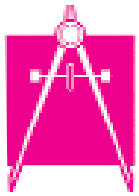

Comparison of alternative modelling techniques in estimating short term effect of air pollution with application to the Italian Meta-analysis data

Confronto di modelli alternativi per la stima degli effetti a breve termine dell'inquinamento atmosferico con applicazione allo studio MISA

Michela Baccini,[1,2] Annibale Biggeri,[1,2] Gabriele Accetta,[2] Corrado Lagazio,[3] Aitana Lerxtundi,[1,4] Joel Schwartz[5]

[1] *Department of Statistics «G. Parenti», University of Florence, Italy*
[2] *Biostatistics Unit, Institute for Cancer Prevention (CSPO), Florence, Italy*
[3] *Department of Statistical Sciences, University of Udine, Italy*
[4] *Research Group on Statistics, Applied Economics and Health (GRECS), University of Girona, Spain*
[5] *Department of Environmental Health, Harvard School of Public Health, Boston, USA*

Corrispondenza: *Annibale Biggeri, Department of Statistic «G. Parenti», University of Florence; e-mail: abiggeri@ds.unifi.it*

Riassunto

Nel 2002 sono state avanzate critiche sull'uso del software statistico standard (Splus, SAS, Stata) per adattare Modelli Additivi Generalizzati (GAM) alle serie temporali di dati epidemiologici. Le critiche si riferiscono a problemi di convergenza dell'algoritmo implementato e all'uso inappropriato di un'approssimazione lineare per il calcolo degli errori standard delle stime dei termini parametrici, tra i quali l'effetto dell'inquinamento atmosferico. Qui noi utilizziamo due metodi alternativi, che non sono influenzati dagli stessi problemi, per lo studio dell'associazione tra PM10 e mortalità/ricoveri ospedalieri nella Metanalisi italiana degli effetti a breve termine degli inquinanti atmosferici (MISA). Il primo approccio proposto è basato sulla stima di GAM con *penalized regression spline* attraverso il metodo diretto implementato nella libreria *mgcv* del software *R* (GAM-R). Il secondo approccio è completamente parametrico, basato sulla specificazione e stima di modelli lineari generalizzati con *spline* di regressione (GLM+NS). Viene fornita un'analisi di sensibilità mediante la variazione del numero di gradi di libertà per la *spline* utilizzata per modellare la stagionalità e della modalità di aggiustamento per l'effetto confondente della temperatura. I risultati sono discussi alla luce della teoria asintotica sviluppata nell'ambito dei modelli additivi e di uno studio di simulazione inteso a spiegare le discrepanze tra le stime GLM+NS e GAM-R. Viene concluso che in generale l'approccio completamente parametrico ha migliori proprietà statistiche del GAM-R, il quale potrebbe portare a stime distorte degli effetti dell'inquinamento atmosferico, a meno che non venga stabilito un certo grado di *undersmoothing* per la *spline* stagionale.
(*Epidemiol Prev* 2006; 30(4-5): 279-88)

Parole chiave: modelli additivi generalizzati, spline di regressione, spline di regressione penalizzate, serie storiche epidemiologiche

Abstract

In 2002, serious criticism was raised about the use of standard statistical software (Splus, SAS, Stata) to fit Generalized Additive Models (GAM) to epidemiological time series data. This criticism concerns convergence problems of the backfitting algorithm and inappropriate use of a linear approximation in estimating standard errors of estimates for parametric terms, such as the effect of air pollution. Here we analysed the association between PM10 and Mortality/Hospital Admissions in the Italian Meta-analysis of Short-term effects of Air pollutants (MISA) using two alternative approaches that are not affected by the same drawbacks: GAM with penalized regression spline fitted by the direct method in R *(GAM-R) software and Generalized Linear Models with natural cubic spline (GLM+NS). A sensitivity analysis is also provided varying number of degrees of freedom for the seasonality spline and modality of adjustment for confounding effect of temperature. Published theoretical results and a simulation study are provided in order to explain discrepancies between GLM+NS and GAM-R estimates. We conclude that in general the fully parametric GLM+NS approach retains better statistical properties than GAM-R that could bring to biased air pollution effect estimates unless a certain degree of undersmoothing for seasonality spline is settled.*
(*Epidemiol Prev* 2006; 30(4-5): 279-88)

Key words: *Generalized Additive Model, regression spline, penalized regression spline, epidemiological time series*

Introduction

Short-term effects of air pollution on health are widely documented and several meta-analyses have been conducted.[1-7] In 2002, a major concern was raised about the numerical accuracy of the estimated pollutant effect obtained fitting Generalized Additive Models (GAM) where seasonal confounding is adjusted by using a non parametric function of time.[8,9] Ramsay et al.[10] and Dominici et al.[11] identified important critical points in the analyses of epidemiological time series using commercial statistical software which fits GAM by back-

Abbreviations used in the text

MISA:	Meta-analysis of Italian studies on Short-term effects of Air pollution
GAM:	Generalized Additive Models
GAM-R:	Generalized Additive Models fitted using the direct method
GLM+NS:	Generalized Linear Models with natural cubic spline(s)
GAM-S:	Generalized Additive Models fitted using the backfitting algorithm
ICD 9:	International Classification of Diseases, Ninth Revision
IRLS:	Iterative Re-weighted Least Squares
GCV:	Generalized Cross Validation

fitting algorithm, such as Splus, SAS, and Stata. In brief:

1. the estimated standard errors of parametric terms obtained fitting GAM in these software packages are biased if a nonparametric term is included in the model;

2. the default convergence criteria of backfitting algorithm defined in the *gam* functions are too lax to assure convergence, and can lead to biased estimates of the pollutant effect.

For the point 1, Splus (but also other statistical packages, for example SAS and STATA) provides an approximation of the variance-covariance matrix, which takes into account only the linear component of the variable that was fit with a smooth function.[12] Then, a bias is expected whenever strong non-linearity and non-horthogonality between parametric and non-parametric terms are present, such as between the smoother for time and the pollutant concentration in epidemiological time series analysis.[10]

The second point is for certain aspects trivial: whenever the magnitude of the effect to be estimated is of the same order of the convergence criteria some degree of numerical instability is expected, which decreases as the effect size increases. More interestingly, if the data exhibit a relevant degree of concurvity («collinearity» among parametric and non-parametric components of the model), convergence of backfitting algorithm can be very slow.[12,13] This exacerbates the consequences of using default convergence criteria, in terms of potential failure to converge, and resultant biased parameter estimates. Dominici et al.[11] found that, when a spline for time and a spline for weather are included in the model, the greater the degree of concurvity, the greater is the overestimation of the pollutant effect.

