Sensitivity of the estimated air pollution-respiratory admissions relationship to statistical model choice

Running Title	Effect of model choice in studies of air pollution and respiratory				
	admissions				
Authors	Bircan Erbas				
	Department of Public Health,				
	The University of Melbourne, VIC 3053, Australia.				
	Rob J Hyndman				
	Department of Econometrics and Business Statistics,				
	Monash University, VIC 3800, Australia.				
Corresponding author:	Bircan Erbas				
	Department of Public Health,				
	The University of Melbourne,				
	Level 2, 723 Swanston Street, Carlton, Victoria 3053, Australia.				
	Email: b.erbas@unimelb.edu.au.				
	Telephone: +613 8344 0635. Fax: +613 9349 5815.				

Abstract

The objective of this study is to demonstrate the methodological shortcomings of currently available analytical methods for single-city time series data.

We analyse daily Chronic Obstructive Pulmonary Disease (COPD) and daily asthma hospital admissions in Melbourne, Australia from July 1989 to December 1992. Air pollution data comprise nitrogen dioxide, ozone and sulphur dioxide and air particles index consistent with particulates between 0.1 and 1μ m in aerodynamic diameter. Statistical analyses were performed using generalized linear models, generalized additive models, Poisson autoregressive models and transitional regression models.

The estimated effect of nitrogen dioxide on COPD hospital admissions was similar across the different statistical models, RR = 1.06 (95% CI 1.01 to 1.11). Similarly the estimated effect nitrogen dioxide on asthma hospital admissions was also consistent, RR = 1.05 (95% CI 1.01 to 1.09). However, the effects of ozone, air particles index and sulphur dioxide were highly sensitive to model specification for both COPD and asthma hospital admissions.

In single-city studies of air pollution and respiratory disease, very different conclusions can be drawn from competing models. Furthermore, real time series data have greater complexity than any of the commonly-used existing models allow. Consequently, single city studies should use several statistical models to demonstrate the stability of estimated effects.

INTRODUCTION

Single city epidemiological studies of the adverse effects of air pollution on respiratory morbidity and mortality have been well documented in many countries of diverse populations, geography and climate. In fact, findings from these studies have shown that increases in respiratory admissions and respiratory mortality suggest adverse effects of air pollutants well below the recommended World Health Organization guidelines (Touloumi et al. 1997). This type of study design is one of the most commonly used ecological study designs, particularly in developing countries where there is limited or no pollution or morbidity/mortality data for large-scale multi-city studies of air pollution and respiratory disease (Wong et al. 2001; Joseph et al. 2003; Berktas and Bircan 2004). Epidemiological studies that provide accurate estimates of the effects of ambient air pollutants on mortality and incidence are necessary for revising public health policy. Thus, meaningful interpretations of these air pollution health effects require implementation of appropriate statistical techniques.

APHEA (Air Pollution and Health, a European Approach), one of the largest groups to study the replicability of findings across different countries produced a standard protocol of analysis (Katsouyanni et al. 1996). Despite this, considerable debate remains with regards to the appropriateness of the choice of statistical method used for analysis, particularly when the health and environmental data are time series. A number of studies have focused on aspects of these issues. Lipfert and Wyzga (1995) provide a comprehensive overview of the technical issues with respect to seasonality, confounding and the modeling of multi-pollutant effects, by synthesizing results from US and foreign mortality studies. Samet et al. (2003) expand on these issues by focusing on quantifying dose-response relations while McMichael et al. (1998) discuss issues related to interpretations from these models.

In 2003, The Health Effects Institute (www.healtheffects.org) released the "Revised Analyses of Time-Series Studies of Air Pollution and Health". This report comprised re-analysis of mortality/morbidity single and multi-city data from various countries of diverse cli-

mate and populations originally published as an article or a report. Re-analysis included generalized additive model with strict convergence criterion, GLM-natural splines, with penalized regression splines to assess the sensitivity of the estimated pollutant effects, particularly particulate matter. Although, differences in effect estimates were observed in many single and multi-city studies of mortality, there was very little hospital admissions data. In addition findings were mixed ranging from a substantial change of the estimated effects to little or no change (Commentary on Revised Analyses of Selected Studies, Health Effects Institute, 2003).

We extend this work on single-city studies by examining the robustness of the estimated effects of air pollutants and respiratory morbidity under different statistical models currently used in practice. The length of the time series is 3.5 years and although not ideal, it is important to demonstrate the methodological shortcomings of what is often done in practice particularly since many ecological single-city studies (e.g., Wong et al. 1999; Wong et al. 2001; Kan and Chen 2003; Chen et al. 2004) comprise time series shorter than that used in this study.

