THRESHOLD DEPENDENCE OF MORTALITY EFFECTS FOR FINE AND COARSE PARTICLES IN PHOENIX, ARIZONA

Richard L. Smith, Yuntae Kim, Montserrat Fuentes and Dan Spitzner

University of North Carolina and North Carolina State University

April 6 2000

Abstract

Daily data for fine ($< 2.5\mu g/m^3$) and coarse (2.5–$10\mu g/m^3$) particles are available for 1995–7 from the EPA research monitor in Phoenix. Mortality effects on the 65 and over population are studied for both Phoenix city and for a region of about 50 miles around Phoenix. Coarse particles in Phoenix are believed to be natural in origin and spatially homogeneous, whereas fine particles are primarily vehicular in origin and concentrated in the city itself. For this reason, it is natural to focus on the city mortality data when considering fine particles, and on the region mortality data when considering coarse particles, and most of the results reported here correspond to those assignments. After allowing for seasonality and long-term trend through a nonlinear (B-spline) trend curve, and also for meteorological effects based on temperature and specific humidity, a regression was performed of mortality on PM, using several different measures for PM. Based on a linear PM effect, we find a statistically significant coefficient for coarse particles, but not for fine particles, contrary to what is widely believed about the effects of coarse and fine particles. An analysis of nonlinear pollution-mortality relationships, however, suggests that the true picture is more complicated than that. For coarse particles, the evidence for any nonlinear or threshold-based effect is slight. For fine particles, we find evidence of a threshold, the most likely values of this being within the range 20–25 $\mu g/m^3$. We also find some evidence of interactions of the PM effects with season and year. The main effect here is an apparent seasonal interaction in the coarse PM effect. An attempt was made to explain this in terms of seasonal variation in the chemical composition of PM, but this led to another counterintuitive result: the PM effect is highest in spring and summer, when the anthropogenic concentration of coarse PM is lowest as determined by a principal components

analysis. There is no evidence of confounding between the fine and coarse PM effects. Although these results are all based on one city and should be considered tentative until replicated in other studies, they do suggest that the prevailing focus on fine rather than coarse particles may be an oversimplification. The study also shows that consideration of nonlinear effects can lead to real changes of interpretation, and raises the possibility of seasonal effects associated with the chemical composition of PM.

Implications

The EPA standard for fine particles introduced in 1997 was based on the widespread belief that the most serious health effects occur for fine rather than coarse particles. The present study shows that coarse particles may still have an effect, and therefore should not be neglected. It also confirms that fine particles have an effect, but in this analysis, it is only observed above a threshold in the region of 20–25 $\mu g/m^3$; there appears to be no effect below 15 $\mu g/m^3$. Since the latter figure is the 1997 EPA standard for long-term average fine PM, this suggests that the standard may possily be more stringent than needed. However the main message of the paper is that more study is needed, both of the comparative effects of coarse and fine PM, of possible threshold or nonlinear relationships, and of the effect of variations in the chemical composition of PM.

About the authors

Yuntae Kim is a graduate student and Montserrat Fuentes is an assistant professor, Department of Statistics, North Carolina State University, Raleigh, NC. Dan Spitzner is a graduate student and Richard L. Smith is a professor, Department of Statistics, University of North Carolina, Chapel Hill, NC 27599-3260. Correspondence about this paper should be addressed to Richard Smith at the above address, or by email at rls@email.unc.edu.

Acknowledgements

This work was supported in part by EPA Cooperative Agreement CR 825173-01-1 to the University of Washington, as a subcontract to the National Institute of Statistical Sciences (NISS), Research Triangle Park, NC, and by EPA Cooperative Agreement CR 827737-01-0 to Smith. The authors would like to thank Merlise Clyde, Peter Guttorp, Garry Norris and Larry Cox for assistance with the data and for advice about the analysis.

Disclaimer

This paper has not been subjected to the Environmental Protection Agency's internal peer review system and no endorsement by the Agency should be implied or inferred.

BACKGROUND AND DATA

In 1997, the Environmental Protection Agency (EPA) introduced a new particulate matter standard based on $PM_{2.5}$ (particulate matter of aerodynamic diameter 2.5 microns or less) to supplement an earlier standard based on PM_{10} (particulate matter of aerodynamic diameter 10 microns or less). The new standard was founded on the widely held belief that $PM_{2.5}$ is more directly injurious to human health than PM_{10}. Nevertheless, although there has been widespread research on the human health effects of PM_{10} based on time series analysis of daily mortality and morbidity counts (see e.g. Pope *et al.* (1995), Samet *et al.* (1995, 1997), Smith *et al.* (1998, 1999)), there has been comparatively little direct comparison of the epidemiological effects of fine PM (i.e. $PM_{2.5}$) and coarse PM (PM_{10}–$PM_{2.5}$). Schwartz *et al.* (1996) compared the effects of fine PM and coarse PM on mortality using data from the Harvard "Six Cities Study", which involved an average of eight years' data at each of the six cities, using dichotomous sampling data collected every other day. In their study, they reported a consistently stronger effect for fine particles than for coarse particles. They also considered the possibility of the "threshold effect" for fine particles by repeating the analysis restricted to days on which the level of $PM_{2.5}$ was below either 25 or 30 μg/m^3, reporting that even within those days there was still a statistically significant association between $PM_{2.5}$ and mortality. However, Lipfert and Wyzga (1997) criticized this study, arguing that the difference in results for coarse and fine PM could have resulted from differential measurement errors in the two series. Another study by Schwartz *et al.* (1999) reported that coarse particles did not have an adverse health effect in Spokane, WA, a fact which could very likely be explained by the fact that coarse particles in Spokane are mostly natural dust and therefore far less toxic than particles of industrial origin. Despite these studies, overall there has been much less direct epidemiological comparison of fine and coarse PM than there has been of cases where the two are combined into PM_{10} or TSP (total suspended particulates). Of course, the main reason for this is the lack of daily $PM_{2.5}$ data at all but a very small number of stations.

The present study makes use of a new data source, from Phoenix, Arizona. From 1995 to early 1998, the EPA located a research monitoring platform in Phoenix, collecting daily data from DFPSS, TEOM and dichotomous samplers, to determine both fine and coarse

PM measurement, as well as particulate carbon and elemental concentration measurement. These data have been described by the PM Research Monitoring Network Data report for Phoenix, Arizona, February 1995 – December 1997, produced by the US EPA National Exposure Research Laboratory, Research Triangle Park, NC. Based on these data, we have calculated daily fine and coarse PM data by averaging hourly measurements from a Tapered Element Oscillating Microbalance (TEOM) monitor. Although the data contained a small number of negative values, these were not removed as it seems likely that they result from the method used to calculate hourly concentrations rather than simple recording error. Because it is common in this field of research to use averages of up to five days' PM data as an exposure measure, rather than just single daily readings, we also calculated the k-day running averages, for $k = 2, 3, 4, 5$, of coarse and fine PM. In performing this calculation, we followed the convention that if at least one but less than k of the daily readings available, we calculate a k-day average using all available days. For most of the analysis to follow we use $k = 3$, and with this convention, we have 1038 available days of fine PM data and 1026 days of coarse PM data.

Climatic data for Phoenix have been downloaded from the web site of the National Climatic Data Center in Asheville, NC — although climatic data are also directly available from the EPA report, we preferred to use the NCDC data because of the wider range of variables available. Specifically, we made use of the following data available on a daily basis: daily maximum temperature, daily minimum temperature, and specific humidity, the latter calculated from dewpoint and pressure. Daily deaths data were obtained from the Arizona Health Services Department. Based on these, we developed two series of daily deaths, one based on the city of Phoenix, and the other from a wider region of around 50 miles around Phoenix, which includes other cities such as Scottsdale, Mesa and Tempe. We refer to this as the "Phoenix region" data set. Both data sets were restricted to residents, which avoids a possible bias due to seasonal influx of temporary residents during the winter.

