# *TIME SERIES ANALYSIS OF OZONE AND MORTALITY IN CALIFORNIA, 1987-2012*

**Richard L Smith**

**University of North Carolina, Chapel Hill**

Environmental Vision Conference

Marriott Marquis Hotel, Washington DC

May 10, 2016

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

ENVISION

# *About me…*

- Professor of Statistics and Biostatistics (at UNC since 1991)
- Director of the Statistical and Applied Mathematical Sciences Institute (SAMSI), a National Science Foundation Institute
- Research interests in environmental statistics, including climate change, climate and weather extremes, modeling air pollution and health effects of air pollution
- Member of Research Committee of the Health Effects Institute, 2010-2015
- Former member of American Statistical Association Committee on Climate Change Policy and Statistics and participant in “Climate Science Days” organized by AAAS

# Citations, Acknowledgements, Disclaimers…

This talk is largely based on:

1. RLS public comment on the 2015 ozone standard, submitted March 17, 2015 (www.unc.edu/~rls and follow "Preprints" tab)
2. "Air Quality and Acute Deaths in California, 2000-2012" by S. Young, R. Smith and K. Lopiano, under revision for *Regulatory Toxicology and Pharmacology*

See also www.unc.edu/~rls/CApollution.html for data and programs.

Thanks for Stan Young for compiling the data. This work has been supported by the American Petroleum Institute. All views expressed are those of the author and do not represent the official position of API, the University of North Carolina, NSF, or any other organization.

# *Background: the NMMAPS study (Bell et al., 2004, 2006; Smith et al. 2009)*

- Daily death data in 108 US cities, 1987-2000
- Each city analyzed for ozone-mortality relationship, correcting for seasonality and long-term trends, day of week effects, meteorology and (in some analyses) co-pollutants
- Coefficients combined across cities through a two-stage hierarchical model
- Results appear to show a tight “national effect,” but this does not fully reflect the very wide variation in the individual-city effects or the spatial variability of the apparent ozone-mortality relationship
- Further analyses included nonlinear concentration-response relationships, and analysis of spatial variations through effect modifiers.

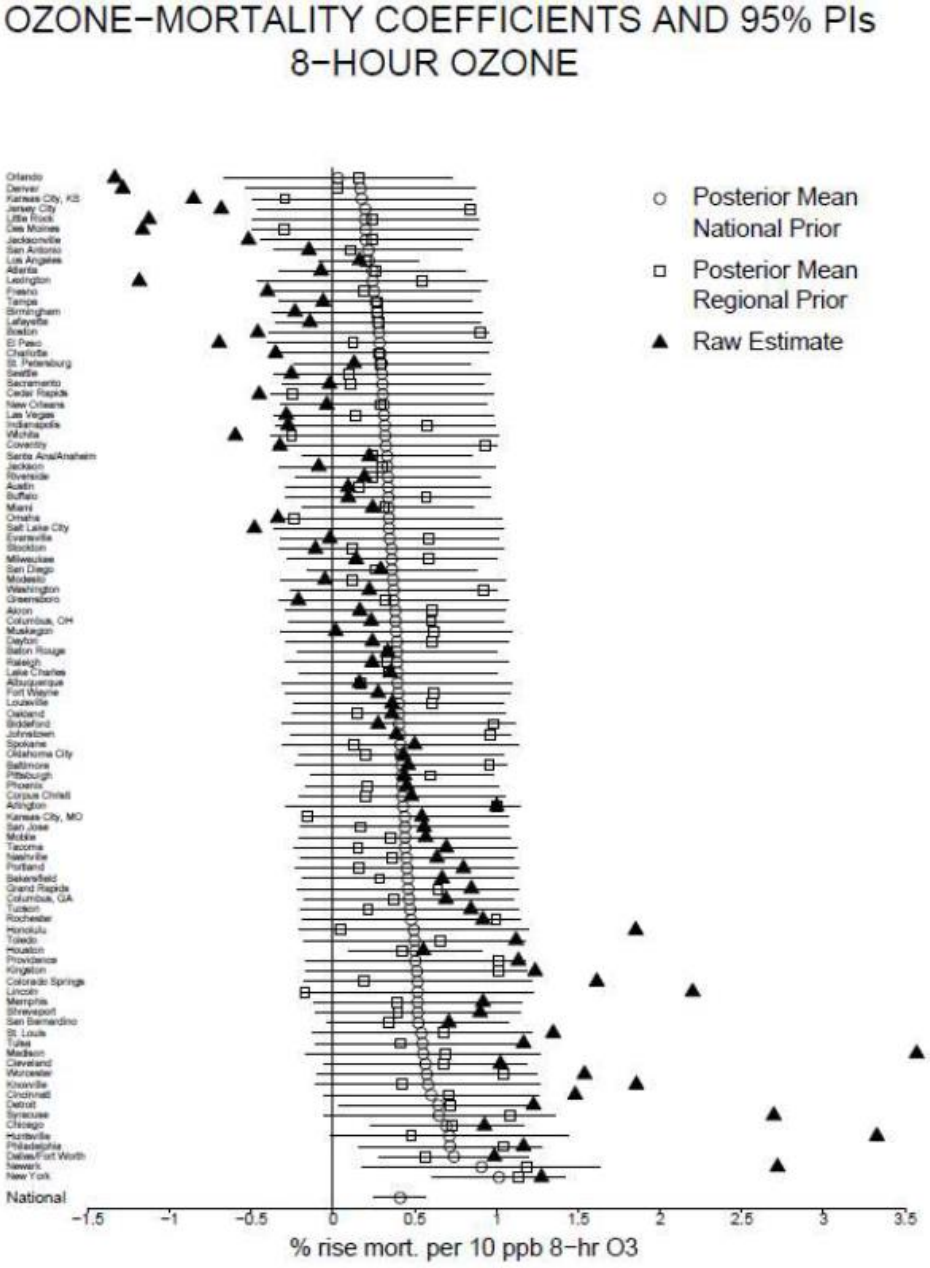


Figure 1 (Figure 4 of Smith et al., 2009). Raw city-specific estimates (triangles), posterior estimates under the "regional" prior (squares) and posterior estimates (circles) and 95% posterior intervals under the "national" prior, for the ozone-mortality coefficient based on 8-hour ozone.

