United States
Environmental Protection
Agency

Environmental Monitoring and Support
Laboratory
Research Triangle Park NC 27711

EPA-600/4-79-070
October 1979

Research and Development

EPA

Statistical Analysis of the Los Angeles Catalyst Study Data

RESEARCH REPORTING SERIES

Research reports of the Office of Research and Development, U S Environmental Protection Agency, have been grouped into nine series. These nine broad categories were established to facilitate further development and application of environmental technology Elimination of traditional grouping was consciously planned to foster technology transfer and a maximum interface in related fields. The nine series are

1. Environmental Health Effects Research
2. Environmental Protection Technology
3. Ecological Research
4. Environmental Monitoring
5. Socioeconomic Environmental Studies
6. Scientific and Technical Assessment Reports (STAR)
7. Interagency Energy-Environment Research and Development
8. "Special" Reports
9. Miscellaneous Reports

This report has been assigned to the ENVIRONMENTAL MONITORING series. This series describes research conducted to develop new or improved methods and instrumentation for the identification and quantification of environmental pollutants at the lowest conceivably significant concentrations. It also includes studies to determine the ambient concentrations of pollutants in the environment and/or the variance of pollutants as a function of time or meteorological factors

STATISTICAL ANALYSIS OF THE LOS ANGELES CATALYST STUDY DATA

by

Johannes Ledolter
George C. Tiao
Spencer B. Graves
Jian-tu Hsieh
Gregory B. Hudak

Department of Statistics
University of Wisconsin
Madison, Wisconsin 53706

Contract No. 68-02-2261

Harold B. Sauls
Project Officer

Analysis and Reports Branch
Environmental Monitoring Systems Laboratory
Environmental Protection Agency
Research Triangle Park, North Carolina 27711

This study was conducted
in cooperation with
The University Of Wisconsin
Madison, Wisconsin 53706

Environmental Monitoring Systems Laboratory
Office of Research and Development
U.S. Environmental Protection Agency
Research Triangle Park, North Carolina 27711

DISCLAIMER

This report has been reviewed by the Environmental Monitoring Systems Laboratory, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

FOREWORD

Measurement and monitoring research efforts are designed to anticipate potential environmental problems, to support regulatory actions by developing an in-depth understanding of the nature and processes that impact health and the ecology, to provide innovative means of monitoring compliance with regulations and to evaluate the effectiveness of health and environmental protection efforts through the monitoring of long-term trends. The Environmental Monitoring Systems Laboratory, Research Triangle Park, North Carolina, has the responsibility for: assessment of environmental monitoring technology and systems; implementation of agency-wide quality assurance programs for air pollution measurement systems; and supplying technical support to other groups in the Agency including the Office of Air, Noise and Radiation, the Office of Toxic Substances and the Office of Enforcement.

To assist in determining and evaluating effects of the catalytic converter upon roadway pollutants, the University of Wisconsin Statistics Department was contracted to conduct time-series analyses, develop appropriate models, and apply other statistical techniques to the data collected in the Los Angeles Catalyst Study. This report documents their analyses and findings.

Thomas R. Hauser, Ph.D.
Director
Environmental Monitoring Systems Laboratory
Research Triangle Park, North Carolina

ABSTRACT

This research program was initiated for the purpose of performing statistical analyses on the data from the Los Angeles Catalyst Study. The objective of the program was to determine the effects of the introduction of the catalytic converter on the atmospheric concentration levels of a number of air pollutants.

This report gives analyses of the CO, Pb, $SO_4^=$, O_3, NO, and NO_2 data covering the period from June 1974 to November 1977. Models were designed to evaluate the freeway contribution to CO and Pb as a function of traffic, windspeed and wind direction. These models were used to assess both the time trend in the pollutant measurements and the pollution concentrations at points near the freeway. In addition, frequency distributions were determined for ambient air quality data.

This report was submitted in fulfillment of Contract No. 68-02-2261 by the University of Wisconsin under the sponsorship of the U. S. Environmental Protection Agency. This report covers the period September 1977 to August 1978, and work was completed as of September 1978.

CONTENTS

FIGURES

Number		Page

TABLES

SECTION 1

INTRODUCTION

This report presents a statistical analysis of the Los Angeles Catalyst Study (LACS) data collected between June 1974 and December 1977.

The primary objective of the study was to determine the impact of catalyst equipped cars on the ambient air quality near freeways. The catalytic converter was adopted by American automobile manufacturers in order to meet federal and state emission standards. It has been used on new cars since the 1975 model year. The converter was designed to reduce the emissions of carbon monoxides (CO) and hydrocarbons (HC). Since catalyst equipped cars are required to use unleaded gasoline, the catalytic converter should have lead to a reduction in lead (Pb) emissions. However, as pointed out by Beltzer, Campion and Peterson (1974) and Bockian, Tsou, Gibbons and Reynolds (1977), the catalytic converter increased the emissions of sulfuric acid, which is measured as sulfate ion ($SO_4^=$).

DATA BASE

The San Diego Freeway in Los Angeles, California (Figure 1-1), was selected as the location for studying the environmental impact of the catalytic converter. Four air monitoring sites, two on each side of the freeway, were established by the Environmental Protection Agency (EPA). The sites were designated A, B, C, and D.

Data on a number of pollutants (including CO, Pb, $SO_4^=$, O_3, O_3, NO_2, and NO) and on meteorological variables (such as wind direction, windspeed, temperature, and relative humidity) were collected from June 1974 to December 1977. Traffic counts and traffic speeds were measured on an hourly basis beginning September 1976. In January 1977, a fifth measurement site (Site F) was added for CO concentrations at the freeway median strip.

The percentage of catalyst equipped cars passing the measurement sites was a critical parameter and has been studied by Parry, Meyer, and Rodes (1977) and Rodes and Evans (1977). It was found that starting in September 1974, the monthly increase rate for registration of catalyst equipped cars was roughly .5 percent. After adjusting for actual miles driven, it was estimated that the percentage of converter equipped cars passing the measurement sites is approximately 30 percent in January 1977 and more than 40 percent by the end of 1977.

PROCEDURES

This report presents the findings of statistical analyses on (1) the daily 4-hr $SO_4^=$ readings from two types of sampling equipment; (2) the daily 4-hr afternoon and 24-hr Pb concentrations; (3) the hourly CO measurements, and (4) the hourly NO, NO_2, and O_3 measurements. $SO_4^=$ readings were collected daily between 3 p.m. and 7 p.m. using a hi-vol (high volume) and a membrane sampler. An hi-vol sampler was also used to collect samples for Pb concentrations.

An empirical mechanistic model was designed to relate the CO concentrations at Sites A, B, C, and F to windspeed, wind direction, traffic count, traffic speed, and the distance of the measurement site from the center of the freeway. This model was used to extrapolate the CO concentrations to other sites near the highway. We also designed an empirical mechanistic model to relate the 4-hr afternoon Pb readings at Site C to windspeed, wind direction, traffic count, and traffic speed. This model was used to assess the trend in Pb emissions.

Frequency distributions for the ambient air quality data were interpreted by an alternative approach. The distribution of pollutant measurements are usually long-tailed and skewed to the right. It has been argued in the literature (Larsen [1969]) that the lognormal distribution provides a good approximation. In our analyses we used a more flexible class of distributions, which are characterized by a power transformation of the observation. Specifically, it transforms the pollutant concentrations to near normality and includes the lognormal as a special case. We considered only high values of the pollutants and used the Pareto distribution to represent the upper tail area of the distribution.

SECTION 2

CONCLUSIONS

PRINCIPAL FINDINGS

CO Findings

CO emissions decreased from 1975 to 1976, but a slight increase occurred in 1977 (traffic also increased in 1977).

The joint model relating the CO concentrations at stations next to the freeway was used to predict CO concentrations for given values of the input variables wind direction, windspeed, traffic count, and traffic speed. The proposed model was capable of predicting the CO concentrations for downwind as well as upwind sites.

Pb Findings

Very little background Pb was present in the studied area; most of the recorded Pb originated from automobile emissions on the freeway. Pb concentrations decreased until the end of 1976. In 1977, however, Pb concentrations were substantially higher. This observed increase in Pb was caused by the additional northbound traffic lane, which reduced afternoon traffic congestions and increased the traffic speed.

Afternoon Pb concentrations are higher on weekends than on weekdays. This weekday-weekend difference was attributed to afternoon weekend traffic speed being higher, resulting in increased Pb emissions.

$SO_4^=$ Findings

From June 1974 to May 1977 hi-vol and membrane $SO_4^=$ readings decreased gradually at all four measurement stations.

The available hi-vol $SO_4^=$ data for June to November 1977 showed a considerable increase. On the other hand, membrane readings increased only slightly during the same period. A comparison of membrane and hi-vol $SO_4^=$ concentrations showed that while seasonal patterns are similar, the hi-vol data were consistently higher.

From 1975 to 1977 the across-the-freeway differences (C-A) for the 4-hr membrane $SO_4^=$ readings increased slightly. Hi-vol $SO_4^=$ (C-A) differences behaved differently; from 1975 to the spring of 1977, a significant reduction was observed. In the summer and fall of 1977 the hi-vol $SO_4^=$ concentrations were slightly higher than the averages of the previous year.

O_3, NO, and NO_2 Findings

From 1976 to 1977, O_3 concentrations decreased slightly at Sites A and C From 1975 to 1977 significant increases in the NO and NO_2 concentrations at Site C and in the across-the-freeway difference (C-A) were noticed.

Analysis of High Pollutant Readings

High CO concentrations were usually recorded during traffic congestions (low traffic speed), but high Pb and $SO_4^=$ readings usually occurred at higher than average traffic speeds.

High Pb readings are associated with high $SO_4^=$ readings and vice versa.

Frequency Distributions for Pollutant Concentrations near the Freeway

The hourly CO and the 4-hr afternoon $SO_4^=$ and Pb concentrations near the freeway (Site C) did not follow a lognormal distribution. A square root transformation of the CO concentrations and a cube root transformation of the 4-hr afternoon $SO_4^=$ readings brought the distributions of the measurements to near normality. For the 4-hr afternoon Pb concentrations, no transformation appeared necessary, and the readings themselves appeared to be distributed normally.

The Pareto distribution gave a good representation for the upper tail area of the distribution of the pollutants studied.

SECTION 3

ANALYSIS OF THE CO DATA

The CO data consisted of hourly observations at Sites A and C from May 1974 to November 1977, and at Sites B and F (median strip) from February 1977 to November 1977.

MONTHLY MEANS AND PERCENTILES

As an initial step in the analysis of the CO data, the afternoon readings (3 p.m. - 7 p.m.) were used to calculate monthly averages. Extreme observations in the data were eliminated according to the following procedure: Let z_i be an observation for a particular month, $\bar{z}$ the average, and s the estimated standard deviation of the z_i's. If $|z_i-\bar{z}|>3s$, z_i will be excluded and $\bar{z}$ and s recalculated. The process was repeated until all observations fell within $\bar{z}\pm3s$.

The monthly means for Sites A and C are given in Figures 3-1a and 3-2a. To see how the observations are distributed within each month, various percentiles of the empirical frequency distribution can be studied; in particular, the 25th, 50th, 75th, and 90th percentiles can be plotted as shown in Figures 3-1b and 3-2b for Sites A and C respectively. The distance between the 25th percentile and the 75th percentile is called the midrange and provides a robust measure for the variation in the data. The 50th percentile (or median) is a measure of the center of the distribution and the 90th percentile describes the upper tail of the distribution.

From these figures we made the following observations:

- The level of CO at Site C was considerably higher than the level at Site A.
- The behavior of CO was markedly seasonal; at both Site A and Site C the CO concentrations were highest during the winter months.
- During the summer months the CO level at Site A was very low since afternoon winds blow predominantly across the freeway towards Site C. In the winter the CO levels at Site A were higher. This may have been caused by the changed meteorology; a larger percentage of winds blew across the freeway towards Site A.

To discern the yearly trend movements more accurately, the strong seasonal effects were blocked out by plotting the 25th, 50th, and 75th percentiles for each month separately. The results are shown in Figures 3-3a and 3-3b.

DIURNAL PLOTS OF CO

In Figures 3-4a and 3-4b diurnal plots of CO for the summer months (May-October) and the winter months (November-April), weekdays (Monday-Thursday) and weekends (Saturday, Sunday) are shown. Note that in all diurnal plots the value for time (t) corresponds to the average of the hour t to t+1, t = 0,1,...,23. From these four diurnal curves we observed that the behavior of CO changed from weekday to weekend, and from summer to winter. The summer weekday curve had two peaks which corresponded roughly to the morning and afternoon rush hour traffic. The summer weekend curve followed a different diurnal pattern, because of the changed weekend traffic pattern. The weekday pattern in the winter did not show a peak during the morning rush hours. This difference can be explained by the change in windspeed and wind direction from summer to winter (Tiao and Hillmer [1977]).

The diurnal plots for Site A are given in Figures 3-5a and 3-5b. During most of the day Site A acted as a background station. In the morning, however, part of the rush-hour contribution was recorded at this site because the windspeed was lower and frequently in the direction of the ocean.

FACTORS INFLUENCING CO CONCENTRATIONS

Traffic and meteorological conditions, particularly windspeed and wind direction, were the major factors which influenced CO concentrations. When studying the effects of the catalytic converter, these variables must be taken into account to properly assess the trend movements of CO. Hourly observations on wind direction and windspeed were available for June 1974 and thereafter. Reliable traffic data were available from September 1976 to December 1976, and September 1977 to November 1977.

TRAFFIC DATA

Hourly traffic counts and average traffic speeds were recorded for every lane on the San Diego Freeway near the monitoring sites. From this information we calculated hourly total traffic counts and average speeds (weighted by the number of cars). Until January 1977 the freeway had four lanes in each direction (northbound and southbound). The lanes were separated by a concrete median strip. In February 1977 a fifth northbound lane was added.

As shown in Figures 3-6a to 3-6d, hourly total traffic counts and average traffic speeds were plotted for weekdays and weekends in 1976 and 1977. Comparing Figures 3-6a and 3-6b, we noticed a difference between weekday and weekend traffic counts and traffic speeds. Weekday traffic counts were characterized by two peaks which occurred during the morning and afternoon rush hours; traffic speed was also considerably reduced during rush hours. On weekends no morning peak was observed; traffic counts peaked around noon and between 4 p.m. and 6 p.m. Two factors explained this difference in peak times. Traffic during the night hours was heavier on weekends than on weekdays, and on weekends the freeway was less congested. The average traffic speed rarely dipped below 50 mph.

A slight increase in traffic volume was noted from 1976 to 1977. A comparison of the traffic data for the two years showed that the additional northbound lane helped in reducing traffic congestion. The average weekday traffic speed for the afternoon rush-hour period increased significantly from 1976 to 1977. However, the weekend traffic speed remained largely unaffected.

Comparing the diurnal pattern of summer weekday CO concentrations with that of weekday traffic counts, Tiao and Hillmer (1977) observed that traffic counts do not reflect the magnitude of CO during the afternoon rush hours. They argued that traffic congestion during the afternoon rush hour period could be a possible reason. Thus, instead of using traffic counts they considered traffic density (TD), which is defined as the ratio of traffic count to average speed. The diurnal curves for the weekday and weekend traffic density in 1976 and 1977 are given in Figures 3-7a and 3-7b. Whereas the weekend traffic density was largely unaffected by the additional northbound lane, a rather substantial decrease in traffic density was recognized during the weekday afternoon rush hours.

When weekday and weekend traffic densitites were compared with the corresponding summer CO concentrations at Site C, the diurnal patterns looked very similar. CO concentration appeared almost linearly related to the traffic density.

In the winter, the weekday CO averages did not exhibit a peak during the morning rush hours. The meteorological variables, wind direction and windspeed, must be taken into account in order to explain this absence of a morning peak because they are important factors for diffusion and transport of the freeway CO contribution.

METEOROLOGICAL VARIABLES

Data on windspeed and wind direction were collected on an hourly basis. Wind direction (WD) was an important factor since it controlled the transport of CO. The freeway is situated 145°, therefore winds coming from 145°-325° transported the freeway contribution towards Site C, while Site A acted as a background station. When winds were from 325°-145° the freeway contribution was recorded at Site A.

Windspeed (WS) is also an important variable since it controls the diffusion of CO as well as its transport. The higher the windspeed, the more CO will be diffused.

Following Phadke, Tiao and Hillmer (1977) and Tiao and Hillmer (1977), the hourly wind vector was divided into two components, one of which was perpendicular ($WS_{\perp}$) and the other parallel ($WS_{\parallel}$) to the freeway:

$$WS_{\perp} = WS \cos(WD-235°)$$

$$WS_{||} = WS \sin(WD-235°)$$

Positive directions of these two components of the wind vector are shown in Figure 1-1.

The diurnal patterns of $WS_{\perp}$ for the winter of 1975/76 and the summer of 1976 are given in Figure 3-8. Notice that in the summer the perpendicular wind component was higher than in the winter. Also notice that in the morning during the winter months (including rush hours) winds blew away from Site C towards Site A. This change in $WS_{\perp}$ could explain the absence of the CO morning peak at Site C during the morning rush hours in the winter.

EMPIRICAL MODELS FOR CO CONCENTRATIONS AT SITES A, B, C, AND F

A new site, F, was added in February 1977 at the median of the freeway. Since that time, hourly data were collected on CO concentrations at the four locations, A, B, C, and F. The diurnal curves of CO for the summer (May-October) weekdays and weekends of 1977 are given in Figures 3-9 and 3-10.

Empirical mechanistic models were used to relate CO concentrations to traffic and meteorological variables. First, models for CO at Sites C and F were considered individually. Then a joint model was proposed which related CO concentrations at all four stations.

A MODEL FOR CO AT SITE C

Tiao and Hillmer (1977) derived a model relating the diurnal behavior of CO at Site C to traffic density and the perpendicular wind component. Their model is given by:

$$CO_t = \alpha + kTD_t e^{-b(WS_{\perp t}-\omega)^2} + a_t \qquad (3.1)$$

where CO_t, TD_t, and $WS_{\perp t}$ are the observed CO concentration, traffic density, and perpendicular wind component at hour t respectively; a_t is the error term; α is a parameter measuring the background CO; k is a parameter proportional to emissions and

$$e^{-b(WS_{\perp t}-\omega)^2}$$

is a diffusion factor involving the perpendicular wind component and two parameters b and ω.

This simple model, which requires only four parameters (α, k, b, ω) to be estimated from the data, was extensively verified on summer and winter, weekday and weekend CO averages from the LACS data base (Tiao and Hillmer [1977]). The validity of this model may be illustrated by using the most recent data on CO, traffic, and $WS_{\perp}$. Employing 1977 weekday summer CO averages at Site C and the corresponding TD and $WD_{\perp}$ values, the parameters in the model were estimated as:

May 1977-October 1977 Weekdays (M-R)

Parameter	Estimate	Standard Error
α	1.27	0.20
k	0.023	0.0014
b	0.019	0.0080
ω	3.51	0.41

The actual CO readings and the initially predicted values for CO are plotted in Figure 3-11 and listed in Table 3-1. Apart from an apparent underprediction during the period from the 8th to the 15th hours, the model (3.1) produced a close agreement between actual and predicted values.

CO trend analysis at Site C

The primary objective of this study was to assess the effect of the catalytic converter on the CO concentrations. An analysis on the CO data was done in terms of Model 3.1. We first investigated whether k, the parameter proportional to the emissions, changed over time.

TABLE 3-1. OBSERVED AND PREDICTED CO AVERAGES*

Hour	Observed CO Average	Predicted CO Average	Hour	Observed CO Average	Predicted CO Average
0	2.36	2.16	12	5.79	5.23
1	1.67	1.74	13	5.47	5.05
2	1.38	1.55	14	5.91	5.42
3	1.18	1.45	15	6.76	6.65
4	1.24	1.49	16	7.00	7.29
5	1.97	2.01	17	7.36	8.01
6	3.43	4.04	18	6.36	7.33
7	5.82	6.12	19	5.03	5.51
8	6.80	6.47	20	3.85	3.89
9	6.63	6.13	21	3.51	3.61
10	6.56	5.70	22	3.27	3.46
11	6.26	5.49	23	3.06	2.85

*Averages for Site C; 1977 Summer Weekdays (Model 3.1)

The entire time span was divided into the following seven periods: (1) June 1974 through October 1974; (2) November 1974 through April 1975; (3) May 1975 through October 1975; (4) November 1975 through April 1976;

(5) May 1976 through October 1976; (6) November 1976 through April 1977; and (7) May 1977 through October 1977. This division corresponded roughly to the "summer" and "winter" periods. To test for possible emission changes, a different emission parameter k was allowed for each period. However, the remaining parameters α, b, and ω were constrained to be the same for all seven periods. Within each 6-month period the 24-hourly averages for weekday CO and WS_1 were calculated for each of the three consecutive 2-month segments This division was chosen to allow for possible changes in the meteorology within each season. Since continuous reliable traffic data for the entire period under study were not available, the following procedure was adopted: (1) for all the 2-month periods prior to January 1977, the same average weekday traffic density figures calculated from the available traffic data for the period September 1976 through December 1976 were employed: and (2) from January/February 1977 onwards, traffic data from the period September 1977 through November 1977 were used.

The CO, WS_1 and TD were then employed to calculate estimates of the parameters α, b, ω, k_1, ..., k_7. A weighted least squares procedure was used, which allowed for the possibility of variances differing from one period to another. The parameter estimates are given in Table 3-2:

TABLE 3-2. CO TREND ASSESSMENT AT SITE C

Period	Parameter	Estimate	Std. Error
	α	1.69	0.069
	b	0.018	0.0021
	ω	2.80	0.13
Summer 74	k_1	0.0175	0.00065
Winter 74/75	k_2	0.0209	0.00070
Summer 75	k_3	0.0228	0.00066
Winter 75/76	k_4	0.0200	0.00082
Summer 76	k_5	0.0197	0.00051
Winter 76/77	k_6	0.0222	0.00091
Summer 77	k_7	0.0209	0.00066

The estimates of k_1 to k_5 were comparable to those obtained by Tiao and Hillmer (1977). Table 3-2 shows that the estimate of k decreased between winter 1974/75 and winter 1975/76, and also between summer 1975 and summer 1976. This decrease can be associated with the increase in the number of catalyst equipped cars. The increase in k from summer 1974 to summer 1975 was at first surprising, but may have been caused by the same TD data (derived from the September-December 1976 period) having been used. Specifically, a major change in the speed limit (from 70 mph to 55 mph) occurred in January 1975. Since TD is the ratio of traffic counts to average speed, the density figures used for summer 1974 were inflated and could account for the increase in k from summer 1974 to summer 1975. Comparing the estimates between 1976 and 1977, a slight increase was observed from summer 1976 to summer 1977 and from winter 1975/76 to winter 1976/77. The evidence, however, was not very strong (increase not significant at .05 level).

A MODEL FOR CO CONCENTRATIONS AT SITE F

Summer weekday data for 1977 were used to design a CO model for the median strip. As a first approximation one might expect that the CO concentrations were linearly related to the traffic densities; $CO = \alpha + kTD$. Wind, however, acts to diffuse CO emissions from automobiles. A plot of the ratio $CO_t/(\hat{\alpha}+\hat{k}TD_t)$ versus $WS_{\perp t}$ is given in Figure 3-12. CO_t, TD_t, and $WS_{\perp t}$ ($t=0, 1, \ldots, 23$) are hourly summer weekday CO averages at Site F, TD, and the perpendicular component of the wind vector, respectively. The least squares estimates of $CO_t = \alpha + kTD_t + a_t$ were $\hat{\alpha}$ and $\hat{k}$.

Figure 3-12 shows that apart from three points (which corresponded to the early morning period 2 a.m. to 5 a.m. and which are most likely influenced by additional meteorological variables) the ratio decreased slowly with increasing $|WS_{\perp}|$. Such a relationship can be approximated by the diffusion factor

$$e^{-\gamma|WS_{\perp}|^{\delta}}$$

where δ is chosen to be .5, and γ is a parameter.

These components led to the following model:

$$CO_t = (\alpha+kTD_t)e^{-\gamma|WS_{\perp t}|^{.5}}+a_t \qquad (3.2)$$

Summer (May-October 1977) weekday CO averages at Site F and corresponding weekday averages of $WS_{\perp}$ and TD were used to estimate the parameters. The results are given below:

May 1977-October 1977 Weekdays (M-R)

Parameter	Estimate	Standard Error
α	2.17	0.30
k	0.047	0.0025
γ	0.16	0.021

The actual CO readings and the predicted values are plotted in Figure 3-13 and listed in Table 3-3. It can be concluded that Model 3.2 produces an excellent fit.

A JOINT MODEL FOR CO CONCENTRATIONS AT SITES A, B, C, AND F

A joint model covering all four stations was designed in addition to the univariate models built for Sites C and F.

As shown in Figure 1-1, Site A was 150 feet and Sites B and C were both 75 feet from the median strip F. Any joint model covering different sites must include the distance from the center of the freeway as an explanatory variable. Since the wind vector controls the transport of the freeway contribution, our model had to reflect that Sites A and B were on opposite sides of the freeway from Site C. Thus, d was chosen to be an indicator variable for the distance such that $d_A = 2$; $d_B = 1$; $d_F = 0$; $d_C = -1$, where one unit corresponded to 75 feet.

