Working Paper Series, N. 2, January 2008

Short term ozone effects on morbidity for the city of Milano, Italy, 1996-2003

Monica Chiogna
Department of Statistical Sciences
University of Padua
Italy

Francesco Pauli
Department of Statistical Sciences
University of Padua
Italy

Abstract: In this paper, we explore a range of concerns that arise in measuring short term ozone effects on health. In particular, we tackle the problem of measuring exposure using alternative daily measures of ozone derived from hourly concentrations. We adopt the exposure paradigm of Chiogna and Bellini (2002), and we compare its performances with respect to traditional exposure measures by exploiting model selection. For investigating model selection stability issues, we then apply the idea of bootstrapping the modelling process.

Keywords: ozone; morbidity; threshold; bootstrap.

Contents

1	**Introduction**	**1**
2	**Data**	**2**
3	**Methods**	**3**
4	**Model Selection**	**6**
5	**Results**	**7**
6	**Discussion**	**13**

Department of Statistical Sciences
Via Cesare Battisti, 241
35121 Padova
Italy

tel: +39 049 8274168
fax: +39 049 8274170
http://www.stat.unipd.it

Corresponding author:
Francesco Pauli
tel: +39 049 827 4127
francesco.pauli@unipd.it
http://www.stat.unipd.it/~pauli

Short term ozone effects on morbidity for the city of Milano, Italy, 1996-2003

Monica Chiogna
Department of Statistical Sciences
University of Padua
Italy

Francesco Pauli
Department of Statistical Sciences
University of Padua
Italy

Abstract: In this paper, we explore a range of concerns that arise in measuring short term ozone effects on health. In particular, we tackle the problem of measuring exposure using alternative daily measures of ozone derived from hourly concentrations. We adopt the exposure paradigm of Chiogna and Bellini (2002), and we compare its performances with respect to traditional exposure measures by exploiting model selection. For investigating model selection stability issues, we then apply the idea of bootstrapping the modelling process.

Keywords: ozone; morbidity; threshold; bootstrap.

1 Introduction

That a high level of ozone is dangerous for health is nowadays an acquired fact. Three recent meta-analyses commissioned by U.S. Environmental Protection Agency, Bell et al. (2005), Ito et al. (2005) and Levy et al. (2005), consider, respectively, 39, 28 and 43 epidemiological studies of the effect of ozone on mortality, finding coherent results.

A number of studies, showing analogous results, consider hospital admissions as well as mortality; Medina-Ramón et al. (2006) is a very recent example of a multicity study, Biggeri et al. (2001) and Biggeri et al. (2004) are examples based on data from Italy.

Results similar to those found for general non-accidental mortality and morbidity are found analyzing events due to respiratory (sometime, particular subclasses) causes Anderson et al. (2004), whereas it is relatively rare to consider cerebrovascular causes.

Epidemiological studies such as those above cited consider as a measure of population exposure to ozone (pollutant, in general) a function of the concentrations of ozone measured, usually hourly, at fixed sites generally located in the city. This function is usually a daily summary such as 1-hour maximum, day average or 8-hours average. This, as evidenciated by WHO (2003), might pose a problem, since such summary may be not representative of actual personal exposure of individuals for various reasons, for example because people do not spend all of their time outdoor or because of a non uniform spatial distribution of the pollutant, the latter remark being particularly true for ozone. Moreover, the usual daily measures may be inadequate summaries of the daily pattern of concentration. Despite these problems, the choice of a daily measure is usually overlooked in published studies, where one of the usual summaries is chosen as if they were equivalent, and no attempt is made to seek a better (in representing exposure) measure.

In order to investigate this aspect, we shall restrict our attention to morbidity and we shall measure exposure by considering alternative daily measures of ozone derived from hourly concentrations, which are described in Section 3. In fact, daily variations of the hourly values of ozone concentration generally show a well-marked daily pattern with the maximum occurring in the early afternoon. By taking into account this behaviour, we adopt the exposure paradigm of Chiogna and Bellini (2002), and we compare its performances with respect to traditional exposure measures by exploiting model selection.

For investigating model selection stability issues, we then apply the idea of bootstrapping the modelling process, as described in Section 4, although we do not attempt to incorporate model uncertainty into the estimates of ozone effects.

Finally, Section 5 describes the application on real data and Section 6 summarizes our findings. All computations have been performed using the freely available statistical software R R Development Core Team (2006).

2 Data

It is customary in investigating the relationship between ozone and mortality or morbidity to limit analysis to the summer period. This is done either because ozone is produced by a chemical reaction driven by solar radiation and so its concentration reaches dangerous levels only during summer when solar radiation is higher; and because indoor concentration is much lower than the outdoor one, so population is more exposed during summer (when more time is spent outdoors). For this reason, we consider data for summer periods only (June-July-August).

