110	40.82	-73.90	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	-	0	76	3
New York	360470011	40.73	-73.95	INDUSTRIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	6	9	171	3
New York	360590005	40.74	-73.59	COMMERCIAL	SUBURBAN	HIGHEST CONCENTRATION	5	27	32	3
New York	360610010	40.74	-73.99	RESIDENTIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	38	38	55	3
New York	360610056	40.76	-73.97	COMMERCIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	10	15	62	3
New York	360810097	40.76	-73.76	RESIDENTIAL	URBAN AND CENTER CITY	GENERAL/BACKGROUND	12	0	197	3
New York	360810098	40.78	-73.85	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	8	6	9	3
New York	360810124	40.74	-73.82	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	-	8	150	3
New York	361030009	40.83	-73.06	RESIDENTIAL	SUBURBAN	UNKNOWN	-	0	116	2
Philadelphia	100031003	39.76	-75.49	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	-	65	189	2
Philadelphia	100031007	39.55	-75.73	AGRICULTURAL	RURAL	OTHER	-	20	144	3
Philadelphia	100032004	39.74	-75.56	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	-	0	82	3
Philadelphia	340070003	39.92	-75.10	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	5	7.6	405	3
Philadelphia	420170012	40.11	-74.88	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	2	12	393	3
Philadelphia	420450002	39.84	-75.37	INDUSTRIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	2	3	413	3
Philadelphia	420910013	40.11	-75.31	RESIDENTIAL	SUBURBAN	UNKNOWN	4	53	630	1
Philadelphia	421010004	40.01	-75.10	RESIDENTIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	7	22	45	3
Philadelphia	421010029	39.96	-75.17	COMMERCIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	11	25	103	3
Philadelphia	421010047	39.94	-75.17	RESIDENTIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	11	21	66	2
Phoenix	040130019	33.48	-112.14	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4.3	333	401	3
Phoenix	040133002	33.46	-112.05	RESIDENTIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	9	339	141	3
Phoenix	040133003	33.48	-111.92	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	5.8	368	78	3
Phoenix	040133010	33.46	-112.12	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4.2	325	7	3

Table A-1. Attributes of location-specific ambient monitors used for air quality characterization and the distance to nearest major roadway.

							Monitor[3]		Roadway[4]	
Location	**ID**	**Lat**	**Long**	**Land Use**	**Location Type[1]**	**Objective[2]**	**Ht (m)**	**Elev (m)**	**Dist (m)**	**Type**
Phoenix	040134005	33.41	-111.93	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	4	352	259	3
Phoenix	040134011	33.37	-112.62	AGRICULTURAL	RURAL	SOURCE ORIENTED	4	258	12	3
Phoenix	040139997	33.50	-112.10	RESIDENTIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	-	346	433	3
Provo	490490002	40.25	-111.66	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	4	1402	353	2
St. Louis	171630010	38.61	-90.16	INDUSTRIAL	SUBURBAN	POPULATION EXPOSURE	4	125	18	4
St. Louis	291830010	38.58	-90.84	AGRICULTURAL	RURAL	UNKNOWN	3	0	340	3
St. Louis	291831002	38.87	-90.23	AGRICULTURAL	RURAL	POPULATION EXPOSURE	4	131	31	3
St. Louis	291890001	38.52	-90.34	RESIDENTIAL	SUBURBAN	UNKNOWN	4	183	161	2
St. Louis	291890004	38.53	-90.38	RESIDENTIAL	SUBURBAN	UNKNOWN	4	183	95	2
St. Louis	291890006	38.61	-90.50	RESIDENTIAL	RURAL	UNKNOWN	4	175	97	3
St. Louis	291893001	38.64	-90.35	COMMERCIAL	SUBURBAN	UNKNOWN	4	161	5	1
St. Louis	291895001	38.77	-90.29	COMMERCIAL	SUBURBAN	UNKNOWN	2	168	421	3
St. Louis	291897002	38.73	-90.38	RESIDENTIAL	SUBURBAN	UNKNOWN	4	168	59	3
St. Louis	291897003	38.72	-90.37	RESIDENTIAL	SUBURBAN	HIGHEST CONCENTRATION	4	0	112	3
St. Louis	295100072	38.62	-90.20	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	14	154	43	4
St. Louis	295100080	38.68	-90.25	RESIDENTIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	4	152	116	3
St. Louis	295100086	38.67	-90.24	RESIDENTIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	4	0	133	3
Washington DC	110010017	38.90	-77.05	COMMERCIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	10	20	54	3
Washington DC	110010025	38.98	-77.02	COMMERCIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	11	91	106	3
Washington DC	110010041	38.90	-76.95	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	-	8	141	4
Washington DC	110010043	38.92	-77.01	COMMERCIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	-	50	278	3
Washington DC	240053001	39.31	-76.47	RESIDENTIAL	SUBURBAN	MAX PRECURSOR EMISSIONS IMPACT	4.6	5	186	3
Washington DC	245100040	39.30	-76.60	RESIDENTIAL	URBAN AND CENTER CITY	HIGHEST CONCENTRATION	4.2	12	14	3
Washington DC	245100050	39.32	-76.58	RESIDENTIAL	URBAN AND CENTER CITY	POPULATION EXPOSURE	4	49	338	2
Washington DC	510130020	38.86	-77.06	COMMERCIAL	URBAN AND CENTER CITY	UNKNOWN	7	171	80	3
Washington DC	510590005	38.89	-77.47	AGRICULTURAL	RURAL	POPULATION EXPOSURE	4	77	315	5
Washington DC	510590018	38.74	-77.08	RESIDENTIAL	SUBURBAN	UNKNOWN	4	11	54	3

Table A-1. Attributes of location-specific ambient monitors used for air quality characterization and the distance to nearest major roadway.										
							Monitor[3]		Roadway[4]	
Location	ID	Lat	Long	Land Use	Location Type[1]	Objective[2]	Ht (m)	Elev (m)	Dist (m)	Type
Washington DC	510591004	38.87	-77.14	COMMERCIAL	SUBURBAN	UNKNOWN	11	110	84	5
Washington DC	510591005	38.84	-77.16	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	-	83.9	50	3
Washington DC	510595001	38.93	-77.20	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	106	18	5
Washington DC	511071005	39.02	-77.49	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	0	75	3
Washington DC	511530009	38.86	-77.64	RESIDENTIAL	SUBURBAN	POPULATION EXPOSURE	4	111	196	2
Washington DC	515100009	38.81	-77.04	RESIDENTIAL	URBAN AND CENTER CITY	UNKNOWN	11	23	83	3

[1] Land use indicates the prevalent land use within 1/4 mile of that site.

[2] Objective Indicates the reason for measuring air quality by the monitor.

[3] Monitor probe height (Ht) and site elevation (Elev) above sea level are given in meters (m).

[4] Distances (Dist) to roadway are given in meters (m). Major road types are defined as: 1=primary limited access or interstate, 2=primary US and State highways, 3=Secondary State and County, 4=freeway ramp, 5=other ramps.

Table A-2. Distance of location-specific ambient monitors used in Tier 1 analyses to stationary sources emitting > 5 tons of NO_x per year, within a 10 kilometer distance of monitoring site.

		Distance (km) to Source emissions >5 tpy and within 10 km								Emissions (tpy) of Sources within 10 km and >5 tpy						
Location	**ID**	**n**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**
Atlanta	130890002	1	4.9		4.9	4.9	4.9	4.9	4.9	34		34	34	34	34	34
Atlanta	130893001	3	7.2	4.0	2.7	2.7	9.2	9.8	9.8	34	2	32	32	34	36	36
Atlanta	131210048	5	6.4	3.3	0.7	0.7	7.3	8.9	8.9	1249	2106	22	22	39	4895	4895
Atlanta	132230003	0														
Atlanta	132470001	0														
Boston	230313002	5	3.5	1.5	1.0	1.0	3.8	4.9	4.9	642	769	31	31	203	1860	1860
Boston	250051005	3	6.7	1.6	5.5	5.5	6.0	8.5	8.5	9	4	5	5	8	14	14
Boston	250092006	12	6.8	2.7	2.5	2.5	7.4	9.9	9.9	439	1083	5	5	21	3794	3794
Boston	250094004	0														
Boston	250095005	10	5.8	2.3	1.7	1.7	6.7	8.6	8.6	201	347	6	6	29	923	923
Boston	250210009	57	5.8	2.5	1.0	1.8	5.9	9.9	9.9	106	283	5	5	9	1155	1419
Boston	250250002	62	4.6	2.4	0.6	1.1	4.3	9.4	9.7	98	273	5	5	9	1155	1419
Boston	250250021	55	6.1	2.3	1.5	1.7	6.5	9.8	9.8	130	304	5	5	11	1155	1419
Boston	250250035	62	5.1	2.6	0.3	0.8	5.1	9.0	9.6	99	273	5	5	9	1155	1419
Boston	250250036	62	5.1	2.6	0.3	0.8	5.1	9.0	9.6	99	273	5	5	9	1155	1419
Boston	250250040	56	5.3	2.4	0.4	0.9	5.6	9.0	9.3	106	286	5	5	9	1155	1419
Boston	250250041	25	7.8	2.0	0.7	0.7	8.2	9.9	9.9	81	206	5	5	11	957	957
Boston	250250042	65	5.3	2.8	0.7	1.0	4.9	10.0	10.0	94	267	5	5	9	1155	1419
Boston	250251003	49	6.4	2.4	0.6	1.0	7.0	9.6	9.6	145	319	5	5	11	1155	1419
Boston	250270020	28	3.7	2.5	0.1	0.1	2.9	8.6	8.6	58	165	5	5	13	868	868
Boston	250270023	28	3.6	2.4	0.4	0.4	3.0	8.4	8.4	58	165	5	5	13	868	868
Boston	330110016	0														
Boston	330110019	0														
Boston	330110020	0														
Boston	330150009	5	3.3	1.0	2.0	2.0	3.3	4.4	4.4	642	769	31	31	203	1860	1860
Boston	330150013	1	8.4		8.4	8.4	8.4	8.4	8.4	29		29	29	29	29	29
Boston	330150014	5	4.0	1.8	1.0	1.0	4.4	5.5	5.5	642	769	31	31	203	1860	1860
Boston	330150015	5	3.1	0.9	1.9	1.9	3.0	4.1	4.1	642	769	31	31	203	1860	1860
Chicago	170310037	17	5.6	2.7	0.7	0.7	5.7	9.5	9.5	18	31	5	5	7	126	126
Chicago	170310063	57	4.9	3.2	0.4	0.5	4.9	9.4	10.0	110	416	5	5	9	1677	2465
Chicago	170310064	33	6.9	2.5	1.2	1.2	6.9	10.0	10.0	94	428	5	5	10	2465	2465
Chicago	170310075	31	7.3	2.7	0.8	0.8	8.4	9.9	9.9	10	7	5	5	7	36	36
Chicago	170310076	46	7.8	2.3	1.3	1.6	8.4	9.8	9.9	170	463	5	5	10	1677	2204
Chicago	170313101	30	6.6	2.2	2.7	2.7	7.2	9.7	9.7	313	1638	5	5	9	8985	8985
Chicago	170313103	30	6.6	2.2	2.7	2.7	7.2	9.7	9.7	313	1638	5	5	9	8985	8985
Chicago	170314002	63	6.7	2.6	0.5	0.5	7.2	9.8	9.9	122	407	5	5	9	1677	2465
Chicago	170314201	7	6.5	1.5	4.0	4.0	6.6	9.0	9.0	8	3	5	5	8	14	14
Chicago	170314201	7	6.5	1.5	4.0	4.0	6.6	9.0	9.0	8	3	5	5	8	14	14

Table A-2. Distance of location-specific ambient monitors used in Tier 1 analyses to stationary sources emitting > 5 tons of NO_x per year, within a 10 kilometer distance of monitoring site.

		Distance (km) to Source emissions >5 tpy and within 10 km								Emissions (tpy) of Sources within 10 km and >5 tpy						
Location	**ID**	**n**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**
Chicago	170318003	63	7.3	2.0	1.7	2.3	8.0	9.6	9.7	361	1201	5	5	18	6216	7141
Chicago	171971011	1	4.0		4.0	4.0	4.0	4.0	4.0	20		20	20	20	20	20
Chicago	180890022	8	5.1	3.8	0.8	0.8	4.1	9.4	9.4	815	1680	8	8	243	4936	4936
Chicago	180891016	8	4.7	2.4	2.1	2.1	4.1	7.6	7.6	815	1680	8	8	243	4936	4936
Cleveland	390350043	5	8.1	1.9	5.2	5.2	8.3	9.9	9.9	673	664	126	126	284	1476	1476
Cleveland	390350060	4	4.1	2.4	1.0	1.0	4.4	6.4	6.4	810	681	165	165	800	1476	1476
Cleveland	390350066	5	8.0	1.9	5.2	5.2	8.3	9.8	9.8	673	664	126	126	284	1476	1476
Cleveland	390350070	5	7.6	1.8	5.5	5.5	7.3	9.7	9.7	673	664	126	126	284	1476	1476
Colorado Springs	080416001	4	5.1	4.4	0.8	0.8	5.1	9.1	9.1	780	1374	16	16	133	2835	2835
Colorado Springs	080416004	10	5.9	2.2	3.5	3.5	5.6	9.8	9.8	48	80	5	5	17	267	267
Colorado Springs	080416005	9	7.5	2.1	3.3	3.3	8.1	9.5	9.5	490	1393	5	5	11	4205	4205
Colorado Springs	080416006	0														
Colorado Springs	080416009	4	5.2	4.3	1.0	1.0	5.3	9.3	9.3	780	1374	16	16	133	2835	2835
Colorado Springs	080416011	14	5.0	2.3	2.0	2.0	5.8	9.6	9.6	345	1113	5	5	22	4205	4205
Colorado Springs	080416013	14	6.3	2.9	2.1	2.1	6.9	9.9	9.9	346	1113	5	5	27	4205	4205
Colorado Springs	080416018	11	6.9	1.7	4.3	4.3	7.1	9.6	9.6	430	1254	5	5	34	4205	4205
Denver	080013001	34	5.3	1.8	1.6	1.6	4.7	9.5	9.5	310	1622	5	5	15	9483	9483
Denver	080050003	19	6.7	3.7	1.0	1.0	9.1	10.0	10.0	313	1233	5	5	17	5404	5404
Denver	080310002	52	5.3	2.5	0.9	0.9	5.8	9.7	9.8	319	1495	5	5	14	5404	9483
Denver	080590006	9	5.9	2.1	2.7	2.7	6.3	8.6	8.6	63	66	11	11	39	182	182
Denver	080590008	9	6.2	2.0	3.7	3.7	6.1	10.0	10.0	59	68	8	8	13	182	182
Denver	080590009	10	6.5	3.2	2.5	2.5	7.0	9.9	9.9	53	66	6	6	13	182	182
Denver	080590010	7	5.5	3.1	1.1	1.1	5.6	9.2	9.2	73	71	12	12	44	182	182
Detroit	260990009	4	4.9	3.2	0.3	0.3	5.7	7.7	7.7	63	70	7	7	46	152	152
Detroit	261630016	51	7.4	2.1	1.3	2.0	7.9	9.8	9.9	387	797	5	6	41	3087	3762
Detroit	261630019	32	6.3	2.2	2.6	2.6	6.5	10.0	10.0	57	168	5	5	12	837	837
El Paso	481410027	22	8.1	1.6	1.5	1.5	8.6	9.3	9.3	99	195	5	5	29	912	912
El Paso	481410028	24	2.2	1.9	0.9	0.9	1.6	9.3	9.3	127	338	5	5	32	1679	1679
El Paso	481410037	15	8.7	2.6	0.1	0.1	9.4	10.0	10.0	135	230	5	5	38	912	912
El Paso	481410044	25	5.9	1.2	4.4	4.4	5.6	9.5	9.5	158	366	5	5	32	1679	1679
El Paso	481410055	24	2.8	1.8	1.6	1.6	2.2	9.6	9.6	127	338	5	5	32	1679	1679

Table A-2. Distance of location-specific ambient monitors used in Tier 1 analyses to stationary sources emitting > 5 tons of NO_x per year, within a 10 kilometer distance of monitoring site.

		Distance (km) to Source emissions >5 tpy and within 10 km								**Emissions (tpy) of Sources within 10 km and >5 tpy**						
Location	**ID**	**n**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**
El Paso	481410057	0														
El Paso	481410058	16	8.8	0.4	8.4	8.4	8.6	9.5	9.5	31	30	5	5	23	106	106
Jacksonville	120310032	20	5.1	3.0	0.7	0.7	5.7	9.6	9.6	201	407	5	5	31	1642	1642
Las Vegas	320030022	7	4.6	0.9	3.8	3.8	3.9	5.6	5.6	175	222	30	30	77	650	650
Las Vegas	320030023	0														
Las Vegas	320030073	0														
Las Vegas	320030078	0														
Las Vegas	320030539	5	6.9	1.2	4.7	4.7	7.2	7.9	7.9	816	760	18	18	851	1665	1665
Las Vegas	320030557	4	9.1	1.2	7.3	7.3	9.7	9.7	9.7	807	877	18	18	772	1665	1665
Las Vegas	320030563	1	7.6		7.6	7.6	7.6	7.6	7.6	84		84	84	84	84	84
Las Vegas	320030601	0														
Las Vegas	320031019	0														
Las Vegas	320032002	1	9.9		9.9	9.9	9.9	9.9	9.9	84		84	84	84	84	84
Los Angeles	060370002	7	3.1	1.1	1.6	1.6	2.9	4.5	4.5	10	4	5	5	9	16	16
Los Angeles	060370016	7	7.5	1.8	4.5	4.5	8.5	8.9	8.9	12	8	5	5	9	29	29
Los Angeles	060370030	35	5.5	2.3	2.1	2.1	5.2	9.8	9.8	23	27	5	5	11	115	115
Los Angeles	060370113	7	4.3	3.1	1.3	1.3	3.2	9.8	9.8	15	10	5	5	13	36	36
Los Angeles	060370206	11	5.6	2.2	2.3	2.3	5.8	9.2	9.2	32	31	6	6	20	109	109
Los Angeles	060371002	18	5.7	2.6	0.1	0.1	6.0	9.9	9.9	47	59	6	6	24	215	215
Los Angeles	060371103	31	6.5	2.7	1.8	1.8	7.2	10.0	10.0	18	21	5	5	10	86	86
Los Angeles	060371201	7	5.1	1.2	3.3	3.3	5.5	6.5	6.5	10	4	6	6	10	15	15
Los Angeles	060371301	45	6.8	2.1	1.2	2.5	7.1	9.7	10.0	22	24	5	5	12	86	115
Los Angeles	060371601	22	6.5	2.3	2.3	2.3	7.2	9.7	9.7	28	33	5	5	12	115	115
Los Angeles	060371701	13	6.1	3.0	1.1	1.1	7.0	9.7	9.7	22	20	5	5	16	70	70
Los Angeles	060372005	10	5.2	3.5	0.2	0.2	5.5	10.0	10.0	12	8	5	5	9	30	30
Los Angeles	060374002	55	6.4	2.3	1.7	2.2	6.2	9.9	9.9	76	159	5	5	16	744	789
Los Angeles	060375001	32	5.1	2.4	0.3	0.3	4.8	9.6	9.6	205	754	6	6	21	4256	4256
Los Angeles	060375005	25	4.6	2.4	1.4	1.4	4.6	9.9	9.9	224	850	6	6	21	4256	4256
Los Angeles	060376002	5	5.6	1.8	3.6	3.6	5.8	7.8	7.8	29	20	8	8	18	54	54
Los Angeles	060376012	6	6.2	2.5	3.0	3.0	6.8	9.7	9.7	26	19	8	8	18	54	54
Los Angeles	060379002	4	7.8	1.0	6.8	6.8	7.7	9.2	9.2	22	28	6	6	9	64	64
Los Angeles	060379033	4	6.3	0.8	5.3	5.3	6.4	7.1	7.1	22	28	6	6	9	64	64
Los Angeles	060590001	17	6.4	2.4	2.8	2.8	7.2	9.4	9.4	14	12	5	5	8	46	46
Los Angeles	060590007	17	6.4	2.4	2.8	2.8	7.2	9.4	9.4	14	12	5	5	8	46	46
Los Angeles	060591003	14	6.1	2.2	2.1	2.1	6.0	9.3	9.3	65	116	5	5	10	434	434
Los Angeles	060595001	16	7.9	1.6	3.4	3.4	8.2	9.5	9.5	19	26	6	6	9	109	109
Los Angeles	060650012	0														
Los Angeles	060655001	0														

Table A-2. Distance of location-specific ambient monitors used in Tier 1 analyses to stationary sources emitting > 5 tons of NO_x per year, within a 10 kilometer distance of monitoring site.

		Distance (km) to Source emissions >5 tpy and within 10 km								Emissions (tpy) of Sources within 10 km and >5 tpy						
Location	**ID**	**n**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**
Los Angeles	060658001	12	7.4	2.2	3.6	3.6	7.4	9.8	9.8	119	358	5	5	10	1254	1254
Los Angeles	060659001	2	4.6	5.9	0.4	0.4	4.6	8.7	8.7	11	9	5	5	11	17	17
Los Angeles	060710001	3	6.9	1.9	5.3	5.3	6.5	9.0	9.0	209	321	10	10	38	579	579
Los Angeles	060710012	0														
Los Angeles	060710014	3	6.0	2.6	3.5	3.5	5.9	8.6	8.6	199	327	6	6	15	577	577
Los Angeles	060710015	3	4.4	4.6	1.7	1.7	1.8	9.7	9.7	752	1045	12	12	296	1948	1948
Los Angeles	060710017	0														
Los Angeles	060710306	3	6.1	2.6	3.6	3.6	5.7	8.9	8.9	199	327	6	6	15	577	577
Los Angeles	060711004	19	7.3	1.7	4.3	4.3	7.4	9.8	9.8	57	120	5	5	18	492	492
Los Angeles	060711234	2	1.6	0.4	1.3	1.3	1.6	1.9	1.9	1122	1168	296	296	1122	1948	1948
Los Angeles	060712002	20	5.7	2.2	2.0	2.0	5.8	9.6	9.6	44	65	5	5	17	250	250
Los Angeles	060714001	1	6.5		6.5	6.5	6.5	6.5	6.5	577		577	577	577	577	577
Los Angeles	060719004	8	5.8	2.5	1.5	1.5	5.7	9.0	9.0	171	438	5	5	10	1254	1254
Los Angeles	061110005	5	6.9	2.5	3.1	3.1	7.7	9.6	9.6	68	118	8	8	19	278	278
Los Angeles	061110007	20	4.7	2.2	1.7	1.7	4.2	9.3	9.3	25	20	5	5	18	76	76
Los Angeles	061111003	0														
Los Angeles	061111004	0														
Los Angeles	061112002	4	6.6	1.0	5.2	5.2	6.8	7.5	7.5	63	113	5	5	7	232	232
Los Angeles	061112003	3	5.5	1.3	4.1	4.1	5.6	6.7	6.7	18	4	14	14	20	22	22
Los Angeles	061113001	7	5.1	2.3	1.9	1.9	5.9	7.4	7.4	35	51	5	5	13	146	146
Miami	120110003	0														
Miami	120110031	0														
Miami	120118002	0														
Miami	120860027	3	4.1	4.2	1.6	1.6	1.8	8.9	8.9	31	19	14	14	27	51	51
Miami	120864002	8	7.0	2.6	1.3	1.3	7.8	9.1	9.1	22	15	8	8	18	51	51
New York	090010113	7	4.4	3.1	1.4	1.4	3.4	8.8	8.8	538	711	48	48	192	1689	1689
New York	090019003	3	6.3	2.0	4.0	4.0	7.4	7.5	7.5	127	179	12	12	37	333	333
New York	090090027	5	2.7	1.0	1.3	1.3	2.7	3.9	3.9	280	484	14	14	86	1144	1144
New York	090091123	6	3.3	2.8	1.2	1.2	2.4	8.9	8.9	234	447	7	7	64	1144	1144
New York	340030001	48	6.5	2.2	2.9	2.9	6.3	9.8	9.9	468	1506	6	7	31	4440	9022
New York	340030005	18	6.8	2.9	0.1	0.1	7.4	10.0	10.0	53	79	6	6	21	307	307
New York	340130011	43	5.4	2.9	0.7	0.8	5.8	9.4	9.5	273	1372	5	5	18	640	9022
New York	340130016	44	5.5	2.8	0.1	1.0	6.3	9.4	9.6	267	1357	5	5	18	640	9022
New York	340131003	32	6.4	2.0	2.1	2.1	6.8	9.3	9.3	77	149	5	5	22	640	640
New York	340170006	42	6.9	2.5	1.1	1.6	7.7	9.5	9.5	369	1420	5	6	24	2213	9022
New York	340210005	8	5.4	1.7	3.2	3.2	5.5	7.3	7.3	115	244	8	8	32	718	718
New York	340230011	20	6.1	2.8	1.0	1.0	7.0	9.5	9.5	95	175	6	6	36	792	792
New York	340273001	1	8.5		8.5	8.5	8.5	8.5	8.5	20		20	20	20	20	20

Table A-2. Distance of location-specific ambient monitors used in Tier 1 analyses to stationary sources emitting > 5 tons of NO_x per year, within a 10 kilometer distance of monitoring site.

		Distance (km) to Source emissions >5 tpy and within 10 km								Emissions (tpy) of Sources within 10 km and >5 tpy						
Location	**ID**	**n**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**
New York	340390004	46	6.3	2.4	0.7	0.9	6.6	9.6	9.7	134	341	5	6	21	594	2213
New York	340390008	12	7.2	2.1	3.2	3.2	8.0	10.0	10.0	23	36	5	5	10	134	134
New York	360050080	54	6.4	2.3	1.8	1.8	6.4	9.9	9.9	241	776	6	6	29	3676	4440
New York	360050083	37	6.0	2.8	1.6	1.6	6.3	9.9	9.9	171	725	6	6	21	4440	4440
New York	360050110	55	5.9	2.2	2.1	2.6	5.7	9.6	9.9	236	769	6	6	29	3676	4440
New York	360470011	56	5.9	2.7	0.7	1.5	5.7	9.7	10.0	296	787	7	7	42	3676	4440
New York	360590005	7	6.3	3.4	1.9	1.9	8.1	9.8	9.8	372	500	7	7	223	1451	1451
New York	360610010	52	5.9	2.5	0.3	1.4	6.1	9.6	9.8	494	1453	5	7	50	4440	9022
New York	360610056	54	5.4	2.6	0.3	1.4	5.5	9.9	10.0	470	1429	7	7	50	4440	9022
New York	360810097	11	6.3	2.1	2.9	2.9	6.9	9.5	9.5	65	77	13	13	26	246	246
New York	360810098	48	7.1	2.3	1.6	2.8	7.8	9.8	9.8	262	820	6	7	31	3676	4440
New York	360810124	24	7.0	2.6	2.1	2.1	8.0	10.0	10.0	436	1136	8	8	26	4440	4440
New York	361030009	3	3.8	3.2	2.0	2.0	2.0	7.6	7.6	537	759	40	40	161	1410	1410
Philadelphia	100031003	39	5.5	2.5	1.6	1.6	6.2	9.7	9.7	282	481	5	5	62	2058	2058
Philadelphia	100031007	11	9.2	0.6	8.0	8.0	9.3	9.8	9.8	323	494	6	6	63	1351	1351
Philadelphia	100032004	32	4.8	1.9	0.7	0.7	4.7	8.4	8.4	223	403	5	5	45	1312	1312
Philadelphia	340070003	69	7.7	2.3	1.8	2.0	8.5	10.0	10.0	87	196	5	5	24	477	1478
Philadelphia	420170012	10	4.1	2.3	1.2	1.2	4.2	9.4	9.4	85	96	11	11	57	275	275
Philadelphia	420450002	30	4.8	2.6	0.2	0.2	5.4	9.5	9.5	504	1055	5	5	73	4968	4968
Philadelphia	420910013	12	5.1	2.5	1.4	1.4	4.3	8.8	8.8	89	232	5	5	12	823	823
Philadelphia	421010004	32	5.9	2.5	1.0	1.0	5.6	9.9	9.9	58	111	5	5	20	571	571
Philadelphia	421010029	74	5.7	2.1	1.1	1.8	5.6	9.7	9.7	74	148	5	5	19	477	1033
Philadelphia	421010047	73	5.2	2.1	0.6	0.8	4.8	9.6	9.7	95	221	5	5	19	1033	1478
Phoenix	040130019	11	6.8	2.2	4.2	4.2	6.7	9.8	9.8	106	313	5	5	10	1049	1049
Phoenix	040133002	6	4.1	2.3	1.3	1.3	4.1	6.9	6.9	21	19	5	5	15	56	56
Phoenix	040133003	10	6.7	1.4	4.1	4.1	6.6	9.0	9.0	50	80	9	9	24	272	272
Phoenix	040133010	10	5.0	0.9	3.5	3.5	4.9	6.6	6.6	115	328	5	5	10	1049	1049
Phoenix	040134005	11	5.8	2.9	0.8	0.8	7.0	9.4	9.4	81	116	6	6	38	350	350
Phoenix	040134011	1	6.4		6.4	6.4	6.4	6.4	6.4	18		18	18	18	18	18
Phoenix	040139997	10	8.5	1.2	5.6	5.6	8.7	9.9	9.9	115	328	5	5	10	1049	1049
Provo	490490002	7	6.6	3.7	1.2	1.2	8.2	9.4	9.4	60	38	7	7	83	102	102
St Louis	171630010	48	7.0	2.8	1.3	1.9	8.0	9.8	9.9	112	178	5	5	17	538	848
St Louis	291830010	1	1.7		1.7	1.7	1.7	1.7	1.7	7821		7821	7821	7821	7821	7821
St Louis	291831002	9	7.5	2.1	4.3	4.3	7.7	9.9	9.9	1868	4704	7	7	8	14231	14231
St Louis	291890001	10	7.7	1.3	6.2	6.2	7.4	9.8	9.8	24	20	5	5	15	60	60
St Louis	291890004	6	8.9	1.5	6.9	6.9	9.8	10.0	10.0	38	37	7	7	28	105	105
St Louis	291890006	8	7.0	1.7	4.2	4.2	7.9	8.7	8.7	25	34	6	6	11	105	105
St Louis	291893001	16	7.3	2.0	3.4	3.4	7.6	9.6	9.6	22	43	5	5	11	181	181

Table A-2. Distance of location-specific ambient monitors used in Tier 1 analyses to stationary sources emitting > 5 tons of NO$_x$ per year, within a 10 kilometer distance of monitoring site.

Location	**ID**	**Distance (km) to Source emissions >5 tpy and within 10 km**								**Emissions (tpy) of Sources within 10 km and >5 tpy**						
		n	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**	**mean**	**std**	**min**	**2.5**	**50**	**97.5**	**max**
St Louis	291895001	11	7.5	1.7	4.3	4.3	7.7	9.7	9.7	46	62	5	5	15	181	181
St Louis	291897002	16	5.7	1.8	2.0	2.0	5.4	9.7	9.7	28	37	5	5	15	143	143
St Louis	291897003	16	6.2	2.0	2.5	2.5	6.0	9.6	9.6	24	33	5	5	15	143	143
St Louis	295100072	46	6.3	2.5	0.7	2.0	6.5	9.9	9.9	77	150	5	5	16	508	848
St Louis	295100080	31	6.9	2.2	0.4	0.4	7.3	10.0	10.0	98	176	5	5	17	848	848
St Louis	295100086	35	6.7	2.3	1.7	1.7	6.6	9.9	9.9	94	168	5	5	17	848	848
Washington DC	110010017	13	5.4	2.4	2.9	2.9	4.5	9.7	9.7	557	1643	11	11	34	6009	6009
Washington DC	110010025	6	6.4	1.0	4.8	4.8	6.5	7.6	7.6	40	35	11	11	26	98	98
Washington DC	110010041	10	6.1	2.4	0.6	0.6	6.1	9.8	9.8	124	137	11	11	66	410	410
Washington DC	110010043	12	5.0	3.2	0.3	0.3	4.6	9.8	9.8	109	129	11	11	46	410	410
Washington DC	240053001	11	7.5	2.1	2.6	2.6	7.9	9.7	9.7	1034	3225	6	6	45	10756	10756
Washington DC	245100040	26	5.0	2.5	0.3	0.3	4.9	9.5	9.5	122	220	6	6	56	1118	1118
Washington DC	245100050	24	6.2	2.1	2.4	2.4	6.0	10.0	10.0	129	227	6	6	56	1118	1118
Washington DC	510130020	14	6.2	2.6	1.5	1.5	5.4	9.8	9.8	558	1579	11	11	46	6009	6009
Washington DC	510590005	2	4.9	4.8	1.4	1.4	4.9	8.3	8.3	13	7	8	8	13	18	18
Washington DC	510590018	6	8.4	0.4	8.0	8.0	8.4	9.2	9.2	1104	2413	9	9	13	6009	6009
Washington DC	510591004	10	7.4	1.6	3.7	3.7	7.8	9.3	9.3	80	173	14	14	19	571	571
Washington DC	510591005	8	6.3	2.0	4.6	4.6	5.5	9.4	9.4	94	193	14	14	19	571	571
Washington DC	510595001	4	6.5	2.8	3.2	3.2	6.8	9.2	9.2	30	19	17	17	22	58	58
Washington DC	511071005	5	7.1	2.3	4.5	4.5	6.5	9.6	9.6	14	8	8	8	12	27	27
Washington DC	511530009	0														
Washington DC	515100009	9	7.0	2.4	1.1	1.1	7.9	8.8	8.8	809	1959	14	14	156	6009	6009

Appendix B. Temporal Air Quality Characterization

Appendix B contains the ambient air quality analysis results by year for each of the named locations. Boxplots were constructed to display the annual average and hourly concentration distributions across years for a single location. The box extends from the 25th to the 75th percentile, with the median shown as the line inside the box. The whiskers extend from the box to the 5th and 95th percentiles. The extreme values in the upper and lower tails beyond the 5th and 95th percentiles are not shown to allow for similar scaling along the y-axis for the plotted independent variables. The mean is plotted as a dot; typically it would appear inside the box, however it will fall outside the box if the distribution is highly skewed. All concentrations are shown in parts per billion (ppb). The boxplots for hourly concentrations were created using a different procedure than for the annual statistics, because of the large number of hourly values and the inability of the graphing procedure to allow frequency weights. Therefore, the appropriate weighted percentiles and means were calculated and plotted as shown, but the vertical lines composing the sides of the box are essentially omitted. Tables are provided that summarize the complete distribution, with percentiles given in segments of 10.

List of Figures

List of Tables

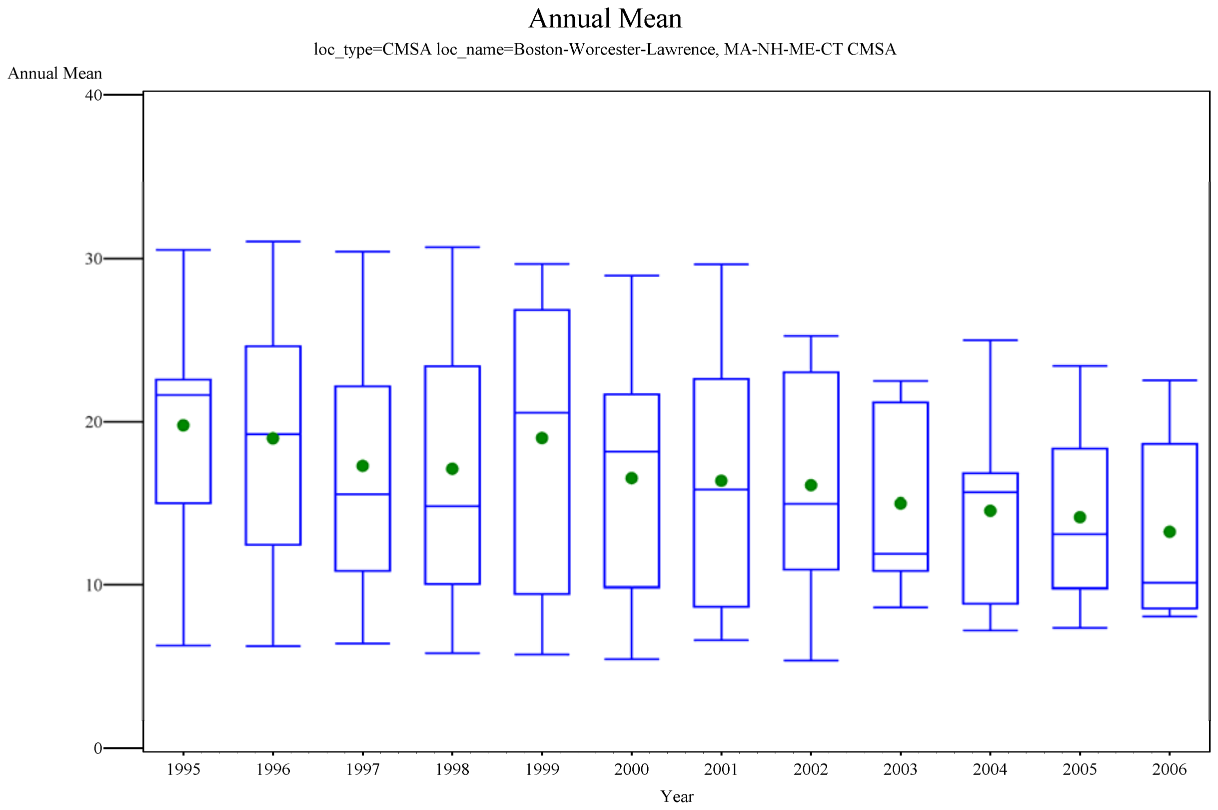


Figure B-1. Temporal distribution of annual average NO_2 ambient concentrations, Boston CMSA, years 1995-2006.

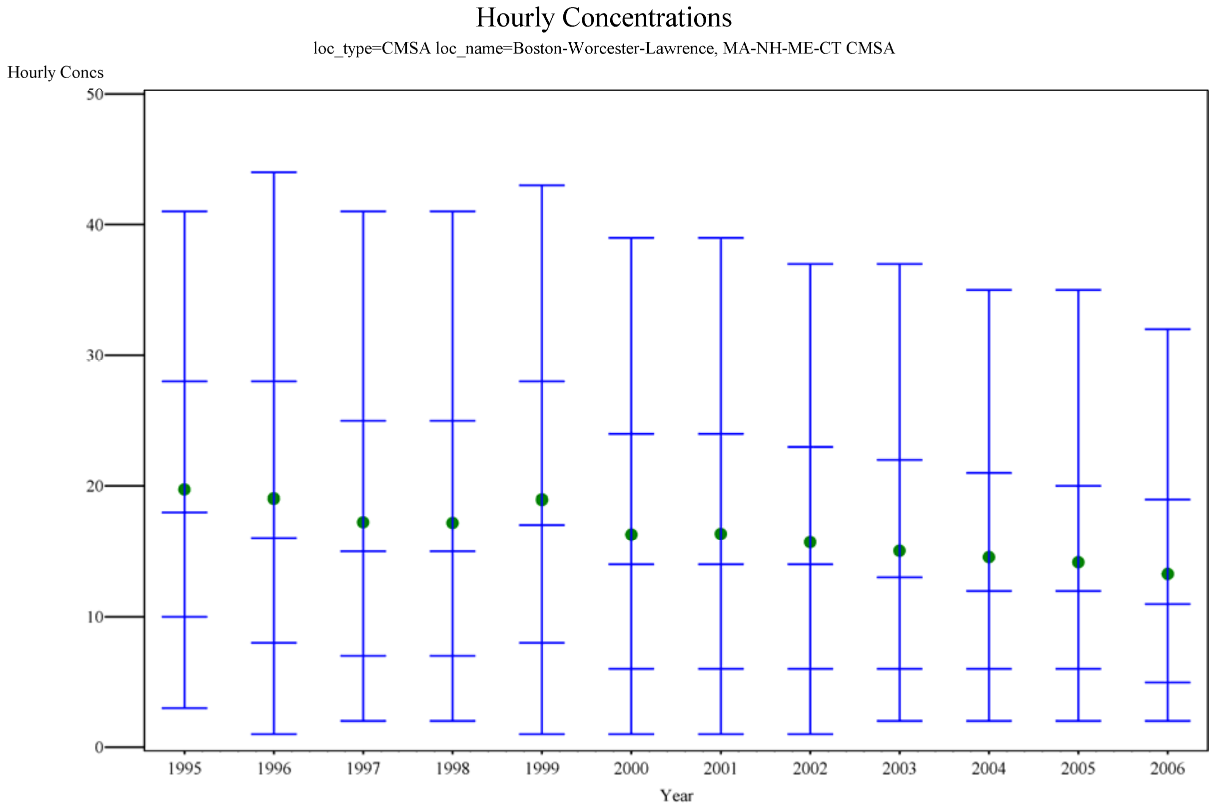


Figure B-2. Temporal distribution of hourly NO_2 ambient concentrations, Boston CMSA, years 1995-2006.

Table B-1. Temporal distribution of annual average NO_2 ambient concentrations, Boston CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	12	20	7	34	6	12	14	16	21	22	22	23	23	27	31
1996	10	19	8	42	6	8	11	14	17	19	21	24	26	29	31
1997	11	17	8	44	6	9	11	13	15	16	19	22	22	27	30
1998	11	17	8	48	6	8	10	12	15	15	19	23	23	28	31
1999	7	19	9	45	6	6	9	20	20	21	21	21	27	30	30
2000	7	17	8	49	5	5	10	11	11	18	20	20	22	29	29
2001	10	16	8	50	7	7	8	10	12	16	20	22	24	28	30
2002	10	16	7	43	5	7	10	12	13	15	19	22	24	25	25
2003	5	15	6	42	9	9	10	11	11	12	17	21	22	22	22
2004	7	15	6	41	7	7	9	12	12	16	16	16	17	25	25
2005	8	14	6	39	7	7	10	10	11	13	15	18	19	23	23
2006	7	13	6	42	8	8	9	10	10	10	15	15	19	23	23

Table B-2. Temporal distribution of hourly NO_2 ambient concentrations, Boston CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	99946	20	12	62	0	5	9	12	15	18	22	26	30	36	100
1996	83541	19	14	72	0	3	7	10	13	16	21	25	30	38	205
1997	90161	17	12	72	0	3	6	9	11	15	18	23	28	35	134
1998	89710	17	13	75	0	3	5	8	11	15	18	23	28	35	112
1999	54043	19	13	70	0	3	7	10	13	17	21	25	30	37	117
2000	56196	16	12	76	0	2	5	7	11	14	18	22	27	34	95
2001	82048	16	13	77	0	2	4	7	10	14	18	22	27	34	114
2002	80472	16	12	75	0	2	5	7	10	14	17	21	26	32	93
2003	41198	15	11	75	0	3	5	7	10	13	16	19	24	31	99
2004	56831	15	10	71	0	3	5	7	10	12	15	19	23	29	96
2005	66244	14	11	75	0	3	5	7	9	12	15	18	23	29	113
2006	57681	13	10	74	0	3	4	6	8	11	14	17	22	28	79

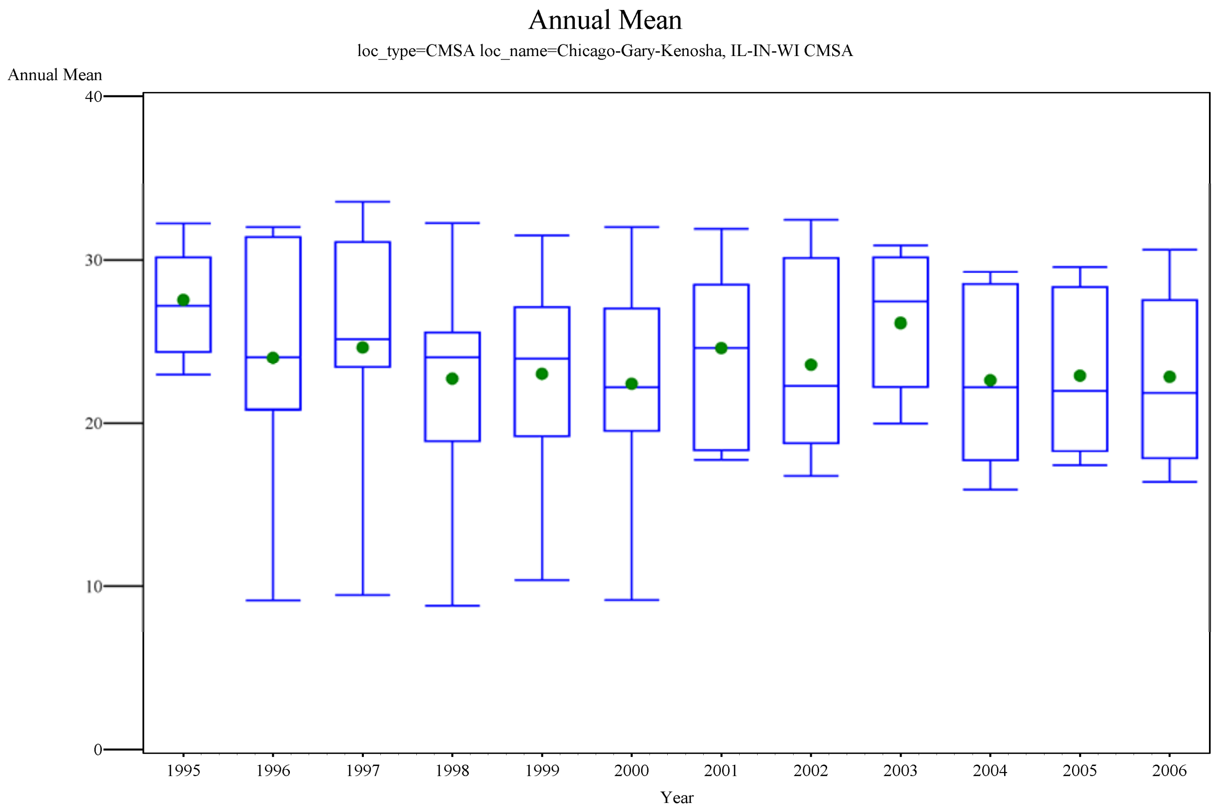


Figure B-3. Temporal distribution of annual average NO_2 ambient concentrations, Chicago CMSA, years 1995-2006.

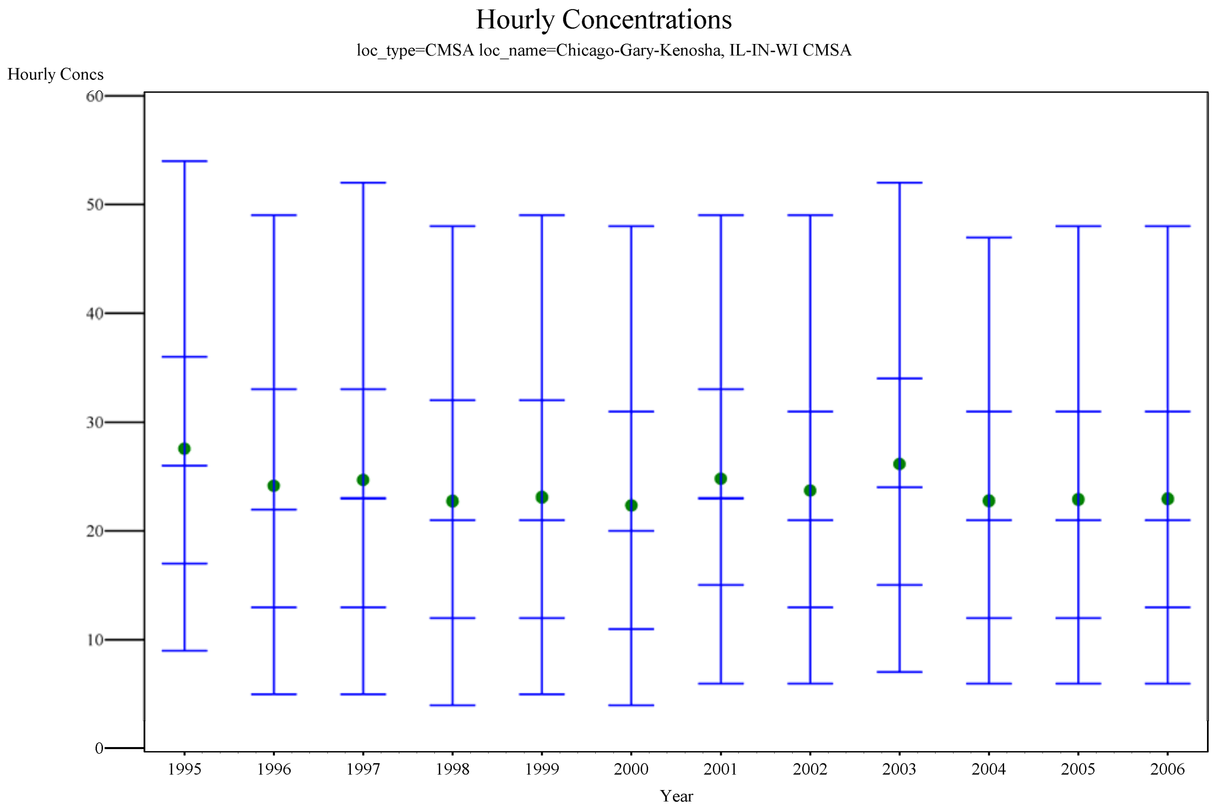


Figure B-4. Temporal distribution of hourly NO_2 ambient concentrations, Chicago CMSA, years 1995-2006.

Table B-3. Temporal distribution of annual average NO_2 ambient concentrations, Chicago CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	7	28	3	12	23	23	24	26	26	27	29	29	30	32	32
1996	7	24	8	32	9	9	21	23	23	24	28	28	31	32	32
1997	6	25	8	34	9	9	23	23	24	25	27	31	31	34	34
1998	9	23	7	32	9	9	17	19	23	24	25	26	31	32	32
1999	9	23	7	29	10	10	17	19	22	24	24	27	31	32	32
2000	9	22	7	30	9	9	18	20	21	22	23	27	29	32	32
2001	7	25	5	21	18	18	18	24	24	25	28	28	28	32	32
2002	7	24	6	24	17	17	19	22	22	22	23	23	30	32	32
2003	5	26	5	19	20	20	21	22	25	27	29	30	31	31	31
2004	6	23	6	25	16	16	18	18	20	22	24	29	29	29	29
2005	6	23	5	23	17	17	18	18	20	22	24	28	28	30	30
2006	5	23	6	27	16	16	17	18	20	22	25	28	29	31	31

Table B-4. Temporal distribution of hourly NO_2 ambient concentrations, Chicago CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	58998	28	14	51	0	11	15	19	22	26	29	33	38	47	113
1996	59447	24	14	58	0	7	11	15	18	22	26	31	36	43	127
1997	51443	25	15	59	0	7	11	15	19	23	27	31	36	44	113
1998	76365	23	14	61	0	6	10	13	17	21	25	29	34	41	112
1999	74985	23	14	61	0	7	10	13	17	21	25	30	35	42	113
2000	75327	22	14	62	0	6	10	13	17	20	24	29	34	41	108
2001	58268	25	13	54	0	9	13	16	20	23	27	31	36	43	114
2002	58383	24	14	59	0	8	12	15	18	21	25	29	34	42	149
2003	42406	26	14	54	0	10	14	17	21	24	28	32	37	45	122
2004	49210	23	13	57	0	8	11	14	18	21	25	28	33	41	101
2005	51043	23	13	59	0	8	11	14	17	21	24	29	34	41	106
2006	42009	23	13	57	0	8	11	14	17	21	25	29	34	41	137

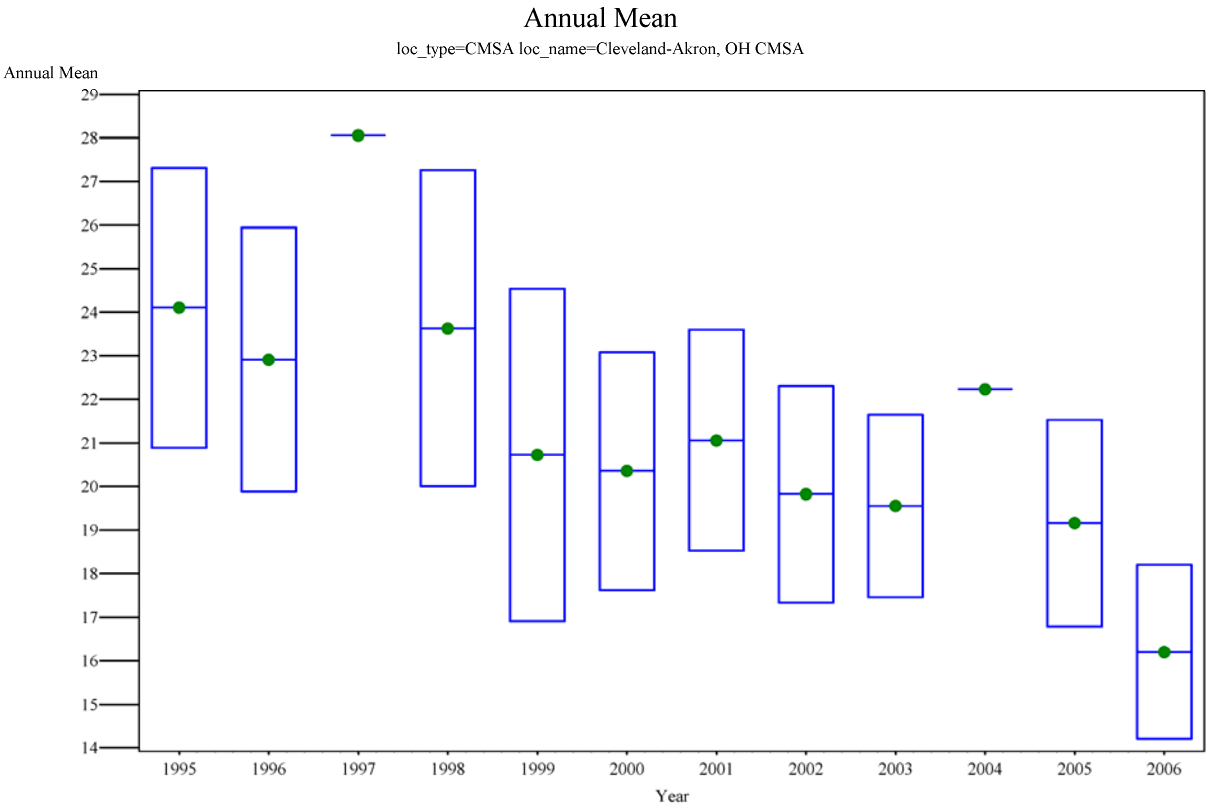


Figure B-5. Temporal distribution of annual average NO_2 ambient concentrations, Cleveland CMSA, years 1995-2006.

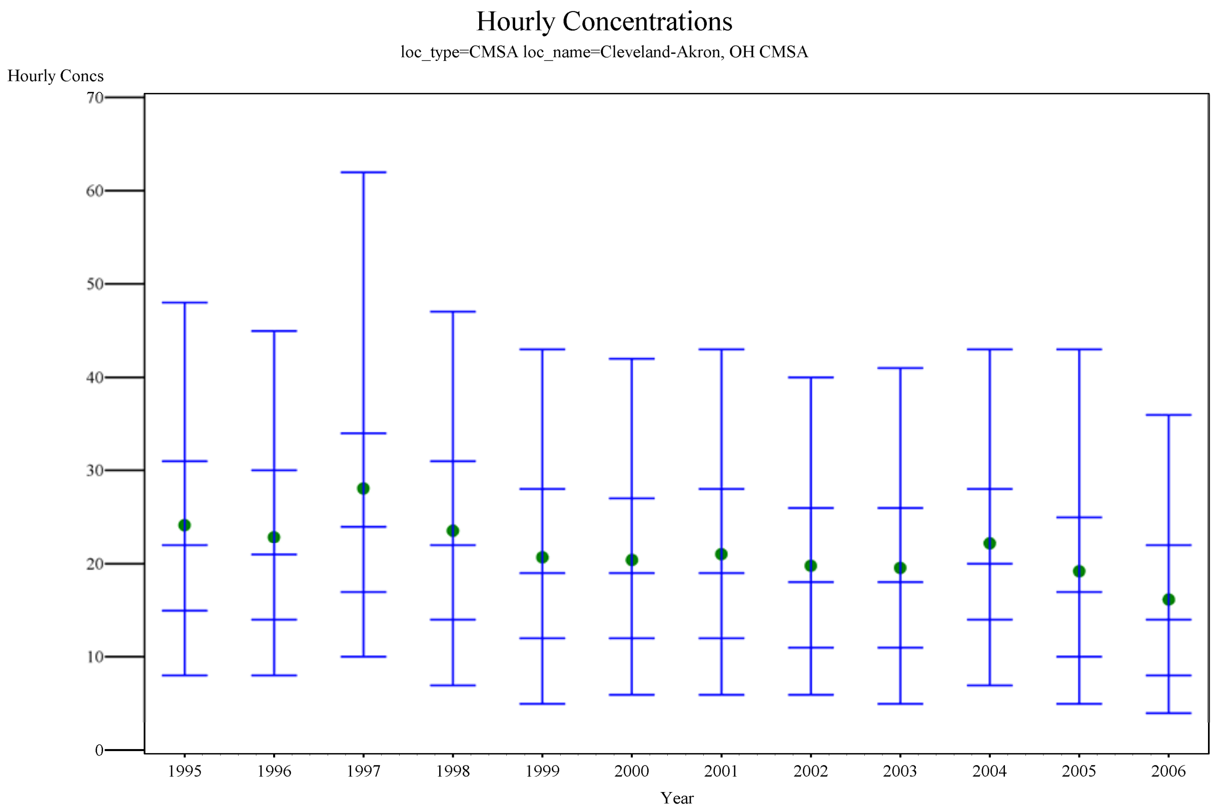


Figure B-6. Temporal distribution of hourly NO_2 ambient concentrations, Cleveland CMSA, years 1995-2006.

Table B-5. Temporal distribution of annual average NO_2 ambient concentrations, Cleveland CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	2	24	5	19	21	21	21	21	21	24	27	27	27	27	27
1996	2	23	4	19	20	20	20	20	20	23	26	26	26	26	26
1997	1	28		0	28	28	28	28	28	28	28	28	28	28	28
1998	2	24	5	22	20	20	20	20	20	24	27	27	27	27	27
1999	2	21	5	26	17	17	17	17	17	21	25	25	25	25	25
2000	2	20	4	19	18	18	18	18	18	20	23	23	23	23	23
2001	2	21	4	17	19	19	19	19	19	21	24	24	24	24	24
2002	2	20	4	18	17	17	17	17	17	20	22	22	22	22	22
2003	2	20	3	15	17	17	17	17	17	20	22	22	22	22	22
2004	1	22		0	22	22	22	22	22	22	22	22	22	22	22
2005	2	19	3	17	17	17	17	17	17	19	22	22	22	22	22
2006	2	16	3	17	14	14	14	14	14	16	18	18	18	18	18

Table B-6. Temporal distribution of hourly NO_2 ambient concentrations, Cleveland CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	16042	24	13	53	2	10	13	16	19	22	25	29	34	41	108
1996	16593	23	12	52	1	9	13	15	18	21	24	28	32	39	148
1997	8300	28	17	59	0	12	15	18	21	24	28	32	38	49	253
1998	16680	24	13	53	0	9	13	16	19	22	25	29	33	40	89
1999	16743	21	12	58	0	7	10	13	16	19	22	26	30	37	86
2000	16399	20	11	55	0	8	10	13	16	19	22	25	30	36	74

2001	16566	21	12	56	0	8	10	13	16	19	22	26	30	37	103
2002	16464	20	11	56	1	8	10	12	15	18	21	24	28	35	88
2003	16948	20	11	57	0	7	10	13	15	18	20	24	28	35	90
2004	8484	22	11	51	0	10	13	15	18	20	23	26	30	37	83
2005	16558	19	12	60	0	7	9	12	14	17	20	23	28	35	85
2006	16853	16	10	64	0	5	8	10	12	14	16	20	24	30	175

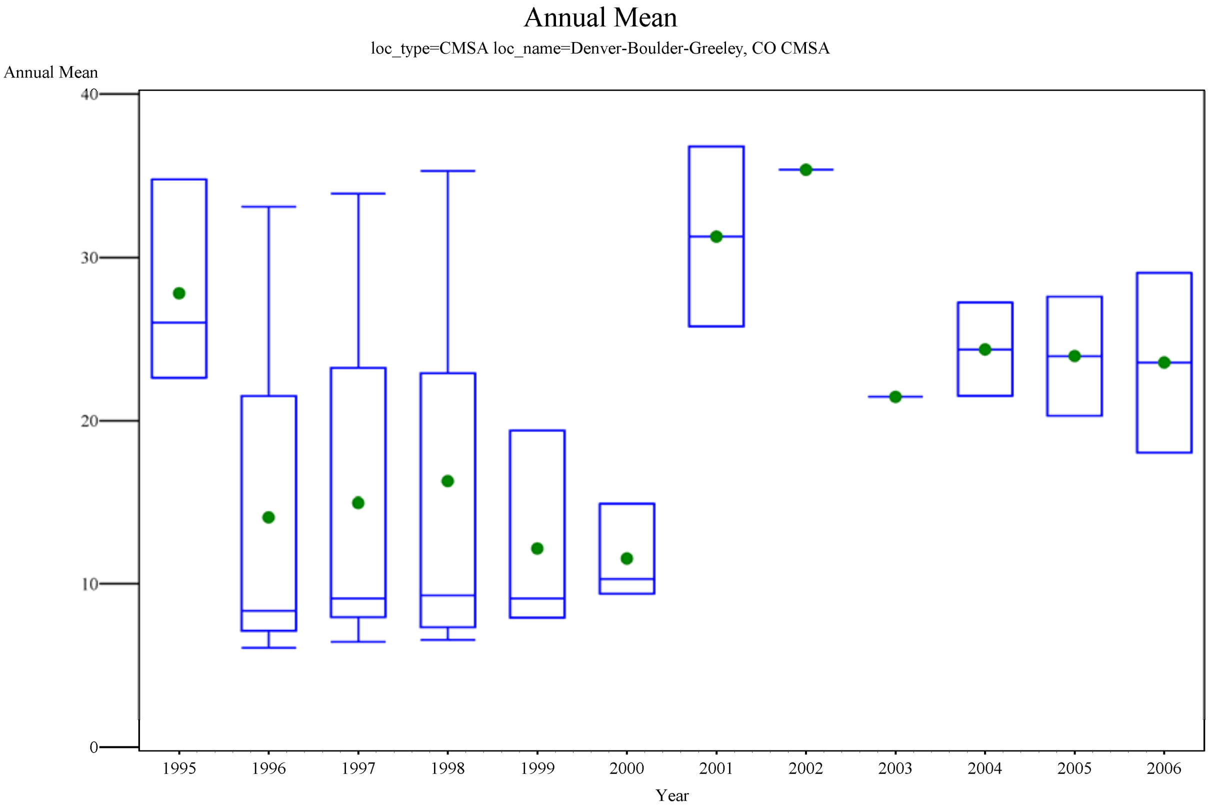


Figure B-7. Temporal distribution of annual average NO_2 ambient concentrations, Denver CMSA, years 1995-2006.

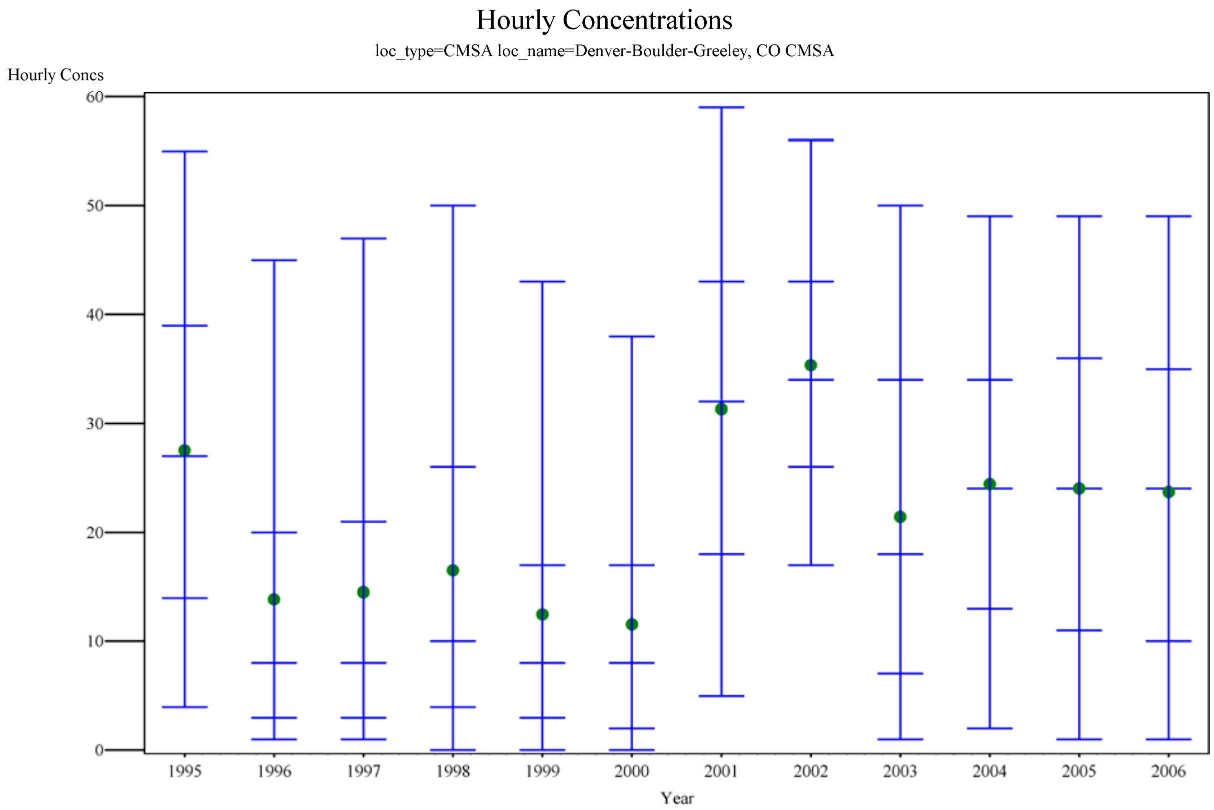


Figure B-8. Temporal distribution of hourly NO_2 ambient concentrations, Denver CMSA, years 1995-2006.

