Multifractal Cascade Analysis on the Nature of Air Pollutants Concentration Time Series Over China

Xin Gao*, Xuan Wang1, Hongbin Shi1

1School of Urban-rural Planning and Landscape Architecture, Xuchang University, Xuchang-461000, China

Abstract

This study investigates the temporal multifractal cascade behaviors of the time series of air pollution in major cities of China, taking one-year data of hourly time series of six pollutants (PM2.5, PM10, CO, NO2, O3, SO2) and Air Quality Index (AQI) as model inputs. Considering the right-skewness with heavy tail widely existed in an air pollution time series, the stable distribution with four parameters is used to fit the frequency distributions of all the data, and the result shows that it can give well fittings over the whole range like the log-normal model. The shared periodicity and nonstationarity are also investigated based on spectral analysis, and the usual cycles such as daily and semi-daily cycles and two dynamical regimes with a crossover at approximately 11 days are widely found in all the pollution series. In the universal multifractal analysis, the underlying significances of the two parameters in air pollution system are addressed by means of the analogues of simulation data. Moreover, an evidence of first order multifractal phase transitions is found through self-organizing criticality analysis on the data of Air Pollution Index (API), which can provide an available estimation interval of the moment order $q_D$ varying between 1.95 and 2.98. Finally, downscaling performances of the multifractal cascade models are numerically investigated with the aim to explore the potential applications in the air pollution field, which demonstrate that log-normal and log-Poisson models, but not $\alpha$, $\beta$, binomial and uniform models, can effectively recover the extreme values.

Keywords: Cascade; Multifractal; Air Pollutant; Time Series; Downscaling

1. Introduction

Air pollution has become one of the most concerned environmental problems in the world because of its significant negative impacts on human health and restrictions on

* Corresponding author. Tel: (86) 13569480352
E-mail address: gxin826@126.com
national development. However, environmental pollution is a complex system, which is essentially the result of a series of physical and chemical interactions between natural and anthropogenic environmental conditions (Mayer, 1999; Kinney, 2008). With the heavy dependence on the fluctuations of a variety of meteorological and emission variables, air pollutant concentrations randomly and irregularly varied with regard to time and space. Knowledge of the structure of the time series of the concentrations of major air pollutants is needed in order to make pollution prevention policies and better understand dynamical mechanisms governing their temporal variabilities (Xepapadeas 1992; Liu et al., 2003; Esposito et al., 2016).

Time series of air pollution data measured at any location usually distributes as a strong right-skewed uni-modal shape (Georgopoulos and Seinfeld 1982). Identifying optimal parent probability functions for these distributions is fundamental for precisely predicting concentration related statistics (Sedek et al., 2006; Yahaya et al., 2011; El-Shanshoury et al., 2017). Several probability density functions (log-normal, gamma, Weibull, beta, log-logistic, Pearson type V, etc.) have recently been frequently and successfully used to fit the measured data of air pollution (Karaca et al., 2005; Gavril et al., 2006; Leiva et al., 2010). Some of these studies estimated the exceedance frequencies of air pollutants concentrations over the air quality standard (AQS) and also predicted the needed emission reduction in order to meet the AQS (Marchant et al., 2013). Among all the applied parent probability models, the log-normal distribution is used the most; however, results from this distribution, while fitting well with the observed data with medium level of pollution, do not fit well to
heavy pollution cases with high concentrations. In the latter case, the type I
two-parameter exponential distribution is a better choice (Maciejewska et al.,
2015). Earlier studies have also suggested that the optimal model depends on
pollutant species of interest, e.g. the lognormal distribution for particulate matter
and the majority of the nitric oxide, oxides of nitrogen and sulphur dioxide, the
gamma distribution for carbon monoxide, nitrogen dioxide and ozone, and the
Weibull distribution for carbon monoxide and ozone (Taylor et al., 1986). In addition,
power law distributions has also been frequently employed to fit concentration values
of air pollutants (Kütchenhoff and Thamerus, 1996).

Compared to the complicated mathematical models with huge computational cost,
the interpretation and regulation about the distributions of concentration values for air
pollution time series through some easy-to-operate physical models are obviously
inspiring and encouraging. In order to give a physical explanation of the lognormality
of pollutant concentrations, Ott (1990) proposed a successive random dilution (SRD)
theory which refers to processes of dilution and mixture at a series of stages by
constantly adding fresh air into newly added containers, and the resulted
concentrations are approximately lognormally distributed. As a counterpart, this
phenomenological idea can be traced back to Richardson’s celebrated poem on
self-similar cascades describing atmospheric dynamics as a cascade process
(Richardson 1965; Schertzer and Lovejoy, 2011). The idea was strictly conceptualized
with help of fractal theory until 1980s, before evolving into multifractal. Since then
several probabilistic cascade models have been built and widely applied such as in the
scaling analysis of fully developed turbulence, downscaling of soil moisture and evaluation of mineral resources (Schertzer et al., 1997; Agterberg 2007; Mascaro et al., 2010), with the simplest monofractal and multifractal models being $\beta$-model and $\alpha$-model (Schertzer and Lovejoy, 2011). Moreover, inspired by simple cascade models, the general forms of multifractal, also called multifractal formalisms, were deduced when scales tend to be infinitesimal. Of the two most famous formalisms, the multifractal spectrum $f(\alpha)$ are adopted for $\alpha$-$f(\alpha)$ model, while $\gamma$-$c(\gamma)$ model lies in the codimension function $c(\gamma)$ respectively. (Evertsz and Mandelbrot, 1992; Schertzer and Lovejoy, 1987). In the past 30 years, techniques of specific calculation and implementation for multifractal formalisms have been proposed, which include box-counting method, histogram method, probability distribution, wavelet analysis and detrended fluctuation analysis (Lovejoy et al., 1987; Meneveau and Sreenivasan, 1989; Lazarev et al., 1994; Muzy et al., 1994; Kantelhardt et al., 2002). Indeed, multifractal analysis has gradually become a standard technique for quantitatively delineating the nonlinear evolution of a complex system and the multiscaling characteristics of physical quantity, and used to aid the understanding of the intrinsic regularity and the mechanism of physical changes (Shen et al., 2015). For instance, Shi et al. (2015) studied the dynamic characteristics of PM$_{2.5}$ series during a heavy haze episode of Chengdu City by constructing a self-organized critical model and showed that the multifractal features of the series were highly correlated with long-range correlation and non-Gaussian distribution. Gao et al. (2016) studied the nonlinear characteristics of the spatial distribution of particulate matter in China,
resulting in a power law distribution with a spatial scale index of $\beta=0.45$ and a mesoscale range of 90-550 km.

