Biostatistics (2001), **2**, 3, *pp.* 337–349
Printed in Great Britain

Respiratory health and air pollution: additive mixed model analyses

BRENT A. COULL*
Department of Biostatistics, Harvard School of Public Health, 655 Huntington Avenue, Boston, Massachusetts 02115, USA
Email: bcoull@hsph.harvard.edu

J. SCHWARTZ
Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts 02115, USA

M. P. WAND
Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115, USA

SUMMARY

We conduct a reanalysis of data from the Utah Valley respiratory health/air pollution study of Pope and co-workers (Pope *et al.*, 1991) using *additive mixed models*. A relatively recent statistical development (e.g. Wang, 1998; Verbyla *et al.*, 1999; Lin and Zhang, 1999), the methods allow for smooth functional relationships, subject-specific effects and time series error structure. All three of these are apparent in the Utah Valley data.

Keywords: Longitudinal data; Nonparametric regression; Peak expiratory flow; Penalized splines; Smoothing.

1. INTRODUCTION

Studies in which day to day changes in air pollution are compared to day to day changes in pulmonary function in a fixed cohort of subjects have become common in the past decade. These studies usually involve approximately 100 daily measurements on each subject over a season. Such studies control for subject characteristics that are constant over the course of the study, as each subject serves as his or her own control.

It is of course possible that characteristics of the subjects affect their response to air pollution. Such interactions have received little attention to date. However, a recent report of the National Research Council (1998), reviewing research needs for the health effects of particulate air pollution, identified the determination of subject characteristics that affected sensitivity as a key research need. Because such cohort studies have the potential to generate large numbers of observations, past computer limitations have resulted in analytical strategies with minimized computational intensity, at the cost of restricting the ability of the study to examine interactions. For example, a widely cited paper by Pope *et al.* (1991) examined peak expiratory flow (PEF) in schoolchildren. Data were collected on each of $n = 41$ children for $T = 109$ consecutive days. For ease of computation, these data were reduced to a data set of T observations by subtracting each child's personal mean peak flow from each of their observations, and

*To whom correspondence should be addressed.

summing those deviations on each day. Such an approach ensures each person is their own control but does not allow an assessment of heterogeneity of response or the sources of heterogeneity. Recently, procedures for fitting mixed models, which allow the assessment of such heterogeneity, have become available in standard statistical packages, and they have begun to be used by epidemiologists in such studies.

In recent years, there has also been increasing attention to the role of seasonality and weather as potential confounding influences on respiratory health. Further, these factors have highly nonlinear influences on respiratory health. Restriction to one season does not eliminate the seasonal pattern; it restricts it to one quarter of the annual cycle. This usually also results in a nonlinear dependence. Recent studies of time series of counts of respiratory events and air pollution have used generalized additive models to address the nonlinear dependence on time and temperature. Hence there is clearly a need for mixed additive models that allow assessment of heterogeneity while addressing nonlinear relationships. To illustrate this approach, we have reanalysed the data of Pope and co-workers.

Additive modelling of longitudinal data is a relatively recent development in the statistical literature. Schwartz (1993) used additive models to analyse daily counts. Altman and Casella (1995) and Staniswalis and Lee (1998) took a nonparametric approach to growth curve analysis. Other nonparametric regression approaches to analysing longitudinal data include those based on kernel smoothers (Zeger and Diggle, 1994), smoothing splines (Anderson and Jones, 1995; Wang and Taylor, 1995; Brumback and Rice, 1998; Wang, 1998; Lin and Zhang, 1999; Verbyla *et al.*, 1999), local likelihood estimation (Betensky, 1997) and penalized splines (Parise *et al.*, 2000). The analyses in this paper build on the latter approach, which has the advantage of being computationally simpler than most other smoothing approaches (Brumback *et al.*, 1999). These additive models describe complex covariate effects on each child's peak expiratory flow while allowing for unexplained population heterogeneity and serial correlation among repeated measurements.

2. UTAH VALLEY STUDY

Pope *et al.* (1991) presented a study investigating the association between respiratory health and respirable particulate pollution in a sample of 41 Utah Valley schoolchildren. For 109 consecutive days beginning on or around December 1, 1990, the study recorded daily measures of PEF on each child, the amount of particulate matter with an aerodynamic diameter less than or equal to 10 μm ($\mathtt{PM}_{10}$), and several weather variables, such as the lowest temperature for that day.

The authors related respiratory performance to air pollution by defining

$$\Delta\mathtt{PEF}_j = \text{average deviation of peak flow from each child's average for day } j,$$

and fitting the linear model

$$\Delta\mathtt{PEF}_j = \beta_0 + \beta_1(\mathtt{PM}_{10})_j + \beta_2 \mathtt{low} \cdot \mathtt{temp}_j + \beta_3 \mathtt{day} \cdot \mathtt{num}_j + \varepsilon_j,$$

while correcting for serial correlation and heteroscedasticity among days in the study. Thus, this approach estimates the effect of particulate matter averaged over the population of children. We consider models in the next section that characterize the effects of particulate matter within each child.

3. MODELS

We reanalyse the Utah Valley data using various *additive mixed models*. Mixed models have a long history in the analysis of data from longitudinal air pollution studies (e.g. Laird and Ware, 1982;

Schwartz *et al.*, 1990), and provide an effective mechanism for accounting for within subject correlation (Laird and Ware, 1982). Additive models (Hastie and Tibshirani, 1990) have enjoyed more recent use in environmental epidemiological research (Schwartz, 1994a,b), stemming from the fact that confounders can be controlled for in a more flexible fashion.

