NISS

Airborne Particulate Matter and Daily Mortality in Birmingham, Alabama

Jerry M. Davis, Jerome Sacks, Nancy Saltzman, Richard L. Smith, and Patricia Styer

Technical Report Number 55
November, 1996

National Institute of Statistical Sciences
19 T. W. Alexander Drive
PO Box 14006
Research Triangle Park, NC 27709-4006
www.niss.org

AIRBORNE PARTICULATE MATTER AND DAILY MORTALITY IN BIRMINGHAM, ALABAMA

by

**Jerry M. Davis [1] , Jerome Sacks , Nancy Saltzman
Richard L. Smith [2] and Patricia Styer**

**National Institute of Statistical Sciences [3]
P.O. Box 14162, Research Triangle Park, NC 27709-4162**

November 20 1996

Abstract

As one of a series of studies on the effect of atmospheric particles on human mortality, Schwartz (1993) analyzed mortality data from Birmingham, Alabama, as a function of small particles in the atmosphere (PM10), meteorology, and systematic time trends. His overall conclusion was that when the meteorological and systematic time-trend effects were adjusted for, there is a statistically significant effect of PM10 on mortality. This is a specific and short-term effect: high particles, typically measured through three-day averages, have an immediate impact on daily mortality rates. The findings of Schwartz (1993) mirrored those of a number of other studies by Schwartz and co-workers, which together have contributed to a widely discussed belief that there is a causal relationship between particles and mortality. An independent study by Samet *et al.* (1995) verified the numerical correctness of the results but did not examine Schwartz's selection of statistical models for this data set.

We re-examine the whole question, using the same initial data as Schwartz but incorporating a wider range of meteorological variables. When we use the same variables as included by Schwartz, we obtain similar results to his. But when we use alternative models we obtain different conclusions. In particular, when humidity is included among the meteorological variables (it is excluded in the analysis by Schwartz), we find that the PM10 effect is not statistically significant.

---

[1] Permanent address: Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, NC 27695-8208

[2] Corresponding author: Department of Statistics, University of North Carolina, Chapel Hill, NC 27599-3260; email rs@stat.unc.edu

[3] Research supported by U.S. Environmental Protection Agency under Cooperative Agreement #CR819638-01-0 and by National Science Foundation Grant DMS-9208758

We perform a series of analyses. The first set of analyses introduce models that fit daily death counts to covariates (including PM10) through a variety of standard linear regression methods. The second set duplicates the methodology in Styer *et al.* (1995) used in the study of Chicago and Salt Lake City. A third set of analyses employs Poisson regression. In all instances, we find no significant effect due to PM10. The results we find here are consistent with those of Styer *et al.*: model selection is critical in making conclusions about the effect of particulates on mortality, requiring consistent, defensible approaches to assure reliable interpretations.

Keywords. Causal inference, model selection, observational data, PM10, Poisson regression, semiparametric modelling, cubic splines.

1 BACKGROUND

In recent years, there has been much interest in the health effects of atmospheric particles. Papers such as Schwartz and Marcus (1990) and Schwartz and Dockery (1992a, 1992b) claimed a statistically significant correlation between the level of particles in the atmosphere and daily mortality in the population of 65 years and over. Although various measures of particles have been used in these studies, particular interest has focussed on so-called PM10, i.e. particles of aerodynamic diameter 10 microns or less. The studies themselves used a variety of regression models with deaths as response variable, together with particle counts and a range of meteorological variables as predictors. They also took into account systematic time-trend and seasonal variation, but in all cases reported, produced a statistically significant particles effect.

Recently Samet *et al.* (1995), in a study commissioned by the Health Effects Institute (HEI), re-examined a number of these papers and studies. Particular issues raised by these authors included the effect of other pollutants such as SO_2, which may have a confounding effect; the whole question of causality; whether the alleged effects are on the whole population or whether they only affect a small sub-population of patients who are already very ill (the "harvesting" effect). However, in spite of these complicating factors, Samet *et al.* did not contradict the overall finding that particles affect mortality.

In related work, Styer *et al.* (1995) examined data from Cook County (Illinois) and Salt Lake County (Utah). The study for Salt Lake County failed to show any significant PM10 effect, while that for Cook County, examining effect by season, showed a seemingly significant effect for the Fall season, and a lesser one for the Spring season, with no effect in Summer or Winter. These findings are hard to reconcile with a consistent causal effect. However, the work by Styer *et al.* was based on different data sets from any of those considered by Schwartz and co-workers, or by the HEI group.

In the present paper, we re-examine a study based on Birmingham, Alabama. This city was studied for mortality by Schwartz (1993) and for hospital admissions by Schwartz (1994). The mortality study was repeated by Samet *et al.* (1995), who confirmed the conclusions, but by repeating the same models as used by Schwartz. In other words, their study verified that the data and numerical results were correct, but did not consider the effect of alternative modeling strategies.

For the present study, we have reconstructed the data from their original sources, and have developed independent analyses. Although some of our analyses roughly parallel those of Schwartz, for most of them we have made our own decisions as to which variables to include and the form of the model. However, the emphasis of the study, focussing on short-term effects as measured by three-day PM10 averages, is the same as in the papers by Schwartz. In some of our analyses, we also find what appears to be a statistically significant PM10 effect, but this is highly sensitive to seemingly arbitrary choices about the form of the model. When all the model uncertainties are taken into account, our overall conclusion is that we do not find any consistent effect due to PM10.

2 DATA

To construct a data set for this problem, three sources were combined: mortality, meteorology and PM10. The limitations of the data set are mainly dictated by the availability of PM10 data.

PM10 data

The sampling period is January 1, 1985 through December 31, 1988, the same as in Schwartz (1993). During this period, PM10 data are available from the Environmental Protection Agency's aerometric data base for 13 monitors in the city of Birmingham. The 13 monitors do not necessarily represent different locations: when the type of monitor or the method of measurement changes, this is treated as if it were a different monitor.

For the first seven months of 1985, data from three monitors are available, but they are only collected every six days (the same sampling days for each monitor). Data of this nature are of limited use for studying daily mortality. However, from August 1985 onwards, there is always at least one monitor collecting daily data, except for a small number of days when the data are missing.

Consistently with previous studies, including Schwartz (1993) and Styer *et al.* (1995), which have pointed towards three-day averages of PM10 as the most reliable predictor of mortality, we have computed three-day averages throughout the time period. In other words, for each day, the PM10 value used in the regression is the average PM10 value for that day and the previous two days. This average was taken over all monitors for which data were available in that period. When this algorithm is followed, there are only three days' missing data (i.e. days on which there are no measurements at all throughout the three-day observation period) during the whole period August 1985–December 1988. Subsequent analyses are based on this time period.

Mortality data

Daily mortality data were available from the National Center for Health Statistics for 1985–1988. The data were classified in four ways: by gender (male/female), by race (black/non-black), by age (under 65/65+) and by cause of death (respiratory/cancer/ circulatory/other disease/accidental). For most of our study, we have worked with "total non-accidental deaths", defined as the sum total of deaths in all categories except accidental. This is consistent with Schwartz (1993). Some parts of the study have also been based on "elderly non-accidental deaths", defined as the sum total of deaths of all individuals aged 65+, excluding accidents. This measure has also been used in a number of previous studies.

