Tracking Air Quality Trends with SAS/GRAPH®

David Mintz, U.S. EPA, Research Triangle Park, NC
Terence Fitz-Simons, U.S. EPA, Research Triangle Park, NC
Michelle Wayland, U.S. EPA, Research Triangle Park, NC

Abstract
Tracking and interpreting ambient air quality trends are the main functions of the Air Quality Trends Analysis Group at EPA. SAS/GRAPH is instrumental in this process, allowing us to study trends using several graphing techniques. This paper discusses graphics we have generated using combinations of the GMAP, GCHART, and GPLOT procedures. We will show how these graphics help us to discover, and to some extent explain, geographical, seasonal, and temporal patterns in the data.

Introduction
As part of the US EPA air program, trends in the quality of the nations air are analyzed and published each year in the *National Air Quality and Emissions Trends Report*. This report highlights trends for six air pollutants for which the EPA has set National Ambient Air Quality Standards (NAAQS). Ten-year trends in carbon monoxide, lead, nitrogen dioxide, (ground-level) ozone, particulate matter, and sulfur dioxide are reported with a national perspective. This paper discusses four graphics we have found to be very useful for displaying these trends, though their use may be extended to other types of data.

Displaying Temporal Patterns

Tile Maps
The tile map is good for displaying data with a seasonal pattern. We call it a tile map because one square, or *tile*, is plotted for each day of the year. The color of the tile shows the level of concentration for that day, lighter shades indicating lower concentrations and darker shades indicating higher concentrations. Due to the arrangement of all the days in a year, you can quickly see any seasonal pattern that may exist. In addition, you may look at year-to-year trends by including multiple tile maps on the same page. Figure 1 shows ozone trends in Los Angeles from 1982 to 1991. You can see that ozone readings are much higher in the summer months. (This is due to the fact that ozone is formed in the presence of abundant sunlight and higher temperatures.) Also, you can see that the number of high ozone days has steadily decreased since 1982. This graphic helps us to show that ozone is a "summer pollutant". The tile map can also be used to display other temporal patterns. For example, displaying diurnal patterns would work much the same way, except each tile would represent an hour of the day. Then day-to-day trends, let's say for one week, could be examined by including multiple tile maps on the same page. The tile map, if carefully constructed, can convey complicated temporal patterns in an easily understood manner.

Boxplots
The boxplot graphic is good for showing trends in the *distributions* of data, as opposed to just one value. This graphic consists of several boxplots, one for each year, connected by a trend line for a summary statistic, usually the mean. This allows you not only to detect trends over time, but also to tell if there are differences in the yearly distributions. For example, Figure 2 shows the national trend in carbon monoxide from 1984 to 1993. You can see a downward trend in the mean over the ten-year period. Looking at the boxplots, you can see that the downward trend is due to decreasing concentrations at the *high* end of the distributions. The boxplot graphic gives you the ability to view entire distributions, often revealing to you why a trend may exist.

Displaying Geographical Patterns

Population-County Maps
This graphic is a combination of two plots - a map and a stacked bar chart. The map shows the levels of concentrations across the country. The adjacent stacked bar chart shows the populations associated with the ranges of concentrations. Combining these two graphs is effective because you can see how many people live in areas at various concentration levels and where those areas are located across the U.S. In Figure 3, for example, there were about 6 to 7 million people living in counties that exceeded the NAAQS for PM-10 (particulate matter measuring ≤ 10 microns in diameter) in 1993. This level is represented by the darkest shading. In addition, you can see where those counties are located on the map. This graphic is very useful for cross-referencing two items of interest - population and geographic location, based on a third variable (in this

case, air quality concentration levels).

Boxplot-Dot Maps

This graphic also combines two plots - a boxplot and a map. The boxplot separates the data into quartiles, each represented by a different color. Arrows are drawn to show the values at each quartile and at other percentiles. Each dot on the adjacent map represents a monitoring site and is color-coded to correspond to the quartile ranges of the boxplot. (In some instances plotting too many points on one map may clutter the graphic, in which case splitting the map into four separate maps may be desired - such as we have done in Figure 4.) This display allows you to see where a particular monitor falls within the distribution. It also helps you to detect any geographical patterns that may exist. Figure 4, for example, shows that areas of highest PM-10 concentrations tend to be located along the west coast and the Ohio River Valley. Boxplot-dot maps can be very useful for cross-referencing spatial and distributional aspects of data.