Table B-7. Temporal distribution of annual average NO_2 ambient concentrations, Denver CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	3	28	6	23	23	23	23	23	26	26	26	35	35	35	35
1996	6	14	11	77	6	6	7	7	8	8	9	22	22	33	33
1997	6	15	11	74	6	6	8	8	9	9	9	23	23	34	34
1998	5	16	13	77	7	7	7	7	8	9	16	23	29	35	35
1999	3	12	6	52	8	8	8	8	9	9	9	19	19	19	19
2000	3	12	3	26	9	9	9	9	10	10	10	15	15	15	15
2001	2	31	8	25	26	26	26	26	26	31	37	37	37	37	37
2002	1	35		0	35	35	35	35	35	35	35	35	35	35	35
2003	1	21		0	21	21	21	21	21	21	21	21	21	21	21
2004	2	24	4	17	21	21	21	21	21	24	27	27	27	27	27
2005	2	24	5	21	20	20	20	20	20	24	28	28	28	28	28
2006	2	24	8	33	18	18	18	18	18	24	29	29	29	29	29

Table B-8. Temporal distribution of hourly NO_2 ambient concentrations, Denver CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	23204	28	17	62	0	6	11	16	22	27	32	36	41	48	286
1996	46816	14	15	108	0	1	2	4	6	8	11	16	25	37	137
1997	45049	15	15	106	0	1	3	4	6	8	12	17	26	39	141
1998	40258	17	17	100	0	1	3	5	7	10	15	22	31	42	148
1999	23164	12	13	108	0	0	2	4	6	8	10	14	21	33	96
2000	24649	12	13	108	0	0	1	3	5	8	10	14	19	30	141
2001	15204	31	17	55	0	8	15	21	27	32	36	41	45	52	157
2002	7688	35	13	36	0	20	24	28	31	34	38	41	45	51	159
2003	6989	21	17	78	0	3	5	8	13	18	25	31	37	44	136
2004	15878	24	15	60	0	4	10	16	20	24	28	32	37	43	115
2005	15467	24	16	65	0	3	8	14	19	24	29	33	38	44	114
2006	13775	24	15	65	0	3	7	13	19	24	28	33	38	44	169

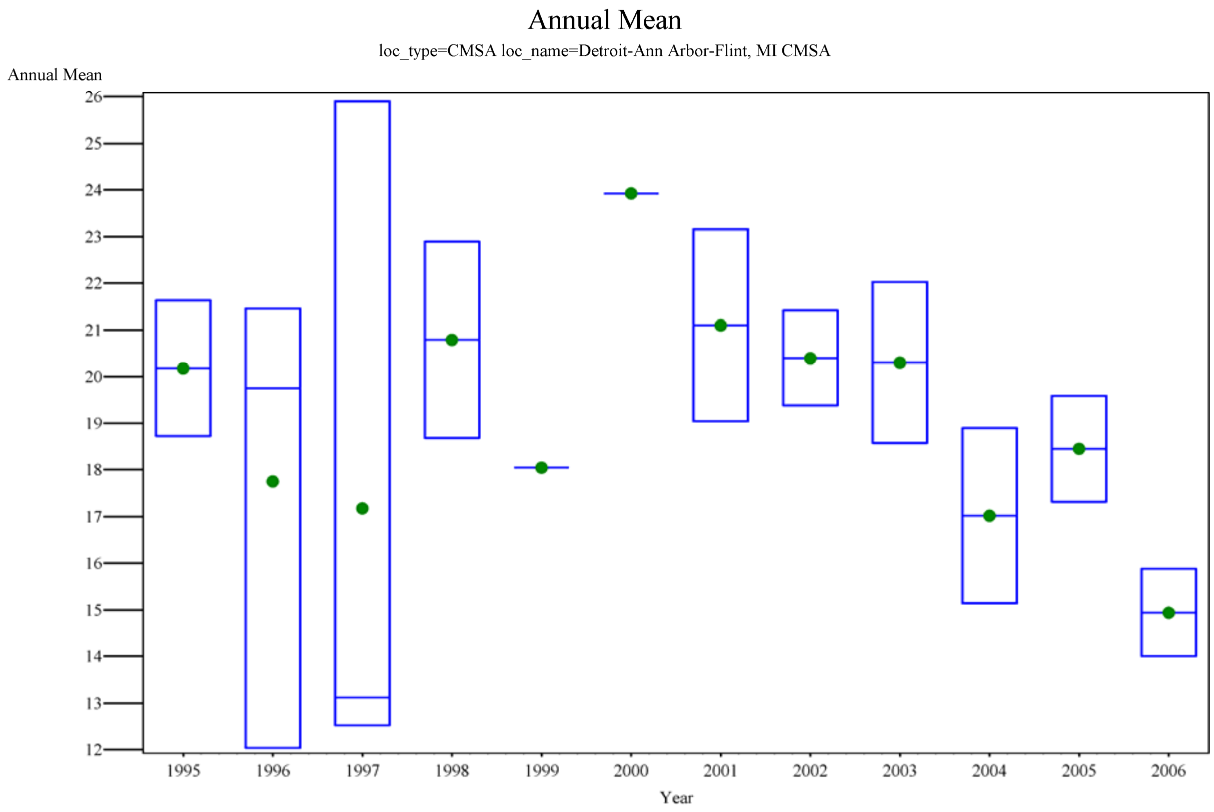


Figure B-9. Temporal distribution of annual average NO_2 ambient concentrations, Detroit CMSA, years 1995-2006.

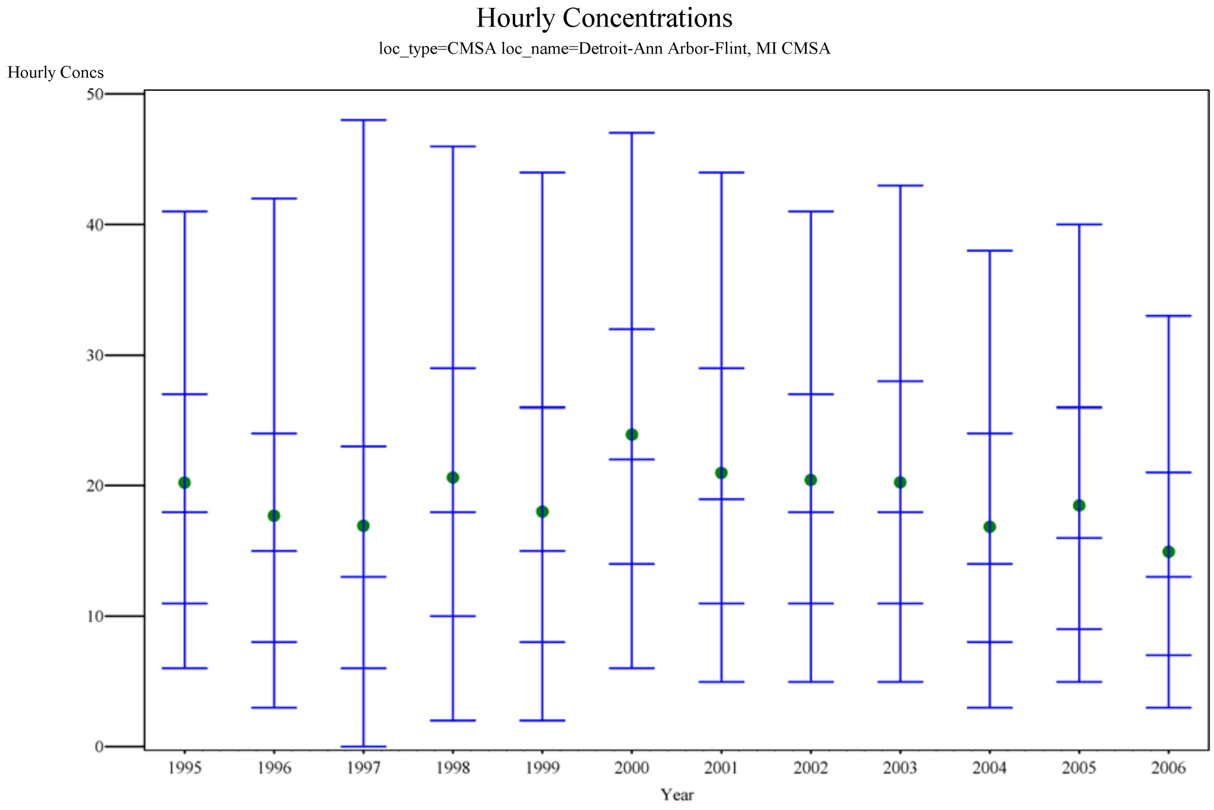


Figure B-10. Temporal distribution of hourly NO_2 ambient concentrations, Detroit CMSA, years 1995-2006.

Table B-9. Temporal distribution of annual average NO_2 ambient concentrations, Detroit CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	2	20	2	10	19	19	19	19	19	20	22	22	22	22	22
1996	3	18	5	28	12	12	12	12	20	20	20	21	21	21	21
1997	3	17	8	44	13	13	13	13	13	13	13	26	26	26	26
1998	2	21	3	14	19	19	19	19	19	21	23	23	23	23	23
1999	1	18		0	18	18	18	18	18	18	18	18	18	18	18
2000	1	24		0	24	24	24	24	24	24	24	24	24	24	24
2001	2	21	3	14	19	19	19	19	19	21	23	23	23	23	23
2002	2	20	1	7	19	19	19	19	19	20	21	21	21	21	21
2003	2	20	2	12	19	19	19	19	19	20	22	22	22	22	22
2004	2	17	3	16	15	15	15	15	15	17	19	19	19	19	19
2005	2	18	2	9	17	17	17	17	17	18	20	20	20	20	20
2006	2	15	1	9	14	14	14	14	14	15	16	16	16	16	16

Table B-10. Temporal distribution of hourly NO_2 ambient concentrations, Detroit CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	16629	20	12	58	0	8	10	12	15	18	21	25	29	35	117
1996	23600	18	13	74	0	4	7	9	12	15	18	22	27	35	167
1997	24117	17	16	94	0	2	5	7	10	13	16	21	26	36	322
1998	14863	21	14	68	0	5	9	12	15	18	22	27	31	39	136
1999	7110	18	13	73	0	4	7	9	12	15	19	24	29	36	104
2000	8590	24	13	56	0	8	12	15	19	22	26	30	35	42	128
2001	15154	21	13	61	0	7	9	12	15	19	23	27	32	38	194
2002	16623	20	15	73	0	7	10	12	15	18	22	25	30	36	443
2003	16569	20	13	62	0	7	9	12	15	18	21	25	30	36	139
2004	14779	17	11	66	0	5	7	9	12	14	17	21	26	33	78
2005	15827	19	12	63	0	6	8	10	13	16	19	23	28	35	84
2006	17273	15	10	64	0	4	6	8	10	13	16	19	23	29	58

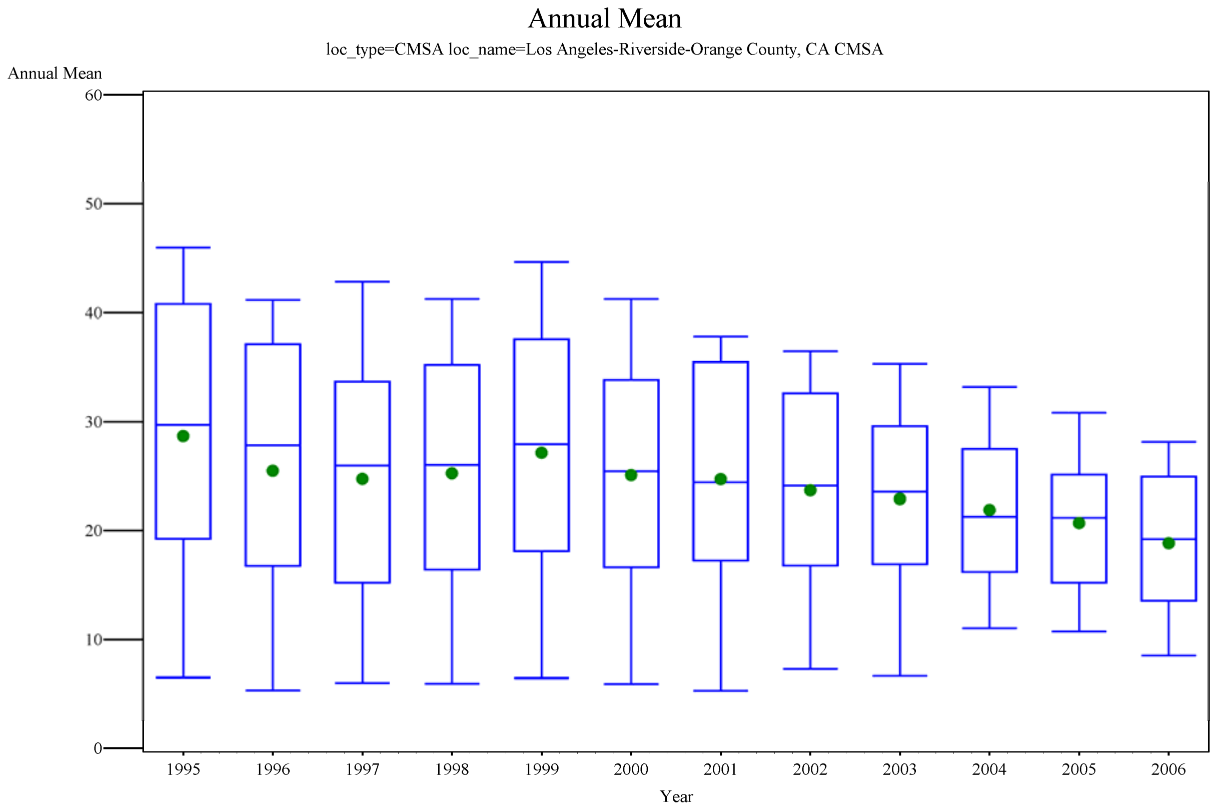


Figure B-11. Temporal distribution of annual average NO_2 ambient concentrations, Los Angeles CMSA, years 1995-2006.

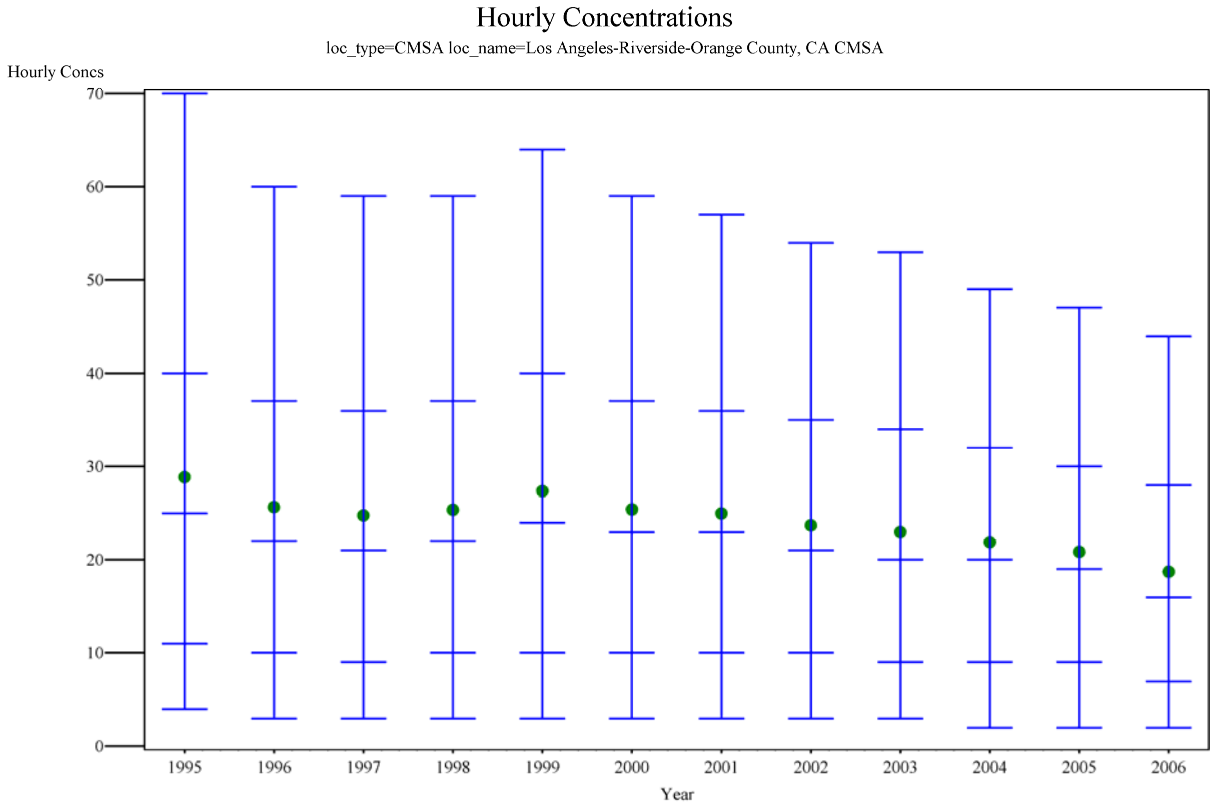


Figure B-12. Temporal distribution of hourly NO_2 ambient concentrations, Los Angeles CMSA, years 1995-2006.

Table B-11. Temporal distribution of annual average NO_2 ambient concentrations, Los Angeles CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	36	29	13	47	5	8	18	20	23	30	37	39	45	46	46
1996	29	25	12	46	4	6	15	17	21	28	31	35	38	41	42
1997	33	25	12	47	4	8	14	16	20	26	29	33	34	42	43
1998	32	25	11	44	4	9	16	19	21	26	33	34	36	39	43
1999	31	27	12	44	5	10	18	20	23	28	32	35	39	39	51
2000	32	25	11	43	4	10	16	20	22	25	28	32	36	39	44
2001	31	25	11	43	4	9	17	19	24	24	27	33	36	37	41
2002	32	24	9	39	5	10	16	18	22	24	25	29	33	36	40
2003	32	23	9	37	5	11	15	18	21	24	26	29	31	34	35
2004	28	22	7	33	5	13	15	17	20	21	24	27	30	31	34
2005	28	21	7	34	5	12	14	16	19	21	22	25	27	31	31
2006	26	19	7	35	5	9	13	15	17	19	20	23	25	27	30

Table B-12. Temporal distribution of hourly NO_2 ambient concentrations, Los Angeles CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	290519	29	22	78	0	6	9	14	19	25	30	37	45	57	239
1996	232203	26	19	74	0	5	8	12	17	22	28	34	40	50	250
1997	263050	25	19	75	0	4	7	11	16	21	27	33	40	50	200
1998	257541	25	19	74	0	5	8	12	17	22	28	34	40	50	255
1999	253401	27	20	73	0	5	8	13	18	24	30	37	43	54	307
2000	263311	25	18	72	0	5	8	12	17	23	28	34	40	50	214
2001	251895	25	18	71	0	5	8	12	17	23	28	33	39	48	251
2002	258452	24	17	71	0	5	8	11	16	21	26	32	38	46	262
2003	259935	23	17	72	0	4	7	11	15	20	25	31	37	45	163
2004	225075	22	15	70	0	4	7	11	15	20	25	29	35	42	157
2005	227769	21	14	69	0	4	7	11	15	19	23	28	33	40	136
2006	184205	19	14	74	0	3	6	9	12	16	20	25	31	38	107

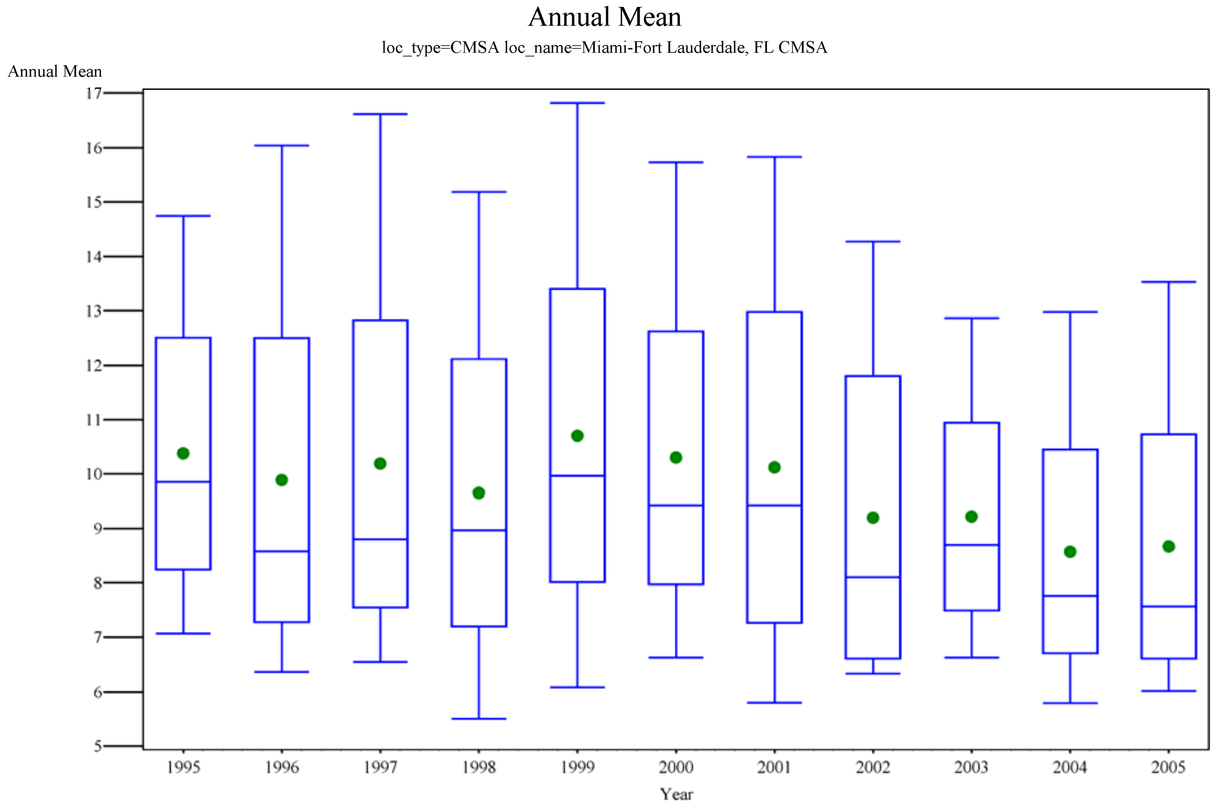


Figure B-13. Temporal distribution of annual average NO_2 ambient concentrations, Miami CMSA, years 1995-2006.

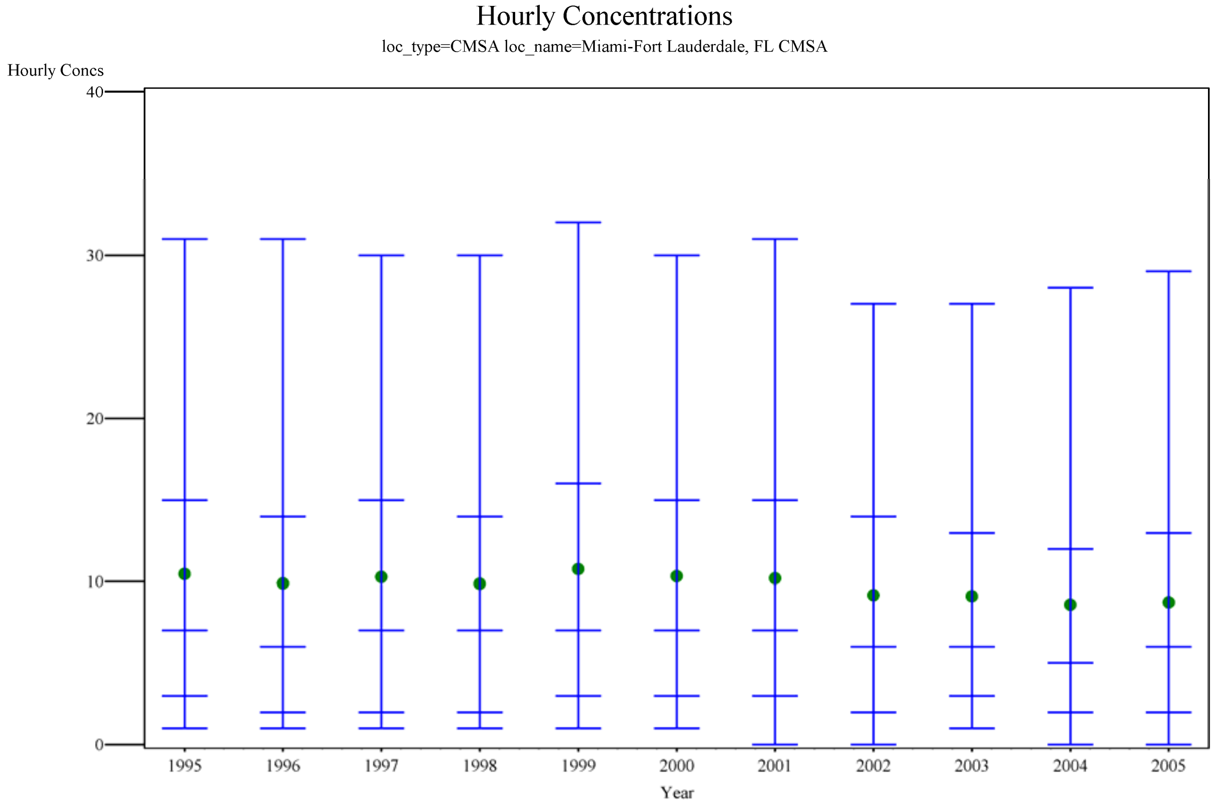


Figure B-14. Temporal distribution of hourly NO_2 ambient concentrations, Miami CMSA, years 1995-2006.

Table B-13. Temporal distribution of annual average NO_2 ambient concentrations, Miami CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	4	10	3	31	7	7	7	9	9	10	10	10	15	15	15
1996	4	10	4	43	6	6	6	8	8	9	9	9	16	16	16
1997	4	10	4	43	7	7	7	9	9	9	9	9	17	17	17
1998	4	10	4	42	6	6	6	9	9	9	9	9	15	15	15
1999	4	11	4	42	6	6	6	10	10	10	10	10	17	17	17
2000	4	10	4	37	7	7	7	9	9	9	10	10	16	16	16
2001	4	10	4	42	6	6	6	9	9	9	10	10	16	16	16
2002	4	9	4	39	6	6	6	7	7	8	9	9	14	14	14
2003	4	9	3	29	7	7	7	8	8	9	9	9	13	13	13
2004	4	9	3	36	6	6	6	8	8	8	8	8	13	13	13
2005	4	9	3	38	6	6	6	7	7	8	8	8	14	14	14

Table B-14. Temporal distribution of hourly NO_2 ambient concentrations, Miami CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	32713	10	10	95	0	1	2	3	5	7	10	13	18	25	75
1996	33086	10	10	103	0	1	2	3	4	6	9	12	17	25	96
1997	32754	10	10	97	0	1	2	3	5	7	10	13	18	25	94
1998	30849	10	10	98	0	1	2	3	5	7	10	12	16	23	69
1999	32721	11	11	99	0	1	2	3	5	7	10	14	18	26	128
2000	31833	10	10	99	0	1	2	4	5	7	10	13	17	24	203
2001	33063	10	10	98	0	1	2	3	5	7	10	13	17	24	86
2002	33755	9	9	96	0	1	2	3	4	6	9	12	16	22	80
2003	31031	9	9	97	0	1	2	3	4	6	8	11	15	21	85
2004	33625	9	10	117	0	1	2	2	4	5	7	10	14	21	417
2005	32342	9	10	109	0	0	1	2	4	6	8	11	15	22	94

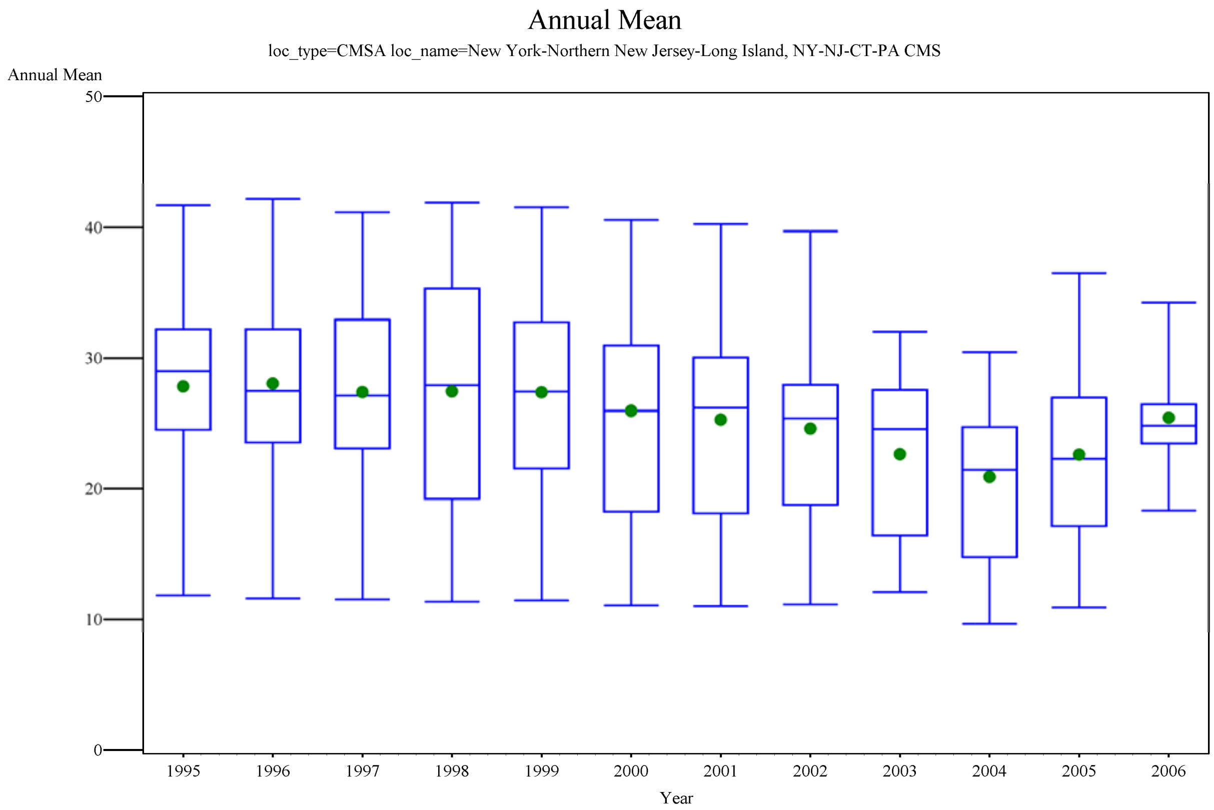


Figure B-15. Temporal distribution of annual average NO_2 ambient concentrations, New York CMSA, years 1995-2006.

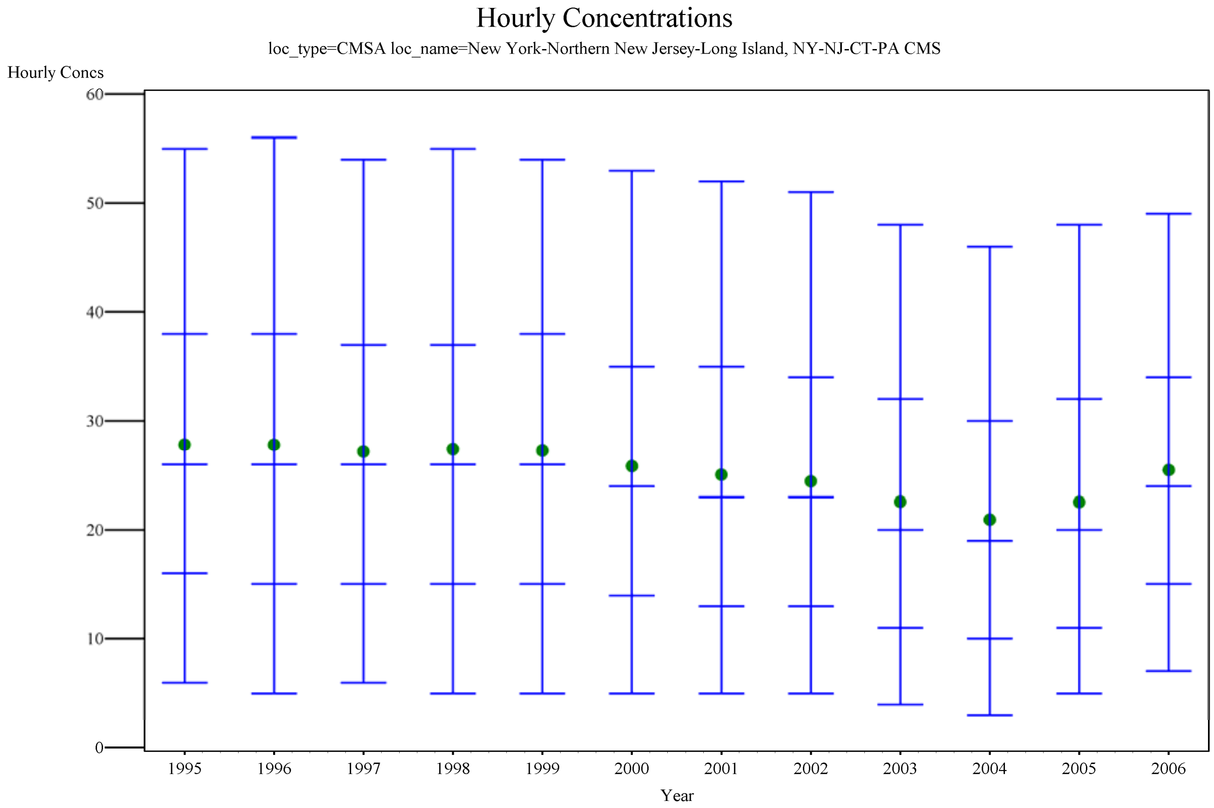


Figure B-16. Temporal distribution of hourly NO_2 ambient concentrations, New York CMSA, years 1995-2006.

Table B-15. Temporal distribution of annual average NO_2 ambient concentrations, New York CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	16	28	8	28	12	16	24	25	26	29	30	31	33	39	42
1996	15	28	8	29	12	17	22	26	27	27	29	32	34	41	42
1997	16	27	8	30	12	17	23	24	26	27	29	31	35	40	41
1998	14	27	9	34	11	15	18	22	27	28	30	33	36	40	42
1999	16	27	9	31	11	17	19	24	26	27	29	33	33	41	42
2000	16	26	8	32	11	16	18	19	25	26	29	30	32	38	41
2001	14	25	8	32	11	17	17	21	24	26	27	27	31	38	40
2002	17	25	8	31	11	16	17	20	22	25	28	28	29	38	40
2003	15	23	6	28	12	14	16	18	21	25	26	27	29	30	32
2004	14	21	7	31	10	13	14	17	20	21	24	24	28	30	30
2005	16	23	7	31	11	13	16	18	22	22	25	27	27	32	36
2006	5	25	6	23	18	18	21	23	24	25	26	26	30	34	34

Table B-16. Temporal distribution of hourly NO_2 ambient concentrations, New York CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	133504	28	16	56	0	9	14	18	22	26	31	35	40	48	162
1996	122074	28	16	57	0	8	13	18	22	26	31	35	40	48	162
1997	131144	27	15	56	0	9	13	17	22	26	30	35	40	47	181
1998	116748	27	16	58	0	8	13	17	22	26	31	35	40	48	240
1999	132646	27	16	57	0	8	13	17	22	26	30	35	40	48	148
2000	134037	26	15	58	0	8	12	16	20	24	28	33	38	46	118
2001	114478	25	15	61	0	7	10	15	19	23	28	33	38	45	142
2002	141480	24	15	60	0	7	11	14	18	23	27	32	37	44	129
2003	122724	23	14	61	0	6	10	13	16	20	25	29	35	42	138
2004	115578	21	13	64	0	5	8	12	15	19	23	27	32	40	156
2005	133856	23	14	63	1	6	9	13	16	20	24	29	35	42	119
2006	42223	25	13	51	0	10	13	17	20	24	28	32	37	43	92

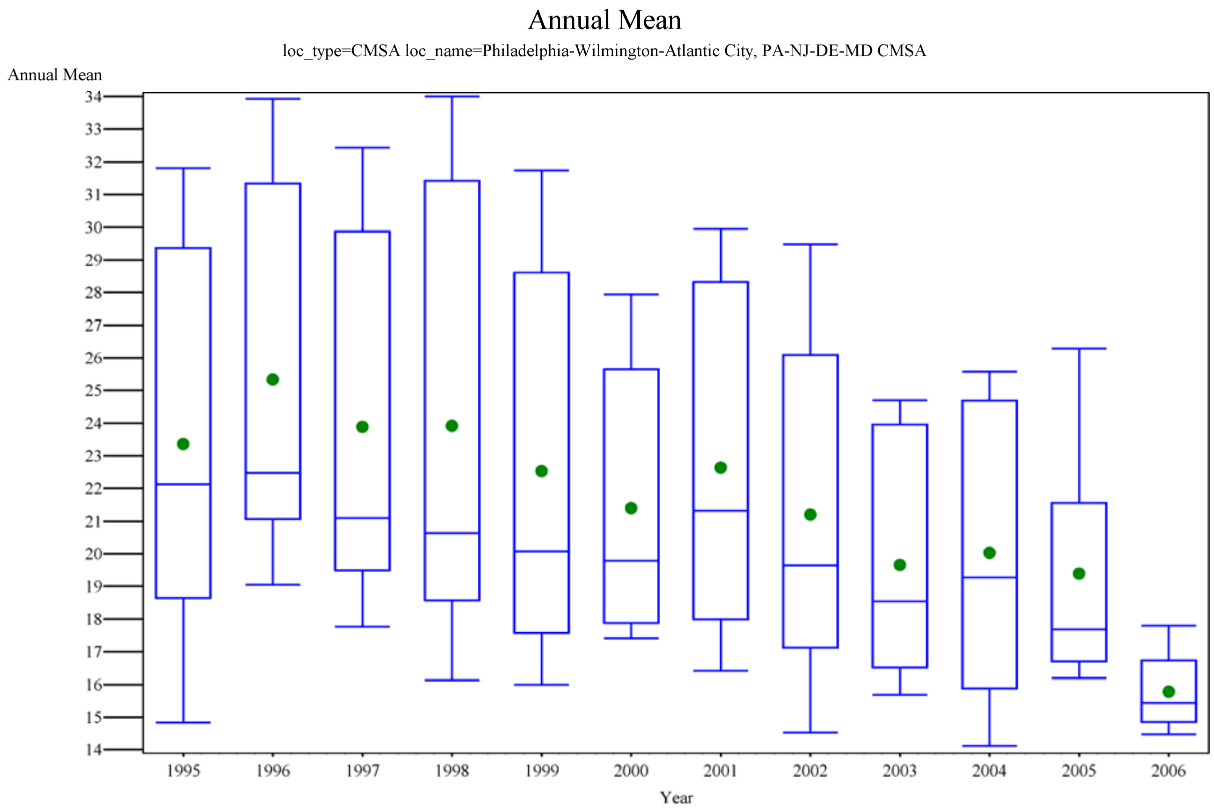


Figure B-17. Temporal distribution of annual average NO_2 ambient concentrations, Philadelphia CMSA, years 1995-2006.

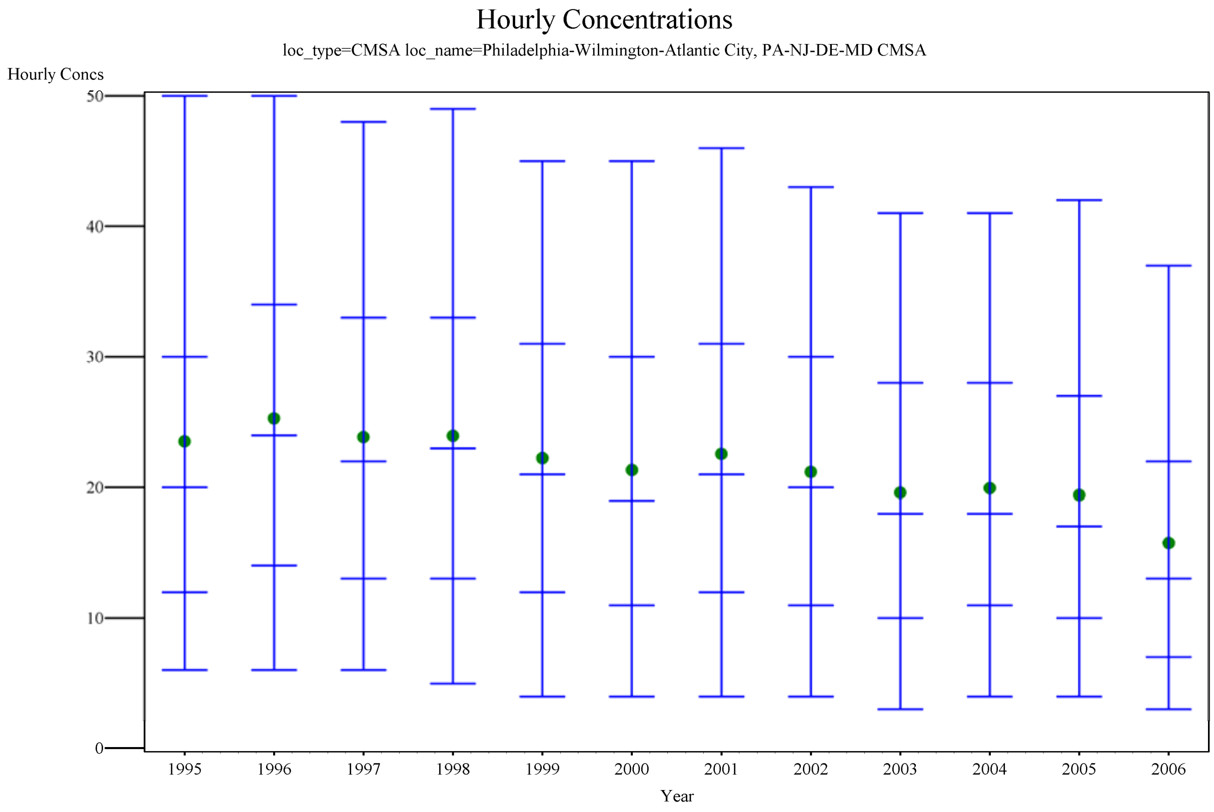


Figure B-18. Temporal distribution of hourly NO_2 ambient concentrations, Philadelphia CMSA, years 1995-2006.

Table B-17. Temporal distribution of annual average NO_2 ambient concentrations, Philadelphia CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	8	23	6	27	15	15	17	20	20	22	24	28	31	32	32
1996	8	25	6	24	19	19	21	21	21	22	24	29	33	34	34
1997	8	24	6	25	18	18	19	20	20	21	22	28	32	32	32
1998	8	24	7	30	16	16	18	19	19	21	22	29	33	34	34
1999	8	23	6	28	16	16	17	18	18	20	22	27	30	32	32
2000	6	21	4	20	17	17	18	18	19	20	20	26	26	28	28
2001	7	23	5	24	16	16	18	19	19	21	26	26	28	30	30
2002	8	21	5	26	15	15	16	18	19	20	20	24	28	29	29
2003	6	20	4	19	16	16	17	17	18	19	19	24	24	25	25
2004	7	20	4	22	14	14	16	18	18	19	23	23	25	26	26
2005	7	19	4	19	16	16	17	17	17	18	20	20	22	26	26
2006	4	16	1	9	14	14	14	15	15	15	16	16	18	18	18

Table B-18. Temporal distribution of hourly NO_2 ambient concentrations, Philadelphia CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	65415	24	14	60	0	8	10	14	19	20	26	30	35	40	140
1996	67989	25	14	55	0	8	11	17	20	24	30	30	40	42	100
1997	68291	24	14	57	0	8	11	15	19	22	26	30	35	42	247
1998	66847	24	14	58	0	7	11	15	19	23	27	31	36	42	97
1999	64813	22	13	59	0	6	10	14	17	21	25	29	33	40	109
2000	51145	21	13	60	0	6	10	13	16	19	23	27	32	39	97
2001	59227	23	13	59	0	6	10	14	17	21	25	29	34	40	96
2002	66779	21	12	59	0	6	10	13	16	20	23	27	32	38	268
2003	49256	20	12	62	0	5	8	11	15	18	22	26	30	36	105
2004	58509	20	12	59	0	6	9	12	15	18	22	26	30	36	101
2005	56459	19	12	62	0	6	9	11	14	17	21	25	29	36	120
2006	32357	16	11	69	0	4	6	8	10	13	16	20	25	31	95

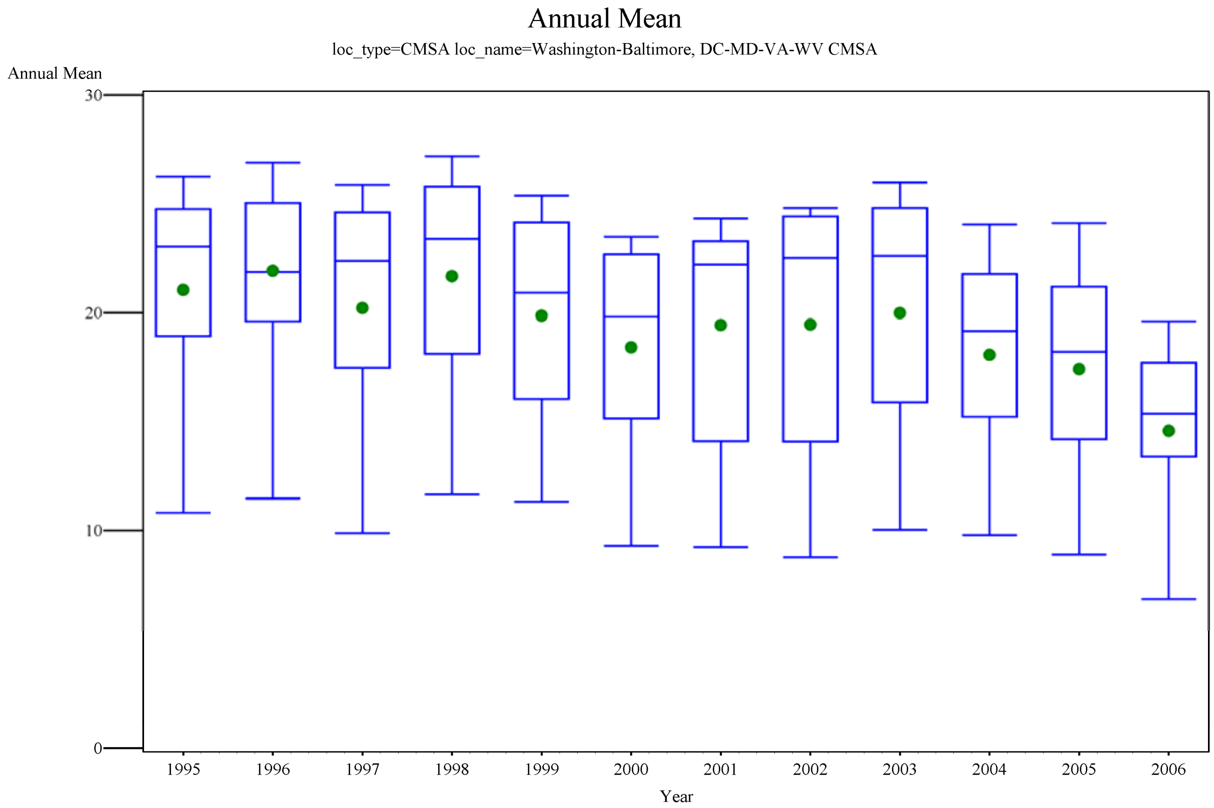


Figure B-19. Temporal distribution of annual average NO_2 ambient concentrations, Washington DC CMSA, years 1995-2006.

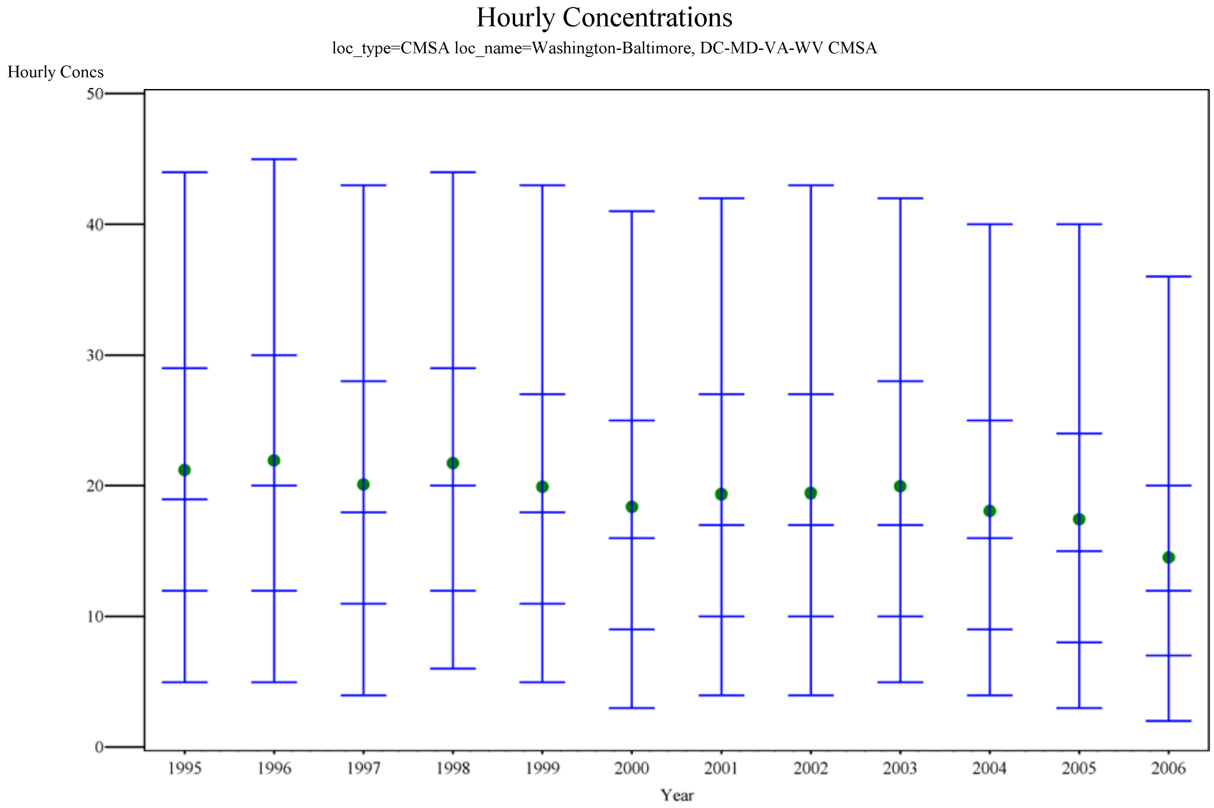


Figure B-20. Temporal distribution of hourly NO_2 ambient concentrations, Washington DC CMSA, years 1995-2006.

Table B-19. Temporal distribution of annual average NO_2 ambient concentrations, Washington DC CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	12	21	5	25	11	11	19	19	22	23	23	25	25	26	26
1996	11	22	4	20	11	20	20	21	22	22	24	24	25	26	27
1997	11	20	5	27	10	11	17	19	21	22	22	24	25	26	26
1998	11	22	5	23	12	15	18	20	22	23	24	25	26	26	27
1999	12	20	5	25	11	12	14	18	20	21	23	24	24	25	25
2000	12	18	5	27	9	10	13	17	18	20	21	23	23	23	23
2001	11	19	5	28	9	11	14	19	20	22	23	23	23	24	24
2002	10	19	6	31	9	10	13	16	20	23	23	24	25	25	25
2003	11	20	6	28	10	12	16	18	18	23	23	23	25	26	26
2004	12	18	5	27	10	10	15	15	17	19	21	21	22	23	24
2005	12	17	5	28	9	10	14	15	17	18	21	21	21	22	24
2006	10	15	4	30	7	7	10	14	15	15	16	17	18	19	20

Table B-20. Temporal distribution of hourly NO_2 ambient concentrations, Washington DC CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	98349	21	13	59	0	7	10	13	16	19	23	27	31	38	145
1996	91551	22	12	57	0	7	11	14	17	20	24	28	32	39	107
1997	87646	20	12	62	0	6	9	12	15	18	21	25	30	37	155
1998	89335	22	12	57	0	8	11	14	16	20	23	27	32	38	285
1999	100112	20	12	61	0	6	9	12	15	18	21	25	30	37	114
2000	101494	18	12	64	0	5	8	11	13	16	19	23	28	35	141
2001	91594	19	12	62	0	6	9	11	14	17	20	24	29	36	89
2002	83969	19	12	64	0	6	9	11	14	17	20	24	30	37	108
2003	93111	20	12	61	0	6	9	12	14	17	21	25	30	37	102
2004	99370	18	11	63	0	5	8	10	13	16	19	23	28	34	115
2005	96396	17	12	68	0	5	7	10	12	15	18	22	27	34	115
2006	83691	15	11	73	0	4	6	7	9	12	14	18	23	30	129

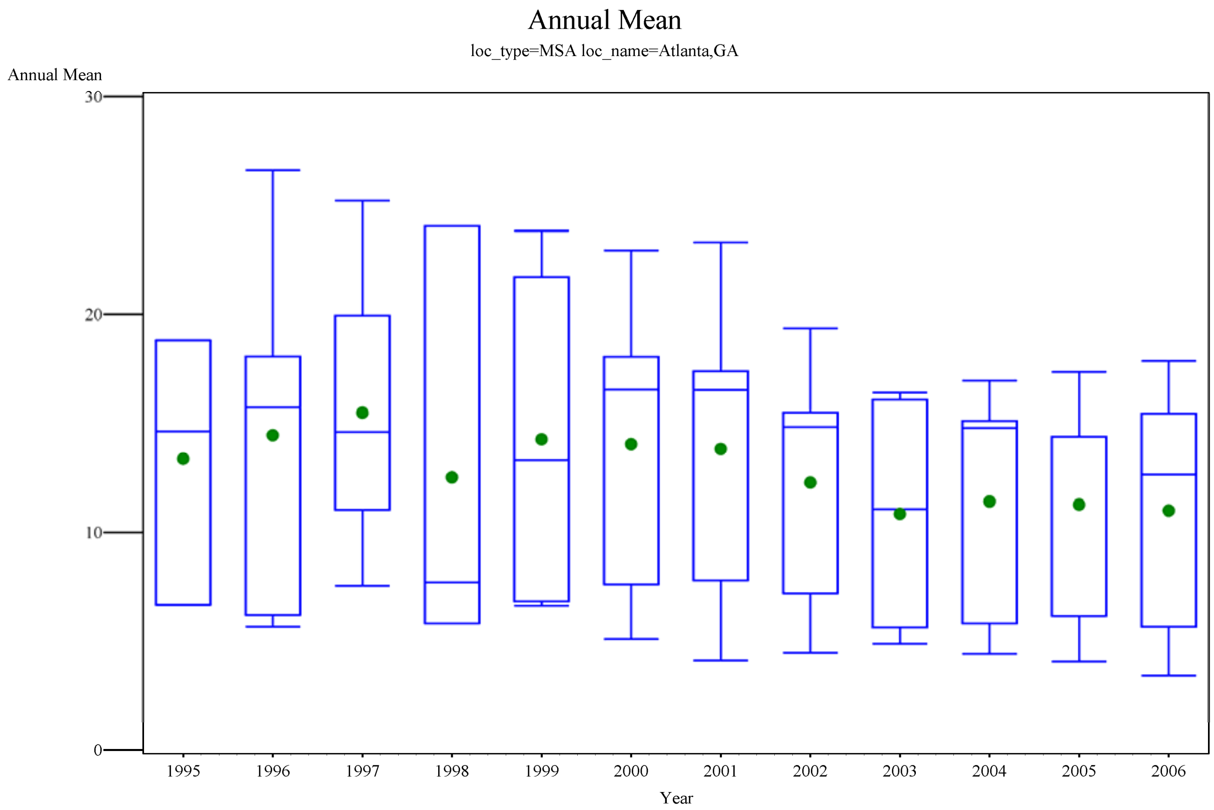


Figure B-21. Temporal distribution of annual average NO_2 ambient concentrations, Atlanta MSA, years 1995-2006.

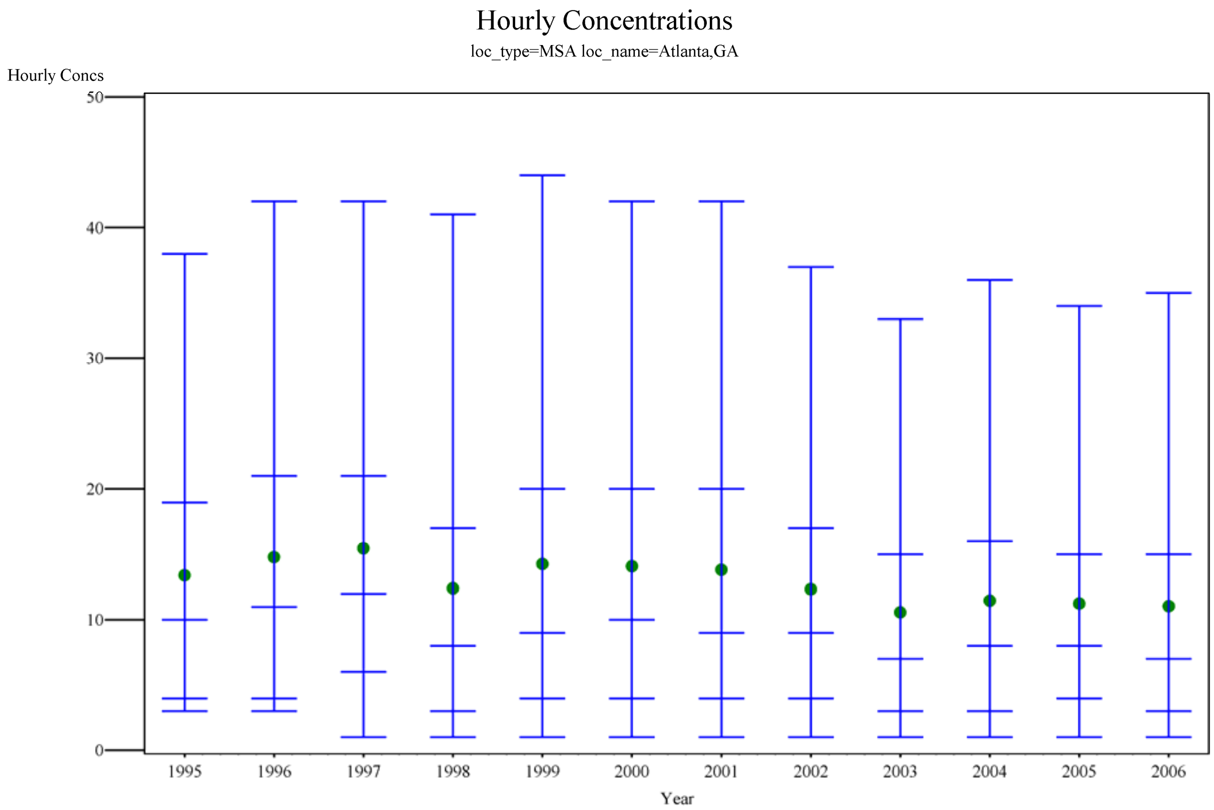


Figure B-22. Temporal distribution of hourly NO_2 ambient concentrations, Atlanta MSA, years 1995-2006.

Table B-21. Temporal distribution of annual average NO_2 ambient concentrations, Atlanta MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	3	13	6	46	7	7	7	7	15	15	15	19	19	19	19
1996	5	14	9	61	6	6	6	6	11	16	17	18	22	27	27
1997	4	15	7	47	8	8	8	15	15	15	15	15	25	25	25
1998	3	13	10	80	6	6	6	6	8	8	8	24	24	24	24
1999	4	14	9	61	7	7	7	7	7	13	20	20	24	24	24
2000	5	14	7	53	5	5	6	8	12	17	17	18	21	23	23
2001	5	14	8	56	4	4	6	8	12	17	17	17	20	23	23
2002	5	12	6	51	4	4	6	7	11	15	15	16	17	19	19
2003	4	11	6	56	5	5	5	6	6	11	16	16	16	16	16
2004	5	11	6	51	4	4	5	6	10	15	15	15	16	17	17
2005	5	11	6	51	4	4	5	6	10	14	14	14	16	17	17
2006	5	11	6	57	3	3	5	6	9	13	14	15	17	18	18

Table B-22. Temporal distribution of hourly NO_2 ambient concentrations, Atlanta MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	25213	13	12	89	1	3	3	5	7	10	13	16	22	30	93
1996	40576	15	13	89	1	3	3	5	8	11	14	18	24	34	122
1997	31069	15	13	86	1	3	5	7	9	12	15	18	23	33	181
1998	24142	12	13	105	0	1	3	4	6	8	11	14	20	30	124
1999	31121	14	14	99	0	2	4	5	7	9	12	17	23	35	242
2000	40584	14	14	97	1	1	3	5	7	10	13	17	23	33	110
2001	42761	14	14	98	1	1	3	5	7	9	13	17	23	33	172
2002	42076	12	12	95	1	1	3	5	6	9	11	15	20	29	136
2003	32215	11	11	101	0	1	2	3	5	7	9	13	17	26	91
2004	42124	11	11	98	1	1	3	4	6	8	10	14	19	28	127
2005	42279	11	11	96	1	1	3	4	6	8	10	13	18	27	97
2006	41052	11	11	98	1	2	3	4	5	7	9	13	18	27	73

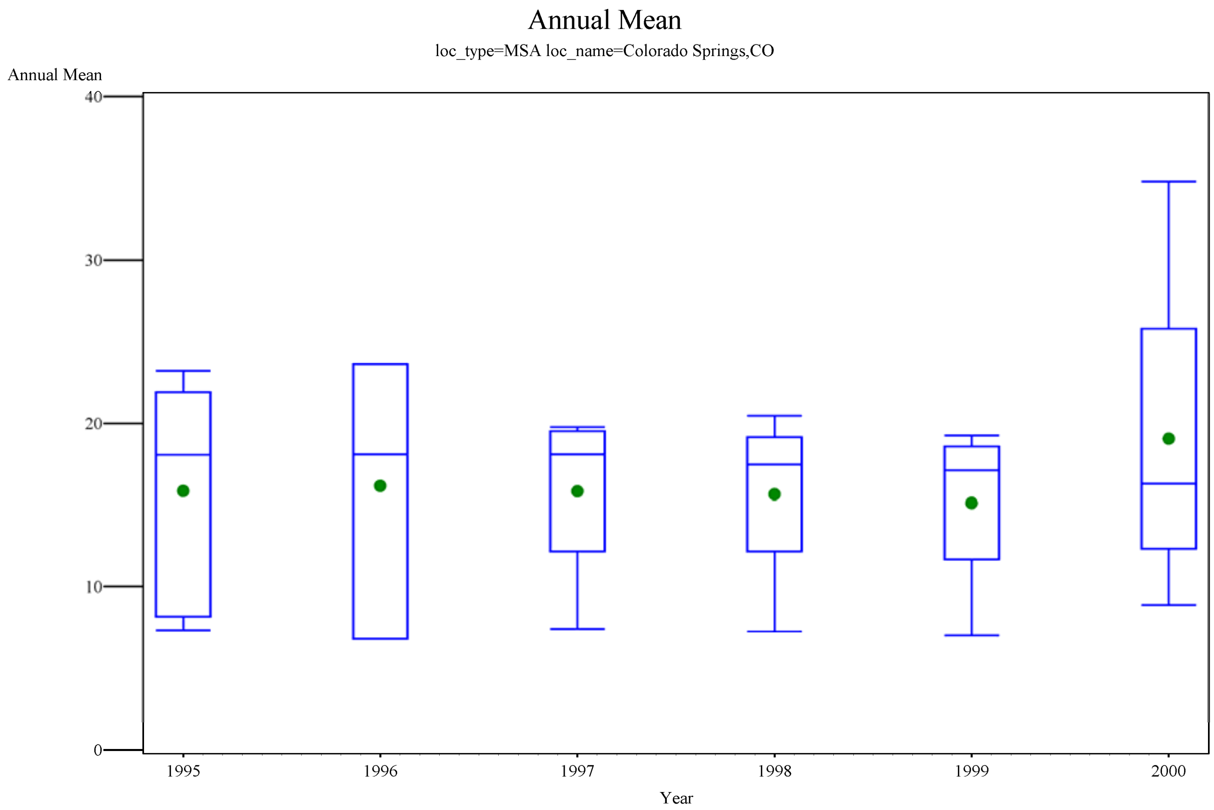


Figure B-23. Temporal distribution of annual average NO_2 ambient concentrations, Colorado Springs MSA, years 1995-2006.

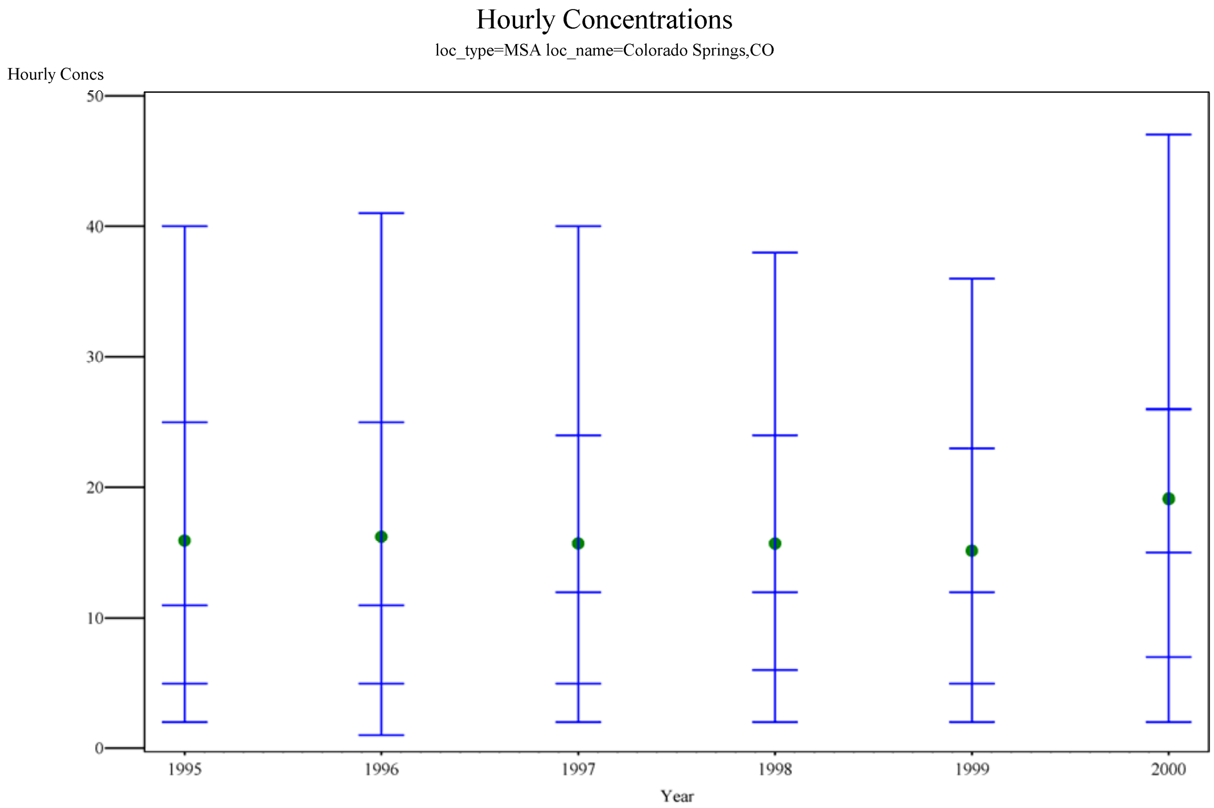


Figure B-24. Temporal distribution of hourly NO_2 ambient concentrations, Colorado Springs MSA, years 1995-2006.

Table B-23. Temporal distribution of annual average NO_2 ambient concentrations, Colorado Springs MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	7	16	7	42	7	7	8	12	12	18	21	21	22	23	23
1996	3	16	9	53	7	7	7	7	18	18	18	24	24	24	24
1997	4	16	6	36	7	7	7	17	17	18	19	19	20	20	20
1998	4	16	6	37	7	7	7	17	17	17	18	18	20	20	20
1999	4	15	6	37	7	7	7	16	16	17	18	18	19	19	19
2000	4	19	11	58	9	9	9	16	16	16	17	17	35	35	35

Table B-24. Temporal distribution of hourly NO_2 ambient concentrations, Colorado Springs MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	58569	16	14	91	0	2	4	6	8	11	16	22	29	36	148
1996	25387	16	16	101	0	2	4	6	8	11	16	21	28	35	246
1997	33469	16	13	80	0	3	5	6	9	12	16	21	27	35	118
1998	34509	16	12	76	0	3	5	7	9	12	16	22	27	34	85
1999	34472	15	12	82	0	3	4	6	9	12	16	21	26	32	230
2000	33956	19	20	106	0	3	6	8	11	15	20	24	28	34	308

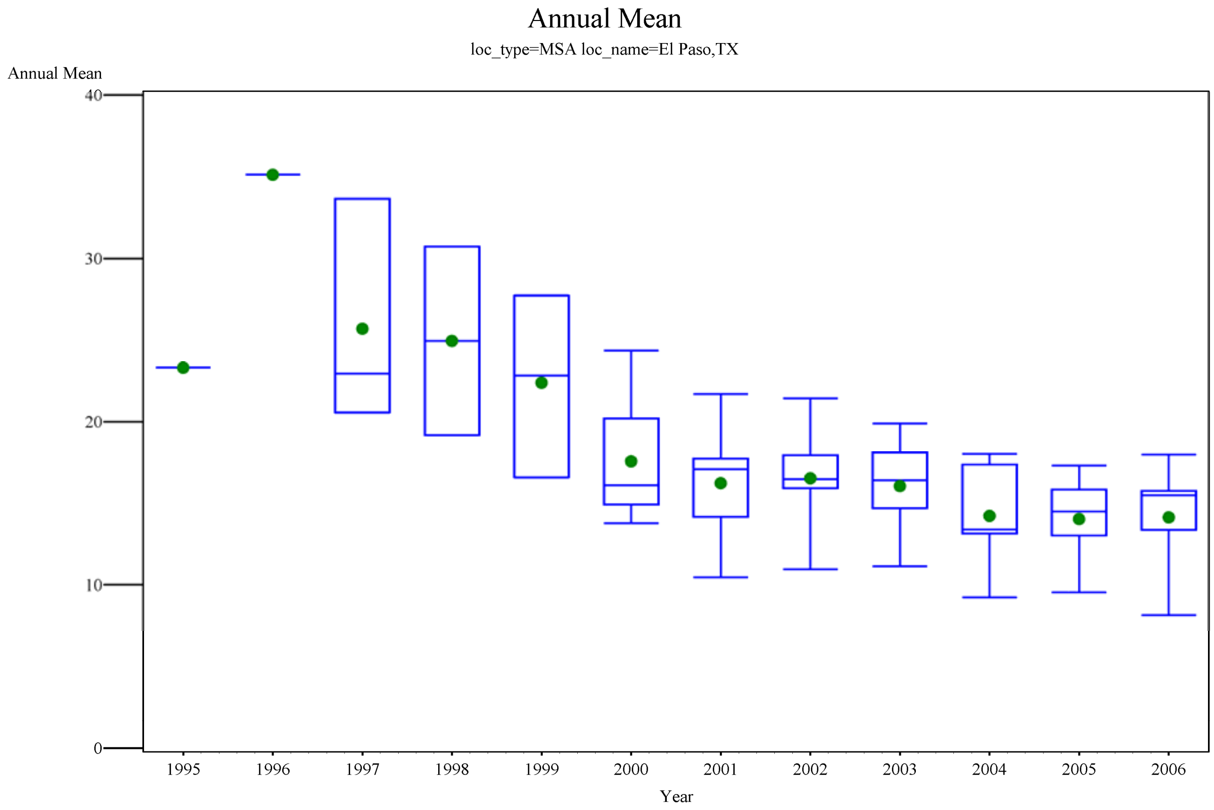


Figure B-25. Temporal distribution of annual average NO_2 ambient concentrations, El Paso MSA, years 1995-2006.