Meteorological data

The Birmingham meteorological data for this study came from the U.S. National Climatic Data Center in Asheville (NC). The data set is TD3210 (Summary of the Day - First Order). The specific data used were the BIRMINGHAM MUNI AP data. The WBAN number for this station is 13876, and it is located at 33.57 deg N and 86.75 deg W at an elevation of 191m.

The meteorological variables (and their one- and two-day lagged values) considered in this study were as follows:–

tmax:	daily maximum temperature (^{o}C)
tmin:	daily minimum temperature (^{o}C)
mntp:	average daily temperature (average of the max. and min. temp.) (^{o}C)
dpnt:	departure from the normal temperature (^{o}C)
dptp:	average daily dew point temperature (^{o}C)
cldg:	cooling degree days (based on 18 ^{o}C)
htdg:	heating degree days (based on 18 ^{o}C)
aths:	apparent temperature for heat stress (^{o}C)
atcs:	apparent temperature for cold stress (^{o}C)
mxrh:	daily maximum relative humidity (%)
mnrh:	daily minimum relative humidity (%)
mnsh:	average daily specific humidity (g/kg)
awnd:	average daily wind speed (m/s)
pres:	average daily station pressure (mb)
grad:	daily total of incoming solar radiation (MJ/meter squared)

Table 2.1: List of meteorological variables

Here, heating degree days are defined to be 18–mntp when this is positive, otherwise it is 0. Similarly, cooling degree days are mntp–18 when this is positive, otherwise 0. The definitions of apparent temperature for heat or cold stress follow Kalkstein and Valimont (1986, 1987).

Some of these variables (e.g., apparent temperature) were derived from other variables. In addition quadratic terms and product terms were considered. Many other data fields are available from TD3210; however, they were not considered relevant to the problem. The solar radiation data came from the U.S. National Climatic Data Center's National Solar Radiation Data Base (1961-1990).

Summary of previous analyses

From the descriptions in Schwartz (1993), it would appear that the mortality and PM10 data sources are essentially the same as ours. However, the meteorology data he used consisted only of daily mean temperature and humidity at Birmingham airport. Moreover, it would appear that the only meteorological variables actually used by Schwartz were daily mean temperature and an indicator of whether daily mean temperature was over 28°C, both lagged one day. Plots of the data showed that the largest part of the variability in the data was due to sytematic seasonal and time trends, which do not appear to be directly related to either particles or meteorology. Schwartz concluded that these trends appear to follow a two-year cycle (based on four years' data), and modeled the cyclical trend by including 24 sine and cosine terms in the regression equation, together with a linear trend term and an indicator variable for year.

The principal method employed by Schwartz was Poisson regression, using the generalized estimating equations approach developed by Liang and Zeger (1986). Numerous variants on this technique, including corrections for overdispersion and autocorrelation, robust estimation, least squares regression and nonlinear regression for the temperature and PM10 components of the model, were used but apparently without greatly affecting the conclusions. The results consistently showed a statistically significant PM10 effect. In Schwartz's papers this has often been characterized in terms of the relative risk associated with an increase in PM10 levels of 100 $\mu g.m^{-3}$. A typical result for this is 1.11, with 95% confidence interval of (1.02,1.20).

The independent study by Samet *et al.* (1995) focussed on six cities corresponding to analyses by Schwartz and his co-authors, one of which was Birmingham. However, in only one case (Philadelphia) did they seek to reconstruct the entire analysis, including recompiling the data from original sources and making their own judgements about the model selection so as to make a completely independent appraisal of the earlier results by Schwartz and Dockery (1992a). In other cases, including Birmingham, they largely accepted the data and models from the original papers, and concentrated on verifying the numerical results. In most cases they found no major discrepancies but our own concern is primarily about model selection issues.

3 LINEAR REGRESSION ANALYSIS

In this paper we consider a variety of methods of fitting the data, including ordinary linear regression, Poisson regression, and various models involving nonparametric or semiparametric components. In this section we consider methods of analysis in which the number of deaths on each day is modelled as a function of covariates, through a direct linear modelling procedure. For the most part, following Schwartz, "deaths" means all deaths from nonaccidental causes, though for some of the analyses, we restrict ourselves to deaths of age 65 and over. Also for applying linear regression, it is often more appropriate to consider some transformation of the death count, e.g. a logarithmic or square root transformation, rather than raw deaths. In particular, for data which are believed to be Poisson distributed, a square root transformation is widely used, because it approximately stabililizes variance. Normal linear regression based on square roots of deaths is approximately equivalent to Poisson regression based on raw deaths, but is easier to apply because the normal regression is just least squares regression, which is computationally straightforward, whereas Poisson regression involves an iterative procedure and is therefore more time consuming. Therefore, in the initial analysis, normal linear regression is employed together with a square root transformation.