Some discussion needs to be given of the reasons for considering separate city and regional data. There were good *a priori* reasons, which the detailed analysis confirms, for expecting fine particles effects to be strongest in the city data, and coarse particles effects in the regional data. Fine particles in Phoenix are primarily vehicular in origin, spatially heterogeneous, and concentrated in urban areas. We would not expect effects due to downtown traffic to affect people living many miles from the city, or in other cities where the PM levels are different from Phoenix. In contrast, coarse PM in this region is believed to be primarily of natural origin, and spatially homogeneous. We therefore expect the effects of coarse PM to be homogeneous over the region, and from a statistical point of view, we can expect to get more precise estimates if we use a larger data set.

INITIAL DATA ANALYSIS

Previous time series studies of air pollution and mortality have made clear that there are both meteorological and long-term trend and seasonality effects which must be taken into account. Naturally, we find the same for the current data set.

Fig. 1 shows daily deaths for both the region and the city, with a smoothed scatterplot smoother running through the data points. The latter was obtained using the `lowess` function in the statistical package SPlus, with $f = 0.05$ (this is a parameter controlling the amount of smoothing). This plot shows a strong seasonal pattern and possibly some additional long-term trend.

Fig. 2 shows levels of three-day averages of PM coarse and PM fine, also with a scatterplot smoother. These are shown in preference to one-day values because the three-day values are the ones used in the more detailed analysis later in the paper. A strong seasonal effect is clear here as well, and possibly an overall increasing trend.

Initial studies of meteorological effects show that both temperature and specific humidity are relevant, but the effects may be nonlinear. With temperatures, some previous studies (e.g. Smith *et al.* 1998, 1999) have suggested that a piecewise linear effect, with different slopes on either side of some threshold value, may fit the data better than a polynomial trend, and initial exploratory regression analyses suggested that this might be true here. Specifically, a model where the dependence on daily maximum temperature *tmax* is of the form

$$\begin{cases} b_1 \; tmax, & tmax < u, \\ b_1 \; tmax + b_2 \; (tmax - u), & tmax > u, \end{cases} \tag{1}$$

with $b_1 < 0$, $b_1 + b_2 > 0$ and some threshold u, appears to be a good fit. Exploration of

$u = 20,\ 25,\ 30$ and 35^{o}C led to $u = 30$ as the threshold chosen for subsequent analysis. Other variables included were daily minimum temperature *tmin*, specific humidity *sh*, and the square of specific humidity, which we denote *shsq*. Including both these and the PM-based variables, a complete list of covariates used in the analysis is contained in Table 1.

day	Number of day (1=Feb 1 1995; 1065=Dec 31 1997)
mortc	Elderly nonaccidental mortality in city
mortr	Elderly nonaccidental mortality in region
tmax	Daily maximum temperature
tmin	Daily minimum temperature
sh	Daily mean specific humidity
$tg30$	Larger of $tmax - 30$ and 0
shsq	Square of *sh*
$p1c$	Daily coarse particles level (PM_{10}−$PM_{2.5}$)
$p2c$	Two-day averages of p1c
$p3c$	Three-day averages of p1c
$p4c$	Four-day averages of p1c
$p5c$	Five-day averages of p1c
$p1f$	Daily fine particles level ($PM_{2.5}$)
$p2f$	Two-day averages of p1f
$p3f$	Three-day averages of p1f
$p4f$	Four-day averages of p1f
$p5f$	Five-day averages of p1f

Table 1. Description of variables used in analysis.

In addition, all the variables in Table 1, except *day* and the mortality variables, are also considered in lagged form: for instance $tmax_m$ means the value of *tmax* lagged m days. We consider $m = 0, 1, 2, 3, 4$; $tmax_0$ is today's maximum temperature, $tmax_1$ is yesterday's, and so on.

For the seasonal and long-term trend, it is obvious from Fig. 1 that a simple polynomial or piece-wise linear function will not be adequate, and the solution adopted here is a B-spline representation, which represents the estimated function as a continuous sequence of cubic polynomials of the form

$$f(day) = \sum_{k=1}^{K} c_k \ B\left(\frac{K \times day}{1065} - k + \frac{1}{2}\right), \tag{2}$$

where $B(\cdot)$ is the B-spline basis function, see e.g. Smith *et al.* (1998), p. 97, or Green and Silverman (1994), pp. 157–8. In effect, what this does is to represent the trend as a linear combination of K independent functions, with coefficients c_k estimated from the data, and whose smoothness may be controlled by varying the value of K. Note that (2) effectively ensures that the "knots" of the B-spline representation, given the value of K, are uniformly distributed throughout the 1065 days for which data are available.

The final "initial data analysis" issue is the form of dependence of mortality on the regression terms. In the present paper, this has been achieved by a linear regression of square root of daily mortality on the long-term trend, meteorological and PM-based variables. The choice of a square root transformation was made after a comparison with a logarithmic transformation and with no transformation, using methods similar to Atkinson (1985), section 6.2. The square root transformation was clearly superior to the other two transformations in every comparison made.

To summarize the results of this section, the model adopted is a linear regression in which the dependent variable is square root of daily mortality in either Phoenix city or Phoenix region, and the linear regression terms are long-term trends modeled by (2), together with a subset of the meteorological and PM-based variables in Table 1, lagged from 0 to 4 days.

DETAILS OF REGRESSION ANALYSIS

For the first part of the analysis, several different values of K, ranging from 8 to 48, were tried in (2), and for each, meteorological variables from Table 1 (and their lagged values) were selected by backward selection, using hypothesis tests with size 0.1 to decide whether to retain meteorological variables (in other words, a meteorological variable was retained whenever the P-value for that variable was smaller than 0.1). The resulting models were compared by a variety of model selection devices, including *PRESS*, *AIC* and *BIC* (see, e.g. Neter *et al.* (1996), or any standard text of linear models). In general, *PRESS* and *AIC* behave similarly and tend to favor models with larger numbers of parameters, while *BIC* selects models with fewer parameters. This behavior was seen here, as the optimal value of K when selected by *PRESS* or *AIC* was 40 for the regional data and 24 for the city data; when selected by *BIC*, it was 16 and 12 respectively. Although this leaves open the question of what K we should actually use, one point in favor of smaller K was that the backward selection procedure in that case selects more meteorological variables, and that seems desirable in principle, so that the resulting model includes both meteorological terms and a long-term trend. Therefore, the *BIC* values were adopted for further analysis, with meteorological variables given in Tables 2 and 3. The subsequent results in the paper are not overly sensitive to the precise model chosen at this stage of the analysis, a point we return to later.

$tmax_1$	$tmin_0$	sh_1	$tg30_1$	$tg30_2$	$shsq_1$	$shsq_3$

Table 2. Meteorological variables used in analysis of Phoenix region data. Suffixes denote lags. The model also included a long-term trend based on (2), with $K = 24$.

$tmax_2$	$tmin_1$	$tmin_3$	sh_1	$tg30_0$	$shsq_1$

Table 3. Meteorological variables used in analysis of Phoenix city data, together with long-term trend basedon (2) with $K = 16$.

After selecting an initial model to represent the long-term trend and meteorological components, different PM variables based on Table 1, together with lagged values, were added to the model one at a time, in an attempt to ascertain what the strongest effect would be. At this point in the analysis, it emerged that taken as linear terms, those based on coarse PM contributed more statistically significant effects than those based on fine PM. For example, using the regional data, any one of $p1c_0$, $p2c_0$ or $p3c_0$ was statistically significant, with t statistics (ratio of parameter estimate to standard error) of 3.5, 3.6, 3.2 respectively. For a large data set such as this, the t distribution effectively coincides with a normal distribution, so any t value larger than 2 is statistically significant at level .05, and the values quoted here are significant at levels .001 or smaller. Results for coarse particles in Phoenix city are similar, but with larger standard errors leading to smaller t values, e.g. the t value for $p3c_0$ is 2.0. In contrast, no analysis based on fine PM, for either the city or region, produced a t statistic larger than 1.2, which is not statistically significant.