Consider the Utah Valley setting of the previous section. Let y_{ij} denote the measured PEF for school child i, $i = 1, \ldots, 41$ on day j, $j = 1, \ldots, 109$, and let x_j, s_j, and t_j be the daily measures of $\texttt{PM}_{10}$, day in season, and low temperature, respectively. A linear mixed model (Laird and Ware, 1982) for PEF is

$$\mathrm{y}_{ij} = \alpha + U_i + \beta x_j + s_j + t_j + \varepsilon_{ij}, \tag{3.1}$$

where U_i i.i.d. $N(0, \sigma_u^2)$ and ε_{ij} i.i.d. $N(0, \sigma_\varepsilon^2)$ with 'i.i.d.' standing for 'independently and identically distributed'. The fixed effect β represents the linear effect of $\texttt{PM}_{10}$ on PEF, after controlling for linear effects of low temperature and seasonality. The random effects U_i reflect population heterogeneity with respect to baseline PEF unexplained by other terms in the model, and induce an exchangeable correlation among observations measured on the same subject (Diggle *et al.*, 1994). Such heterogeneity is common in respiratory health studies (Schwartz *et al.*, 1988), and is evident in the Utah Valley data with mean PEF for the 41 subjects ranging from approximately 100 to 400 l min^{-1}. In matrix notation, model (3.1) takes the form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}, \tag{3.2}$$

where $\mathbf{y} = (y_{11}, y_{12}, \ldots, y_{nT})^\mathsf{T}$, $\boldsymbol{\beta} = (\alpha, \beta, \delta, \gamma)^\mathsf{T}$, $\mathbf{u} = (U_1, \ldots, U_n)^\mathsf{T}$,

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & s_1 & t_1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_T & s_T & t_T \\ 1 & x_1 & s_1 & t_1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_T & s_T & t_T \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_1 & s_1 & t_1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_T & s_T & t_T \end{bmatrix}, \quad \text{and} \quad \mathbf{Z} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}.$$

One can now fit models having form (3.2) using standard statistical software, such as `SAS` procedure `PROC MIXED` (Littell *et al.*, 1996) or `S-PLUS` function `lme()` (Pinheiro and Bates, 2000).

Model (3.1) has three inflexibilities for the purpose of analysing respiratory health data. First, the model specifies a constant effect, β, of air pollution across the schoolchildren population. In the current context, this lack of subject$\times \texttt{PM}_{10}$ interaction assumption may be violated since there are likely to be unmeasured characteristics that affect a child's respiratory sensitivity to inhaled particulate matter. Second, as noted in the Introduction, recent environmental epidemiological research (e.g. Schwartz, 1994a) has shown the effects of seasonality and weather to have highly nonlinear effects on respiratory health. Thus, the assumption of linear temperature and day effects is likely to be overly simplistic, potentially biasing the estimated effect of $\texttt{PM}_{10}$. These considerations motivate the *additive mixed model*

$$y_{ij} = \alpha + U_i + (\beta + V_i)x_j + f(s_j) + g(t_j) + \varepsilon_{ij}, \tag{3.3}$$

where α, U_i, β, and ε_{ij} are as defined above, f and g are arbitrary smooth functions, and V_i i.i.d. $N(0, \sigma_v^2)$. The random effects V_i allow for heterogeneity in the population with respect to the effect of $\texttt{PM}_{10}$, so that fixed β now represents the average $\texttt{PM}_{10}$ effect in the population of schoolchildren. The smooth functions $f(s_j)$ and $g(t_j)$ relax the linearity assumption for the temperature and day effects in model (3.1).

In order to fit model (3.3), consider first only the $f(s_j)$ term in the model. Let $\kappa_1^s, \ldots, \kappa_{K_s}^s$ be a set of distinct numbers, or *knots*, inside the range of the s_j. The knots are usually taken to be relatively 'dense' among the observations in an attempt to capture the curvature in f. A reasonable allocation rule is one knot for every 4–5 observations, up to a maximum of about 40 knots. Ruppert and Carroll (2000) described an algorithm for choosing the number of knots, and demonstrated its effectiveness through simulation. Now consider an analogous set of K_t knots $\kappa_1^t, \ldots, \kappa_{K_t}^t$ associated with the t_j. Let $a_+ = \max(0, a)$. A linear penalized spline model (Eilers and Marx, 1996) for (3.3) is

$$y_{ij} = \alpha + U_i + (\beta + V_i)x_j + \delta s_j + \sum_{k=1}^{K_s} b_k^s (s_j - \kappa_k^s)_+ + \gamma t_j + \sum_{k=1}^{K_t} b_k^t (t_j - \kappa_k^t)_+ + \varepsilon_{ij}, \tag{3.4}$$

subject to the constraints

$$\sum_{k=1}^{K_s} (b_k^s)^2 < B_1, \qquad \sum_{k=1}^{K_t} (b_k^t)^2 < B_2, \tag{3.5}$$

for some constants B_1 and B_2.

