Article

Complexity analysis in particulate matter measurements

Luciano Telesca[1], Michele Lovallo[2]
[1]Consiglio Nazionale delle Ricerche, Istituto di Metodologie per l'Analisi Ambientale, C.da S.Loja, 85050 Tito (PZ), Italy
[2]ARPAB, 85100 Potenza, Italy
E-mail: luciano.telesca@imaa.cnr.it

Received 14 June 2011; *Accepted 20 July 2011*; *Published online 1 September 2011*
IAEES

Abstract
We investigated the complex temporal fluctuations of particulate matter data recorded in London area by using the Fisher-Shannon (FS) information plane. In the FS plane the PM10 and PM2.5 data are aggregated in two different clusters, characterized by different degrees of order and organization. This results could be related to different sources of the particulate matter.

Keywords particulate matter; Fisher information measure; Shannon entropy.

1 Introduction

Particulate matter (PM) has been recognized to cause harmful effects on human health and to be correlated with pollution problems and global climate change. Evidences about the exposition risk to this pollutant by population were given; in particular, PM particles may be inhaled into the lungs where the chemicals ranging from metal compounds to acid droplets can severely affect health, with respiratory problems especially in children and elderly people living in large metropolitan areas. Recent epidemiological studies have shown that concentration of fine particles are strongly related with human mortality, bronchitis and reduced lung function (WHO, 2003; Martuzevicius et al., 2004; Pope et al., 2002; Shendriker and Steinmetz, 2003). Furthermore, the deposition of PM particles on soils and in coastal waters could be potentially risky the ecosystem health (Gao et al., 2002). Recent studies have also shown that aerosol particles can influence the Earth's energy budget directly and indirectly, with clear consequences on climate. The direct effect consists in scattering and absorbing both solar and infrared radiation in the atmosphere. The indirect effect is linked to the action of aerosol particles as cloud condensation nuclei affecting cloud physical and radiative properties (Yu et al., 2002). Different sources (natural and anthropogenic), the chemical-physical and morphological composition, the spatial and temporal distribution increase the level of complexity of PM (Chandra Mouli et al., 2006; IPCC, 2001), whose variability is affected by several different processes. In this context it is important to characterize the dynamics of PM time series using methodologies able to gain insight into the complex nature of such measurements.

2 The Fisher-Shannon Information Plane

The Fisher Information Measure (FIM) is a powerful tool to investigate complex and nonstationary signals, and quantifies the degree of organization of a system. The FIM was introduced by Fisher (1925) in the context of statistical estimation. In a seminal paper Frieden (1990) has shown that FIM is a versatile tool to describe the evolution laws of physical systems. FIM allows to accurately describe the behavior of dynamic systems, and to characterize the complex signals generated by these systems (Vignat and Bercher, 2003). This approach has been used by Martin et al. to characterize the dynamics of EEG signals (Martin et al., 1999). Martin et al. (2001) have shown the informative content of FIM in detecting significant changes in the behavior of nonlinear dynamical systems disclosing, thus, FIM as an important quantity involved in many aspects of the theoretical and observational description of natural phenomena. The FIM was used in studying several geophysical and environmental phenomena, revealing its ability in describing the complexity of a system (Balasco et al., 2008) and suggesting its use as to reveal reliable precursors of critical events (Telesca et al., 2009a; Telesca et al., 2010).

The Shannon entropy is the well-known magnitude to quantify the degree of disorder in dynamical systems. The Shannon entropy can be used to define the degree of uncertainty involved in predicting the output of a probabilistic event (Shannon, 1948; Hilborn, 1994). For discrete distributions, this means that if one predicts the outcome exactly before it happens, the probability will be a maximum value and, as a result, the Shannon entropy will be a minimum. If one is absolutely able to predict the outcome of an event, the Shannon entropy will be zero. For distributions (probability densities) on a continuous variable, ranging e.g. over the real line, the Shannon entropy can reach any arbitrary value, positive or negative. Therefore, the use of the Shannon entropy power (defined through the exponential of the Shannon entropy) avoids the difficulty of dealing with negative information measures. Shannon entropy provides a scientific method to understand the essential state of things (de Araujo et al., 2003; Fuhrman et al., 2000).