At this stage, therefore, our conclusion is that there may be a significant result due to coarse PM, but there is no sign of any due to fine PM.

NONLINEAR DEPENDENCE FOR COARSE AND FINE PM

The picture becomes considerably more complicated, however, if the possibility of a nonlinear PM response is taken into consideration. For most of the following discussion, for reasons explained earlier, we use regional data when looking at coarse PM and city data when looking at fine PM. The results for coarse PM on the city data are similar to those for the region data, but with much wider confidence bands making it harder to characterize the form of the relationship. In contrast, it was impossible to find any effect, linear or nonlinear, relating fine PM to regional mortality. As noted already, however, this is only to be expected based on what is known about the sources of PM, so it seems reasonable to concentrate on the city data when looking for fine PM effects.

At this stage of the analysis, it was decided to concentrate on $p3c_0$ and $p3f_0$ (three-day averages with no time lags; in other words, the average of today's value, yesterday's, and the day before yesterday's) as the main PM variables of interest. This decision was based both on the results of the previous section, and on general experience of this field of research, which has shown that two-day or three-day averages of PM very often give a more reliable indication of epidemiological effects than single-day values.

There are, however, numerous ways in which we can look for nonlinear effects, and in this section we consider three of them. All models fitted include the same meteorological and long-term trend terms as in the preceding section.

Piecewise Linear Analysis

The first nonlinear method tried was a piecewise-linear method, similar to the model adopted in (1) for temperature. For each of several possible thresholds, separate linear trends were estimated below and above the threshold, together with 95% confidence intervals. An early example of the idea of looking at separate linear trends below and above a threshold was the paper of Ostro (1984), in which he applied this method to data from London in the 1950s and 1960.

Fig. 3 shows the result of this analysis, in which the response to PM as a piecewise linear function is plotted, together with the confidence bands. The plots are shown for coarse PM for the region data and for fine PM for the city data. The results show a contrast between the cases of coarse and fine PM. For coarse PM (right-hand half of the plot), there is no significant change in slope either side of the threshold, except for threshold 10 which is rather meaningless in view of the shortage of coarse PM data (and very wide confidence bands on the coefficient) below this threshold. In contrast, for fine PM, although there is no significant effect when represented as a linear term, when piecewise linear terms are selected, there are some significant results. In particular, the plots for thresholds 20 and 25 show that for either of these, the PM-fine effect above the threshold is statistically significant, though not the effect below the threshold. The t statistics for the effect above the threshold were 2.4 and 2.7, respectively for thresholds 20 and 25.

B-spline analysis

The second nonlinear method tried was to represent the PM (coarse or fine) effect as a B-spline representation, similar to the formula given in (2) for the time-dependent effect. For this analysis, the number of knots was fixed at $K = 4$ which is large enough to display a nonlinear effect if there is one; for a much larger K than that, the randomness in estimating individual c_k coefficients would be so large as to render the results meaningless. The meteorological and long-term trend terms in the model were the same as in the linear analysis.

Results were expressed as relative risks (RR), using the long-term mean PM value as a reference level for which $RR = 1$. This was 13.0 in the case of fine PM, 33.6 for coarse PM. Pointwise 95% confidence bands were computed using the standard errors and covariances of parameter estimates in the regression analysis. Unfortunately, confidence bands computed by this method tend to be very wide, but they are included in the plots because they give at least a rough indication of to what extent various nonlinearities in the plot may be regarded as true effects ratherthan just artifacts of the data. Fig. 4 shows these plots, for both coarse PM (regional data) and fine PM (city data).

In plot (a), for coarse PM in the region, it can be seen that the sharpest increase in the nonlinear curve occurs between 20 and 40 $\mu g/m^3$, though there also appears to be a second sharp rise over 60 $\mu g/m^3$. Overall, however, the results of this plot do not contradict a simple overall linear effect. In plot (b) for fine PM, in contrast, there appears to be a clear change in the slope somewhere in the region of 20 $\mu g/m^3$, and the accompanying confidence bands show that this is statistically significant. This may be evidence of a threshold effect, or at least, of a significantly nonlinear relationship which contrasts sharply with our earlier finding of no relationship at all.

Another issue raised by Fig. 3 is whether, for any of the thresholds considered, they give serious evidence against a linear effect. This can be tested, with the null hypothesis of a linear PM effect, and the alternative hypothesis of a piecewise linear effect with threshold as shown. For the five plots based on coarse PM, the P values of the test statistics are all in excess of 0.5, indicating no threshold effect. For the five plots based on fine PM, the P values (top to bottom) are .06, .05, .007, .005, .33. For thresholds 20 and 25, in particular, this provides strong evidence that a piecewise linear fit improves on a simple linear fit.

Diagnostics for the B-spline analysis

A number of the standard regression diagnostics (see, e.g., Neter *et al.* 1996) were computed for the fitted models with the B-spline representation for the PM effect. These serve as a check on whether the model is a reasonable fit to the data. Such diagnostics could have been computed for all the models fitted, but we focus on this one because of all the models considered, the one involving a non-linear PM effect seems the closest to what we really want.

One issue raised by our decision to concentrate on linear regression with a square-root transformation of deaths is whether this approach copes adequately with the problems of overdispersion and serial correlation which sometimes arise in studies involving Poisson regression (Samet *et al.* 1995). Overdispersion refers to the property that variances of the observed responses are larger than those which would hold if the data were truly independent Poisson counts. In the case of a square root transformation, the variance is very nearly stabilized to a constant value 0.25. An observed variance larger than that is therefore an indication of overdispersion. Other studies typically indicate an overdispersion in the region of 1.05 to 1.1 (i.e. 5% to 10% larger than the Poisson variance).

For fine particles, the estimated residual variance was 0.2715. Dividing by 0.25, this therefore corresponds to an overdispersion of 1.09. For coarse particles, the corresponding residual variance is 0.2816, or an overdispersion of 1.13. These results are therefore at the high end of the accepted range of overdispersions, which may indicate some additional source of variability which has not been taken into account.

Serial correlations have been calculated based on the studentized residuals. For fine particles, the first three values are .034, .035, .103. To judge the significance of these, a common rule of thumb is to compare them with $2/\sqrt{N}$, where N is the sample size on which the serial correlations are based. In this case, $2/sqrtN \approx .062$, which means that the third-order autocorrelation is significant. None of the other autocorrelations is significant, however. The results for the model based on coarse particles fitted to the region data are similar: serial correlations .055, .037, .084,... so that the third value is again significant, but none of the others. We do not have a ready explanation for this.

There are also a number of diagnostics aimed at determining whether any of the observations are particularly influential on the final results. There are several of these which tend to work in a similar way, so we concentrate on one, namely DFFITS (Neter *et al.* 1999, chapter 9, or Atkinson 1985, chapter 3). According to criteria originally given by Belsley, Kuh and Welsch (1980), DFFITS indicates an influential observation at a value $2\sqrt{p/N}$, where N is the sample size and p is the number of parameters in the model. In the case of the fine particles analysis, this is 0.295, and there are no fewer than 65 values (out of the 1014 for which DFFITS could be calculated) which exceed that in absolute value, the largest in absolute value being $-.6196$. This is a little hard to interpret, but we have examined whether outlying values of DFFITS correspond to outlying observations of PM, but no clear pattern emerges. The picture for coarse particles is similar, with 55 values of DFFITS exceeding the cutoff value of .327, the largest in absolute value being

$-.5853$, but extreme values of DFFITS again do not correspond to extreme value of the particulate matter variable of interest.

Overall, there are a number of features here which might justify further exploration, but none which casts serious doubt on the correctness of the model.