The proposed joint model took the form:

$$CO_{jt} = \left\{ \alpha + kTD_t \frac{e^{-\eta|d_j|(WS_{\perp t} - \mathrm{sgn}(d_j)\sqrt{|d_j|}\omega)^2/(1+\beta|d_j|)}}{\sqrt{1+\beta|d_j|}} \right\} e^{-\gamma|WS_{\perp t}|^{.5}} + a_{jt} \qquad (3.3)$$

where

- CO_{jt} = the observed CO concentration for Site j(A,B,F or C) at hour t
- TD_t = the traffic density at hour t.
- d_j = the distance of each location from the median strip as defined above.
- $WS_{\perp t}$ = the perpendicular component of the wind vector at hour t.
- sgn = sign function; $\mathrm{sgn}(x) = \begin{cases} 1 & \text{if } x>0 \\ -1 & \text{if } x<0 \end{cases}$
- a_{jt} = error terms, and
- $\alpha, k, \eta, \beta, \omega$ and γ are unknown parameters to be estimated from the data.

TABLE 3-3. OBSERVED AND PREDICTED CO AVERAGES*

Hour	Observed CO Average	Predicted CO Average	Hour	Observed CO Average	Predicted CO Average
0	4.88	4.26	12	7.56	7.57
1	3.19	3.30	13	7.25	7.36
2	2.41	2.87	14	8.06	8.04
3	1.87	2.57	15	10.38	10.01
4	2.20	2.63	16	11.61	10.91
5	4.68	3.95	17	12.34	11.82
6	10.36	9.04	18	9.19	10.91
7	14.50	14.32	19	7.44	8.71
8	13.73	14.49	20	6.65	6.72
9	11.58	11.67	21	6.91	6.93
10	9.80	9.20	22	6.96	7.00
11	8.46	8.07	23	6.41	6.04

*Averages for Site F; 1977 summer weekdays (Model 3.2)

Model 3.3 was fitted using the 1977 summer weekday CO averages for the four sites and the corresponding $WS_{\perp}$ and TD values. The parameter estimates are given below:

Joint CO model - Summer weekday 1977

Parameter	Estimate	Standard Error
α	1.58	0.11
k	0.051	0.0015
η	0.088	0.014
β	1.29	0.19
ω	-4.12	0.32
γ	0.17	0.016

The correlation matrix between errors at Sites A, B, F, and C is given by:

$$P = \begin{bmatrix} 1.00 & .53 & .28 & .34 \\ & 1.00 & .58 & .32 \\ & & 1.00 & .25 \\ & & & 1.00 \end{bmatrix}$$

Actual CO averages and predicted values are plotted in Figures 3-14a through 3-14d and listed in Table 3-4. Model 3.3 is quite parsimonious in the parameters; apart from the covariances for the errors only six parameters have to be estimated from the data. It produced a satisfactory overall fit at all four locations.

TABLE 3-4. OBSERVED AND PREDICTED CO AVERAGES*

Hour	Site A Averages		Site B Averages		Site F Averages		Site C Averages	
	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.
0	2.01	1.73	2.53	2.34	4.88	3.87	2.36	2.25
1	1.73	1.66	1.94	1.97	3.19	2.83	1.67	1.98
2	1.50	1.63	1.66	1.82	2.41	2.35	1.39	1.82
3	1.33	1.57	1.41	1.70	1.87	2.03	1.18	1.70
4	1.34	1.57	1.47	1.71	2.20	2.11	1.24	1.72
5	2.00	1.72	2.51	2.22	4.68	3.54	1.97	2.19
6	2.94	2.30	4.81	4.20	10.36	9.04	3.43	3.97
7	3.43	2.92	6.12	6.27	14.50	14.73	5.82	5.81
8	3.25	2.66	5.05	5.82	13.73	14.92	6.80	6.39
9	2.68	1.90	3.99	3.89	11.58	11.83	6.63	6.08
10	1.97	1.38	2.63	2.31	9.80	9.17	6.56	5.83
11	1.33	1.16	1.43	1.52	8.46	7.96	6.26	5.63
12	1.01	1.10	1.13	1.31	7.56	7.43	5.79	5.23
13	0.86	1.06	0.92	1.18	7.25	7.20	5.47	4.86
14	0.76	1.04	0.83	1.14	8.06	7.91	5.91	5.07
15	0.73	1.04	0.74	1.16	10.38	9.97	6.76	6.24
16	0.59	1.05	0.62	1.21	11.61	10.92	7.00	7.01
17	0.49	1.10	0.54	1.39	12.34	11.89	7.36	8.05
18	0.64	1.19	0.64	1.73	9.19	10.96	6.36	7.59
19	0.97	1.33	0.68	2.11	7.44	8.65	5.03	5.66
20	1.52	1.49	2.21	2.35	6.65	6.53	3.85	3.91
21	1.91	1.74	2.71	2.87	6.91	6.76	3.51	3.61
22	2.31	1.86	3.09	3.08	6.96	6.83	3.27	3.50
23	2.43	1.92	3.08	2.95	6.41	5.79	3.06	3.02

*Averages for Sites A, B, F, and C; 1977 summer weekdays (Model 3.3)

Interpretation of the joint model

- If d=0 (median) Model 3.3 reduces to Model 3.2, (the univariate model for CO concentrations at Site F).

- For any fixed d≠0, the first factor in Model 3.3 corresponds to the univariate model for CO concentrations at Site C (Model 3.1). Moreover, in the joint model the background concentration α and the freeway contribution were diffused by the factor

$$e^{-\gamma |WS_{\perp}|^{.5}}.$$

- For fixed $WS_{\perp}$, CO concentrations decreased with increasing distance from the freeway. This decrease was less for the downwind sites.

- One advantage of the joint model (3.3) was it could be used to predict CO concentrations at any given distance from the freeway median.

It is worth noting that this model was different from other models considered in the literature, notably the HIWAY model (Zimmerman and Thompson [1975]). The HIWAY model can be used only to predict pollutant concentrations in the downwind direction from the source. In future research it would be worthwhile to compare the predictive accuracies of those models using the LACS data and data from other regions.

SECTION 4

ANALYSIS OF THE Pb DATA

Since catalyst equipped cars are required to use unleaded gasoline, the catalytic converter should lead to a reduction in Pb emissions. A statistical analysis of ambient Pb concentrations was conducted in order to test this assumption. The Pb data used consisted of 4-hr afternoon (3 p.m. to 7 p.m.) and 24-hr hi-vol Pb concentrations at Sites A, B, C, and D.

ANALYSIS OF MONTHLY MEANS

Figures 4-1a through 4-1d show the monthly averages of the 4-hr afternoon readings, for Sites A, B, C, and D respectively. Monthly averages of the afternoon (3 p.m.-7 p.m.) windspeed are shown in Figure 4-2a. Figure 4-2b shows the monthly relative frequency of the afternoon wind from the direction (145°-325°), i.e., when the wind was blowing across the freeway towards Site C. The following observations were recorded:

- The levels of Pb at downwind Sites C and D were substantially higher than those at upwind Sites A and B.

- In the summer months when the afternoon winds were blowing predominantly across the freeway towards Site C (Figure 4-2a and 4-2b), very little Pb was seen at Sites A and B. This absence of Pb indicated there was almost no background Pb in this area, and that most of the recorded Pb came from automobile emissions on the freeway. In the winter months when the proportion of winds from 145°-325° was smaller, some of the freeway contribution was noted at Sites A and B.

- The 4-hr Pb concentrations at Sites C and D decreased gradually until the end of 1976. The summer (May-October) means of the 4-hr afternoon readings are given in Table 4-1. At Site C the decrease from summer 1975 to summer 1976 was approximately 17 percent; at Site D the average annual decrease from 1975 to 1976 was 11 percent.

- Starting in February 1977, Pb concentrations increased sharply, especially at Site C, which was closest to the freeway on the downwind side. Table 4-1 shows that from summer 1976 to summer 1977 Pb concentrations increased 70 percent. The increase at Site D was approximately 40 percent (less data were available for Site D since hi-vol sampling for Pb at this station was discontinued in June 1977). The addition of a fifth lane (northbound) to the freeway explained this sudden increase. The opening of the extra lane increased afternoon traffic speed on lanes closest to Sites C and D (Figure 1-1). Laboratory studies (Hirschler, et al. [1957]) have shown that Pb emissions increase with traffic speed.

- The increase in Pb at Site D was smaller than the increase at Site C. This discrepancy may have occurred because most of the larger sized Pb particles, which are emitted at higher speeds, settled near the freeway.

24-hr Pb averages

The monthly means of the 24-hr Pb data at Sites C and D are plotted in Figures 4-3a and 4-3b. The summer (May-October) averages for these data are listed in Table 4-1a. Notice at both sites the 24-hr Pb concentrations decreased gradually until 1976. From 1976 to 1977, because of the change in traffic speed, the 24-hr Pb readings increased at both locations. The increase observed in the 24-hr Pb averages, however, was small compared to the increase seen in the 4-hr afternoon data, because the additional northbound lane affected the northbound traffic speed mainly in the afternoon.

Comparison of weekday and weekend afternoon Pb

To further explore the argument that higher speeds cause higher Pb emissions, weekday and weekend Pb readings were compared. Phadke, Tiao and Hillmer (1977) and Tiao and Hillmer (1977) reported that the afternoon Pb readings were about 60 percent higher on weekends than on weekdays, even though on weekends fewer cars passed the measurement stations.

This difference between weekday and weekend Pb readings was attributed to the afternoon traffic speed being higher on weekends than on weekdays, resulting in increased weekend Pb emissions. A comparison of the traffic data for 1976 and 1977 shows the average afternoon traffic speed on weekdays increased from 37 mph in 1976 to 47 mph in 1977. The speed of northbound traffic closest to Site C increased even more; from 27 mph in 1976 to 47 mph in 1977. At the same time there was very little change in the weekend speed between 1976 and 1977. Therefore, the speed difference between weekdays and weekends decreased in 1977 and thus the difference between weekday and weekend afternoon Pb should have been less prominent. This prominent reduction was confirmed by Figure 4-4a (Site C) and Figure 4-4b (Site D).

A simple model for the 4-hr afternoon Pb data at Site C

Daily afternoon (3 p.m. to 7 p.m.) averages of traffic count, traffic speed, $WS_\perp$, and Pb from the periods August 27-December 27, 1976, and September 14-November 30, 1977, were used to build an empirical model for the concentrations. Since reliable data on both traffic count and traffic speed were available for these two periods, there was no need to distinguish between weekdays and weekends.

The average of the 4-hr afternoon Pb readings for the 1976 period was 7.91 $\mu g/m^3$. For the 1977 data (after the additional northbound lane was opened) the average rose to 10.87 $\mu g/m^3$. The proposed model helped to determine whether this increase could be explained by changed meteorological and/or traffic conditions.

TABLE 4.1. OBSERVATIONS FOR SUMMER 4-HR Pb DATA

4-Hr Pb Concentrations (3 p.m. to 7 p.m.)					
Site	Year(s)	Mean	Standard Deviation	Number of Observations	t-statistic
C	1975	8.16	2.78	174	N/A*
	1976	6.77	2.25	166	N/A
	1977	11.41	2.14	92	N/A
D	1974	5.64	1.29	151	N/A
	1975	4.90	1.00	176	N/A
	1976	4.40	0.96	166	N/A
	1977	6.18	1.18	20	N/A
C	1976-1975	N/A	N/A	N/A	-5.05
	1977-1976	N/A	N/A	N/A	16.14
D	1975-1974	N/A	N/A	N/A	-5.83
	1976-1975	N/A	N/A	N/A	-4.71
	1977-1976	N/A	N/A	N/A	7.63

*Not Applicable

TABLE 4.1a. OBSERVATIONS FOR SUMMER 24-Hr Pb DATA

24-Hr Pb Concentrations					
Site	Year(s)	Mean	Standard Deviation	Number of Observations	t-statistic
C	1974	8.19	1.83	143	N/A*
	1975	7.98	1.48	173	N/A
	1976	7.00	1.60	166	N/A
	1977	7.43	1.19	74	
D	1974	5.00	1.02	46	N/A
	1975	4.19	1.01	57	N/A
	1976	3.75	0.73	57	N/A
	1977	4.16	0.70	63	N/A
C	1975-1974	N/A	N/A	N/A	-1.13
	1976-1975	N/A	N/A	N/A	-5.86
	1977-1976	N/A	N/A	N/A	2.07
D	1975-1974	N/A	N/A	N/A	-4.03
	1976-1975	N/A	N/A	N/A	-2.67
	1977-1976	N/A	N/A	N/A	3.14

*Not Applicable

As noted earlier (Hirshler, et al, [1957]), an increase in traffic speed produces an increase in Pb emissions. This relationship is illustrated in Figure 4-5a, where the ratio of afternoon Pb to the total afternoon traffic count (TC) is plotted against the average afternoon traffic speed. Averages of Pb/TC for successive traffic speed intervals were plotted instead of individual observations. This plot shows that the ratio (Pb/TC) increases proportional to traffic speed, implying a relationship of the form:

$$Pb/TC = kTS \text{ or } Pb = kTC \times TS. \qquad (4.1)$$

The wind affects Pb concentrations by means of transport and diffusion. These influences are shown in Figure 4-5b, where we plotted Pb/(TCxTS) against $WS_{\perp}$ which is to take Pb/(TCxTS) proportional to the dispersion factor

$$e^{-b(WS_{\perp}-\omega)^2} \qquad (4.2)$$

where b and ω are appropriate constants.

Combining 4.1 and 4.2 the relationship can be written as

$$Pb = kTC \times TS e^{-b(WS_{\perp}-\omega)^2} \qquad (4.3)$$

The parameters $k = k_1$, for 1976 and $k = k_2$ for 1977 were introduced to test whether a change occurred in the Pb emissions from 1976 to 1977. After having modeled the dependence of Pb on traffic and meteorological variables, it was possible to test whether the parameter k, which was proportional to the Pb emissions, changed over time.

Since a larger contribution was expected from the northbound traffic lane (because it was closest to the receptor at Site C), the effects of the northbound and the southbound traffic were separated. The next step was to design the model:

$$Pb_t = k_i(\alpha TC_t^N TS_t^N + (1-\alpha) TC_t^S TS_t^S) e^{-b(WS_{\perp t}-\omega)^2} + a_t \qquad (4.4)$$

where: Pb_t = the observed 4-hr afternoon Pb reading at Site C for day t,

TC_t^N, TC_t^S = total afternoon northbound (southbound) traffic count for day t,

TS_t^N, TS_t^S = average afternoon northbound or southbound traffic speed for day t,

$WS_{\perp t}$ = average afternoon perpendicular component of the wind vector,

a_t = error term,

k_1, k_2 = parameters proportional to the Pb emissions; $k = k_1$ for the 1976 period and $k = k_2$ for the 1977 period,

α = parameters measuring the relative contribution of the northbound and southbound lane,

b, ω = parameters in the dispersion factor involving the perpendicular wind component.

Using the set of data on Pb, traffic, and $WS_{\perp}$ described above, the parameters in Model 4.4 were estimated by nonlinear least squares. The parameter estimates and their standard errors are given below:

Parameter	Estimate	Standard Error
$10^6 k_1$ (1976)	8.77	0.39
$10^6 k_2$ (1977)	8.75	0.52
α	0.75	0.10
b	0.011	0.0035
ω	2.19	0.56

Various diagnostic checks such as residual plots failed to reveal any apparent inadequacy in the model. In particular, a plot of residuals against traffic speed showed no apparent relationship, thus indicating that the model adequately described the relationship between Pb and traffic speed.

Two important observations were made from the parameter estimates:

- The estimate for α (.75) showed that the northbound lane, which is closest to the receptor Site C, had a much larger contribution to the Pb concentrations at this site than the southbound traffic lane. That the heavier Pb particles settle fairly quickly explained the larger contribution of the northbound lane.

- The estimated emission constants k_1 and k_2 for 1976 and 1977 did not differ significantly, indicating that the 37 percent increase in the means from 1976 to 1977 was caused by the change in traffic speed.

SECTION 5

ANALYSIS OF $SO_4^=$ DATA

The catalytic converter was designed to reduce the levels of hydrocarbon and carbon monoxide emissions. Although, it successfully reduces the emissions, the catalyst also converts the sulfur in gasoline to sulfate ion ($SO_4^=$). This feature of the converter caused alarm since an increase in ambient sulfate concentrations from catalyst equipped automobiles could have harmful health effects.

The $SO_4^=$ data analyzed consisted of daily 4-hr afternoon (3 p.m. to 7 p.m.) concentrations from both the hi-vol and membrane sampler. The difference in the 4-hr concentrations between Sites C and A reflected the contribution from the freeway, since in the afternoon the wind blew predominantly in a direction perpendicular to the freeway (Evans and Rodes [1977]).

In analyzing the $SO_4^=$ data daily afternoon readings were used to calculate monthly averages at the different sites.

MONTHLY $SO_4^=$ AVERAGES FROM HI-VOL SAMPLERS

The monthly $SO_4^=$ averages from the hi-vol sampler at Sites A, B, C, and D are given by the solid lines in Figures 5-1a through 5-1d. It was noticed that $SO_4^=$ decreased gradually at all four stations until May 1977

Available data for June 1977 to November 1977 at Sites A and C showed a considerable increase in $SO_4^=$. No data were available to confirm this increase for Sites B and D.

Monthly means at Sites C and D downwind from the freeway were slightly higher than those for the background Sites A and B, indicating a contribution to $SO_4^=$ from the freeway. The contribution, however, was very small.

MONTHLY $SO_4^=$ AVERAGES FROM MEMBRANE SAMPLERS

The broken lines in Figures 5-1a and 5-1c show the monthly membrane sampler $SO_4^=$ averages from June 1974 to November 1977. A slight reduction in membrane $SO_4^=$ was noticed at both Sites A and C until May 1977. As on the hi-vol readings a slight increase occurred for the last part of the data (June-November 1977).

A comparison of membrane and hi-vol $SO_4^=$ concentrations, showed that although seasonal patterns were similar, the hi-vol data were consistently higher than the membrane readings.

AN ADDITIONAL GRAPHICAL REPRESENTATION OF THE $SO_4^=$ DATA

In Figures 5-1a through 5-1d the monthly averages of the hi-vol and membrane $SO_4^=$ are plotted. These were useful for a preliminary trend assessment. The minimum, the 25th, 50th, and 75th percentiles, and the maximum were also plotted to describe the variations of the $SO_4^=$ samples. Since every month (or quarter) did not have the same number of data points box plots (Tukey [1977], McGill, Tukey, and Larsen [1978]) were used to incorporate this information into the plots. The box plots are illustrated in Figures 5-2a and 5-2b for both the hi-vol and membrane $SO_4^=$ readings at Site C. The data are grouped into 3-month periods corresponding to the seasons: winter (December, January, February); spring (March, April May); summer (June, July, August), and fall (September, October, November). The lower line in each box corresponds to the 25th percentile (lower quartile); the broken line gives the median and the upper line gives the 75th percentile (upper quartile) of all observations for each season. The length of the box measures the interquartile range and the width of the box indicates the group size. It is chosen proportional to the square root of the number of observations in each group. This choice was made because most measures of variation (e.g. the sample standard deviation) are proportional to the square root of the group size. Figures 5-2a and 5-2b show the maximum and the minimum $SO_4^=$ readings within each season.

ACROSS-THE-FREEWAY DIFFERENCES

It was pointed out earlier that the difference between the 4-hr afternoon readings at Sites C and A can be represented as the contribution from the freeway. Monthly means and quarterly box plots for the (C-A)$SO_4^=$ differences between these sites are given in Figures 5-3a through 5-4b. A slight increase was observed in both monthly means and quarterly medians of the membrane $SO_4^=$ differences. These observations were confirmed by the means of the summer differences shown in Tables 5-1 and 5-1a. The increase in the summer differences, however, was not statistically significant. Figure 5-4b indicates that for the first year of the study little data on membrane $SO_4^=$ were available.

Hi-vol $SO_4^=$ (C-A) differences behaved quite differently from the corresponding membrane readings. From 1975 to the spring of 1977 a significant reduction in the hi-vol $SO_4^=$ differences was observed. At the same time the membrane differences increased slightly. The hi-vol $SO_4^=$ differences for the summer and fall of 1977 were slightly higher than the averages for the previous year, but were still significantly smaller than those in 1975 (Table 5-1). The discrepancy between the hi-vol and membrane readings raised an important question: Which, if either, of the two measurement methods was giving the correct pollutant reading? One possible conjecture was put forward by Tiao and Hillmer (1977). They argued (1) that the 4-hr hi-vol difference was caused by SO_2 reacting on the filter to form $SO_4^=$, and (2) that this artifact formation also occurred on the membrane filter, but to a lesser extent. Although this theory explained the observed discrepancies fairly well, further research is necessary to determine the precise nature of these two filters.

TABLE 5-1. OBSERVATIONS FOR SUMMER 4-HR $SO_4^=$ HI VOL DATA

Year(s)	Mean	Standard Deviation	Number of Observations	t-statistic
1975	4.87	3.98	164	N/A*
1976	2.53	3.86	156	N/A
1977	3.29	5.49	87	N/A
1976-1975	N/A	N/A	N/A	-5.34
1977-1976	N/A	N/A	N/A	1.26
1977-1975	N/A	N/A	N/A	-2.61

*Not Applicable

TABLE 5-1a. OBSERVATIONS FOR SUMMER 4-HR $SO_4^=$ MEMBRANE

Year(s)	Mean	Standard Deviation	Number of Observations	t-statistic
1975	.14	4.53	46	N/A
1976	.67	2.04	150	N/A
1977	.99	2.15	134	N/A
1976-1975	N/A	N/A	N/A	1.12
1977-1976	N/A	N/A	N/A	1.29
1977-1975	N/A	N/A	N/A	1.69

*Not Applicable

SECTION 6

ANALYSIS OF HIGH POLLUTANT CONCENTRATIONS

Ancillary variables such as traffic, meteorology, and the concentrations of other pollutants which accompanied high concentrations of CO, Pb, and $SO_4^=$ (hi-vol) at Site C were investigated. The upper five percent of each of the above three pollutants were identified. They were augmented by the corresponding concentrations of the remaining two pollutants, by meteorological variables (wind direction, windspeed, relative humidity and temperature), and by traffic count and traffic speed. Earlier studies of this kind using aerometric data from other regions can be found in Cleveland, Kleiner, and Warner (1976) and Tiao, Box, and Hamming (1974).

Since the Pb and $SO_4^=$ data considered were daily 4-hr afternoon readings, the accompanying meteorological and traffic variables were averaged over the 3 p.m. to 7 p.m. period. For the hourly CO data the daily maxima were considered.

Several tables which present data for the upper five percent of the highest daily CO readings are given. Table 6-1 shows a dot diagram indicating in which months the high values occurred. Dot diagrams of the corresponding Pb and $SO_4^=$ 4-hr afternoon readings in terms of their percentiles and the proportion of corresponding winds from the direction 145°-325° (blowing across the freeway towards Site C) are given in Table 6-2. Table 6-3 displays the proportions of the corresponding measurements on windspeed, relative humidity, temperature, traffic count and traffic speed which are above/below their overall average for that particular hour of the day.

CO FINDINGS

High concentrations of CO usually occurred during the fall and winter period (September-March). When traffic speed was low and winds blew across the freeway towards Site C with below average windspeeds, high concentrations of CO were also recorded.

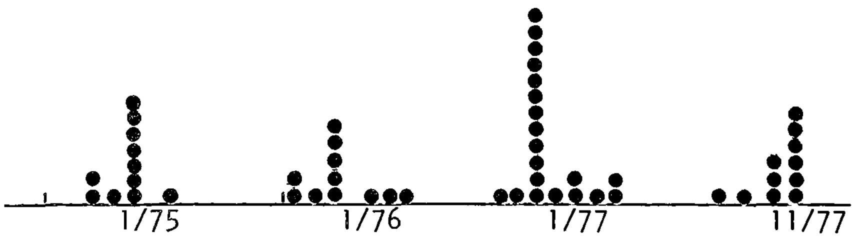
Pb FINDINGS

High Pb readings were recorded in the summer of 1975 and in 1977 when a fifth lane was added in the northbound direction. The new lane reduced traffic congestion and increased traffic speed. High Pb readings also occur when traffic speed was above average and when winds blew perpendicular to the freeway at lower than average speed. Furthermore, high Pb readings were recorded when both relative humidity and temperature rose above average High $SO_4^=$ readings were associated with the upper five percent of the Pb readings.

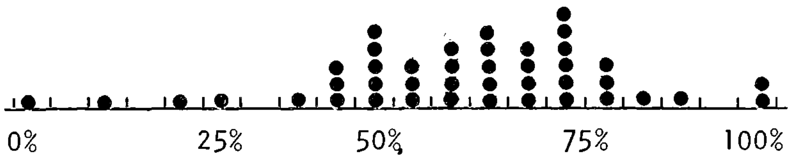
$SO_4^=$ FINDINGS

High $SO_4^=$ readings occurred throughout the year, with most of the high $SO_4^=$ concentrations recorded in 1975. On days of high traffic speed and across-the-freeway winds blowing towards Site C, high $SO_4^=$ readings were observed. Furthermore, high $SO_4^=$ values usually occurred on days with high relative humidity. A very large proportion of high $SO_4^=$ readings was accompanied by high Pb concentrations.

