PUBLIC COMMENT ON EPA OZONE STANDARD: NEW TIME SERIES ANALYSES OF THE RELATIONSHIP BETWEEN OZONE AND SHORT-TERM MORTALITY

DOCKET NUMBER EPA-HQ-OAR-2008-0699

RICHARD L. SMITH

MARCH 17, 2015

This comment presents some new time series analyses of the ozone-mortality association, both based on the well-established NMMAPS dataset, and a new 2000-2012 dataset from California. These results are relevant to the proposed changes in the ozone standard for several reasons: time series analyses are by far the most widely studied and definitive results relating ozone directly to mortality, and have been extensively used in the EPA's own Regulatory Impact Analysis. Our overall conclusion is that the results are substantially dependent on various subjective choices of data and statistical models, and therefore, their use to support the proposed new standard must still be regarded as tentative.

Introduction and Summary

The papers by Bell and co-authors (2004, 2006) were the first multi-city studies of the relationship between ozone and short-term mortality, using the National Morbidity, Mortality and Air Pollution (NMMAPS) dataset. This dataset consisted of daily death counts and many associated covariates (including meteorological and air pollution covariates) for 108 US cities for 5,114 days, from January 1, 1987 through December 31, 2000. It remains by far the largest air pollution dataset to have been distributed publicly and widely analyzed by many different authors.

Smith et al. (2009) re-analyzed the NMMAPS dataset and critically examined several of the conclusions relating ozone and short-term mortality. Amongst other things, this paper recast all the results in terms of daily maximum 8-hour average ozone, which is the definition of ozone used in the current EPA standard. In contrast, the earlier papers of Bell et al. used 24-hour average ozone. It is not the intention of the present note to review this material in detail, but we highlight two results as background for the discussion that follows. First, Figure 1 of the present note (reproducing Figure 4 of Smith et al.) shows the variation of ozone-mortality coefficients by city, as well as Bayesian estimates under two priors that were used to combine the results across cities: the "national prior" which treats all cities in the entire database as drawn from a single common distribution, and the "regional prior" which treats cities as from a common distribution within each of seven regions. The figure highlights the very great variation in the raw estimates, and shows how even the Bayesian posterior estimates are sensitive to the choice of prior. Second, Figure 2 of the present note (Figure 6 of Smith et al.) is a smoothed map of the ozone-mortality coefficients as they vary over the country, highlighting the systematic variation over space – in particular, it appears that the ozone-mortality relationship is highest in the north-east of the country (with a few high spots elsewhere, e.g. Texas) while there are many parts of the country for which the coefficient is much smaller and arguably insignificant.

Some time after the publication of Smith et al. (2009), scientists from the EPA requested a file of the city-specific coefficients, which was provided. In the course of preparing the present note, this file has been placed online, as well as the 8-hour ozone data (which was derived from NMMAPS data, but not a part of the original NMMAPS dataset). See www.unc.edu/~rls/ExplanatoryNote.docx for an explanatory note and links to both the 8-hour ozone and mortality coefficients data. These analyses have been heavily cited in the proposed rulemaking documents, especially in the Regulatory Impact Analysis. It is therefore especially important to understand still further how sensitive such results are to both model specification and data. The present note gives further analyses that have been completed within the last few months, both of NMMAPS and other datasets. Papers based on these analyses are currently being prepared for submission to refereed journals.

Major points of the current analyses include:

1. New datasets from California, covering the period 2000-2012, do not show any dependence between short-term mortality and either ozone or $PM_{2.5}$. The mortality data were obtained by Dr. Stan Young and will be both made publicly available and analyzed further in a forthcoming paper of Lopiano, Smith and Young (2015). At the present time, we do not have post-2000 data from other states.
2. The results for 2000-2012 California data are in fact consistent with the results from the 1987-2000 NMMAPS dataset, restricted to the state of California. To establish this point, we have repeated the methods used for the 2000-2012 California dataset (which used daily maximum temperature, daily minimum temperature and daily mean relative humidity as meteorological covariates, different from those used in most published NMMAPS analyses) and applied them to the earlier NMMAPS data. This shows that there is no overall contradiction with previously published results, but this naturally raises the question of whether other NMMAPS results would still hold when applied to more recent data.
3. We have also studied the possibility of nonlinear concentration-response curves more systematically than in earlier NMMAPS papers. For the post-2000 California data, results based on nonlinear concentration-response functions show essentially the same result as for linear concentration-response functions – no overall effect. For the NMMAPS data, the results also do not show a consistent effect across regions, the “national average” result being the only one that shows an increasing and statistically significant mortality risk across a wide range of ozone levels. We have also examined the possible role of either confounding or effect modification by PM_{10} in these analyses. (Some analyses have also been done with other criteria pollutants as the possible confounder, but these lead to similar conclusions.) The results do not contradict the claims of Bell et al. (2004, 2007), who stated that confounding by PM_{10} does not play a major role in the ozone-mortality relationship. However, we have also attempted an effect modification analysis by essentially splitting the data for each city into two parts, either below or above a pre-defined PM_{10} threshold, and examining the ozone-mortality association separately for each part. The results for this analysis are mixed: the basic problem is that as soon as one attempts any analysis beyond the simple national-average ozone-only analysis, the confidence bands become too wide to make any definitive conclusions.

In the author’s view, the robustness of the ozone-mortality association (to changes in the time and spatial region covered by the data, to the covariates included in the analysis, and to the form of the statistical model) is still far from well-established. Therefore, claims about the benefits of the proposed change in the standard are speculative.

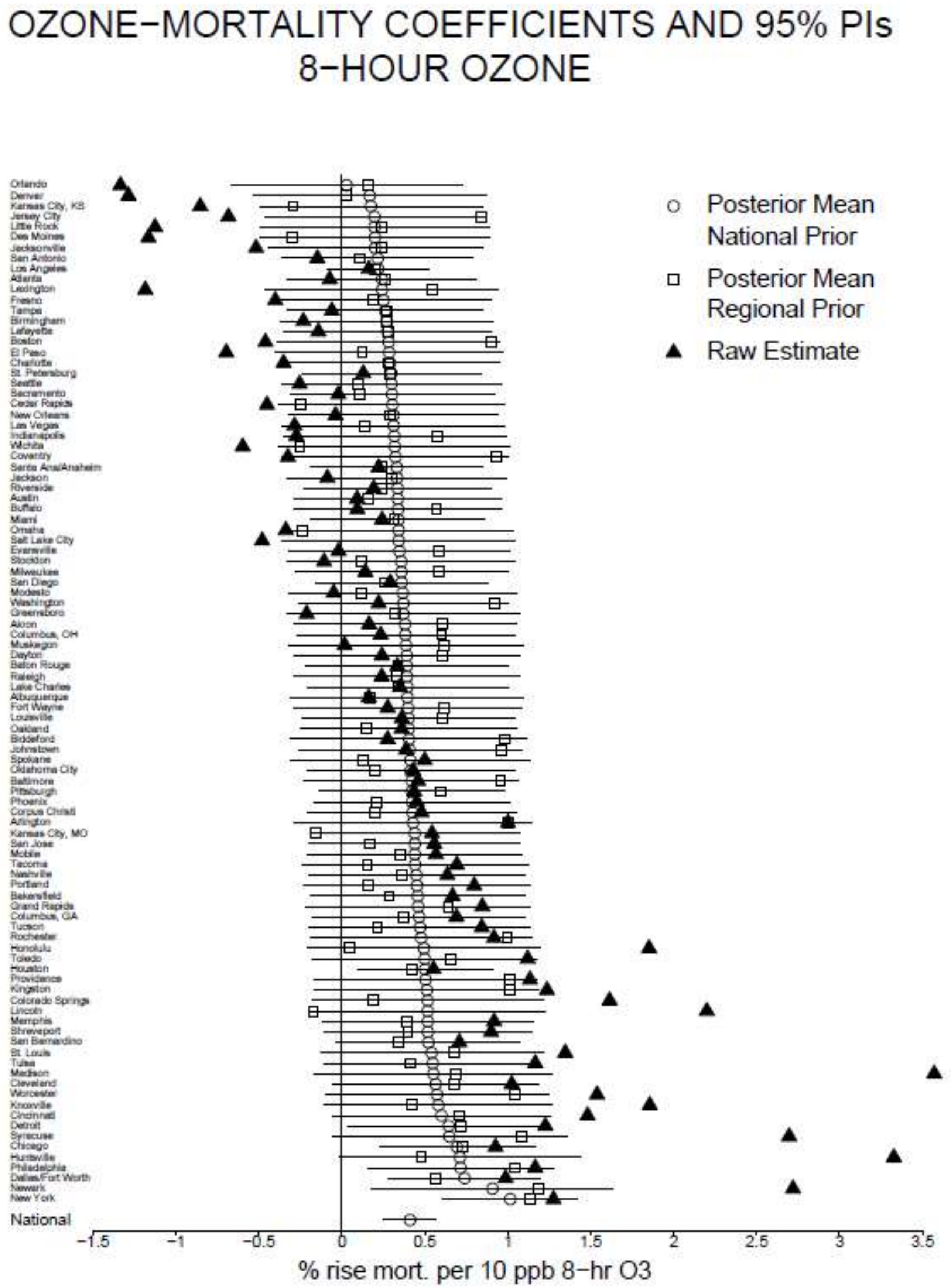

Figure 1 (Figure 4 of Smith et al., 2009). Raw city-specific estimates (triangles), posterior estimates under the "regional" prior (squares) and posterior estimates (circles) and 95% posterior intervals under the "national" prior, for the ozone-mortality coefficient based on 8-hour ozone.

