Trends in Monitored Concentrations of Carbon Monoxide

Jo Ellen Brandmeyer, Peter Frechtel, and Margaret Z. Byron
RTI International, P.O. Box 12194, Research Triangle Park, NC 27709-2194

Joe Elkins and James Hemby
Air Quality Trends Analysis Group
U.S. Environmental Protection Agency, Research Triangle Park, NC 27711

Venkatesh Rao
Office of Air and Radiation
U.S. Environmental Protection Agency, Ann Arbor, MI 48105

Abstract

Carbon monoxide (CO) is one of the criteria pollutants regulated under the Clean Air Act. Numerous metropolitan areas instituted oxygenated gasoline (oxyfuel) programs during winter months to reduce CO emissions from motor vehicles, but some have since discontinued these requirements. This paper demonstrates a screening method for determining monitoring stations of potential interest. Monitoring stations with at least 8 years of relevant data during the period from 1990 through 2000 were screened for either an upward linear trend or upward inflection. Statistical tests assessed the trend in the annual second maximum nonoverlapping 8-hour average of CO for each monitor over the 11-year period. Of the 433 sites analyzed, 34 showed a statistically significant overall upward trend or statistically significant upward curvature. This analysis method can be used to screen for sites with increasing CO concentrations. The identified sites should then be examined further to determine the magnitude of the concentrations as compared to the existing standard. Because some areas have changed their fuel requirements within the last few years of the analysis, we recommend repeating this test annually.

Introduction

Carbon monoxide (CO) is a colorless, odorless, and poisonous gas produced by incomplete burning of carbon in fuels. Approximately 75% of nationwide CO emissions are from transportation sources. The largest emissions contribution comes from highway motor vehicles. Thus, the focus of CO controls as well as CO monitoring has been on traffic-oriented sites in urban areas where the main source of CO is motor vehicle exhaust. Other CO sources include wood-burning stoves, incinerators, and other heavy industrial sources.

The National Ambient Air Quality Standard (NAAQS) for carbon monoxide is 9 ppm for an 8-h average not to be exceeded more than once per year. The EPA motor vehicle program has achieved considerable success in reducing CO emissions. EPA standards in the early 1970s prompted automakers to improve basic engine design. By 1975, most new cars were equipped with catalytic convertors designed to convert CO to carbon dioxide. In the 1980s, automakers introduced more sophisticated converters plus on-board computers and oxygen sensors to help optimize the efficiency of the catalytic converter.

CO emissions from automobiles increase dramatically in cold weather because cars need more fuel to start at cold temperatures, and some emission control devices operate less efficiently when they are cold. Until 1994, vehicles were tested for CO emissions only at 75°F. But, recognizing the effect of cold weather, the 1990 Clean Air Act (the Act) calls for 1994 and later cars and light trucks to meet a carbon monoxide standard at 20°F as well.

The Act also stipulates expanded requirements for inspection and maintenance programs. These routine emission system checks should help identify malfunctioning vehicles that emit excessive levels of CO and other pollutants (the so-called "high emitters"). The inspections will be complemented by requirements for onboard warning devices to alert drivers when their emission control systems are not working properly.

Yet another strategy to reduce CO emissions from vehicles is to add oxygen-containing compounds to gasoline. This has the effect of "leaning-out" the air-to-fuel ratio, thereby promoting more complete fuel combustion. The most common oxygen additives are ethers and alcohols. Several western and northern U.S. cities have employed wintertime oxygenated gasolines for many years. The Act expands this concept and requires that oxygenated gasolines be used during the winter months in certain metropolitan areas with high CO levels.

With these control programs and technology improvements, today's passenger cars and light-duty trucks are capable of emitting 90% to 95% less CO over their lifetimes than their uncontrolled counterparts of the 1960s. As a result, ambient CO levels have dropped, despite large increases in the number of vehicles on the road and the number of miles they travel. However, in recent months, with continued heavy increase in vehicle travel, there have been indications that CO levels are climbing again in certain parts of the country. The objective of this work is to examine those areas of the country where mobile-source activity is heavy (in CO nonattainment and problem areas) and/or where CO air quality has been a persistent problem and determine whether CO levels are increasing.

Experimental Methods

CO concentration data were extracted for 858 monitoring sites from EPA's Aerometric Information Retrieval System (AIRS) on March 14, 2002. To meet the completeness requirement for this analysis, at least 8 years of data must have been available for the years 1990 to 2000, inclusive. Statistical analyses were performed for the 433 sites that met this requirement.

The Metropolitan Statistical Area (MSA) code was also downloaded for each site. The codes were linked to the most recent list of areas that employ or have discontinued oxyfuel requirements.[1] This information was used to group the sites (oxyfuel ended vs. no change in oxyfuel requirements) and to interpret the results of the analyses.