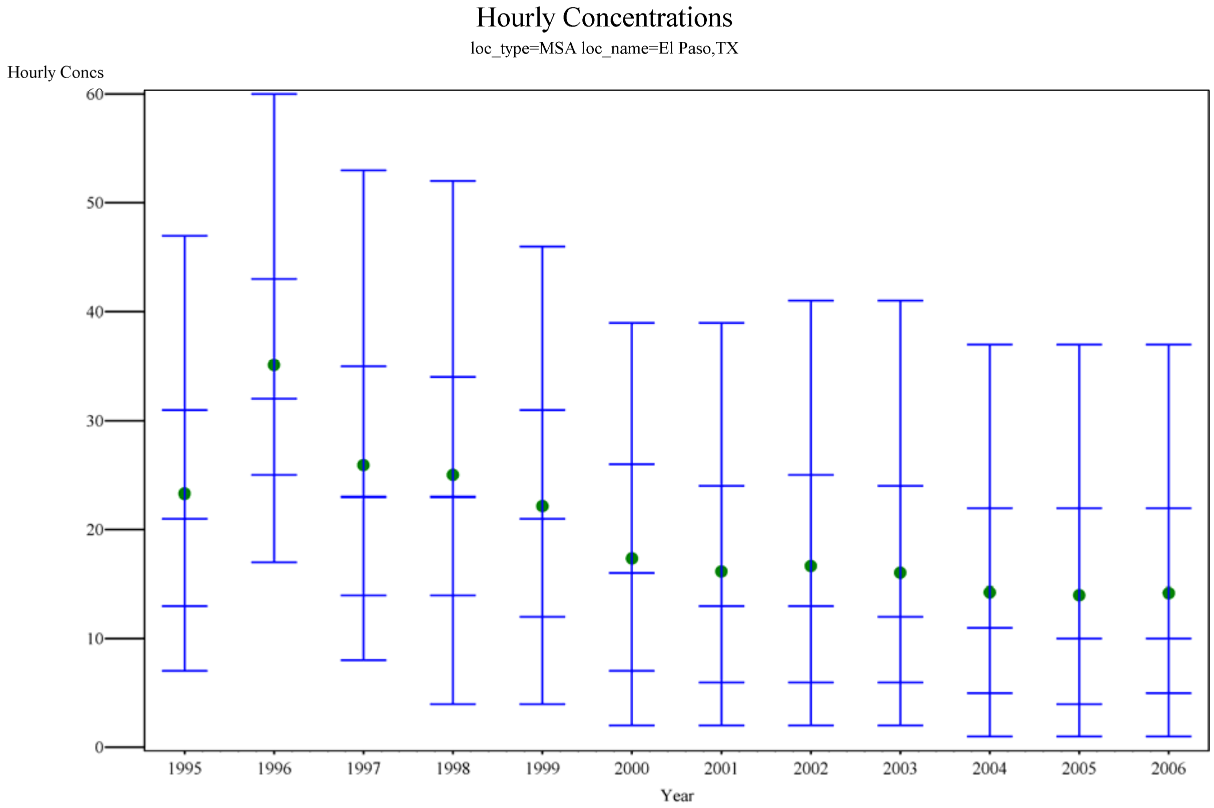


Figure B-26. Temporal distribution of hourly NO_2 ambient concentrations, El Paso MSA, years 1995-2006.

Table B-25. Temporal distribution of annual average NO_2 ambient concentrations, El Paso MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	1	23		0	23	23	23	23	23	23	23	23	23	23	23
1996	1	35		0	35	35	35	35	35	35	35	35	35	35	35
1997	3	26	7	27	21	21	21	21	23	23	23	34	34	34	34
1998	2	25	8	33	19	19	19	19	19	25	31	31	31	31	31
1999	3	22	6	25	17	17	17	17	23	23	23	28	28	28	28
2000	4	18	5	26	14	14	14	16	16	16	16	16	24	24	24
2001	5	16	4	26	10	10	12	14	16	17	17	18	20	22	22
2002	5	17	4	23	11	11	13	16	16	16	17	18	20	21	21
2003	5	16	3	21	11	11	13	15	16	16	17	18	19	20	20
2004	5	14	4	25	9	9	11	13	13	13	15	17	18	18	18
2005	5	14	3	21	10	10	11	13	14	15	15	16	17	17	17
2006	5	14	4	26	8	8	11	13	14	15	16	16	17	18	18

Table B-26. Temporal distribution of hourly NO_2 ambient concentrations, El Paso MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	6960	23	13	58	3	9	12	14	17	21	25	29	34	41	113
1996	6627	35	15	43	2	20	23	27	29	32	36	40	46	54	219
1997	22888	26	15	58	0	10	13	16	20	23	28	32	38	45	174
1998	15523	25	15	61	0	7	12	15	19	23	27	32	37	45	166
1999	23447	22	13	60	0	6	10	14	17	21	25	28	33	40	108
2000	30772	17	13	72	0	3	5	8	12	16	20	24	28	34	125
2001	38020	16	12	77	0	3	5	7	10	13	16	21	27	34	102
2002	41466	17	13	77	0	4	5	7	10	13	17	22	28	35	153
2003	39968	16	13	80	0	3	5	7	9	12	16	21	27	35	106
2004	41952	14	12	83	0	2	4	6	8	11	14	19	25	32	97
2005	41496	14	12	86	0	2	4	5	7	10	14	19	24	31	87
2006	37203	14	12	84	0	2	4	6	8	10	14	19	25	32	99

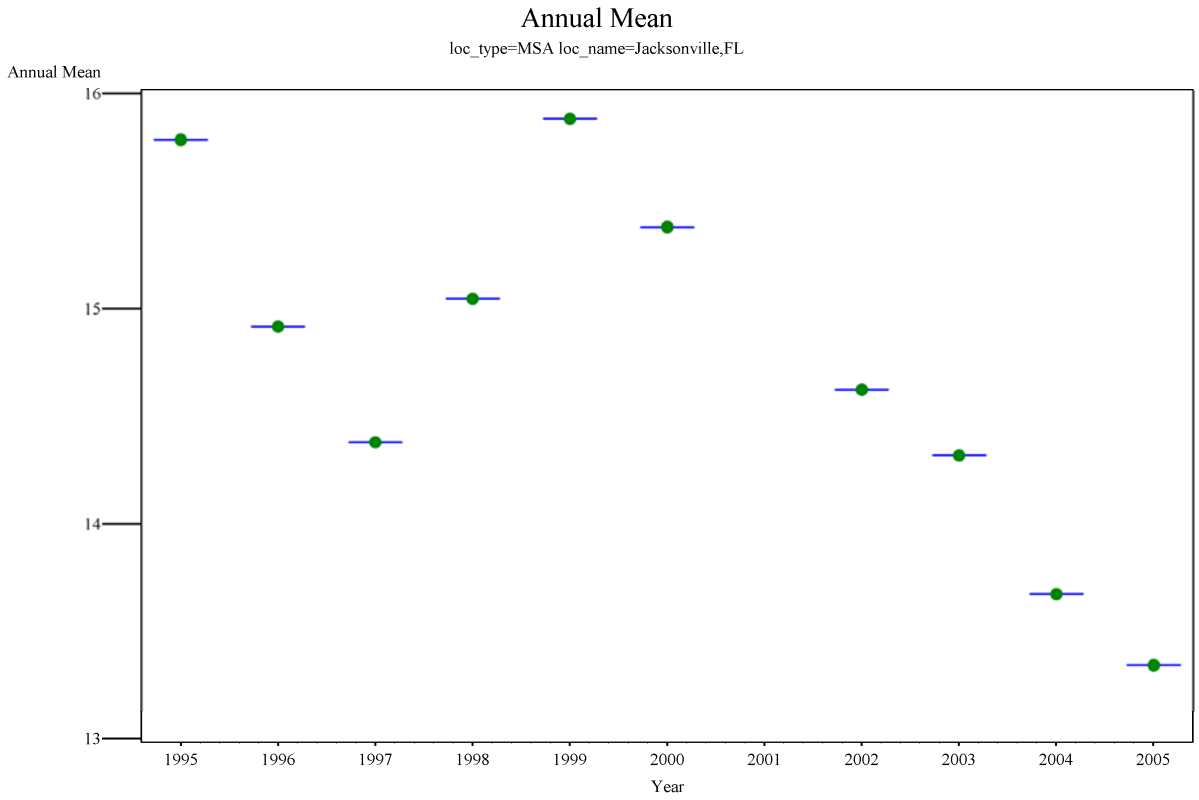


Figure B-27. Temporal distribution of annual average NO_2 ambient concentrations, Jacksonville MSA, years 1995-2006.

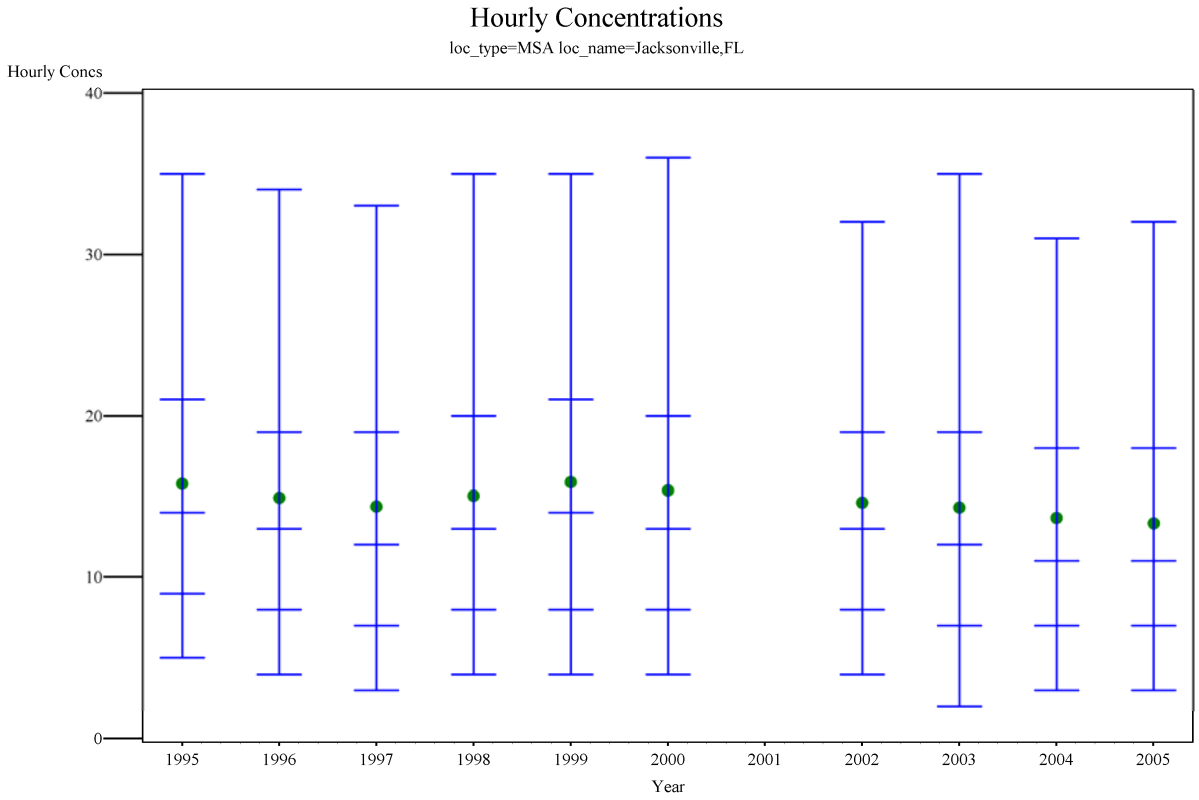


Figure B-28. Temporal distribution of hourly NO_2 ambient concentrations, Jacksonville MSA, years 1995-2006.

Table B-27. Temporal distribution of annual average NO_2 ambient concentrations, Jacksonville MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max

1995	1	16		0	16	16	16	16	16	16	16	16	16	16	16
1996	1	15		0	15	15	15	15	15	15	15	15	15	15	15
1997	1	14		0	14	14	14	14	14	14	14	14	14	14	14
1998	1	15		0	15	15	15	15	15	15	15	15	15	15	15
1999	1	16		0	16	16	16	16	16	16	16	16	16	16	16
2000	1	15		0	15	15	15	15	15	15	15	15	15	15	15
2002	1	15		0	15	15	15	15	15	15	15	15	15	15	15
2003	1	14		0	14	14	14	14	14	14	14	14	14	14	14
2004	1	14		0	14	14	14	14	14	14	14	14	14	14	14
2005	1	13		0	13	13	13	13	13	13	13	13	13	13	13

Table B-28. Temporal distribution of hourly NO_2 ambient concentrations, Jacksonville MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	7755	16	10	60	0	6	8	9	11	14	16	19	23	29	76
1996	8148	15	10	64	0	5	7	9	11	13	15	18	21	28	80
1997	8326	14	9	65	0	5	6	8	10	12	15	17	21	27	92
1998	8211	15	10	65	0	5	7	9	11	13	15	18	22	28	66
1999	7795	16	10	61	0	5	7	9	12	14	16	20	24	30	63
2000	7661	15	10	67	0	5	7	9	11	13	15	18	23	30	72
2002	7944	15	10	66	0	5	7	9	11	13	15	17	21	27	294
2003	7041	14	10	71	0	4	6	8	10	12	14	17	21	28	76
2004	7451	14	11	83	0	4	6	7	9	11	13	16	20	26	201
2005	7890	13	9	67	0	4	6	8	9	11	13	16	20	26	64

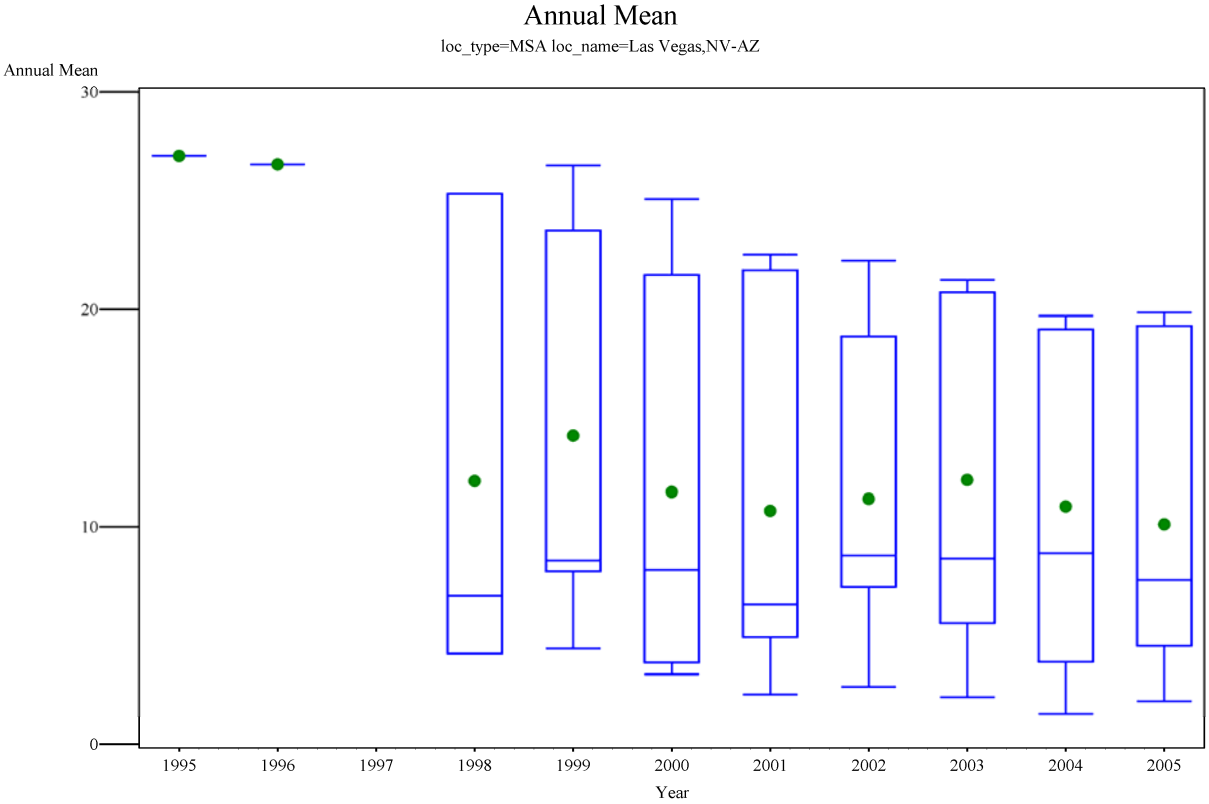


Figure B-29. Temporal distribution of annual average NO_2 ambient concentrations, Las Vegas MSA, years 1995-2006.

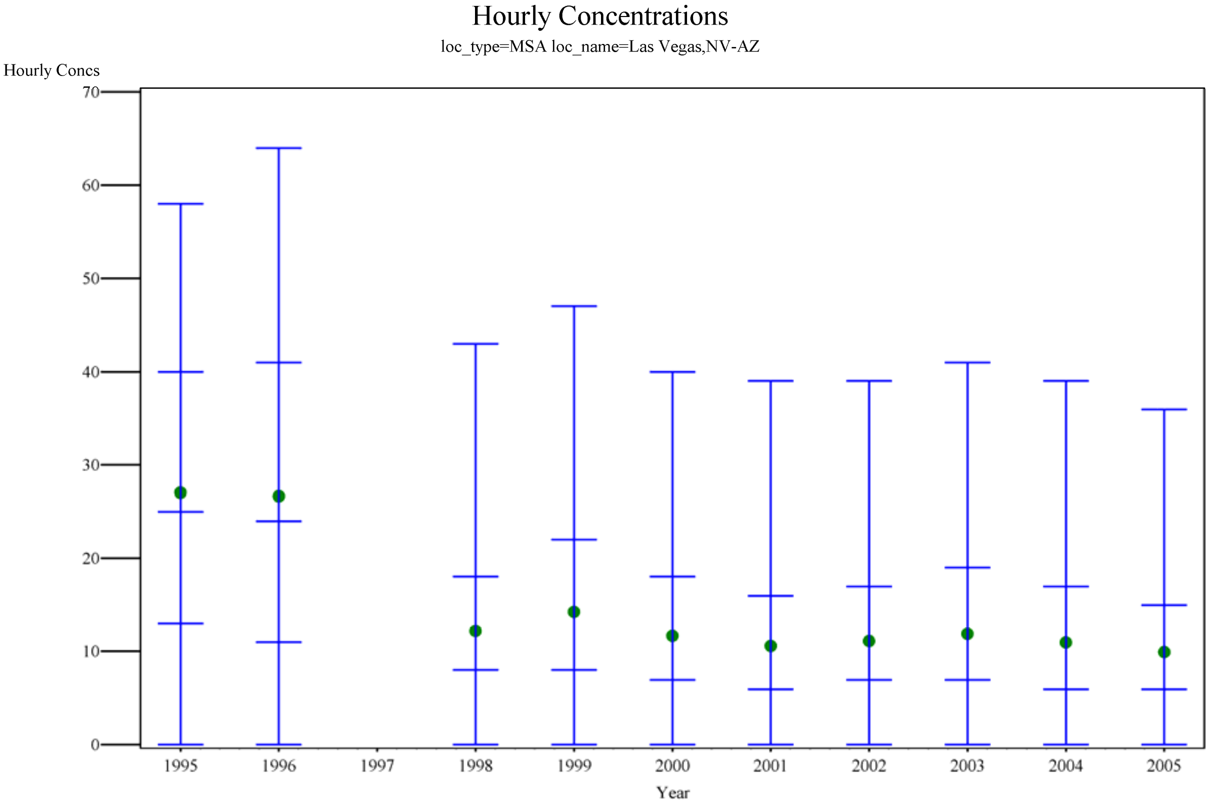


Figure B-30. Temporal distribution of hourly NO_2 ambient concentrations, Las Vegas MSA, years 1995-2006.

Table B-29. Temporal distribution of annual average NO_2 ambient concentrations, Las Vegas MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	1	27		0	27	27	27	27	27	27	27	27	27	27	27
1996	1	27		0	27	27	27	27	27	27	27	27	27	27	27
1998	3	12	12	95	4	4	4	4	7	7	7	25	25	25	25
1999	5	14	10	71	4	4	6	8	8	8	16	24	25	27	27
2000	6	12	9	81	3	3	4	4	8	8	8	22	22	25	25
2001	6	11	9	84	2	2	5	5	6	6	7	22	22	23	23
2002	9	11	8	68	3	3	3	7	7	9	10	19	22	22	22
2003	7	12	8	66	2	2	6	8	8	9	19	19	21	21	21
2004	7	11	8	73	1	1	4	5	5	9	19	19	19	20	20
2005	6	10	8	76	2	2	5	5	6	8	9	19	19	20	20

Table B-30. Temporal distribution of hourly NO_2 ambient concentrations, Las Vegas MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	7951	27	20	74	0	0	11	15	20	25	31	37	42	50	410
1996	8723	27	22	81	0	0	9	12	17	24	31	38	44	54	149
1998	25234	12	14	118	0	0	0	0	5	8	10	14	23	35	103
1999	43110	14	16	110	0	0	0	5	6	8	12	18	28	39	110
2000	46403	12	14	119	0	0	0	0	5	7	10	15	23	34	100
2001	49734	11	14	128	0	0	0	0	0	6	8	13	21	33	104
2002	74814	11	13	117	0	0	0	0	5	7	10	14	21	32	87
2003	58398	12	14	119	0	0	0	0	5	7	10	15	24	35	103
2004	57484	11	13	120	0	0	0	0	0	6	9	14	23	33	73
2005	48911	10	12	123	0	0	0	0	0	6	9	12	18	30	75

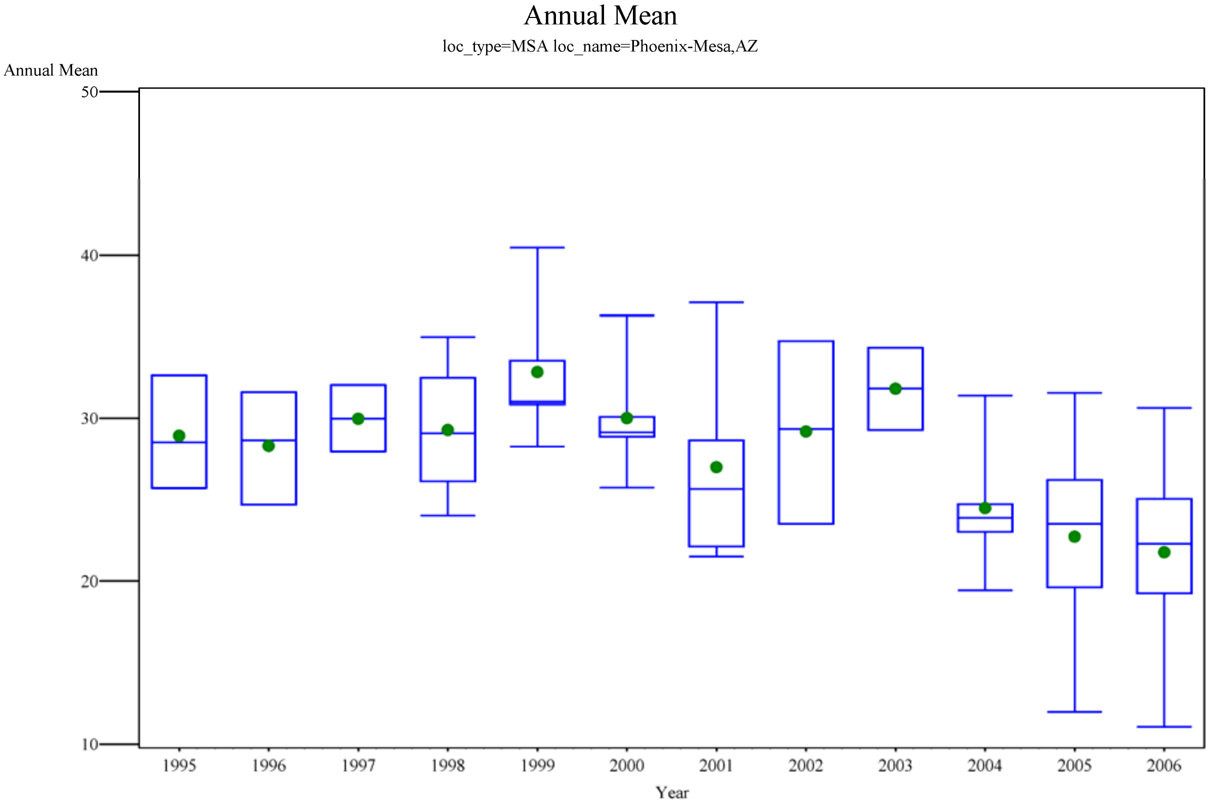


Figure B-31. Temporal distribution of annual average NO_2 ambient concentrations, Phoenix MSA, years 1995-2006.

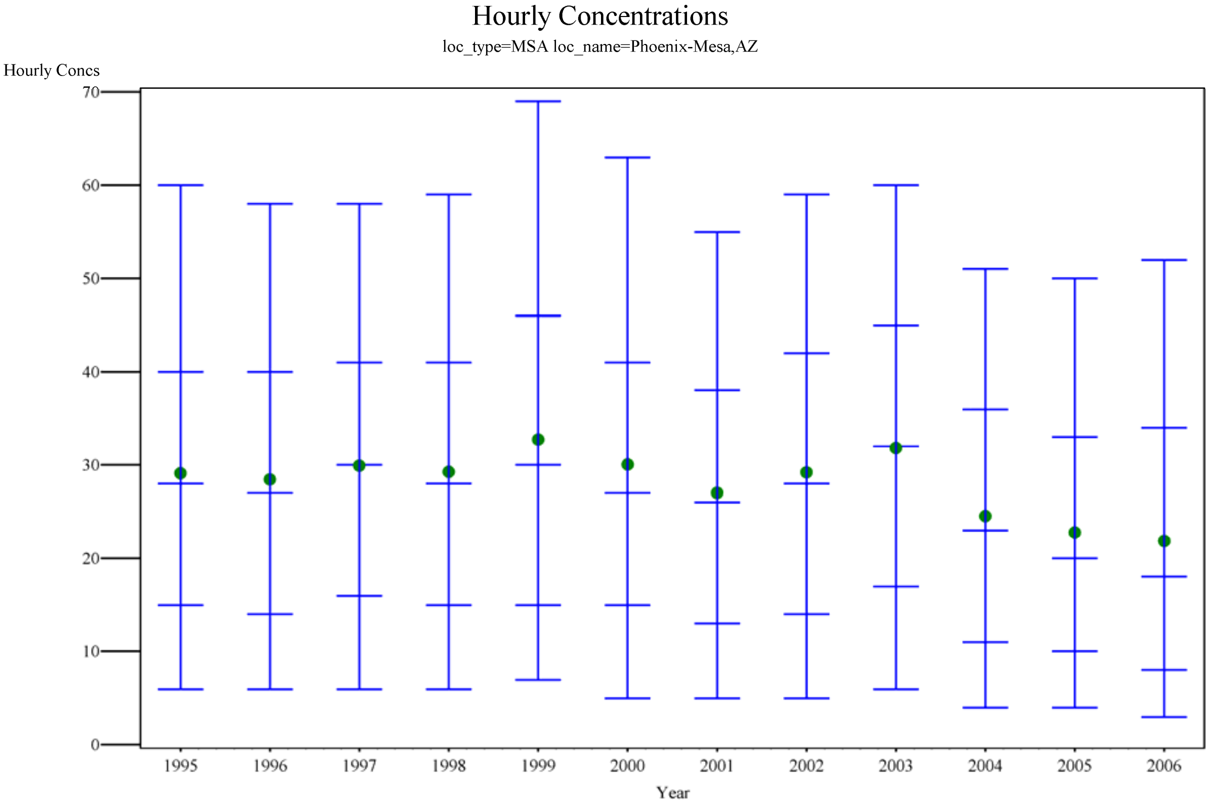


Figure B-32. Temporal distribution of hourly NO_2 ambient concentrations, Phoenix MSA, years 1995-2006.

Table B-31. Temporal distribution of annual average NO_2 ambient concentrations, Phoenix MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	3	29	3	12	26	26	26	26	29	29	29	33	33	33	33
1996	3	28	3	12	25	25	25	25	29	29	29	32	32	32	32
1997	2	30	3	10	28	28	28	28	28	30	32	32	32	32	32
1998	4	29	5	15	24	24	24	28	28	29	30	30	35	35	35
1999	5	33	5	14	28	28	30	31	31	31	32	34	37	40	40
2000	5	30	4	13	26	26	27	29	29	29	30	30	33	36	36
2001	5	27	6	23	22	22	22	22	24	26	27	29	33	37	37
2002	3	29	6	19	24	24	24	24	29	29	29	35	35	35	35
2003	2	32	4	11	29	29	29	29	29	32	34	34	34	34	34
2004	5	25	4	18	19	19	21	23	23	24	24	25	28	31	31
2005	6	23	7	29	12	12	20	20	24	24	24	26	26	32	32
2006	6	22	7	30	11	11	19	19	21	22	24	25	25	31	31

Table B-32. Temporal distribution of hourly NO_2 ambient concentrations, Phoenix MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	23196	29	17	59	0	8	12	17	23	28	33	37	44	53	128
1996	23598	28	17	59	0	8	12	17	22	27	32	37	43	51	115
1997	14629	30	16	55	0	8	13	18	25	30	35	39	44	52	114
1998	32078	29	17	58	0	8	12	17	23	28	33	38	44	52	116
1999	40996	33	22	66	0	9	13	18	24	30	36	42	49	60	198
2000	41686	30	21	71	0	8	12	17	22	27	32	38	45	54	267
2001	40463	27	16	59	1	7	11	15	21	26	31	36	41	49	118
2002	25028	29	17	59	0	7	12	17	23	28	34	39	45	53	108
2003	14195	32	17	55	0	8	14	20	27	32	37	42	48	55	101
2004	42176	25	15	62	0	6	9	13	18	23	28	33	39	45	104
2005	50583	23	15	66	0	5	8	12	16	20	25	31	36	44	131
2006	48791	22	16	73	0	4	7	10	13	18	24	30	37	46	111

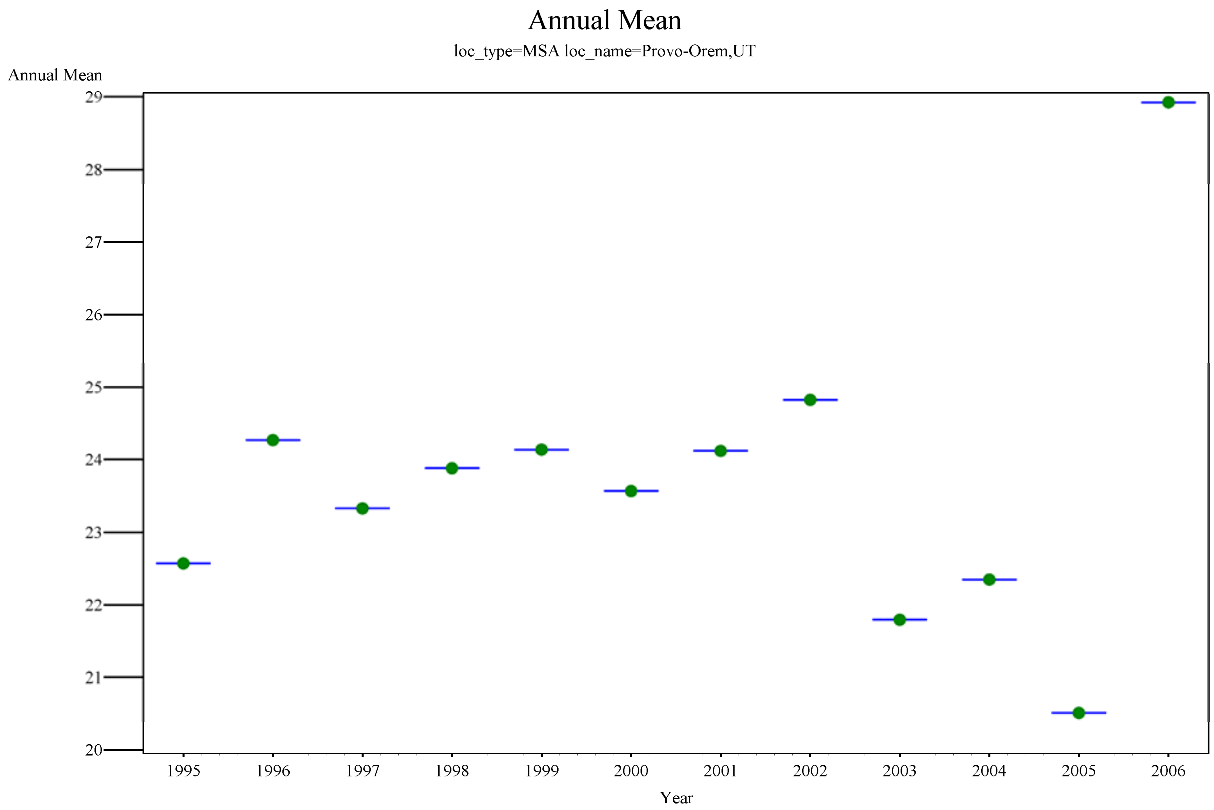


Figure B-33. Temporal distribution of annual average NO_2 ambient concentrations, Provo MSA, years 1995-2006.

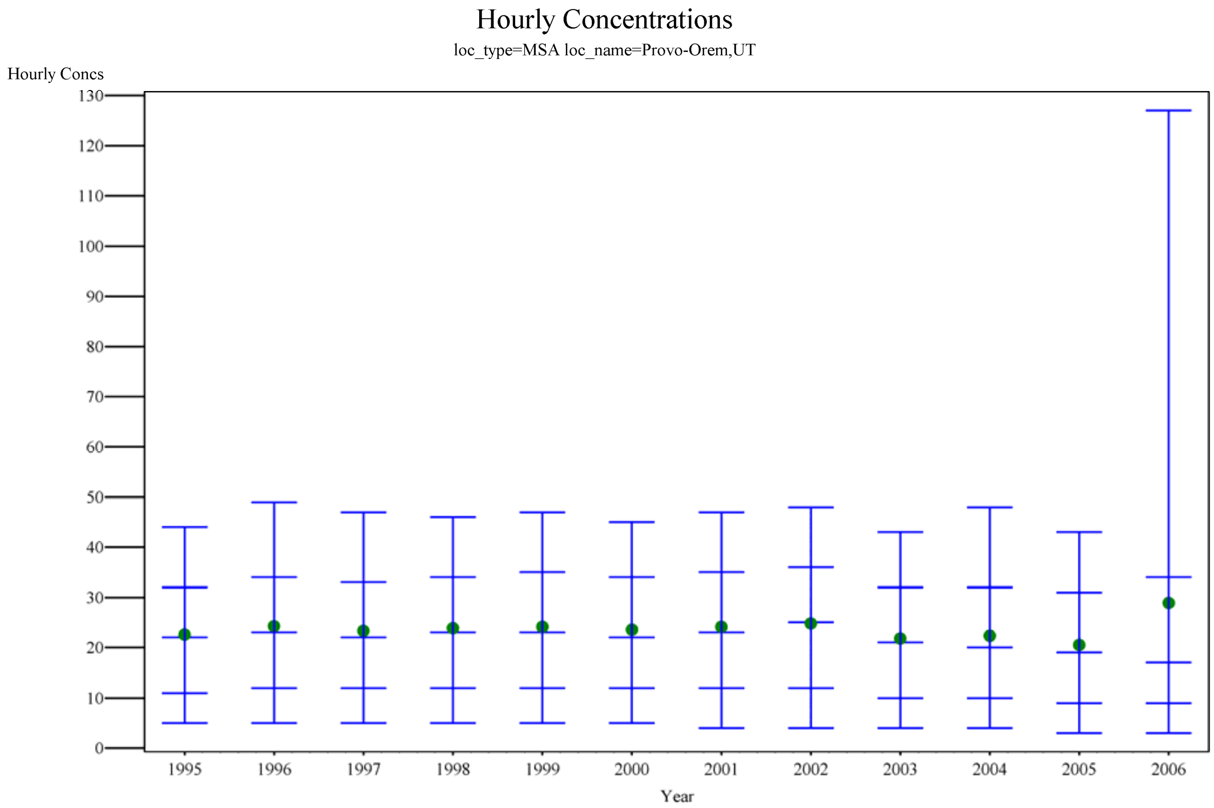


Figure B-34. Temporal distribution of hourly NO_2 ambient concentrations, Provo MSA, years 1995-2006.

Table B-33. Temporal distribution of annual average NO_2 ambient concentrations, Provo MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	1	23		0	23	23	23	23	23	23	23	23	23	23	23
1996	1	24		0	24	24	24	24	24	24	24	24	24	24	24
1997	1	23		0	23	23	23	23	23	23	23	23	23	23	23
1998	1	24		0	24	24	24	24	24	24	24	24	24	24	24
1999	1	24		0	24	24	24	24	24	24	24	24	24	24	24
2000	1	24		0	24	24	24	24	24	24	24	24	24	24	24
2001	1	24		0	24	24	24	24	24	24	24	24	24	24	24
2002	1	25		0	25	25	25	25	25	25	25	25	25	25	25
2003	1	22		0	22	22	22	22	22	22	22	22	22	22	22
2004	1	22		0	22	22	22	22	22	22	22	22	22	22	22
2005	1	21		0	21	21	21	21	21	21	21	21	21	21	21
2006	1	29		0	29	29	29	29	29	29	29	29	29	29	29

Table B-34. Temporal distribution of hourly NO_2 ambient concentrations, Provo MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	8002	23	13	55	0	7	10	13	17	22	26	30	34	40	67
1996	8430	24	15	61	0	7	10	14	18	23	28	32	37	43	97
1997	7034	23	13	57	0	7	10	14	18	22	26	31	35	41	81
1998	8210	24	13	56	0	7	10	14	18	23	28	32	37	42	78
1999	8563	24	13	55	0	7	11	14	19	23	28	33	37	42	77
2000	8406	24	13	56	0	7	10	14	18	22	27	32	37	42	74
2001	8501	24	14	57	0	6	10	14	19	23	28	33	38	43	72
2002	8200	25	14	57	0	6	10	15	20	25	30	34	38	43	80
2003	7730	22	13	59	0	6	8	12	16	21	26	30	34	39	72
2004	8302	22	15	66	0	5	8	12	16	20	25	30	35	42	90
2005	8502	21	13	62	0	5	8	11	15	19	23	28	33	39	64
2006	6993	29	34	118	0	5	7	10	13	17	22	30	38	61	164

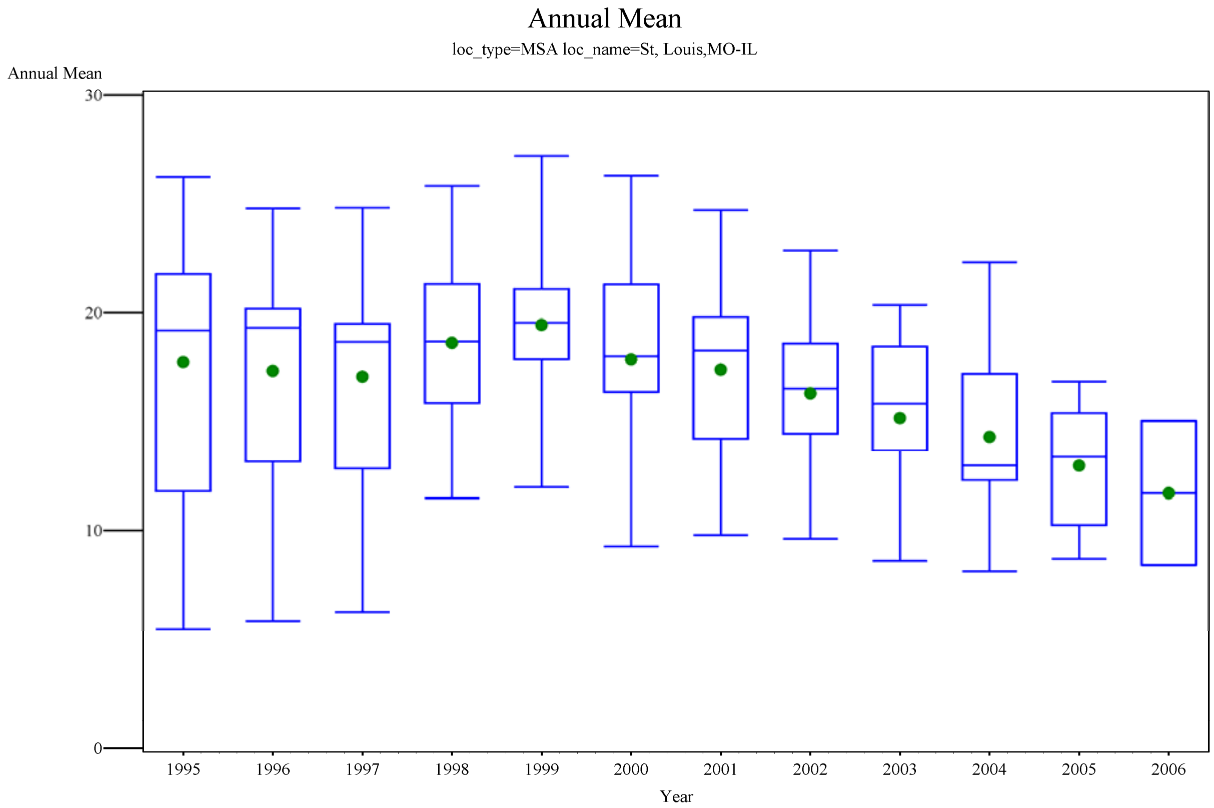


Figure B-35. Temporal distribution of annual average NO_2 ambient concentrations, St. Louis MSA, years 1995-2006.

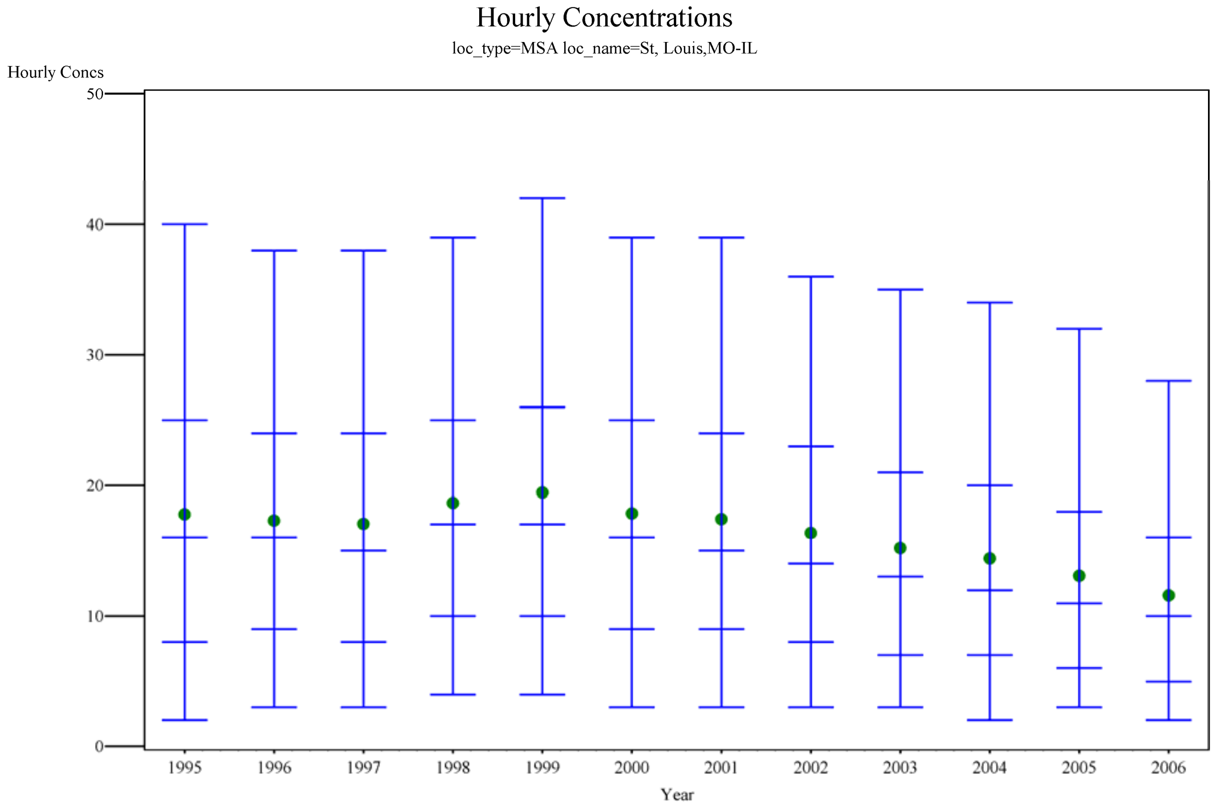


Figure B-36. Temporal distribution of hourly NO_2 ambient concentrations, St. Louis MSA, years 1995-2006.

Table B-35. Temporal distribution of annual average NO_2 ambient concentrations, St. Louis MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	10	18	6	35	5	8	12	15	19	19	20	22	22	24	26
1996	10	17	6	33	6	8	12	16	19	19	20	20	21	23	25
1997	10	17	6	32	6	8	12	16	19	19	19	19	21	23	25
1998	8	19	5	25	11	11	13	18	19	19	19	20	22	26	26
1999	9	19	5	24	12	12	14	18	18	20	21	21	24	27	27
2000	9	18	5	29	9	9	12	16	17	18	19	21	21	26	26
2001	8	17	5	28	10	10	12	17	17	18	19	20	20	25	25
2002	9	16	4	26	10	10	11	14	15	16	17	19	21	23	23
2003	9	15	4	26	9	9	10	14	14	16	16	18	19	20	20
2004	9	14	4	31	8	8	10	12	13	13	16	17	18	22	22
2005	6	13	3	24	9	9	10	10	12	13	15	15	15	17	17
2006	2	12	5	40	8	8	8	8	8	12	15	15	15	15	15

Table B-36. Temporal distribution of hourly NO_2 ambient concentrations, St. Louis MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	85072	18	12	68	0	4	7	10	13	16	19	23	28	34	103
1996	86085	17	11	65	0	4	7	10	13	16	19	22	26	32	84
1997	86314	17	11	67	0	4	7	10	12	15	18	22	26	33	274
1998	68308	19	11	58	0	6	9	12	14	17	20	23	28	33	97
1999	77611	19	12	61	0	6	9	12	14	17	20	24	29	36	99
2000	77327	18	11	64	0	5	8	10	13	16	19	22	27	34	85
2001	67871	17	11	64	0	5	7	10	13	15	19	22	27	33	95
2002	76693	16	11	65	0	5	7	9	12	14	17	21	25	31	124
2003	77543	15	10	67	0	4	6	8	11	13	16	19	23	29	123
2004	75493	14	10	69	0	4	6	8	10	12	15	18	22	28	130
2005	49948	13	9	70	0	4	5	7	9	11	13	16	20	26	70
2006	16688	12	8	70	0	3	5	6	8	10	12	15	18	23	53

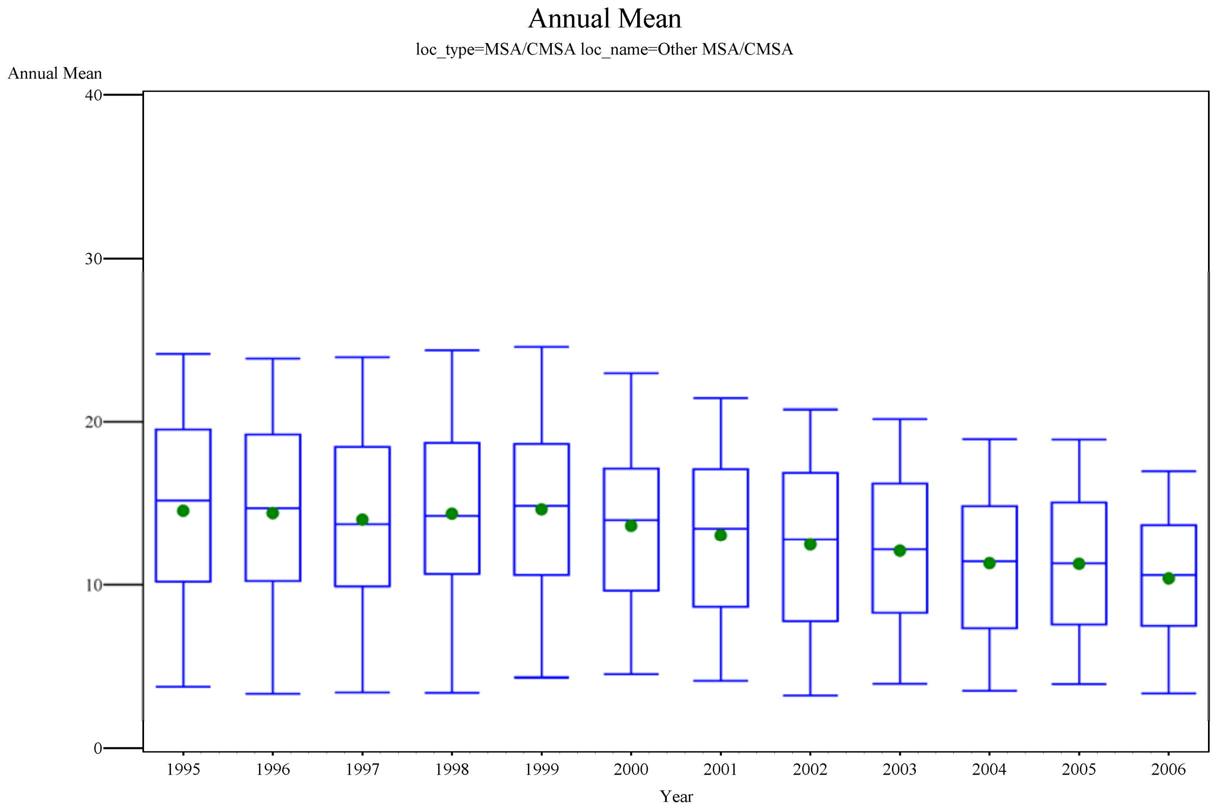


Figure B-37. Temporal distribution of annual average NO_2 ambient concentrations, Other MSA/CMSA, years 1995-2006.

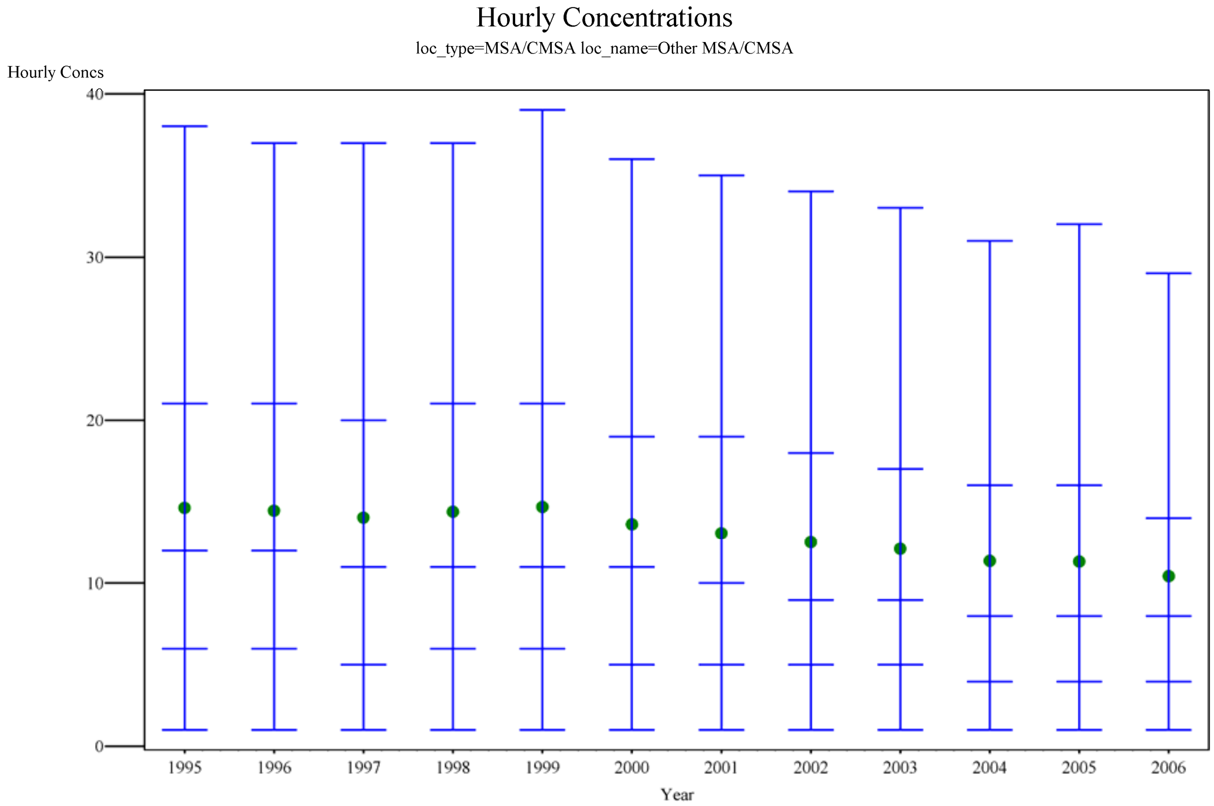


Figure B-38. Temporal distribution of hourly NO_2 ambient concentrations, Other MSA/CMSA, years 1995-2006.

Table B-37. Temporal distribution of annual average NO_2 ambient concentrations, Other MSA/CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	186	15	6	44	1	5	8	11	13	15	17	18	21	22	32
1996	186	14	6	43	1	5	9	11	13	15	16	18	20	22	30
1997	187	14	6	43	2	5	9	11	12	14	16	18	19	22	29
1998	185	14	6	43	1	5	10	11	13	14	16	18	20	22	31
1999	192	15	6	42	1	6	9	11	14	15	16	18	20	23	29
2000	199	14	6	41	1	5	8	11	12	14	16	17	18	21	26
2001	201	13	6	43	1	5	7	10	12	13	15	17	18	20	27
2002	209	12	6	45	1	5	7	9	11	13	14	16	17	20	27
2003	202	12	5	42	1	5	7	9	11	12	14	15	17	18	26
2004	211	11	5	44	1	5	7	9	10	11	13	14	16	17	25
2005	207	11	5	43	1	5	7	9	10	11	12	14	16	17	24
2006	147	10	4	41	1	4	6	9	9	11	12	13	14	16	18

Table B-38. Temporal distribution of hourly NO_2 ambient concentrations, Other MSA/CMSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	186	15	6	44	1	5	8	11	13	15	17	18	21	22	32
1996	1520743	14	12	81	0	2	5	7	9	12	15	18	23	31	336
1997	1520290	14	11	82	0	2	4	6	9	11	14	18	23	30	313
1998	1503051	14	11	80	0	2	5	7	9	11	15	18	23	31	300
1999	1560074	15	12	83	0	3	5	7	9	11	14	18	24	32	172
2000	1630060	14	11	81	0	2	4	6	8	11	13	17	22	29	289
2001	1648640	13	11	84	0	2	4	6	8	10	13	16	21	29	193
2002	1713558	13	11	85	0	2	4	5	7	9	12	15	20	28	158
2003	1661992	12	10	84	0	2	4	5	7	9	12	15	19	26	148
2004	1738133	11	10	87	0	2	3	5	7	8	11	14	18	25	160
2005	1706730	11	10	87	0	2	3	5	6	8	11	14	18	25	153
2006	1168444	10	9	87	0	2	3	5	6	8	10	13	17	23	240

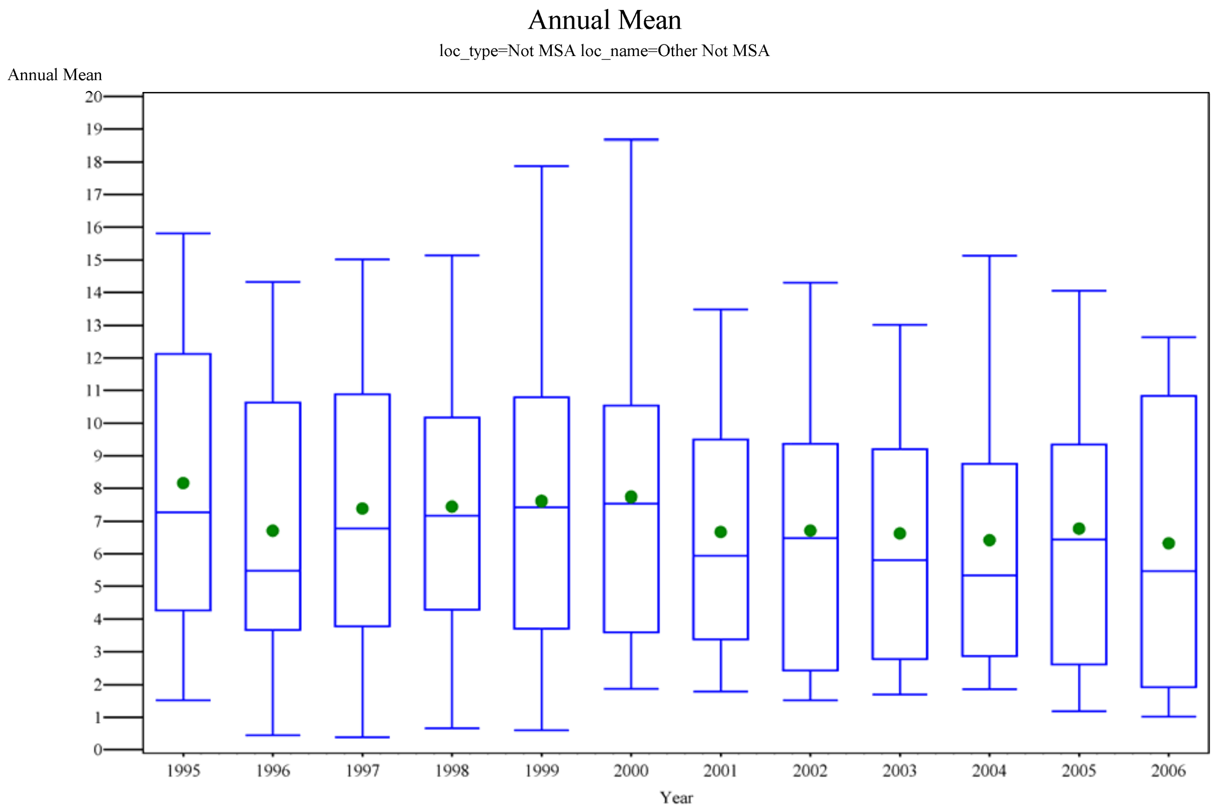


Figure B-39. Temporal distribution of annual average NO_2 ambient concentrations, Other Not MSA, years 1995-2006.

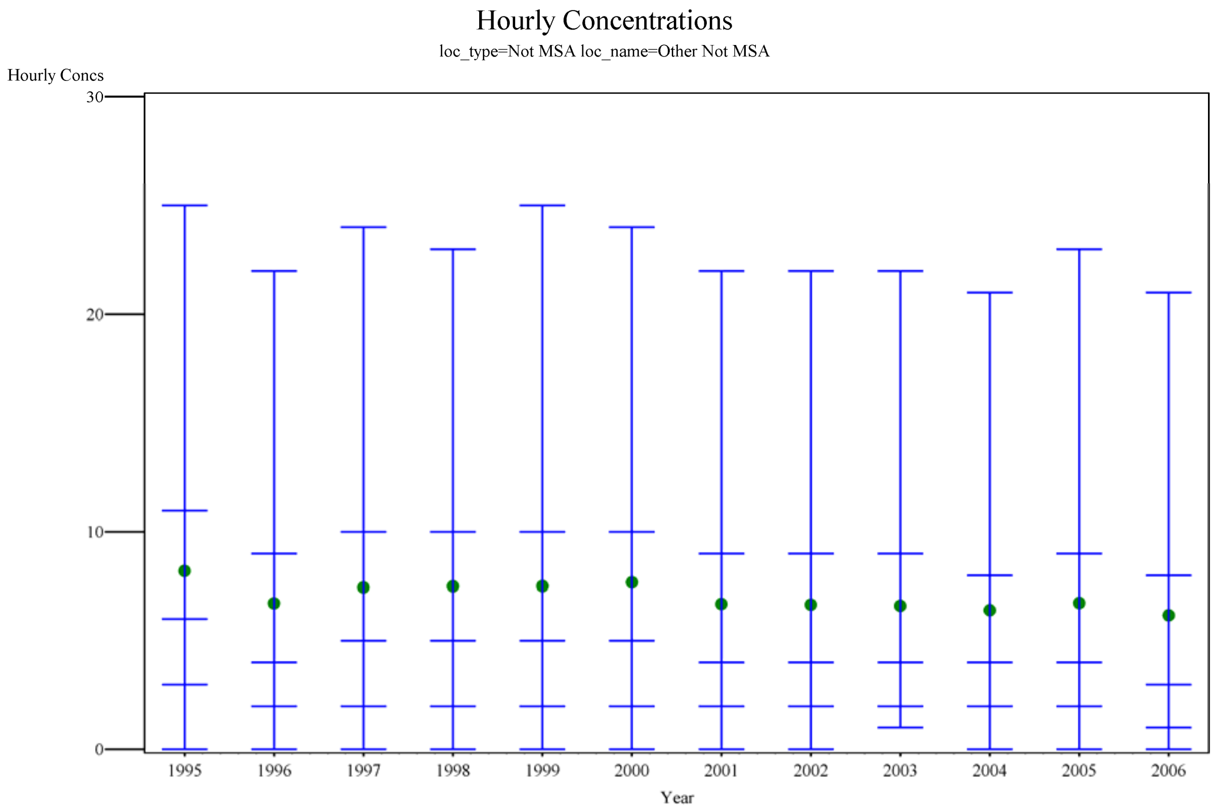


Figure B-40. Temporal distribution of hourly NO_2 ambient concentrations, Other Not MSA, years 1995-2006.

Table B-39. Temporal distribution of annual average NO_2 ambient concentrations, Other Not MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	28	8	5	59	1	2	4	5	7	7	8	10	13	15	19
1996	29	7	5	71	0	0	2	4	5	5	7	10	13	14	14
1997	35	7	5	67	0	1	3	4	5	7	9	10	12	14	20
1998	33	7	5	62	1	1	3	4	5	7	7	10	12	14	19
1999	36	8	5	67	0	1	3	4	5	7	8	9	12	16	20
2000	39	8	4	57	2	2	3	5	6	8	8	10	11	14	19
2001	41	7	4	60	1	2	3	4	5	6	8	9	10	13	17
2002	42	7	4	65	1	2	2	3	4	6	8	8	10	13	16
2003	44	7	4	61	1	2	3	3	4	6	8	9	11	13	15
2004	47	6	4	64	2	2	2	3	4	5	7	8	11	13	16
2005	43	7	4	63	1	2	2	3	5	6	8	9	11	12	17
2006	26	6	5	71	1	1	2	2	3	5	8	10	11	12	16

Table B-40. Temporal distribution of hourly NO_2 ambient concentrations, Other Not MSA, years 1995-2006.

Year	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1995	225810	8	9	104	0	0	2	3	4	6	7	10	13	19	217
1996	234628	7	8	118	0	0	1	2	3	4	6	8	11	17	164
1997	278906	7	8	113	0	0	1	2	3	5	6	9	12	18	207
1998	264015	8	8	105	0	1	2	3	4	5	7	9	12	18	181
1999	290382	8	9	113	0	0	2	2	3	5	6	9	12	18	286
2000	316568	8	8	104	0	1	2	3	4	5	7	9	12	18	192
2001	328407	7	7	109	0	1	1	2	3	4	6	8	11	16	139
2002	340873	7	7	112	0	1	1	2	3	4	5	8	11	17	267
2003	351652	7	7	110	0	1	2	2	3	4	5	7	10	16	201
2004	375716	6	7	115	0	1	1	2	3	4	5	7	10	16	285
2005	353229	7	8	114	0	1	1	2	3	4	6	8	11	17	262
2006	207114	6	7	119	0	0	1	2	2	3	5	7	10	16	101

Appendix C. Spatial NO_2 Air Quality Characterization

Appendix C contains the ambient air quality analysis results by monitoring site within each of the named locations. Boxplots were constructed to display the annual average and hourly concentration distributions across sites for a single location. The box extends from the 25th to the 75th percentile, with the median shown as the line inside the box. The whiskers extend from the box to the 5th and 95th percentiles. The extreme values in the upper and lower tails beyond the 5th and 95th percentiles are not shown to allow for similar scaling along the y-axis for the plotted independent variables. The mean is plotted as a dot; typically it would appear inside the box, however it will fall outside the box if the distribution is highly skewed. All concentrations are shown in parts per billion (ppb). The boxplots for hourly concentrations were created using a different procedure than for the annual statistics, because of the large number of hourly values and the inability of the graphing procedure to allow frequency weights. Therefore, the appropriate weighted percentiles and means were calculated and plotted as shown, but the vertical lines composing the sides of the box are essentially omitted. Tables are provided that summarize the complete distribution, with percentiles apportioned in segments of 10.

List of Figures

List of Tables

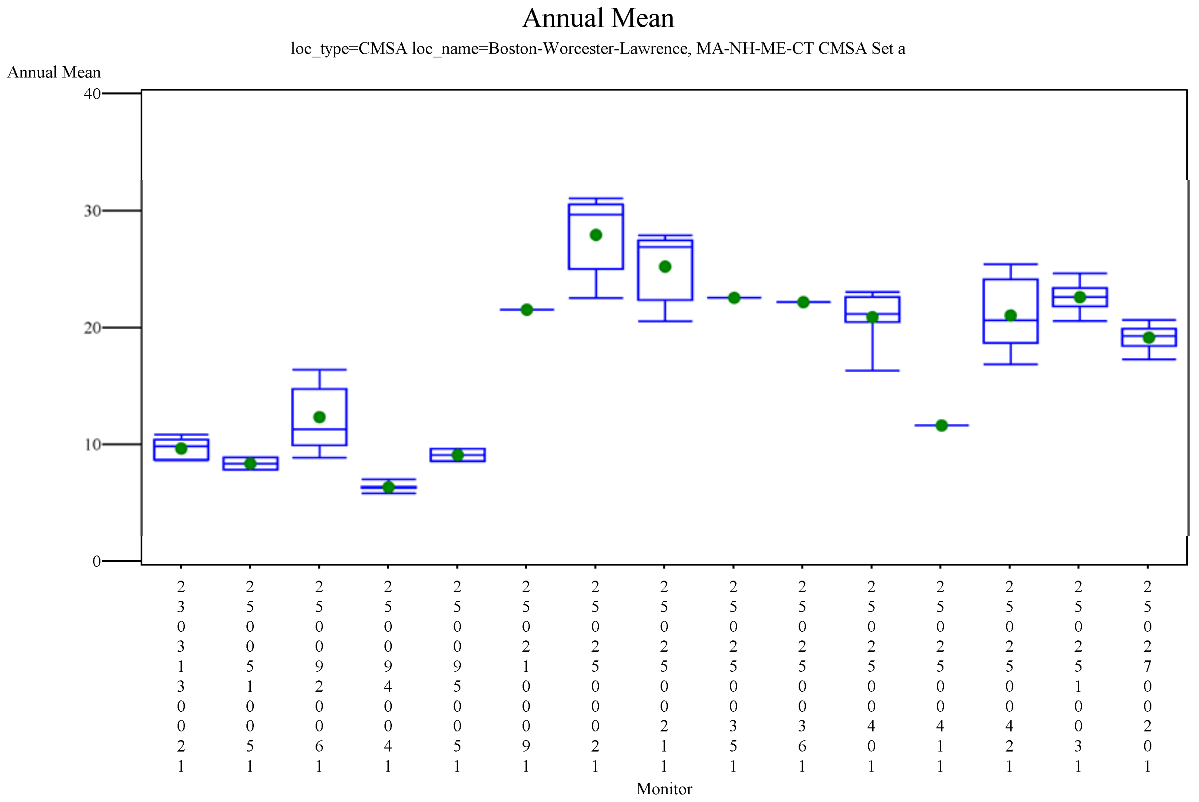


Figure C-1. Spatial distribution of annual average NO_2 concentration, Boston CMSA set a, years 1995-2006.