In order to match population exposure and collected data, which are measured at stations in the city center, we consider hospital admissions for people resident in the town of Milano only. Moreover, since it is generally believed that high pollutant levels can adversely affect health conditions of population groups which are already weakened, like the elderly, the youngest and those with chronic diseases, we focus the analysis on people aged more than 75 years at the time of the event and we exclude hospital admissions due to all accidental causes (codes ICD-IX 800-999).

Daily hospital admissions data for the period 1995-2003 were obtained from the Regional Health Informative System, for all hospitals located in Milano (Figure 1). In order to consider relevant events, only admissions not required by the general practitioner, not related to a surgical event and not scheduled to last less than one day were selected. We excluded events for which the reason for admission was not specified.

Meteorological and environmental data for the same period were obtained from the Regional Agency for Environmental Protection (ARPA) of Lombardia, which collects hourly data on temperature, rain, wind velocity and direction and, from year 2000, humidity. The same agency collects data about air pollution. The monitoring network consists of eight stations (see `http://www.arpalombardia.it/qaria/` for details). Concentration of

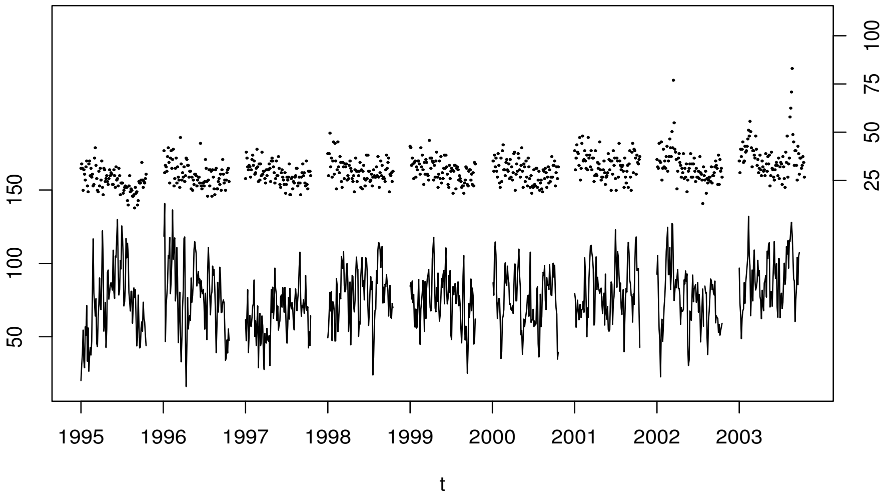


Figure 1: Summer time series of daily average ozone concentration (continuous line, left y axis), and daily number of hospital admissions (dots, right y axis).

Ozone (Figure 1) is measured at three sites (called Juvara, Parco Lambro, Verziere after the toponomymy). Missing data are present due to temporary inactivity of monitoring stations. In the analysis only days with at least 75% of hourly data were considered.

3 Methods

Let $Y_{k,t}$ be the number of hospital admissions, in day t for age class k, where, upon preliminary analysis, we consider classes 74-89 and 90-ω. We assume

$$Y_{k,t} \sim \text{Poisson}(\lambda_{k,t}), \tag{1}$$

with

$$\log(\lambda_{k,t}) = \beta_{0,k} + \text{confounding}(k,t) + \text{ozone}(t), \tag{2}$$

where confounding(k,t) is a term including all variables relevant to confounding control, and ozone(t) is a function of ozone concentration measuring the effect of the pollutant.

As for confounding, we let confounding$(k,t) = f(t) + g_k(T_t) + z_t + \gamma h(t) + \alpha w(t)$, where $f(t)$ is a smooth function of time, T_t is the mean of day temperature in the previous three days and $g_k(T_t)$ is a age class-specific smooth function of it, z_t is the daily average of PM_{10} concentration in day t, $h(t)$ is a holiday indicator, and $w(t)$ associates to t an integer value corresponding to the day of the week (monday=1, …, saturday=6, sunday is base value). Such confounding control is quite customary. Temperature is traditionally present in models for health impact of pollution and has a significant effect on health conditions of the elderly. Similarly, particulate matter (PM_{10}) is usually included as effect modifier, whereas other pollutants are generally ignored (Bell et al. (2005), Ito et al. (2005), Levy et al. (2005)).

It is worth saying that, at first, more complicated formulations were tried. Complications included both a finer stratification (age classes: 75-80, 80-85, 85-90, 90-ω) and the use of other explanatory variables such as humidity or temperature difference between the current day and the mean lagged over the previous three days. Such complications, however, did not significantly improve goodness of fit, so they were not considered in order to reduce the computational burden. In particular, we ignored humidity because of lack of data (available from 2000) and its non significance in the available periods.

Traditionally, effect of ozone in the linear predictor is considered to be proportional to the daily average concentration or to the maximum of the daily hourly concentrations, i.e. $\text{ozone}(t) = \beta o_t$, where o_t is the chosen daily indicator. In this work, aside of these standard summaries of ozone daily pattern, we consider other alternatives, which try to measure other features of the daily exposure, potentially more suitable to capture effects on the health status. By using a daily average or the maximum, in fact, we summarize in one figure the daily behaviour of ozone, thus potentially ignoring relevant aspects such as excursions or persistence of high values of the concentration.

