T-3626

A Stochastic Model of Eight-Hour Running Average Carbon Monoxide Concentrations in Denver, Colorado.

by

William L. Monson

T-3626

A thesis submitted to the Faculty and Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Master of Science (Mathematics).

Golden, Colorado

Date: 6/26/89

Signed: William L. Monson
William L. Monson

Approved: Aldo V. Vecchia
Dr. Aldo V. Vecchia
Thesis Advisor

Golden, Colorado

Date: 6\26\89

Ardel Boes
Dr. Ardel Boes
Professor and
Department Head
Mathematics

ABSTRACT

This thesis develops a stochastic model of the eight-hour running average carbon monoxide (CO) concentration in Denver. The model is used to generate single and joint-station predictions and simulations for two Denver monitors. The single and joint-station prediction errors are analyzed to evaluate the differences between the two predictors. A method of determining confidence intervals for the predictions is presented. Simulation results are used to calculate probability estimates of violations of the national air quality standard (NAQS).

Data for the model consist of hourly carbon monoxide and wind speed for the winter of 1987 and 1986. The model contains three major components; seasonal, meteorologic, and auto-regressive integrated moving average (ARIMA) (Box-Jenkins 1964). The first two components are functions of related variables and time, the third consists of lagged carbon monoxide and random perturbations.

Modeling results show that single-station models were appropriate at both sites for both years. The single-station models preformed favorably under medium levels of carbon monoxide, but seem to have problems detecting higher levels. When higher levels of CO existed, the joint-station model gave better predictions than its single station counterpart. The

largest difference between the predictors occurred when high levels of carbon monoxide (6+) were present or under periods of missing values.

TABLE OF CONTENTS

ACKNOWLEDGEMENT

I would like to thank Dr. Skip Vecchia for all the help that he offered during the development and writing of this thesis.

Personally, I would like to thank my parents for encouraging my education and deep thanks to my wife Michalyn for putting up with all that has gone into tring to finish this work.

Figure LIST OF FIGURES Page

Figure LIST OF FIGURES Page

Table | LIST OF TABLES | Page

LIST OF SYMBOLS

P = probability operation

C8 = eight-hour CO

w = wind speed

F = generic function

K = non-negative constant

Q = sample covariance matrix

j = station index

tr = average traffic volume

b = vector of parameters

a = vector of parameters

t = time index

e = errors from ARIMA component

<> = not equal

TR = traffic volume component

Ln() = natural logarithm

s = variance of errors

T = Box-Cox transformed C8

ih = inversion height

dt = temperature change

rh = relative humidity

L = Box-Cox parameter

W = differenced C8 series

C = hourly CO

m = month index

B = backward operator

M = met. component

S = seasonal component

n = number of samples

p = pie 3.141257

h = hourly index

f = frequency

T = Box-Cox transform

l = predictor lag

' = predictor operator

1.0 INTRODUCTION

The highest levels of carbon monoxide in the Denver area are recorded at the Camp, Carriage, and Welby monitors within the South Platte river basin. There are currently five monitors within the industrial South Platte basin, which are all in north-central Denver. The proximity to pollution sources (downtown traffic and airport) and meteorologic conditions of the area contribute to the elevated levels found from November to February every year. During 1986, as many as twenty readings over the eight-hour NAQS (nine ppm) were recorded with eight sites having two or more violations of the standard. This is in contrast to the hourly NAQS which was seldom violated during the same period. Thus the problem is clearly not hourly emissions but, rather the transportation of pollutants or the lack of it.

Most studies of Denver's carbon monoxide have concentrated on modeling the behavior of the hourly carbon monoxide. An exception to this is the special study that was conducted in 1986/1987 by the Colorado Department of Health (CDH) where summaries of the behavior of the eight-hour series are presented. The special study looked at behavior of Camp's eight hour series over several years. They modeled the second highest reading from year to year. They noted that Camp's second maximum was significantly higher during the special study winter than in past years, namely, 24 ppm and 18 ppm, respectively. Questions arose concerning the representivity of 1986/1987 being a typical pollution season.

This thesis develops a stochastic model of eight-hour running average carbon monoxide concentrations. Unlike other studies of Denver carbon monoxide, this study is solely concerned with the properties of Denver metro eight-hour running average series. The model was developed to demonstrate how stochastic models can be employed to measure and utilize similarities between sites. The model is used to generate single and joint-station predictions and simulations for two Denver monitors. The single and joint-station prediction errors are analyzed to measure the similarities between sites. Summary statistics of the simulated series are used to evaluate the stability of each site.

1.1 MODELING APPROACH

It has been known for some time that carbon monoxide levels are influenced by meteorologic and pollution sources. Meisel and Zeldin (1978) point out that, depending on geographic location, carbon monoxide may be correlated with inversion height, mixing height, wind speed, and surface temperatures. In order to better understand carbon monoxide levels, several people have developed stochastic models that represent concentration levels as a function of three separate elements: time of day, traffic volume, and meteorology. The strength of this approach is that it allows each element's contribution to carbon monoxide to be evaluated independently. Once the contributions are isolated, the carbon monoxide can be adjusted for meteorologic, and/or traffic volume changes that occurred. The most common application of the adjusted data is to evaluate control strategies. But in addition, the adjusted data can be used to evaluate the year to year similarities between monitors.

To separate the elements and their interactions, several meteorologic filters have been developed by the EPA over the past ten years. Two of the most widely used procedures are regression analysis, and categorical trees. Multivariate regression analysis is used to determine and estimate parameters that relate meteorologic and pollution variables to air quality. The variables are determined by their significance to the overall regression variance. Once the parameters are determined, the regression equation can be used to estimate potential carbon monoxide levels given the proper set of input variables.

The categorical approach determines carbon monoxide potential from subjectively identifying and classifying key meteorologic and pollution variables (Meisel and Zeldin 1978). This set of variables is usually determined from a knowledge of local meteorologic and pollution sources. While this procedure can often produce more meaningful results than regression, it requires an extensive knowledge of each monitor's environment. The regression approach will be used in the model presented here.

1.2 OTHER MODELS

Lon-Mu, Hudak, and Tiao (1987) developed a regression model for relating hourly carbon monoxide as a function of time of day, meteorology and traffic volume. The general form of their model is

$$F_1(C_t) = F_2(TR_t) + F_3(M_t) + F_4(S_t) + F_5(C_{t-s}) + e_t. \qquad (1.0)$$

where

$$F_1(C_t) = Ln(C_t + K) \qquad (1.1a)$$

$$F_2(TR_t) = d_1 \; Ln(tr_{t-1}) \qquad (1.1b)$$

$$F_3(M_t) = d_2 \; Ln(dt_{t-1}) + d_3 \; Ln(1/w_{t-1}) \qquad (1.1c)$$

$$F_4(S_t) = d_4 + d_5 \sin(2pt/24) \qquad (1.1d)$$

$$F_5(C_{t-s}) = d_6 \; Ln(C_{t-1} + K) \qquad (1.1e)$$

The model has five components: a transformation of carbon monoxide F_1, traffic volume F_2, meteorology F_3, daily cycle F_4, and ARIMA F_5. The model derives transformed carbon monoxide from the addition of selected variables of meteorology, traffic volume, and lagged carbon monoxide. The model was applied over a series of years to determine the effect of local meteorologic and traffic volume changes on Phoenix carbon monoxide levels. Once determined, the effects were removed and carbon monoxide trend was determined over the control period.

To determine which variables to include in each of the components Lon-Mu, Hudak, and Tiao correlated major meteorologic, traffic, and cyclic variables with carbon monoxide. They found that inverse of wind speed and changes in surface temperature were significantly correlated with Phoenix carbon monoxide levels. Their model included lagged variables of both wind speed and surface temperature change as well as traffic volume. The seasonal component of daily frequency was arrived at by fitting a series of sine curves to the hourly averages.

Their model follows from equations 1.0 and 1.1,

$$Ln(C_t + K) = d_4 + d_6 \, Ln(C_{t-1} + K) + d_1 \, Ln(tr_{t-1}) \qquad (1.2)$$
$$+ d_3 \, Ln(w_{t-1}^{-1} + K) + d_2 \, Ln(dt_{t-1}) + d_5 \sin(2pt/24) + e_t \, .$$

The assumptions of the model are that all logarithmic transforms are normal (carbon monoxide, traffic counts, wind speed, and temperature change), and error terms are normal and independently distributed with constant variance. The constant K has been added to avoid taking a log of a zero value. This type of transform is common among quality control processes where measurement sensitivity is an issue. The natural logarithmic transform is part of a larger set of possible Box-Cox transforms (Box-Cox 1964),

$$T_L(C_t) = \begin{cases} [(C_t + K)^L - 1] / L, & \text{if } L <> 0 \\ Ln(C_t + K), & \text{if } L = 0 \end{cases} \quad . \tag{1.3}$$

The value of L is chosen based on which of the transformed series more closely follows a normal distribution. Once transformed, normal theory can be applied including regression analysis with normal confidence interval projections. However, a true transformation to normality is often difficult to find for environmental data. Large changes in mean levels of outdoor pollutants are usually correlated with seasonal changes. In the case of carbon monoxide, mean levels are significantly higher in the Winter than Summer in most parts of the country. A method of dealing with seasonal differences is to separate pollutants into partitions of similar mean levels. This reduces variances within each group and leads to predictions with significantly smaller confidence intervals. For example in Lou et al. (1986) the data is processed separately by Winter, Summer, Spring, and Fall. The grouping was arrived at by examining mean differentials for several group classifications. i.e. weekend versus weekday, day of week, months of the year. To estimate the parameters in (1.2) involves solving a system of normal equations. Two of the most commonly used methods are general least square (GLS) and maximum likelihood (ML).

1.3 EIGHT-HOUR SERIES

In this section the issues of modeling the eight-hour running average series are discussed. Some of the issues are common to all multivariate modeling and others are unique to the eight-hour series. Contrasts are made with hourly series models when possible.

The eight-hour running average carbon monoxide series is defined as

$$C8_t = (C_{t-7}+C_{t-6}+C_{t-5}+C_{t-4}+C_{t-3}+C_{t-2}+C_{t-1}+C_t) \ / \ n_t \qquad (1.4)$$

where $n_t = 8 -$ (number of missing C_{t-s})

As a result of the averaging process, the eight-hour series is much smoother and less noisy than the hourly series. The smoother eight-hour series has two advantages over the original hourly series. First, it contains fewer missing values and secondly, sites have stronger cross correlation. If the correlation is passed to modeling residuals, the information may be used to improve prediction, even though the single-station residuals are white noise approximations.

A common problem to all multiple regression modeling is that of missing data in the modeled series or independent variables. The eight-hour series will have fewer missing values then the hourly series, but will not be without missing values. The missing data can be caused by several mechanisms including different sampling rates, recording equipment failure or calibration problems. If the missing data is generated because of different sampling rate, a solution is to aggregate over time and produce paired data. The aggregated variables are not necessarily a value of the original series but are often combinations of several variables. Equation (1.4) produces an aggregated hourly series using a moving averaging filter. In the case of meteorologic variables, averaging may not makes sense particularly to wind speed and direction. An alternative to averaging meteorologic variables, is to categorize the meteorologic variables according to pollution potential (Zeldin and Meisel 1978). While aggregation and categorization will provide paired values under different sampling rates, if there was equipment failure there will still be periods of missing values. These remaining missing values are particularly troublesome if the model requires cyclic or lagged parameters to be estimated. To examine the cyclic or lagged behavior of a series requires producing either the autocorrelation function (ACF) and/or spectral density functions (SDF). Both functions require series that have no missing values. A common approach is to use parts of the original series that contain no missing values to produce ACF estimates. The alternative to limiting the ACF to continuous periods, would be to estimate all the missing values and then proceed with ACF calculations. For the model presented here estimates of all eight-hour series ACF and SDF are based on continuous periods of five days or more. A disadvantage of the eight-hour model is that the number of samples that produce the eight-hour values may vary, i.e., $n_t <> n_{t+1}$. Therefore, a model of the eight-hour series must account for these difference in sampling numbers. The problem of different sample numbers is a question of determining if sample size is related to carbon monoxide levels and if so how. A procedure to evaluate this relationship preceeds as follows. First, transform the carbon monoxide to meet normality criteria, and then perform analysis of variance (ANOVA) by sample size. If the differences between mean values are significant, a weighted adjustment to prediction could be assessed. This would let the sample size determine a ratio of the expected value of carbon monoxide series. An alternative to weighting predictions, and the approach used here, is to weight trend errors. The trend residuals errors are weighted inversely proportional to the sample size number ($1/n_t$). This allows non-missing values of the eight-hour series to contribute more to trend estimates.

There are disadvantages associated with modeling any linear combination and the eight-hour series is no exception. The eight-hour series is serially correlated by definition rather than by chance. This serial correlation complicates the lagged component of equation 1.2. Using equation 1.4 and some algebra, the difference in preceding values under continuous sampling period is

$$C8_t - C8_{t+1} = (C_{t-7} - C_{t+1}) / 8 . \qquad (1.5)$$

If the hourly values are consistent for long periods the eight-hour differenced value can be close to zero. Consequently, the first lag ACF correlation becomes close to unity. The strong correlation at lag one can cause the series to be nonstationary.

1.4 SINGLE STATION MODEL

In this section, a general model for the eight-hour carbon monoxide series is presented. The model incorporates some existing methodology of hourly models that were described earlier in section 1.0. The assumptions are similar to the hourly model assumptions, namely, both meteorologic and pollution variables can be transformed to normality ,and residuals are normal and not serially correlated. The eight-hour model was limited to three components; seasonal, meteorologic, and ARIMA,

$$F_1(C8_t) = F_3(M_t) + F_4(S_t) + F_5(C8_{t-s}) + e_t . \qquad (1.6)$$

The functions have been kept basically the same as the model presented earlier. The major changes to the model are the dropping of the traffic components, adding frequencies to the seasonal component, and including different variables in both the meteorologic and ARIMA components. The dropping of the traffic volume parameter was necessary because neither hourly nor average daily traffic volume data was available for both study years. In the future, if the data becomes available it could be added. The seasonal component was modified to reflect the additional correlation at frequencies twelve and six,

$$F_4(S_t) = a_0 + a_1 \sin(2pt/24) + a_2 \cos(2pt/24) + a_3 \sin(2pt/12) + a_4 \cos(2pt/12) + a_5 \sin(2pt/6) \quad (1.7)$$

The meteorologic components contain only functions of lagged wind speed,

$$F_3(M_t) = a_6 \, Ln(w_t + K) + a_7 \, Ln(w_{t-1} + K) \quad (1.8)$$

Surface temperature and inversion height were dropped from the meteorologic component because of lack of data.The additional frequencies and wind parameters were detected by performing multiple regression analysis over series of cosines and sines and lagged wind speeds, respectively. This is the same approach employed by Lon-Mu et al. (1986).

The ARIMA component was changed from the hourly model of AR(1) to a ARIMA(1,1,1),

$$F_5(C8_{t-s}) = (1 + b_0)\, Z_{t-1} - b_0\, Z_{t-2} + b_1 \, e_{t-1} \quad (1.9)$$

where, $Z_t = Ln(C8_t + K)$.

This assumes that the eight-hour series is nonstationary with a correlated random component. The ARIMA(1,1,1) model was selected after a visual check of the ACF and PACF of the eight-hour series indicated that the differenced eight-hour series could be model as a ARMA(1,1). Figures 1.1 and 1.2 show the ACF and PACF of the differenced series for the Welby site in November of 1987. Note the strong correlation at lag one for both functions.

To increase the seasonal accuracy, two mean adjustment schemes are proposed:

$u_j = \text{mean}(c_t)$ where j indicates Nov.,Dec.,Jan.,Feb (1.10)

$u_j = \text{mean}(c_t)$ where j indicates monday,... sunday . (1.11)

While Lou et. al. (1986) evaluated several adjustment schemes, here only monthly and daily schemes are considered. The choice of equation 1.10 or 1.11 will be based on which produces the smaller prediction variance. By accounting for different means, the residual variance within groups will be reduced and lead to more consistent predictions. Hence, the first component of (1.6) is

$$F_1(C8_t) = \text{Ln}(C8_t - u_j + K) \qquad (1.12)$$

and Z_t in (1.9) becomes $Z_t = \text{Ln}(C8_t - u_j + K)$.

The parameters for (1.6) are estimated in stages. First, the eight-hour series are transformed according to 1.12. Next, the a_j's are estimated using a GLS routine, i.e, GLM in SAS. The residuals of these estimates are weighted inversely proportional to the number of samples. The seasonal and meteorologic contributions are then removed from (1.6) by subtracting their estimated values. Finally the b_j's are estimated using a ML routine (Bockwell and Davis 1987).

Figure 1.1 Autocorrelation of Differenced Eight-Hour Carbon Monoxide for Welby November 1987

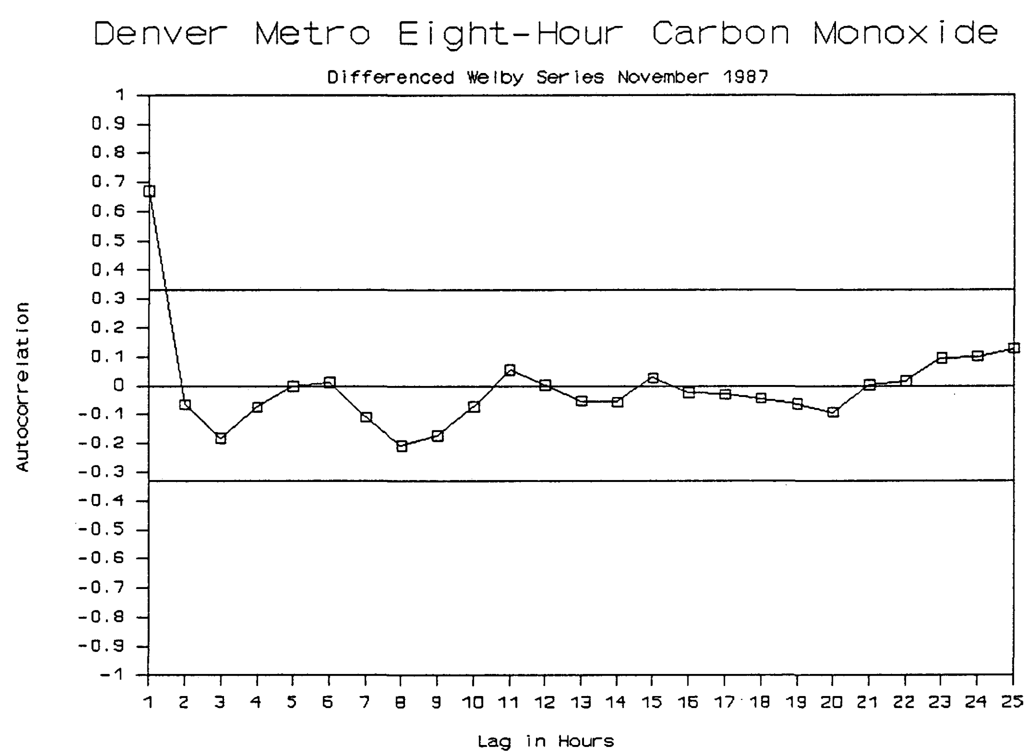
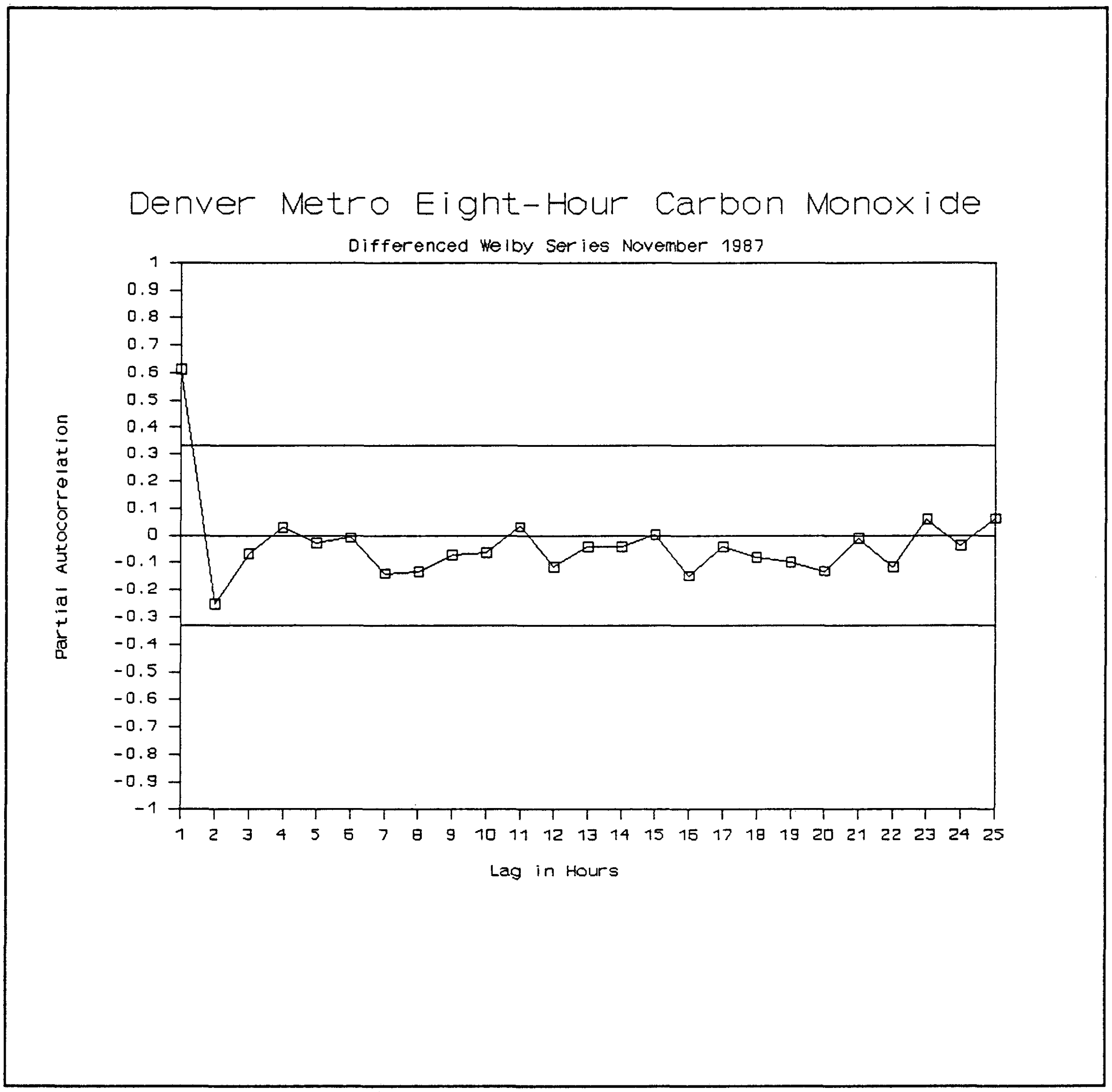