The effects of meteorology on ambient CO concentrations were not examined in this study. For example, certain meteorological parameters (e.g., mixing height and windspeed) need to be considered when comparing emissions to ambient concentration measurements.[3,4] However, the Glen et al. study[3] concluded that seasonal fluctuations in CO concentrations are explained by the variations in these meteorological parameters, whereas the long-term trend is primarily due to the trend in emissions. Although the current analysis did not account for interannual meteorological changes, the same overall downward trend was identified.

The analysis used the second maximum nonoverlapping 8-h average CO concentration (SECMX) for each year. This statistic was selected for analysis because it coincides with the 8-h NAAQS for CO. Missing values (i.e., years without a SECMX value for a monitor) were not filled in; that is, linear interpolation or some other method was not employed to fill in missing data. The data for each site were then analyzed independently of all other sites; that is, no spatial averaging was performed to obtain annual average values for each MSA.

Although the SECMX values form the basis of the annual CO trends published by EPA's Air Quality Trends Analysis Group in the Trends Report,[2] the methodology employed in this study differed in three basic ways:

- The Trends Report fills in missing data, whereas this study used only the data that were available from AIRS.
- The Trends Report aggregates data and analyzes results for each MSA, whereas this study performed the data analysis separately for each monitor.
- The analysis for the Trends Report used only the nonparametric Theil test, whereas this study also used two linear regression models.

The three analyses that were performed for each site were the Theil test, first-order linear regression, and quadratic (second-order) linear regression. Each of these analyses included a statistical hypothesis test that computes a p-value for each monitor. If the p-value is less than a critical value n between 0 and 1, then the test has a result that is "significant at $\alpha = n$." A smaller value for α indicates a greater likelihood that the data truly possess the detected trend.

Every test was two-sided, meaning that the α-level used to detect an increasing or a decreasing trend was $\alpha/2$. Therefore, if a monitor exhibited an increasing trend, then the p-value for the test would have to be less than $\alpha/2$ for the increasing trend to be significant. For example, if a monitor exhibited an upward trend that was significant at $\alpha = 0.01$, then the probability of seeing as extreme an upward trend as this monitor under the null hypothesis of no trend is less than 0.005 (0.5%).

The Theil test and both regression models are discussed below.

Theil Test

The Theil test[5] is a nonparametric statistical test that can be used instead of regression-based methods for discerning a monotonic trend. It examines whether the concentration from year to year tends to increase or decrease consistently, making it a test of monotonicity. This test is not concerned with the magnitude of the year-to-year differences. The null hypothesis is that there is no monotonic trend in the data.

The first step in the test is to examine all possible [n(n-1)/2] pairs of data points from a given monitor, where n = 8, 9, 10, or 11. Next, a count is taken of all the pairs that show an increasing or decreasing trend. The null hypothesis will be rejected and the test results will indicate a significant monotonic increasing (or decreasing) trend if this count of the data point pairs is greater than (or less than) a certain critical value. A large positive value indicates a positive trend, and a large negative value indicates a negative trend.

The Theil test was applied for two reasons. First, it is appropriate when the errors from a linear regression are not normally, or close to normally, distributed. The data here may not meet the normality assumption. Second, this test was recommended to EPA for determining whether an area has a significant trend.[6] Therefore, this test is used in EPA's annual Trends Reports.

Choice of Urban and Rural Sites

Unlike the Theil test, linear regression is a parametric test. All linear regression models incorporate three basic assumptions: (1) the data are normally distributed, (2) the variance is constant at each time, and (3) no autocorrelation exists between time periods.

A first-order linear regression was performed using PROC REG in SAS.[7] The linear regression model used SECMX as the dependent variable. To make the results less dependent on the magnitude of the year, a transformation was performed on the value of the year by subtracting 1989 (i.e., 1 less than the minimum year in the dataset):

$$YR' = YEAR - 1989 \qquad (1)$$

YR' was the only independent variable in the regression model. PROC REG includes a hypothesis test for a nonzero slope. The *p*-value from this hypothesis test is presented in the results tables.

Quadratic Regression

A second linear regression was also performed using PROC REG. This test was a quadratic (second-order) linear regression that used both (YR´) and $(YR')^2$ as independent variables. The *p*-value from the test for a nonzero coefficient on the squared term is presented in the results tables. A significant *p*-value for this test indicates significant curvature in the regression line. That is, an upward trend suggests that the slope has increased from the early years to the recent years.