A key issue which arises is controlling for seasonal variation in respiratory disease and air pollution. Fourier terms of sine and cosine pairs with varying periods have been accepted as a method to control for seasonal variation in the outcome variable (Hoek **et al**. 1997). However, few studies have controlled for possible seasonality in the pollution and/or weather variables (Schwartz 1993; Kelsall **et al**. 1997; Samet **et al**. 2000). The sensitivity of the observed effects may change with inclusion of confounding effects for seasonality in model specification. Therefore, it is necessary to control for possible confounding that may induce spurious pollution effects.

To assess the strength and magnitude of seasonal variation in pollutants and weather variables, we use a method of seasonal adjustment called STL (Seasonal-Trend decomposition based on Loess smoothing)(Cleveland and Terpenning 1982). Variables exhibiting strong seasonality were adjusted with this method and the resulting seasonally adjusted series were used in subsequent analysis.

We explore the robustness of the pollution-respiratory disease relation using a variety of regression-based approaches, controlling for secular trends, seasonality, and confounding effects of weather variables. These models include: (1) Generalized Linear Models (GLM); (2) Generalized Additive Models (GAM); (3) Parameter Driven Poisson Regression Models (PDM); and (4) Transitional Regression Models (TRM). In each case, we consider models based on a Poisson distribution, incorporating over-dispersion and serial correlation where possible.

The objectives of this study are: (1) to demonstrate that very different conclusions can be drawn from competing models applied to real time series data from a single city; and (2) to show that real time series data have greater complexity than any of the commonly-used existing models allow.

METHODS

Data

COPD (Chronic Obstructive Pulmonary Disease) and asthma hospital admissions from all short-stay acute public hospitals in Melbourne, registered on a daily basis by the Department of Human Services (State Government of Victoria), were used as the outcome variables for the period 1 July 1989 to 31 December 1992. International Classification of Disease (ICD) codes for COPD (490–492, 494, 496) and asthma (493) were used to define COPD and asthma.

Melbourne is the second largest city in Australia, with the main source of air pollution emissions from motor vehicles. In Melbourne, levels of sulfur dioxide are relatively low due to the absence of sulfur-emitting industries. Particulate pollution is highest during autumn and winter, due to the widespread use of wood fires. High levels of nitrogen dioxide (about 65% of total emissions from motor vehicles) and ozone are major constituents of air pollution in this city (EPA 1999).

Air pollution data were obtained from the Environment Protection Authority (EPA). Maximum hourly values were averaged each day across nine monitoring stations in Melbourne, for nitrogen dioxide, sulfur dioxide, and ozone, all measured in parts-perhundred-million (pphm). Particulate matter was measured by a device which detects back-scattering (B_{scat}) of light by visibility-reducing particulates between 0.1 and $1\mu m$ in aerodynamic diameter. Air particles index (API) were derived from $B_{scat} \times 10^{-4}$. Maximum hourly API values were averaged across 9 monitoring stations each day. Weather variables include three hourly maximum daily levels of relative humidity, dry bulbs temperature and dew point temperature. The measures were averaged across four monitoring stations in the Melbourne area.

Statistical Methods

The confounding effects of seasonality in each of the pollution and meteorological variables were filtered out prior to fitting a model. If seasonality is not adequately controlled, models can exhibit spurious estimates of air pollution effects due to the presence of seasonality in both hospital admissions and air pollution series. In addition, the confounding effects of weather data also exhibit strong seasonality, compounding the problem. We utilize the Seasonal-Trend decomposition procedure based on Loess (STL) method. STL consists of a sequence of applications of the loess smoother that gives a decomposition that is highly resistant to extreme observations. When decomposing each series we assume a periodic seasonal component with varying lengths for the windows in the loess smooth for the seasonal and frequency components. We do not use STL to remove the long wavelength patterns in the hospital admissions data because it is not normally distributed and any adjustment will result in a loss of the underlying distribution. Thus, we use Fourier series to measure the seasonality in the hospital admissions data, after controlling for seasonality in all explanatory variables via STL.