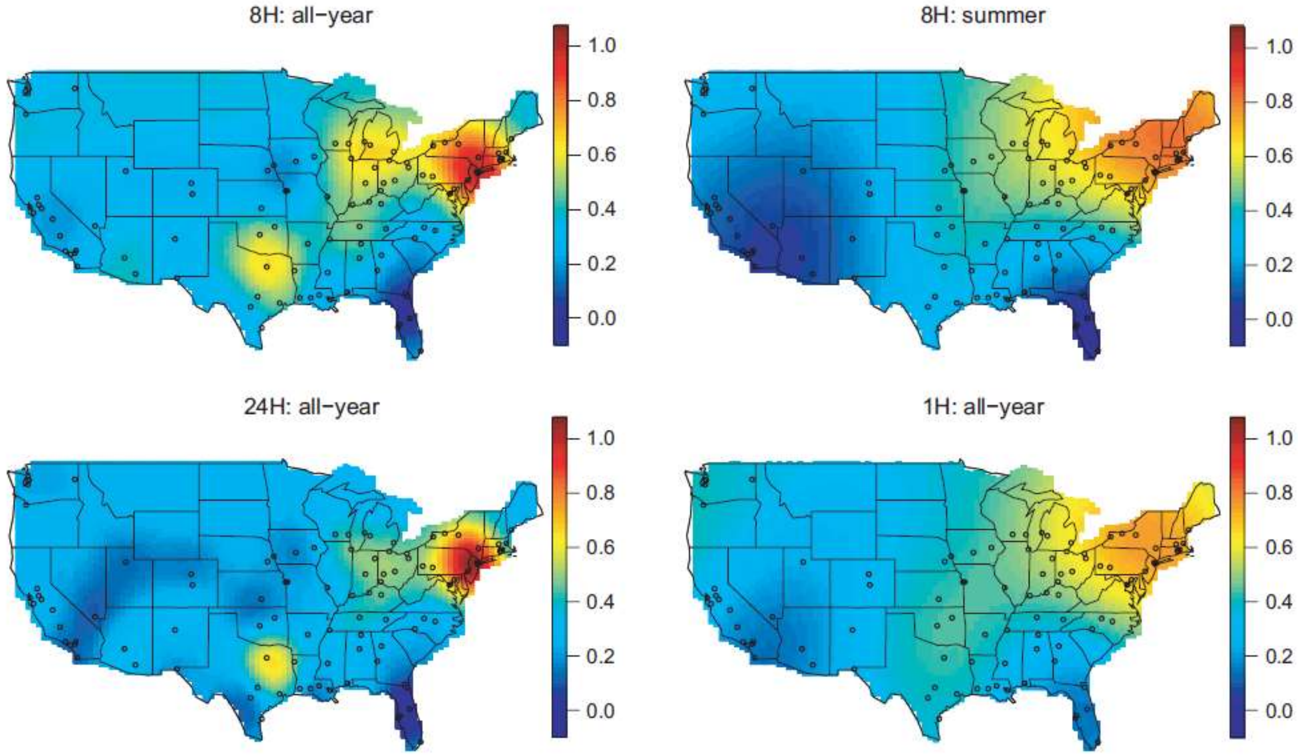

Figure 2 (Figure 6 of Smith et al., 2009). Map of spatially dependent ozone-mortality coefficients for 8-hour ozone (all-year data), 8-hour ozone (summer data), 24-hour ozone (all-year data) and 1-hour ozone (all-year data).

The remainder of this discussion presents an outline of the new analyses that support the above points.

Time Series Analysis of California Mortality Data

The time series model is adapted from models previously used for the National Morbidity, Mortality and Air Pollution (NMMAPS) data series, see in particular Dominici et al. (2003), Bell et al. (2004) and Smith et al. (2009). The code used for the results in the present paper is at www.unc.edu/~rls/EpiTimeSeriesCodeRLS.txt.

The basic model is of the form

$$\begin{aligned}\mathrm{Log}(\mu_t) = {} & \text{Overall Mean} + \mathrm{DLM}(l_1,\ldots,l_k) + s(t, n_{yr}*df_0) + \mathrm{DOW} \\ & + s(M_1(t); df_1) + s((M_1(t-1) + M_1(t-2) + M_1(t-3))/3; df_2) \\ & + \ldots + s(M_p(t); df_1) + s((M_p(t-1) + M_p(t-2) + M_p(t-3))/3; df_2) \qquad (1)\end{aligned}$$

where

μ_t is expected number of deaths on day t;

$\mathrm{DLM}(l_1,\ldots,l_k)$ refers to a (linear) distributed lag model for the air pollution variable; this includes regression terms $\beta_1(X(t- l_1) + X(t- l_2) + \ldots + X(t- l_k))/k + \beta_2\{ X(t- l_2) - (X(t- l_1) + X(t- l_2) + \ldots + X(t- l_k))/k\} + \ldots + \beta_k\{ X(t- l_k) - (X(t- l_1) + X(t- l_2) + \ldots + X(t- l_k))/k\}$ where the lead coefficient β_1 represents the mean rise in mortality per one unit rise in air pollution X, distributed over lags $l_1,\ldots,l_k$;

$s(t, n_{yr}*df_0)$ refers to a natural spline on time variable t over n_{yr} years with df_0 degrees of freedom per year; this represents the long-term trend (including seasonal component);

DOW is a day of week component (treated as a factor variable);

$s(M_1(t); df_1)$ represents a nonlinear trend on current-day value of met variable M_1 with df_1 degrees of freedom;

$s((M_1(t-1) + M_1(t-2) + M_1(t-3))/3; df_2)$ represents a nonlinear trend on the average of the three previous days of met variable M_1 with df_2 degrees of freedom;

other met variables $M_2,\ldots,M_p$ are treated similarly to M_1;

days with missing data are omitted from the analysis;

the model is fitted as a generalized linear model with log link and *quasipoisson* mean-variance structure; this is similar to assuming a Poisson distribution but with an additional parameter representing overdispersion.

In addition to fitting model (1) with all the variables, we have also fitted the model without the air pollution component and dropping some of the meteorological terms. A likelihood ratio test is conducted when each of the terms from $s(M_1(t); df_1)$ to $s((M_p(t-1) + M_p(t-2) + M_p(t-3))/3; df_2)$ is dropped; this is an additional check that the selection of meteorological terms is appropriate.

For the California analysis, we have used three meteorological variables: daily maximum temperature, daily minimum temperature and daily mean relative humidity. Previous NMMAPS studies including Smith et al (2009) have used two meteorological variables, daily mean temperature and dewpoint, but otherwise the same model form as above. For the degrees of freedom, in previous studies df_0 has been typically taken between 7 and 12, df_1 and df_2 between 3 and 6; we have varied these by trial and error to understand the sensitivity of the analysis to these choices.

The most critical component of the model (1) is the selection of lags $l_1,\ldots,l_k$ to represent the air pollution component. The NMMAPS analyses in Bell et al. (2004) and Smith et al. (2009) used lags 0, 1, 2,…,6 (in some cases with an additional refinement, the *constrained distributed lag* model in which some of the coefficients $\beta_2,\ldots,\beta_k$ are constrained to be equal; however, this usually has only minor impact on the important coefficient β_1). Other common approaches use any of lags 0, 1, 2 in a single-lag model, or averages over any combination of lags 0, 1, 2, 3. For the present study, we have tried different combinations of lags to look for the lag combination that best represents the air pollution effect. We believe this approach to be justified in view of the weak evidence for any air pollution effect in these dataset; however, in view of the selection bias inherent in such an approach, we caution against over-interpretation of such results, especially in cases where the p-value is over 0.01 or the result highly depend on the selection of a particular combination of lags.

Figure 3: Map of California Air Basins (Source: Webpage of the California Air Resources Board)

South Coast Air Basin

The approach outlined in the previous section is applied to data from each of eight California air basins (Fig. 1). Because they are the two most populated air basins, we concentrate initially on the South Coast air basin (which includes Los Angeles, Orange, Riverside and San Bernardino Counties) and the San Francisco Bay air basin (San Francisco, Marin, Sonoma, Napa, Solano, Contra Costa, Alameda, Santa Clara and San Mateo counties). A map of all the air basins is in Figure 3. For the response variable, we use total non-accidental mortality among people aged 65 and over.

Fitting the meteorological model alone, in Table 1 we tabulate the p-value associated with dropping each of the six terms in turn. Five of the meteorological variables are very highly significant; the only exception is current-day relative humidity. This result is based on the particular choices df_0=7, df_1=df_2=6, but the overall conclusion is robust against alternative values of those three degree of freedom parameters.

Variable	Lags	p-value
Daily Max Temperature	Current day 0	<1 e-16
Daily Max Temperature	Mean of 1,2,3	4.6 e-7
Daily Min Temperature	Current day 0	2.5 e-4
Daily Min Temperature	Mean of 1,2,3	2.4 e-5
Mean Daily Relative Humidity	Current day 0	0.18
Mean Daily Relative Humidity	Mean of 1,2,3	1.5 e-10

Table 1: Statistical significance of meteorological components: based on model (1) without air pollution component and with df_0=7, df_1=df_2=6, fitted to nonaccidental mortality for ages 65 and up, South Coast air basin.

In subsequent analyses, we have retained all six meteorology components; this is to ensure consistency across different air basins and to avoid the analysis being biased by overuse of statistical significance tests; however, Table 1 is evidence that we have identified appropriate meteorological variables for the overall analysis.

We now consider addition air pollution variables to the meteorological model in Table 1. Initially, we concentrate on ozone. Table 2 shows the coefficient estimates, standard error (SE), t-value and p-value associated with ozone at various combination of lags. The units here are percent rise in mortality per 10 ppb rise in ozone. The strongest positive coefficient is based on lags 0, 1, 2 and 3, for which the model predicts a 0.1% rise in mortality per 10 ppb rise in ozone. However, neither this nor any of the other values in the table comes anywhere close to being statistically significant. This is for 13 years of data over one of the most densely populated areas of the US – if there is an ozone-mortality effect in California, we ought to see it here.