Interpretation of Statistical Results

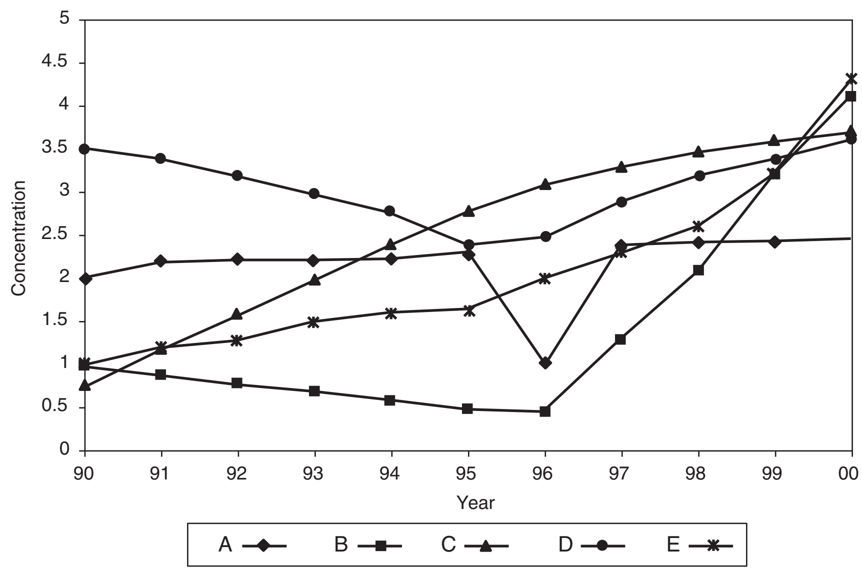
These three statistical tests are complementary in that each examines the data differently. The Theil test looks for a monotonic trend, first-order linear regression applies normality theory for a linear trend, and quadratic regression applies normality theory for a nonlinear trend. All three spotlight sites that may be of interest to policy makers, but no single test will detect all interesting sites. They can be used together, however, to discern patterns in the data. Consider the following five trends, as illustrated in Figure 1.

Trend A

This site has a consistent, upward trend that is not dramatic. However, 1996 was a very "clean" year at the site, with a SECMX value lower than the rest of the years.

The Theil test undoubtedly will detect a significant upward trend at site A. The first-order regression model may not find a significant trend at site A for two reasons. First, the anomalous point in 1996 inflates the variance. Second, the slope estimate will not be much greater than zero because the increasing trend is only slight. The quadratic regression model may or may not be significant for this site.

Figure 1. Examples of trends A through E.



Site A may be of interest to policy makers. For example, upon examination of associated data such as temperature, they may find a meteorological reason that 1996 was such a clean year (e.g., warm winter) and decide that the true pattern is a consistent increase in CO concentration.

Trend B

From 1990 to 1996, the concentrations at site B decreased slightly. The concentrations then increased dramatically from 1997 to 2000.

At site B the Theil test may not detect a trend because of a lack of a consistent pattern in the early years. It also will not be influenced by the explosive pattern in the recent years. However, the first-order regression model will certainly detect an increasing trend. The high concentrations in the later years will increase the slope of the regression line. If the increase is more dramatic in the very recent years, the quadratic regression model may also detect a significant upward inflection.

Site B also would likely be of interest to policy makers, because the most recent years show a dramatic increase in concentration.

Trend C

The concentrations at site C increased dramatically from 1990 to 1995. The rate of increase then slowed from 1996 to 2000, although the concentrations continued to increase.

At site C, both the Theil test and the first-order regression model will detect an increasing trend. However, the quadratic regression model might detect a downward curvature.

This may be a site where population growth is explosive, but the state or local government has taken drastic steps to reduce emissions per capita. This pattern is likely to interest policy makers because the site is showing improvement via slower concentration growth, although the concentration at the site is still increasing.

Trend D

The concentrations at site D decreased from 1990 to 1995 but increased from 1996 to 2000. The concentrations in 1990 and 2000 were similar to each other.

At site D, both the Theil test and the first-order regression model likely will fail to detect a trend. The Theil test will have about the same number of increasing and decreasing pairs. The slope of the first-order linear regression line likely will be nearly zero. The quadratic regression model, however, will detect a significant upward curvature.

This site may be of interest to policy makers because the pattern suggests that the concentrations will continue to increase. This pattern may be prevalent where the oxyfuels program was discontinued.

Trend E

The concentrations at site E increased from 1990 to 1995. The increase became more pronounced from 1996 to 2000.

At site E, all three tests will produce significant results. This site exhibits a consistent increase in concentrations, and it merits special vigilance.

Results and Discussion

This study analyzed data for the 433 sites that met the completeness test. One or more statistical tests revealed significance at 79% of the sites at the $\alpha = 0.10$ level. This result was expected due to the effects of fleet turnover.