Reanalyses of several published papers were done using approaches alternative to Splus implementation of GAM.[14] The present paper analyses the data of the Italian Meta-analysis of Short-term Effects of Air Pollution (MISA),[15-17] using alternative approaches for modelling seasonality which are not affected by the same drawbacks of GAM via backfitting algorithm: GLM with natural cubic spline(s) and GAM with penalized regression splines fitted by the *gam* function implemented for R software by Wood.[18-20] Both these approaches estimate the variance covariance matrix correctly and are less sensitive to the definition of convergence criteria.

Sensitivity of results to changes in number of degrees of freedom was checked and a simulation study was performed in order to make clear the nature of possible discrepancies in city-specific results under the two approaches. The interpretation of our findings at the light of published theoretical results[21] was discussed.

Finally in order to provide an all-in valuation of the modelling strategy adopted, different approaches in adjusting for confounding effect of temperature were adopted.[22]

Methods for data analysis

Spline-based approaches in brief

The characteristics of epidemiological time series data require statistical methods able to control for temporal trends over multiple years. Since mortality and other health indicators tend to rise and fall each year with the season, this trend is an inherently nonlinear confounding effect of temporal trend. One approach to dealing with temporal trend is to divide the time span of the study into shorter periods and fit separate polynomials within each range. This allows different shape curves to fit different ranges. Natural cubic splines are a form of this parametric approach. While flexible, they can still be sensitive to the position of break points between the time periods (knots). To avoid this most of the air pollution studies used more flexible semi-parametric approaches, specifying models with smoothing splines or locally weighted regressions in moving ranges of the data (loess). These models, belonging to the class of GAM, are those implemented in Splus (and SAS and Stata) by backfitting algorithm.[12]

A third approach consists in defining penalized regression splines. Penalized regression splines are a middle way between parametric splines and smoothing splines. They use separate polynomials in each range (as parametric splines do), but they reduce the sensitivity to knots location by using many of them and avoid excessively wiggly curves by constraining the coefficients not to change too much between one break point and another.[20] The use of penalized regression splines eliminates the need to implement backfitting algorithm, while still providing the flexibility of smoothing splines. This semi-parametric approach is implemented in the *gam* function of R.[19]

Principal characteristics of the MISA study

The MISA study,[15-17] a planned meta-analysis of epidemiological time series of eight Italian cities (1990-1999), investigated mortality for all natural, cardiovascular and respiratory causes and hospital admissions for cardiovascular and respiratory diseases. Health data were collected from Local Health Authorities and regional files. Daily pollutant concentrations were obtained from Regional Environmental Protection Agencies or local sources. The same procedure for