Figure 1.2 Partial Autocorrelation of Differenced Eight-Hour Carbon Monoxide for Welby November 1987



2.0 MODEL APPLICATION - DENVER 1986 AND 1987

In this section, the single station model for the eight-hour carbon monoxide levels described in the above section is applied to two Denver sites: Welby and Carriage. In addition, a joint model is presented using single station residuals. The results of this section include evaluation of the differences between monthly and daily mean adjustment schemes, measuring the predictability of each station, and comparison between single and joint-station predictions.

2.1 MODELING SCOPE

The two sites were chosen from among seven possible state operated CO monitors (Table 2.1). The two stations were isolated because of their high levels and strong correlation. In fact the two sites have the strongest hourly correlation at lag zero from year to year of any two sites within the Denver metro area over the study years. Site descriptions from Macrae (1988) for both the study sites follow.

Carriage (080580014F01): 2325 Irving

The Carriage site derives its name from the location in the middle of a city block (i.e., along the alley at mid-block) on a plot of ground owned by Denver and known as a "carriage lot". This site is at an elevation of 5300 feet MSL and 2.5 miles west of the downtown CBD. The terrain is relatively flat and residential housing surrounds the site at distances of approximately 75 to 100 feet. The site is isolated enough that it is unaffected by localized traffic patterns. Topographically the site is located along a ridge out of the Platte River Valley that cuts though the CBD area. However, it is only 80-100 feet above the valley floor and well within the influence of major traffic influx into and out of the CBD.

Welby (080220001F01):

Sited in July 1973, the location (8 miles north of the Denver CBD) is ideal to measure nightime drainage of the Denver air mass and thermally driven midday up-river reversal. The site is within 50 feet of the South Platte River which is the lowest area of the City. The area is very open and predominantly agricultural. High levels of all pollutants monitored have occurred. CO monitoring was initiated in November, 1986 as a special study paremeter. Upon completion of the study and examination of the database, the value of maintaining a CO effort as a SLAMS location was confirmed. This site has proven to be even more valuable than orginally anticipated and monitoring is projected to continue at least at the present effort.

Before performing site comparisons, it is important to consider the compatibility of the sites. Bailey (1987) points out several situations with environmental monitoring, where site comparisons are inappropriate. However, in this case the comparisons seem to be justified since both sites are the same monitor type (neighborhood), located in the same topographic regimes (South Platter River basin), and sensitive to the same major pollution sources and wind patterns. The single-station model was applied using two years of winter data from the two monitors. The four winter months of November, December, January, and February correspond to Denver's "CO high pollution season" in which elevated levels of carbon monoxide are expected. The study was limited to two years of four months each because of the vast volume of data that is involved. The total number of readings used in this study was approximately twenty-four thousand. All the data was retrieved from the Environmental Protection Agency's air quality retrieval system (AIRS).

The comparison between monthly and daily adjustment schemes (1.10 and 1.11) was conducted as a preliminary stage to producing the single-station projections. The adjustment schemes were evaluated on a portion of the data set and the optimal adjustment used to produce the remaining single-station estimates. The portion of data used for the adjustment comparisons included Welby for November 1987.

The joint-station predictions were constructed from the single station residuals. Because the stations have many similarities and are geographically near each other, they are assumed to have similar pollution patterns. Therefore, if the error in predicting carbon monoxide at any one station (Carriage) is large, then the errors at the other stations should be large. To explain the relationship between the single and joint-station models, the single-station model is written as

$$Z_t = U_t + (1 + b_0)\ Z_{t-1} - b_0\ Z_{t-2} + b_1\ e_{t-1} + e_t\ , \qquad (2.1)$$

where $U_t = F_3(M_t) + F_4(S_t)$ and $Z_t = Ln(C8_t - u_j + K)$ (see 1.7, 1.8, 1.9 and 1.12) .

To produce the joint estimates from the single station residuals, a multivariate structure was added to equation 2.1 (Vecchia 1983). To construct the needed structure, let $e_t^{(1)}$ and $e_t^{(2)}$ be the residuals from Welby and Carriage respectively. From the assumption of multivariate normality of e_t, the sequence $[e_t^{(1)}, e_t^{(2)}]$ are independent and identically distributed bivariate normal random vectors, with mean zero and covariance matrix Q. The procedure of updating the single station predictions for Welby is as follows:

(1) Determine single-station predictions.

(2) Calculate error of Carriage : $e_t^{(2)}$

(3) Find $h = E[e_t^{(1)} | e_t^{(2)}]$

(4) Let new prediction be $Z'^{(1)}_t = Z_t^{(1)} + h$

Another advantage of defining the error covariance structure, is that it can be used to simulate joint-station series. The simulation data are possible realizations of each monitor. They represent alternative series that could have resulted under the same given meteorologic conditions.

There are many ways to use the simulations to make statements of both the model and the monitors. One way of evaluating the model adequacy is to determine how well it reproduces monthly and daily means and standard deviation of recorded data. In addition, the means and standard deviations of the simulations can reflect the stability of the monitors. The procedure for generating simulated values of the joint-station series at time t is as follows: (1) generate a bivariate normal random vector $[e_t^{(1)}, e_t^{(2)}]$ out of the appropriate joint-station error distribution, (2) add components U_t and $F_5(C_{t-s})$ to the residuals to obtain $Z_t^{(1)}$ and $Z_t^{(2)}$, and 3) untransform $Z_t^{(\cdot)}$ to original units $C8_t^{(\cdot)}$.

2.2 PRELIMINARY DATA ANALYSIS

In this section, some preliminary statistics and the results from the single station model are presented. The results include comparisons between the monthly and daily adjustment schemes, the single-station predictions for Welby and Carriage, and joint-station and single station predictions for Welby.

Before proceeding with the fitting of the model and comparisons, some preliminary statistics were produced to explore the behavior of the study sites. The original eight-hour carbon monoxide data are graphed in figures 2.1 though 2.6. Basic statistic by month and day are presented for both wind speed and CO in figures 2.7 though 2.16. Evidence of the translation of the hourly series to the eight-hour series is shown in figures 2.17 and 2.12. Some of the preliminary analysis contains data from the Camp site. This was done to contrast the similarities between the Welby and Carriage sites. Both the Welby and Carriage sites appear to have the same cyclic behavior in both hourly and eight-hour series, namely two peaks. The eight-hour series difference from the hourly series in that the peaks are smoother and lagged. Table 2.2 shows how the correlation coefficient of the hourly series is lower than its eight-hour counterpart.

The higher carbon monoxide levels of 1986 compared to 1987 correspond with the lower wind speed from 1986 to 1987. This is due to the inverse correlation between carbon monoxide and wind speed. The most pronounced feature from year-to-year is the similar shape of the carbon monoxide versus the time of day curve.(Figure 2.11 and Figure 2.12) This similarity would suggest a stable process year-to-year with changes in mean levels. If this is the case the single-station model might have overlapping parameter estimates year-to-year. Another indicator of similarity is the correlation coefficient.

To test the assumptions of the normality, histograms of the transformed wind speed and carbon monoxide data were compared with normal distributions. A chi square goodness of fit test indicated that transformed data most closely approximated normality with the Box-Cox transform of L = 0 (i.e. natural logrithmic transform). Figures 2.18 and 2.19 show an example of before and after transformed carbon monoxide data set.

Assuming that normality assumptions have been met, the next step is to produce the single station predictions. To produce estimates of the a_j's and b_j's involved in equations 1.7, 1.8, and 1.9, the components are solved in stages as follows.

Steps to Single Station Model

1. Transform data to normal
2. Mean adjust data
3. Estimate met. and seasonal parameters
4. Estimate ARIMA parameters
5. Untransform predictions to original units.

2.3 MODELING RESULTS

As discussed in section 2.2, the evaluation of monthly vs daily adjustments is considered for the Welby site in November for both study years. The difference between the adjustment schemes can be measured by comparing their single station predictions. The difference between the monthly and daily adjustment single station predictions occur at step two and step five. At step two the proper mean is subtracted from the transformed data. Then at step five the proper mean is added to the prediction to reverse the process.

Results showed very little difference between parameter estimates based on monthly and those on daily adjustments. Table 2.3 displays the estimates for Welby in November 1987 for the two adjustment schemes. Figure 2.20 shows the predictions for Welby for November 1987 using the monthly and daily adjustment. Prediction results indicated little difference between the two adjustment schemes. Because of the lack of difference between the adjustments, the monthly adjustment was selected for the rest of the study.

The seasonal and meteorological components were estimated using a GLS routine (the GLM routine in SAS, see appendix), while the ARIMA component was estimated with a ML procedure (PEST Bockwell and Davis Copyright 1980). Missing values in the meteorologic data were ignored, where as the ARIMA component was estimated using periods of non missing values. Because there was a lack of a period of two-hundred or more for the months of January and February in both years, the estimates for those months are incomplete. Table 2.4 shows the parameters estimates for both study years.

Residuals from the single-station model were examined to ensure that the following assumptions were not violated: (1) the residuals are mutually independent; and (2) they are normally distributed with mean 0 and constant variance. A visual inspection of graphs of the residuals versus time indicated no abnormalities. In addition, graphs of the autocorrelations of the residuals showed that the residuals had no significant serial correlations. Graphs of the ACF of the residual from Welby for November 1986 and 1987, along with 95-percent confidence limits assuming white noise, are shown in Figures 2.21 and Figure 2.22 respectively. Histograms of the residuals showed that the residuals appeared to be normally distributed, but that some differences exist in the mean and standard deviations among the months. An example of the histogram differences for the Carriage site are shown in figure 2.23 and figure 2.24. The differences were considered small enough for the model to be consisted appropriate. After the model was determined for each station, correlation of the residuals were calculated between Welby and Carriage for each month. The sample size to calculate the ARIMA components for November and December of 1986 and 1987 were 144, 121, 145, and 168 respectively. The correlation for November, December of 1986 and 1987 were 0.71, 0.68, 0.81, and 0.68, respectively. When compared to white noise the correlations are significantly high and therefore can be used to produce better joint-station predictions (figure 2.25). Following the procedure outlined in section 2.2 joint-station and single-station simulations were calculated. One method of evaluating the model adequacy is to determine how well the models reproduces the violation probabilities. Table 2.5 depicts average violations probabilities of the single and joint simulations. The single-station violation predictions are consistently lower than the actual recorded values. On the other hand, the joint-station model seems to represent the violation probabilities more accurately.

Table 2.1 Denver Metro CO Network for 1986-87

Denver Metro Carbon Monoxide
Sampling Network for 1986-87

Site Name	City	Class
Welby	Denver	Neighborhood
Camp	Denver	Micro
Carriage	Denver	Neighborhood
Auraria	Denver	Middle
Aurora	Aurora	Neighborhood
Highlands	Littleton	Neighborhood
Boulder	Boulder	Micro
Arvada	Arvada	Neighborhood

Table 2.2 Correlation Matrix for Winter 1987

Carbon Monoxide

	Welby	Camp	Carriage
Hourly			
Welby	1	0.53	0.81
Camp	0.53	1	0.66
Carriage	0.81	0.66	1
Eight hour			
Welby	1	0.66	0.90
Camp	0.66	1	0.76
Carriage	0.90	0.76	1

Table 2.3 Monthly and Daily Parameter Estimates

Welby November 1987

Parameter	Monthly Adjusted	Daily Adjusted
a0	4.6790	4.6826
a1	0.0023	0.0022
a2	0.0111	0.0110
a3	0.0079	0.0080
a4	0.0051	0.0051
a5	-0.0027	-0.0028
a6	0.0409	0.0387
a7	0.1540	0.1436

Table 2.4 Welby and Carriage Parameter Estimates
Winter 1986 and 1987

	November 1987		December 1987	
Parameter	Welby	Carriage	Welby	Carriage
a0	4.6790	6.2252	4.7376	6.6501
a1	0.0023	0.0025	0.0025	0.0064
a2	0.0111	0.0000	0.0081	-0.0027
a3	0.0079	0.0155	-0.0012	0.0052
a4	0.0051	0.0067	0.0101	0.0138
a5	-0.0027	-0.0011	-0.0008	0.0020
a6	0.0409	0.1680	0.0673	0.2103
a7	0.1540	0.4244	0.1435	0.4903
b0	0.102	0.134	0.122	0.156
b1	0.365	0.224	0.556	0.345
	November 1986		December 1986	
a0	4.3096	4.8011	3.8726	6.6711
a1	0.0050	0.0083	0.0039	0.0038
a2	0.0046	0.0002	0.0212	0.0089
a3	0.0080	0.0118	0.0066	0.0139
a4	0.0052	0.0095	0.0109	0.0158
a5	-0.0022	-0.0053	0.0006	-0.0002
a6	0.1435	-0.0392	-0.0991	0.2077
a7	0.0873	0.2664	0.0894	0.4977
b0	0.103	0.216	0.673	0.125
b1	0.476	0.002	-0.302	0.012

Table 2.4 (CON'T)

Parameter	January 1987 Welby	January 1987 Carriage	February 1987 Welby	February 1987 Carriage
a0	4.6301	6.3208	4.3946	5.0419
a1	-0.0050	-0.0036	-0.0922	-0.0081
a2	-0.0096	-0.0172	-0.0006	-0.0021
a3	-0.0030	0.0008	-0.0049	-0.0036
a4	0.0082	0.1291	0.0038	0.0057
a5	-0.0014	0.0010	-0.0004	-0.0010
a6	0.0692	0.1973	0.0316	0.0832
a7	0.1151	0.4187	0.0919	0.2058
b0	-	-	-	-
b1	-	-	-	-

Parameter	January 1986 Welby	January 1986 Carriage	February 1986 Welby	February 1986 Carriage
a0	4.3619	4.9685	4.3778	5.0582
a1	-0.0041	-0.0092	-0.0070	-0.0022
a2	0.0046	-0.0053	0.0010	0.0019
a3	-0.0016	-0.0001	-0.0041	-0.0036
a4	0.0063	0.0036	0.0040	0.0085
a5	0.0004	-0.0055	0.0007	0.0000
a6	0.0080	0.1244	0.0248	0.0671
a7	0.1083	0.1459	0.0943	0.2262
b0	-	-	-	-
b1	-	-	-	-

Table 2.5 Percent of Readings Over NAQS for Welby and Carriage for Winter 1986 and 1987

Actual Percent Over NAQS

	Welby		Carriage	
	1987	1986	1987	1986
November	0.1	0	0.4	0.3
December	0.9	0.5	2.5	0.8
January	0.1	0	0.5	0
February	0	0	0	0
Total	1.1	0.5	3.4	1.1

Single-station Simulated Percent Over NAQS

	Welby		Carriage	
	1987	1986	1987	1986
November	0	0	0.25	0.1
December	0.5	0.27	3.1	0.6
January	-	-	-	-
February	-	-	-	-