Lags Included	Estimate	SE	t-value	p-value
0	0.0870	0.1135	0.77	0.44
1	-0.0472	0.1136	-0.42	0.68
2	0.0471	0.1141	0.41	0.68
0,1	0.0266	0.1315	0.20	0.84
1,2	0.0002	0.1330	0.00	1.00
0,1,2	0.0825	0.1507	0.55	0.58
0,1,2,3	0.1222	0.1673	0.73	0.46
0,1,2,3,4	0.0941	0.1802	0.52	0.60
0,1,2,3,4,5	0.0096	0.1905	0.05	0.96
0,1,2,3,4,5,6	-0.0479	0.1992	-0.24	0.81

Table 2: Statistical significance of ozone component with various combinations of lags: based on model (1) df_0=7, df_1=df_2=6. Estimate is percent rise in mortality for 10 ppb rise in ozone. South Coast air basin; response variable is non-accidental mortality aged 65 and over.

The model fit is illustrated pictorially in Figure 4. In this plot, we show residuals from the parts of the model that do not include air pollution (long-term trend, day of week and meteorology). If there was a strong ozone effect, this ought to be apparent from the plot: it is not. This is reinforced by the fitted straight line and associated confidence bands, showing very little dependence on ozone.

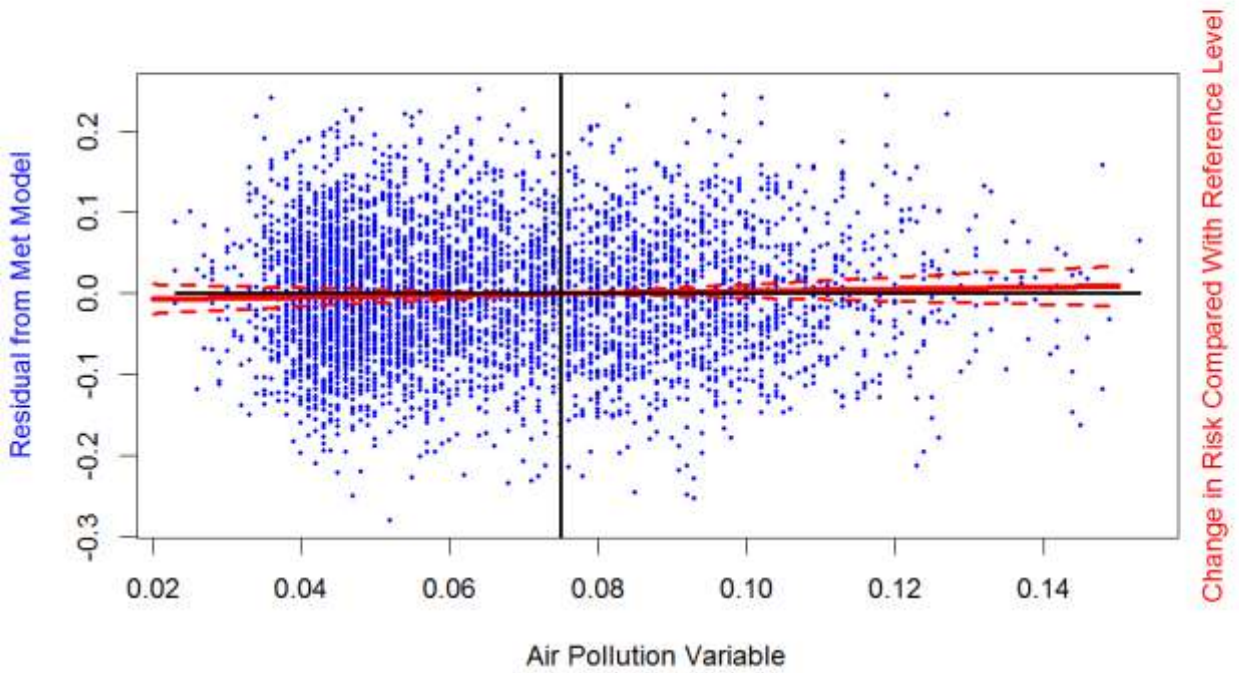

Figure 4. South Beach air basin. Blue dots: residuals from the model that includes long-term trends, day of week and meteorology, plotted against the air pollution variable (ozone). Red solid and dashed lines: implied change of relative risk with respect to ozone level 0.075 ppm (the current ozone standard), with pointwise 95% confidence bands.

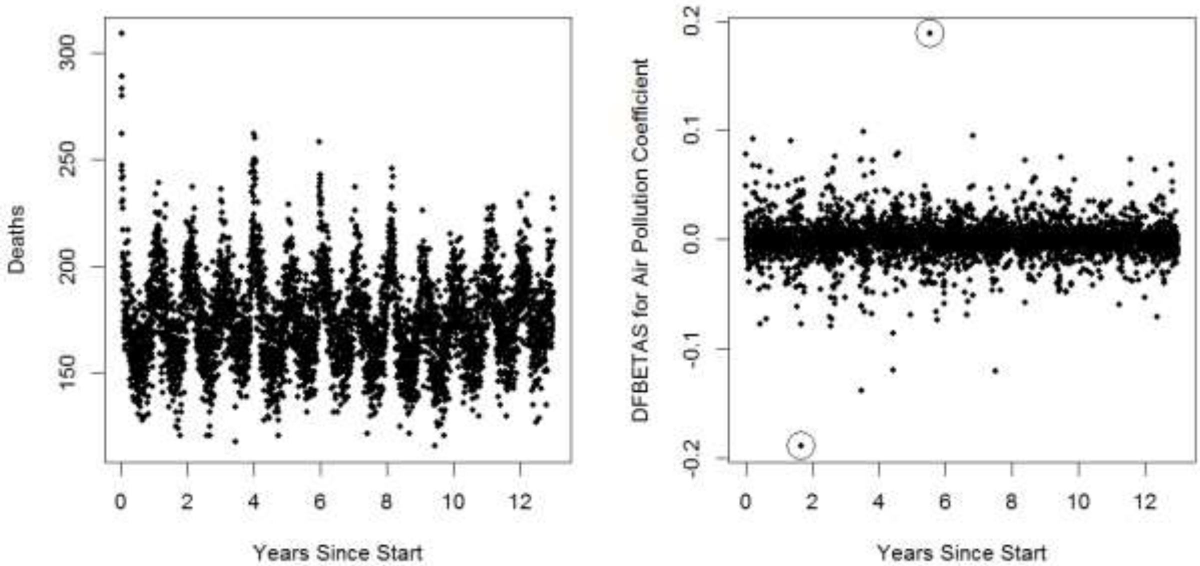

Figure 5: (Left) Plot of daily deaths in South Coast air basin. The seasonal cycle is evident. (Right) Plot of DFBETAs for the leading air pollution coefficient. The two circled points are the ones with the most extreme value of DFBETAs, implying most influence on the air pollution coefficient.

Figure 5 illustrates another kind of plot. The left hand plot is a plot of the raw daily mortality data against time: this makes obvious the very strong seasonal effect. The right hand plot is for the regression diagnostic DFBETAS (Belsley et al., 1980) which is a common diagnostic for determining the influence of each individual observation on any of the fitted coefficients. In this case, DFBETAS is plotted for the main air pollution regression coefficient of interest. The most extreme observations, in terms of their influence on the air pollution coefficient, are observation 626 at the lower and observation 2026 at the upper end (circled on plot). This suggests trying to refit the model with either of these observations omitted. If this is done, the regression coefficient 0.1222 reported in Table 2 changes to 0.1512 if observation 626 is omitted and 0.0879 if observation 2026 is omitted; the standard error remains at about 0.167 in all three cases. Thus, although the regression coefficient does indeed change if one of these influential observations is omitted, the coefficient is still not large enough to produce a statistically significant result.

The same analysis was tried using $PM_{2.5}$ in place of ozone, with results shown in Table 3. In this case, several of the estimates appear to be statistically significant with a p-value <0.05 (smallest value 0.017), but all the statistically significant values are negative, which is not biologically plausible. We conclude that either the small p-values are an artifact of the selection effect already mentioned, or there is some other biological mechanism, such as confounding by some other pollutant, that explains these results.

Lags Included	Estimate	SE	t-value	p-value
0	0.1261	0.0998	1.26	0.21
1	-0.1966	0.0990	-1.99	0.05
2	-0.2121	0.0995	-2.13	0.03
0,1	-0.0425	0.1144	-0.37	0.71
1,2	-0.2720	0.1151	-2.36	0.018
0,1,2	-0.1133	0.1294	-0.88	0.38
0,1,2,3	-0.1636	0.1409	-1.16	0.25
0,1,2,3,4	-0.1611	0.1499	-1.07	0.28
0,1,2,3,4,5	-0.2609	0.1582	-1.65	0.10
0,1,2,3,4,5,6	-0.2435	0.1659	-1.47	0.14

Table 3: Statistical significance of $PM_{2.5}$ component with various combinations of lags: based on model (1) df_0=7, df_1=df_2=6. Estimate is percent rise in mortality for 10 μg/m^3 rise in $PM_{2.5}$. South Coast air basin; response variable is non-accidental mortality aged 65 and over.

In these analyses, the overdispersion parameter was of the order of 1.07 – in other words, the variance of the mortality variables is inflated by a factor of 1.07 compared with the Poisson distribution. This is typical for this kind of analysis and does not indicate a problem. A much larger overdispersion parameter could indicate some important missing covariates.

San Francisco Bay Air Basin

So far, we have only considered one air basin. The second most populated is San Francisco Bay, which has substantially different weather patterns and demographics from the Los Angeles area. Therefore, the entire analysis has been repeated for this air basin, as a test of how robust the analyses are for different regions of the state.

