TIME SERIES ANALYSIS OF MORTALITY AND ASSOCIATED WEATHER AND POLLUTION EFFECTS IN LOS ANGELES COUNTY

By

R. H. Shumway[1]
A. S. Azari[1]
Y. Pawitan[2]

[1]Division of Statistics
University of California
Davis CA 95616

[2]Department of Biostatistics
University of Washington
Seattle WA 98195

Final Report
Contract ARB-A5-152-33
September 15, 1988

Submitted to:

Research Division
California Air Resources Board
P.O. Box 2815
Sacramento CA 95812

TABLE OF CONTENTS

ABSTRACT

Parametric and non-parametric time-series analyses of ten years of daily mortality, pollution, and weather data from Los Angeles County were performed. State-space modeling and time and frequency domain regressions were used to modify the database and to isolate significant weather factors and pollutants associated with increased daily mortality.

Mortality was significantly related to pollution and temperature. The relationship with temperature was parabolic, with minimum mortality at about 75 degrees. Mortality had a significant and increasing relationship with concentrations of three highly correlated pollutants -- carbon monoxide, total hydrocarbons, and particulates -- whose individual contributions to the increased pollution associated with increased mortality could not be distinguished. Nitrogen dioxide and sulfur dioxide were much less strongly related to mortality, and models relating mortality to both ozone and temperature could not distinguish their respective contributions.

A parametric nonlinear time series model involving linear and squared terms in temperature and the logarithm of pollution provides a reasonable predictive model. Minimum and maximum predicted mortalities differed by approaximately ten percent over the ranges of temperature and pollution in the data set. The parametric and non-parametric regressions yielded similar relationships.

1. Introduction

The possibility of significant relationships between short-term or long-term levels of air pollution and mortality is of concern to environmental agencies responsible for setting health-based standards for ambient air pollution. Environmental agencies must choose the levels of these standards to safeguard the general population, including sensitive subgroups, against adverse health effects.

One can generally attempt to answer two separate questions relating to the possible effects of air pollution levels on mortality. The first is that of determining the extent and nature of the association between pollutants and mortality levels in the presence of possible environmental contributors such as weather while taking account of the fact that the observations made over time are inherently correlated. A second question is that of defining the nature of a dose-response relation for use in predicting levels of mortality as a function of pollution and weather effects. This second question is of critical interest to the various regulatory agencies responsible for establishing health-based standards for pollutants.

It is generally accepted that very severe pollution episodes, such as the London smog of 1952, did cause significant excess mortality. Such episodes of exceedingly high pollution do not presently occur in the United States, but a number of large metropolitan areas, including Los Angeles, have in recent years persistently experienced levels of ambient air pollution exceeding the current health-based standards. Numerous analyses have attempted to determine whether there is an increasing relationship between mortality and pollution levels in large metropolitan areas in the absence of severe episodes. The possible existence of thresholds below which pollutants have no discernible effect on mortality has been of special interest in these analyses.

A classic database that has been the subject of a number of investigations contains daily observations on mortality, temperature, relative humidity, sulfur dioxide levels and British smoke levels, measured over 14 winters in the London Metropolitan Area. A review of a number of such studies, including a groundbreaking British contribution by Martin and Bradley (1960) is given in Ware et al (1981) who describe their results in terms of standard regressions relating mortality to the primary pollutants for the winter of 1958, ignoring the effect of weather. Later studies, such as Mazumdar et al (1982) included all 14 years and considered both linear and nonlinear temperature and pollution effects. Ostro (1984) considered year-by-year regressions with a change in slope at a fixed breakpoint. The latter two studies also adjusted mortality statistics for epidemics by subtracting a 15-day moving average from each value. Schimmel (1978) reports a similar analysis of mortality in New York City.

The use of various time series techniques to allow for correlated errors in the regressions and possible lagging relations between mortality and the weather and pollution effects has also been considered by a number of investigators. An early study by Wyzga (1978) used distributed lag models and assumed a first-order autoregressive structure for the errors. A more recent study by Schwartz and Marcus (1987) used instantaneous linear regressions and assumed a second-order autoregressive structure for the errors. They investigated possible nonlinear dose-response relations using an heuristic local averaging technique. Shumway et al (1983) used frequency domain methods to investigate the significance of the mortality-pollution relationship in various frequency bands and to estimate the lagged regression coefficients; the mortality and pollution variables were transformed to logarithms.

The above studies involving the London data all find significant associations between mortality and British smoke, and between mortality and sulfur dioxide, whether temperature and relative humidity are included in the model or not. All studies find some evidence of a nonlinear dose-response relation between mortality and pollution levels. Shumway et al (1983) showed that the best two models of those considered involved instantaneous logarithms of either British smoke or sulfur dioxide combined with instantaneous and two-day lagged temperature. The primary effects seemed to lie in the 7-21 day period band. Nearly 50% of the variability in total mortality was explained by the best model in this period range. The two pollutants were highly correlated in all frequency ranges. Models of the same type explained somewhat lower percentages of the variability of cardiovascular and respiratory mortality.

Another population of great interest that has been persistently exposed to ambient pollution concentrations exceeding health-based standards is that of the greater Los Angeles area. Because the pollutants of primary importance are those commonly associated with automobile emissions, the effects of several pollutants besides particulate matter and sulfur dioxide need to be investigated. Weather effects also differ greatly from those of the London Metropolitan area. The availability for the years 1970-1979 of both continuous daily monitoring of a number of pollutants and an extensive database including all deaths occurring in Los Angeles County makes an analysis similar to that of the London data feasible.

This paper considers the problem of modeling total and cardiovascular mortality in Los Angeles County as functions of weather and pollution factors. There were a relatively small number of daily deaths from respiratory causes and this series was not analyzed in detail. The analysis uses time and frequency domain regression techniques to identify the pollutants that contribute significantly to mortality -- carbon monoxide, hydrocarbons, and particulates -- and then estimates three nonlinear dose-response functions that can be used to predict the number of daily deaths as functions of temperature and each of these pollutants. We begin by describing how the data we analyzed were extracted from the available databases.

2. Initial Data Reduction

The raw data used in this analysis consisted of eleven series measured daily in Los Angeles County during the ten years 1970-79: three mortality series, two weather series, and the levels of six air pollutants. The 11 underlying series and the abbreviations that we will sometimes use for the pollutants are listed in Table 1.

The three mortality series were extracted from an extensive mortality file including all deaths other than those caused by accidents, poisoning and violence of Los Angeles residents and non-residents in Los Angeles County during the ten-year period and also the deaths of Los Angeles County residents which occurred in some other locality. The International Classification of Disease(ICD) Codes, 8th and 9th revisions, were used to classify total mortality into respiratory and cardiovascular mortality.

The two weather series were extracted from a file consisting of maximum daily temperature and average relative humidity at Downtown Los Angeles and at four airports -- Los Angeles International, Long Beach, Burbank, and Ontario. The average of the weather variables over the five stations was used in this study.

The six pollutants indicated in Table 1 were measured daily at six monitoring stations located in Azusa, Burbank, Downtown Los Angeles, North Long Beach, Reseda and West Los Angeles. The single daily value for each pollutant used in our analysis was the average of its daily maxima at the six stations. Missing values of the daily maxima were interpolated from those of nearby stations by state-space smoothing methods, as described below. Data series for all pollutants except particulates (KM) were relatively complete. The KM data was often available for at most three stations. All the pollutants are familiar except the KM measurement of particulate concentrations. The KM monitors drew ambient air through a segment of porous tape during two-hour intervals and then measured the amount of light transmitted through the tape. The KM measurement method is considered to be quite sensitive to particles small enough to penetrate deep into the human lung. The adverse health effects of inhaling these extremely small particles are of special concern. A nice survey of the general character of pollution in Los Angeles and its relation to weather factors is given in Tiao et al (1975).

The averaging of pollution values is justified by a geographic consideration of station locations and their contribution to the overall pollution in the Los Angeles area. Figure 1 shows the geographical locations of the air quality and weather monitoring stations on a map of Los Angeles County. Another primary reason for averaging was that mortality was cumulated over the entire area. It should also be noted that averaging weather and pollution variables reduces the overall file from 36 pollution and 10 weather series containing 3652 days each to 6 pollution and 2 weather series containing 3652 days each.

A number of data values in the pollution series were missing, particularly in the particulate (KM) series. Missing values were interpolated using state-space smoothing methods as, for example, is described in Shumway (1988, Section 3.4). The vector of pollutants $\underline{X}_t{}' = (X_{1t}, X_{2t}, \ldots, X_{pt})$ for p stations at time point t is assumed to satisfy the first-order autoregressive model

$$\underline{X}_t = \Phi \underline{X}_{t-1} + \underline{W}_t \qquad (1)$$

where Φ is a p x p transition matrix that summarizes the interdependence among the lagged pollution values and the $\underline{W}_t$ are uncorrelated p x 1 vectors of errors with common covariance matrix Q. The observations are assumed to be generated as a linear function of the unknown vector of pollutants plus an observation noise, say

$$\underline{y}_t = A_t \underline{X}_t + \underline{V}_t \qquad (2)$$

where A_t is a pxp matrix that converts the pollutant vector into the observed series $\underline{y}_t$. If the jth pollutant value is missing the jth row of A_t is taken to be 0; otherwise, A_t is the identity matrix. The uncorrelated vectors $\underline{V}_t$ are assumed to have common covariance matrix R. The parameters in the model are the regression matrix Φ , the noise covariances Q and R, and the initial mean of the pollutant vector.