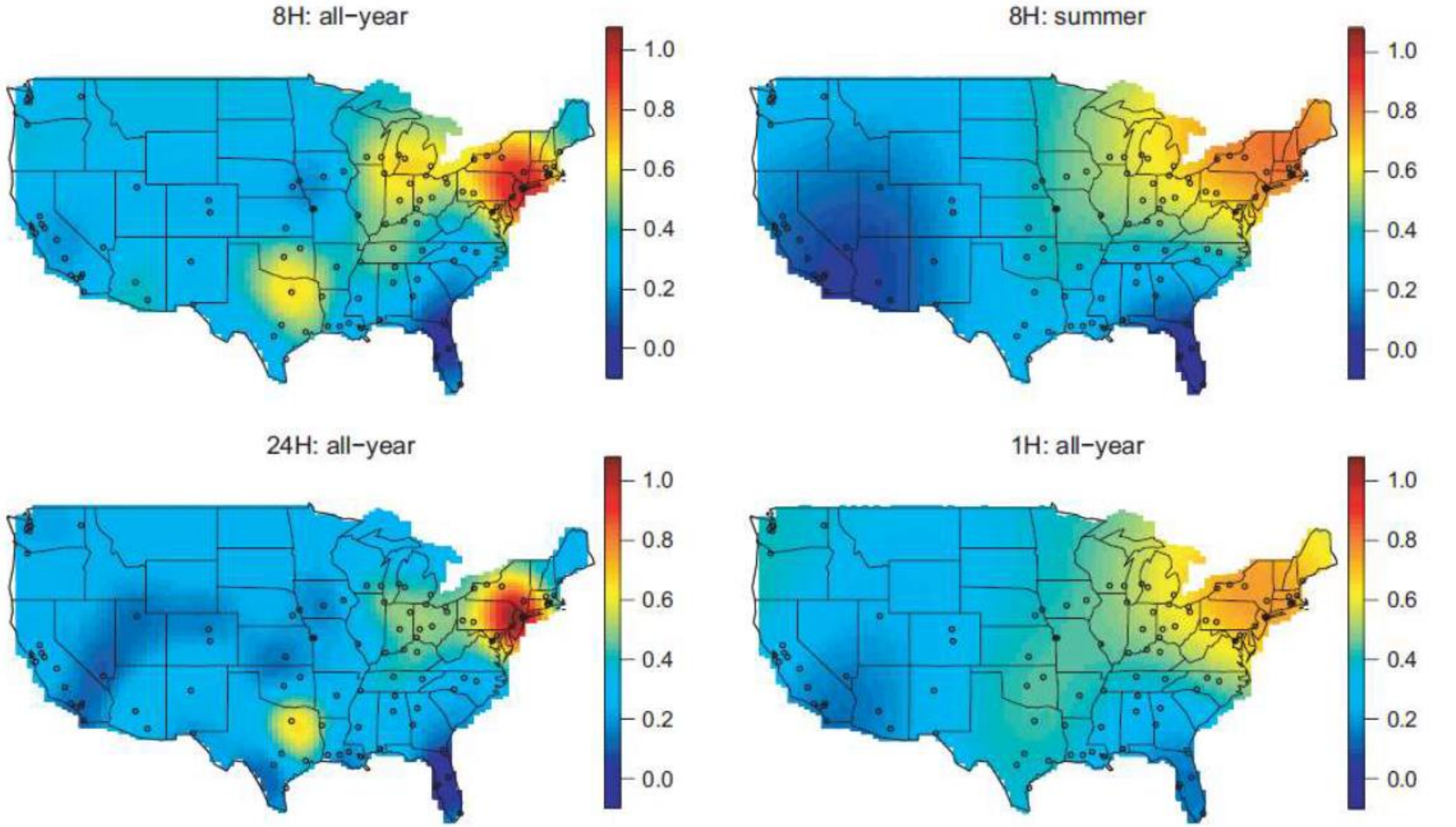


Figure 2 (Figure 6 of Smith et al., 2009). Map of spatially dependent ozone-mortality coefficients for 8-hour ozone (all-year data), 8-hour ozone (summer data), 24-hour ozone (all-year data) and 1-hour ozone (all-year data).

# *New Dataset*

- Daily mortality (by age group and type of death) for 2000-2012, eight air basins in California
- Air quality: daily max 8-hour average ozone, daily average PM2.5
- Meteorology: daily max and min temperature, daily max relative humidity
- Data and programs online at:

[www.unc.edu/~rls/CApollution.html](www.unc.edu/~rls/CApollution.html)

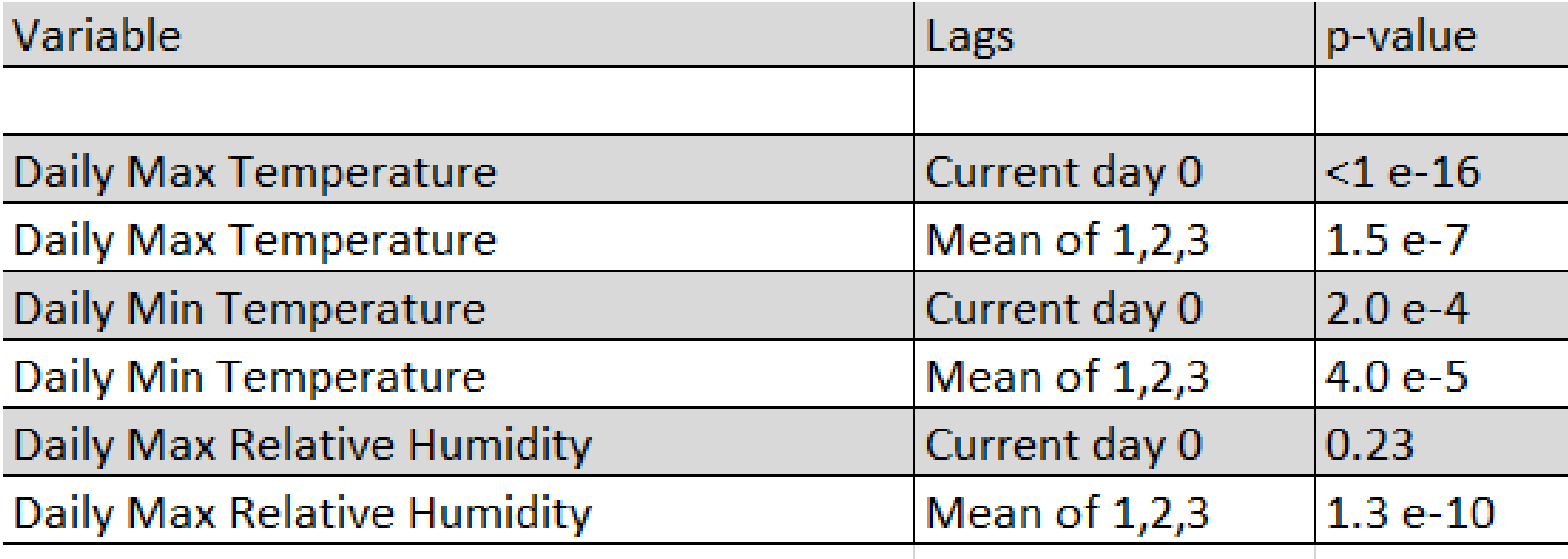
# *Analysis Strategy*

Log mean daily death count =
Nonlinear function of date[1] +
Day of week effect +
Nonlinear function of metorology[2] +
Linear distributed lag model for ozone[3]