To date most studies focused on characterizing multifractal properties that reflect tempo-spatial structure complexities inherent in the air pollution systems, and investigating relationships between air pollutants and metrological factors for seeking the direct evidences of mechanical variations in fractality. Thus, more research is needed on the multifractal formalisms based on the probabilistic cascade models representing physical processes in air pollution systems. For example, scale invariance mechanism provides a creative idea for straightly establishing a relationship to connect the concentration values at different scales, making it possible to easily obtain the concentration values at infinite finer scales. Such a tool is useful for calculating frequencies at finer scales exceeding air quality standards so as to estimate more accurate reduction in favor of providing more better and effective decision-making for local government. Besides, in-depth discussions about the physical meanings of multifractal parameters in air pollution systems are still lacking, thus limiting their applications in pollution assessment studies. In addition, the continuity of a multifractal spectrum also needs to be addressed, which may provide some useful information for extreme analysis. This is especially the case in the determination of critical values of the moment order $q$ for representing typical concentrations when Monte Carlo simulation is used.

The present study tries to fill some knowledge gaps discussed above. Firstly, the basic statistical analysis of the confirmation of right-skewed, fat-tailed, long-range
dependent and nonstationary features for air pollution time series is made with
spectral analysis and frequency distribution fitting methods. Secondly, the feasibility
of using multifractal parameters as an evaluation tool of air pollution is discussed, i.e.,
using the calculated multifractal parameters to evaluate the pollution levels and
identify the main pollutants in seven cities, in addition, the continuities of
codimension functions of air pollution time series are discussed to determine the
estimation intervals for the moment order $q$ and the types of discontinuity points.
Finally, the downscaling capability in air pollution is discussed based on several
probabilistic cascade modes and multifractal characteristics.

2. Methodology

Data collection and quality assurance and control

Six air pollutants including CO (mg/m$^3$), NO$_2$ (µg/m$^3$), O$_3$ (µg/m$^3$), PM$_{10}$ (µg/m$^3$),
PM$_{2.5}$ (µg/m$^3$), and SO$_2$ (µg/m$^3$) were measured using automated monitoring systems
widely distributed in cities at county level and above. SO$_2$, NO$_2$ and O$_3$ were
measured by the ultraviolet fluorescence method, the chemiluminescence method, and
the UV-spectrophotometry method, respectively; CO was measured by the
non-dispersive infrared absorption method and the gas filter correlation infrared
absorption method; and PM$_{2.5}$ and PM$_{10}$ were measured by the micro oscillating
balance method and the $\beta$ absorption method.

AQI is a dimensionless index to quantitatively describe air quality. An air quality
index is defined as the maximum of the individual air pollutant quality indices
standardized to the corresponding concentrations of the six pollutants separately based on AQI Equation and air quality standards. Seven cities in China, including Beijing, Shenyang, Shanghai, Wuhan, Chengdu, Guangzhou and Xi'an, representing different regions as well as climatic and topographic conditions are selected for investigation (Fig. 1). Time series of the hourly concentration of six pollutants and AQI from January 2016 to December 2016 were used in the analysis (available from the China National Environmental Monitoring Center, http://106.37.208.233:20035/). In addition, 13 years (2000-2012) daily API data are also used for phase transitions and self-organized criticality analysis. Linear piecewise interpolation method is used for filling the data gaps caused by power outages, equipment failures and other situations.

Methods

For a scale invariant measure $\varepsilon$, which can be a value of rainfall intensity or pollutant concentration, a fractional Brownian motion (fBm) can be constructed by incremental variances of the measures at different scales. Moreover, the parameter $H_u$, also called the Hurst index, is commonly used to express the intermittency and variation for a fractional Brownian motion. Obviously, a fractional Brownian motion is an addition process, in which each point is generated by the addition of the variances of two adjacent points and a random number, and $H_u$ can be calculated using the probability density function method and R/S analysis. The corresponding to the addition process is the multiplication process, also called cascade process. For example, if we consider the 1D cascade process, it is assumed that there is an amount $\varepsilon_1$ with a certain support of unit length. Firstly, the unit is divided into two subunits of
equal length, which are each multiplied by a random variable $W_1$ that satisfies a
certain probability distribution; these values are denoted as $\varepsilon_{2,1}$ and $\varepsilon_{2,2}$ with $<W_1>=1$,
where $W$ is the weight and $<>$ means expectation. Secondly, according to the same
operation, the resulting subunits of the segmentation in the previous step are further
divided, multiplied by the same distributed random variable $W_2$, and this process is
repeated $n$ times to obtain a series $\varepsilon_n$. For $\varepsilon_n$, the geometrical supports become $r=1/2^n$,
meaning the finest scale is $1/2^n$, and $1/r$ is $\lambda$, where $\lambda$ is the scale ratio. From the
generation process shown in Fig. 2, it can be seen that as the scale ratio increases, $\varepsilon_n$
gradually becomes more intermittent and has a greater oscillation.