The Fisher-Shannon information plane is a plane whose coordinate axes are the Fisher Information Measure and the Shannon entropy power and is used to investigate the dynamics of a time series. The use of the Fisher-Shannon Information plane has revealed its potential in discriminating dynamical processes, contributing in deepening the understanding of underlying mechanisms (Telesca et al., 2011).

Let us introduce the relevant Fisher- and Shannon-associated quantities (Martin et al., 2001). Let *f(x)* be a probability density of the variable *x*. Fisher's quantity of information associated to *f* is defined as the (possibly infinite) non-negative number *I*

$$I = \int_{-\infty}^{+\infty} \left(\frac{d}{dx} f(x) \right)^2 \frac{dx}{f(x)} \qquad (1)$$

The Shannon entropy is given by the following formula (Vignat and Bercher, 2003):

$$H_X = -\int_{-\infty}^{+\infty} f(x) \log f(x) dx \qquad (2)$$

For convenience the alternative notion of Shannon entropy power (Dembo et al., 1991):

$$N_X = \frac{1}{2\pi e} e^{2H_X} \qquad (3)$$

will be used rather than the entropy H_X. The use of the power entropy N_X instead of the Shannon one H_X arises from the so-called 'isoperimetric inequality' (Dembo et al., 1991; Romera and Dehesa, 2004; Angulo et al., 2008; Esquivel et al., 2010), a lower bound to the Fisher-Shannon product which reads as $IN_X \geq d$, where d is the dimension of the space. The ‘isoperimetric inequality’ suggests that the FIM and the Shannon entropy are intrinsically linked, so that the dynamical characterization of signals should be improved when analyzing them in the so called Fisher-Shannon (FS) information plane (Vignat and Berchet, 2003), in which the *y*- and *x*-axis are the FIM and the Shannon entropy (as outlined above, instead of the Shannon entropy we will use the entropy power N_X). Vignat and Berchet (2003) showed that the simultaneous examination of both Shannon entropy and FIM through the FS plane could improve the characterization of the non-stationary behavior of complex signals, like the Tsallis and the power exponential signals. The product IN_X can be considered as a statistical measure of complexity (Angulo et al., 2008). The line $IN_X=1$ separates the FS plane in two parts: one allowed ($IN_X>1$) and one not allowed ($IN_X<1$), and the distance of a signal point from the ‘isocomplexity line’ $IN_X=1$ can measure the degree of complexity of the signal.

Eqs. 1 and 2 involve the calculation of the probability density function (pdf) f(x). An estimation of the pdf f(x) may be obtained by means of the kernel density estimator technique (Devroye, 1987; Janicki and Weron, 1994). The kernel density estimator provides an approximate value of the density in the form

$$\hat{f}_M(x) = \frac{1}{Mb} \sum_{i=1}^{M} K\left(\frac{x - x_i}{b}\right) \qquad (4)$$

where M is the number of data and $K(u)$ is the kernel function, which is a continuous non-negative and symmetric function satisfying

$$K(u) \geq 0 \text{ and } \int_{-\infty}^{+\infty} K(u)du = 1 \qquad (5)$$

where b is the bandwidth. In our estimation procedure the kernel used is the Gaussian of zero mean and unit variance. In this case

$$\hat{f}_M(x) = \frac{1}{M\sqrt{2\pi b^2}} \sum_{i=1}^{M} e^{-\frac{(x-x_i)^2}{2b^2}} \qquad (6)$$

The Gaussian kernel allows to evaluate the kernel density estimator and the bandwidth with a low computational complexity (Raykar and Duraiswami, 2006).