Brumback *et al.* (1999) pointed out that model (3.4), subject to constraints (3.5), is equivalent to the mixed model

$$y_{ij} = \alpha + U_i + (\beta + V_i)x_j + \delta s_j + \sum_{k=1}^{K_s} b_k^s (s_j - \kappa_k^s)_+ + \gamma t_j + \sum_{k=1}^{K_t} b_k^t (t_j - \kappa_k^t)_+ + \varepsilon_{ij}, \tag{3.6}$$

where b_k^s i.i.d. $N(0, \sigma_s^2)$ and b_k^t i.i.d. $N(0, \sigma_t^2)$. Like the linear version (3.1), we can express model (3.6) in matrix form (3.2) with $\mathbf{X}$ and $\boldsymbol{\beta}$ defined as above and $\mathbf{u}$ and $\mathbf{Z}$ now defined as

$$\mathbf{u} = [U_1, \ldots, U_n, V_1, \ldots, V_n, b_1^s, \ldots, b_{K_s}^s, b_1^t, \ldots, b_{K_t}^t]^\mathsf{T}$$

and

$$\mathbf{Z} = [\mathbf{Z}_U | \mathbf{Z}_V | \mathbf{Z}_b],$$

where $\mathbf{Z}_U$ and $\mathbf{Z}_V$ are the corresponding covariate matrices for the random intercepts $(U_1, \ldots, U_N)^\mathsf{T}$ and random slopes $(V_1, \ldots, V_N)^\mathsf{T}$, respectively, and

$$\mathbf{Z}_b = \begin{bmatrix} (s_1 - \kappa_1^s)_+ & \ldots & (s_1 - \kappa_{K_s}^s)_+ & (t_1 - \kappa_1^t)_+ & \ldots & (t_1 - \kappa_{K_t}^t)_+ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (s_T - \kappa_1^s)_+ & \ldots & (s_T - \kappa_{K_s}^s)_+ & (t_T - \kappa_1^t)_+ & \ldots & (t_T - \kappa_{K_t}^t)_+ \\ (s_1 - \kappa_1^s)_+ & \ldots & (s_1 - \kappa_{K_s}^s)_+ & (t_1 - \kappa_1^t)_+ & \ldots & (t_1 - \kappa_{K_t}^t)_+ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (s_T - \kappa_1^s)_+ & \ldots & (s_T - \kappa_{K_s}^s)_+ & (t_T - \kappa_1^t)_+ & \ldots & (t_T - \kappa_{K_t}^t)_+ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (s_1 - \kappa_1^s)_+ & \ldots & (s_1 - \kappa_{K_s}^s)_+ & (t_1 - \kappa_1^t)_+ & \ldots & (t_1 - \kappa_{K_t}^t)_+ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (s_T - \kappa_1^s)_+ & \ldots & (s_T - \kappa_{K_s}^s)_+ & (t_T - \kappa_1^t)_+ & \ldots & (t_T - \kappa_{K_t}^t)_+ \end{bmatrix}.$$

Thus, the additive mixed model (3.6) also falls within the linear mixed model framework, with the special case $\sigma_v^2 = \sigma_s^2 = \sigma_t^2 = 0$ corresponding to the simpler model (3.1). As a result, model fitting for this more complex model is no more difficult than for the linear mixed model.

Finally, previous respiratory health studies have shown that peak flows measured repeatedly on the same subject are likely to be serially correlated (Neas *et al.*, 1995). That is, it is likely that observations taken on consecutive days are much more strongly correlated than observations taken at the beginning and the end of the 109 day study period. This correlation may not be completely explained by the predictors, resulting in correlation among the errors. To check for serial correlation in the Utah Valley data, we inspected the residuals from the fit of preliminary model (3.6) assuming independent errors. A partial auto-correlation plot of the mean residuals for each day suggested the presence of first-order serial correlation. Thus, we fit model (3.6) under the first-order auto-regressive (AR(1)) error structure

$$\varepsilon_{ij} = \rho\varepsilon_{i,j-1} + \xi_{ij},$$

where ξ_{ij} i.i.d. $N(0, \sigma_\xi^2)$. See Diggle *et al.* (1994, Section 5.2) for additional examples of models containing both random intercepts and serial correlation.

Because the mixed model framework easily handles correlated errors, adaptation of model (3.6) to include an AR(1) error structure is also straightforward. The following section reports model fits based on restricted maximum likelihood estimation (REML) of variance components σ_u^2, σ_v^2, σ_s^2, σ_t^2 and AR(1) parameter ρ, unless noted otherwise.

Upon fitting the model, the calculation of predicted curves $\widehat{f}$ and $\widehat{g}$ is straightforward. One can use fixed effect estimates $\widehat{\delta}$ and $\widehat{\gamma}$ and the best linear unbiased predictors (BLUP)

$$\widehat{\mathbf{u}} = [\widehat{U}_1, \ldots, \widehat{U}_n, \widehat{V}_1, \ldots, \widehat{V}_n, \widehat{b}_1^s, \ldots, \widehat{b}_{K_s}^s \widehat{b}_1^t, \ldots, \widehat{b}_{K_t}^t]^\mathsf{T}$$

of the random effects (Robinson, 1991) to calculate

$$\widehat{f}(s_j) = \widehat{\delta} s_j + \sum_{k=1}^{K_s} \widehat{b}_k^s (s_j - \kappa_k^s)_+$$

and

$$\widehat{g}(t_j) = \widehat{\gamma} t_j + \sum_{k=1}^{K_t} \widehat{b}_k^t (t_j - \kappa_k^t)_+.$$