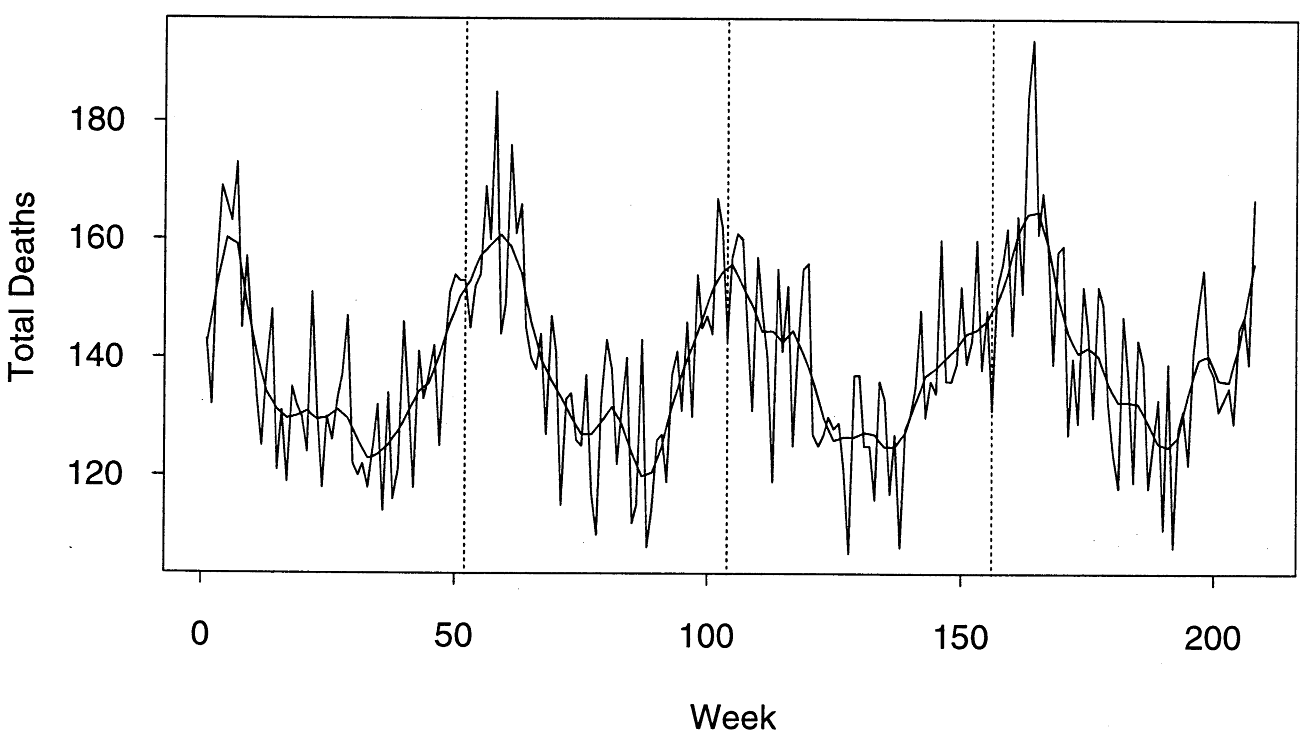


Figure 3.1: Birmingham weekly deaths and smoothed trend

The covariates employed include a variety of meteorological variables, as well as PM10. Before we fit any model, however, it is important to note that there is a very strong trend in these data. Fig. 3.1 displays weekly total deaths (for each week throughout the four-year period, 1985–8) together with a smooth curve obtained via the `lowess` function in SPlus (Becker *et al.*, 1988). (The three dotted vertical lines represent the end of each year.) It is evident that there is a strong seasonal trend, but the pattern is not exactly the same in each year. For example, in both 1986 and 1988 the peak occurred in late February, whereas in 1987 it occured in early January. Also the 1987 peak is smaller than that for 1986 and 1988. It may not suffice simply to model this variation as an annual seasonal cycle.

Schwartz (1993), citing the well-known phenomenon that epidemics tend to follow a two-year cycle, chose to model this systematic variation via a set of 24 pairs of sine-cosine terms, in effect, treating it as cyclic variation with a period of two years. The justification for this, however, seems doubtful when the evidence for a two-year periodocity is based on only four years' data. An alternative is simply to model the trend as a smooth function of time, without any inherent periodicity, and this is the approach which has been adopted in the first part of our analysis here. The basic approach to such analysis is via cubic spline representations, see e.g. Green and Silverman (1994). This topic is reviewed in Appendix A.

Linear regression analysis

We take square roots of total non-accidental deaths as our y variable, and fit a linear model incorporating a smooth trend and meteorological variables (Appendix A, equation A.1). The study is confined to the period August 1985 – December 1988 for which daily PM10 data are available.

Plots of the data, together with previous experience such as the work of Styer *et al.* (1995), suggested that temperature and humidity are the most important variables to be considered, but that it is important to include one- or two-day lagged values of these variables, and also to allow for a nonlinearity in the deaths vs. temperature relationship — for low to moderate temperatures, deaths tend to decrease as temperature rises, but at high temperatures, deaths increase with increasing temperature. Plots of the data suggested an increasing deaths vs. temperature relationship up to around 30 ^{o}C, and then a decreasing relationship. To allow such a relationship to be modelled as a combination of two straight lines, a new variable tg30 was introduced, defined by

$$\mathrm{tg30} = \begin{cases} \mathrm{tmax} - 30 & \text{if tmax} > 30, \\ 0 & \text{if tmax} \leq 30, \end{cases}$$

where all temperatures are in ^{o}C. By including this variable, we seem to avoid the need to build in separate temperature regressions for each season, as was done in Styer *et al.* (1995).

We denote 1- or 2-day lagged variables by suffixes 1 or 2, e.g. mnsh_2 denotes the mean specific humidity from 2 days prior to the current day.

The initial analysis used the four meteorological variables tmax, tmin, mnsh and tg30 together with their 1- and 2-day lagged values, plus 41 spline terms corresponding to "knots" being placed at one-month intervals. Routine variable selection techniques on the meteorological terms quickly showed that we could eliminate all the meteorological variables except for $tmax_1$, mnsh, $mnsh_2$ and $tg30_1$. At this stage, we did not attempt to eliminate any of the spline terms, leaving this for later consideration of the number of knots to include.

Curiously, both the mnsh and $mnsh_2$ terms are statistically significant, though they are of opposite sign. For example, we can test whether the mnsh term is significant, by omitting this term and calculating the deviance statistic (see Appendix A). This leads to a deviance of 8.4 with one degree of freedom, which is highly significant against the χ_1^2 null distribution (P-value 0.004). One possible explanation for this is that the true effect might be seasonal. This was tested by introducing an indicator variable for each season (fall=September, October, November, winter=December, January, February, etc.) and including season$\times$humidity interaction terms, but these did not explain the effect which had been observed.

The next thing to consider is the number of knots in the spline function. Using the 41-knot function as just described, a plot of the trend function (not shown here) shows many minor fluctuations which do not appear to correspond to real variations in the data. In other words, the curve is undersmoothed. One way to remedy this is to reduce the number of knots. The analysis was therefore repeated with 21 and then with 14 equally-spaced knots. In each case the broad pattern of the trend is preserved, but the 21-knot curve is much smoother than the 41-knot curve, and the 14-knot curve is much smoother again. Formal tests based on the deviance suggest that it would be valid to reduce the model to the smoothest, i.e. 14-knot, curve, but as a compromise between smoothness of the regression function and the desire to reproduce the observed fluctuations in the data, we have adopted the 21-knot spline function as our principal model in all subsequent analyses based on this approach.

Source	DF	Deviance
Trend	20	158.57
Meteorology	4	17.38
pmmean3	1	0.70

Table 3.1. Analysis of Deviance table for the linear regression model based on square root total deaths

With this model, we obtain the "analysis of deviance" table given in Table 3.1, based on a 21-knot spline to model the trend, four meteorological variables, and pmmean3. This table has been constructed for an analysis of deviance, rather than the more familiar analysis of variance, primarily so as to facilitate direct comparison with the Poisson analysis of Section 5, where deviance is the most natural measure of model fit. The principal parameter values, along with their standard errors and t values (estimate divided by standard error) are given in Table 3.2. We should also note that weekday effects and meteorology $\times$ season interaction terms were considered, but did not appear to be statistically significant against the model of Table 3.2.

Parameter	Estimate	Standard Error	t value
$tmax_1$	–.0010	.0046	–2.19
mnsh	.011	.007	1.57
$mnsh_2$	–0.16	.007	–2.18
$tg30_1$	.032	.015	2.21
pmmean3	.00086	.00103	0.84

Table 3.2. Parameter estimates, standard errors and t values

The most striking feature about these tables, particularly in view of Schwartz's results, is that the pmmean3 coefficient does not appear to be statistically significant. Translated into the relative risk (RR) values associated with a rise in pmmean3 of 100 $\mu g/m^3$ (the measure most commonly used by Schwartz), we obtain an estimated RR=1.04 but with 95% confidence interval (0.95, 1.14). The point estimate is substantially smaller than Schwartz's, and the confidence interval includes values less than 1, which indicates that the effect is not significant.

It can be seen that the current and two-day-lagged values of specific humidity, mnsh and $mnsh_2$, are both present in this model, with coefficients of opposite sign. This aspect of the model seems hard to explain in terms of any direct effects, but it is a very persistent feature of our analyses, as will be seen throughout the paper. In the case of table 3.2, it appears that the coefficient of mnsh (with a t value of 1.57) could be dropped from the analysis, and in fact if this is done, there is very little change in the conclusion about PM10. The mnsh term is included in Table 3.2 largely for consistency with later analyses where it is significant. On the other hand, if we drop both mnsh and $mnsh_2$ from Table 3.2, the results change considerably. The increase in deviance, when these two terms are dropped, is 5.93. This is almost precisely at the 5% upper-tail probability value of the χ^2_2 distribution, suggesting that the humidity terms are right on the borderline between being significant or not. However, if we drop these terms, then the value of the pmmean3 coefficient rises to .00161 (standard error .00092, t value 1.75) which corresponds to RR=1.07 and a 95% confidence interval of (0.99, 1.16) — still not as large as Schwartz's values, but much closer.