Bayesian analysis for threshold selection

One can take the analysis of the previous section somewhat further by looking specifically for a threshold effect of the form

$$\begin{cases} 0 & \text{if } p \le u \\ b_1(p-u) & \text{if } p \ge u \end{cases} \tag{3}$$

where p is a PM-variable (coarse or fine) and u is the threshold. The purpose of this section is to see how far we can go towards formally selecting the best value of u to be consistent with formula (3).

Conditionally on u, the dependence between the vector of responses Y and the matrix of covariates X (which includes long-term trend, meteorology and PM variables) is a linear model of the form

$$\mathrm{E}\{Y\} = X^{(u)}b^{(u)}, \qquad \mathrm{Cov}\{Y\} = v^{(u)}I, \tag{4}$$

in which the matrix of covariates $X^{(u)}$, the regression parameters $b^{(u)}$ and the residual variance $v^{(u)} > 0$, all depend on the threshold parameter u. If we take a Bayesian point of view, assuming a joint prior density for $(u, b^{(u)}, v^{(u)})$ of the form

$$\pi(u, b^{(u)}, v^{(u)}) \propto \frac{1}{v^{(u)}}, \qquad 0 \le u \le u_{max},\ v^{(u)} > 0, \tag{5}$$

for some upper bound u_{max} on the permissible values of u, then by combining (4) and (5) and integrating out $b^{(u)}$ and $v^{(u)}$, the marginal density of Y given u is of the form

$$f(Y \mid u) \propto G(u)^{n-q}, \quad 0 \le u \le u_{max}. \tag{6}$$

Here, n is the number of observations, q is the number of regressors in the linear model (4), and $G^2(u)$ is the conventional error sum of squares for the linear regression model (4) with u treated as fixed. Bayesian inference for u may therefore be based directly on the conditional density (6), renormalizing the probabilities so that the posterior density of u integrates to 1.

In practice, we have assumed u to lie on a discrete grid (10, 11, 12,...,35 for PM-fine, 0, 2, 4, ..., 70 for PM-coarse) and have computed posterior densities by summing the values of (6) over this grid, renormalizing so that the overall sum of probabilities is 1. The results are shown in Fig. 5. These results may be interpreted as an overall probabilistic statement about the location of the threshold, based on the data available.

We have already seen in Fig. 4 that the strongest evidence for a threshold is in plot (b), for fine PM in Phoenix city, with less strong evidence in plot (a) (coarse PM in the region). Results in Fig. 5 confirm this, but also give new insights into the strength of evidence for the existence of a threshold.

For coarse PM, plot (a) shows a peak in the posterior density around 20 μg/m^3, but it is not a very strong peak, and the posterior density does not tend to 0 near $u = 0$, which suggests that there may in fact be no threshold at all.

For fine PM, plot (b) shows a very clear peak in the posterior density near $u = 22$ μg/m^3, with the posterior density near 0 outside the range 15–30. Although the results have been calculated on the assumption of a uniform prior distribution for u, the general form of this plot (with a much higher posterior density in the range 15–30 than outside that range) will not be very sensitive to this, provided a prior density is adopted which is consistent with reasonable prior belief over a wide range of values of u. Thus in this case, we deduce strong evidence in favor of the existence of a threshold.

SENSITIVITY ANALYSES

As a check on the sensitivity of the main results in THE paper to some of the modeling assumptions made at the beginning, they were repeated with the following changes: (i) the choice of K (number of knots in the B-spline representation) was made by AIC rather than BIC; (ii) the size of the hypothesis tests performed at the backward selection stage was increased from .10 to .15 (the effect of this will be to include more meteorological variables in the analysis); (iii) the meteorological modeling was confined to temperature-based variables (no humidity), i.e. $tmax$, $tmin$ and $tg30$ (see Table 1), together with their lagged values.

We shall not present very detailed results of this, but following are the main conclusions. An AIC-based selection of K led to $K = 40$ for the region data and $K = 20$ for the city data. With these changes to the model, both the coarse and fine PM effects are a little weaker than those in the preceding analysis, but the qualitative results are the same — there is a significant linear effect for coarse PM in the region and a significant nonlinear effect for fine PM in the city.

We should note, however, that the trend and seasonal variation must not be over-modeled. During the AIC analysis for the city data, it was noted that $K = 48$ gives an AIC value not very different from the optimal value $K = 20$, but when the fine particles analysis was repeated with this value of K, the results were entirely different, several of the linear coefficients appearing significantly negative. The interpretation of this result would appear to be that local fluctuations of the order of three weeks (the interval between knots in this analysis) are short enough to be confounded with the fine particles effect, leading to incorrect estimates for the latter. This serves as a warning against indiscriminate reliance on AIC or indeed any "black box" statistical criterion.

INTERACTIONS

Another question we have considered is the possibility of an interaction between the PM effect and either season or year. If there are different effects in different season or years, this could be an indication that the true relationship is more complicated than simple cause and effect.

As an example, a seasonal interaction model for coarse PM was defined as follow. Instead of a single regressor for coarse PM, four variables were defined, one for each season. For example, "winter coarse PM" is the coarse PM value during the winter months (December, January, February), 0 the rest of the year. Spring, summer and fall coarse PM values were defined similarly. The main regression analysis of the paper, for the region data, was rerun, producing the results in Table 4. Also shown for comparison are the mean levels of coarse PM for each season.

Season	Mean Coarse PM	Estimate	S.E.	t statistic	p-value
Winter	33.6	0.0036	0.0023	1.5	0.13
Spring	28.9	0.0139	0.0026	5.3	0.0001
Summer	31.6	0.0063	0.0026	2.4	0.018
Fall	39.3	0.0023	0.0022	1.0	0.3

Table 4. Interactions of coarse PM and season.

It can be seen from Table 4 that the coarse PM effect is really only significant during the spring and summer months. It is possible that this could be an artifact of what is really a nonlinear relationship between PM and mortality, but we doubt this because (a) our previous studies showed no sign of this, and (b) if it were due to a nonlinear relationship it appears to be the wrong way round — the high PM coefficient is during seasons when the mean coarse PM level is low. Below, we shall offer another explanation for the seasona variation in the PM coefficient.

The possibility of an interaction by year was suggested by Fig. 2, in which it can be seen that both coarse and fine PM were unusually low during the first half of 1995. To test whether this had any influence on the results, four "years" were defined corresponding to Feb–June 1995 (year 1), July 1995–June 1996 (year 2), July 1996–June 1997 (year 3) and july–December 1997 (year 4). A year effect was estimated using exactly analogous methodology to that just described for the season effect. In the case of coarse PM, it indeed turns out that the effect is lowest in year 1, but when the overall significance of the coarse PM $\times$ year interaction was tested using an F test, it was not significant (p-value .07). In contrast, the f test for a seasonal interaction was significant with p-value .004.

In the case of fine PM, there was again evidence of a seasonal interaction when modeled as a linear effect, but in this case, it does appear to be a proxy for the threshold dependence noted earlier in the paper. When the seasonal interaction model was fitted based on a threshold model, with separate effects below and above a threshold of 25 μg/m^3, the seasonal effect disappeared. A year interaction effect was noted even in the threshold model, with a significant negative coefficient below the threshold in year 1. In this case, an F test for the overall presence of a year interaction effect in the threshold model was significant with a p-value of .016. However, this is not at significant as the previously noted seasonal effect for coarse PM, and since it is rather hard to explain a negative dependence between fine PM and mortality, we feel this is much more likely to be an artifact of some kind.