An example for p=2 stations measuring ozone for Downtown Los Angeles and West Los Angeles is shown in Figure 2. In this case, the original ozone series for the Downtown Los Angeles Station had four values missing, whereas the West Los Angeles Station was completely observed during the short time period used in the interpolation. The parameters in the state-space model (1) and (2) are estimated by maximum likelihood using the EM algorithm. The interpolated values shown in the smoothed Downtown series are the Kalman smoothed estimators (see Shumway(1988,Example 3.20)). In general, because of constraints on computing time, a relatively small number (15) of days was used for each interpolation.

Figure 3 shows portions of the final collection of series, which consisted of 3652 days of recorded values for each of the 11 series in Table 1, measured over the ten year period 1970-79. The main feature in the three displayed series, which cover roughly three years is the long cycle with a period of approximately one year period which appears in each of the series. This seems to be due to the occurrence of higher mortality levels during the winter when temperatures are lower; pollution levels also seem to cycle in parallel with mortality.

Although the data are correlated over time, it is still useful to examine scatterplots relating pollution and temperature to mortality and this is done in Figure 4. Here, the negative relation with temperature is definitely confirmed; there is some nonlinear tendency. The pollution seems to be positively associated with elevated mortality and again there is some graphical evidence of nonlinear behavior.

The preceding indicates that a more systematic investigation of all pollutants and weather effects would be advisable. The next section considers the problem of identifying linear contributors as a function of frequency. We also reduce the data further by linear filtering.

3. Preliminary Identification of Significant Contributors

Since the underlying data series do not exhibit any serious departures visually from what might be interpreted as stationary behavior, it is useful to begin with an analysis of the periodicities evident in the main series. This is done with the idea that eventually the number of points in the file can be reduced to a more manageable size by filtering and subsampling.

An initial spectral analysis of the 11 series showed a dominant period at one year and some minor peaks in some of the pollution series at one month and at one week. The ordinary squared coherences between pollution and all the inputs except sulfur dioxide and relative humidity were uniformly high in this long period band. Table 2 shows a summary of the values of squared coherence for a number of combinations involving one and two inputs and total mortality as an output. The squared coherence can be interpreted as the multiple correlation or percentage of variation accounted for at a particular period. For a description of the methods and equations used for lagged regression of multiple inputs on a single output see Brillinger (1981, Chapter 8) or Shumway (1988, Section 4.2). The best model, involving temperature and carbon monoxide, accounts for about 82% of the variation in the yearly period band. By comparison, for the London data Shumway et al (1983) found that about 50% of the variation in mortality at a 10-day period band could be accounted for by sulfur dioxide and British smoke.

On the basis of the coherences and spectral analyses which indicated that the significant activity was in the low frequency or long period band, it seemed to be sensible to consider filtering the series into the band containing periods longer than 10 days. A lowpass filter was designed to filter frequencies above .10 cycles per day and the resulting filtered series (original and filtered mortality shown in Figure 5) were subsampled weekly beginning on the first day of 1970 to yield a final series for analysis consisting of 508 points for each of the 11 series (Figure 6). These final series can be interpreted as smoothed weekly measures of daily deaths, weather effects and pollution levels. Scatter diagrams relating filtered temperature and carbon monoxide to filtered total mortality do not differ qualitatively from the original scatterplots shown in Figure 4. We therefore conclude that essentially no information has been lost by the filtering and subsampling procedure.

The filtered and subsampled long-period data can be used to develop linear regression models relating the significant inputs to the output mortality measures. In all of the regressions, cardiovascular and total mortality are found to behave in an almost parallel manner. Therefore, the linear and time series regression results will be reported for total mortality.

As a first step, autocorrelations and cross-correlations were computed relating all series. As expected, these functions are dominated by the one-year period. Table 3 shows ordinary correlations computed at zero lag for all variables. Of special interest are the high correlations between total or cardiovascular mortality and four variables--temperature and the pollutants CO, HC and KM. The correlation with temperature is negative, i.e. lower temperatures raise mortality, higher temperatures decrease mortality up to a point. A nonlinearity appears at the upper temperatures which will be discussed in Section 5 (see also Oechsli and Buechley (1970) or Macfarlane(1978)). The near collinearity of temperature and ozone levels should be noted as well as a near collinearity involving the three pollutants CO, HC and KM.

A number of ordinary least squares regression models were run in a stepwise manner to isolate the subsets of variables which seemed to do the best job of predicting total mortality. It is recognized that these estimators will be less efficient than those which take into account the possible correlation in the residuals as has been done by Wyzga(1978), Schwartz and Marcus(1987) and in Section 6 of this paper. It is likely that using the standard Cochran-Orcutt approach with a second-order autoregression fitted to the residuals will satisfactorily emulate the yearly period and yield a fairly efficient estimator under the assumption that the inputs are not random. A summary of the best linear models is shown in Table 4 and it is clear that those involving temperature and one of the pollutants carbon monoxide, hydrocarbons and particulates do the best job. Adding any other dependent variable to the regressions explains negligible additional variability. The regression coefficients under the various models are relatively stable. Of course the quoted standard errors (in parentheses) are computed under the assumption of no correlation in the residuals. We use the results only to indicate that models involving temperature and one of the three pollutants are reasonable candidates for a time series regression or for the nonparametric kernel smoothing or time series regression techniques discussed in Sections 5 and 6.

The above preliminary analysis can be legitimately criticized for its failure to account either for possible nonlinearities in the regression relation or for correlations due to observations being adjacent in time. The next section considers a nonparametric approach to dealing with the nonlinearity question. This leads to reasonable approximations for the nonlinear dose-response profiles relating temperature and the three pollutants to mortality.

4. Nonlinear Dose-Response Relations

The scatterplots shown in Figure 4 raise the distinct possibility that mortality may depend nonlinearly on temperature or pollution or on both temperature or pollution. If we assume that the linear analysis of the preceding sections has been relatively successful in identifying significant contributors, we can try a nonparametric approach to fitting a nonlinear dose-response surface in temperature and one of the significant pollutants. Such methods do not require uncorrelated data but they also do not provide any information about the variability of the dose-response surface.

In order to get an idea as to the general form for the dose-response surface, we consider applying several nonparametric smoothing techniques. The linear analysis of the previous section suggests that it may be reasonable to limit models to those which explain mortality in terms of temperature and one of the pollutants carbon monoxide (CO), hydrocarbons (HC) or particulate levels (KM).

As a first attempt, a nearest neighbor kernel smoothing was applied that replaces each mortality point on a grid by a weighted average of its 10 nearest neighbors. The weights used were inversely proportional to the squared distance to the grid point (see, for example, Ripley(1981, Section 4.2)). The profiles resulting from this smoothing are shown on the lefthand side of Figure 6 for CO and KM (All surfaces and contours were drawn using the Golden Software (1987) computing package). The hydrocarbon surfaces were virtually identical to those for KM and are not shown. A large part of the variability in the scatterpolots of Figure 4 has clearly been reduced and we begin to see distinct patterns for the temperature-mortality and pollution-mortality profiles. For example, in the case of CO, one observes high daily mortality at both low and high temperatures with a minimum number of deaths at approximately 76-80 degrees. The relation along the pollution scale appears to have some curvature and to be strictly increasing.

In order to determine whether the temperature-mortality association might possibly be modeled by some smoother nonparametric or parametric form, rectangular smoothing over 7x3 and 3x5 grids was applied to the CO and KM nearest neighbor profiles. The resulting smoothed surfaces, shown on the righthand side of Figure 6, indicate that the mortality response might be modeled as quadratic in temperature. The CO and KM mortality profiles are smooth and increase at a rate which suggest a logarithmic transformation.

The interpretation of the nonparametric profiles is somewhat hindered by the fact that the large amount of averaging required to produce the final surfaces also introduces "end-effects" which complicate any dose-response predictions for high and low temperature and pollution levels. Also, the nonparametric profiles do not yield standard errors without applying some further computer intensive resampling technique such as the bootstrap. Therefore, it appears that a fairly simple parametric modeling in terms of a linear and quadratic term in temperature and the logarithm of the pollution level would be satisfactory. The procedure could be modified for the effects of time correlation and would produce standard error estimates for the predicted values.

5. Nonlinear Time Series Regression

The relatively smooth nonparametric surfaces shown in Figure 6 suggest a model with linear and quadratic terms for temperature where mortality is varying linearly with the logarithm of pollution. This implies a model of the form

$$M_t = \alpha_0 + \alpha_1 T_t + \alpha_2 T_t^2 + \beta_1 \ln P_t + X_t \tag{3}$$

where M_t is smoothed mortality for week t, expressed in deaths per day. The independent variables are temperature T_t and pollution P_t. The additive correlated errors X_t are assumed to satisfy an autoregressive model of unspecified order; the model that eventually provides the best fit is of order two, i.e.,

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t \tag{4}$$

where ϕ_1 and ϕ_2 are autoregressive coefficients and W_t are uncorrelated zero-mean normal variables with error variance σ^2.