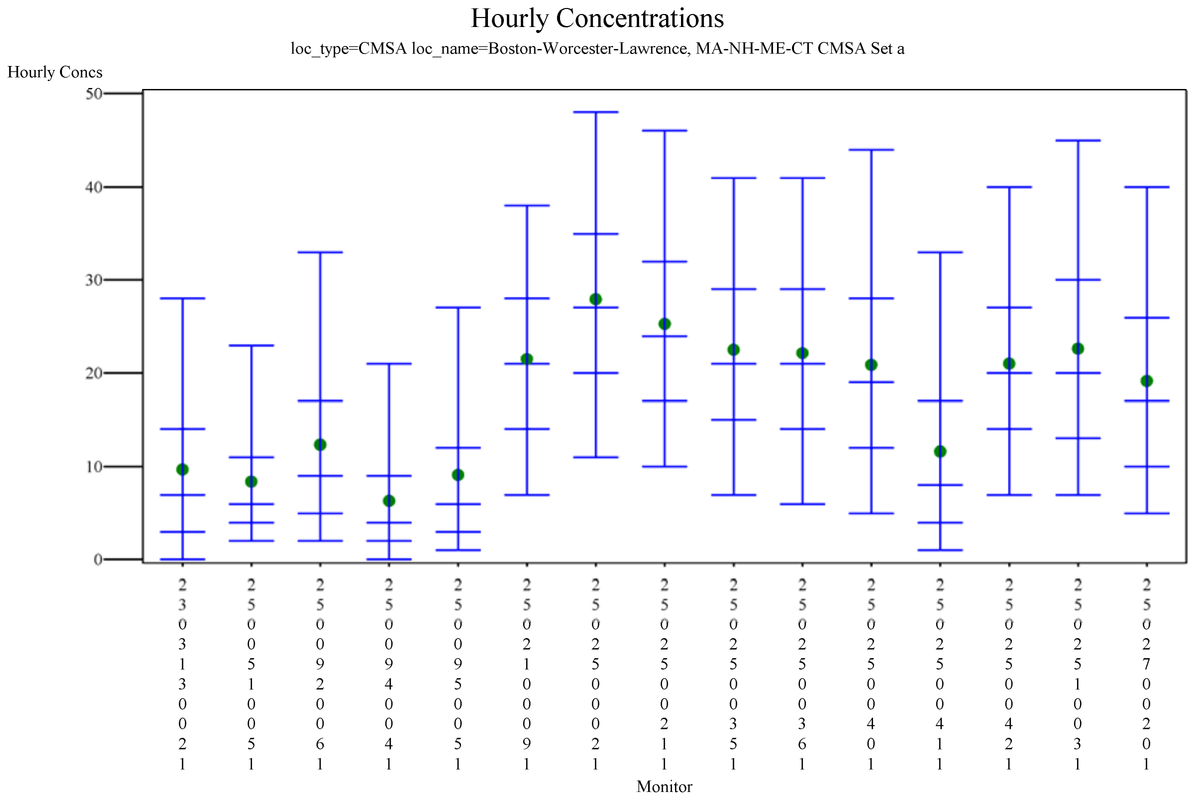


Figure C-2. Spatial distribution of hourly NO_2 concentration, Boston CMSA set a, years 1995-2006.

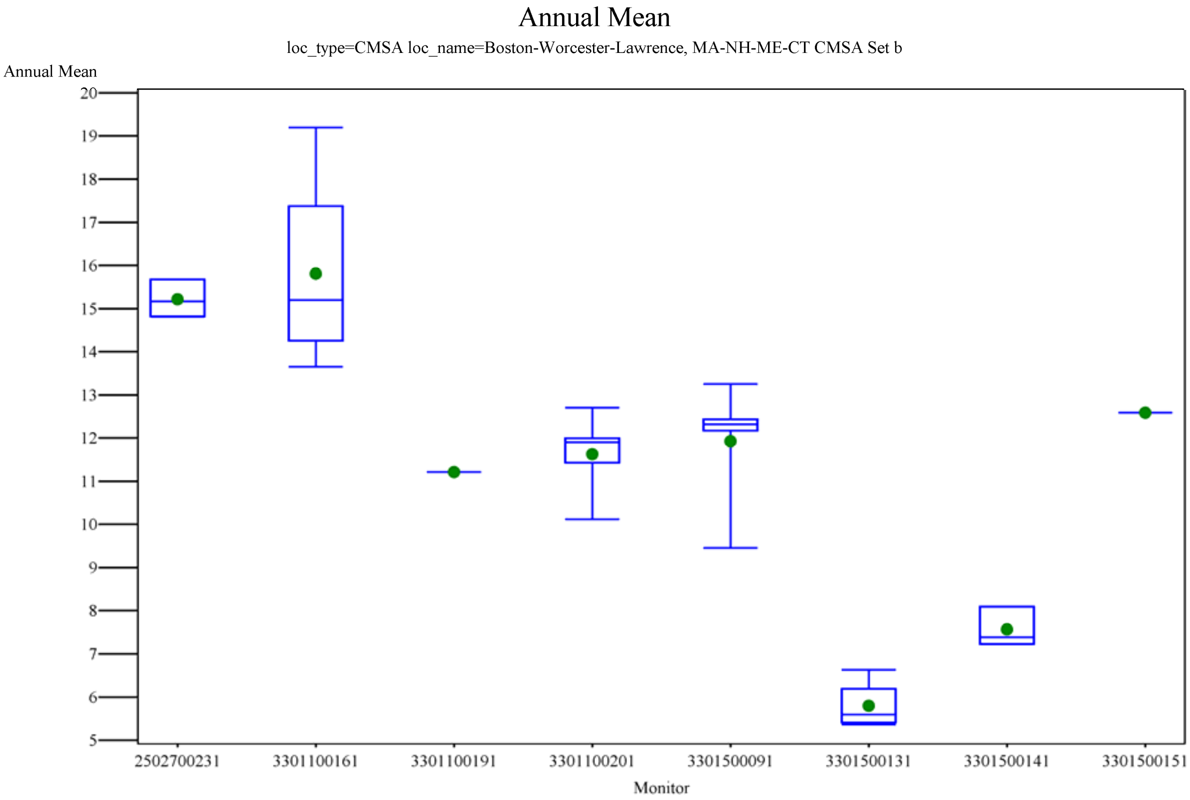


Figure C-3. Spatial distribution of annual average NO_2 concentration, Boston CMSA set b, years 1995-2006.

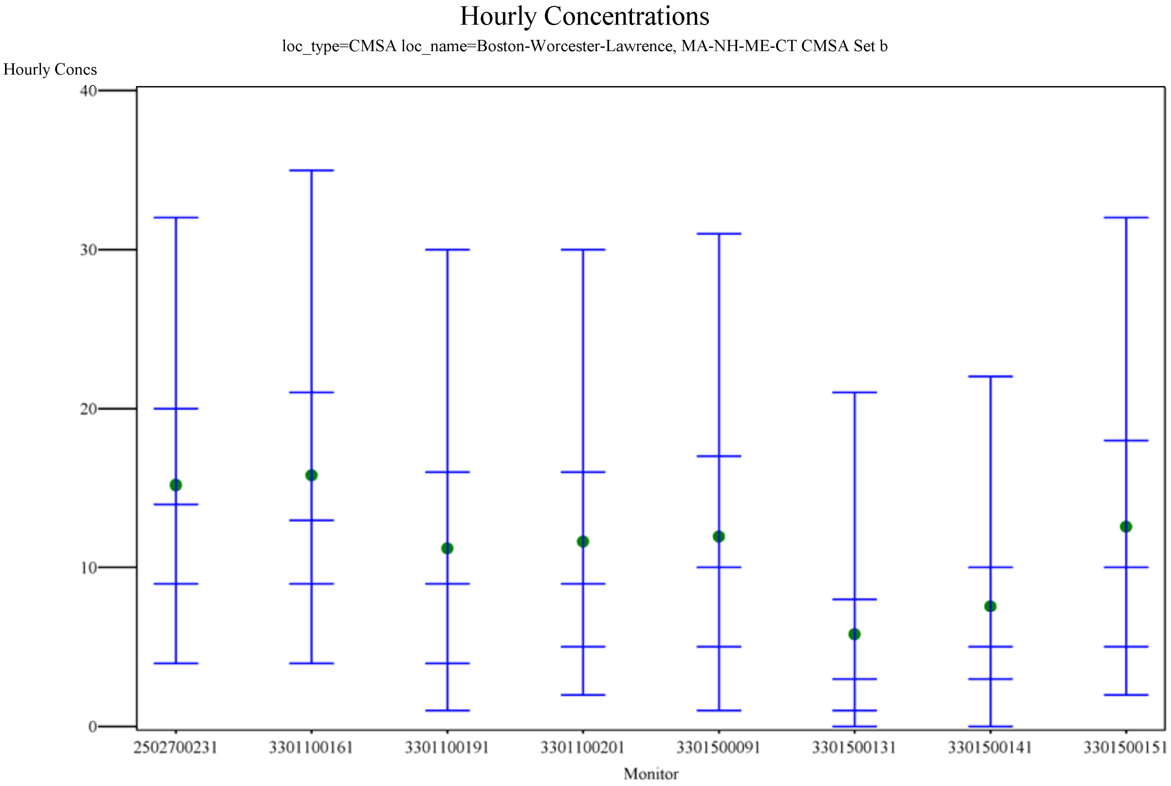


Figure C-4. Spatial distribution of hourly NO_2 concentration, Boston CMSA set b, years 1995-2006.

Table C-1. Spatial distribution of annual average NO_2 concentration, Boston CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
2303130021	7	10	1	9	9	9	9	9	9	10	10	10	10	11	11
2500510051	2	8	1	9	8	8	8	8	8	8	9	9	9	9	9
2500920061	10	12	3	22	9	9	10	10	11	11	13	15	15	16	16
2500940041	5	6	0	7	6	6	6	6	6	6	6	6	7	7	7
2500950051	2	9	1	8	9	9	9	9	9	9	10	10	10	10	10
2502100091	1	22			22	22	22	22	22	22	22	22	22	22	22
2502500021	11	28	3	11	23	23	25	25	29	30	30	30	31	31	31
2502500211	8	25	3	12	21	21	22	23	27	27	27	27	28	28	28
2502500351	1	23			23	23	23	23	23	23	23	23	23	23	23
2502500361	1	22			22	22	22	22	22	22	22	22	22	22	22
2502500401	11	21	2	10	16	18	20	21	21	21	22	22	23	23	23
2502500411	1	12			12	12	12	12	12	12	12	12	12	12	12
2502500421	6	21	3	16	17	17	19	19	19	21	22	24	24	25	25
2502510031	5	23	2	7	21	21	21	22	22	23	23	23	24	25	25
2502700201	8	19	1	6	17	17	18	19	19	19	19	20	20	21	21
2502700231	3	15	0	3	15	15	15	15	15	15	15	16	16	16	16
3301100161	4	16	2	15	14	14	14	15	15	15	16	16	19	19	19
3301100191	1	11			11	11	11	11	11	11	11	11	11	11	11
3301100201	5	12	1	8	10	10	11	11	12	12	12	12	12	13	13
3301500091	5	12	1	12	9	9	11	12	12	12	12	12	13	13	13
3301500131	4	6	1	10	5	5	5	5	5	6	6	6	7	7	7
3301500141	3	8	0	6	7	7	7	7	7	7	7	8	8	8	8
3301500151	1	13			13	13	13	13	13	13	13	13	13	13	13

Table C-2. Spatial distribution of hourly NO_2 concentration, Boston CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
2303130021	58123	10	9	94	0	1	2	4	5	7	9	12	16	23	100
2500510051	16732	8	7	81	0	2	3	4	5	6	8	10	13	18	50
2500920061	80761	12	10	80	0	3	4	6	7	9	12	15	20	27	90
2500940041	41337	6	7	108	0	0	1	2	3	4	6	7	10	16	70
2500950051	16228	9	8	91	0	2	3	4	5	6	8	11	14	22	51
2502100091	8546	22	10	46	0	9	13	15	18	21	23	27	30	35	75
2502500021	87534	28	11	40	0	14	18	21	24	27	30	33	37	43	134
2502500211	63990	25	11	45	0	13	16	18	21	24	26	30	34	40	205
2502500351	8539	23	10	47	0	10	13	16	19	21	24	27	31	37	74
2502500361	8542	22	11	49	0	9	12	15	19	21	24	28	31	36	100
2502500401	91196	21	12	59	1	7	10	13	16	19	22	26	31	38	113
2502500411	8319	12	10	89	0	2	3	5	6	8	11	15	19	27	81
2502500421	48078	21	10	48	0	9	12	15	17	20	22	25	29	35	79
2502510031	40775	23	12	54	0	9	12	14	17	20	24	28	33	40	94
2502700201	63836	19	11	59	0	6	9	11	14	17	21	24	29	35	95
2502700231	24267	15	9	58	0	5	8	10	12	14	16	19	22	27	93
3301100161	33436	16	10	64	0	6	8	9	11	13	16	19	23	29	158
3301100191	8022	11	9	81	0	2	3	5	7	9	11	14	18	24	54
3301100201	41325	12	9	75	0	3	4	6	7	9	11	14	18	25	62
3301500091	40978	12	9	77	0	2	4	6	8	10	12	15	19	25	63
3301500131	33536	6	7	118	0	0	1	2	2	3	5	7	10	15	50
3301500141	25372	8	7	94	0	1	2	3	4	5	7	9	12	17	48
3301500151	8599	13	9	75	0	3	5	6	8	10	12	16	20	27	65

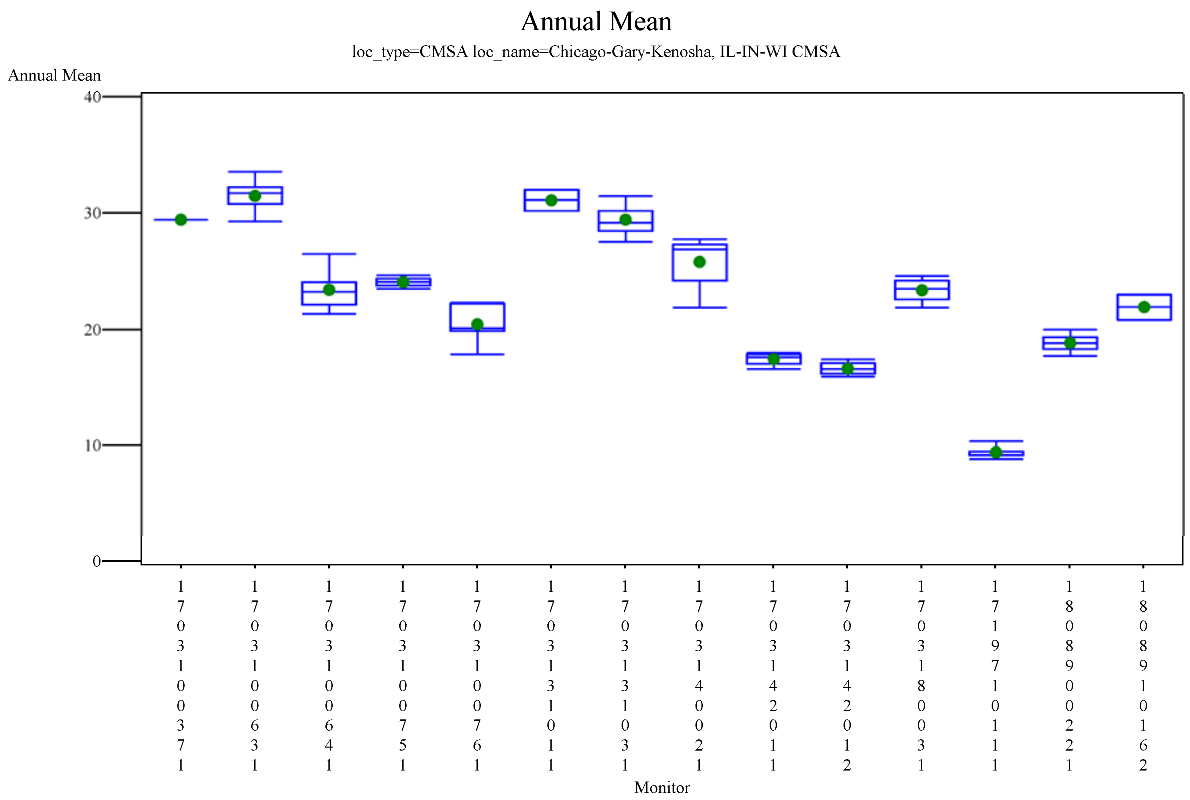


Figure C-5. Spatial distribution of annual average NO_2 concentration, Chicago CMSA, years 1995-2006.

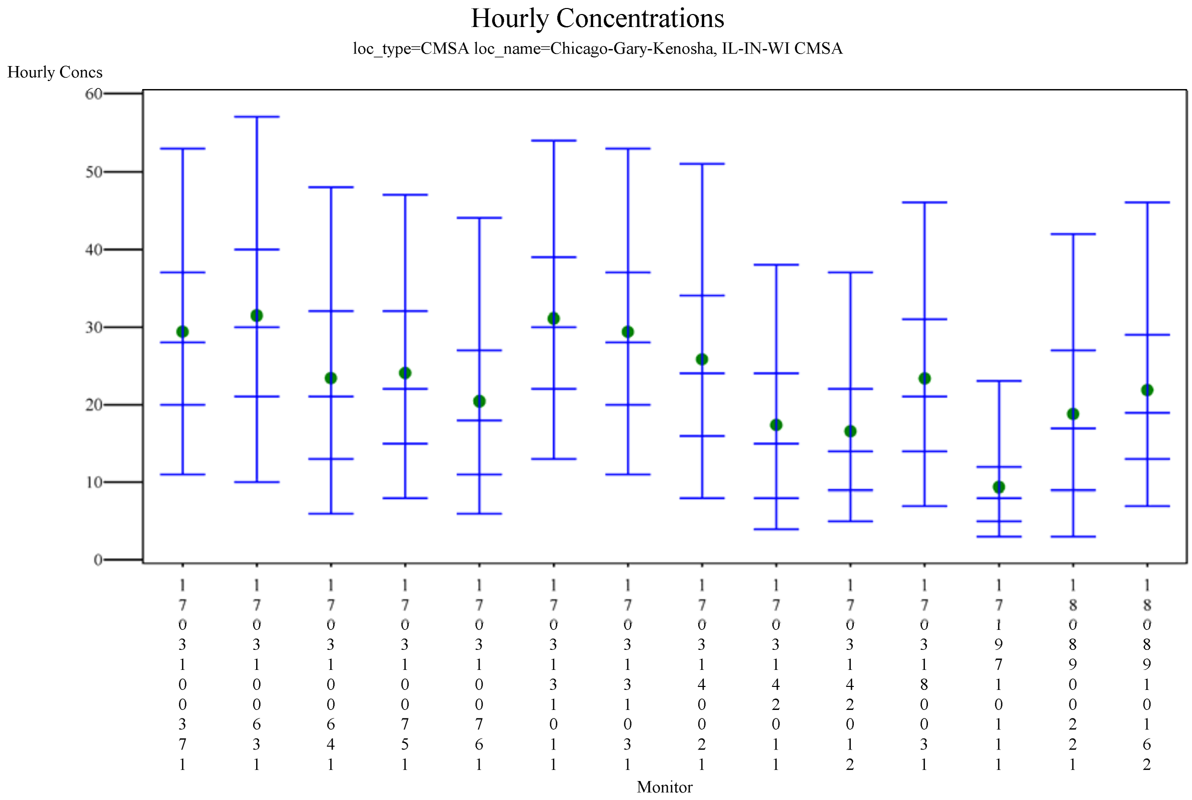


Figure C-6. Spatial distribution of hourly NO_2 concentration, Chicago CMSA, years 1995-2006.

Table C-3. Spatial distribution of annual average NO_2 concentration, Chicago CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1703100371	1	29			29	29	29	29	29	29	29	29	29	29	29
1703100631	12	31	1	4	29	30	31	31	31	32	32	32	32	32	34
1703100641	6	23	2	8	21	21	22	22	23	23	24	24	24	26	26
1703100751	4	24	0	2	23	23	23	24	24	24	24	24	25	25	25
1703100761	5	20	2	9	18	18	19	20	20	20	21	22	22	22	22
1703131011	3	31	1	3	30	30	30	30	31	31	31	32	32	32	32
1703131031	9	29	1	5	28	28	28	28	29	29	30	30	31	31	31
1703140021	12	26	2	8	22	23	24	24	26	27	27	27	27	28	28
1703142011	4	17	1	4	17	17	17	17	17	18	18	18	18	18	18
1703142012	4	17	1	4	16	16	16	16	16	17	17	17	17	17	17
1703180031	8	23	1	4	22	22	22	23	23	23	24	24	24	25	25
1719710111	5	9	1	6	9	9	9	9	9	9	9	9	10	10	10
1808900221	8	19	1	4	18	18	18	18	19	19	19	19	20	20	20
1808910162	2	22	2	7	21	21	21	21	21	22	23	23	23	23	23

Table C-4. Spatial distribution of hourly NO_2 concentration, Chicago CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1703100371	8630	29	13	44	0	15	19	22	25	28	31	35	39	47	113
1703100631	101935	31	15	46	0	13	19	23	27	30	34	38	43	51	137
1703100641	52139	23	13	57	0	8	11	15	18	21	25	29	34	41	127
1703100751	34028	24	12	52	0	10	13	16	19	22	26	29	34	41	113
1703100761	42946	20	12	59	0	7	10	12	15	18	21	25	30	37	98
1703131011	25141	31	13	41	3	16	20	23	27	30	33	37	41	48	105
1703131031	75061	29	13	44	0	14	18	22	25	28	31	35	39	47	149
1703140021	102779	26	13	51	0	11	14	17	20	24	27	31	36	44	106
1703142011	32625	17	11	64	0	5	7	10	12	15	19	22	27	33	77
1703142012	32552	17	10	62	0	6	8	10	12	14	17	20	25	31	70
1703180031	68952	23	12	53	0	9	12	15	18	21	25	29	33	40	97
1719710111	41227	9	6	69	0	3	4	5	6	8	9	11	13	18	52
1808900221	63295	19	12	66	0	4	7	10	13	17	20	25	29	36	131
1808910162	16574	22	12	56	3	9	12	14	16	19	22	26	31	39	125

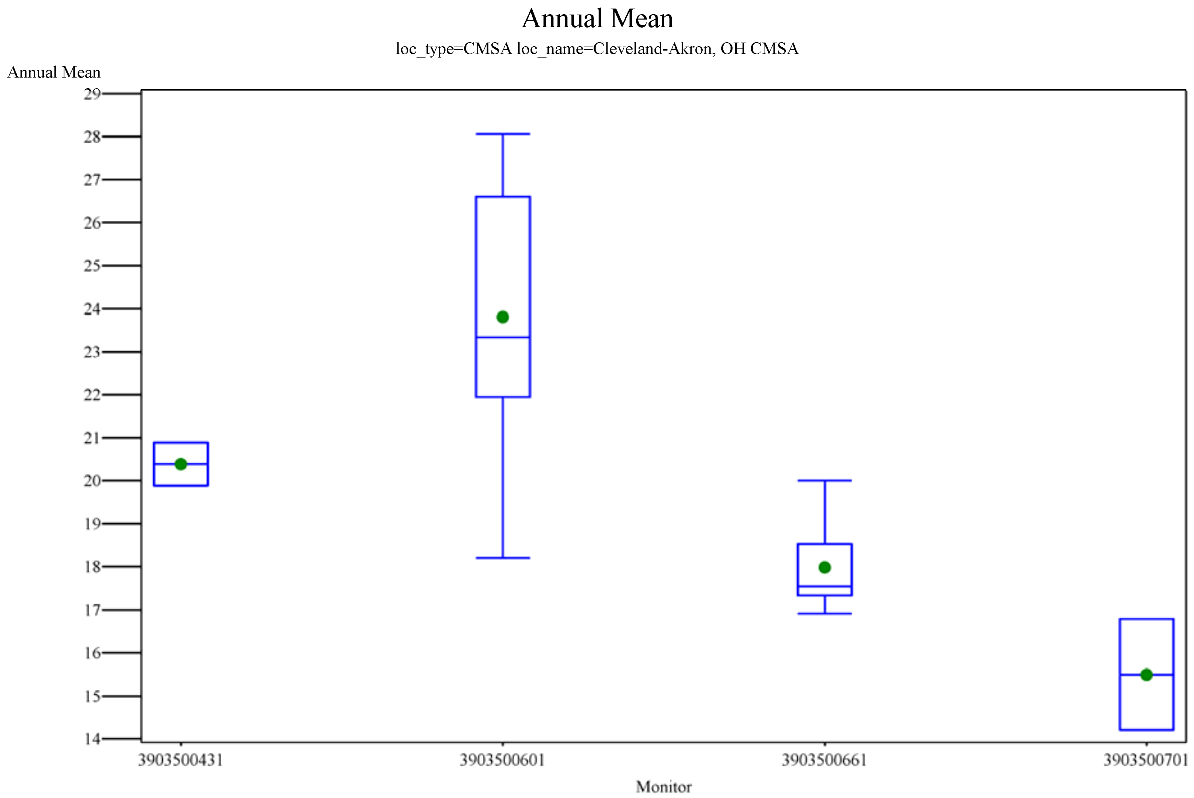


Figure C-7. Spatial distribution of annual average NO_2 concentration, Cleveland CMSA, years 1995-2006.

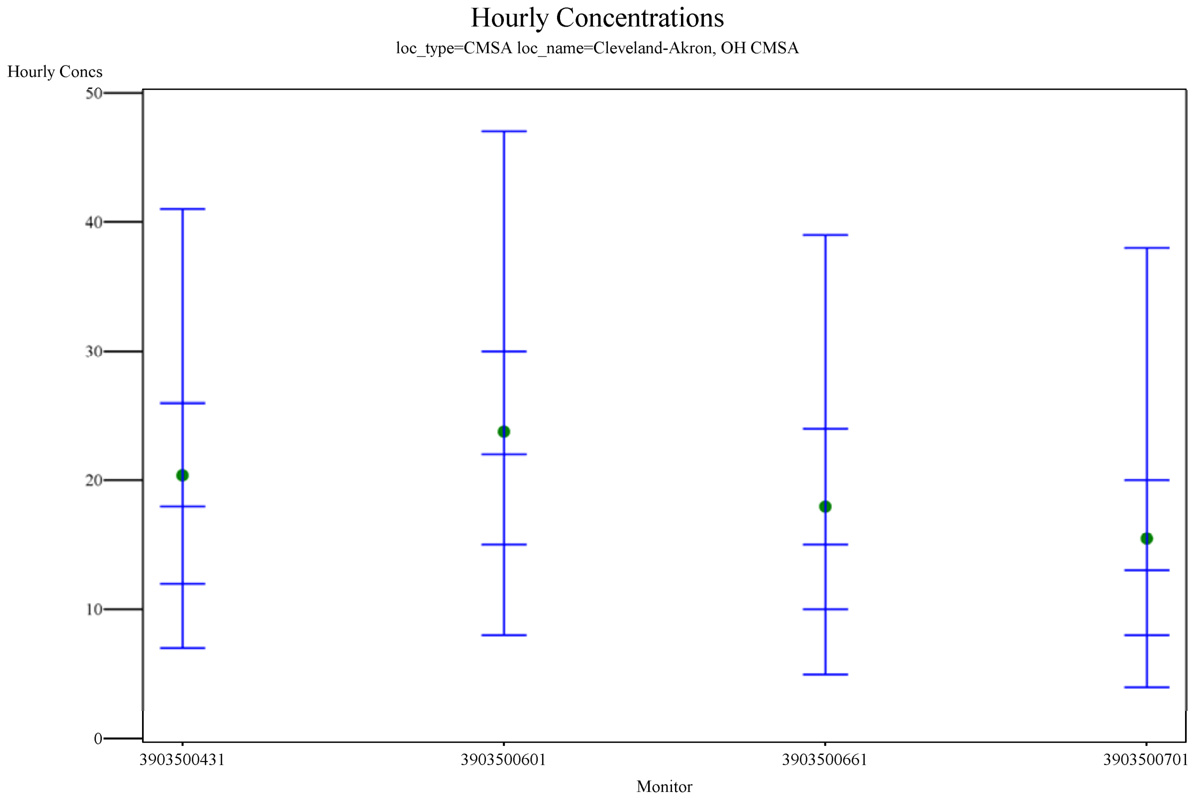


Figure C-8. Spatial distribution of hourly NO_2 concentration, Cleveland CMSA, years 1995-2006.

Table C-5. Spatial distribution of annual average NO_2 concentration, Cleveland CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
3903500431	2	20	1	4	20	20	20	20	20	20	21	21	21	21	21
3903500601	12	24	3	12	18	22	22	22	22	23	25	26	27	27	28
3903500661	6	18	1	6	17	17	17	17	17	18	18	19	19	20	20
3903500701	2	15	2	12	14	14	14	14	14	15	17	17	17	17	17

Table C-6. Spatial distribution of hourly NO_2 concentration, Cleveland CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
3903500431	16215	20	11	54	1	8	11	13	16	18	21	24	28	35	92
3903500601	99696	24	13	53	0	10	13	16	19	22	25	28	33	40	253
3903500661	50100	18	11	60	0	7	9	11	13	15	18	22	26	33	103
3903500701	16619	15	11	70	0	5	7	9	10	13	15	18	23	30	175

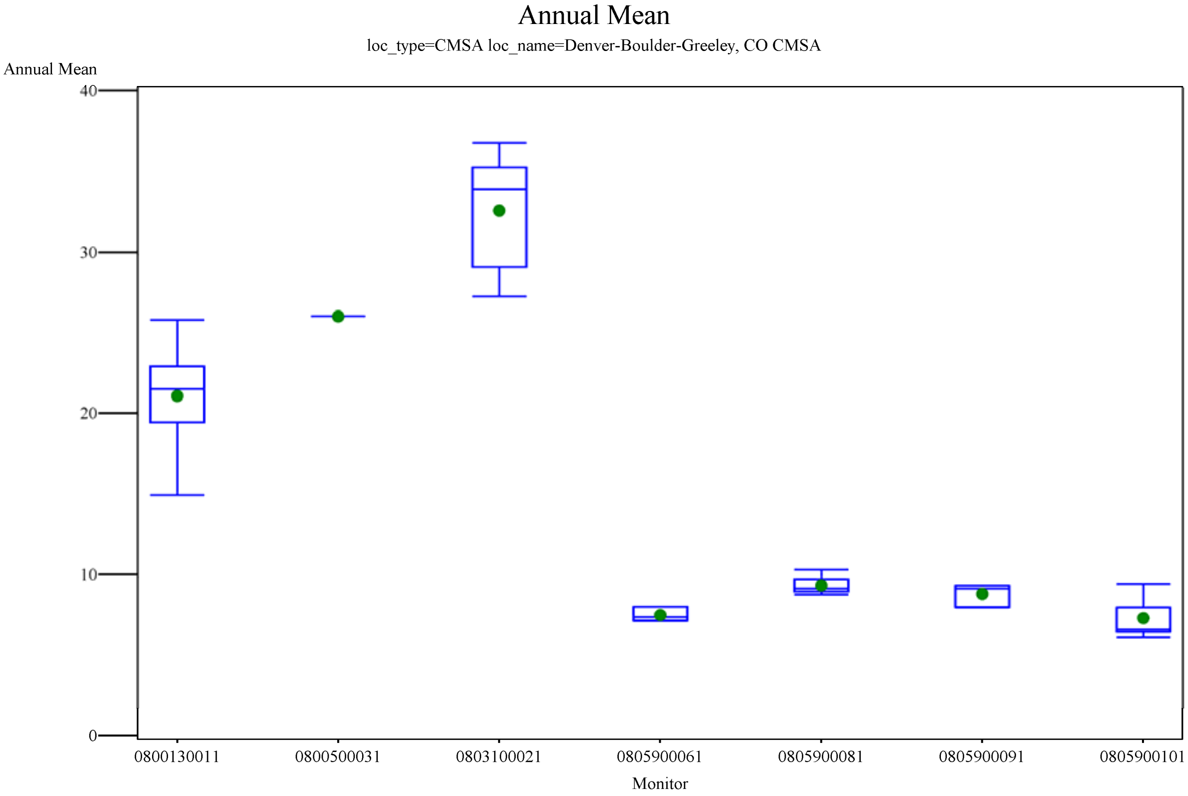


Figure C-9. Spatial distribution of annual average NO_2 concentration, Denver CMSA, years 1995-2006.

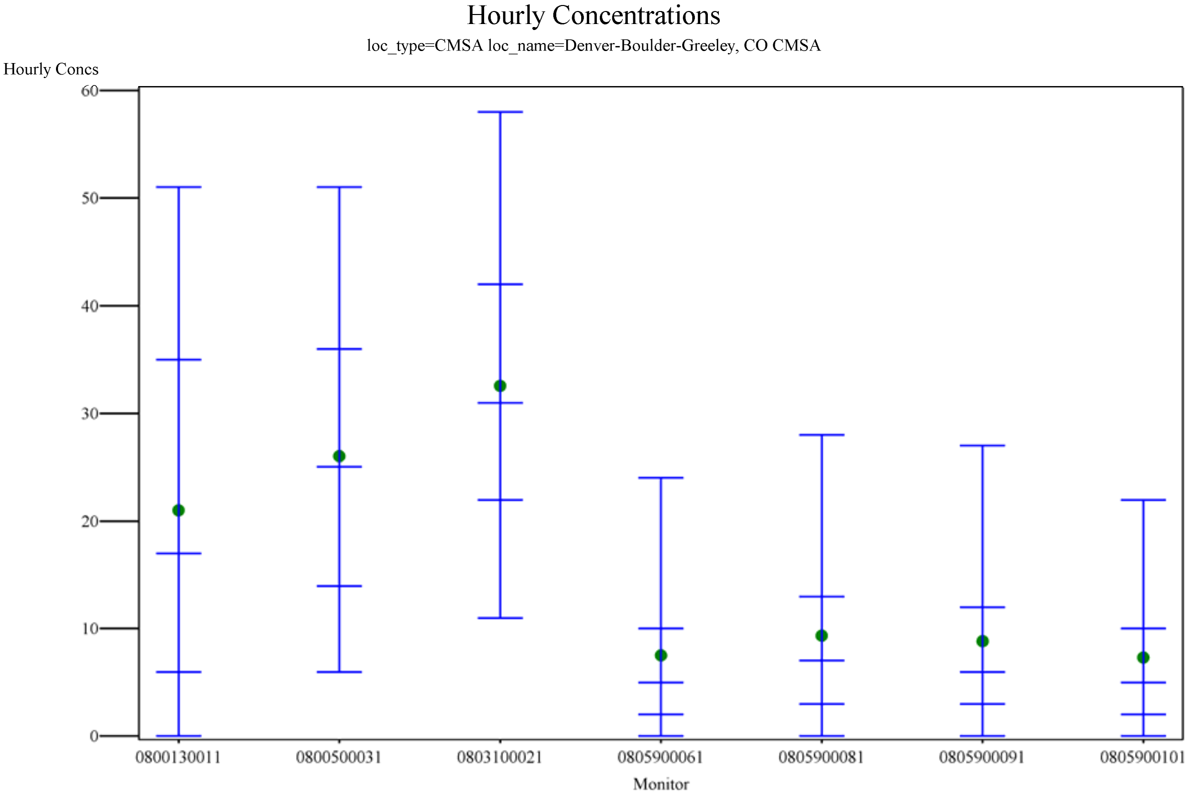


Figure C-10. Spatial distribution of hourly NO_2 concentration, Denver CMSA, years 1995-2006.

Table C-7. Spatial distribution of annual average NO_2 concentration, Denver CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
0800130011	11	21	3	14	15	18	19	20	21	21	22	23	23	23	26
0800500031	1	26			26	26	26	26	26	26	26	26	26	26	26
0803100021	9	33	4	11	27	27	28	29	33	34	35	35	35	37	37
0805900061	3	7	0	6	7	7	7	7	7	7	7	8	8	8	8
0805900081	4	9	1	7	9	9	9	9	9	9	9	9	10	10	10
0805900091	3	9	1	8	8	8	8	8	9	9	9	9	9	9	9
0805900101	5	7	1	19	6	6	6	6	7	7	7	8	9	9	9

Table C-8. Spatial distribution of hourly NO_2 concentration, Denver CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
0800130011	83703	21	17	82	0	2	4	7	11	17	25	32	38	45	239
0800500031	7790	26	15	57	0	8	12	16	20	25	29	34	39	45	176
0803100021	68630	33	15	46	0	15	20	24	28	31	35	39	44	51	286
0805900061	22077	7	8	109	0	1	1	3	4	5	6	9	12	18	66
0805900081	32449	9	9	97	0	0	2	3	5	7	9	12	15	22	68
0805900091	24368	9	9	100	0	1	2	3	5	6	8	10	14	20	88
0805900101	39124	7	8	106	0	1	2	2	4	5	6	9	12	17	98

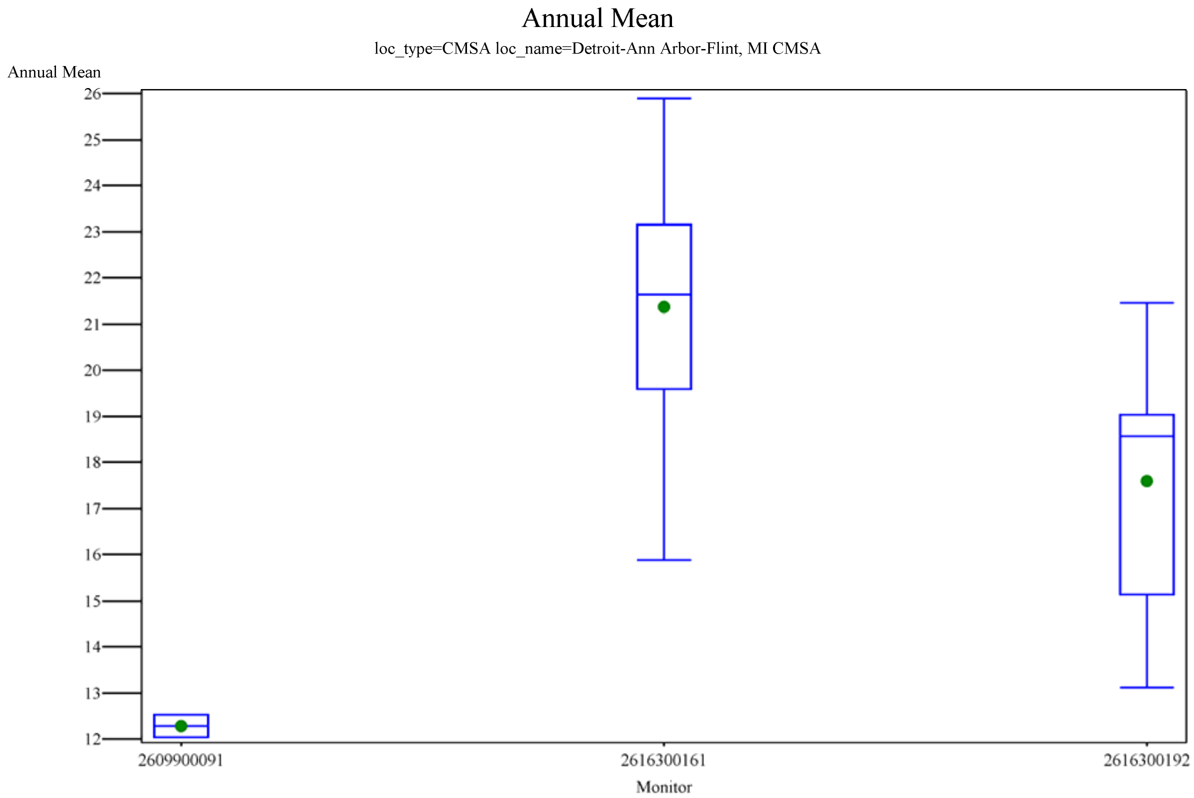


Figure C-11. Spatial distribution of annual average NO_2 concentration, Detroit CMSA, years 1995-2006.

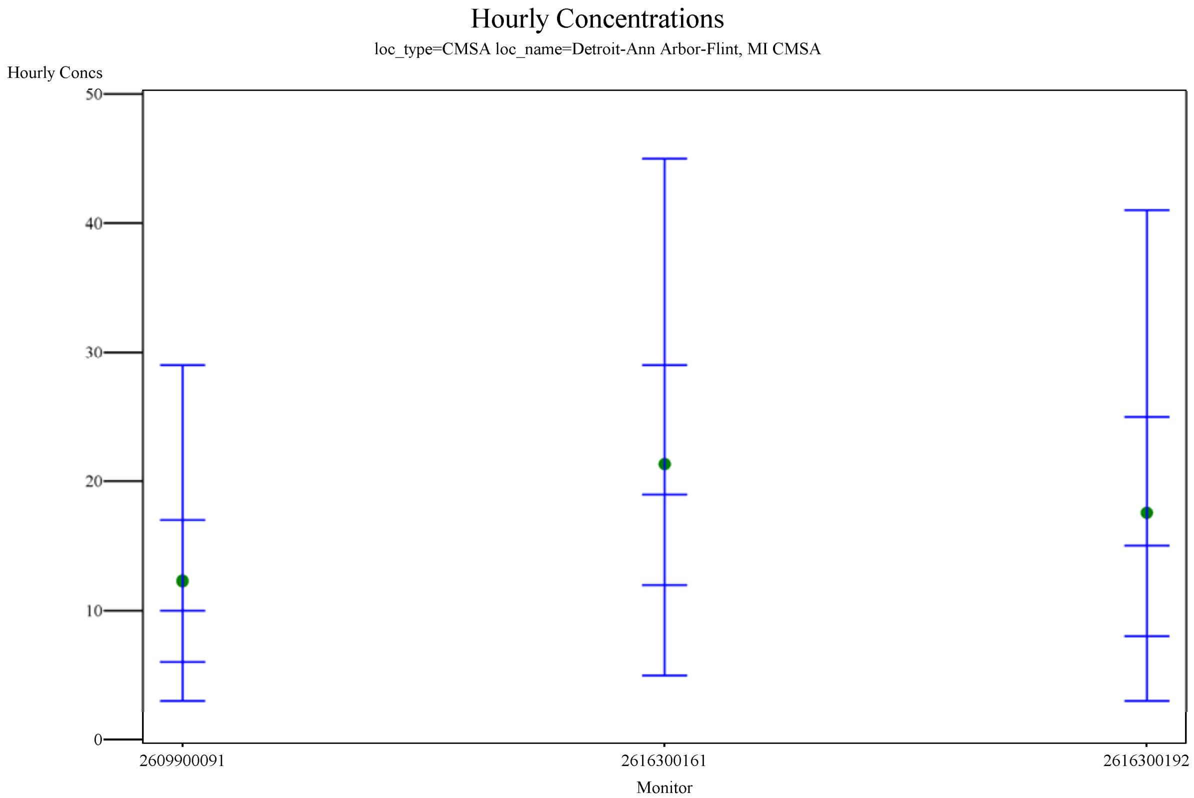


Figure C-12. Spatial distribution of annual average NO_2 concentration, Detroit CMSA, years 1995-2006.

Table C-9. Spatial distribution of annual average NO_2 concentration, Detroit CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
2609900091	2	12	0	3	12	12	12	12	12	12	13	13	13	13	13
2616300161	11	21	3	13	16	19	20	20	21	22	22	23	23	24	26
2616300192	11	18	3	14	13	14	15	17	18	19	19	19	19	19	21

Table C-10. Spatial distribution of annual average NO_2 concentration, Detroit CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
2609900091	16523	12	9	75	0	3	5	6	8	10	12	15	19	25	322
2616300161	86487	21	13	62	0	7	10	13	16	19	23	26	31	38	244
2616300192	88124	18	13	75	0	5	7	9	12	15	18	22	27	35	443

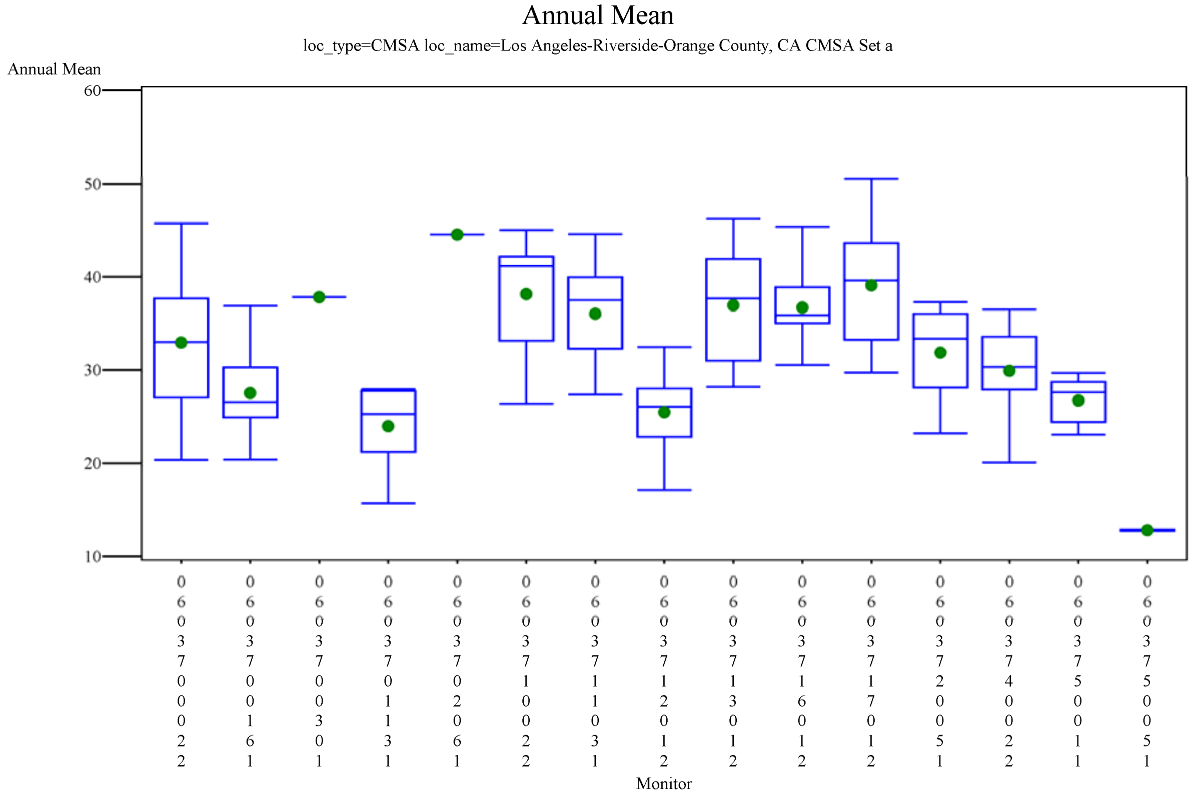


Figure C-13. Spatial distribution of annual average NO_2 concentration, Los Angeles CMSA set a, years 1995-2006.

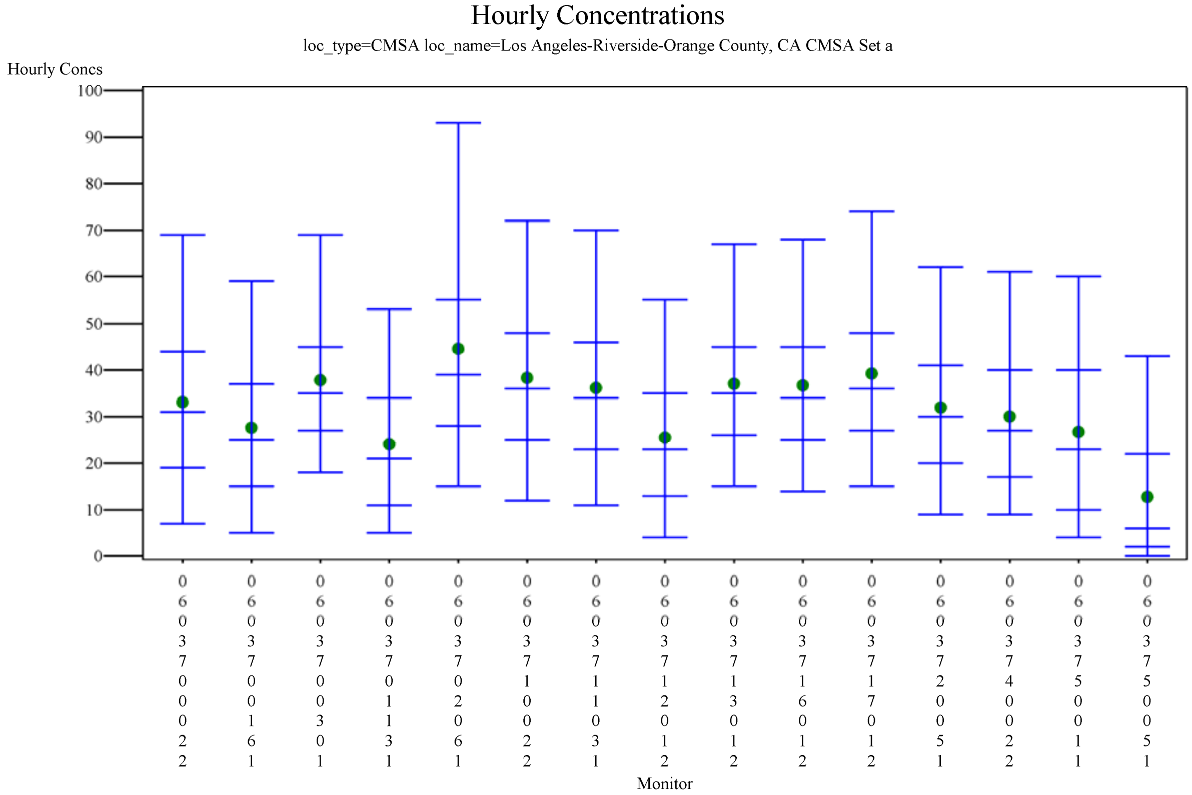


Figure C-14. Spatial distribution of hourly NO_2 concentration, Los Angeles CMSA set a, years 1995-2006.

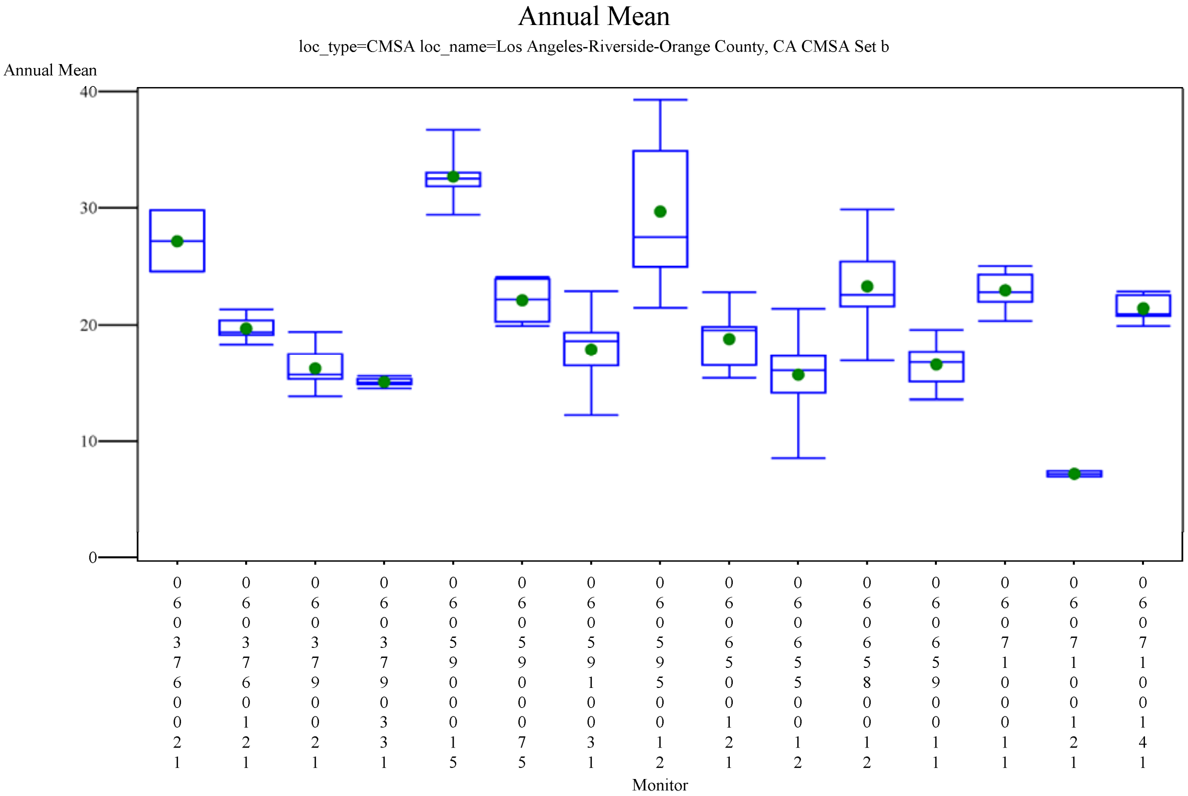


Figure C-15. Spatial distribution of annual average NO_2 concentration, Los Angeles CMSA set b, years 1995-2006.

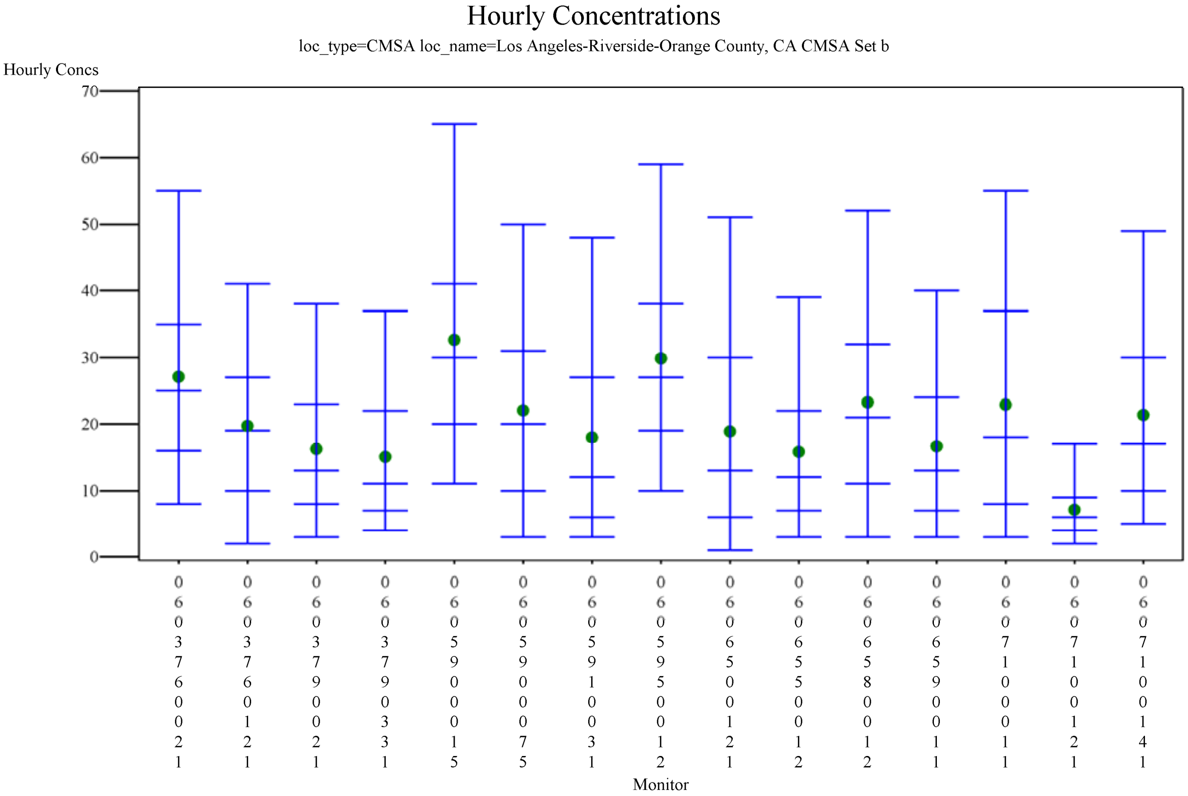


Figure C-16. Spatial distribution of hourly NO_2 concentration, Los Angeles CMSA set b, years 1995-2006.

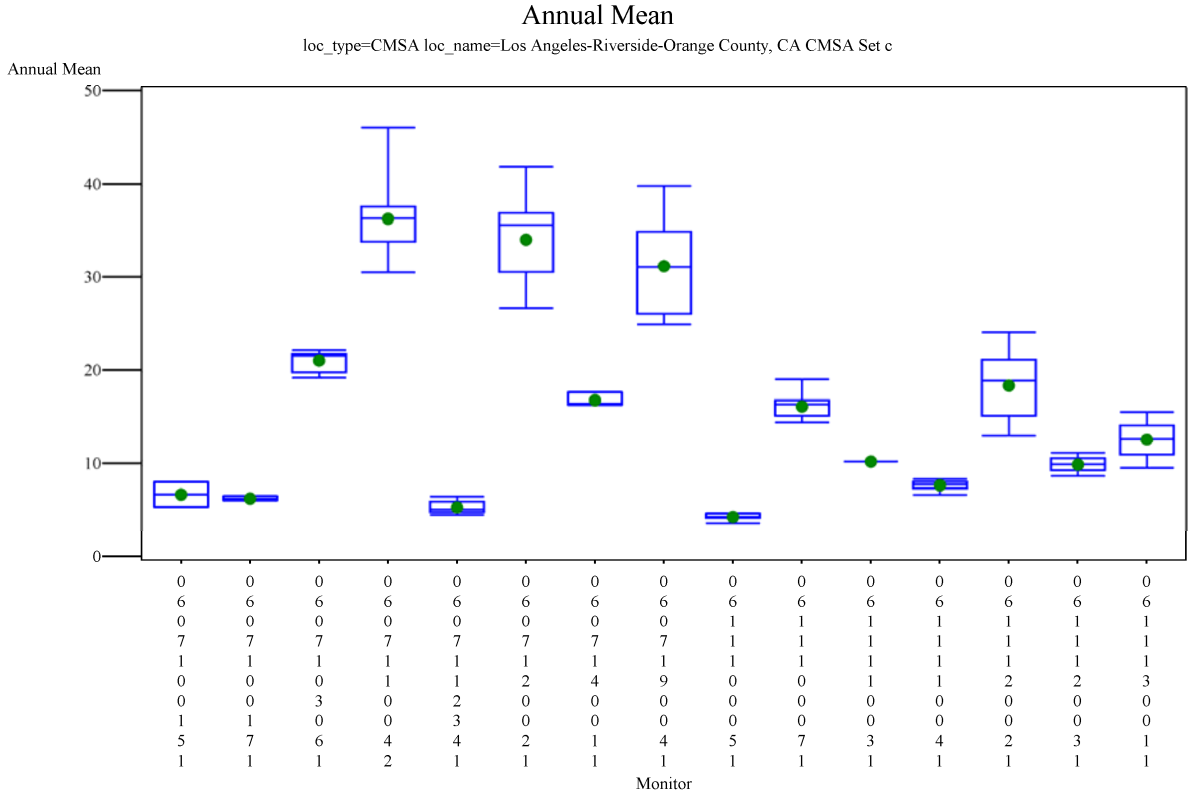


Figure C-17. Spatial distribution of annual average NO_2 concentration, Los Angeles CMSA set c, years 1995-2006.

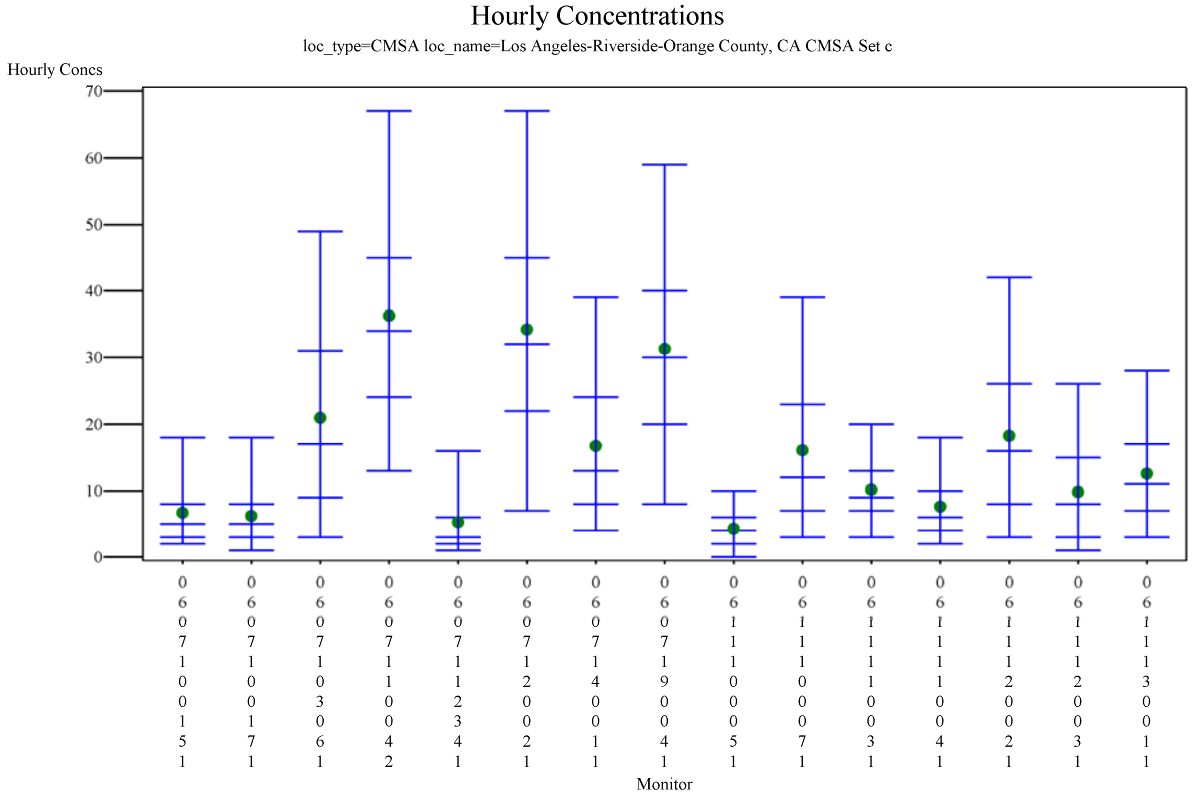


Figure C-18. Spatial distribution of hourly NO_2 concentration, Los Angeles CMSA set c, years 1995-2006.