Further, one can obtain variability bands for $\widehat{f}$ and $\widehat{g}$ by adding and subtracting twice the estimated standard error of the estimated function (e.g. Bowman and Azzalini, 1997, pp. 75–76). Bias aside, these bands can be interpreted as approximate pointwise confidence intervals (Hastie and Tibshirani, 1990). They are also useful for detection of leverage and display of inherent variability. For additive models in the mixed model framework, one can easily derive standard errors using standard multivariate statistical manipulations after obtaining an estimate of $\mathrm{Cov}([\hat{\boldsymbol{\beta}}^\mathsf{T} \widehat{\mathbf{u}}^\mathsf{T}]^\mathsf{T} | \mathbf{b})$, where $\mathbf{b} = [b_1^s, \ldots, b_{K_s}^s b_1^t, \ldots, b_{K_t}^t]^\mathsf{T}$ (Zhang *et al.*, 1998). Although this quantity is currently not supported by the mixed model packages, direct computation based on the closed-form solution (Robinson, 1991)

$$\widehat{\boldsymbol{\theta}} = (\mathbf{C}^\mathsf{T}\mathbf{R}^{-1}\mathbf{C} + \mathbf{D})^{-1}\mathbf{C}^\mathsf{T}\mathbf{R}^{-1}\mathbf{y},$$

for $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\beta}}^\mathsf{T}, \widehat{\mathbf{u}}^\mathsf{T})^\mathsf{T}$, where $\mathbf{C} = [\mathbf{X}|\mathbf{Z}_b]$, $\mathbf{D} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} \end{bmatrix}$, $\mathbf{G} = \mathrm{Cov}(\mathbf{b})$, and $\mathbf{R} = \mathrm{Cov}(\mathbf{y}|\mathbf{b})$, is straightforward.

4. INFERENCE

To assess the significance of the estimate of interest $\widehat{\beta}$ and the estimate of the serial correlation parameter $\widehat{\rho}$, one can use the normality of these estimates and compare the magnitude of these estimates to their corresponding standard errors. Other questions of interest relate to the variance components in the model. For instance, both the null hypothesis of a homogeneous $\mathtt{PM}_{10}$ effect (across the population of schoolchildren) and the null hypotheses of linearity for the smooth functions f and g are special cases of model (3.6) in which variance components are zero.

Specifically, consider the seasonality effect in model (3.6). The test of linearity corresponds to

$$\begin{aligned} H_0 &: \sigma_s^2 = 0 \\ H_1 &: \sigma_s^2 > 0. \end{aligned} \tag{4.7}$$

Denote the maximum likelihoods under model (3.6) and that model with $\sigma_s^2 = 0$ as L_1 and L_0 respectively. Because the null hypothesis places the value of the variance component on the boundary of the parameter space, likelihood ratio theory for (4.7) is nonstandard. At first glance, it would appear that one could apply the asymptotic theory of Self and Liang (1987) that states that under H_0 and normal homoscedastic errors, the distribution of the likelihood ratio statistic $l = -2(\ln L_0 - \ln L_1)$ is the same as that of

$$Z^2 I(Z > 0),$$

where Z is a standard normal random variable and I is the indicator function. That is, the distribution of the likelihood ratio statistic is a 50 : 50 mixture of 0 and a χ_1^2 distribution. This asymptotic theory, however, is not directly applicable in the additive mixed model context because this work is based on independent observations, and the random effects $\{b_k^s\}$ and $\{b_k^t\}$ induce dependence among observations on different subjects. Thus, theory on hypothesis testing in additive models is not well developed, and is an area of active research.

As a result, we use the parametric bootstrap (Efron and Tibshirani, 1993) to assess the significance of variance components relating to the questions of interest in the Utah Valley study. In particular, we generate observations from a null distribution $\widehat{F}_0$ for the data under H_0. Given generated data sets $\mathbf{y}_b$, $b = 1, \ldots, B$, we compute the likelihood ratio statistic, l_b, for each data set and estimate the significance level of l by

$$\widehat{p}_{boot} = \frac{\sum_{i=1}^{B} I(l_b \geqslant l)}{B}.$$

As noted by Efron and Tibshirani (1993), the choice for $\widehat{F}_0$ should be a distribution that obeys H_0 and is most reasonable for the observed data. In the additive mixed model context, we generate data from the model (3.6) under H_0 with the maximum likelihood estimates under the null model plugged in for the unknown parameters; that is, for testing (4.7), we generate data from the model

$$y_{ij} = \tilde{\alpha} + U_i + (\tilde{\beta} + V_i)x_j + \tilde{\delta} s_j + \tilde{\gamma} t_j + \sum_{k=1}^{K_t} b_k^t (t_j - \kappa_k^t)_+ + \varepsilon_{ij}, \tag{4.8}$$

where U_i i.i.d. $N(0, \tilde{\sigma}_u^2)$, V_i i.i.d. $N(0, \tilde{\sigma}_v^2)$, b_k^t i.i.d. $N(0, \tilde{\sigma}_t^2)$, and

$$\varepsilon_{ij} = \tilde{\rho}\varepsilon_{i,j-1} + \xi_{ij},$$

with ξ_{ij} i.i.d. $N(0, \tilde{\sigma}_\xi^2)$. Here, the tilde represents the MLE under the null model. Methods for testing other variance components in the model follow analogously.

Table 1. *Results of likelihood ratio tests between pairs of models for testing stated hypotheses*

Null hypothesis	Alternative hypothesis	$-2\log(\text{LR})$	P-value	
			Self and Liang	Bootstrap
f Linear	f Smooth	0.02	0.444	0.275
g Linear	g Smooth	1.30	0.127	0.045
$\sigma_v^2 = 0$	$\sigma_v^2 \neq 0$	5.38	0.010	0.005

5. Analysis of Utah Valley data

In view of the considerations described Section 3 we now consider the model

$$y_{ij} = \alpha + U_i + (\beta + V_i)x_j + f(s_j) + g(t_j) + \varepsilon_{ij},$$
$$\begin{bmatrix} U_i \\ V_i \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}\right), \qquad \varepsilon_{ij} = \rho\varepsilon_{i,j-1} + \xi_{ij},$$

for the Utah Valley data. We fit this model and various sub-models to assess the presence of

(1) nonlinear temperature and seasonal effects,
(2) autocorrelation,
(3) subject specific $\texttt{PM}_{10}$ effects.