Table 5 shows the estimated autoregressive coefficients for a sequence of nonlinear models of the form (3) involving temperature and CO. It is clear that autoregressive coefficients for models containing more than two terms are small and that the goodness of fit criteria AIC (Akaike(1973)) and BIC (Schwarz (1978)) both go through local minima for a second-order model. Both measures essentially penalize the logarithm of the error variance by a function that is proportional to the number of parameters and the sample size (See section 3.3.6 of Shumway (1988)).

The complete regression results for the three pollutants are shown in Table 6 and we note the similarity between the estimated coefficients relating particulates (KM) and hydrocarbons (HC) to mortality. The error variances are reduced substantially by using the autoregressive error terms and the quoted standard errors indicate that all terms are statistically significant. It is interesting to observe that the coefficients applied to temperature and pollution levels are substantially smaller for the correlated time series regressions. This means, for example, that the predicted number of deaths due to incremental variations in temperature and pollution levels will be smaller for the correlated model.

Figure 7 compares the surfaces fitted by ordinary and time series nonlinear regressions to those obtained by the original nonparametric smoothing techniques as shown in Figure 6. The ordinary least squares surface has the same asymmetry in temperature as the nonparametric surfaces, where the minimum number of deaths seems to occur for temperatures in the neighborhood of 80 degrees. The weighted least squares produces more symmetrical contours, with the minimum response closer to 75 degrees.

These qualitative observations can be enhanced by examining contour plots of the predicted values implied by the estimated parameters in Table 6. Figure 8 shows contour plots for the nonparametric smoothing compared to parametric predictions for the expected mortality computed using ordinary and weighted least squares and the model given in equations (3) and (4). Again, we note that ordinary least squares contours are skewed and predict a higher mean number of deaths.

For the weighted least squares analysis at the ideal temperature of 75 degrees, predicted mortality increases by 14 deaths per day as CO levels increase from 1 to 20, with the steepest rates of increase at the lower levels. At a fixed level of CO, say 10, the average number of deaths increases from 170 at the ideal temperature of 75 degrees to 182 at the lowest and highest temperatures 50 degrees and 98 degrees respectively. The standard errors of the predicted mean mortality values were high, ranging from 20 deaths per day at the lower temperature to 54 deaths per day at the highest temperature.

Figure 9 shows comparable contour plots for the particulate (KM) levels. Again, the ordinary least squares contours are skewed and in fairly close agreement with the nonparametric smoothing contours. The most comfortable temperature from the weighted least squares analysis is about 71 degrees whereas ordinary least squares and nonparametric regression predicts in the neighborhood of 83 degrees. At the most comfortable temperature, predicted daily deaths increased from about 160 at a KM level of 20 to about 172 at a KM level of 80 for a dynamic range of about 12 deaths per day. At a KM level of 52 the 168 deaths per day at the most comfortable temperature of 72 degrees would increase to 184 deaths per day at 98 degrees leading to a dynamic range of about 16 deaths per day in the hotter direction. Deaths per day increased to 176 as temperatures decreased to 52 degrees for a dynamic predicted range of 8 deaths per day.

The model used in this analysis is obviously a specialized one although it is nonlinear and includes the effects of time correlation. For example, equation (3) predicts that the mortality is an additive nonlinear function of temperature and pollution, whereas there may be significant interactions present, especially when low or high temperatures are combined with high pollution levels. One can see some evidence of interaction in the nonparametric surfaces in Figure 6; the rate of increase in mortality as a function of the pollution level at low temperatures may be somewhat greater than at the higher temperatures. Furthermore, at higher pollution levels, the temperature relation departs somewhat from that which would be predicted by a quadratic relation. The problem of taking these possible interactions into account is complicated by the fact that a limited number of data points are available in the regions of extreme temperatures and pollution levels.

Oechsli and Buechley (1970) propose an exponential model for relating temperature and age to mortality. Their results are limited to temperatures above 75 degrees where the exponential and quadratic models would be essentially equivalent. Since the mortality curve begins increasing below 70 degrees, the quadratic function fits better over the entire range.

Oechsli and Buechley were also able to examine age-specific heat mortality which is not done in this study. Our results are primarily oriented toward evaluating the overall effects of pollution in the presence of temperature. Hence, partitioning the original population by age or by any one of a number of other factors such as might be contributed by differing local weather conditions and pollution readings was not attempted. We did find that total mortality and cardiovascular mortality were essentially collinear with correlation .92; respiratory mortality tended to be low and the associations were more tentative.

6. Conclusions

Time series regression and nonparametric smoothing techniques have been applied to a large database containing ten years of daily data from Los Angeles County -- mortality temperature, relative humidity, and the concentrations of six major air pollutants. Possible dose-response relations were examined using nonparametric kernel smoothing and found to be nonlinear in both temperature and each of the significant pollutants. A parametric nonlinear time series model was fitted which could be used to predict average daily mortality during a week as the sum of linear and quadratic functions of average temperature and a logarithmic function of any one of the three significant pollutants. Contour plots are presented that can be used to define acceptable thresholds for the pollutants contributing significantly to mortality.

Mortality fluctuations were associated strongly in the yearly period hand with fluctuations in temperature combined with levels of carbon monoxide, hydrocarbons, or particulates. The relationships may be briefly summarized by: (1) mortality was consistently related to pollution levels and to temperature; (2) mortality increased with increasing pollution at any fixed temperature; (3) for any set of fixed levels of pollution, mortality decreased as temperature increased from minimum values to about 75 degrees and thereafter increased with increasing temperature.

The details pertaining to the dependence of mortality upon levels of individual pollutants are rather complex. Carbon monoxide, particulates, and total hydrocarbons were each consistently and significantly related to mortality, and their individual relationships with mortality were almost equally strong. The nearly equal strengths of the three relationships and the high correlations between the three pollutants make it impossible to attribute the increased mortality associated with increased pollution to any one of them. Hence each of these three pollutants is a surrogate for a general condition of higher pollution associated with increased mortality.

Of the other pollutants studied, nitrogen dioxide and sulfur dioxide were much less strongly related to mortality than were these three pollutants. Although ozone is a major pollutant in Los Angeles, it was highly correlated with temperature. When temperature was included in the model, ozone did not have a significant effect.

The relationships between mortality, pollution, and temperature predict that mortality will increase by about 10% if the significantly related pollutants increase from their minimum levels to the highest levels recorded, with temperature held constant. Approximately the same percentage increases in mortality are predicted for the lowest and highest temperatures, compared to the temperature with minimum mortality, assuming pollution to be constant. Nonsignificant contributors to mortality were sulfur dioxide, nitrogen dioxide, ozone, and relative humidity.

7. Acknowledgements

The authors are indebted to Dr. John Moore of the Air Resources Board for numerous helpful suggestions, particularly during the painstaking data extraction phase. The revision was strongly influenced by the comments made by Professor Will Gersch of the University of Hawaii in his discussion at the Annual Statistical Meetings in San Francisco, August 17-20, 1987.

REFERENCES

Akaike, H. (1973). Information theory and an extension of the maximum likelilhood principle. 2nd. International Symp. on Information Theory ed. B. N. Petrov and F. Csaki, 267-281. Budapest: Akademiai Kiade.

Brillinger, D.R. (1981). Time Series: Data Analysis and Theory. 2nd ed. San Francisco: Holden-Day.

Golden Software (1987). Surfer, Golden Software, Inc., P.O. Box 281, Golden Colorado 80402.

Hall, S.R. (1952). Evaluation of Particulate Concentrations with Collecting Apparatus. Anal. Chem. 24 996-1000.

Macfarlane, A. (1978). Daily mortality and environment in English Conurbations II. Deaths during summer hot spells in greater London. Environmental Research 15 332-341.

Martin, A.E. and W. H. Bradley (1960). Mortality, fog and atmospheric pollution: An investigation during the winter of 1958-59. Mon. Bull. Minist. Health 19 56-72.

Mazumdar, S., H. Schimmel and I.T.T. Higgins (1982). Relation of daily mortality to air pollution: An analysis of 14 London winters, 1958.59-1971/72. Arch. Environ. Health 37 213-220.

Oechsli, F.W. and R.W. Buechley (1970). Excess mortality associated with three Los Angeles Sepstember hot spells. Environmental Research 3 277-284.

Ostro, B. (1984). A search for a threshold in the relationship of air pollution to mortality: A reasanlysis of data on London winters. Environ. Health Perspectives 58 397-399.

Ripley, B.D. (1981). Spatial Statistics. New York: Wiley.

Schimmel, H. (1978). Evidence for possible acute health effects of ambient air pollution from time series analysis: Methodological questions and some new results based on New York City daily mortality, 1963-1976. Bull. New York Academy of Medicine 54 1052-1108.

Schwartz, J. and A. Marcus (1987). Mortality and air pollution in London: A time series analysis. Dept. of Biostatistics, School of Public Health, Harvard University, Boston.

Schwarz, G. (1978). Estimating the dimension of a model. Ann. Statist. 6 461-464.

Shumway, R.H. (1988). Applied Statistical Time Series Analysis. Englewood Cliffs: Prentice Hall.

Shumway, R.H., R. Tai, L. Tai, and Y. Pawitan (1983). Statistical analysis of daily London mortality and associated weather and pollution effects. Technical Report 53, Division of Statistics, University of California, Davis.

Tiao, G.C., G.E.P. Box and W.J. Hamming (1975). Analysis of Los Angeles photochemical smog data: A statistical overview. J. Air Pollution Control Assoc. 25 260-268.