Table C-11. Spatial distribution of annual average NO_2 concentration, Los Angeles CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
0603700022	12	33	7	22	20	25	25	29	33	33	36	36	39	41	46
0603700161	12	28	5	17	20	22	24	26	26	27	28	29	32	33	37
0603700301	1	38			38	38	38	38	38	38	38	38	38	38	38
0603701131	12	24	4	18	16	17	20	23	24	25	26	28	28	28	28
0603702061	1	45			45	45	45	45	45	45	45	45	45	45	45
0603710022	11	38	6	16	26	29	33	35	40	41	41	41	42	45	45
0603711031	11	36	6	16	27	27	32	33	34	37	39	39	40	43	45
0603712012	12	26	4	17	17	20	21	24	25	26	26	28	28	31	32
0603713012	12	37	6	16	28	30	31	31	36	38	39	41	43	43	46
0603716012	10	37	4	11	31	33	35	35	35	36	37	38	39	42	45
0603717012	12	39	7	17	30	31	31	35	36	40	43	43	44	46	51
0603720051	12	32	5	15	23	24	27	29	32	33	34	35	37	37	37
0603740022	11	30	5	16	20	24	28	29	29	30	32	33	34	34	37
0603750011	9	27	2	9	23	23	23	24	27	28	28	29	29	30	30
0603750051	2	13	0	1	13	13	13	13	13	13	13	13	13	13	13
0603760021	2	27	4	14	25	25	25	25	25	27	30	30	30	30	30
0603760121	5	20	1	6	18	18	19	19	19	19	20	20	21	21	21
0603790021	6	16	2	12	14	14	15	15	16	16	16	18	18	19	19
0603790331	5	15	0	3	15	15	15	15	15	15	15	15	15	16	16
0605900015	5	33	3	8	29	29	31	32	32	33	33	33	35	37	37
0605900075	4	22	2	10	20	20	20	21	21	22	24	24	24	24	24
0605910031	12	18	3	16	12	13	16	17	18	19	19	19	20	20	23
0605950012	11	30	6	19	21	25	25	25	27	28	33	34	35	35	39
0606500121	9	19	3	14	15	15	16	17	18	20	20	20	22	23	23
0606550012	12	16	3	22	9	12	13	15	16	16	16	17	18	20	21
0606580012	12	23	4	16	17	19	21	22	22	23	24	25	26	29	30
0606590011	12	17	2	11	14	14	15	15	17	17	17	18	18	19	20
0607100011	12	23	1	6	20	21	22	22	22	23	24	24	24	25	25
0607100121	2	7	0	5	7	7	7	7	7	7	7	7	7	7	7
0607100141	5	21	1	6	20	20	20	21	21	21	22	23	23	23	23
0607100151	2	7	2	28	5	5	5	5	5	7	8	8	8	8	8
0607100171	3	6	0	4	6	6	6	6	6	6	6	7	7	7	7
0607103061	7	21	1	5	19	19	20	21	21	22	22	22	22	22	22
0607110042	11	36	4	12	31	31	34	34	36	36	37	38	38	39	46
0607112341	9	5	1	12	5	5	5	5	5	5	5	6	6	6	6
0607120021	12	34	5	13	27	27	30	31	33	36	36	36	38	38	42
0607140011	3	17	1	4	16	16	16	16	16	16	16	18	18	18	18
0607190041	12	31	5	16	25	26	26	26	29	31	33	34	35	38	40
0611100051	7	4	0	8	4	4	4	4	4	4	4	4	5	5	5
0611100071	9	16	1	9	14	14	14	15	16	16	16	17	17	19	19
0611110031	1	10			10	10	10	10	10	10	10	10	10	10	10
0611110041	7	8	1	7	7	7	7	8	8	8	8	8	8	8	8
0611120021	12	18	4	20	13	14	15	15	17	19	20	20	22	22	24
0611120031	9	10	1	8	9	9	9	9	9	10	10	11	11	11	11
0611130011	12	13	2	16	9	10	11	11	11	13	14	14	14	15	16

Table C-12. Spatial distribution of hourly NO_2 concentration, Los Angeles CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
0603700022	97734	33	20	59	0	11	16	21	26	31	35	41	47	58	223
0603700161	97838	28	18	63	0	8	13	17	21	25	29	34	40	50	196
0603700301	6817	38	17	44	8	21	25	28	32	35	38	42	48	57	160
0603701131	97124	24	16	67	0	7	9	12	16	21	26	32	37	45	201
0603702061	7604	45	25	56	0	19	25	30	34	39	45	51	60	75	208
0603710022	88656	38	19	49	0	17	23	28	32	36	41	45	52	62	262
0603711031	88425	36	19	52	0	15	20	25	30	34	38	43	49	60	239
0603712012	96922	26	16	64	0	7	11	15	19	23	28	33	38	47	163
0603713012	97352	37	17	45	0	19	24	28	31	35	39	43	48	57	250
0603716012	81411	37	18	48	0	17	23	27	31	34	38	42	48	58	225
0603717012	98551	39	18	47	0	19	25	29	33	36	40	45	52	63	184
0603720051	98151	32	17	54	0	13	18	22	26	30	34	38	44	52	225
0603740022	88730	30	17	58	0	12	16	19	23	27	31	37	43	52	208
0603750011	74014	27	19	72	0	5	9	12	17	23	30	37	43	51	178
0603750051	15047	13	15	114	0	0	1	2	4	6	10	17	26	36	91
0603760021	16534	27	15	57	0	10	14	18	21	25	28	32	37	46	159
0603760121	39399	20	12	61	0	4	9	12	16	19	22	25	30	36	120
0603790021	46871	16	11	69	0	5	7	9	11	13	17	21	26	32	140
0603790331	40341	15	11	73	0	5	6	7	9	11	14	18	25	32	103
0605900015	40987	33	17	53	0	14	19	22	26	30	34	38	44	55	175
0605900075	33847	22	15	70	0	5	9	10	14	20	23	30	36	42	127
0605910031	97546	18	15	85	0	4	6	7	9	12	16	23	31	40	183
0605950012	88510	30	16	54	0	12	17	20	24	27	31	35	41	50	192
0606500121	69857	19	17	91	0	3	5	7	10	13	18	25	34	43	307
0606550012	95624	16	12	73	0	4	6	8	10	12	15	19	25	33	82
0606580012	95642	23	16	67	0	6	10	13	17	21	25	30	35	44	150
0606590011	95010	17	13	75	0	4	6	8	10	13	17	22	27	34	127
0607100011	94741	23	17	76	0	5	7	9	12	18	25	33	40	48	196
0607100121	14753	7	5	69	0	2	4	4	5	6	7	8	10	14	57
0607100141	39719	21	14	67	0	7	9	11	14	17	22	27	33	41	113
0607100151	15531	7	6	82	0	2	3	3	4	5	6	7	10	14	60
0607100171	23713	6	5	84	0	2	3	3	4	5	6	7	9	13	73
0607103061	56831	21	15	70	0	5	8	11	13	17	22	28	34	42	100
0607110042	88766	36	17	48	0	17	22	26	30	34	38	43	49	58	199
0607112341	69325	5	5	103	0	1	2	2	3	3	4	5	7	12	62
0607120021	95054	34	18	54	0	12	19	24	28	32	37	42	48	58	170
0607140011	24587	17	11	68	0	6	7	9	11	13	16	21	27	34	86
0607190041	97785	31	16	51	0	12	18	22	26	30	33	38	43	51	162
0611100051	54034	4	4	89	0	0	1	3	3	4	5	5	6	8	81
0611100071	73031	16	12	74	0	4	6	8	10	12	16	20	26	33	123
0611110031	8240	10	5	52	0	4	6	7	8	9	10	12	14	16	61
0611110041	56869	8	5	66	0	3	4	5	6	6	7	9	11	14	66
0611120021	94238	18	13	70	0	4	7	9	12	16	19	24	29	36	124
0611120031	70332	10	8	85	0	1	2	4	6	8	10	13	17	21	93
0611130011	95263	13	8	65	0	4	6	7	9	11	13	15	18	23	127

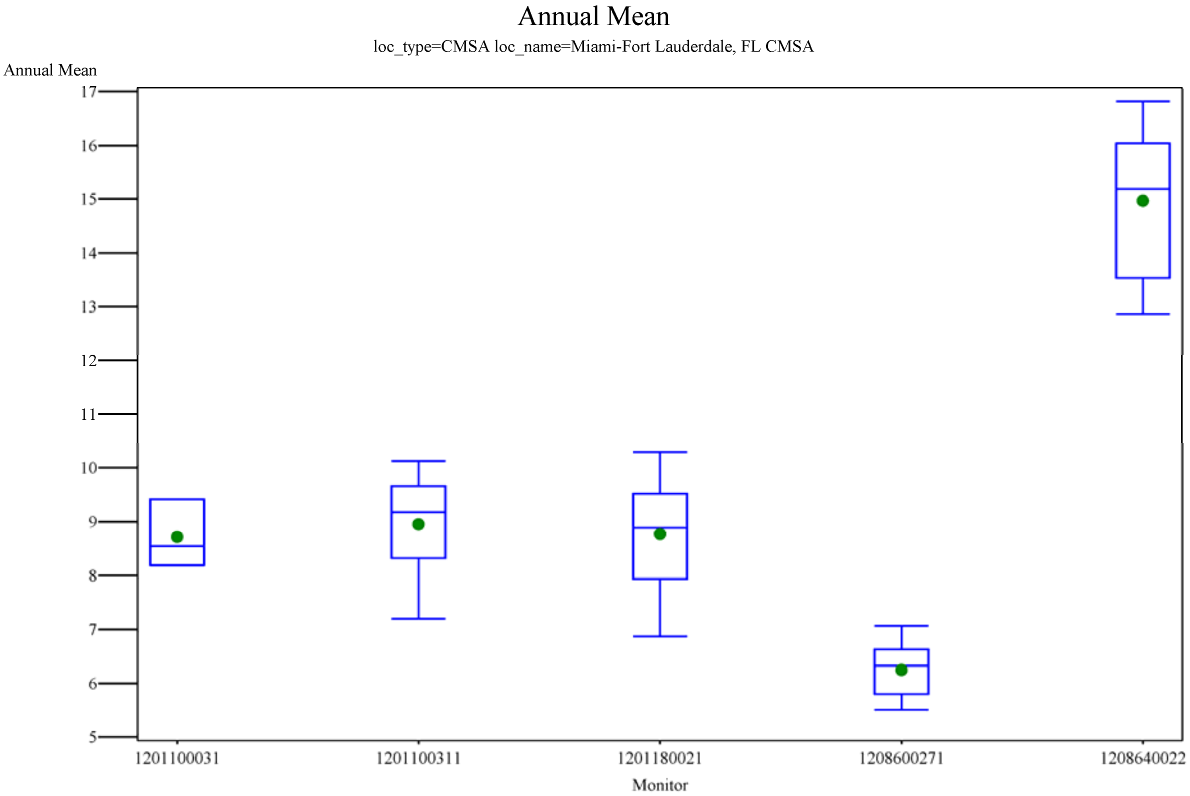


Figure C-19. Spatial distribution of annual average NO_2 concentration, Miami CMSA, years 1995-2006.

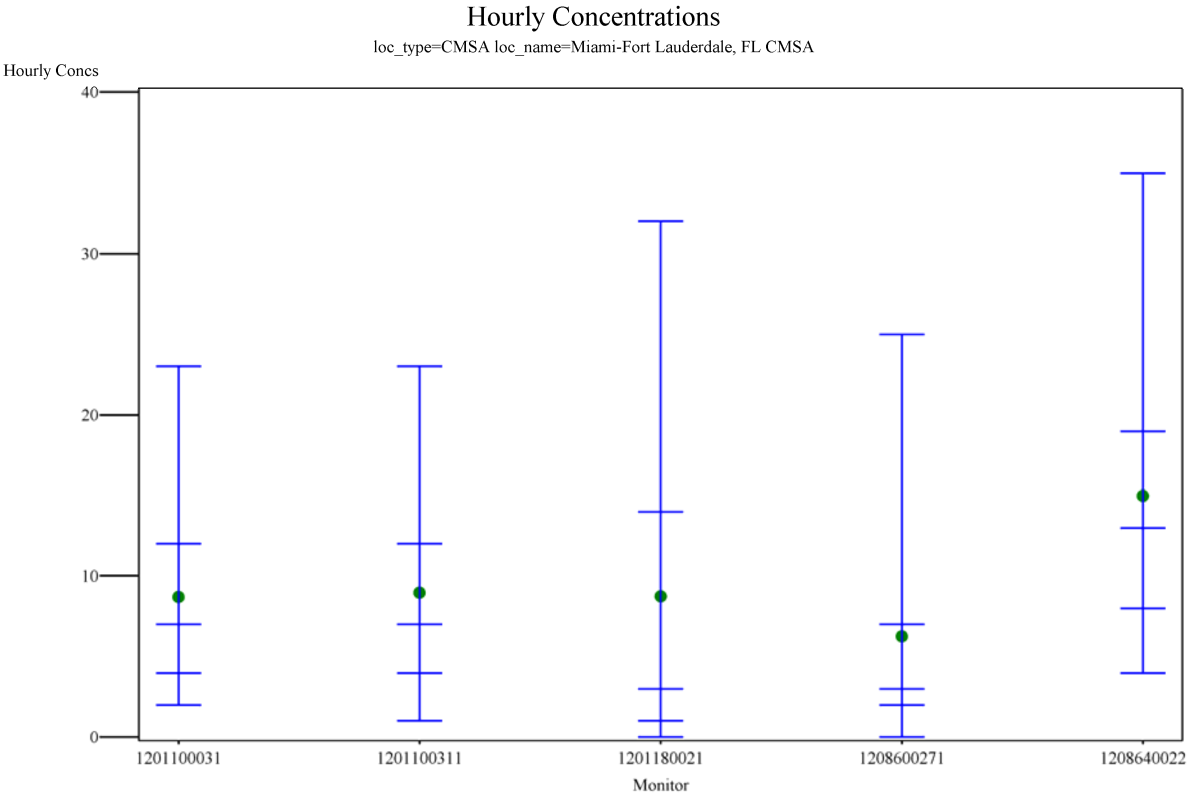


Figure C-20. Spatial distribution of hourly NO_2 concentration, Miami CMSA, years 1995-2006.

Table C-13. Spatial distribution of annual average NO_2 concentration, Miami CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1201100031	3	9	1	7	8	8	8	8	9	9	9	9	9	9	9
1201100311	8	9	1	12	7	7	8	9	9	9	9	9	10	10	10
1201180021	11	9	1	11	7	8	8	8	9	9	9	9	10	10	10
1208600271	11	6	0	7	6	6	6	6	6	6	6	7	7	7	7
1208640022	11	15	1	9	13	13	14	14	15	15	16	16	16	17	17

Table C-14. Spatial distribution of hourly NO_2 concentration, Miami CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1201100031	24440	9	7	81	0	2	3	4	5	7	8	10	13	18	65
1201100311	63306	9	7	78	0	2	3	5	6	7	9	11	14	18	64
1201180021	92241	9	11	128	0	0	1	1	2	3	5	11	18	26	128
1208600271	87068	6	8	132	0	1	1	2	2	3	4	5	9	17	75
1208640022	90717	15	10	67	0	5	7	9	11	13	15	18	22	28	417

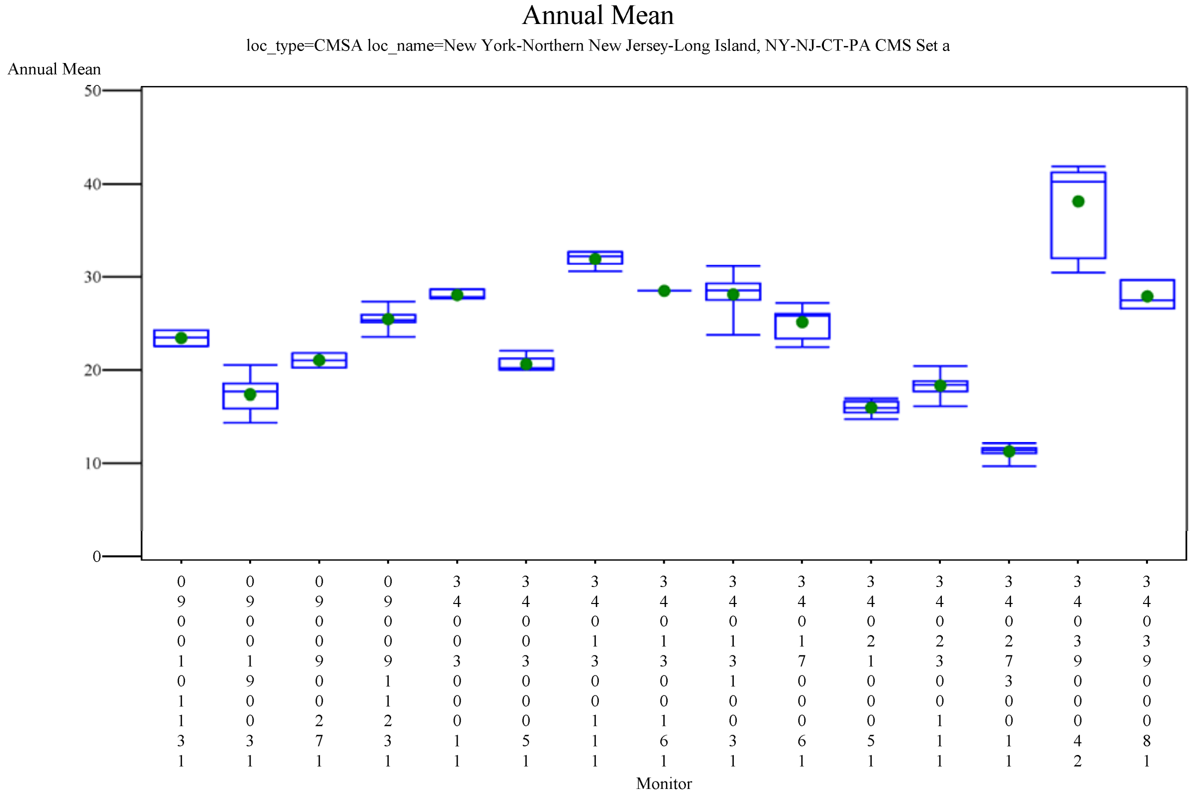


Figure C-21. Spatial distribution of annual average NO_2 concentration, New York CMSA set a, years 1995-2006.

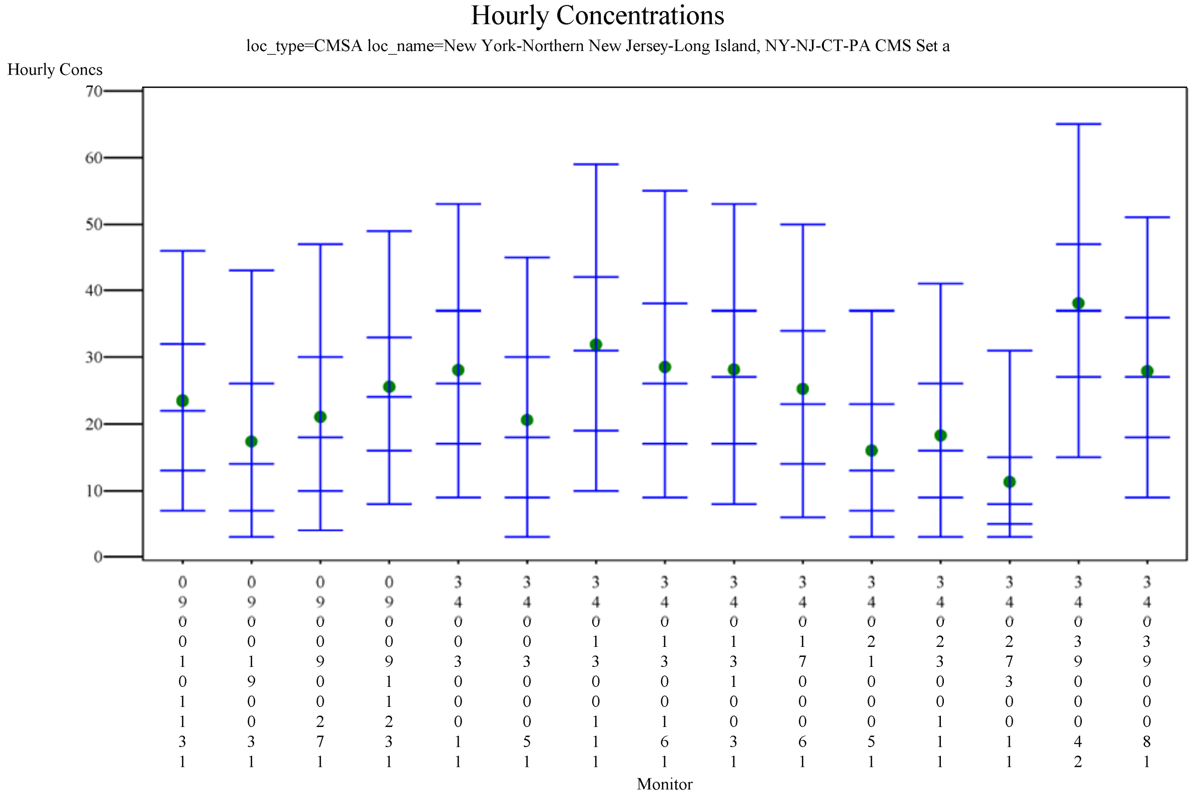


Figure C-22. Spatial distribution of hourly NO_2 concentration, New York CMSA set a, years 1995-2006.

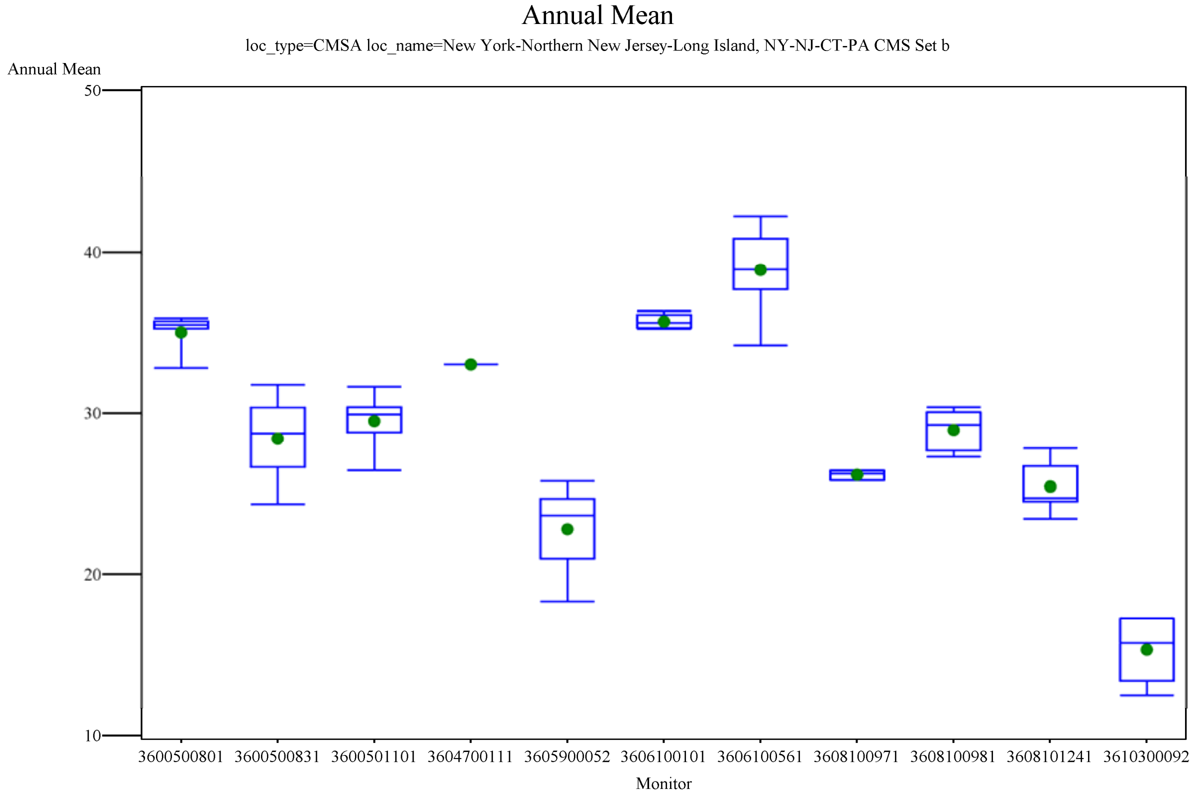


Figure C-23. Spatial distribution of annual average NO_2 concentration, New York CMSA set b, years 1995-2006.

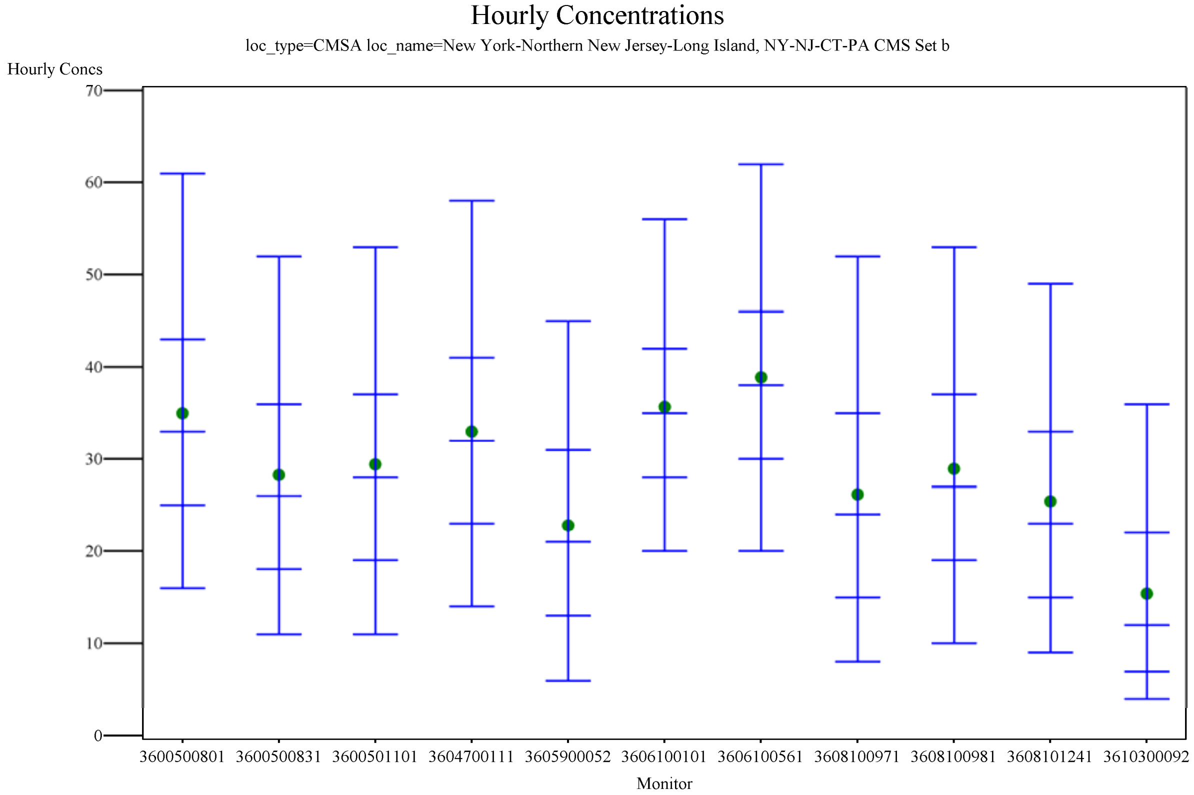


Figure C-24. Spatial distribution of hourly NO_2 concentration, New York CMSA set b, years 1995-2006.

Table C-15. Spatial distribution of annual average NO_2 concentration, New York CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
0900101131	3	23	1	4	23	23	23	23	24	24	24	24	24	24	24
0900190031	8	17	2	11	14	14	15	16	18	18	18	18	19	21	21
0900900271	2	21	1	5	20	20	20	20	20	21	22	22	22	22	22
0900911231	9	26	1	4	24	24	25	25	25	25	26	26	27	27	27
3400300011	3	28	1	2	28	28	28	28	28	28	28	29	29	29	29
3400300051	4	21	1	5	20	20	20	20	20	20	20	20	22	22	22
3401300111	5	32	1	3	31	31	31	31	32	32	32	33	33	33	33
3401300161	1	29			29	29	29	29	29	29	29	29	29	29	29
3401310031	11	28	2	7	24	26	27	28	28	29	29	29	29	29	31
3401700061	11	25	2	6	22	23	23	25	26	26	26	26	26	27	27
3402100051	11	16	1	4	15	15	15	16	16	16	16	17	17	17	17
3402300111	11	18	1	6	16	17	18	18	18	18	19	19	19	19	20
3402730011	11	11	1	6	10	11	11	11	11	11	11	12	12	12	12
3403900042	11	38	4	12	30	32	32	39	40	40	41	41	41	42	42
3403900081	3	28	2	6	27	27	27	27	27	27	27	30	30	30	30
3600500801	5	35	1	4	33	33	34	35	35	35	36	36	36	36	36
3600500831	12	28	2	9	24	25	27	27	28	29	30	30	31	31	32
3600501101	6	30	2	6	26	26	29	29	30	30	30	30	30	32	32
3604700111	1	33			33	33	33	33	33	33	33	33	33	33	33
3605900052	11	23	2	10	18	20	21	22	22	24	24	24	25	25	26
3606100101	4	36	1	1	35	35	35	35	35	36	36	36	36	36	36
3606100561	10	39	2	6	34	35	37	38	38	39	40	40	41	42	42
3608100971	3	26	0	1	26	26	26	26	26	26	26	26	26	26	26
3608100981	7	29	1	4	27	27	28	28	28	29	30	30	30	30	30
3608101241	5	25	2	7	23	23	24	25	25	25	26	27	27	28	28
3610300092	6	15	2	14	13	13	13	13	14	16	17	17	17	17	17

Table C-16. Spatial distribution of hourly NO_2 concentration, New York CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
0900101131	25148	23	13	55	0	9	12	15	18	22	25	29	34	40	109
0900190031	67123	17	13	75	0	4	6	8	10	14	18	23	29	36	103
0900900271	16002	21	14	65	0	6	8	11	14	18	22	27	33	40	101
0900911231	76418	26	13	50	0	11	14	17	20	24	27	31	36	43	240
3400300011	25620	28	14	50	3	11	15	19	23	26	31	35	40	47	119
3400300051	34090	21	14	66	3	5	8	11	14	18	22	27	33	40	124
3401300111	41642	32	16	50	3	12	17	21	26	31	35	40	45	53	148
3401300161	8368	29	15	52	3	11	15	18	22	26	31	36	41	49	103
3401310031	93578	28	14	51	3	11	15	19	23	27	31	35	40	47	150
3401700061	93886	25	14	56	2	9	12	16	19	23	27	32	37	44	147
3402100051	94591	16	11	67	2	4	7	8	11	13	16	20	25	32	79
3402300111	94366	18	12	65	3	5	8	10	13	16	19	23	28	35	99
3402730011	92642	11	9	82	0	3	3	5	7	8	10	13	17	24	95
3403900042	92472	38	15	41	3	19	25	29	33	37	41	45	50	58	225
3403900081	23611	28	13	47	3	11	16	20	24	27	30	34	38	44	122
3600500801	41120	35	14	40	0	19	23	26	30	33	37	40	45	54	181
3600500831	95448	28	13	47	0	13	17	20	23	26	30	34	39	46	136
3600501101	46299	29	13	45	0	14	18	21	24	28	31	35	40	47	119
3604700111	8300	33	14	41	3	17	21	25	28	32	35	39	43	51	155
3605900052	89801	23	13	56	0	8	11	14	18	21	25	29	34	40	162
3606100101	30694	36	11	31	0	23	27	29	32	35	37	40	44	50	118
3606100561	81341	39	13	33	0	24	28	32	35	38	41	44	48	55	162
3608100971	24104	26	14	54	0	10	13	17	20	24	28	33	38	45	95
3608100981	56186	29	13	46	0	13	17	20	24	27	31	35	40	47	114
3608101241	39406	25	13	50	0	11	14	17	20	23	27	31	36	43	144
3610300092	48236	15	10	67	0	5	7	8	10	12	15	19	24	31	86

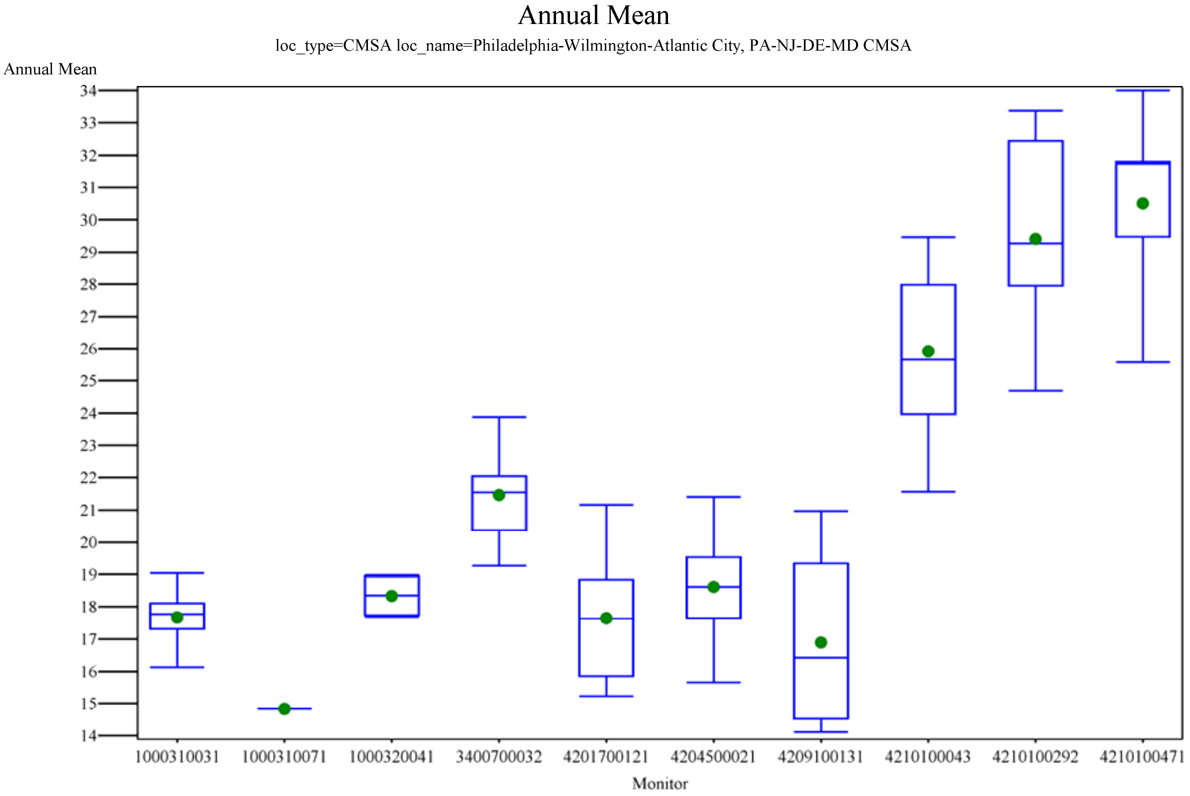


Figure C-25. Spatial distribution of annual average NO_2 concentration, Philadelphia CMSA, years 1995-2006.

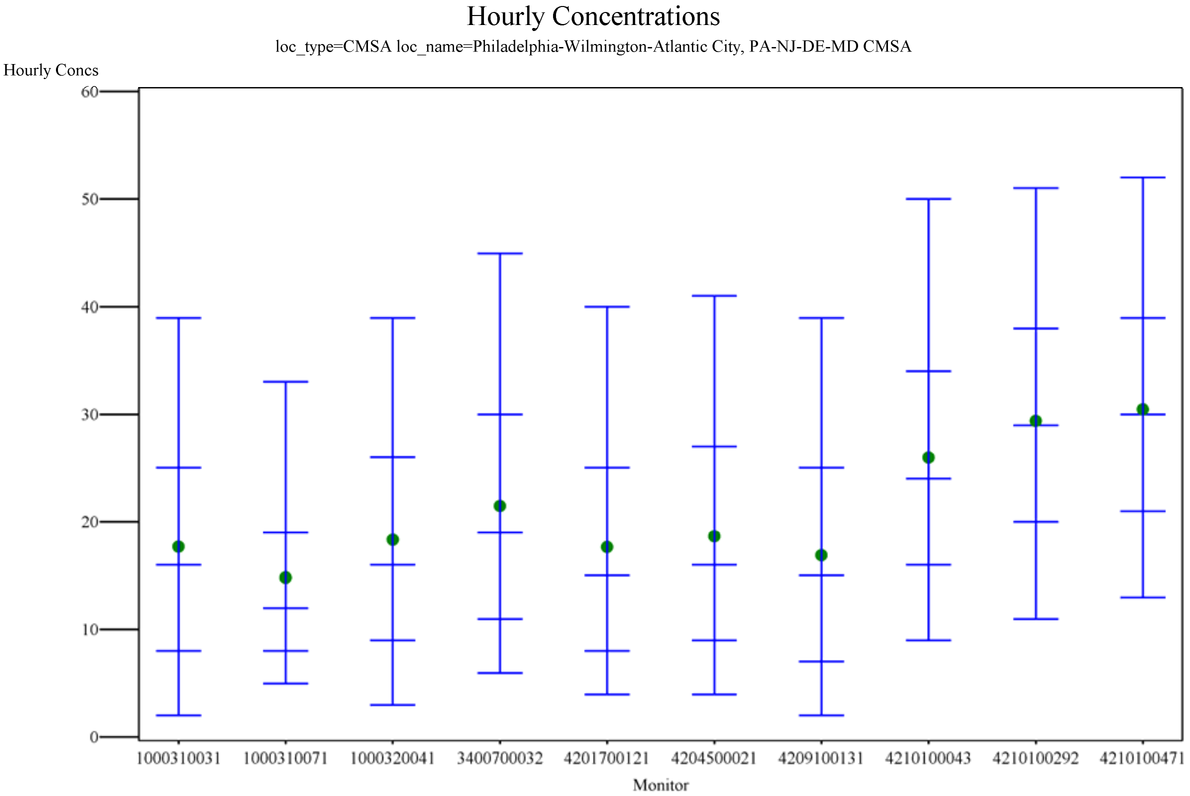


Figure C-26. Spatial distribution of hourly NO_2 concentration, Philadelphia CMSA, years 1995-2006.

Table C-17. Spatial distribution of annual average NO_2 concentration, **Philadelphia** CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1000310031	5	18	1	6	16	16	17	17	18	18	18	18	19	19	19
1000310071	1	15			15	15	15	15	15	15	15	15	15	15	15
1000320041	4	18	1	4	18	18	18	18	18	18	19	19	19	19	19
3400700032	10	21	1	7	19	20	20	20	21	22	22	22	23	24	24
4201700121	12	18	2	11	15	16	16	16	17	18	18	18	20	20	21
4204500021	12	19	2	8	16	17	17	18	18	19	19	19	20	20	21
4209100131	11	17	2	13	14	14	15	16	16	16	17	18	19	19	21
4210100043	11	26	3	10	22	23	24	24	26	26	27	28	28	29	29
4210100292	10	29	3	11	25	25	26	28	28	29	31	32	33	33	33
4210100471	9	31	3	10	26	26	26	29	30	32	32	32	34	34	34

Table C-18. Spatial distribution of hourly NO_2 concentration, Philadelphia CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1000310031	40363	18	12	69	0	4	7	10	12	16	19	23	28	34	247
1000310071	6611	15	9	62	1	6	7	9	10	12	15	17	21	28	69
1000320041	31615	18	12	63	0	5	8	11	13	16	20	23	28	34	115
3400700032	84603	22	13	59	3	7	10	13	16	19	23	27	32	39	114
4201700121	102584	18	12	67	0	5	7	9	12	15	19	23	28	34	106
4204500021	100344	19	12	64	0	5	8	10	13	16	20	24	29	36	268
4209100131	93572	17	12	69	0	4	6	9	11	15	18	22	27	33	99
4210100043	90975	26	13	49	0	10	14	18	20	24	28	31	37	43	190
4210100292	81218	29	13	43	0	15	19	21	25	29	30	35	40	46	120
4210100471	75202	31	12	40	0	16	20	23	26	30	31	36	40	47	140

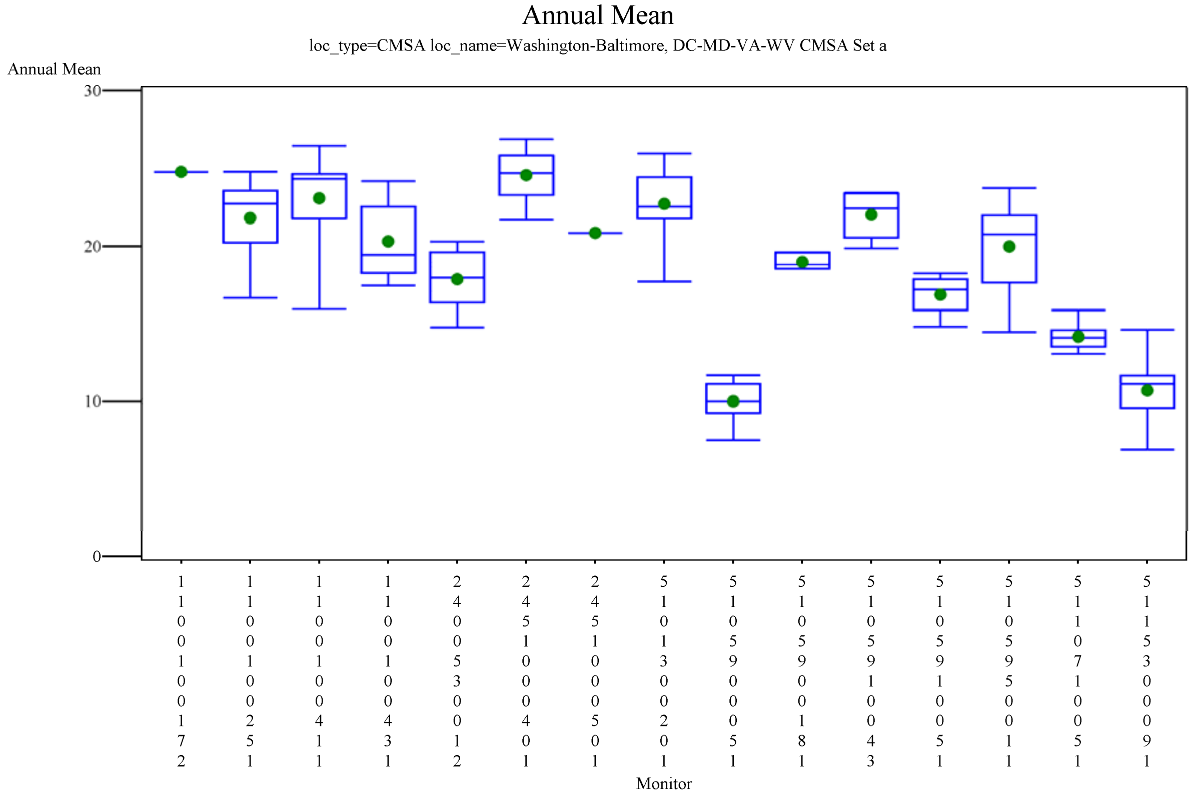


Figure C-27. Spatial distribution of annual average NO_2 concentration, Washington DC CMSA set a, years 1995-2006.

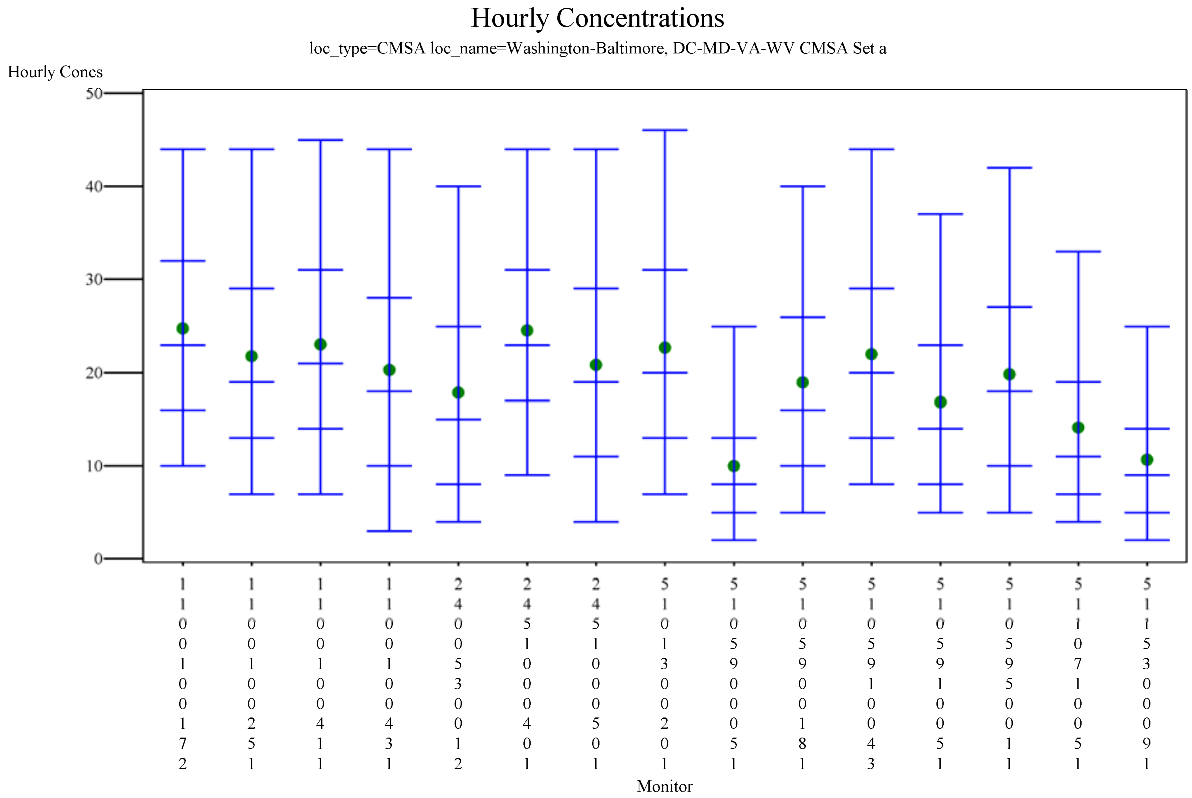


Figure C-28. Spatial distribution of hourly NO_2 concentration, Washington DC CMSA set a, years 1995-2006.

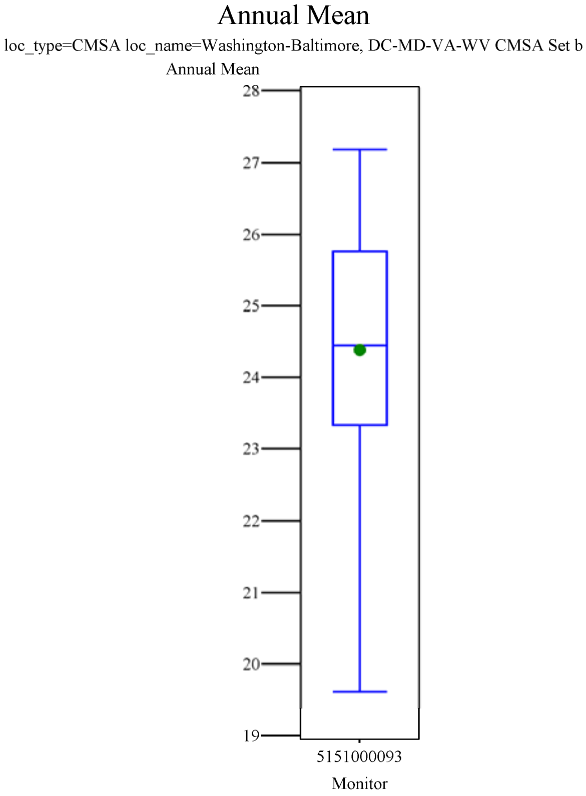


Figure C-29. Spatial distribution of annual average NO_2 concentration, Washington DC CMSA set b, years 1995-2006.

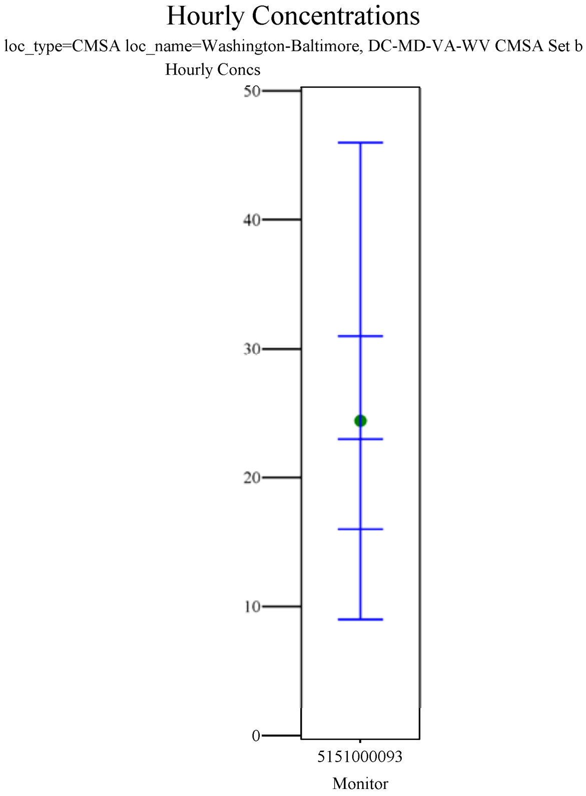


Figure C-30. Spatial distribution of hourly NO_2 concentration, Washington DC CMSA set b, years 1995-2006.

Table C-19. Spatial distribution of annual average NO$_2$ concentration, Washington DC CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1100100172	1	25			25	25	25	25	25	25	25	25	25	25	25
1100100251	12	22	2	11	17	19	20	21	22	23	23	23	24	24	25
1100100411	12	23	3	12	16	21	21	23	23	24	24	25	25	25	26
1100100431	12	20	2	12	17	18	18	18	19	19	21	22	23	23	24
2400530012	8	18	2	11	15	15	15	17	18	18	18	19	20	20	20
2451000401	11	25	2	7	22	23	23	23	24	25	26	26	26	26	27
2451000501	1	21			21	21	21	21	21	21	21	21	21	21	21
5101300201	12	23	2	10	18	21	21	22	22	23	23	24	25	25	26
5105900051	11	10	1	12	7	9	9	10	10	10	10	11	11	11	12
5105900181	3	19	1	3	19	19	19	19	19	19	19	20	20	20	20
5105910043	6	22	2	7	20	20	21	21	22	22	23	23	23	23	23
5105910051	4	17	1	9	15	15	15	17	17	17	17	17	18	18	18
5105950011	10	20	3	15	14	16	17	19	20	21	22	22	22	23	24
5110710051	8	14	1	6	13	13	13	14	14	14	14	14	15	16	16
5115300091	12	11	2	18	7	9	9	10	10	11	11	11	12	12	15
5151000093	12	24	2	8	20	23	23	23	24	24	25	26	26	26	27

Table C-20. Spatial distribution of hourly NO$_2$ concentration, Washington DC CMSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1100100172	8584	25	11	45	4	12	15	18	20	23	27	30	33	39	113
1100100251	102444	22	12	55	0	9	11	14	16	19	23	27	32	39	285
1100100411	103173	23	12	53	0	9	12	15	18	21	24	28	33	39	141
1100100431	102217	20	13	64	0	6	9	12	15	18	22	26	31	38	258
2400530012	63983	18	12	65	0	5	7	10	12	15	19	23	28	34	114
2451000401	89589	25	11	44	0	12	15	18	21	23	26	29	33	39	108
2451000501	7872	21	12	60	0	6	9	12	16	19	23	27	32	38	75
5101300201	97517	23	13	56	0	8	11	14	17	20	24	28	34	41	110
5105900051	89964	10	7	73	0	3	4	5	6	8	10	12	15	20	101
5105900181	22689	19	11	60	0	6	9	11	13	16	20	24	29	36	89
5105910043	50294	22	11	52	0	10	12	14	17	20	23	27	31	38	91
5105910051	34022	17	11	63	0	6	8	9	12	14	17	21	26	32	129
5105950011	79051	20	12	61	0	6	9	12	14	18	21	25	30	36	155
5110710051	65327	14	9	65	0	5	7	8	10	11	14	17	21	28	64
5115300091	101671	11	7	68	0	3	5	6	7	9	11	13	16	21	84
5151000093	98221	24	12	48	0	11	14	17	20	23	26	29	34	40	115

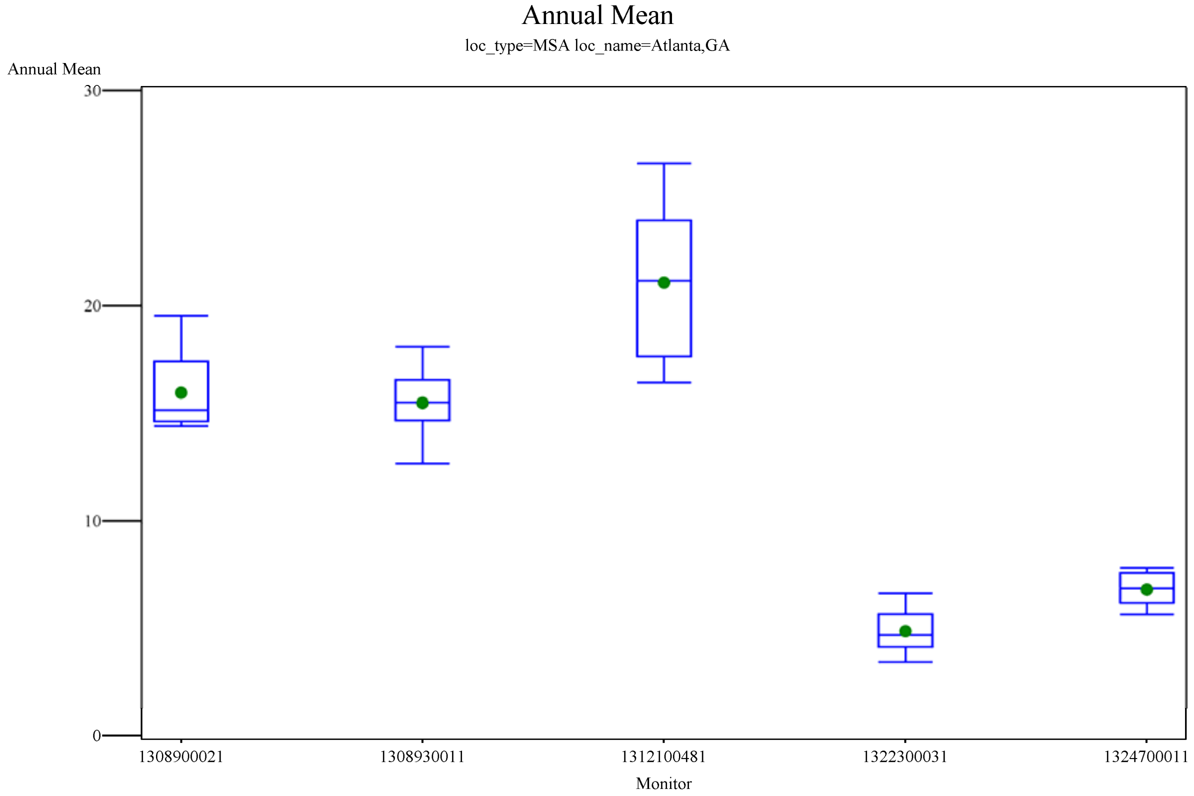


Figure C-31. Spatial distribution of annual average NO_2 concentration, Atlanta MSA, years 1995-2006.

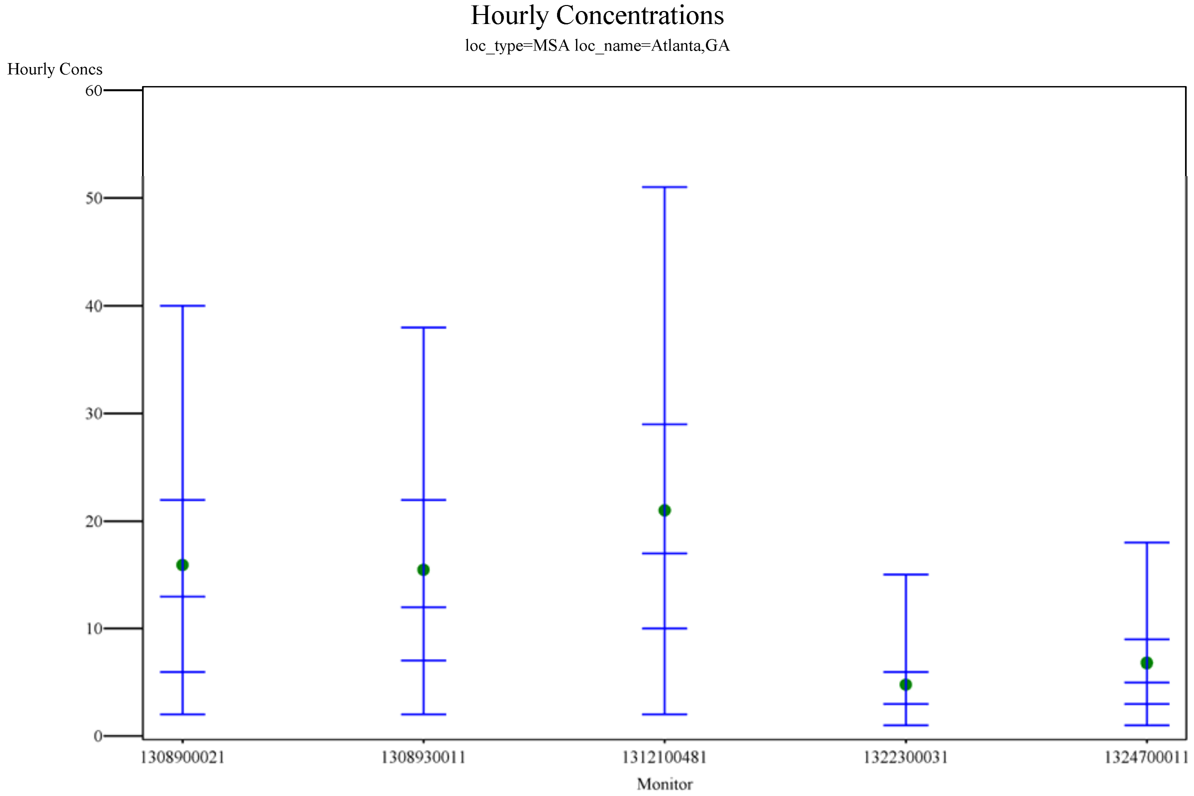


Figure C-32. Spatial distribution of hourly NO_2 concentration, Atlanta MSA, years 1995-2006.

Table C-21. Spatial distribution of annual average NO_2 concentration, Atlanta MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1308900021	10	16	2	11	14	14	15	15	15	15	16	17	18	19	20
1308930011	9	15	2	10	13	13	14	15	15	16	16	17	17	18	18
1312100481	12	21	4	17	16	17	17	18	19	21	23	24	24	25	27
1322300031	10	5	1	20	3	4	4	4	4	5	5	5	6	6	7
1324700011	12	7	1	11	6	6	6	6	6	7	7	8	8	8	8

Table C-22. Spatial distribution of hourly NO_2 concentration, Atlanta MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1308900021	83891	16	12	77	0	3	5	8	10	13	16	20	25	33	139
1308930011	72029	15	11	73	1	4	6	8	10	12	15	19	24	32	95
1312100481	98975	21	15	73	0	5	8	11	14	17	21	26	33	43	181
1322300031	80168	5	5	108	0	1	1	2	3	3	4	5	7	11	70
1324700011	100149	7	6	81	0	2	3	3	4	5	6	8	10	14	242

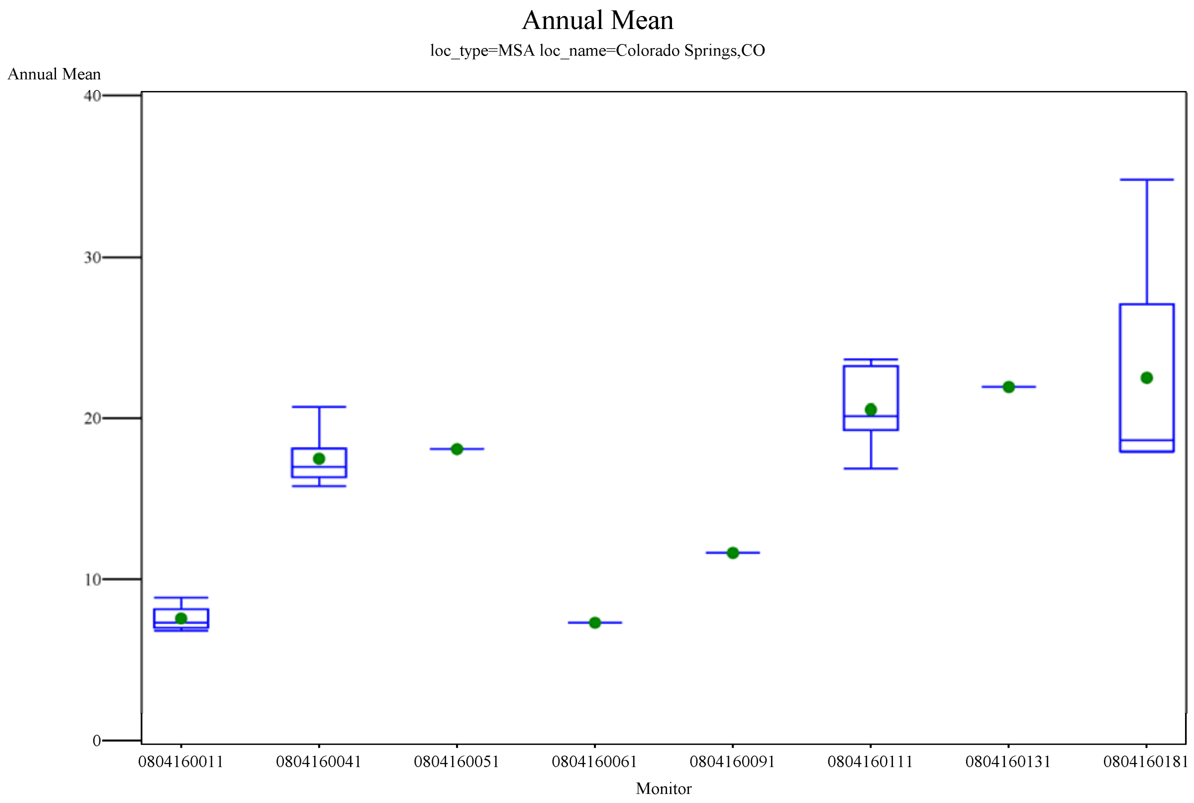


Figure C-33. Spatial distribution of annual average NO_2 concentration, Colorado Springs MSA, years 1995-2006.

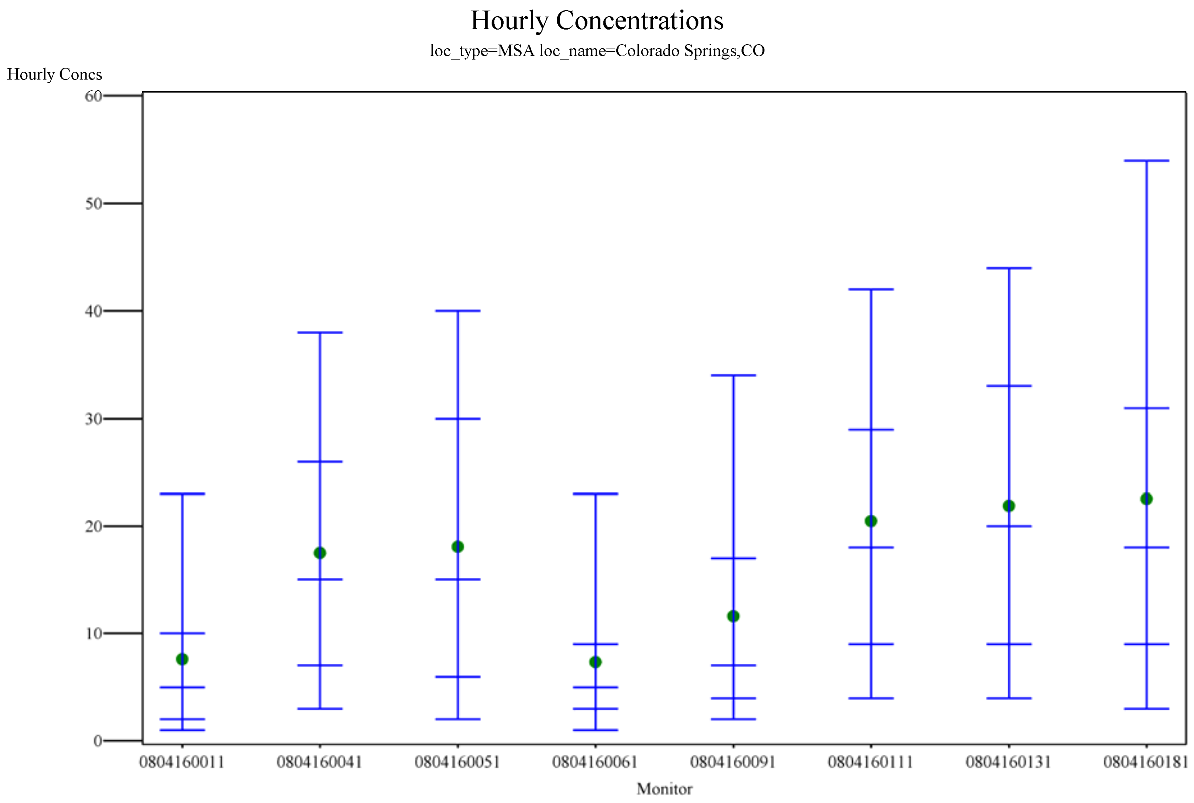


Figure C-34. Spatial distribution of hourly NO_2 concentration, Colorado Springs MSA, years 1995-2006.

Table C-23. Spatial distribution of annual average NO_2 concentration, Colorado Springs MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
0804160011	6	8	1	10	7	7	7	7	7	7	7	8	8	9	9
0804160041	6	17	2	10	16	16	16	16	17	17	17	18	18	21	21
0804160051	1	18			18	18	18	18	18	18	18	18	18	18	18
0804160061	1	7			7	7	7	7	7	7	7	7	7	7	7
0804160091	1	12			12	12	12	12	12	12	12	12	12	12	12
0804160111	6	21	3	12	17	17	19	19	20	20	20	23	23	24	24
0804160131	1	22			22	22	22	22	22	22	22	22	22	22	22
0804160181	4	22	8	37	18	18	18	18	18	19	19	19	35	35	35

Table C-24. Spatial distribution of hourly NO_2 concentration, Colorado Springs MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
0804160011	51373	8	7	94	0	1	2	3	4	5	7	9	12	18	59
0804160041	51288	17	11	66	0	4	6	9	12	15	20	24	28	34	115
0804160051	8345	18	13	74	1	3	5	7	10	15	21	27	32	36	143
0804160061	7993	7	7	99	0	1	2	3	4	5	6	8	11	16	49
0804160091	8282	12	10	89	0	2	3	4	6	7	10	14	20	29	56
0804160111	50707	21	16	77	0	5	7	10	14	18	23	27	31	37	246
0804160131	8637	22	14	62	0	5	8	11	15	20	26	31	36	41	87
0804160181	33737	23	21	94	0	5	7	10	14	18	23	28	33	41	308

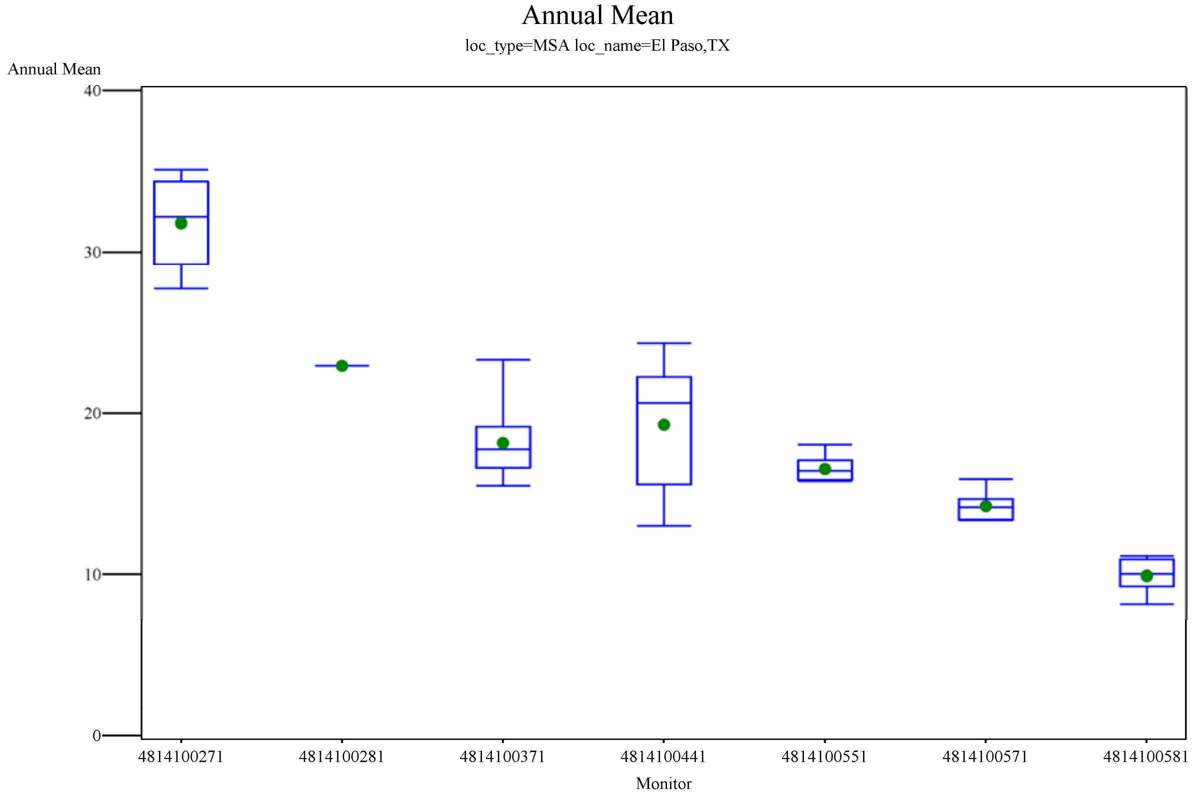


Figure C-35. Spatial distribution of annual average NO_2 concentration, El Paso MSA, years 1995-2006.

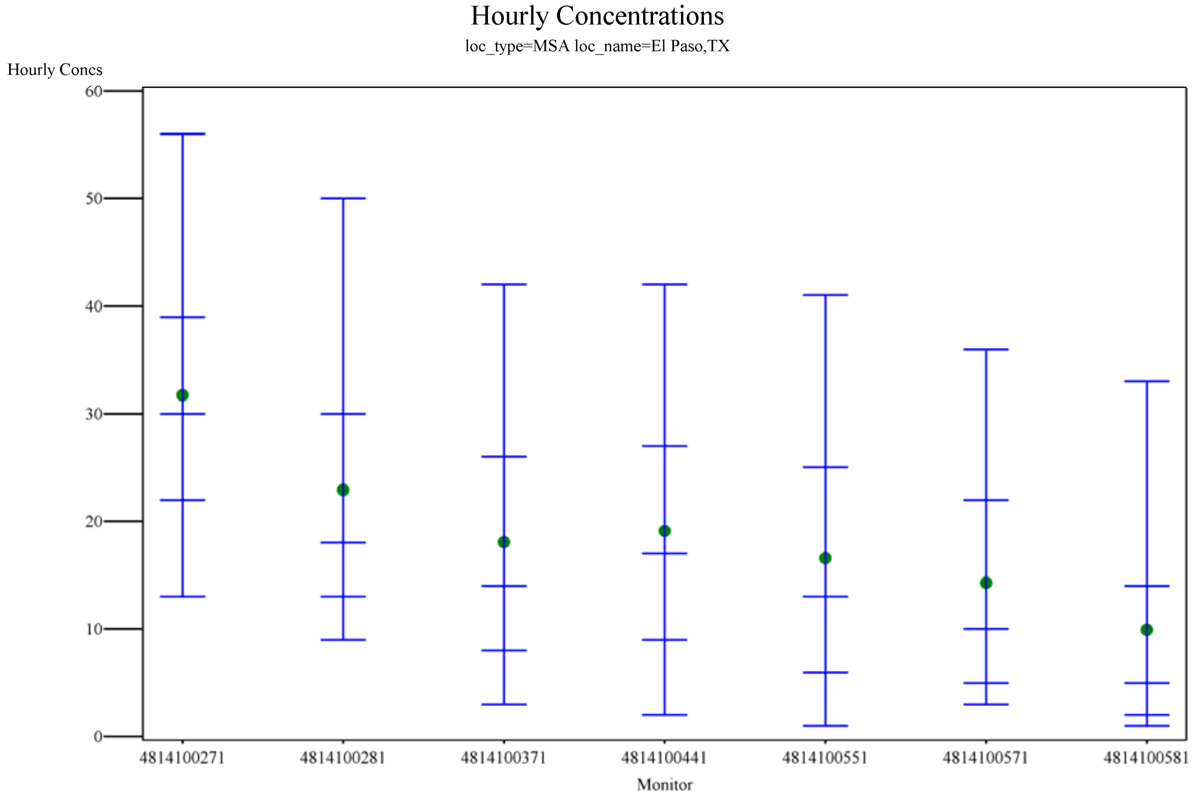


Figure C-36. Spatial distribution of hourly NO_2 concentration, El Paso MSA, years 1995-2006.

