Time scale effects in acute association between air-pollution and mortality

Myrto Valari, Lucio Martinelli, Edouard Chatignoux, James Crooks and Val Garcia

Myrto Valari:
NRC Post-doc at the U.S Environmental Protection Agency, Atmospheric Modeling Division
U.S. Environmental Protection Agency
109 T.W. Alexander Drive
Research Triangle Park, NC 27711
Valari.Myrto@epa.gov

Lucio Martinelli:
Laboratoire de Physique de la Matière Condensée
Ecole Polytechnique, 91128 Palaiseau Cedex, France

Edouard Chatignoux:
Obseravatoire Régional de Santé Ile-de-France
21/23, rue Miollis 75732 Paris Cedex 15, France
e.chatignoux@ors-idf.org

James Crooks:
U.S Environmental Protection Agency, Atmospheric Modeling Division
U.S. Environmental Protection Agency
109 T.W. Alexander Drive
Research Triangle Park, NC 27711
Val Garcia:
U.S Environmental Protection Agency, Atmospheric Modeling Division
U.S. Environmental Protection Agency
109 T.W. Alexander Drive
Research Triangle Park, NC 27711
Garcia.Val@epa.gov
Abstract

We used wavelet analysis and generalized additive models (GAM) to study timescale effects in the acute association between mortality and air-pollution. Daily averages of measured NO$_2$ concentrations in the metropolitan Paris area are used as surrogates of human exposure from 2000 to 2004. The NO$_2$ time series was decomposed with wavelet analysis to six independent variables representing different durations of population exposure. We used these variables as predictors in a mortality regression model and compared the coefficients estimated for the different timescales. We found a strong dependency of the exposure-response function on the duration of the air-pollution event. In contrast to previous studies that showed a monotone increase in the relationship between exposure to air-pollution and mortality from shorter to longer timescales, our results show a non-linear response suggesting that the overall acute effect consists of two discrete patterns: a short-term response (2 to 15 days) where mortality relative risks decrease to near null values with the duration of the air-pollution event; an intermediate timescale pattern (16 to 55 days) where mortality relative risk climbs back up to positive levels. The revealed pattern shows that the overall acute effect of air-pollution on mortality does not reflect only a short-term mortality displacement in a population already at high death risk due to chronic conditions but also the transition into this pool from the healthy population.
Introduction

Epidemiological evidence relating air-pollution to mortality at metropolitan areas [Katsouyanni et al., 2001 and Samet et al., 2000] has been widely interpreted as the response in a population of individuals with fragile health, such as the elderly and persons with chronic cardiac or respiratory diseases [Environmental Protection Agency, 1996]. This means that part of the epidemiological association reflects the shortening of life expectancy by only a few days. This effect, termed as mortality displacement [Schimmel and Murawski, 1976] has been, in some cases, suggested as the only interpretation of the acute association between air-pollution and mortality and has led to reluctance to enhance emission controls [Lipfert and Wyzga, 1995]. Such policy implications, stress the need to quantify the shortening of life expectancy implied by acute epidemiological studies.

In response to this, the epidemiological research community developed several methods to estimate air pollution-mortality associations at various timescales. If the statistical association between exposure and mortality only reflected a few days shortening in the life expectancy at a population group already at high death risk, the days following the air pollution episode should be marked with mortality below baseline levels [Zanobetti et al., 2002]. However, previous investigations provided counter evidence of this short-term mortality displacement hypothesis by showing that associations between air-pollution and mortality increase with the duration of the exposure [Dominici et al., 2003; Kelsall et al., 1999; Zeger et al., 1999 and Schwartz, 2000].

Here we extended the methodology of Dominici et al. [2003] to study the mortality displacement effect on a five-year dataset (2000-2004) in the metropolitan Paris area. Similarly to Dominici et al., [2003], we used cutoff frequencies in the power spectrum to isolate the temporal variability of the exposure variable corresponding to a set of discrete timescales. Each wavelength specific component was back transformed to the time domain and the complete set of independent variables was used as co-variates in the regression to mortality. Here, we used an orthogonal wavelet decomposition of the exposure variable instead of Fourier analysis [Dominici et al., 2003]. Wavelets preserve local features
of the time-series [Farge, 1992] and therefore, the decomposed exposure variables contain an
additional layer of information compared to the Fourier decomposition on the moment in time when
specific frequency events occur. This enhancement in locality is bound to help capturing the transition
between a state where short-term mortality displacement is the major component of the
epidemiological association to a timescale where refill from the healthy population overcomes the
depletion rate of the susceptible pool.