This it appears that if humidity is included — for which there is strong, though not overwhelming, evidence — then the PM10 effect is not significant, but if it is omitted, then the evidence for a PM10 effect is much stronger.

Other aspects of the model fit have been examined. The analysis was repeated using both untransformed and log-transformed value of total nonaccidental death count as the dependent variable, but the square root transformation is easily the best of those three, as measured by log likelihood. Another issue is that of overdispersion, i.e. is the variance of daily death counts greater than what would be expected for a Poisson distribution? One way to examine this question is as follows: assuming the underlying death counts are indeed Poisson distributed, then it can easily be verified that the variance of square roots deaths is approximately 0.25 independently of the Poisson mean. For this regression we obtained a residual variance estimate of 0.2615 with 1,222 degrees of freedom. The ratio of the observed to the theoretical variance is thus 0.2615/0.25 =1.046; a formal test of the hypothesis that the true variance is 0.25 could be based on the statistic obtained by multiplying the variance ratio by 1,222, which has a $\chi^2_{1,222}$ null distribution. In this case the statistic takes on the value 1,278, and the P-value of the associated test is 0.13, i.e. not significant and any commonly used level of significance. Perhaps more importantly than the formal test, an increase of only 5% in the residual variance has negligible influence on the standard errors of the parameter estimates, which is the main reason for being concerned about overdispersion. Based on these calculations, we conclude that it would be valid to perform a Poisson regression analysis without correction for overdispersion, and this is done later in the paper, in Section 5.

We also examined serial correlation among the residuals of the model. The first few serial correlations were –0.032, –0.006, –0.006,..., and again these are very small in magnitude and not statistically significant (the usual test of statistical significance for serial correlations is based on $2/\sqrt{N}$, where N is the number of observations; here this comes to 0.057). This pattern has been repeated in all the analyses of this paper: the serial correlations are not statistically significant, though curiously, they are always negative, which could be indicative of a "harvesting" effect. From the point of view of whether serial correlations affect the standard errors of the parameter estimates, however, there appears to be no effect.

Source	DF	Deviance
Trend	20	135.56
Meteorology	3	15.67
pmmean3	1	1.34

Table 3.3. Analysis of Deviance table for the linear regression model based on square root elderly nonaccidental deaths

The same analysis was repeated using just the elderly (65 and over) nonaccidental deaths as dependent variable, again with a square root transformation. In this case a backwards variable selection resulted in tmin, mnsh and mnsh_2 as the significant meteorological variables. The analysis of deviance table is in Table 3.3, and the parameter estimates are in Table 3.4. All three meteorological variables are statistically significant, but pmmean3 is not. In this case the estimated RR is 1.06, with 95% confidence interval (0.96, 1.16).

Suppose we repeat this analysis using all the temperature-based covariates, but omitting humidity altogether. In this case it seems reasonable that we should repeat the model-selection part of the analysis, and when we do that, using standard forward and backward variable selection techniques, we find that there are two models, each with two meteorological variables, that produce almost identically good fits. The first of these has $tmax_1$ and $tg30_1$ as the two meteorological variables, while the second has $tmin_2$ and $tg30_1$. The second of these is *very slightly* the better as judged by residual mean squared error, but the first seems more logical because it is based primarily on one predictor variable (the previous day's maximum temperature) and is consistent with previous results including Schwartz's. If we add pmmean3 to this model, the resulting model has an estimated RR=1.11, 95% confidence interval from 1.01 to 1.22, whereas if we add pmmean3 to the model based on $tmin_2$ and $tg30_1$, the corresponding figures are RR=1.05, 95% confidence interval from 0.95 to 1.16. Thus the first of these analyses produces results almost identical to those of Schwartz, but the second is quite different, and once again highlights the sensitivity of the results to seemingly arbitrary choices among models. The results of the first analysis are given in detail in Tables 3.5 and 3.6.

Parameter	Estimate	Standard Error	t value
tmin	–0.13	.006	–2.19
mnsh	.025	.011	2.31
$mnsh_2$	–.017	.007	–2.57
pmmean3	.00109	.00095	1.16

Table 3.4. Model for elderly deaths: Parameter estimates, standard errors and t values

Source	DF	Deviance
Trend	20	135.56
Meteorology	2	6.75
pmmean3	1	1.34

Table 3.5. Analysis of Deviance table for the linear regression model based on square root elderly nonaccidental deaths, humidity excluded

Parameter	Estimate	Standard Error	t value
tmax_1	−0.010	.004	−2.70
tg30_1	.024	.015	1.61
pmmean3	.00214	.00094	2.29

Table 3.6. Model excluding humidity for elderly deaths: Parameter estimates, standard errors and t values

In this case we have again tested the model for overdispersion and serial correlation. The results for serial correlation are very similar to those using total nonaccidental deaths: there is a persistent negative pattern, but they are very small in magnitude, and not statistically significant. There is slightly stronger evidence of overdispersion: compared with the anticipated residual variance of 0.25, the actual residual variance is inflated by a factor of 1.083, for which the test based on the $\chi^2_{1,222}$ distribution produces a P-value of 0.022. So this is significant at, say, the 5% level, but it is still not a very strong effect, and we ignore it in subsequent Poisson-based analyses.

Further examination of meteorological effects

The previous analysis is still rather restrictive in the way it has handled meteorological effects, so we now report an alternative analysis in which these were examined in more detail. This analysis was carried out largely independently of the one just reported, and used an approach to the seasonal trend terms much closer to that of Schwartz (1993). Specifically, 12 pairs of sine-cosine terms were included in the model, of the form

$$c_j(t) = \cos\left(\frac{2\pi jt}{730}\right), \quad s_j(t) = \cos\left(\frac{2\pi jt}{730}\right), \quad j = 1, 2, ..., 12,$$

to correspond to a biennial seasonal effect. Each pair of sine-cosine terms was retained if either one of them was significant at the 10% level. According to this criterion, the terms with $j = 1, 2, 5, 7, 9$ were significant for the analysis based on overall nonaccidental deaths, and those with $j = 1, 2, 3, 5, 7, 9$ for a parallel analysis based on nonaccidental deaths in the 65+ age range. In both cases, a square root transformation was applied prior to fitting standard linear models.

At this stage, numerous variables from Table 2.1, together with their lagged values, were added to the model. The covariate with the largest absolute t value was included in the model and the whole procedure repeated until no new covariates were included. As with the preceding analysis, PM10 was not included in the analysis at this stage.

For a y variable consisting of the square root of total nonaccidental deaths, the only variables that turned out to be significant by this procedure were mnsh_2, mnpr (product of mnsh and pres) and mnshsq (square of mnsh). The t values associated with these variables were respectively –3.6, –2.8 and 3.5 (a minus sign indicating that the fitted coefficient was negative). Even with this model the R^2 value is only 0.133, indicating substantial variability not explained by any of the regression terms.

Now if we add pmmean3 to this model, the t values just reported do not change very much (to –3.2, –2.9, 3.5) but the pm10 coefficient is 0.00052 (corresponding to a relative risk of about 1.02) with a standard error 0.00087 ($t = 0.6$) and so clearly not significant.