Table C-25. Spatial distribution of annual average NO_2 concentration, El Paso MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
4814100271	4	32	3	10	28	28	28	31	31	32	34	34	35	35	35
4814100281	1	23			23	23	23	23	23	23	23	23	23	23	23
4814100371	11	18	2	12	15	16	17	17	17	18	18	18	19	21	23
4814100441	8	19	4	22	13	13	13	18	20	21	21	22	23	24	24
4814100551	7	17	1	5	16	16	16	16	16	16	16	16	17	18	18
4814100571	7	14	1	6	13	13	13	14	14	14	15	15	15	16	16
4814100581	6	10	1	11	8	8	9	9	10	10	10	11	11	11	11

Table C-26. Spatial distribution of hourly NO_2 concentration, El Paso MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
4814100271	29730	32	14	45	1	16	20	24	27	30	33	37	42	49	219
4814100281	8045	23	14	60	5	10	12	13	15	18	22	27	34	42	117
4814100371	87748	18	13	71	0	5	7	9	12	14	18	23	29	36	153
4814100441	62362	19	13	67	0	5	8	11	14	17	21	25	30	36	125
4814100551	53960	17	13	78	0	3	5	7	10	13	18	23	28	35	87
4814100571	57229	14	11	79	0	3	4	6	8	10	14	19	25	31	85
4814100581	47248	10	11	109	0	1	2	3	4	5	7	11	18	27	84

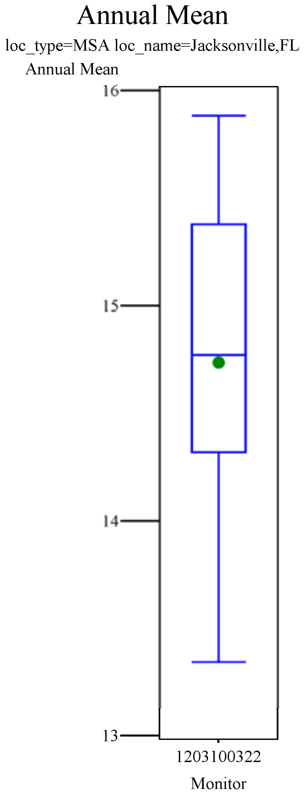


Figure C-37. Spatial distribution of annual average NO_2 concentration, Jacksonville MSA, years 1995-2006.

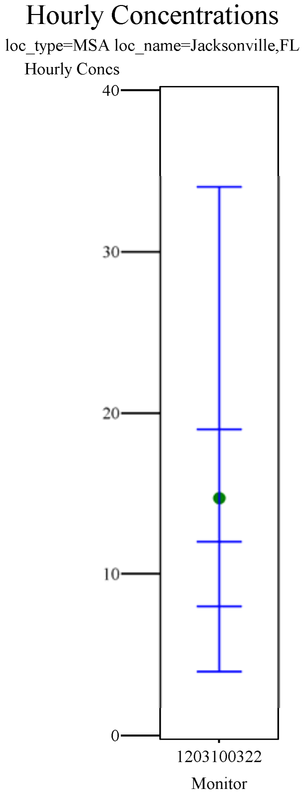


Figure C-38. Spatial distribution of hourly NO_2 concentration, Jacksonville MSA, years 1995-2006.

Table C-27. Spatial distribution of annual average NO_2 concentration, Jacksonville MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1203100322	10	15	1	6	13	14	14	14	15	15	15	15	16	16	16

Table C-28. Spatial distribution of hourly NO_2 concentration, Jacksonville MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1203100322	78222	15	10	67	0	5	7	9	10	12	15	18	22	28	294

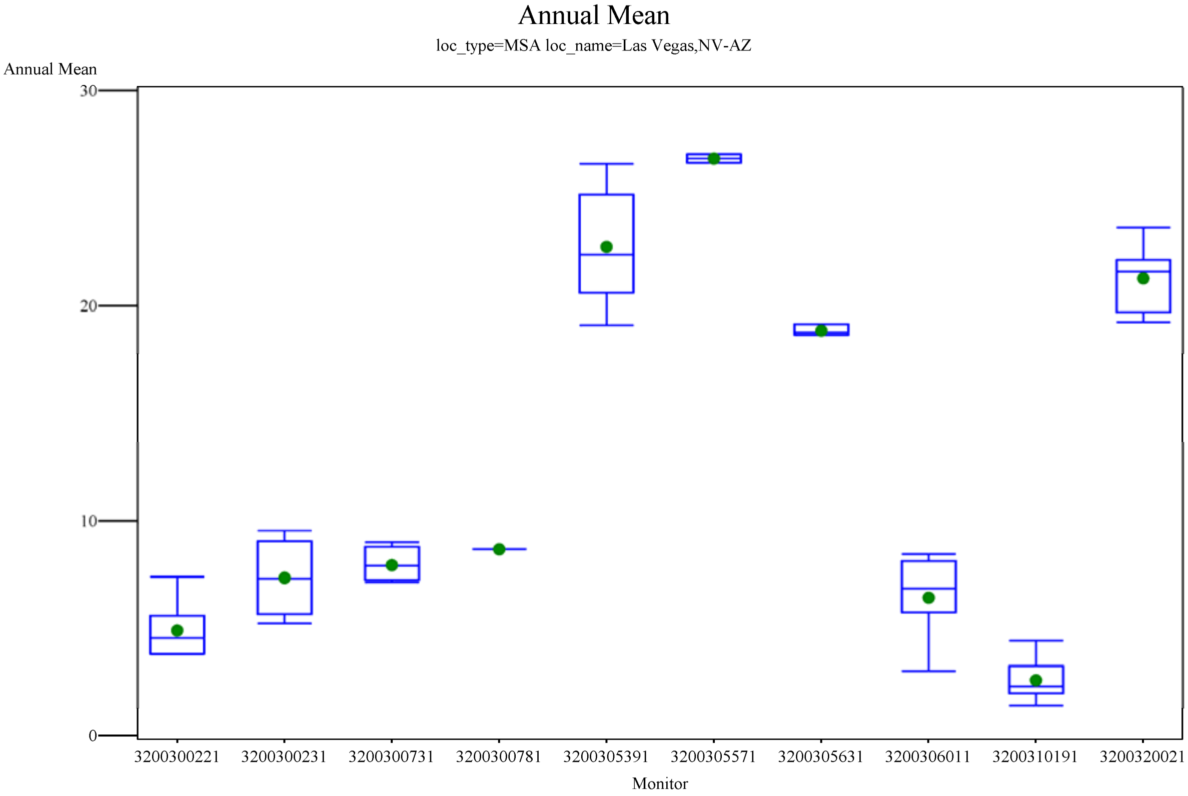


Figure C-39. Spatial distribution of annual average NO_2 concentration, Las Vegas MSA, years 1995-2006.

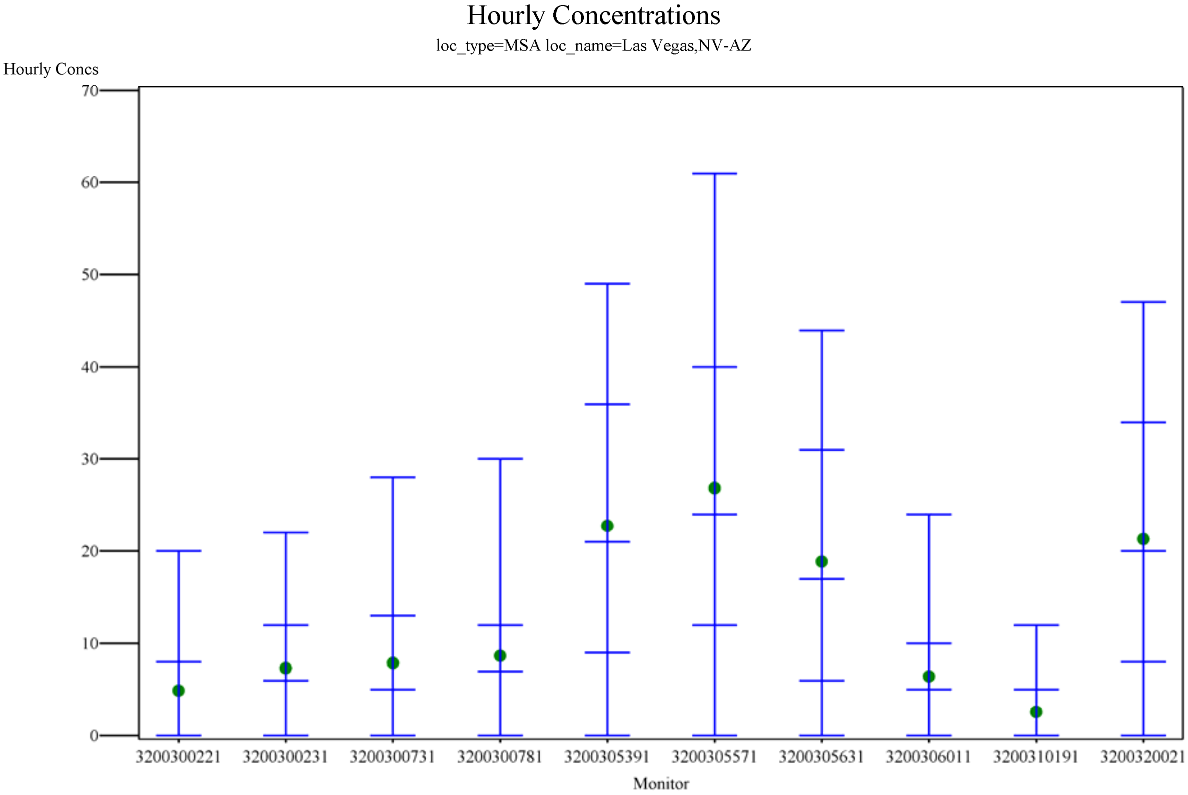


Figure C-40. Spatial distribution of hourly NO_2 concentration, Las Vegas MSA, years 1995-2006.

Table C-29. Spatial distribution of annual average NO_2 concentration, Las Vegas MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
3200300221	7	5	1	26	4	4	4	4	4	5	5	5	6	7	7
3200300231	4	7	2	28	5	5	5	6	6	7	9	9	10	10	10
3200300731	7	8	1	9	7	7	7	8	8	8	8	8	9	9	9
3200300781	1	9			9	9	9	9	9	9	9	9	9	9	9
3200305391	8	23	3	12	19	19	20	21	22	22	23	25	25	27	27
3200305571	2	27	0	1	27	27	27	27	27	27	27	27	27	27	27
3200305631	3	19	0	1	19	19	19	19	19	19	19	19	19	19	19
3200306011	5	6	2	34	3	3	4	6	6	7	7	8	8	8	8
3200310191	7	3	1	38	1	1	2	2	2	2	3	3	3	4	4
3200320021	7	21	2	7	19	19	20	21	21	22	22	22	22	24	24

Table C-30. Spatial distribution of hourly NO_2 concentration, Las Vegas MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
3200300221	58087	5	7	152	0	0	0	0	0	0	5	7	10	15	91
3200300231	34550	7	8	105	0	0	0	0	5	6	8	10	13	18	52
3200300731	56906	8	10	124	0	0	0	0	0	5	8	11	15	22	104
3200300781	8672	9	10	115	0	0	0	0	5	7	8	10	14	22	87
3200305391	64921	23	16	70	0	5	7	10	14	21	28	33	38	44	103
3200305571	16674	27	21	78	0	0	10	14	19	24	31	37	43	52	410
3200305631	25061	19	15	78	0	0	5	7	11	17	23	28	33	39	87
3200306011	42417	6	8	124	0	0	0	0	0	5	7	8	12	18	51
3200310191	57230	3	5	186	0	0	0	0	0	0	0	0	6	9	71
3200320021	56244	21	16	73	0	0	6	9	13	20	27	32	36	42	110

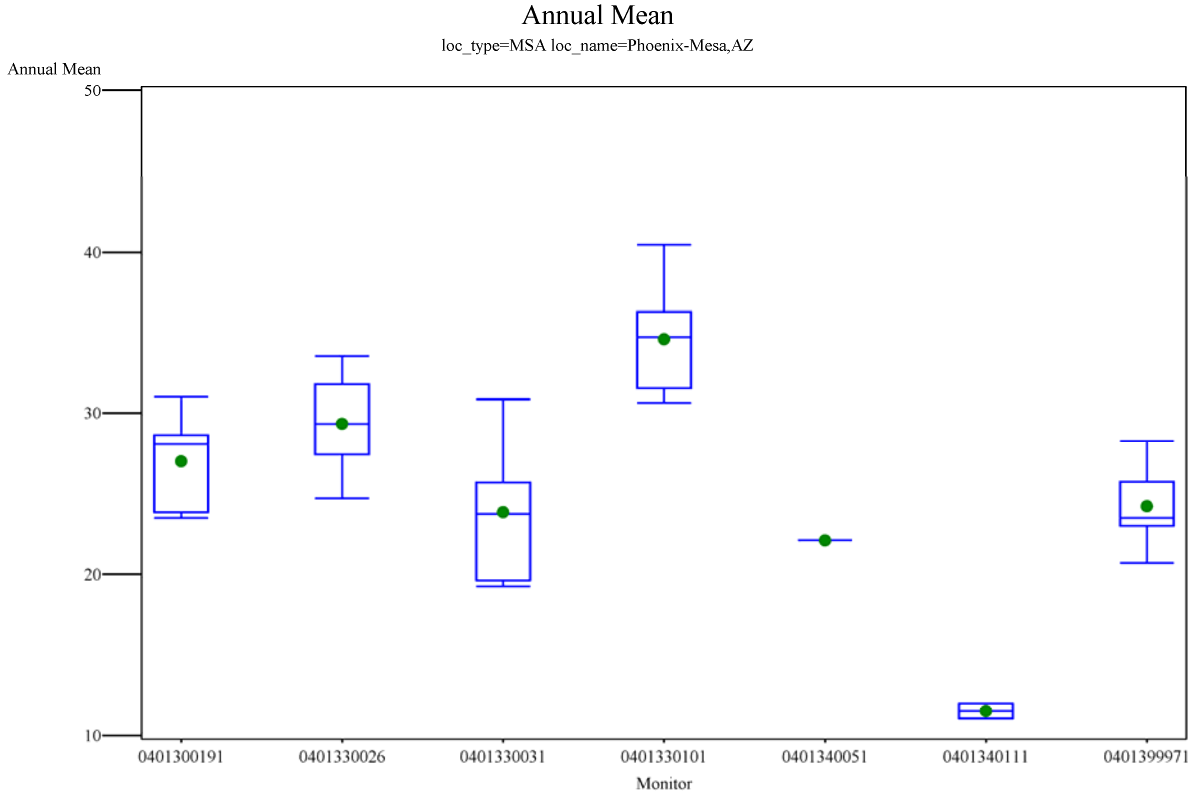


Figure C-41. Spatial distribution of annual average NO_2 concentration, Phoenix MSA, years 1995-2006.

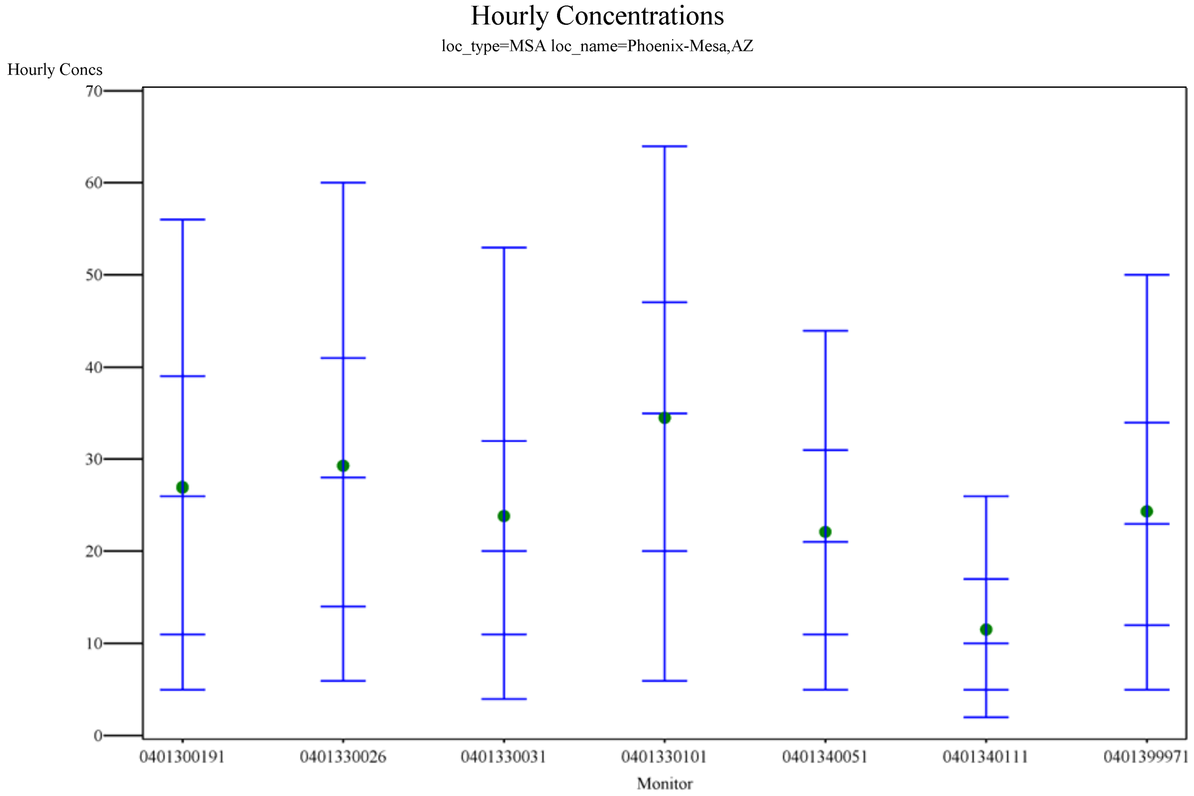


Figure C-42. Spatial distribution of hourly NO_2 concentration, Phoenix MSA, years 1995-2006.

Table C-31. Spatial distribution of annual average NO_2 concentration, Phoenix MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
0401300191	10	27	3	10	24	24	24	25	27	28	28	29	29	30	31
0401330026	12	29	3	10	25	25	26	29	29	29	30	32	32	33	34
0401330031	10	24	4	17	19	19	20	21	23	24	24	25	28	30	31
0401330101	9	35	3	9	31	31	31	32	34	35	35	36	37	40	40
0401340051	1	22			22	22	22	22	22	22	22	22	22	22	22
0401340111	2	12	1	6	11	11	11	11	11	12	12	12	12	12	12
0401399971	5	24	3	12	21	21	22	23	23	24	25	26	27	28	28

Table C-32. Spatial distribution of hourly NO_2 concentration, Phoenix MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
0401300191	81411	27	17	63	0	6	9	14	20	26	32	37	42	50	148
0401330026	97376	29	17	59	0	8	12	17	23	28	33	38	44	53	151
0401330031	80162	24	19	78	0	6	9	12	16	20	25	30	35	45	267
0401330101	73070	35	18	53	0	9	16	23	30	35	40	45	50	58	164
0401340051	7420	22	13	58	2	7	9	13	17	21	25	29	33	39	99
0401340111	16459	12	8	69	0	2	4	6	8	10	13	16	18	22	53
0401399971	41521	24	15	60	0	7	10	14	19	23	27	32	37	45	131

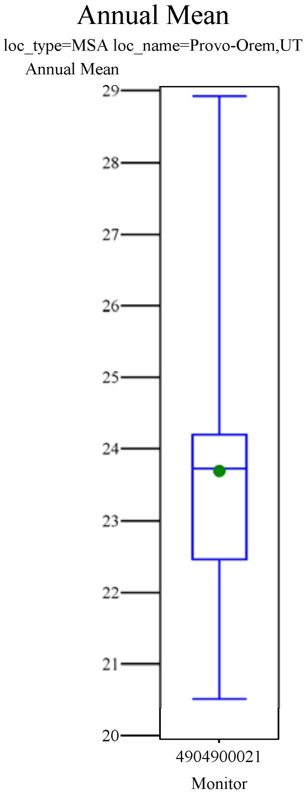


Figure C-43. Spatial distribution of annual average NO_2 concentration, Provo MSA, years 1995-2006.

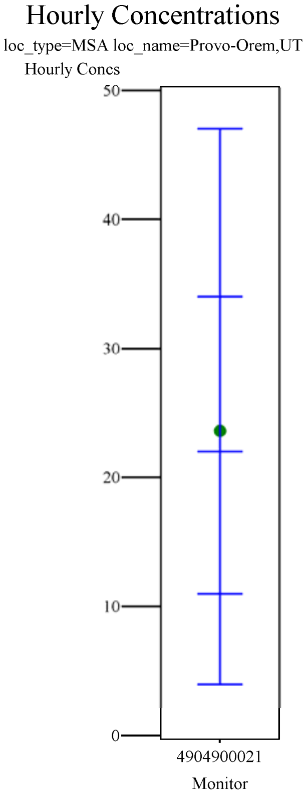


Figure C-44. Spatial distribution of hourly NO_2 concentration, Provo MSA, years 1995-2006.

Table C-33. Spatial distribution of annual average NO_2 concentration, Provo MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
4904900021	12	24	2	9	21	22	22	23	23	24	24	24	24	25	29

Table C-34. Spatial distribution of hourly NO_2 concentration, Provo MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
4904900021	96873	24	16	68	0	6	9	13	17	22	27	31	36	42	164

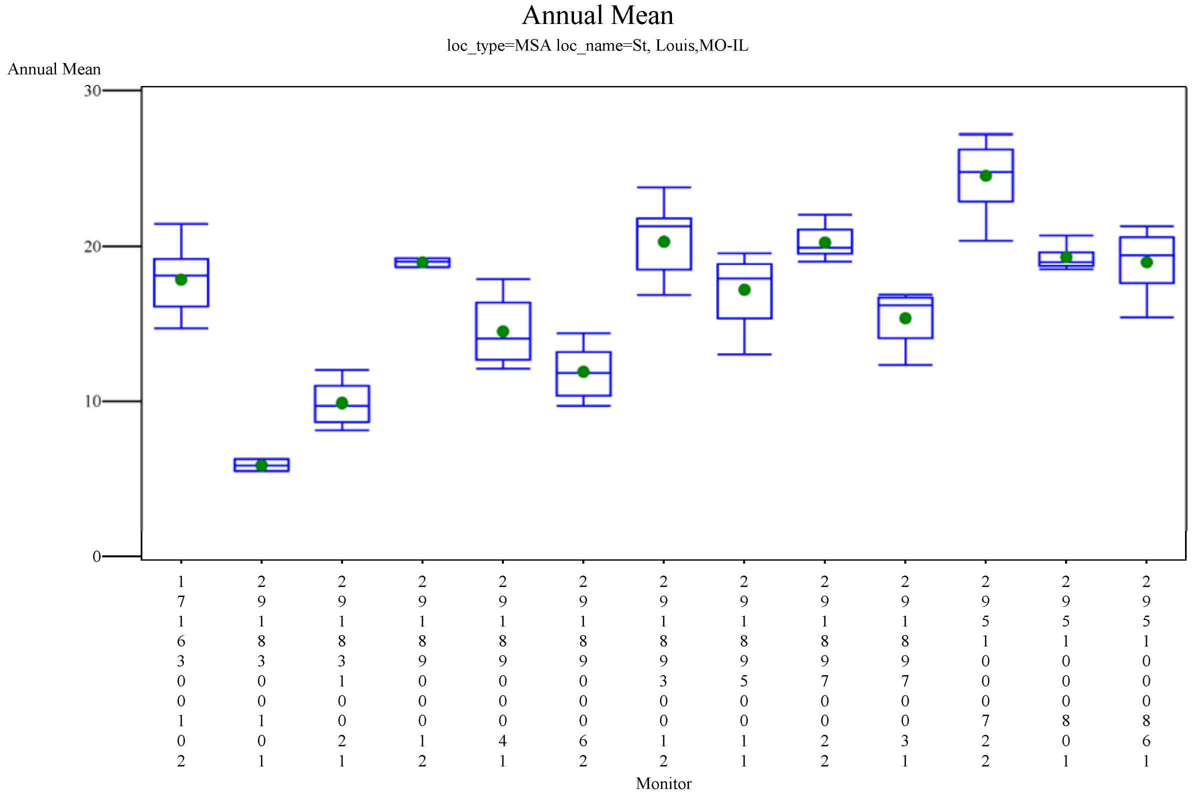


Figure C-45. Spatial distribution of annual average NO_2 concentration, St. Louis MSA, years 1995-2006.

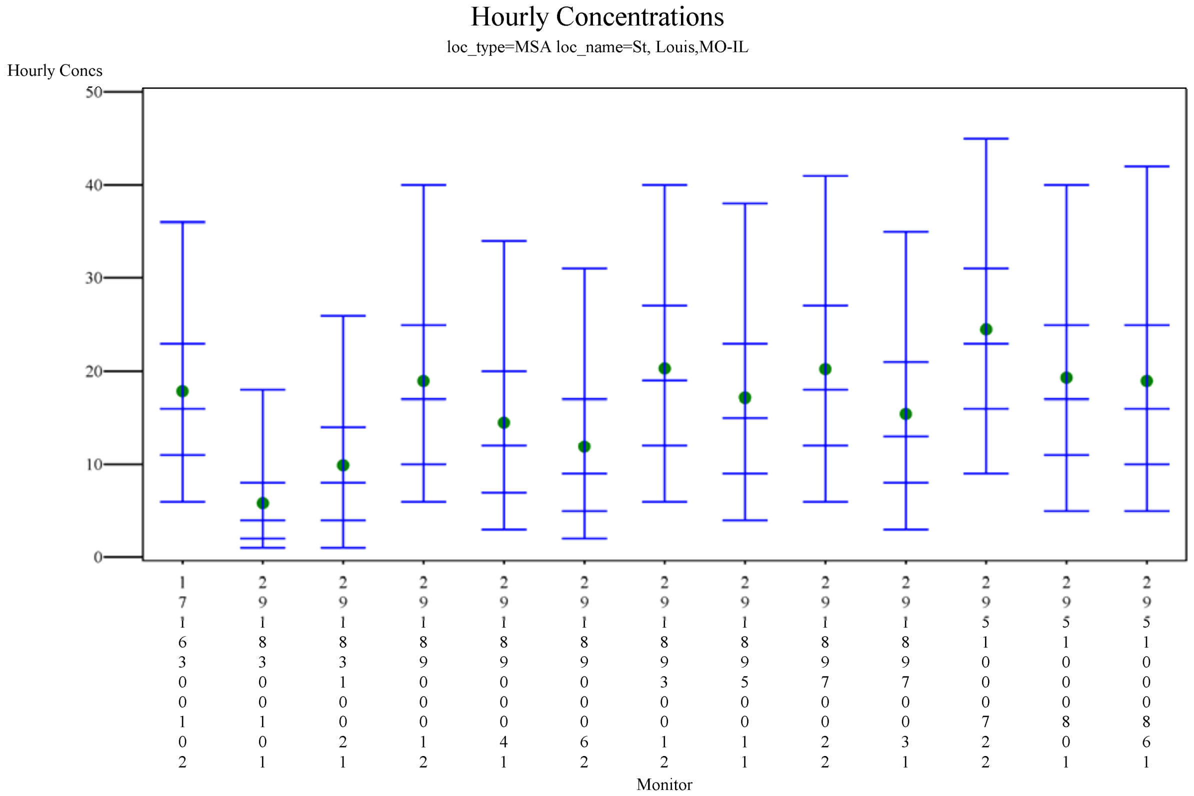


Figure C-46. Spatial distribution of hourly NO_2 concentration, St. Louis MSA, years 1995-2006.

Table C-35. Spatial distribution of annual average NO_2 concentration, St. Louis MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1716300102	12	18	2	12	15	15	16	16	17	18	18	19	19	20	21
2918300101	3	6	0	7	5	5	5	5	6	6	6	6	6	6	6
2918310021	12	10	1	13	8	8	9	9	9	10	11	11	11	11	12
2918900012	3	19	0	2	19	19	19	19	19	19	19	19	19	19	19
2918900041	6	15	2	15	12	12	13	13	14	14	14	16	16	18	18
2918900062	11	12	1	12	10	10	10	11	12	12	12	13	13	13	14
2918930012	11	20	2	11	17	17	18	19	20	21	22	22	22	22	24
2918950011	10	17	2	13	13	14	15	16	17	18	19	19	19	19	20
2918970022	6	20	1	6	19	19	20	20	20	20	20	21	21	22	22
2918970031	4	15	2	14	12	12	12	16	16	16	16	16	17	17	17
2951000722	10	25	2	9	20	21	23	24	25	25	25	26	26	27	27
2951000801	5	19	1	5	19	19	19	19	19	19	19	20	20	21	21
2951000861	6	19	2	11	15	15	18	18	19	19	20	21	21	21	21

Table C-36. Spatial distribution of hourly NO_2 concentration, St. Louis MSA, years 1995-2006.

Monitor ID	n	Mean	SD	COV	Min	p10	p20	p30	p40	p50	p60	p70	p80	p90	Max
1716300102	101236	18	9	52	0	8	10	12	14	16	19	21	25	31	123
2918300101	25873	6	6	98	0	1	2	2	3	4	5	7	9	13	51
2918310021	99623	10	8	81	0	2	3	4	6	8	10	12	16	21	73
2918900012	25801	19	11	58	0	7	9	12	14	17	20	23	28	34	89
2918900041	51987	15	10	68	0	4	6	8	10	12	15	18	22	29	80
2918900062	93770	12	9	79	0	3	4	5	7	9	12	15	19	25	79
2918930012	95589	20	11	52	0	8	11	13	16	19	22	25	29	35	101
2918950011	86912	17	11	62	0	6	8	10	12	15	18	21	26	32	124
2918970022	51777	20	11	54	0	8	11	13	16	18	21	25	29	36	103
2918970031	32235	15	10	66	0	4	7	9	11	13	16	19	24	30	64
2951000722	85643	25	11	46	0	11	15	18	20	23	26	29	33	40	130
2951000801	42884	19	11	59	0	7	10	12	15	17	20	23	28	34	274
2951000861	51623	19	12	62	0	6	9	11	14	16	19	23	28	36	87

Appendix D. Technical Memorandum on Regression Modeling

This appendix provides a technical memorandum submitted to EPA by ICF International. The memo is as submitted, with the exception of modified page numbering and addition of borders around each table.

MEMORANDUM

To: Stephen Graham, US EPA

From: Jonathan Cohen and Arlene Rosenbaum

Date: February 15, 2008

Re: Regression Modeling of NO_2 Exceedances of 150 ppb versus Annual Mean

---

SUMMARY

This document describes our regression analyses of 1995 to 2006 NO_2 hourly concentration data. Regression was used to estimate the annual number of exceedances of 150 ppb from the annual mean, in 20 locations (mostly large urban areas). Exposures to concentrations above certain thresholds may be associated with adverse health effects. These models were applied in an as-is scenario to estimate the annual exceedances at sites with annual means equal to the 1995-2006 current average for their location. These models were also applied in a current-standard scenario to predict the annual exceedances at sites with annual means equal to the current NO2 standard of 53 ppb. The current-standard scenario is an extrapolation to higher annual means than currently observed; the maximum annual mean across all complete site-years was 51 ppb, in Los Angeles.

We found these results unsatisfactory, both because the regression models did not show a strong relationship between the annual means and the exceedances, and because the predicted numbers of exceedances for the current-standard scenario were in many cases extremely high and quite uncertain. For this reason we decided not to apply the regression modeling to the other concentration levels of interest (200, 250, and 300 ppb) but instead decided to develop empirical exceedance estimates, as described elsewhere.

DATA

All of the 1995 to 2006 NO_2 hourly concentration data from AQS were compiled and annual summary statistics for each site-year combination were computed. Of particular interest is the long-term air quality measured by the annual mean and the short-term air quality measured by the annual numbers of hourly exceedances of selected levels 150, 200, 250 and 300 ppb. Exposures to concentrations above these thresholds may be associated with adverse health effects. To make the results temporally representative, we restricted the analyses to the 20 percent of site-years that were 75 % complete, as defined by having data for 75 % of the hours in a year and having data for at least 75 % of the hours in a day (i.e., 18 hours or more) on at least 75 % of the days in a year. We also spatially grouped the data into 18 urban areas with high annual means and high exceedances; these locations were all CMSAs or MSAs either with at least one site-year annual mean above 25.72 ppb (the 90th percentile) or with at least one exceedance of 200 ppb, as follows.

- Boston
- Cleveland

- Denver
- Detroit
- Los Angeles
- New York
- Philadelphia
- Washington DC
- Atlanta
- Colorado Springs
- El Paso
- Las Vegas
- Phoenix
- St. Louis
- Chicago
- Miami
- Jacksonville
- Provo.

The remaining site-years were analyzed as two additional location groups: “Other MSA/CMSA” site-years in an MSA or CMSA, and “Other Not MSA” site-years not in an MSA. Thus we have a total of 20 “locations.”

REGRESSION MODELS

The regression modeling of the 1995-2006 NO2 data continues the analyses by McCurdy (1994)[1] of the 1988-1992 data. A regression model is used to estimate the mean number of exceedances from the annual mean. McCurdy (1994) assumed normally distributed exceedances and an exponential link function to estimate exceedances of 150, 200, 250, and 300 ppb based on the 1988-1992 data. In this section we present the results of the regression analyses for exceedances of 150 ppb using eight alternative models based on the 1995-2006 data. Throughout this discussion, “exceedances” will refer to annual numbers of hourly exceedances of 150 ppb, unless otherwise stated.

Of the eight models, the two selected regression models were the Poisson exponential model and the normal linear model, stratified by location. The Poisson exponential model is of the form:

- Number of exceedances has a Poisson distribution.
- Mean exceedances = exp(a + b × annual mean).
- The intercept a, and slope b, depend on the location.

The normal linear model is of the form:

- Number of exceedances has a normal distribution with standard deviation s.
- Mean exceedances = a + b × annual mean.
- The intercept a, slope b, and s all depend on the location.

---

[1] McCurdy TR (1994). Analysis of high 1 hour NO2 values and associated annual averages using 1988-1992 data. Report to the Office of Air Quality Planning and Standards, Durham NC.

The first issue to be resolved was to decide whether to apply the regression analyses to the means and exceedances for each season separately or to each year. We examined the exceedance data for Colorado Springs, which had the highest maximum number of annual exceedances of 200 ppb, 69, which occurred at site 804160181 in 2000. Of these 69 exceedances, 34 occurred in the winter on January 18-20, 2000, and 35 occurred in the summer on June 12-14, 2000. This limited analysis suggests that there is no clear pattern of seasonality in the exceedances. We decided to apply the regression modeling to the annual means and annual exceedances.

Table 1 describes the eight regression models fitted. As described shortly, we fitted two distributions (normal and Poisson), two link functions (identity and exponential), and two stratifications (all data and stratified by location). The McCurdy (1994) analysis used a normal distribution, an exponential link, and stratified by location into Los Angeles and Not Los Angeles.

We fitted generalized linear models where the number of exceedances has a given distribution (we fitted normal and Poisson distributions) and where the mean number of exceedances is a given function g of the annual mean. The function g(x) is called the link function. We can also define the link by defining the inverse link, i.e., the solution for x of the equation g(x) = y.

We fitted two link functions, an identity link g(x) = x and a logarithmic link g(x) = log(x), where "log" denote the natural logarithm. The corresponding inverse links are the identity link, which we also call the "linear" function, and the exponential function. Thus, the linear inverse link models are of the form:

$$\text{Mean exceedances} = a + b \times \text{annual mean}.$$

The exponential inverse link models are of the form:

$$\text{Mean exceedances} = \exp(a + b \times \text{annual mean}).$$

Table 1. Goodness-of-fit statistics for eight generalized linear models.

Distribution	**Inverse Link**	**Strata (a separate model is fitted in each stratum)**	**R squared for all data**	**Min R squared among locations**	**Max R squared among locations**	**Log-Likelihood**	**Number of strata in final model**
Normal	Linear	All	0.033			-11527	1
Normal	Linear	Location	0.244	0.006	0.616	-6065	13**
Normal	Exponential	All	0.066			-11438	1
Normal	Exponential	Location	0.401	0.005	0.981	-8734	11***
Poisson	Linear	All	0.025			-4737	1
Poisson	Linear	Location	Not Shown*	Not Shown*	Not Shown*	Not Shown*	Not Shown*
Poisson	Exponential	All	0.064			-3660	1
Poisson	Exponential	Location	0.406	0.004	0.976	-2694	13**

* Model converged for only Cleveland, Atlanta, and "Other Not MSA" locations. Results are not shown since the model failed to converge for the "Other MSA" location, so the overall goodness-of-fit is not comparable to the other seven models.

** "Other MSA" includes Chicago, Detroit, Philadelphia, Jacksonville, Las Vegas, Provo, St. Louis.

*** "Other MSA" includes Chicago, Cleveland, Detroit, Philadelphia, Jacksonville, Las Vegas, Phoenix, Provo, St. Louis.

For each link function we fitted models using the normal distribution and the Poisson distribution. The normal model is at best an approximation since the numbers of exceedances must be positive or zero integers, but the normal distribution is continuous and includes negative values. The Poisson model takes the form:

$$\text{Prob(y exceedances)} = (M^y/y!)e^{-M},\ y = 0, 1, 2, \ldots,$$

where M is the mean exceedances.

We fitted these four models (two links, two distributions) either to all the data or stratified by location. Thus the model fitted to all the data assumes that a and b have the same value for all site-years, and the model fitted by location assumes that a and b have the same value for all site-years at the same location but these values may vary between locations. For the normal models, the variance of the number of exceedances is assumed to be the same for all site-years in each stratum. For the Poisson models, the variance equals the mean number of exceedances.

The models stratified by location were fitted in two steps. First, each model was separately fitted to each of the 20 locations. For several models and locations, there were problem cases where the algorithm either failed to converge to a solution, predicted a negative slope for the annual mean, or had only zero or one site-year with at least one exceedance. In the second case, if the slope is negative, then the model implies that exceedances decrease when the annual mean increases, which is unexpected and could lead to inconsistent results for projecting exceedances to the current-standard scenario. In the third case, there would be zero degrees of freedom and the model would be over-fitted for that location. To deal with these problem cases, we re-allocated all the problem locations into the “Other MSA” combined location and refitted the models. The results in Table 1 stratified by location are for the refitted models. The re-allocated locations are listed in the footnotes.

Table 1 gives R squared and log-likelihood goodness-of-fit summary statistics. The R squared statistic is the squared Pearson correlation coefficient between the observed number of exceedances and the predicted mean number of exceedances. Negative predicted means are replaced by zero for this calculation. Values close to 1 indicate a good fit and values close to zero indicate a poor fit. For the models stratified by location, it is evident that the R squared value has a wide range across the locations, varying from a very poor fit at some locations to a very good fit at other locations.

For these models the log-likelihood is a better overall goodness-of-fit statistic. The log-likelihood is defined as the logarithm of the fitted joint density function to all 4,177 site-years. The better-fitting models are those with the highest values of the log-likelihood. (The log-likelihood can only be used to compare different models; its value for a single statistical model is not meaningful). Of the various normal models, the best-fitting is stratified by location and uses a linear inverse link. Of the various Poisson models, the best-fitting is stratified by location and uses an exponential inverse link. The Poisson models fit better than the normal models, which is to be expected since the actual data are positive or zero discrete count data and the numbers of exceedances are frequently zero, implying a very small mean.

We selected the Poisson exponential model stratified by location and the normal linear model stratified by location. The estimated parameter values for these models are displayed in Tables 2 and 3, respectively.

The fitted models for the CMSA locations only are displayed in Figures 1 to 3 and are shown for all locations in the attached file “corrplots.selected models.doc.” Figure 1 and the first three attached plots shows the number of exceedances plotted against the annual mean. These plots clearly show how weak the relationship between the exceedances and the annual mean is. Figure 2 and the next three attached plots are for the Poisson exponential model, plotting predicted versus observed exceedances. Figure 3 and the final three attached plots are for the normal linear model, plotting predicted versus observed exceedances (negative predictions were replaced by zero). Comparing the normal and Poisson model predictions, the normal model tends to under-predict the higher numbers of observed exceedances.

The extensive Tables 7 and 8 at the end of this document and the attached Excel file "predictions.selected models.xls" contain predicted values and 95 percent confidence and prediction intervals for the number of exceedances at given mean levels. Table 7 is for the Poisson exponential model. Table 8 is for the normal linear model. Each table gives calculated predictions at annual mean values of 20, 30, 40, 50, 53, and 60 ppb and at the minimum, mean, and maximum annual mean value for each location. The predicted value is the estimated mean number of exceedances.

Tables 4 and 5 are shorter tables showing only the predictions for a mean of 53 ppb and for the mean annual mean. The predictions for a mean of 53 ppb estimate the number of exceedances for a hypothetical site-year with the highest annual mean concentration under the current-standard scenario, i.e., when the highest annual mean site-year for a given location just meets the annual standard. The predictions for a mean equal to the mean annual mean estimate the number of exceedances for the typical “as-is” scenario, i.e., for a hypothetical site-year with an annual mean that is the average annual mean for that location.

Table 2. Parameters for Poisson exponential model stratified by location.

Location Type	**Location Name**	**Parameter***	**Estimate**	**Standard Error**	**Lower Confidence Bound**	**Upper Confidence Bound**	**P-value (Chi-square test that parameter = 0)**
CMSA	Boston-Worcester-Lawrence, MA-NH-ME-CT CMSA	Intercept	-6.887	2.832	-14.693	-2.757	0.02
CMSA	Boston-Worcester-Lawrence, MA-NH-ME-CT CMSA	mean	0.144	0.116	-0.061	0.430	0.22
CMSA	Boston-Worcester-Lawrence, MA-NH-ME-CT CMSA	Scale	1.000	0.000	1.000	1.000	_
CMSA	Cleveland-Akron, OH CMSA	Intercept	-14.209	4.374	-25.210	-7.312	0.00
CMSA	Cleveland-Akron, OH CMSA	mean	0.548	0.164	0.283	0.952	0.00
CMSA	Cleveland-Akron, OH CMSA	Scale	1.000	0.000	1.000	1.000	_
CMSA	Denver-Boulder-Greeley, CO CMSA	Intercept	-4.399	1.186	-7.182	-2.435	0.00
CMSA	Denver-Boulder-Greeley, CO CMSA	mean	0.137	0.038	0.070	0.222	0.00
CMSA	Denver-Boulder-Greeley, CO CMSA	Scale	1.000	0.000	1.000	1.000	_
CMSA	Los Angeles-Riverside-Orange County, CA CMSA	Intercept	-5.628	0.253	-6.134	-5.142	0.00
CMSA	Los Angeles-Riverside-Orange County, CA CMSA	mean	0.181	0.006	0.169	0.194	0.00
CMSA	Los Angeles-Riverside-Orange County, CA CMSA	Scale	1.000	0.000	1.000	1.000	_
CMSA	Miami-Fort Lauderdale, FL CMSA	Intercept	-5.780	1.641	-9.774	-3.068	0.00
CMSA	Miami-Fort Lauderdale, FL CMSA	mean	0.342	0.114	0.138	0.606	0.00
CMSA	Miami-Fort Lauderdale, FL CMSA	Scale	1.000	0.000	1.000	1.000	_
CMSA	New York-Northern New Jersey-Long Island, NY-NJ-CT-PA CMS	Intercept	-6.800	1.269	-9.560	-4.537	0.00
CMSA	New York-Northern New Jersey-Long Island, NY-NJ-CT-PA CMS	mean	0.147	0.037	0.079	0.224	0.00
CMSA	New York-Northern New Jersey-Long Island, NY-NJ-CT-PA CMS	Scale	1.000	0.000	1.000	1.000	_
CMSA	Washington-Baltimore, DC-MD-VA-WV CMSA	Intercept	-6.559	3.054	-14.610	-2.054	0.03
CMSA	Washington-Baltimore, DC-MD-VA-WV CMSA	mean	0.145	0.135	-0.073	0.482	0.28
CMSA	Washington-Baltimore, DC-MD-VA-WV CMSA	Scale	1.000	0.000	1.000	1.000	_
MSA	Atlanta,GA	Intercept	-5.081	1.917	-9.975	-2.139	0.01

Table 2. Parameters for Poisson exponential model stratified by location.

Location Type	**Location Name**	**Parameter***	**Estimate**	**Standard Error**	**Lower Confidence Bound**	**Upper Confidence Bound**	**P-value (Chi-square test that parameter = 0)**
MSA	Atlanta,GA	mean	0.140	0.099	-0.040	0.363	0.16
MSA	Atlanta,GA	Scale	1.000	0.000	1.000	1.000	_
MSA	Colorado Springs,CO	Intercept	-4.846	0.401	-5.675	-4.097	0.00
MSA	Colorado Springs,CO	mean	0.284	0.012	0.261	0.309	0.00
MSA	Colorado Springs,CO	Scale	1.000	0.000	1.000	1.000	_
MSA	El Paso,TX	Intercept	-10.436	2.455	-16.783	-6.664	0.00
MSA	El Paso,TX	mean	0.350	0.074	0.233	0.538	0.00
MSA	El Paso,TX	Scale	1.000	0.000	1.000	1.000	_
MSA	Phoenix-Mesa,AZ	Intercept	-1.568	0.400	-2.363	-0.798	0.00
MSA	Phoenix-Mesa,AZ	mean	0.106	0.013	0.081	0.131	0.00
MSA	Phoenix-Mesa,AZ	Scale	1.000	0.000	1.000	1.000	_
MSA/CMSA	Other MSA/CMSA	Intercept	-5.137	0.222	-5.580	-4.711	0.00
MSA/CMSA	Other MSA/CMSA	mean	0.152	0.010	0.132	0.172	0.00
MSA/CMSA	Other MSA/CMSA	Scale	1.000	0.000	1.000	1.000	_
Not MSA	Other Not MSA	Intercept	-4.672	0.467	-5.654	-3.818	0.00
Not MSA	Other Not MSA	mean	0.227	0.036	0.158	0.300	0.00
Not MSA	Other Not MSA	Scale	1.000	0.000	1.000	1.000	_

* Using the report notation, a = "Intercept", and b = "mean." "Scale" equals 1, by definition, for this model.

Table 3. Parameters for normal linear model stratified by location.

Location Type	Location Name	Parameter*	Estimate	Standard Error	Lower Confidence Bound	Upper Confidence Bound	P-value (Chi-square test that parameter = 0)
CMSA	Boston-Worcester-Lawrence, MA-NH-ME-CT CMSA	Intercept	-0.023	0.034	-0.090	0.043	0.49
CMSA	Boston-Worcester-Lawrence, MA-NH-ME-CT CMSA	mean	0.003	0.002	-0.001	0.006	0.17
CMSA	Boston-Worcester-Lawrence, MA-NH-ME-CT CMSA	Scale	0.135	0.009	0.119	0.156	_
CMSA	Cleveland-Akron, OH CMSA	Intercept	-3.259	2.127	-7.617	1.098	0.13
CMSA	Cleveland-Akron, OH CMSA	mean	0.176	0.099	-0.027	0.378	0.08
CMSA	Cleveland-Akron, OH CMSA	Scale	1.755	0.265	1.341	2.436	_
CMSA	Denver-Boulder-Greeley, CO CMSA	Intercept	-0.439	0.383	-1.211	0.332	0.25
CMSA	Denver-Boulder-Greeley, CO CMSA	mean	0.044	0.018	0.008	0.080	0.01
CMSA	Denver-Boulder-Greeley, CO CMSA	Scale	1.097	0.129	0.885	1.408	_
CMSA	Los Angeles-Riverside-Orange County, CA CMSA	Intercept	-3.301	0.620	-4.519	-2.083	0.00
CMSA	Los Angeles-Riverside-Orange County, CA CMSA	mean	0.194	0.023	0.148	0.240	0.00
CMSA	Los Angeles-Riverside-Orange County, CA CMSA	Scale	4.723	0.174	4.402	5.085	_
CMSA	Miami-Fort Lauderdale, FL CMSA	Intercept	-0.496	0.384	-1.265	0.273	0.20
CMSA	Miami-Fort Lauderdale, FL CMSA	mean	0.070	0.037	-0.005	0.144	0.06
CMSA	Miami-Fort Lauderdale, FL CMSA	Scale	0.828	0.088	0.681	1.036	_
CMSA	New York-Northern New Jersey-Long Island, NY-NJ-CT-PA CMS	Intercept	-0.230	0.104	-0.435	-0.024	0.03
CMSA	New York-Northern New Jersey-Long Island, NY-NJ-CT-PA CMS	mean	0.013	0.004	0.005	0.020	0.00
CMSA	New York-Northern New Jersey-Long Island, NY-NJ-CT-PA CMS	Scale	0.407	0.022	0.368	0.454	_
CMSA	Washington-Baltimore, DC-MD-VA-WV CMSA	Intercept	-0.032	0.069	-0.167	0.104	0.64
CMSA	Washington-Baltimore, DC-MD-VA-WV CMSA	mean	0.003	0.003	-0.004	0.010	0.35
CMSA	Washington-Baltimore, DC-MD-VA-WV CMSA	Scale	0.208	0.013	0.186	0.236	_
MSA	Atlanta,GA	Intercept	-0.041	0.069	-0.178	0.096	0.55
MSA	Atlanta,GA	mean	0.008	0.005	-0.002	0.017	0.11

Table 3. Parameters for normal linear model stratified by location.

Location Type	Location Name	Parameter*	Estimate	Standard Error	Lower Confidence Bound	Upper Confidence Bound	P-value (Chi-square test that parameter = 0)
MSA	Atlanta,GA	Scale	0.226	0.022	0.189	0.277	_
MSA	Colorado Springs,CO	Intercept	-36.358	11.812	-60.391	-12.326	0.00
MSA	Colorado Springs,CO	mean	2.689	0.674	1.318	4.061	0.00
MSA	Colorado Springs,CO	Scale	22.519	3.123	17.551	30.362	_
MSA	El Paso,TX	Intercept	-2.017	0.440	-2.898	-1.135	0.00
MSA	El Paso,TX	mean	0.131	0.024	0.083	0.178	0.00
MSA	El Paso,TX	Scale	0.920	0.098	0.757	1.151	_
MSA	Phoenix-Mesa,AZ	Intercept	-7.102	15.545	-38.177	23.974	0.65
MSA	Phoenix-Mesa,AZ	mean	0.423	0.557	-0.689	1.536	0.45
MSA	Phoenix-Mesa,AZ	Scale	22.513	2.274	18.697	27.828	_
MSA/CMSA	Other MSA/CMSA	Intercept	-0.100	0.051	-0.201	0.000	0.05
MSA/CMSA	Other MSA/CMSA	mean	0.013	0.003	0.006	0.019	0.00
MSA/CMSA	Other MSA/CMSA	Scale	1.098	0.015	1.069	1.128	_
Not MSA	Other Not MSA	Intercept	-0.064	0.049	-0.160	0.031	0.19
Not MSA	Other Not MSA	mean	0.021	0.006	0.009	0.032	0.00
Not MSA	Other Not MSA	Scale	0.549	0.018	0.514	0.587	_

* Using the report notation, a = “Intercept”, b = “mean”, and standard deviation = “Scale.”

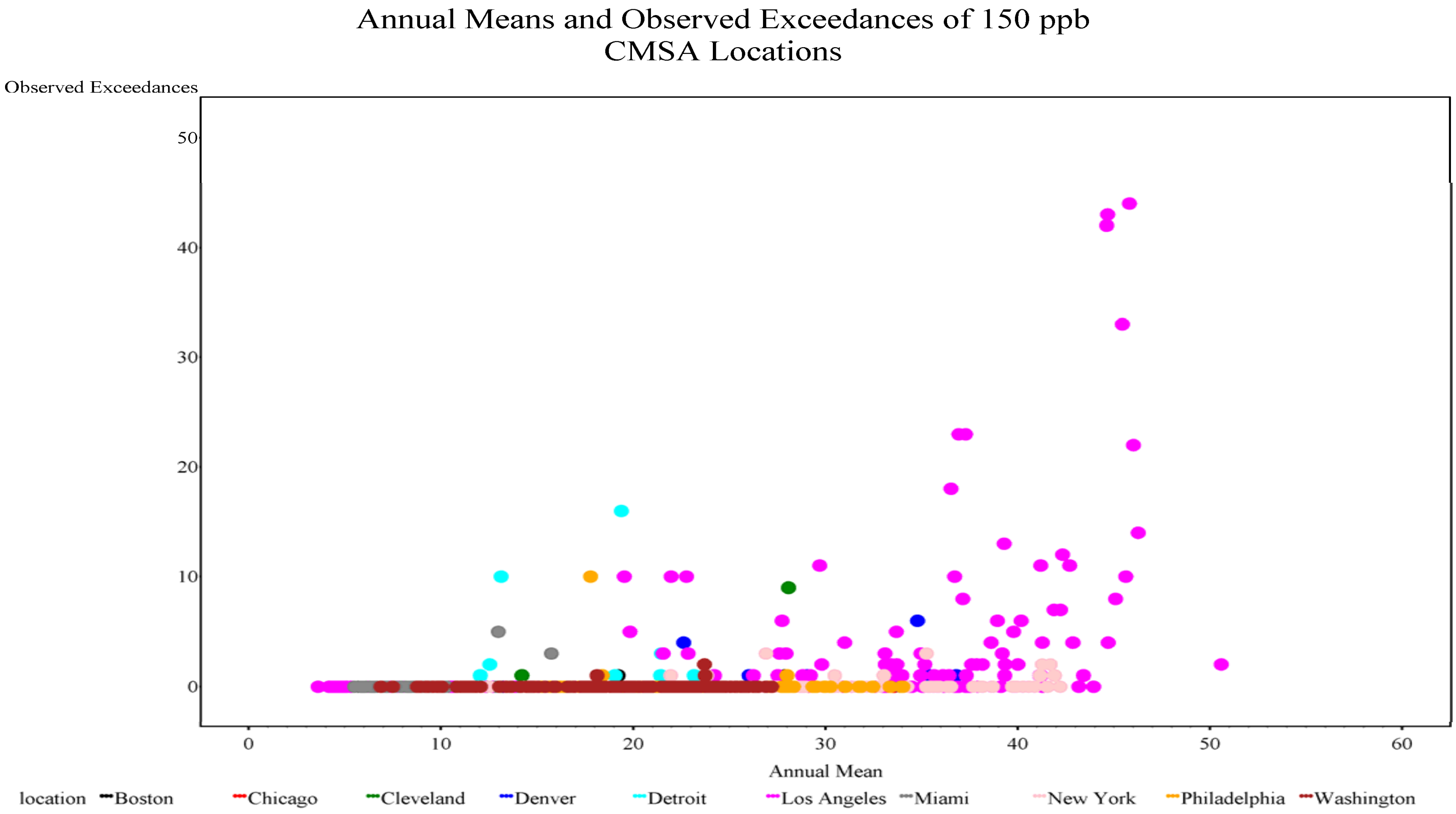


Figure 1. Exceedances of 150 ppb versus annual means for CMSA locations.

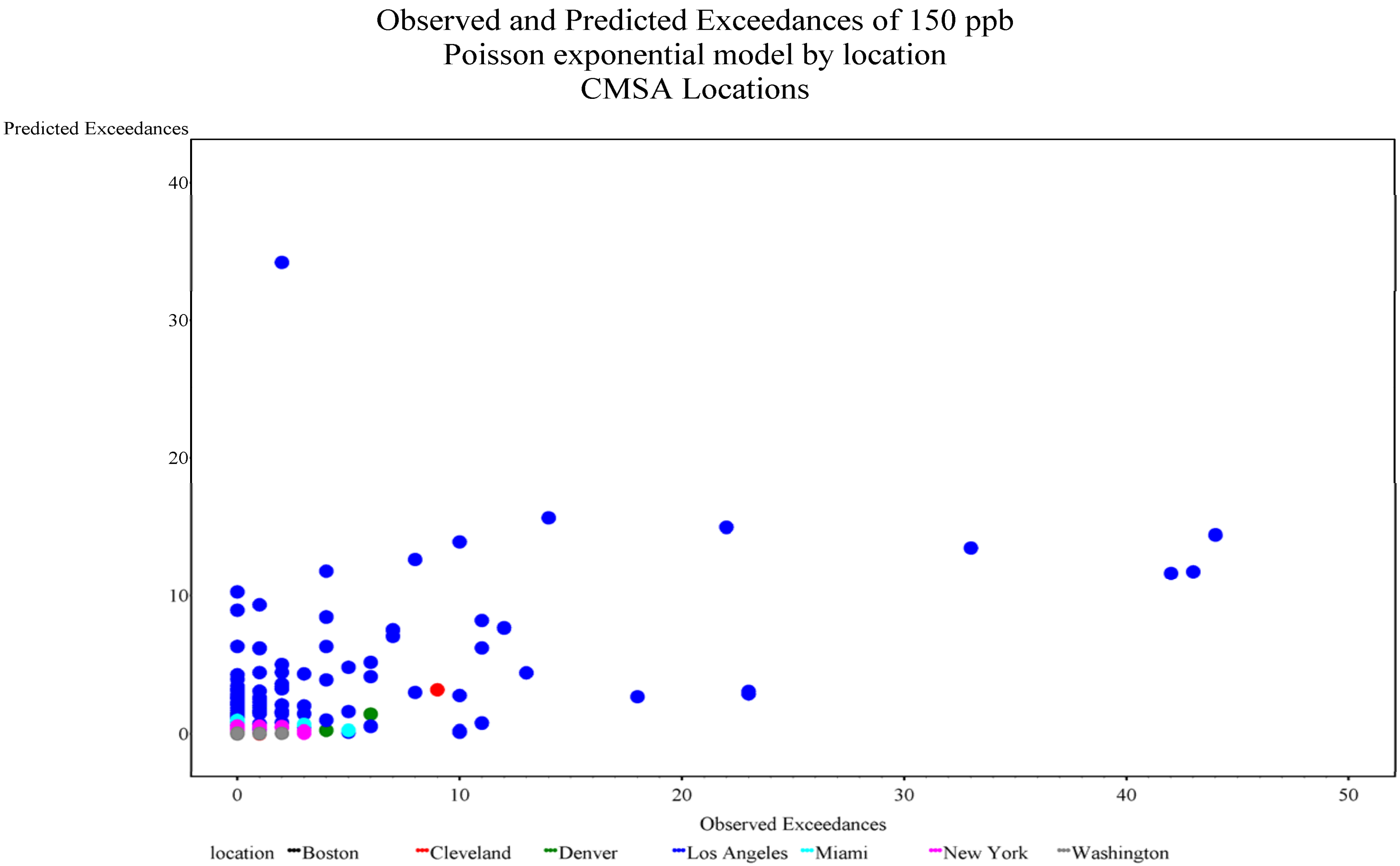


Figure 2. Predicted and observed exceedances for CMSA locations using Poisson exponential model.

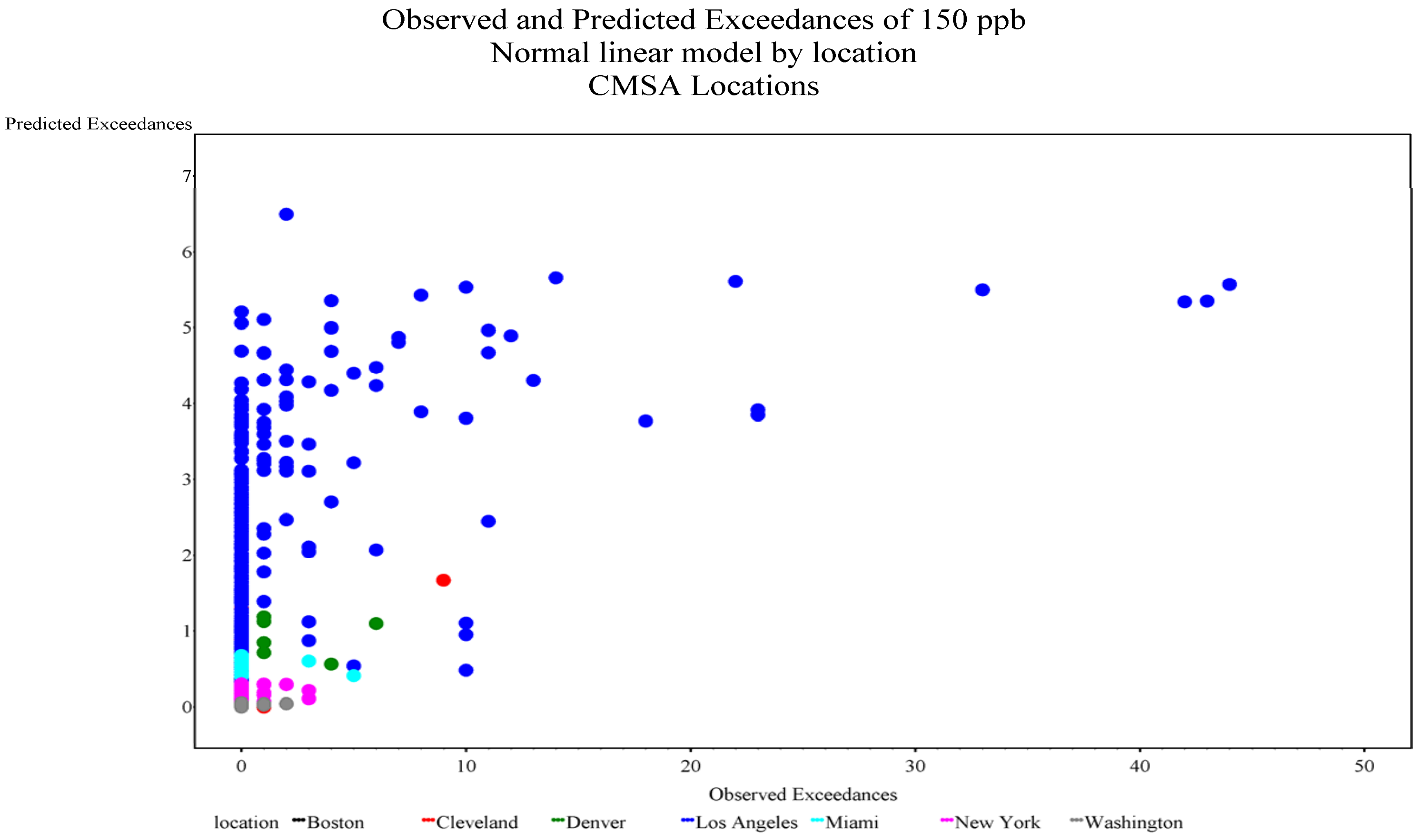

Observed and Predicted Exceedances of 150 ppb
Normal linear model by location
CMSA Locations

Figure 3. Predicted and observed exceedances for CMSA locations using normal linear model

Table 4. As-is and current-standard scenario predictions for Poisson exponential model, with separate coefficients for each location.

					95% Confidence Interval for Mean Number of Exceedances		**95% Prediction Interval for Number of Exceedances**	
Location	**Annual Mean**	**Observed Mean Exceed-ances**	**Observed Max Exceed-ances**	**Predicted Exceed-ances**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
Boston	53.0	0.019	1	2.081	0.002	1000.000	0	1000
Boston	16.8	0.019	1	0.011	0.001	0.091	0	0
Cleveland	53.0	0.455	9	1000.000	578.253	1000.000	364	1000
Cleveland	21.2	0.455	9	0.073	0.011	0.474	0	1
Denver	53.0	0.389	6	17.140	2.958	99.308	2	98
Denver	18.7	0.389	6	0.158	0.057	0.438	0	1
Los Angeles	53.0	1.403	44	53.244	44.092	64.297	37	73
Los Angeles	24.3	1.403	44	0.293	0.238	0.360	0	2
Miami	53.0	0.182	5	1000.000	35.520	1000.000	29	1000
Miami	9.7	0.182	5	0.086	0.026	0.281	0	1
New York	53.0	0.092	3	2.737	0.646	11.604	0	13
New York	25.5	0.092	3	0.048	0.022	0.104	0	1
Washington	53.0	0.030	2	3.038	0.001	1000.000	0	1000
Washington	19.4	0.030	2	0.023	0.007	0.082	0	0
Atlanta	53.0	0.057	1	10.242	0.012	1000.000	0	1000
Atlanta	12.9	0.057	1	0.038	0.008	0.181	0	1
Colorado Springs	53.0	7.346	143	1000.000	1000.000	1000.000	1000	1000
Colorado Springs	16.3	7.346	143	0.792	0.528	1.189	0	3
El Paso	53.0	0.295	7	1000.000	177.602	1000.000	156	1000
El Paso	17.7	0.295	7	0.015	0.001	0.142	0	1
Phoenix	53.0	4.469	147	56.901	31.702	102.130	26	106
Phoenix	27.3	4.469	147	3.760	3.221	4.389	0	8
Other MSA/CMSA	53.0	0.079	39	18.369	9.388	35.940	7	41
Other MSA/CMSA	13.9	0.079	39	0.048	0.040	0.058	0	1
Other Not MSA	53.0	0.081	7	1000.000	85.717	1000.000	75	1000
Other Not MSA	7.0	0.081	7	0.046	0.028	0.075	0	1

Table 5. As-is and current-standard scenario predictions for Normal linear model, with separate coefficients for each location.								
					95% Confidence Interval for Mean Number of Exceedances		**95% Prediction Interval for Number of Exceedances**	
Location Name	**Annual Mean**	**Observed Mean Exceed-ances**	**Observed Max Exceed-ances**	**Predicted Exceed-ances**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
Boston	53.0	0.019	1	0.111	0.000	0.245	0.000	0.412
Boston	16.8	0.019	1	0.019	0.000	0.045	0.000	0.289
Cleveland	53.0	0.455	9	6.046	0.000	12.267	0.000	13.612
Cleveland	21.2	0.455	9	0.455	0.000	1.188	0.000	4.198
Denver	53.0	0.389	6	1.906	0.645	3.168	0.000	4.490
Denver	18.7	0.389	6	0.389	0.031	0.747	0.000	2.648
Los Angeles	53.0	1.403	44	6.965	5.561	8.369	0.000	16.360
Los Angeles	24.3	1.403	44	1.403	0.921	1.884	0.000	10.703
Miami	53.0	0.182	5	3.199	0.024	6.375	0.000	6.871
Miami	9.7	0.182	5	0.182	0.000	0.426	0.000	1.871
New York	53.0	0.092	3	0.439	0.220	0.658	0.000	1.272
New York	25.5	0.092	3	0.092	0.031	0.152	0.000	0.897
Washington	53.0	0.030	2	0.136	0.000	0.364	0.000	0.608
Washington	19.4	0.030	2	0.030	0.000	0.065	0.000	0.443
Atlanta	53.0	0.057	1	0.360	0.000	0.739	0.000	0.957
Atlanta	12.9	0.057	1	0.057	0.000	0.117	0.000	0.514
Colorado Springs	53.0	7.346	143	106.169	56.853	155.486	36.477	175.862
Colorado Springs	16.3	7.346	143	7.346	0.000	16.002	0.000	54.709
El Paso	53.0	0.295	7	4.902	3.249	6.555	2.384	7.421
El Paso	17.7	0.295	7	0.295	0.024	0.567	0.000	2.172
Phoenix	53.0	4.469	147	15.339	0.000	44.043	0.000	69.369
Phoenix	27.3	4.469	147	4.469	0.000	10.773	0.000	50.219
Other MSA/CMSA	53.0	0.079	39	0.584	0.324	0.844	0.000	2.752
Other MSA/CMSA	13.9	0.079	39	0.079	0.037	0.120	0.000	2.232
Other Not MSA	53.0	0.081	7	1.036	0.505	1.566	0.000	2.238
Other Not MSA	7.0	0.081	7	0.081	0.030	0.132	0.000	1.161

The 95% confidence interval gives the uncertainty of the expected value, i.e., of the average number of exceedances over hypothetically infinitely many site-years with the same annual mean. The 95% prediction interval gives the uncertainty of the value for a single site-year, taking into account both the uncertainty of the estimated parameters and the variability of the number of exceedances in a given site-year about the overall mean. All prediction intervals were truncated to be greater than or equal to zero and less than or equal to 1,000. The maximum possible number of exceedances in a year is the maximum number of hours in a leap year, 8,784. The maximum observed exceedances in a year was 69.

For annual means within the range of the data, the predicted numbers of exceedances are generally within the range of the observed numbers of exceedances. The normal model predictions tend to be lower than the Poisson model predictions. At annual mean levels above the range of the data, the Poisson model with the exponential inverse link sometimes gives extremely high estimates, well beyond the truncation limit of 1,000. This is mainly due to the exponential link; each increase of the annual mean by 1 ppb increases the predicted exceedances by a multiplicative factor of exp(b), where $b > 0$. The upper bounds of the normal linear model prediction intervals are at most a more reasonable 202, but these predictions are less reliable because the Poisson model with an exponential inverse link fits the data much better. For the normal linear model, each increase of the annual mean by 1 ppb increases the predicted exceedances by b ppb.