1. Splines with 7 df per year
2. Additive spline (3-6 df) representation for daily max temp, daily min temp and daily max RH, plus averages lagged 1-3 days of same variables
3. Separate terms for ozone in lags $l_1,l_2,\ldots,l_k$, result expressed as overall net effect

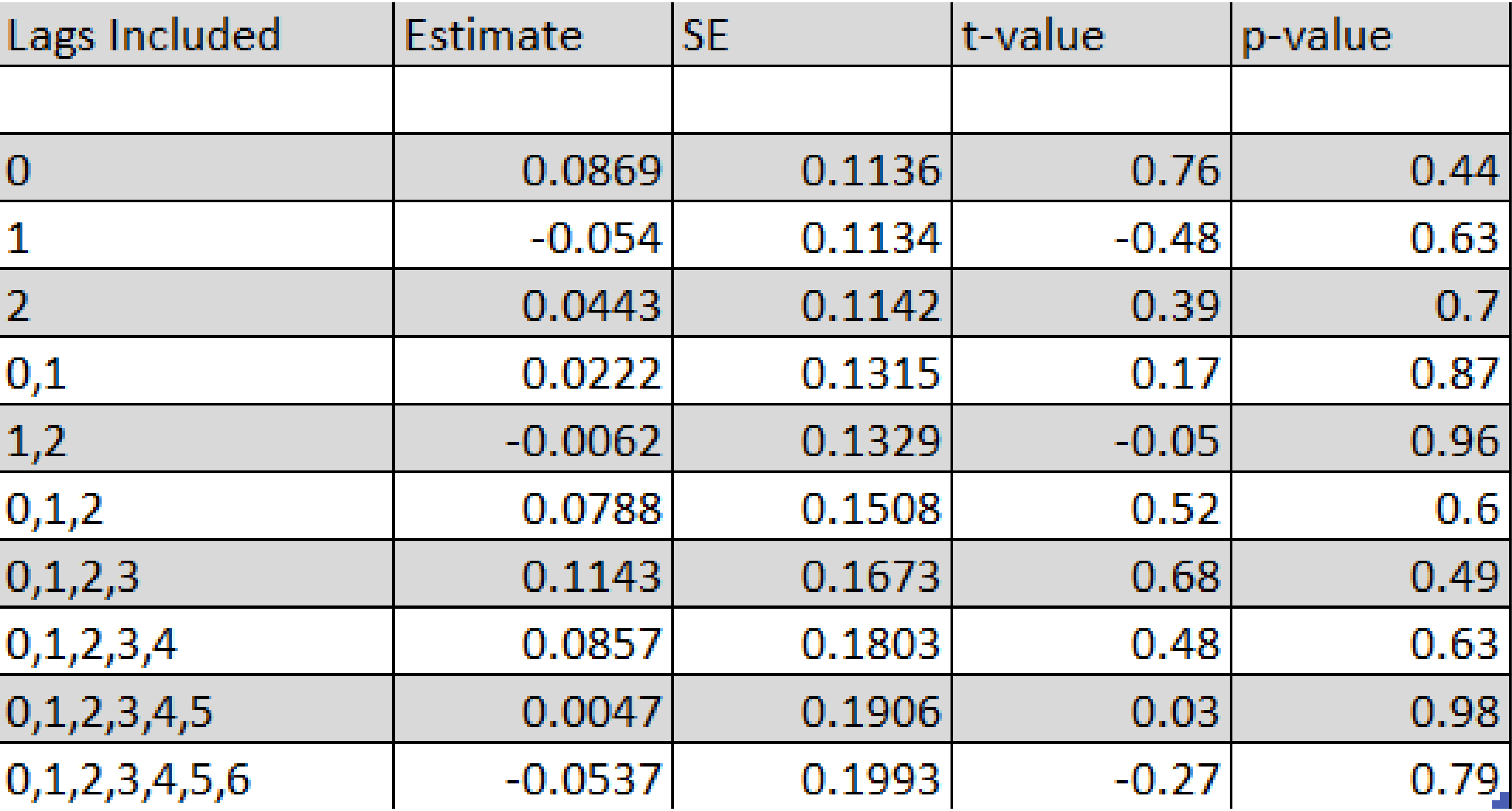
# *Extensions/Comments*

1. Nonlinear concentration-response (C-R) curves: introduce a “nonlinear distributed lag” model in which the net effect is a nonlinear function of ozone
2. Combine results across cities or air basins (similar to NMMAPS analysis)
3. Comparisons with NMMAPS analysis strategy:
    a. Different meteorological variables (NMMAPS used daily mean temperature and dewpoint plus their lagged 1-3 days average values)
    b. No age group x long-term trend terms



| Variable | Lags | p-value |
|---|---|---|
| | | |
| Daily Max Temperature | Current day 0 | <1 e-16 |
| Daily Max Temperature | Mean of 1,2,3 | 1.5 e-7 |
| Daily Min Temperature | Current day 0 | 2.0 e-4 |
| Daily Min Temperature | Mean of 1,2,3 | 4.0 e-5 |
| Daily Max Relative Humidity | Current day 0 | 0.23 |
| Daily Max Relative Humidity | Mean of 1,2,3 | 1.3 e-10 |