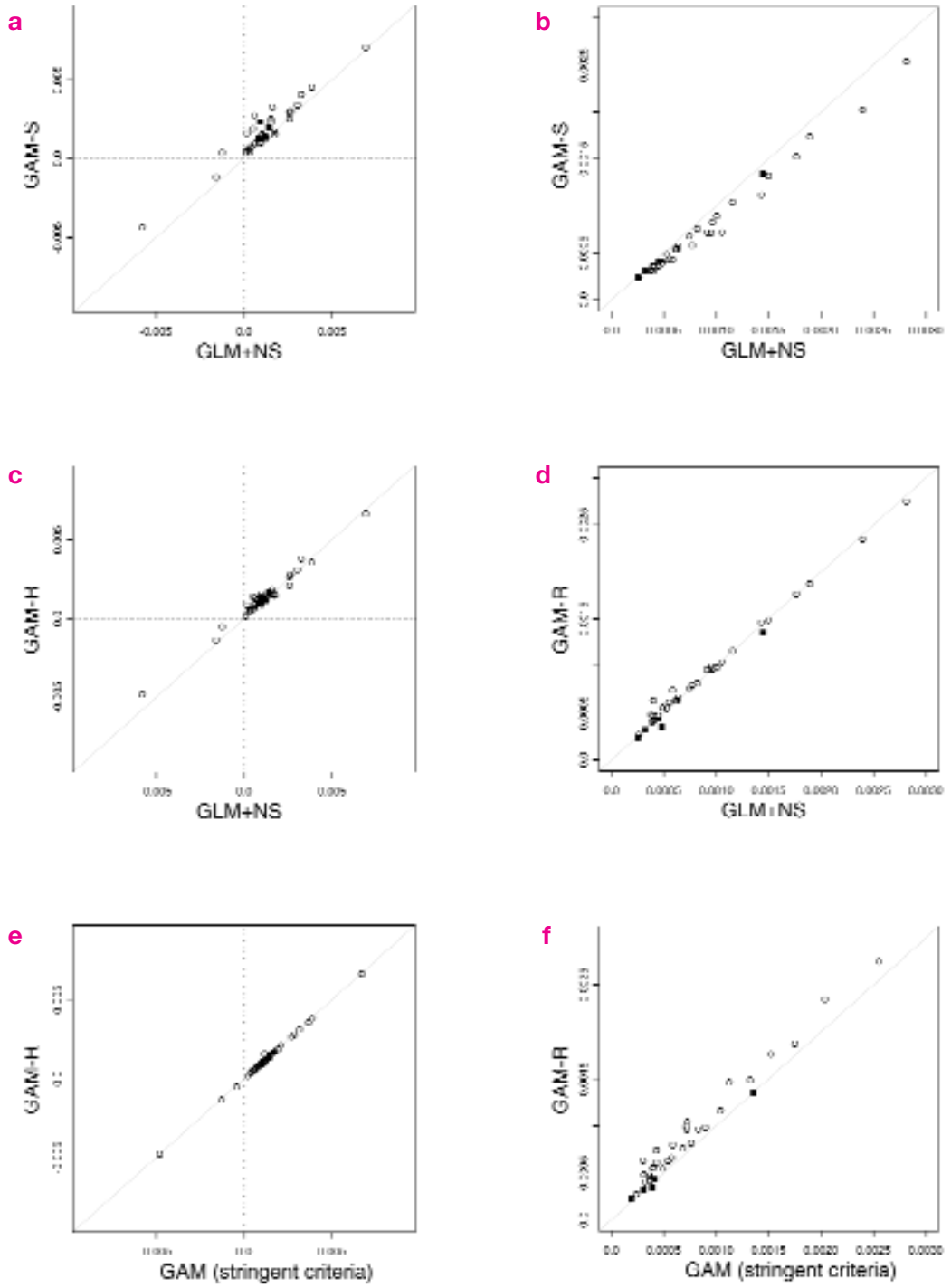

Figure 1. Italian Meta-analysis of Short-term Effects of Air Pollution. MISA 1995-1999. ***(a)*** *City specific and Meta-analytic (in square bold) Effect Estimates (log Relative Risk) for PM10 (increase of 10 μg/m^3) by fitting GAM-default settings (Y-axis) vs GLM-NS (X-axis).* ***(b)*** *City specific and Meta-analytic (in square bold) Standard Error Estimates for PM10 effects by fitting GAM-default settings (Y-axis) and GLM-NS (X-axis).* ***(c)*** *City specific and Meta-analytic (in square bold) Effect Estimates (log Relative Risk) for PM10 (increase of 10 μg/m^3) by fitting GAM-direct method in R Language (Y-axis) vs GLM-NS (X-axis).* ***(d)*** *City specific and Meta-analytic (in square bold) Standard Error Estimates for PM10 effects by fitting GAM-direct method in R Language (Y-axis) and GLM-NS (X-axis).* ***(e)*** *City specific and Meta-analytic (in square bold) Effect Estimates (log Relative Risk) for PM10 (increase of 10 μg/m^3) by fitting GAM-direct method in R Language (Y-axis) vs GAM-stringent convergence criteria (X-axis).* ***(f)*** *City specific and Meta-analytic (in square bold) Standard Error Estimates for PM10 effects by fitting GAM-direct method in R Language (Y-axis) and GAM-stringent convergence criteria (X-axis).*

Method	Mortality						Hospital Admissions			
	All Natural Causes		Cardiovascular		Respiratory		Cardiac		Respiratory	
	fixed	random	fixed	random	fixed	random	fixed	random	fixed	random
GAM-S default	1.12 (0.82;1.42)	1.24 (0.63;1.86)	1.23 (0.76;1.69)	1.43 (0.62;2.25)	2.24 (1.09;3.41)	1.96 (-0.69;4.68)	1.23 (0.93;1.53)	1.30 (0.83;1.78)	2.13 (1.76;2.50)	2.35 (1.52;3.18)
GAM-S stringent	0.92 (0.62;1.22)	1.06 (0.46;1.66)	1.03 (0.57;1.50)	1.24 (0.43;2.06)	1.96 (0.81;3.13)	1.69 (-0.97;4.42)	0.99 (0.69;1.29)	1.02 (0.64;1.39)	1.26 (0.89;1.63)	1.42 (0.66;2.20)
GAM-R	0.90 (0.55;1.25)	1.04 (0.41;1.67)	1.05 (0.52;1.58)	1.26 (0.41;2.12)	1.92 (0.64;3.21)	1.70 (-0.96;4.43)	0.95 (0.50;1.40)	0.95 (0.50;1.40)	1.34 (0.84;1.86)	1.34 (0.64;2.03)
GLM+NS	0.85 (0.52;1.18)	0.98 (0.35;1.61)	0.97 (0.45;1.50)	1.21 (0.32;2.10)	1.74 (0.44;3.05)	1.41 (-1.41;4.32)	0.77 (0.40;1.15)	0.82 (0.32;1.32)	0.73 (0.27;1.20)	0.91 (-0.04;1.86)

Table 1. Italian Meta-analysis of Short-term Effects of Air Pollution. MISA 1995-1999.
Combined meta-analytic estimates of percentage increase in outcome (95% CI) associated to a PM10 increase of 10 μg/m^3 by fixed and random effects models. City specific estimates obtained by GAM via backfitting with default convergence criteria of Splus 2000, GAM via backfitting with stringent convergence criteria, GAM via direct method in R Software, GLM with natural cubic spline.

collecting data was used in all participating cities (Turin, Milan, Verona, Ravenna, Bologna, Florence, Rome, Palermo). MISA used a common model for the city specific analyses. The analysis was age-adjusted (0-64; <65-74; 75+). We controlled for time-related confounding including in the model spline terms, whose number of degrees of freedom was a priori specified (5 per year for mortality only for the third age class, since indicator variables for season were used for the first two; 6-5-6 per year for hospital admissions for cardiac diseases for the three age classes, respectively; 7-5-6 per year for hospital admissions for respiratory diseases). Two linear terms constrained to joint in 21°C for temperature and linear and quadratic terms for relative humidity were defined.[23] We controlled also for day of the week, holidays and influenza epidemics by appropriate dummy variables.

Comparing methods

In this paper the following spline-based modelling approaches were proposed as alternative to GAM via backfitting algorithm for the city-specific analyses of MISA:

- Generalized Linear Models (GLM) with natural cubic spline(s) with fixed pre-specified knots,[24] fitted by the standard iteratively reweighted least squares (IRLS) algorithm (GLM+NS);
- GAM with penalized regression spline(s), fitted by direct method implemented in the *mgcv* R library[19] (GAM-R).

GLM+NS is a fully parametric alternative to GAM. Once the number and position of knots has been defined and an appropriated design matrix has been build, the maximum likelihood estimates of the coefficients of GLM+NS can be obtained using standard algorithms for the estimation of Generalized Linear Models.[25] In this analysis, knots were placed evenly throughout of covariate values.

The function *gam* of R allows inclusion in the model of penalized regression splines whose smoothing parameters are fixed to obtain the desired number of degrees of freedom. This function maximizes the penalized likelihood by a direct method which avoids the iterative process nested in the backfitting algorithm.[13] The GAM implementation in R correctly calculates the variance-covariance matrix.

To complement methods comparison, we fit also GAM with smoothing cubic spline(s) by the *gam* function of Splus with default (of the order $\varepsilon<10^{-3}$) and stringent ($\varepsilon<10^{-14}$) convergence criteria (GAM-S), despite this approach is affected by the previously described drawbacks.