We use a number of regression type approaches to describe the associations between ambient levels of air pollution and respiratory disease. We begin with a Generalized Linear

Models (GLM) analysis. For a GLM with a log link function, we specify the expectation of a random variable Y_t as

$$E(Y_t|X_t) = \exp\left(\beta_0 + \sum_{i=1}^r \beta_i X_{t,i}\right). \tag{1}$$

Here Y_t denotes daily counts of respiratory disease and air pollution and $X_t = (X_{t,1}, \dots, X_{t,r})'$ denotes the explanatory variables at time t. We assume an overdispersed Poisson model, estimated using a quasi-likelihood approach. Akaike's Information Criterion or AIC (Akaike 1973), was used for variable selection.

A nonparametric alternative to the parametric GLM is the Generalized Additive Model (GAM). GAMs allow non-linear relationships between the response variable and each explanatory variable (Hastie and Tibshirani 1990). For a GAM, we assume

$$E(Y_t|X_t) = \exp\left(\beta_0 + \sum_{i=1}^r g_i(X_{t,i})\right)$$
 (2)

where each g_i is a smooth, possibly non-linear, univariate function. Any of the g_i can be made linear to obtain a semi-parametric model. As with a GLM, we use quasi-likelihood estimation. Cubic smoothing splines were used to estimate the non-parametric functions g_i . We fix the smoothing parameter to be that value for which \hat{g}_i has four "degrees of freedom".

To adequately control for the confounding effects of weather variables, nonlinear functions between daily morbidity/mortality and weather, particularly temperature has been frequently documented (Saez et al. 1995; Sunyer et al. 1997; Rossi et al. 1999; Lee et al. 2000). Although a similar approach hasn't been taken to model the exposure-response functional form, a handful of studies have explored this possible phenomenon (Schwartz 1994; Kelsall et al 1997; Hagen et al 2000). At present, the shape of the exposure-response relationship between pollutants and morbidity/mortality is yet to be clarified (Abrahamowicz et al. 2003). Therefore, we have used both parametric and nonparametric statistical methodology to examine the functional form between hospital admissions and

each pollutant and each climatic variable.

The AIC was used to determine if each variable was to enter linearly or nonlinearly. Thus, only where there was statistical evidence of nonlinearity was a relationship allowed to have curvature.

A step-wise selection procedure was used in S-Plus (1999) to determine the inclusion and functional form of each of the pollutants and meteorological variables in a GAM. Step-wise selection is a traditional approach to analysis. The inclusion and thus importance of a variable is determined by statistical criteria. If met, the variable is included otherwise deleted. This iterative procedure of deletion, refitting and verification continues until all insignificant variables are eliminated from the final model. Step-wise selection has major limitations, particularly in epidemiological analysis where variables are significant confounders to the main exposure variable of interest. This procedure may overlook marginally significant variables thus biasing the results of the final multivariable models. Nevertheless this approach was chosen to illustrate the practical, although not necessarily optimal, method of using step-wise procedures available in many statistical packages and frequently adopted in many single city studies.

For each of the different models we began with a base model comprising a temporal time trend, day-of-week effects, and sine and cosine pairs for the seasonal variation in the outcome variable. Each of the pollutants, with confounding meteorological variables, were entered based on its significance and AIC criterion.

A parameter driven model (PDM) by Liang and Zeger (1986) was used to incorporate the presence of serial correlation, where the serial correlation is set up through an unobservable latent process.

A Poisson regression model has conditional mean

$$E(Y_t|\varepsilon_t, X_t) = \exp(X_t'\beta + \varepsilon_t), \tag{3}$$

where β denotes a vector of parameters, and ε_t is a latent process allowing both overdis-

persion and autocorrelation in Y_t . We allow ε_t to follow a first-order autoregressive process.

In addition, we utilized Transitional Regression Models (TRM) introduced by Brumback **et al**. (2000). Here, we present a special case of a TRM, defined as GLM with time series errors. For a Poisson with AR(1) errors the conditional mean is defined as

$$\mu_t = \exp(X_t'\beta) + \psi_1 e_{t-1} \sqrt{v_t},\tag{4}$$

where $e_t = (Y_t - v_t)/\sqrt{v_t}$ and $v_t = \exp(X_t'\beta)$.

Here e_t is scaled to give constant variance. Note that $e_t = \psi_1 e_{t-1} + \delta_t$ where $\{\delta_t\}$ is an independent series with zero mean.

Both the PDM and the TRM are adjuncts to the core regression models. However, the PDM and TRM differ with respect to modelling the serial correlation. The GLM with AR(1) errors, a special case of the TRM, is an observation driven model where the serial correlation is handled by including past values of the response variable as a covariate, whereas the serial correlation in a PDM is set up through an unobservable latent process. GLM, GAM, PDM and TRM all assume a Poisson distribution for both the COPD and asthma hospital admissions data.