Table 4 shows the statistical significance of the individual meteorology components, analogous to Table 1 for the South Coast air basin. The main difference from Table 1 is that the component due to relative humidity is not statistically significant. (Although not reported in the table, if both relative humidity components – current day and the average of lags 1,2,3 – are dropped together, rather than one at a time, we also do not get a statistically significant component due to relative humidity.) In the following analyses, to maintain consistency of analysis methods across different air basins, the main results are still reported including relative humidity, but to assess the sensitivity to this component, some of the analyses have been repeated omitting relative humidity altogether.

Variable	Lags	p-value
Daily Max Temperature	Current day 0	9.05E-11
Daily Max Temperature	Mean of 1,2,3	0.0071
Daily Min Temperature	Current day 0	0.0019
Daily Min Temperature	Mean of 1,2,3	0.043
Mean Daily Relative Humidity	Current day 0	0.41
Mean Daily Relative Humidity	Mean of 1,2,3	0.32

Table 4: Statistical significance of meteorological components: based on model (1) without air pollution component and with df_0=7, df_1=df_2=6, fitted to nonaccidental mortality for ages 65 and up, San Francisco Bay air basin.

Table 5 shows the results when ozone is added to the analysis. As with our earlier analyses for the South Coast air basin, none of the estimates of the ozone effect at various lags is statistically significant at the 0.05 level. However, two of the analyses (with lag 0 alone, and with lags 0 and 1 together) are statistically significant with a p-value of about .02 if the relative humidity component is omitted. This result illustrates the principle that if enough different models are tried, it is usually possible to find some model that gives a statistically significant result: it does not imply that the result is significant in any practical sense. It should also be noted, however, that all the coefficients of models that include lag 0 are similar in magnitude (between 0.3 and 0.6): the variation in p-values is mostly due to their standard errors.

Lags Included	RH included?	Estimate	SE	t-value	p-value
0	yes	0.4464	0.2471	1.81	0.071
1	yes	0.1889	0.2413	0.78	0.43
2	yes	-0.1560	0.2442	-0.4	0.52
0,1	yes	0.4909	0.3030	1.62	0.11
1,2	Yes	0.0225	0.2947	0.08	0.94
0,1,2	Yes	0.3281	0.3502	0.94	0.35
0,1,2,3	Yes	0.4210	0.3927	1.07	0.28
0,1,2,3,4	Yes	0.4716	0.4167	1.13	0.26
0,1,2,3,4,5	Yes	0.4703	0.4310	1.09	0.28
0,1,2,3,4,5,6	Yes	0.3325	0.4448	0.75	0.45
0	No	0.4838	0.2121	2.28	0.023
0,1	No	0.5948	0.2604	2.28	0.022

Table 5: Statistical significance of ozone component with various combinations of lags: based on model (1) df_0=7, df_1=df_2=6. Relative humidity is omitted from some of the analyses. Estimate is percent rise in mortality for 10 ppb rise in ozone. San Francisco Bay air basin; response variable is non-accidental mortality aged 65 and over.

Table 6 shows the corresponding results for $PM_{2.5}$, where again relative humidity has been omitted from some of the analyses to illustrate the sensitivity to this component. Our conclusions are similar: some rows of this table show a statistically significant effect with a p-value of the order 0.02, but taking account of the number of models examined in order to achieve this result, it is unlikely to be of practical significance.

Lags Included	RH included?	Estimate	SE	t-value	p-value
0	Yes	0.3031	0.2362	1.28	0.20
1	Yes	0.1235	0.2373	0.52	0.60
2	Yes	0.3769	0.2312	1.63	0.10
0,1	Yes	0.3968	0.2700	1.47	0.14
1,2	Yes	0.4614	0.2679	1.72	0.09
0,1,2	Yes	0.5903	0.3067	1.92	0.05
0,1,2,3	Yes	0.5688	0.3297	1.72	0.08
0,1,2,3,4	Yes	0.5042	0.3482	1.45	0.15
0,1,2,3,4,5	Yes	0.5500	0.3634	1.51	0.13
0,1,2,3,4,5,6	Yes	0.4884	0.3767	1.30	0.19
0,1,2,3	No	0.5712	0.3123	1.83	0.07
0,1,2,3,4	No	0.6518	0.3341	1.95	0.05
0,1,2,3,4,5	No	0.8169	0.3535	2.31	0.021
0,1,2,3,4,5,6	No	0.7737	0.3702	2.09	0.037

Table 6: Statistical significance of $PM_{2.5}$ component with various combinations of lags: based on model (1) df_0=7, df_1=df_2=6. Relative humidity is omitted from some of the analyses. Estimate is percent rise in mortality for 10 µg/m^3 rise in $PM_{2.5}$. San Francisco Bay air basin; response variable is non-accidental mortality aged 65 and over.

To give some pictorial assessment of the model fit and its sensitivity, in this case, we have chosen the model with lags 0,1,2,3,4,5 and no RH. Once again, we should point out that the interpretation of such a model must take account of the deliberate selection of the model to give the largest air pollution effect – we have followed this strategy in order to reinforce the difficulty of finding any model with a truly significant $PM_{2.5}$ coefficient.

The figures analogous to Figures 4 and 5 are now Figures 6 and 7. Figure 6 illustrates that even with this deliberate choice of model, the increase in modeled relative risk over the range of the data is slight. The DFBETAs plot on the right side of Figure 7 identifies observations 373 and 1071 as most influential; when these are omitted, the regression coefficient changes from 0.8169 to 0.8847 (observation 373 omitted) or 0.7308 (observation 1071 omitted). In neither case does the new regression coefficient change any of our substantive conclusions about the data.

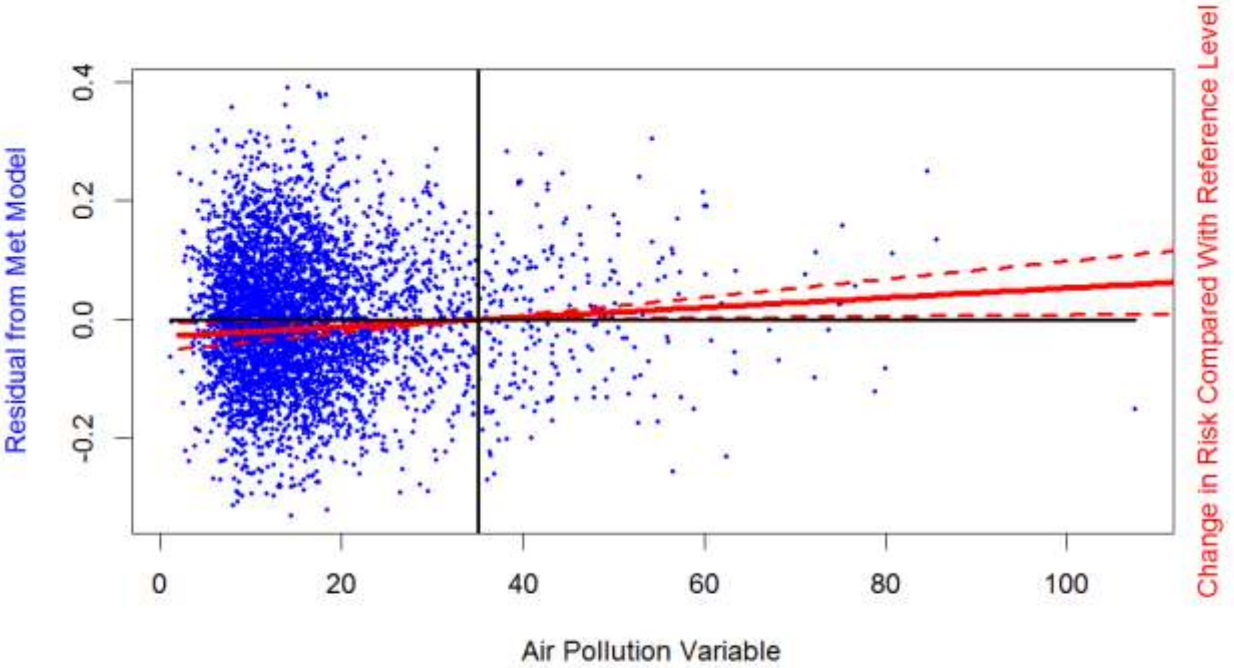

Figure 6. San Francisco Bay air basin. Blue dots: residuals from the model that includes long-term trends, day of week and meteorology, plotted against the air pollution variable ($PM_{2.5}$). Red solid and dashed lines: implied change of relative risk with respect to $PM_{2.5}$ level 35 $\mu g/m^3$ (the current daily $PM_{2.5}$ standard), with pointwise 95% confidence bands.

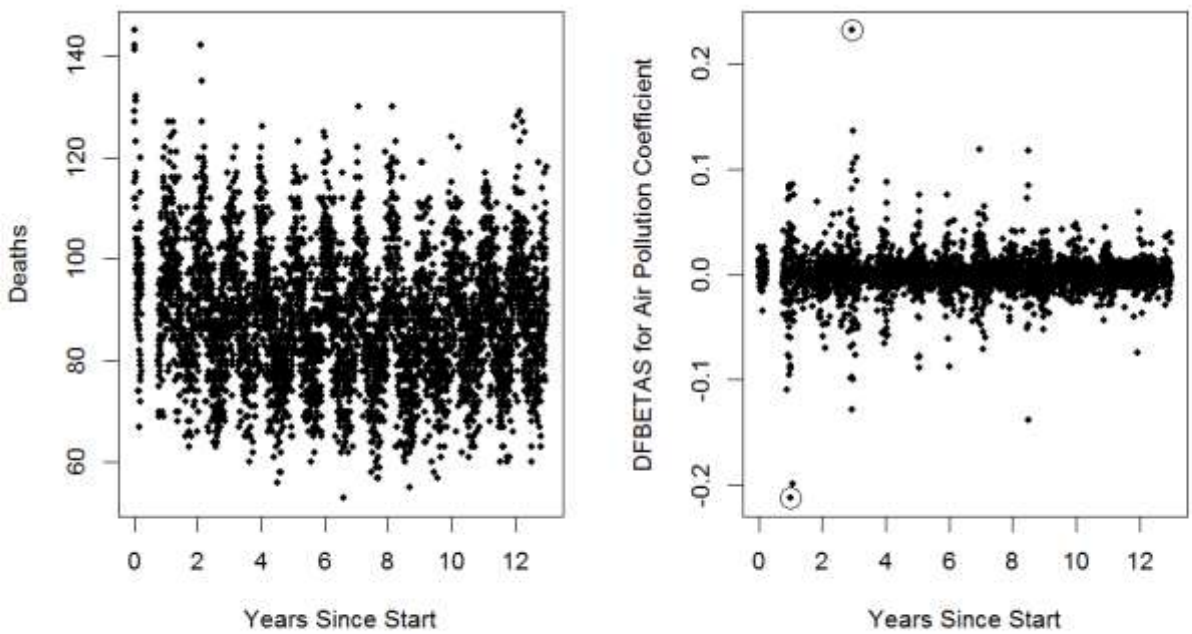

Figure 7: (Left) Plot of daily deaths in San Francisco Bay air basin. (Right) Plot of DFBETAs for the leading air pollution coefficient. The two circled points are the ones with the most extreme value of DFBETAs, implying most influence on the air pollution coefficient.