Not shown here are the results for the normal model with an exponential inverse link, which was the model formulation selected by McCurdy (1994). That model gives roughly similar predictions to the Poisson model with the exponential inverse link.

We can compare these predictions with the predictions for Los Angeles from McCurdy (1994) based on 1988-1992 data. Table 6 gives the McCurdy (1994) exceedance estimates for exceedances of 150 ppb together with our estimates for the 1995-2006 data based on the Poisson exponential model (see Table 7) and the normal linear model (see Table 8). It is easily seen that the McCurdy (1994) estimates agree reasonably well with our Poisson exponential model predictions, with predicted exceedances being a little lower for annual means up to 53 ppb, but a little higher at 60 ppb. The McCurdy (1994) model predicts 75 exceedances at 53 ppb, compared to our Poisson exponential model prediction of 53 exceedances. However, the McCurdy (1994) estimates are all much higher than our normal linear model predictions. For example, the McCurdy (1994) model predicts 75 exceedances at 53 ppb, compared to our normal linear model prediction of 7 exceedances. These findings are primarily due to the fact that McCurdy also used an exponential link function.

Table 6. Comparison of predicted exceedances of 150 ppb using McCurdy (1994) for 1988-1992 data and the Poisson exponential and normal linear models for 1995-2006 data.

Annual Mean (ppb)	**Predicted Exceedances of 150 ppb**		
	McCurdy (1994) normal exponential model. 1988-1992 data.	**Poisson exponential model. 1995-2006 data.**	**Normal linear model. 1995-2006 data.**
20	4	0	1
30	9	1	3
40	33	5	4
50	57	31	6
53	75	53	7
60	142	189	8

CONCLUSION

These analyses found a poor relationship between the annual means and the exceedances of 150 ppb, as well as frequently unrealistically high predictions of exceedances of 150 ppb for the current-standard scenario. The uncertainty at higher exceedance threshold concentration levels (200 to 300 ppb) would be expected to be even higher because the numbers of site-years with non-zero exceedances are even lower (which implies a much weaker numerical relationship between the annual mean and the annual exceedances). For example, for Los Angeles, the maximum number of exceedances of 150 ppb was 44, but the maximum number of exceedances of 200 ppb was only 5. Therefore we chose not to continue the regression analyses to higher exceedance threshold concentration levels.

DETAILED PREDICTION TABLES

Table 7. Predictions for Poisson exponential model, with separate coefficients for each location.								
					95% Confidence Interval for Mean Number of Exceedances		**95% Prediction Interval for Number of Exceedances**	
Location	**Annual Mean**	**Observed Mean Exceed-ances**	**Observed Max Exceed-ances**	**Predicted Exceed-ances**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
Boston	20.0	0.019	1	0.018	0.004	0.090	0	1
Boston	30.0	0.019	1	0.076	0.010	0.576	0	1
Boston	40.0	0.019	1	0.321	0.006	17.564	0	14
Boston	50.0	0.019	1	1.352	0.003	661.873	0	680
Boston	53.0	0.019	1	2.081	0.002	1000.000	0	1000
Boston	60.0	0.019	1	5.692	0.001	1000.000	0	1000
Boston	5.4	0.019	1	0.002	0.000	0.175	0	0
Boston	16.8	0.019	1	0.011	0.001	0.091	0	0
Boston	31.0	0.019	1	0.089	0.010	0.801	0	1
Cleveland	20.0	0.455	9	0.039	0.004	0.358	0	1

Table 7. Predictions for Poisson exponential model, with separate coefficients for each location.

					95% Confidence Interval for Mean Number of Exceedances		95% Prediction Interval for Number of Exceedances	
Location	**Annual Mean**	**Observed Mean Exceed-ances**	**Observed Max Exceed-ances**	**Predicted Exceed-ances**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
Cleveland	30.0	0.455	9	9.244	2.693	31.732	2	32
Cleveland	40.0	0.455	9	1000.000	29.509	1000.000	23	1000
Cleveland	50.0	0.455	9	1000.000	291.652	1000.000	184	1000
Cleveland	53.0	0.455	9	1000.000	578.253	1000.000	364	1000
Cleveland	60.0	0.455	9	1000.000	1000.000	1000.000	1000	1000
Cleveland	14.2	0.455	9	0.002	0.000	0.092	0	0
Cleveland	21.2	0.455	9	0.073	0.011	0.474	0	1
Cleveland	28.1	0.455	9	3.193	1.490	6.845	0	9
Denver	20.0	0.389	6	0.189	0.074	0.482	0	2
Denver	30.0	0.389	6	0.740	0.438	1.251	0	3
Denver	40.0	0.389	6	2.902	1.201	7.014	0	9
Denver	50.0	0.389	6	11.376	2.426	53.350	1	53
Denver	53.0	0.389	6	17.140	2.958	99.308	2	98
Denver	60.0	0.389	6	44.600	4.659	426.973	4	454
Denver	6.1	0.389	6	0.028	0.004	0.186	0	1
Denver	18.7	0.389	6	0.158	0.057	0.438	0	1
Denver	36.8	0.389	6	1.871	0.925	3.786	0	6
Los Angeles	20.0	1.403	44	0.135	0.104	0.174	0	1
Los Angeles	30.0	1.403	44	0.825	0.713	0.954	0	3
Los Angeles	40.0	1.403	44	5.050	4.632	5.505	1	10
Los Angeles	50.0	1.403	44	30.917	26.439	36.154	20	44
Los Angeles	53.0	1.403	44	53.244	44.092	64.297	37	73
Los Angeles	60.0	1.403	44	189.281	144.681	247.629	138	260
Los Angeles	3.6	1.403	44	0.007	0.004	0.011	0	0
Los Angeles	24.3	1.403	44	0.293	0.238	0.360	0	2
Los Angeles	50.6	1.403	44	34.208	29.084	40.236	22	48
Miami	20.0	0.182	5	2.882	0.636	13.069	0	13
Miami	30.0	0.182	5	88.023	2.282	1000.000	2	1000
Miami	40.0	0.182	5	1000.000	7.591	1000.000	7	1000
Miami	50.0	0.182	5	1000.000	24.900	1000.000	33	1000
Miami	53.0	0.182	5	1000.000	35.520	1000.000	29	1000
Miami	60.0	0.182	5	1000.000	81.274	1000.000	40	1000
Miami	5.5	0.182	5	0.020	0.003	0.154	0	1
Miami	9.7	0.182	5	0.086	0.026	0.281	0	1
Miami	16.8	0.182	5	0.970	0.380	2.475	0	4
New York	20.0	0.092	3	0.021	0.007	0.065	0	0
New York	30.0	0.092	3	0.092	0.052	0.163	0	1
New York	40.0	0.092	3	0.403	0.211	0.773	0	2
New York	50.0	0.092	3	1.760	0.507	6.107	0	7
New York	53.0	0.092	3	2.737	0.646	11.604	0	13

Table 7. Predictions for Poisson exponential model, with separate coefficients for each location.

					95% Confidence Interval for Mean Number of Exceedances		95% Prediction Interval for Number of Exceedances	
Location	**Annual Mean**	**Observed Mean Exceed-ances**	**Observed Max Exceed-ances**	**Predicted Exceed-ances**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
New York	60.0	0.092	3	7.677	1.121	52.548	0	53
New York	9.7	0.092	3	0.005	0.001	0.028	0	0
New York	25.5	0.092	3	0.048	0.022	0.104	0	1
New York	42.2	0.092	3	0.557	0.260	1.193	0	3
Washington	20.0	0.030	2	0.026	0.008	0.081	0	1
Washington	30.0	0.030	2	0.109	0.011	1.044	0	2
Washington	40.0	0.030	2	0.463	0.004	55.438	0	57
Washington	50.0	0.030	2	1.968	0.001	1000.000	0	1000
Washington	53.0	0.030	2	3.038	0.001	1000.000	0	1000
Washington	60.0	0.030	2	8.368	0.000	1000.000	0	1000
Washington	6.9	0.030	2	0.004	0.000	0.256	0	1
Washington	19.4	0.030	2	0.023	0.007	0.082	0	0
Washington	27.2	0.030	2	0.072	0.014	0.366	0	1
Atlanta	20.0	0.057	1	0.102	0.032	0.327	0	1
Atlanta	30.0	0.057	1	0.412	0.034	4.953	0	5
Atlanta	40.0	0.057	1	1.665	0.023	122.647	0	103
Atlanta	50.0	0.057	1	6.735	0.014	1000.000	0	1000
Atlanta	53.0	0.057	1	10.242	0.012	1000.000	0	1000
Atlanta	60.0	0.057	1	27.243	0.008	1000.000	0	1000
Atlanta	3.4	0.057	1	0.010	0.000	0.230	0	0
Atlanta	12.9	0.057	1	0.038	0.008	0.181	0	1
Atlanta	26.6	0.057	1	0.257	0.037	1.770	0	3
Colorado Springs	20.0	7.346	143	2.295	1.662	3.168	0	6
Colorado Springs	30.0	7.346	143	39.206	33.759	45.531	26	53
Colorado Springs	40.0	7.346	143	669.766	526.509	852.001	523	870
Colorado Springs	50.0	7.346	143	1000.000	1000.000	1000.000	1000	1000
Colorado Springs	53.0	7.346	143	1000.000	1000.000	1000.000	1000	1000
Colorado Springs	60.0	7.346	143	1000.000	1000.000	1000.000	1000	1000
Colorado Springs	6.8	7.346	143	0.054	0.029	0.102	0	1
Colorado Springs	16.3	7.346	143	0.792	0.528	1.189	0	3
Colorado Springs	34.8	7.346	143	153.247	130.906	179.401	121	189
El Paso	20.0	0.295	7	0.032	0.005	0.230	0	1
El Paso	30.0	0.295	7	1.075	0.536	2.156	0	4

Table 7. Predictions for Poisson exponential model, with separate coefficients for each location.

					95% Confidence Interval for Mean Number of Exceedances		**95% Prediction Interval for Number of Exceedances**	
Location	**Annual Mean**	**Observed Mean Exceed-ances**	**Observed Max Exceed-ances**	**Predicted Exceed-ances**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
El Paso	40.0	0.295	7	35.703	11.290	112.906	11	119
El Paso	50.0	0.295	7	1000.000	95.081	1000.000	94	1000
El Paso	53.0	0.295	7	1000.000	177.602	1000.000	156	1000
El Paso	60.0	0.295	7	1000.000	757.520	1000.000	634	1000
El Paso	8.2	0.295	7	0.001	0.000	0.020	0	0
El Paso	17.7	0.295	7	0.015	0.001	0.142	0	1
El Paso	35.1	0.295	7	6.447	3.454	12.036	1	14
Phoenix	20.0	4.469	147	1.731	1.287	2.329	0	5
Phoenix	30.0	4.469	147	4.988	4.367	5.698	1	10
Phoenix	40.0	4.469	147	14.375	10.922	18.919	7	24
Phoenix	50.0	4.469	147	41.422	24.843	69.066	21	71
Phoenix	53.0	4.469	147	56.901	31.702	102.130	26	106
Phoenix	60.0	4.469	147	119.362	55.901	254.864	56	254
Phoenix	11.1	4.469	147	0.673	0.404	1.119	0	3
Phoenix	27.3	4.469	147	3.760	3.221	4.389	0	8
Phoenix	40.5	4.469	147	15.110	11.361	20.098	7	25
Other MSA/CMSA	20.0	0.079	39	0.122	0.107	0.140	0	1
Other MSA/CMSA	30.0	0.079	39	0.559	0.442	0.707	0	2
Other MSA/CMSA	40.0	0.079	39	2.552	1.681	3.874	0	6
Other MSA/CMSA	50.0	0.079	39	11.648	6.317	21.480	4	25
Other MSA/CMSA	53.0	0.079	39	18.369	9.388	35.940	7	41
Other MSA/CMSA	60.0	0.079	39	53.171	23.650	119.541	20	116
Other MSA/CMSA	0.5	0.079	39	0.006	0.004	0.010	0	0
Other MSA/CMSA	13.9	0.079	39	0.048	0.040	0.058	0	1
Other MSA/CMSA	34.0	0.079	39	1.025	0.756	1.391	0	4
Other Not MSA	20.0	0.081	7	0.878	0.459	1.681	0	3
Other Not MSA	30.0	0.081	7	8.514	2.297	31.556	1	32
Other Not MSA	40.0	0.081	7	82.532	11.133	611.822	10	573
Other Not MSA	50.0	0.081	7	799.989	53.545	1000.000	57	1000
Other Not MSA	53.0	0.081	7	1000.000	85.717	1000.000	75	1000
Other Not MSA	60.0	0.081	7	1000.000	256.785	1000.000	226	1000
Other Not MSA	0.3	0.081	7	0.010	0.004	0.025	0	0
Other Not MSA	7.0	0.081	7	0.046	0.028	0.075	0	1

Table 7. Predictions for Poisson exponential model, with separate coefficients for each location.

					95% Confidence Interval for Mean Number of Exceedances		**95% Prediction Interval for Number of Exceedances**	
Location	**Annual Mean**	**Observed Mean Exceed-ances**	**Observed Max Exceed-ances**	**Predicted Exceed-ances**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
Other Not MSA	19.7	0.081	7	0.823	0.438	1.547	0	3

Table 8. Predictions for Normal linear model, with separate coefficients for each location.								
					95% Confidence Interval for Mean Number of Exceedances		**95% Prediction Interval for Number of Exceedances**	
Location Name	**Annual Mean**	**Observed Mean Exceed-ances**	**Observed Max Exceed-ances**	**Predicted Exceed-ances**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
Boston	20.0	0.019	1	0.027	0.000	0.056	0.000	0.297
Boston	30.0	0.019	1	0.052	0.000	0.107	0.000	0.327
Boston	40.0	0.019	1	0.078	0.000	0.166	0.000	0.361
Boston	50.0	0.019	1	0.103	0.000	0.226	0.000	0.399
Boston	53.0	0.019	1	0.111	0.000	0.245	0.000	0.412
Boston	60.0	0.019	1	0.128	0.000	0.287	0.000	0.441
Boston	5.4	0.019	1	0.000	0.000	0.039	0.000	0.263
Boston	16.8	0.019	1	0.019	0.000	0.045	0.000	0.289
Boston	31.0	0.019	1	0.055	0.000	0.113	0.000	0.330
Cleveland	20.0	0.455	9	0.252	0.000	1.019	0.000	4.003
Cleveland	30.0	0.455	9	2.008	0.141	3.874	0.000	6.173
Cleveland	40.0	0.455	9	3.763	0.035	7.492	0.000	9.163
Cleveland	50.0	0.455	9	5.519	0.000	11.163	0.000	12.553
Cleveland	53.0	0.455	9	6.046	0.000	12.267	0.000	13.612
Cleveland	60.0	0.455	9	7.275	0.000	14.846	0.000	16.125
Cleveland	14.2	0.455	9	0.000	0.000	0.769	0.000	3.243
Cleveland	21.2	0.455	9	0.455	0.000	1.188	0.000	4.198
Cleveland	28.1	0.455	9	1.667	0.140	3.194	0.000	5.673
Denver	20.0	0.389	6	0.446	0.085	0.807	0.000	2.706
Denver	30.0	0.389	6	0.888	0.353	1.424	0.000	3.185
Denver	40.0	0.389	6	1.331	0.499	2.163	0.000	3.720
Denver	50.0	0.389	6	1.773	0.613	2.934	0.000	4.306
Denver	53.0	0.389	6	1.906	0.645	3.168	0.000	4.490
Denver	60.0	0.389	6	2.216	0.716	3.716	0.000	4.933
Denver	6.1	0.389	6	0.000	0.000	0.402	0.000	2.136
Denver	18.7	0.389	6	0.389	0.031	0.747	0.000	2.648
Denver	36.8	0.389	6	1.189	0.458	1.920	0.000	3.543
Los Angeles	20.0	1.403	44	0.573	0.053	1.093	0.000	9.876
Los Angeles	30.0	1.403	44	2.510	1.962	3.058	0.000	11.814
Los Angeles	40.0	1.403	44	4.447	3.579	5.315	0.000	13.776
Los Angeles	50.0	1.403	44	6.384	5.109	7.660	0.000	15.760
Los Angeles	53.0	1.403	44	6.965	5.561	8.369	0.000	16.360
Los Angeles	60.0	1.403	44	8.321	6.612	10.031	0.000	17.766
Los Angeles	3.6	1.403	44	0.000	0.000	0.000	0.000	6.747
Los Angeles	24.3	1.403	44	1.403	0.921	1.884	0.000	10.703
Los Angeles	50.6	1.403	44	6.492	5.193	7.792	0.000	15.871
Miami	20.0	0.182	5	0.899	0.108	1.689	0.000	2.757
Miami	30.0	0.182	5	1.596	0.092	3.099	0.000	3.873
Miami	40.0	0.182	5	2.293	0.065	4.521	0.000	5.131
Miami	50.0	0.182	5	2.990	0.034	5.947	0.000	6.463

Table 8. Predictions for Normal linear model, with separate coefficients for each location.

					95% Confidence Interval for Mean Number of Exceedances		95% Prediction Interval for Number of Exceedances	
Location Name	**Annual Mean**	**Observed Mean Exceedances**	**Observed Max Exceedances**	**Predicted Exceedances**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
Miami	53.0	0.182	5	3.199	0.024	6.375	0.000	6.871
Miami	60.0	0.182	5	3.687	0.001	7.373	0.000	7.834
Miami	5.5	0.182	5	0.000	0.000	0.281	0.000	1.607
Miami	9.7	0.182	5	0.182	0.000	0.426	0.000	1.871
Miami	16.8	0.182	5	0.677	0.103	1.250	0.000	2.449
New York	20.0	0.092	3	0.023	0.000	0.096	0.000	0.829
New York	30.0	0.092	3	0.149	0.079	0.218	0.000	0.955
New York	40.0	0.092	3	0.275	0.148	0.401	0.000	1.088
New York	50.0	0.092	3	0.401	0.204	0.598	0.000	1.228
New York	53.0	0.092	3	0.439	0.220	0.658	0.000	1.272
New York	60.0	0.092	3	0.527	0.256	0.798	0.000	1.375
New York	9.7	0.092	3	0.000	0.000	0.028	0.000	0.707
New York	25.5	0.092	3	0.092	0.031	0.152	0.000	0.897
New York	42.2	0.092	3	0.302	0.161	0.444	0.000	1.118
Washington	20.0	0.030	2	0.032	0.000	0.067	0.000	0.445
Washington	30.0	0.030	2	0.063	0.000	0.143	0.000	0.483
Washington	40.0	0.030	2	0.095	0.000	0.237	0.000	0.531
Washington	50.0	0.030	2	0.127	0.000	0.335	0.000	0.589
Washington	53.0	0.030	2	0.136	0.000	0.364	0.000	0.608
Washington	60.0	0.030	2	0.158	0.000	0.432	0.000	0.654
Washington	6.9	0.030	2	0.000	0.000	0.081	0.000	0.412
Washington	19.4	0.030	2	0.030	0.000	0.065	0.000	0.443
Washington	27.2	0.030	2	0.054	0.000	0.117	0.000	0.471
Atlanta	20.0	0.057	1	0.110	0.020	0.201	0.000	0.573
Atlanta	30.0	0.057	1	0.186	0.015	0.357	0.000	0.672
Atlanta	40.0	0.057	1	0.262	0.001	0.522	0.000	0.787
Atlanta	50.0	0.057	1	0.337	0.000	0.689	0.000	0.916
Atlanta	53.0	0.057	1	0.360	0.000	0.739	0.000	0.957
Atlanta	60.0	0.057	1	0.413	0.000	0.857	0.000	1.055
Atlanta	3.4	0.057	1	0.000	0.000	0.092	0.000	0.452
Atlanta	12.9	0.057	1	0.057	0.000	0.117	0.000	0.514
Atlanta	26.6	0.057	1	0.161	0.019	0.303	0.000	0.637
Colorado Springs	20.0	7.346	143	17.426	7.454	27.398	0.000	65.075
Colorado Springs	30.0	7.346	143	44.318	24.197	64.439	0.000	95.397
Colorado Springs	40.0	7.346	143	71.210	38.662	103.758	13.462	128.958
Colorado Springs	50.0	7.346	143	98.102	52.682	143.522	31.411	164.793
Colorado Springs	53.0	7.346	143	106.169	56.853	155.486	36.477	175.862

Table 8. Predictions for Normal linear model, with separate coefficients for each location.								
					95% Confidence Interval for Mean Number of Exceedances		**95% Prediction Interval for Number of Exceedances**	
Location Name	**Annual Mean**	**Observed Mean Exceed-ances**	**Observed Max Exceed-ances**	**Predicted Exceed-ances**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
Colorado Springs	60.0	7.346	143	124.994	66.550	183.438	47.873	202.115
Colorado Springs	6.8	7.346	143	0.000	0.000	0.000	0.000	31.109
Colorado Springs	16.3	7.346	143	7.346	0.000	16.002	0.000	54.709
Colorado Springs	34.8	7.346	143	57.235	31.241	83.228	3.296	111.173
El Paso	20.0	0.295	7	0.594	0.303	0.886	0.000	2.474
El Paso	30.0	0.295	7	1.900	1.270	2.529	0.000	3.866
El Paso	40.0	0.295	7	3.205	2.140	4.270	1.049	5.361
El Paso	50.0	0.295	7	4.511	2.994	6.027	2.085	6.936
El Paso	53.0	0.295	7	4.902	3.249	6.555	2.384	7.421
El Paso	60.0	0.295	7	5.816	3.844	7.789	3.065	8.568
El Paso	8.2	0.295	7	0.000	0.000	0.000	0.000	0.981
El Paso	17.7	0.295	7	0.295	0.024	0.567	0.000	2.172
El Paso	35.1	0.295	7	2.567	1.719	3.416	0.516	4.619
Phoenix	20.0	4.469	147	1.367	0.000	11.546	0.000	47.846
Phoenix	30.0	4.469	147	5.601	0.000	12.546	0.000	51.449
Phoenix	40.0	4.469	147	9.835	0.000	25.027	0.000	57.734
Phoenix	50.0	4.469	147	14.069	0.000	39.591	0.000	66.390
Phoenix	53.0	4.469	147	15.339	0.000	44.043	0.000	69.369
Phoenix	60.0	4.469	147	18.303	0.000	54.495	0.000	76.880
Phoenix	11.1	4.469	147	0.000	0.000	16.406	0.000	46.824
Phoenix	27.3	4.469	147	4.469	0.000	10.773	0.000	50.219
Phoenix	40.5	4.469	147	10.035	0.000	25.696	0.000	58.093
Other MSA/CMSA	20.0	0.079	39	0.158	0.100	0.216	0.000	2.311
Other MSA/CMSA	30.0	0.079	39	0.287	0.173	0.401	0.000	2.442
Other MSA/CMSA	40.0	0.079	39	0.416	0.239	0.593	0.000	2.576
Other MSA/CMSA	50.0	0.079	39	0.545	0.304	0.786	0.000	2.711
Other MSA/CMSA	53.0	0.079	39	0.584	0.324	0.844	0.000	2.752
Other MSA/CMSA	60.0	0.079	39	0.674	0.368	0.980	0.000	2.848
Other MSA/CMSA	0.5	0.079	39	0.000	0.000	0.003	0.000	2.061
Other MSA/CMSA	13.9	0.079	39	0.079	0.037	0.120	0.000	2.232
Other MSA/CMSA	34.0	0.079	39	0.339	0.200	0.477	0.000	2.495

Table 8. Predictions for Normal linear model, with separate coefficients for each location.

					95% Confidence Interval for Mean Number of Exceedances		**95% Prediction Interval for Number of Exceedances**	
Location Name	**Annual Mean**	**Observed Mean Exceed-ances**	**Observed Max Exceed-ances**	**Predicted Exceed-ances**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
Other Not MSA	20.0	0.081	7	0.351	0.193	0.508	0.000	1.440
Other Not MSA	30.0	0.081	7	0.558	0.290	0.827	0.000	1.669
Other Not MSA	40.0	0.081	7	0.766	0.384	1.148	0.000	1.910
Other Not MSA	50.0	0.081	7	0.973	0.477	1.469	0.000	2.161
Other Not MSA	53.0	0.081	7	1.036	0.505	1.566	0.000	2.238
Other Not MSA	60.0	0.081	7	1.181	0.571	1.791	0.000	2.421
Other Not MSA	0.3	0.081	7	0.000	0.000	0.035	0.000	1.024
Other Not MSA	7.0	0.081	7	0.081	0.030	0.132	0.000	1.161
Other Not MSA	19.7	0.081	7	0.345	0.190	0.499	0.000	1.434

Appendix E. Technical Memorandum on Land Use and Surface Analysis

Figures E-1 to E-5 show the manually created land-use sectors around each application site; in each case a 1.9 mile (3 km) radius circle was used. The city centers are also labeled. Data in each case are from the NLCD92 database. Prior to the release of AERSURFACE, the user was required to manually pull values of Bowen ratio (β_0), albedo (α), and surface roughness (z_0) per season and per land-use sector from look-up tables in the *AERMET User's Guide*. Using the look-up tables, values of these three surface characteristics vary by the four seasons and by eight basic land-use categories. Furthermore, the *AERMOD Implementation Guide* was somewhat ambiguous about whether Bowen ratio values should also vary with wind direction sector, as does the surface roughness. AERSURFACE resolves these issues by providing a uniform methodology for calculation of surface effects on dispersion; it also only varies surface roughness by wind direction.

Before AERSURFACE, without an automated algorithm to determine land-use patterns, it was simplest for the user to visually estimate land usage by sector. With AERSURFACE, the land-use is automatically determined. The proximity of the meteorological site to an airport and whether the site was located in an arid region were previously not explicitly accounted for as they now are in AERSURFACE. Snow cover, too, is critical for determination of α, but was largely left to user's discretion regarding its presence. With AERSURFACE, the lookup tables have separate columns for winter without much snow and for winter with abundant snow. The user determines if winter at a particular location contains at least one month of continuous snow cover, and AERSURFACE will pull values of the surface characteristics from the appropriate winter column.

We conducted a sensitivity test to evaluate the impacts of using this new tool on the present analysis. Figure F-6 shows a sample comparison of surface roughness values at the Philadelphia site with and without the use of AERSURFACE. In the Figure, estimated surface roughness values using visual land-use estimations and look-up table values are shown in muted shades and AERSURFACE values in dark shades. Monthly season definitions are the same in both cases. However, in the AERSURFACE case, winter was specified as having a one-month period of snow cover. Also, in the AERSURFACE case the site was specified as being at an airport.

In this case, z_0 values are much lower with AERSURFACE than with a visual estimation of land-use. In the AERSURFACE tool, Philadelphia was noted as being at an airport, tending to represent the lower building heights in the region and the inverse distance weighting implemented in the tool. Thus, lower z_0 values were obtained over most developed-area sectors in this scenario. The indication that at least one month of continuous snow cover is present also tends to lower wintertime z_0 values. In addition to these systematic differences, the automated AERSURFACE land-use analysis for Philadelphia tended to identify less urban coverage and more water coverage, lowering roughness values, but it also tended to identify more forest cover and less cultivated land cover than our visual analysis, increasing some z_0 values.

β_0 and α also varied significantly between the scenarios. However, this was largely due to two practical matters: First, the independence of these variables of wind direction in the AERSURFACE case and secondly the use of monthly-varying moisture conditions in one test case and not another. Thus we have not presented those results here.

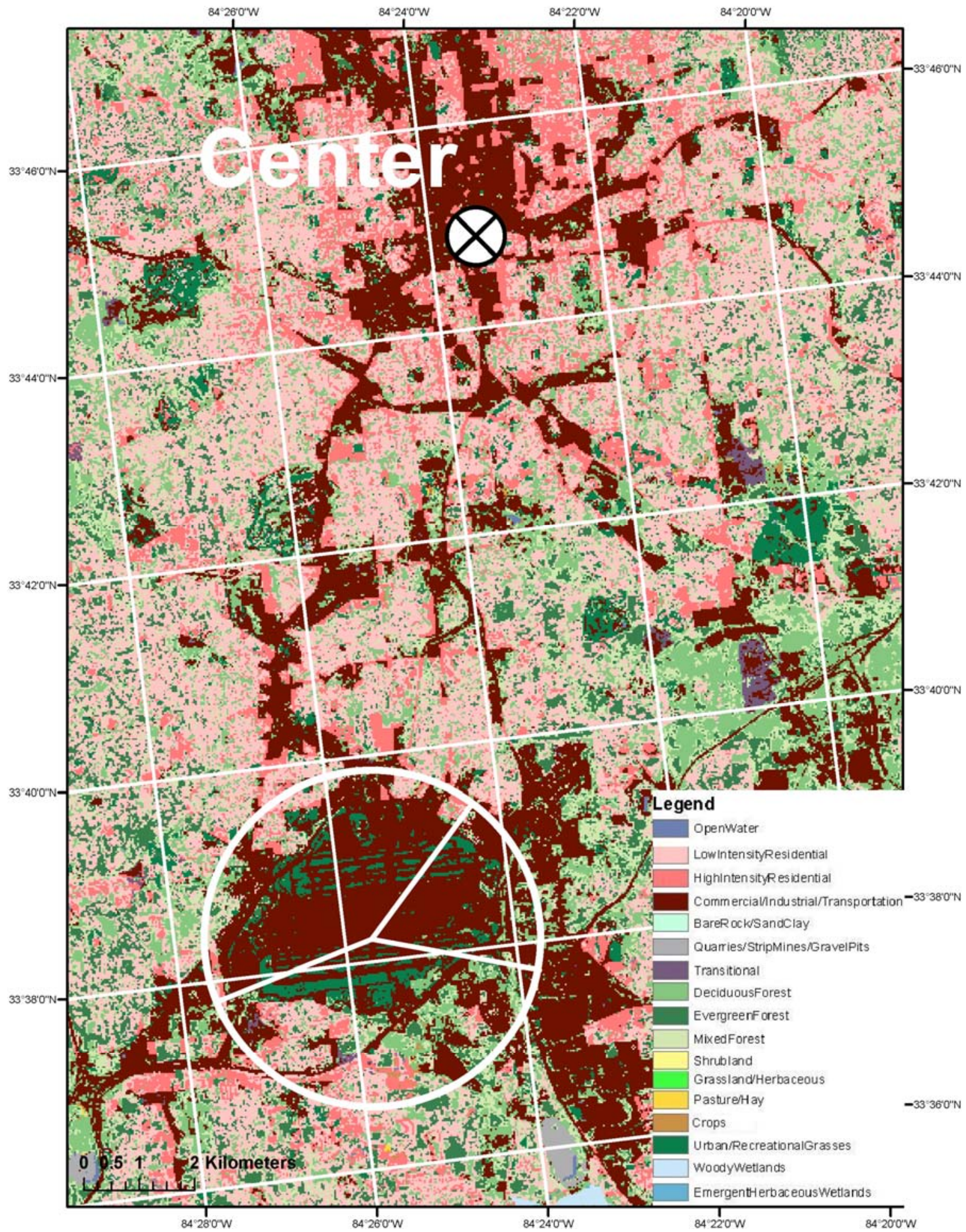


Figure E-1. Land-use and sectors around the Atlanta-area surface meteorological station (KATL). Sector borders are 43, 104, and 255 degrees from geographic North. Atlanta city center is labeled.

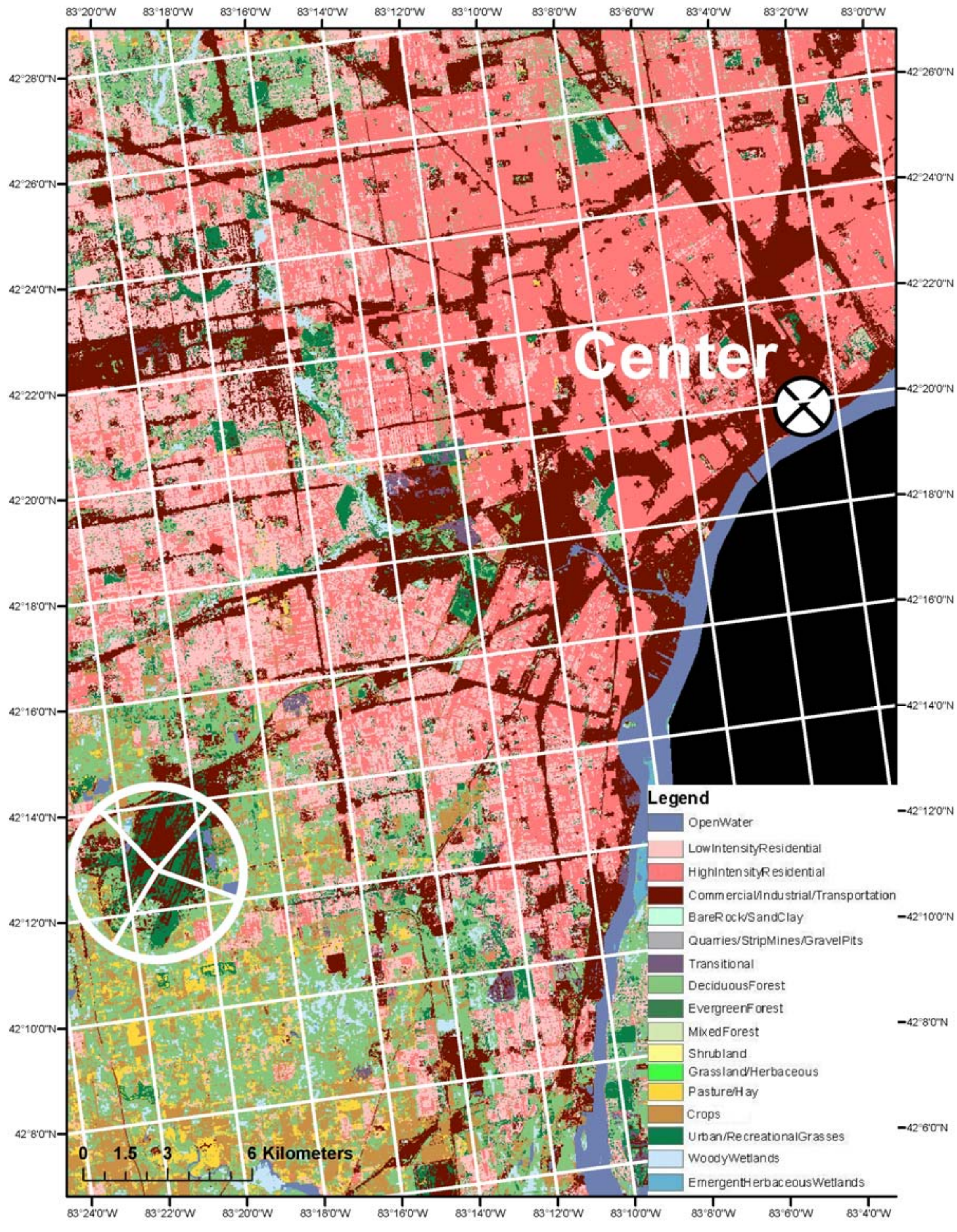


Figure E-2. Land-use and sectors around the Detroit-area surface meteorological station (KDTW). Sector borders are 49, 117, 217, and 322 degrees from geographic North. Detroit city center is labeled.

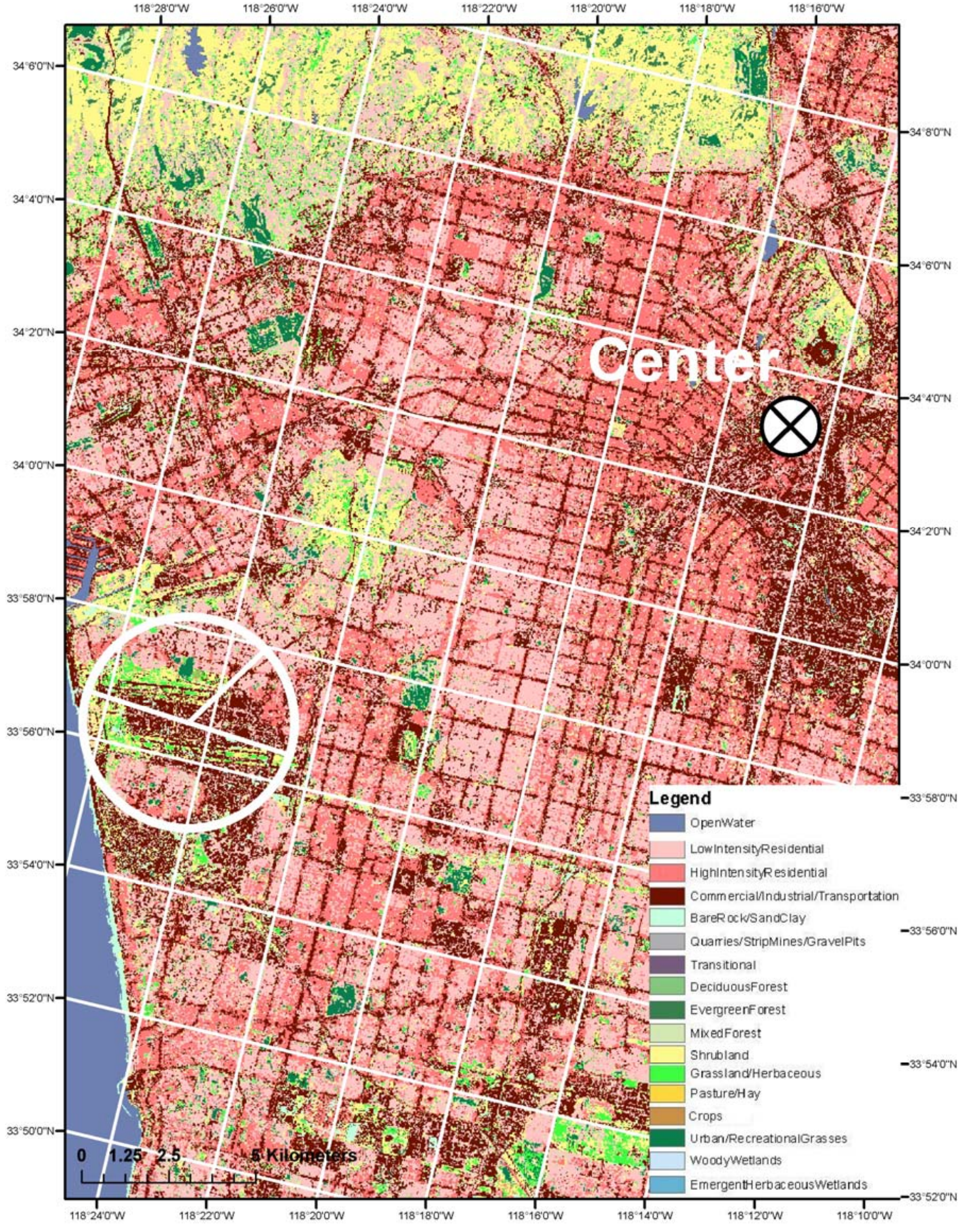


Figure E-3. Land-use and sectors around the Los Angeles-area surface meteorological station (KLAX). Sector borders are 34, 96, and 275 degrees from geographic North. Los Angeles city center is labeled.

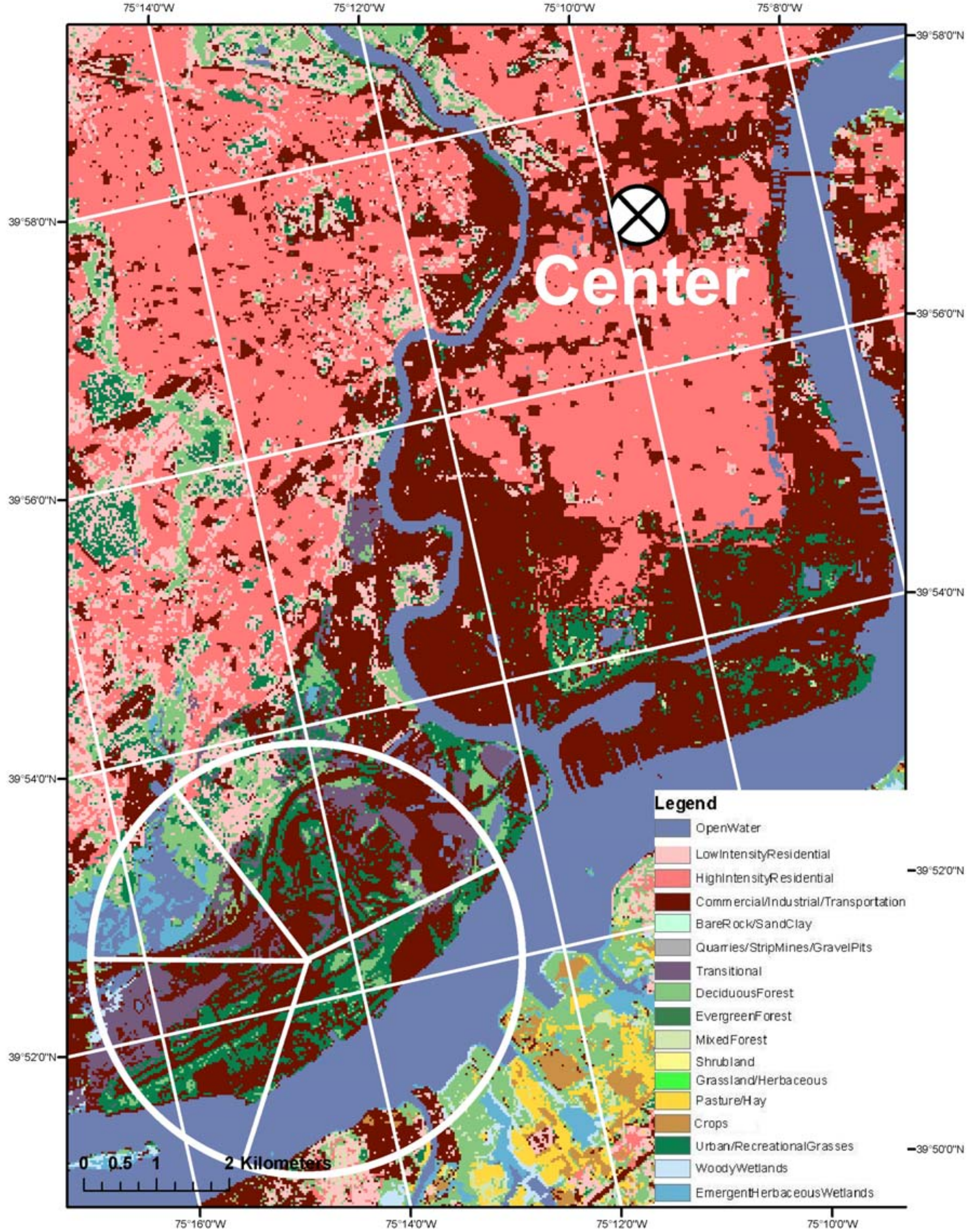


Figure E-4. Land-use and sectors around the Philadelphia-area surface meteorological station (KPHL). Sector borders are 80, 184, 262, and 312 degrees from geographic North. Philadelphia city center is labeled.

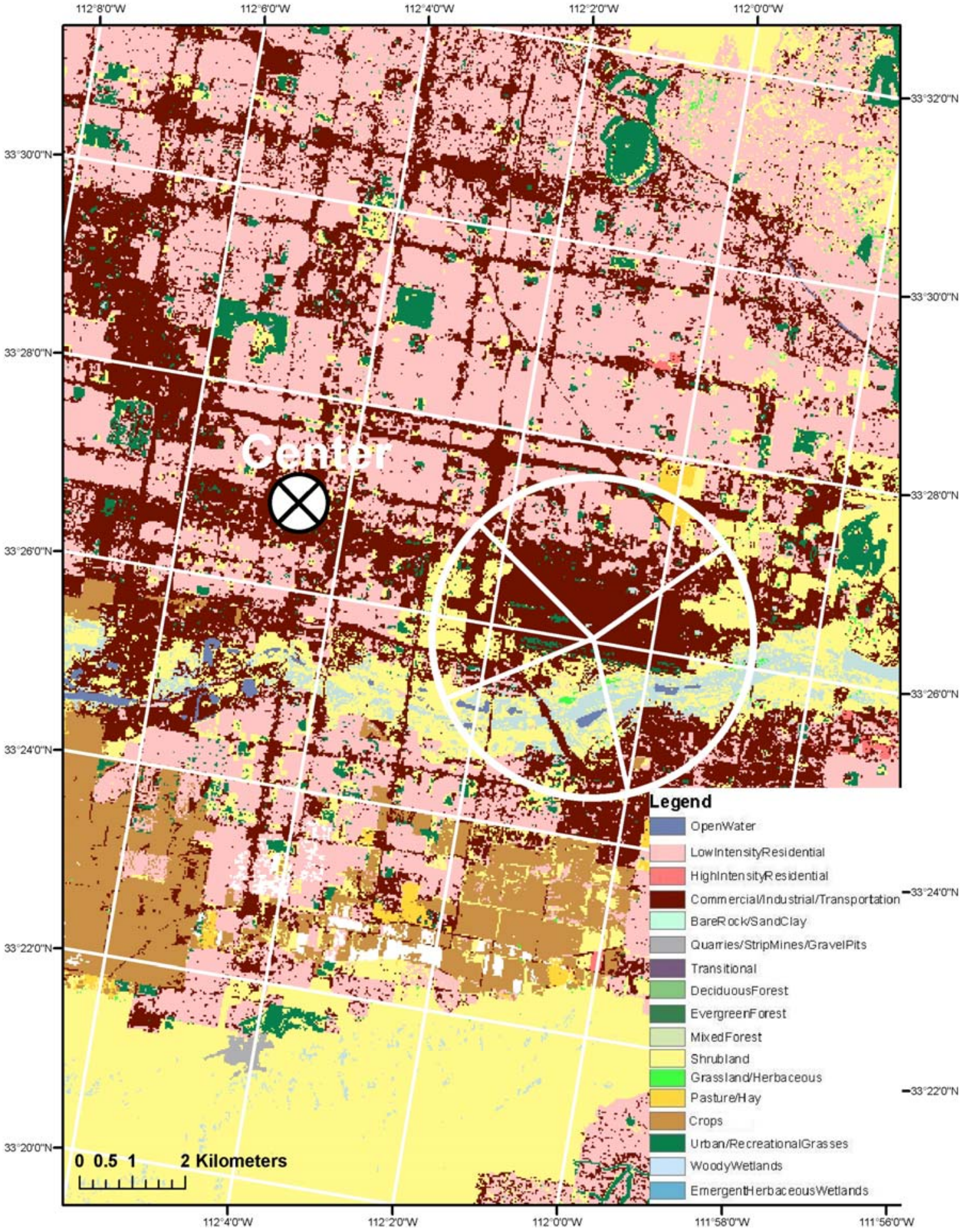


Figure E-5. Land-use and sectors around the Phoenix-area surface meteorological station (KPHX). Sector borders are 47, 153, 233, and 304 degrees from geographic North. Phoenix city center is labeled.

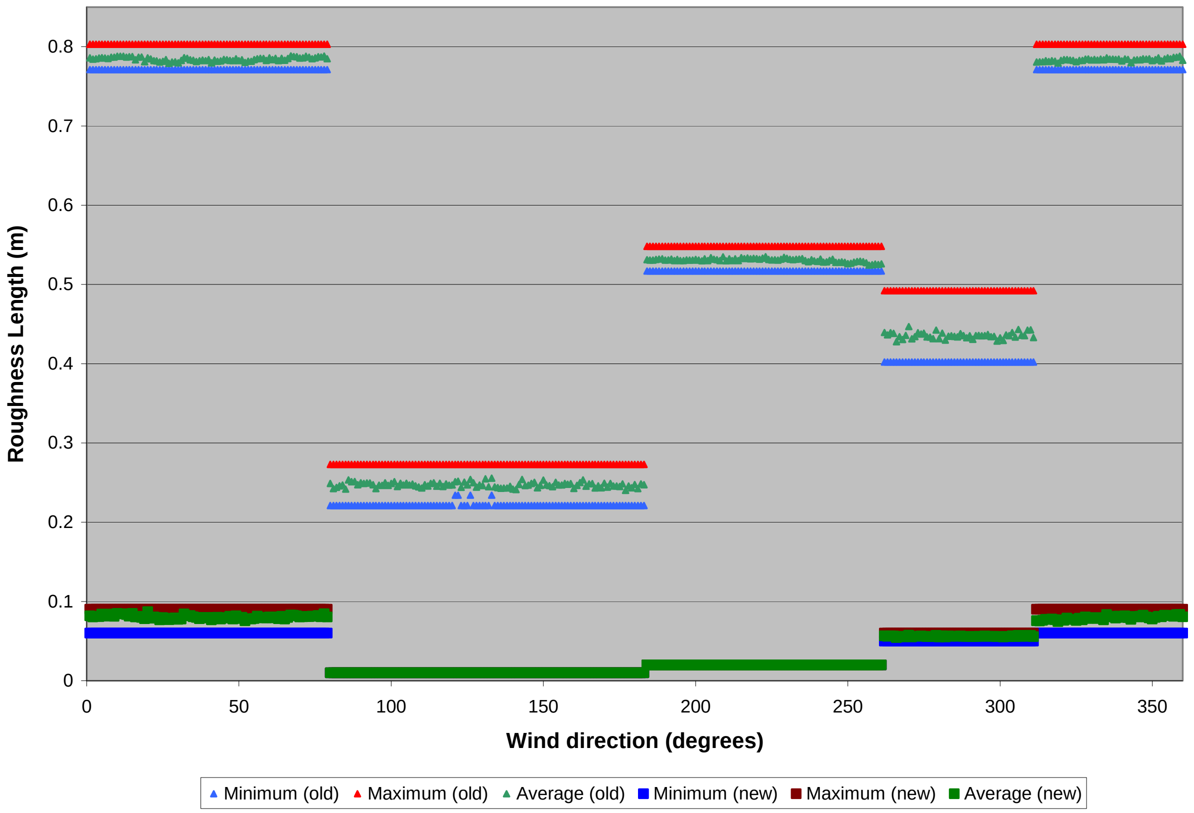


Figure E-6. Estimated z_0 Values for the Philadelphia Scenario Using Visual and AERSURFACE Land-Use Estimations.

Appendix F. Technical Memorandum on Longitudinal Diary Construction Approach

TECHNICAL MEMORANDUM

TO: Stephen Graham and John Langstaff, US EPA
FROM: Arlene Rosenbaum
DATE: February 29, 2008
SUBJECT: **The Cluster-Markov algorithm in APEX**

Background

The goals of population exposure assessment generally include an accurate estimate of both the average exposure concentration and the high end of the exposure distribution. One of the factors influencing the number of exposures at the high end of the concentration distribution is time-activity patterns that differ from the average, e.g., a disproportionate amount of time spent near roadways. Whether a model represents these exposure scenarios well depends on whether the treatment of activity pattern data accurately characterizes differences among individuals.

Human time-activity data for population exposure models are generally derived from demographic surveys of individuals' daily activities, the amount of time spent engaged in those activities, and the ME locations where the activities occur. Typical time-activity pattern data available for inhalation exposure modeling consist of a sequence of location/activity combinations spanning a 24-hour duration, with 1 to 3 records for any single individual. But modeling assessments of exposure to air pollutants typically require information on activity patterns over long periods of time, e.g., a full year. For example, even for pollutant health effects with short averaging times (e.g., ozone 8-hour average) it may be desirable to know the frequency of exceedances of a threshold concentration over a long period of time (e.g., the annual number of exceedances of an 8-hour average ozone concentration of 0.07 ppm for each simulated individual).

Long-term activity patterns can be estimated from daily ones by combining the daily records in various ways, and the method used for combining them will influence the variability of the long-term activity patterns across the simulated population. This in turn will influence the ability of the model to accurately represent either long-term average high-end exposures, or the number of individuals exposed multiple times to short-term high-end concentrations.

A common approach for constructing long-term activity patterns from short-term records is to re-select a daily activity pattern from the pool of data for each day, with the implicit assumption that there is no correlation between activities from day to day for the simulated individual. This approach tends to result in long-term activity patterns that are very similar across the simulated population. Thus, the resulting exposure estimates are likely to underestimate the variability across the population, and therefore, underestimate the high-end concentrations.

A contrasting approach is to select a single activity pattern (or a single pattern for each season and/or weekday-weekend) to represent a simulated individual's activities over the modeling period. This approach has the implicit assumption that an individual's day to day activities are perfectly correlated. This approach tends to result in long-term activity patterns that are very different across the simulated population, and therefore may over-estimate the variability across the population.

The Cluster-Markov algorithm

Recently, a new algorithm has been developed and incorporated into APEX that attempts to more realistically represent the day-to-day correlation of activities for individuals. The algorithms first use cluster analysis to divide the daily activity pattern records into groups that are similar, and then select a single daily record from each group. This limited number of daily patterns is then used to construct a long-term sequence for a simulated individual, based on empirically-derived transition probabilities. This approach is intermediate between the assumption of no day-to-day correlation (i.e., re-selection for each time period) and perfect correlation (i.e., selection of a single daily record to represent all days).

The steps in the algorithm are as follows.

- For each demographic group (age, gender, employment status), temperature range, and day-of-week combination, the associated time-activity records are partitioned into 3 groups using cluster analysis. The clustering criterion is a vector of 5 values: the time spent in each of 5 microenvironment categories (indoors – residence; indoors – other building; outdoors – near road; outdoors – away from road; in vehicle)..
- For each simulated individual, a single time-activity record is randomly selected from each cluster.
- Next the Markov process determines the probability of a given time-activity pattern occurring on a given day based on the time-activity pattern of the previous day and cluster-to-cluster transition probabilities. The cluster-to-cluster transition probabilities are estimated from the available multi-day time-activity records. (If insufficient multi-day time-activity records are available for a demographic group, season, day-of-week combination, then the cluster-to-cluster transition probabilities are estimated from the frequency of time-activity records in each cluster in the CHAD data base.)

Attachment 1 presents the Cluster-Markov algorithm in flow chart format.

Evalaution against observations (Rosenbaum and Cohen 2004)

The Cluster-Markov algorithm is also incorporated into the Hazardous Air Pollutant Exposure Model (HAPEM). The algorithm in HAPEM was tested using multi-day diary data sets collected as part of the Harvard Southern California Chronic Ozone Exposure Study (Xue et al. 2005, Geyh et al. 2000). In this study, 224 children in ages between 7 and 12 yr were followed for 1 year from June 1995 to May 1996, for 6 consecutive days each month. The subjects resided in two separate areas of San Bernardino County: urban

Upland CA, and the small mountain towns of Lake Arrowhead, Crestline, and Running Springs, CA.

For purposes of clustering the activity pattern records were characterized according to time spent in each of 5 aggregate microenvironments: indoors-home, indoors-school, indoors-other, outdoors, and in-transit. For purposes of defining diary pools and for clustering and calculating transition probabilities the activity pattern records were divided by day type (i.e., weekday, weekend), season (i.e., summer or ozone season, non-summer or non-ozone season), age (7-10 and 11-12), and gender.

Week-long sequences (Wednesday through Tuesday) for each of 100 people in each age/gender group for each season were simulated.

To evaluate the algorithm the following statistics were calculated for the predicted multi-day activity patterns and compared them with the actual multi-day diary data.

- For each age/gender group for each season, the average time in each microenvironment
- For each simulated person-week and microenvironment, the average of the within-person variance across all simulated persons (The within-person variance was defined as the variance of the total time per day spent in the microenvironment across the week.)
- For each simulated person-week the variance across persons of the mean time spent in each microenvironment.

In each case the predicted statistic for the stratum was compared to the statistic for the corresponding stratum in the actual diary data. The mean normalized bias for the statistic, which is a common performance measure used in dispersion model performance and was also calculated as follows.

$$NBIAS = \frac{100}{N}\sum_{1}^{N}\frac{(predicted - observed)}{observed}$$

The predicted time-in-microenvironment averages matched well with the observed values. For combinations of microenvironment/age/gender/season the normalized bias ranges from –35% to +41%. Sixty percent of the predicted averages have bias between –9% and +9%, and the mean bias across any microenvironment ranges from -9% to +4%. Fourteen predictions have positive bias and 23 have negative bias.

For the variance across persons for the average time spent in each microenvironment, the bias ranged from –40% to +120% for any microenvironment/age/gender/season. Sixty-five percent of the predicted variances had bias between –22% and +24%. The mean normalized bias across any microenvironment ranged from –10% to +28%. Eighteen predictions had positive bias and 20 had negative bias.

For the within-person variance for time spent in each microenvironment, the bias ranged from –47% to +150% for any microenvironment/age/gender/season. Seventy percent of the predicted variances had bias between –25% and +30%. The mean normalized bias across any microenvironment ranged from –11% to +47%. Twenty-eight predictions had positive bias and 12 had negative bias, suggesting some tendency for overprediction of this variance measure.

The overall conclusion was that the proposed algorithm appeared to be able to replicate the observed data reasonably well. Although some discrepancies were rather large for some of the “variance across persons” and “within-person variance” subsets, about two-thirds of the predictions for each case were within 30% of the observed value.

A detailed description of the evaluation is presented in Attachment 2.

Comparison with other algorithms (US EPA 2007)

As part of the application of APEX in support of US EPA’s recent review of the ozone NAAQS several sensitivity analyses were conducted. One of these was to make parallel simulations using each of the three algorithms for constructing multi-day time-activity sequences that are incorporated into APEX.

Table 1 presents the results for the number of persons in Atlanta population groups with moderate exertion exposed to 8-hour average concentrations exceeding 0.07 ppm. The results show that the predictions made with alternative algorithm Cluster-Markov algorithm are substantially different from those made with simple re-sampling or with the Diversity-Autocorrelation algorithm (“base case”). Note that for the cluster algorithm approximately 30% of the individuals with 1 or more exposure have 3 or more exposures. The corresponding values for the other algorithms range from about 13% to 21%.

Table 2 presents the results for the mean and standard deviation of number of days/person with 8-hour average exposures exceeding 0.07 ppm with moderate or greater exertion. The results show that although the mean for the Cluster-Markov algorithm is very similar to the other approaches, the standard deviation is substantially higher, i.e., the Cluster-Markov algorithm results in substantially higher inter-individual variability.

Table 1. Sensitivity to longitudinal diary algorithm: 2002 simulated counts of Atlanta general population and children (ages 5-18) with any or three or more 8-hour ozone exposures above 0.07 ppm concomitant with moderate or greater exertion (after US EPA 2007).

	One or more exposures			Three or more exposures		
Population Group	**Simple re-sampling**	**Diversity-Autocorrelation**	**Cluster-Markov**	**Simple re-sampling**	**Diversity-Autocorrelation**	**Cluster-Markov**
General Population	979,533	939,663 (-4%)	668,004 (-32%)	124,687	144,470 (+16%)	188,509 (+51%)
Children (5-18)	411,429	389,372 (-5%)	295,004 (-28%)	71,174	83,377 (+17%)	94,216 (+32%)

Table 2. Sensitivity to longitudinal diary algorithm: 2002 days per person with 8-hour ozone exposures above 0.07 ppm concomitant with moderate or greater exertion for Atlanta general population and children (ages 5-18) (after US EPA 2007).

	Mean Days/Person			Standard Deviation		
Population Group	**Simple re-sampling**	**Base case**	**Cluster**	**Simple re-sampling**	**Base case**	**Cluster**
General Population	0.332	0.335 (+1%)	0.342 (+3%)	0.757	0.802 (+6%)	1.197 (+58%)
Children (5-18)	0.746	0.755 (+1%)	0.758 (+2%)	1.077	1.171 (+9%)	1.652 (+53%)

References

Geyh, AS, Xue, J, Ozkaynak, H, and Spengler, JD. 2000. The Harvard Southern California chronic ozone exposure study: Assessing ozone exposure of grade-school-age children in two Southern California communities. *Environ Health Persp.* 108:265-270.

Rosenbaum, AS, and Cohen JP. 2004. *Evaluation of a multi-day activity pattern algorithm for creating longitudinal activity patterns.* Memorandum prepared for Ted Palma. USEPA OAQPS by ICF International.

US EPA . 2007. *Ozone Population Exposure Analysis for Selected Urban Areas*. EPA-452/R-07-010 http://www.epa.gov/ttn/naaqs/standards/ozone/data/2007-01_o3_exposure_tsd.pdf.

Xue J, Liu SV, Ozkaynak H, Spengler J. 2005. Parameter evaluation and model validation of ozone exposure assessment using Harvard Southern California Chronic Ozone Exposure Study Data. *J. Air & Waste Manage. Assoc.* **55**:1508–1515.

ATTACHMENT 1

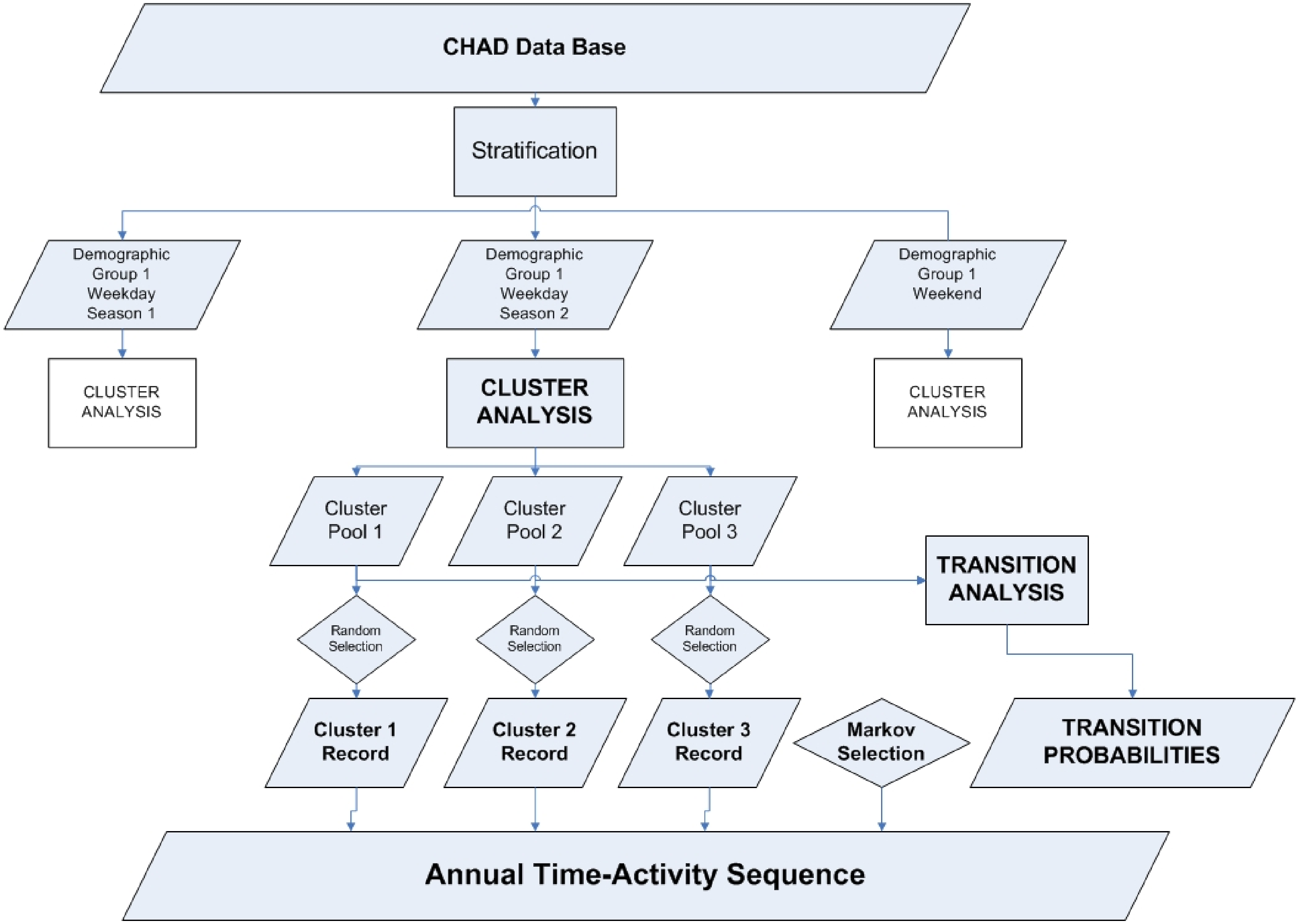


Figure A1-1. Flow chart of Cluster-Markov algorithm.

ATTACHMENT 2

TECHNICAL MEMORANDUM

TO: Ted Palma, US EPA
FROM: Arlene Rosenbaum and Jonathan Cohen, ICF Consulting
DATE: November 4, 2004
SUBJECT: Evaluation of a multi-day activity pattern algorithm for creating longitudinal activity patterns.

BACKGROUND

In previous work ICF reviewed the HAPEM4 modeling approach for developing annual average activity patterns from the CHAD database and recommended an approach to improve the model's pattern selection process to better represent the variability among individuals. This section summarizes the recommended approach. (For details see the memorandum of July 23, 2002 from ICF Consulting to Ted Palma.)

Using cluster analysis, first the CHAD daily activity patterns are grouped into either two or three categories of similar patterns for each of the 30 combinations of day type (summer weekday, non-summer weekday, and weekend) and demographic group (males or females; age groups: 0-4, 5-11, 12-17, 18-64, 65+). Next, for each combination of day type and demographic group, category-to-category transition probabilities are defined by the relative frequencies of each second-day category associated with each given first-day category, where the same individual was observed for two consecutive days. (Consecutive day activity pattern records for a single individual constitute a small subset of the CHAD data.)

To implement the proposed algorithm, for each day type and demographic group, one daily activity pattern per category is randomly selected from the corresponding CHAD data to represent that category. That is, if there are 3 cluster categories for each of 3 day types, 9 unique activity patterns are selected to be averaged together to create an annual average activity pattern to represent an individual in a given demographic group and census tract.

The weighting for each of the 9 activity patterns used in the averaging process is determined by the product of two factors. The first is the relative frequency of its day type, i.e., 0.18 for summer weekdays, 0.54 for non-summer weekdays, and 0.28 for weekends.
The second factor in the weighting for the selected activity pattern is determined by simulating a sequence of category-types as a one-stage Markov chain process using the transition probabilities. The category for the first day is selected according to the relative frequencies of each category. The category for the second day is selected according to the category-to-category transition probabilities for the category selected for the first day. The category for the third day is selected according to the transition probabilities for the category selected for the second day. This is repeated for all days in the day type (65 for summer weekdays, 195 for non-summer weekdays, 104 for weekends), producing a sequence of daily

categories. The relative frequency of the category-type in the sequence associated with the selected activity pattern is the second factor in the weighting.

PROPOSED ALGORITHM STEPS

The proposed algorithm is summarized in Figure 1. Each step is explained in this section.

Data Preparation

Step 1: Each daily activity pattern in the CHAD data base is summarized by the total minutes in each of five micro-environments: indoors – residence; indoors – other building; outdoors – near road; outdoors – away from road; in vehicle. These five numbers are assumed to represent the most important features of the activity pattern for their exposure impact.

Step 2: All CHAD activity patterns for a given day-type and demographic group are subjected to cluster analysis, resulting in 2 or 3 cluster categories. Each daily activity pattern is tagged with a cluster category.

Step 3: For each day-type and demographic group, the relative frequency of each day-type in the CHAD data base is determined.

Step 4: All CHAD activity patterns for a given day-type and demographic group that are consecutive days for a single individual, are analyzed to determine the category-to-category transition frequencies in the CHAD data base. These transition frequencies are used to calculate category-to-category transition probabilities.

For example, if there are 2 categories, A and B, then

P_{AA} = the probability that a type A pattern is followed by a type A pattern,
P_{AB} = the probability that a type A pattern is followed by a type B pattern ($P_{AB} = 1 - P_{AA}$),
P_{BB} = the probability that a type B pattern is followed by a type B pattern, and
P_{BA} = the probability that a type B pattern is followed by a type A pattern ($P_{BA} = 1 - P_{BB}$).

Activity Pattern Selection

For each day-type and demographic group in each census tract

Step 5: One activity pattern is randomly selected from each cluster category group (i.e., 2 to 3 activity patterns)

Creating Weights for Day-type Averaging

For each day-type and demographic group in each census tract

Step 6: A cluster category is selected for the first day of the day-type sequence, according to the relative frequency of the cluster category days in the CHAD data set.

Step 7: A cluster category is selected for each subsequent day in the day-type sequence day by day using the category-to-category transition probabilities.

Step 8: The relative frequency of each cluster category in the day-type sequence is determined.