Joint-station Simulated Percent Over NAQS

	Welby	
	1987	1986
November	0	0.1
December	0.7	0.6
January	-	-
February	-	-

Figure 2.1 Eight-Hour Series for Welby and Carriage November 1987

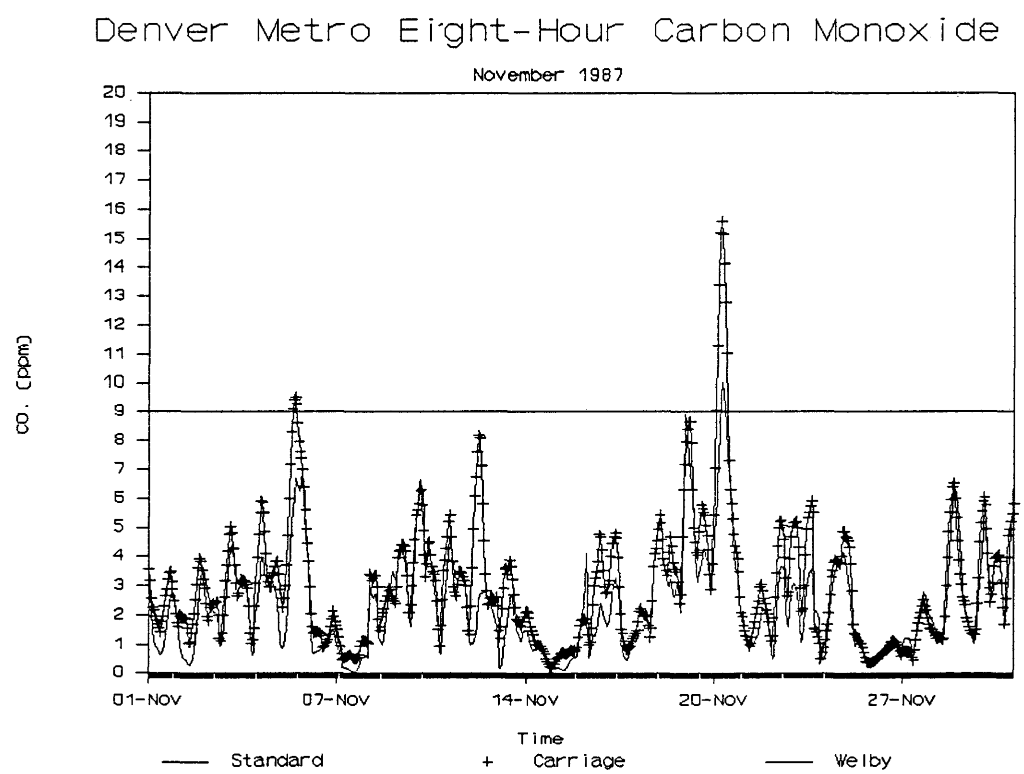


Figure 2.2 Eight-Hour Series for Welby and Carriage December 1987

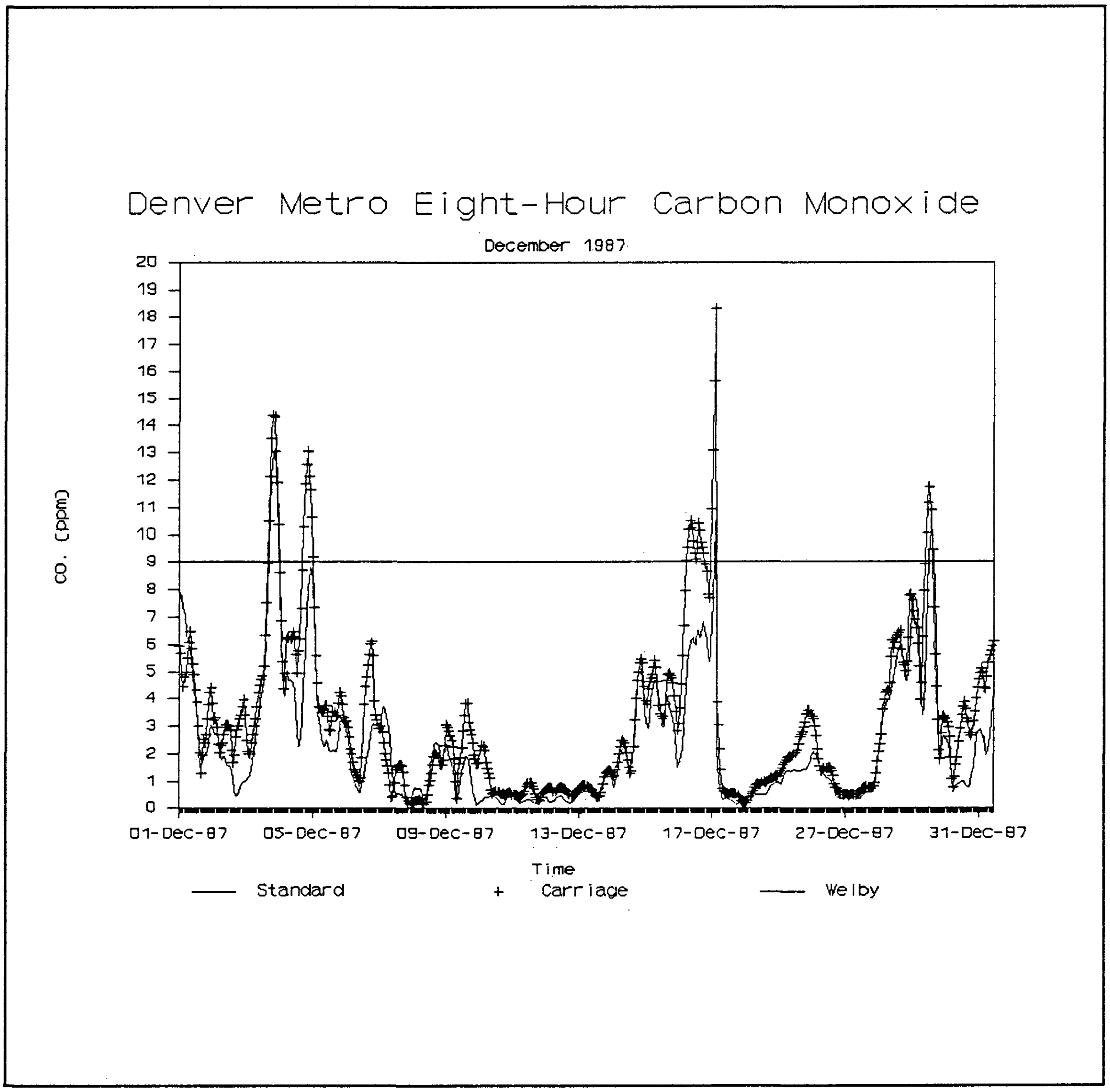


Figure 2.3 Eight-Hour Series for Welby and Carriage
Janurary 1988

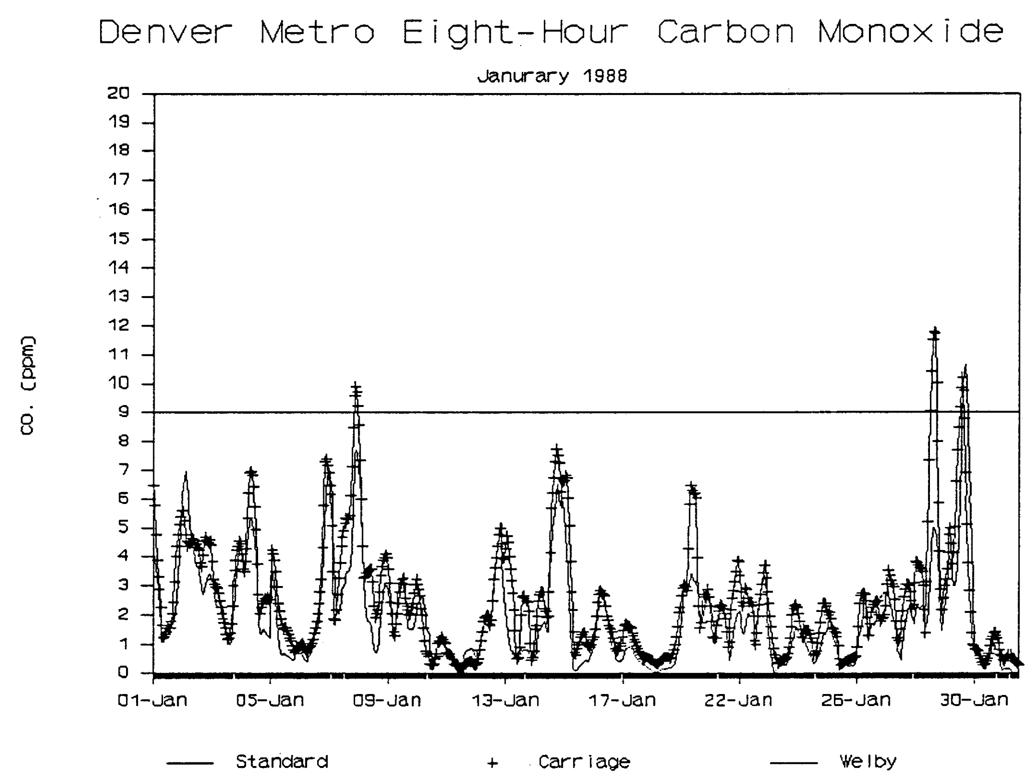


Figure 2.4 Eight-Hour Series for Welby and Carriage February 1988

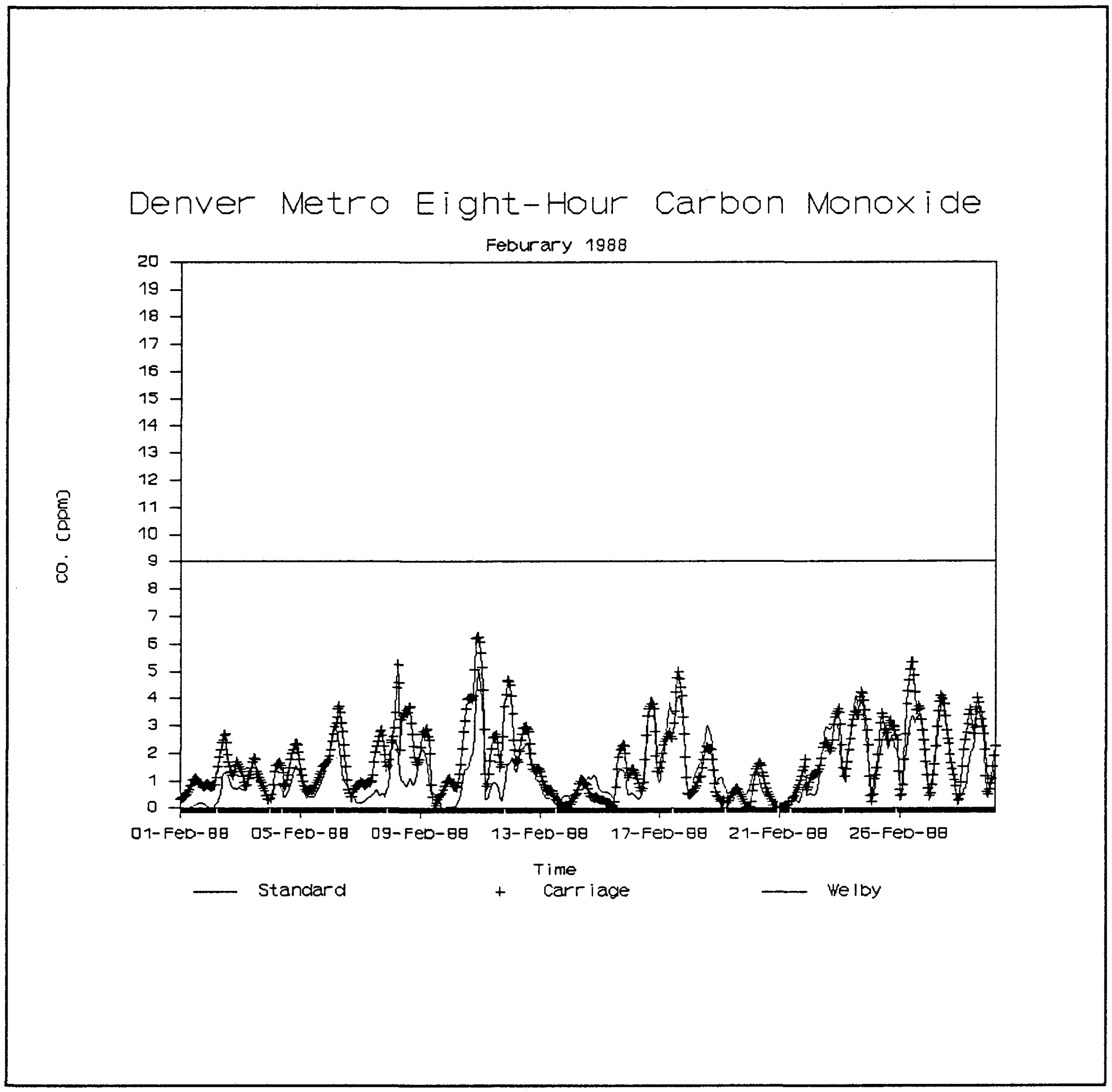


Figure 2.5 Eight-Hour Series for Welby and Carriage November 1986

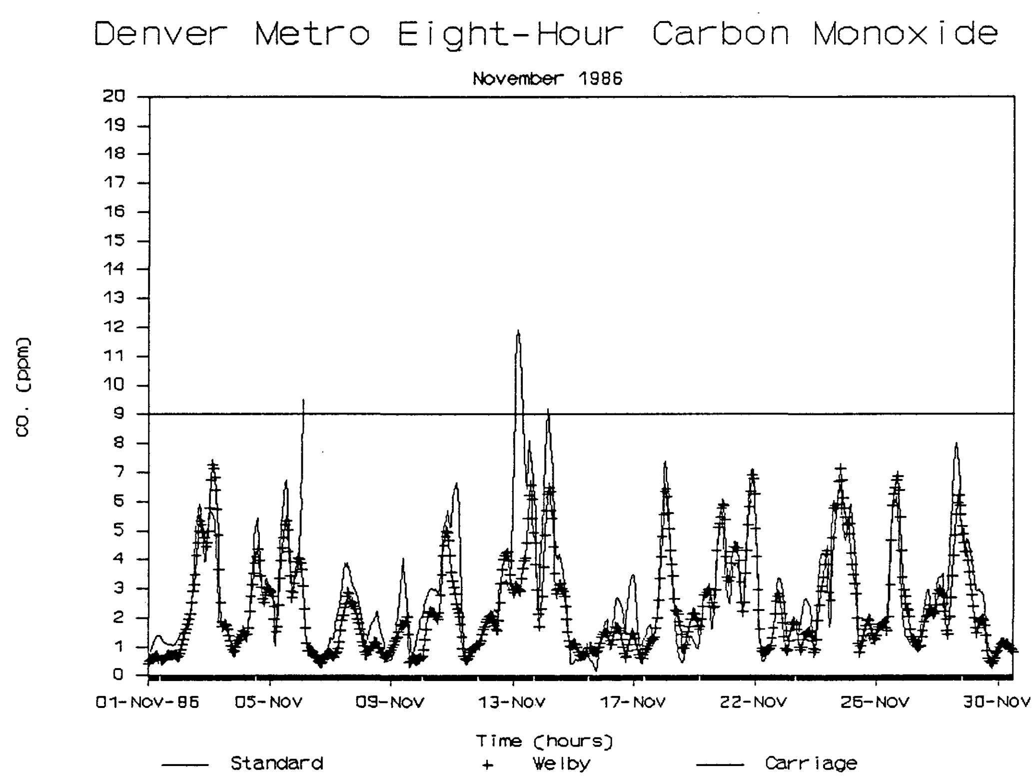


Figure 2.6 Eight-Hour Series for Welby and Carriage December 1986

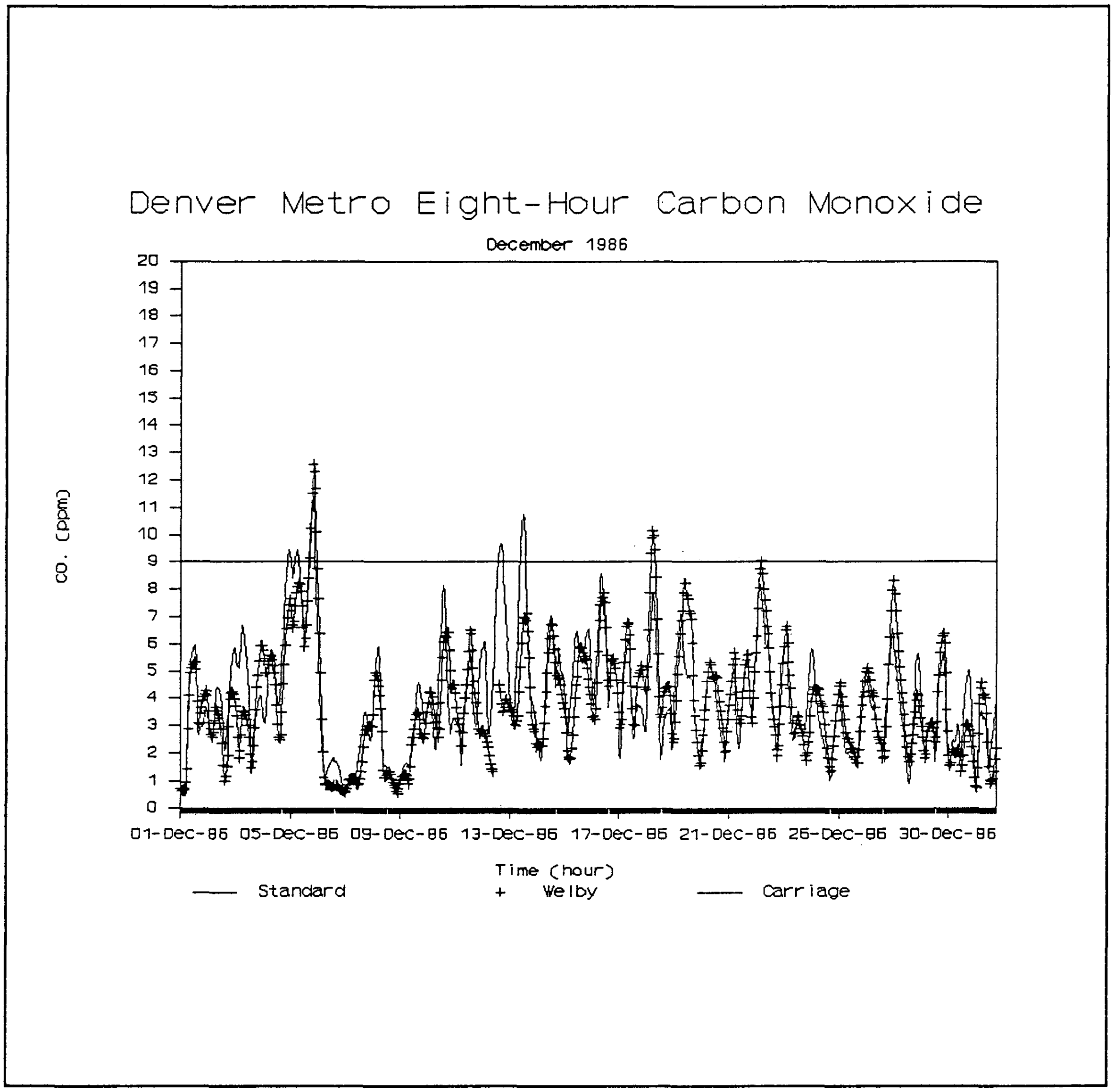


Figure 2.7 Eight-Hour Daily Averages for Welby, Camp, and Carriage Winter 1986

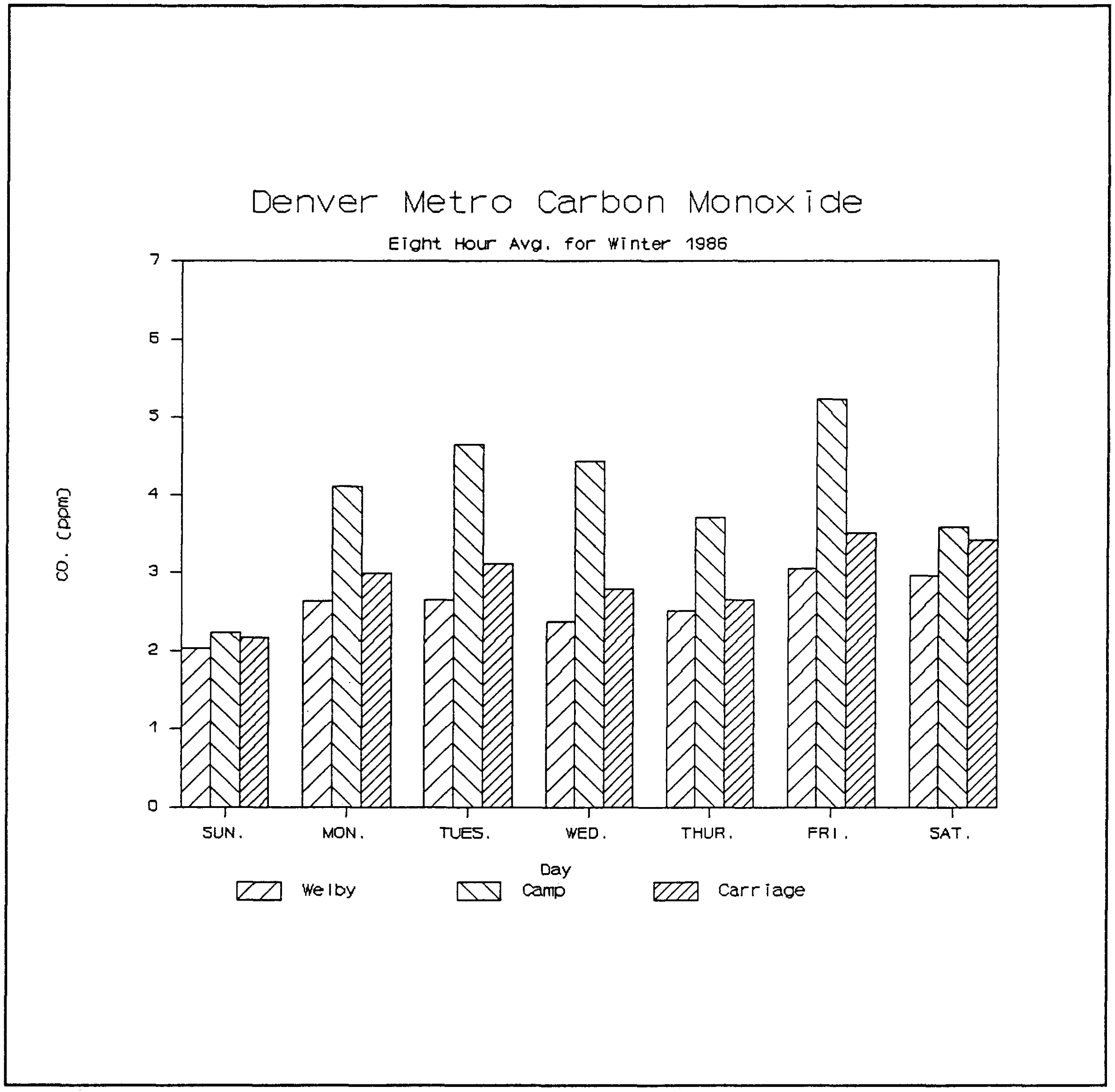


Figure 2.8 Eight-Hour Daily Averages for Welby, Camp, and Carriage Winter 1987

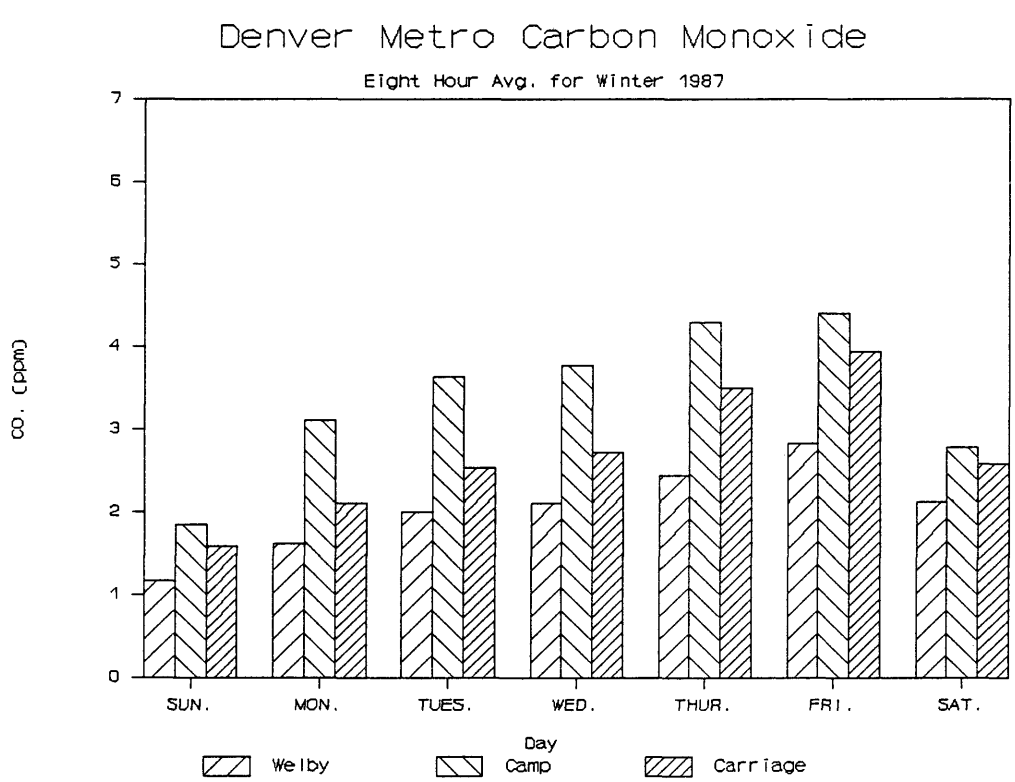


Figure 2.9 Eight-Hour Monthly Averages for Welby, Camp, and Carriage Winter 1986

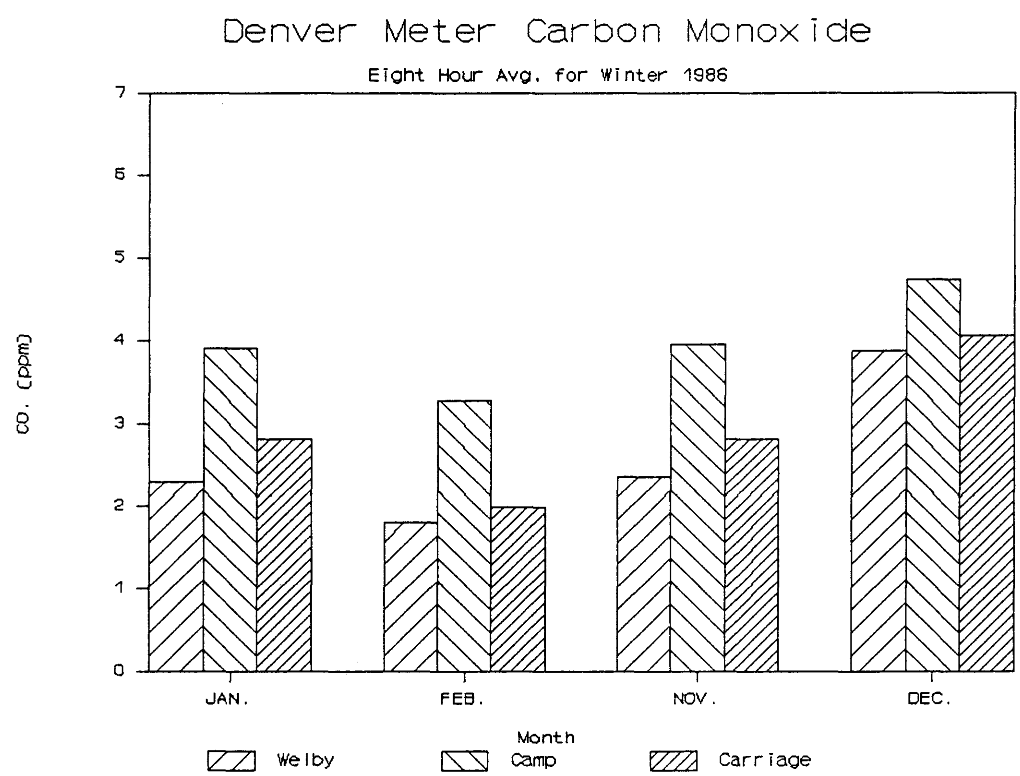


Figure 2.10 Eight-Hour Monthly Averages for Welby, Camp, and Carriage Winter 1987