The overdispersion parameter for these analyses was around 1.05.

Combining Results Across Air Basins

In the NMMAPS papers on ozone (Bell et al 2004, Smith et al 2009), the single-city analyses were repeated for up to 98 US cities for which ozone and mortality data were available. They were then combined across cities using a hierarchical model analysis, based on an algorithm originally due to Everson and Morris (2000) and code by Roger Peng into the R function "tlnise". The same method is used here to produce estimates that are combined across all eight air basins in our study.

It would not be practicable (or interpretable) to repeat all the analyses for every combination of meteorological variables, lags of the pollutant variable, or degrees of freedom for the spline components of the model. Therefore, some choices were made, guided by the analyses already conducted for the South Coast and San Francisco Bay air basins, as follows:

1. All analyses used all six meteorological variables.
2. The degree of freedom parameters were set to be respectively 7, 6 and 6, for df_0, df_1 and df_2.
3. For both ozone and $PM_{2.5}$, only certain certain combinations of lags were tried.

The results of this analysis are shown in Table 7. None of the analyses show a statistically significant effect when combined across all eight air basins.

Variable	Lags	Estimate	SE	t-value	p-value
Ozone	0,1	0.3376	0.2434	1.39	0.17
Ozone	0,1,2	0.3165	0.2466	1.28	0.20
Ozone	0,1,2,3	0.4149	0.3260	1.28	0.20
PM2.5	0,1	0.0126	0.2034	0.06	0.95
PM2.5	0,1,2,3	-0.0006	0.2464	0.00	1.00
PM2.5	0,1,2,3,4,5	0.0689	0.2799	0.25	0.81

Table 7: Combined results across all eight air basins.

All the analyses in this paper so far are based on total non-accidental mortality for ages 65 and up. The analysis was repeated using (a) total non-accidental mortality for all ages, (b) respiratory deaths aged 65 and up, (c) circulatory deaths aged 65 and up, (b) combined respiratory and circulatory deaths aged 65 and up. None of these produced a statistically significant result in the combined analyses.

The results of Table 7 were also repeated with the choices df_0=7, df_1=6, df_2=6 replaced by (a) df_0=10, df_1=6, df_2=6, (b) df_0=7, df_1=3, df_2=3, (c) df_0=10, df_1=3, df_2=3. The analysis of Table 7 was also repeated with relative humidity omitted from the analysis. None of these changes produced a statistically significant result in any of the combined analyses.

Comparisons with NMMAPS

We have pointed out that the statistical methods of this paper are similar to those of the NMMAPS study (see in particular Bell et al. 2004, Smith et al. 2009), but they are not identical. Those papers also included an interaction effect between age and long-term trend, and the meteorological variables were daily mean temperature and dewpoint, rather than those of the present paper. What happens if we use exactly the same methods for the two datasets?

To investigate this question, we recompiled the NMMAPS dataset but using tmax, tmin and daily max relative humidity as the meteorological variables. (Those variables are all in the NMMAPS dataset, but were not used in the previously cited papers.) The dataset was analyzed using the same computer code as the other analyses in this paper, applied to deaths aged 65 and over analyzed as a single age group (no interactions). We took $df_0=7$, $df_1=df_2=6$ as in most of the analyses in this paper, and the distributed lag structure based on lags 0 through 6.

City	Estimate	SE	t-value	p-value
Bakersfield	0.7031	0.9970	0.71	0.48
Fresno	0.1577	0.9520	0.17	0.87
Los Angeles	0.1941	0.2199	0.88	0.38
Modesto	0.3027	1.5057	0.20	0.84
Oakland	0.8943	1.0210	0.88	0.38
Riverside	0.0255	0.6019	0.04	0.97
Sacramento	-0.0913	0.8334	-0.11	0.91
San Bernardino	0.7358	0.6330	1.16	0.25
San Diego	0.1080	0.4717	0.23	0.82
San Jose	-0.0481	0.9756	-0.05	0.96
Santa Ana Anaheim	0.1231	0.4815	0.26	0.80
Stockton	0.9981	1.3775	0.72	0.47
All CA	0.2485	0.2307	1.08	0.28
National	0.2873	0.0915	3.14	0.0017

Table 8: Estimates for the ozone effect in 12 California cities from the NMMAPS study (San Francisco omitted because of lack of ozone data). Also shown are the combined results from all 12 cities under "All CA", and the combined results of all 98 US cities included in the NMMAPS ozone study. Applied to all deaths aged 65 and up, using tmax, tmin and maximum relative humidity as the three meteorological variables, and a distributed lag model for ozone covering lags 0-6.

Since the first part of this note is concerned with California data, we concentrated on the California cities in the NMMAPS database. Table 8 shows results for each city, and the combined result for all 12 California cities. Also shown in Table 8 is the national result, in which the 12 California cities were combined with 86 other US cities, reanalyzed using the software of the present paper.

The last result shows a combined estimate of 0.287 (percent rise in mortality per 10 ppb rise in 8-hour daily max ozone) and a standard error (more precisely, posterior standard deviation) of 0.0915. By comparison, the result quoted in Smith et al. (2009) was a combined estimate of 0.411 and a posterior standard deviation of 0.080. Just to make a further comparison with the results of Smith et al (2009), the method of the present paper was repeated with mortality data from all age groups 55 and up (the same as in the original NMMAPS analyses) – in this case our estimated combined national coefficient, using the meteorological model of the present paper, rises only very slightly, from 0.287 to 0.300. Therefore, the difference in combined estimates compared with Smith et al (2009) appears to be due to the different meteorological variables used and not to the different treatments of age groups. It seems plausible that the treatment of meteorology in the present paper (in particular, the separate use of tmax and tmin) is superior to the treatment in the earlier NMMAPS papers, resulting in a lower estimate of the ozone effect because of less confounding by weather.

Nonlinear Distributed Lag Models

In this section, we consider an extension of the preceding analyses that allows for the leading air pollution term to be nonlinear.

Specifically, where we have previously defined the DLM($l_1,\ldots,l_k$) to have components

$$\beta_1(X(t-l_1) + X(t-l_2) + \ldots + X(t-l_k))/k + \beta_2\{X(t-l_2) - (X(t-l_1) + X(t-l_2) + \ldots + X(t-l_k))/k\} + \ldots + \beta_k\{X(t-l_k) - (X(t-l_1) + X(t-l_2) + \ldots + X(t-l_k))/k\}, \quad (2)$$

we now replace the term $\beta_1(X(t-l_1) + X(t-l_2) + \ldots + X(t-l_k))/k$ by a nonlinear term of form

$$s((X(t-l_1) + X(t-l_2) + \ldots + X(t-l_k))/k; df_3), \quad (3)$$

in other words, a nonlinear spline in the average air pollution variable over all k lags, with df_3 degrees of freedom. As in previous discussions of nonlinear spline terms, there is no hard and fast rule for choosing df_3, but in subsequent discussion we have generally set it equal to 6 since this is large enough in most cases to show a clear nonlinearity but not so large that the model is distorted by evident overfitting.

The proposed model (3) is not as general as that of Gasparrini et al. (2010), who have proposed a tensor product representation that, in effect, allows all the components in (2) to be nonlinear at the same time. Such a model may be appropriate for examining the temperature-mortality dependence (the main focus of the cited paper) where, as we have seen, there is much stronger evidence for the direct influence of the covariate, but for the present study, where the covariate of interest is either ozone or $PM_{2.5}$, such a model would surely be overfitting, given the paucity of evidence for even a linear effect. On the other hand, by retaining the linear terms $\beta_2,\ldots,\beta_k$ in (2), the model is more general than that of Bell et al. (2006), who used a spline term in the average ozone at lags 0 and 1, but without any "distributed lag" component.

As a first example of this method, we develop the nonlinear analog of Figures 4 and 6. The model formed by combining (1) and (3), based on lags 0,1,2, and 3, was fitted to all eight California air basins for both ozone and $PM_{2.5}$. Selected results are in Figures 8 and 9; results for other air basins are similar. In neither case is there any evidence of a systematic increase in risk with ozone or $PM_{2.5}$.