Of greater interest to this study, however, was that a statistically significant upward trend or curvature was revealed at 34 sites. Table 1 lists the results of the three statistical models for all sites where at least one model revealed a significant upward trend or positive quadratic component. Seven pieces of information are included for each site: (1) MSA containing the site, (2) ending date for the oxyfuel program (if applicable), (3) monitor ID in AIRS, (4) number of years of data used in the analysis, (5) results of the Theil test, (6) results of a hypothesis test that the slope of the line from the first-order linear regression model is nonzero, and (7) results of a hypothesis test that the coefficient associated with the squared term is nonzero for the quadratic regression model. Of the sites listing dates ending the oxyfuel program, all either are located in a federal reformulated gasoline area or have an oxyfuel requirement in their contingency plan.

Figure 2 shows the locations of the monitoring sites with at least one statistical model showing a statistically significant upward trend or positive quadratic component. Only those sites located within the coterminous United States are included in this map.

A plot of the SECMX vs. year was generated for each of the 433 sites in this analysis. For each plot the concentration values are shown as stars. The solid line represents the quadratic regression line, and the dashed lines represent the 95% confidence bands around the regression line. That is, there is a 95% probability that the true trend lies within the area bounded by the dashed lines and only a 5% probability that the true trend lies outside this area. Examples of patterns found in these plots are included as Figures 3 through 7.

Table 1. Carbon Monoxide Monitoring Sites Where at Least One Statistical Test Shows Increasing Concentration

MSA	Ending Date Oxyfuel Requirement	Monitor ID	Years of Data	Theil Test	1st Order Regression Model	2nd Order Regression Model
—	—	370770001421011	8	NS	UP10	NS
—	—	410350006421011	11	DOWN01	DOWN01	UP01
Charlotte, NC	—	371190038421011	11	DOWN01	DOWN01	UP05
Charlotte, NC	—	371191009421011	8	UP05	UP05	NS
Kansas City, MO	—	290470009421011	10	DOWN05	DOWN05	UP10
Los Angeles, CA	—	060371201421011	11	DOWN01	DOWN01	UP10
Los Angeles, CA	—	060379002421011	11	DOWN01	DOWN01	UP05
Louisville, KY	—	211110046421011	11	DOWN05	DOWN01	UP05
Minneapolis–St. Paul, MN	—	271230865421011	8	DOWN05	DOWN05	UP05
Modesto, CA	6/1/1998*	060990005421011	11	DOWN05	DOWN05	UP01
Oakland, CA	—	060010003421011	10	DOWN05	DOWN01	UP10
Oakland, CA	—	060130002421011	11	DOWN01	DOWN01	UP05
Oakland, CA	—	060133001421011	11	DOWN01	DOWN01	UP05
Vancouver, WA	10/21/1996*	530110010421011	11	DOWN01	DOWN01	UP01
Provo, UT	—	490490002421011	11	DOWN01	DOWN01	UP05
Reno, NV	—	320311005421011	11	DOWN01	DOWN01	UP05
Sacramento, CA	6/1/1998*†	060170010421011	9	DOWN05	DOWN01	UP01
Sacramento, CA	6/1/1998*†	060170011421011	8	DOWN01	DOWN01	UP10
Sacramento, CA	6/1/1998*†	060670006421011	11	DOWN01	DOWN01	UP10
Sacramento, CA	6/1/1998 *†	060670007421011	11	DOWN01	DOWN01	UP01
San Diego, CA	6/1/1998*†	060730003421011	10	DOWN01	DOWN01	UP05
San Diego, CA	6/1/1998*†	060731007421011	11	DOWN01	DOWN01	UP01
San Francisco, CA	6/1/1998*	060811001421011	11	DOWN01	DOWN01	UP10
San Jose, CA	—	060850004421011	11	DOWN05	DOWN01	UP01
San Jose, CA	—	060850004421012	11	DOWN05	DOWN01	UP01
San Juan, PR	—	721270002421011	11	DOWN05	DOWN05	UP10
San Luis Obispo, CA	—	060792002421011	11	DOWN01	DOWN01	UP10
Santa Rosa, CA	—	060970003421011	11	DOWN05	DOWN05	UP10
Seattle, WA	10/11/1996*	530610012421011	11	DOWN01	DOWN01	UP05
Stockton, CA	6/1/1998*	060770008421011	11	DOWN05	DOWN05	UP05
Stockton, CA	6/1/1998*	060771002421011	11	DOWN05	DOWN01	UP05
Tampa, FL	—	120571045421011	8	DOWN05	DOWN01	UP10
Vallejo, CA	—	060950004421011	11	DOWN01	DOWN01	UP05
Yuba City, CA	—	061010003421011	10	DOWN01	DOWN01	UP05

*Oxyfuel program retained as contingency measure.
†Federal reformulated gasoline program area.