3 Data Analysis

We applied the Fisher-Shannon information plane analysis to data of particulate matter concentration (PM10 and PM2.5) collected in several sites in the London area (www.londonair.org.uk). In particular we analysed 73 PM10 and 16 PM2.5 daily time series. The largest time span is between 1993 and 2010. The time variability of PM10 and PM2.5 is shown in Fig. 1 as an example. The measured data present gaps, but this does not influence the correct application of the methods, because the FIM and the Shannon entropy use the probability density function of the available data.

For each particulate matter time series, the FIM and the Shannon entropy power were calculated and the results plotted in Fig. 2, where each symbol represents the position of the time series in the FS information plane. The FIM of the PM10 is on the average lower than that of the PM2.5; on the contrary its Shannon entropy is higher. Thus, a clear relationship between the complexity parameters (FIM and Shannon entropy)

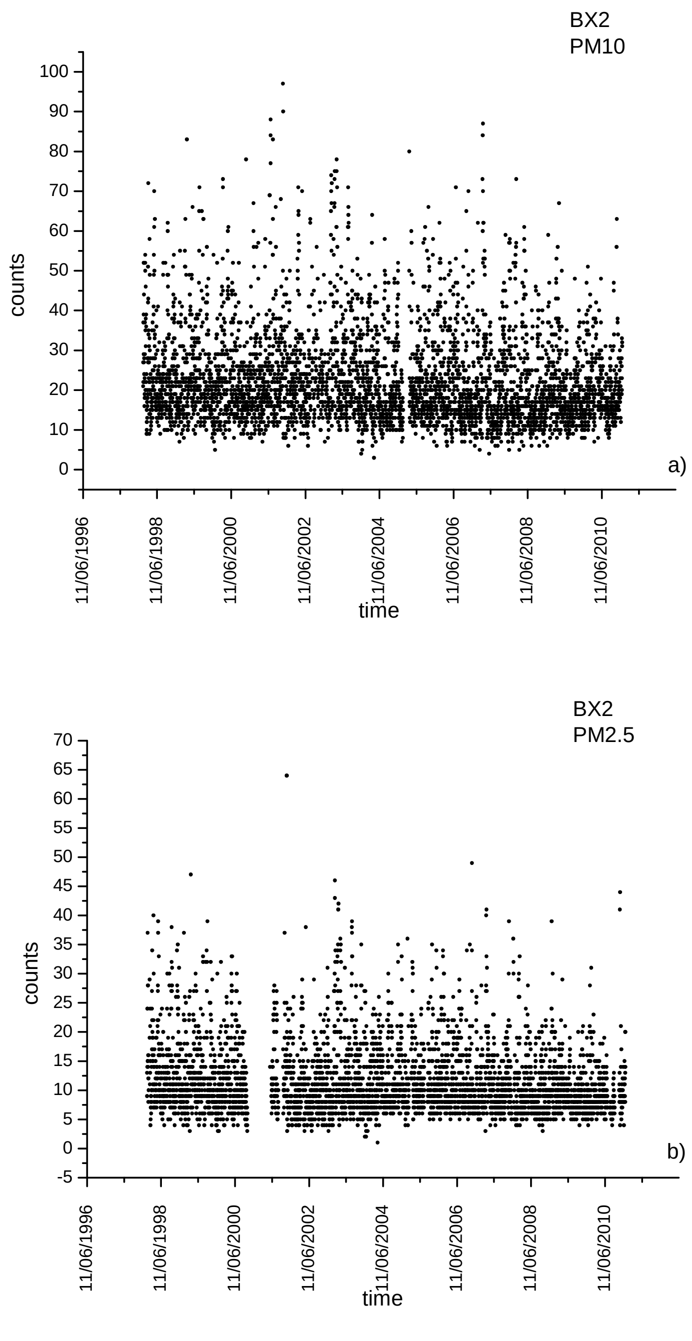

Fig. 1 Examples of time variation of PM10 (a) and PM2.5 (b).