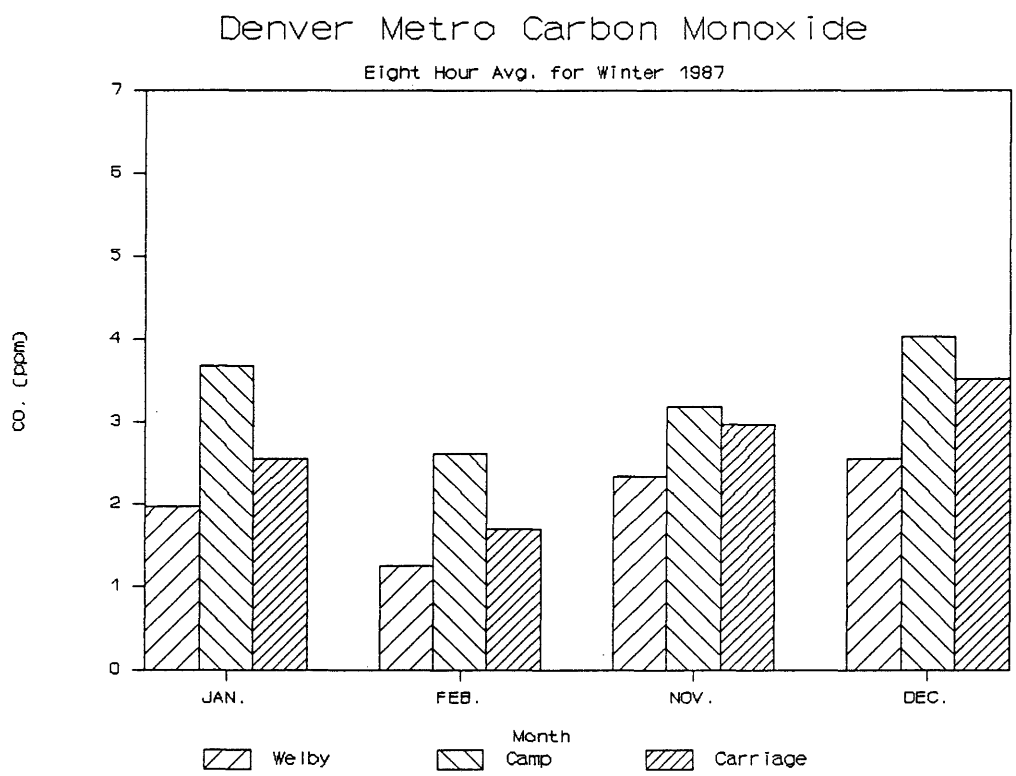


Figure 2.11 Eight-Hour Hourly Averages for Welby, Camp, and Carriage Winter 1986

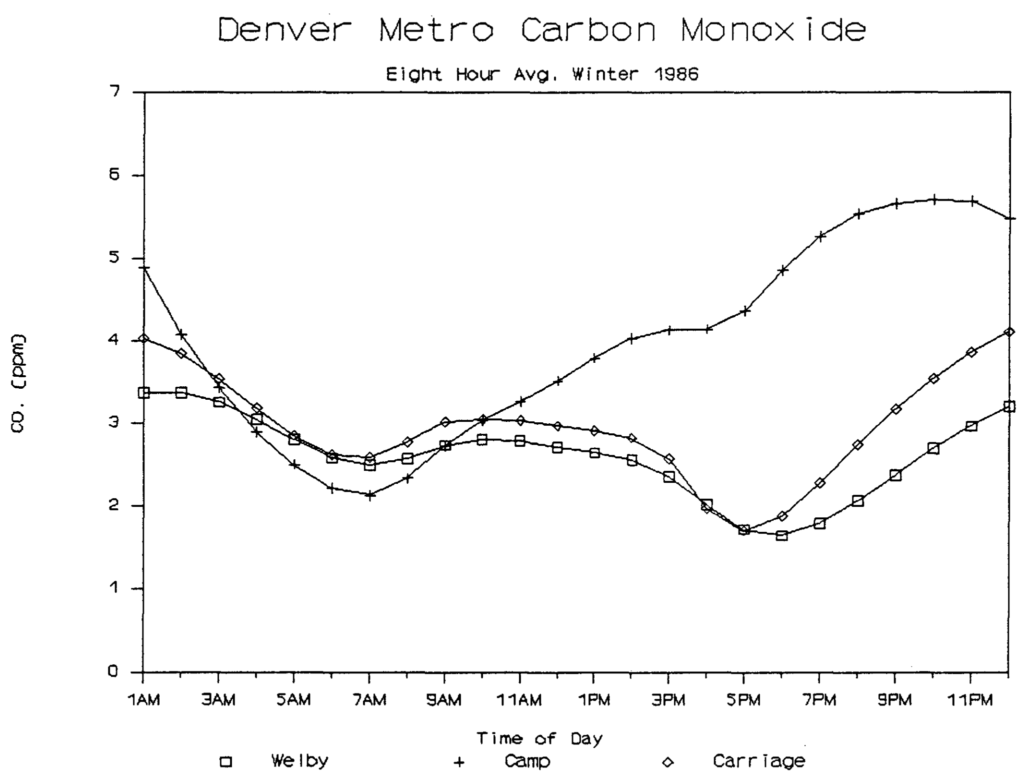


Figure 2.12 Eight-Hour Hourly Averages for Welby, Camp, and Carriage Winter 1987

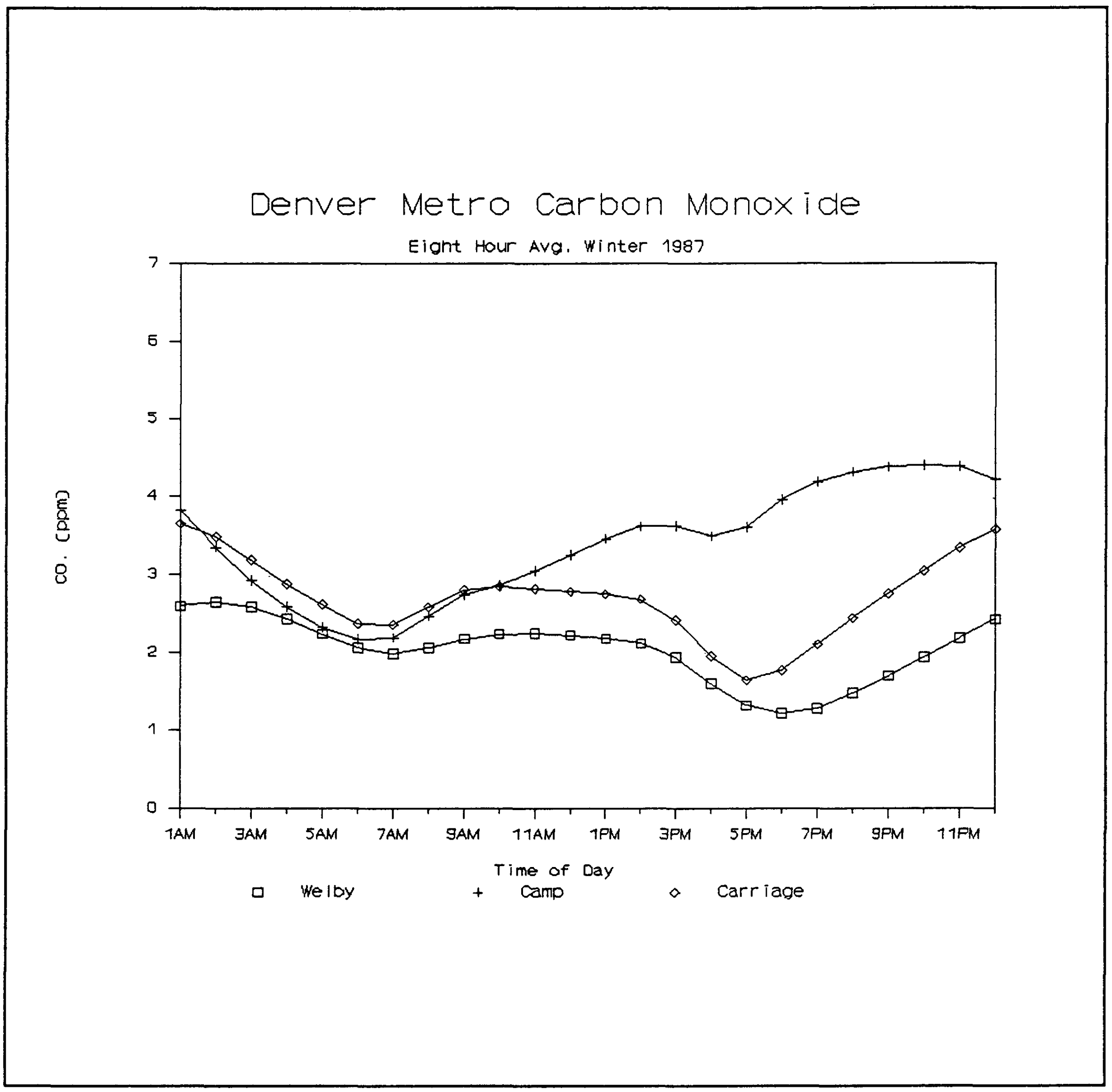


Figure 2.13 Daily Wind Speed Averages for Welby, Camp, and Carriage Winter 1986

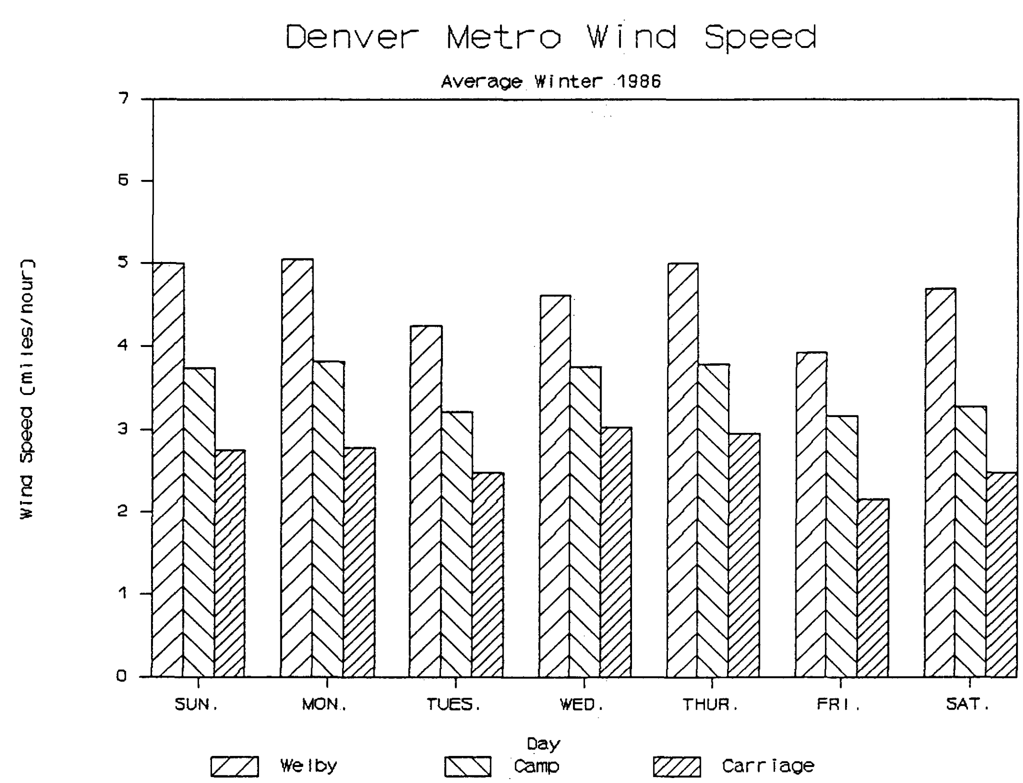


Figure 2.14 Daily Wind Speed Averages for Welby, Camp, and Carriage Winter 1987

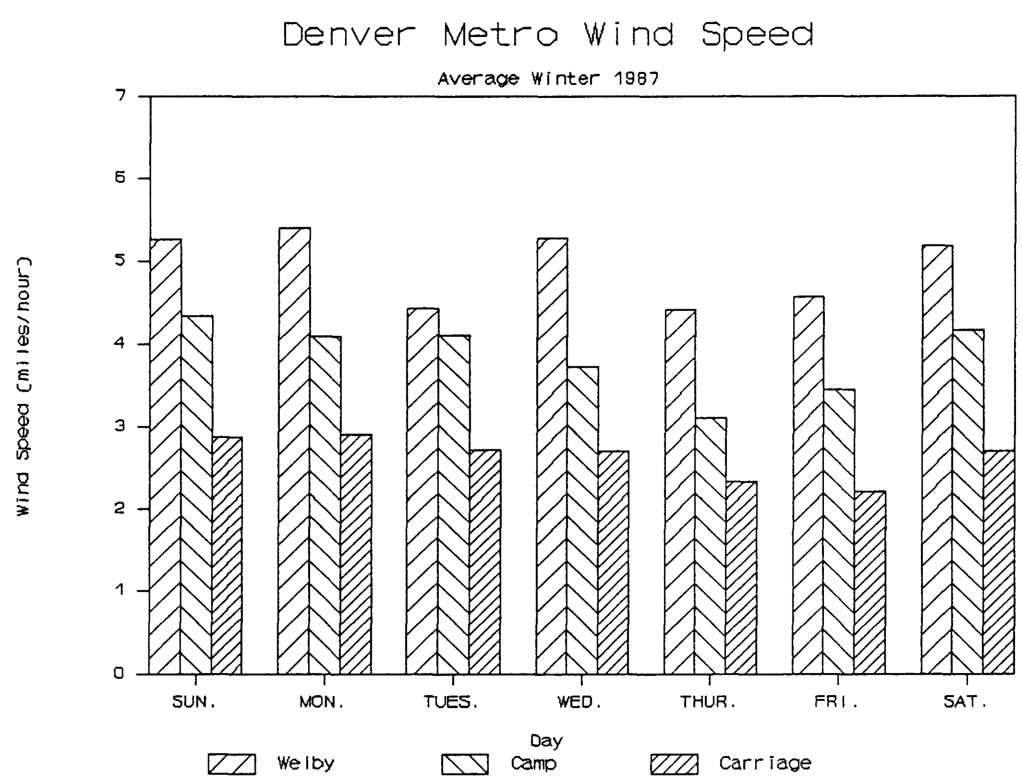


Figure 2.15 Hourly Wind Speed Averages for Welby, Camp, and Carriage Winter 1986

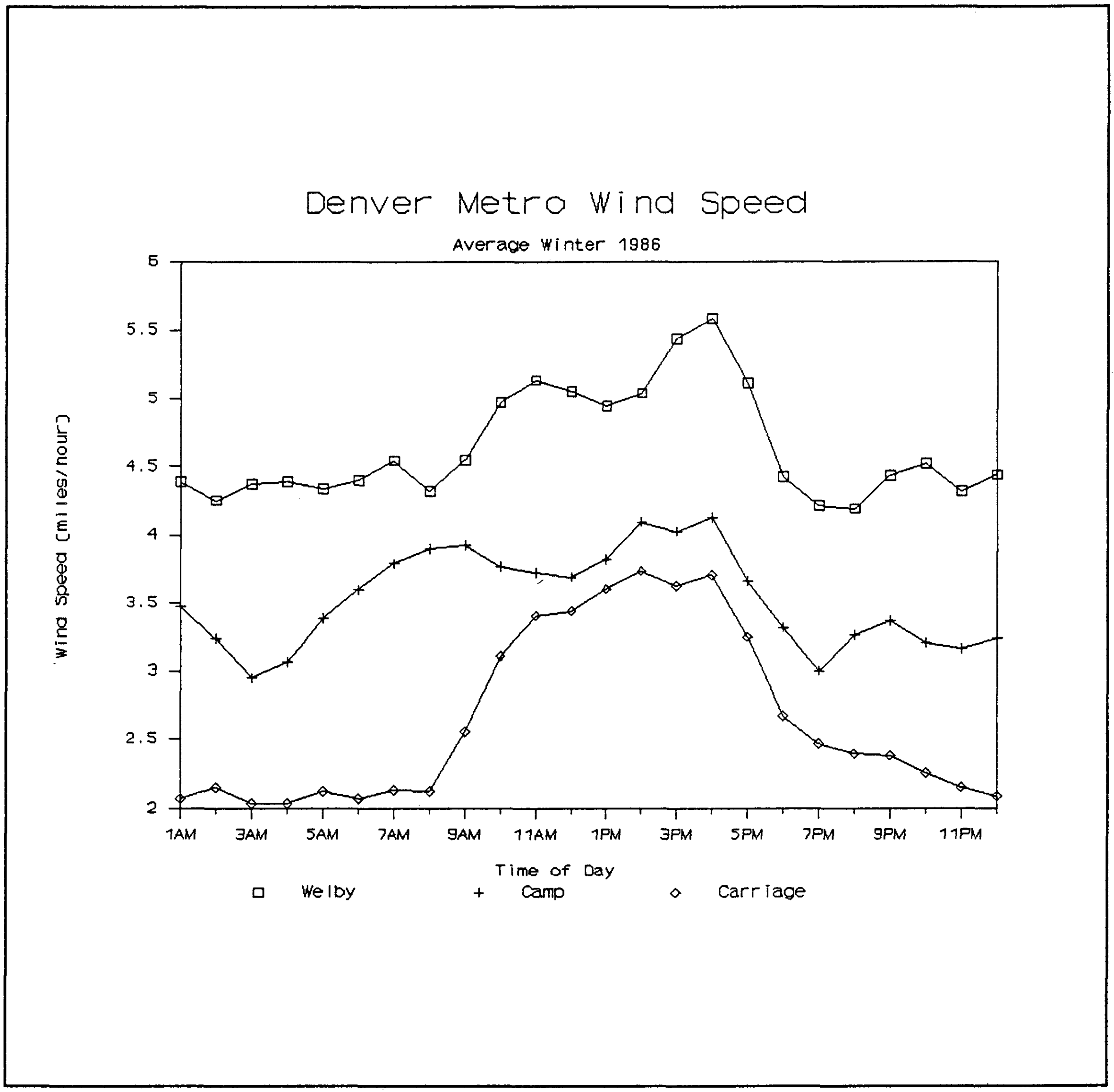


Figure 2.16 Hourly Wind Speed Averages for Welby, Camp, and Carriage Winter 1987

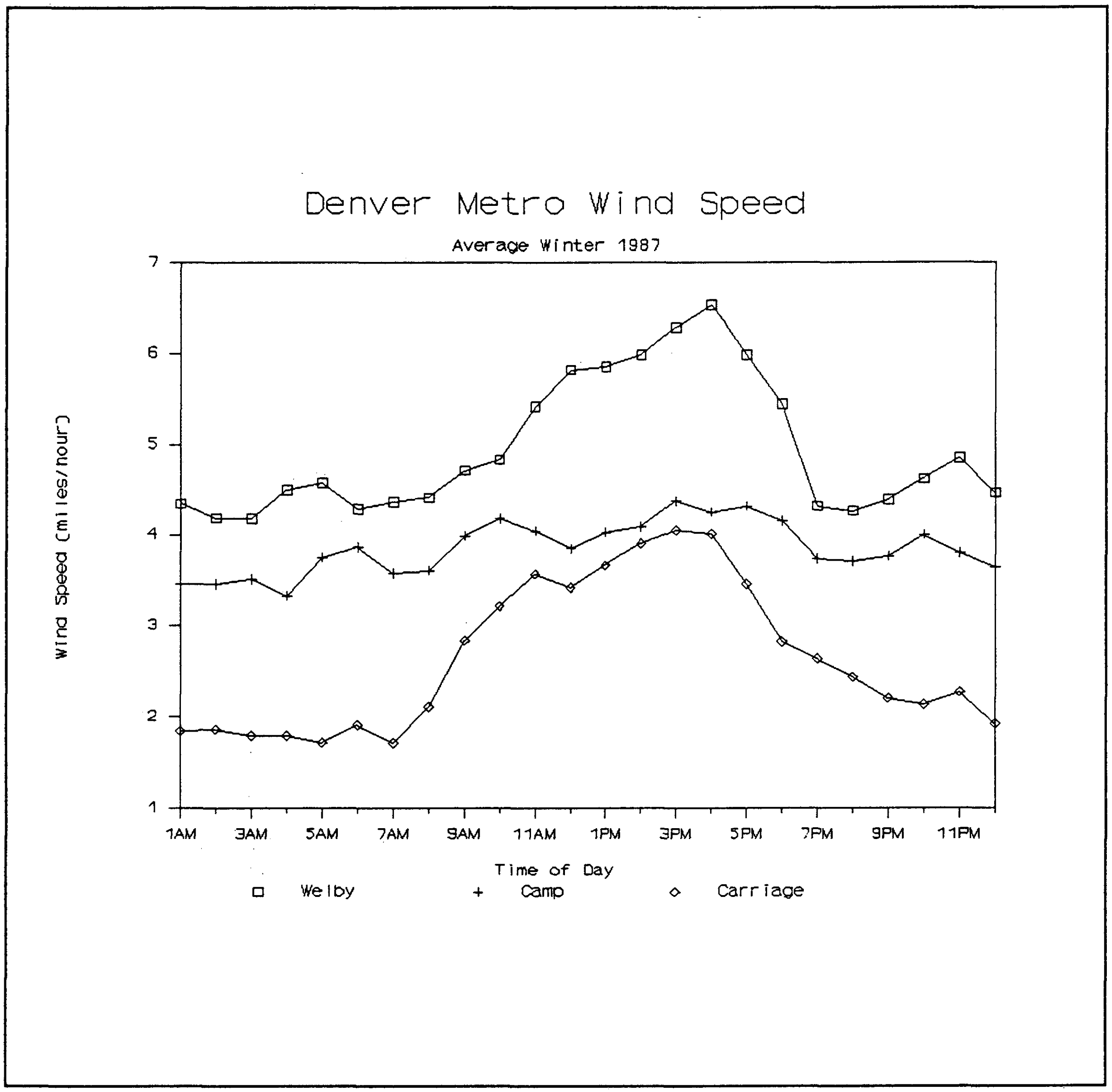


Figure 2.17 Hourly Carbon Monoxide Hourly Averages for Welby and Carriage Winter 1987