Summary

We developed all the graphics described above with SAS/GRAPH software in the MVS environment. In addition, most of the graphics code is in macro form and is available upon request.

Acknowledgments

We would like to thank Nicole G. Mintz for her patient proofreading and thoughtful comments.

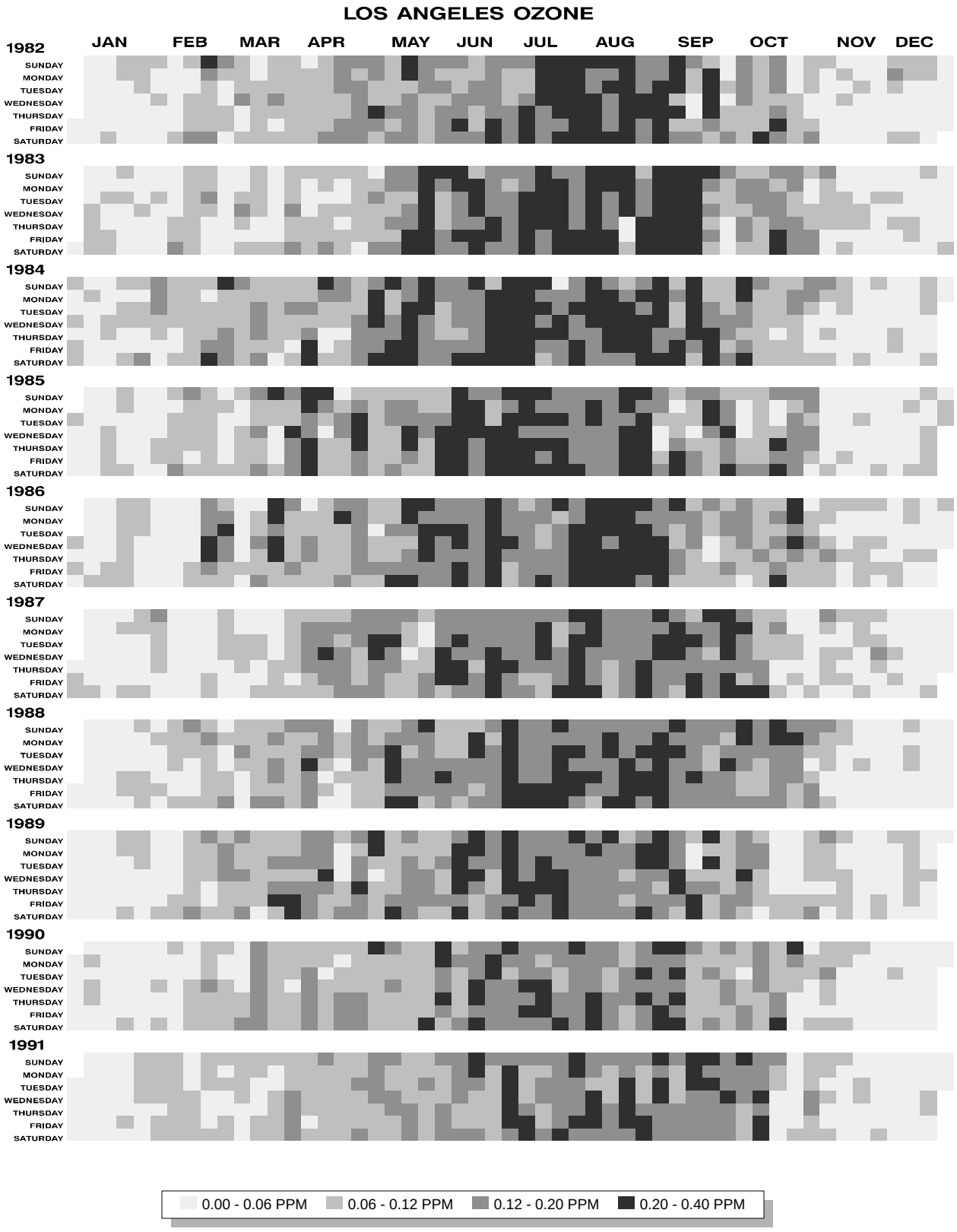
References

National Air Quality and Emissions Trends Report, 1991, EPA-450-R-92-001, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, October 1992