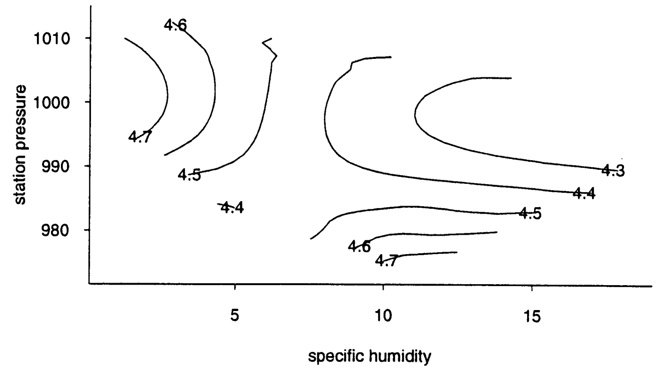
For a corresponding analysis based on nonaccidental deaths in the 65+ age range, the significant meteorological variables were mnsh_2, mnsh and mnshsq with t values –4.1, 3.7, –3.0 respectively. Adding pmmean3 produced an estimated coefficient of 0.00063 (RR=1.03), standard error .00089, t value 0.7.

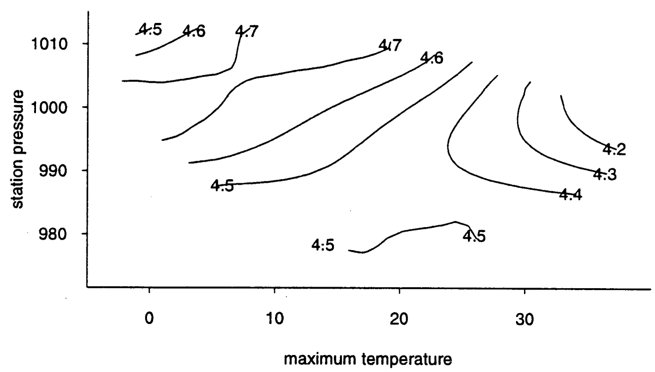
The choice of meteorological variables requires some explanation — especially, the omission of any temperature-based variable from either model. There are a number of high correlations among the meteorological variables. For example, the correlation for tmax$\times$ mnsh is 0.82, for tmin$\times$mnsh 0.94, for dptp$\times$mnsh 0.97, etc. Thus there are likely to be strong confounding effects. Some understanding of this is possible by studying contour or perspective plots of deaths against various pairs of meteorological variables. For example, Fig. 3.2 shows contour plots of square root total deaths against three pairs of variables, (a) pressure and specific humidity, (b) pressure and daily maximum temperature, (c) specific humidity and daily maximum temperature. The model was based on natural splines with 2 degrees of freedom.

The plots show that the highest death counts occurs when pressure is high, and temperature and specific humidity are low. This combination of events represents a winter cold air outbreak scenario for northern Alabama. While not an unusual event for this region, strong cold air outbreaks do not occur with great frequency.

Another peak in the death count occurs when pressure is low, and specific humidity and temperature are moderate. These conditions would generally be found in conjunction with the movement of a moderately strong low pressure system through the region during the spring or fall months. The meteorological conditions associated with these events would normally not be considered stress producing in the usual meteorological sense.

Low death counts are associated with high pressure, temperature and specific humidity. This combination represents the usual summer conditions in northern Alabama. People who have lived in this region for some time have adapted to these conditions, and thus they do not produce the high death counts that similar conditions might produce in a northern city such as Chicago.





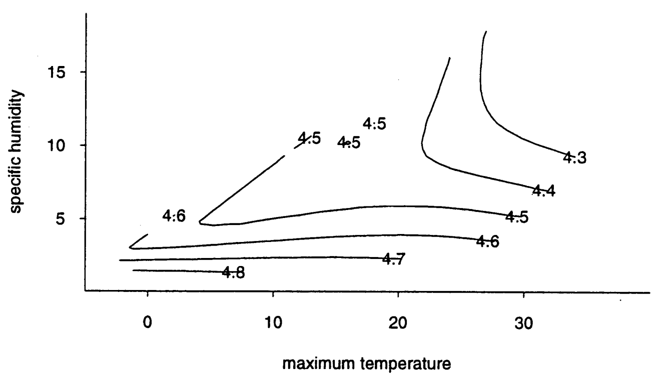


Figure 3.2: Contour plots for square root deaths vs. various pairs of meteorological variables.

4 SEMIPARAMETRIC ANALYSIS

This section approaches the problem of variable selection in a totally different way, which has been previously used by Styer *et al.* (1995) based on foundational work by Sacks *et al.* (1989) and Welch *et al.* (1992). The essence of the approach is semiparametric and Bayesian, and is employed as the basis for an alternative Poisson regression analysis in Section 5. Particular features of this method are that it simultaneously uncovers essential interactions and nonlinearities in the covariates while selecting the important factors. At the same time it imposes weak restrictions on the class of possible functional relationships between response (mortality) and factors. In addition to selection of models, the method can also be used to estimate an overall PM10 effect, though as in other analyses presented in this paper, we did not find any significant PM10 effect. The dependent variable here was taken to be nonaccidental deaths in the 65+ age group.

Separate models were fitted by month and by 3-month season. We used a total of twelve variables, including pmmean3, nine meteorological variables (mntp, mnsh, pres and their 1- and 2-day lagged values) the day of month or season, and the year. The response variable is taken as the square root of non-accidental elderly mortality.

Five of the nine meteorological variables were active factors in at least one of the models. Specific humidity and its lagged values were consistently significant predictors of mortality, with a linear relationship. As in the regression analysis, mnsh had a positive association with mortality while the lagged variables mnsh_1 and mnsh_2 showed negative association. We found that mnsh was active in August, mnsh_1 was active in June, and both were active in summer. Also, mnsh_2 was active in December and in winter. Other active meteorological variables were pres, active in June and December, pres_1, active in March, and mntp_2, active in April, June, and December. The only active non-meteorological variable was day, active in spring. PM10 was not an active factor in any month or season. Possible nonlinear functional forms appeared only for the variables pres_1 and mnsh_2 in March and winter, respectively. There were no apparent variable interactions.

A description of the method appears in Appendix B.

5 POISSON REGRESSION ANALYSIS

We now return to the subject of Poisson regression analysis. Daily deaths are treated as Poisson counts and we include terms for season, the day of the week, weather, and particulate matter. Specifically, we assume that the non-accidental daily death counts are Poisson distributed and that

$$\log EY = X\beta$$

where EY is the expected daily death count on a given day and X contains components for a seasonal trend, modeled using cubic splines, a factor indicating the day of the week on which the death occurred, several meteorological variables, and pmmean3. The day of week has been included for comparison with other studies in which it has been shown to be significant. For some of our analyses, the meteorological variables have been taken to be the same as in Section 3: this permits direct comparison between the Poisson and linear regression approaches. Other analyses have been based on a wider selection of variables (specifically, these were specific humidity, barometric pressure, and the departure of mean daily temperature from the climatic normal mean temperature, together with lagged values of these variables at one and two days) which have then been reduced using the semiparametric modeling results described in Section 4. Many different models have been tried, and only a selection of the results is presented here.