In Figure 2(a) we show the daily pattern of hourly concentration by plotting the distributions of hourly concentration for a summer month (August). Given this average pattern, day to day variability can be observed, as seen in Figure 2(b), where two hypothetical daily patterns of concentration are shown.

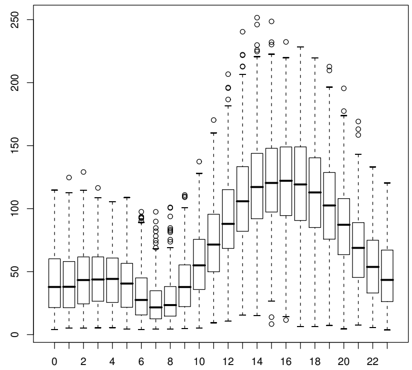
Taking into account such behaviour and following Chiogna and Bellini (2002), the idea that we pursue here is to approximate exposure to ozone by substituting the one-figure summary with a set of indicators allowing to grasp some aspects of the shape of the concentration curve.

In particular, given a threshold S, we consider three measures, denoted as d, i and m, where d is the number of hours during which the daytime hourly concentration is above S, i is the difference between the daytime maximum hourly concentration and S, m is the nighttime average concentration. Figure 2(c) shows a typical daily pattern and the values of the three measures for that day. In the figure, daytime is defined to be between 8 am to 9 pm. As it can be seen, i and d measure the extent and the duration of the exceedance of the threshold S: it is clearly seen that they allow to approximate the area of the portion of the concentration curve above the threshold. Using the new measures, we are now able to distinguish between the two patterns shown in Figure 2(b), where the maximum is the same but the behaviour of the pollutant is clearly different.

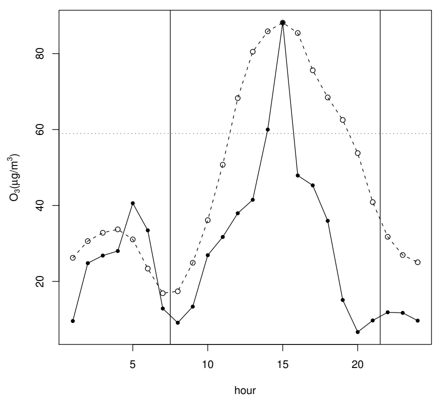
The new measures are included as linear contribution in the term $\text{ozone}(t)$; furthermore, we consider also substituting the pair d, i by the product of the twos. For such measures, we consider also the lagged values over the previous three days, conventionally denoted by the suffix (lag). In Table 1, all alternative formulations for the term $\text{ozone}(t)$ are listed.

As a final remark, it is worth noting that the use of d and i implies a non linear (threshold) effect of ozone on health, as such measures are null for days with concentrations below the threshold. This makes such indicators particularly interesting. Some authors (see, for example, Kim et al. (2004)) explicitely question the linearity assumption, to conclude that a threshold model does in fact perform better (threshold for summer period is at 40ppb). Nevertheless, the existence of a threshold for short term ozone effect is still an open issue.

(a)



(b)



(c)

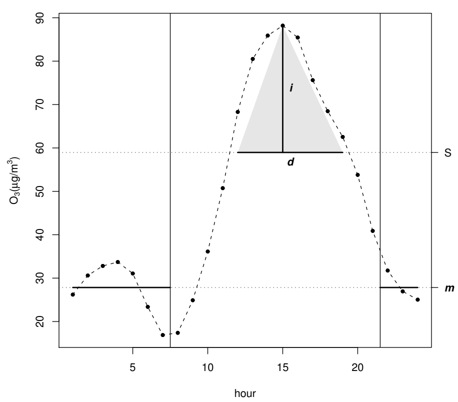


Figure 2: (a): distributions of hourly concentrations for the month of August; (b): comparison of possible one day pattern implying the same maximum but different population exposure; (c): typical one day pattern of O_3 concentration (black dots connected by dashed line) and representation of measures d, i and m corresponding to threshold S; vertical lines limit the daytime hours.

	model
1	$base$
2	ave
3	max
4	$AVE^{(lag)}$
5	$MAX^{(lag)}$
6	m
7	$m^{(lag)}$
8	i
9	d
10	di
11	$i^{(lag)}$
12	$d^{(lag)}$
13	$di^{(lag)}$
14	$i+m$
15	$d+m$
16	$di+m$
17	$i+m^{(lag)}$
18	$d+m^{(lag)}$
19	$di+m^{(lag)}$
20	$i^{(lag)}+m$
21	$d^{(lag)}+m$
22	$di^{(lag)}+m$
23	$i^{(lag)}+m^{(lag)}$
24	$d^{(lag)}+m^{(lag)}$
25	$di^{(lag)}+m^{(lag)}$

Table 1: Ozone related model component, in other words, terms to be included in ozone(t). ave stands for daily average, max for daily maximum.