In all the models the same number of degrees of freedom was used, according to the MISA protocol (see the previous paragraph). The issue of undersmoothing when fitting GAM models is addressed in the simulation study below.

The combined meta-analytic estimates were calculated using fixed and random effects models.[26]

Sensitivity study

A sensitivity analysis was performed varying the number of degrees of freedom for the natural cubic splines and penalized regression splines. Finally a sensitivity analysis evaluated the impact of non parametric modelling of temperature on pollutant effect estimates.

Simulation study

Statistical theory indicates that, provided the knots number and position are correctly specified, the estimates of the exposure effect obtained from a model with a natural cubic spline for the temporal trend are affected by negligible bias. In contrast, semi-parametric modelling can lead to biased results, unless concurvity is small or a certain degree of undersmoothing for a given dataset is specified (that is, more degrees of freedom are used to fit the term than for the natural spline model). This result goes back to a theorem by Rice[21] in the context of Additive Models, which states the asymptotic behaviour of the parametric coefficients estimators when a non parametric function is in the model. The advantage of the GAM model, then, is that the correct position of the knots is not known, and in this case its greater flexibility may be useful. In order to shed light on this point, a simple simulation analysis was performed, considering the

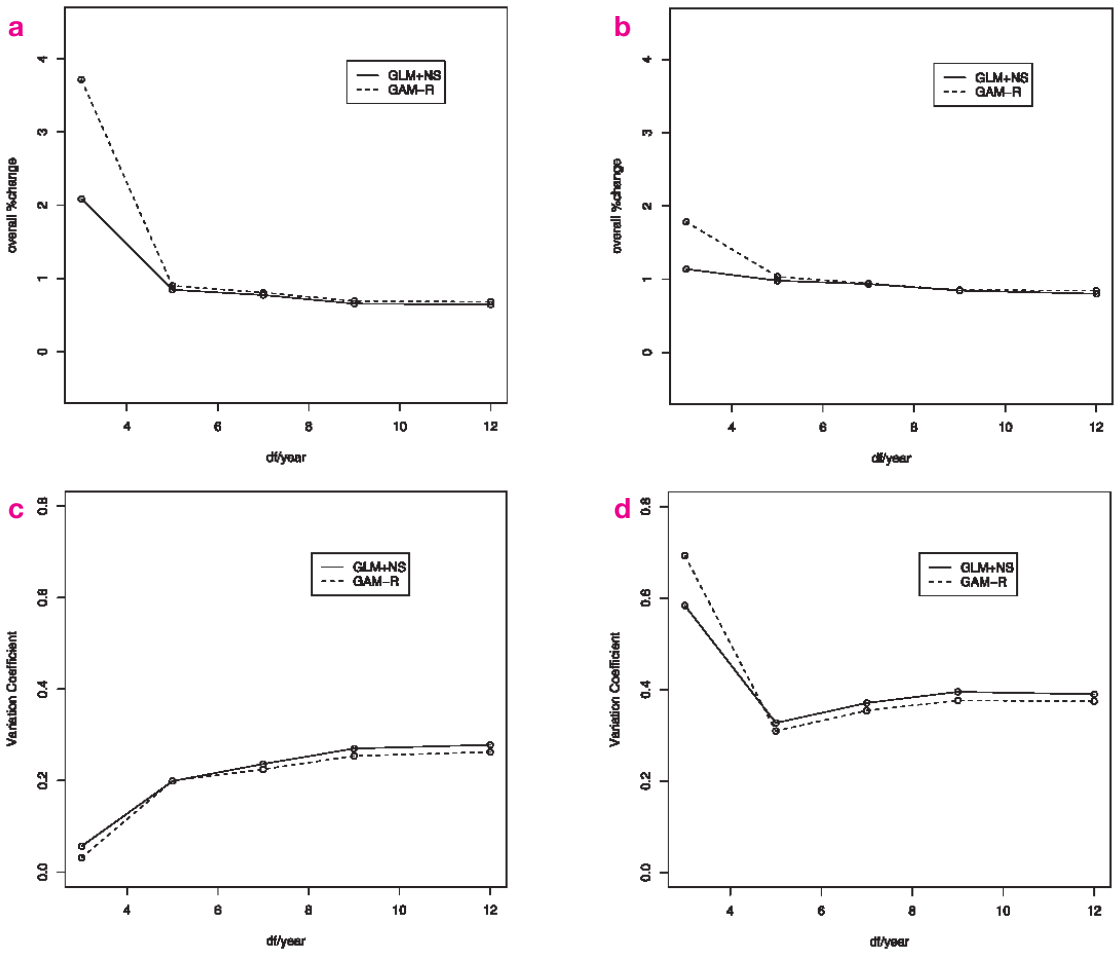

Figure 2. Italian Meta-analysis of Short-term Effects of Air Pollution. MISA 1995-1999. ***(a)*** *Fixed effect overall estimates for the effect of PM10 on total mortality under GLM+NS and GAM-R, varying the number of degrees of freedom for the seasonality spline.* ***(b)*** *Random effect overall estimates for the effect of PM10 on total mortality under GLM+NS and GAM-R, varying the number of degrees of freedom for the seasonality spline.* ***(c)*** *Variation coefficients of fixed effect overall estimates for PM10 on total mortality under GLM+NS and GAM-R, varying the number of degrees of freedom for the seasonality spline.* ***(d)*** *Variation coefficients of random effect overall estimates for PM10 on total mortality under GLM+NS and GAM-R, varying the number of degrees of freedom for the seasonality spline.*