A parameter is needed to describe the fBm scale invariance and intermittence,
while the latter requires multiple fractal dimensions to be described. Assuming there
is a real value $\gamma$, as $\gamma$ increases, there is a relationship between the probability
distributions of the values larger than $\lambda^\gamma$ and the scale rations: $\Pr (\varepsilon_n>\lambda^\gamma) = \lambda^{-c(\gamma)}$, where
$c(\gamma)$ is a function of $\gamma$ and is the characterization of the support sets corresponding to
each value $\gamma$ for $\varepsilon_n$. The function $c(\gamma)$ is also called the codimension function and
corresponds to the dimension function $f(\alpha)$ proposed by Evertsz and Mandelbrot
(1992), where $\gamma$ and $\alpha$ are singularity indexes. The $\gamma$-$c(\gamma)$ model, a multifractal
formalism widely used in the fields of turbulence, rainfall, weather, stocks, and
pollution, was first proposed by Schertzer and Lovejoy (1987). To obtain $c(\gamma)$, we
need to calculate the statistical properties of the higher moments $<\varepsilon_n^q>$ of $\varepsilon_n$, and there
is a relationship between $<\varepsilon_n^q>$ and the scales: $<\varepsilon_n^q>=\lambda^{K(q)}$, where $K(q)$ is the scaling
function and there are the following relationships: $K(q)=\min(q\gamma-c(\gamma))$ and
\(c(\gamma) = \min(q^\gamma - K(q))\), with \(c(\gamma)\) determined by the Lengdre transformation. The above \(\gamma - c(\gamma)\) model is based on probability theory, while the \(\alpha - f(\alpha)\) model is based on measurement theory, and there is also the following relation between the two models:

\(f(\alpha) = 1 - c(\gamma)\), \(\alpha = 1 - \gamma\). More comparisons between the two models can be referred to Cheng and Agterberg (1996). Compared to one parameter of mono-fractals, \(K(q)\) and \(c(\gamma)\) contain an infinite number of parameters. For simplicity, Schertzer derived the universal multifractal model by nonlinear mixing and densification of different cascade processes with the following formulas:

\[
K(q) = \frac{C_1}{\alpha - 1} (q^\alpha - q), 0 \leq \alpha \leq 2
\]

(1)

\[
c(\gamma) = C_1 (\frac{\gamma}{C_2 \alpha^2} + \frac{1}{\alpha})^\alpha, \frac{1}{\alpha}, \frac{1}{\alpha^2} = 1
\]

(2)

where \(\alpha\) denotes the multifractal intensity, and \(C_1\) is the codimension of the information dimension (Schertzer et al., 1987).

### 3. Results and discussions

**The statistical features of the air pollution time series**

The possible divergence of high moments for statistical quantities about heavy-tailed random variables indicates the inefficiency and instability on the estimation or analysis by using traditional probability theory, and thus a new probabilistic model is needed to assign. For the extremes or outliers in heavy tail, it is theoretically considered that their variances are infinite, and the sums of them will no longer be normally distributed, instead they will be random variables of a stable distribution. The probability density function of this model commonly is defined
based on 4 parameters, including stability parameter $\alpha_s$ allowing the tail shape of PDF to be adjusted, skewness $\beta$, scale parameter $c$, and position parameter $\mu$. For a stable distribution, there are three known special cases, when $\alpha_s=2$, it reduces to a normal distribution, while $\alpha_s=1$, $\beta=0$ and $\alpha_s=1/2$, $\beta=1$, it corresponds to a Cauchy distribution and Levy distribution respectively. Taking the PM$_{2.5}$ time series of Beijing as an example, the fitting results by using log-normal distribution, stable distribution, normal distribution and kernel function models are shown in Fig. 3, which exhibit well fittings for log-normal and stable models over the whole range, also reveal that the time series is right-skewed with heavy tail. The features can be extended to the rest of the pollution series of 7 cities, and Table 1 and Table 2 list the estimates of $\alpha_s$ and $\beta$ for all the pollution time series.

For a relation of a spectrum versus wave numbers in a periodogram of a time series, which can be built under the Fourier transformation, the curve can reflect a distribution of spectral energy at different scales, and also allows to detect periodic characteristics for the time series. Generally speaking, in the curve an energy peak represent one cycle, clearly having higher value than surrounding areas. Like Fig. 4, the daily and semi-daily cycles are experimentally confirmed to be commonly existed in all of the time series, and like Fig. 5, not a few time series are also characterized by the presence of lower energy peaks at 8h, 6h and even less. It is evident that the daily and semi-daily cycles are closely associated with the forces driven by the Sun, which can deeply affect human activities such as various anthropogenic sources and meteorological conditions such as diffusion and deposition factors. However, it is
difficult to explain the other energy peaks, which may be involved with the other influencing factors such as pollutants properties (physical or chemical) or complex interactions between natural and anthropogenic environmental conditions, and it is noted that the cycles should not be ignored from their comparison of amplitudes with the daily or semi-daily cycles.