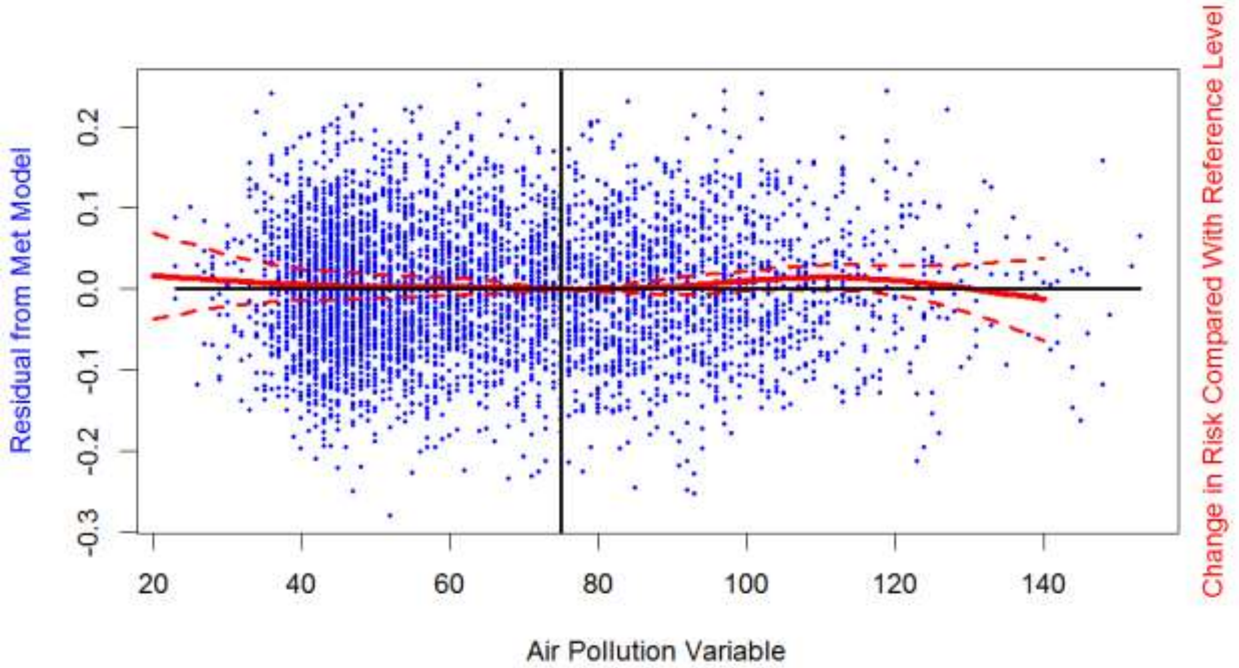

Figure 8. Nonlinear dependence of mortality on ozone for South Beach air basin. Analogous to Figure 4, but using a nonlinear model for the ozone-mortality relationship.

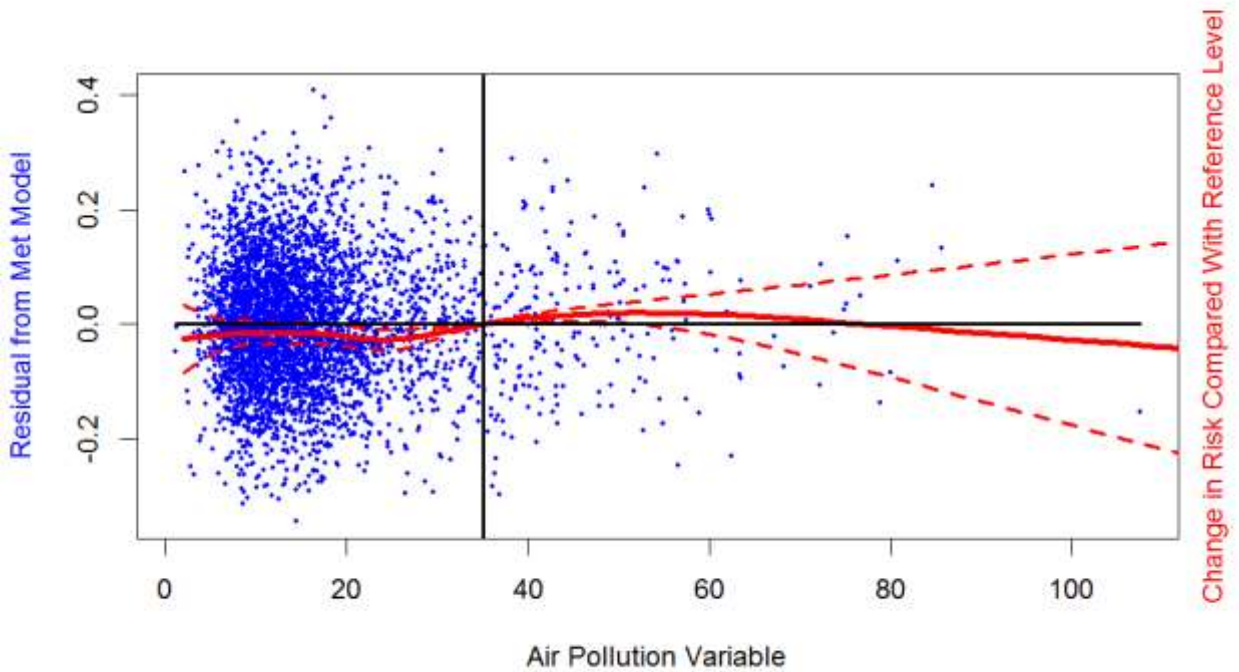

Figure 9. Nonlinear dependence of mortality on $PM_{2.5}$ for San Francisco Bay air basin. Analogous to Figure 6, but using the full meteorological model (including relative humidity), and a nonlinear model for the ozone-mortality relationship.

Next, we have computed a combined curve for all eight California air basins, using the "tlnise" algorithm to combine the estimated relative risks and their standard errors for each value of ozone or $PM_{2.5}$. The results are in Figure 10 for ozone and Figure 11 for $PM_{2.5}$. We have not attempted to show the residuals from the individual meteorological analyses since it would be too cumbersome to try to display results from eight air basins on one plot.

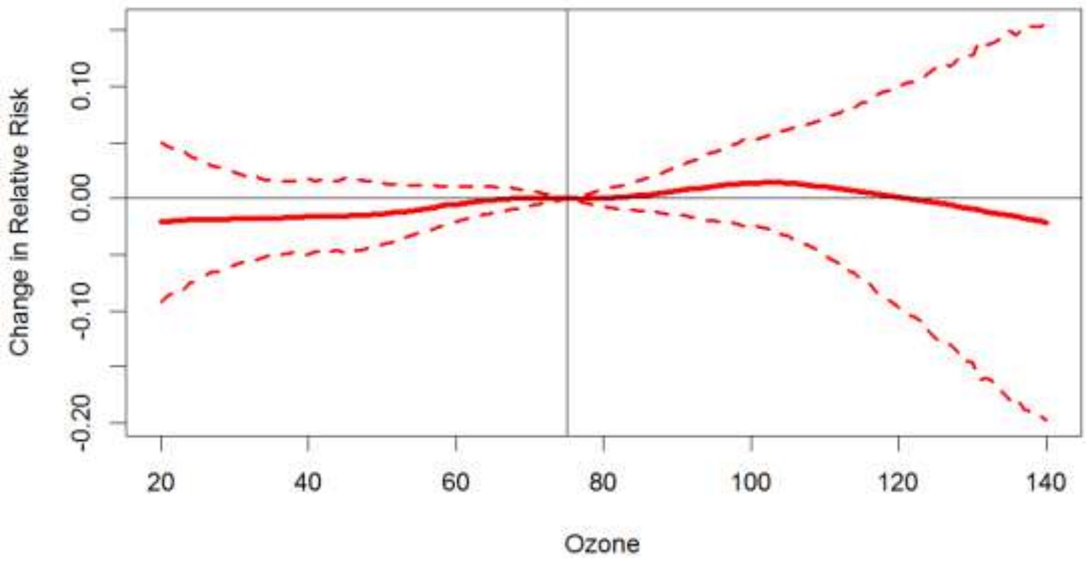

Figure 10: Combined ozone result for eight California air basins.

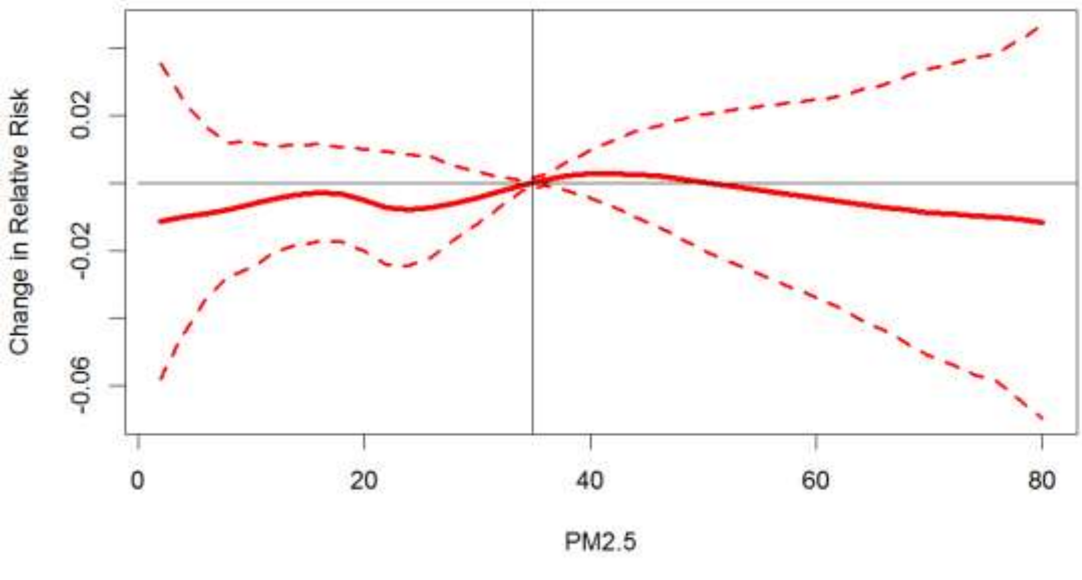

Figure 11: Combined $PM_{2.5}$ result for eight California air basins.

