

Timescale effect estimation in time-series studies of air pollution and health: A Singular Spectrum Analysis approach

Massimo Bilancia, *Department of Statistical Sciences*
Girolamo Stea, *Department of Statistical Sciences*

Abstract

A wealth of epidemiological data suggests an association between mortality/morbidity from pulmonary and cardiovascular adverse events and air pollution, but uncertainty remains as to the extent implied by those associations although the abundance of the data. In this paper we describe an SSA (Singular Spectrum Analysis) based approach in order to decompose the time-series of particulate matter concentration into a set of exposure variables, each one representing a different timescale. We implement our methodology to investigate both acute and long-term effects of PM_{10} exposure on morbidity from respiratory causes within the urban area of Bari, Italy.

AMS 2000 subject classifications: Primary 62P12; secondary 62J99.

Keywords: Airborne particulate matter, PM_{10} , Singular Spectrum Analysis - SSA, Generalized additive models - GAM.



Full Text: [PDF](#)

Bilancia, Massimo, Stea, Girolamo, Timescale effect estimation in time-series studies of air pollution and health: A Singular Spectrum Analysis approach, *Electronic Journal of Statistics*, 2, (2008), 432-453 (electronic). DOI: 10.1214/07-EJS123.

References

- [1] Anderson HR, Atkinson RW, Peacock JL, Marston L, Konstantinou K (2004) Meta-analysis of time-series studies and panel studies of particulate matter (PM) and ozone (O_3): report of a WHO task group, Copenhagen, WHO Regional Office for Europe, document EUR/04/5042688.
- [2] The APHEA project (1996) Short term effects of air pollution on health: a European approach using epidemiological time series data, *Journal of Epidemiology and Community Health*, 50, Supplement(1).
- [3] Bell ML, Samet JM, Dominici F. (2004) Time-series studies of particulate matter, *Annual Review of Public Health*, 25, 247–80.
- [4] Biggeri A, Bellini P, Terracini B (2004) Metaanalisi italiana degli studi sugli effetti a breve termine dell'inquinamento atmosferico 1996–2002 [Meta-analysis of the Italian studies on short-term effects of air pollution 1996–2002], *Epidemiologia e Prevenzione*, 28(S2), 4–100 (in Italian).
- [5] Broomhead DS, King GP (1986) Extracting qualitative dynamics from experimental data, *Physica D* 20, 217–236. [MR0859354](#)
- [6] Buchstaber VM (1994) Time series analysis and grassmanians. In: Gindikin S (Ed.) *Applied Problems of Radon Transform*. Volume 162 of AMS Transaction - Series 2,