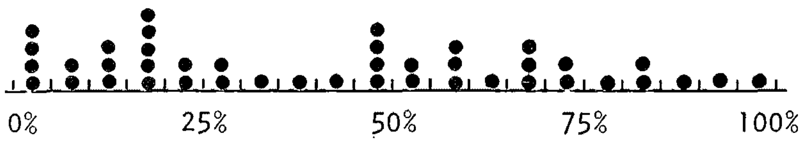
time of occurrence



corresponding
4-hour afternoon
Pb



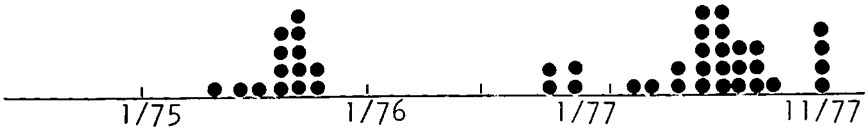
corresponding
4-hour afternoon
SO_4



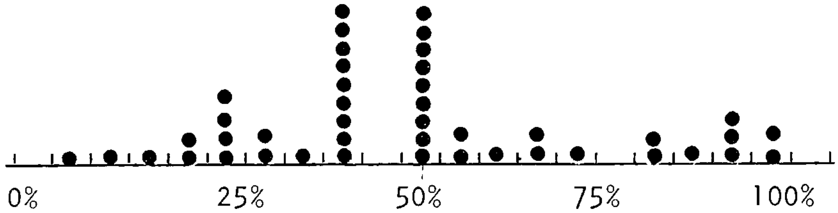
wind direction	145°-325° 77.6%	325°-145° 22.4%
	below average	above average
windspeed	85.4%	14.6%
relative humidity	56.8%	43.2%
temperature	48.8%	51.2%
traffic count	79.2%	20.8%
traffic speed	95.8%	4.2%

TABLE 6-1. DESCRIPTION OF ANCILLARY DATA ACCOMPANYING THE UPPER FIVE PERCENT OF THE CO CONCENTRATIONS.

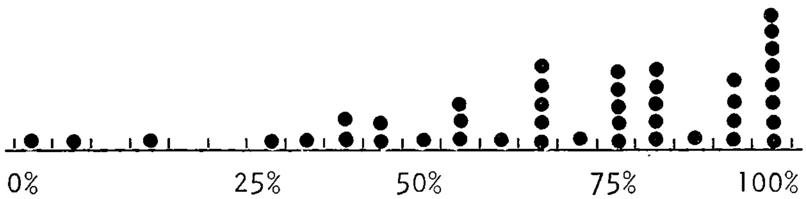
time of occurrence



corresponding
daily max.
CO



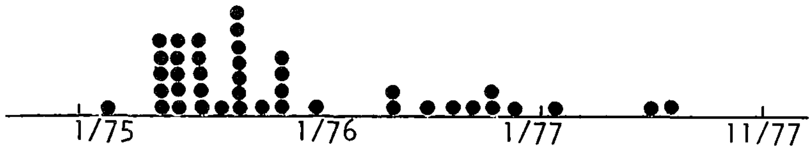
corresponding
4-hour afternoon
SO_4



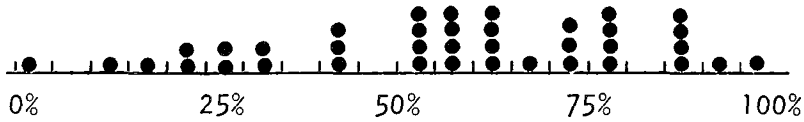
wind direction	145°-325° 92.9%	325°-145° 7.1%
	below average	above average
windspeed	67.4%	32.6%
relative humidity	22.6%	77.4%
temperature	30.0%	70.0%
traffic count	50.0%	50.0%
traffic speed	33.3%	66.7%

TABLE 6-2. DESCRIPTION OF ANCILLARY DATA ACCOMPANYING THE UPPER FIVE PERCENT OF THE Pb CONCENTRATIONS

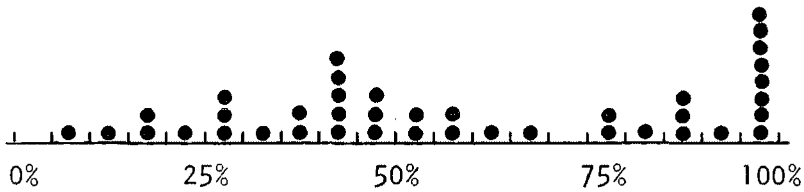
time of occurrence



corresponding daily max. CO



corresponding 4-hour afternoon SO_4



wind direction	145°-325° 94.7%	325°-145° 5.3%
	below average	above average
windspeed	45.9%	54.1%
relative humidity	13.5%	86.5%
temperature	41.7%	58.3%
traffic count	71.4%	28.6%
traffic speed	28.6%	71.4%

TABLE 6-3. DESCRIPTION OF ANCILLARY DATA ACCOMPANYING THE UPPER FIVE PERCENT OF THE $SO_4^=$ (HI-VOL) CONCENTRATIONS

SECTION 7

FREQUENCY DISTRIBUTIONS FOR HOURLY AND 4-HOURLY POLLUTANT CONCENTRATIONS

CLASS OF POWER TRANSFORMATIONS (MODEL 7.1)

In Section 3 the monthly 25th, 50th, 75th, and 90th percentiles of the hourly CO measurements were plotted. Quarterly percentiles of the 4-hr afternoon $SO_4^=$ concentrations were presented in Section 5. The distributions of the pollutant measurements appear to be long tailed and skewed to the right. Various models for such distributions have been put forward in the literature, in particular the lognormal distribution, Aitchinson and Brown (1957).

There is substantial literature, Zimmer and Larsen (1965), Larsen (1969, 1973), Michels (1971), and Kornreich (1974), which indicates that the lognormal distribution is generally useful in representing ambient air quality data. Theoretical reasons for believing so have been put forward by Singpurwalla (1972), and Kahn (1973). However, cases have been reported in the literature, for example, by Kalpasanov and Kurchatova (1976), DeNevers and Lee (1977), in which the data deviate significantly from the lognormal hypothesis.

In this section a more general class of distributions, which includes the lognormal distribution as a special case, is considered. This particular class of distributions (which are sometimes called Box-Cox transformations, due to the work by Box and Cox [1964]) arises from power transformations. Specifically, letting z_i be the original observation, it is assumed that the transformed variable x_i,

$$x_i = \frac{z_i^{\lambda}-1}{\lambda} \qquad (7.1)$$

is normally distributed with mean μ and variance σ^2.

In Model 7.1, λ is the transformation parameter. If $\lambda = 1$, the original z_i is normally distributed; if $\lambda \to 0$, $\lim \frac{z_i^{\lambda}-1}{\lambda} = \log z_i$ is normally distributed, i.e., z_i follows a lognormal distribution; if $\lambda = 1/2$, the square root of the original observation is normally distributed.

The introduction of the additional parameter λ makes it possible to check whether the lognormal distribution provides a good overall representation of the frequency distribution or whether other power transformations (such as square root or cube root) are more suitable. The distribution of the z_i can be written as:

$$f(z_i|\lambda,\mu,\sigma^2) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} z_i^{\lambda-1} \exp\{-\frac{1}{2\sigma^2}[\frac{z_i^{\lambda}-1}{\lambda} - \mu]^2\} & \text{for } \lambda \neq 0 \\ \text{or} & \\ \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{z_i} \exp\{-\frac{1}{2\sigma^2}[\log z_i - \mu]^2\} & \text{for } \lambda = 0 \end{cases} \quad (7.2)$$

Maximum likelihood estimates for the parameters λ, μ, and σ^2 can be derived and the interested reader is referred to the appendix for technical details.

CO Analysis

This approach is illustrated for the hourly CO measurements in Table 7-1a, for the 4-hr afternoon $SO_4^=$ (hi-vol and membrane) measurements and for the Pb data shown in Tables 7-2a through 7-4a (all for Site C).

The hourly CO concentrations were divided into consecutive 3-month periods to block out possible seasonal and/or trend effects. The estimates for λ are given in the first column of Table 7-1a. Notice that λ varied closely around .50, indicating that the square roots of the hourly CO concentrations followed a normal distribution. In columns 2 and 3 of Table 7-1a. estimates for μ and σ^2 are given for the various seasons by constraining λ to be the same over the entire time span. In this case the maximum likelihood estimate of λ was .50, further supporting that the square root was the appropriate transformation to consider, and that the hourly CO data at Site C did not follow a lognormal distribution.

SO_4 Analysis

In column 1 of Tables 7-2a and 7-3a estimates are given of λ for consecutive 6-month periods of the daily 4-hr afternoon hi-vol and membrane readings. Compared with the values of λ for CO in Table 7-1a the variability of λ across the periods was considerably greater for the $SO_4^=$ readings. Columns 2 and 3 present estimates of μ and σ^2 for the various periods by constraining λ to be the same for all seasons. The estimates were .40 and .35 for $SO_4^=$ hi-vol and $SO_4^=$ membrane readings respectively. These estimates indicated the cube root transformed the observations to near normality.

Pb Analysis

The same analysis is given for the Pb data at Site C. No transformation was necessary for the daily 4-hr afternoon Pb readings (overall estimate of λ was .85). The readings themselves appeared to be distributed normally.

DISTRIBUTIONS FOR THE UPPER TAIL

In studying air pollutant data, interest often centers on high readings. Although the class of power transformations may provide a good *overall* description of the frequency distribution, this class of distributions does not necessarily always give a good fit. The upper tail of the distribution is of the main interest.

Thus focus was placed on the high values of the ambient air quality data and a distribution was proposed which provided an adequate description of the upper tail probabilities. In particular, the observations exceeding some threshold value z_0 were assumed to follow a Pareto distribution

$$f(z_i|\alpha) = \frac{\alpha - 1}{z_0^{1-\alpha}} z_i^{-\alpha} \quad \text{for } z_i \geq z_0 > 0 \quad \text{and } \alpha > 1 \qquad (7.3)$$

Given z_0, maximum likelihood estimates of the parameter α were easily derived (details are given in the appendix).

In deciding on the threshold value z_0, excessive values were avoided since a sufficient number of observations have to be available for an efficient estimation of the parameter σ.

For the hourly CO data at Site C, 8.75 ppm was chosen as the value of z_0. This value was approximately the 95th percentile of all hourly CO measurements. Again the parameter α was estimated separately for consecutive 3-month periods to block out possible seasonal and/or trend variation.

The estimates of α are given in the first column of Table 7-1b. The second column compares the empirical frequency distribution with the theoretical Pareto distribution with parameter $\hat{\alpha}$. Using the null hypothesis which states that data are from a Pareto distribution, the statistic in column 2 follows approximately a Chi-square distribution with eight degrees of freedom (nine separate intervals were considered). Comparing these entries with the tabulated values from a Chi-square distribution with eight degrees of freedom ($\chi_8^2(.05) = 15.5$; $\chi_8^2(.01) = 20.1$) only one was particularly large (for September 1975 through November 1975.)

In Tables 7-2b through 7-4b estimates of α and Chi-square goodness-of-fit statistics are given for 4-hr afternoon hi-vol and membrane $SO_4^=$ and Pb. In order to obtain enough observations for estimation, the cutoff point z_0 in the Pareto distribution was chosen as the 75th percentile of all 4-hr afternoon data. The agreement with the Pareto distribution was excellent in most cases.

AN APPLICATION OF THE PARETO DISTRIBUTION TO DETERMINE TAIL PROBABILITIES

The Pareto distribution was used to assess the probability of exceeding a certain standard.

The first step was to let z_s be the air quality standard for hourly CO concentration, and the assumption that $z_s > z_0$ be the threshold value of the Pareto distribution. The probability that $z \geq z_s$ then becomes:

$$P(z \geq z_s) = P(z \geq z_0)P(z \geq z_s | z \geq z_0) \tag{7.4}$$

The first factor in Model 7.4 was determined by the choice of the threshold value z_0; the second term depended on the parameter in the Pareto distribution.

Specifically,

$$P(z \geq z_s | z \geq z_0) = \frac{z_s}{z_0}^{1-\alpha} \tag{7.5}$$

As an illustration we considered July, August, and September 1977. The estimate of α was $\hat{\alpha} = 8.30$ for this period. Furthermore, $P(z \geq 8.75 \text{ ppm})$ was equal to .0509.

Twenty ppm was assumed an important standard for hourly CO concentrations. The probability of exceeding 20 ppm was then

$$P(z \geq 20 \text{ ppm}) = P(x \geq 8.75 \text{ ppm})P(z \geq 20 \text{ ppm} | z \geq 8.75 \text{ ppm}) = .0001219.$$

We let x be the number of times 20 ppm was exceeded. For the roughly 2000 hourly CO measurements which were collected over a period of three months, the model shows:

$$P(x = 0) = 0.7837$$
$$P(x = 1) = 0.1910$$
$$P(x = 2) = 0.0233$$
$$P(x \geq 3) = 0.0020$$

The expected number of exceedances was $E(x) = .2437$.

These calculations are for illustration only. However, studies of this kind can be useful in the formulation and assessment of air quality standards.

TABLE 7-1a. ESTIMATES OF λ, μ, σ^2 FOR BOX-COX TRANSFORMATION: LOS ANGELES HOURLY CO DATA, SITE C

Time Period	$\hat{\lambda}$	$\hat{\lambda}$ = .50	
		$\hat{\mu}$	$\hat{\sigma}^2$
June, July, August (1974)	0.65	1.86	1.56
September, October, November (1974)	0.50	2.09	1.85
December, January, February (1975)	0.35	1.70	3.70
March, April, May (1975)	0.55	1.94	1.36
June, July, August (1974)	0.60	2.22	1.23
September, October, November (1975)	0.45	2.57	1.82
December, January, February (1976)	0.45	2.06	2.30
March, April, May (1976)	0.55	1.71	1.51
June, July, August (1976)	0.45	2.07	1.15
September, October, November (1976)	0.50	2.13	2.36
December, January, February (1977)	0.55	2.03	2.33
March, April, May (1977)	0.55	1.73	1.53
June, July, August (1977)	0.60	2.18	1.42

TABLE 7-1b. ESTIMATES OF α FOR PARETO DISTRIBUTION WITH z_0 = 8.75: LOS ANGELES HOURLY CO DATA, SITE C

Time Period	$\hat{\alpha}$	χ^2_8
June, July, August (1974)	10.33	2.04
September, October, November (1974)	6.29	20.06
December, January, February (1975)	4.84	16.42
March, April, May (1975)	8.69	17.31
June, July, August (1974)	10.18	9.15
September, October, November (1975)	5.53	29.83
December, January, February (1976)	6.16	19.31
March, April, May (1976)	7.35	9.26
June, July, August (1976)	10.77	1.34
September, October, November (1976)	5.23	17.28
December, January, February (1977)	5.84	8.14
March, April, May (1977)	8.35	9.64
June, July, August (1977)	8.30	2.43

TABLE 7-2a. ESTIMATES OF λ, μ, σ^2 FOR BOX-COX TRANSFORMATION: LOS ANGELES 4-HR AFTERNOON $SO_4^=$ DATA (HI-VOL), SITE C

Time Period	$\hat{\lambda}$	$\hat{\lambda}$ = .40	
		$\hat{\mu}$	$\hat{\sigma}^2$
November 1974 - April 1975	0.75	4.81	1.76
May 1975 October 1975	0.15	6.21	1.52
November 1975 - April 1976	-0.10	5.04	1.73
May 1976 October 1976	-0.10	5.19	1.27
November 1976 April 1977	0.45	3.83	2.44
May 1977 - September 1977	0.75	5.41	1.67

TABLE 7-2b. ESTIMATES OF α FOR PARETO DISTRIBUTION WITH z_0 AS THE 75th PERCENTILE: LOS ANGELES 4-HR AFTERNOON $SO_4^=$ DATA (HI-VOL), SITE C

Time Period	$\hat{\alpha}$	χ^2_8
November 1974 - April 1975	7.40	13.20
May 1975 - October 1975	4.54	7.07
November 1975 - April 1976	4.78	8.82
May 1976 - October 1976	4.54	16.64
November 1976 - April 1977	4.84	9.89
May 1977 - September 1977	6.41	6.31

TABLE 7-3a. ESTIMATES OF λ, μ, σ^2 FOR BOX-COX TRANSFORMATION: LOS ANGELES 4-HR AFTERNOON $SO_4^=$ DATA (MEMBRANE), SITE C.

Time Period	$\hat{\lambda}$	$\hat{\lambda}$ = .35	
		$\hat{\mu}$	$\hat{\sigma}^2$
November 1974 - April 1975	-0.70	3.94	2.41
May 1975 - October 1975	-0.15	4.58	1.75
November 1975 - April 1976	0.25	2.49	1.58
May 1976 - October 1976	0.35	3.55	2.48
November 1976 - April 1977	0.30	2.52	2.56
May 1977 - September 1977	0.55	3.52	2.01

TABLE 7-3b. ESTIMATES OF α FOR PARETO DISTRIBUTION WITH z_0 AS THE 75th PERCENTILE: LOS ANGELES 4-HR AFTERNOON $SO_4^=$ DATA (HI-VOL), SITE C

Time Period	$\hat{\alpha}$	χ_8^2
November 1974 - April 1975	3.06	1.73
May 1975 - October 1975	3.19	5.88
November 1975 - April 1976	too few data points above overall 75th percentile	
May 1976 - October 1976	3.61	5.80
November 1976 - April 1977	4.16	16.45
May 1977 - September 1977	4.47	12.44

TABLE 7-4a. ESTIMATES OF λ, μ, σ^2 FOR BOX-COX TRANSFORMATION: LOS ANGELES 4-HR AFTERNOON Pb DATA, SITE C

Time Period	$\hat{\lambda}$	$\hat{\lambda}$ = .85	
		$\hat{\mu}$	$\hat{\sigma}^2$
November 1975 April 1975	1.00	5.00	4.50
May 1975 - October 1975	0.50	5.78	4.04
November 1975 - April 1976	1.15	4.30	2.81
May 1976 - October 1976	0.00	4.76	2.78
November 1976 April 1977	0.85	6.21	6.31
May 1977 - September 1977	1.00	7.90	2.84

TABLE 7-4b. ESTIMATES OF α FOR PARETO DISTRIBUTION WITH z_0 AS THE 75th OVERALL PERCENTILE: LOS ANGELES 4-HR AFTERNOON Pb DATA, SITE C.

Time Period	α	χ^2_8
November 1974 - April 1975	8.81	7.31
May 1975 - October 1975	5.85	14.29
November 1975 April 1976	11.00	1.75
May 1976 - October 1976	10.88	10.82
November 1976 - April 1977	8.41	15.98
May 1977 - September 1977	6.04	24.07

SECTION 8

ANALYSIS OF O_3, NO, AND NO_2 DATA

In the LACS program, chemiluminescence instruments for monitoring nitric oxide (NO) and nitrogen dioxide (NO_2) were added in January 1975 and instruments for measuring ozone (O_3) were added in January 1976. Hourly observations for these three pollutants were made at Sites A and C.

A preliminary analysis of the 4-hr afternoon (3 p.m. to 7 p.m.) O_3, NO, and NO_2 measurements was conducted. The afternoon period was of particular interest for trend assessment since during the afternoon hours the wind usually blew roughly perpendicular to the freeway. Thus the difference in the readings at Sites C and A should have reflected the contribution from the freeway traffic.

MONTHLY MEANS OF 4-HR AFTERNOON O_3, NO, AND NO_2

The monthly averages of the 4-hr afternoon readings of O_3, NO, and NO_2 at Sites A and C as well as their across-the-freeway difference (C-A) are plotted in Figures 8-1a through 8-3c. From these plots we made the following observations:

- O_3: The level of O_3 at the background site (A) was considerably higher than O_3 at Site C. Since the NO emitted from the automobiles reacted very rapidly with the available O_3 to form NO_2, the background O_3 was essentially consumed before it reached Site C.

  O_3 was highly seasonal. High concentrations occurred in the summer months because of increased solar radiation and strong and persisent night and daytime inversions (Tiao, Box, and Hamming [1975]), and Tiao, Phadke, and Box [1976]).

  At both Sites A and C, the concentrations of O_3 decreased from 1976 to 1977; a slight reduction also occurred in the across-the-freeway difference. This observation was confirmed by the sample means and standard deviations of (C-A) for the periods May-October 1976 and May-October 1977 (Table 8-1a).

- NO: The NO concentration at Site A practically vanished during the summer months when winds blew from the ocean towards Site C, indicating there was very low NO background in this area. During the winter months we observed some NO at Site A since in the winter period winds occasionally blew towards the sea.

While Figure 8-2a through 8-2c exhibits very little change in the background level, it does show a significant increase from 1975 to 1977 in the NO concentrations at Site C and in the across-the-freeway difference (C-A). Means and standard deviations for May-October 1975, 1976, and 1977 are given in Table 8-1.

- $\underline{NO_2}$: The monthly means at Site C were considerably higher than those at Site A, indicating a substantial contribution to NO_2 from the freeway. Although NO_2 is not considered a primary pollutant from the automobile, it quickly formed between Sites A and C through the reaction of NO with the available O_3.

As in the NO data, very few yearly changes were noticed in NO_2 at the background station. On the other hand, a significant increase from 1975 to 1977 in the concentrations at Site C and in the across-the-freeway differences was observed. The averages and standard deviations of (C-A) for the period May-October are given in Table 8-1.

TABLE 8-1. OBSERVATIONS FOR SUMMER ACROSS-THE-FREEWAY DIFFERENCES FOR POLLUTANTS

Pollutant	Year(s)	Mean	Standard Deviation	Number of Observations	t-statistic
O_3	1976	-5.39	3.43	493	N/A
	1977	-6.29	3.07	676	N/A
	1977-1976	N/A	N/A	N/A	-4.71
NO	1975	0.273	0.192	560	N/A
	1976	0.314	0.109	576	N/A
	1977	0.435	0.102	668	N/A
	1976-1975	N/A	N/A	N/A	4.44
	1977-1976	N/A	N/A	N/A	20.21
	1977-1975	N/A	N/A	N/A	18.86
NO_2	1975	0.0530	0.0241	596	N/A
	1976	0.0623	0.0358	556	N/A
	1977	0.0728	0.0308	688	N/A
	1976-1975	N/A	N/A	N/A	5.20
	1977-1976	N/A	N/A	N/A	5.60
	1977-1975	N/A	N/A	N/A	12.69

DIURNAL DIAGRAMS FOR O_3, NO, and NO_2

Monthly means of the afternoon (3 p.m. to 7 p.m.) period (Figures 8-1a through 8-3c) were used to assess the overall trend in O_3, NO, and NO_2. Summer (May-October), weekday (Monday-Thursday) 1977 diurnal plots of these three pollutants at Sites A and C are given in Figures 8-4a through 8-4c. From these plots the following observations were made:

- Ozone at the background station peaks around 2 p.m. in the afternoon. As the wind transports ozone across the freeway it is consumed by NO to form NO_2. This reaction is very fast and, thus, only low O_3 concentrations are recorded at Site C.

- NO concentrations at Site A are essentially zero from 10 a.m. through 8 p.m. since the wind is transporting the NO emissions from the automobiles away from Site A towards Site C. Since part of NO reacts with O_3, NO at Site C peaks after the ozone is depleted. There is apparently not enough O_3 at the freeway to convert all of the NO automobile emissions to NO_2.

- NO_2 concentrations at Site C follow a pattern similar to that of ozone at Site A. The formation of NO_2 at Site C reaches its peak at the same time when O_3 peaks at Site A and gradually diminishes with the depletion of ozone.

APPENDIX A

MAXIMUM LIKELIHOOD ESTIMATES OF λ, μ, AND σ^2 FOR THE DISTRIBUTIONS INVOLVING POWER TRANSFORMATIONS

It is assumed that the transformed variables $x_i = \frac{z_i^{\lambda}-1}{\lambda}$ $(1 \leq i \leq n)$ are *i.i.d. normal* (μ, σ^2) *where* z_1, z_2, ..., z_n *are hourly pollution measurements.* Therefore:

$$f(x_1 x_2, \ldots, x_n | \lambda, \mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\} \qquad (A.1)$$

Making the transformation to the z_i, we get

$$f(z_1, z_2, \ldots, z_n | \lambda, \mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} (\prod_{i=1}^{n} z_i^{\lambda-1}) \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\}$$

(A.2)

The log likelihood function is

$$\ell(\lambda, \mu, \sigma^2 | z_1, z_2, \ldots, z_n) = k - n\log + (\lambda-1) \sum_{i=1}^{n} \log z_i \quad \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

where k is a constant. (A.3)

For given λ, the maximum likelihood estimates of μ and σ^2 are given by:

$$\hat{\mu}(\lambda) = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad (A.4)$$

$$\sigma^2(\lambda) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu}(\lambda))^2$$

and the concentrated log likelihood function of λ is

$$\ell_c(\lambda | z_1, \ldots, z_n) = k - n\log\hat{\sigma}(\lambda) + (\lambda-1) \sum_{i=1}^{n} \log z_i . \qquad (A.5)$$

The ML estimate of λ can be derived by evaluating (A.5) for various values of λ.