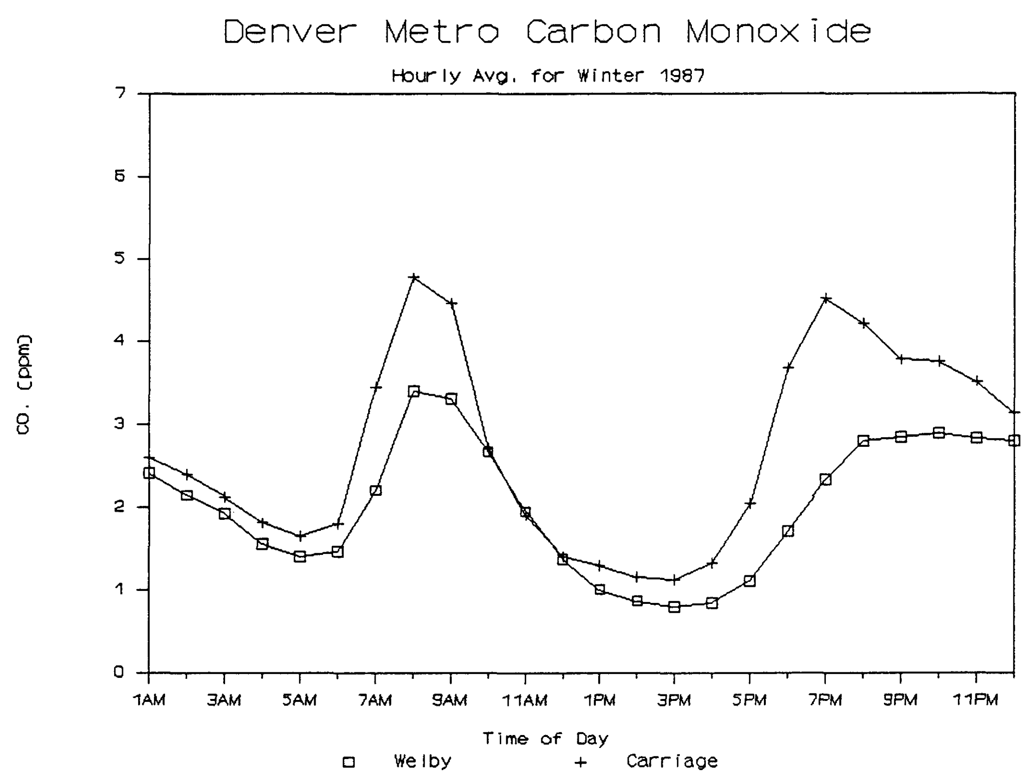


Figure 2.18 Histogram of Welby Untransformed Eight-Hour Carbon Monoxide

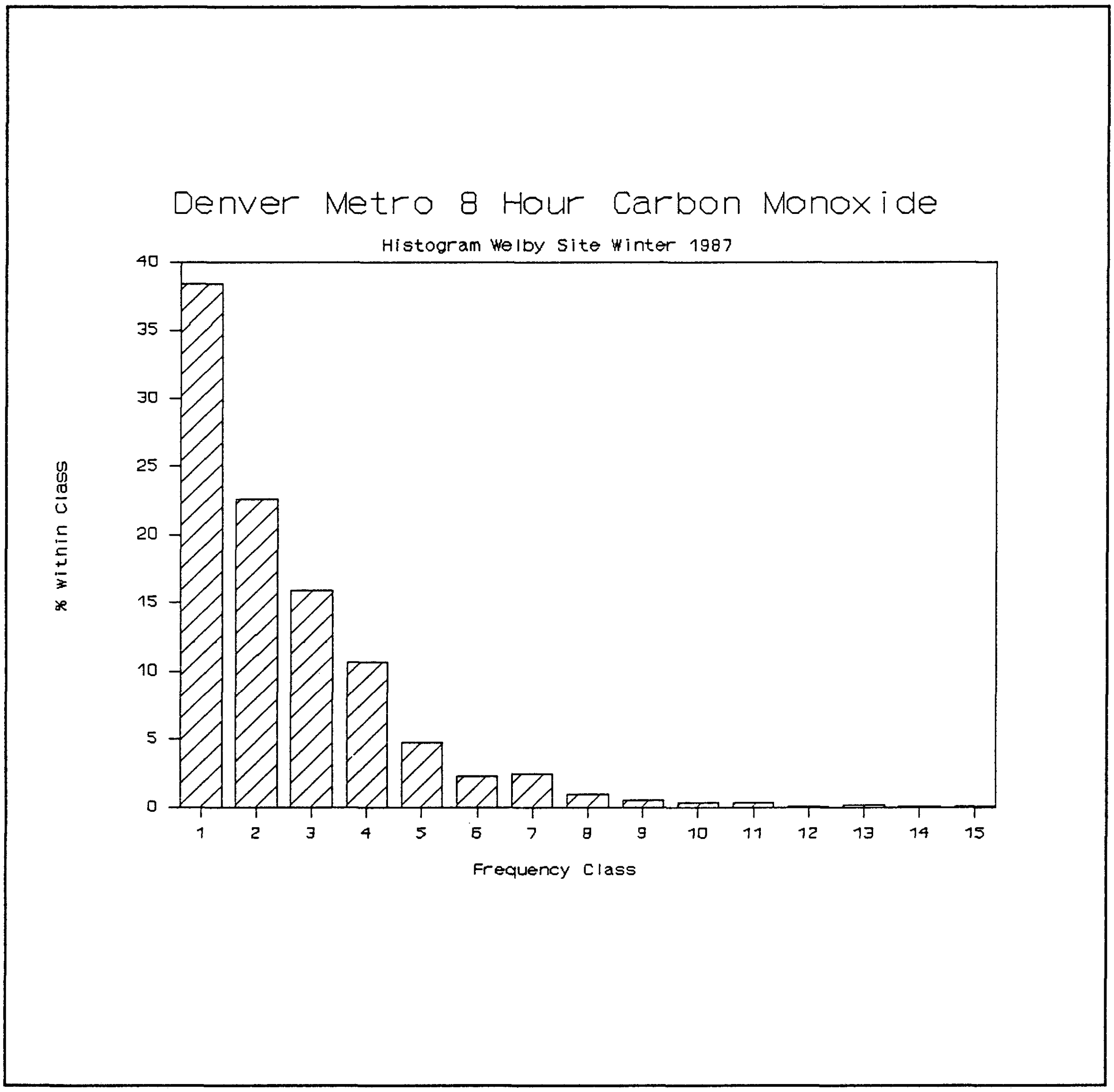


Figure 2.19 Histogram of Welby Transformed Eight-Hour Carbon Monoxide

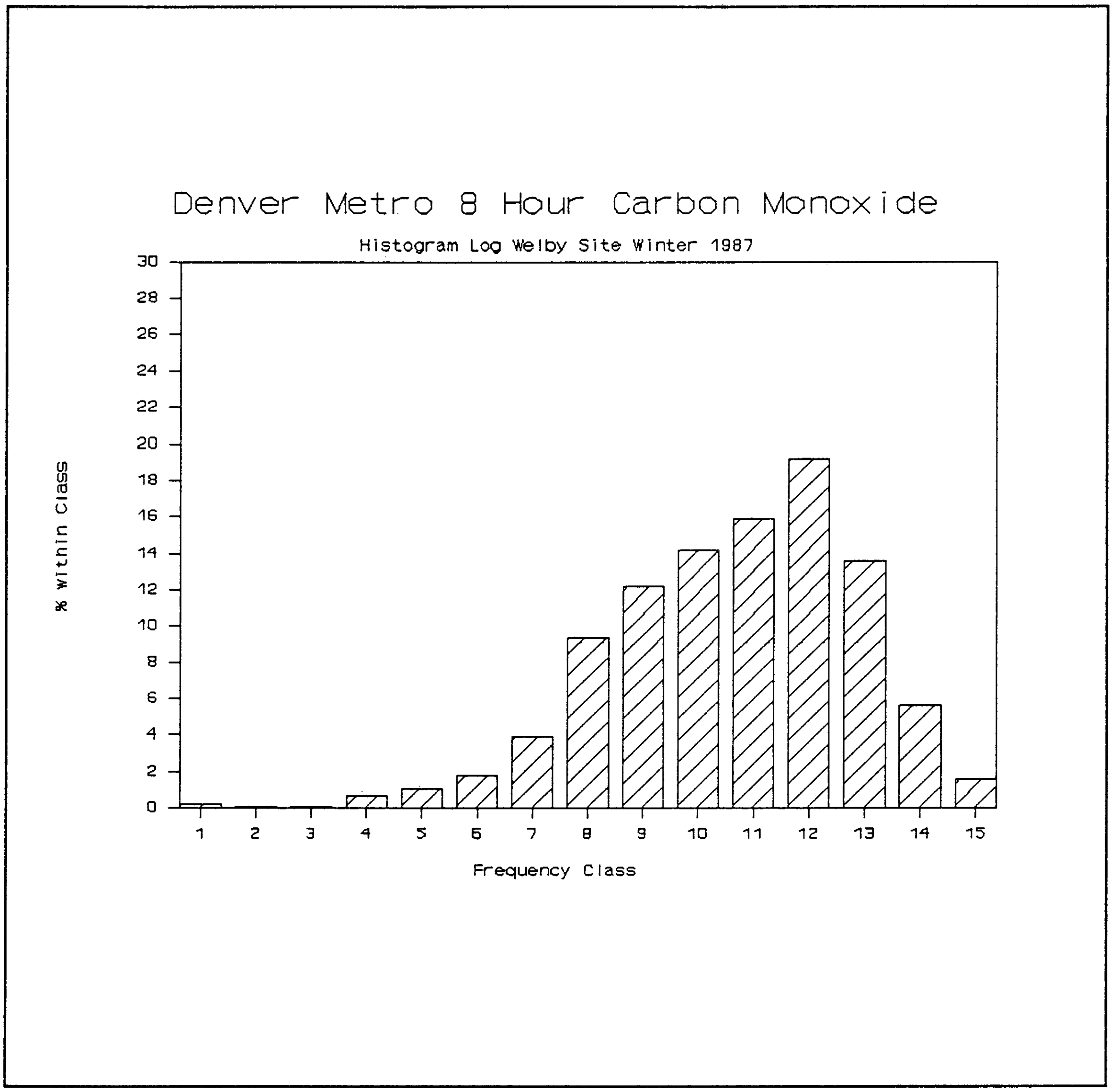


Figure 2.20 Daily vs Monthly Prediction for Welby November 1987

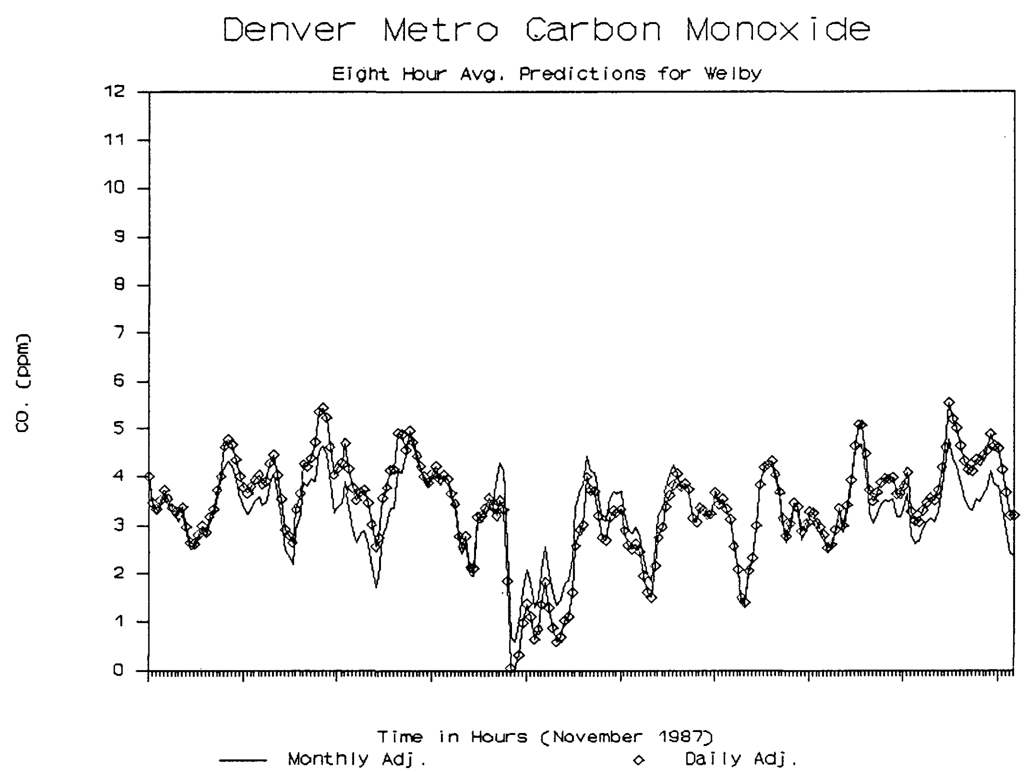


Figure 2.21 Autocorrelation of Residuals for Welby November 1986

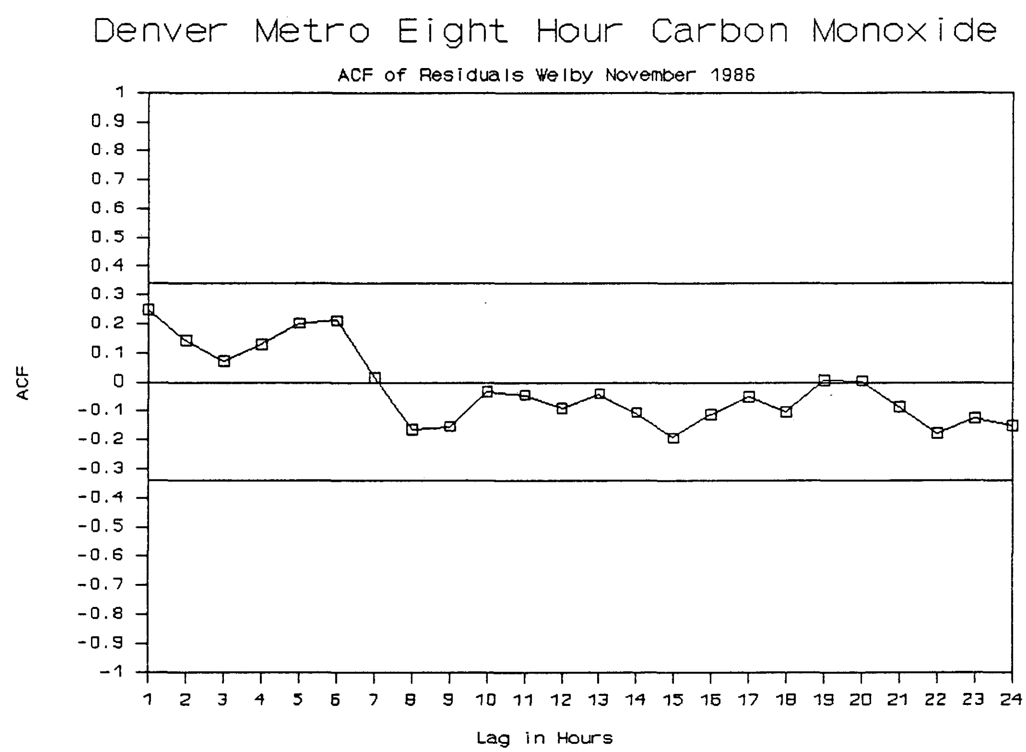


Figure 2.22 Autocorrelation of Residuals for Welby November 1987

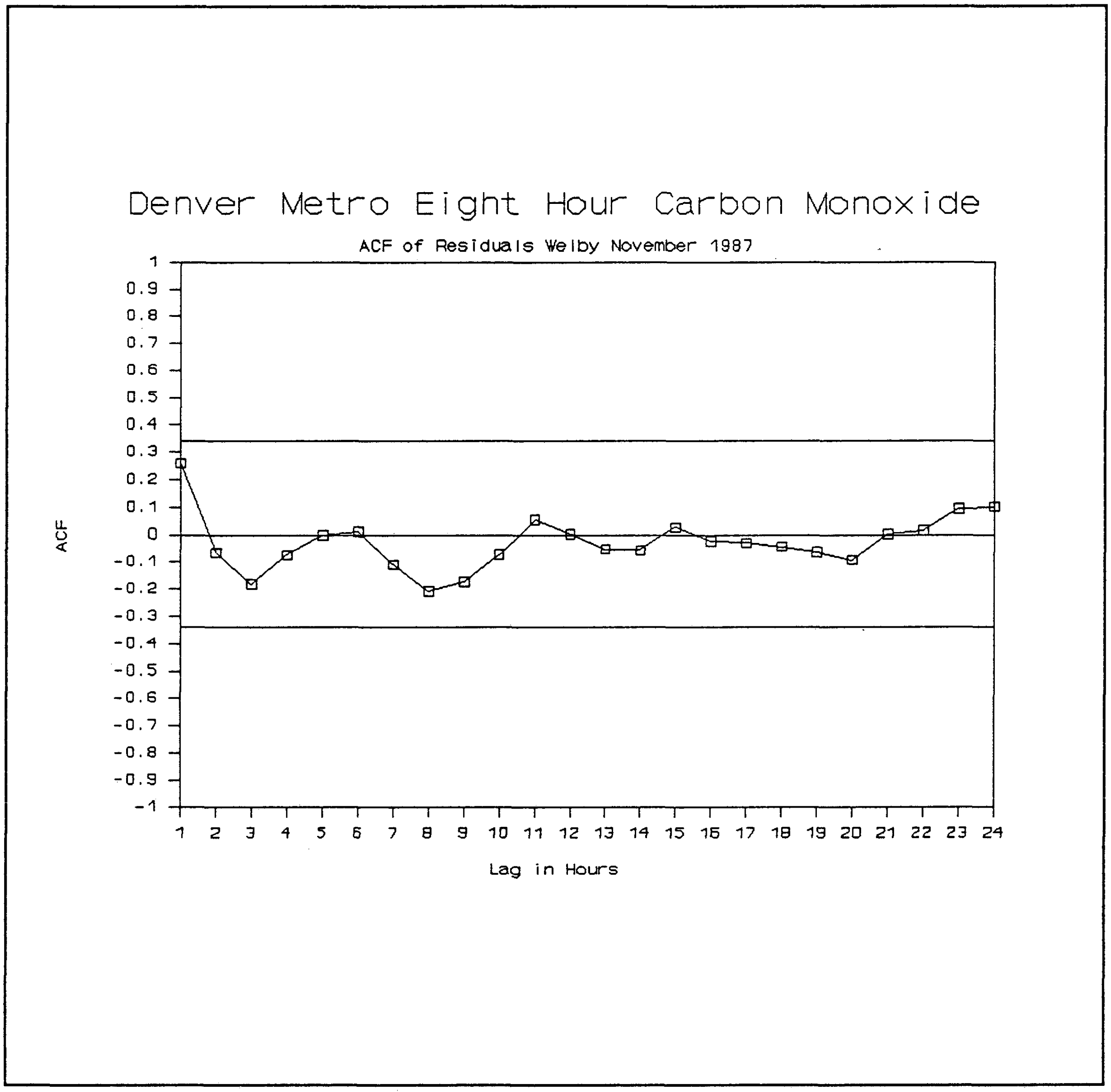


Figure 2.23 Histogram of Residuals for Carriage November 1986

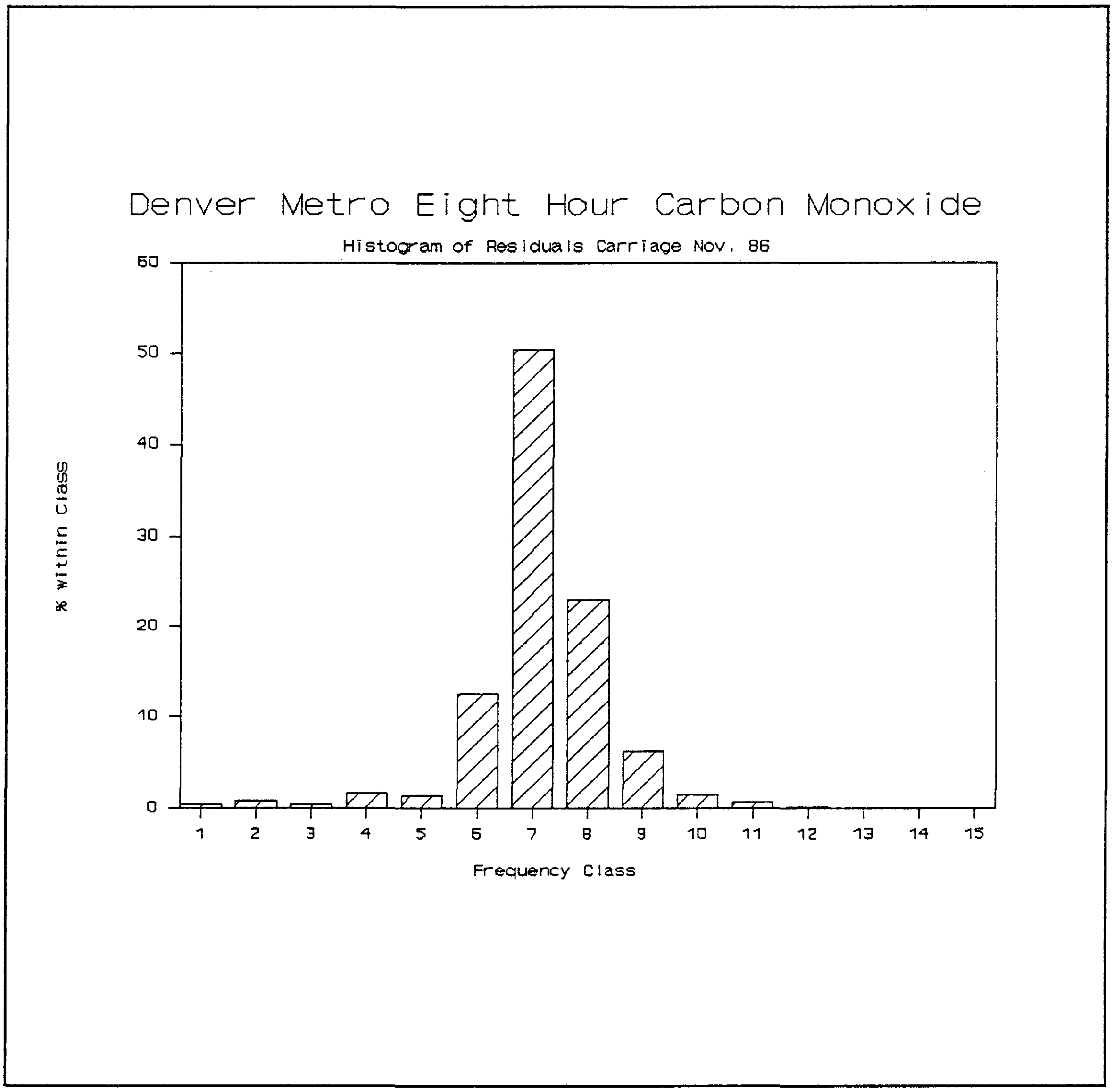


Figure 2.24 Histogram of Residuals for Carriage November 1987

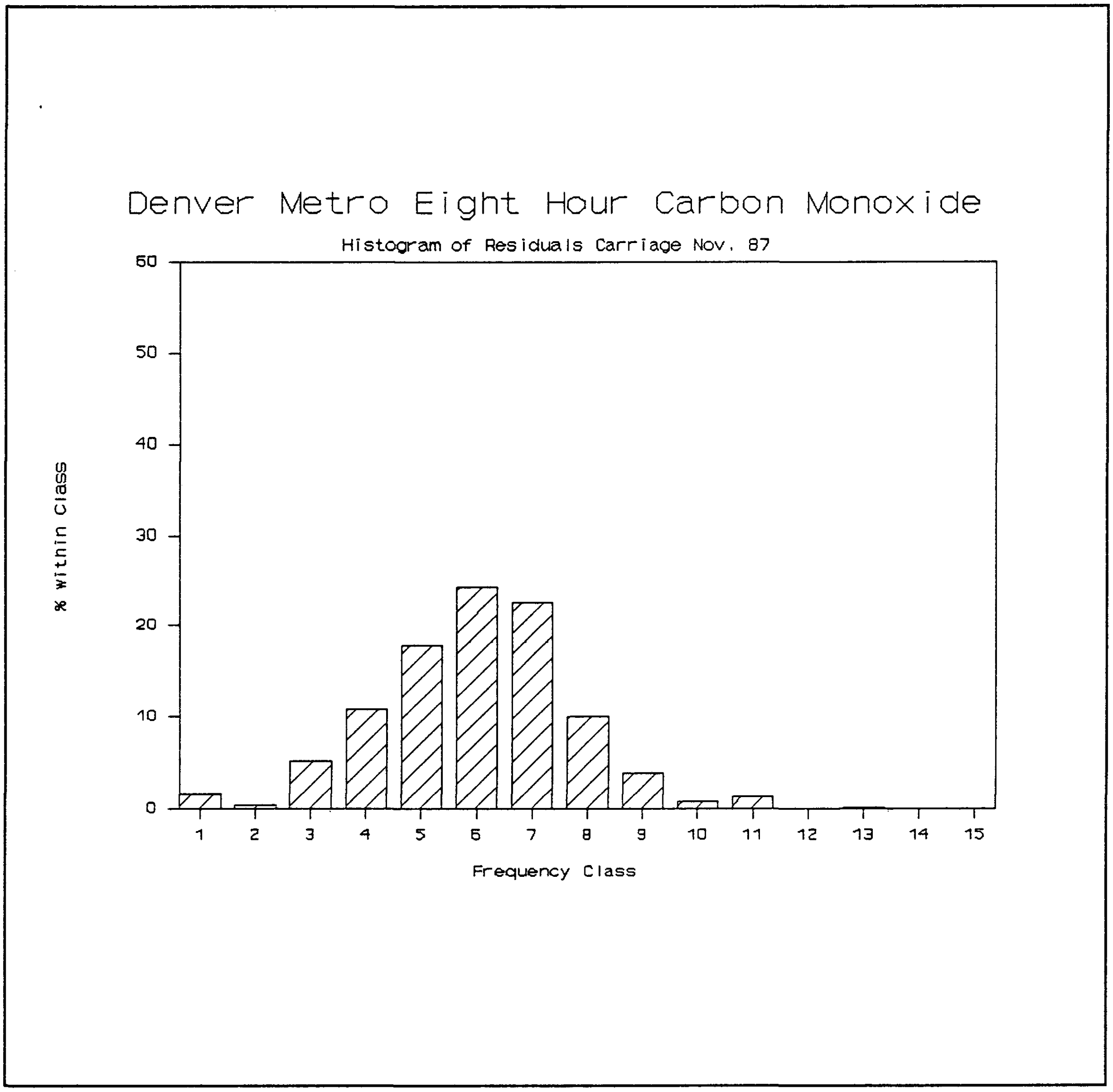
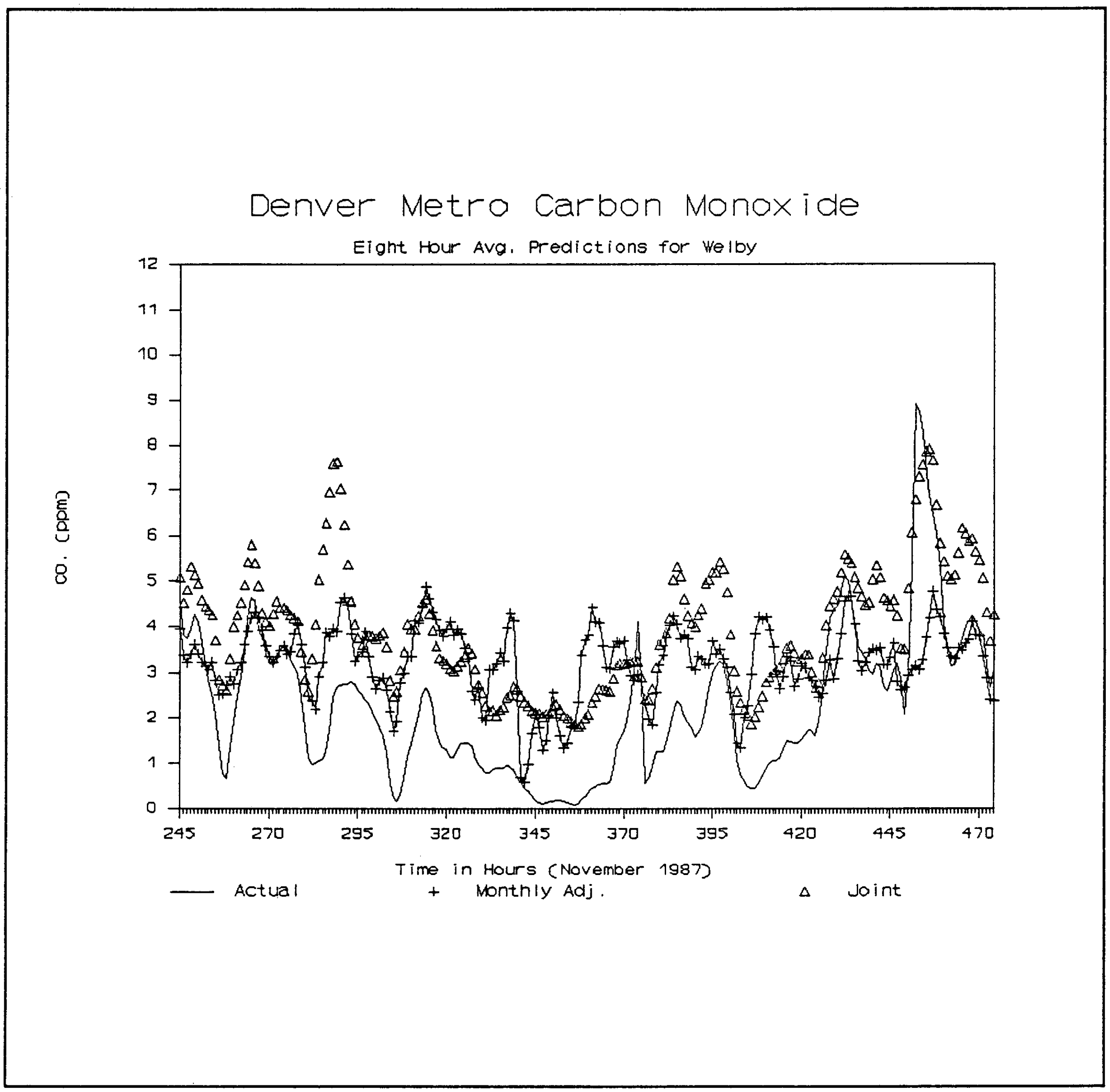


Figure 2.25 Joint and Single-Station Predictions for Welby November 1987



CONCLUSIONS AND FURTHER WORK

This thesis develops a stochastic model of eight-hour running average carbon monoxide concentrations (Eq. 2.1). The model is applied to two Denver metro monitors using two years of winter data. Procedures for producing single-station and joint-station simulations from the model are presented .

Parameter results showed that the single-station model was appropriate for both sites over both years. The robustness of the model can be attributed to similar behavior the eight-hour series exhibited over the study period (Figure 2.11 and 2.12). In the case of a change in pollution pattern, the model would have to be adjusted from year to year.

The single-station predictions were respectable when eight-hour levels were between three and seven ppm. However when the levels were above seven or below two, the model had problems. Simulated data showed the joint-station prediction preformed significantly better than the single-station counterparts at detecting violation of the NAQS (Table 2.5).If the study were expanded to contain all the Denver CO monitors, the joint-station predictions would change depending on which sites were paired together. An analysis of the change in prediction errors e_t could be used as a tool to evaluate the monitoring network.

The single-station model has the potential to be used as a real time predictor of violations of the NAQS if the predictions at higher levels could be improved. One improvement would be to add the traffic and temperature parameters described in section 1.0. Another improvement might be to classify Denver's meteorologic behavior and mean adjust by classification rather than monthly or daily means. A possible meteorologic classification scheme based on inversion height has been proposed by Bill Neff of National Oceananic and Atmospheric Administration (NOAA).