4 Model Selection

Two typical criteria to select among models of different complexity are the Un-Biased Risk Estimate (UBRE) and the Generalized Cross Validation (GCV), which are suggested in Wood (2000) in the context of Generalized Additive Models (GAMs) for choosing the degree of smoothness of non linear components and for variable selection. The first, in particular, is suggested when the scale parameter of the GAM model is known, as in the Poisson case. The formula for the UBRE score, given that the scale parameter is 1, is

$$\frac{D}{n} + 2\frac{p}{n}, \tag{3}$$

where D is the deviance, i.e., twice the difference between the log-likelihood for the saturated model and the log-likelihood for the present model, and p is the total degrees of freedom (including the estimated d.o.f. of the smooth functions). With the same notation, the GCV criterion is given by

$$\frac{nD}{(n-p)^2}. \tag{4}$$

The two criteria are implemented in the package mgcv within the software R.

It should be noted that, conditionally on the choice of one such criterion, the model selection process itself is a source of uncertainty in final estimates, especially if cardinality of the set of alternative models is high. In other words, in a setting like ours, where we consider one model not involving ozone and tens of models including some ozone function, we can not rule out the possibility that a model involving ozone is selected over the model not involving it only by chance.

In order to investigate the uncertainty due to model selection (in our case, uncertainty in the estimate of the selection criterion), we adopt an approach mutuated from Sauerbrei (1999), who suggests to replicate the model selection process on bootstrap resamples and to consider as an indicator of the worthiness of a model its frequency of selection. A similar approach has also been taken by other authors such as Buckland et al. (1997), Veall (1992) and Freedman and Navidi (1986). Despite the lack of a formal justification for such a procedure, it seem to us an effective method to attach a measure of reliability to the outcome of the model selection process, a need which has been emphasized by many authors in different contexts (Chatfield (1995), Clyde (2000)).

5 Results

For the selected summer window in years 1995-2003, we considered 2 outcomes, i.e., all admissions and respiratory admissions. To compute the exposure measures, we considered as daytime the interval 8am - 9pm and we computed the exposure measures d and i for S equal to: the mean of hourly ozone concentrations (58.96$\mu g/m^3$), denoted by M, their third empirical quartile (85.90$\mu g/m^3$), denoted by $3Q$, and the law alarm threshold (120$\mu g/m^3$), denoted by L. Table 2 shows the number of days for which measures d and i are non zero, i.e., the number of exceedances.

	None	M	3Q	L
Threshold value ($\mu g/m^3$)		58.96	85.90	120
Admission data	956	942	896	696

Table 2: Number of observations for which measures d and i are greater than zero for each threshold.

To compute the indicators, hourly data from different stations were first aggregated over space (only stations showing no more than 25% missing data were considered).

Table 3 shows for threshold M pairwise collinearity among variables entering ozone(t), which is useful to discharge redundant variables. It is worth noting that variables d and i are highly correlated (this is an effect of the typical behaviour of ozone during the day), so inclusion of both may lead to an issue. This is the reason why we never linearly insert both indicators in the term ozone(t) (see Table 1).

The model space included 61 models: seven model (1-7 in Table 1) involving traditional measures or m, which does not depend on the threshold S; the remaining 54 models (8-25 in Table 1) are defined for each possible threshold (M, 3Q, L in pictures). Selection of smoothness for non linear terms in the predictor was made using the UBRE score; model

	d	i	m	max	$d^{(lag)}$	$i^{(lag)}$	$m^{(lag)}$	$max^{(lag)}$	di	$di^{(lag)}$	ave	$ave^{(lag)}$
d	–	.59	.41	.35	.23	.30	.04	.47	.69	.45	.49	.52
i	.59	–	.18	.61	.28	.51	$-.06$	.79	.97	.66	.68	.62
m	.41	.18	–	$-.02$	.36	.43	.37	.40	.28	.45	.24	.62
max	.35	.61	$-.02$	–	.11	.31	$-.14$	.51	.59	.40	.79	.35
$d^{(lag)}$	.23	.28	.36	.11	–	.66	.49	.55	.29	.57	.20	.71
$i^{(lag)}$	.30	.51	.43	.31	.66	–	.28	.89	.52	.83	.39	.85
$m^{(lag)}$	.04	$-.06$	.37	$-.14$	.49	.28	–	.11	$-.03$	.13	$-.03$	.47
$max^{(lag)}$	.47	.79	.40	.51	.55	.89	.11	–	.80	.91	.58	.87
di	.69	.97	.28	.59	.29	.52	$-.03$	.80	–	.69	.69	.67
$di^{(lag)}$	.45	.66	.45	.40	.57	.83	.13	.91	.69	–	.49	.85
ave	.49	.68	.24	.79	.20	.39	$-.03$	.58	.69	.49	–	.48
$ave^{(lag)}$	.52	.62	.62	.35	.71	.85	.47	.87	.67	.85	.48	–

Table 3: Correlation matrix of covariates computed with threshold M.

selection was performed using both UBRE and GCV. As the two criteria led to the same model for all outcomes, and taking into account that in bootstrap resampling differences in frequencies were ignorable, in what follows we shall always refer to UBRE.