Source	DF	Deviance
	(a)	
Trend	20	180.3
Meteorology	3	15.94
pmmean3	1	0.48
	(b)	
Trend	20	180.3
Meteorology	2	9.33
pmmean3	1	3.06

Table 5.1. Analysis of Deviance table for the Poisson regression model based on square root total deaths, (a) including mnsh_2, (b) excluding mnsh_2.

Tables 5.1 and 5.2 present the Poisson regression analog of the results in Tables 3.1 and 3.2, in which total nonaccidental deaths were taken as the Y variable. The meteorological variables here were taken to be tmax_1, tg30_1 and mnsh_2 — in Section 3, the variable mnsh was also included but this was not found to be statistically significant and it is omitted here. This selection of meteorological variables resulted from a new search, using backwards variable selection, taking tmax, tmin, tg30 and mnsh, together with their one-day and two-day lagged versions, as an initial suite of meteorological variables, and dropping them one at a time. The variable selection was then repeated, from scratch, omitting humidity from the analysis. This again led to tmax_1 and tg30_1 as the significant meteorological variables. Both forms of model, with and without humidity, are reported in Tables 5.1 and 5.2. All runs of the model included a 21-knot spline component to represent the trend, and pmmean3. Weekday effects were also tried but were found not to be significant, and so were excluded from the final model. The pmmean3 coefficient here has a quite different interpretation from Section 3, because the regression function applies on a log scale rather than a square root scale, but it is possible to interpret both sets of results in terms of RR (the relative risk associated with a rise of 100 $\mu g/m^3$) and this is the simplest way to compare the results. From Table 5.2(a), we calculate RR=1.03, with a 95% confidence intervals (0.94, 1.13), for the model including $mnsh_2$. The corresponding results from Table 5.2(b), in which $mnsh_2$ has been omitted, are a point estimate of 1.07 and a confidence interval of (0.99, 1.16). The P-value associated with $mnsh_2$ itself is 0.045. The results are all very consistent with those of Section 3.

All of these results have been obtained using maximum likelihood to estimate the parameters. In several previous papers including those of Schwartz (1993) and Samet *et al.* (1995), a great deal has been made of the distinction between maximum likelihood analyses and various forms of quasi-likelihood (or generalized estimating equations) analysis. However the main motivation for considering the latter type of estimator is that they may provide better protection against overdispersion and serial correlation in the data. As already discussed in Section 3, we do not believe that these are significant factors in the present data set, and for this reason we do not see any reason to depart from standard maximum likelihood analysis.

Parameter	Estimate	Standard Error	t value
		(a)	
$tmax_1$	−.0029	.0018	−1.65
$tg30_1$	.0140	.0066	2.13
$mnsh_2$	−.0060	.0030	−2.01
pmmean3	.00031	.00045	0.70
		(b)	
$tmax_1$	−.0045	.0016	−2.82
$tg30_1$	.0144	.0066	2.19
pmmean3	.00070	.00040	1.75

Table 5.2. Parameter estimates, standard errors and t values, (a) including $mnsh_2$, (b) excluding $mnsh_2$.

One departure which has been tried is to incorporate serial dependence in the model, by including the residual from the previous day's death count as an additional covariate. This gives a form of autoregressive model, but the effect of the additional term is not statistically significant, and the change in the pmmean3 coefficient is minuscule.

For the analysis of elderly mortality, we have again repeated the analyses of Section 3 using Poisson regression, obtaining very similar results. Instead of reporting these in detail, we give instead an alternative analysis which used the semiparametric analyses of Section 4 as the basis for meteorological variable selection. This analysis, however, again identified mnsh and $mnsh_2$ as key variables, along with dpnt (departure from normal temperature).

We considered polynomial and interaction terms for these variables, but found the simpler model to be adequate, especially since the semiparametric model also showed no evidence of important interactions nor of nonlinear relationships for these variables. Fig. 5.1 shows the fitted seasonal trend for both total and elderly mortality. The coefficients for PM10 are not statistically significant in either case. Results for elderly mortality are given in Tables 5.3 and 5.4.

Source	DF	Deviance
Season	20	157.2
Meteorology	3	14.8
pmmean3	1	2.65

Table 5.3. Analysis of Deviance for Poisson regression for elderly nonaccidental deaths

Parameter	Estimate	Standard Error	t value
dpnt	–0.00587	0.00312	–1.89
mnsh	0.00951	0.00506	1.88
$mnsh_2$	–0.00996	0.00337	–2.96
pmmean3	0.00080	0.00049	1.63

Table 5.4. Estimated coefficients, standard errors and t values

Change (from Table 5.4)	New pmmean3 coeff.	New s.e.	New t value
16 DF for season	0.00062	0.00049	1.27
24 DF for season	0.00074	0.00050	1.48
Drop $mnsh_2$	0.00128	0.00046	2.78

Table 5.5. Changes in PM10 coefficient for elderly mortality with changes in the terms of the model

