Journal of Oceanology and Limnology
Vol. 37 No. 2, P. 448-460, 2019
**https://doi.org/10.1007/s00343-019-8006-3**

# Wave height statistical characteristic analysis*

LIU Guilin[1], CHEN Baiyu[2, **], WANG Liping[3], ZHANG Shuaifang[4], ZHANG Kuangyuan[5], LEI Xi[3]

[1] *College of Engineering, Ocean University of China, Qingdao 266100, China*

[2] *College of Engineering, University of California Berkeley, CA 94720, USA*

[3] *College of Mathematical Science, Ocean University of China, Qingdao 266100, China*

[4] *Department of Mechanical Engineering, University of Florida, Gainesville, USA*

[5] *Department of Economics, Penn State University, State College, USA*

Received Jan. 7, 2018; accepted in principle Mar. 13, 2018; accepted for publication May 6, 2018


**Abstract** When exploring the temporal and spatial change law of ocean environment, the most common method used is using smaller-scale observed data to derive the change law for a larger-scale system. For instance, using 30-year observation data to derive 100-year return period design wave height. Therefore, the study of inherent self-similarity in ocean hydrological elements becomes increasingly important to the study of multi-year return period design wave height derivation. In this paper, we introduced multifractal to analyze the statistical characteristics of wave height series data observed from oceanic hydrological station. An improvement is made to address the existing problems of the multifractal detrended fluctuation analysis (MF-DFA) method, where trend function showed a discontinuity between intervals. The improved MF-DFA method is based on signal mode decomposition, replacing piecewise polynomial fitting used in the original method. We applied the proposed method to the wave height data collected at Chaolian Island, Shandong, China, from 1963 to 1989 and was able to conclude the wave height sequence presented weak multi-fractality. This result provided strong support to the past research on the derivation of multi-year return period design wave height with observed data. Moreover, the new method proposed in this paper also provides a new perspective to explore the intrinsic characteristic of data.

**Keyword**: wave height; partition function; multifractal spectrum; multifractal detrended fluctuation analysis (MF-DFA); signal mode decomposition

## 1 INTRODUCTION

The term *fractal*, according to Mandelbrot, founder of fractal analysis, refers to "a shape made of parts similar to the whole in some way," in which "similar ... in some way" can be self-similarity, self-affinity, or statistical self-similarity, in terms of time or physical space. Fractal phenomena exist widely in nature (e.g., the shapes of coastlines, river networks, and the shapes of trees) and in physics and chemistry (such as fractal noise and soil particle distribution (Li and Zhang, 2005; Kou et al., 2015; Chen et al., 2016a, b). The nature of the fractal analysis is searching for innate law in seemingly unordered, high-dimension and uncertain data by exploring different fractal dimensions. By varying the fractal dimensions, one is enabled to model the dynamic changing law of the environment, and therefore optimize distribution model. Currently (Cai and Gouveia, 2013; Chen et al., 2016c)., the concept and idea of fractal analysis have been abstracted into a methodology, and there have been significant achievements made in data science

* Supported by the NSFC-Shandong Joint Fund "Study on the Disaster-Causing Mechanism and Disaster Prevention Countermeasures of Multi-Source Storm Surges" (No. U1706226), the National Natural Science Foundation of China "Coastal Engineering and Risk Assessment Based on a Four-Layer Nested Multi-Objective Probability Model" (No. 51379195), the Natural Science Foundation of Shandong Province "Three-Layer Nested Multi-Objective Probability Prediction and Risk Assessment in Coastal Engineering" (No. ZR2013EEM034), and the Program of Promotion Plan for Postgraduates' Educational Quality "Paying More Attention to the Study on the Cultivation Mode of Mathematical Modeling for Engineering Postgraduates" (No. 861801232417).
** Corresponding author: baiyu@berkeley.edu

based on the fractal analysis (Liu et al., 2015; Cao et al., 2016a, b). Fractal analysis has become a new perspective to study complicated systems and bridges the parts and the overall system (Chen et al., 2008, 2017b; Li and Burgueño, 2010; Ponce-López et al., 2016; Chen and Wang., 2017).

Previously, the method for calculating wave heights of a certain return period was to choose a probability distribution model (Gumbel and Max-Entropy Distribution) as the probability distribution for annual extremes of wave heights (Wang et al., 2013a). The observed annual extreme wave heights were used to determine the parameters of the probability distribution, after which a cumulative rate determined the wave height that only occurs once in many years (a century or more) (Wang et al., 2016, 2017). Such a calculation method has three drawbacks: 1) it assumes that all annual extreme wave heights follow the same probability distribution while disregarding whether the time period is short term or long term (Wang et al., 2013b; Chen et al., 2017a); 2) the statistical characteristics of the data obtained over a short period of time and those obtained over a long period (a century or more) are both strictly self-similar(Wang et al., 2014); and 3) during the calculation, only annual extreme wave heights are used (only one data point each year) while most observational data are disregarded (Wang et al., 2015). All of these drawbacks lead to unconvincing calculation results. In fact, the long-term evolution of wave heights at an observation point is a very complex system, and only by conducting multifractal analysis to reveal the relationship between the short-term observational wave height sequences (parts) and the long-term observational wave height sequence (overall) can the latter be inferred from the former (Kantelhardt et al., 2002; Lu, 2004; Yang et al., 2018), thus providing a good foundation for more reasonable estimations of wave heights that only occur once in many years.