In each case, the estimated relative risk curve is slightly increasing with increasing values of the pollution variable (up to about 100 ppb ozone, or about 40 μg/m^3 $PM_{2.5}$), but the wide confidence intervals cast clear doubt on the statistical significance of that relationship. Overall, the results reinforce our earlier claims that there is no systematic evidence of a mortality association with either ozone or $PM_{2.5}$ in California.

Let us now relate these to national results based again on the NMMAPS dataset. Apart from a brief discussion in Smith et al. (2009), the main national nonlinear result for ozone is that of Bell et al.

(2006), but this suffers from several limitations: it uses 24-hour ozone averages rather than 8-hour daily maxima; they only calculate the "national" curve without any regional results; and it is based on the average ozone level at lags 0 and 1 without any "distributed lag" component. The present analysis uses the nonlinear distributed lag model (3), based on the full range of lags from 0 through 6 days, and is based on 8-hour daily maxima to be consistent with the measure used for the ozone standard. The meteorological variables are daily maximum temperature, daily minimum temperature and daily average relative humidity, as throughout this paper. As well as the "national" average, results are computed for each of the seven NMMAPS regions as used in previous NMMAPS papers such as Dominici et al. (2003) and Smith et al. (2009).

Results are shown in Figure 12. For five of the regions (North West, Upper Midwest, South East, Southern California, South West) there is no consistent relationship between ozone and mortality when the width of the confidence bands is taken into account. For the North East and Industrial Midwest, there is a steady increase of relative risk with increasing level of ozone though the upper confidence band is at or slightly above zero throughout the 20-75 ppb range. Only when all the results are combined into a single "national" curve do we see an increasing ozone-mortality relationship that is clearly statistically significant as judged by the confidence bands.

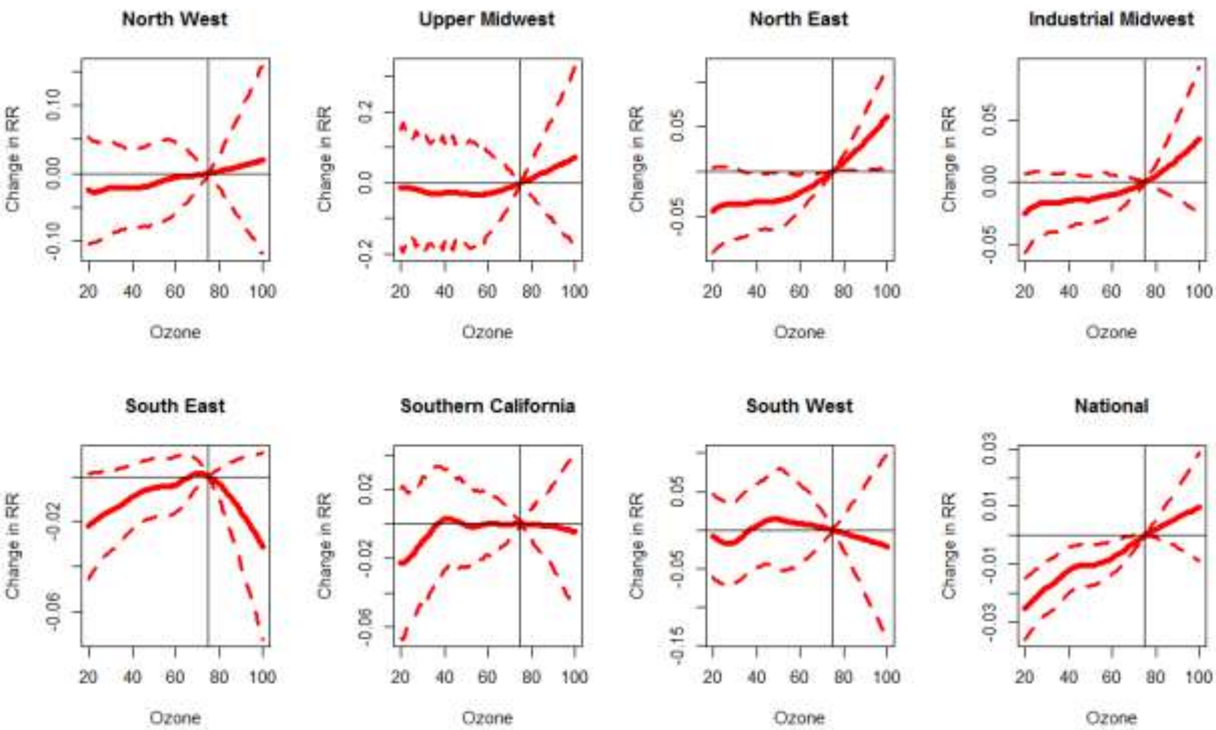

Figure 12: Combined ozone result for the NMMAPS dataset in each of seven regions of the US and nationally.

We now examine the role of a possible PM_{10} confounding effect.

As noted by Bell et al. (2007), only about 25% of all days in the NMMAPS dataset have both ozone and PM_{10} data (and for $PM_{2.5}$, the proportion is far smaller, since regular $PM_{2.5}$ monitoring only started in 1999). The present analysis is based on the nonlinear distributed-lag ozone-mortality response function using lags 0 and 1 for ozone. A day is defined as "available" for the ozone-mortality analysis only if both lags 0 and 1 of ozone are available for that day. Among all available days, about 30% also have PM_{10} available for the same day (aggregated over all days in all cities). However, that 30% rises to about 57% if we allow for PM_{10} at any of lags 0, 1 and 2 (only one of the three days being required to make the day available for analysis). We have followed that strategy here, and have defined the PM_{10} variable used for analysis as the average of available values across all three lags. This strategy has the advantage of increasing the number of days for which both ozone and PM_{10} are available, but the use of a more flexible definition of PM_{10} could also have the disadvantage of not fully capturing the confounding effect of PM_{10} if indeed there is one. Another aspect of the current analysis is that we only allow PM_{10} to enter linearly in the regression equation, whereas we are still assuming a nonlinear distributed lag model for ozone. For this reason as well, the analysis may not be fully capturing the confounding effect of PM_{10}.

Figure 13 shows the result of this analysis, when the resulting ozone-mortality curves are combined nationally using the same hierarchical analyses as for Figure 12.The result does not show any notable contrast between the two curves: in fact the blue curve (that does include the PM_{10} effect) shows a slightly stronger increase in risk across the range of ozone below 75 ppb than the red curve (without PM_{10}). While the result is far from definitive, this strengthens the argument that PM_{10} is most likely not a confounder in the ozone-mortality relationship.

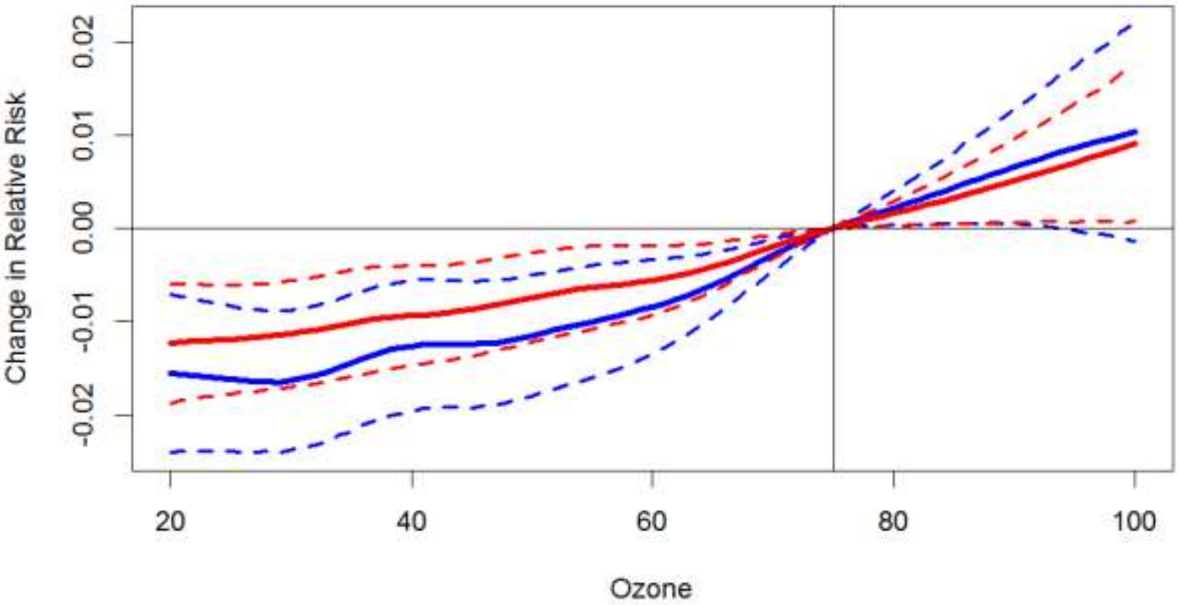

Figure 13: Red curves: Nonlinear relative risk curve (with pointwise 95% confidence bands) for national ozone effect based on (0,1) lags. Blue curves: corresponding curves calculated taking PM_{10} into account for all days where PM_{10} is available for at least one of lags 0, 1 or 2.

Another kind of analysis, however, is an effect modifier analysis, in which we examine the possibility that the ozone-mortality relationship might itself vary according to the underlying level of PM_{10}. The motivation for this is provided by a result in Smith et al. (2009), that included the following result (Table 2). For the all-year analysis and the standard (tlnise) method of combining results across stations, the linear ozone-mortality coefficient, restricted to days in which PM_{10} is below the median for each city, is 0.070 (posterior standard deviation 0.206). The same coefficient, restricted to days in which PM_{10} is above the median for each city, is 0.271 (posterior standard deviation 0.139). In other words, the ozone coefficient is statistically significant for high PM_{10} but not for low PM_{10}. One possible interpretation of that result is that it might be more efficient to control for PM_{10} than ozone.