situation in which only a spline for time trend and a linear term for air pollutant effect were included in the model.
We generated pseudo data using the daily number of hospital admissions for respiratory diseases and the mean daily concentration of NO_2 from Barcelona (1995-1999).[27]
We obtained a pseudo curve for seasonality, $f_0(t)$, fitting a Poisson Additive Model on the daily number of events with a penalized regression spline for time trend with pre-defined number of degrees of freedom (d_0). Then, we generated a pseudo air pollution time series, X_t, fitting a penalized regression spline for time trend with df_0 degrees of freedom on the daily air pollution data and adding to the fitted curve normal error terms. Specifying the value of the error variance, we controlled the amount of concurvity in data. Finally, we simulated outcome time series sampling from the following Poisson distribution:
$Y_t \sim Poisson\ (\mu_{0t})\ \log \mu_{0t} = \alpha_0 + f_0(t) + \beta_0 X_t$
where α_0 and β_0 are the «true» coefficients of the model.
Different simulation analyses were performed varying the number of degrees of freedom df_0 used to obtain the pseudo seasonality curve: df_0= 3,4,5,7,9 per year. The results reported in Table 2 were obtained fixing the concurvity amount to 0.45 and the true effect size β_0 to 0.0006. This parameters choice correspond to a realistic situation which is close to those usually observed in epidemiological time series analysis.
For each choice of df_0, we sampled 3000 outcome time series and we analysed them fitting three different models:
1) a GAM with a penalized regression spline for time trend with df_0 degrees of freedom;
2) a GLM with a natural cubic regression spline for time trend with df_0 degrees of freedom;
3) a GAM with a penalized regression spline for time trend whose degrees of freedom were selected by Generalized Cross

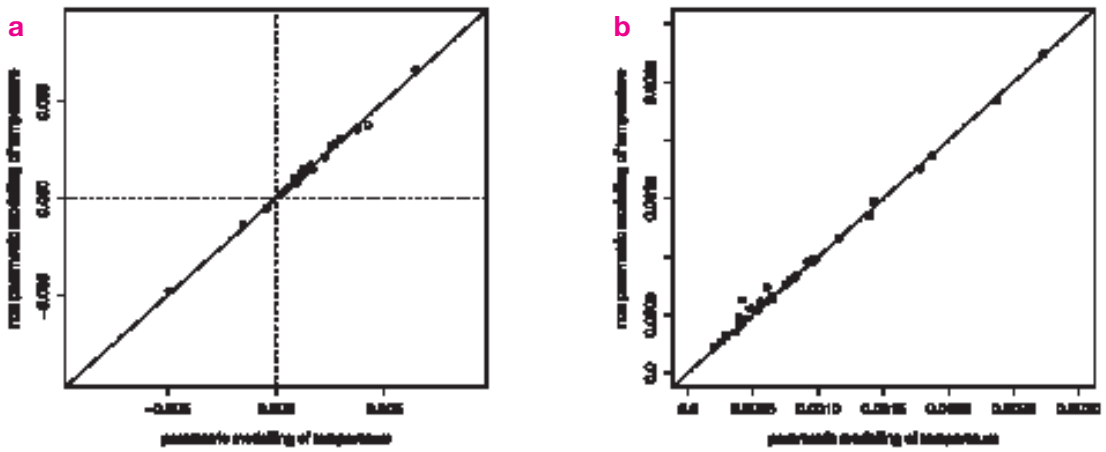

Figure 3. Italian Meta-analysis of Short-term Effects of Air Pollution. MISA 1995-1999. ***(a)*** *City specific and random effects meta-analytic (in square bold) effect estimates (log Relative Risk) for PM10 (increase of 10 μg/m³) by fitting GAM-R with non-parametric modelling of temperature (Y-axis) vs GAM-R with parametric modelling of temperature (X-axis).* ***(b)*** *City specific and random effects meta-analytic (in square bold) Standard Error Estimates for PM10 effects by fitting GAM-R with non-parametric modelling of temperature (Y-axis) vs GAM-R with parametric modelling of temperature (X-axis)*

Validation.[12] This choice was motivated by the fact that GCV is known to undersmooth.[28] According to the asymptotic theorem by Rice, less biased effect estimates have to be expected using this model selection method.

Moreover, in order to better address the problem related to the need of undersmoothing under the GAM-R approach, we performed also a simulation analysis fitting models with more than df_0 degrees of freedom for the seasonal penalized regression spline. In particular, after having fixed df_0=3 and 5 per year, we explored performances of GAM-R assuming different percentage of undersmoothing: 20, 40, 60, 80, 100 and 140 percent.

It should be noticed that, in order to evaluate the bias under GAM-R in a conservative way, we generated simulated data using pseudo curves fitted by penalized regression splines. In principle, this choice could bias the results of the simulation analysis in favour of the semi-parametric approach.

Results from MISA data

We present below the results for the effects of PM10 in the calendar period 1995-1999. For mortality (available data from 6 cities) the exposure variable was defined as the mean of PM10 concentrations in the current and previous days (lag 0-1), while for hospital admissions (7 cities) lag 0-3 was used.[15-17]

As expected, we found relevant disagreements between the city-specific results obtained from standard GAM-S, with

Table 2. Results of simulation analysis varying the number of degrees of freedom in generating pseudo seasonality curve (df_0=3,4,5,7,9 per year, β_0=0.0006, concurvity=0.45).

df_0 per year	GLM+NS		GAM-R		GAM-R + GCV	
	% Bias	Real Coverage	% Bias	Real Coverage	% Bias	Real Coverage
3	2.7	95.1	155.6	5.9	16.6	93.5
4	1.6	94.7	70.5	60.0	12.2	94.2
5	0.9	95.6	15.5	93.4	4.1	95.5
7	3.3	94.8	6.3	95.0	3.0	94.8
9	4.1	94.9	5.0	94.8	3.1	94.8

Table 3. Results of simulation analysis using fixed percentages of undersmoothing under GAM-R approach (df_0 = 3,5 per year, β_0 = 0.0006, concurvity = 0.45).

Percentage of undersmoothing	df_0 = 3 per year		df_0 = 5 per year	
	% Bias	Real Coverage	% Bias	Real Coverage
0 %	155.6	5.9	15.5	93.4
20 %	76.4	49.0	8.6	94.5
40 %	44.2	78.9	4.4	95.3
60 %	26.4	89.6	2.5	95.3
80 %	16.8	93.0	1.5	95.3
100 %	11.5	94.2	1.0	95.2
140 %	5.9	94.6	0.4	95.3

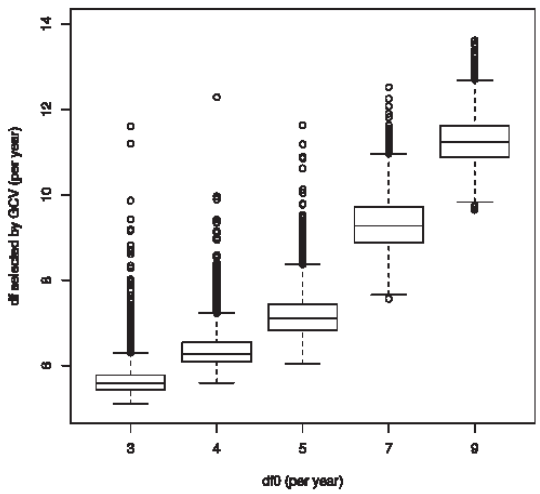

Figure 4. Boxplot of number of degrees of freedom selected by GCV, under different specification of degrees of freedom for the spline in pseudo data (concurvity=0.45, β_0=0.0006).

default or more stringent criteria, and the two alternative approaches, GLM+NS and GAM-R.