- [7] Chatfield C. (2004) *The Analysis of Time Series: an Introduction*, CRC Press, Boca Raton, FL. [MR2001158](#)
- [8] Diggle PJ (1990) *Time series: a Biostatistical Introduction*, Oxford Science Publications, Oxford. [MR1055357](#)
- [9] Dominici F, McDermott A, Zeger SL, Samet JM (2002) On the use of Generalized Additive Models in Time-Series Studies of Air Pollution and Health, *American Journal of Epidemiology*, 156(3), 1–11.
- [10] Dominici F, McDermott A, Zeger SL, Samet JM (2003) Airborne particulate matter and mortality: timescale effects in four US cities, *American Journal of Epidemiology*, 12, 1055–1065.
- [11] Dominici F, Sheppard L, Clyde M. (2003) Health effects of air pollution: a statistical review, *International Statistical Review*, 71, 243–276.
- [12] Gantmacher FR (1998) *The Theory of Matrices*, Chelsea, Providence (RI). [MR1657129](#)
- [13] Ghil M, Allen MR, Dettinger MD, Ide K, Kondrashov D, Mann ME, Robertson AW, Saunders A, Tyan Y, Varadi F, Yiou P (2002) Advanced spectral methods for climatic time series, *Review of Geophysics* 40(1), 1–41.
- [14] Golyandina N, Nekrutin A, Zhigljavsky A (2001) *Analysis of Time Series Structure: SSA and related techniques*, Chapman & Hall/CRC, New York. [MR1823012](#)
- [15] Golyandina N, Osipov N (2007) The Caterpillar-SSA method for analysis of time series with missing values, *Journal of Statistical Planning and Inference*, 137:2642–2653. [MR2326115](#)
- [16] Hastie TJ, Tibshirani RJ (1990) *Generalized Additive Models*, Chapman & Hall, London. [MR1082147](#)
- [17] Hastie TJ, Tibshirani R, Friedman J. (2001) *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, Springer, New York. [MR1851606](#)
- [18] Heckman NE, Zamar RH (2000) Comparing the shapes of regression functions, *Biometrika*, 87, 135–144. [MR1766834](#)
- [19] Martin AE, (1964) Mortality and morbidity statistics and air pollution, *Proceeding of the Royal Medical Society*, 57, 969–75.
- [20] McKeown M, Hansen L, Sejnowski T (2003) Independent component analysis of functional MRI: what is signal and what is noise? *Current Opinion in Neurobiology*, 13, 620–29.
- [21] Morgan G, Lincoln D, Lumley T, Sheppard V, Beard JF, Jalaludin B, Corbett S (2003) Time scale effects of particulate air pollution and mortality in Sydney, Australia, 1994 to 2000, *Epidemiology*, 14(5)- S111, ISEE-569.
- [22] Peng RD, Dominici F, Louis TA (2006) Model choice in time series of air pollution and mortality, *Journal of the Royal Statistical Society A*, 169(2), 179–203. [MR2225539](#)
- [23] Peng RD, Dominici F, Zeger SL (2006) Reproducible epidemiologic research, *American Journal of Epidemiology*, 163, 783–789.
- [24] R Development Core Team (2007) *R: a Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>.
- [25] Rosseeuw PJ, Van Driessen K (1999) A fast algorithm for the minimum covariance determinant estimator, *Technometrics*, 41, 212–223.

- [26] Samet J, Zeger S, Kelsall J, Xu J, Kalkstein L (1998) Does weather confound or modify the association of particulate air pollution with mortality? *Environmental Research Section A*, 77, 9–19.
- [27] Saucier A (2005) Construction of data-adaptive orthogonal wavelet bases with an extension of principal component analysis, *Applied Computational and Harmonic Analysis*, 18, 300–328. [MR2131290](#)
- [28] Schwartz J (2000) Harvesting and long-term exposure effects in the relation between air pollution and mortality, *American Journal of Epidemiology*, 151, 440–448.
- [29] Vautard R, Ghil M (1989) Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series, *Physica D* 35, 395–424. [MR1004204](#)
- [30] Wood SN (2000) Modelling and smoothing parameter estimation with multiple quadratic penalties, *Journal of the Royal Statistical Society B*, 62(2), 413–428. [MR1749600](#)
- [31] Wood SN (2004) Stable and efficient multiple smoothing parameter estimation for generalized additive models, *Journal of the American Statistical Association*, 99, 673–686. [MR2090902](#)
- [32] Wood SN (2004) *Generalized Additive Models: an Introduction with R*, Chapman & Hall/CRC, Boca Raton. [MR2206355](#)
- [33] Wood SN, Augustine NH (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modelling, *Ecological Modelling*, 157, 157–177.
- [34] Vautard R, Yiou P, Ghil M (1992) Singular-spectrum analysis: a toolkit for short, noisy chaotic signals, *Physica D* 58, 95–126.
- [35] Yiou P, Sornette D, Ghil M (2000) Data-adaptive wavelets and multi-scale singular spectrum analysis, *Physica D* 142, 254–290. [MR1771905](#)
- [36] Zanobetti A, Wand MP, Schwartz J, Ryan LM (2000) Generalized additive distributed lag models: quantifying mortality displacement, *Biostatistics*, 1, 279–291.