A similar analysis was conducted, again using the NMMAPS data, but by the analysis methods of the present note, in particular, the nonlinear distributed lag model in ozone (for lags 0 and 1). The results are dependent on various modeling choices including the assumed threshold value of PM_{10}. One typical result is shown in Figure 14, where the PM_{10} threshold was taken as 25 µg/m^3., close to the median over the entire database. Note that both risk curves are (as in earlier plots) standardized to the risk level at 75 ppb: our objective is to study possible changes in the shape of the relative risk curves, not their level relative to one another.

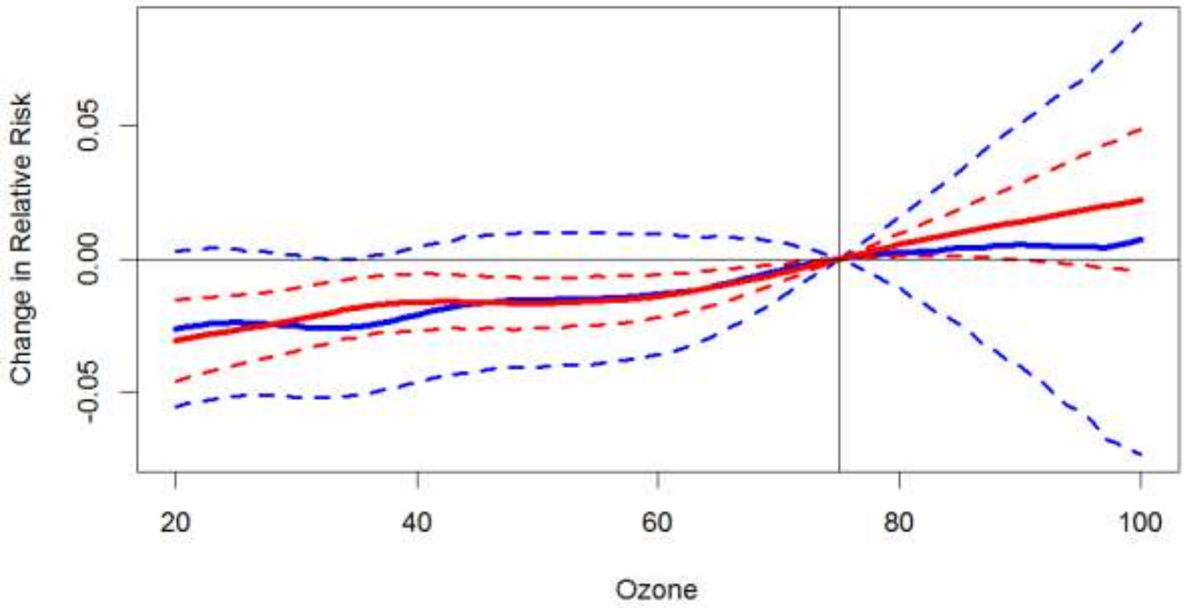

Figure 14: Red curves: Nonlinear relative risk curve (with pointwise 95% confidence bands) for the national ozone effect based on (0,1) lags, restricted to days with PM_{10} above 25 µg/m^3. Blue curves: same but for PM_{10} below 25 µg/m^3.

The result is again somewhat ambiguous. A strong "effect modifier" result would arise if there was a clear separation between the red and blue curves, which there is not. On the other hand, the width of the confidence bands for the blue curve (below the PM_{10} threshold) is substantially wider than for the red curve (above the PM_{10} threshold), implying a lack of statistical significance in this case. Overall, the result reinforces the difficulty of deriving any substantive results that go beyond the single-pollutant nationally averaged concentration-response curve.

Discussion

The first objective of this commentary was to examine the influence of ozone and $PM_{2.5}$ on mortality in eight California air basins, 2000-2012. The method of analysis is one that has become standard in studies of this nature. Daily death counts are fitted using a generalized linear model with log link and a "quasipoisson" distributional form. That model assumes the same mean-variance structure as the Poisson distribution but with one additional parameter representing overdispersion. In the analyses of this paper, the overdispersion parameter is typically of the order of 1.05 to 1.07, indicating only a slight departure from the Poisson distribution. However, the standard errors are slightly larger when the overdispersion parameter is included and this is therefore recommended as a conservative approach.

The covariates included in the analysis are long-term trend, day of week and three meteorological variables: daily maximum temperature, daily minimum temperature and relative humidity. The well-known NMMAPS dataset used daily mean temperature and dewpoint as the two meteorological variables, but in other respects the analysis is the same. In particular, for each meteorological variable we have included the current day's value and the average of the three previous days' values, each modeled nonlinearly through a smoothing spline. The degrees of freedom of these smoothing splines, as well as the one involving long-term trend, have been varied to allow us to study the sensitivity to that parameter.

The trickiest aspect of the analysis is in deciding which lag or lags of the air pollution variable to include. Unlike the choice of degrees of freedom in the smoothing spline, the choice of lags for the air pollution analysis does appear to have a substantial effect on the estimated coefficients. In these analyses, we have tried a total of ten different combinations of lags for the single-air basin analyses, and then used the results of that to guide the choice of lags for the combined analyses. Trying a wide variety of different lag structures and only using the one that gives the largest coefficient or the most statistically significant result has a flavor of "data snooping" and could lead to biased results. We have tried to mitigate that effect by only considering results for which the p-value is well under the standard 0.05, but there are not many such cases where even this mild criterion is satisfied.

For ozone, we were unable to find any significant result for the South Coast air basin, and only for San Francisco Bay when the humidity variable was excluded, and with a relatively mild p-value (0.018). When the results are combined across all air basins, there is no effect of ozone on mortality.

For $PM_{2.5}$, there were several statistically significant results in South Coast but with negative coefficients, which does not make sense biologically. In San Francisco Bay, there were some statistically significant results with a positive coefficient, but only for models in which relative humidity was excluded. Given the inconsistent results and the relatively mild p-values associated with them, it seems likely that these results are a case of spurious statistical significance. When combined across all air basins, there is no statistically significant effect.

To establish a direct comparison with the analyses of the NMMAPS dataset in Smith et al (2009), the results of that paper have been re-derived using the exact same statistical method used in the present paper (in particular, the same meteorological variables). Consistent with the results of the present paper, none of the individual-city results for California using the NMMAPS dataset showed a statistically significant effect, nor did the combined result of all California cities. However, the present method of analysis still shows a statistically significant "national" effect when applied to the whole NMMAPS dataset, although smaller than that reported in Smith et al. (2009). The fact that the national estimate is smaller using tmax, tmin and relative humidity as the meteorological variables could imply that these variables do a better job than the original NMMAPS variables (daily mean temperature and dewpoint) of capturing the confounding effect of ozone with meteorology. Moreover, Smith et al (2009) noted spatial variation of the ozone-mortality coefficient across cities, possible explanatory variables being the percentage of residences with central air conditioning (high in California) and the use of public transportation (high use of public transportation could correspond to high exposure to ambient pollutants, but use of public transportation is generally low in California). Because of this spatial variation, Smith et al questioned whether it made sense to compute a "national" estimate in the face of clear evidence that the effect is not, in fact, national in scope. The results of the present paper in no way contradict that discussion, and indeed strengthen it by confirming that California data for 2000-2012, most of which lies after the end of the NMMAPS data period, still does not show any statistically significant relationship between mortality and either ozone or $PM_{2.5}$.

These analyses have been extended to allow a nonlinear concentration-response curve through the nonlinear distributed lag model. When restricted to the California dataset from 2000-2012, the results are similar to the linear analyses, and show little evidence of mortality dependence on either ozone or PM2.5. When the same methods are applied the NMMAPS dataset, we obtain new results that considerably extend those in previous papers such as Bell et al. (2006), Smith et al. (2009). Results averaged by region do not show a consistent pattern: only for the nationally averaged result is there a clear increasing mortality risk with ozone. We have also examined the possible role of confounding or effect modification by PM_{10}. While the analyses do not provide support that either confounding or effect modification has an important role on the results, they also highlight the limitations of the data for providing further insight beyond the simple single-pollutant analyses.

References

Bell ML, McDermott A, Zeger SL, Samet JM, Dominici F (2004) Ozone and Short-term Mortality in 95 US Urban Communities, 1987-2000, *Journal of the American Medical Association*, **292**(19):2372-2378.

Bell, ML, Kim, JY and Dominici, F (2007), Potential confounding of particulate matter on the short-term association between ozone and mortality in multisite time-series studies. *Environmental Health Perspectives* **115**, 1591-1595.

Bell ML, Peng RD, Dominici F (2006) The exposure-response curve for ozone and risk of mortality and the adequacy of current ozone regulations. *Environmental Health Perspectives* **114**, 532-536.

Belsley DA, Kuh E and Welsch RE (1980), *Regression Diagnostics*. New York: Wiley.

Dominici F, McDermott A, Daniels M, Zeger SL and Samet JM (2003), Mortality among residents of 90 cities. In *Revised Analyses of the National Morbidity, Mortality and Air Pollution Study, Part II*. Health Effects Institute, Cambridge, MA, pp. 9--24.

Everson PJ and Morris, CN (2000), Inference for multivariate normal hierarchical models. *Journal of the Royal Statistical Society B*, **62**, 399-412.
Gasparrini A, Armstrong B and Kenward MG (2010), Distributed lag nonlinear models. *Statistics in Medicine* **29**, 2224-2234.

Lopiano KK, Smith RL and Young SS (2015), Air quality and acute deaths in California, 2000-2012. In preparation for journal submission.

R Core Team (2014), *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org.

Smith RL, Xu B, Switzer P. (2009) Reassessing the relationship between ozone and short-term mortality in U.S. urban communities. *Inhal Toxicol* **29 (S2)**, 37–61.

Acknowledgements

This work was financially supported by the American Petroleum Institute. I am grateful to Stan Young for numerous conversations and for access to the California dataset.