**Table. S1:** 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.

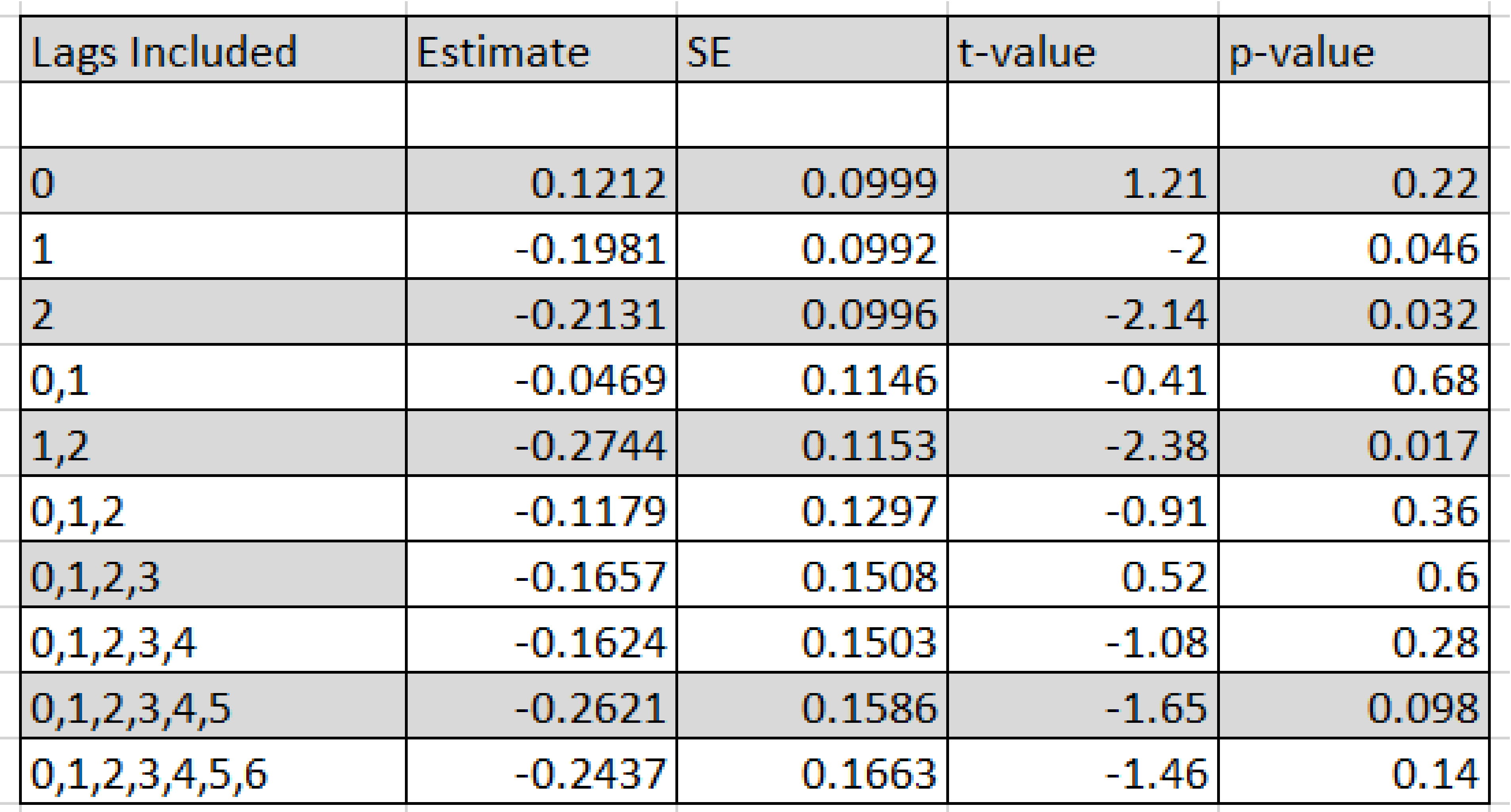


| Lags Included | Estimate | SE | t-value | p-value |
|---|---|---|---|---|
| | | | | |
| 0 | 0.0869 | 0.1136 | 0.76 | 0.44 |
| 1 | -0.054 | 0.1134 | -0.48 | 0.63 |
| 2 | 0.0443 | 0.1142 | 0.39 | 0.7 |
| 0,1 | 0.0222 | 0.1315 | 0.17 | 0.87 |
| 1,2 | -0.0062 | 0.1329 | -0.05 | 0.96 |
| 0,1,2 | 0.0788 | 0.1508 | 0.52 | 0.6 |
| 0,1,2,3 | 0.1143 | 0.1673 | 0.68 | 0.49 |
| 0,1,2,3,4 | 0.0857 | 0.1803 | 0.48 | 0.63 |
| 0,1,2,3,4,5 | 0.0047 | 0.1906 | 0.03 | 0.98 |
| 0,1,2,3,4,5,6 | -0.0537 | 0.1993 | -0.27 | 0.79 |

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.

| 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.



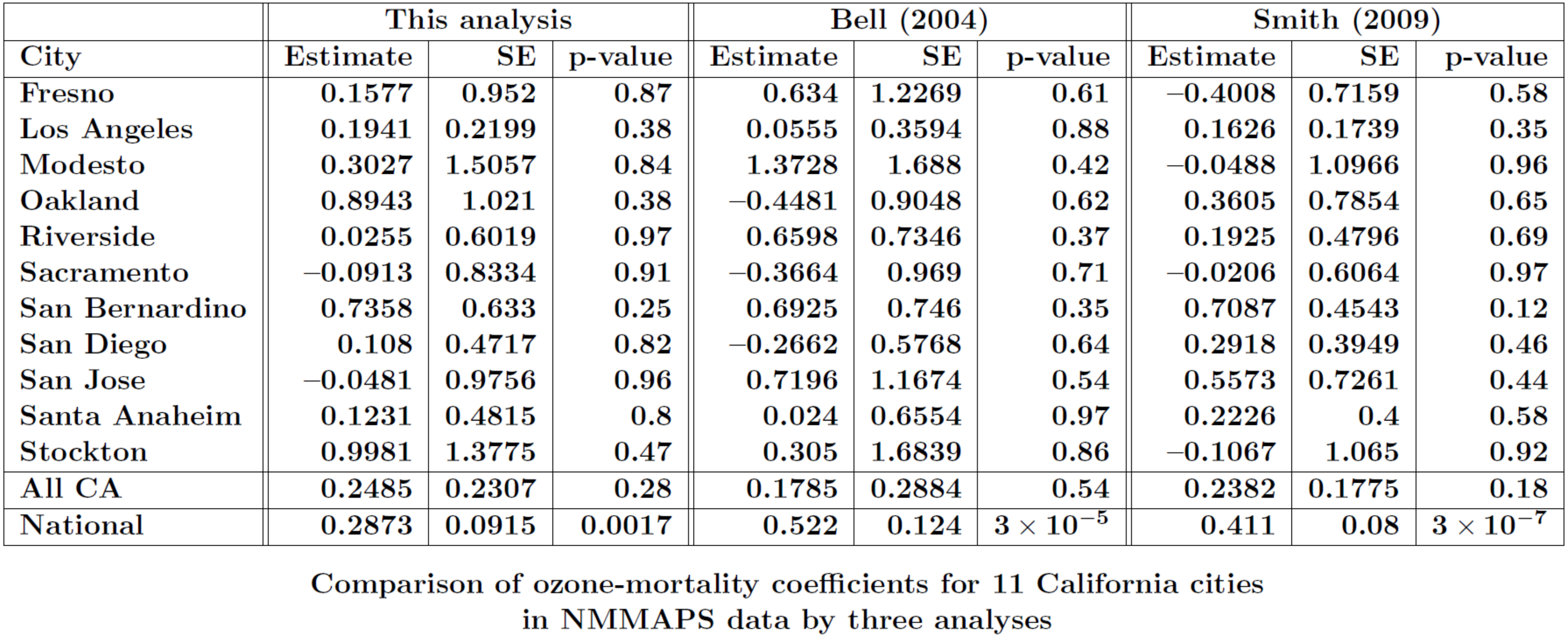
| Lags Included | Estimate | SE | t-value | p-value |
|---|---|---|---|---|
| | | | | |
| 0 | 0.1212 | 0.0999 | 1.21 | 0.22 |
| 1 | -0.1981 | 0.0992 | -2 | 0.046 |
| 2 | -0.2131 | 0.0996 | -2.14 | 0.032 |
| 0,1 | -0.0469 | 0.1146 | -0.41 | 0.68 |
| 1,2 | -0.2744 | 0.1153 | -2.38 | 0.017 |
| 0,1,2 | -0.1179 | 0.1297 | -0.91 | 0.36 |
| 0,1,2,3 | -0.1657 | 0.1508 | 0.52 | 0.6 |
| 0,1,2,3,4 | -0.1624 | 0.1503 | -1.08 | 0.28 |
| 0,1,2,3,4,5 | -0.2621 | 0.1586 | -1.65 | 0.098 |
| 0,1,2,3,4,5,6 | -0.2437 | 0.1663 | -1.46 | 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 $\mu g/m^3$ rise in $PM_{2.5}$. South Coast air basin; response variable is non-accidental mortality aged 65 and over.