The following notation was used for the statistical results:

DOWN01 = downward trend, significant at α level 0.01
DOWN05 = downward trend, significant at α level 0.05
DOWN10 = downward trend, significant at α level 0.10
NS = no significant trend
UP01 = upward trend, significant at α level 0.01
UP05 = upward trend, significant at α level 0.05
UP10 = upward trend, significant at α level 0.10

Figure 3 illustrates a site that was screened out by this analysis; none of the three tests revealed an upward trend. The statistical results were DOWN01, DOWN01, and NS for the Theil test, first-order linear regression, and quadratic regression, respectively.

The Theil test revealed a statistically significant upward trend at only one site. Its data and quadratic regression results are shown in Figure 4. The first-order linear regression model also revealed an upward trend at this site. Both these tests were significant at the $\alpha = 0.05$ level. The second-order linear regression found no significant trend at this site. This pattern is similar to Trend C, described above.

Figure 4 also demonstrates how this analysis method should be used to screen monitoring sites. Although two statistical tests revealed an upward trend, this site is not of immediate concern because the concentrations are far below the NAAQS value of 9 ppm. If this site is located in an area of high population growth, then it should be reevaluated in the future.

Figures 5 through 7 illustrate patterns that are similar to Trend D, described above. The site in Figure 5 apparently experienced minimum CO concentrations during the period 1995 to 1997. The concentrations increased after that period. For this site, the Theil test revealed a downward pattern at the $\alpha = 0.05$ level, and the first-order linear regression model revealed a downward pattern at the $\alpha = 0.01$ level. However, the quadratic regression model revealed an upward pattern at the $\alpha = 0.01$ level. Also, the lower bound of the 95% confidence limit is increasing, and concentrations are not low like those shown in Figure 4.

Figure 2. Locations of monitoring sites in the coterminous United States with at least one statistical model showing a significant upward trend. Circles represent sites that have stopped an oxygenated gasoline requirement. Diamonds represent other sites.

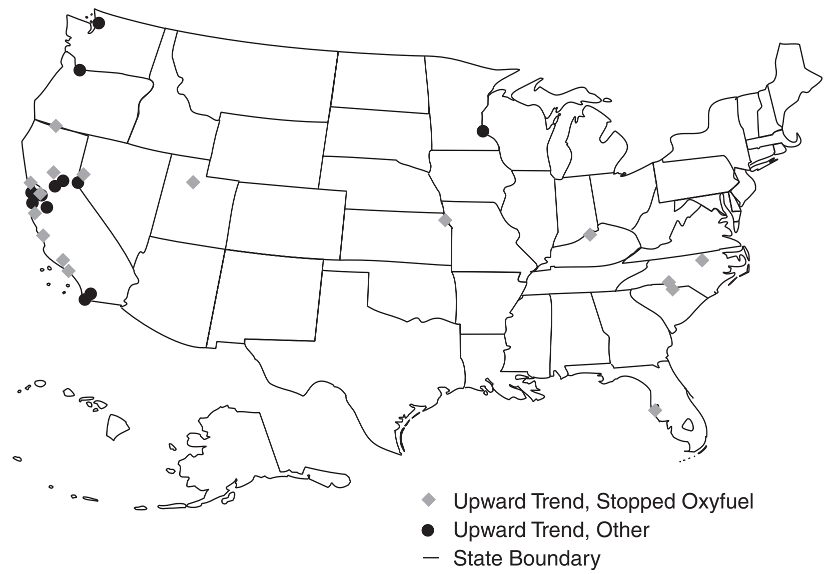


Figure 3. Example of a site screened out by the combined statistical models.