For the PDM a SAS macro (GLIMMIX) was used to estimate both the regression parameters and covariance parameters. For the latent process we assumed an AR(1) structure. For the TRM we used an S-PLUS function provided by Brumback (personal communication).

Each of the four models was fitted to the asthma and COPD hospital admissions data. To simplify the analysis of seasonality, we excluded the leap days of 29 February 1992 in each series.

Prior to fitting the regression models we developed a base model for both COPD and asthma hospital admissions comprising Fourier series functions of $\sin(2\pi jt/365)$ and

 $\cos(2\pi jt/365)$ for $j=1,2,\ldots,J$. The value of J was chosen using the AIC. For COPD admissions, J=4 and for asthma admissions, J=10. The base model also included a time trend (a quadratic for GLM, PDM, TRM and a smooth term for GAM), and day-of-week effects. Dry bulbs temperature and relative humidity were seasonally adjusted and considered as potential confounders. Seasonally adjusted nitrogen dioxide, ozone, and non-seasonally adjusted sulphur dioxide, and API were considered as exposure variables in the models. As there is insufficient epidemiological or biological information upon which to base an a priori hypothesis of the lag effect of pollutants and weather on respiratory disease (Anderson **et al.** 1997) we have investigated instantaneous lags of up to 5 days for each of the pollutants and weather variables.

To allow comparison across models we use the following goodness of fit indices:

- Mean square error (MSE) = mean $\{(Y_t \hat{Y}_t)^2\}$, where \hat{Y}_t are the (inverse link transformed) fitted values.
- Mean square proportional error (MSPE) = mean $\{(Y_t \hat{Y}_t)^2 / \hat{Y}_t\}$.
- AIC = $n \log(\sigma^2) + 2p$, where σ^2 is the variance of the raw residuals (response minus fitted values), and p is the number of degrees of freedom in each model.

Results are presented as relative risks for an increase from the 10^{th} to 90^{th} percentile of the respective pollutant.

RESULTS

COPD hospital admissions in Melbourne, Australia from 1 July 1989 to 31 December 1992

INSERT FIGURE 1 ABOUT HERE

Table 1 displays results from the analyses of COPD hospital admissions, using different statistical methods. Where a variable has been included in a linear function, the relative risk is shown. For the GAM, variables which were included using a smoothing spline are denoted by g(). Daily COPD hospital admissions increased significantly with increased ambient outdoor levels of same day nitrogen dioxide (NO₂). The estimated nitrogen dioxide coefficients from the models in Table 1 were consistent, replicated across different models, and were statistically significant (p < 0.5), which is indicative of a robust relationship between nitrogen dioxide and COPD hospital admissions. The observed effects for ambient outdoor levels of ozone, particulates and sulfur dioxide were all highly sensitive to model specification. A GAM analysis showed a nonlinear relationship between sulfur dioxide and COPD hospital admissions in Melbourne, Australia (see Figure 1).

INSERT FIGURE 2 ABOUT HERE

After fitting each model, a plot of the autocorrelation function (ACF) of the residuals with corresponding 95% critical values (Figure 2) is a useful tool to visualise any remaining correlation in the data. The ACF of the residuals from a GLM is inadequate because of significant serial correlation. The other three models fare better, although the ACF of the residuals from a parameter driven model and GAM show some significant correlations. Of these three models, the TRM is ranked highest on the basis of AIC and MSE, and the PDM is best on the basis of MSPE. However, the GAM has the important advantage that it explains more of the variation in COPD admissions through the structure of the covariates than through the correlation terms.

Asthma hospital admissions in Melbourne, Australia from 1 July 1989 to 31 December 1992

INSERT FIGURE 3 ABOUT HERE

Table 2 displays results from the analyses of asthma hospital admissions, using different statistical methods. Similar results were found for asthma hospital admissions and same day nitrogen dioxide. The observed effects for same day nitrogen dioxide in Table 2 were robust to different model specifications, but lagged 1 day effects were not agreeable. The observed effects for ambient outdoor levels of ozone and API were highly sensitive to model specification. A GAM analysis of asthma hospital admissions showed a significant nonlinear effect of ozone lagged 2 days. This result is displayed in Figure 3.

