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

Crooks.James@epa.gov

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 indicators of human exposure to urban air pollution 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 suggests that the overall acute effect of air-pollution on mortality reflects not 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 could help capture 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. The area of the study contains Paris and the three surrounding departments Hauts-de-Seine, Seine-Saint-Denis and Val-de-Marne. The area encompasses 124 communes (municipalities) over a surface of 762 Km^2 with a population of more than 6 164 418 inhabitants at a mean population density of 8090 inhabitants/Km^2 (1999 census). In early August 2003 France sustained an unprecedented heat wave marked with exceptionally high temperatures and pollutant levels. Mortality in the metropolitan Paris area during this period exceeded the mean value by 475 deaths. To avoid extrapolation of any global pattern that may apply to the rest of the dataset to such an extreme event data from August 1st 2003 to August 20th 2003 are excluded from the analysis. 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 15$\mu g/m^3$) 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). The number of NO_2 monitors finally used in the analysis during the study period varied from 12 to 17. 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).

In Paris, like other urban areas around the world, motor vehicle exhaust is the major source of nitrogen dioxide and primary particulate matter (PM) emission (50 and 31% of the total emitted mass respectively). Under such conditions, the correlation between NO_2 and particulate matter, especially ultra-fine particles emitted from diesel engine combustion), may be so strong (here r^2=0.7) that NO_2 acts as indicator of human exposure to PM [*Brook et al.*, 2007]. Because NO_2 data are available for the entire study period and the number of NO_2 monitors in the area of our study is higher than for PM (1 to 4 available monitors) we used NO_2 as indicator of population exposure to the overall mix of traffic-related air-pollution.

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]:

$$log(\mu_t) = \sum_k (X_{kt}\beta_k) + S(time) + confounders \qquad (1)$$ Eq 1

where k indicates the discrete timescales of wavelet decomposition, t the day index, μ_t the quasi-poisson mean of the output mortality distribution at day t, X_k the exposure variable representing temporal variations within the k^{th} wavelength range. S and β 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.

With wavelet analysis the exposure time series is decomposed into the time-frequency space simultaneously. Basis functions used for the transform (mother wavelets) differ from the trigonometric functions used in Fourier analysis (sines and cosines) in that they are finite in both time and frequency domains. Wavelet transform consists of fitting scaled and translated versions of the mother wavelet (wavelets) to generate a set of coefficients. The advantage compared to Fourier analysis is that the generated coefficients contain information on both the amplitude of any "periodic" signals within the time series and on how this amplitude varies with time. We used Daubechies 5 mother wavelets for the

decomposition of the NO_2 time series into six scales representing different exposure durations. The scales of a wavelet decomposition represent wavelength (or frequency) ranges whose bounds are set through cutoff periods (or frequencies) on a dyadic grid to satisfy orthogonality [*Farge*, 1992]. The smallest scale must be greater than $2*\Delta x$ (Δx=data resolution) to satisfy the Nyquist theorem. Here we set the cutoff periods at 3, 6, 12, 24, 48 and 96 days. The relationship between the equivalent Fourier period (pseudo-period) and wavelet scale can be derived by convolution of the mother wavelet with a cosine wave of known wavelength and computing the scale at which the wavelet power spectrum reaches its maximum [*Meyers*, 1993]. For the wavelet decomposition performed here the six timescales were defined as: <4, 4 to 8, 9 to 15, 16 to 28, 29 to 55, and > 55 days. The back transform to the time domain provides the orthogonal set of exposure variables representing variability with the corresponding timescales.

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 5 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 three 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 Fourier analysis which pre-assumes a global pattern of periodicity across time series.

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 to15 days time scale. This is consistent with the short-term mortality displacement hypothesis [*Schwartz*, 2000]. At the 9 to15 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. The decrease in ERR at this timescale suggests a sort of competition for explanatory power with the spline of time.

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]), challenging 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

Brook J.R., R.T. Burnet, T.F. Dann, S. Cakmak, M.S. Goldberg, X.Fan and A.J. Wheeler (2007), Further interpretation of the acute effect of nitrogen dioxide observed in Canadian time-series studies, *J Expos Sci Environ Epidemiol*, *17*(S2), S36-S44.

Dockery, D.W. (2009), Health effects of particulate air pollution, *Ann Epidemiol*, *19*, 257-263.