MAXIMUM LIKELIHOOD ESTIMATE OF THE PARAMETER α IN THE PARETO DISTRIBUTION

Suppose we have n independent observations $z_1,\ldots,z_n$ from the Pareto distribution with density

$$f(z_i|\alpha) = \frac{\alpha-1}{z_0^{1-\alpha}} z_i^{-\alpha}, \quad z_i \geq z_0 > 0 \tag{A.6}$$

The joint probability density of $z_1,\ldots,z_n$ is

$$f(z_1,z_2,\ldots,z_n|\alpha) = \left[\frac{\alpha-1}{z_0^{1-\alpha}}\right]^n \prod_{i=1}^{n} z_i^{-\alpha}, \quad z_i \geq z_0 \text{ for all } 1 \leq i \leq n \tag{A.7}$$

The log likelihood function is given by

$$\ell(a|z_1,\ldots,z_n) = n\log(\alpha-1)-n(1-\alpha)\log z_0 - \alpha\sum_{i=1}^{n}\log z_i \tag{A.8}$$

Thus,

$$\frac{d\ell}{d\alpha} = \frac{n}{\alpha-1} + n\log z_0 - \sum_{i=1}^{n}\log z_i \tag{A.9}$$

The ML estimate of α, which is derived by setting the derivative equal to zero, is given by

$$\hat{\alpha} = 1 + \frac{1}{\log\dot{r}} \tag{A.10}$$

where $\dot{r} = \left[\prod_{i=1}^{n} \frac{z_i}{z_0}\right]^{\frac{1}{n}}$.

REFERENCES

Aitchison, J., and J.S.C. Brown. 1957. The Lognormal Distribution. Cambridge University Press, New York, New York.

Beltzer, M., R. J. Campion, and W. L. Peterson. 1974. Measurement of Vehicle Particulate Emissions. Society of Automotive Engineers Paper 740286, February 1974.

Bockian, A. H., G. Tsou, D. Gibbons, and R. Reynolds. 1977. Sulfate Concentrations at Two Los Angeles Freeways. The Los Angeles Catalyst Study Symposium. EPA-600/4-77-034, U.S. Environmental Protection Agency, Washington, D. C., p. 281.

Box, G.E.P., and D. R. Cox. 1964. An Analysis of Transformations. J. Royal Stat. Soc., Series B, 26: 211.

Cleveland, W. S., B. Kleiner, and J. L. Warner. 1976. Robust Statistical Methods and Photochemical Air Pollution Data. J. Air Pollut. Control Assoc., 26:36.

DeNevers, N., and K. W. Lee. 1977. Extreme Values in TSP Distribution Functions. J. Air Pollut. Control Assoc., 27:995.

Hirschler, D. A., L. F. Gilbert, F. W. Lamb, and L. M. Niebylski. 1957. Particulate Lead Compounds in Automobile Exhaust Gas. Ind. Eng. Chem., 49:1131.

Kahn, H. 1973. Distribution of Air Pollutants. J. Air Pollut. Control Assoc., 23:973.

Kalpasanov, Y., and G. Kurchatova. 1976. A Study of the Statistical Distribution of Chemical Pollutants in Air. J. Air Pollut. Control Assoc., 26:981

Kornreich, L. D. ed. 1974. Proceedings of the Symposium on Statistical Aspects of Air Quality Data. EPA-650/4-74-038, U.S. Environmental Protection Agency, Washington, D.C.

Larsen, R. I. 1969. A New Mathematical Model of Air Pollutant Concentration Averaging Time and Frequency. J. Air Pollut. Control Assoc., 19:24.

Larsen, R. I. 1973. An Air Quality Data Analysis System for Interrelating Effects, Standards and Needed Source Reductions. J. Air Pollut. Control Assoc., 23:993.

Michels, D. E. 1971. Lognormal Distribution for Plutonium in the Outdoors. In: Proceedings of Environmental Plutonium Symposium. E. B. Fowler, R. W. Henderson, and M. F. Milligan, eds., Rept. LA-4756, Los Alamos Scientific Laboratory.

McGill, R., J. W. Tukey, and W. A. Larson. 1978. Variations of Box Plots. The Am. Statist. 32:12.

Parry, E. P., R. A. Meyer, and C. E. Rodes. 1977. Determination of Percentage of Diesel Trucks and Catalyst Equipped Cars. In: The Los Angeles Catalyst Study Symposium, EPA-600/4-77-034, U.S. Environmental Protection Agency, Washington, D. C., p. 147.

Phadke, M. S., G. C. Tiao, and S. C. Hillmer. 1977. Statistical Evaluation of the Environmental Impact of the Catalytic Converter. Paper presented at the 1977 Annual Meeting of the Air Pollution Control Association, Toronto, Canada, June 1977.

Rodes, C. E., and G. F. Evans. 1977. Summary of LACS Integrated Pollutant Data. In: The Los Angeles Catalyst Study Symposium, EPA-600/4-77-034, U. S. Environmental Protection Agency, Washington, D. C., p.301.

Singpurwalla, N. D. 1972. Extreme Values from a Lognormal Law with Applications to Air Pollution Problems. Technometrics, 14:703.

Tiao, G. C., G.E.P. Box, and W. J. Hamming. 1974. A Statistical Analysis of the Los Angeles Ambient Carbon Monoxide Data. Technical Report No. 265 Department of Statistics, University of Wisconsin, Madison, Wisconsin.

Tiao, G. C., G.E.P. Box, and W. J. Hamming. 1975. Analysis of the Los Angeles Photochemical Smog Data: A Statistical Overview. J. Air Pollut. Control Assoc., 25:260.

Tiao, G. C., and S. C. Hillmer. 1977. Statistical Analysis of the Los Angeles Catalyst Study Data - Rationale and Findings. In: The Los Angeles Catalyst Study Symposium. EPA-600/4-77-034, U.S. Environmental Protection Agency, Washington, D. C. p. 415. Also in Environ. Science and Technol., 12:820, July 1978.

Tiao, G. C., M. S. Phadke, and G.E.P. Box. 1976. Some Empirical Models for the Los Angeles Photochemical Smog Data. J. Air Pollut. Control Assoc. 26:485.

Tukey, J. W. 1977. Exploratory Data Analysis. Addison-Wesley, Reading, Massachusetts.

Zimmer, C. E., and R. I. Larsen. 1965. Calculating Air Quality and its Control. J. Air Pollut. Control Assoc., 15:565.

Zimmerman, J. R., and R. S. Thompson. 1975. Users Guide for HIWAY, a Highway Pollution Model. USEPA, February 1975.

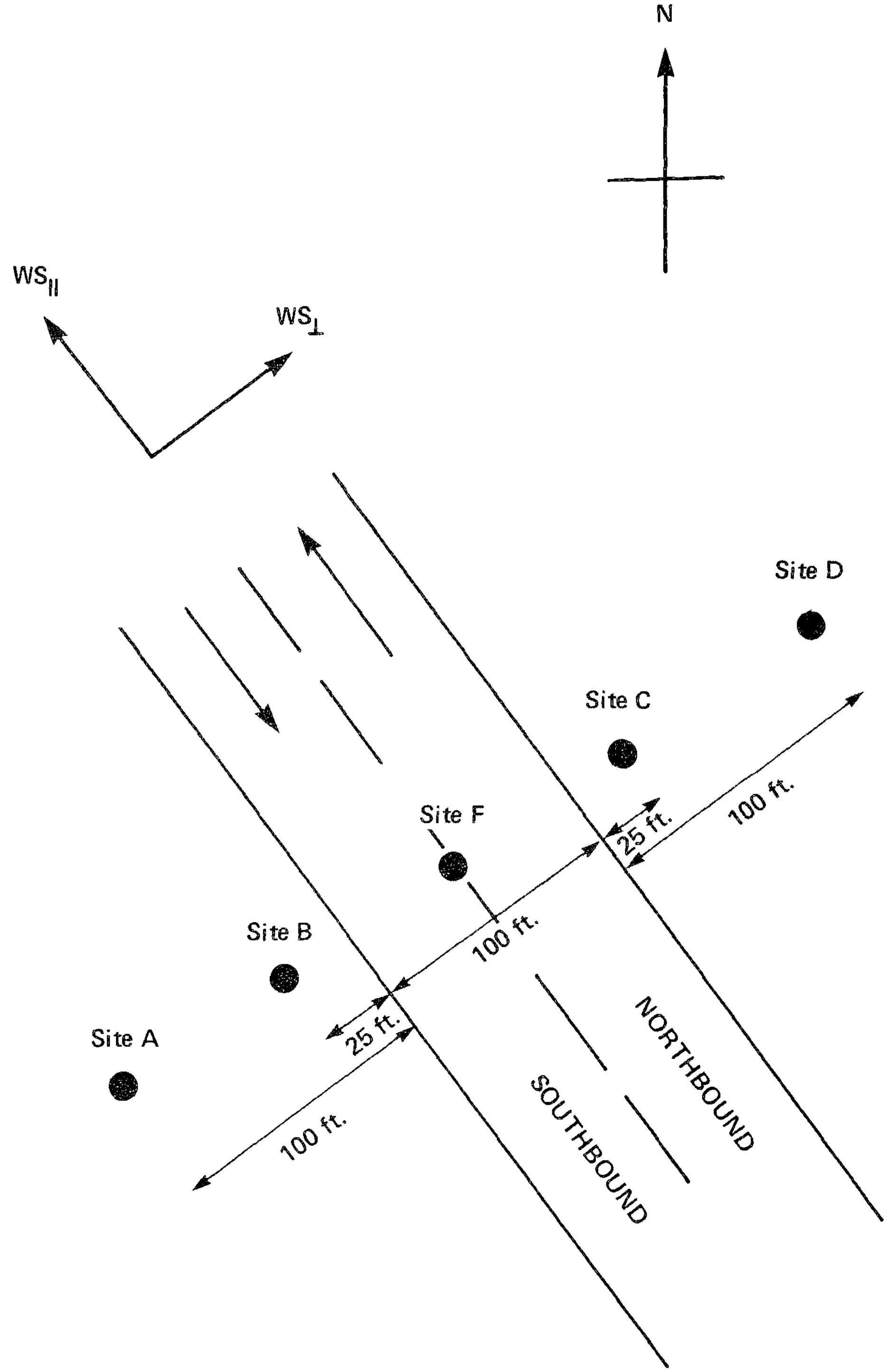


Figure 1-1. Map of Los Angeles air monitoring stations at San Diego Freeway.

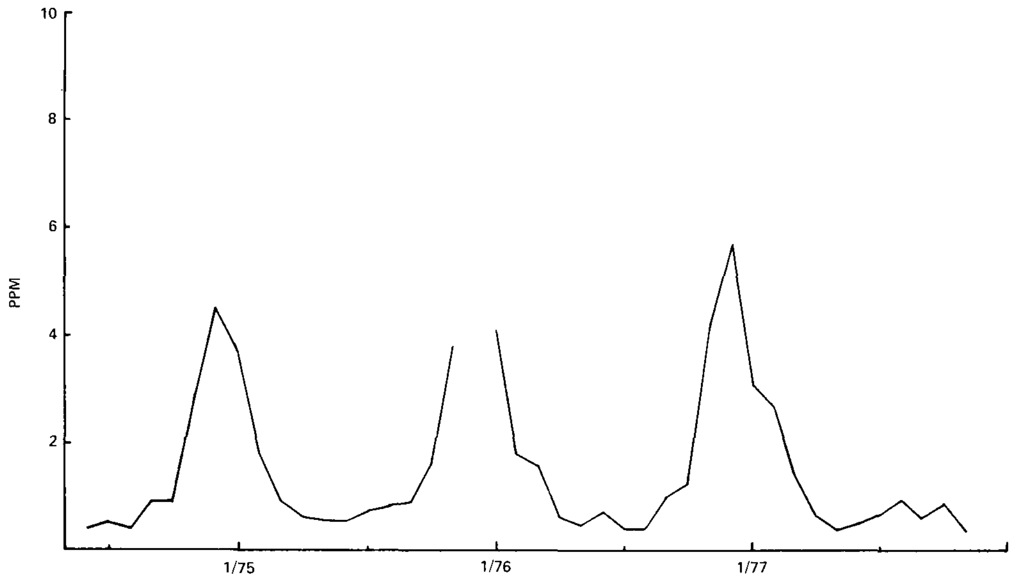


Figure 3-1a. Monthly means of 4-hour afternoon (3 p.m. to 7 p.m.) CO concentrations, Site A.

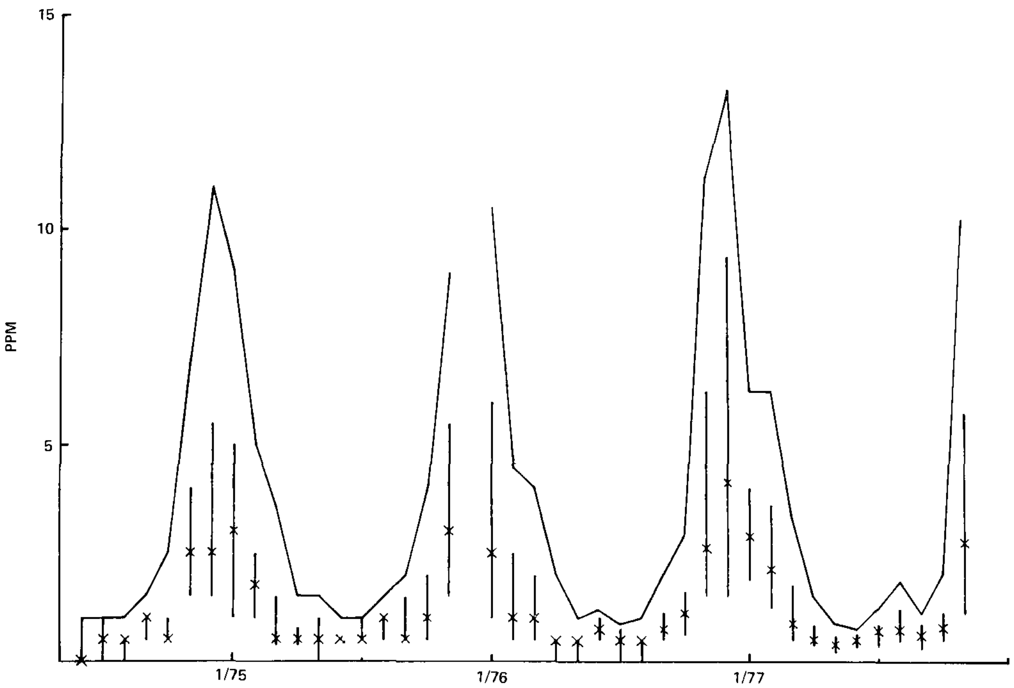


Figure 3-1b. Monthly 25th, 50th, 75th, and 90th percentiles of 4-hour afternoon (3 p.m. to 7 p.m.) CO concentrations, Site A.

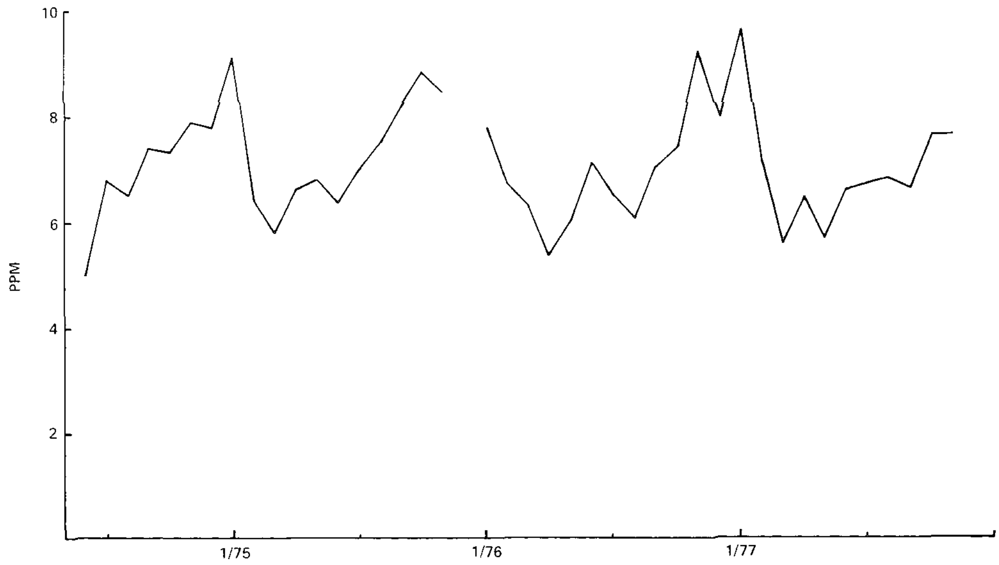


Figure 3-2a. Monthly means of 4-hour afternoon (3 p.m. to 7 p.m.) CO concentrations, Site C.

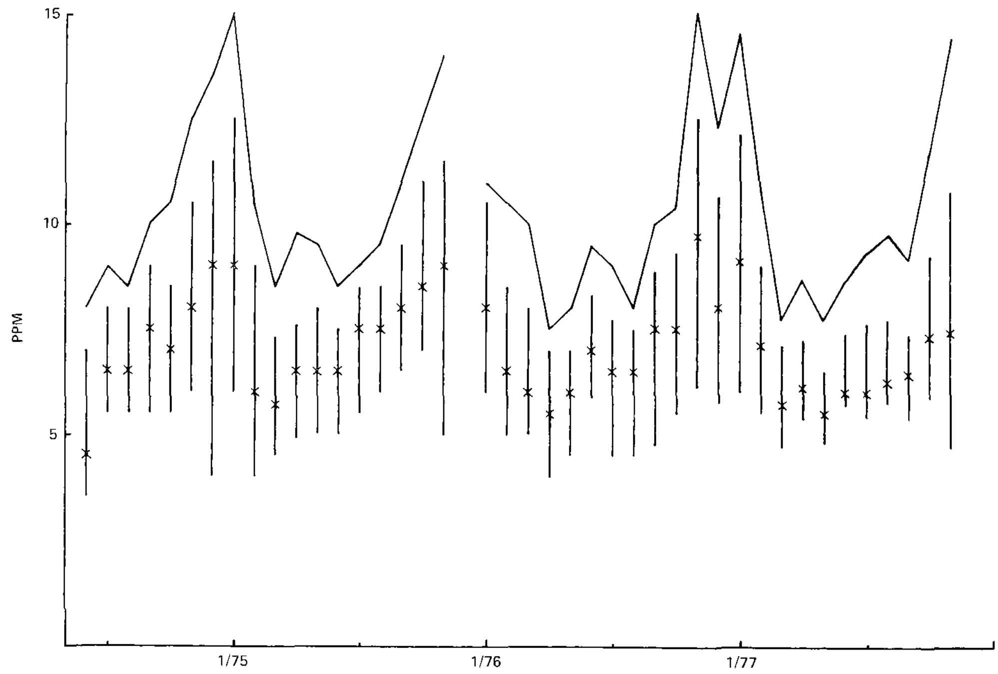


Figure 3-2b. Monthly 25th, 50th, 75th, and 90th percentiles of 4-hour afternoon (3 p.m. to 7 p.m.) CO concentrations, Site C.

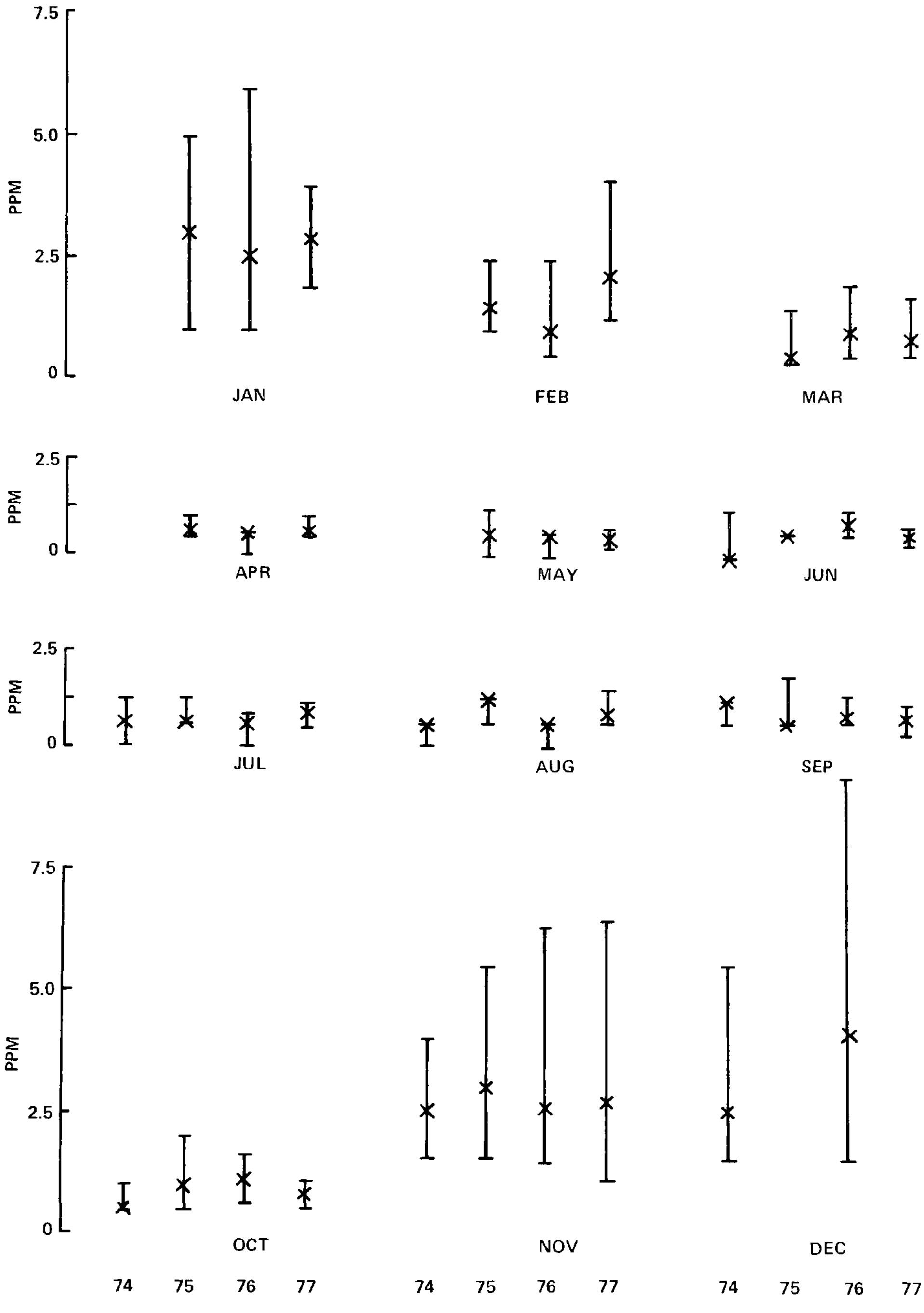


Figure 3-3a. Plot of 25th, 50th, and 75th percentiles of 4-hour afternoon (3 p.m. to 7 p m.) CO concentrations, Site A.

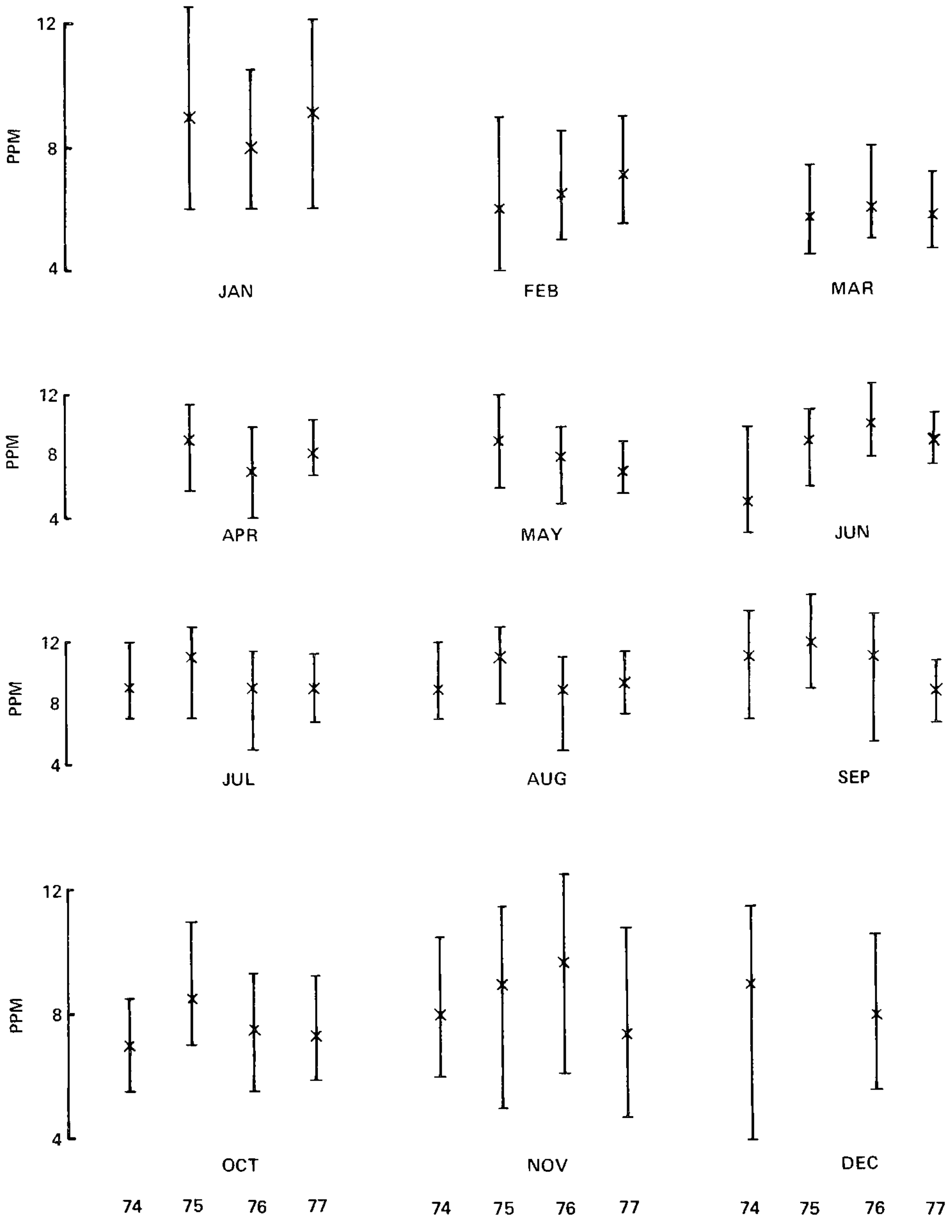


Figure 3-3b. Plot of 25th, 50th, and 75th percentiles of 4-hour afternoon (3 p.m. to 7 p.m.) CO concentrations, Site C.