| 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 4:** Combined results across all eight air basins.

# Comparison with NMMAPS data (1987-2000)



| | This analysis | | | Bell (2004) | | | Smith (2009) | | |
|---|---|---|---|---|---|---|---|---|---|
| City | Estimate | SE | p-value | Estimate | SE | p-value | Estimate | SE | p-value |
| Fresno | 0.1577 | 0.952 | 0.87 | 0.634 | 1.2269 | 0.61 | −0.4008 | 0.7159 | 0.58 |
| Los Angeles | 0.1941 | 0.2199 | 0.38 | 0.0555 | 0.3594 | 0.88 | 0.1626 | 0.1739 | 0.35 |
| Modesto | 0.3027 | 1.5057 | 0.84 | 1.3728 | 1.688 | 0.42 | −0.0488 | 1.0966 | 0.96 |
| Oakland | 0.8943 | 1.021 | 0.38 | −0.4481 | 0.9048 | 0.62 | 0.3605 | 0.7854 | 0.65 |
| Riverside | 0.0255 | 0.6019 | 0.97 | 0.6598 | 0.7346 | 0.37 | 0.1925 | 0.4796 | 0.69 |
| Sacramento | −0.0913 | 0.8334 | 0.91 | −0.3664 | 0.969 | 0.71 | −0.0206 | 0.6064 | 0.97 |
| San Bernardino | 0.7358 | 0.633 | 0.25 | 0.6925 | 0.746 | 0.35 | 0.7087 | 0.4543 | 0.12 |
| San Diego | 0.108 | 0.4717 | 0.82 | −0.2662 | 0.5768 | 0.64 | 0.2918 | 0.3949 | 0.46 |
| San Jose | −0.0481 | 0.9756 | 0.96 | 0.7196 | 1.1674 | 0.54 | 0.5573 | 0.7261 | 0.44 |
| Santa Anaheim | 0.1231 | 0.4815 | 0.8 | 0.024 | 0.6554 | 0.97 | 0.2226 | 0.4 | 0.58 |
| Stockton | 0.9981 | 1.3775 | 0.47 | 0.305 | 1.6839 | 0.86 | −0.1067 | 1.065 | 0.92 |
| All CA | 0.2485 | 0.2307 | 0.28 | 0.1785 | 0.2884 | 0.54 | 0.2382 | 0.1775 | 0.18 |
| National | 0.2873 | 0.0915 | 0.0017 | 0.522 | 0.124 | $3 \times 10^{-5}$ | 0.411 | 0.08 | $3 \times 10^{-7}$ |

Comparison of ozone-mortality coefficients for 11 California cities in NMMAPS data by three analyses