Dominici F, A. McDermott, S.L. Zeger and J.M. Samet (2003), Airborne particulate matter and mortality: timescale effects in four US cities, *Am J Epidemiol, 157,*1055-1065.

Environmental Protection Agency (1996), Office of Air Quality Planning and Standards. Review of the National Ambient Air Quality Standards for Particulate Matter: policy assessment of scientific and technical information. OAQPS Staff Paper. Research Triangle Park, NC: Environmental Protection

Agency. (Publication no. EPA-452\R-96-013).

Farge M. Wavelet Transforms and their Applications to Turbulence (1992), *Annual Review of Fluid Mechanics, 24*, 395-458.

Host S, S. Larrieu, L. Pascal, M. Blanchard, C. Declercq, P. Fabre, J-F Jusot, B. Chardon, A. Le Tertre , V. Wagner, H. Prouvost and A. Lefranc (2008), Short-term associations between fine and coarse particles and hospital admissions for cardiorespiratory diseases in six French cities, *Occup Environ Med*, *65*(8), 544-551, doi:10.1136/oem.2007.036194.

Jerrett M, R.T. Burnett, CA3 Pope, K. Ito, G. Thurston, D. Krewski, Y. Shi, E. Calle and M. Thun (2009), Long-term ozone exposure and mortality. *N Engl J Med*, *360*, 1085-1095.

Katsouyanni K, G. Touloumi, E. Samoli, A. Gryparis, A. Le Tertre, Y. Monopolis, G. Rossi, D. Zmirou, F. Ballester, A. Boumghar, H.R. Anderson, B. Wojtyniak, A. Paldy, R. Braunstein, J. Pekkanen, C. Schindler and J. Schwartz (2001), 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*, *12*(5), 521-531.

Kelsall J, S. Zeger and J. Samet (1999), Frequency domain log-linear models: air pollution and mortality. *Appl Stat*, *48,* 331-344.

Larrieu S, A. Lefranc, G. Gault, E. Chatignoux, F. Couvy, B. Jouves and L. Filleul (2009), Are the short-term effects of air pollution restricted to cardiorespiratory diseases?, *Am J Epidemiol*, *169,* 1201-1208.

Lipfert FW and R.E. Wyzga (1995), Air pollution and mortality: issues and uncertainties, *J Air Waste Manage Assoc, 45,* 949-966.

Meyers S. D., B. G. Kelly, and J. J. O'Brien (1993), An introduction to wavelet analysis in oceanography and meteorology: With application to the dispersion of Yanai waves., *Mon. Wea. Rev., 121***,** 2858-2866.

Pope, CA3 (2007), Mortality effects of longer term exposures to fine particulate air pollution: review of recent epidemiological evidence, *Inhal Toxicol*, *19*(1), 33-38.

Samet J, S. Zeger, J. Kelsall, J. Xu and L. Kalkstein (1998), Does weather confound or modify the association of particulate air pollution with mortality? An analysis of the Philadelphia data, 1973-1980, *Environ Res*, *77,* 9-19.

Samet, J. M., F. Dominici, F. C. Curriero, I. Coursac, and S. L. Zeger (2000), Fine particulate air pollution and mortality in 20 U.S. cities, 1987-1994, *N. Engl. J. Med*, *343*(24), 1742-1749, doi:10.1056/NEJM200012143432401.

Schimmel H and T.J. Murawski (1976). Proceedings: the relation of air pollution to mortality, *J Occup Med*, *18*, 316-333.

Schwartz J. (2000), Harvesting and long term exposure effects in the relationship between air pollution and mortality, *Am J Epidemiol*, *151*,440-448.

Touloumi, G., R. Atkinson, A. Le Tertre, E. Samoli, J. Schwartz, C. Schindler, J.M. Vonk, G. Rossi, M. Saez, D. Rabszenko, and K. Katsouyanni (2004), Analysis of health outcome time series data in epidemiological studies, *Environmetrics, 15,* 101-117.

Touloumi G, E. Samoli, M. Pipikou, A. Le Tertre, R. Atkinson and K. Katsouyanni (2006), Seasonal confounding in air pollution and health time-series studies: effect on air pollution effect estimates, *Stat Med, 25,* 4164-4178.

Wood S. (2006), Generalized additive models: An Introduction with R. Boca Raton, *Chapman & Hall/CRC*.