For a time or space series, the nonstationarity means the values differ with respect to time and location, and requires that when making a forecasting or an analysis the underlying trend need to be removed. The multifractal cascade framework provides a new effective idea to distinguish and identify the feature. As argued by Davis et al. (1993), when $\beta<1$, it indicates that the process is stationary, and $\beta>1$ means the process is non-stationary, while for $1<\beta<3$, the increment of the process is stationary. Here, using the PM$_{2.5}$ series of Beijing as an example, it can be seen from Fig. 6 that the spectrum can be divided into two parts at the point of 11 days, which is consistent with the conversion scale between low-frequency weather and the weather with a slope $\beta$ that is close to 0 at large scales and 5/3 at small scales in the turbulence field (Lovejoy and Schertzer, 2012). Meanwhile, the valley value appears at approximately 60 days, and from the performance of $\beta$, the series is found to be non-stationary. Through the further investigation, non-stationarity and the crossover between the two regimes corresponding to different scales also generally exist in the rest of pollutants series. Another parameter related to $\beta$ is the Hurst index $H_u$, and there is a relation of $\beta = 2H_u + 1$, where $H_u$ can be used to determine the positive or negative long-range correlation; specifically, when $0<H_u<1/2$, the series is positive, and when $1/2<H_u<1$, ...
Multifractal analysis has gradually become a standard high moments statistical method, which can relate the values of different scales through a scaling function and describe the scale invariance for a time series through multiple dimensions. For the estimation of the multifractal formalism, here we adopted a method called double-trace moment involving two moment operations to a scaling quantity, which has been proven to be more effective and stable than moment method (Lavallée, 1991). As proposed by Schertzer (1987), firstly, a non-stationary series must be converted to a stationary series using the relation $\Delta v \approx \varepsilon^{1/3} \kappa^{1/3}$, where $\Delta v$ is a difference in velocities at two arbitrary points with an certain interval in a turbulence field, and $\varepsilon$ and $\kappa$ are the conservative energy and scaling variables respectively. Here, the concentration values are observed directly and act as substitutes for the velocity $v$.

The specific estimation steps are as follows: firstly, first-order differences are performed on the minimum scale for a time series; then, the absolute values are used instead and divided by the mean value, thus the processed time series is normalized to obtain a new and conservative time series $\varepsilon_i = |\Delta \mu|/\langle \Delta \mu \rangle$, where $\mu$ denotes the concentration values of an air pollution time series; finally, a series of statistical quantities at various scales defined based on the average for the above result $\varepsilon_i$, are subjected to power operations two times with the orders of $\eta$ and $q$ respectively to obtain $Tr\lambda(\varepsilon_i^\eta)^q$, which are related to $\lambda$ by $Tr\lambda(\varepsilon_i^\eta)^q \approx \lambda^{K(q, \eta)}$, where the functions $K(q, \eta)$
are the scaling functions of double moments and are related to the moment scaling function \( K(q) \) as \( K(q, \eta) = \eta^\alpha K(q) \). The double trace moment estimators are functions with three variables, and thus the quantities are all three-dimensional arrays. Fig. 7 shows the relations of \( \text{Tr}\lambda(\epsilon^q\eta)^q \) versus \( \lambda \) of the PM\(_{10}\) series of Beijing, and we can see that there are good linear relationships between them; in other words, the scale invariance remains within certain intervals, and the values of \( K(q, \eta) \) are accurately estimated by the least square method. Once the values of \( K(q, \eta) \) are known, then \( \alpha \) is estimated by the relation \( K(q, \eta) = \eta^\alpha K(q) \), and \( C_1 \) can also be estimated by the Equation 1 and 2.

The universal multifractal formalism, where \( \alpha \) is a Levy variable, in general gives a noise with such parameter as a generator to form an universal multifractal process. For a process, four types are categorized with respect to \( \alpha \). When \( \alpha = 0 \), it is a mono-fractal; when \( 0 < \alpha < 2 \) except \( \alpha = 1 \), it is a log-Levy multifractal, and in the other cases when \( \alpha = 2 \) and \( \alpha = 1 \), it corresponds to a log-normal and log-Cauchy multifractal respectively. In the formalism \( \alpha \) is commonly used as a characterization for multifractal intensity and reflect hierarchical structure for a process, while \( C_1 \) is a first-order statistic, also called information co-dimension which reflects sparseness of mean process. For \( C_1 \), having a relation \( C_1 = 1 - D_1 \) with information dimension \( D_1 \), which to some extent can represent the background information, or give an overall description for a process. A more intuitive understanding of the physical meanings for the two parameters can be obtained from Fig. 8, which presents four simulated universal multifractal process with (a) \( \alpha = 0.6, C_1 = 0.12 \), (b) \( \alpha = 1.6, C_1 = 0.12 \), (c) \( \alpha = 1.2 \), (d) \( \alpha = 0.2 \).
\( C_1 = 0.12 \) and (d) \( \alpha = 1.2, \ C_1 = 0.52 \). (URL of the specific simulation method: http://www.physics.mcgill.ca/~gang/software/index.html). According to Fig. 8a and 8b, compared to the lower \( \alpha \) having more strong peaks, the higher \( \alpha \) owns less but stronger upward peaks, and it indicates that extreme pollution risk will disappear but instead a long term higher pollution when \( \alpha \) decreases. According to Fig. 8c and 8d, compared with lower \( C_1 \), for higher \( C_1 \) the degree of sparseness increases, and some extreme outliers will occur, it seems to make a compression and stretch for the data as a whole.