# *Nonlinear C-R Curves*

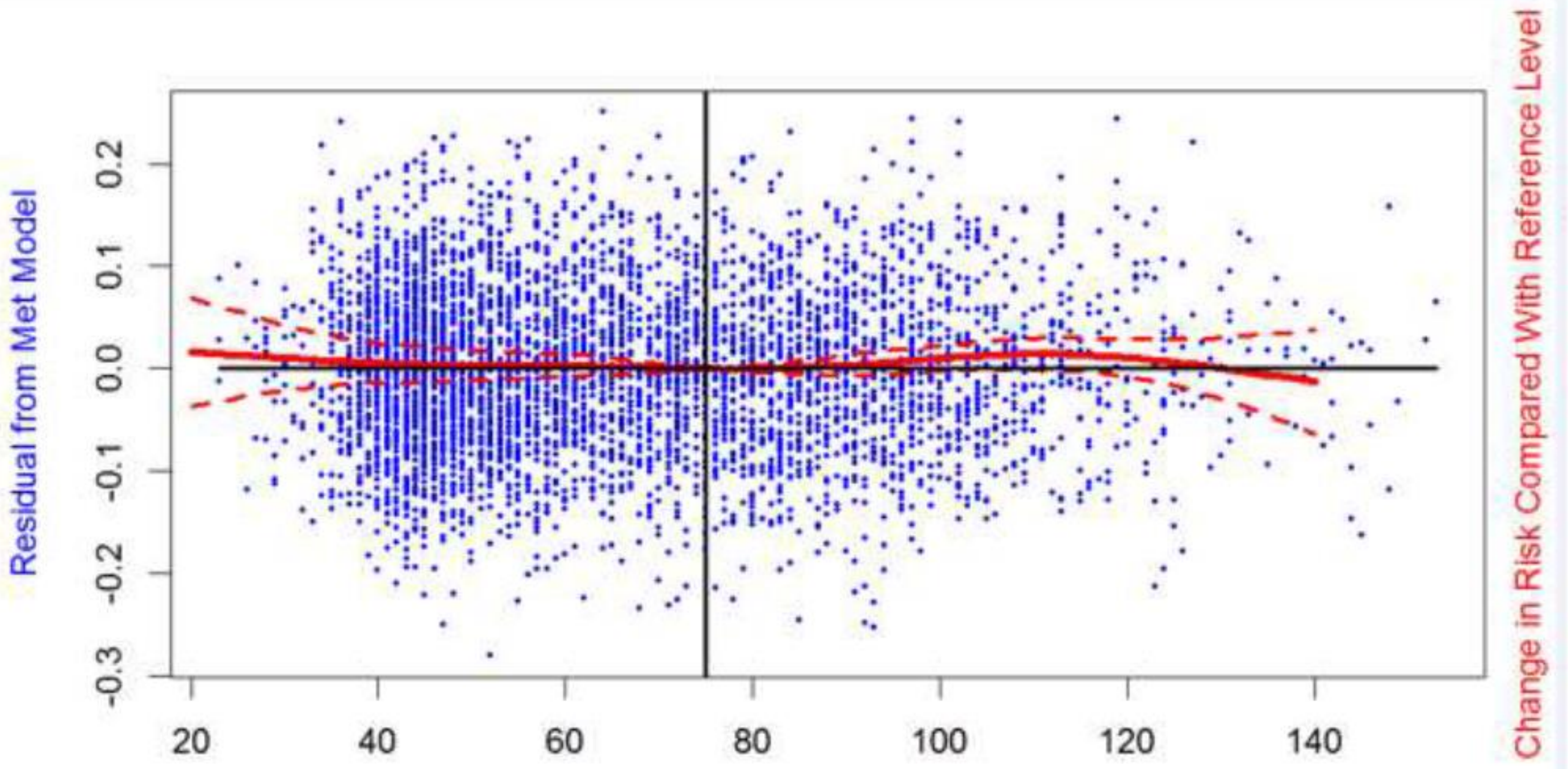
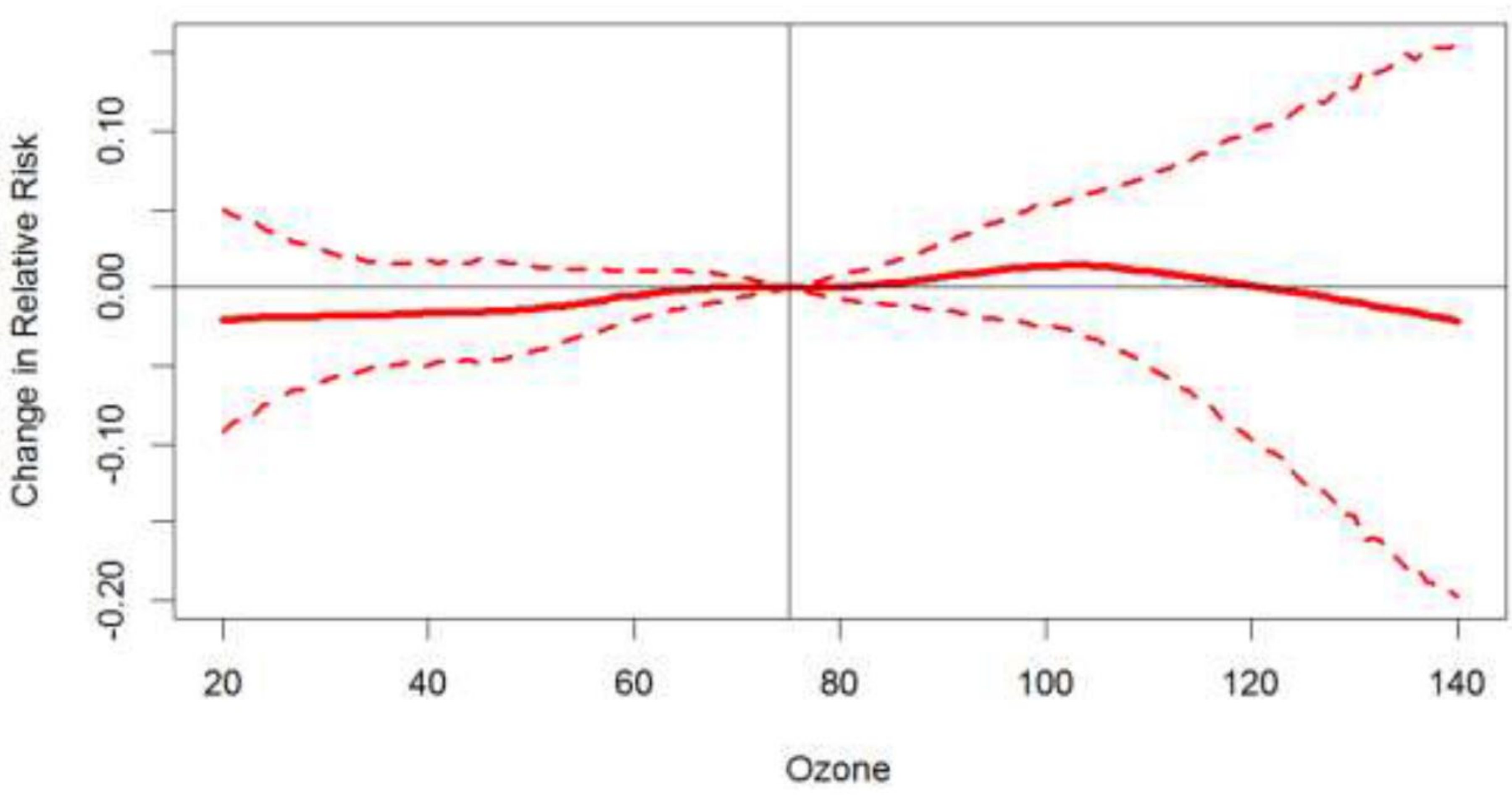


Fig. 4. Nonlinear dependence of mortality on ozone for South Coast air basin. Blue dots: residuals from the model that includes long-term trends, day of week and meteorology, plotted against the air pollution variable (ozone). Red solid and dashed curves: implied change of relative risk with respect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

# Nonlinear C-R Curves



Combined ozone result for eight California air basins.