Zanobetti A, J. Schwartz, E. Samoli, A. Gryparis, G. Touloumi, R. Atkinson, A. Le Tertre, J. Bobros, M. Celko, A. Goren, B. Forsberg, P. Michelozzi, D. Rabczenko, E. Aranguez Ruiz and K. Katsouyanni (2002), The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement, *Epidemiology, 13,* 87-93.

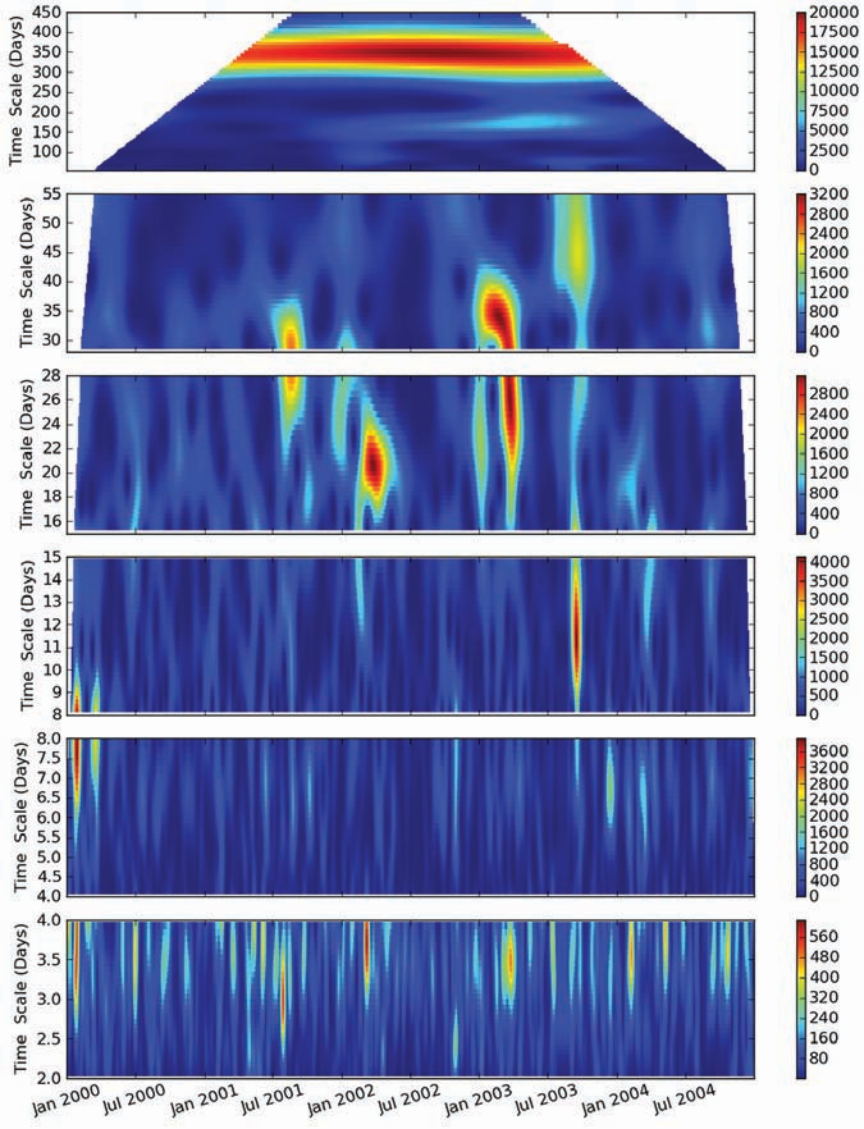
Zeger SL, F. Dominici and J. Samet (1999), Harvesting-resistant estimates of air pollution effects on mortality, *Epidemiology*, *10,* 171-175.