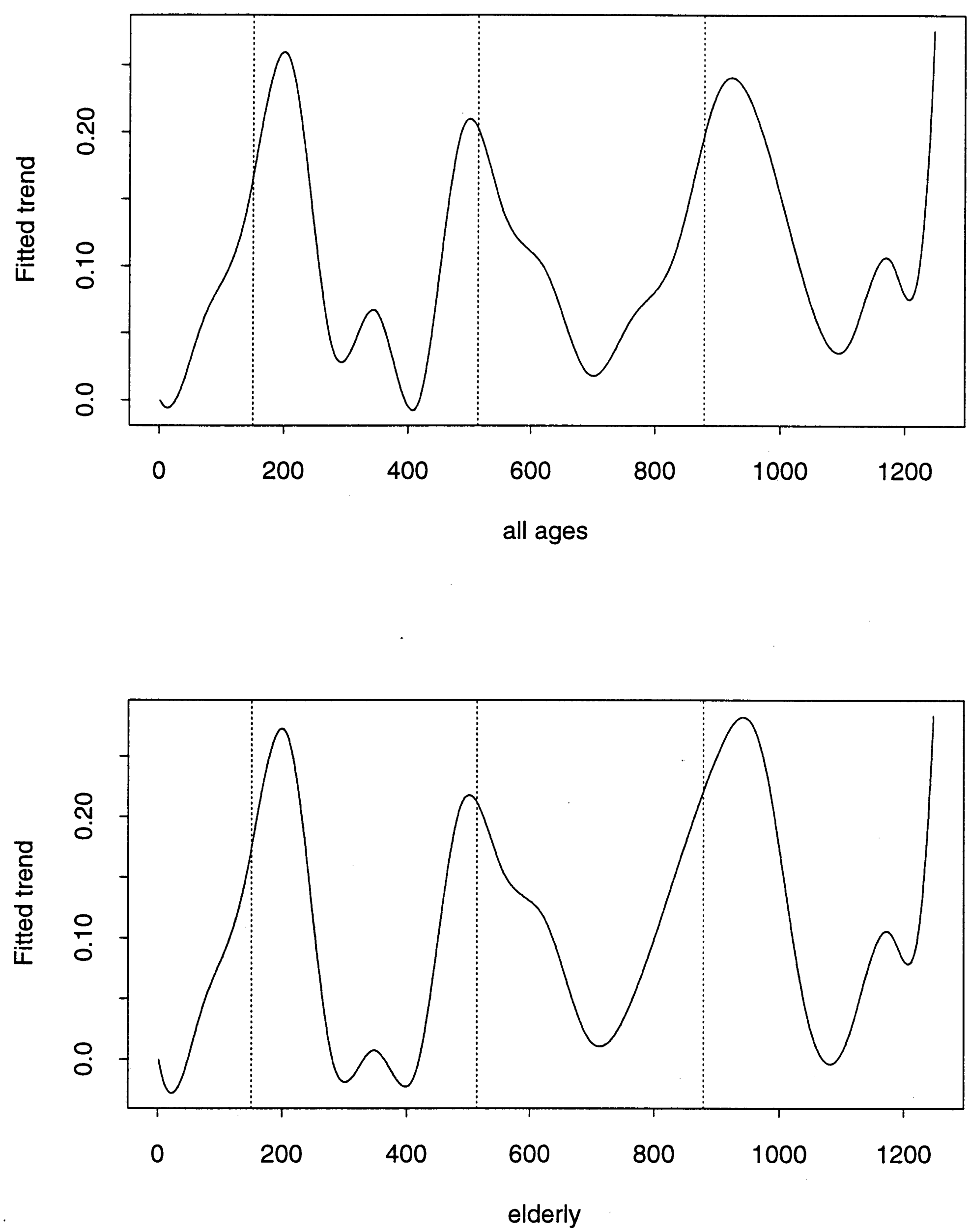


Figure 5.1: Fitted trends in models for total nonaccidental deaths (top plot) and elderly deaths (bottom plots)

We found that small variations in the model can lead to substantial differences in conclusions about the size and significance of the PM10 coefficient. For each mortality response, we altered the specification of the seasonal trend, by either increasing or decreasing the number of knots in the cubic spline, and we dropped the lag-2 specific humidity term. The entries in Table 5.5 show how the estimates for the PM10 coefficient for elderly mortality change with these changes in the model. Each line describes a single change in the base model described above, and the corresponding PM10 estimate. The pmmean3 coefficient varies considerably with different numbers of knots in the spline representation, but the only case when it appears clearly to be statistically significant is when $mnsh_2$ is dropped from the model. In all models examined, the specific humidity term is itself statistically significant, and the inclusion of this term represents the major departure between our analysis and the previous work of Schwartz (1993).

The general use of using multiple regression techniques to quantify the effect of particulate air pollution is open to question because of problems of collinearity among the covariates. General levels of mortality and particulate pollution vary by season, and both can depend on the same meteorological conditions. By making relatively small changes in the specification of the model, the estimated PM10 effect varies considerably. This variability in the estimated coefficients needs to be considered in any conclusions about the possible effects of particulate air pollution.

6 CONCLUSIONS

It is becoming commonplace to use statistical methods based on regression to search for effects in large-scale epidemiological data bases. Many variants of the basic methodology may be tried, such as ordinary least-squares regression (usually applied to some transformation of the response variable, here square root of deaths), maximum-likelihood Poisson regression, Zeger-Liang regression, etc., but it is rare, with data sets such as we have considered here, that it makes much difference which of these methodologies is adopted. Corrections for overdispersion, and for autocorrelations in the daily data, can also be applied but do not seem to be very important. However, what does make a big difference is deciding which covariates to put into the model. In the case of the Birmingham data, these are of two types, (i) a trend component representing seasonal or long-term effects, and (ii) meteorological variables. As far as the trend is concerned, our main departure from Schwartz is to use splines with varying numbers of knots to represent trends of varying smoothness, and (in the case of Section 4) doing separate analyses by month as well as including a year effect. However, although we believe the spline representation leads to a more satisfactory characterization of the trend than the alternative analysis based on trigonometric functions, this does not have a great influence on the most important quantity under study, namely the effect of PM10. On the other hand, the selection of meteorological variables can have a major effect. We have outlined a number of different approaches to this selection, but they all have in common the inclusion of a humidity (and/or lagged humidity) term, which is absent in Schwartz' models. As a result we do not find a significant pm10 component in any model, except in the case when (to imitate

Schwartz) we omit the humidity component. Even in this case the effect is only barely significant when applied to elderly mortality, which is the variable which has been most widely studied in analyses of the PM10 effect.

This sensitivity to the inclusion or exclusion of certain meteorological variables is indicative of the difficulties inherent in this kind of analyses. The reported coefficients and their standard errors differ widely from one model to another, and there are no absolute grounds for choosing any one model as the "right" model. The fact that the model selected for Birmingham is of a quite different nature from models that have been used in previous papers studying Chicago, Philadelphia and many other U.S. cities only adds to the confusion created by the confounding of meteorological effects. In our view there are enormous difficulties inherent in any claim that a variable such as PM10 has a consistent and statistically significant effect over a wide geographical area.

APPENDIX A: CUBIC SPLINE MODELS FOR SMOOTH TRENDS

A simple example of a model which includes both a trend and meteorological covariates is

$$y_t = f(t) + \sum_{j=1}^{p} \beta_j x_{jt} + \epsilon_t \qquad (A.1)$$

where y_t denotes square root deaths on day t, x_{jt} denotes the j'th covariate on day t, and $\{\epsilon_t\}$ are random errors. Here, the covariates represent meteorological variables, PM10, day of week, etc. The function $f(t)$, representing the deterministic trend, is modelled through B-spline basis functions, by

$$f(t) = \alpha_0 + \sum_{k=1}^{K-1} \alpha_k \delta_k(t) \qquad (A.2)$$

where $\delta_1(\cdot), ..., \delta_K(\cdot)$ are K basis functions given by

$$\delta_k(t) = B\left\{\frac{K}{T}(t-\tau_k)\right\} - \frac{1}{K}, \quad 1 \le t \le T, \qquad (A.3)$$

Here $B(\cdot)$ is the (cubic) B-spline basis function, which is a non-negative function, zero outside the interval $(-2, 2)$, with continuous derivatives up to the third order, and total integral 1 (Green and Silverman 1994, pp. 157–8). Also T is the total length of the series and $\tau_k = Tk/K$ is the k'th "knot". The reason for subtracting $1/K$ in (A.3) is so that each of the K basis functions sums to 0, thus allowing α_0 in (A.2) to be interpreted as a main effect and the α_k, $1 \le k \le K-1$ as "treatment effects" corresponding to the first $K-1$ knots. The reason that the sum in (A.2) stops at $k = K-1$, rather than $k = K$, is to avoid indeterminacy — the $K+1$ spline terms represented by $\delta_0(t) \equiv 1$ and $\delta_k(t)$, $1 \le k \le K$ are linearly dependent, but leaving out the term corresponding to $k = K$ removes this degeneracy.

An extension of the model (A.1) is to write

$$y_t = f(t) + \sum_{j=1}^{p} f_j(t) x_{jt} + \epsilon_t \qquad (A.4)$$

where $f_j(t)$ is another smooth function of time represented as a linear combination of B-spline basis functions; $f_j(t) \equiv \beta_j$ is a special case of this. This allows for time $\times$ covariate interaction terms — in particular, if we suspected that the PM10 effect was different at different times (cf. Styer *et al.* 1995), we could model that effect by including such cross-product terms. None of the final models included terms of this nature but several were considered during the model-fitting procedure.

Equation (A.1) defines a special case of the generalized additive model (Hastie and Tibshirani 1990) which is implemented in S or Splus (Chambers and Hastie 1993). The extension given by (A.4) is discussed by Chambers and Hastie (1993) but not currently implemented in S. However, the basis function representation easily allows it to be fitted with standard regression software.