Table 1 shows model comparisons based on the likelihood ratio tests of Section 4. P-values are based on both naive application of the asymptotic theory of Self and Liang (1987) and a bootstrap hypothesis test with $B = 200$. We note that in all cases the naive asymptotic approach is conservative, an observation that agrees with results of Verbyla *et al.* (1999) in the smoothing spline context. The test for linearity of g provides moderate evidence that, after controlling for other factors in the model, the effect of low temperature on a child's PEF is nonlinear, whereas the analogous test for f suggests that the effect of day in season is linear. Figure 1 shows a plot of the predicted lowest temperature and seasonality effects on a child's PEF and 95% pointwise confidence bands.

Table 2 shows the estimated mean $\texttt{PM}_{10}$ slope, AR(1) correlation, and variance components and their associated standard errors. The large value for $\widehat{\sigma}_u^2$ reflects the large variability in the mean PEF values for the 41 subjects. The estimate $\widehat{\beta}$ indicates a modest yet statistically significant negative mean effect of $\texttt{PM}_{10}$ on children's PEF, a result that is consistent with the population-averaged analyses of Pope *et al.* (1991). However, results from the likelihood ratio test $H_0 : \sigma_v = 0$ in Table 1 yields strong evidence of population heterogeneity with respect to $\texttt{PM}_{10}$ response. Figure 2, which shows a kernel density estimate of the predicted values $\widehat{V}_i$, demonstrates this heterogeneity. This plot and Figure 3, which shows normal probability plots of the predicted values $\widehat{U}_i$ and $\widehat{V}_i$, suggest that the distribution of the random slopes is skewed with three subjects exhibiting a strong negative reaction to PM_{10}. The estimated AR(1) parameter $\widehat{\rho}$ and associated standard error confirm the presence of residual correlation above that modelled via $\{U_i\}$ and $\{V_i\}$.

Finally, we examine the sensitivity of conclusions to model assumptions. First, to check the plausibility of a linear PM_{10} effect on PEF, we fit more general models leaving the form of this effect unspecified. REML estimation for this model selected the linear form for the PM_{10} effect, with estimated variance component associated with this smooth term equal to zero. Second, we remark that, since both the smooth term for $f(s_j)$ and the autoregressive process $\varepsilon_{ij} = \rho\varepsilon_{i,j-1} + \xi_{ij}$ reflect PEF associations over time, estimates of these processes are not independent. For the Utah Valley data, for instance, Figure 4 shows the predicted curve $\widehat{f}(\cdot)$ given model (3.6) and independent errors.

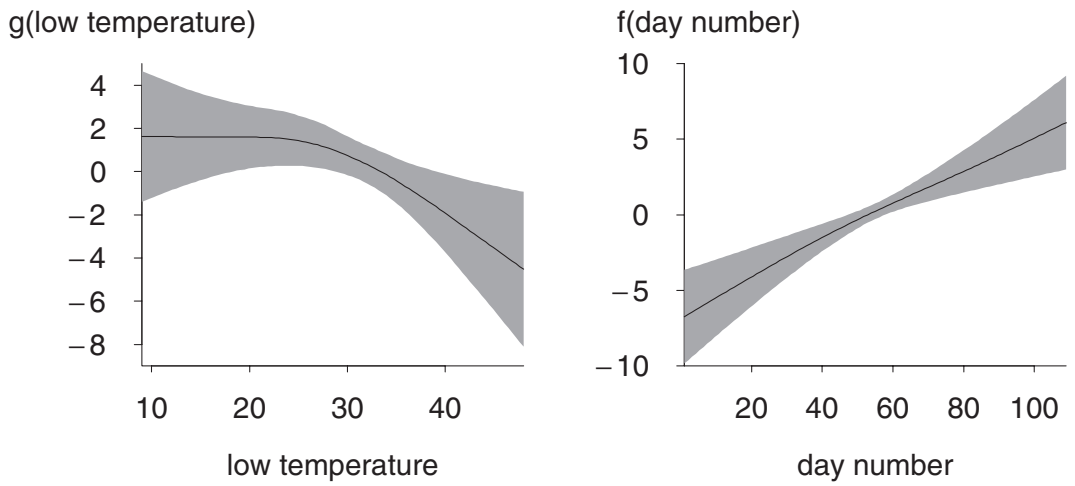

Fig. 1. Fitted low temperature and day number effects for model (3.6) with AR(1) errors.

Table 2. *Results of fitting full model to Utah Valley data*

Parameter	Estimate	Standard error
β	−0.070	0.025
ρ	0.510	0.014
σ_u^2	2630	598
σ_v^2	0.009	0.005
σ_s^2	0.007	0.011
σ_t^2	< 0.001	0.004

The difference between this plot and the second plot in Figure 1 is the effect of the correlated error assumption on the estimated day in season effect. Although we report the estimate from the AR(1) model (based on the significance of $\widehat{\rho}$), the choice has little practical implication in the current study as the difference in the estimate of interest, $\widehat{\beta}$, is minimal (-0.070 versus -0.086) and does not change the substantive conclusions of the analysis. Third, one might expect a child's baseline PEF level and susceptibility to particulate matter to be correlated, suggesting the formulation

$$\begin{bmatrix} U_i \\ V_i \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} \sigma_u^2 & \sigma_u\sigma_v\rho_{u,v} \\ \sigma_u\sigma_v\rho_{u,v} & \sigma_v^2 \end{bmatrix}\right)$$

for the random effects. SAS memory requirements made fitting this model using REML infeasible. However, minimum variance quadratic unbiased estimation (Searle *et al.*, 1992) of the variance components yields an estimated mean PM_{10} effect of $\widehat{\beta} = -0.068$ (SE $= 0.025$), demonstrating that conclusions are also robust to the assumption of independence for the random effects.