With the underlying significance in a process for the two parameters, next let's discuss the overall characteristics of all the air pollution time series in the seven cities. As seen from Fig. 9 and Fig. 10, in addition to Chengdu, the values of \( C_1 \) and \( \alpha \) of the three pollutants PM\(_{2.5}\), PM\(_{10}\) and AQI observed in the northern cities are generally greater than those in the southern cities due to the heating and easily calmed weather in the north China during winter, and also indirectly indicates that the north is the main polluted area of China in winter. It is noted that the maximum values of \( \alpha \) for PM\(_{2.5}\), PM\(_{10}\), and AQI occur in a western city, Chengdu, which imply that there are more stronger pollution peak areas than the other cities during 2016. This result may be caused by Chengdu's unique natural and geographical environment. Chengdu is located in the Sichuan Basin in Western China, a very closed geographic unit, and is perennially characterized by low wind speed, high cloudiness and humidity, and more neutral and stable weather conditions, and so forth; therefore, the inversion occurs frequently with high intensity, and the pollutants diffuse poorly. With the higher
values of $C_{1,PM_{2.5}}$, $C_{1,PM_{10}}$ and $C_{1,AQI}$ for Beijing, it can be concluded from the time series that the extreme pollutions yielded from the sparseness occurred on February 8 and May 6, 2016. The former is caused by the calm weather during the heating period, which results in a worsening of pollutant diffusion conditions and is afterwards coupled with the migration from other areas, yielding the extreme pollution; moreover, the latter is caused by the dust storms and reaches extreme values in a short time. The pollution for the three pollutants in Shanghai, Guangzhou, and Wuhan are relatively less serious; however, Guangzhou and Wuhan need to face with certain pollution risks from PM$_{2.5}$ and PM$_{10}$ respectively. Shenyang has the largest $\alpha_{SO_2}$ and the lowest $C_{1,SO_2}$, i.e., it has the strongest pollution peaks and the smallest sparse degree for SO$_2$, indicating that the pollutant consistently maintain at a high level of pollution, and an irrational energy structure, especially coal-fired heating, is the main reason. Wuhan has a highest level of O$_3$ pollution, which is also identified from the measured series with high concentration values in July and August, and it is mainly caused by automobiles and marine exhausts, volatile organics used in spray paintings, dyes, and petrochemical productions, which are processes that also lead to the larger value for $C_{1,CO}$, while the pollution level of O$_3$ in Beijing is relatively low. The maximum values of $\alpha_{CO}$ and $\alpha_{NO_2}$ occur in Xi’an, and the maximum value of $C_{1,NO_2}$ occurs in Shenyang, and the pollutants mainly come from vehicle exhaust, coal-fired boiler heating and thermal power generation. Contrary to the other pollutants, $\alpha_{CO}$ and $C_{1,CO}$ roughly show a negative correlation, indicating that there are fewer pollution peaks in southern cities, such as Shanghai, Wuhan, and Guangzhou, but they largely deviate
from normal levels, i.e., meaning a rapid increase of CO in a short time against a stable background, and the case could also cause greater harm to human health and ecosystem. This is also the reason why some southern cities have listed automobile exhaust as their primary pollutant.

Although the above scaling behaviors of the pollutants were observed through the two parameters $\alpha$ and $C_1$, among which $C_1$ is a first-order moment parameter and although $\alpha$ reflects a high-order moment characteristic of the time series, it does not intuitively provide the scaling features of all the moments. It should be noted that the codimension function $c(\gamma)$ refers to structures that can be described as a collection of interwoven fractal subsets that exhibit power-law behavior with a range of scaling exponents. A comparison of the estimates of $c(\gamma)$ for the AQI series among different cities is shown in Fig.11, where the axis of abscissa denotes singularity index $\gamma$, representing a sample space with the concentration values constantly increasing from left to right, on the other hand, the ordinate axis shows the estimates of codimensions $c(\gamma)$, i.e., the higher it is, the larger the sparseness degree of the time series is. From the figure, it can be concluded that the highest peak of AQI occurs in Beijing by the maximum estimates of $\gamma$, indicating that it is exposed to the most serious extreme pollution episode, and in contrast, the risks of extreme pollution in Guangzhou, Shanghai, and Wuhan maintain at a relatively low level. Similarly, through the comparisons of the other pollutants among different cites, more major and prominent pollutants can be identified and deduced for each city: Beijing, Chengdu, and Xi'an each have three pollutants, i.e., PM$_{2.5}$-PM$_{10}$-SO$_2$, NO$_2$-PM$_{10}$-SO$_2$, and O$_3$-PM$_{10}$-SO$_2$. 
respectively; Wuhan has high extreme pollution risks of CO and O₃, while the risks in
Guangzhou and Shenyang are mainly associated with CO and PM₁₀ pollution,
respectively.

Due to limitations regarding resolutions and sample sizes, \( c(\gamma) \) is not always
continuous in the domain of definition, and there are two kinds of discontinuities.
From the perspective of the self-organized criticalities of the physical systems, they
are connected with the phase changes in the physical processes, and a first order and a
second order phase transition were proposed by Schertzer and Lovejoy (1993)
through the study of the cascade processes. For a cascade process, when the sample
size is large enough, the discontinuity point \( q_d \) can be revealed, which is also called
the first order phase transition, and the corresponding theoretical limit values \( \gamma_{s,d} \) and
\( c(\gamma_{s,d}) \) can also be obtained. When the sample size is not sufficient, the discontinuity \( q_d \)
can’t be reached, but the sample discontinuity \( q_s \) can be obtained, which is called the
second order phase transition. Therefore, the sample size is very important in the
determination of the theoretical limits of \( \gamma \), \( c(\gamma) \) and other parameters especially the
moment order \( q \). Next the quantitative analysis of the API data of seven cities within
the period from 2000 to 2012 is conducted to observe the multifractal phase
transitions in air pollution systems.

Before calculation, each series needs to be divided into two time segments, which
are set to be 4 years and 13 years respectively in the scenario. The specific calculation
process is as follows: taking a series as an example, firstly, the long-tail behavior of
the survival function for a conservative measures differenced from the series at the
minimum scale is fitted to obtain $q_d$, and then the scaling moment function $K(q)$ is fitted using the formula $K(q) = \gamma_{\text{max}} q - \Delta s$, where $\gamma_{\text{max}}$ denotes the maximum of singular values, and $\Delta s$ is the effective dimension, which correspond to the slope and intercept of the equation respectively. Secondly, the parameters $\gamma_D$, $q_s$, and $\gamma_s$ are calculated using the formulas $\gamma_D = K'(q_d)$, $q_s = (\Delta s/C_1)^{1/\alpha}$, and $\gamma_s = K'(q_s)$. The results analyzed as a representative of five cities selected for the two different lengths of data are listed in Table 3. For each city, $q_s > q_D$ and $\gamma_s > \gamma_D$ indicate that the temporal variations of the API series observed in air pollution systems belong to the first multifractal phase transitions. Theoretically, the difference of both $q_D$ and $\gamma_D$ under the different sample sizes are quite small, and $\gamma_s$ increases as sample sizes increase. However, for $q_D$ and $\gamma_D$, a remarkable change appears in Xi’an and Shanghai, which means that the concentrations vary drastically with respect to the changes in time; in contrast, $q_s$ and $\gamma_s$ only have a few minor changes in all cities. Moreover, the theoretical extreme values $\gamma_{s,d}$ of these data are equal to the fitted values $\gamma_{\text{max}}$ of the empirical data. The behaviors on the intervals $[\gamma_D, \gamma_{s,d}]$, between the first discontinuity points and theoretical maximums are fully characterized, and it also has been proven that $c(\gamma_{s,d}) \approx \Delta s$ and oscillates around the theoretical value of $1 + \log(13)/\log(365) = 1.435$. The downscaling behaviors of multifractal cascade analysis applied in the air pollutants concentration data