Ware, J.H., L.A. Thibodeau, F.E. Speizer, S. Colome and B.G. Ferris, Jr. (1981). Assessment of the health effects of atmospheric sulfur oxides and particulate matter: Evidence from observational studies. Environ. Health Perspectives 41 255-275.

Wyzga, R.E. (1978). The effect of air pollution upon mortality: A consideration of distributed lag models. Jour. Amer. Statistit. Assoc. 73 463-472.

Table 1: Summary of Time Series Measure in Los Angeles (3652 days beginning January 1, 1970 and ending December 31, 1979)

MORTALITY	1. Total(T)
	2. Respiratory(R)
	3. Cardiovascular(C)
WEATHER	4. Temperature(Tm)
	5. Relative Humidity(RH)
POLLUTION	6. Carbon Monoxide(CO) in parts per 10^6
	7. Sulfur Dioxide(SO2) in parts per 10^8
	8. Nitrogen Dioxide(NO2) in parts per 10^8
	9. Hydrocarbons(HC) in parts per 10^6
	10. Ozone(OZ) in parts per 10^8
	11. Particulates(KM)[1]

[1]One KM unit is defined as that deposit of particulate matter which produces a light abosrbance of 0.1 when a volume of one cubic meter of air has been passed through one square centimeter of the filter (Hall, 1952).

Table 2: Squared Coherence and Multiple Coherence Relating Mortality to the Weather and Pollution Inputs ($\underline{\quad}$ indicates significance at the .001 level)

	Long Period (in wks)			Short Period (in da)							
Period	64	32	4.3	26	21	18	16	14	11	7	4
Variables											
1 vs 4	$.\underline{74}$	$.\underline{60}$	.29	$.\underline{34}$	$.\underline{23}$	$.\underline{21}$	$.\underline{28}$	$.\underline{25}$	$.\underline{22}$	.05	.04
1 vs 5	$.\underline{25}$	$.\underline{37}$	.26	$.\underline{18}$	$.\underline{19}$	$.\underline{19}$	$.\underline{24}$	$.\underline{24}$	$.\underline{15}$	.01	.08
1 vs 6	$.\underline{78}$	$.\underline{82}$	.18	$.\underline{21}$	.09	.04	$.\underline{15}$	$.\underline{18}$	$.\underline{17}$	.10	.07
1 vs 7	.24	.20	.19	$.\underline{20}$	.07	.06	$.\underline{15}$	$.\underline{13}$	.07	.03	.03
1 vs 8	$.\underline{67}$	$.\underline{60}$	.18	$.\underline{24}$	.10	.05	$.\underline{19}$	$.\underline{18}$	$.\underline{14}$	.12	.12
1 vs 9	$.\underline{66}$	$.\underline{64}$	$.\underline{31}$	$.\underline{30}$	.10	.04	$.\underline{19}$	$.\underline{19}$	.10	.01	.06
1 vs 10	$.\underline{67}$	$.\underline{61}$	$.\underline{39}$	$.\underline{34}$	.11	.09	$.\underline{20}$	$.\underline{13}$	.10	$.\underline{13}$	.02
1 vs 11	$.\underline{74}$	$.\underline{78}$	.19	$.\underline{20}$	.07	.03	$.\underline{14}$	$.\underline{16}$	$.\underline{15}$	$.\underline{20}$	.05
1 vs 4,5	.77	.67	.48	.36	.26	.25	.33	.30	.25	.14	.08
1 vs 4,6	.80	$.\underline{82}$	.29	.34	.26	.26	.29	.27	.23	.21	.07
1 vs 4,7	.74	$.\underline{73}$	.29	.34	.25	.22	.31	.28	.23	.12	.05
1 vs 4,8	.79	$.\underline{71}$	.31	.34	.25	.24	.31	.30	.22	.25	.14
1 vs 4,9	.78	$.\underline{69}$	.33	.36	.25	.24	.31	.30	.22	.09	.07
1 vs 4,10	.74	.66	.41	.40	.24	.22	.32	.29	.22	.14	.05
1 vs 4,11	.79	$.\underline{79}$	.29	.35	.27	.23	.29	.27	.23	.27	.04

Table 3: Correlation Matrix (Zero-lag) of Weekly Filtered Series

	Mortality			Weather		Pollutants					
	T	R	C	Tm	RH	CO	SO2	NO2	HC	OZ	KM
T	1.00	.78	.92	-.36	-.23	.50	.07	.20	.41.	-.38	.48
R		1.00	.68	-.33	-.07	.31	-.00	.06	.22	-.32	.27
C			1.00	-.46	-.22	.48	.04	.14	.36	-.48	.46
Tm				1.00	-.30	-.08	.50	.44	.13	.85	-.01
RH					1.00	-.44	-.27	-.37	-.43	-.05	-.43
CO						1.00	.43	.70	.85	-.27	.89
SO2							1.00	.79	.61	.47	.50
NO2								1.00	.82	.35	.74
HC									1.00	-.01	.81
OZ										1.00	-.12
KM											1.00

Table 4: Results of Ordinary Linear Least Squares Analysis; Dependent variable is Total Mortality. Independent variarables are Temperature(Tm), Carbon Monoxide(CO), Hydrocarbons(HC) and Suspended Particulates(KM). 508 Observations

Dependent	Error SS	Regression Coeff.	R-Sq.
Mean	101,985		
Tm	88,754	-.567(.065)	.13
CO	74,956	2.100(.139)	.27
Tm,CO	60,715	-.486(.054) 1.986(.130)	.40
Tm,HC	64,766	-.654(.056) .512(.037)	.36
Tm,KM	66,421	-.554(.057) .438(.034)	.35

Table 5: Determination of Model Order for Autoregressive Errors in Nonlinear Least Squares

Model Order	Autoregressive Coeff.				AIC	BIC
1	.653				4.2718	4.3134
2	.404	.408			4.0932	4.1432
3	.382	.386	.049		4.0938	4.1521
4	.388	.418	.081	-.084	4.0905	4.1571

Table 6: Summary of Parameter Estimates From Ordinary(OLS) and Time Series(WLS) Nonlinear Regressions. Standard errors are in parentheses.

	Carbon Mon.		Hydrocarbons		Particulates	
Est.	OLS	WLS	OLS	WLS	OLS	WLS
Constant	371.66	266.87 (17.37)	330.40	238.65 (17.93)	337.26	240.35 (17.97
Tm(lin)	-5.89	-3.25 (.46)	-6.38	-3.16 (.46)	-6.07	-3.04 (.46)
Tm(quad)	.0362	.0218 (.003)	.0383	.0218 (.003)	.0368	.0214 (.003)
Poll(log)	16.14	10.73 (1.62)	25.12	11.04 (2.12)	20.03	9.01 (1.88)
AR1	.00	.404 (.041)	.00	.427 (.040)	.00	.418 (.040)
AR2	.00	.408 (.041)	.00	.427 (.040)	.00	.441 (.040)
Error Variance	103.65	59.67	112.04	60.33	115.06	60.77