Step 9: The activity patterns selected for each cluster category (Step 5) are averaged together using the cluster category frequencies (Step 8) as weights, to create a day-type average activity pattern.

Creating Annual Average Activity Patterns

For each demographic group in each census tract

Step 10: The day-type average activity patterns are averaged together using the relative frequency of day-types as weights, to create an annual average activity pattern.

Creating Replicates

For each demographic group in each census tract

Step 11: Steps 5 through 10 are repeated 29 times to create 30 annual average activity patterns.

EVALUATING THE ALGORITHM

The purpose of this study is to evaluate how well the proposed one-stage Markov chain algorithm can reproduce observed multi-day activity patterns with respect to demographic group means and inter-individual variability, while using one-day selection.

In order to accomplish this we propose to apply the algorithm to observed multi-day activity patterns provided by the WAM, and compare the means and variances of the predicted multi-day patterns with the observed patterns.

Current APEX Algorithm

Because the algorithm is being considered for incorporation into APEX, we would like the evaluation to be consistent with the approach taken in APEX for selection of activity patterns for creating multi-day sequences. The APEX approach for creating multi-day activity sequences is as follows.

Step1: A profile for a simulated individual is generated by selection of gender, race (not implemented?), age group, and home sector from a given set of distributions consistent with the population of the study area.

Step 2: A specific age within the age group is selected from a uniform distribution.

Step 3: The employment status is simulated as a function of the age.

Step 4: For each simulated day, the user defines an initial pool of possible diary days based on a user-specified function of the day type (e.g., weekday/weekend) and temperature.

Step 5: The pool is further restricted to match the target gender and employment status exactly and the age within 2A years for some parameter A. The diary days within the pool are assigned a weight of 1 if the age is within A years of the target age and a weight of w (user-defined parameter) if the age difference is between A and 2A years. For each simulated day, the probability of selecting a given diary day is equal to the age weight divided by the total of the age weights for all diary days in the pool for that day.

Approach to Incorporation of Day-to-Day Dependence into APEX Algorithm

If we were going to incorporate day-to-day dependence of activity patterns into the APEX model, we would propose preparing the data with cluster analysis and transition probabilities as described in Steps 1-4 for the proposed HAPEM 5 algorithm, with the following modifications.

- For Step 2 the activity patterns would be divided into groups based on day-type (weekday, weekend), temperature, gender, employment status, and age, with cluster analysis applied to each group. However, because the day-to-day transitions in the APEX activity selection algorithm can cross temperature bins, we would propose to use broad temperature bins for the clustering and transition probability calculations so that the cluster definitions would be fairly uniform across temperature bins. Thus we would probably define the bins according to season (e.g., summer, non-summer).

- In contrast to HAPEM, the sequence of activity patterns may be important in APEX. Therefore, for Step 4 transition probabilities would be specified for transitions between days with the same day-type and season, as in HAPEM, and also between days with different day-types and/or seasons. For example, transition probabilities would be specified for transitions between summer weekdays of each category and summer weekends of each category.

Another issue for dividing the CHAD activity records for the purposes of clustering and calculating transition probabilities is that the diary pools specified for the APEX activity selection algorithm use varying and overlapping age ranges. One way to address this problem would be to simply not include consideration of age in the clustering process, under the assumption that cluster categories are similar across age groups, even if the frequency of each cluster category varies by age group. This assumption could be tested by examination of the cluster categories stratified by age group that were developed for HAPEM5. If the assumption is found to be valid, then the cluster categories could be pre-determined for input to APEX, while the transition probabilities could be calculated within APEX during the simulation for each age range specified for dairy pools.

If the assumption is found to be invalid, then an alternative approach could be implemented that would create overlapping age groups for purposes of clustering as follows. APEX age group ranges and age window percentages would be constrained to some maximum values. Then a set of overlapping age ranges that would be at least as large as the largest possible dairy pool age ranges would be defined for the purposes of cluster analysis and transition probability calculation. The resulting sets of cluster categories and transition probabilities would be pre-determined for input into APEX and the appropriate set used by APEX for each diary pool used during the simulation.

The actual activity pattern sequence selection would be implemented as follows. The activity pattern for first day in the year would be selected exactly as is currently done in APEX, as described above. For the selecting the second day’s activity pattern, each age weight would be multiplied by the transition probability P_{AB} where A is the cluster for the first day’s activity pattern and B is the cluster for a given activity pattern in the available pool of diary days for day 2. (Note that day 2 may be a different day-type and/or season than day 1.) The probability of selecting a given diary day on day 2 is equal to the age weight times P_{AB} divided by the total of the products of age weight and P_{AB} for all diary days in the pool for day 2. Similarly, for the transitions from day 2 to day 3, day 3 to day 4, etc.

Testing the Approach with the Multi-day Data set

We tested this approach using the available multi-day data set. For purposes of clustering we characterized the activity pattern records according to time spent in each of 5 microenvironments: indoors-home, indoors-school, indoors-other, outdoors (aggregate of the 3 outdoor microenvironments), and in-transit.

For purposes of defining diary pools and for clustering and calculating transition probabilities we divided the activity pattern records by day type (i.e., weekday, weekend), season (i.e., summer or ozone season, non-summer or non-ozone season), age (6-10 and 11-12), and gender. Since all the subjects are 6-12 years of age and all are presumably unemployed, we need not account for differences in employment status. For each day type, season, age, and gender, we found that the activity patterns appeared to group in three clusters.

In this case, we simulated week-long sequences (Wednesday through Tuesday) for each of 100 people in each age/gender group for each season, using the transition probabilities. To evaluate the algorithm we calculated the following statistics for the predicted multi-day activity patterns for comparison with the actual multi-day diary data.

- For each age/gender group for each season, the average time in each microenvironment

- For each age/gender group, season, and microenvironment, the average of the within-person variance across all simulated persons (We defined the within-person variance as the variance of the total time per day spent in the microenvironment across the week.)

- For each age/gender group, season, and microenvironment, the variance across persons of the mean time spent in that microenvironment

In each case we compared the predicted statistic for the stratum to the statistic for the corresponding stratum in the actual diary data.[1]

We also calculated the mean normalized bias for the statistic, which is a common performance measure used in dispersion model performance and which is calculated as follows.

$$NBIAS = \frac{100}{N}\sum_{1}^{N}\frac{(predicted - observed)}{observed} \quad \%$$

RESULTS

Comparisons of simulated and observed data for time in each of the 5 microenvironments are presented in Tables 1 – 3 and Figures 2-5.

Average Time in Microenvironment

Table 1 and Figure 2 show the comparisons for the average time spent in each of the 5 microenvironments for each age/gender group and season. Figure 3 shows the comparison for all the microenvironments except indoor, home in order to highlight the lower values.

---

[1] For the diary data, because the number of days per person varies, the average of the within-person variances was calculated as a weighted average, where the weight is the degrees of freedom, i.e., one less than the number of days simulated. Similarly, the variance across persons of the mean time was appropriately adjusted for the different degrees of freedom using analysis of variance.

Table 1 and the figures show that the predicted time-in-microenvironment averages match well with the observed values. For combinations of microenvironment/age/gender/season the normalized bias ranges from –35% to +41%. Sixty percent of the predicted averages have bias between –9% and +9%, and the mean bias across any microenvironment ranges from -9% to +4%. Fourteen predictions have positive bias and 23 have negative bias. A Wilcoxon signed rank test that the median bias across the 40 combinations = 0 % was not significant (p-value = 0.40) supporting the conclusion of no overall bias.

Variance Across Persons

Table 2 and Figure 4 show the comparisons for the variance across persons for the average time spent in each microenvironment. In this case the bias ranges from –40% to +120% for any microenvironment/age/gender/season. Sixty-five percent of the predicted variances have bias between –22% and +24%. The mean normalized bias across any microenvironment ranges from –10% to +28%. Eighteen predictions have positive bias and 20 have negative bias. Figure 4 suggests a reasonably good match of predicted to observed variance in spite of 2 or 3 outliers. A Wilcoxon signed rank test that the median bias across the 40 combinations = 0 % was not significant (p-value = 0.93) supporting the conclusion of no overall bias.

Within-Person Variance for Persons

Table 3 and Figure 5 show the comparisons for the within-person variance for time spent in each microenvironment. In this case the bias ranges from –47% to +150% for any microenvironment/age/gender/season. Seventy percent of the predicted variances have bias between –25% and +30%. The mean normalized bias across any microenvironment ranges from –11% to +47%. Twenty-eight predictions have positive bias and 12 have negative bias, suggesting some tendency for overprediction of this variance measure. And indeed a Wilcoxon signed rank test that the median bias across the 40 combinations = 0 % was very significant (p-value = 0.01) showing that the within-person variance was significantly overpredicted. Still, Figure 4 suggests a reasonably good match of predicted to observed variance in most cases, with a few overpredicting outliers at the higher end of the distribution. So although the positive bias is significant in a statistical sense (i.e., the variance is more likely to be overpredicted than underpredicted), it is not clear whether the bias is large enough to be important.

CONCLUSIONS

The proposed algorithm appears to be able to replicate the observed data reasonably well, although the within-person variance is somewhat overpredicted.

It would be informative to compare this algorithm with the earlier alternative approaches in order to gain perspective on the degree of improvement, if any, afforded by this approach.

Two earlier approaches were:

1. Select a single activity pattern for each day-type/season combination from the appropriate set, and use that pattern for every day in the multi-day sequence that corresponds to that day-type and season.
2. Re-select an activity pattern for each day in the multi-day sequence from the appropriate set for the corresponding day-type and season.

Goodness-of-fit statistics could be developed to compare the three approaches and find which model best fits the data for a given stratum.

Table 1. Average time spent in each microenvironment: comparison of predicted and observed.

Microenvironment	Demographic Group	Season	Observed (hours/day)	Predicted (hours/day)	Normalized Bias
Indoor, home	Girls, 6-10	Summer	15.5	16.5	6%
		Not Summer	15.8	15.5	-2%
	Boys, 6-10	Summer	15.7	15.2	-3%
		Not Summer	15.8	16.4	4%
	Girls, 11-12	Summer	16.2	15.3	-5%
		Not Summer	16.5	16.5	0%
	Boys, 11-12	Summer	16.0	15.6	-3%
		Not Summer	16.2	16.1	-1%
	MEAN				-1%
Indoor, school	Girls, 6-10	Summer	0.7	0.7	-9%
		Not Summer	2.3	2.5	7%
	Boys, 6-10	Summer	0.8	0.5	-34%
		Not Summer	2.2	2.2	0%
	Girls, 11-12	Summer	0.7	0.7	6%
		Not Summer	2.1	2.4	13%
	Boys, 11-12	Summer	0.6	0.9	38%
		Not Summer	2.4	2.7	11%
	MEAN				4%
Indoor, other	Girls, 6-10	Summer	2.9	2.4	-14%
		Not Summer	2.4	2.7	13%
	Boys, 6-10	Summer	2.2	2.7	21%
		Not Summer	1.9	1.8	-3%
	Girls, 11-12	Summer	2.2	1.6	-25%
		Not Summer	2.2	2.1	-2%
	Boys, 11-12	Summer	2.3	2.2	-5%

		Not Summer	1.9	2.0	4%
	MEAN				-2%
Outdoors	Girls, 6-10	Summer	3.7	3.5	-6%
		Not Summer	2.5	2.5	0%
	Boys, 6-10	Summer	4.1	4.3	4%
		Not Summer	3.1	2.7	-12%
	Girls, 11-12	Summer	3.7	5.2	41%
		Not Summer	2.3	2.1	-5%
	Boys, 11-12	Summer	3.9	4.3	9%
		Not Summer	2.6	2.4	-7%
	MEAN				3%
In-vehicle	Girls, 6-10	Summer	1.1	0.9	-20%
		Not Summer	1.0	0.9	-13%
	Boys, 6-10	Summer	1.1	1.3	13%
		Not Summer	1.0	0.9	-16%
	Girls, 11-12	Summer	1.2	1.1	-12%
		Not Summer	0.9	0.8	-15%
	Boys, 11-12	Summer	1.1	1.0	-5%
		Not Summer	0.9	0.8	-7%
	MEAN				-9%

Table 2. Variance across persons for time spent in each microenvironment: comparison of predicted and observed.

Microenvironment	Demographic Group	Season	Observed (hours/day)^2	Predicted (hours/day)^2	Normalized Bias
Indoor, home	Girls, 6-10	Summer	70	42	-40%
		Not Summer	67	60	-9%
	Boys, 6-10	Summer	54	49	-9%
		Not Summer	35	30	-12%
	Girls, 11-12	Summer	56	47	-17%
		Not Summer	42	38	-10%
	Boys, 11-12	Summer	57	63	12%
		Not Summer	39	42	8%
	MEAN				-10%
Indoor, school	Girls, 6-10	Summer	6.0	5.2	-13%
		Not Summer	9.5	5.9	-38%
	Boys, 6-10	Summer	5.6	3.8	-32%
		Not Summer	5.3	8.2	53%
	Girls, 11-12	Summer	4.9	5.5	11%
		Not Summer	5.4	5.3	-1%
	Boys, 11-12	Summer	5.6	6.0	6%
		Not Summer	9.2	11	23%
	MEAN				1%
Indoor, other	Girls, 6-10	Summer	46	32	-30%
		Not Summer	44	46.	6%
	Boys, 6-10	Summer	34	33	-4%
		Not Summer	23	16	-27%
	Girls, 11-12	Summer	21	18	-15%
		Not Summer	28	22	-22%

	Boys, 11-12	Summer	33	31	-6%
		Not Summer	30	30	0%
	MEAN				-12%
Outdoors	Girls, 6-10	Summer	17	23	37%
		Not Summer	9.3	6.8	-27%
	Boys, 6-10	Summer	17	18	3%
		Not Summer	8.3	7.6	-8%
	Girls, 11-12	Summer	22	22	0%
		Not Summer	9.0	9.1	1%
	Boys, 11-12	Summer	13	29	120%
		Not Summer	10	11	8%
	MEAN				17%
In-vehicle	Girls, 6-10	Summer	1.9	2.3	24%
		Not Summer	1.8	1.6	-11%
	Boys, 6-10	Summer	2.5	4.7	93%
		Not Summer	1.5	1.6	9%
	Girls, 11-12	Summer	3.5	4.7	34%
		Not Summer	2.8	2.0	-28%
	Boys, 11-12	Summer	3.2	5.4	69%
		Not Summer	1.3	1.7	35%
	MEAN				28%

Table 3. Average within person variance for time spent in each microenvironment: comparison of predicted and observed.

Microenvironment	Demographic Group	Season	Observed (hours/day)^2	Predicted (hours/day)^2	Normalized Bias
Indoor, home	Girls, 6-10	Summer	20	29	49%
		Not Summer	18	23	25%
	Boys, 6-10	Summer	17	30	75%
		Not Summer	15	24	64%
	Girls, 11-12	Summer	22	42	93%
		Not Summer	22	25	13%
	Boys, 11-12	Summer	21	24	16%
		Not Summer	17	24	38%
	MEAN				47%
Indoor, school	Girls, 6-10	Summer	2.3	2.4	5%
		Not Summer	7.3	6.4	-12%
	Boys, 6-10	Summer	2.0	1.5	-25%
		Not Summer	6.7	5.8	-14%
	Girls, 11-12	Summer	1.7	2.1	29%
		Not Summer	7.4	7.6	3%
	Boys, 11-12	Summer	1.4	2.9	101%
		Not Summer	7.3	7.8	6%
	MEAN				12%
Indoor, other	Girls, 6-10	Summer	14	14	-4%
		Not Summer	14	18	30%
	Boys, 6-10	Summer	12	17	42%
		Not Summer	10	13	26%
	Girls, 11-12	Summer	10	10	1%
		Not Summer	14	15	7%

	Boys, 11-12	Summer	11	14	26%
		Not Summer	12	13	7%
	MEAN				17%
Outdoors	Girls, 6-10	Summer	8.4	9.5	13%
		Not Summer	3.4	3.2	-3%
	Boys, 8-10	Summer	6.7	9.5	42%
		Not Summer	3.4	4.4	28%
	Girls, 11-12	Summer	10	25	150%
		Not Summer	4.0	4.5	11%
	Boys, 11-12	Summer	9.2	7.4	-20%
		Not Summer	4.3	3.7	-15%
	MEAN				26%
In-vehicle	Girls, 6-10	Summer	1.0	0.90	-13%
		Not Summer	0.90	0.48	-47%
	Boys, 6-10	Summer	1.1	1.4	31%
		Not Summer	0.81	0.71	-12%
	Girls, 11-12	Summer	1.3	1.3	4%
		Not Summer	1.3	1.1	-16%
	Boys, 11-12	Summer	2.4	1.6	-34%
		Not Summer	0.85	0.85	1%
	MEAN				-11%

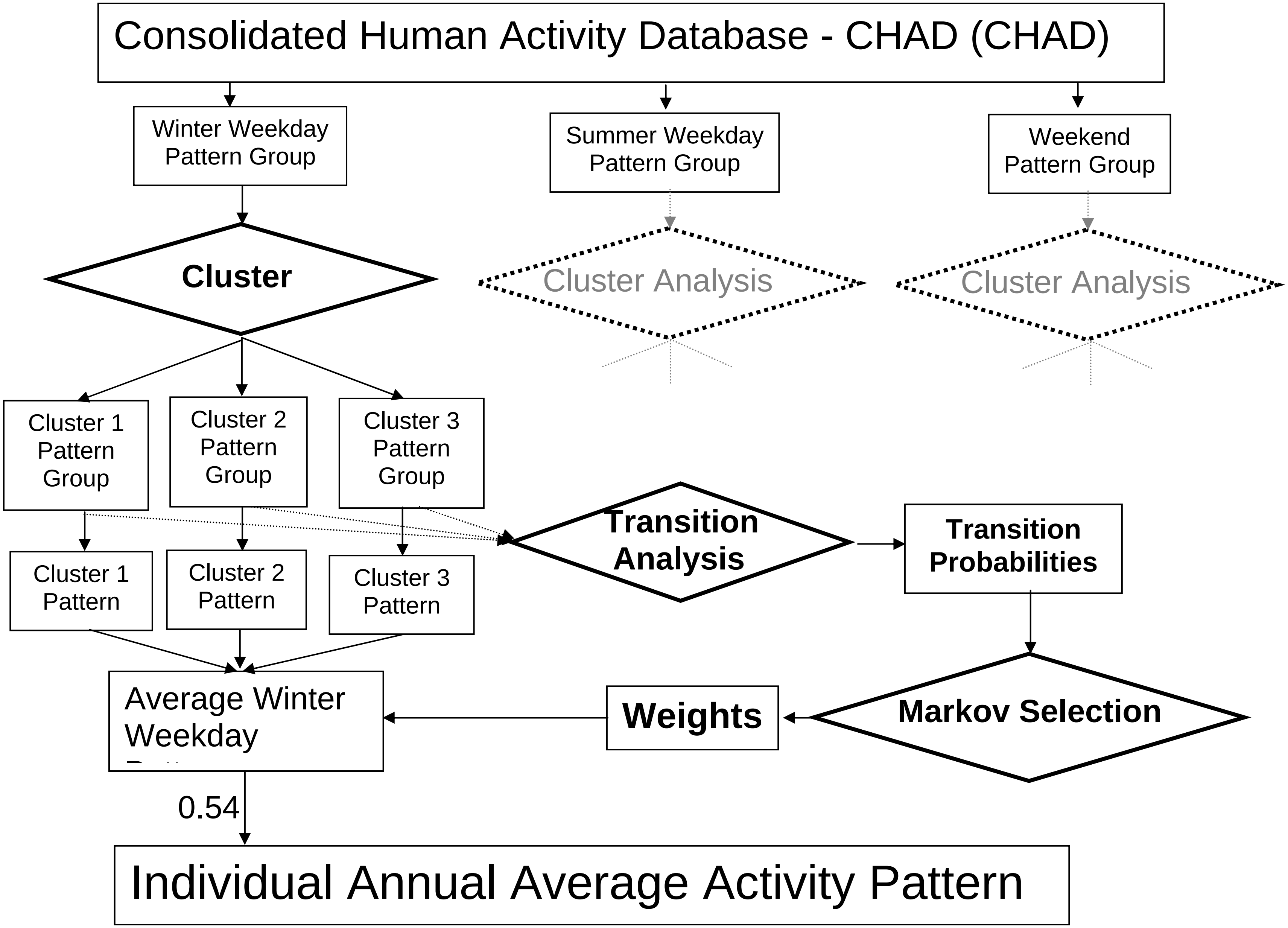


Figure 1. Flow diagram of proposed algorithm for creating annual average activity patterns for HAPEM5.

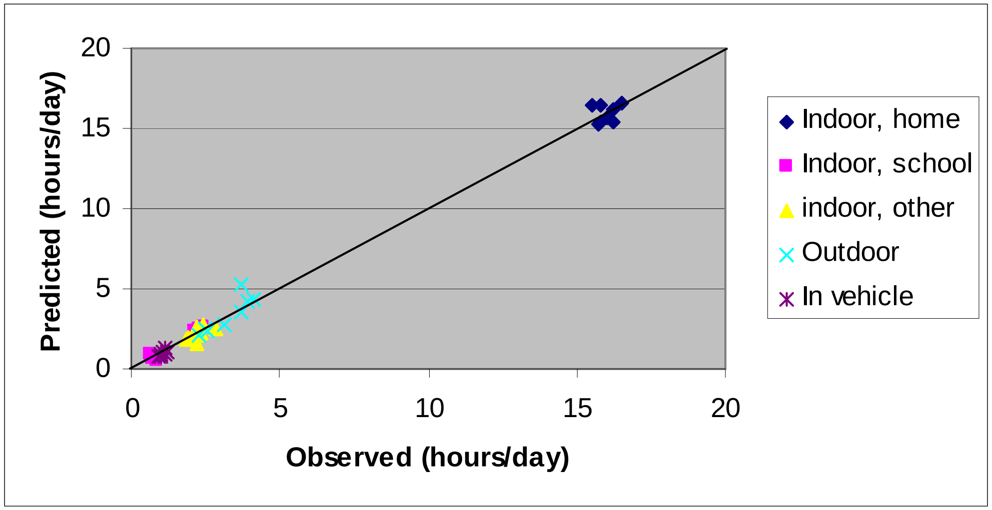


Figure 2. Comparison of predicted and observed average time in each of 5 microenvironments for age/gender groups and seasons.

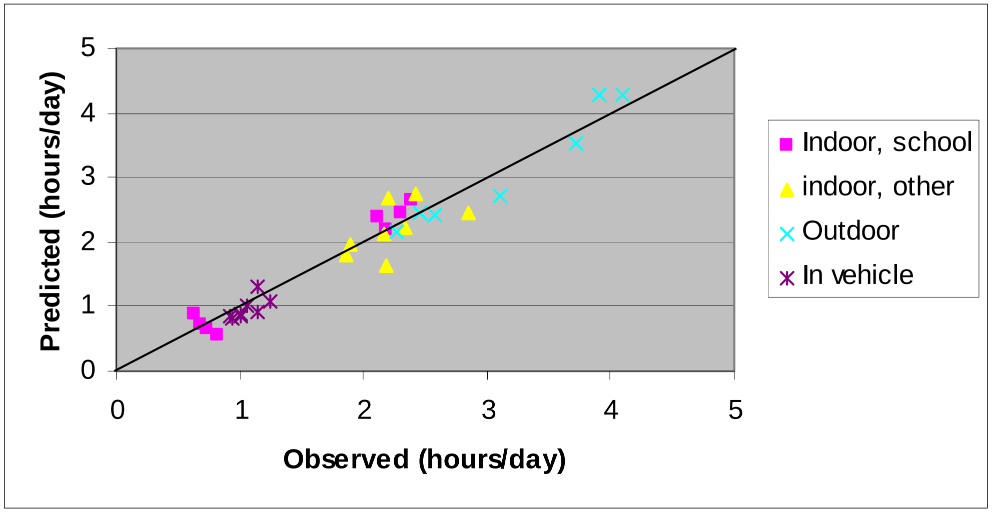


Figure 3. Comparison of predicted and observed average time in each of 4 microenvironments for age/gender groups and seasons.

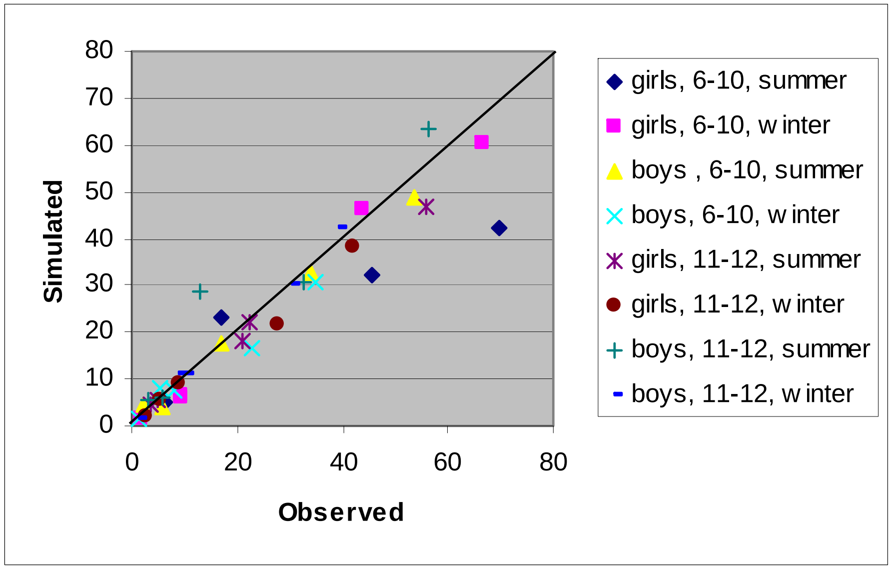


Figure 4. Comparison of predicted and observed variance across persons for time spent in each of 5 microenvironments for age/gender groups and seasons.

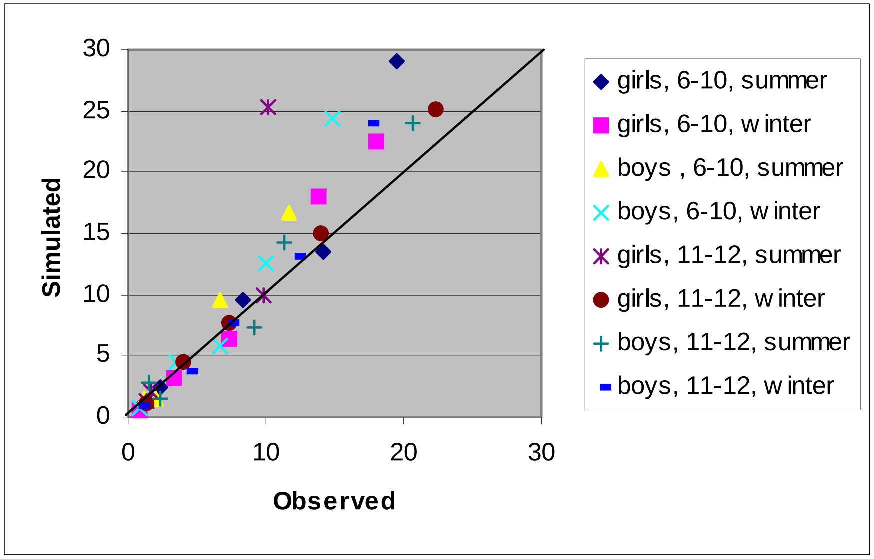


Figure 5. Comparison of predicted and observed the average within-person variance for time spent in each of 5 microenvironments by age/gender groups and seasons.

Appendix G – Exposure Risk Results for Asthmatics and Asthmatic Children

This Appendix provides supplemental exposure and risk characterization results for two subpopulations, all asthmatics and asthmatic children. The data are presented in series of summary tables and figures across each of the scenarios investigated (i.e. with modeled air quality as is and simulating just meeting the current standard), with and without modeled indoor sources (i.e., gas stoves), for each of the potential health effect benchmark levels (i.e., 200, 250, 300 ppb 1-hour), and across three years of modeled air quality (i.e., 2001 to 2003). Repeated exposures are presented only for the lowest potential health effect benchmark level (i.e., 200 ppb 1-hour).

G.1 All Asthmatics

Table 1. Estimated number of asthmatics in Philadelphia County exposed at or above potential health effect benchmark levels (1 to 6 times per year), using modeled air quality (as is) and with just meeting the current standard (std), and with and without indoor sources.

Year (AQ)	Indoor Source	Level (ppb)	Persons with Number of Repeated Exposures					
			1	2	3	4	5	6
2001 (as is)	Yes	200	49796	19544	8959	4516	2666	1732
		250	4867	1414	658	381	265	157
		300	1388	404	157	108	59	39
	No	200	10544	2577	1230	795	520	422
		250	2584	765	413	295	186	118
		300	1013	344	177	98	39	29
2001 (std)	Yes	200	128147	96119	70079	50253	35965	26167
		250	49632	18322	8523	4808	3095	2152
		300	16805	4480	1828	1219	866	638
	No	200	90211	51600	31720	19805	12899	8938
		250	40466	14362	6155	3225	2141	1414
		300	15100	3590	1595	1003	755	569
2002 (as is)	Yes	200	47652	17720	8056	4170	2662	1765
		250	4430	1173	530	274	166	127
		300	1240	393	147	88	69	49
	No	200	9505	2411	1240	706	401	323
		250	2276	778	332	185	117	88
		300	975	304	137	59	49	49
2002 (std)	Yes	200	133524	102861	77512	57152	42473	31800
		250	53367	20737	9855	5784	3489	2623
		300	18828	5220	2324	1447	925	648
	No	200	98849	60056	36913	23238	15850	10875
		250	43972	16367	7370	4066	2680	1734
		300	16693	4389	1950	1131	766	510
2003 (as is)	Yes	200	52639	22084	11950	7441	4863	3457
		250	14407	5040	2599	1577	935	650
		300	6568	1892	887	512	335	245
	No	200	26120	10007	5857	3783	2609	1842
		250	11142	3927	2040	1261	777	550
		300	5605	1627	778	462	285	206
2003 (std)	Yes	200	132640	102034	76909	58857	44719	34990
		250	73387	38505	22953	15416	11101	8499
		300	39283	16213	9280	6175	4374	3259
	No	200	109726	73489	51133	36551	27509	21181
		250	65437	33096	18948	12710	8964	6862
		300	35948	14502	8474	5654	4098	2935

Table 2. Estimated percent of asthmatics in Philadelphia County exposed at or above potential health effect benchmark levels (1 to 6 times per year), using modeled air quality (as is) and with just meeting the current standard (std), and with and without indoor sources.

Year (AQ)	Indoor Source	Level (ppb)	Percent (%) of Persons With Repeated Exposures					
			1	**2**	**3**	**4**	**5**	**6**
2001 (as is)	Yes	200	31	12	6	3	2	1
		250	3	1	0	0	0	0
		300	1	0	0	0	0	0
	No	200	6	2	1	0	0	0
		250	2	0	0	0	0	0
		300	1	0	0	0	0	0
2001 (std)	Yes	200	79	59	43	31	22	16
		250	31	11	5	3	2	1
		300	10	3	1	1	1	0
	No	200	55	32	20	12	8	5
		250	25	9	4	2	1	1
		300	9	2	1	1	0	0
2002 (as is)	Yes	200	29	11	5	3	2	1
		250	3	1	0	0	0	0
		300	1	0	0	0	0	0
	No	200	6	1	1	0	0	0
		250	1	0	0	0	0	0
		300	1	0	0	0	0	0
2002 (std)	Yes	200	82	63	48	35	26	20
		250	33	13	6	4	2	2
		300	12	3	1	1	1	0
	No	200	61	37	23	14	10	7
		250	27	10	5	2	2	1
		300	10	3	1	1	0	0
2003 (as is)	Yes	200	32	14	7	5	3	2
		250	9	3	2	1	1	0
		300	4	1	1	0	0	0
	No	200	16	6	4	2	2	1
		250	7	2	1	1	0	0
		300	3	1	0	0	0	0
2003 (std)	Yes	200	81	63	47	36	27	21
		250	45	24	14	9	7	5
		300	24	10	6	4	3	2
	No	200	67	45	31	22	17	13
		250	40	20	12	8	6	4
		300	22	9	5	3	3	2

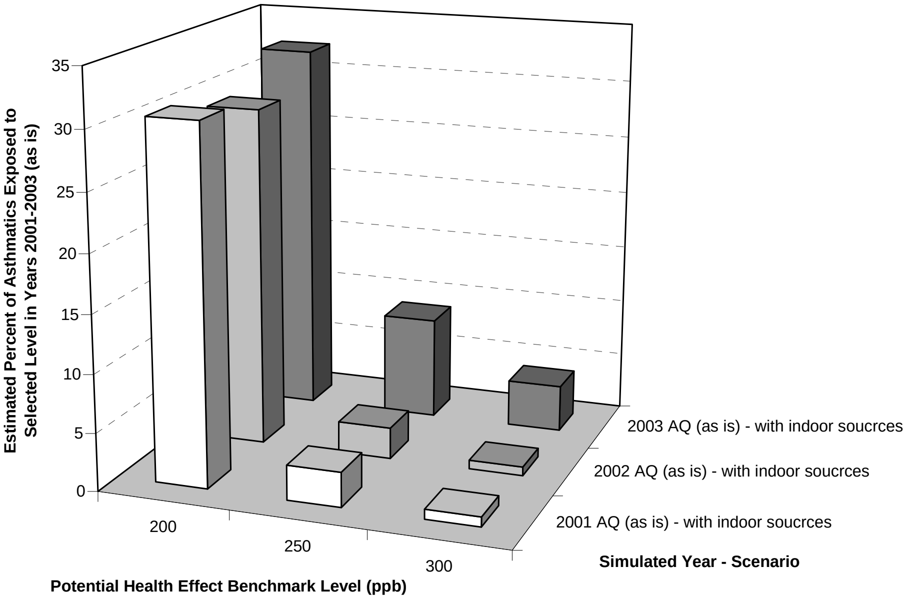


Figure 1. Estimated percent of all asthmatics in Philadelphia County with at least on NO_2 exposure at or above potential health effect benchmark level, using 2001-2003 modeled air quality (as is), with modeled indoor sources.

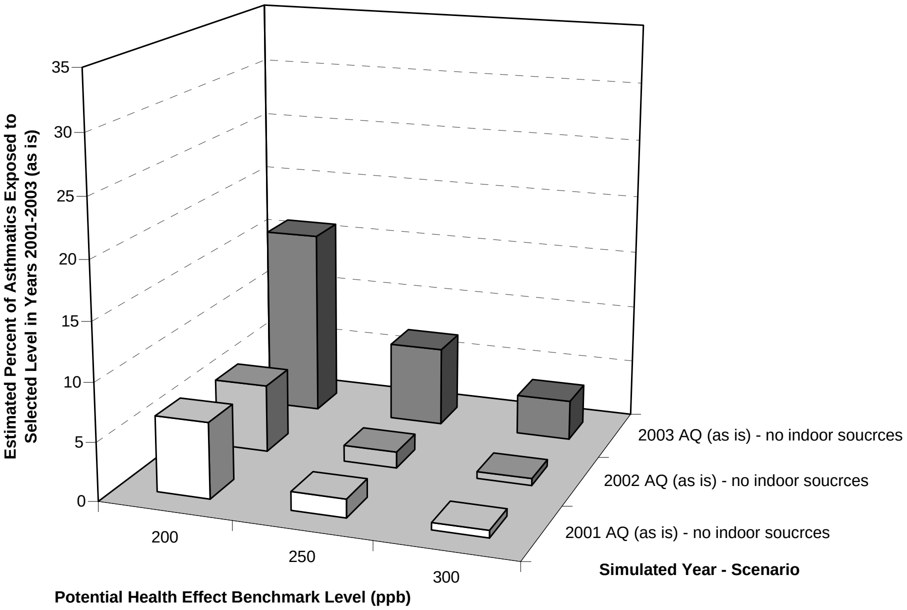


Figure 2. Estimated percent of all asthmatics in Philadelphia County with at least on NO_2 exposure at or above potential health effect benchmark level, using 2001-2003 modeled air quality (as is), with no indoor sources.

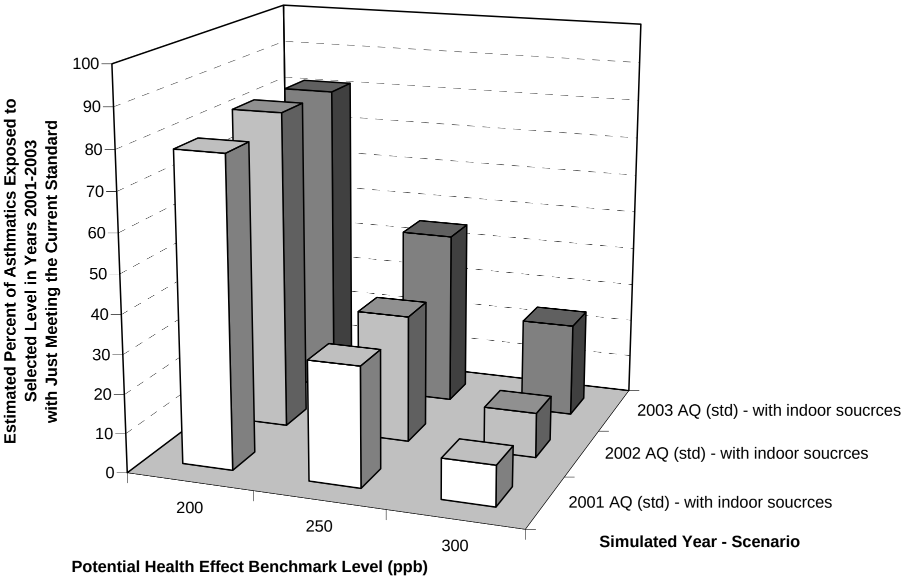


Figure 3. Estimated percent of all asthmatics in Philadelphia County with at least on NO_2 exposure at or above potential health effect benchmark level, using 2001-2003 modeled air quality just meeting the current standard (std), with modeled indoor sources.

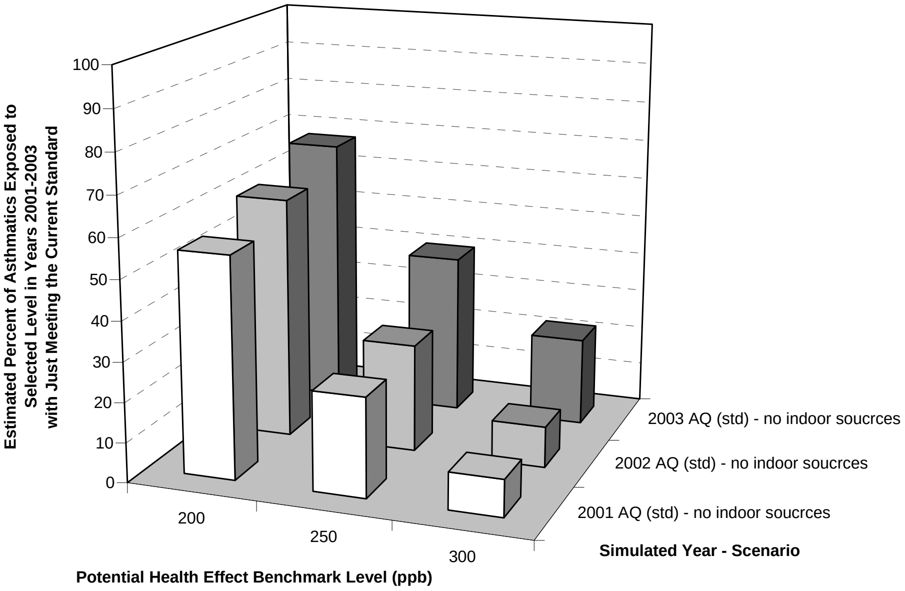


Figure 4. Estimated percent of all asthmatics in Philadelphia County with at least on NO_2 exposure at or above potential health effect benchmark level, using 2001-2003 modeled air quality just meeting the current standard (std), with no indoor sources.

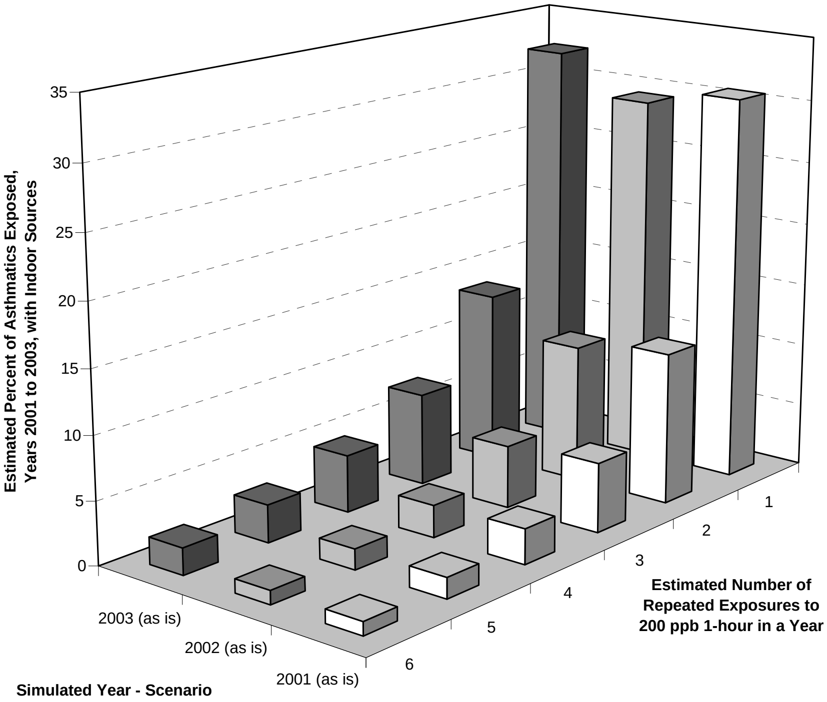


Figure 5. Estimated percent of all asthmatics in Philadelphia County with repeated NO_2 exposures at or above 200 ppb 1-hr, using 2001-2003 modeled air quality (as is), with modeled indoor sources.

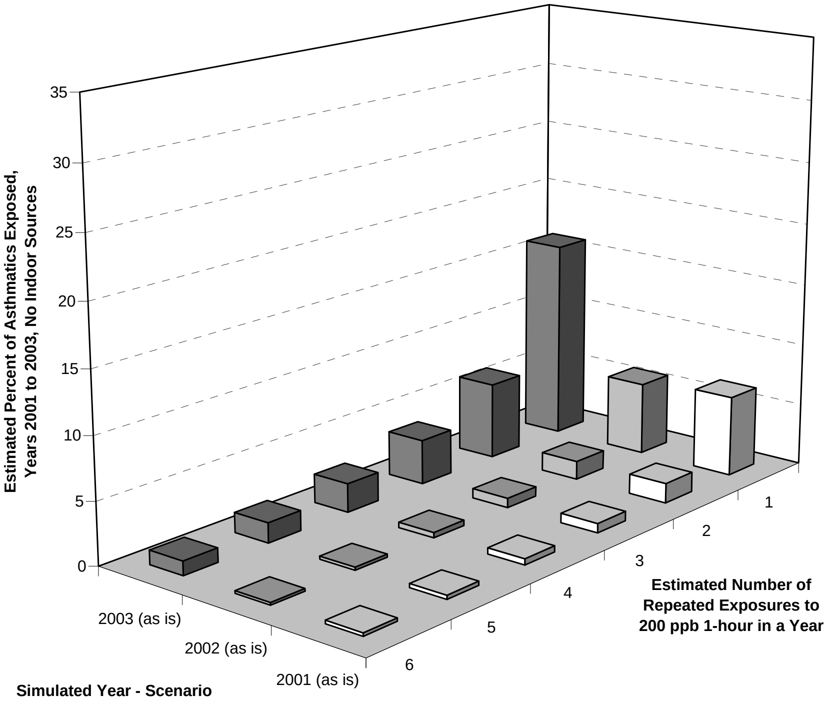


Figure 6. Estimated percent of all asthmatics in Philadelphia County with repeated NO_2 exposures at or above 200 ppb 1-hr, using 2001-2003 modeled air quality (as is), without indoor sources.

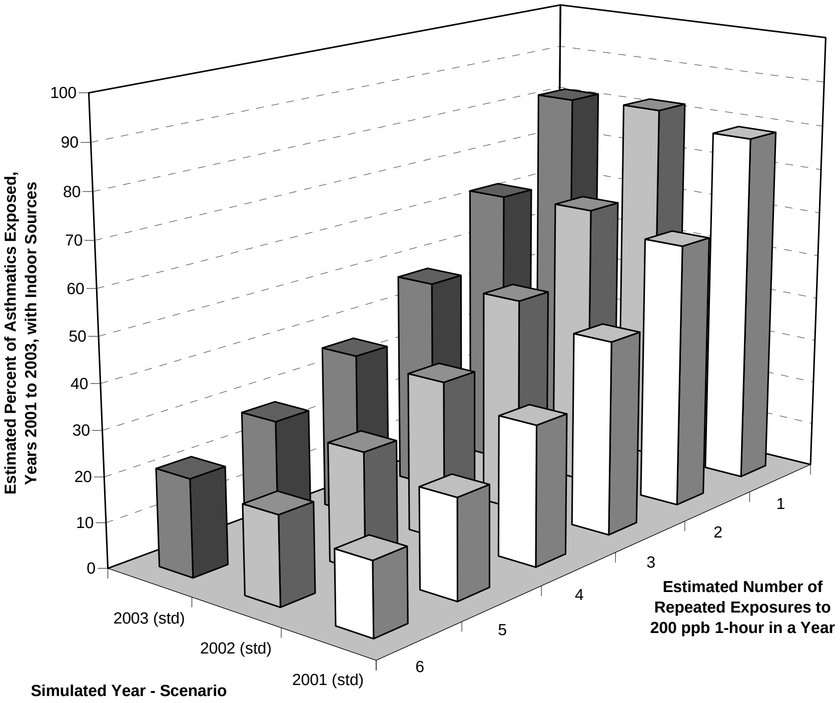


Figure 7. Estimated percent of all asthmatics in Philadelphia County with repeated NO_2 exposures at or above 200 ppb 1-hour, using 2001-2003 modeled air quality just meeting the current standard (std), with modeled indoor sources.

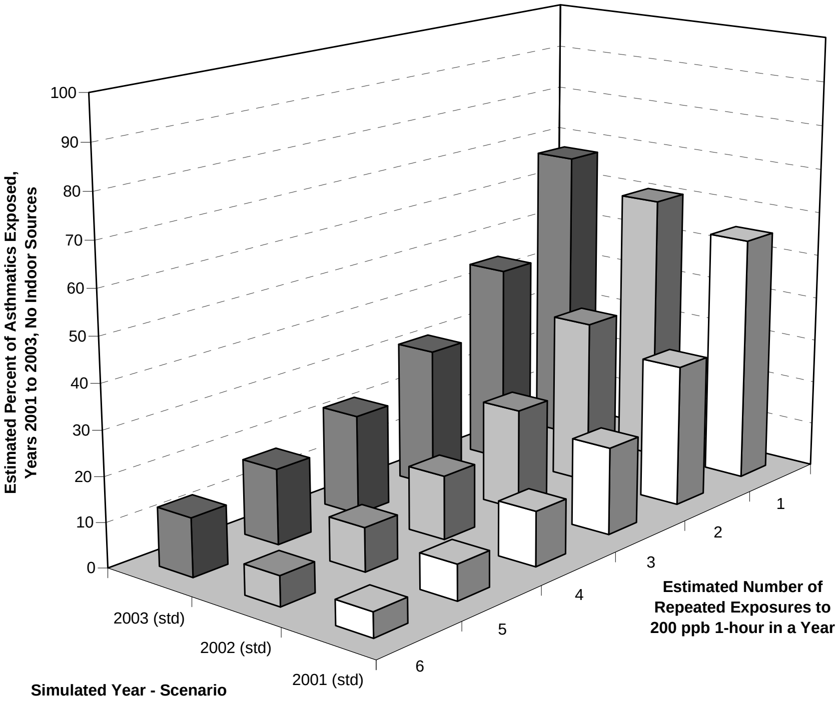


Figure 8. Estimated percent of all asthmatics in Philadelphia County with repeated NO_2 exposures at or above 200 ppb 1-hour, using 2001-2003 modeled air quality just meeting the current standard (std), with no indoor sources.

G-2 Asthmatic Children

Table 3. Estimated number of asthmatic children in Philadelphia County exposed at or above potential health effect benchmark levels (1 to 6 times per year), using modeled air quality (as is) and with just meeting the current standard (std), and with and without indoor sources.

	Indoor	**Level**	**Persons With Number of Repeated Exposures**					
Year (AQ)	**Source**	**(ppb)**	**1**	**2**	**3**	**4**	**5**	**6**
2001 (as is)	Yes	200	11351	3649	1418	709	424	267
		250	709	167	68	49	20	10
		300	128	49	10	10	0	0
	No	200	2329	401	147	98	58	58
		250	393	97	39	20	0	0
		300	97	29	10	10	0	0
2001 (std)	Yes	200	36656	26353	18272	12133	8271	5783
		250	13543	4530	1877	926	533	295
		300	3909	768	236	187	128	88
	No	200	27511	16067	9890	6094	3757	2430
		250	11282	3735	1413	500	333	197
		300	3440	638	187	128	109	79
2002 (as is)	Yes	200	10636	3338	1439	800	494	346
		250	692	139	49	30	0	0
		300	70	10	0	0	0	0
	No	200	1771	315	158	79	10	0
		250	158	49	20	10	0	0
		300	30	10	0	0	0	0
2002 (std)	Yes	200	38834	28678	20840	14308	10063	6996
		250	14855	4887	1978	1086	652	514
		300	4203	947	336	228	119	79
	No	200	30548	18685	11394	7063	4336	2782
		250	12487	3775	1288	738	493	365
		300	3736	670	276	158	99	39
2003 (as is)	Yes	200	12525	4693	2736	1712	1100	797
		250	3541	1240	678	423	247	178
		300	1545	423	237	138	89	39
	No	200	6724	2526	1515	984	708	492
		250	2784	1032	531	335	188	128
		300	1368	355	208	119	69	39
2003 (std)	Yes	200	37931	28305	20344	15230	11013	8483
		250	20044	9893	6016	4088	2888	2253
		300	10562	4100	2381	1643	1211	906
	No	200	32066	21662	14938	10326	7647	6018
		250	18770	8897	4974	3371	2388	1859
		300	9547	3704	2223	1496	1072	817

Table 4. Estimated percent of asthmatic children in Philadelphia County exposed at or above potential health effect benchmark levels (1 to 6 times per year), using modeled air quality (as is) and with just meeting the current standard (std), and with and without indoor sources.

Year (AQ)	**Indoor Source**	**Level (ppb)**	**Percent (%) of Persons With Repeated Exposures**					
			1	**2**	**3**	**4**	**5**	**6**
2001 (as is)	Yes	200	23	8	3	1	1	1
		250	1	0	0	0	0	0
		300	0	0	0	0	0	0
	No	200	5	1	0	0	0	0
		250	1	0	0	0	0	0
		300	0	0	0	0	0	0
2001 (std)	Yes	200	75	54	38	25	17	12
		250	28	9	4	2	1	1
		300	8	2	0	0	0	0
	No	200	57	33	20	13	8	5
		250	23	8	3	1	1	0
		300	7	1	0	0	0	0
2002 (as is)	Yes	200	22	7	3	2	1	1
		250	1	0	0	0	0	0
		300	0	0	0	0	0	0
	No	200	4	1	0	0	0	0
		250	0	0	0	0	0	0
		300	0	0	0	0	0	0
2002 (std)	Yes	200	81	60	43	30	21	15
		250	31	10	4	2	1	1
		300	9	2	1	0	0	0
	No	200	64	39	24	15	9	6
		250	26	8	3	2	1	1
		300	8	1	1	0	0	0
2003 (as is)	Yes	200	26	10	6	4	2	2
		250	7	3	1	1	1	0
		300	3	1	0	0	0	0
	No	200	14	5	3	2	1	1
		250	6	2	1	1	0	0
		300	3	1	0	0	0	0
2003 (std)	Yes	200	79	59	43	32	23	18
		250	42	21	13	9	6	5
		300	22	9	5	3	3	2
	No	200	67	45	31	22	16	13
		250	39	19	10	7	5	4
		300	20	8	5	3	2	2

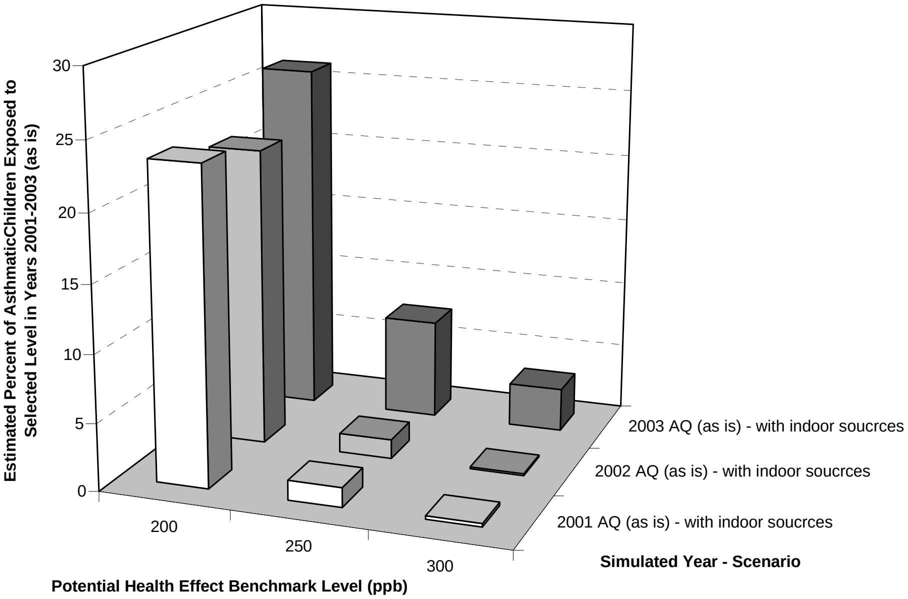


Figure 9. Estimated percent of asthmatic children in Philadelphia County with at least on NO_2 exposure at or above potential health effect benchmark level, using 2001-2003 modeled air quality (as is), with modeled indoor sources.

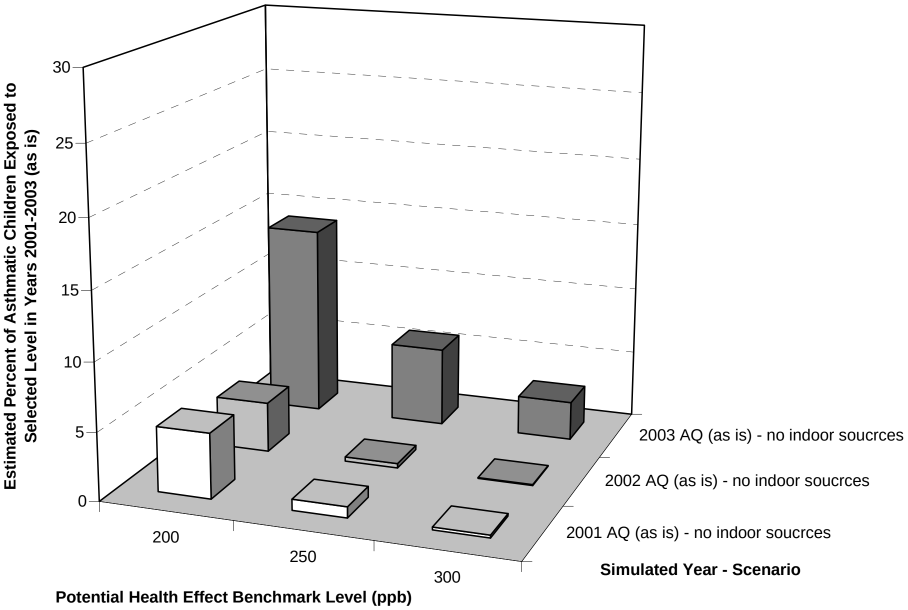


Figure 10. Estimated percent of asthmatic children in Philadelphia County with at least on NO_2 exposure at or above potential health effect benchmark level, using 2001-2003 modeled air quality (as is), with no indoor sources.

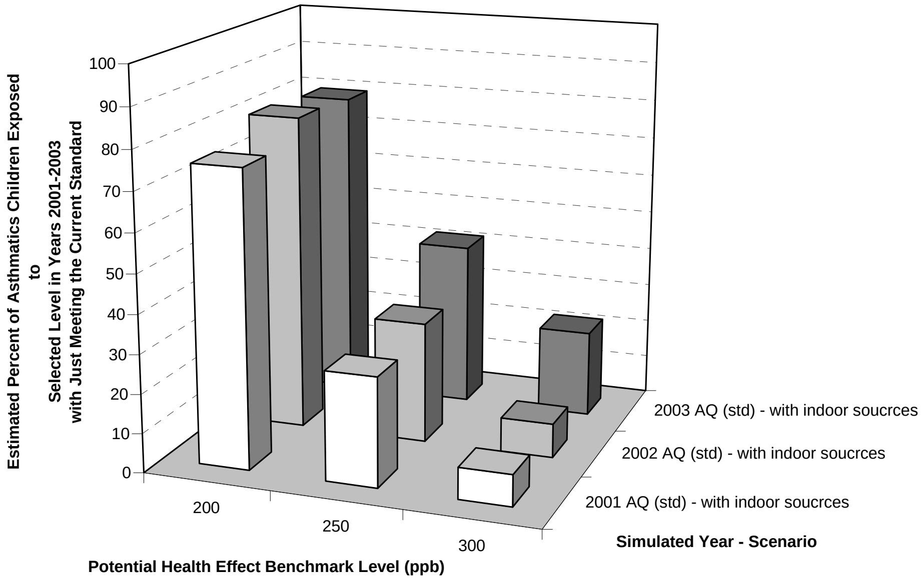


Figure 11. Estimated percent of asthmatic children in Philadelphia County with at least on NO_2 exposure at or above potential health effect benchmark level, using 2001-2003 modeled air quality just meeting the current standard (std), with modeled indoor sources.

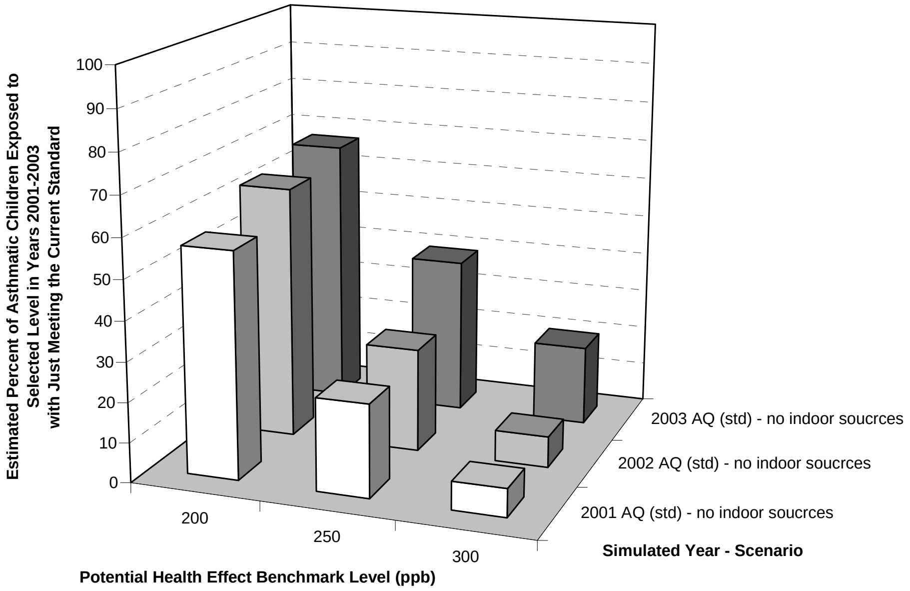


Figure 12. Estimated percent of asthmatic children in Philadelphia County with at least on NO_2 exposure at or above potential health effect benchmark level, using 2001-2003 modeled air quality just meeting the current standard (std), with no indoor sources.

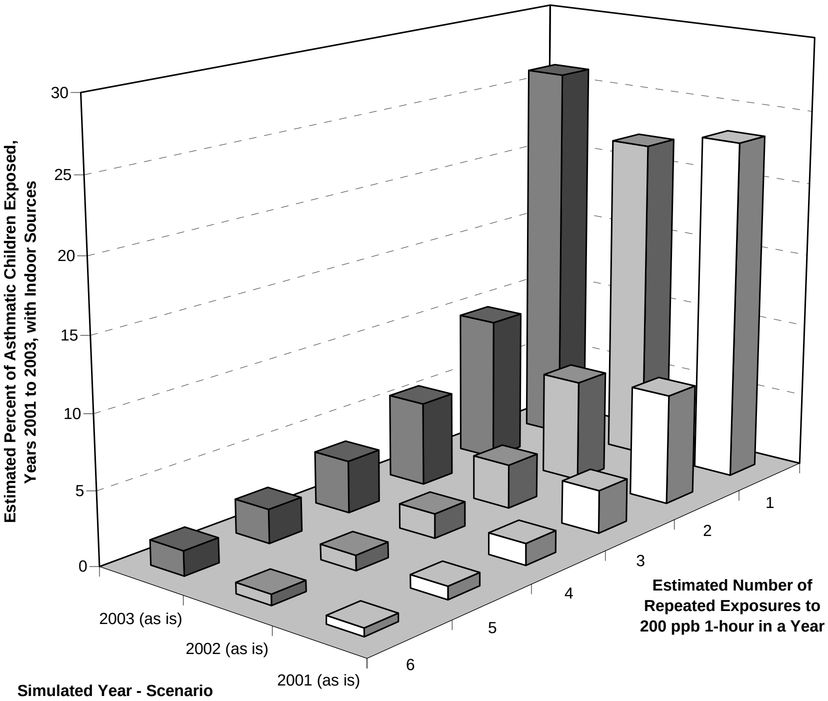


Figure 13. Estimated percent of asthmatic children in Philadelphia County with repeated NO_2 exposures at or above 200 ppb 1-hr, using 2001-2003 modeled air quality (as is), with modeled indoor sources.

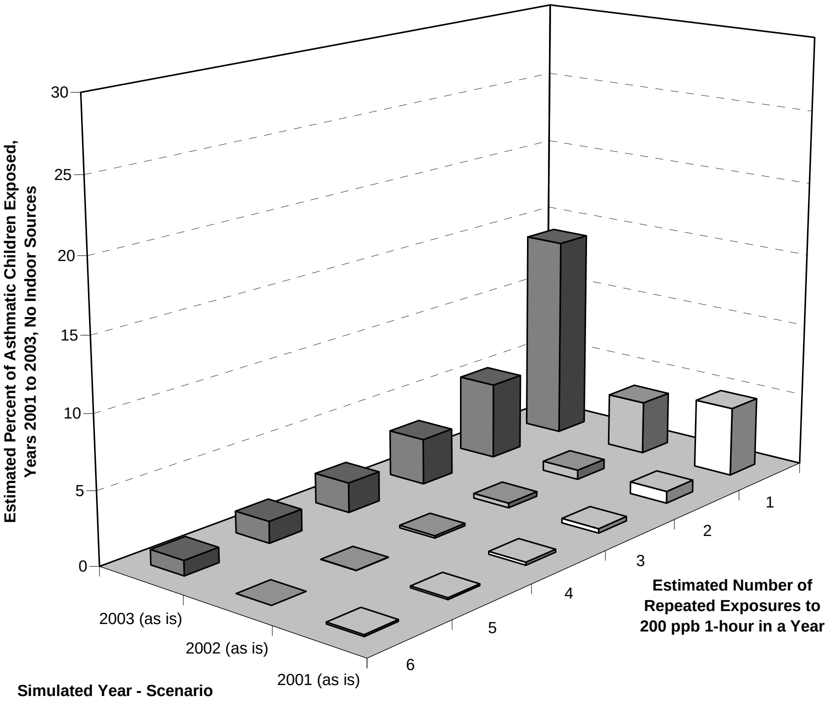


Figure 14. Estimated percent of asthmatic children in Philadelphia County with repeated NO_2 exposures at or above 200 ppb 1-hr, using 2001-2003 modeled air quality (as is), with no indoor sources.

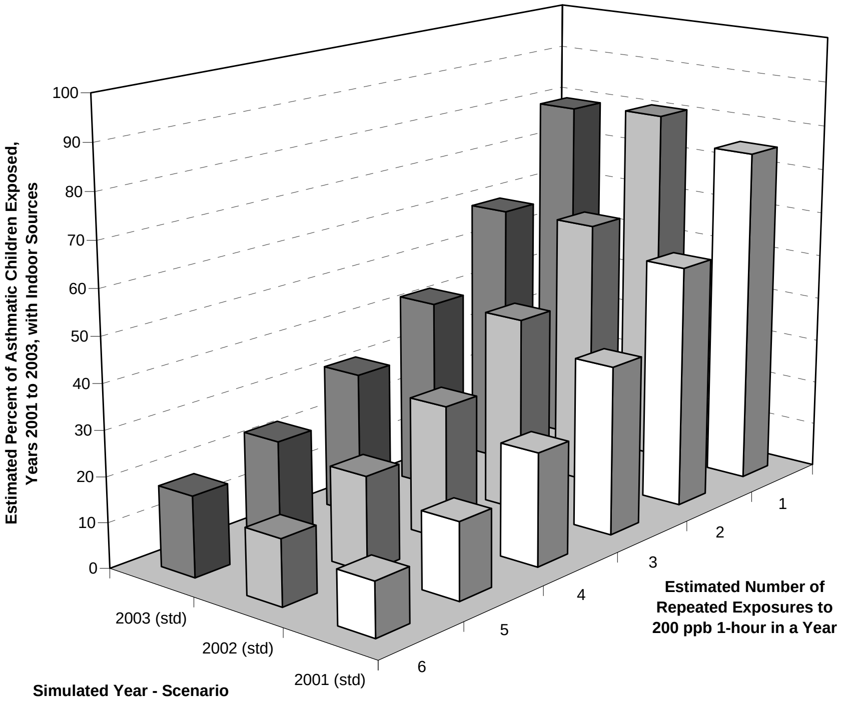


Figure 15. Estimated percent of asthmatic children in Philadelphia County with repeated NO_2 exposures at or above 200 ppb 1-hr, using 2001-2003 modeled air quality meeting the current standard (std), with modeled indoor sources.

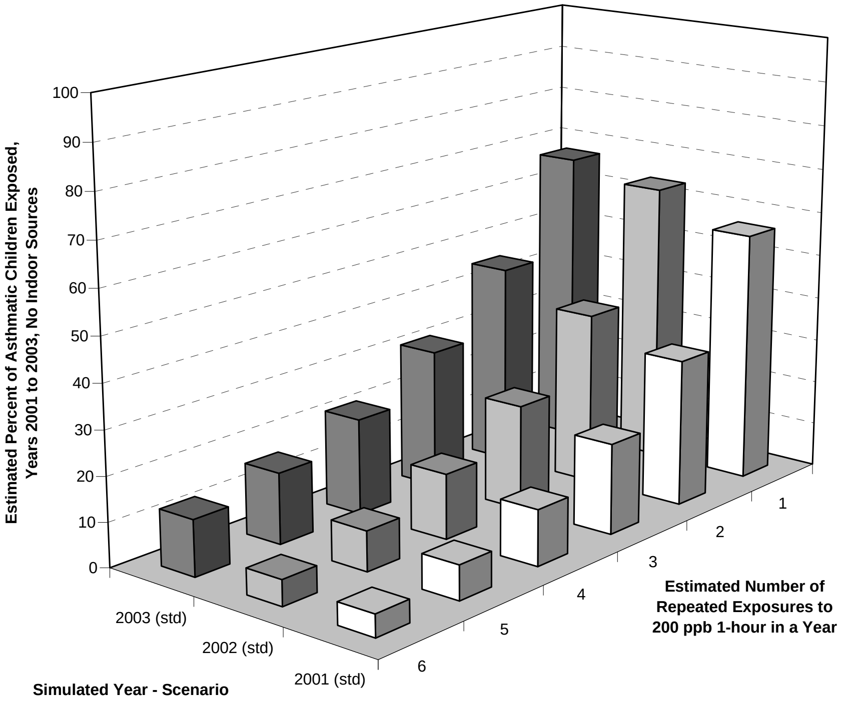


Figure 16. Estimated percent of asthmatic children in Philadelphia County with repeated NO_2 exposures at or above 200 ppb 1-hr, using 2001-2003 modeled air quality meeting the current standard (std), with no indoor sources.

United States
Environmental Protection
Agency

Office of Air Quality Planning and Standards
Air Quality Strategies and Standards Division
Research Triangle Park, NC

EPA-452/P-08-002
April 2008