

Modelling short-term, air pollution health effects, using air quality indices⁽¹⁾

Effetti a breve termine dell'inquinamento sulla salute ed indici di qualità dell'aria

Francesca Bruno, Rossella Miglio
Dipartimento di Scienze Statistiche "P. Fortunati"
Università degli Studi di Bologna
bruno@stat.unibo.it, miglio@stat.unibo.it

Riassunto: Negli ultimi anni molti studi sono stati condotti al fine di valutare la presenza di una eventuale associazione tra inquinamento dell'aria e mortalità. La relazione tra le concentrazioni di inquinanti e mortalità (o morbosità) nel medesimo giorno, o nei giorni successivi, viene usualmente specificata mediante modelli di regressione in cui la variabile di risposta è il numero di morti (o ricoveri) giornalieri e in cui le variabili esplicative, in aggiunta agli inquinanti, sono variabili ambientali e meteorologiche. Generalmente per evitare problemi di collinearità le analisi vengono condotte per i singoli inquinanti. Nel presente lavoro si indaga come sia possibile tener conto dell'effetto di tutti gli inquinanti mediante l'uso di indici di qualità dell'aria come variabile esplicativa, in alternativa o in aggiunta al singolo inquinante.

Keywords: Air Quality Indices, Poisson Models, Air Pollution Short-Term Effects

1. Introduction

Epidemiological studies over the past decade suggest the existence of a persistent statistical association between increases in daily air pollutant concentration and increases in mortality (particularly, non-accidental mortality in elderly people).

Typically, generalized additive models for Poisson data are used to model daily mortality, adjusting for non-linear trends using smoothing splines of other semi-parametric approaches, and including smoothed or parametric function and lags of meteorological variables and pollution variables.

In the context of a single time-series analysis it is difficult to make reliable inference about individual pollutants. Complicating factors in distinguishing between individual pollutants effects, include high correlation between the pollutant and the presence of substantial, possibly correlated, measurement errors. More appropriate statistical methods are needed to evaluate the health effects of each air pollutant under several conditions: the pollutant acting alone, the pollutant acting in the presence of a variety of mixtures of other air pollutants. Some studies examine the effects of a specified pollutant stratified by the level of other pollutants (Schwartz et al., 1996), but in using this method it is only possible to study two pollutants at the same time. In order to consider models with several pollutants, avoiding collinearity problems, possible

⁽¹⁾ The present paper is financially supported by a 2000 grant (project n.26, sector: Economics and Statistics) for researches of national interest by the Italian Ministry of the University and Scientific and Technological Research.

alternative methods include: grouping or combining pollutants as defined by source profiles, or, grouping or combining pollutants as defined by principal component or factor analysis (Cox, 2000).

Our proposal is to consider pollutant combination as an air quality index. This index could be considered alone or with a single pollutant. In the first case, the possibility of making some inferences about different health effects of different pollutants is lost, even if useful information about overall effects may still be found and could be particularly important for intervention in public health policy. In the other case, it could be possible to control, by means of the index, the effects of the pollutant for different levels of air pollution due to the other pollutants. In this paper, data from Bologna are analysed, by using generalized additive models and air quality indices as predictor with, or in alternative to, a single pollutant. In section 2 we briefly describe the air quality indices. In section 3 additional information on the data and variables are given and we also compare the results obtained with a single pollutant and/or the air quality index.

2. Air Quality Indices

In this study we used the kind of index contained in the general class of air quality indices proposed in Bruno and Cocchi (2002). This class of indices can be constructed by simplifying the dimensions of the data through successively aggregating functions. Usually, the three dimensions of the air pollution data are: time, space and type of pollutant.

The first synthesis is computed on time, since the reference time is not the day. For this reason we propose to use the syntheses established by the Italian Law for the environment: daily mean, hourly maximum per day and daily maximum of the 8 mobile-averages. Using the same notation as Bruno and Cocchi (2002) we obtain:

$$X_{q\{ij\}} = q_h(X_{ijh}), \quad (1)$$

where q are the different aggregating functions over the time h ($=1, \dots, H$); i ($=1, \dots, I$) indexes the monitoring sites (space) and j ($=1, \dots, J$) indexes the type of pollutant.

After this step, two other dimensions exist: pollutants and monitoring sites. It will be necessary, therefore, to establish the most suitable order of aggregation. In Bruno and Cocchi (2002) the interpretation of the two different orders of aggregation is discussed at length. Since the aim of this work is to study short-term health effects, the aggregating function, either between pollutants, or between monitoring sites, is the maximum. The final indices obtained by the two orders of aggregation, in the case where the aggregating functions are the maxima, have the same values.

For simplicity, we choose the first method of aggregation: monitoring sites- pollutants. The final index is:

$$I_{R,M(M)} = \max_j f_R(X_{M\{j\}}), \quad (2)$$

where: $X_{M\{j\}} = \max_i (X_{q\{ij\}})$ is the synthesis among monitoring sites, and f_R is the standardizing function necessary when the aggregation is computed over the pollutants.

Since our interest is in a continuous final index scale, we consider a simple transformation as:

$$f_R(Y) = \frac{b_{c+1} - b_c}{a_{(c+1)j} - a_{cj}} (Y - a_{cj}) + b_c, \text{ for } a_{cj} < Y < a_{(c+1)j}, c = 1, \dots, C, \forall j = 1, \dots, J, \quad (3)$$

where the considered thresholds are contained in Table 1.

