Environmental Pollution xxx (2016) 1–14

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Multifractal behavior of an air pollutant time series and the relevance to the predictability☆

Qingli Dong [a], Yong Wang [a, b, *], Peizhi Li [a]

[a] School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China
[b] Postdoctoral Research Station of Dongbei University of Finance and Economics, Dalian 116025, China

ARTICLE INFO

Article history:
Received 1 August 2016
Received in revised form
21 November 2016
Accepted 30 November 2016
Available online xxx

Keywords:
Multifractality
Air pollutants
Predictability
Spectrum analysis

ABSTRACTS

Compared with the traditional method of detrended fluctuation analysis, which is used to characterize fractal scaling properties and long-range correlations, this research provides new insight into the multifractality and predictability of a nonstationary air pollutant time series using the methods of spectral analysis and multifractal detrended fluctuation analysis. First, the existence of a significant power-law behavior and long-range correlations for such series are verified. Then, by employing shuffling and surrogating procedures and estimating the scaling exponents, the major source of multifractality in these pollutant series is found to be the fat-tailed probability density function. Long-range correlations also partly contribute to the multifractal features. The relationship between the predictability of the pollutant time series and their multifractal nature is then investigated with extended analyses from the quantitative perspective, and it is found that the contribution of the multifractal strength of long-range correlations to the overall multifractal strength can affect the predictability of a pollutant series in a specific region to some extent. The findings of this comprehensive study can help to better understand the mechanisms governing the dynamics of air pollutant series and aid in performing better meteorological assessment and management.

1. Introduction

Serious air pollution has been witnessed in China due to increased energy consumption and rapid industrialization in recent decades. Combined with social and economic issues, air pollution can largely threaten the framework of sustainable development (Thatcher and Hurley, 2010). According to a study by the World Bank, China has 16 of the 20 most polluted cities in the world (López et al., 2011), and pollution has had serious effects on public health, resulting in various diseases, such as cardiopulmonary and respiratory diseases (Chudnovsky et al., 2014). Discharged material, such as SO_2, NO_2, CO, O_3, $PM_{2.5}$ and PM_{10}, is assessed by many air quality monitoring systems (AQMS) (Zhao et al., 2015) to inform the public about the level of air pollution. It is of critical importance to investigate the time structure of pollutant series not only to better understand the dynamic mechanisms (Bandowe et al., 2014) and design more efficient early warning systems but also to provide management with more information to characterize the behavior of pollutants, such as diffusion, dilution and coagulation (Xue et al., 2015).

However, air pollution systems are complex, being affected by local interactions and correlations between various factors. The diverse sources of air pollutants include industrial processes, vehicular emissions and energy production from power stations, coupled with complicated physical and chemical processes (Feng et al., 2015). Air pollution systems have various components, such as meteorological factors, atmosphere self-purification and solar radiation, all of which have certain influences on the evaluation of air pollution concentrations (Shen et al., 2016). In an open and dissipative system, all the factors are correlated and interact with each other on different timescales, making it difficult to analyze the structure and temporal variation of air pollutants. Therefore, more attention is being paid to the scaling formalism to analyze the structural characteristics (Gokhale and Khare, 2007).

Based on the concept of a fractal, in recent decades, the scaling behavior of the process of air pollution has attracted increased attentions. Various methods, including a computational fluid

☆ This paper has been recommended for acceptance by Dr. Hageman Kimberly Jill.

* Corresponding author. School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China.
E-mail address: ywang@dufe.edu.cn (Y. Wang).

http://dx.doi.org/10.1016/j.envpol.2016.11.090
0269-7491/© 2016 Elsevier Ltd. All rights reserved.

dynamics model (Chu et al., 2005) and statistical approaches (Gokhale and Khare, 2007), have been adopted to study the dynamical behavior of air pollutants. Specifically, linear unbiased estimates, correlation analyses, power spectrum analysis and so on are sophisticated approaches that have been widely used to analyze and describe the distribution, periodicity and trends of pollutant concentrations (Xue et al., 2015). By using rescaled range analysis, detrended fluctuation analysis (DFA) and spectral analysis, Kai et al. (2008). analyzed three pollution series and the daily air pollution index (API) of Shanghai. They found that there are two different power laws in these series, which indicated different self-origanized critical states. Continuous wavelet transform fractal analysis was introduced by Yuval and Broday (2010) to assess the predictability of pollutant time series. They found that the predictability of air pollutants are consistent with that of meteorological variables in the short scale. Based on mono-fractal analysis, Lee et al. (2006). analyzed the scaling behavior of a one-year series of hourly average O_3 and found that scaling invariance exists in the studied series and the box dimension is a decreasing function.