Data
Mortality and Health Predictors
We used daily non-accidental mortality (ICD-10 A00-R99) counts for the metropolitan Paris area from
2000 to 2004. Data from August 1st 2003 to August 15th 2003 are excluded from the analysis because
of the exceptionally high temperatures and pollutant levels (heat wave of summer 2003 over Europe).
The approach we followed (see Methods below) relates daily mortality counts to a set of health
predictors. Air-pollution predictors were derived from ambient NO$_2$ measurements at several central
monitors in the area on hourly basis. To assure the quality of NO$_2$ data only monitors with less than
25% of missing hourly values were used. Monitors were selected to represent only background air-
pollution levels. This was achieved by applying two criteria: (i) coherence between data measured at
different monitors (overlap of the interquartiles and differences between mean values lower than
15µg/m$^3$) and (ii) sensitivity to the addition of each individual monitor (variability of the mean value
less than 15% and correlation coefficient higher than 0.8). Data were spatially aggregated across the
selected monitors and temporally averaged over twenty four hours to represent the daily exposure
experienced by the mean metropolitan population (see the study of Host et al., [2008] for more details
on data).

Methods
Epidemiological model
The daily data were analyzed with time-series methods, using generalized additive Poisson regression models allowing for overdispersion [Wood, 2006]. Possible confounders, including long-term trend, seasonality, days of the week, holidays, influenza epidemics, minimum temperature of the current day and maximum temperature of the previous day, were controlled using Air Pollution and Health: A European Approach 2 (APHEA-2) methodology [Touloumi et al., 2004]. Long-term trend and seasonality were modeled using a penalised regression spline of time. A large set of basis functions (equal to 50 per year) was used, and smoothing was used to remove autocorrelation of the model's residuals (by minimising the absolute value of the sum of the partial autocorrelation function of the model’s residuals) [Touloumi et al., 2006]. Dummy variables for days of the week and holidays were included as other independent variables. Temperature and influenza terms were modeled using parametric splines with 3 degrees of freedom. Both minimal and maximal temperatures were taken into account because other than minimum or maximum temperatures acting alone, their combination has been also associated with health effects [Samet, 1998]. All analyses were performed using the MGCV package in R software (R 2.11.1).

Spectral analysis and timescale decomposition

Using wavelet analysis, the time series of the exposure variable is decomposed into a set of six independent variables each representing a different timescale of exposure. All six wavelength components are used as linear mortality co-predictors in the same Poisson regression following the model developed by Dominici et al., [2003]:

\[
\log(\mu_t) = \sum_k (X_{kt}\beta_k) + S(time) + \text{confounders} \tag{1}
\]

where \( k \) indicates the discrete timescales of wavelet decomposition, \( t \) the day index, \( \mu_t \) the quasi-poisson mean of the output mortality distribution at day \( t \), \( X_k \) the exposure variable representing
temporal variations within the \( k^{\text{th}} \) wavelength range. \( S \) and \( \beta \) are determined by the regression and they represent a smooth function of calendar time and the linear coefficients reflecting correlation between mortality and air-pollution respectively.

The wavelet transform of the exposure variable for each pollutant returns a two dimensional set of coefficients representing the wavelet power spectrum at each point in time. The clear advantage of this method compared to the Fourier transform is that the power spectrum is computed separately at each point of the time-series. Its computation depends only on a finite window around that point and not on the whole time series, therefore accounting for transient features of periodicity instead of pre-assuming a global pattern \([Cazelles et al., 2007]\). The time series decomposition consists of two steps: first, the time series is analyzed following a multi-resolution scheme, providing a set of decomposition coefficients at different scales \([Farge, 1992]\). For this study, we performed a five level analysis using orthogonal Daubechies (order 5) mother functions. In the frequency domain, we discretized the \( \text{NO}_2 \) power spectra into six timescales: \(<4, 4 \text{ to } 8, 9 \text{ to } 15, 15 \text{ to } 28, 29 \text{ to } 55, \text{ and } > 55 \text{ days}\). Then the coefficients of each discrete time scale are back transformed to the time domain.

**Results and discussion**

Wavelet power spectra for each timescale are shown in Figure 1. Note that for visualization reasons the power spectrum shown in Figure 1 has been computed using continuous wavelets (Morlet mother functions) instead of the orthogonal Daubechies to obtain a smoother pattern. To respect the Nyquist theorem, and given that the resolution of our data is one day, the highest frequency considered in the analysis is every two days. For the longest wavelength component (i.e. events of duration longer than 55 days) the dominant feature is the annual cycle at 365 days that is distributed across the study period in a near-uniform pattern. At intermediate time scales sporadic weather and air-pollution events are captured, such as particularly cold months (January 2002 and 2003) or the heat-wave of August 2003. At the shortest timescales weekly or few days long anomalies are represented. We note here that
these transient features would not have been detected with a Fourier analysis that has maximum resolution in the frequency domain but null resolution in the time domain.

We applied the back transform separately to each timescale to obtain the six orthogonal exposure variables at the time domain (Figure 2). Used as independent co-predictors in the epidemiological model (Eq 1), these variables are related to mortality through six regression coefficients (Figure 3). Risk-estimates are expressed as percentages of mortality increase due to a 10-unit increase in exposure (%ERR). The analysis provides evidence of a strong time-scale dependency of a highly non-linear pattern. At timescales shorter than 15 days, mortality relative risk decreases with the duration of the air-pollution events, with an ERR close to 0 for the 9 to 15 days time scale. This is consistent with the short-term mortality displacement hypothesis [Schwartz, 2000]. At the 9 to 15 days time scale, the increase in mortality due to air-pollution is counterbalanced by the depletion of the sub-population at high death risk due to chronic conditions, age etc. However at longer time-scales (from 15 to 55 days) risk estimates climb back up to similar and even higher levels than for the short-term. This provides evidence of an impact not only to a susceptible pool of the population but also to the generally healthy individuals that exposed to high levels of air-pollution for sufficient amount of time (more than 15 days) may develop chronic conditions and enter the susceptible group. The interpretation of ERR for timescales longer than 55 days is ambiguous, because of confounding effects of seasonality and other long-term trends.