Estimated coefficients for pollutant concentrations in models selected as best according to UBRE criterion are reported in Table 4. It is immediate to see that the best models included the new exposure measures.

For each considered outcome, we then generated 500 bootstrap replications and for each of them we selected the best model. The frequencies of selection for models in Table 1 is reported in Figure 3. Overall, for all admissions threshold L was selected 48% times, $3Q$ 5% times and M 47% times; for respiratory admissions the percentages were, respectively: 55% , 7% , 38% .

Figure 3 shows the selection frequencies of models appearing in bootstrap replications. The bootstrap selection led to the model selected as best on the original data for either admissions due to all causes and admissions due to respiratory causes. The model without ozone was never the best in any bootstrap replication.

In Table 4 we report the coefficients for the models selected on the original data and on bootstrap resampling and for models involving traditional measures (daily average and daily maximum). For respiratory admissions, we notice that, contrary to what happens to the new indicators, the effects of the traditional measures are not significantly different from zero when modelling.

To perform some residuals analysis on final models, we considered normal probability plots of randomized residuals, that is, of quantities

$$r_i = (1 - u_i)F(y_i - 1; \hat{\lambda}_i) + u_i F(y; \hat{\lambda}_i), \tag{5}$$

where u_i are IID uniform r.v. on $[0, 1]$, F is the Poisson distribution function and $\hat{\lambda}_i$ its estimated parameter. The normal probability plots (see Figure 4) do not evidenciate neither significant departures from normality, nor differences among models.

In studies on short term health effects of air pollution, it is customary to report final results on pollution effects in terms of relative risk (RR) corresponding to a fixed increase in the pollutant concentration. The RR is computed as the ratio of average number of episodes for the increased pollutant concentration and the average number of episodes for the base concentration level; for the special case of the Poisson model, the RR is simply

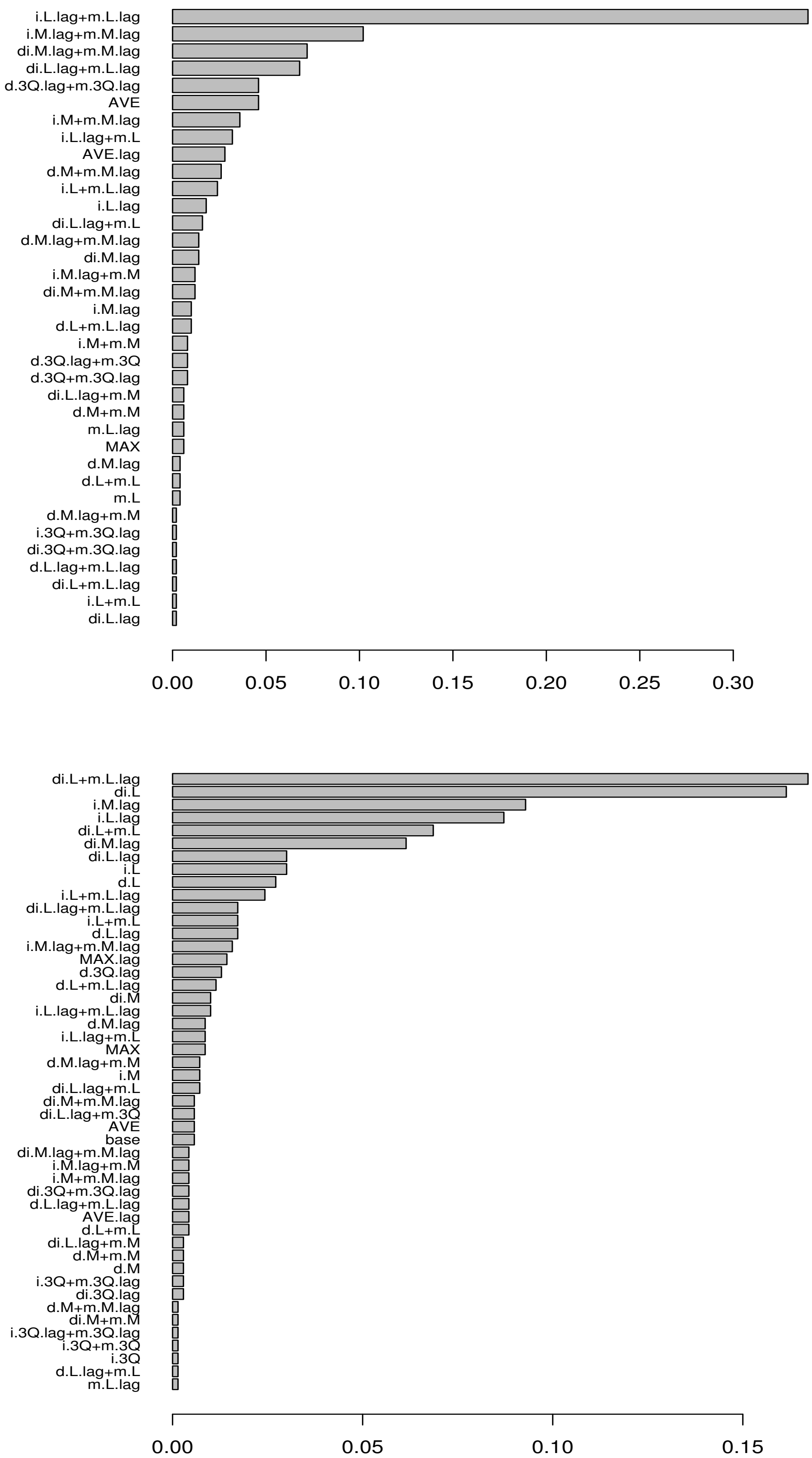
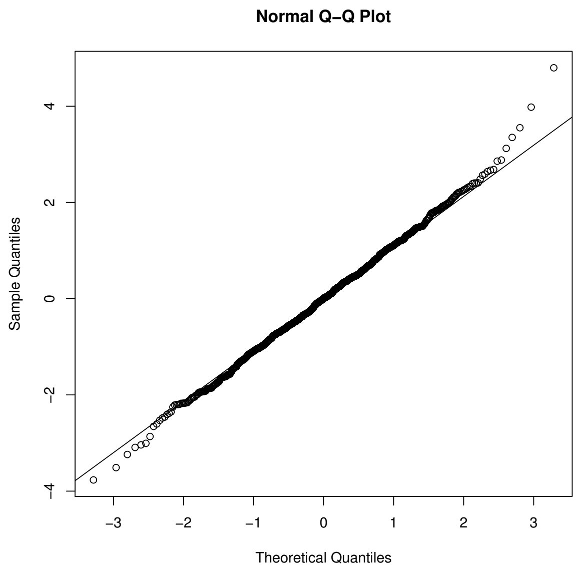