6. DISCUSSION

The analyses in this paper represent flexible methods for accounting for both the usual correlations encountered in longitudinal data and complex covariate effects. The use of penalized splines affords

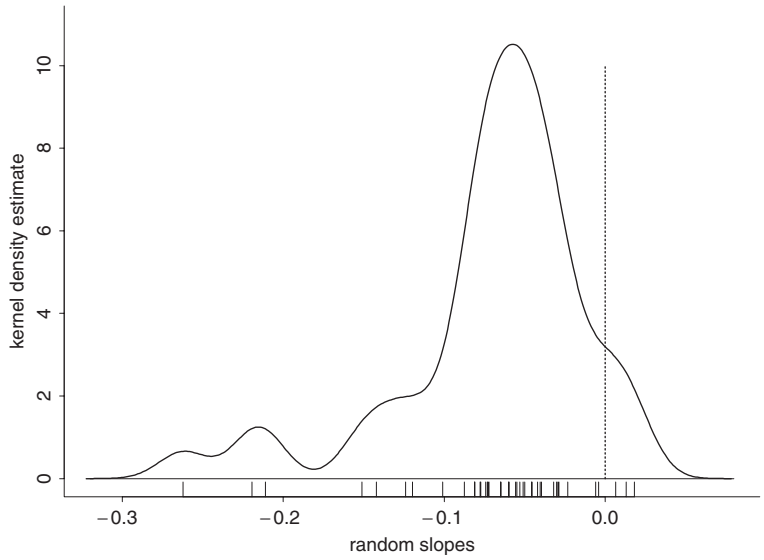

Fig. 2. Kernel density estimate of $\widehat{\beta}_1 + \widehat{V}_i$ values for model (3.6).

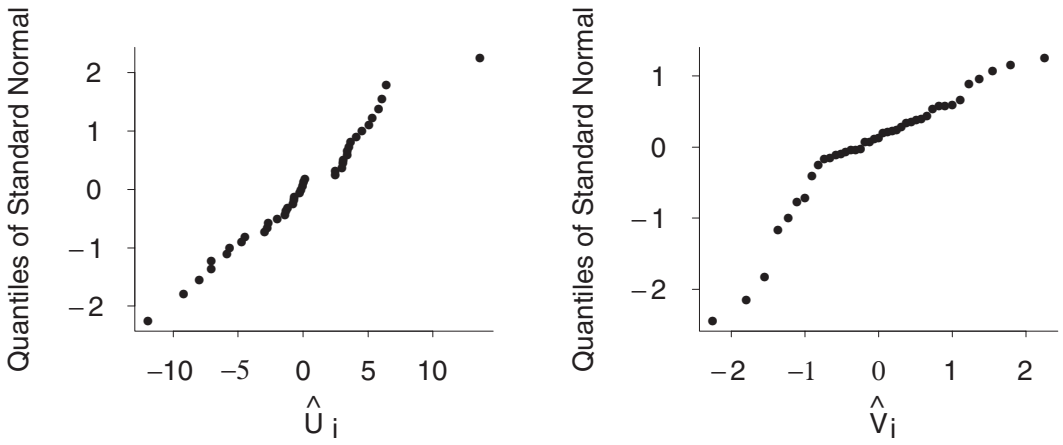

Fig. 3. Normal probability plots of predicted values $\widehat{U}_i$ and $\widehat{V}_i$.

several advantages over smoothing splines. In particular, penalized splines yield (1) a simpler mixed model formulation, (2) low-rank smoothers that allow one to fit the models to much larger problems, and (3) direct additive model fitting without the backfitting algorithm. One might question the subjectivity involved in the choice of the number and placement of knots. However, it has been our experience that penalized regression spline fitting is quite insensitive to the choice of knots. For instance, in the current respiratory health application, refitting the models after both decreasing and increasing the number of knots by 50% resulted in virtually no change in any of the model statistics in Tables 1 and 2. In the context of the Utah Valley study, this penalized spline approach yields moderate evidence of a nonlinear effect of low temperature on pulmonary function.

As noted in the Introduction, the National Research Council has made identifying which persons are

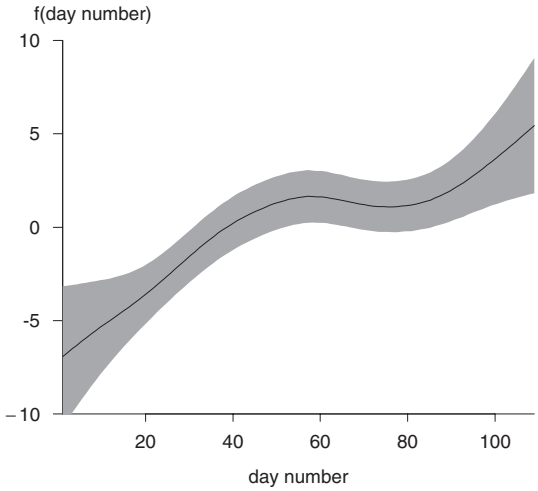

Fig. 4. Fitted day number effect for model (3.6) with independent errors.

particularly susceptible to particulate air pollution a high priority. The analyses presented in this article directly contribute toward this goal by indicating that there is in fact population heterogeneity in response to PM_{10}. In particular, these results show that there are three subjects in the Utah Valley study who exhibited a particularly acute response to PM_{10}. This finding of heterogeneity in pulmonary responses to particulate air pollution is new. A previous investigation of this question (Brunekreef *et al.*, 1991) reported heterogeneity in responses to ozone, but not to particles. However, this result was based on analyses that involved only a few measurements of lung function and particulate air pollution per subject. In addition, these analyses assessed heterogeneity using the approach of Korn and Whittemore (1979), whereby one fits fixed, subject-specific slopes and then examines the resulting estimates. This approach is less efficient than the mixed model approach used here.