Our analysis provides evidence of larger effects at longer time-scales (i.e. one to two months), which is in agreement with previous investigators that studied the epidemiological association between air-pollution and mortality as a function of the time-scale [Schwartz, 2000 and Dominici et al., 2003]. This is also consistent with studies focusing on associations between mortality and air-pollution at even longer-term exposures (i.e. several years) [Pope, 2007; Dockery, 2009 and Jerett et al., 2009]. On the other hand, the pattern of the association across time-scales revealed from our study implies that short term exposure to air-pollution is not only responsible for a few days shortening of life expectancy but
also for the transition from a healthy population in the susceptible group. This is in agreement with many studies highlighting adverse effects of exposure to air-pollution on less severe outcomes than mortality (e.g. hospitalization [Host et al., 2008] or medical visits [Larrieu et al., 2009]), discarding the hypothesis that air pollution effects are limited to the pool of very frail people. However, our results provide evidence of some degree of short-term mortality displacement during the first two weeks following air-pollution events. Similar patterns were observed by [Schwartz, 2000] for acute pneumonia, but not for all causes mortality.

**Conclusion**

With the present study we find evidence of strong and highly non-linear time-dependencies in the association between exposure to air-pollution and mortality. A low degree of short-term mortality displacement is found for the first couple of weeks after air-pollution events. However, our analysis strongly suggests that the acute impact of air-pollution on mortality does not reflect only the precipitation of deaths that would occur shortly afterwards regardless pollution; larger effects are estimated for longer-term exposures, which is consistent with the hypothesis that air-pollution is responsible for the development of chronic conditions at healthy individuals.

**References**


**Figure captions**

**Figure 1:** Wavelet power spectrum (square modulus of wavelet coefficients) of NO$_2$ daily averaged concentration time-series measured and spatially aggregated across several central monitors in the city of Paris from January 1, 2000 to December 31, 2004. Continuous Morlet wavelets are used as mother
functions for the decomposition. The spectrum is divided in six sections to isolate variability in the set
of the six discrete timescales (top to bottom) 55 days, 29 to 55, 16 to 28, 9 to 15, 4 to 8, <4. The y-axis
is the Fourier period (scale) in days and the x-axis is the date in the time domain. The colormap

**Figure 2:** Decomposition into a six-component series of data on NO$_2$ (µg/m$^3$) for Paris, from January 1,
2000 to December 21, 2004. On each plot the overall (i.e. before decomposition) time series is plotted
for reference (black line) and on top of it each wavelength component (colored line). Time series 1 to 6
(top to bottom) are the timescale decompositions from the longer to the shortest time-scales (> 55 days,
29 to 55, 16 to 28, 9 to 15, 4 to 8, <4)

**Figure 3:** Estimated % of increase in mortality due to increase of the NO$_2$ concentration by 10 µg/m$^3$.
The point at the left extreme of the plot represents the overall effect (i.e. without spectral analysis) and
is added as a reference. Points 1 to 6 (left to right) are the excess relative risks estimated for each
timescale component of the orthogonal wavelet decomposition from the shortest to the longest
wavelengths.
Figure 1: Wavelet power spectrum (square modulus of wavelet coefficients) of NO$_2$ daily averaged concentration time-series measured and spatially aggregated across several central monitors in the city of Paris from January 1, 2000 to December 31, 2004. Continuous Morlet wavelets are used as mother functions for the decomposition. The spectrum is divided in six sections to isolate variability in the set of the six discrete timescales (top to bottom) 55 days, 29 to 55, 16 to 28, 9 to 15, 4 to 8, <4. The y-axis is the Fourier period (scale) in days and the x-axis is the date in the time domain. The colormap corresponds to the power of the wavelet spectrum. The absolute value of power increases for long periods due to the corresponding wavelet broadening.
Figure 2: Decomposition into a six-component series of data on NO$_2$ (µg/m$^3$) for Paris, from January 1, 2000 to December 21, 2004. On each plot the overall (i.e. before decomposition) time series is plotted for reference (black line) and on top of it each wavelength component (colored line). Time series 1 to 6 (top to bottom) are the timescale decompositions from the longer to the shortest time-scales (> 55 days, 29 to 55, 16 to 28, 9 to 15, 4 to 8, <4)
Figure 3: Estimated % of increase in mortality due to increase of the NO$_2$ concentration by 10 µg/m$^3$. The point at the left extreme of the plot represents the overall effect (i.e. without spectral analysis) and is added as a reference. Points 1 to 6 (left to right) are the excess relative risks estimated for each timescale component of the orthogonal wavelet decomposition from the shortest to the longest wavelengths.