Figure 3: Frequency of model selection in bootstrap resampling for admissions for all causes (upper panel) and admissions for respiratory causes (lower panel).

(a)



(b)

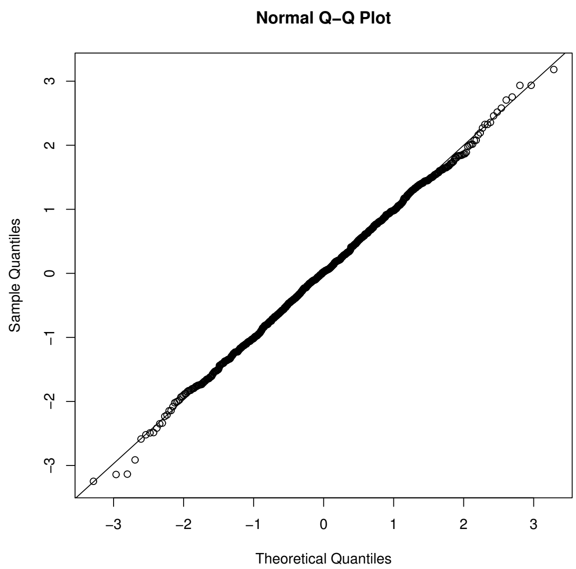
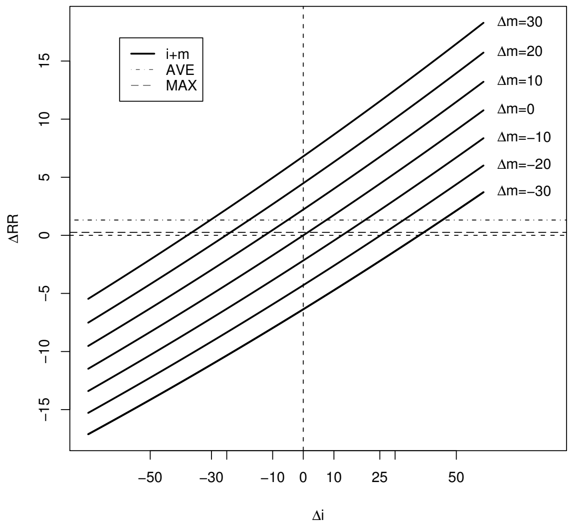


Figure 4: Normal probability plots of randomized residuals for models *(a)* admissions for all causes, *(b)* admissions for respiratory causes.