Table 1: Air pollution thresholds (a_{cj}) for the different air quality classes and corresponding thresholds (b_c) of the standardised index.

	$a_{c1} - a_{(c+1)1}$	$a_{c2} - a_{(c+1)2}$	$a_{c3} - a_{(c+1)3}$	$a_{c4} - a_{(c+1)4}$	$a_{c5} - a_{(c+1)5}$	$b_c - b_{c+1}$
	SO ₂	NO ₂	CO	O ₃	PTS	Index
Good	0-50	0-50	0-2.5	0-60	0-40	0-50
Average	51-125	51-200	2.6-15	61-180	41-150	51-100
Poor ⁽²⁾	126-250	201-400	15.1-30	181-360	151-300	101-200
Extreme ⁽³⁾	>250	>400	>30	>360	>300	>200

Source: Emilia Romagna Regional Agency for Environmental Protection (A.R.P.A)

3. An empirical study in the urban area of Bologna

The data used in this analysis are based on daily measurements from 1996 to 1999 of non-accidental mortality (from the AUSL of Bologna), analysing total suspended particulate (TSP), ozone (O₃), sulphur dioxide (SO₂), nitrogen dioxide (NO₂), carbon monoxide (CO) and meteorological variables (from the Emilia Romagna Regional Agency for Environmental Protection). During the period under study (1996-1999) the monitoring network of air quality changed. Therefore, we consider here the only monitoring sites present for the entire period: Giardini Margherita, Fiera and Malpighi.

The considered models are Generalized Additive Models for Poisson data. Effects estimates were adjusted for age and the usual confounding factors were considered: long term trend, seasons, days of the week, holidays, influenza epidemics, meteorological conditions (temperature and humidity). Each pollutant was introduced in the models as linear effect. With respect to models that consider the index as an alternative of a single pollutant, we have considered three different situations: a linear effect, two linear terms constrained at 100 value and a dicomization of the index with respect to the 100 value. We have also considered models where each pollutant is present in conjunction with the index in the tree different from previously described forms. Model adequacy was checked by the analysis of the partial autocorrelation function and the analysis of residuals. We have considered, as usual, for all models lags 0-3 and some distributed lags (0-1, 1-2).

In the following we report, as an example, the results obtained for distributed lag (0-1) for CO and SO₂; while the index is expressed in the dichotomized version and as linear effect. Lag structure in air pollution models is complicated by the fact that each component of air pollution may have different lag structure, for example ozone and

⁽²⁾ The term “poor” corresponds to the attention threshold for each pollutant.

⁽³⁾ The term “extreme” corresponds to the alarm threshold for each pollutant.

particulate matter. Hence, it may not be possible to sort out all of the possible combinations of lags. Distributed lag models may be useful in this regard. When we consider models with the pollutant and the index together it may be preferable to reconstruct the index without the pollutant under study. With respect to the two considered models this problem is reduced by the fact that NO₂ and O₃ are the two pollutants that most contributed to determine high values of the index, particularly values in the "poor" class.

In table 2 we have reported the effects, estimated in the different models, expressed as percentage increase in risk (and 95% confidence interval) for a 10 µg/m³ (1 mg for CO) increase in pollutant concentration.

The obtained results show how it is possible through the air quality index to evaluate a global effect of air pollution on health. On the other hand, it is possible to control the effects of the pollutant for a different level of air pollution (due to the other pollutants) by means of the index. When we consider pollutant and index together we obtain different results. Carbon monoxide seems not to be influenced by the level of the other pollutant; while there is a change in the effect on mortality rates due to sulphur dioxide: probably, in this case, the other pollutants act as confounders. This is due to the fact that SO₂ is a pollutant that in Bologna reaches low levels (average 7.63) belonging to the "good" class of Table 1. For the same reasons, analogous considerations can be made when we consider the model with two linear terms, constrained at 100 value, for the index.

The use of the air quality index in the epidemiologic time series seems promising. In future works it will be interesting to consider, by the index, a different characterization of the exposure to the overall pollutants, by means of the duration of the exposition as suggested in Chiogna, Bellini (2002).

Table 2: Relative Risks (RR) [95% confidence interval]

	SO ₂	CO	Index qual	Index lin. effect
Single effect	5.25(1.50;9.14)	2.47(1.20;3.76)	5.30(1.23;9.54)	1.59 (0.52;2.66)
Pollutant and index	4.23(0.44;8.16)		4.32(0.19;8.63)	
	3.48(-0.39;7.49)			1.30(0.18;2.43)
		2.19(0.88;3.51)	3.00(-1.13;7.29)	
		2.03(0.67;3.41)		0.83(-0.31;1.97)

References

- Bruno F., Cocchi D. (2002). A unified strategy for building simple air quality indices, *Environmetrics*, in Press.
- Chiogna M., Bellini P. (2002). Alternative air pollution measures for detecting short-term health effects in epidemiological studies, *Environmetrics*, 13, 55-69.
- Cox L. H. (2000). Statistical issues in the study of air pollution involving airborne particulate matter, *Environmetrics*, 11, 611-626.
- Schwartz, J., Spix C., Touloumi, G., Bacharova L., Le Tertre A., Piekarski T., Ponce de Leon A., Ponka A., Rossi G., Saez M., Schouten J.P. (1996). Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions, *Journal of Epidemiology and Community Health*, 50 (Suppl. 1), S3-S11.