It remains to see whether there is any natural explanation for the seasonal interaction effect that was found for coarse PM. One possibility is that this might be associated with seasonal variations in chemical composition of PM. In addition to the TEOM data used throughout this paper, we have available a breakdown of the air pollution data into 44 chemical elements (excluding carbon) that are constituents of coarse PM. We remove elements that are typically below the detection limit. This analysis is based on about 300 days' data and the elements used in the study are: Al, Si , S, Ci, K, Ca, Ti, Mn, Fe, Cu , Zn, Pb. A principal components analysis of the constituent elements of coarse PM shows that the crustal elements (Al, Si, K, Ca, Ti, Mn, Fe) explain 55% of the variation of Pm coarse, the anthropogenic elements (Fe, Cu, Zn, Pb) explain 30%, and the elements of marine origin: Cl (NaCl, Na was not measured) explain 5%. Table 5 shows a breakdown by season of the means of three principal components corresponding to each of these groups.

Season	Crustal	Anthropogenic	Marine
Winter	–.144	.503	–.589
Spring	–.278	–.323	.073
Summer	.004	–.483	.41
Fall	.245	.222	.03

Table 5. Breakdown by season of mean level of each of the three principal groups of elements (standardized to overall mean 0 for each component)

The results in Table 5 suggests that the composition of coarse PM differs throughout the year, with the crustal elements highest in spring and summer and the anthropogenic elements lower. If this were the explanation for the seasonal interaction, however, the implication would be that crustal, rather than anthropogenic, elements were responsible for the PM-mortality associations! This result seems clearly counterinituitive and suggests that we have not got to the bottom of the seasonal interactions.

A final effect that has been examined is the possibility of confounding between coarse and fine particles. In all models studied up to now, coarse and fine particles have been treated separately, putting in one or the other, but not both at the same time. If the regression coefficients were to change dramatically when both pollutants were included in the same model, that would further complicate the interpretation of the results. Fortunately, the evidence on this point is that the coefficients do not change very much. To make a specific comparison, piecewise linear effects were fitted for both fine and coarse particles (separately) based on threshold $u = 25$. They were then all put in together, to examine how the coefficients changed.

Primary pollutant	Est.	S.E.	Est.	S.E.
	with co-pollutant		without co-pollutant	
Fine particles, below threshold	–.006	.005	–.010	.005
Fine particles, above threshold	.050	.018	.042	.019
Coarse particles, below threshold	.0003	.0062	.0008	.0003
Coarse particles, above threshold	.0065	.0019	.0068	.0021

Table 6. Interactions of fine and coarse PM.

In table 6, regression parameter estimates and their standard errors are shown both for fine and coarse particles, below and above the threshold. As with earlier studies, the results where fine particles are the primary pollutant are for the city data, and the results where coarse particles are the primary pollutant are for the region data. Then, however, whichever of fine or coarse was the primary pollutant, the other was also included as a co-pollutant, and the coefficient of the primary pollutant re-estimated in this case. The results are in the last two columns of Table 4. In no case does the estimate for the primary pollutant change significantly as a result of including the co-pollutant.

The last conclusion is reassuring in that it is consistent with fine and coarse particles being essentially separate pollutants having distinct effects. Note, however, that we have not studied possible confounding of either fine or coarse particles with gaseous pollutants such as ozone or sulfur dioxide, and since past studies have suggested confounding between particulate matter and gaseous co-pollutants (see e.g. Samet *et al.* 1997), it would seem worthwhile to consider that aspect as well.

DISCUSSION

There are a number of aspects of this analysis which raise further points for discussion.

The study is based on only three years of data at a single site: many other studies are based on either a much longer series or on combining data from many sites (Dominici, Samet and Zeger 2000). This is a limitation, pointing towards the need for more studies of these issues.

Other recent studies have examined both of the primary issues in this paper, the existence of thresholds and the comparisons of fine and coarse particles. In particular, ours is not the only study to suggest that the effect of coarse particles may be equal to or greater than that of fine particles — Ostro and Lipsett (2000) have made a similar claim for California data, as have Castillejos *et al.* (2000) for data from Mexico City. There are also other recent studies on thresholds. Cakmak *et al.* (1999) have considered the possible effect of measurement error on the estimation of a threshold. Daniels *et al.* (2000) have examined the existence of a threshold in PM_{10} data across the 20 largest cities of the US, the same data base as in Dominici *et al.* (2000). Their preliminary results suggest the absence of a threshold in PM_{10} data for all-cause mortality, though there is clearly a need for more detailed research on the best way to combine data from different cities.

Lipfert and Wyzga (1997) discussed the possible role of differential measurement error in the attribution of mortality effects to a single pollutant. Specifically, they argued that the results of Schwartz *et al.* (1996), which claimed a stronger effect for fine particles, could be the result of fine particles being more accurately measured than coarse particles in the six-cities data set. We have no direct evidence on measurement error in the Phoenix data set, but we have no reason to think that it acts differentially in favor of fine particles as suggested by Lipfert and Wyzga.

The question of measurement error also arises in the difference between ambient monitor measurements and the personal exposure of individuals. There have been some studies of the effect of imputing personal exposures — for example, Dominici, Zeger and Samet (2000) have proposed a mathematical modeling approach to this, but it also appears from their paper that currently available data on personal exposure are quite limited, and do not distinguish between fine and coarse particles. Another issue related to this topic is the effect of spatial variation. For example, Lipfert and Wyzga (1997) reported on various studies in the eastern US in which fine particles were more homogeneously distributed than coarse particles. As noted at the beginning of this paper, we believe that in Phoenix, coarse particles are more homogeneously distributed than fine particles. Direct data to support this point are limited, but we do have data on fine and coarse particles from Phoenix city and from four other locations within the Phoenix region used in this paper. Measuring correlations of logarithms of PM concentration to improve numerical stability of the results, we find that the spatial correlations between the Phoenix downtown site and four other sites in the region (Higley, Tempe, ASU West and Estrella Park) are respectively .85, .88, .93, .76 for coarse PM and .64, .90, .91, .74 for fine PM. Thus in the case of Higley, the correlation of coarse PM with the Phoenix station is clearly higher than that of fine PM, while for the other three stations, the correlations are about the same for coarse and fine PM. This is, inevitably, inconclusive about whether coarse particles are indeed more homogeneously distributed than fine particles, but the results are qualitatively very different from those reported by Lipfert and Wyzga (1997) for Philadelphia, for instance.

CONCLUSIONS

The original purpose of this study was to compare the effects of coarse and fine PM on mortality in Phoenix. Knowledge of the dominant origins of PM (natural dust for coarse, vehicular emissions for fine) suggested that the effects would be primarily concentrated on Phoenix city for fine PM, but would be apparent throughout the region for coarse PM, and this was largely confirmed by the statistical analysis. Linear regressions for coarse and fine PM, taking into account meteorological and trend/seasonal effects, led us to conclude that there is a significant effect for coarse PM but not for fine PM, contrary to the prevailing orthodoxy in this field. The results were rather different, however, when nonlinear effects were taken into consideration.

Three different methods were used to study nonlinear effects: (a) a piecewise linear effect below and above a threshold, (b) a smooth nonlinear effect based on a cubic spline representation, (c) formal selection of a threshold by Bayesian means. None of the three methods led to any conclusions that contradicted a linear effect for coarse PM, but in the case of fine PM, there was clear evidence for a change of slope somewhere in the region of 20–25μg/m^3. The conclusion is that fine PM may indeed have an effect at high levels, but only above the current EPA standard for the long-term mean, of 15μg/m^3.

Additional analyses suggested there could be significant interactions in the PM effect with season and year. The strongest effect was a seasonal interaction for coarse PM, the effect being significant only in spring and summer. An attempt was made to explain this in terms of the chemical constitutents of coarse PM, and it was found that crustal elements of coarse PM were highest, and anthropogenic elements lowest, in the spring and summer. If interpreted causally, however, this result would imply that crustal and not anthropogenic sources of PM are primarily responsible for deaths, which does not seem a very plausible conclusion. A more reassuring conclusion was that there was no evidence of any confounding between fine and coarse PM.