INSERT FIGURE 4 ABOUT HERE

Figure 4 is a plot of the autocorrelation function of the residuals with corresponding 95% critical values for each of the models in Table 2. Although, the GAM is best on all three goodness-of-fit criteria, all models were inadequate due to the strong and significant correlation pattern observed in the autocorrelation function of the residuals.

DISCUSSION

There remains much debate regarding the adequacy of the statistical methods used to quantify the short term associations between air pollution and respiratory admissions within a single city study. The key methodological issues are the time series nature of the health and environmental data, non-normality and dose-relationships. We extend recent epidemiological single city studies by assessing the robustness of the observed pollution-respiratory admission relation to different statistical models with various underlying methodological assumptions. We address the time series nature of the environmental data by seasonally adjusting pollutants and weather variables where appropriate. None of the statistical methods adopted in this study are new, all are variations of classical regression methods currently used in practice.

In this study we are not arguing for any particular method of analysis; instead, our find-

ings demonstrate that for single-city studies, the conclusions drawn are highly dependent on the choice of model used. Single-city studies are common in practice, and the length of series we use here is typical of what is used in practice. Consequently, our study has important implications for the type of analysis that is commonly carried out in single city studies of air pollution and respiratory disease.

The findings from this study show that the relationship between ambient outdoor concentrations of nitrogen dioxide and both COPD and asthma hospital admissions is robust to different statistical methodology. The positive result of nitrogen dioxide is consistent with an Australian study (Morgan et al. 1998) and a number of European single city studies of COPD and asthma hospital admissions (Anderson et al. 1997; Vigotti 1999; Fusco et al. 2001).

The observed effects of ozone on both COPD and asthma hospital admissions were highly sensitive to model specification. Although our findings on the effects of ozone are consistent with another Australian study (Morgan et al. 1998), they are not in agreement with other single city studies of COPD and asthma hospital admissions. Studies in European cities (Vigotti 1999; Fusco et al. 2001) report associations between ozone and hospital admissions as the strongest and most replicable.

The relationship between particulates and both COPD and asthma hospital admissions were non-robust. However, the negative association in our study is similar to the Australian study by Morgan et al. (1998), who report negative effects of particulates (measured similar to the Melbourne study) on asthma hospital admissions. Both Vigotti et al. (1999) and Fusco et al. (2001) have also shown less consistent associations with particulates and hospital admissions for respiratory disease. The relationship between sulphur dioxide and both COPD and asthma admissions were non-robust.

The statistical methods presented in this study were inadequate in addressing all the methodological issues common to studies of respiratory disease and air pollution. For COPD hospital admissions, an observation driven model with AR(1) errors (labelled TRM) performed best, although a GAM was in some ways preferable because of its su-

perior ability to explain variation through the structure of the covariates. For asthma hospital admissions, no model was completely satisfactory in representing the strong correlation structure in the residuals. These results suggest that GAMs (commonly used for single city studies), are not completely adequate in modelling residual correlational in the mortality, morbidity variable, after adjusting for the common confounding variables.

The reliability of the convergence criterion of the backfitting algorithm used in fitting generalised additive models in S-PLUS has received some attention in recent years with regards to biased estimation of regression parameters and corresponding standard errors (Dominici et al. 2002; Katsouyanni et al. 2002b; Dominici et al. 2003). In light of this, S-PLUS has updated the software to include a more stringent convergence criterion. There is also the R implementation of the S language, with a GAM function that uses penalized regression splines for the smoothing terms and computed using penalized iteratively re-weighted least squares (Wood 2000) thereby removing potential estimation errors encountered in earlier findings from a GAM analysis of single city study associations. The GAM function is part of the mgcv package for R. The function not only avoids back-fitting but also does automatic bandwidth selection, thereby, allowing more flexibility in its utilization.

A potential limitation of the study is the use of a step-wise procedure for inclusion of variables in a model. In epidemiological studies variables not included in the model, due to inclusion criteria within the step-wise procedure, may be a potential confounder and often decisions are made as to which confounding variables are to be included prior to modelling. However, our preference was to use a stepwise procedure for each model as this is typically how analysis is carried out in practice. Generally, analysts use only one model and will use a stepwise selection procedure (or something similar) to determine the variables to include or otherwise. Thus, we are demonstrating the flaws in existing practice, rather than recommending this as the best approach.

An often ideal approach for model comparison is to simulate data for emission scenarios that may reflect real processes. However, one of the main objectives of this study is show that real time series data have greater complexity than any of the commonly-used existing models allow. We would not be able to demonstrate this using simulation techniques.