APPENDIX-B

This appendix contains a series of SAS program that were used to produce estimates of seasonal and meteorological componets for both the single station and paired models.

```
/*********************************************************
This is a SAS program that reads in the winter 1987 CO. and
wind speed data from a ascii file. It then mean adjust both
the CO. and Wind data by monthly means. Following this is
logrithm transform of the data and final estimates both the
seasonal and meterorlogic componets for the single station
model.
/*********************************************************

FILENAME COSTUDY 'THES87.DAT';
FILENAME P85OUT  '[ZMB.THESIS]D85RES.DAT';
FILENAME P89OUT  '[ZMB.THESIS]D89RES.DAT';
FILENAME P811OUT '[ZMB.THESIS]D811RES.DAT';
DATA CSTUDY;
INFILE COSTUDY;
INPUT CUMTIME MONTH DAY TIME C5 C9 C11 W5 W9 W11 C85 C89 C811
NC85 NC89 NC811;
 IF C5=9983 THEN C5=. ;
 IF C9=9983 THEN C9=. ;
 IF C11=9983 THEN C11=. ;
 IF W5=9983 THEN W5=.;
 IF W9=9983 THEN W9=.;
 IF W11=9983 THEN W11=.;
 IF C85=9983 THEN C85=. ;
 IF C89=9983 THEN C89=. ;
 IF C811=9983 THEN C811=. ;
 IF NC85=9983 THEN NC85=. ;
 IF NC89=9983 THEN NC89=. ;
 IF NC811=9983 THEN NC811=. ;
 IF NC89=0 THEN NC89=.;
 IF NC85=0 THEN NC85=.;
 IF NC811=0 THEN NC811=.;
RUN;
DATA DSTUDY;
 SET CSTUDY;
 MEAN1=2.32;
 MEAN2=1.45;
 MEAN11=2.16;
 MEAN12=2.50;
 IF MONTH=1 THEN C85=C85-MEAN1;
 IF MONTH=2 THEN C85=C85-MEAN2;
 IF MONTH=11 THEN C85=C85-MEAN11;
```

```
IF MONTH=12 THEN C85=C85-MEAN12;
MEAN1=3.85;
MEAN2=2.36;
MEAN11=2.91;
MEAN12=3.93;
IF MONTH=1 THEN C89=C89-MEAN1;
IF MONTH=2 THEN C89=C89-MEAN2;
IF MONTH=11 THEN C89=C89-MEAN11;
IF MONTH=12 THEN C89=C89-MEAN12;
MEAN1=2.89;
MEAN2=1.63;
MEAN11=2.72;
MEAN12=3.45;
IF MONTH=1 THEN C811=C811-MEAN1;
IF MONTH=2 THEN C811=C811-MEAN2;
IF MONTH=11 THEN C811=C811-MEAN11;
IF MONTH=12 THEN C811=C811-MEAN12;
MEAN1=4.65;
MEAN2=6.05;
MEAN11=5.37;
MEAN12=5.14;
IF MONTH=1 THEN W5=W5-MEAN1;
IF MONTH=2 THEN W5=W5-MEAN2;
IF MONTH=11 THEN W5=W5-MEAN11;
IF MONTH=12 THEN W5=W5-MEAN12;
MEAN1=3.42;
MEAN2=4.04;
MEAN11=1.94;
MEAN12=5.32;
IF MONTH=1 THEN W9=W9-MEAN1;
IF MONTH=2 THEN W9=W9-MEAN2;
IF MONTH=11 THEN W9=W9-MEAN11;
IF MONTH=12 THEN W9=W9-MEAN12;
MEAN1=2.61;
MEAN2=3.11;
MEAN11=2.94;
MEAN12=2.79;
IF MONTH=1 THEN W11=W11-MEAN1;
IF MONTH=2 THEN W11=W11-MEAN2;
IF MONTH=11 THEN W11=W11-MEAN11;
IF MONTH=12 THEN W11=W11-MEAN12;
N5=1/NC85;
N9=1/NC89;
N11=1/NC811;
ZL85=LOG(C85+50);
ZL89=LOG(C89+50);
ZL811=LOG(C811+50);
WL5=LOG(1/(W5+50));
WL9=LOG(1/(W9+50));
```

```
 WL11=LOG(1/(W11+50));
 CT=CUMTIME+1;
 COS24=COS((CT*2*3.1459)/24);
 SIN24=SIN((CT*2*3.1459)/24);
 COS12=COS((CT*2*3.1459)/12);
 SIN12=SIN((CT*2*3.1459)/12);
 COS6=COS((CT*2*3.1459)/6);
 SIN6=SIN((CT*2*3.1459)/6);
 COS4=COS((CT*2*3.1459)/4);
 SIN4=SIN((CT*2*3.1459)/4);
 COS3=COS((CT*2*3.1459)/3);
 SIN3=SIN((CT*2*3.1459)/3);
RUN;
DATA ESTUDY;
 SET DSTUDY;
 WY5=LAG1(WL5);
 WY9=LAG1(WL9);
 WY11=LAG1(WL11);
 ZY85=LAG8(ZL85);
 ZY89=LAG8(ZL89);
 ZY811=LAG8(ZL811);
 ZD85=DIF1(ZL85);
 ZD89=DIF1(ZL89);
 ZD811=DIF1(ZL811);
RUN;
PROC SORT DATA=ESTUDY OUT=FSTUDY;
 BY MONTH;
RUN;
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N5;
 OUTPUT OUT=P85 P=YHAT R=RES5;
 MODEL ZL85=COS24 SIN24 COS12 SIN12 COS6 WL5 WY5;
RUN;
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N9;
 OUTPUT OUT=P89 P=YHAT R=RES9;
 MODEL ZL89=COS24 SIN24 COS12 SIN12 COS6 WL9 WY9 ;
RUN;
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N11;
 OUTPUT OUT=P811 P=YHAT R=RES11;
 MODEL ZL811=COS24 SIN24 COS12 SIN12 COS6 WL11 WY11;
RUN;
DATA _NULL_;
 SET P85;
 FILE P85OUT;
```

```
 IF YHAT  = . THEN YHAT=9983;
 IF RES5 = . THEN RES5=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES5;
RUN;
DATA _NULL_;
 SET P89;
 FILE P89OUT;
 IF YHAT = . THEN YHAT=9983;
 IF RES9 = . THEN RES9=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES9;
RUN;
DATA _NULL_;
 SET P811;
 FILE P811OUT;
 IF YHAT = . THEN YHAT=9983;
 IF RES11 = . THEN RES11=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES11;
RUN;

FILENAME RESOUT  '[ZMB.THESIS]FILTERA7.DAT';
FILENAME PREOUT  '[ZMB.THESIS]CWINDA7.DAT';
DATA C5STUD;
INFILE P85OUT;
INPUT CUMTIME MONTH DAY TIME YHAT5 RESD5 ;
 RUN;
DATA C9STUD;
INFILE P89OUT;
INPUT CUMTIME MONTH DAY TIME YHAT9 RESD9 ;
 RUN;
DATA C11STUD;
INFILE P811OUT;
INPUT CUMTIME MONTH DAY TIME YHAT11 RESD11 ;
 RUN;
PROC SORT DATA=C5STUD OUT=S5STUD;
 BY CUMTIME;
RUN;
PROC SORT DATA=C9STUD OUT=S9STUD;
 BY CUMTIME;
RUN;
PROC SORT DATA=C11STUD OUT=S11STUD;
 BY CUMTIME;
RUN;
DATA RES1;
 MERGE S5STUD S9STUD;
RUN;
DATA RES2;
 MERGE RES1 S11STUD;
RUN;
```

```
DATA _NULL_;
 SET RES2;
 FILE RESOUT;
 PUT CUMTIME MONTH DAY TIME RESD5 RESD9 RESD11 YHAT5 YHAT9
YHAT11;
RUN;
DATA _NULL_;
 SET RES2;
 MEAN1=2.32;
 MEAN2=1.45;
 MEAN11=2.16;
 MEAN12=2.50;
 IF MONTH=1 THEN RESZ5=EXP(YHAT5) - 50 + MEAN1;
 IF MONTH=2 THEN RESZ5=EXP(YHAT5) - 50 + MEAN2;
 IF MONTH=11 THEN RESZ5=EXP(YHAT5) - 50 + MEAN11;
 IF MONTH=12 THEN RESZ5=EXP(YHAT5) - 50 + MEAN12;
 MEAN1=3.85;
 MEAN2=2.36;
 MEAN11=2.91;
 MEAN12=3.93;
 IF MONTH=1 THEN RESZ9=EXP(YHAT9) - 50 + MEAN1;
 IF MONTH=2 THEN RESZ9=EXP(YHAT9) - 50 + MEAN2;
 IF MONTH=11 THEN RESZ9=EXP(YHAT9) - 50 + MEAN11;
 IF MONTH=12 THEN RESZ9=EXP(YHAT9) - 50 + MEAN12;
 MEAN1=2.89;
 MEAN2=1.63;
 MEAN11=2.72;
 MEAN12=3.45;
 IF MONTH=1 THEN RESZ11=EXP(YHAT11) - 50 + MEAN1;
 IF MONTH=2 THEN RESZ11=EXP(YHAT11) - 50 + MEAN2;
 IF MONTH=11 THEN RESZ11=EXP(YHAT11) - 50 + MEAN11;
 IF MONTH=12 THEN RESZ11=EXP(YHAT11) - 50 + MEAN12;
 IF RESZ5 = . THEN RESZ5=9983;
 IF RESZ9 = . THEN RESZ9=9983;
 IF RESZ11 = . THEN RESZ11=9983;
 FILE PREOUT;
 PUT CUMTIME MONTH DAY TIME RESZ5 RESZ9 RESZ11;
RUN;

/************************************************************
This is a SAS program that reads in the winter 1986 CO. and
wind speed data from a ascii file. It then mean adjust the
CO. by daily means and wind data by monthly means. Following
this is logrithmic transform of the data and final estimates
of both the seasonal and meterorlogic componets for the single
station model.
/************************************************************
```

```
FILENAME COSTUDY 'THES86.DAT';
FILENAME P85OUT  '[ZMB.THESIS]D85RES.DAT';
FILENAME P89OUT  '[ZMB.THESIS]D89RES.DAT';
FILENAME P811OUT '[ZMB.THESIS]D811RES.DAT';
DATA CSTUDY;
INFILE COSTUDY;
INPUT CUMTIME MONTH DAY TIME C5 C9 C11 W5 W9 W11 C85 C89 C811
NC85 NC89 NC811;
 IF C5=9983 THEN C5=. ;
 IF C9=9983 THEN C9=. ;
 IF C11=9983 THEN C11=. ;
 IF W5=9983 THEN W5=.;
 IF W9=9983 THEN W9=.;
 IF W11=9983 THEN W11=.;
 IF C85=9983 THEN C85=. ;
 IF C89=9983 THEN C89=. ;
 IF C811=9983 THEN C811=. ;
 IF NC85=9983 THEN NC85=. ;
 IF NC89=9983 THEN NC89=. ;
 IF NC811=9983 THEN NC811=. ;
 IF NC89=0 THEN NC89=.;
 IF NC85=0 THEN NC85=.;
 IF NC811=0 THEN NC811=.;
RUN;
DATA DSTUDY;
 SET CSTUDY;
 MEAN1=2.05;
 MEAN2=2.62;
 MEAN3 =2.74;
 MEAN4 =2.30;
 MEAN5 =2.55;
 MEAN6 =3.33;
 MEAN7 =2.63;
 IF DAY=1 THEN C85=C85-MEAN1;
 IF DAY=2 THEN C85=C85-MEAN2;
 IF DAY=3  THEN C85=C85-MEAN3;
 IF DAY=4  THEN C85=C85-MEAN4;
 IF DAY=5  THEN C85=C85-MEAN5;
 IF DAY=6  THEN C85=C85-MEAN6;
 IF DAY=7  THEN C85=C85-MEAN7;
 MEAN1=2.17;
 MEAN2=4.15;
 MEAN3 =4.97;
 MEAN4 =4.06;
 MEAN5 =3.90;
 MEAN6 =5.50;
 MEAN7 =3.21;
 IF DAY=1 THEN C89=C89-MEAN1;
```

```
IF DAY=2 THEN C89=C89-MEAN2;
IF DAY=3  THEN C89=C89-MEAN3;
IF DAY=4  THEN C89=C89-MEAN4;
IF DAY=5  THEN C89=C89-MEAN5;
IF DAY=6  THEN C89=C89-MEAN6;
IF DAY=7  THEN C89=C89-MEAN7;
MEAN1=2.12;
MEAN2=2.93;
MEAN3 =3.35;
MEAN4 =2.58;
MEAN5 =2.87;
MEAN6 =3.88;
MEAN7 =2.95;
IF DAY=1 THEN C811=C811-MEAN1;
IF DAY=2 THEN C811=C811-MEAN2;
IF DAY=3  THEN C811=C811-MEAN3;
IF DAY=4  THEN C811=C811-MEAN4;
IF DAY=5  THEN C811=C811-MEAN5;
IF DAY=6  THEN C811=C811-MEAN6;
IF DAY=7  THEN C811=C811-MEAN7;
MEAN1=4.65;
MEAN2=6.05;
MEAN11=5.37;
MEAN12=5.14;
IF MONTH=1 THEN W5=W5-MEAN1;
IF MONTH=2 THEN W5=W5-MEAN2;
IF MONTH=11 THEN W5=W5-MEAN11;
IF MONTH=12 THEN W5=W5-MEAN12;
MEAN1=3.42;
MEAN2=4.04;
MEAN11=1.94;
MEAN12=5.32;
IF MONTH=1 THEN W9=W9-MEAN1;
IF MONTH=2 THEN W9=W9-MEAN2;
IF MONTH=11 THEN W9=W9-MEAN11;
IF MONTH=12 THEN W9=W9-MEAN12;
MEAN1=2.61;
MEAN2=3.11;
MEAN11=2.94;
MEAN12=2.79;
IF MONTH=1 THEN W11=W11-MEAN1;
IF MONTH=2 THEN W11=W11-MEAN2;
IF MONTH=11 THEN W11=W11-MEAN11;
IF MONTH=12 THEN W11=W11-MEAN12;
N5=1/NC85;
N9=1/NC89;
N11=1/NC811;
ZL85=LOG(C85+50);
```

```
 ZL89=LOG(C89+50);
 ZL811=LOG(C811+50);
 WL5=LOG(1/(W5+50));
 WL9=LOG(1/(W9+50));
 WL11=LOG(1/(W11+50));
 CT=CUMTIME+1;
 COS24=COS((CT*2*3.1459)/24);
 SIN24=SIN((CT*2*3.1459)/24);
 COS12=COS((CT*2*3.1459)/12);
 SIN12=SIN((CT*2*3.1459)/12);
 COS6=COS((CT*2*3.1459)/6);
 SIN6=SIN((CT*2*3.1459)/6);
 COS4=COS((CT*2*3.1459)/4);
 SIN4=SIN((CT*2*3.1459)/4);
 COS3=COS((CT*2*3.1459)/3);
 SIN3=SIN((CT*2*3.1459)/3);
RUN;
DATA ESTUDY;
 SET DSTUDY;
 WY5=LAG1(WL5);
 WY9=LAG1(WL9);
 WY11=LAG1(WL11);
 ZY85=LAG8(ZL85);
 ZY89=LAG8(ZL89);
 ZY811=LAG8(ZL811);
 ZD85=DIF1(ZL85);
 ZD89=DIF1(ZL89);
 ZD811=DIF1(ZL811);
RUN;
PROC SORT DATA=ESTUDY OUT=FSTUDY;
 BY MONTH;
RUN;
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N5;
 OUTPUT OUT=P85 P=YHAT R=RES5;
 MODEL ZL85=COS24 SIN24 COS12 SIN12 COS6 WL5 WY5;
RUN;
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N9;
 OUTPUT OUT=P89 P=YHAT R=RES9;
 MODEL ZL89=COS24 SIN24 COS12 SIN12 COS6 WL9 WY9 ;
RUN;
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N11;
 OUTPUT OUT=P811 P=YHAT R=RES11;
```

```
 MODEL ZL811=COS24 SIN24 COS12 SIN12 COS6 WL11 WY11;
RUN;
DATA _NULL_;
 SET P85;
 FILE P85OUT;
 IF YHAT  = . THEN YHAT=9983;
 IF RES5 = . THEN RES5=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES5;
RUN;
DATA _NULL_;
 SET P89;
 FILE P89OUT;
 IF YHAT = . THEN YHAT=9983;
 IF RES9 = . THEN RES9=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES9;
RUN;
DATA _NULL_;
 SET P811;
 FILE P811OUT;
 IF YHAT = . THEN YHAT=9983;
 IF RES11 = . THEN RES11=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES11;
RUN;

FILENAME RESOUT  '[ZMB.THESIS]FILTERB6.DAT';
FILENAME PREOUT  '[ZMB.THESIS]CWINDB6.DAT';
DATA C5STUD;
INFILE P85OUT;
INPUT CUMTIME MONTH DAY TIME YHAT5 RESD5 ;
 RUN;
DATA C9STUD;
INFILE P89OUT;
INPUT CUMTIME MONTH DAY TIME YHAT9 RESD9 ;
 RUN;
DATA C11STUD;
INFILE P811OUT;
INPUT CUMTIME MONTH DAY TIME YHAT11 RESD11 ;
 RUN;
PROC SORT DATA=C5STUD OUT=S5STUD;
 BY CUMTIME;
RUN;
PROC SORT DATA=C9STUD OUT=S9STUD;
 BY CUMTIME;
RUN;
PROC SORT DATA=C11STUD OUT=S11STUD;
 BY CUMTIME;
RUN;
DATA RES1;
 MERGE S5STUD S9STUD;
```

```
RUN;
DATA RES2;
 MERGE RES1 S11STUD;
RUN;
DATA _NULL_;
 SET RES2;
 FILE RESOUT;
 PUT CUMTIME MONTH DAY TIME RESD5 RESD9 RESD11 YHAT5 YHAT9
YHAT11;
RUN;
DATA _NULL_;
 SET RES2;
 MEAN1=2.05;
 MEAN2=2.62;
 MEAN3 =2.74;
 MEAN4 =2.30;
 MEAN5 =2.55;
 MEAN6 =3.33;
 MEAN7 =2.63;
 IF DAY=1 THEN RESZ5=EXP(YHAT5) - 50 + MEAN1;
 IF DAY=2 THEN RESZ5=EXP(YHAT5) - 50 + MEAN2;
 IF DAY=3  THEN RESZ5=EXP(YHAT5) - 50 + MEAN3;
 IF DAY=4 THEN RESZ5=EXP(YHAT5) - 50 + MEAN4;
 IF DAY=5 THEN RESZ5=EXP(YHAT5) - 50 + MEAN5;
 IF DAY=6 THEN RESZ5=EXP(YHAT5) - 50 + MEAN6;
 IF DAY=7 THEN RESZ5=EXP(YHAT5) - 50 + MEAN7;
 MEAN1=2.17;
 MEAN2=4.15;
 MEAN3 =4.97;
 MEAN4 =4.06;
 MEAN5 =3.90;
 MEAN6 =5.50;
 MEAN7 =3.21;
 IF DAY=1 THEN RESZ9=EXP(YHAT9) - 50 + MEAN1;
 IF DAY=2 THEN RESZ9=EXP(YHAT9) - 50 + MEAN2;
 IF DAY=3  THEN RESZ9=EXP(YHAT9) - 50 + MEAN3;
 IF DAY=4 THEN RESZ9=EXP(YHAT9) - 50 + MEAN4;
 IF DAY=5 THEN RESZ9=EXP(YHAT9) - 50 + MEAN5;
 IF DAY=6 THEN RESZ9=EXP(YHAT9) - 50 + MEAN6;
 IF DAY=7 THEN RESZ9=EXP(YHAT9) - 50 + MEAN7;
 MEAN1=2.12;
 MEAN2=2.93;
 MEAN3 =3.35;
 MEAN4 =2.58;
 MEAN5 =2.87;