These results, being based on a single city and for a comparatively short time period, cannot be regarded as definitive. Nevertheless, they carry clear implications, which contradict those of the (very few) previous studies of these kinds of questions, in particular, the paper of Schwartz *et al.* (1996). The story about the comparative effects of coarse and fine PM is by no means concluded, and this paper also shows that it is worthwhile to consider nonlinear or threshold-based effects, as well as the possibility of seasonal interaction.

REFERENCES

Atkinson, A.C. (1985), *Plots, Transformations and Regression.* Oxford University Press.

Belsley, D.A., Kuh, E. and Welsch, R.E. (1980) *Regression Diagnostics.* New York: Wiley.

Cakmak, S., Burnett, R.T. and Krewski, D. (1999), Methods for detecting and estimating population threshold concentrations for air pollution-related mortality with exposure measurement error. *Risk Analysis* **19**, 487–496.

Castillejos, M., Borja-Aburto, V., Dockery, D.W., Gold, D.R. and Loomis, D. (2000), Airborne coarse particles and mortality. *Inhalation Toxicology* **12** (Supplement 1), 61–72.

Daniels, M.J., Dominici, F. and Samet, J. (2000), Estimating PM10-mortality dose response curves and threshold levels: an analysis of daily time series for the 20 largest US cities. Preprint, Johns Hopkins University.

Dominici, F., Samet, J.M. and Zeger, S.L. (2000), Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. *J.R. Statist. Soc. A* **163**, to appear.

Dominici, F., Zeger, S.L. and Samet, J.M. (2000), A measurement error model for time series studies of air pollution and mortality. *Biostatistics* **1**, to appear.

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

Lipfert, F. and Wyzga, R. (1997), Daily mortality and size-fractionated particulate matter in six US metropolitan areas: the implications of measurement and modeling uncertainties. *Journal of the Air and Waste Management Association* **47**, 517–523.

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

Ostro, B. and Lipsett, M.J. (2000), Coarse and fine particles and daily mortality in the Coachella Valley, California: A follow-up study. Presented at PM2000 conference.

Neter, J., Kutner, M.H., Nachtsheim, C.J. and Wasserman, W. (1996), *Applied Linear Statistical Models*, Fourth Edition. Irwin, Chicago.

Pope, C.A., Bates, D.V. and Raizenne, M. (1995), Health effects of particulate air pollution: time for reassessment? *Environmental Health Perspectives* **103**, 472–480.

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.

Samet, J.M., Zeger, S.L., Kelsall, J.E., Xu, J. and Kalkstein, L.S. (1997), Air Pollution, Weather and Mortality in Philadelphia, 1973–1988. In *Particulate Air Pollution and Daily Mortality: Analyses of the Effects of Weather and Multiple Air Pollutants. The Phase IB Report of the Particle Epidemiology Evaluation Project.* Health Effects Institute, Cambridge MA, pp. 1–29.

Schwartz, J., Dockery, D.W. and Neas, L.M. (1996), Is daily mortality associated specifically with fine particles? *J. Air and Waste Manage. Assoc.* **46**, 927–939.

Schwartz, J., Norris, G., Larson, T., Sheppard, L., Caiborne C. and Koenig, J. (1999), Episodes of high coarse particles concentration are not associated with increased mortality. *Environ. Health Perspect.* **107**, 339-342.

Smith, R.L., Davis, J.M. and Speckman, P. (1998), Airborne particles and mortality. Chapter 6 of *Case Studies in Environmental Statistics*, edited by D. Nychka, W. Piegorsch and L.H. Cox. Springer Lecture Notes in Statistics, number 132, Springer Verlag, New York, pp. 91–120.

Smith, R.L., Davis, J.M. and Speckman, P. (1999), Human health effects of environmental pollution in the atmosphere. Chapter 6 of *Statistics in the Environment 4: Statistical Aspects of Health and the Environment*, edited by V. Barnett, A. Stein and F. Turkman. John Wiley, Chichester, 91-115.

CAPTIONS OF FIGURES

Fig. 1. Daily deaths in Phoenix city and Phoenix region for the three years of the study, with a fitted smooth curve.

Fig. 2. Three-day averages of coarse PM and fine PM (measured in μg/m^3), with a fitted smooth curve.

Fig. 3. Piecewise linear estimates of the coarse PM effect for the region and the fine PM effect for the city, plotted for different thresholds.

Fig. 4. Nonlinear estimates of relative risk (relative to the mean PM variable), together with pointwise 95% confidence bands. (a): Coarse PM effect - regional deaths data. (b): Fine PM effect - city deaths data.

Fig. 5. Posterior densities for threshold. (a): Coarse PM effect - regional deaths data. (b): Fine PM effect - city deaths data.