In Figures 1 (a)-(b) we compare coefficients estimates and estimated standard errors from GAM-S with default convergence criteria and GLM+NS. Each point corresponds to a city-specific estimate; points marked as bold squares represent combined random effects meta-analytic estimates (3 mortality outcomes, 2 hospital admission outcomes). The GLM+NS coefficients estimates were generally lower and the estimated standard errors were greater, proportionally to their magnitude, than those wrong obtained from GAM-S with default convergence criteria.

Using more stringent convergence criteria, GAM-S provided point estimates very close to those obtained from GAM-R (Figure 1 (e)). This is an expected results, when a large number of knots (here 150) is defined for the penalized regression splines.[19-21] However even if appropriate convergence criteria were defined, performance of GAM-S in terms of estimated precisions did not improve (Figure 1 (f)).

Figures 1 (c)-(d) compare GAM-R with GLM+NS. The results appeared similar, even if point estimates from GLM+NS were usually lower than those obtained from GAM-R.

Focussing on the meta-analyses results (Table 1), we found that qualitatively the main conclusions did not change using the different approaches: for most of the outcomes, effects were statistically significant. However GAM-S with bland convergence criteria resulted in overestimated effects and mistakenly small confidence intervals. The random effects meta-analyses had smaller differences in the confidence intervals between GAM-S with strict convergence criteria and GLM+NS. This is expected, since the smaller within city confidence intervals produced by GAM-S result in larger between city estimates of variance in the random effect meta-analysis. As expected on the basis of city-specific results, the overall estimates under GAM-R were slightly higher than under GLM+NS. For example, using GLM+NS, the overall estimated percent increase of total mortality for natural causes for 10 μg/m^3 increases of PM10 was 0.98 (95 percent confidence interval: 0.35,1.61, random effect model) in the calendar period 1995-1999, for a lag time of 0-1 day. Using GAM-R, the overall estimated percent increase was 1.04 (0.41,1.67). These compare with the wrong estimate of 1.24 (0.63,1.86) from GAM-S with default convergence criteria.

A sensitivity analysis of meta-analytic results to change degrees of freedom for the splines in GLM+NS and in GAM-R were conducted. Results for total mortality analysis are reported in Figure 2. Figures 2 (a)-(b) show the overall estimates of the PM10 effect using 3, 5 (reference), 7, 9 and 12 degrees of freedom per year under fixed effects and random effects meta-analysis models. Both GAM-R and GLM+NS appeared robust to increasing number of degrees of freedom for the spline. On the contrary, using 3 degrees of freedom per year, both approaches (but in particular GAM-R), brought to higher overall effect estimates, in particular when fixed effects meta-analysis was used.

In Figures 2 (c)-(d) the coefficients of variation for fixed effects and random effects meta-analytic estimates are reported, varying the degrees of freedom for the spline. The precision of city-specific estimates usually decreased as the number of degrees of freedom for temporal trend increased (points in the scatter plot), with city-specific confidence intervals for the PM10 effect obtained using 3 degrees of freedom being the narrowest. This explain the plot for the fixed effects meta-analysis.

Combining the city-specific results by random effects meta-analysis, a different behaviour was observed. The estimated variance decreased then increased, with minimum around 5 degrees of freedom per year. When few degrees of freedom for the spline were used, the lower within city variance estimates were balanced by a larger among cities variability.

Finally we performed a sensitivity analysis changing the modelling strategies for temperature in GAM-R. We compared the model proposed in MISA (with two constrained linear terms for temperature) with a model where a penalized regression spline for temperature with 7 degrees of freedom was introduced. The results that we obtained were very similar both in terms of point estimates and precision (Figure 3).

Results of simulation study

As described in the previous paragraph, even if GLM+NS and GAM-R gave consistent results both in city-specific analyses and in meta-analysis, usually the estimated exposure effect under GLM+NS was lower than under GAM-R with the same degrees of freedom for trend. This difference can be attributed to a different asymptotic behaviour of the