Generally, the pollutants concentration time series are always measured at fixed resolutions, and normally the concentration over a sampling interval varies with resolutions. However, people are often more interested in observing and obtaining the...
infinite fine-scale concentrations. In general, the concentrations at coarser resolutions are averagely combined to generate from corresponding values at finer resolutions, where the smoothness undoubtedly results in the losses of some extreme information, but the extremes are not easily recovered and regained by using the traditional imputation and interpolation methods. As mentioned earlier, fractals or multifractals have natural advantages in describing the extreme values based on a hierarchically generating mechanism realized by a power law relation of scaling moment function, and it indicates that they can provide a scale-independent random downscaling tool. From the generation process of a one-dimensional cascade, the parent node is multiplied by a random variable to obtain twice its number of child nodes. With the different theoretically distributed forms of random variables, multiple cascade models were proposed with the specific formulas of the function $K(q)$ shown in Table 4 (Gupta et al., 1993). Obviously, the key to downscaling is to calibrate the parameters in the theoretical models through the empirical data and then generate the downscaled series using the most appropriate one from the well calibrated models.

In order to examine the dependence on the distributions of the simulations using the six models, Fig. 12 shows a comparison of the fittings to the histogram of the Shanghai PM$_{10}$ series for hourly simulated time series of one year using Monte Carlo method. It can be seen that the log-normal and log-Poisson models, especially the latter, relatively agree well with the original series in the whole, but there are also some deficiencies. For example, the heavy-tailed portrayal of the log-normal distribution model is clearly insufficient, while the slight oscillation appears over the
tail for the log-Poisson model. The \( \alpha \) and binomial models show greater oscillation than the log-Possion model over almost the whole range, but their envelope curves are approximately consistent with the frequency distribution of the empirical data. For the uniform model, most of the simulation data are close to 0, but a few extreme outliers will occur occasionally. However, due to the limitation of the assignment of the values for the parameters, the \( \beta \) model which is characterized by two peaks deviates largely from the original series.

Downscaling is as the same as an interpolation method with the common purpose of constructing new data, and to be different, the former mainly focuses on inferring high-resolution information from known low-resolution data, while the latter is more about the points in the range of a discrete set of known data points. However, generally the technique of the former is mainly based on dynamical and statistical approaches, while the latter mostly depends on the linear or nonlinear function relations built upon the known neighboring data. Next let us examine the evolution and changes of distributions of air pollutant concentrations across scales based on multifractal downscaling tools and a comparison in relation to the interpolation method, and the Shanghai PM\textsubscript{10} series is also employed. Firstly, the data with four coarser resolutions of 16 h, 8 h, 4 h and 2 h, are generated by merging adjacent sample points and averaging; then, the above six models and piecewise linear method are used to generate hourly data respectively, the estimated probability density functions using kernel function of which are shown in Fig.13. The goodness-of-fit of the fitted distributions can quantitatively show the similarities between the original
and downscaled series, namely, it can determine which type of the cascade models is the most appropriate to downscale the air pollutants time series and the common test methods include mean bias error (MBE), root mean square error (RMSE), Kolmogorov-Smirnov (K-S) and $\chi^2$ methods (Gavriil, et al., 2006). Here the root mean squared error is used to test the accuracy of the downscaled series, and the specific implementation are as follows: firstly, for each time series at the studied resolutions, a quantile of each empirical concentration is obtained from its estimated probability density function; secondly, the concentration values of the downscaled series corresponding to the quantile values in the first step are calculated based on their estimated probability density functions; then the RMSE results are calculated by using the two concentration pairs. As shown in Table 5, the piecewise linear interpolation has the best overall performance compared with the six cascade models, however, from Fig.13a and Fig.13b it has a poor fitting on the range of larger values while the six models except the uniform model can behave quite well, therefore, it can be inferred that the fittings of the range of the high concentrations determine the sizes of the errors. Of the six cascade models, the log-normal model has the highest accuracy, although the $\alpha$, $\beta$, log-Poisson and binomial models have the similar accuracy. Except the above comparison, as seen in Fig.14 and Fig.15, from $\gamma-c(\gamma)$ functions and the overlay show we also get the linear interpolation cannot reach the information of extreme values or outliers at the small scales, but these values can be obtained from the downscaling methods. Therefore, we conclude that the multifractal cascade model are ideal and effective downscaling models, and when people need the
singular information eliminated because of combining and averaging, or unable to
observe owing to experimental conditions, and it can be inferred or recovered using
the scaling invariance tools.

4. Conclusions

The universal multifractal model is a framework which is defined on the infinite
densifications and multiplications of many independent cascade processes at various
scales, which are assumed to follow the stable distribution identically, and this is also
why we selected the stable distribution as a candidate for the fittings to the air
pollution time series. The presence of the lower energy peaks at 8h, 6h and even less
is worthy studying on account of the ignorable amplitudes, however, it is hard to
explain, and it may be related with the influencing factors of air pollution such as
pollutants properties (physical or chemical) or complex interactions between natural
and anthropogenic environmental conditions. The segments of the two regimes
correspond to the low-frequency weather and weather scales with a slope $\beta$ close to 0
for larger scales and 5/3 for smaller scales respectively, and thus can provide a direct
evidence of self-similarity and nonstationarity. As mean and variance, the general
forms $\alpha$ and $C_1$ are also viewed as standard statistical quantities which are capable of
describing the whole properties of the air pollution series for providing a new
perspective. Considering the well behaviors in the descriptions of extreme
information and peak cluster, the two parameters can be used in the evaluation and
studies of air pollution quality. In view of the excellent downscaling behaviors in the
modeling of air pollution processes, a multiplicative cascade process can be recognized as a representative of the air pollution series which make it possible to further explore the temporal variations and scaling evolutions at finer scales.