 MEAN6 =3.88;
 MEAN7 =2.95;
 IF DAY=1 THEN RESZ11=EXP(YHAT11) - 50 + MEAN1;
```

```
 IF DAY=2 THEN RESZ11=EXP(YHAT11) - 50 + MEAN2;
 IF DAY=3  THEN RESZ11=EXP(YHAT11) - 50 + MEAN3;
 IF DAY=4 THEN RESZ11=EXP(YHAT11) - 50 + MEAN4;
 IF DAY=5 THEN RESZ11=EXP(YHAT11) - 50 + MEAN5;
 IF DAY=6 THEN RESZ11=EXP(YHAT11) - 50 + MEAN6;
 IF DAY=7 THEN RESZ11=EXP(YHAT11) - 50 + MEAN7;
 IF RESZ5 = . THEN RESZ5=9983;
 IF RESZ9 = . THEN RESZ9=9983;
 IF RESZ11 = . THEN RESZ11=9983;
 FILE PREOUT;
 PUT CUMTIME MONTH DAY TIME RESZ5 RESZ9 RESZ11;
RUN;

/************************************************************
*
This is a SAS program that reads in the winter 1987 CO. and
wind speed data from a ascii file. It then mean adjust both
the CO. and Wind data by monthly means. Following this is
logrithm transform of the data and final estimates both the
seasonal and meterorlogic componets for the paried model.
/************************************************************

FILENAME COSTUDY 'THES87.DAT';
FILENAME P85OUT  '[ZMB.THESIS]D85RES.DAT';
FILENAME P89OUT  '[ZMB.THESIS]D89RES.DAT';
FILENAME P811OUT '[ZMB.THESIS]D811RES.DAT';
DATA CSTUDY;
INFILE COSTUDY;
INPUT CUMTIME MONTH DAY TIME C5 C9 C11 W5 W9 W11 C85 C89 C811
NC85 NC89 NC811;
 IF C5=9983 THEN C5=. ;
 IF C9=9983 THEN C9=. ;
 IF C11=9983 THEN C11=. ;
 IF W5=9983 THEN W5=.;
 IF W9=9983 THEN W9=.;
 IF W11=9983 THEN W11=.;
 IF C85=9983 THEN C85=. ;
 IF C89=9983 THEN C89=. ;
 IF C811=9983 THEN C811=. ;
 IF NC85=9983 THEN NC85=. ;
 IF NC89=9983 THEN NC89=. ;
 IF NC811=9983 THEN NC811=. ;
 IF NC89=0 THEN NC89=.;
 IF NC85=0 THEN NC85=.;
 IF NC811=0 THEN NC811=.;
RUN;
DATA DSTUDY;
 SET CSTUDY;
```

```
MEAN1=2.32;
MEAN2=1.45;
MEAN11=2.16;
MEAN12=2.50;
IF MONTH=1 THEN C85=C85-MEAN1;
IF MONTH=2 THEN C85=C85-MEAN2;
IF MONTH=11 THEN C85=C85-MEAN11;
IF MONTH=12 THEN C85=C85-MEAN12;
MEAN1=3.85;
MEAN2=2.36;
MEAN11=2.91;
MEAN12=3.93;
IF MONTH=1 THEN C89=C89-MEAN1;
IF MONTH=2 THEN C89=C89-MEAN2;
IF MONTH=11 THEN C89=C89-MEAN11;
IF MONTH=12 THEN C89=C89-MEAN12;
MEAN1=2.89;
MEAN2=1.63;
MEAN11=2.72;
MEAN12=3.45;
IF MONTH=1 THEN C811=C811-MEAN1;
IF MONTH=2 THEN C811=C811-MEAN2;
IF MONTH=11 THEN C811=C811-MEAN11;
IF MONTH=12 THEN C811=C811-MEAN12;
MEAN1=4.65;
MEAN2=6.05;
MEAN11=5.37;
MEAN12=5.14;
IF MONTH=1 THEN W5=W5-MEAN1;
IF MONTH=2 THEN W5=W5-MEAN2;
IF MONTH=11 THEN W5=W5-MEAN11;
IF MONTH=12 THEN W5=W5-MEAN12;
MEAN1=3.42;
MEAN2=4.04;
MEAN11=1.94;
MEAN12=5.32;
IF MONTH=1 THEN W9=W9-MEAN1;
IF MONTH=2 THEN W9=W9-MEAN2;
IF MONTH=11 THEN W9=W9-MEAN11;
IF MONTH=12 THEN W9=W9-MEAN12;
MEAN1=2.61;
MEAN2=3.11;
MEAN11=2.94;
MEAN12=2.79;
IF MONTH=1 THEN W11=W11-MEAN1;
IF MONTH=2 THEN W11=W11-MEAN2;
IF MONTH=11 THEN W11=W11-MEAN11;
IF MONTH=12 THEN W11=W11-MEAN12;
```

```
 N5=1/NC85;
 N9=1/NC89;
 N11=1/NC811;
 ZL85=LOG(C85+50);
 ZL89=LOG(C89+50);
 ZL811=LOG(C811+50);
 WL5=LOG(1/(W5+50));
 WL9=LOG(1/(W9+50));
 WL11=LOG(1/(W11+50));
 CT=CUMTIME+1;
 COS24=COS((CT*2*3.1459)/24);
 SIN24=SIN((CT*2*3.1459)/24);
 COS12=COS((CT*2*3.1459)/12);
 SIN12=SIN((CT*2*3.1459)/12);
 COS6=COS((CT*2*3.1459)/6);
 SIN6=SIN((CT*2*3.1459)/6);
 COS4=COS((CT*2*3.1459)/4);
 SIN4=SIN((CT*2*3.1459)/4);
 COS3=COS((CT*2*3.1459)/3);
 SIN3=SIN((CT*2*3.1459)/3);
RUN;
DATA ESTUDY;
 SET DSTUDY;
 WY5=LAG1(WL5);
 WY9=LAG1(WL9);
 WY11=LAG1(WL11);
 ZY85=LAG8(ZL85);
 ZY89=LAG8(ZL89);
 ZY811=LAG8(ZL811);
 ZD85=DIF1(ZL85);
 ZD89=DIF1(ZL89);
 ZD811=DIF1(ZL811);
RUN;
PROC SORT DATA=ESTUDY OUT=FSTUDY;
 BY MONTH;
RUN;
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N5;
 OUTPUT OUT=P85 P=YHAT R=RES5;
 MODEL ZL85=COS24 SIN24 COS12 SIN12 COS6 WL5 WY5 ZL89 WL9;
RUN;
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N5;
 OUTPUT OUT=P89 P=YHAT R=RES9;
 MODEL ZL85=COS24 SIN24 COS12 SIN12 COS6 WL5 WY5 ZL89 ZY89
WL9 WY9;
RUN;
```

```
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N5;
 OUTPUT OUT=P811 P=YHAT R=RES11;
 MODEL ZL85=COS24 SIN24 COS12 SIN12 COS6 WL5 WY5 ZL811 WL11;
RUN;
DATA _NULL_;
 SET P85;
 FILE P85OUT;
 IF YHAT  = . THEN YHAT=9983;
 IF RES5 = . THEN RES5=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES5;
RUN;
DATA _NULL_;
 SET P89;
 FILE P89OUT;
 IF YHAT = . THEN YHAT=9983;
 IF RES9 = . THEN RES9=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES9;
RUN;
DATA _NULL_;
 SET P811;
 FILE P811OUT;
 IF YHAT = . THEN YHAT=9983;
 IF RES11 = . THEN RES11=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES11;
RUN;

FILENAME RESOUT  '[ZMB.THESIS]FILTERC7.DAT';
FILENAME PREOUT  '[ZMB.THESIS]CWINDC7.DAT';
DATA C5STUD;
INFILE P85OUT;
INPUT CUMTIME MONTH DAY TIME YHAT5 RESD5 ;
 RUN;
DATA C9STUD;
INFILE P89OUT;
INPUT CUMTIME MONTH DAY TIME YHAT9 RESD9 ;
 RUN;
DATA C11STUD;
INFILE P811OUT;
INPUT CUMTIME MONTH DAY TIME YHAT11 RESD11 ;
 RUN;
PROC SORT DATA=C5STUD OUT=S5STUD;
 BY CUMTIME;
RUN;
PROC SORT DATA=C9STUD OUT=S9STUD;
 BY CUMTIME;
```

```
RUN;
PROC SORT DATA=C11STUD OUT=S11STUD;
 BY CUMTIME;
RUN;
DATA RES1;
 MERGE S5STUD S9STUD;
RUN;
DATA RES2;
 MERGE RES1 S11STUD;
RUN;
DATA _NULL_;
 SET RES2;
 FILE RESOUT;
 PUT CUMTIME MONTH DAY TIME RESD5 RESD9 RESD11 YHAT5 YHAT9
YHAT11;
RUN;
DATA _NULL_;
 SET RES2;
 MEAN1=2.32;
 MEAN2=1.45;
 MEAN11=2.16;
 MEAN12=2.50;
 IF MONTH=1 THEN RESZ5=EXP(YHAT5) - 50 + MEAN1;
 IF MONTH=2 THEN RESZ5=EXP(YHAT5) - 50 + MEAN2;
 IF MONTH=11 THEN RESZ5=EXP(YHAT5) - 50 + MEAN11;
 IF MONTH=12 THEN RESZ5=EXP(YHAT5) - 50 + MEAN12;
 MEAN1=3.85;
 MEAN2=2.36;
 MEAN11=2.91;
 MEAN12=3.93;
 IF MONTH=1 THEN RESZ9=EXP(YHAT9) - 50 + MEAN1;
 IF MONTH=2 THEN RESZ9=EXP(YHAT9) - 50 + MEAN2;
 IF MONTH=11 THEN RESZ9=EXP(YHAT9) - 50 + MEAN11;
 IF MONTH=12 THEN RESZ9=EXP(YHAT9) - 50 + MEAN12;
 MEAN1=2.89;
 MEAN2=1.63;
 MEAN11=2.72;
 MEAN12=3.45;
 IF MONTH=1 THEN RESZ11=EXP(YHAT11) - 50 + MEAN1;
 IF MONTH=2 THEN RESZ11=EXP(YHAT11) - 50 + MEAN2;
 IF MONTH=11 THEN RESZ11=EXP(YHAT11) - 50 + MEAN11;
 IF MONTH=12 THEN RESZ11=EXP(YHAT11) - 50 + MEAN12;
 IF RESZ5 = . THEN RESZ5=9983;
 IF RESZ9 = . THEN RESZ9=9983;
 IF RESZ11 = . THEN RESZ11=9983;
 FILE PREOUT;
 PUT CUMTIME MONTH DAY TIME RESZ5 RESZ9 RESZ11;
RUN;
```

```
/**********************************************************
This is a SAS program that reads in the winter 1986 CO. and
wind speed data from a ascii file. It then mean adjust both
the CO. and Wind data by monthly means. Following this is
logrithm transform of the data and final estimates both the
seasonal and meterorlogic componets for the paried model.
/**********************************************************

FILENAME COSTUDY 'THES86.DAT';
FILENAME P85OUT  '[ZMB.THESIS]D85RES.DAT';
FILENAME P89OUT  '[ZMB.THESIS]D89RES.DAT';
FILENAME P811OUT '[ZMB.THESIS]D811RES.DAT';
DATA CSTUDY;
INFILE COSTUDY;
INPUT CUMTIME MONTH DAY TIME C5 C9 C11 W5 W9 W11 C85 C89 C811
NC85 NC89 NC811;
 IF C5=9983 THEN C5=. ;
 IF C9=9983 THEN C9=. ;
 IF C11=9983 THEN C11=. ;
 IF W5=9983 THEN W5=.;
 IF W9=9983 THEN W9=.;
 IF W11=9983 THEN W11=.;
 IF C85=9983 THEN C85=. ;
 IF C89=9983 THEN C89=. ;
 IF C811=9983 THEN C811=. ;
 IF NC85=9983 THEN NC85=. ;
 IF NC89=9983 THEN NC89=. ;
 IF NC811=9983 THEN NC811=. ;
 IF NC89=0 THEN NC89=.;
 IF NC85=0 THEN NC85=.;
 IF NC811=0 THEN NC811=.;
RUN;
DATA DSTUDY;
 SET CSTUDY;
 MEAN1=2.42;
 MEAN2=1.79;
 MEAN11=2.43;
 MEAN12=3.80;
 IF MONTH=1 THEN C85=C85-MEAN1;
 IF MONTH=2 THEN C85=C85-MEAN2;
 IF MONTH=11 THEN C85=C85-MEAN11;
 IF MONTH=12 THEN C85=C85-MEAN12;
 MEAN1=4.15;
 MEAN2=3.25;
 MEAN11=4.04;
 MEAN12=4.66;
 IF MONTH=1 THEN C89=C89-MEAN1;
```

```
IF MONTH=2 THEN C89=C89-MEAN2;
IF MONTH=11 THEN C89=C89-MEAN11;
IF MONTH=12 THEN C89=C89-MEAN12;
MEAN1=2.74;
MEAN2=1.94;
MEAN11=2.86;
MEAN12=4.03;
IF MONTH=1 THEN C811=C811-MEAN1;
IF MONTH=2 THEN C811=C811-MEAN2;
IF MONTH=11 THEN C811=C811-MEAN11;
IF MONTH=12 THEN C811=C811-MEAN12;
MEAN1=5.86;
MEAN2=4.78;
MEAN11=4.46;
MEAN12=4.11;
IF MONTH=1 THEN W5=W5-MEAN1;
IF MONTH=2 THEN W5=W5-MEAN2;
IF MONTH=11 THEN W5=W5-MEAN11;
IF MONTH=12 THEN W5=W5-MEAN12;
MEAN1=3.98;
MEAN2=3.44;
MEAN11=3.77;
MEAN12=3.10;
IF MONTH=1 THEN W9=W9-MEAN1;
IF MONTH=2 THEN W9=W9-MEAN2;
IF MONTH=11 THEN W9=W9-MEAN11;
IF MONTH=12 THEN W9=W9-MEAN12;
MEAN1=3.06;
MEAN2=2.55;
MEAN11=2.80;
MEAN12=2.32;
IF MONTH=1 THEN W11=W11-MEAN1;
IF MONTH=2 THEN W11=W11-MEAN2;
IF MONTH=11 THEN W11=W11-MEAN11;
IF MONTH=12 THEN W11=W11-MEAN12;
N5=1/NC85;
N9=1/NC89;
N11=1/NC811;
ZL85=LOG(C85+50);
ZL89=LOG(C89+50);
ZL811=LOG(C811+50);
WL5=LOG(1/(W5+50));
WL9=LOG(1/(W9+50));
WL11=LOG(1/(W11+50));
CT=CUMTIME+1;
COS24=COS((CT*2*3.1459)/24);
SIN24=SIN((CT*2*3.1459)/24);
COS12=COS((CT*2*3.1459)/12);
```

```
 SIN12=SIN((CT*2*3.1459)/12);
 COS6=COS((CT*2*3.1459)/6);
 SIN6=SIN((CT*2*3.1459)/6);
 COS4=COS((CT*2*3.1459)/4);
 SIN4=SIN((CT*2*3.1459)/4);
 COS3=COS((CT*2*3.1459)/3);
 SIN3=SIN((CT*2*3.1459)/3);
RUN;
DATA ESTUDY;
 SET DSTUDY;
 WY5=LAG1(WL5);
 WY9=LAG1(WL9);
 WY11=LAG1(WL11);
 ZY85=LAG8(ZL85);
 ZY89=LAG8(ZL89);
 ZY811=LAG8(ZL811);
 ZD85=DIF1(ZL85);
 ZD89=DIF1(ZL89);
 ZD811=DIF1(ZL811);
RUN;
PROC SORT DATA=ESTUDY OUT=FSTUDY;
 BY MONTH;
RUN;
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N5;
 OUTPUT OUT=P85 P=YHAT R=RES5;
 MODEL ZL85=COS24 SIN24 COS12 SIN12 COS6 WL5 WY5 ZL89 WL9;
RUN;
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N5;
 OUTPUT OUT=P89 P=YHAT R=RES9;
 MODEL ZL85=COS24 SIN24 COS12 SIN12 COS6 WL5 WY5 ZL89 ZY89
WL9 WY9;
RUN;
PROC GLM DATA=FSTUDY;
 BY MONTH;
 WEIGHT N5;
 OUTPUT OUT=P811 P=YHAT R=RES11;
 MODEL ZL85=COS24 SIN24 COS12 SIN12 COS6 WL5 WY5 ZL811 WL11;
RUN;
DATA _NULL_;
 SET P85;
 FILE P85OUT;
 IF YHAT  = . THEN YHAT=9983;
 IF RES5 = . THEN RES5=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES5;
```

```
RUN;
DATA _NULL_;
 SET P89;
 FILE P89OUT;
 IF YHAT = . THEN YHAT=9983;
 IF RES9 = . THEN RES9=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES9;
RUN;
DATA _NULL_;
 SET P811;
 FILE P811OUT;
 IF YHAT = . THEN YHAT=9983;
 IF RES11 = . THEN RES11=9983;
 PUT CUMTIME MONTH DAY TIME YHAT RES11;
RUN;

FILENAME RESOUT  '[ZMB.THESIS]FILTERC6.DAT';
FILENAME PREOUT  '[ZMB.THESIS]CWINDC6.DAT';
DATA C5STUD;
INFILE P85OUT;
INPUT CUMTIME MONTH DAY TIME YHAT5 RESD5 ;
 RUN;
DATA C9STUD;
INFILE P89OUT;
INPUT CUMTIME MONTH DAY TIME YHAT9 RESD9 ;
 RUN;
DATA C11STUD;
INFILE P811OUT;
INPUT CUMTIME MONTH DAY TIME YHAT11 RESD11 ;
 RUN;
PROC SORT DATA=C5STUD OUT=S5STUD;
 BY CUMTIME;
RUN;
PROC SORT DATA=C9STUD OUT=S9STUD;
 BY CUMTIME;
RUN;
PROC SORT DATA=C11STUD OUT=S11STUD;
 BY CUMTIME;
RUN;
DATA RES1;
 MERGE S5STUD S9STUD;
RUN;
DATA RES2;
 MERGE RES1 S11STUD;
RUN;
DATA _NULL_;
 SET RES2;
 FILE RESOUT;
```

```
 PUT CUMTIME MONTH DAY TIME RESD5 RESD9 RESD11 YHAT5 YHAT9
YHAT11;
RUN;
DATA _NULL_;
 SET RES2;
 MEAN1=2.42;
 MEAN2=1.79;
 MEAN11=2.43;
 MEAN12=3.80;
 IF MONTH=1 THEN RESZ5=EXP(YHAT5) - 50 + MEAN1;
 IF MONTH=2 THEN RESZ5=EXP(YHAT5) - 50 + MEAN2;
 IF MONTH=11 THEN RESZ5=EXP(YHAT5) - 50 + MEAN11;
 IF MONTH=12 THEN RESZ5=EXP(YHAT5) - 50 + MEAN12;
 MEAN1=4.15;
 MEAN2=3.25;
 MEAN11=4.04;
 MEAN12=4.66;
 IF MONTH=1 THEN RESZ9=EXP(YHAT9) - 50 + MEAN1;
 IF MONTH=2 THEN RESZ9=EXP(YHAT9) - 50 + MEAN2;
 IF MONTH=11 THEN RESZ9=EXP(YHAT9) - 50 + MEAN11;
 IF MONTH=12 THEN RESZ9=EXP(YHAT9) - 50 + MEAN12;
 MEAN1=2.74;
 MEAN2=1.94;
 MEAN11=2.86;
 MEAN12=4.03;
 IF MONTH=1 THEN RESZ11=EXP(YHAT11) - 50 + MEAN1;
 IF MONTH=2 THEN RESZ11=EXP(YHAT11) - 50 + MEAN2;
 IF MONTH=11 THEN RESZ11=EXP(YHAT11) - 50 + MEAN11;
 IF MONTH=12 THEN RESZ11=EXP(YHAT11) - 50 + MEAN12;
 IF RESZ5 = . THEN RESZ5=9983;
 IF RESZ9 = . THEN RESZ9=9983;
 IF RESZ11 = . THEN RESZ11=9983;
 FILE PREOUT;
 PUT CUMTIME MONTH DAY TIME RESZ5 RESZ9 RESZ11;
RUN;
```