SPlus also has an in-built function (called `bs`) for generating B-spline basis functions, which is not identical to (A.3). We have used both methods of generating these basis functions, with similar though not identical results for the meteorological and pmmean3 estimates.

For comparison between models, it is customary to use residual sum of squares but, to facilitate later comparisons with other methods such as Poisson regression, we use an equivalent likelihood criterion. If M_0 and M_1 are two nested models containing respectively p_0 and p_1 parameters ($p_0 < p_1$) and if L_0 and L_1 denote the values of the negative log likelihood associated with the two models, then the likelihood ratio test of M_0 against M_1 is based on the "deviance" statistic

$$D = 2(L_0 - L_1).$$

Under the null hypothesis that M_0 is correct, D has an approximately $\chi^2_{p_1 - p_0}$ distribution. This test has been used repeatedly in the analysis.

APPENDIX B: THE SEMIPARAMETRIC METHOD

The model starts with the specification of mortality on day t, denoted y_t, as a realization of a stochastic process, $Y_t(x_t)$:

$$Y_t(x_t) = \beta_{year_t} + Z_t(x_t) + \epsilon_t$$

where $\beta_1, \beta_2, \ldots,$ are constants for each year and $x_t = (x_{1t}, \ldots, x_{pt})$ is a vector of covariates for day t. We assume the ϵ_t to be independent normal with mean 0 and variance

σ_ϵ^2. We assume that $Z_t(x_t)$ is a zero mean Gaussian process with covariance function $\text{cov}[Z_t(x_t), Z_t(x_t')] = \sigma_Z^2 R(x_t, x_t')$ where

$$\sigma_Z^2 R(x_t, x_t') = \sigma_Z^2 \exp\left(-\sum_{j=1}^{p} \theta_j |x_{jt} - x_{jt}'|^{p_j}\right).$$

In our case $p = 11$ (corresponding to nine meteorological variables, one "day" variable, and pmmean3), and $1 \leq p_j \leq 2$.

We have data $\{(x_t, y_t),\ 1 \leq t \leq T\}$, where $T = 1,247$ for the period from 3 August 1985 to 31 December 1988. If σ_Z, σ_ϵ, and $R(\cdot,\cdot)$ are known, then the best linear unbiased predictor $\hat{y}^*(x)$ at a new point x in year k can be written as

$$\hat{y}^* = \hat{\beta}_k + \hat{Z}(x) = \hat{\beta}_k + r(x)^T C^{-1}(y - F\hat{\beta}),$$

where $y = (y_1, y_2, \ldots, y_T)^T$, C is the correlation matrix of y, i.e. $(\sigma_Z^2/\sigma^2)R + (\sigma_\epsilon^2/\sigma^2)I$ where $\sigma^2 = \sigma_Z^2 + \sigma_\epsilon^2$, $R = \{R(x_s, x_t), 1 \leq s, t \leq T\}$ is the $T \times T$ matrix of correlations among values of Z at the data points, $r(x) = (\sigma_Z^2/\sigma^2)[R(x_1, x), \ldots, R(x_T, x)]^T$,

$$F = \begin{pmatrix} \mathbf{1}_{T_1} & \mathbf{0}_{T_2} & \ldots & \mathbf{0}_{T_q} \\ \mathbf{0}_{T_1} & \mathbf{1}_{T_2} & \ldots & \mathbf{0}_{T_q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{T_1} & \mathbf{0}_{T_2} & \ldots & \mathbf{1}_{T_q} \end{pmatrix}_{T \times q}$$

corresponding to q years of data with $T_1, T_2, ..., T_q$ days' data in each year, and $\hat{\beta} = (\hat{\beta}_1, ..., \hat{\beta}_q)^T = (F^T C^{-1} F)^{-1} F^T C^{-1} y$ is the usual GLS estimator of $\beta = (\beta_1, ..., \beta_q)^T$. Here $\mathbf{1}_{T_k}$ and $\mathbf{0}_{T_k}$ denote column vectors of ones and zeros of lengths T_k.

As in Styer *et al.* (1995), the parameters σ_Z, σ_ϵ, θ, and $p_1, ..., p_{11}$ are fitted by the method of maximum likelihood. Cross-validation is used to assess variability of estimates. The smoothness of the response surface as a function of the corresponding variables is indicated by p_j ($p_j = 2$ means a high level of smoothness). Relative importance of variables is indicated by their corresponding values of θ when the variables are on normalized scales, with larger values of θ indicating greater importance. During the covariate selection procedure, those coefficients θ which are zero are the factors not included; the others are selected.

REFERENCES

Becker, R.A., Chambers, J.M. and Wilks, A.R. (1988), *The New S Language: A Progamming Environment for Data Analysis and Graphics.* Wadsworth and Brooks/Cole, Pacific Grove, California.

Chambers, J.M. and Hastie, T.J. (eds.) (1993), *Statistical Models in S.* Chapman and Hall, London.

Green, P.J. and Silverman, B.J. (1994), *Nonparametric Regression and Generalized Linear Models: A roughness penalty approach.* Chapman and Hall, London.

Hastie, T.J. and Tibshirani, R.J. (1990), *Generalized Additive Models.* Chapman and Hall, London.

Kalkstein, L.S. and Valimont, K.M. (1986), An evaluation of summer discomfort in the United States using a relative climatological index. *Bulletin of the American Meteorological Society* **67**, 842–848.

Kalkstein, L.S. and Valimont, K.M. (1987), An evaluation of winter weather severity in the United States using the weather stress index. *Bulletin of the American Meteorological Society* **68**, 1535–1540.

Liang, K.Y. and Zeger, S.L. (1986), Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Sacks, J., Welch, W.J., Mitchell, T.J. and Wynn, H.P. (1989), Design and analysis of computer experiments. *Statistical Science* **4**, 409–435.

Samet, J.M., Zeger, S.L. and Berhane, K. (1995), The Association of Mortality and Particulate Air Pollution. In *Particulate Air Pollution and Daily Mortality: Replication and Validation of Selected Studies. The Phase I Report of the Particle Epidemiology Evaluation Project.* Health Effects Institute, Cambridge MA, pp. 1–104.

Schwartz, J. (1993), Air pollution and daily mortality in Birmingham, Alabama. *American Journal of Epidemiology* **137**, 1136–1147.

Schwartz, J. (1994), Air pollution and hospital admissions for the elderly in Birmingham, Alabama. *American Journal of Epidemiology* **139**, 589–598.

Schwartz, J. and Dockery, D.W. (1992a), Increased mortality in Philadelphia associated wuth daily air pollution concentrations. *Am. Rev. Respir. Dis.* **145**, 600-604.

Schwartz, J. and Dockery, D.W. (1992b), Particulate air pollution and daily mortality in Steubenville, Ohio. *Am. J. Epidemiology* **135**, 12-19.

Schwartz, J. and Marcus, A. (1990), Mortality and air pollution in London: A time series analysis. *Am. J. Epidemiology* **131**, 185-194.

Styer, P., McMillan, N., Gao, F., Davis, J. and Sacks, J. (1995), The effect of outdoor airborne particulate matter on daily death counts. *Environmental Health Perspectives* **103**, 490–497.

Welch, W.J., Buck, R.J., Sacks, J., Wynn, H.P., Mitchell, T.J. and Morris, M.D. (1992), Screening, predicting and computer experiments. *Technometrics* **34**, 15–25.