and the size of the particulate matter can be observed. Finer particulate is characterized by less disorder and more organization than coarser particulate. This different complex behaviour shown by the two particulate size could be put in relationship with the complex dynamical interactions affecting differently the PM fraction. Fine particles are, in general, formed by means of nucleation processes (i.e., gas molecules coming together to form a new particle), condensation of gases onto existing particles and reaction in the liquid phase. Finer particles are mainly influenced by anthropic activities, which are generally well defined and located in a limited area. By contrast, most of the coarse particles are directly released into the atmosphere mainly from mechanical disruption such as crushing, grinding, or suspensions of dust from construction and agricultural operations (Telesca et al., 2009b). Furthermore, they are also affected by long range transport events like volcanic eruptions, forest fire, sea spray and desert dust (Querol, 2004; Alastuey, 2005). Moreover, the more uniformly dispersion in small and large geographic region of fine particles along with their longer atmospheric lifetime (i.e., days to weeks), can explain such more organization and order respect to larger particles, which generally deposit more rapidly than small particles (i.e., few days) and are less uniformly concentrated than fine particles.

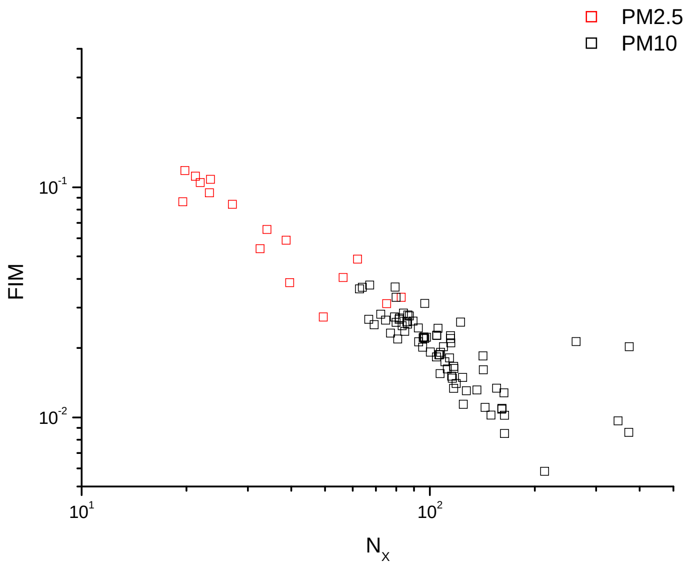

Fig. 2 Fisher-Shannon information plane results for the PM10 and PM2.5 time series.

4 Conclusions

We approached the analysis of several PM10 and PM2.5 time series recorded in London area by means of the Fisher Information Measure and the Shannon entropy, which act as detectors of order/disorder in dynamical behavior of complex systems. In the Fisher-Shannon information plane the two particulate fractions are well distinguished, suggesting that the different dynamical interactions affecting the two PM fractions could be linked with the order and organization degree. In the future studies, some advanced methods, like wavelet analysis (Gaucherel, 2011), may also be used to detect order/disorder in dynamical behavior of complex systems.

References

Angulo JC, Antolin J, Sen KD. 2008. Fisher–Shannon plane and statistical complexity of atoms. Physics Letters A, 372: 670-674

Balasco M, Lapenna V, Lovallo M, et al. 2008. Fisher Information Measure analysis of Earth's apparent resistivity. International Journal of Nonlinear Science, 5: 230-236

Chandra Mouli S, Venkata Mohan V, Balaram M, et al. 2006. A study on trace elemental composition of atmospheric aerosols at a semi-arid urban site using ICP-MS technique. Atmospheric Environment, 40: 136-146

de Araujo DB, Tedeschi W, Santos AC, Elias JrJ, Neves UPC, Baffa O. 2003. Shannon entropy applied to the analysis of event-related fMRI time series. NeuroImage, 20: 311–317

Dembo A, Cover TA, Thomas JA. 1991. Information theoretic inequalities. IEEE Transactions on Information Theory, 37: 1501-1518

Devroye L. 1987. A Course on Density Estimation. Birkhauser, Boston, USA

Esquivel RO, Angulo JC, Antolin J, et al. 2010. Analysis of complexity measures and information planes of selected molecules in position and momentum spaces. Physical Chemistry Chemical Physics, 12: 7108-7116