Admissions due to all causes



Admissions due to respiratory causes

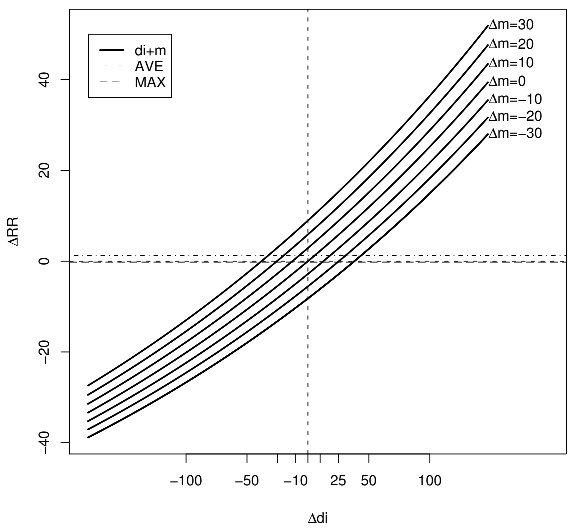


Figure 5: Change in relative risk due to a variation in ozone concentrations measures. Each of the thick lines represents the percentage change in RR as a function of the variation in i or di measure (for all admissions and respiratory admissions respectively) for a fixed variation in m measure; horizontal lines represent the change in RR according to models based on average and maximum value.

Variable	Estimate	s.e.	z-stat	p-value
	all admissions			
PM_{10}	0.0034439	0.0007272	4.736	0.0000
$i_L^{(lag)}$	0.0017362	0.0005519	3.146	0.0017
$m^{(lag)}$	0.0021925	0.0005971	3.672	0.0002
PM_{10}	0.0034439	0.0007272	4.736	0.0000
$i_L^{(lag)}$	0.0017362	0.0005519	3.146	0.0017
$m^{(lag)}$	0.0021925	0.0005971	3.672	0.0002
PM_{10}	0.0026390	0.0007584	3.480	0.0005
ave	0.0013092	0.0005074	2.580	0.0099
PM_{10}	0.0028594	0.0008195	3.489	0.0005
max	0.0002530	0.0002613	0.968	0.3329
	respiratory admissions			
PM_{10}	0.0026387	0.0018993	1.389	0.1647
di_L	0.0002510	0.0000913	2.749	0.0060
$m^{(lag)}$	0.0028568	0.0015327	1.864	0.0623
PM_{10}	0.0026387	0.0018993	1.389	0.1647
di_L	0.0002510	0.0000913	2.749	0.0060
$m^{(lag)}$	0.0028568	0.0015327	1.864	0.0623
PM_{10}	0.002366	0.0018986	1.246	0.2126
di_L	0.000225	0.0000902	2.495	0.0126
PM_{10}	0.0027482	0.0019469	1.412	0.1581
ave	0.0012376	0.0013243	0.935	0.3500
PM_{10}	0.0035933	0.0021082	1.704	0.0883
max	−0.0002017	0.0006769	−0.298	0.7657

Table 4: For each outcome, estimated coefficients for the best model on the original sample, best bootstrap choice, models based on traditional measures.

the exponential of the increase. Using the above described non-standard measures, we can not univocally associate an increase of ozone concentration to a relative risk, because the increase in ozone concentration might differently act on the ozone indicators, increasing only one of the three measures, or two of them, or all three of them.

RRs can, nevertheless, be obtained from the estimates of the coefficients of the linear predictor. In Figure 5, we report, for the best models according to bootstrap, the percentage change in risk due to changes in the indicators of ozone concentration. For a fixed variation in m, Δm, with $\Delta m = \pm 10, \pm 20, \pm 30 \mu g/m^3$, each of the thick lines represents the change in RR, ΔRR, as a function of the variation in i, Δi, or in di, Δdi. As a term of comparison, for models based on the traditional ozone indicators, the changes in RR due to an increase of $10 \mu g/m^3$ have been added to the plots (see the horizontal dotted lines). Such plots seem to us particularly interesting. For admissions due to all causes, for example, they show that the same variation in RR happens if the night average stays the same and i increases of $10 \mu g/m^3$, or if the night average drops down of $10 \mu g/m^3$and i increases of roughly $25 \mu g/m^3$. In other words, they allow to explore which changes in the daily patterns of concentration are relevant and how they affect health.

6 Discussion

In this paper, we have attempted to explore a range of concerns that arise in measuring short term ozone effects on morbidity. In particular, we have tackled measuring of exposure, as discussed in Section 3, and we have tried to address model selection uncertainty, as discussed in Section 4.

For measuring exposure, we have used three indicators, which grasp some aspects of the ozone concentration curve. Advantages of the new measures include the following. First, they allow to easily incorporate a threshold effect of ozone. This can be particularly important. According to Martuzzi et al. (2006), there is at present no evidence of such a threshold; the use of a threshold model (at 35ppb, approximatively $70 \mu g/m^3$) is, however, suggested in order to get round the uncertainty on the shape of the concentration response curve at low concentration levels (UNECE (2004)). Second, and perhaps most important, they allow to grasp which variations of concentration are relevant for health. The fact that, for respiratory admissions, only the new indicators lead to an estimated ozone effect which is significantly different from zero, may be a confirmation that the proposed exposure paradigm effectively reflects important aspects of exposure.