two estimators for the PM10 effect under the semi-parametric and the parametric approach.[21,29-31]
Both estimators are «consistent», in the sense that bias and variance of them tend to zero as the sample size increases. This means that a certain bias is ever to be expected, but it can be negligible if the sample size is sufficiently large. However, Rice[21] demonstrated the bias of the estimator of the parametric component (i.e. air pollution effect) in a semi-parametric model (i.e. GAM-R) converges to zero slower than in a fully parametric model (i.e. GLM+NS). As a consequence, for a given sample size, we should expect smaller bias of the estimator under GLM+NS than under GAM-R and comparable variances. On the basis of this result we argue that the observed discrepancy between point estimations under the two approaches can be due to a slightly overestimate the true effect of PM10 by GAM-R, and may be corrected by using more degrees of freedom for trend.
The results of the simulation study can help us to better understand the meaning of the previous statement. Table 2 reports the results of simulation analyses varying the number of degrees of freedom used for generating the pseudo seasonality curve. Inference on the air pollution coefficient β_0 under fully parametric GLM+NS and semi-parametric GAM-R appeared different. Even if the number of degrees of freedom used for fitting data was correctly specified (we used the same number of degrees of freedom for generating pseudo data and in model specification), the GAM-R estimator resulted strongly biased for high amounts of smoothing (relative bias=155.57% for df_0=3 per year, relative bias=70.48% for df_0=4 per year). The bias decreased as df_0 increased. On the contrary, the percent relative bias under GLM+NS ranged from 0.93 to 4.11. The real coverage of the 95% confidence intervals was ever close to 95% under GLM+NS, while it appeared unsatisfactory under GAM-R, in particular for low values of df_0 (5.9% and 61.55% when df_0 =3 and 4 per year, respectively).
This outcome could indicate a certain tendency of semi-parametric approach to be more appropriate in presence of strong seasonality in data. The beneficial effect of undersmoothing on the inference of the parametric component is emphasized by the improved performance of GAM-R if combined with GCV, which is well-known to result in undersmoothing.[32] With reference to this, it should be noticed that the number of degrees of freedom selected by GCV in the simulation analysis was always higher than the true one (Figure 4).
In a more extensive simulation analysis,[33,34] we found that in general bias increased as concurvity increased, but the size of bias depended on the modelling approach. For example, for a value of concurvity around 0.70 (not unusual in real datasets), β_0=0.0006 and df_0=5 per year, GAM-R strongly overestimated the effect of air pollutant (relative bias = 81.76%) and produced bad estimates of confidence interval (real coverage = 81.07%), while negligible bias (-4.24%) and good coverage of confidence interval (95.6%) were found under the fully parametric GLM+NS approach.
The performances of estimators depend also by the true size of air pollutant effect. For β_0=0.006, corresponding to an unrealistically strong increase of 6.2%, we found negligible bias under both approaches, but using GAM-R the bias quickly increased as the effect size decreased (not reported). Even in these situations, if models with a certain amount of undersmoothing were fit or GCV was used for selecting the number of degrees of freedom, performances of GAM-R improved. Finally, Table 3 shows percent bias of air pollution effect estimates and coverage of 95% confidence interval obtained using pre-defined percentages of undersmoothing under the semi-parametric GAM-R approach, for df_0 =3,5 (situations where strong and moderated amount of bias were observed, see Table 2), concurvity = 0.45 and β_0=0.0006. In general, the performances of GAM-R improved as the percentage of undersmoothing increased, with larger percentages needed when df_0 = 3 per year.

Discussion

Modelling epidemiological time series presents difficulties due to: (a) the small order of magnitude of the effects; and (b) the strong confounding effect of seasonality/time trend and weather.
Since the 1990s, the use of GAM became common, allowing flexible and local non-parametric modelling of confounders.[35] The statistical software commonly used to fit GAMs was based on backfitting algorithm, which can be affected by problems of convergence and produce biased effect estimates, depending on the degree of concurvity in data.[11-13] Moreover, coherently with Ramsay et al.[10] we found that estimated standard errors from GAM fitted by Splus (and SAS, results not shown), even with stringent convergence criteria, are invalid. These drawbacks can bring to bad inference, in particular when single cities analyses are conducted. Consequences on combined meta-analytic results are expected to be less serious.
Several alternatives to GAM with smoothing spline(s) via backfitting are possible and should be preferred.[36] GLM with parametric natural spline(s) can be specified. This solution is exempt of problems of convergence and variance covariance matrix estimation, but it requires that the number and position of knots to be specified a priori. This could be in principle a limitation for the applicability of such approach to many situations. To reduce this problem we used only one spline for time. Since time is an equally spaced covariate (it indexes days under study) and the number of knots is not small (between 5 to 7 per years), we do not expect relevant differences from different knots specifications.
A second alternative to GAM-S consists in fitting GAM with penalized regression spline(s) by the direct method implemented for R by Wood.[19] After having fixed appropriate convergence criteria and specified an appropriate large number of knots for the penalized regression spline, we found that the backfitting algorithm and the direct method provided similar

point estimates of pollutant effect. However, the direct method offers the advantages of requiring less computation for standard errors, that can be estimated without using approximation procedures, and being less sensitive to the choice of algorithm convergence criteria.
In the MISA data, under both approaches, meta-analytic overall point estimates did not appear sensitive to changing number of degrees of freedom for the spline, unless a very small number of degrees of freedom was specified (3 per year). This robustness was more evident when a random effects meta-analysis model was used and city-specific parametric models were specified.
Under the random effects models, a trade-off between overall effect and variance was observed, as documented also by Daniels et al.[37] and our prior choice of degree of freedom for the smoother balanced accuracy and precision of the effect estimate.
Differences between the estimated effects under GLM+NS and GAM-R were observed. These differences are consistent with published theoretical results and with the results of the simulation analysis.
The modelling approach adopted to adjust for temperature effect did not appear relevant, parametric «V» shaped function being a good alternative to non parametric modelling of the relationship between temperature and mortality (to address in detail this issue is outside the purpose of this paper; for recent approach see Welty and Zeger.[22]

Conclusion

Both GAM with penalized regression spline(s) and GLM with natural regression spline(s) are appropriate for analysis of short term effect of air pollution in epidemiological time series context. However, as documented in statistical literature,[21] the parametric approach retains better finite sample properties than the semi-parametric one in estimating the air pollutant effect (i.e. the parametric part of the model). In our example, GAM-R point estimates did not appear affected by relevant bias compared with GLM+NS, but major consequences could be observed in presence of large amount of concurvity in the data or if a small number of degrees of freedom for the seasonality spline(s) was used, as shown by the simulation analysis. In this sense the fully parametric approach is to be preferred. As an alternative, choosing the amount of smoothing by GCV or using a large number of degrees of freedom can reduce bias when semi-parametric models are specified.
In any case, a sensitivity analysis changing the number of degrees of freedom for the splines is to be recommended, regardless of modelling approach adopted. If we are interested in overall estimates and a random effects meta-analysis model is specified, a certain robustness of results is to be expected,[37] but very small number of degrees of freedom should be ever avoided.
The main drawback to using GLM+NS is the dependence of the fitted curve on the knots position. In theory, the use of GLM+NS can be critical if regression splines are defined for not equally spaced covariates, like meteorological variables. In this case, inference could be sensitive to knots placement. We advise to avoid use of complex spline-based modelling if simpler alternative are possible or perform sensitivity analyses changing positions of knots.