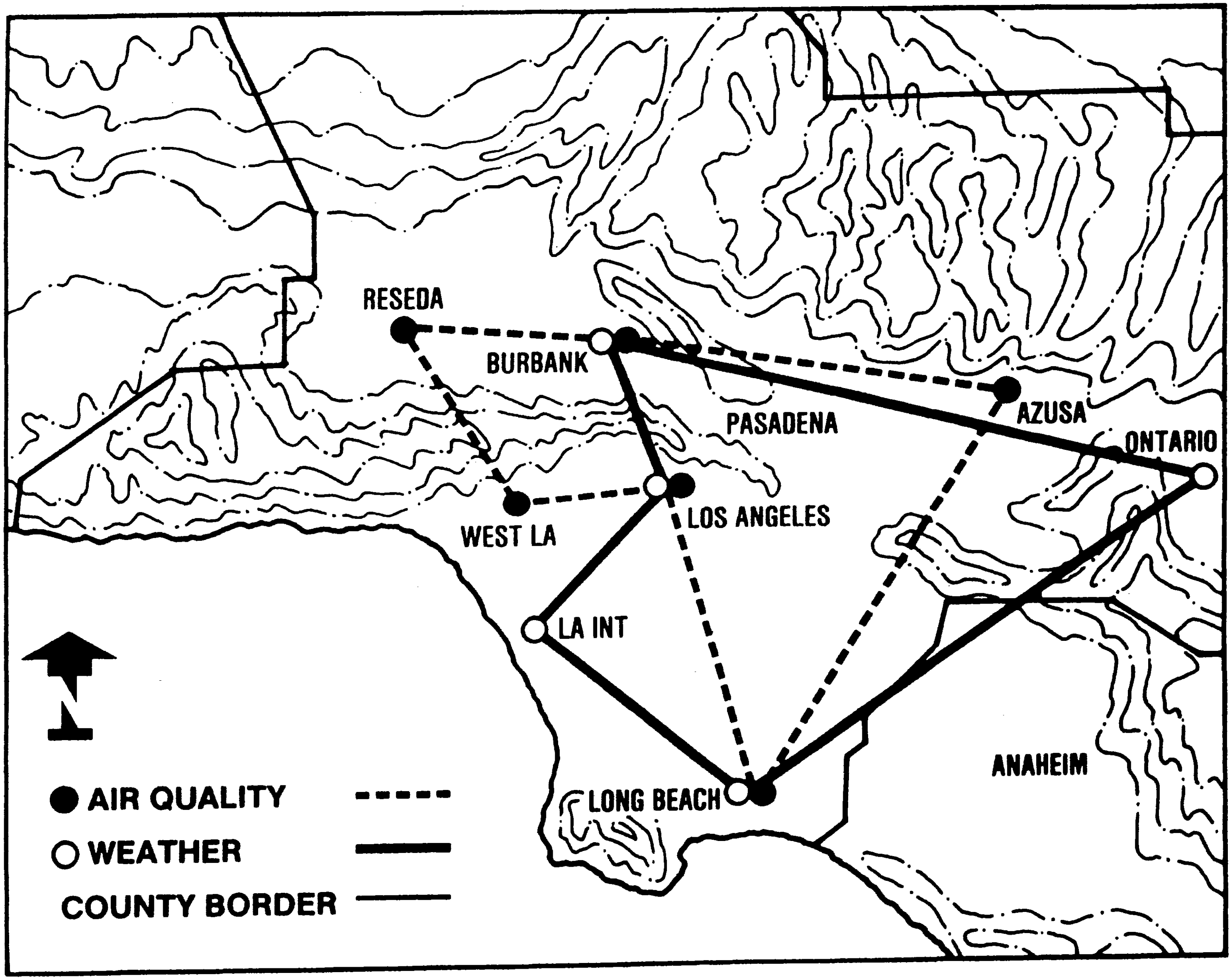


Figure 1: Map of Los Angeles County With Air Quality and Weather Stations

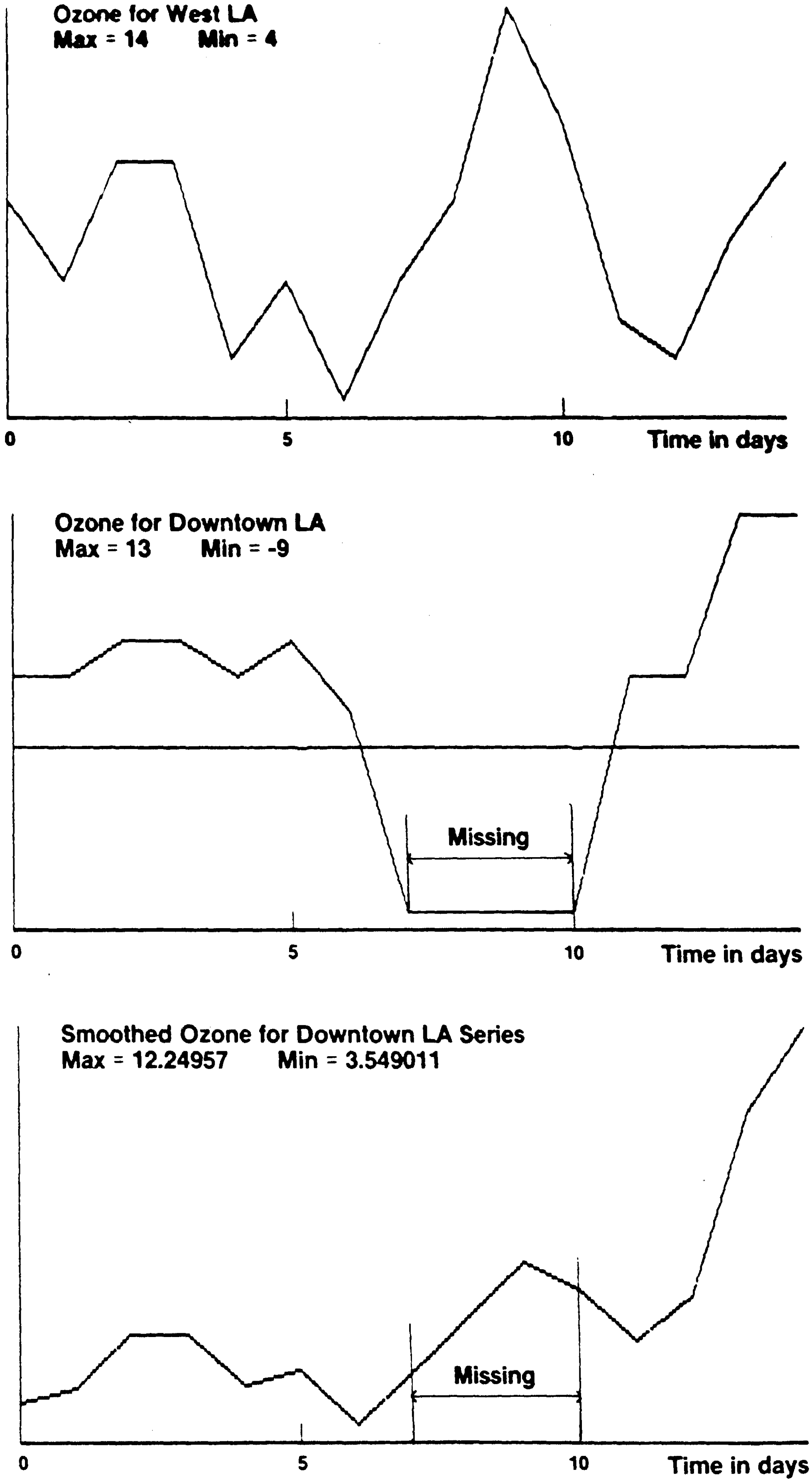


Figure 2: Illustration of Interpolating Missing Ozone Levels Using the Kalman Smoother

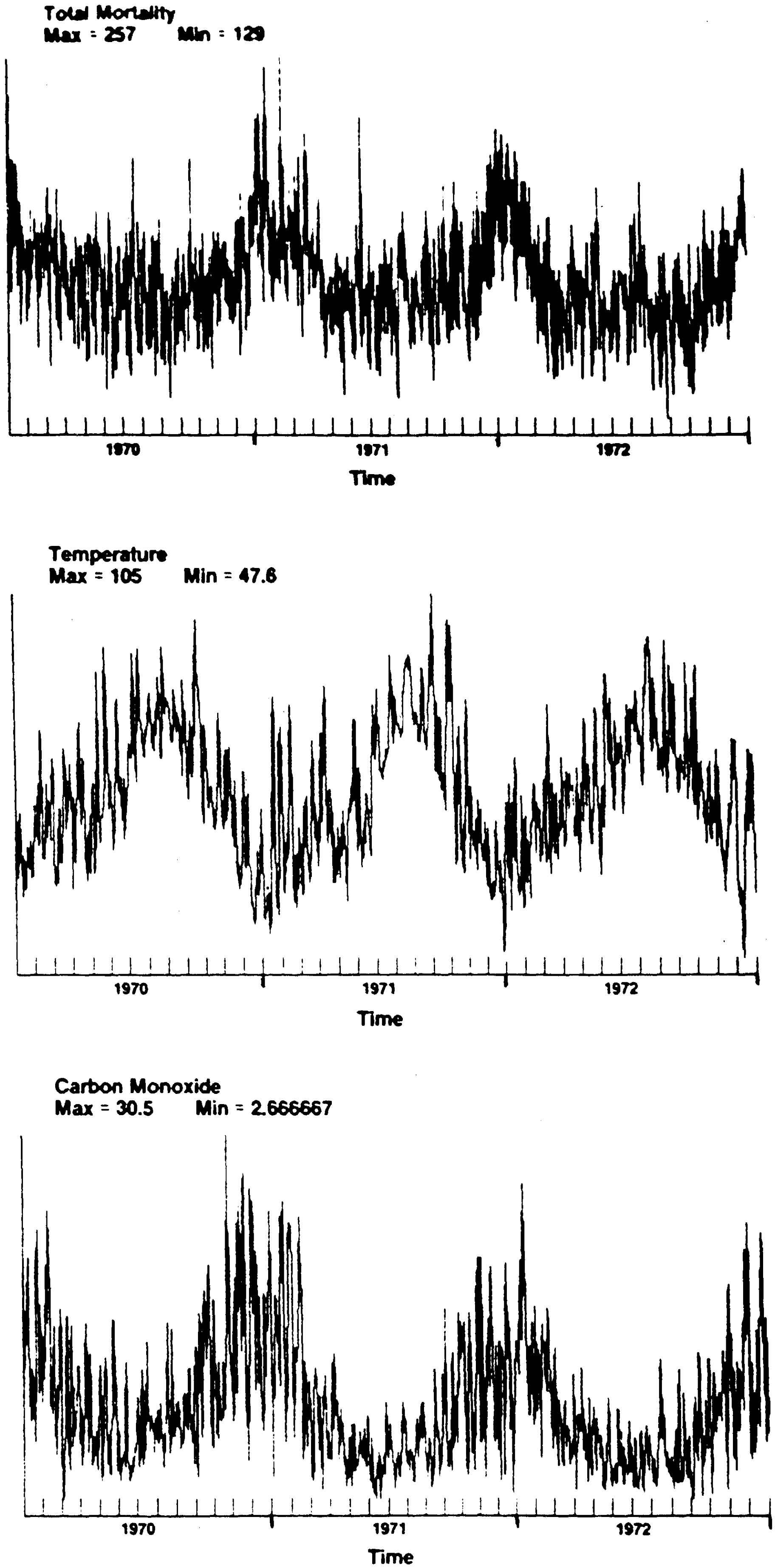
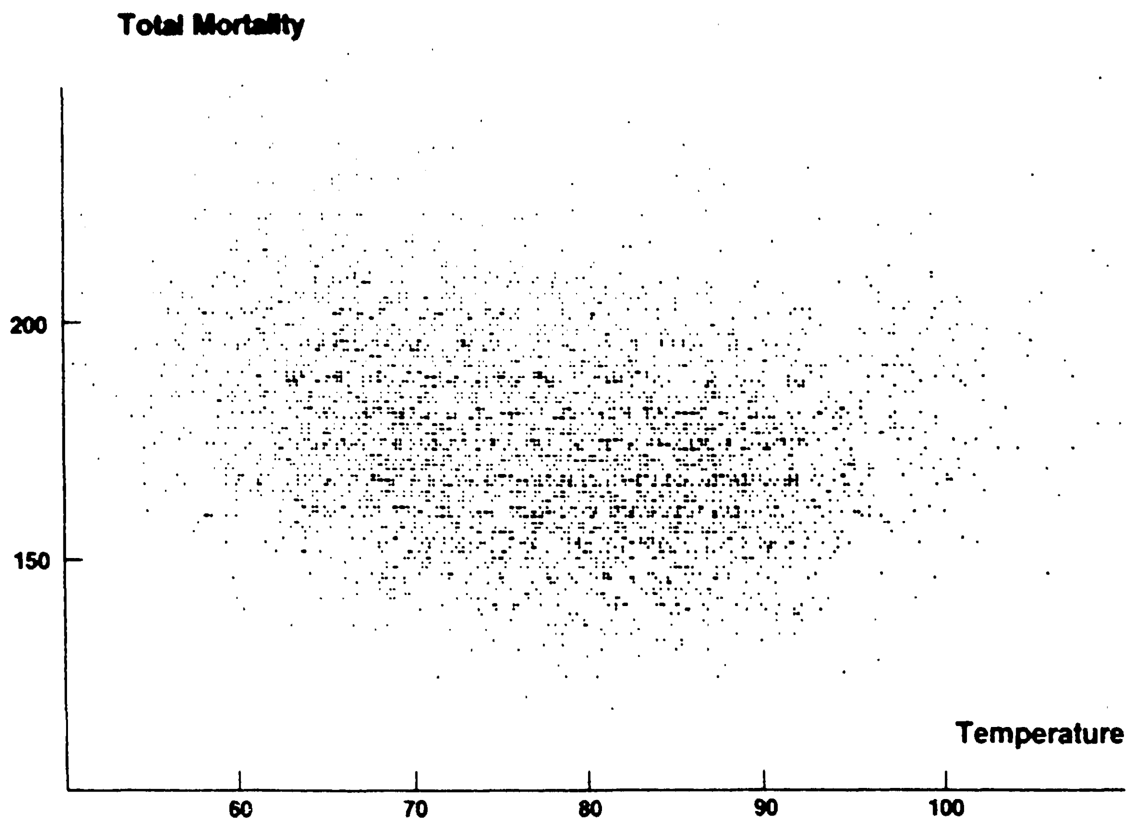