**DISCLAIMER**

The authors declare that they have no competing financial interests.

**REFERENCES**


Esposito, E., De Vito, S., Salvato, M., Bright, V., Jones, R.L., and Popoola, O. (2016). Dynamic neural network architectures for on field stochastic calibration of
indicative low cost air quality sensing systems. *Sens. Actuators B* 231: 701-713.


Yahaya, A.S., Ramli, N.A., Ahmad, F., Mohd, N., Muhammad, N., and Bahrim, N.H.

Table 1: The $\alpha$ estimates of the six pollutants and AQI time series at the seven stations

<table>
<thead>
<tr>
<th></th>
<th>AQI</th>
<th>CO</th>
<th>NO$_2$</th>
<th>O$_3$</th>
<th>PM$_{10}$</th>
<th>PM$_{2.5}$</th>
<th>SO$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beijing</td>
<td>1.3899</td>
<td>1.135</td>
<td>1.7151</td>
<td>1.4955</td>
<td>1.4453</td>
<td>1.3221</td>
<td>0.4773</td>
</tr>
<tr>
<td>Shenyang</td>
<td>1.3418</td>
<td>1.6149</td>
<td>1.9687</td>
<td>1.7773</td>
<td>1.4632</td>
<td>1.3486</td>
<td>1.4092</td>
</tr>
<tr>
<td>Shanghai</td>
<td>1.4903</td>
<td>1.4285</td>
<td>1.6308</td>
<td>1.7961</td>
<td>1.5186</td>
<td>1.4619</td>
<td>1.2269</td>
</tr>
<tr>
<td>Wuhan</td>
<td>1.6495</td>
<td>1.7886</td>
<td>1.617</td>
<td>1.5231</td>
<td>1.8013</td>
<td>1.5444</td>
<td>1.197</td>
</tr>
<tr>
<td>Guangzhou</td>
<td>1.6282</td>
<td>1.701</td>
<td>1.5821</td>
<td>1.1431</td>
<td>1.5691</td>
<td>1.6162</td>
<td>1.6332</td>
</tr>
<tr>
<td>Chengdu</td>
<td>1.629</td>
<td>1.7709</td>
<td>1.9432</td>
<td>1.3248</td>
<td>1.6002</td>
<td>1.6311</td>
<td>1.3596</td>
</tr>
<tr>
<td>Xi’an</td>
<td>1.2074</td>
<td>1.3767</td>
<td>1.938</td>
<td>1.1909</td>
<td>1.5672</td>
<td>1.1342</td>
<td>1.4482</td>
</tr>
</tbody>
</table>

Table 2: The $\beta$ estimates of the six pollutants and AQI time series at the seven stations

<table>
<thead>
<tr>
<th></th>
<th>AQI</th>
<th>CO</th>
<th>NO$_2$</th>
<th>O$_3$</th>
<th>PM$_{10}$</th>
<th>PM$_{2.5}$</th>
<th>SO$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beijing</td>
<td>1</td>
<td>0.9994</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.7249</td>
</tr>
<tr>
<td>Shenyang</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Shanghai</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Wuhan</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Guangzhou</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Chengdu</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Xi’an</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3: The phase transition parameters of five cities based on API data for two kinds of length data where $l$ is the length

<table>
<thead>
<tr>
<th></th>
<th>Beijing</th>
<th>Guangzhou</th>
<th>Shanghai</th>
<th>Shenyang</th>
<th>Xi’an</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l=4$</td>
<td>$l=13$</td>
<td>$l=4$</td>
<td>$l=13$</td>
<td>$l=4$</td>
<td>$l=13$</td>
</tr>
<tr>
<td>$\Delta s$</td>
<td>0.997</td>
<td>1.175</td>
<td>0.841</td>
<td>0.948</td>
<td>1.578</td>
</tr>
<tr>
<td>$\gamma_{\max}$</td>
<td>0.562</td>
<td>0.661</td>
<td>0.424</td>
<td>0.478</td>
<td>0.826</td>
</tr>
<tr>
<td>$q_D$</td>
<td>2.607</td>
<td>2.984</td>
<td>2.346</td>
<td>2.537</td>
<td>1.732</td>
</tr>
<tr>
<td>$\gamma_D$</td>
<td>0.465</td>
<td>0.613</td>
<td>0.252</td>
<td>0.311</td>
<td>0.294</td>
</tr>
<tr>
<td>$q_s$</td>
<td>2.959</td>
<td>3.080</td>
<td>3.820</td>
<td>3.939</td>
<td>3.216</td>
</tr>
<tr>
<td>$\gamma_s$</td>
<td>0.517</td>
<td>0.624</td>
<td>0.444</td>
<td>0.515</td>
<td>0.797</td>
</tr>
<tr>
<td>$\gamma_{sd}$</td>
<td>0.556</td>
<td>0.649</td>
<td>0.448</td>
<td>0.491</td>
<td>0.979</td>
</tr>
<tr>
<td>$c(\gamma_{sd})$</td>
<td>0.952</td>
<td>1.130</td>
<td>0.806</td>
<td>0.908</td>
<td>1.542</td>
</tr>
</tbody>
</table>
Table 4: The theoretical expressions of $K(q)$ for six cascade downscaling models