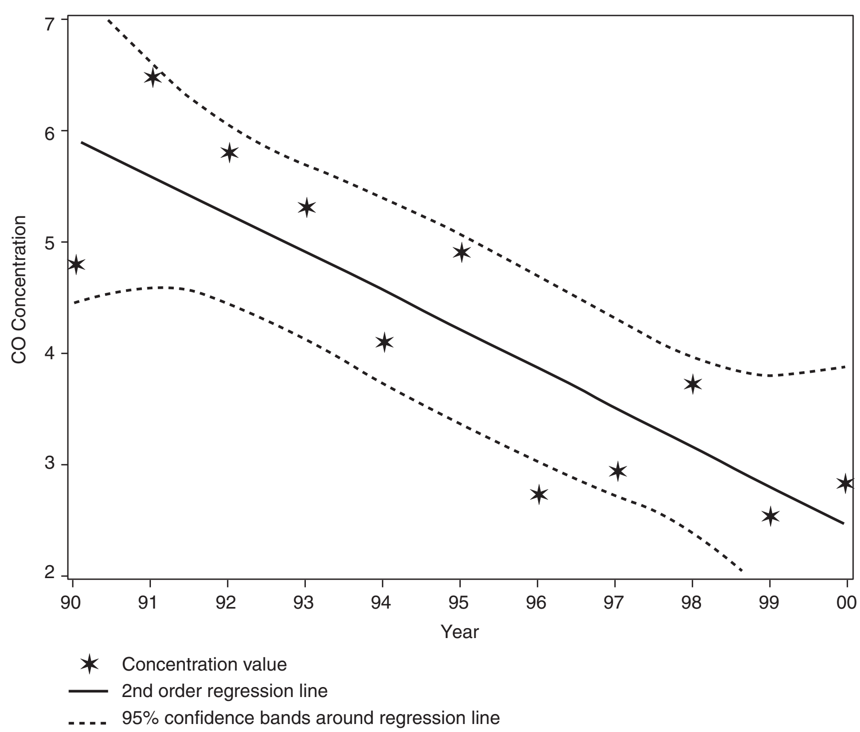


Figure 4. Example of a site with increasing trend. This site did not have data for the years 1990 through 1992.

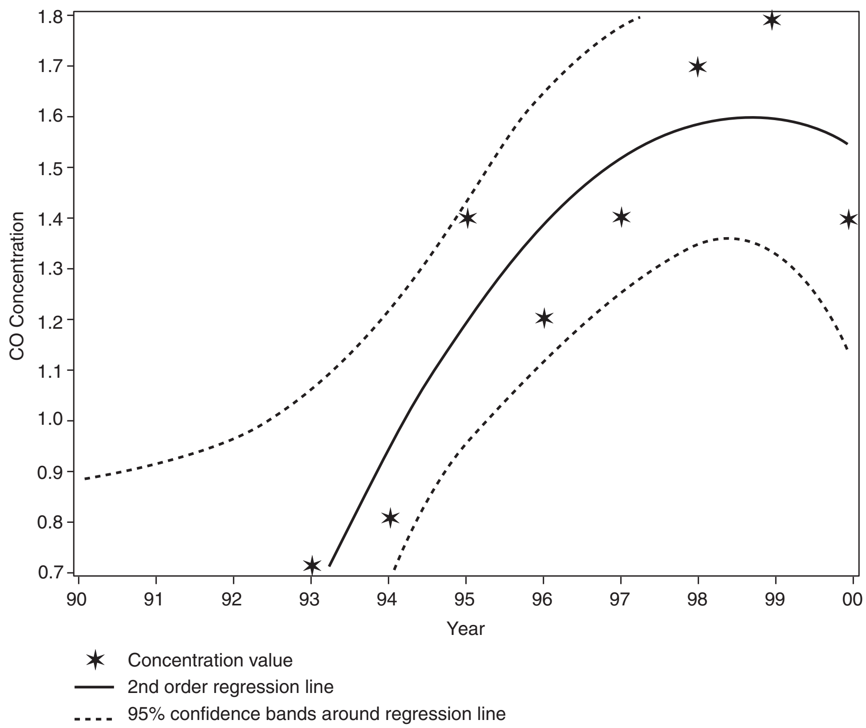
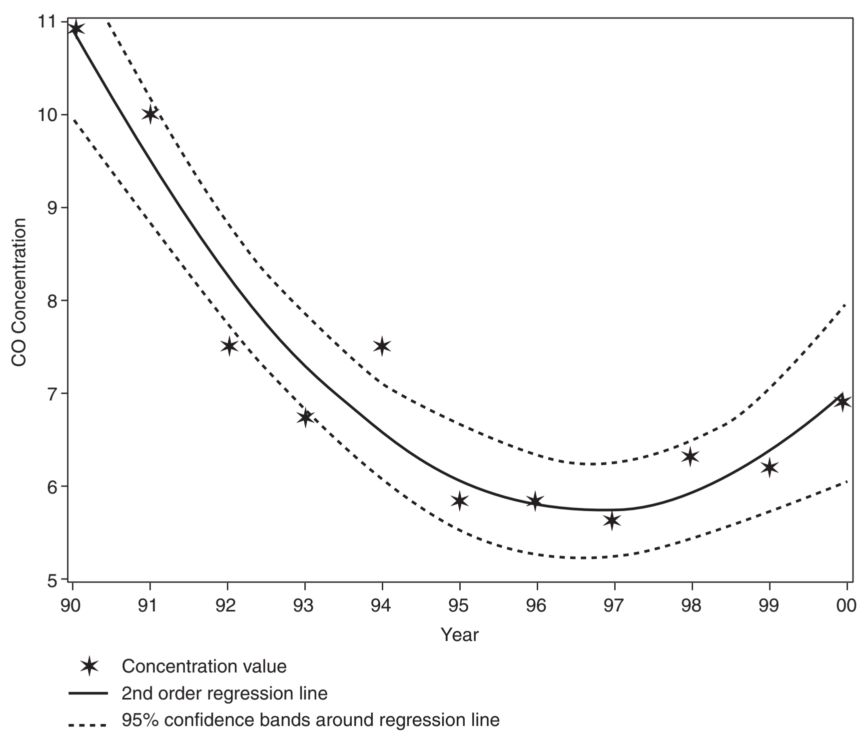


Figure 5. Example of a site with increasing trend in recent years.