Figure 3: Original Temperature, Carbon Monoxide and Total Mortality Series for 1024 Days (1970-1972)



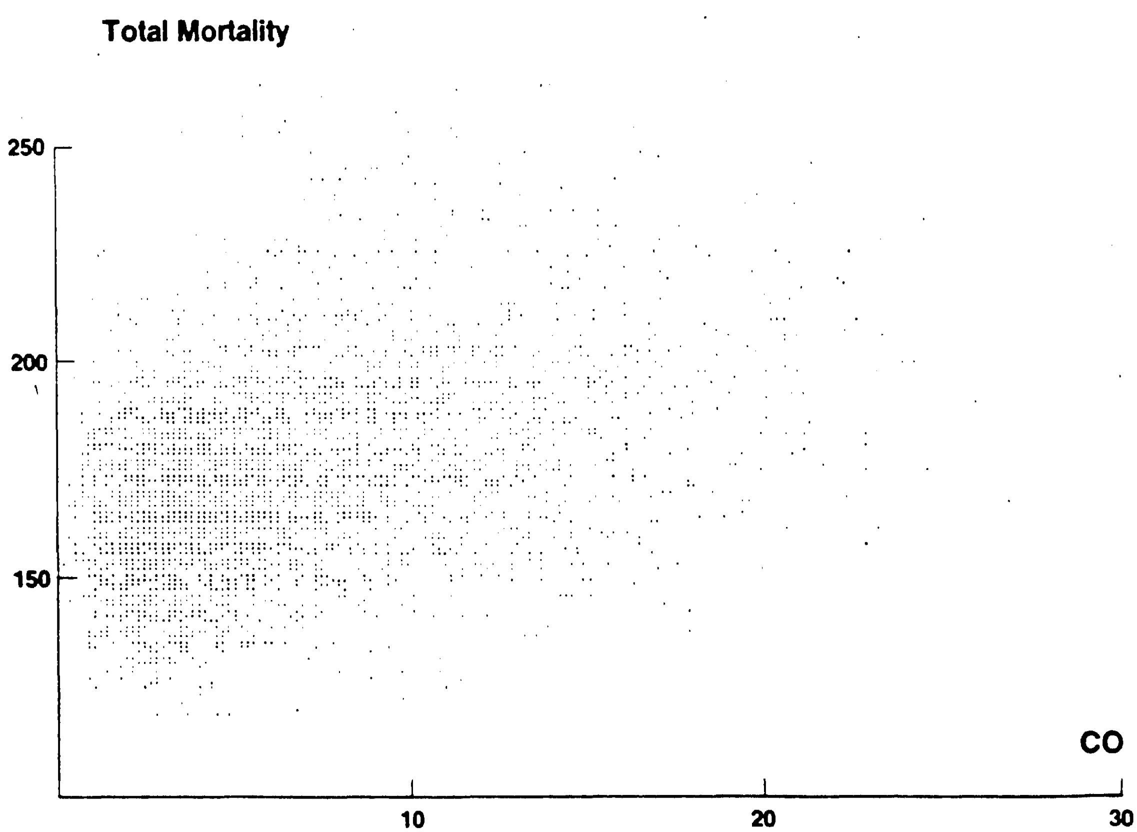
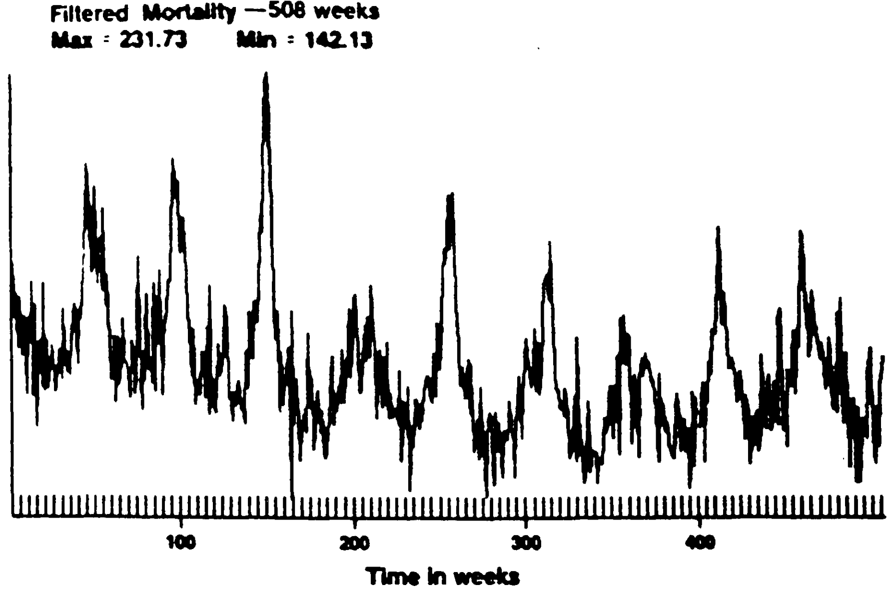
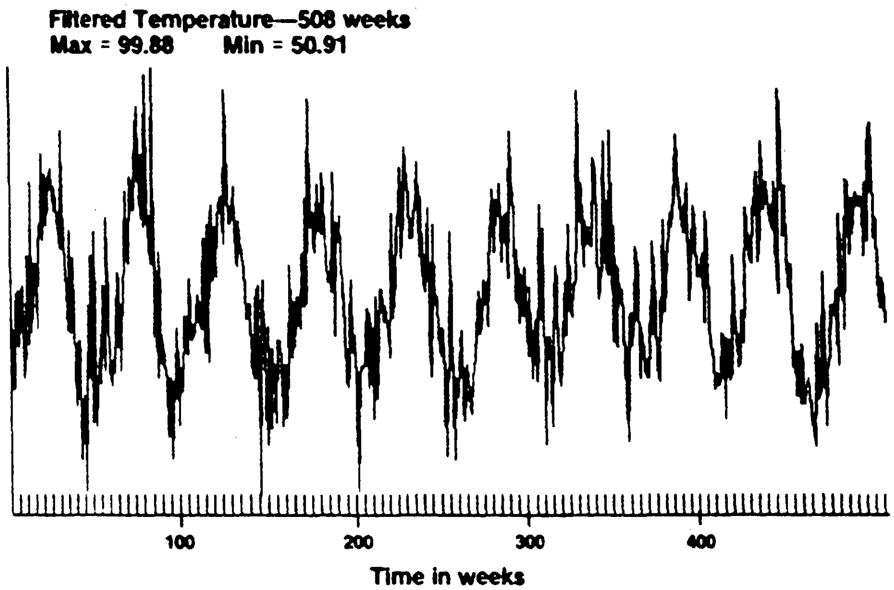