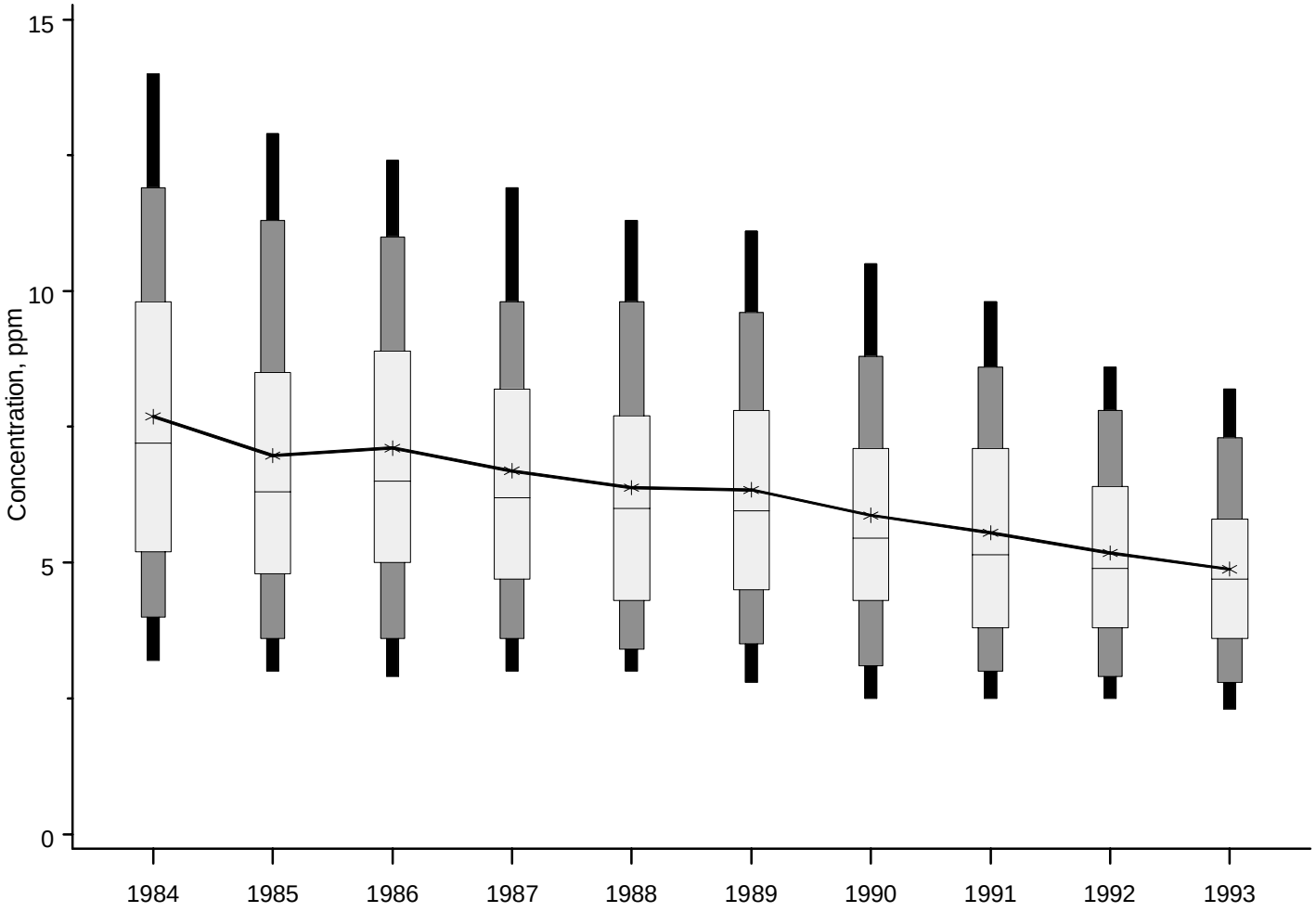
National Air Quality and Emissions Trends Report, 1993, EPA-454/R-94-026, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, October 1994

Further Information

If you have questions or comments, please contact David Mintz at mintz.david@epamail.epa.gov or the following mailing address:
USEPA/OAQPS/EMAD/AQTAG
Mail Drop 14
Research Triangle Park, NC 27711