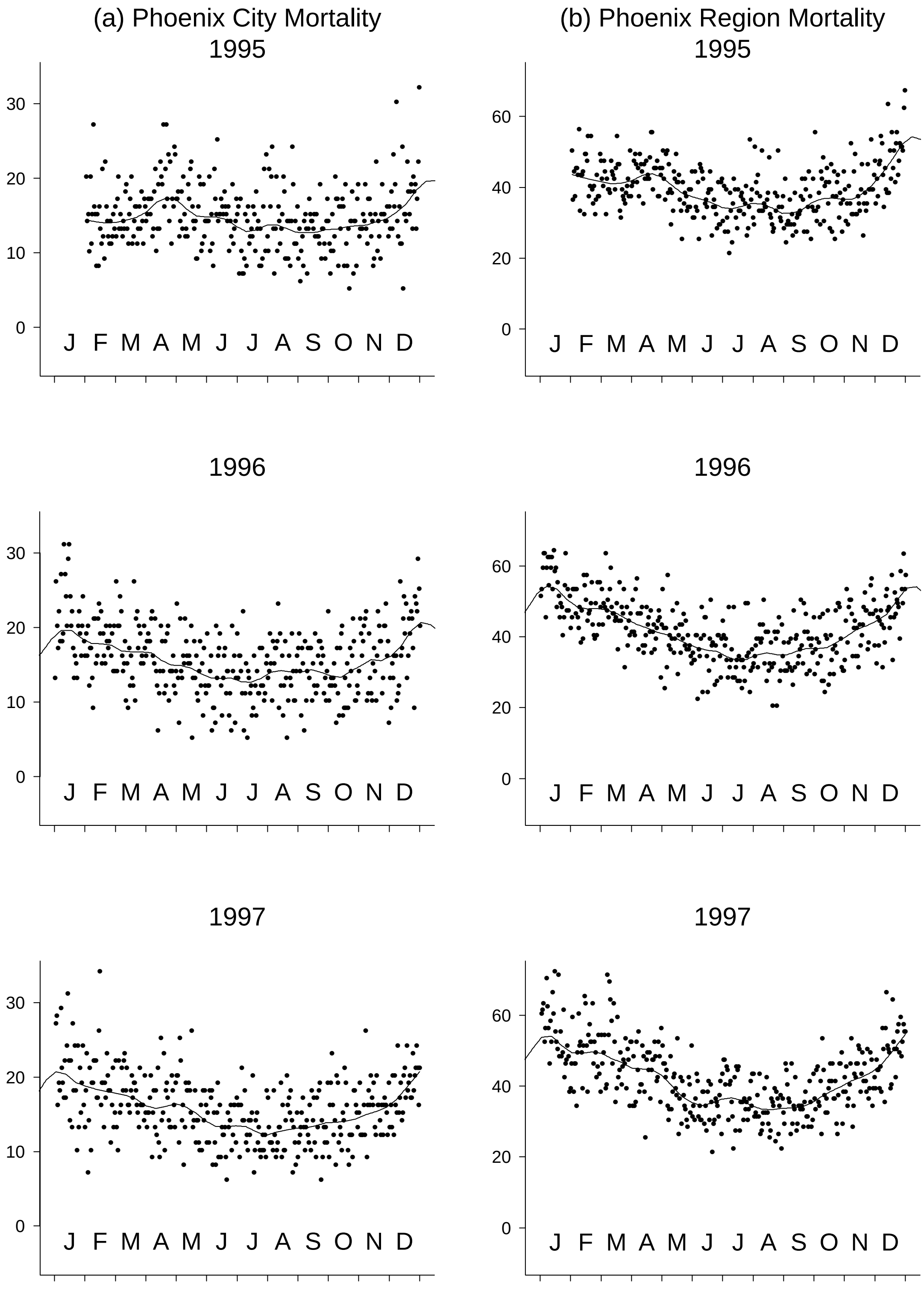
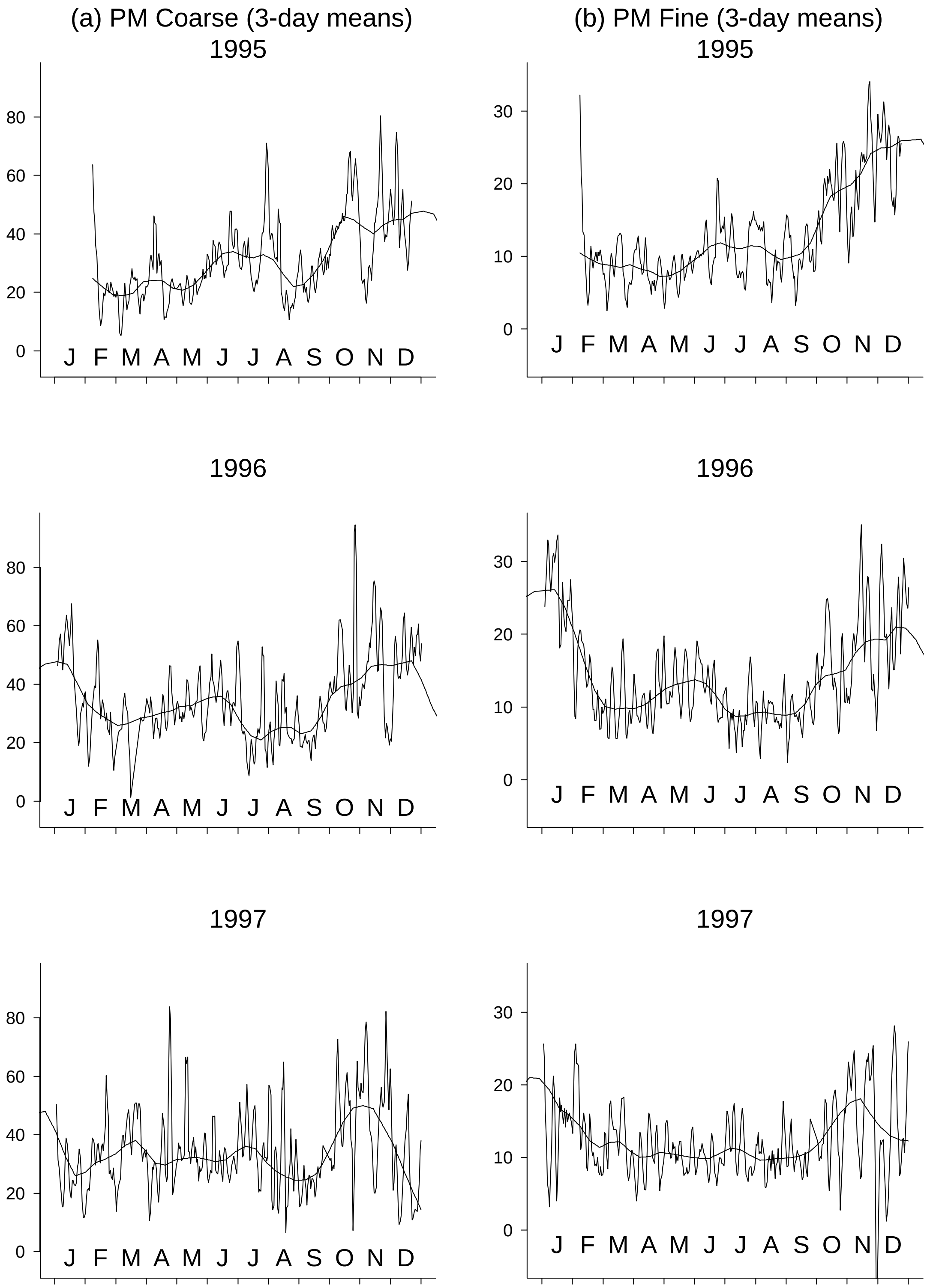

(a) Phoenix City Mortality
(b) Phoenix Region Mortality
1995
1996
1997
0
10
20
30
40
60
J F M A M J J A S O N D

Fig. 1



(a) PM Coarse (3-day means)
(b) PM Fine (3-day means)
1995
1995
1996
1996
1997
1997
0
20
40
60
80
10
30
J F M A M J J A S O N D