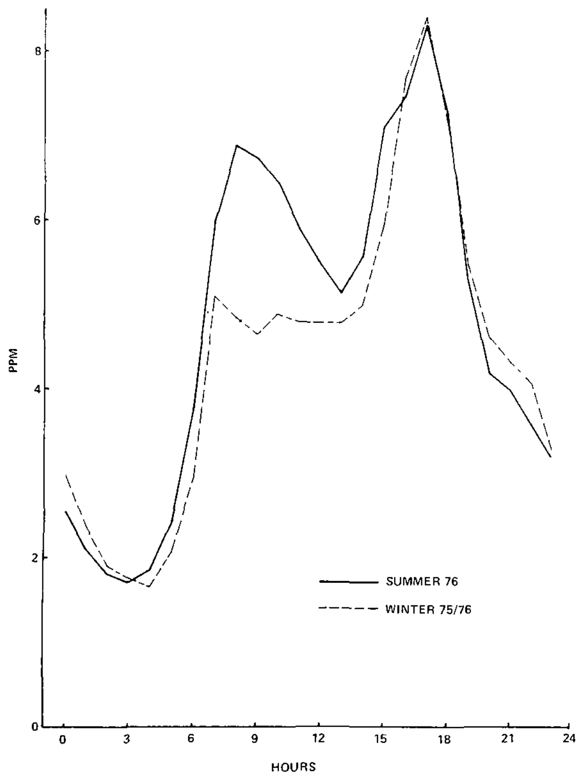


Figure 3-4a. Diurnal plots of CO at Site C; winter - summer comparison (weekday).

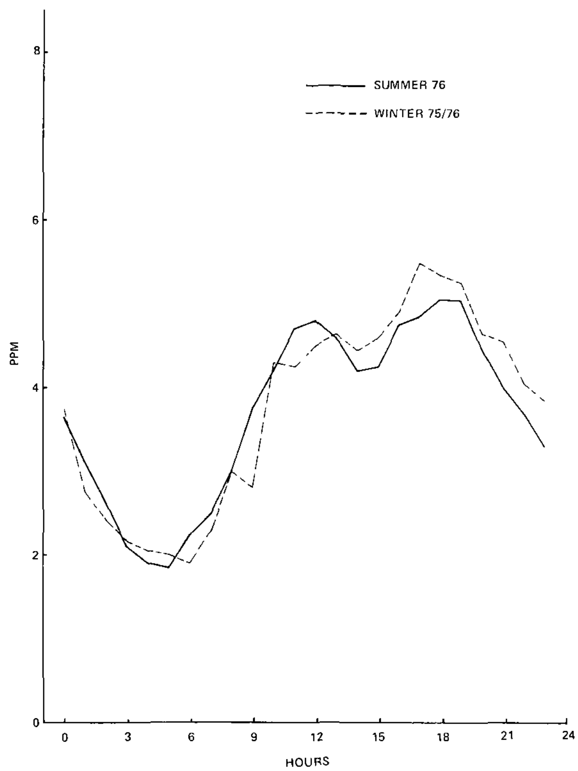


Figure 3-4b. Diurnal plots of CO at Site C; winter - summer comparison (weekend).

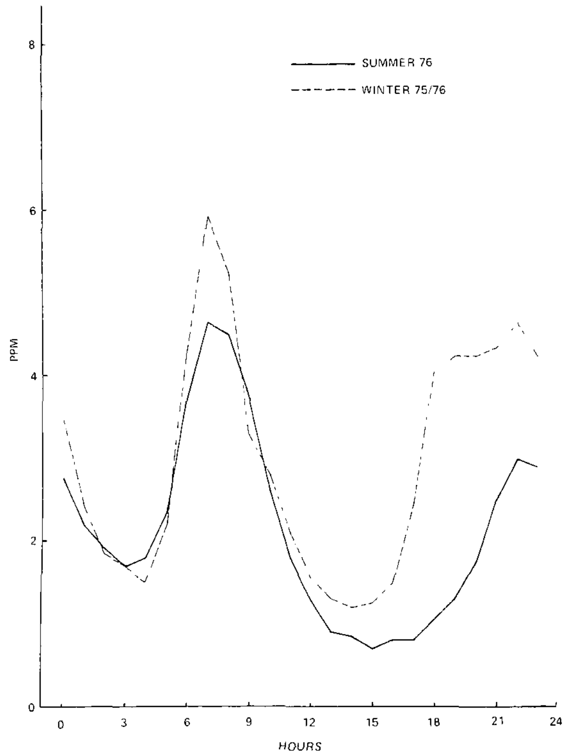


Figure 3-5a. Diurnal plots of CO at Site A; winter - summer comparison (weekday).

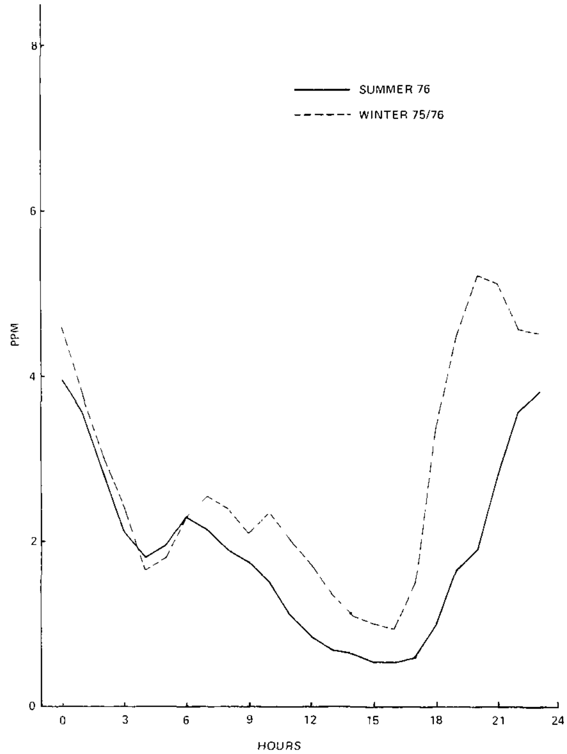


Figure 3-5b. Diurnal plots of CO at Site A; winter summer comparison (weekend).

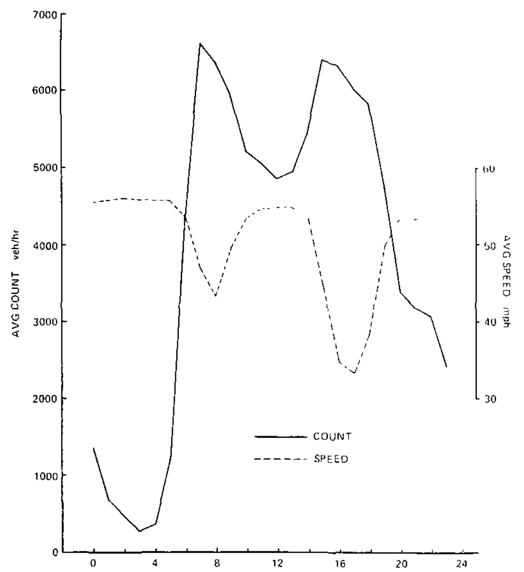


Figure 3-6a. Diurnal plot of traffic counts and traffic speed, weekday 1976.

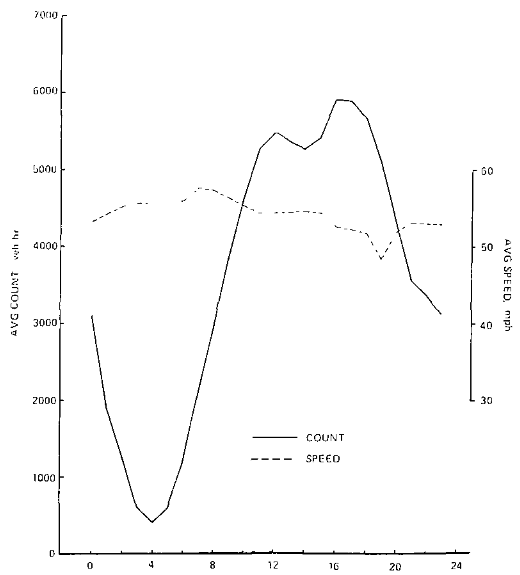


Figure 3-6b. Diurnal plot of traffic counts and traffic speed, weekend 1976.

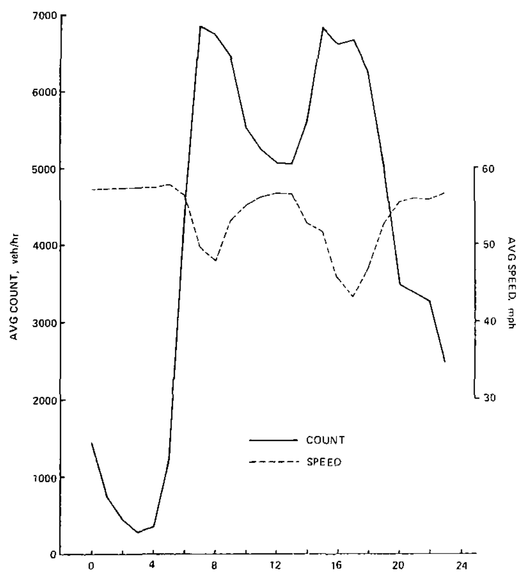


Figure 3-6c. Diurnal plot of traffic counts and traffic speed, weekday 1977.

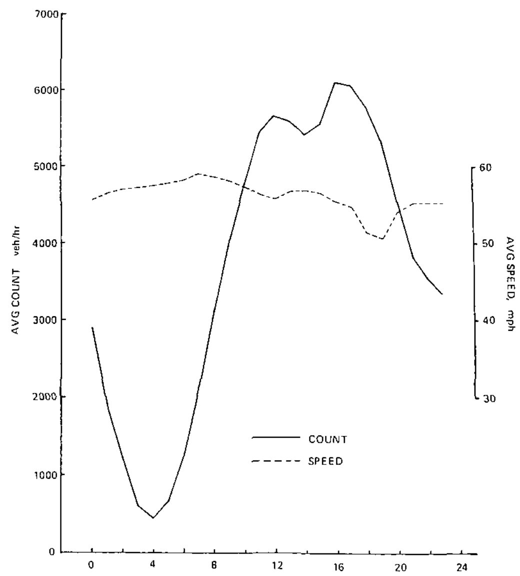


Figure 3-6d. Diurnal plot of traffic counts and traffic speed, weekend 1977.

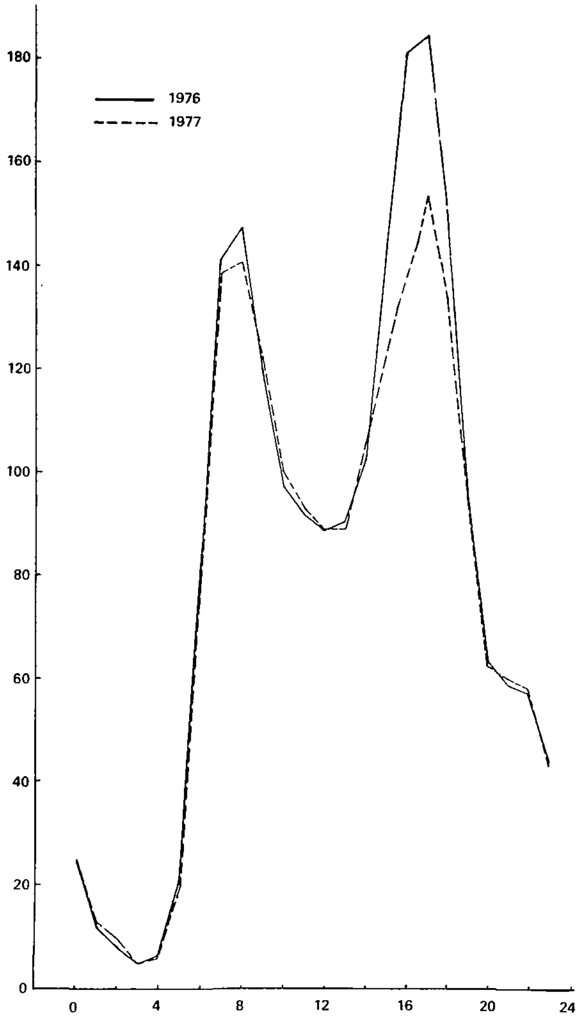


Figure 3-7a. Diurnal plot of traffic density (count/speed) weekday.

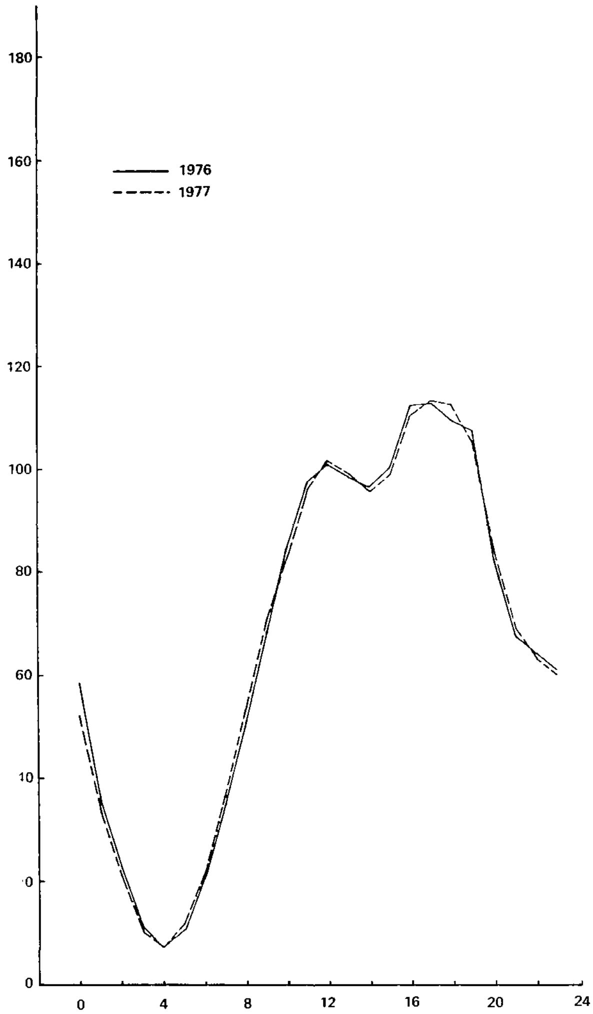


Figure 3-7b. Diurnal plot of traffic density (count/speed) weekend.

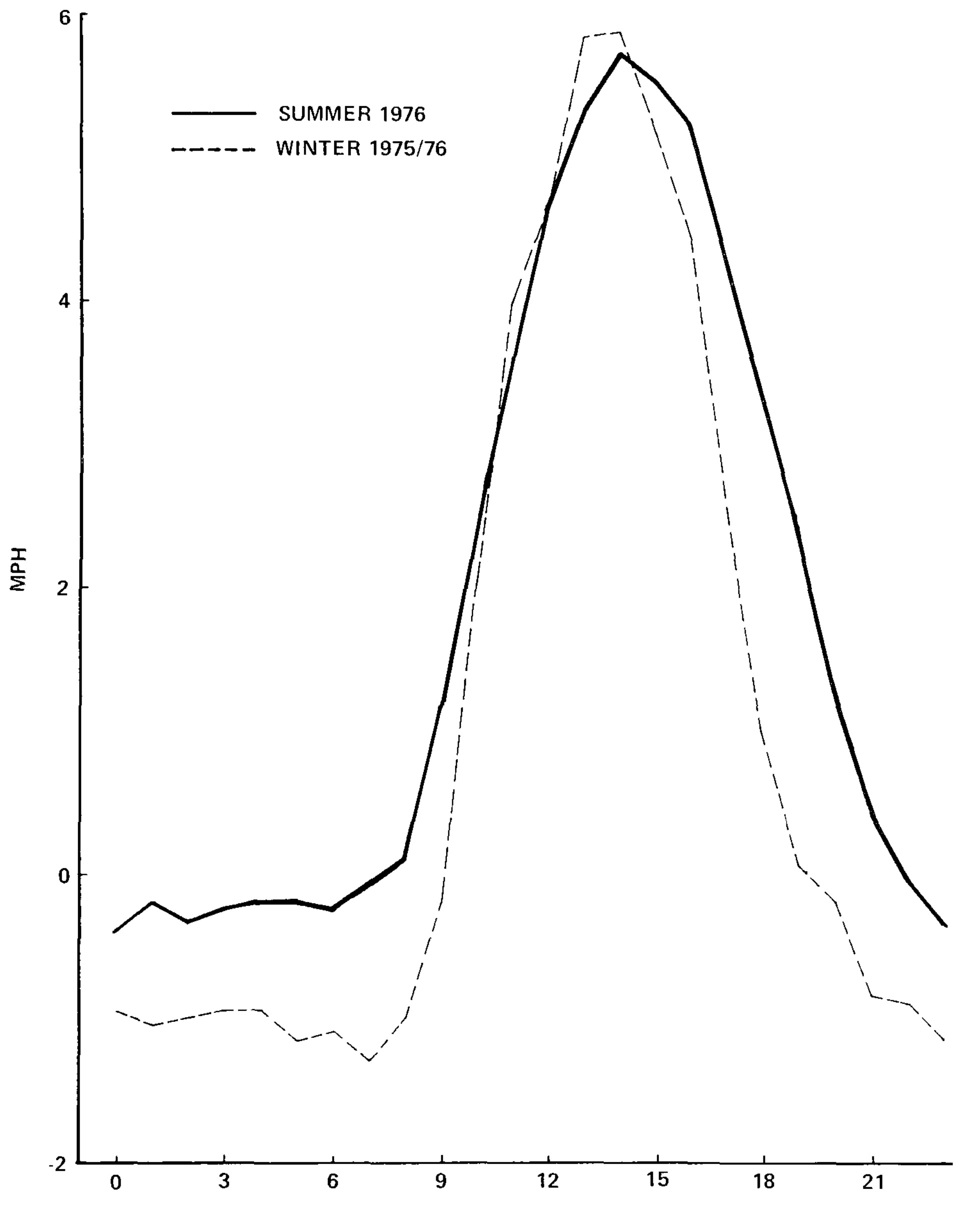


Figure 3-8. Diurnal plots of $WS_{\perp}$.

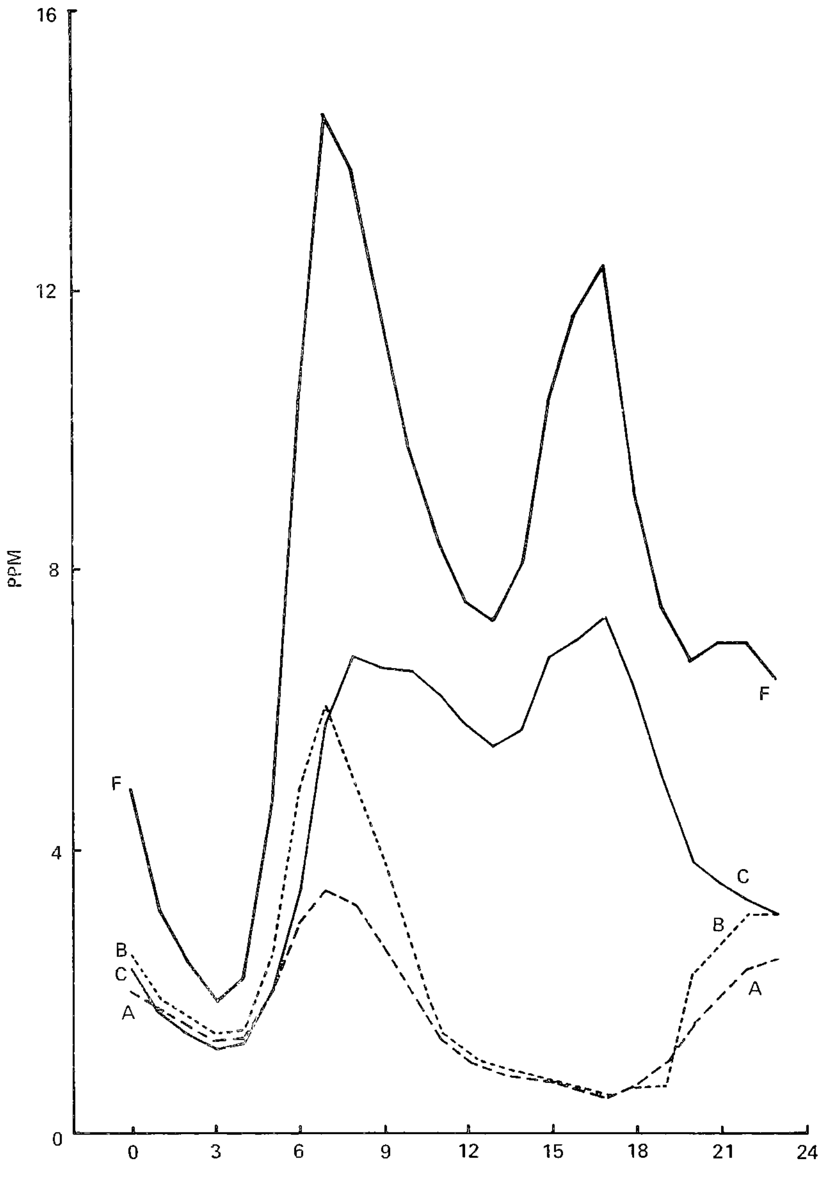


Figure 3-9. Diurnal plot of CO; Sites A, B, F, and C; summer weekday.

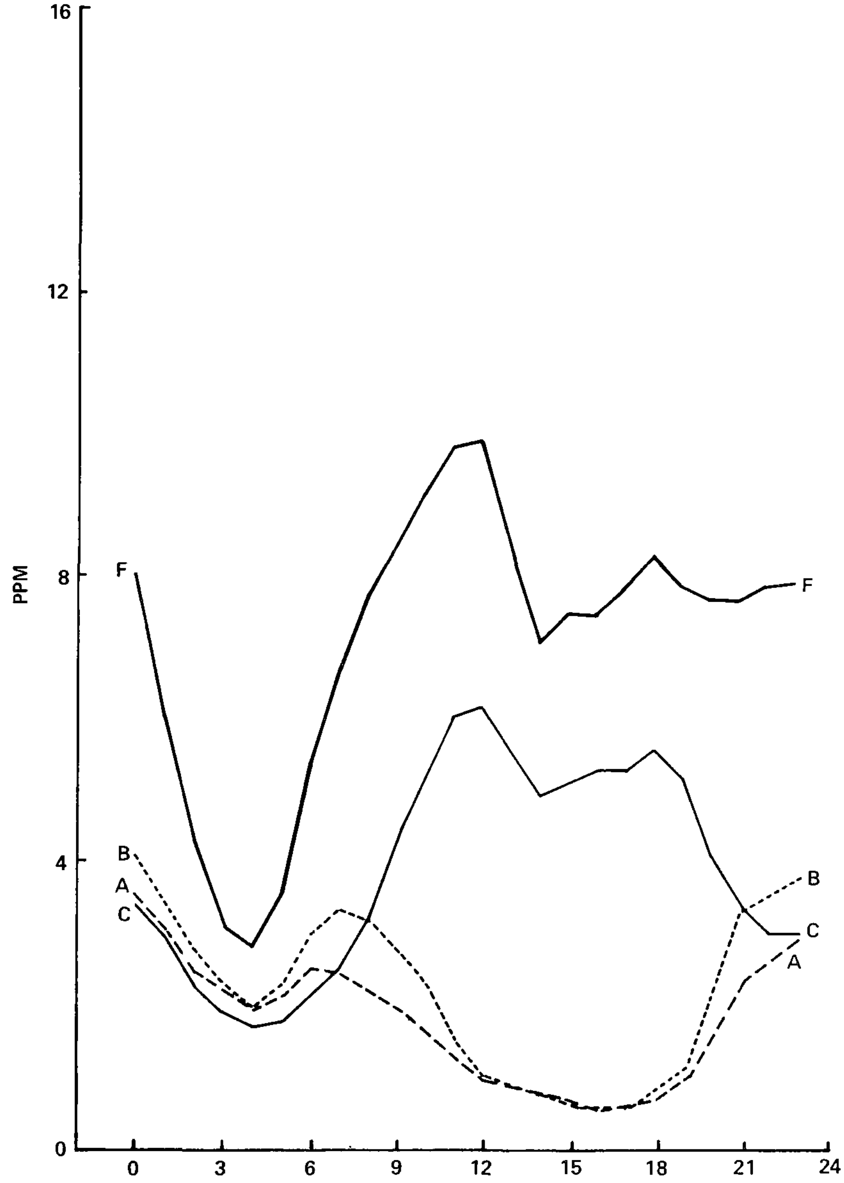


Figure 3-10. Diurnal plot of CO; Sites A, B, F, and C; summer weekend.