Figure 4: Scatter Diagrams Relating Total Mortality to Temperature and Carbon Monoxide Levels (3652 Days)





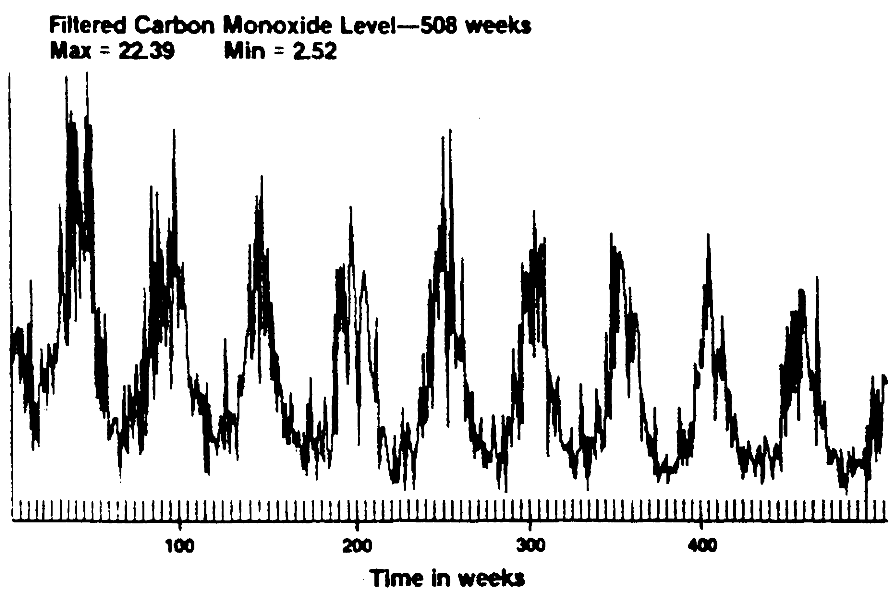


Figure 5: Long Period Mortality, Temperature and Carbon Monoxide Levels

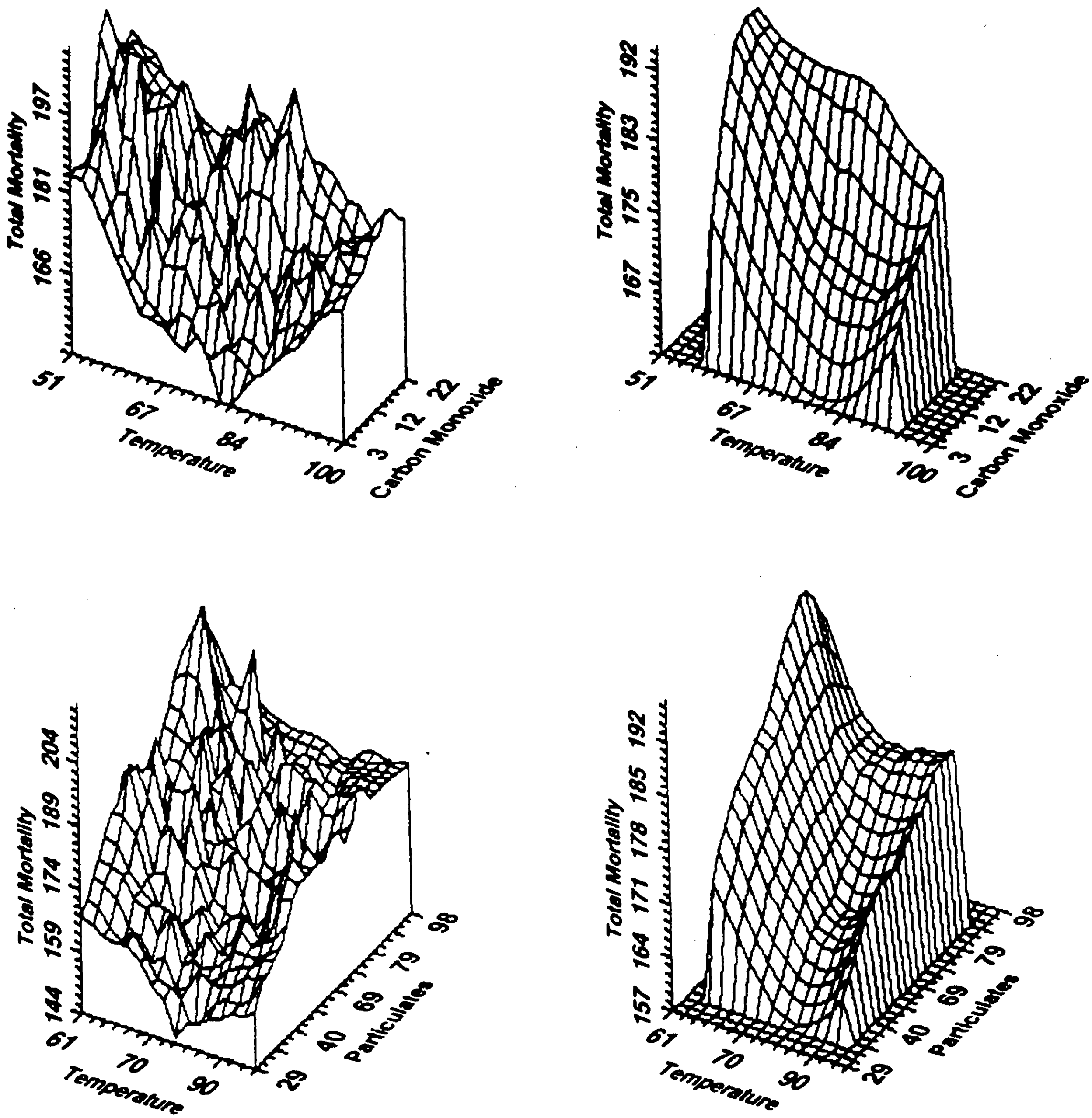


Figure 6: Nonparametric Surfaces for Carbon Monoxide (upper) and Particulates (lower). Lefthand sides are 10-point nearest neighbor smooths weighted by inverted squared distance. Righthand sides are 7x3 and 3x5 rectangular smooths of the nearest neighbor grids.