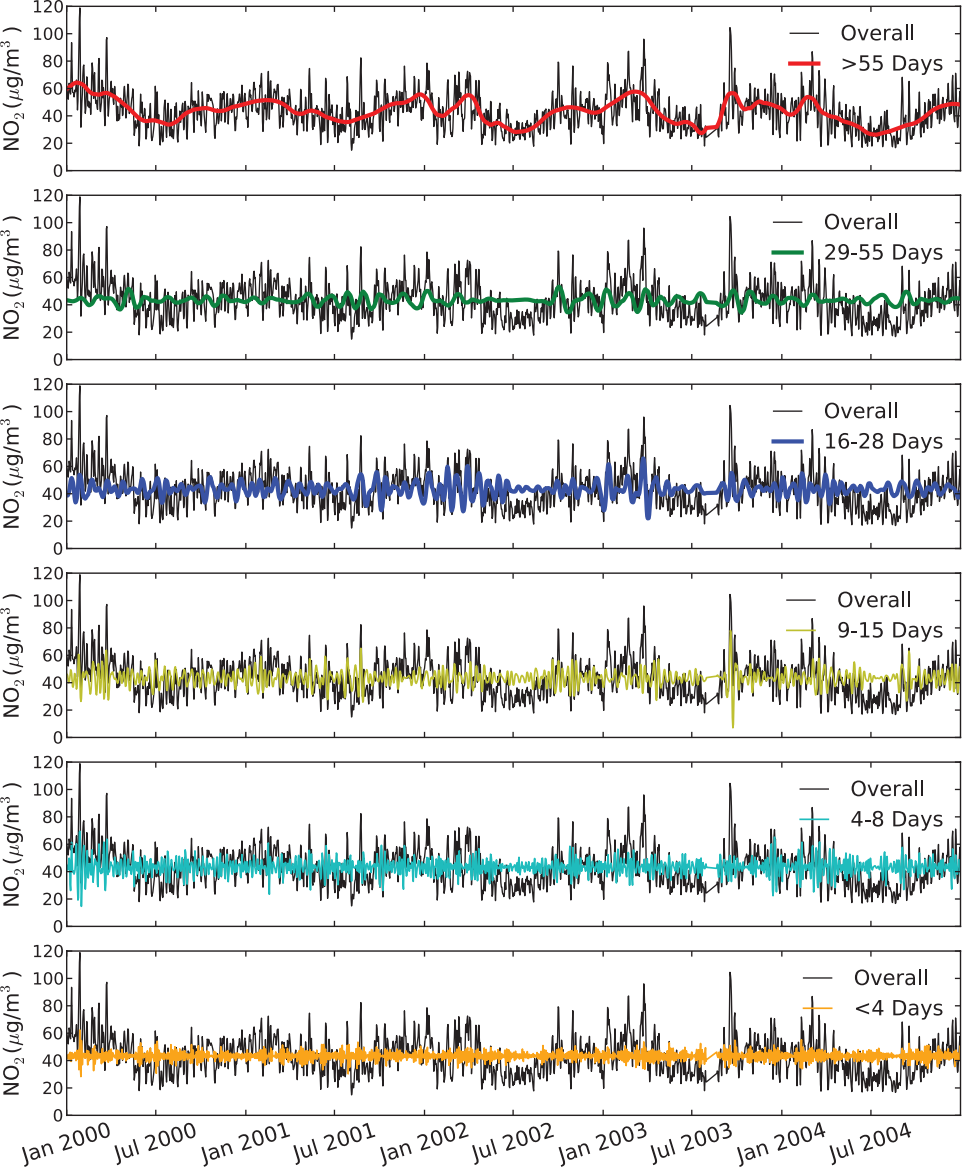
Figure captions

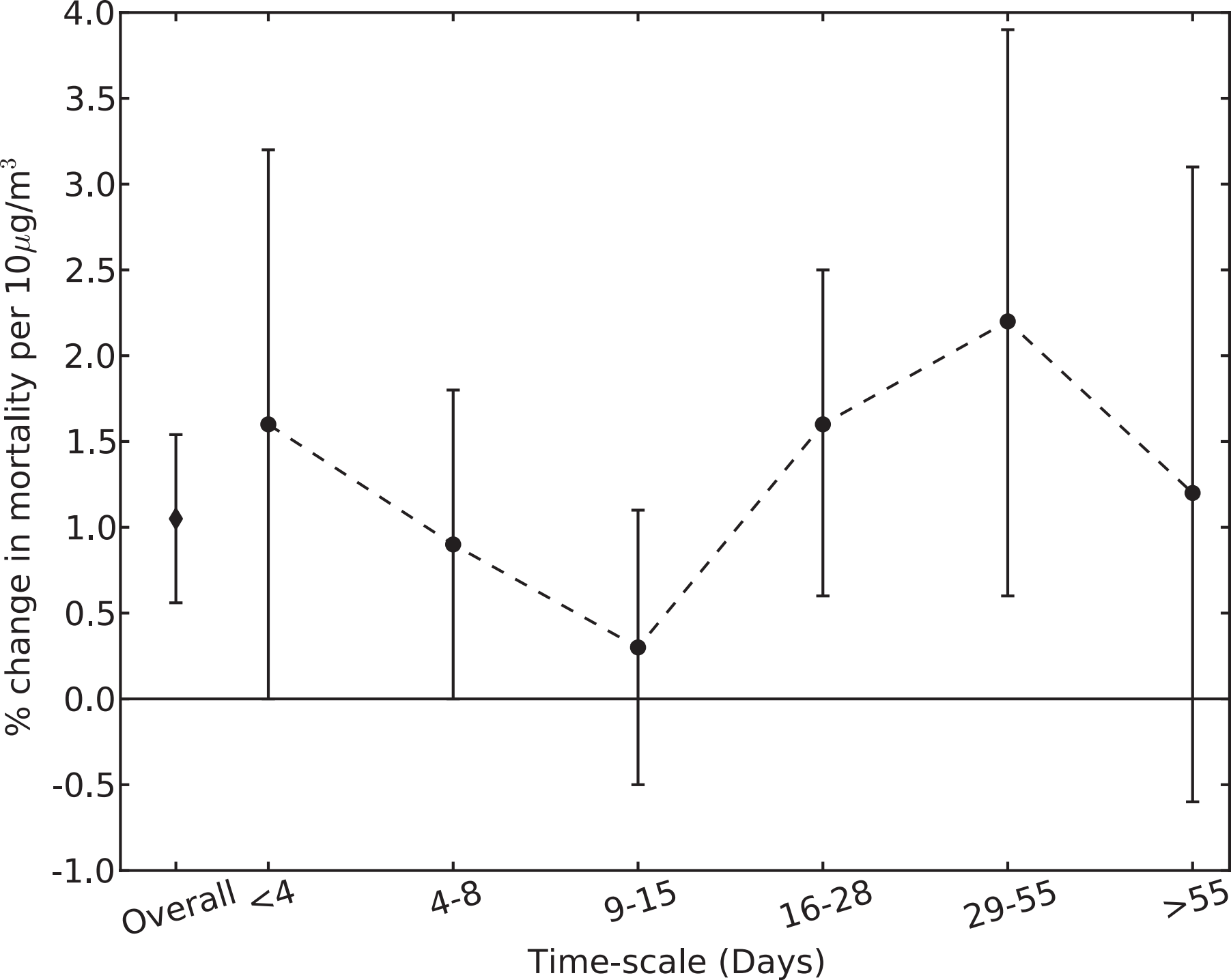
Figure 1: Wavelet power spectrum (square modulus of wavelet coefficients) of NO_2 daily averaged concentration time-series measured and spatially aggregated across several central monitors in the city of Paris from January 1, 2000 to December 31, 2004. Continuous Morlet wavelets are used as mother functions 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 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 ($\mu g/m^3$) 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). Data from August 1st 2003 to August 20th 2003 (heat-wave of summer 2003) are excluded from the analysis.

Figure 3: Estimated % of increase in mortality due to increase of the NO_2 concentration by 10 $\mu g/m^3$. 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.


Time Scale (Days)
450
400
350
300
250