The site in Figure 6 discontinued its oxyfuel requirements as of October 21, 1996; the vertical line at Year = 1996 indicates the year that this requirement ended. However, the data do not include whether the second highest concentration for 1996 occurred during or after the oxyfuel program. For this site, both the Theil test and the first-order linear regression model revealed a downward pattern at the $\alpha = 0.01$ level. However, the second-order linear regression model revealed an upward pattern at the $\alpha = 0.01$ level. The pattern of the 95% confidence limits of the second-order linear regression line indicates a high probability of nearly stable to rapidly increasing concentration.

The site in Figure 7 discontinued its oxyfuel requirements as of June 1, 1998, more recently than the site in Figure 6. Because of the increased scatter of the data around the regression line, the 95% confidence region is larger and the patterns not as statistically significant as those for the site in Figure 6. For this site, both the Theil test and the first-order linear regression model revealed a downward pattern at the $\alpha = 0.05$ level, whereas the quadratic regression model revealed an upward pattern at the $\alpha = 0.05$ level.

This study demonstrates the utility of using more than one statistical test to determine patterns in ambient concentration data. The Theil test is a nonparametric, monotonic test that measures numbers of pairs of data that increase vs. decrease. First-order linear regression examines the significance of the slope of the least-squares line through all the available data. Quadratic regression examines the significance of the coefficient of the second-order term in the least-squares regression. Although interpolation cannot be used to extrapolate beyond the range of the data, the significance of the second-order term provides a measure of the curvature (i.e., change in the trend) of the regression line. This additional information is useful in locating sites with recent increasing concentrations, even when the overall trend is downward or not significant.

Unlike the Trends Report,[2] which examines trends for regions based on MSA, this study looked for trends associated with individual monitors. Trends in more localized areas, therefore, could be discovered because areal averaging was not performed. Uncovering localized trends is important when one part of an MSA experiences rapid population growth with the associated rapid growth in vehicular emissions.

Conclusions

This analysis revealed relatively few sites with statistically significant upward trends or inflection in CO concentrations during the period 1990 to 2000. By combining regression models with the Theil test, 34 of 433 sites were identified for further analysis. Because this study demonstrated that the simpler Theil test performed nearly as well as the first-order linear regression in identifying upward linear trends, we do not recommend performing first-order linear regression on these relatively short data sets in the future. However, this study showed that the quadratic regression model successfully identifies sites where the concentration has increased in recent years, thereby identifying potential problem areas earlier than the Theil test. Because this method is to be used to identify sites of potential interest, we further recommend using $\alpha = 0.10$ and a one-sided

Figure 6. Example of a site with increasing trend in recent years. The vertical line indicates the year that the oxygenated gasoline requirement ended.