Conflitti di interesse: nessuno

References

1 Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery D, Schwartz J, Zanobetti A. *The National Morbidity, Mortality, and Air Pollution Study (HEI Project No. 96-7): morbidity and mortality from air pollution in the United States*. Health Effects Institute, Cambridge, MA, 2000.
2 Atkinson RW, Anderson HR, Sunyer J, et al. Acute effects of particulate air pollution on respiratory admissions: results from APHEA 2 project. Air Pollution and Health: a European Approach. *Am J of Respir Crit Care Med* 2001; 164: 1860-66.
3 Katsouyanni K, Touloumi G, Samoli E, Gryparis A, Le Tertre A, Monopolis Y et al. Confounding and effect modification in the short-term effects of ambient particles on total mortality: results from 29 European cities within the APHEA2 project. *Epidemiology* 2001; 12: 521-531.
4 Burnett RT, Cakmak S, Brook JR. The effect of the urban ambient air pollution mix on daily mortality rates in 11 Canadian cities. *Can J Public Health* 1998; 89: 152-56.
5 Ballester Diez F, Saez Zafra M, Perez-Hoyos S, et al. The EMECAM project: a discussion of the results in the participating cities. Estudio Multicentrico Espanol sobre la Relacion entre la Contaminacion Atmosferica y la Mortalidad. *Rev Esp Salud Publica* 1999; 73: 303-14 (spanish).
6 Saez M, Ballester F, Barcelo MA, Perez-Hoyos S, Bellido J, Tenias JM et al. A combined analysis of the short term effects of photochemical air pollutants on mortality within the EMECAM project. *Environ Health Perspect* 2002; 110: 221-28.
7 Hoek G, Brunekreef B, Verhoeff A, van Wijnen J, Fischer P. Daily mortality and air pollution in The Netherlands. *J Air Waste Manag Assoc 2000*; 50: 1380-89.
8 Kaiser J. Software Glitch Threw Off Mortality Estimates. *Science* 2002; 296: 1945-46.
9 Statistical error leaves pollution data up in the air. *Nature* 2002; 417: 677.
10 Ramsay TO, Burnett RT, Krewski D. The Effects of Concurvity in Generalized Additive Models Linking Mortality to Ambient Particulate Matter. *Epidemiology* 2003; 14: 18-23.
11 Dominici F, McDermott A, Zeger SL, Samet J. On the use of Generalized Additive Models in Time-series Studies of Air Pollution and Health. *Am J of Epidemiol* 2002; 156, 193-203.
12 Hastie TJ, Tibshirani RJ. *Generalized Additive Models*. London, Chapman & Hall, 1990.
13 Green PJ, Silverman BW. *Nonparametric Regression and Generalized Linear Models*. London, Chapman & Hall, 1994.
14 HEI Special report: *Revised Analyses of Time Series Studies of Air Pollution on Health*. Health Effects Institute, May 2003.
15 Biggeri A, Bellini P, Terracini B. Meta-analysis of the italian studies on short-term effects of air pollution. *Epidemiol Prev* 2001; 25(suppl): 1-72 (italian).
16 Biggeri A, Baccini M, Accetta G, Lagazio C. Estimates of short-term effects of air pollutants in Italy. *Epidemiol Prev* 2002; 26(4): 203-05 (italian).

17 Biggeri A., Baccini M, Bellini P, Terracini B. Meta-analysis of the Italian studies on short-term effects of air pollution (MISA) 1990-1999. *Int J Occup Environ Health* 2005; 11(1): 107-22.
18 the R Development Core Team. 2006. R: a language and environment for statistical computing. ISBN 3-900051-07-0. http://www.r-project.org
19 Wood SN. Modelling and Smoothing Parameter Estimation with Multiple Quadratic Penalties. *J R Stat Soc [Ser B]* 2000; 62:413-28.
20 Currie ID, Durban M. Flexible smoothing with P-splines: a unified approach. *Statistical Modelling* 2002; 4: 333-49.
21 Rice J. Convergence rates for partially splined models. *Statist. Probabil. Letters* 1986; 4, 203-08.
22 Welty LJ, Zeger SL. *Flexible Distributed Lag Models: Are the Acute Effects of PM10 on Mortality the Result of Inadequate Control for Weather and Season?* Johns Hopkins University Department of Biostatistics Working Papers, Working Paper 38, 2004.
23 Kelsall J, Samet J, Zeger S. Air pollution and mortality in Philadelphia, 1974-1988. *Am J of Epidemiol* 1997; 146: 750-62.
24 de Boor C. *A Practical Guide to Splines*. New York, Springer Verlag, 1978.
25 McCullagh P, Nelder JA. *Generalized Linear Models* (2nd edition). London, Chapman Hall, 1989.
26 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7: 177-88.
27 Saurina C, Barcelo MA, Saez M, Tobias A. The short-term impact of air pollution on the mortality. Results of the EMECAM project in the city of Barcelona, 1991-1995. *Rev Esp Salud Publica* 1999; 73: 199-207 (in Spanish).
28 Hurvich CM, Simonoff JS, Tsai CL. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *J R Stat Soc [Ser B]* 1998; 60: 271-93.
29 Heckman N. Spline Smoothing in partly linear model. *J R Stat Soc [Ser B]* 1986; 48, 244-48.
30 Cuzick J. Semiparametric additive regression. *J R Stat Soc [Ser B]* 1992; 54: 831-43.
31 Speckman, P. Kernel smoothing in partial linear models. *J R Stat Soc [Ser B]* 1988; 50, 413-36.
32 Hurvich CM, Simonoff JS, Tsai CL. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *J R Stat Soc [Ser B]* 1998; 60: 271-93.
33 Baccini M, Biggeri A, Lertxundi A, Saez M. Semi-parametric and parametric approaches in the analysis of short term effects of air pollution on health. *Epidemiology* 2003; 14: S64.
34 Baccini M, Biggeri A, Lagazio C, Lerxtundi A, Saez M. Parametric and Semi-Parametric approaches in the analysis of short-term effects of air pollution on health. *Comput Stat Data Anal* (in press).
35 Schwartz J. The use of generalized additive models in epidemiology. *Proceedings in XVII International Biometric Society Conference, Hamilton, Ontario* 1994; 55-80.
36 Lumley T, Sheppard L. Time series analyses of air pollution on health: straining at gnats and swallowing camels? *Epidemiology* 2003; 14: 13-14.
37 Daniels MJ, Dominici F, Samet S. Underestimation of standard errors in multisites time series studies. *Epidemiology* 2004; 15: 57-62.