In developing countries, increasing urbanisation combined with increasing vehicle usage leads to greater daily exposure to outdoor gaseous pollutants making this an important area of research particularly in terms of immediate health outcomes of the young and elderly. Further development of statistical methodology is necessary to provide accurate and reliable risk estimates for susceptible populations. It is essential to understand the methodological short-comings and problems of currently available statistical methods used for single studies of air pollution and respiratory disease. None of the methods considered here were completely adequate in addressing complexities inherent in time series data. Consequently, results from single-city studies should be interpreted with caution, particularly if only one method of analysis has been used to demonstrate these effects. Key recommendations should reflect support for extensive research into further developing the statistical methods to quantify the short-term impact of air pollution on public health.

Acknowledgements: This study was funded by a Public Health Ph.D. Postgraduate Scholarship from the National Health and Medical Research Council. Part of this research appears in the proceedings of the 16^{th} International Workshop on Statistical Modelling.

REFERENCES

- Abrahamowicz M, Schopflocher T, Leffondre K, du Berger R, Krewski D. (2003) Flexible modeling of exposure-response relationship between long-term average levels of particulate air pollution and mortality in the American Cancer Society study. *Toxicol Environ Health A.* **10**, 1625–54.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In: *2nd International Symposium on Information Theory*. B.N. Petrov & F. Csaki (eds), Adademiai Kidao, Budapest; pp. 267–281.
- Anderson, H.R., Spix, C., Medina, S., Schouten, J., Castellsague, J., Rossi, G., Zmirou, D.,
 Touloumi, G., Wojtyniak, B., Ponka, A., Bacharova, L., Schwartz, J., & Katsouyanni,
 K. (1997) Air pollution and daily admissions for chronic obstructive pulmonary
 disease in 6 European cities: results from the APHEA project. *European Respiratory Journal*, 10, 1064–1071.
- Berktas BM, Bircan A.(2003)Effects of atmospheric sulphur dioxide and particulate atter concentrations on emergency room admissions due to asthma in Ankara. *Tuberk Toraks* **51**, 231–8.
- Brumback, B.A., Ryan, L.M., Schwartz, J.D., Neas, L.M., Stark, P.C., & Burge, H.A. (2000)

 Transitional regression models, with application to environmental time series. *Journal of the American Statistical Association*, **95**, 16–27.
- Chen Y, Yang Q, Krewski D, Shi Y, Burnett RT, McGrail K. (2004) Influence of relatively low level of particulate ar pollution on hospitalization for COPD in elderly people. *Inhal Toxicol* **16**, 21–5.
- Cleveland, W.S., & Terpenning, I.J. (1982) Graphical methods for seasonal adjustment. *Journal of the American Statistical Association*, 77, 52–62.
- Dominici F , McDermott A, and Hastie T. 2003 Issues in semi-parametric regression with applications in time series regression models for air pollution and mortality. http://www.ihapss.jhsph.edu

- Dominici F, McDermott A, Zeger S.L, Samet J.M. (2002) On Generalized Additive Models in time series studies of air pollution and health. *American Journal of Epidemiology* **156**, 1–11.
- EPA, "EPA Air Quality Information" (1999) Environment Protection Authority, Victoria, Australia. http://www.epa.vic.gov.au/aq/info/
- Fusco D, Forastlere F, Michelozzi P, Spadea T, Ostro B, Arca M, Perucci CA. (2001) Air pollution and hospital admissions for respiratory conditions in Rome, Italy. *European Respiratory Journal* 17, 1143–50
- Hagen, J.A, Nafstad, P., Skrondal, A., Bjorkly, S., & Magnus, P. (2000) Associations between outdoor air pollutants and hospitalization for respiratory disease. *Epidemiology* **11**, 136–140.
- Hastie, T., & Tibshirani, R.J. (1990) Generalized additive models. London: Chapman and Hall.
- Health Effects Institute (2003) Revised Analyses of Time-Series Studies of Air Pollution and Health. www.healtheffects.org
- Hoek, G., Schwartz, J., Groot, B., & Eilers, P. (1997) Effects of ambient particulate matter and ozone on daily mortality in Rotterdam, The Netherlands. *Archives of Environmental Health*, **52**, 455–463.
- Joseph A, Ad S, Srivastava A. (2003) PM(10) and its impacts on health a case study in Mumbai. *Int J Environ Health Res* **13** 207–14.
- Kan H, Chen B. (2003) Air pollution and daily mortality in Shanghai: a time-series study. *Arch Environ Health* **58** 360–7.
- Katsouyanni, K., Schwartz, J., Spix, C., Touloumi, G., Zmirou, D., Zanobetti, A., Wojtyniak, B., Vonk, J.M., Tobias, A., Ponka, A., Medina, S., Bacharova, L. & Anderson, H.R. (1996) Short term effects of air pollution on health: a European approach using epidemiologic time series data: the APHEA protocol. *Journal of Epidemiology and Community Health, Supplement*, **50**, S12–S18.