# Nonlinear C-R Curves

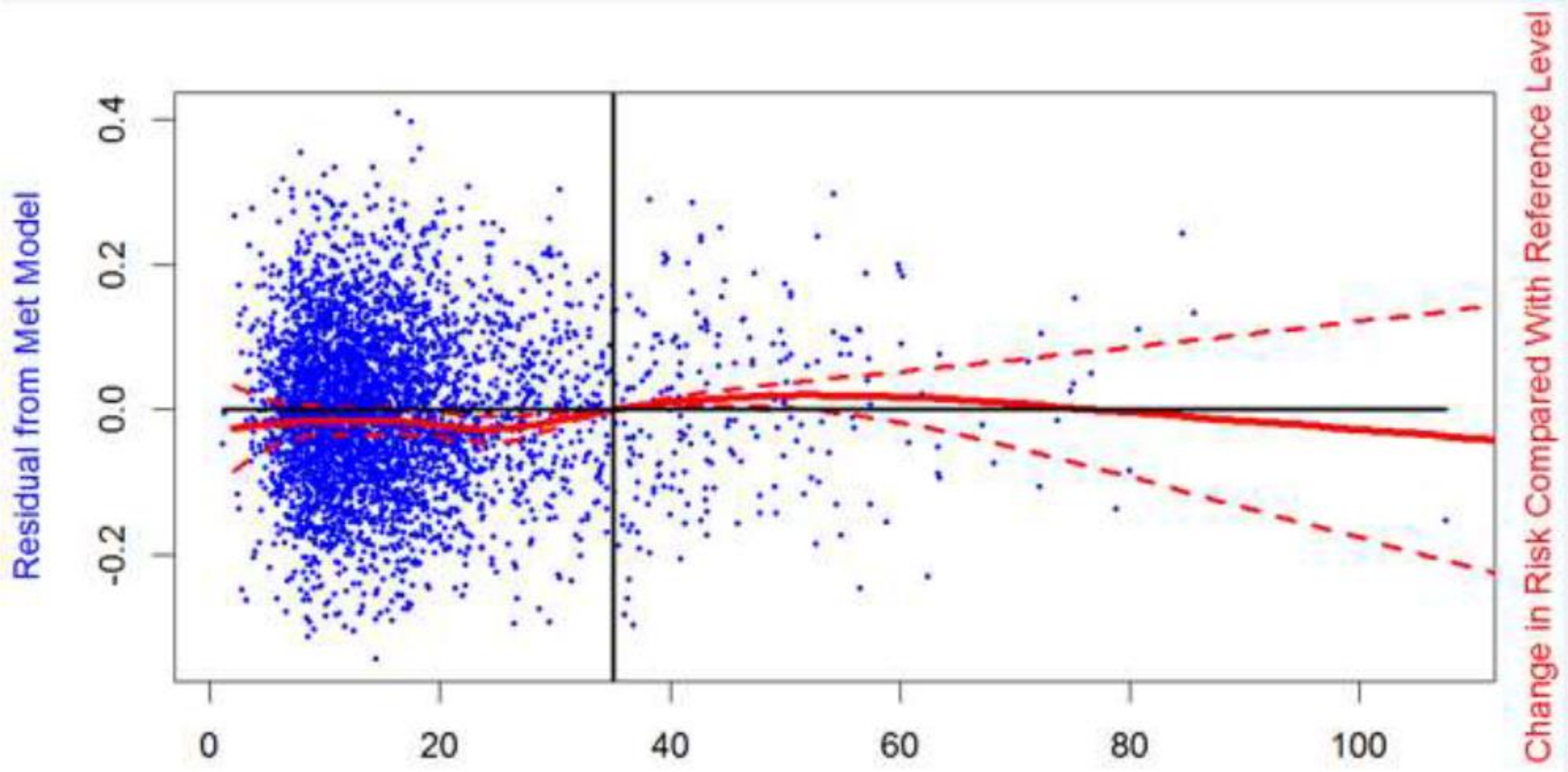
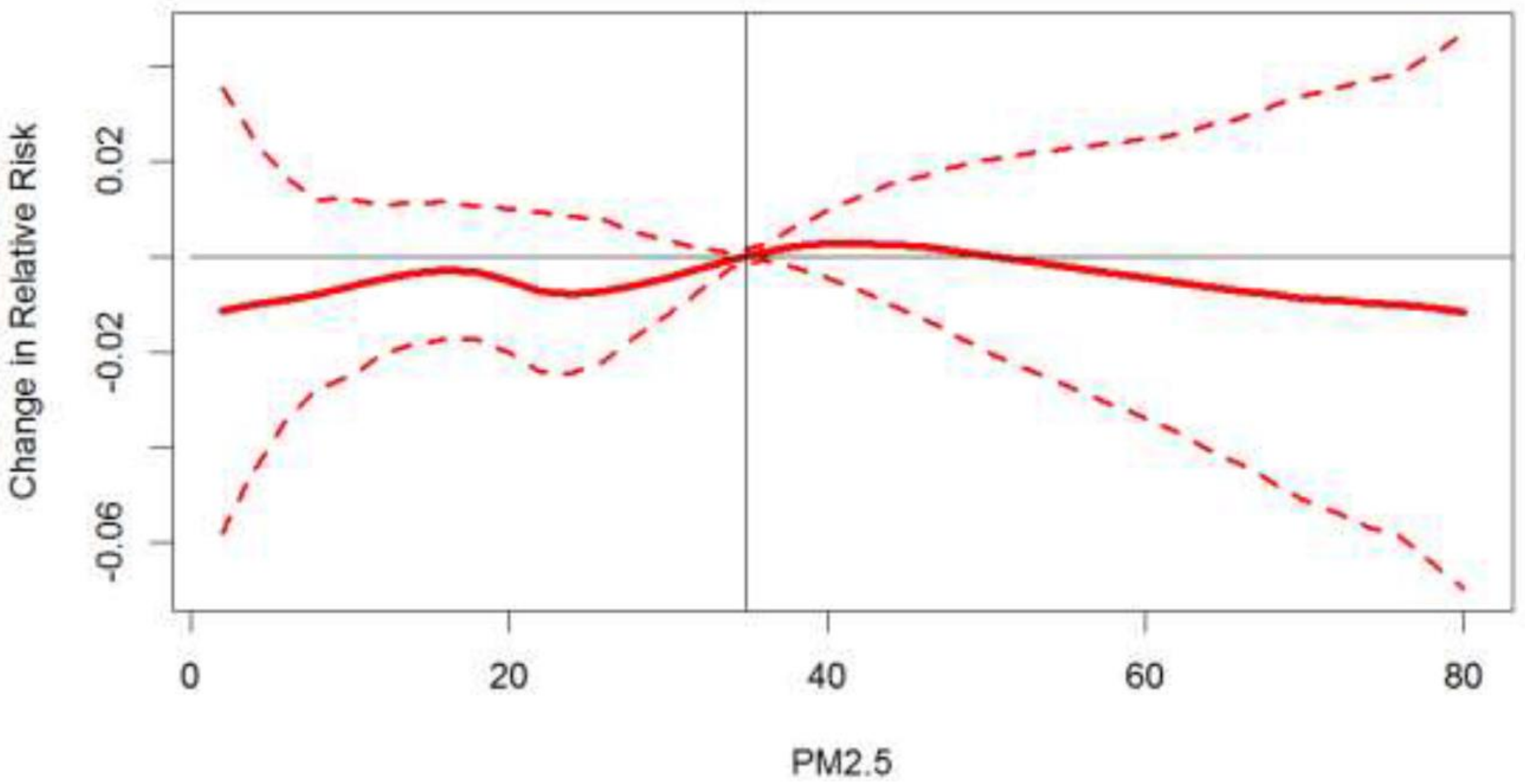