Fisher RA. 1925. Theory of statistical estimation. Proceedings of the Cambridge Philosophical Society, 22: 700-725

Frieden BR. 1990. Fisher information, disorder, and the equilibrium distributions of physics. Physical Review A, 41: 4265-4276

Fuhrman S, Cunningham MJ, Wen X, et al. 2000. The Application of Shannon Entropy in the Identification of Putative Drug Targets. Biosystems, 55: 5-14

Gao CY, Nelson ED, Field MP, et al. 2002. Characterization of atmospheric trace elements on PM2.5 particulate matter over the New York–New Jersey harbor estuary. Atmospheric Environment, 36: 1077-1086

Gaucherel C. 2011. Wavelet analysis to detect regime shifts in animal movement. Computational Ecology and Software, 1(2): 69-85

Hilborn RC. 1994. Chaos and Nonlinear Dynamics. Oxford University Press, Oxford, UK

IPCC (Intergovernmental Panel on Climate Change). 2011. Climate Change 2001: The Scientific Basis (JT Houghton et al. eds). Cambridge Univ. Press, New York, USA

Janicki A, Weron A. 1994. Simulation and Chaotic Behavior of Stable Stochastic Processes. Marcel Dekker, New York, USA

Martin MT, Pennini F, Plastino A. 1999. Fisher's information and the analysis of complex signals. Physics Letters A, 256: 173–180

Martin MT, Perez J, Plastino A. 2001. Fisher information and nonlinear dynamics. Physica A, 291: 523–532

Martuzevicius D, Grinshpun SA, Reponen T, et al. 2004. Spatial and temporal variations of PM2.5 concentration and composition throughout an urban area with high freeway density—the Greater Cincinnati study. Atmospheric Environment, 38: 1091-1105

Pope CA, Burnett R, Thun MJ, et al. 2002. Lung cancer, cardiopulmonary mortality, and long term exposure to fine particulate air pollution. Journal of the American Medical Association, 287: 1132-1141

Raykar VC, Duraiswami R. 2006. Fast optimal bandwidth selection for kernel density estimation. In: Proceedings of the sixth SIAM International Conference on Data Mining. Bethesda, 524-528

Romera E, Dehesa JS. 2004. The Fisher–Shannon information plane, an electron correlation tool. Journal of Chemical Physics, 120: 8906-8912

Shannon CE. 1948. A Mathematical Theory of Communication. Bell System Technical Journal, 27: 379-423, 623-656

Shendriker AD, Steinmetz WK. 2003. Integrating nepholometer measurements for air-borne fine particulate matter (PM2.5) mass concentration. Atmospheric Environment, 37: 1383-1392

Telesca L, Lovallo M, Carniel R. 2010. Time-dependent Fisher Information Measure of volcanic tremor before 5 April 2003 paroxysm at Stromboli volcano, Italy. Journal of Volcanology and Geotermal Research, 195: 78-82

Telesca L, Lovallo M, Hsu HL, et al. 2011. Analysis of dynamics in magnetotelluric data by using the Fisher-Shannon method. Physica A, 390: 1350-1355

Telesca L, Lovallo M, Ramirez-Rojas A, et al. 2009a. A nonlinear strategy to reveal seismic precursory signatures in earthquake-related self-potential signals. Physica A, 388: 2036-2040

Telesca L, Caggiano R, Lapenna V, et al. 2009b. Analysis of dynamics in Cd, Fe, and Pb in particulate matter by using the Fisher–Shannon method. Water, Air and Soil Pollution, 201: 33–41

Vignat C, Bercher JF. 2003. Analysis of signals in the Fisher-Shannon information plane. Physics Letters A, 312: 27-33

WHO. 2003. Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide. World Health Organization

Yu H, Liu SC, Dickinson RE. 2002. Radiative effects of aerosols on the evolution of the atmospheric boundary layer. Journal of Geophysical Research, 107: 4142