LOS ANGELES OZONE
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
SUNDAY
MONDAY
TUESDAY
WEDNESDAY
THURSDAY
FRIDAY
SATURDAY
0.00 - 0.06 PPM
0.06 - 0.12 PPM
0.12 - 0.20 PPM
0.20 - 0.40 PPM

Carbon Monoxide Trend, 1984-93



PM-10 Air Quality Concentrations, 1993

Highest Second Max 24-Hour Average

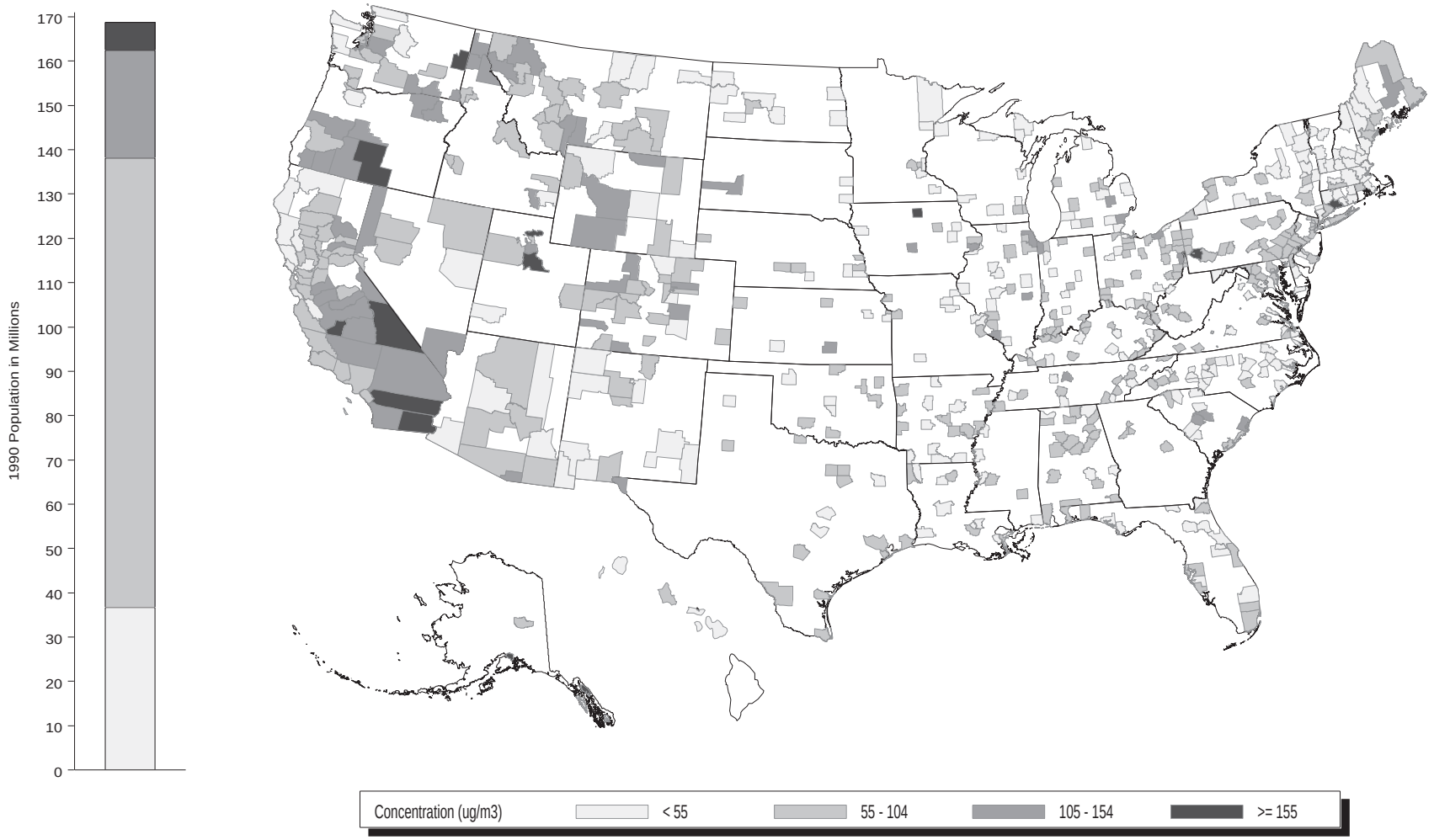
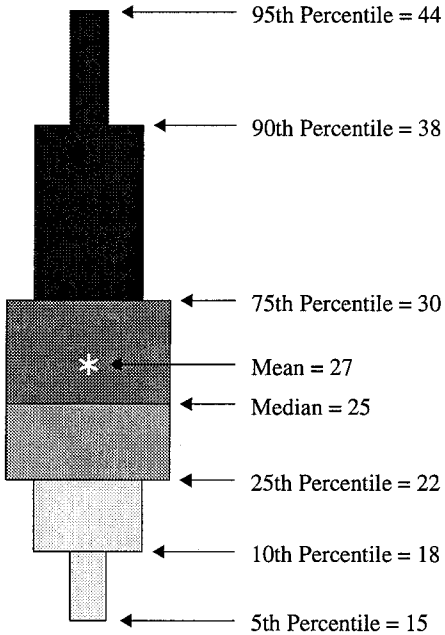