<table>
<thead>
<tr>
<th>Models</th>
<th>Probability density functions</th>
<th>Theoretical expressions of $K(q)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binomial</td>
<td>$0 &lt; q &lt; p &lt; 1$, $p + q = 1$, $P(W = 2p) = 0.5$</td>
<td>$K(q) = \log_2 (\rho^q + q^q) + (q - 1)$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$0 &lt; p &lt; 1$, $P(W = p^{-1}) = p$, $P(W = 0) = 1 - p$</td>
<td>$K(q) = -(q - 1) \log_2 (2p) + (q - 1)$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$P(W = a_1) = p$, $P(W = a_2) = 1 - p$, $0 \leq a_1 &lt; 1 &lt; a_2$</td>
<td>$K(q) = \log_2 (a_1^q + a_2^q (1 - p))$</td>
</tr>
<tr>
<td>Log-normal</td>
<td>$W = \exp (\sigma X - \sigma^2/2)$, where $X$ is a standard normal distribution variable</td>
<td>$K(q) = \frac{\sigma^2}{2 \log 2} q^2 - \left( \frac{\sigma^2}{2 \log 2} + 1 \right) q + q$</td>
</tr>
<tr>
<td>Uniform</td>
<td>$W$ is a uniform distribution variable between [0, 2]</td>
<td>$K(q) = \log_2 \left( \frac{2^q}{q + 1} \right)$</td>
</tr>
<tr>
<td>Log-Poisson</td>
<td>$W = ApY$, $P(Y = m) = c^m \exp(-c)/m!$, $A = \exp(c(1-\beta))$</td>
<td>$K(q) = c \frac{q(1-\beta) + \beta^n - 1}{\log 2}$</td>
</tr>
</tbody>
</table>

Table 5: Root mean square errors of the four kinds of downscaled series using six models and piecewise linear interpolation method for the Shanghai PM$_{10}$ series (unit: $\mu g/m^3$).

<table>
<thead>
<tr>
<th>Models</th>
<th>16 h</th>
<th>8 h</th>
<th>4 h</th>
<th>2 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binomial</td>
<td>4.2343</td>
<td>7.0591</td>
<td>9.9952</td>
<td>11.2413</td>
</tr>
<tr>
<td>$\beta$</td>
<td>4.2327</td>
<td>7.8274</td>
<td>10.4920</td>
<td>13.7244</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>4.5117</td>
<td>8.0639</td>
<td>7.9205</td>
<td>12.2475</td>
</tr>
<tr>
<td>Log-normal</td>
<td>4.1194</td>
<td>6.9743</td>
<td>7.0969</td>
<td>9.2089</td>
</tr>
<tr>
<td>Log-Poisson</td>
<td>4.8051</td>
<td>8.7424</td>
<td>10.3885</td>
<td>12.7417</td>
</tr>
<tr>
<td>Piecewise linear</td>
<td>0.9046</td>
<td>2.6420</td>
<td>4.6451</td>
<td>8.6796</td>
</tr>
<tr>
<td>interpolation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure Captions

Fig. 1. The distribution of the cities in China that are used for the study.

Fig. 2. The cascade generation process.

Fig. 3. The fitting results of the probability density function of the PM$_{2.5}$ series observed in Beijing using log-normal distribution, normal distribution, kernel function and Levy distribution.

Fig. 4. The spectral density function of the PM$_{2.5}$ series observed in Xi’an versus wave numbers.

Fig. 5. The spectral density function of the O$_3$ series observed in Beijing versus wave numbers.

Fig. 6. The spectral density function of the PM$_{2.5}$ series in Beijing, averaged over logarithmically spaced bins, 10 per order of magnitude.

Fig. 7. The relation of the logarithm moments of the orders of $q$ with the values from 0.1 to 3 with the step of 0.1 versus the logarithm scales for the Beijing PM$_{2.5}$ series.

Fig. 8. The comparison of cascade simulations with different values of $\alpha$ and $C_1$: (a) $\alpha=0.6$, $C_1=0.12$; (b) $\alpha=1.6$, $C_1=0.12$; (c) $\alpha=1.2$, $C_1=0.12$; (d) $\alpha=1.2$, $C_1=0.52$.

Fig. 9. The comparison of the $\alpha$ values of six pollutants and AQI series in seven cities.

Fig. 10. The comparison of the $C_1$ values of six pollutants and AQI series in seven cities.

Fig. 11. The estimated $c(\gamma)$ functions of the AQI series in the seven cities.

Fig. 12. The comparisons of fittings of PDF to the Shanghai PM$_{10}$ series by six kinds of simulations generated from different cascade models.

Fig. 13. The comparison of PDF estimated by kernel function for the downscaled series based on the Shanghai PM$_{10}$ series by using the six cascade models and piecewise linear interpolation method.
Fig. 14. The comparison of $c(\gamma)$ functions of the original series (the Shanghai PM$_{10}$ series), averaged series and downscaled series at four scales based on the log-normal random cascade model.

Fig. 15. The overlay display of the downscaled series using the log-normal random cascade model, original series (the Shanghai PM$_{10}$ series) and the interpolated series.
Fig. 1.
a. Firstly, there is a measure $\varepsilon$ with a geometrical support

b. Secondly, two subunits are generated by the first division to the measure

c. Thirdly, the second division to the measures generated in the above step

d. The division processes continue

**Fig. 2**
Fig. 3
Fig. 4
Fig. 5
Fig. 6
Fig. 7
Fig. 8
Fig. 9
Fig. 10
Fig. 11
a. The original series versus the binomial, $\alpha$ and $\beta$ models

b. The original series versus the log-normal and log-Poisson models

c. The original series versus the uniform model

Fig. 12
a. The comparison of the six cascade models and piecewise linear interpolation method

b. The magnified display of the region from 120 to 210 for concentration values

Fig. 13
Fig. 14
Fig. 15