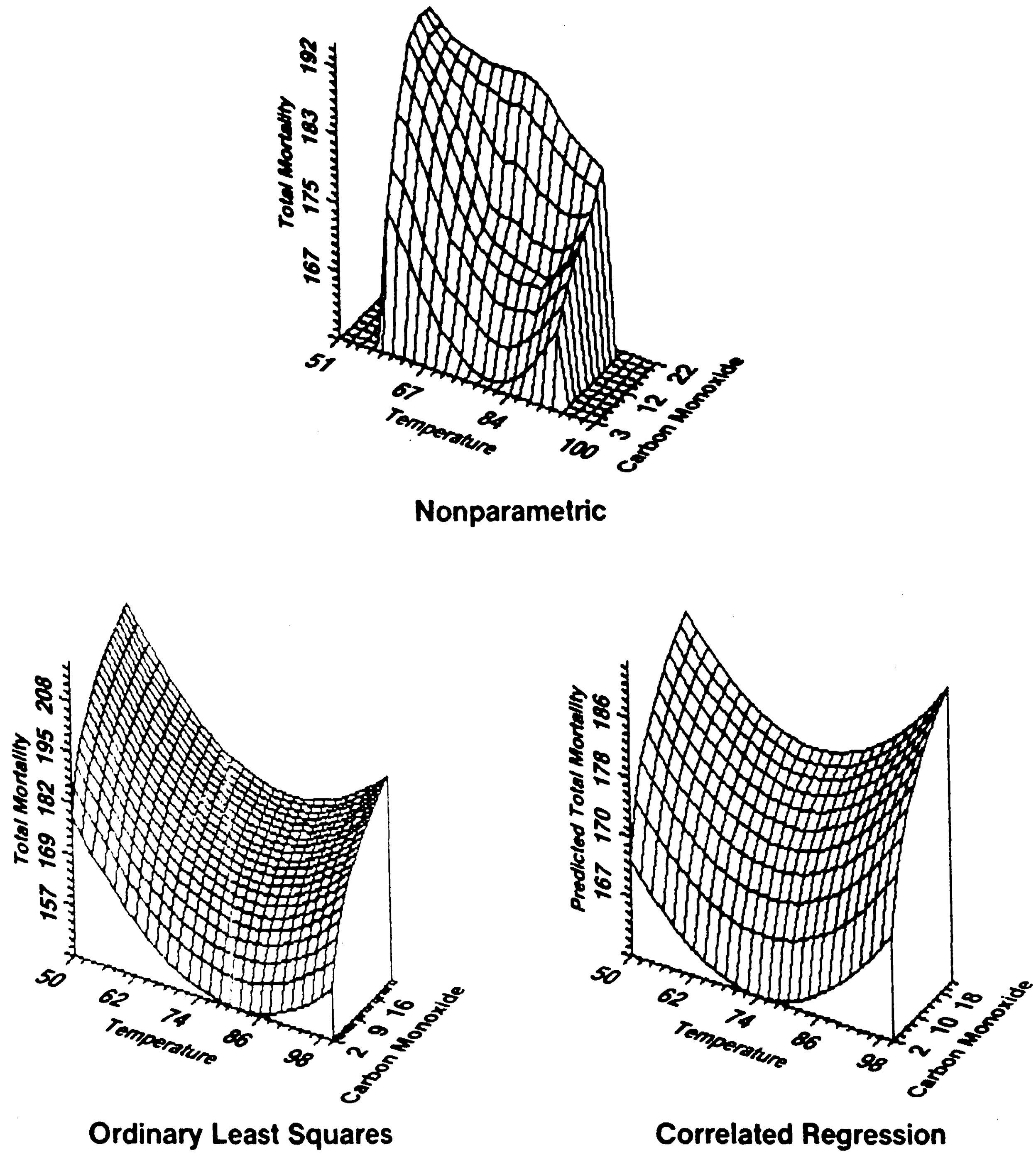
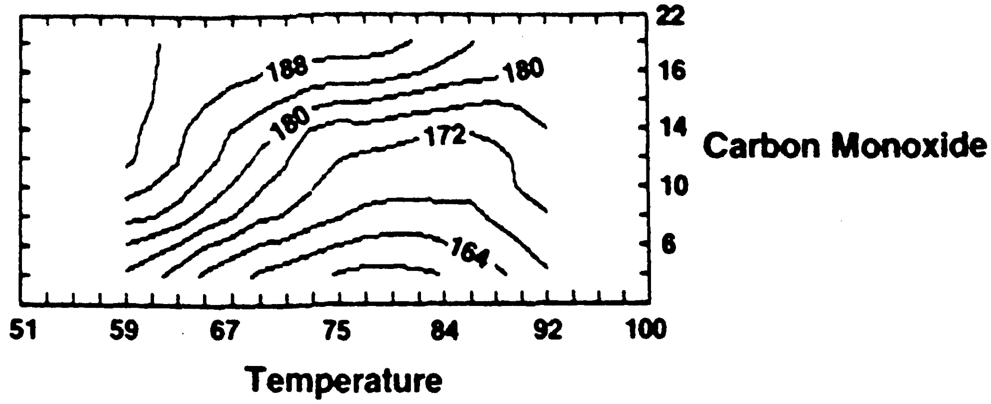
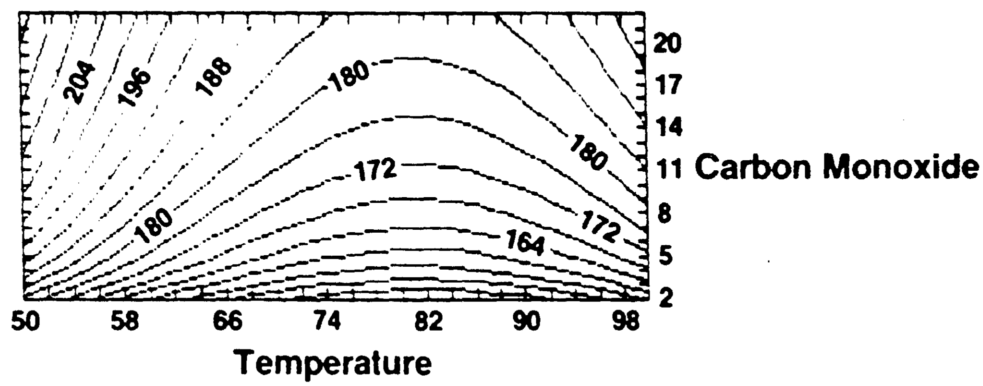


Figure 7: Nonparametric (upper) Compared with Parametric Nonlinear Surfaces (lower). The lefthand side is the ordinary least squares surface. The righthand side is the weighted least squares surface assuming a second order autoregression for the errors.

Nonparametric



Ordinary Least Squares



Correlated Regression

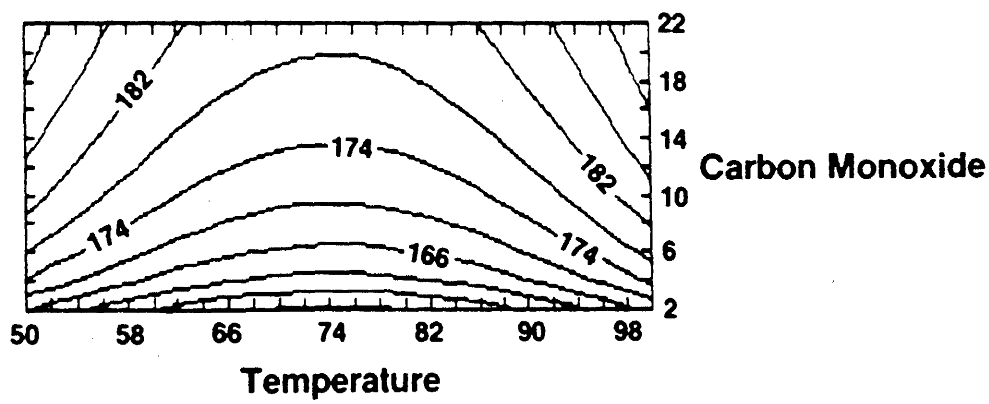


Figure 8: Nonparametric and Parametric Contours for Carbon Monoxide and Temperature. The surfaces are in Figure 7.

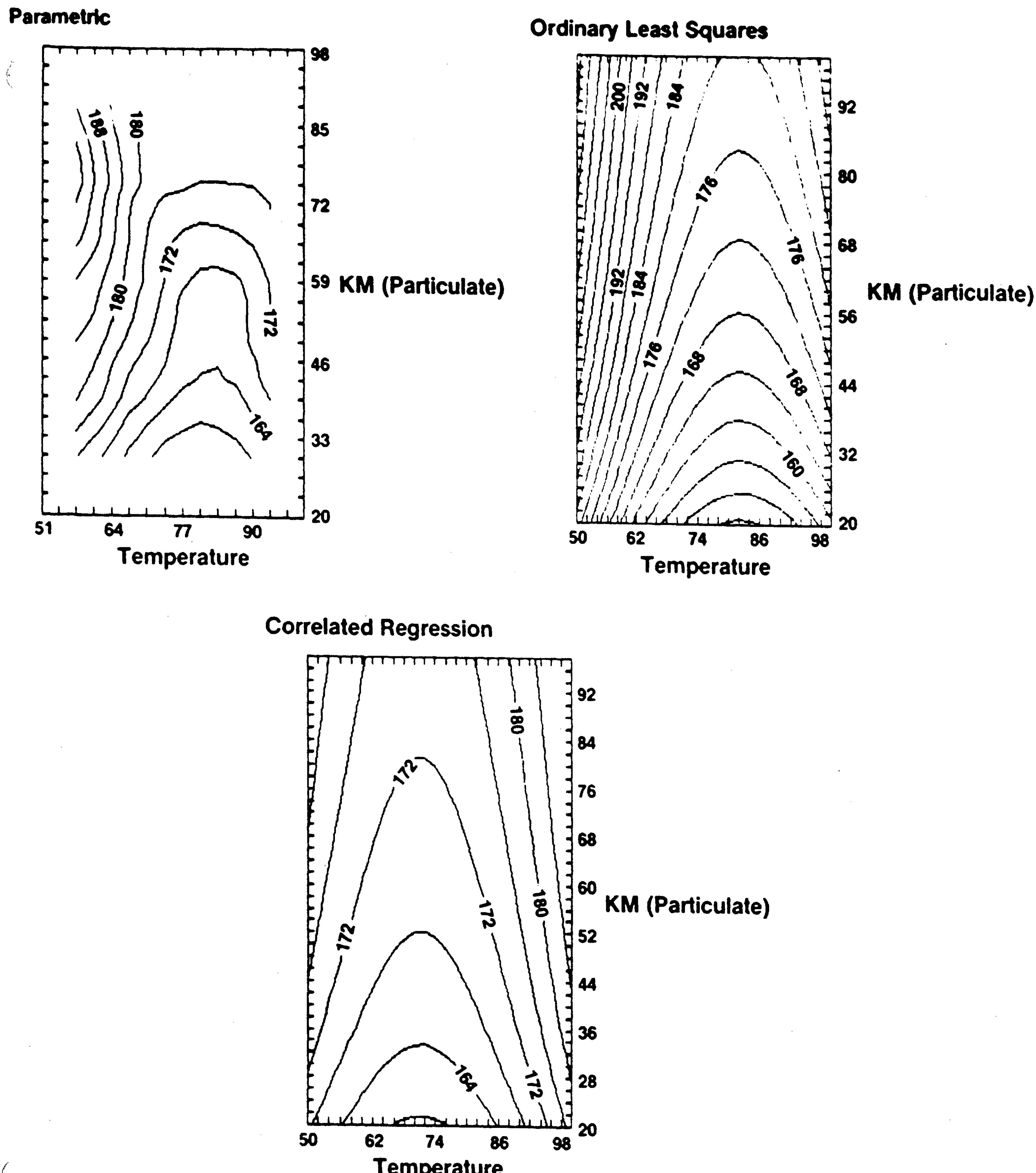


Figure 9: Nonparametric and Parametric Contours for Particulate Level and Temperature