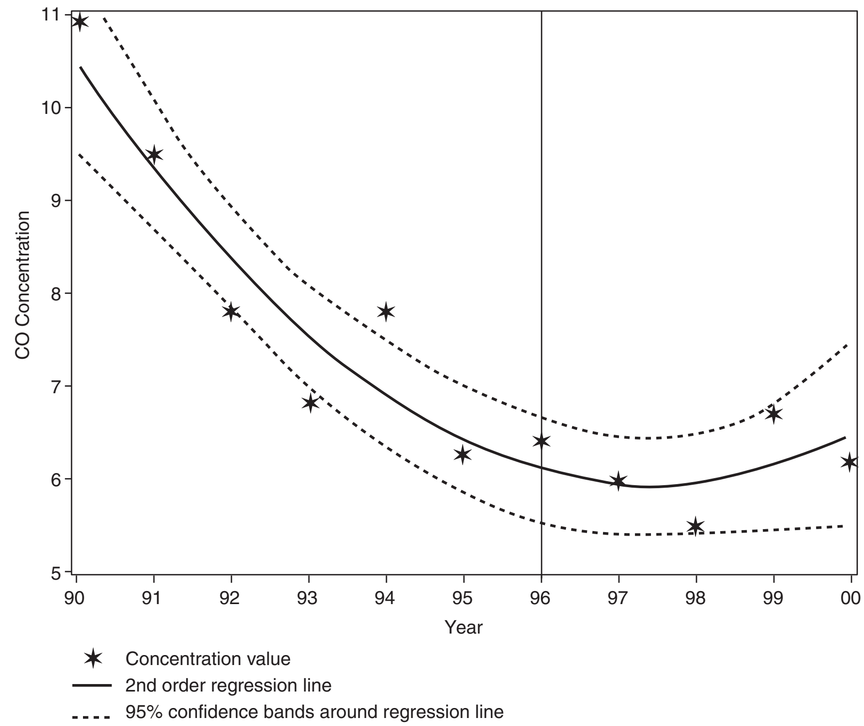
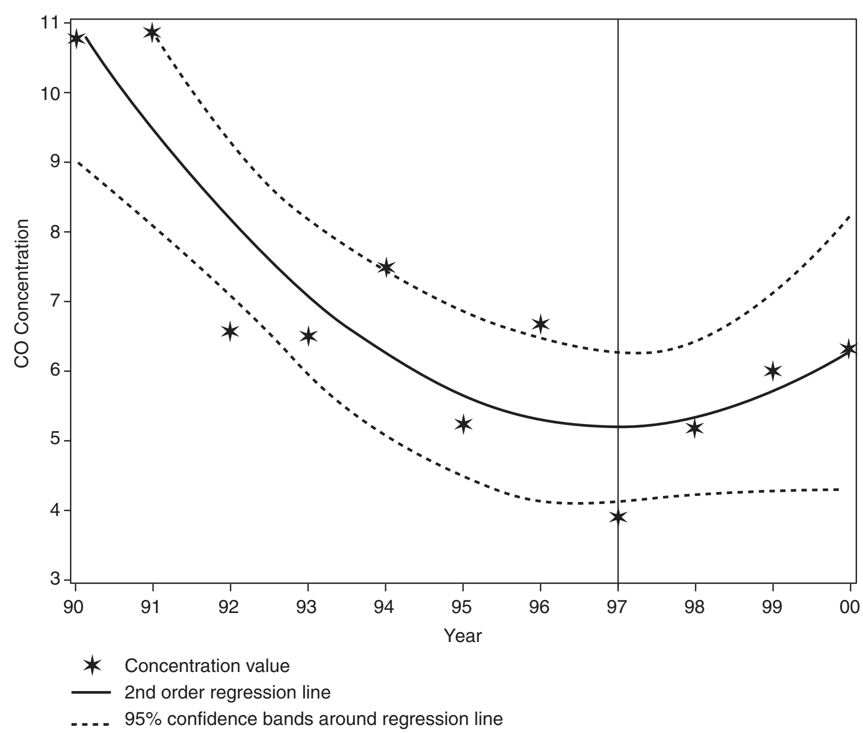


Figure 7. Example of a site with increasing trend that discontinued oxygenated gasoline requirements more recently. The trends for this site are not as significant as those shown in Figure 5.



hypothesis test to reduce the number of false negative results.

This method was designed to be an automated screening method for potential problem areas. Because both vehicle-miles traveled and the vehicle mix in fleets are changing with time, we recommend repeating this analysis annually to determine sites that warrant further analysis.

Acknowledgments

Support for this project was provided by the U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, under contract number GS-10F-0283K(SIN 899-2). We are grateful to Michael Riggs, senior statistician at RTI, for his insightful comments and technical review of this manuscript.

References

1. U.S. Environmental Protection Agency, Air and Radiation, Office of Transportation and Air Quality. *State Winter Oxygenated Fuel Program Requirements for Attainment or Maintenance of CO NAAQS*; October 2001. Available from **http://www.epa.gov/otaq/regs/fuels/oxy-area.pdf** (accessed March 2002).

2. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Emissions Monitoring and Analysis Division, Air Quality Trends Analysis Group. *National Air Quality and Emissions Trends Report*, *1999*; EPA 454/R-01-004; Research Triangle Park, NC, 2001.

3. Glen, W. Graham; Zelenka, Michael P.; Graham, Richard C. Relating Meteorological Variables and Trends in Motor Vehicle Emissions to Monthly Urban Carbon Monoxide Concentrations. *Atmospheric Environment* **1996**, *30*, 4225-4232.

4. Flaum, Jennifer B.; Rao, S. Trivikrama; Zurbenko, Igor G. Moderating the Influence of Meteorological Conditions on Ambient Ozone Concentrations. *JAWMA* **1996**, *46*, 35-46.

5. Hollander, M.; Wolfe, D.A. *Nonparametric Statistical Methods*; John Wiley & Sons: New York, 1973; pp 200–204.

6. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. *U.S. Environmental Protection Agency Intra-Agency Task Force Report on Air Quality Indicators*; EPA-450/4-81-015; Research Triangle Park, NC, February 1981.

7. *SAS/STAT Users' Guide Version 8*; SAS Institute: Cary, NC, 1999.