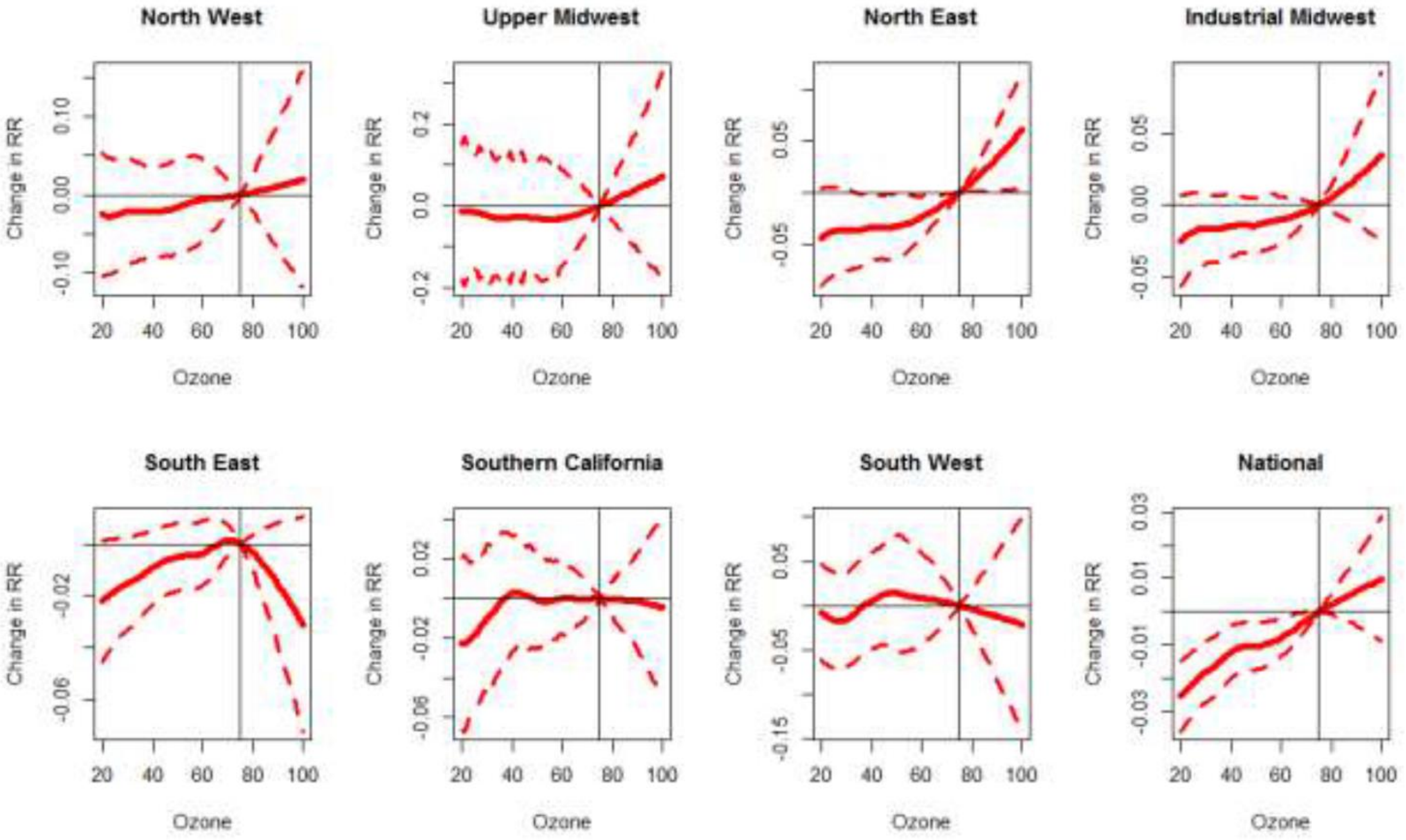
Fig 5. Nonlinear dependence of mortality on $PM_{2.5}$ for San Francisco Bay air basin. Analogous to Figure S4, using the full meteorological model (including relative humidity), and a nonlinear model for the relationship between $PM_{2.5}$ and mortality. The relative risk was computed with respect to a reference level of 35 μg/m$^3$, the current standard for daily max of $PM_{2.5}$.

# Nonlinear C-R Curves



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

# Nonlinear C-R Curves for Ozone (all US)



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

# *Summary and Conclusions*

1. For the CA 2000-2012 dataset:
    a. Very strong effect due to meteorology (especially daily max temp)
    b. No consistent effect due to either ozone or $PM_{2.5}$ for the 8 CA air basins, either individually or combined
    c. The results are robust to different combinations of lags for the distributed lag models, or degrees of freedom for the nonlinear spline terms
    d. Nonlinear C-R curves also show no consistent effect
2. Comparisons with NMMAPS 1987-2000 dataset:
    a. When the same methods for ozone are applied to the CA cities in the NMMAPS dataset, we get very similar conclusions. This is true both using the analysis methods of this paper and repeating the original analysis methods for NMMAPS.
    b. When the same methods are applied to the whole of the NMMAPS dataset, we again see "national effects" but only if the only country is treated as homogeneous

# *The Big Picture*

- Large literature on health effects of ozone
- Only a comparatively small part of that literature deals specifically with mortality
- Within that literature, the NMMAPS study has had an outsize influence
- The present study goes part way towards extending NMMAPS by using more recent data (may extend to other states later)
- This study shows very little evidence of an adverse mortality effect in California. Whether other states will show different results remains to be seen.