For addressing model selection uncertainty, we have relied on bootstrap validation. Beside substantially confirming results on the original samples, bootstrap analysis indicates that there are ozone effects and these are better grasped by the new measures. By selecting a model from a (relatively) high number of alternatives as is done above, one still runs the risk of obtaining estimates of pollution effect which are upward biased. One possibility is to limit alternatives by considering a (relatively) reduced number of models (as Gryparis et al. (2004), for example, who decide to explore only a limited number of lags). Thus, one avoids bias but renounces exploring alternative models which are legitimate. Final results would depend, however, on the definition of the model space.

Finally, it may be interesting to compare results here obtained with findings in Biggeri et al. (2004), being Milano one of the cities considered by the authors. Estimates of RR change due to an increase of $10\mu g/m^3$ in O_3 concentration in hot period in Milano (see Figure 7a and 7b in the study) are approximately +2% for admissions due to respiratory causes, whereas estimates for admissions due to all causes are not available. This value is comparable with RR reported in Figure 5 for admissions. In doing the comparison it must, however, be kept in mind that the summer period in Biggeri et al. (2004) included also May and September, which are instead excluded here.

References

Anderson, H., R. Atkinson, J. Peacock, L. MArston, and K. Konstantinou (2004). *Meta-analysis of time-series studies and panel studies of Particulate Matter (PM) and Ozone (O_3)*. World Health Organization EUR/04/5042688.

Bell, M., F. Dominici, and J. Samet (2005). A meta-analysis of time-series studies of ozone and mortality with comparison to the national morbidity, mortality, and air pollution study. *Epidemiologys 16*(4), 436–445.

Biggeri, A., P. Bellini, and B. Terracini (2001). Meta-analysis of the italian studies on short-term effects of air pollution. *Epidemiologia e prevenzione 25*(2).

Biggeri, A., P. Bellini, and B. Terracini (2004). Meta-analysis of the italian studies on short-term effects of air pollution 1996-2002. *Epidemiologia e prevenzione 28*(4).

Buckland, S., K. Burnham, and N. Augustin (1997). Model selection: an integral part of inference. *Biometrics 53*, 603–618.

Chatfield, C. (1995). Model uncertainty, data mining and statistical inference. *Journal of Royal Statistical Society A 158*(3), 419–466.

Clyde, M. (2000). Model uncertainty and healht effect studies for particulate matter. *Environmetrics 11*, 745–763.

Freedman, D. and W. Navidi (1986). Regression models for adjusting the 1980 census. *Statistical Science 1*(1), 3–39.

Gryparis, A., B. Forsberg, K. Katsouyanni, T. G. Analitis, A., J. Schwartz, E. Samoli, S. Medina, H. Anderson, E. Niciu, H. Wichmann, B. Kriz, M. Kosnik, J. Skorkovsky, J. Vonk, and Z. Dörtbudak (2004). Acute effects of ozone on mortality from the “air pollutiion and health: a European approach” project. *American journal of respiratory critical care medicine 170*, 1080–1087,.

Ito, K., S. De Leon, and M. Lippmann (2005). Associations between ozone and daily mortality analysis and meta-analysis. *Epidemiologys 16*(4), 446–457.

Kim, S., J. Lee, Y. Hong, K. Ahn, and H. Kim (2004). Determining the threshold effect of ozone on daily mortality: an analysis of ozone and mortality in Seoul, Korea, 1995-1999. *Environmental research 94*, 113–119.

Levy, J., S. Chemerynski, and J. Sarnat (2005). Ozone exposure and mortality an empiric bayes metaregression analysis. *Epidemiologys 16*(4), 458–468.

Martuzzi, M., F. Mitis, I. Iavarone, and M. Serinelli (2006). *Health impact of* PM_10 *and ozone in 13 italian cities*. World Health Organization.

Medina-Ramón, M., A. Zanobetti, and J. Schwartz (2006). The effect of ozone and PM10 on hospital admissions for pneumonia and chronic obstructive pulmonary disease: A national multicity study. *American Journal of Epidemiology 163*(6), 579–588.

R Development Core Team (2006). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Sauerbrei, W. (1999). The use of resampling methods to simplify regression models in medical statistics. *Applied Statistics 48*, 313–329.

UNECE (2004). *Modelling and assessment of the health impact of particulate matter and ozone.* United Nations Economic Commission for Europe, Geneva.

Veall, M. (1992). Bootstrapping the process of model selection: an econometric example. *Journal of Applied Econometrics 7*, 93–99.

WHO (2003). *Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide: report on a WHO working group*. WHO Bonn, Germany 13-15 January 2003. Copenhagen, WHO Regional Office for Europe.

Wood, S. N. (2000). Modelling and smoothing parameter estimation with multiple quadratic penalties. *Journal of the Royal Statistical Society B 62*, 413–428.