This finding of response heterogeneity suggests that further research seeking to identify the cause of this susceptibility may be worthwhile. Such research could provide valuable insight into the mechanisms underlying the effects of particles. This determination of mechanism could in turn help explain additional sources of response heterogeneity, since impaired pulmonary function may be life-threatening in persons with certain diseases (e.g. Sunyer *et al.*, 2000). In this vein, one can use results from the additive mixed model fits to investigate what subject characteristics might explain response heterogeneity. Specifically, one can model the empirical Bayes estimates of subject-specific PM_{10} slopes as a function of other demographic data (Waternaux *et al.*, 1989). Analysing the Utah Valley PM_{10} slopes in this way suggests that the limited demographic information available in the Utah Valley study (i.e. age, gender, medication) does not explain the heterogeneity in PM_{10} sensitivity among the schoolchildren population, but illustrates the approach that would be available in a richer data set.

We also note that the Utah Valley study recorded only outdoor measurements of PM_{10}, and not personal exposures. Mage *et al.* (1999) showed that because such measures are correlated only with personal exposure to PM_{10} of *outdoor* origin, one must interpret such outdoor measurements as surrogates for only the ambient portion of personal PM_{10} exposure. Even with this careful interpretation, the use of such a surrogate for the true personal exposure to ambient particles results in a type of measurement error. When interest lies in the population averaged effect of air pollution, this measurement error is well known to bias the estimate of the overall effect toward the null (Carroll *et al.*, 1995). Based on recent studies (Mage *et al.*, 1999; Zeger *et al.*, 2000), we expect this bias to be of the order of 25%.

Because we are also interested in the random slopes for $\mathtt{PM}_{10}$, the situation is more complicated in that some of the subject heterogeneity in response to $\mathtt{PM}_{10}$ may arise from variation in subject-specific slopes between personal exposure to ambient $\mathtt{PM}_{10}$ and the outdoor concentrations. However, three factors in the Utah Valley study serve to reduce, but not eliminate, this variation in slopes between personal and outdoor concentrations. First, since subjects in the study are students at the same school, they share the same indoor environment for much of the day. Second, the schoolchildren spend more time outdoors than adults, and this environment is also homogeneous for all children in the study. Finally, Sarnat *et al.* (2000) showed that the principal determinant of this variation in slopes between outdoor and personal concentrations is the degree of ventilation in the indoor environment, and that such ventilation is heavily influenced by the amount of time the windows are open. Since the study was conducted in a mountain community in the winter, windows were likely to be closed at all times in most if not all of the children's homes.

ACKNOWLEDGEMENTS

We gratefully acknowledge the generosity of Professors Arden Pope and Douglas Dockery for allowing us to use the Utah Valley data, and thank the associate editor, a referee Alexandros Gryporis, Ken Kleinman and Louise Ryan for helpful comments. This research was partially supported by NIH grants ES05860, ES00002, and 1P01 ES 09825-01, and EPA grant R827353.

REFERENCES

ALTMAN, N. S. AND CASELLA, G. (1995). Nonparametric empirical Bayes growth curve analysis. *Journal of the American Statistical Association* **90**, 508–515.

ANDERSON, S. J. AND JONES, R. H. (1995). Smoothing splines for longitudinal data. *Statistics in Medicine* **14**, 1235–1248.

BETENSKY, R. A. (1997). Local estimation of smooth curves for longitudinal data. *Statistics in Medicine* **16**, 2429–2445.

BOWMAN, A. AND AZZALINI, A. (1997). *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. Oxford: Oxford University Press.

BRUMBACK, B. A., RUPPERT, D. AND WAND, M. P. (1999). Comment to 'variable selection and function estimation in additive nonparametric regression using a data-based prior'. *Journal of the American Statistical Association* **94**, 794–797.

BRUMBACK, B. A. AND RICE, J. A. (1998). Smoothing spline models for the analysis of nested and crossed samples of curves. *Journal of the American Statistical Association* **93**, 961–976.

BRUNEKREEF, B., KINNEY, P. L., WARE, J. H., DOCKERY, D. W., SPEIZER, F. E., SPENGLER, J. D. AND FERRIS, B. G. (1991). Sensitive subgroups and normal variation in pulmonary function response to air pollution episodes. *Environmental Health Perspectives* **90**, 189–193.

CARROLL, R. J., RUPPERT, D. AND STEFANSKI, L. A. (1995). *Measurement Error in Nonlinear Models*. London: Chapman & Hall.

DIGGLE, P. J., LIANG, K.-Y. AND ZEGER, S. L. (1994). *Analysis of Longitudinal Data*. Oxford: Clarendon Press.

EFRON, B. AND TIBSHIRANI, R. J. (1993). *Introduction to the Bootstrap*. London: Chapman & Hall.

EILERS, P. H. C. AND MARX, B. D. (1996). Flexible smoothing with B-splines and penalties (with discussion). *Statistical Science* **89**, 89–121.

HASTIE, T. J. AND TIBSHIRANI, R. J. (1990). *Generalized Additive Models*. London: Chapman & Hall.

KORN, E. L. AND WHITTEMORE, A. S. (1979). Methods of analyzing panel studies for discrete and continuous outcomes. *Biometrics* **35**, 795–803.