- Katsouyanni K, Toloumi G, Samoli E, Gryparis A, Monopolis Y, LeTertre A, Boumghar A, Rossi G, Zmirou D, Ballester F, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J, Schindler C, Schwartz J. (2002b) Different convergence parameters applied to the S PLUS gam function. *Epidemiology*, **13**, 742.
- Kelsall, J.E., Samet, J.M., Zeger, S.L., & Xu, J. (1997) Air pollution and mortality in Philadelphia, 1974–1988. *American Journal of Epidemiology*, **146**, 750–762.
- Lee JT, Kim H, Hong YC, Kwon HJ, Schwartz J, Christiani DC. (2000) Air pollution and daily mortality in seven major cities of Korea, 1991-1997. *Environ Res* **84** 247–54.
- Liang, K.Y., & Zeger, S. (1986) Longitudinal data analysis using generalized linear models *Biometrika*, **73**, 13–22.
- Lipfert F.W., and Wyzga R.E. (1995) Air pollution and mortality: Issues and uncertainties. *Journal of Air and Waste Management Association*, **45**: 949–966
- McMichael A.J., Anderson R.H., Brunekreef B., Cohen, A.J. (1998) Inappropriate use of daily mortality analyses to estimate longer-term mortality effects of air pollution. *International Journal of Epidemiology*, **27**: 450–453.
- Morgan, G., Corbett, S., Wlodarcyzk, J. (1998) Air pollution and hospital admissions in Sydney, Australia, 1990-1994. *American Journal of Public Health*, **88**, 1761–1766.
- Rossi G, Vigotti MA, Zanobetti A, Repetto F, Gianelle V, Schwartz J. (1999) Air pollution and cause-specific mortality in Milan, Italy, 1980-1989. *Arch Environ Health* **54** 158–64.
- Saez, M., Sunyer, J., Castellsague, J., Murillo, C., & Anto, J.M. (1995) Relationship between weather temperature and mortality: a time series analysis approach in Barcelona. International Journal Of Epidemiology 24, 576–581.
- Samet, J.M., Dominici, F., Curriero, F.C., Coursac, I., & Zeger, S.L. (2000) Fine particulate air pollution and mortality in 20 U.S. cities, 1987–1994. *The New England Journal of Medicine*, **343**, 1742–1749.

- Samet J.M., Dominici F., McDermott A., & Zeger, S.L. (2003) New problems for an old design: Time series analyses of air pollution and health. *Epidemiology*, **14**, 11–2.
- Schwartz, J (1993) Air pollution and daily mortality in Birmingham, Alabama. *American Journal of Epidemiology*, **137**, 1136–1147.
- Schwartz, J. (1994) Air Pollution & Hospital admissions for the elderly in Birmingham, Alabama. *American Journal of Epidemiology***139**, 589–598.
- Sunyer, J., Spix, C., Quenel, P., Ponce-de-Leon, A., Ponka, A., Barumandzadeh, T., Touloumi, G., Bacharaova, L., Wojtyniak, B., Vonk, J., Bisanti, L., Schwartz, J., & Katsouyanni, K. (1997) Urban air pollution and emergency admissions for asthma in four European cities: The APHEA project. *Thorax*52, 760–765.
- S-PLUS (1999). *Modern statistics and advanced graphics*. Seattle, Washington: MathSoft, Inc.
- Touloumi, G., Katsouyanni, K., Zmirou, D., Schwartz, J., Spix, C., Ponce de Leon, A., Tobias, A., Quennel, P., Rabczenko, D., Bacharova, L., Bisanti, L., Vonk, J.M., & Ponka, A. (1997) Short-term effects of ambient oxidant exposure on mortality: a combined analysis within the APHEA project. *American Journal of Epidemiology*, 146, 177–185.
- Vigotti MA. (1999) Short-term effects of exposure to urban air pollution on human health in Europe. The APHEA Projects (Air Pollution and Health: a European Approach. *Epidemiological preview*, **23**, 408–15.
- Wong TW, Lau TS, Yu TS, Neller A, Wong SL, Tam W, Pang SW. (1999) Air pollution and hospital admissions for respiratory and cardiovascular diseases in Hong Kong. *Occup Environ Med* **56**, 679–83.
- Wong GWK, Ko FWS, Lau TS, et al. (2001) Temporal relationship between air pollution and hospital admissions for asthmatic children in Hong Kong. *Clin Exp Allergy* **31**, 565–569.