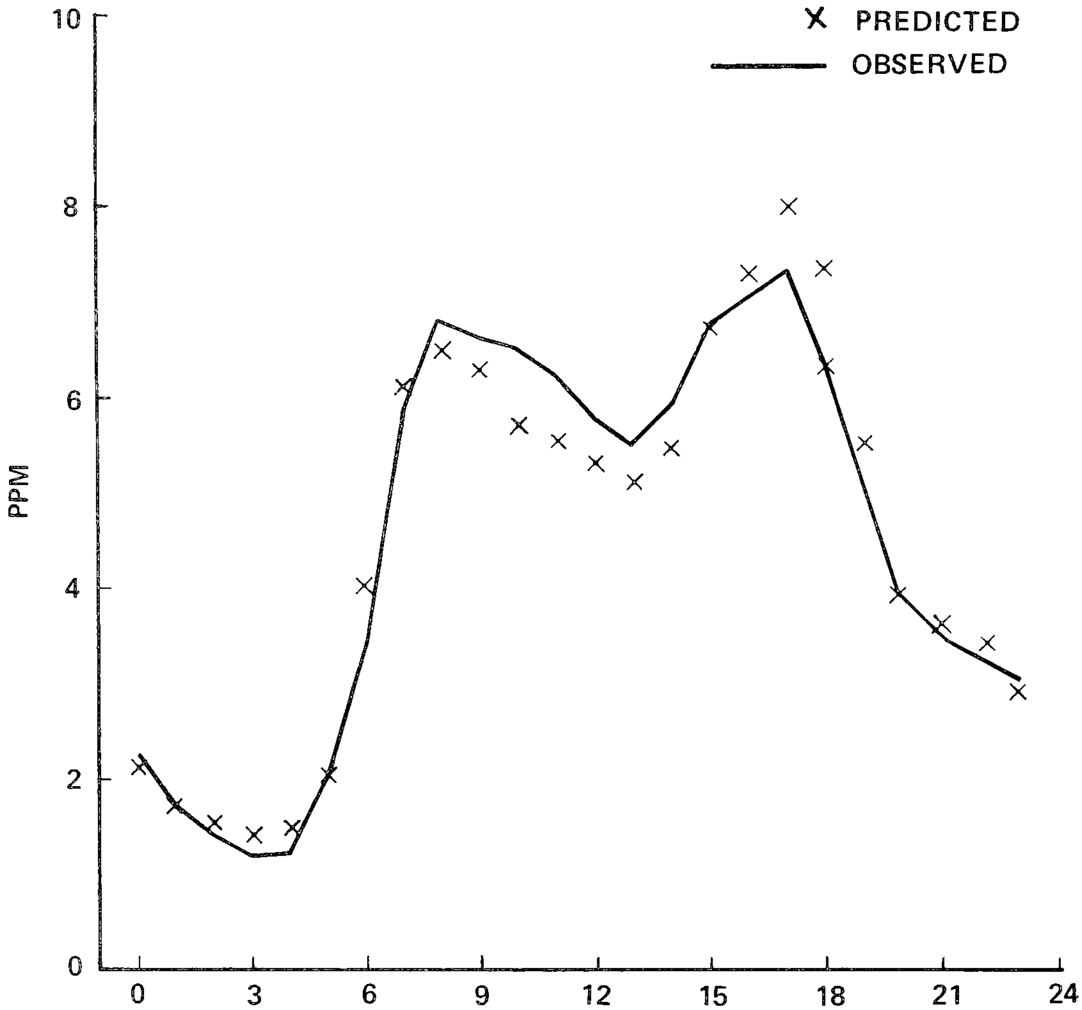


Figure 3-11. CO model 1977 summer weekday fit, Site C.

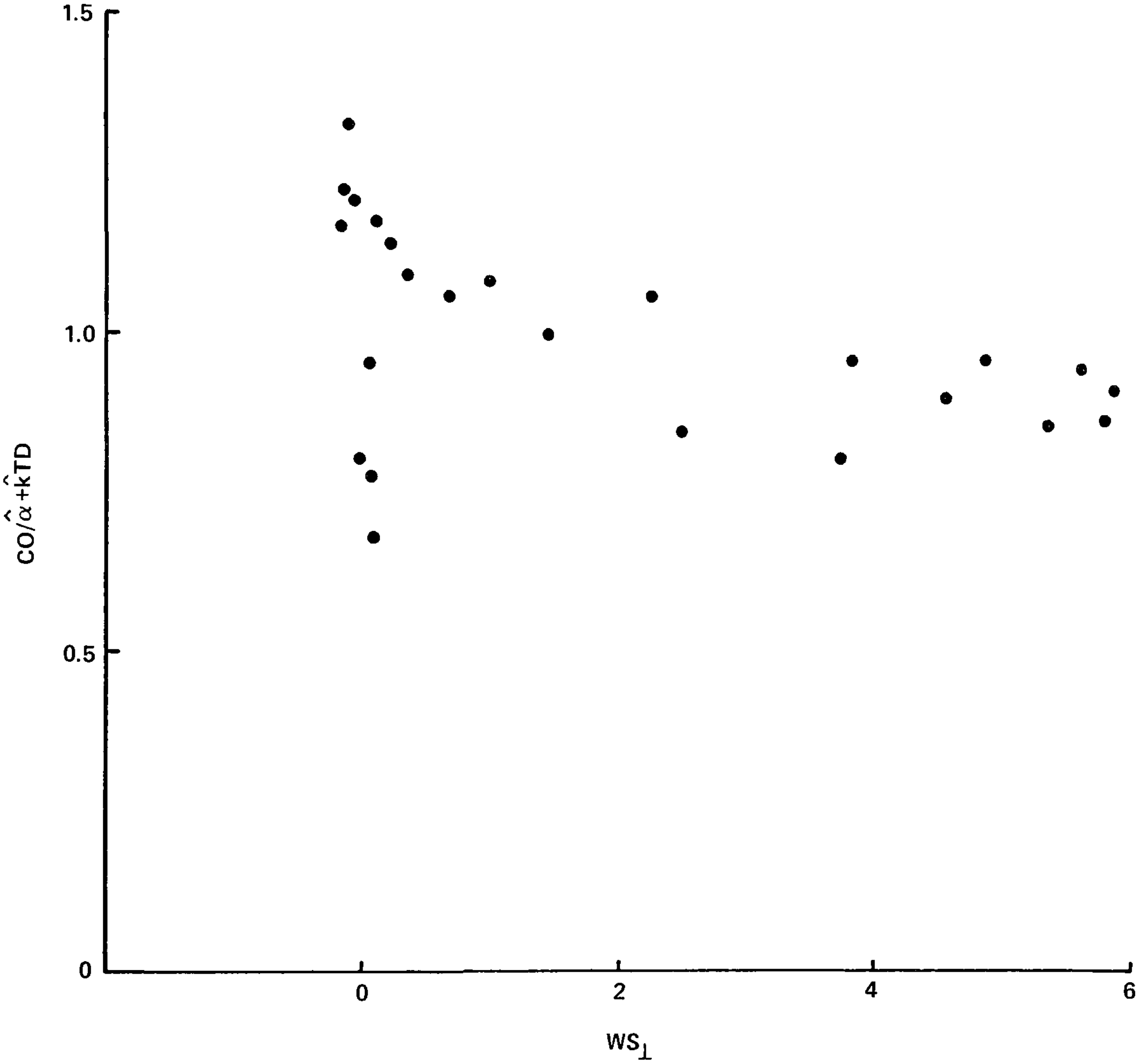


Figure 3-12. Plot of the ratio CO/ $(\hat{\alpha} + \hat{k}TD)$ versus $WS_{\perp}$; Site F.

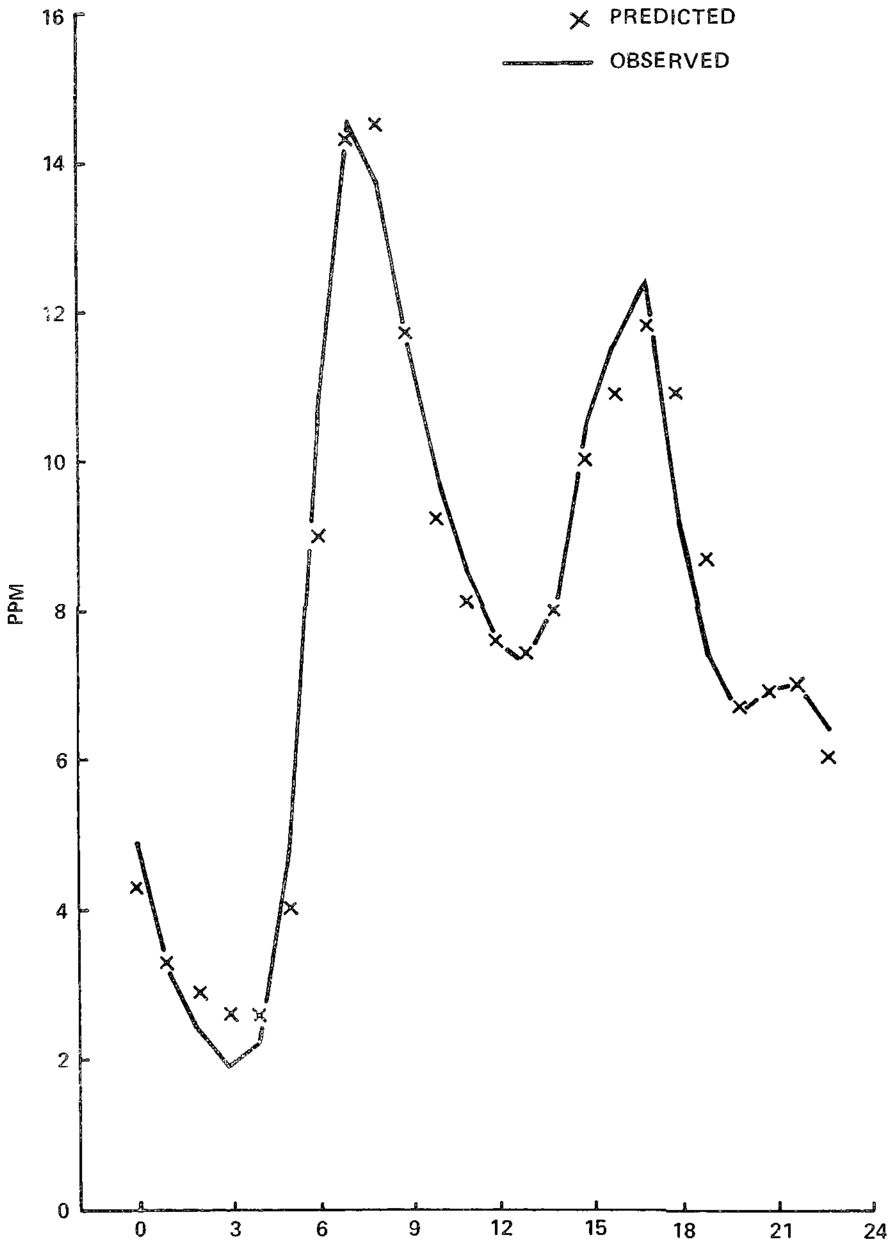


Figure 3-13. CO model 1977 summer weekday fit, Site F (median strip).

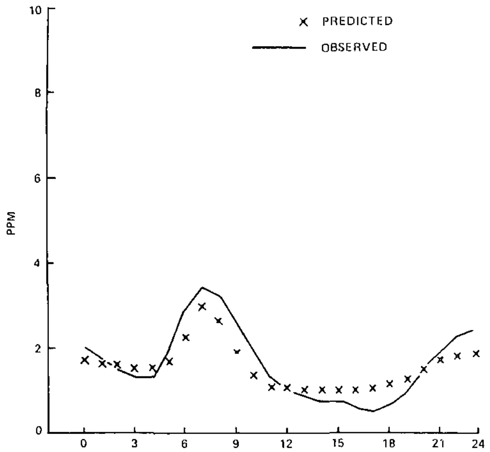


Figure 3-14a. Joint CO model fit summer weekday, 1977, Site A.

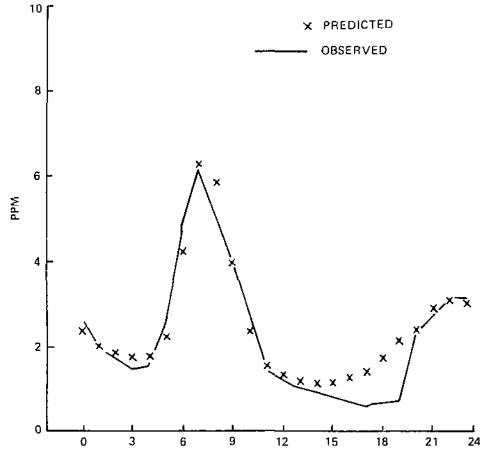


Figure 3-14b. Joint CO model fit summer weekday, 1977, Site B.

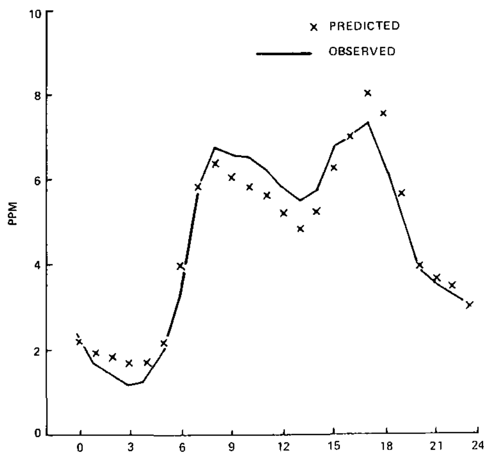


Figure 3-14c. Joint CO model fit summer weekday, 1977, Site C.

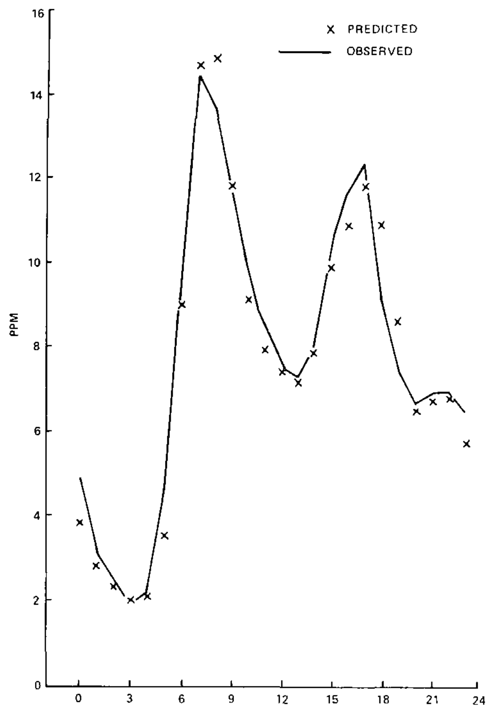


Figure 3-14d. Joint CO model fit summer weekday, 1977, Site F.

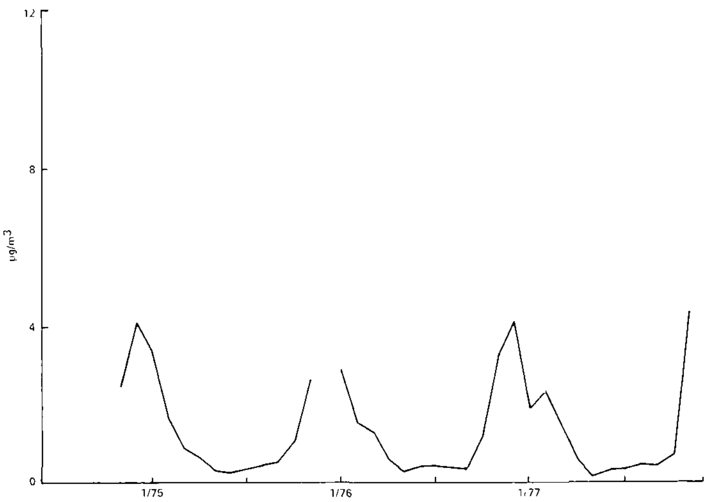


Figure 4-1a. Monthly means of 4-hour afternoon Pb concentrations, Site A.

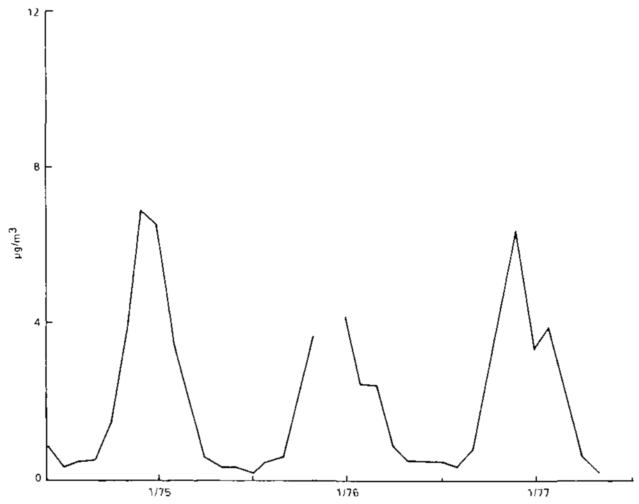


Figure 4-1b. Monthly means of 4-hour afternoon Pb concentrations, Site B.

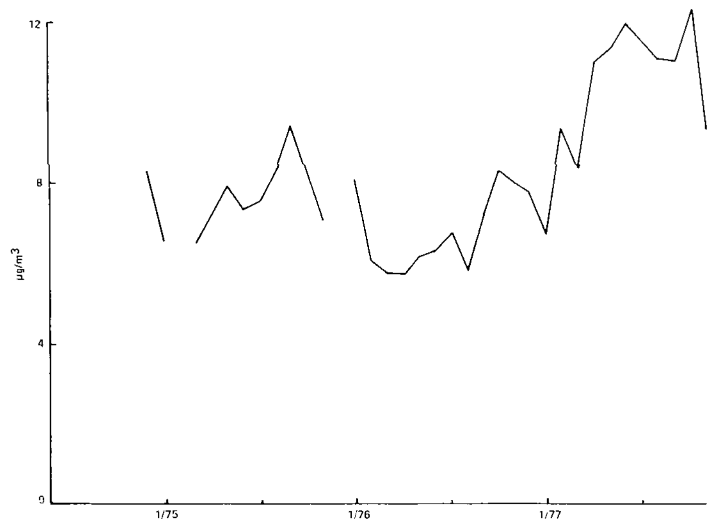


Figure 4-1c. Monthly means of 4-hour afternoon Pb concentrations, Site C.

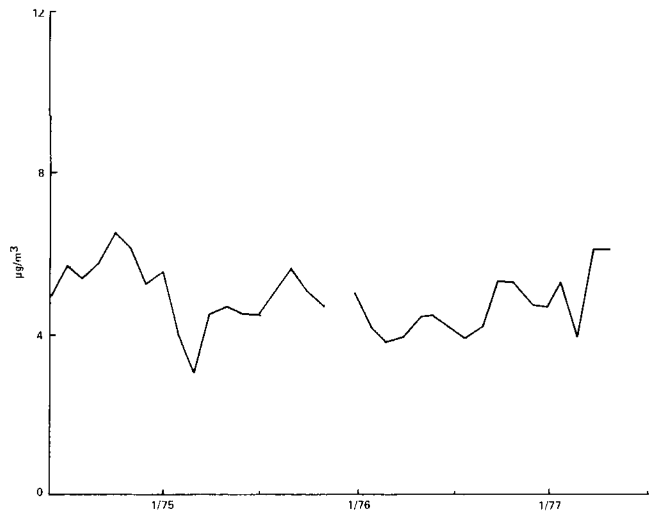


Figure 4-1d. Monthly means of 4-hour afternoon Pb concentrations, Site D.

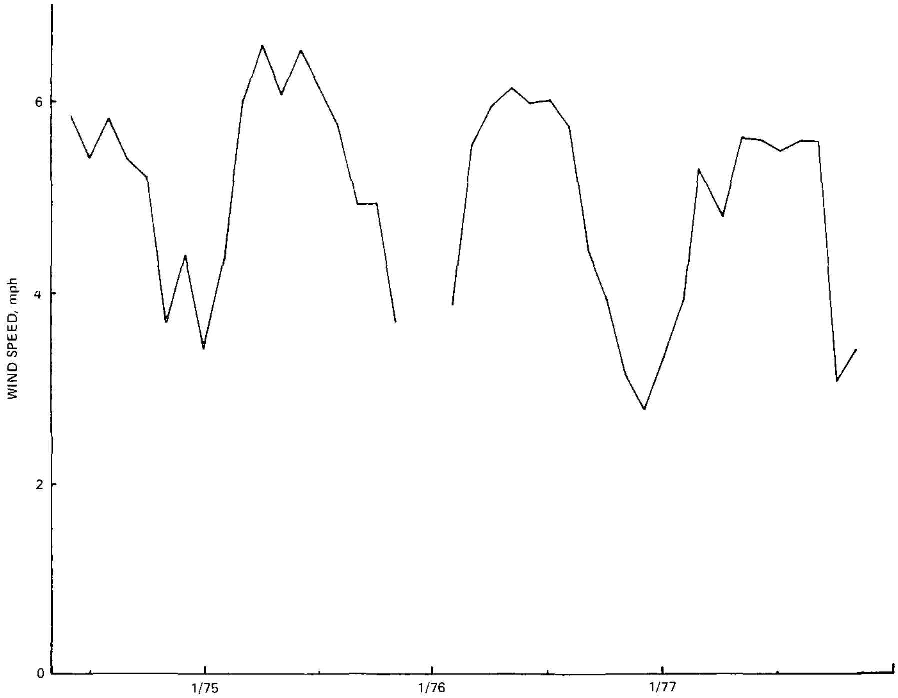


Figure 4-2a. Monthly means of afternoon (3 p.m. to 7 p.m.) windspeed.

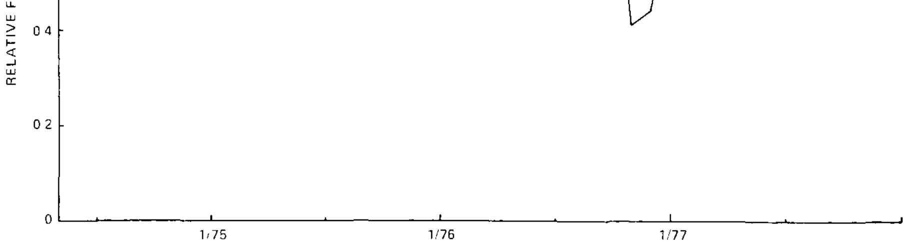


Figure 4-2b. Monthly means of afternoon (3 p.m. to 7 p.m.) winds from 145°- 325°.

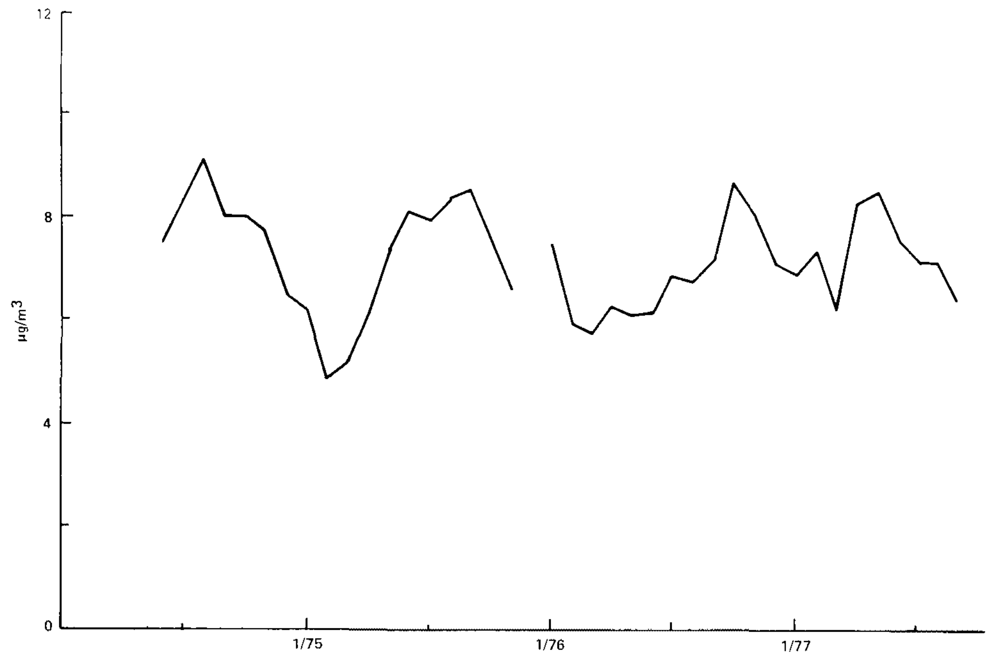


Figure 4-3a. Monthly means of 24-hour Pb concentrations, Site C.