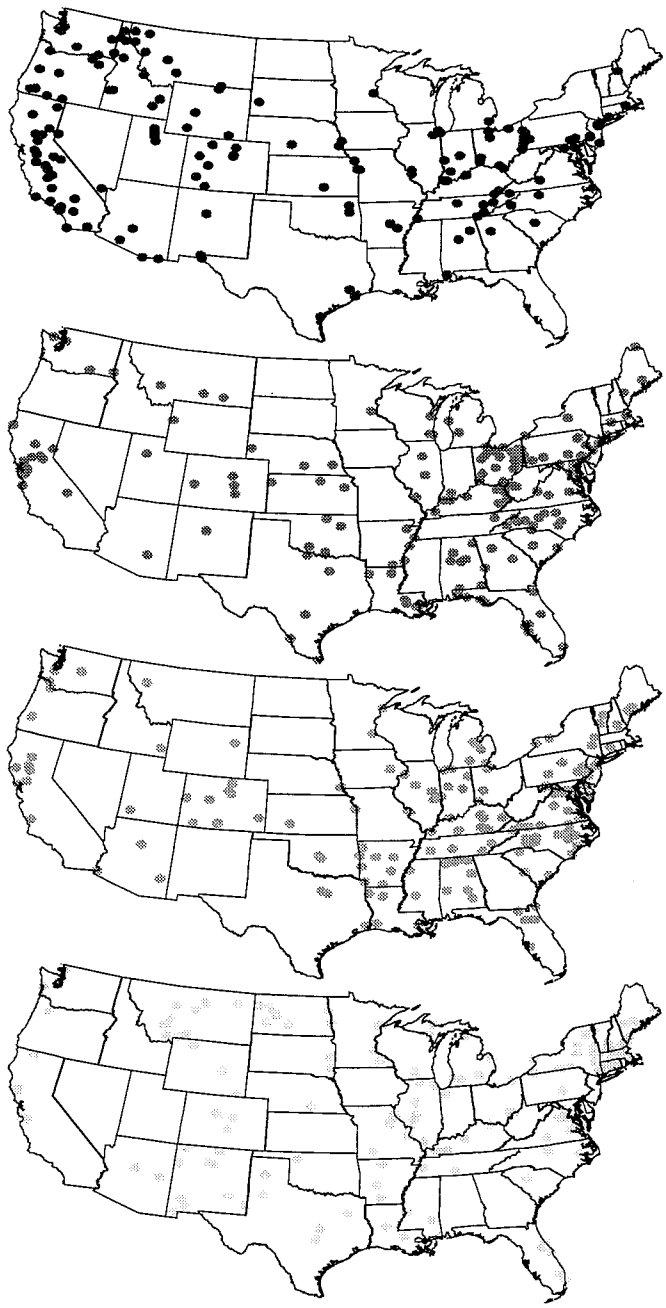
Figure 4

United States Environmental Protection Agency | Office of Air Quality Planning and Standards Research Triangle Park NC 27711 | EPA 454/R-94-026 October 1994

Air

National Air Quality and Emissions Trends Report, 1993





-1993 Particulate Matter-
Highest County Annual Mean
(micrograms per cubic meter)