LAIRD, N. M. AND WARE, J. H. (1982). Random-effects models for longitudinal data. *Biometrics* **38**, 963–974.

LIN, X. AND ZHANG, D. (1999). Inference in generalized additive mixed models using smoothing splines. *Journal of the Royal Statistical Society,* Series B **61**, 381–400.

LITTELL, R. C., MILLIKEN, G. A., STROUP, W. W. AND WOLFINGER, R. D. (1996). *SAS System for Mixed Models*. Cary, NC: SAS Institute, Inc.

MAGE, D., WILSON, W., HASSELBLAD, V. AND GRANT, L. (1999). Assessment of human exposure to ambient particulate matter. *Journal of the Air & Waste Management Association* **49**, 1280–1291.

NATIONAL RESEARCH COUNCIL, (1998). *Research Priorities for Airborne Particulate Matter*. Washington, DC: National Academy Press.

NEAS, L. M., DOCKERY, D. W., KOUTRAKIS, P., TOLLERUD, D. J. AND SPEIZER, F. E. (1995). The association of ambient air pollution with twice daily peak expiratory flow rate measurements in children. *American Journal of Epidemiology* **141**, 111–122.

PARISE, H., RUPPERT, D., RYAN, L. M. AND WAND, M. P. (2000). Incorporation of historical controls using semiparametric mixed models. *Journal of the Royal Statistical Society,* Series C, to appear.

PINHEIRO, J. C. AND BATES, D. M. (2000). *Mixed Effects Models in S and S-plus*. New York: Springer.

POPE, C. A., DOCKERY, D. W., SPENGLER, J. D. AND RAIZENNE, M. E. (1991). Respiratory health and PM_{10} pollution: a daily time series analysis. *American Review of Respiratory Disease* **144**, 668–674.

ROBINSON, G. K. (1991). That BLUP is a good thing; the estimation of random effects (with discussion). *Statistical Science* **6**, 15–51.

RUPPERT, D. AND CARROLL, R. J. (2000). Spatially-adaptive penalties for spline fitting. *Australian and New Zealand Journal of Statistics* **42**, 205–224.

SARNAT, J. A., KOUTRAKIS, P. AND SUH, H. (2000). Assessing the relationship between personal particulate and gaseous exposures of senior citizens living in Baltimore, MD. *Journal of the Air & Waste Management Association* **50**, 1184–1198.

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

SCHWARTZ, J. (1994a). Nonparametric smoothing in the analysis of air pollution and respiratory disease. *Canadian Journal of Statistics* **22**, 471–487.

SCHWARTZ, J. (1994b). The use of generalized additive models in epidemiology. *In Proceedings of the International Biometrics Society Biannual Meeting, Hamilton, Canada*. pp. 55–60.

SCHWARTZ, J., KATZ, S., FEGLEY, R. AND TOCKMAN, M. (1988). Analysis of spirometric data from a national sample of healthy 6–24 year olds. *American Review of Respiratory Disease* **138**, 1405–1414.

SCHWARTZ, J., WYPIJ, D., DOCKERY, D., WARE, J., ZEGER, S., SPENGLER, J. AND FERRIS, B. (1990). Daily diaries of respiratory symptoms and air pollution: methodological issues and results. *Environmental Health Perspectives* **90**, 181–187.

SEARLE, S. R., CASELLA, G. AND MCCULLOCH, C. E. (1992). *Variance Components*. New York: Wiley.

SELF, S. G. AND LIANG, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* **82**, 605–610.

STANISWALIS, J. G. AND LEE, J. J. (1998). Nonparametric regression analysis of longitudinal data. *Journal of the American Statistical Association* **93**, 1403–1417.

SUNYER, J., SCHWARTZ, J., TOBIAS, A., MACFARLANE, D., GARCIA, J. AND ANTO, J. M. (2000). Patients with chronic obstructive pulmonary disease are at increased risk of death associated with urban particle air pollution: a case-crossover analysis. *American Journal of Epidemiology* **151**, 50–56.

VERBYLA, A. P., CULLIS, B. R., KENWARD, M. G. AND WELHAM, S. J. (1999). The analysis of designed experiments and longitudinal data using smoothing splines. *Journal of the Royal Statistical Society,* Series C **48**, 269–311.

WANG, Y. AND TAYLOR, J. M. G. (1995). Inference for smooth curves in longitudinal data with application to an AIDS clinical trial. *Statistics in Medicine* **14**, 1205–1218.

WANG, Y. (1998). Smoothing spline models with correlated random errors. *Journal of the American Statistical Association* **93**, 341–348.

WATERNAUX, C., LAIRD, N. M. AND WARE, J. H. (1989). Methods for analysis of longitudinal data: blood-lead concentrations and cognitive development. *Journal of the American Statistical Association* **84**, 33–41.

ZEGER, S. L. AND DIGGLE, P. J. (1994). Semiparametric models for longitudinal data with application to CD4 cell numbers in HIV seroconverters. *Biometrics* **50**, 689–699.

ZEGER, S. L., THOMAS, D., DOMINICI, F., SAMET, J., SCHWARTZ, J., DOCKERY, D. AND COHEN, A. (2000). Exposure assessement error in time series studies of air pollution: concepts and consequences. *Environmental Health Perspectives* **108**, 419–426.

ZHANG, Z., LIN, X., RAZ, J. AND SOWERS, M. (1998). Semiparametric stochastic mixed models for longitudinal data. *Journal of the American Statistical Association* **93**, 710–719.

[*Received May 18, 2000; revised November 3, 2000; accepted for publication December 4, 2000*]