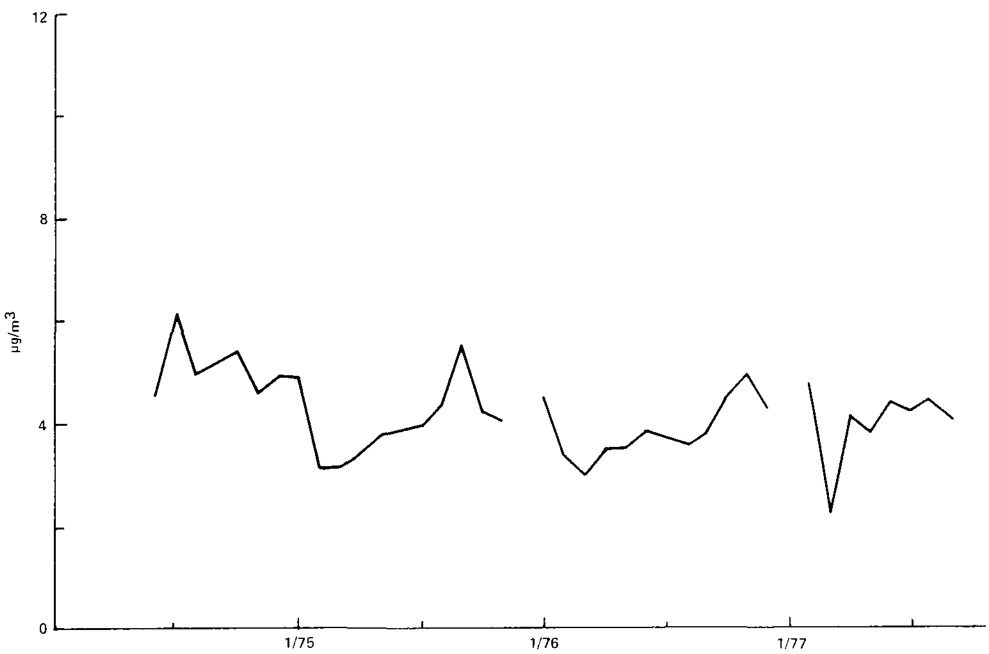


Figure 4-3b. Monthly means of 24-hour Pb concentrations, Site D.

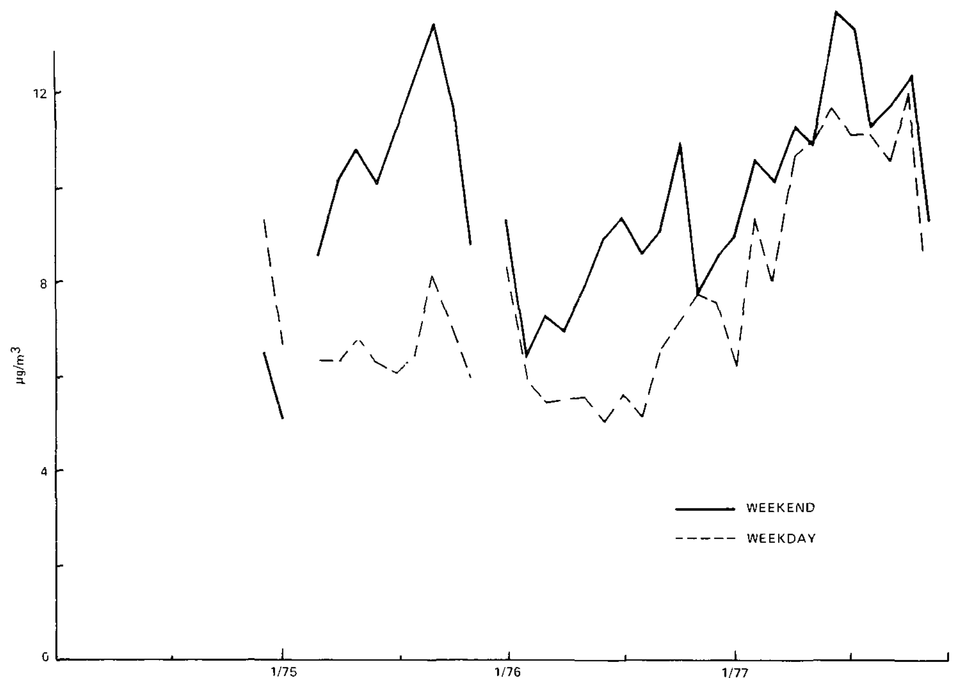


Figure 4-4a. Monthly means of 4-hour (3 p.m. to 7 p.m.) Pb readings weekday weekend comparison, Site C.

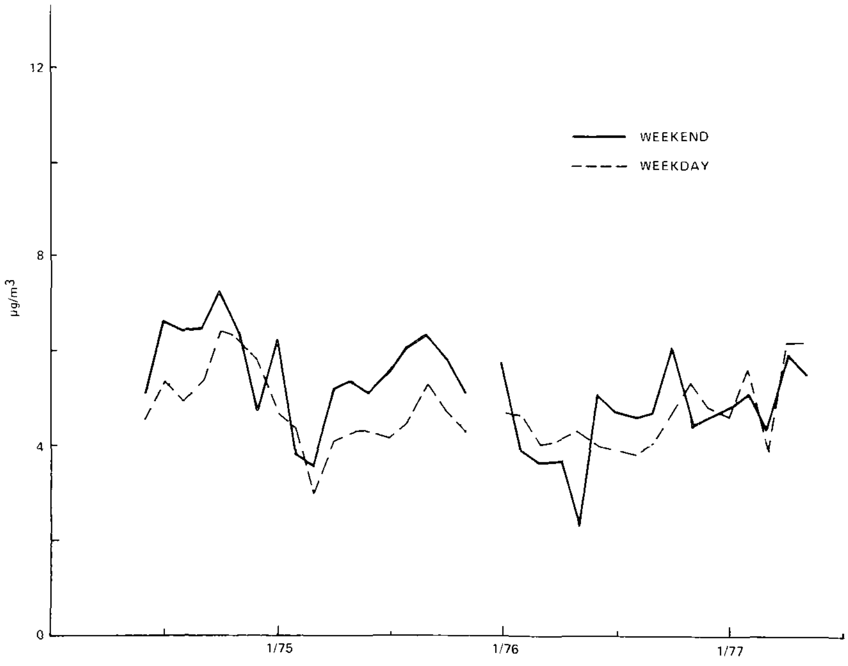


Figure 4-4b. Monthly means of 4-hour (3 p.m. to 7 p.m.) Pb readings weekday - weekend comparison, Site D.

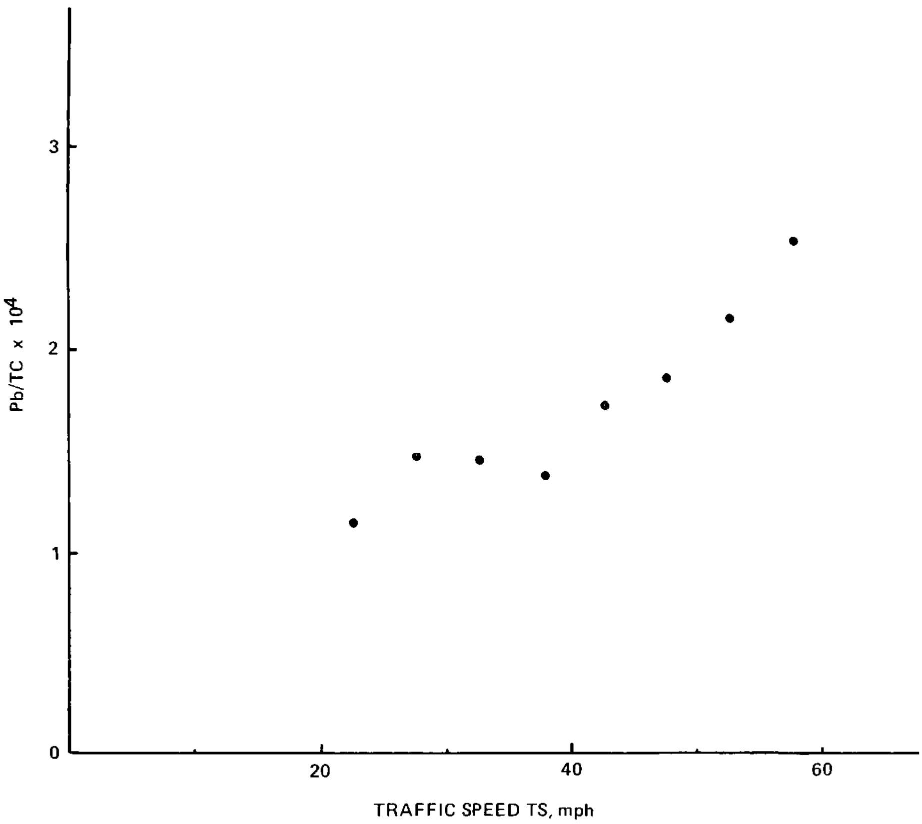


Figure 4-5a. Plot of Pb/TC vs TS.

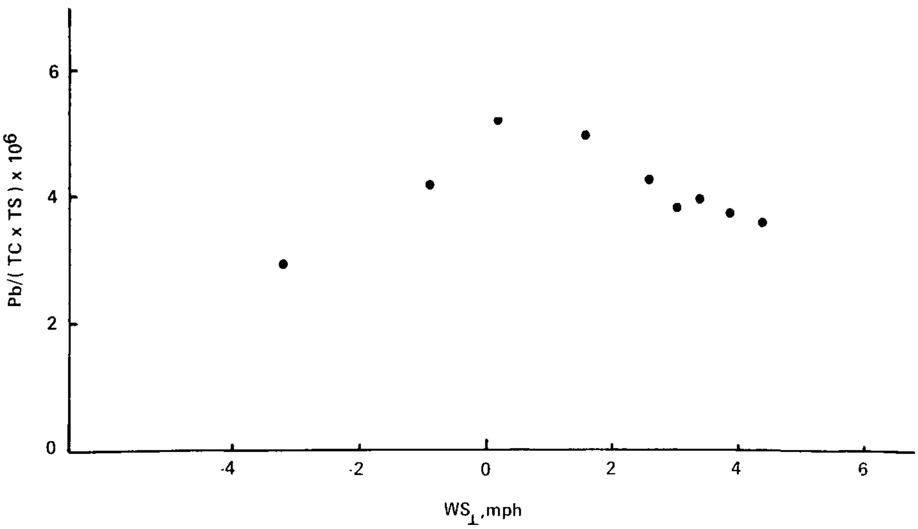


Figure 4-5b. Plot of Pb/TC x TS vs $WS_\perp$.

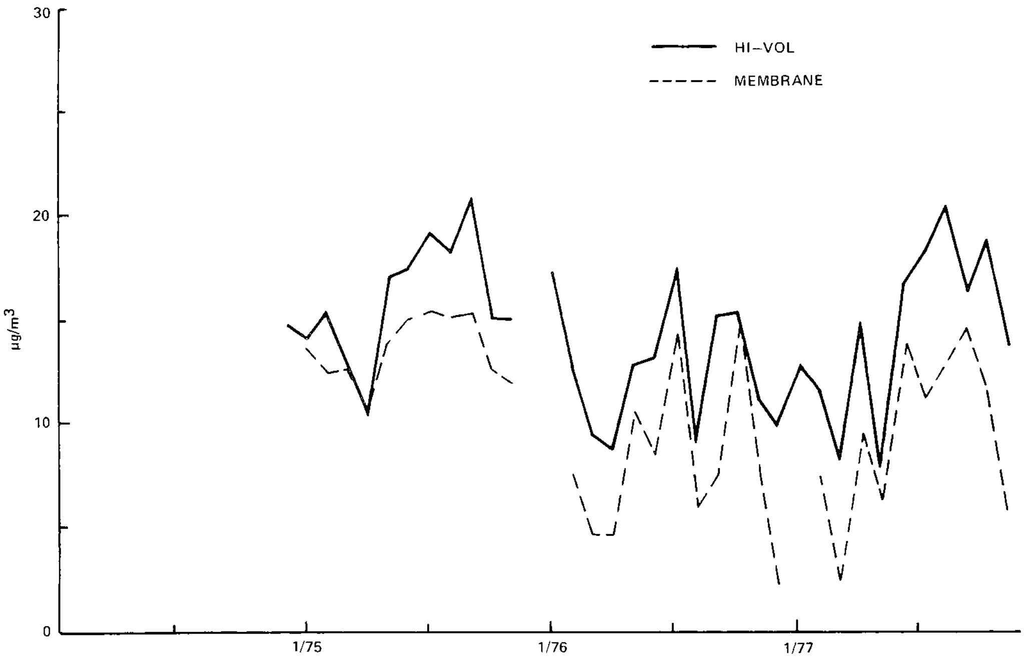


Figure 5-1a. Monthly means of 4-hour hi-vol and membrane afternoon readings of $SO_4^=$; Site A.

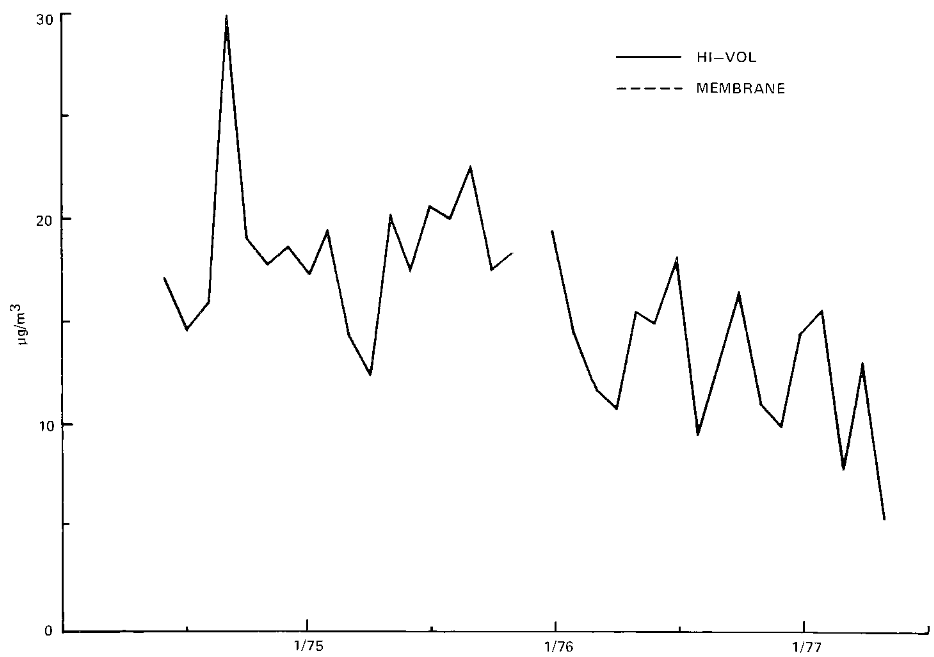


Figure 5-1b. Monthly means of 4-hour hi-vol afternoon readings of $SO_4^=$; Site B.

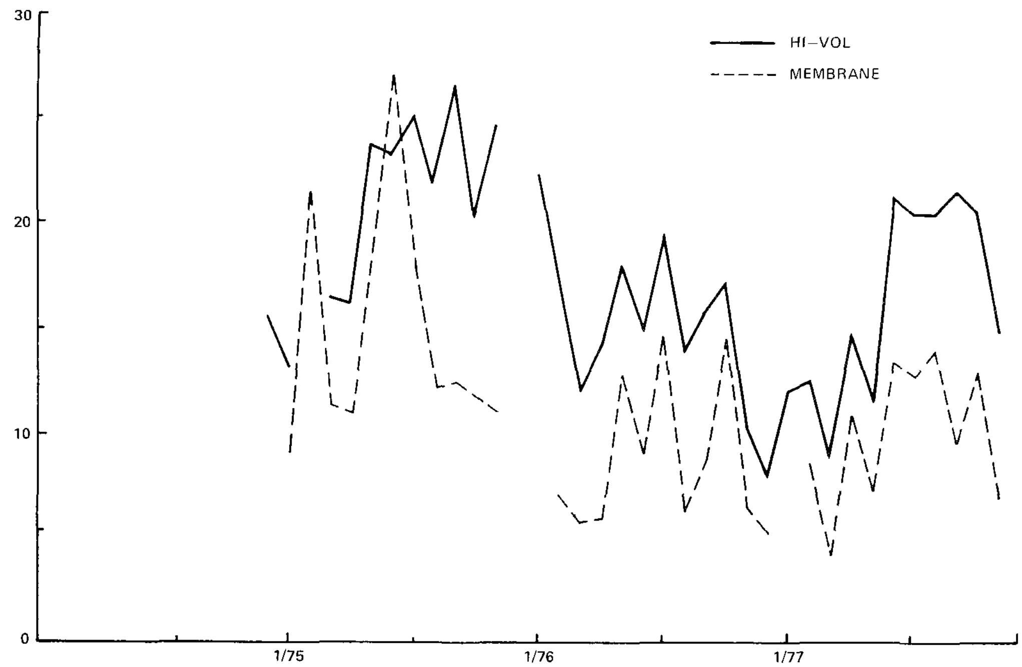


Figure 5-1c. Monthly means of 4-hour hi-vol and membrane readings of $SO_4^=$; Site C

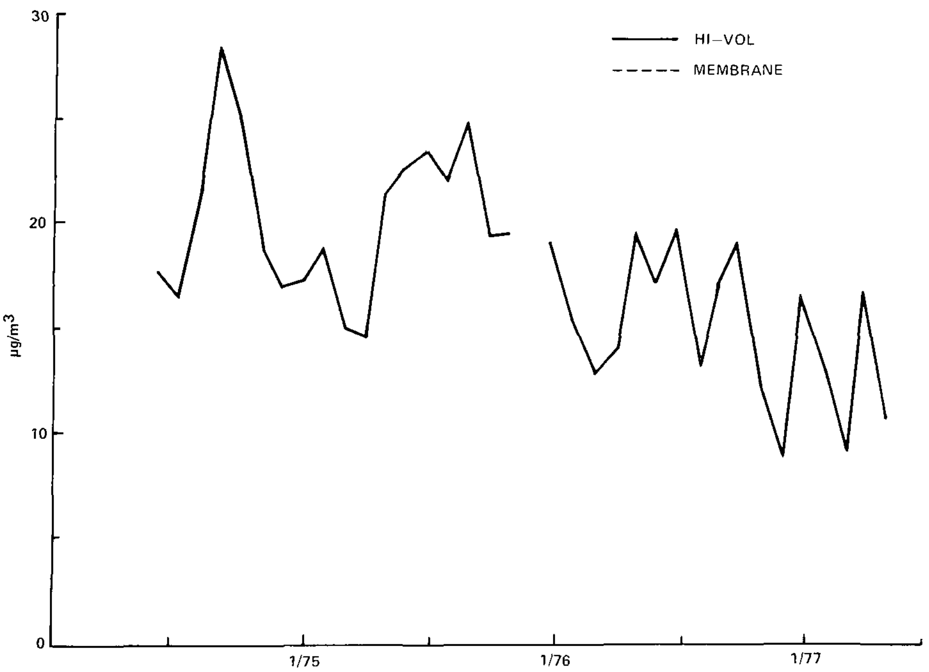


Figure 5-1d. Monthly means of 4-hour hi-vol afternoon readings of $SO_4^=$; Site D.

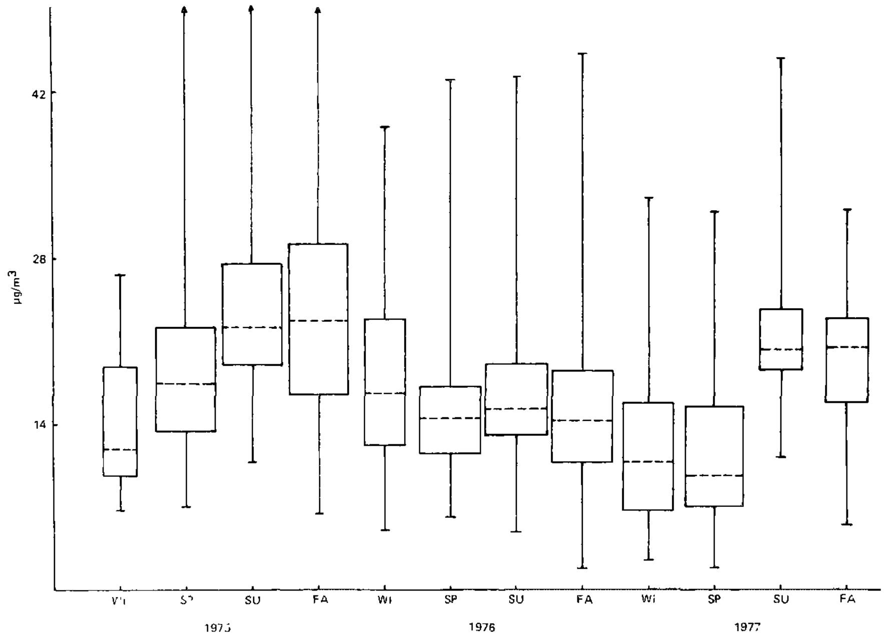


Figure 5-2a. Plot of quarterly 25th, 50th, and 75th percentiles, minimum and maximum of 4-hour hi-vol afternoon $SO_4^=$ concentrations at Site C.

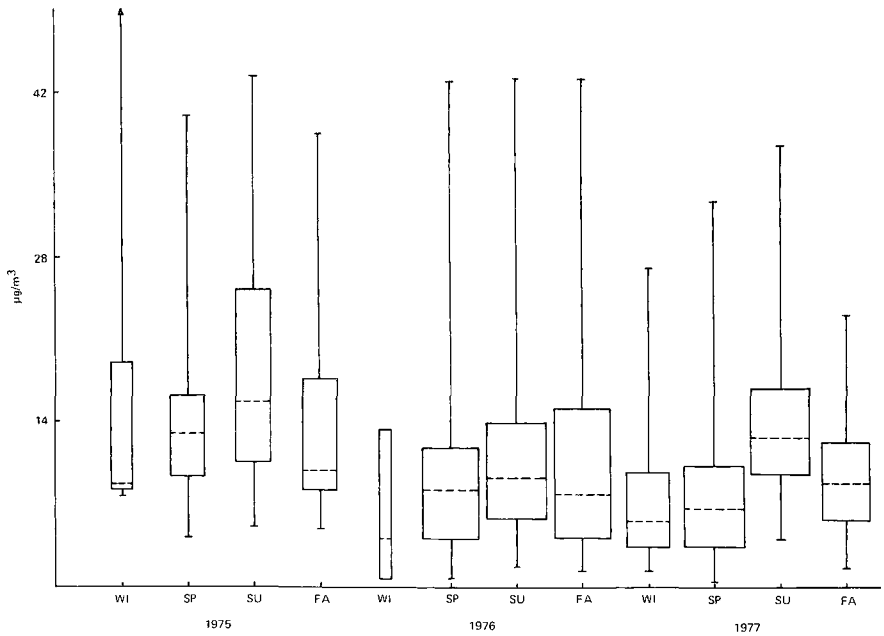


Figure 5-2b. Plot of quarterly 25th, 50th, and 75th percentiles, minimum and maximum of 4-hour membrane afternoon $SO_4^=$ concentrations at Site C.

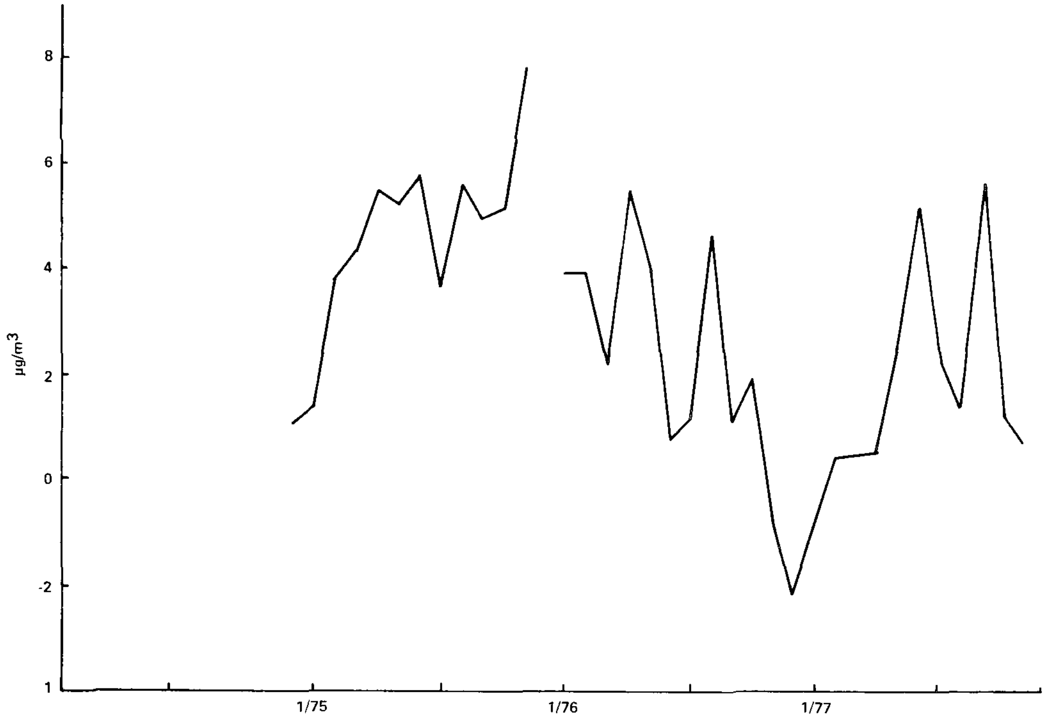


Figure 5-3a. Monthly means of 4-hour hi-vol $SO_4^=$ (C - A) differences.