LIST OF REFERENCES

Bailey C. R., "Statistical Comparison of Sites: Discussion", ASA/EPA Conference on Interpretation of Environmental Data, pp. 73-4, May 1987.

Box G.E.P. and Jenkins G.M., Time Series Analysis: Forcasting and Control, San Franscisco: Holden Day. 1976

Brockwell P.S. and Davis R.A., Time Series: Theory and Methods Springer-Verlay, New York, 1987.

Lon-Mu Liu, Hudak G., and Tiao G.C., "A Statistical Assessment of the Effect of the Arizona Car Inspection / Maintenance Program on Ambient CO Air Quality in Phoenix, Arizona", EPA Report #68-023-3513, March 1986.

Macrae G., "Denver's Carbon Monoxide and Meteorology", Interviewed November 1988 Environmental Protection Agency.

Meisel W.S. and Zeldin M.D., "Use of Meteorological Data in Air Quality Trend Analysis", EPA Report #68-02-2318, May 1978.

"Special Study - Winter 1986/1987 Denver Carbon Monoxide", Air Pollution Control Division Colorado Department of Health, May 1987.

Vecchia, A.V., "A Stochastic Streamflow Model and Precipitation Model for the Platte River from Gothengurg to Grand Island, Nebraska" Hydrologic and Geomorphic Studies of the Platte Basin-U.S.G.S. Professional Paper 1277, 1983.