Fig. 2

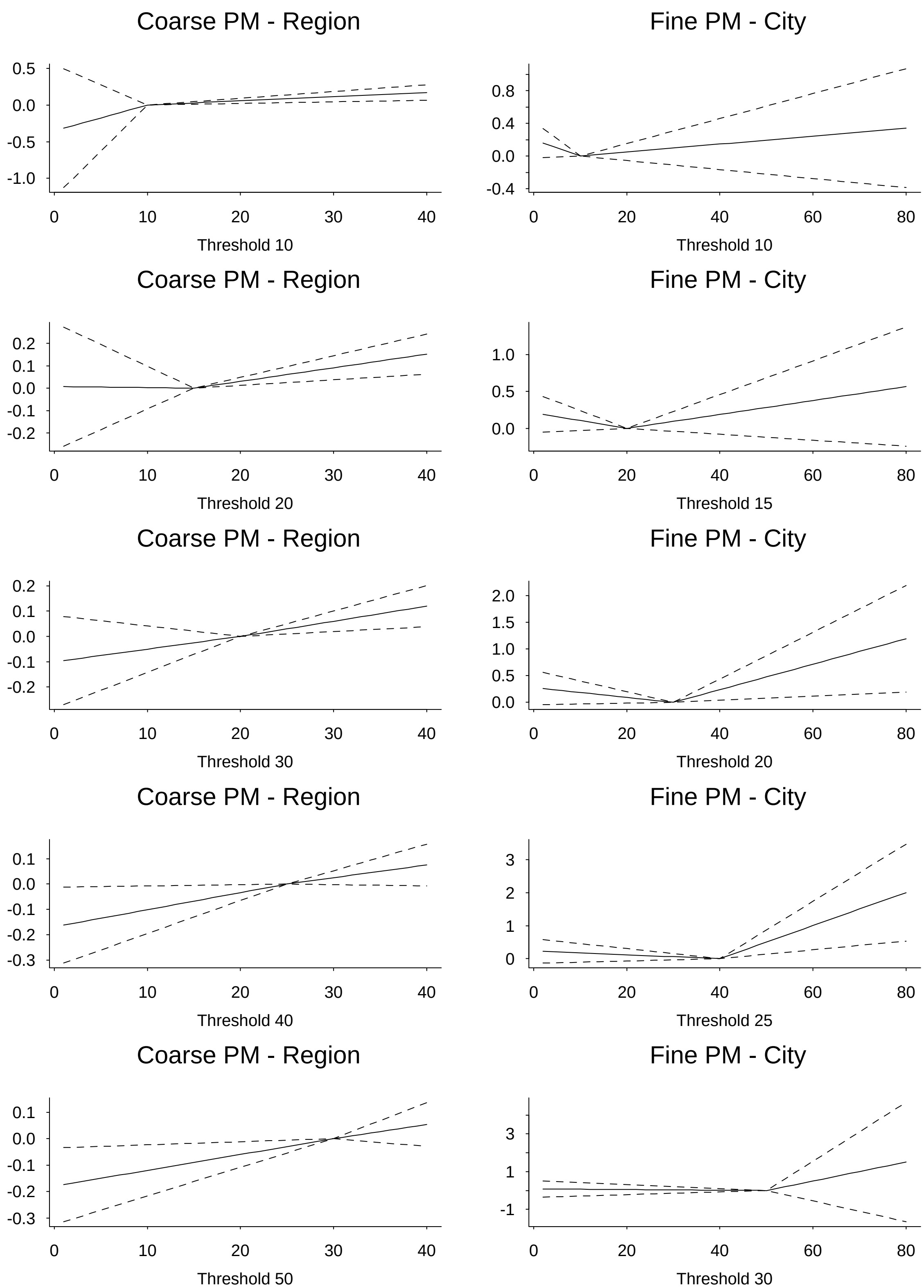
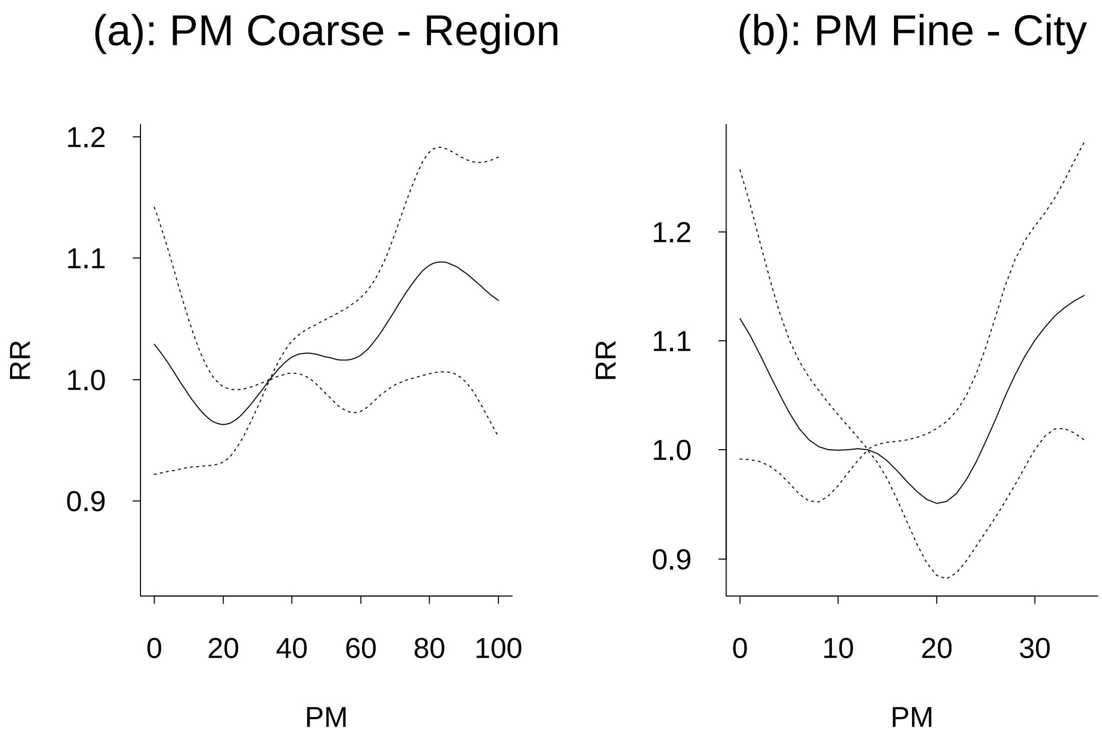
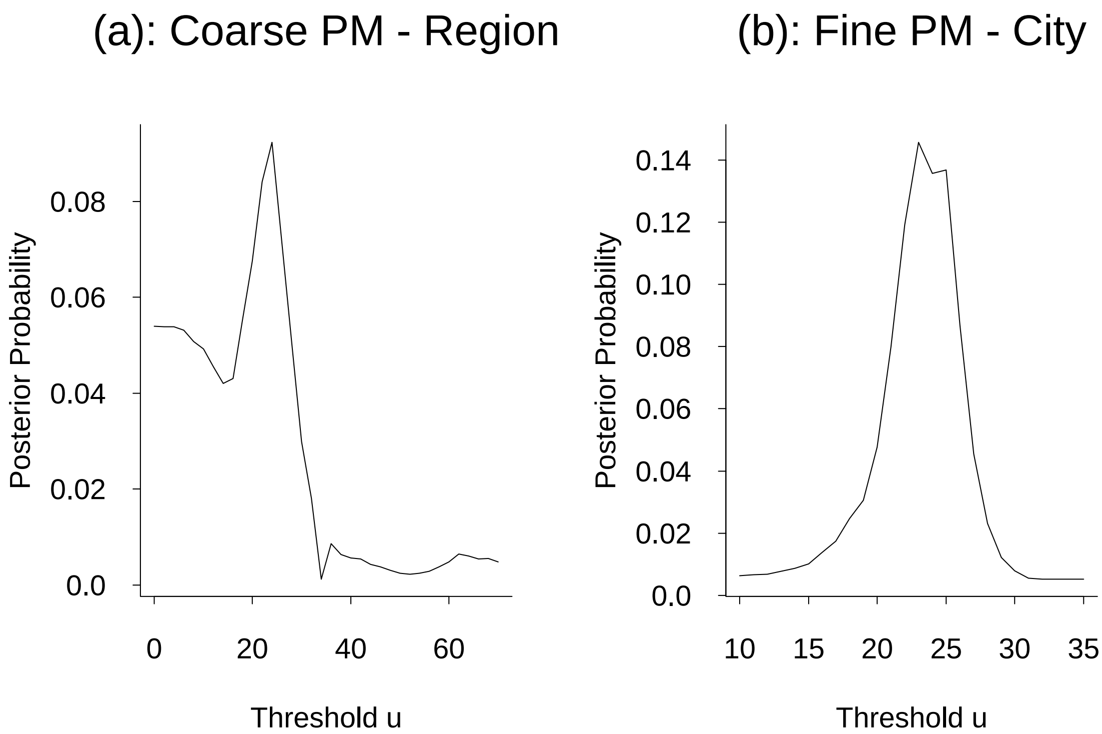

Coarse PM - Region
0.5
0.0
-0.5
-1.0
0 10 20 30 40
Threshold 10
Fine PM - City
0.8
0.4
0.0
-0.4
0 20 40 60 80
Threshold 10
Coarse PM - Region
0.2
0.1
0.0
-0.1
-0.2
0 10 20 30 40
Threshold 20
Fine PM - City
1.0
0.5
0.0
0 20 40 60 80
Threshold 15
Coarse PM - Region
0.2
0.1
0.0
-0.1
-0.2
0 10 20 30 40
Threshold 30
Fine PM - City
2.0
1.5
1.0
0.5
0.0
0 20 40 60 80
Threshold 20
Coarse PM - Region
0.1
0.0
-0.1
-0.2
-0.3
0 10 20 30 40
Threshold 40
Fine PM - City
3
2
1
0
0 20 40 60 80
Threshold 25
Coarse PM - Region
0.1
0.0
-0.1
-0.2
-0.3
0 10 20 30 40
Threshold 50
Fine PM - City
3
1
-1
0 20 40 60 80
Threshold 30

Fig. 3



(a): PM Coarse - Region
(b): PM Fine - City
RR
PM
1.2
1.1
1.0
0.9
0
20
40
60
80
100
10
30

Fig. 4



(a): Coarse PM - Region
Posterior Probability
0.0
0.02
0.04
0.06
0.08
0
20
40
60
Threshold u
(b): Fine PM - City
Posterior Probability
0.0
0.02
0.04
0.06
0.08
0.10
0.12
0.14
10
15
20
25
30
35
Threshold u

Fig. 5