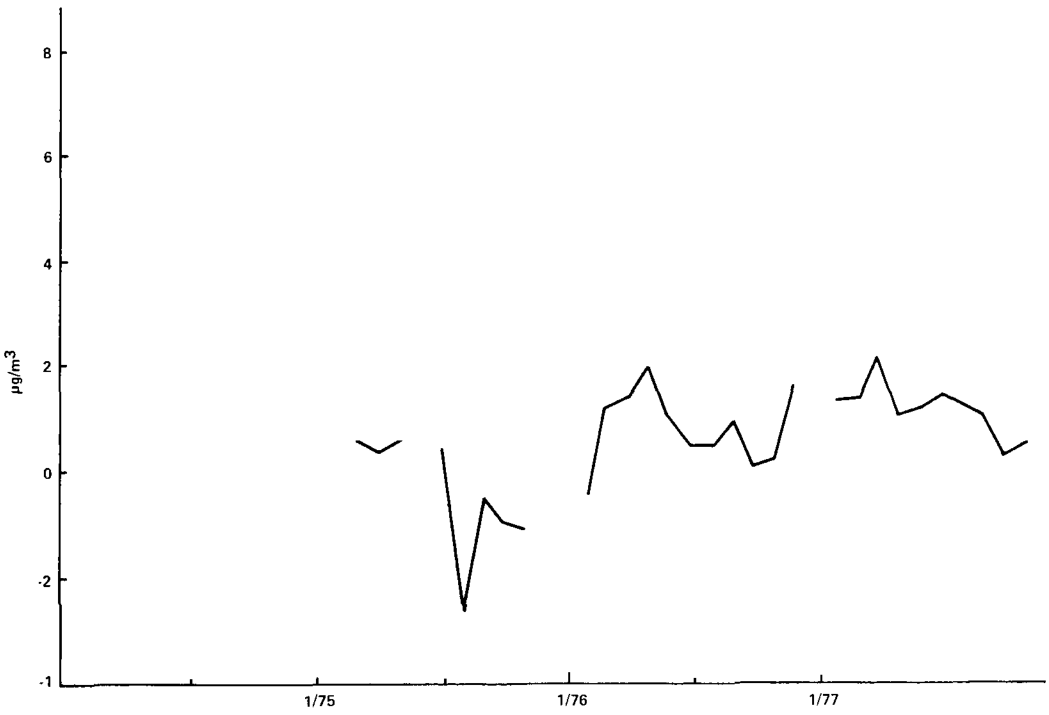


Figure 5-3b. Monthly means of 4-hour membrane $SO_4^=$ (C - A) differences.

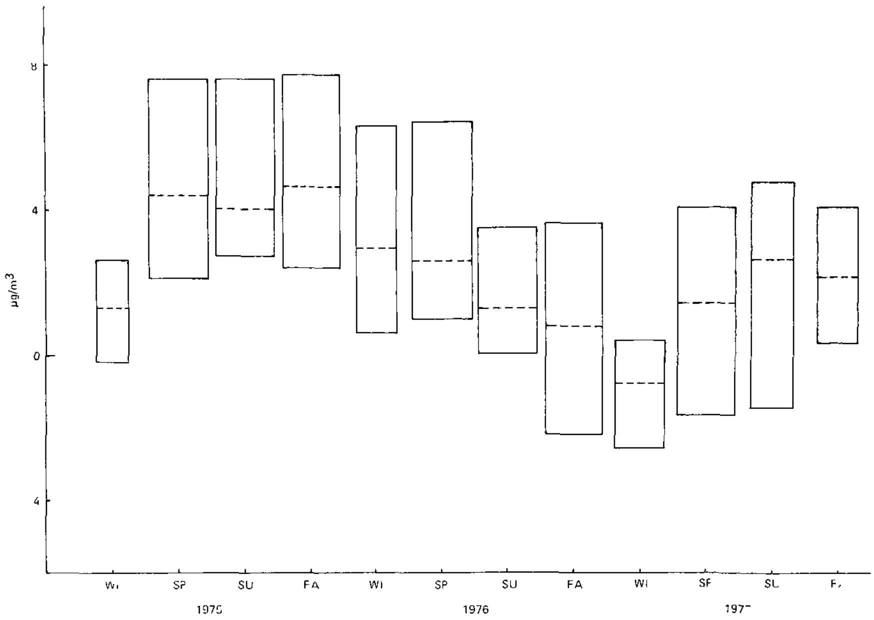


Figure 5-4a. Plot of quarterly 25th, 50th, and 75th percentiles of 4-hour hi-vol afternoon $SO_4^=$ (C - A) differences.

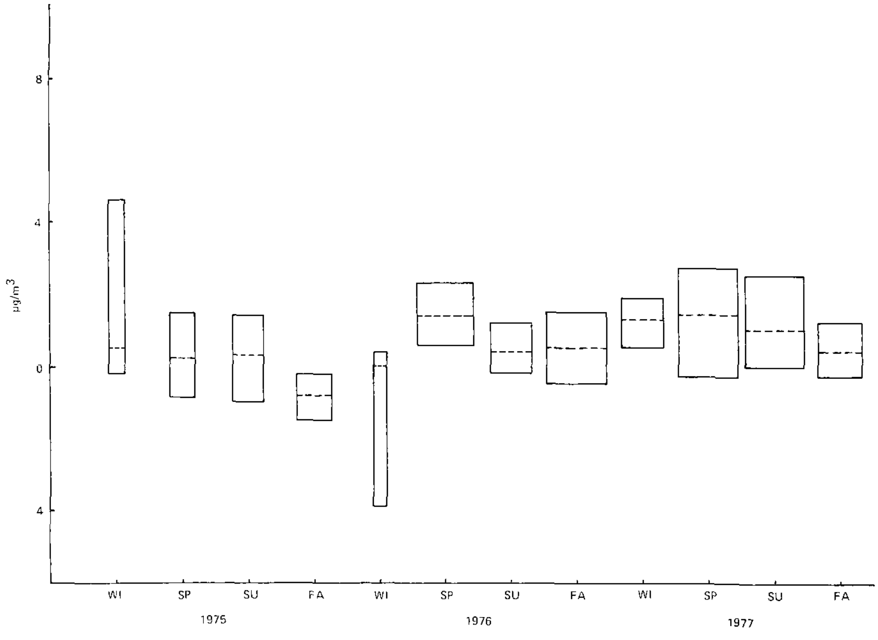


Figure 5-4b. Plot of quarterly 25th, 50th, and 75th percentiles of 4-hour membrane afternoon $SO_4^=$ (C - A) differences.

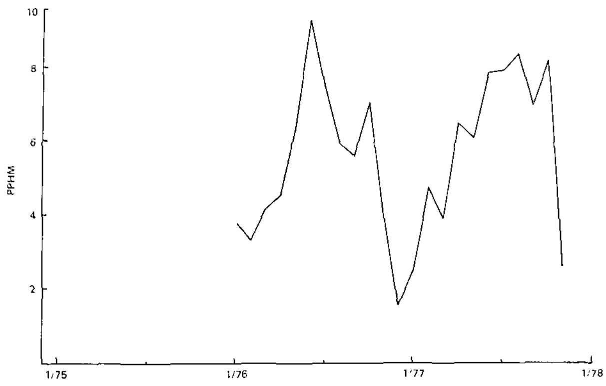


Figure 8-1a. Monthly means of 4-hour (3 p.m. to 7 p.m.) O_3 at Site A.

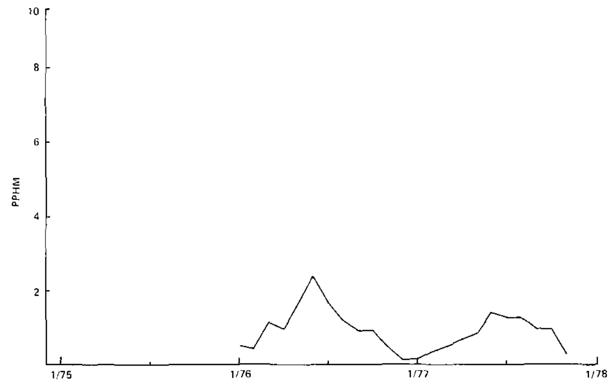


Figure 8-1b. Monthly means of 4-hour (3 p.m to 7 p.m.) O_3 at Site C.

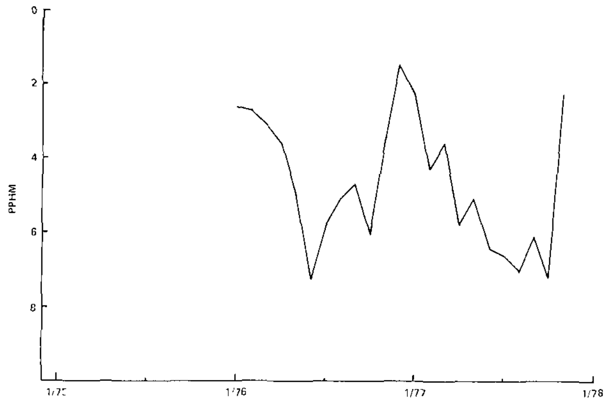


Figure 8-1c. Monthly means of 4-hour (3 p.m. to 7 p.m.) O_3 (C A).

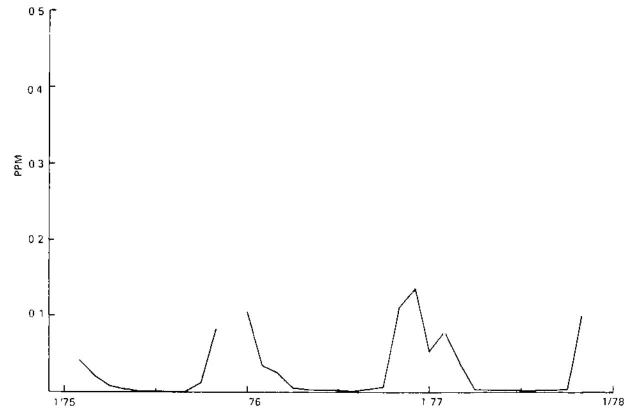


Figure 8-2a. Monthly means of 4-hour (3 p.m. to 7 p.m.) NO at Site A.

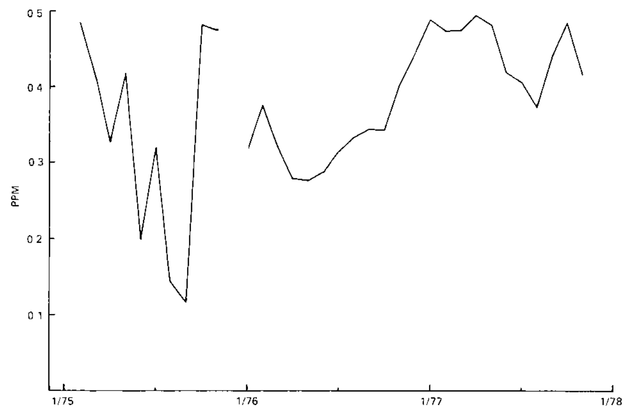


Figure 8-2b. Monthly means of 4-hour (3 p.m. to 7 p m.) NO at Site C.

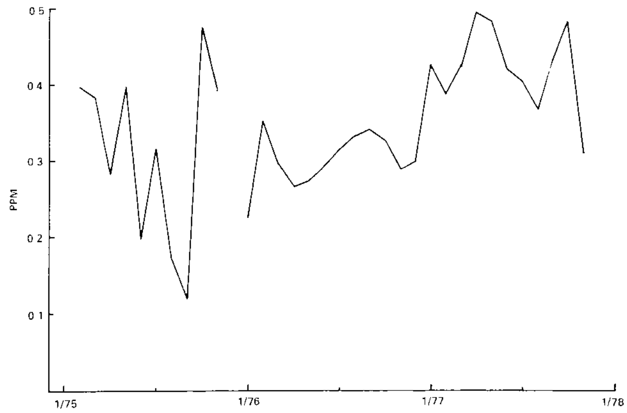


Figure 8-2c. Monthly means of 4-hour (3 p.m. to 7 p.m.) NO (C - A).

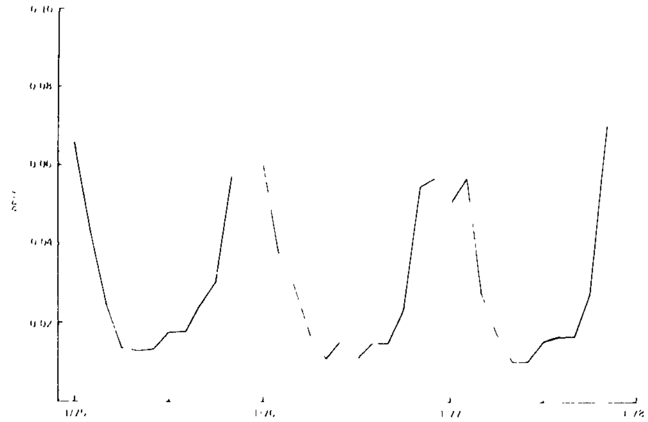


Figure 8-3a. Monthly means of 4-hour (3 p.m. to 7 p.m.) NO_2 at Site A.

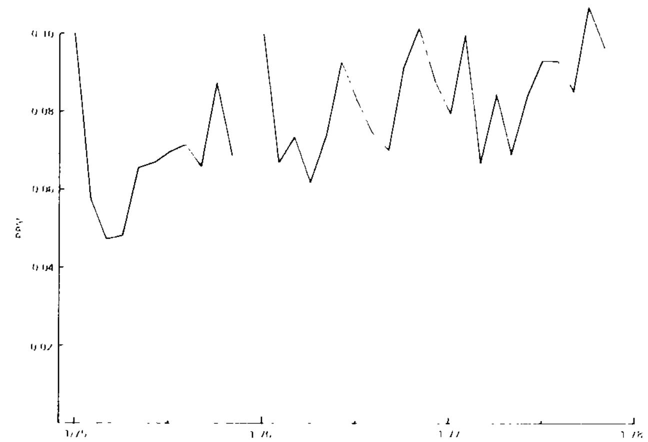


Figure 8-3b. Monthly means of 4-hour (3 p.m. to 7 p.m.) NO_2 at Site C.

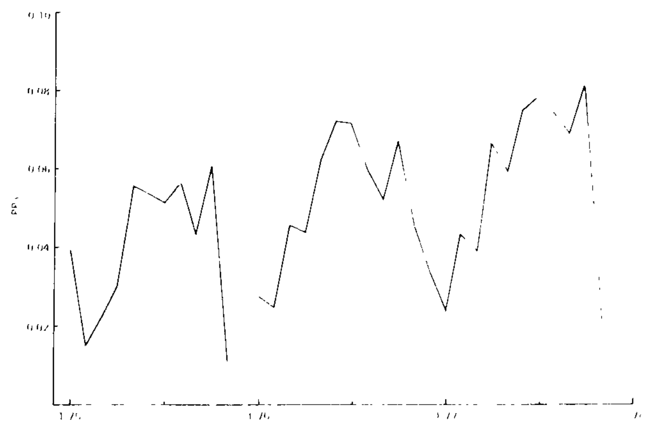


Figure 8-3c. Monthly means of 4-hour (3 p.m. to 7 p.m.) NO_2 (C A).

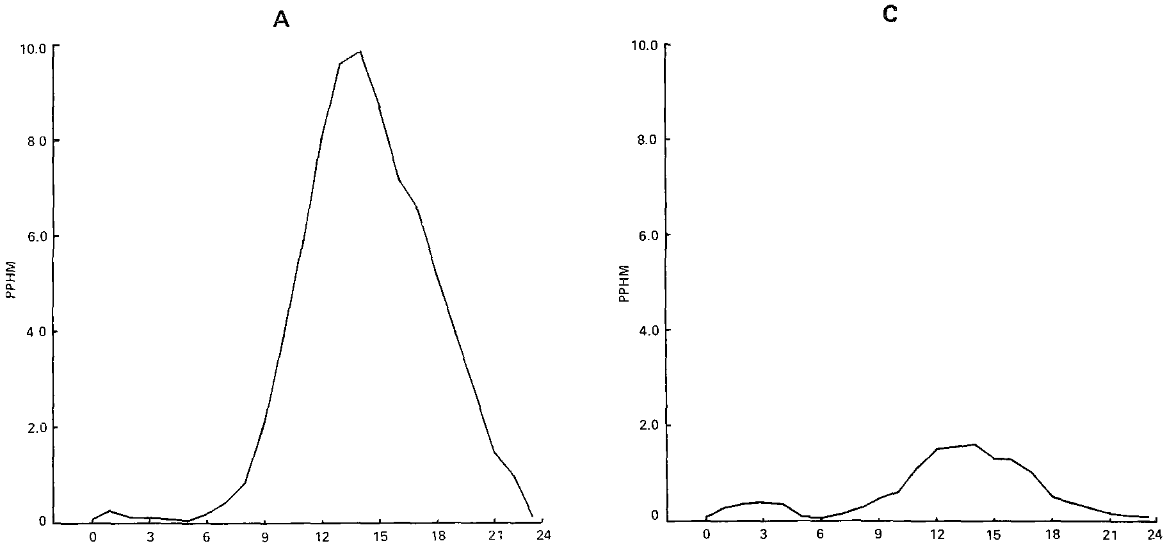


Figure 8-4a. Diurnal plots of O_3 (Monday - Thursday) at Sites A and C.

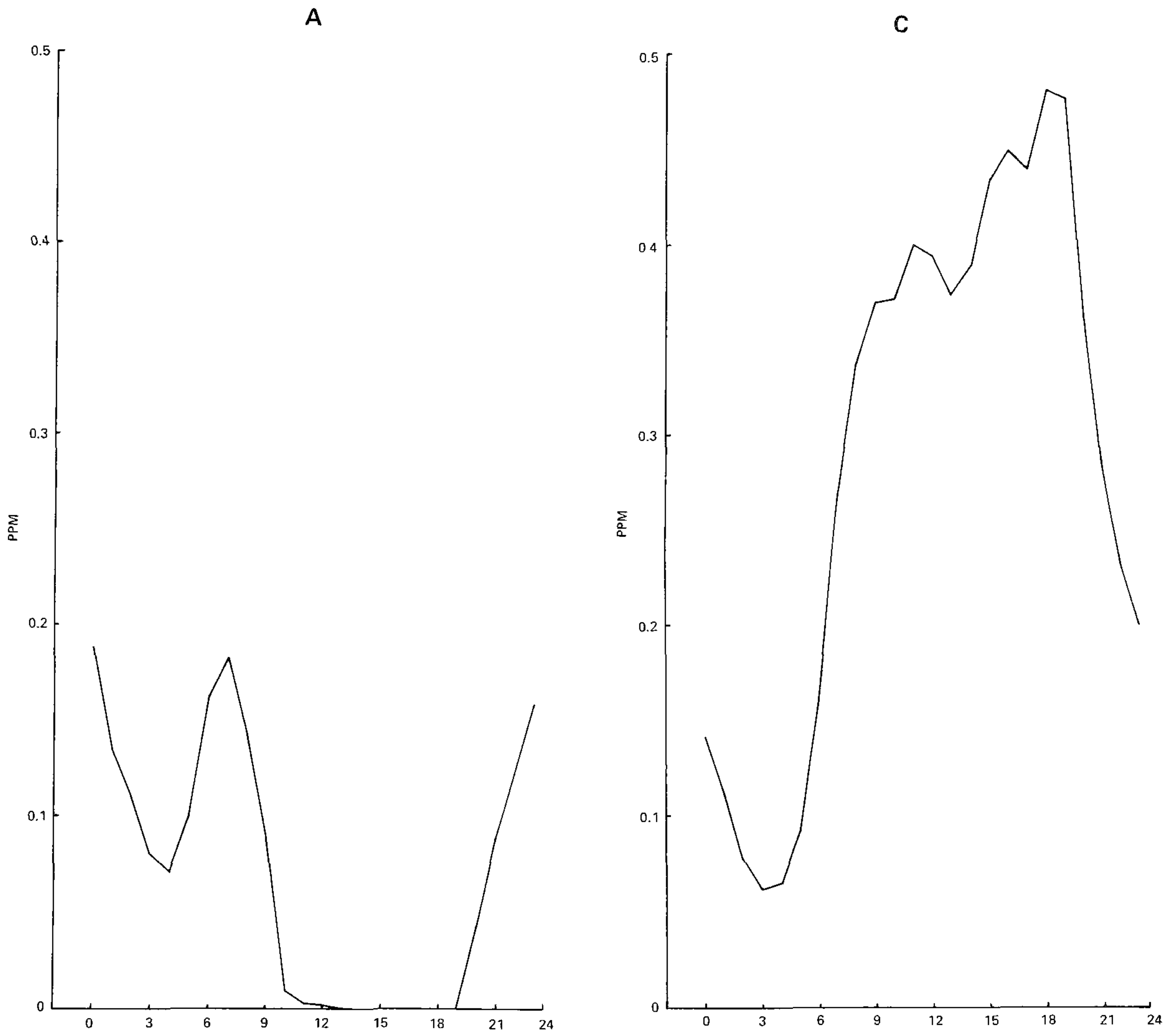


Figure 8-4b. Diurnal plots of NO (Monday - Thursday) at Sites A and C.

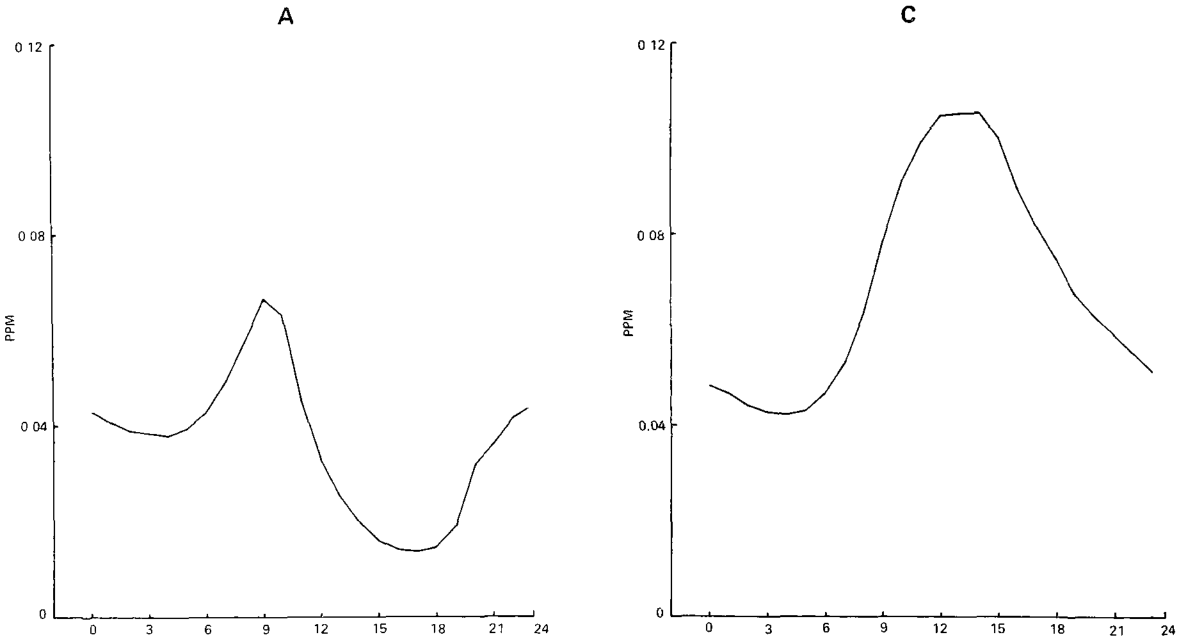


Figure 8-4c. Diurnal plots of NO_2 (Monday - Thursday) at Sites A and C.

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA 600/4-79-070	2.	3. RECIPIENT'S ACCESSION NO.
4. TITLE AND SUBTITLE STATISTICAL ANALYSIS OF THE LOS ANGELES CATALYST STUDY DATA		5. REPORT DATE October 1979
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Johannes Ledolter, George C. Tiao, Spencer B. Graves, Jian-tu Hsieh, Gregory B. Hudak (U. of Wisc) Harold B. Sauls, EPA/EMSL, RTP, N.C.		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics University of Wisconsin Madison, Wisconsin 53706		10. PROGRAM ELEMENT NO.
		11. CONTRACT/GRANT NO. 68-02-2261
12. SPONSORING AGENCY NAME AND ADDRESS Environmental Monitoring and Support Laboratory Office of Research and Development U.S. Environmental Protection Agency Research Triangle Park, N.C. 27711		13. TYPE OF REPORT AND PERIOD COVERED
		14. SPONSORING AGENCY CODE EPA-600/08

15. SUPPLEMENTARY NOTES

16. ABSTRACT

This research was initiated to perform statistical analyses of the data from the Los Angeles Catalyst Study. The objective is to determine the effects of the introduction of the catalytic converter upon the atmospheric concentration levels of a number of air pollutants.

This report gives an analysis of the CO, Pb, SO_4, O_3, NO and NO_2 data covering the period from June 1974 to November 1977. Models are built to evaluate the freeway contribution to CO and Pb as a function of traffic, windspeed and wind direction. These models are used to assess both the time trend in the pollutant measurements and the pollution concentrations at points near the freeway. Furthermore frequency distributions for ambient air quality data near freeways are discussed.

This report was submitted in fulfillment of U.S. E.P.A. Contract No. 68-02-2261 with the University of Wisconsin. It covers work done during the period September 1977 through August 1978.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
mobile source pollution freeway pollution mobile pollution modeling CO modeling mobile source monitoring freeway monitoring catalyst effects	Los Angeles catalytic converter ozone lead sulfate carbon monoxide	43F 68A

18. DISTRIBUTION STATEMENT Release to public	19. SECURITY CLASS (This Report) Unclassified	21. NO. OF PAGES 77
	20. SECURITY CLASS (This page) Unclassified	22. PRICE

EPA Form 2220-1 (9-73)