Wood S.N, (2000) Modelling and Smoothing Parameter Estimation with Multiple Quadratic Penalties. *Journal of the Royal Statistical Society B*, **62**, 413–428.

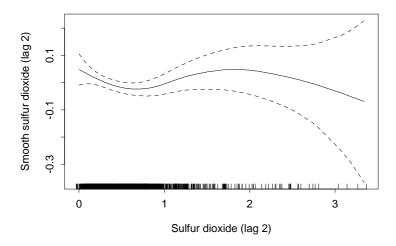


Figure 1: The nonlinear function for sulfur dioxide (lagged 2 days). Dashed lines represent pointwise 95% confidence intervals.

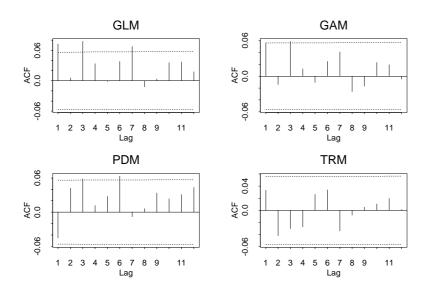


Figure 2: The autocorrelation function for the deviance residuals from the models applied to COPD hospital admissions in Table 1.

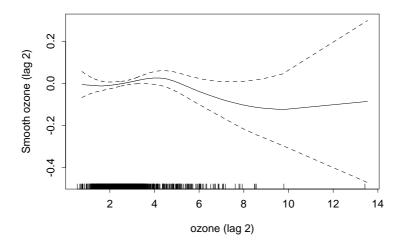


Figure 3: The nonlinear function for ozone (lagged 2 days). Dashed lines represent pointwise 95% confidence intervals.

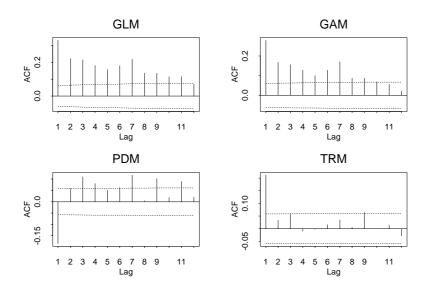


Figure 4: The autocorrelation function for the deviance residuals from the models applied to asthma hospital admissions in Table 2.

Table 1: Relative Risk and 95% CI of COPD hospital admissions for an increase from the 10th to 90th percentile for levels of pollutants, generated using different statistical models.

COPD								
	GLM		GAM		PDM		TRM	
Pollutant	RR	95%CI	RR	95% CI	RR	95%CI	RR	95% CI
$NO_{2,t}$	1.06	1.00-1.12	1.06	1.01-1.11	1.05	1.00-1.11	1.05	1.00-1.10
$O_{3,t-2}$			1.06	1.00-1.11				
API_{t-2}			0.95	0.91 - 1.00				
$SO_{2,t-2}$			g()					
MSE		13.23	12.76		12.88		12.29	
MSPE		1.24		1.19		1.13		1.16
AIC	3	340.42	3	292.19	3	252.84	3	243.09

Table 2: Relative Risk and 95% CI of asthma hospital admissions for an increase from the 10th to 90th percentile for levels of pollutants, generated using different statistical models.

Asthma			-			0 77		
	GLM		GAM		PDM		TRM	
Pollutant	RR	95%CI	RR	95% CI	RR	95%CI	RR	95% CI
$NO_{2,t}$	1.05	1.01-1.08	1.05	1.01-1.09	1.04	1.01-1.08	1.05	1.02-1.08
$NO_{2,t-1}$			0.96	0.92-0.99				
$O_{3,t}$			0.97	0.93 - 1.00				
$O_{3,t-1}$	0.96	0.93-0.99			0.97	0.94-1.09	0.97	0.95-0.99
$O_{3,t-2}$			g()					
API_t								
$SO_{2,t}$								
MSE	57.81		53.68		56.05		55.8	
MSPE		1.75		1.61		1.70		1.69
AIC	5	244.03	5	153.41	5	207.56	5	206.92