In all the studies mentioned above, the pollutant series are analyzed and modeled as monofractal series, which are more suited to model homogeneous time series because there is only one scaling component and the properties are constant (Stanley et al., 1996). To advance the investigation of the dynamics of more heterogeneous and complex series, it is necessary to analyze more scaling exponents, including through fractal and multifractal analyses, especially when the original series is composed of many interwoven fractal subsets (Pamuła and Grech, 2014). Multifractality is regarded as the inherent property of complex and composite systems, which have attracted much attention in recent years.

Multifractality theory is widely used to quantitatively delineate the nonlinear evolution of a complicated system and the multiscale characteristics of physical quantities. It can also aid in understanding the intrinsic regularity and mechanism of physical changes (Windsor and Toumi, 2001). Shen et al. (2016). described and analyzed the multifractal characteristics of the air pollution index (API) in China, which provided a basis for further probing into the complexity of API series. Liu et al. (2015). used multifractal detrended fluctuation analysis (MF-DFA) to characterize the temporal fluctuations of the API record and three common pollution indexes of Shanghai in China. The results showed that the temporal scaling behaviors in studied pollutant series present different power-law relationships. Applying Chhabra and Jensen's multifractal formalism, Munoz Diosdado et al (Munoz Diosdado et al., 2013). analyzed an atmospheric pollutant concentration series from 1990 to 2005 with multifractal analysis. They confirmed the existence of multifractality in this series, which indicated a new level of complexity distinguished by the wide range of necessary fractal dimensions to characterize the dynamics of the air pollutant time series.

Although the fractal and multifractal characteristics of pollutant series in some regions have been investigated, the mechanism of variation in fractality and multifractal sources and the relevance to the predictability have not be fully clarified. By studying the multifractal nature of the measured data, we intend to investigate their persistence properties and heterogeneity features, which can help us understand the structural complexity of pollutant records and provide theoretical support in pollutant series forecasting. The novelty of our work along with the acquired nontrivial findings can be concluded as follow:

- The power spectra of different pollutant series are computed by adopting the basic modulus-squared of the discrete Fourier transform, and different frequency regimes are found. The distinct decreasing trends at low and high frequencies testify to the existence of power-law behaviors and long-range correlations at these frequencies.
- Compared with the traditional method of detrended fluctuation analysis that was widely used in previous studies, this research use spectral analysis together with the multifractal detrended fluctuation analysis, which are based on the stochastic processes, chaos theory and time series analysis, to investigate the statistical self-affinity and multifractality of a pollutant series.
- The time dynamics of pollutant series can be explored at lower timescales with higher frequency datasets. The datasets used in this research is the measured hourly observations of pollutant series that are collected from five important cities in China.
- It is found that different long-range temporal correlations for small and large fluctuations and a fat-tailed probability distribution of variations are two major sources for multifractality, and the fat-tailed probability distribution contributes more to the multifractality.
- A rough relationship between the forecasting accuracy of pollutant series and the parameter of multifractality is found in this paper, which means the contribution of the long-range correlations' multifractal strength to the whole multifractal strength can affect the predictability of pollutant series in a specific region to some extent. The conclusion about the relationship is in agreement with Yuval and Broday (2010). However, this relationship is quantified and unveiled by quantitative analysis and comparison, which has never been obtained so far to the best of our knowledge.

The findings of this study can help to better understand the mechanisms governing the dynamics of air pollutant series and aid in performing better meteorological assessment and management in early-warning system.

The reminder of this paper is organized as follows. The data acquisition and analysis are described in section 2. Section 3 provides the spectral analysis briefly. In section 4, detailed discussions of the empirical results and extended analyses are presented. Lastly, section 5 presents the conclusions of this paper.

2. Data acquisition and analysis

Original hourly records of particulate matter concentrations are measured and published on the air quality publishing platform of China (http://www.aqistudy.cn/historydata/index.php) and the Ministry of Environmental Protection of the People's Republic of China (http://english.mep.gov.cn/), which are our main data resources. In addition, parts of the hourly pollutant data were collected by our self-developed Java programs from May 13, 2014 to Nov 12, 2016. Because the quality of collected records is controlled, the fraction of missing data is small. Specifically, for the record of $PM_{2.5}$ and PM_{10} in Shanghai, the number of entry/missing data is 20876/146 and 20876/549, respectively. In this paper, we adopt the cubic spline interpolation approach (Carrizosa et al., 2013) to handle the missing data.

Contour plots of the originally hourly $PM_{2.5}$ concentrations and PM_{10} concentrations are depicted in Fig. 1. To identify the features of air pollutant regimes and estimate the PM emission characteristics, the lognormal distribution can usually be utilized. In addition, it is an important part for PM forecasting and early warning systems. The probability density function (PDF) and the cumulative distribution function (CDF) of the Lognormal distribution are given by the following formula:

Please cite this article in press as: Dong, Q., et al., Multifractal behavior of an air pollutant time series and the relevance to the predictability, Environmental Pollution (2016), http://dx.doi.org/10.1016/j.envpol.2016.11.090