200
150
100
20000
17500
15000
12500
10000
7500
5000
2500
0
Time Scale (Days)
55
50
45
40
35
30
3200
2800
2400
2000
1600
1200
800
400
0
Time Scale (Days)
28
26
24
22
20
18
16
2800
2400
2000
1600
1200
800
400
0
Time Scale (Days)
15
14
13
12
11
10
9
8
4000
3500
3000
2500
2000
1500
1000
500
0
Time Scale (Days)
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3600
3200
2800
2400
2000
1600
1200
800
400
0
Time Scale (Days)
4.0
3.5
3.0
2.5
2.0
560
480
400
320
240
160
80
Jan 2000
Jul 2000
Jan 2001
Jul 2001
Jan 2002
Jul 2002
Jan 2003
Jul 2003
Jan 2004
Jul 2004


NO2 (μg/m3)
Overall
>55 Days
29-55 Days
16-28 Days
9-15 Days
4-8 Days
<4 Days
0
20
40
60
80
100
120
Jan 2000
Jul 2000
Jan 2001
Jul 2001
Jan 2002
Jul 2002
Jan 2003
Jul 2003
Jan 2004
Jul 2004


4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
% change in mortality per 10μg/m^3
Overall
<4
4-8
9-15
16-28
29-55
>55
Time-scale (Days)