We apply multifractal analysis for the first time in the characterization of a wave height sequence observed in marine and hydrological stations, by conducting a comprehensive analysis on a part of the observed data using methods of statistical physics, from which the characteristics of the whole dataset can be calculated. Currently, two multifractal analysis methods have been proposed: the partition function method (Jiang and Zhou, 2008) and the multifractal detrended fluctuation analysis (MF-DFA) method (Hu and Song, 2003; Wang and Tian, 2006; Xu, 2011).

The detrending of fluctuations is an important step of the MF-DFA method; however, currently, there are some drawbacks associated with the calculations performed during this step: 1) the fitted polynomial function becomes discontinuous at the connection points of adjacent segments, which leads to the generation of new pseudo-fluctuation errors (Zhuang, 2012); 2) the selection of the polynomial order is highly subjective—low-order polynomials cannot accurately represent data fluctuations, whereas higher order polynomials can cause overfitting problems. These drawbacks will lead to erroneous analysis results on observed data.

In this study, we use the method of signal decomposition (Cai et al., 2011a, b, 2016; Xu and Wang, 2011; Chen and Wang, 2012; Hu et al., 2013; Ding and Fan, 2015, 2016; Escalante et al., 2016) to obtain a more reasonable fluctuation function. An improvement has been made to the detrending procedure of the MF-DFA method, which is described in Section 2. In Section 3, we use the MF-DFA method and its improved version to analyze the wave height sequence (significant wave heights) of Chaolian Island, Shandong, China, waves measured over a 27-year period. We make a comparison between the analysis results, including a detailed comparison between the fluctuation function obtained from polynomial fitting and that obtained from mode decomposition, which demonstrates the advantages of the improved method.

## 2 IMPROVEMENT ON THE CONVENTIONAL MF-DFA METHOD

The MF-DFA method is effective in verifying the multifractality of a non-stationary time series, which describes the multifractality of the object using the generalized Hurst exponent $H(q)$. A key step of the MF-DFA method is to calculate the trend function in each segment, determining the fitting polynomial function using the least-squares method. The degree of the polynomial determines the type of detrended fluctuation, which leads to several problems: 1) the fitted polynomial function is discontinuous at connection points between adjacent segments, which leads to the generation of new pseudo-fluctuation errors (Koag et al., 2014; Wei et al., 2017; Zhang et al., 2017b, 2018; Barrs and Chen, 2018; Bhimani et al., 2018); 2) the fitted polynomial can be linear, quadratic, or of higher orders (Liu et al., 2015; Zhang and Kleit, 2016; Xu et al., 2017a, b; Zhang et al., 2017a; Zhe et al., 2017). Lower order polynomials

cannot accurately represent the trend in data fluctuation, whereas higher order polynomials can cause overfitting problems. Selection of the specific type depends on the specific issues being studied, and this is a highly subjective procedure.

Based on the previous problems, this paper proposes using mode decomposition of signals to solve the trend function, which can better meet the requirements of solving fitted trends while avoiding the disadvantages of the conventional MF-DFA method.

### 2.1 Definition of multifractal and an improved MF-DFA method

Multifractal can be described by generalized fractal dimension or the spectral function; the definition is as following:

Definition 1: Suppose the function sequence of time $\{X(t)\}$ is a stable stochastic process, if it satisfies:

$$E[|X(t+\Delta t)|^q]=c(q)(\Delta t)^{\tau(q)+1}. \quad (1)$$

We say that it has multifractal, of which $t\in T$, $0\in T$; $q\in Q$, $[0,1]\in Q$; $T$, $Q$ is a real interval with positive length; $\tau(q)$, $c(q)$ is a function on $Q$, $\Delta t$ is the increment in time, $\tau(q)$ is the multifractal mass index.

The two key parameters of multifractal are the generalized Hurst index and the mass index respectively.

Definition 2: Generalized Hurst index $H$ is defined as:

$$\{E[|X(t+\Delta t)-X(t)|^q]\}^{\frac{1}{q}}=c(q)(\Delta t)^{H(q)}. \quad (2)$$

Definition 3: Definition of multifractal spectrum $f(\alpha)$: If the fractal object is divided into $N$ intervals with $\varepsilon$ as the linearity of each interval, the probability density of the $i$-th interval shall be:

$$P_i(\varepsilon)\sim\varepsilon^{\alpha}. \quad (3)$$

Of which $\alpha$ is called singularity index. Let the number of intervals with the same $\alpha$ value be $N_\alpha(\varepsilon)$, then

$$N_\alpha(\varepsilon)\sim\varepsilon^{-f(\alpha)}, \quad (4)$$

$f(\alpha)$ is the multifractal spectrum, indicating fractal characteristics of sub-intervals with different $\alpha$ values within the fractal object.

Definition 4: Mass index is defined as:

$$\tau(q)=\frac{\ln \chi_q(\varepsilon)}{\ln \varepsilon}, \varepsilon\to 0. \quad (5)$$

Of which $\chi_q(\varepsilon)$ is the partition function, indicating the weighted summation of probabilities of all intervals:

$$\chi_q(\varepsilon)=\sum_{i=1}^{N}P_i^q(\varepsilon)=\varepsilon^{\tau(q)}. \quad (6)$$

Zhang (2006) adapted the fast-bandpass signal-filtering method proposed by Xu and Wang (personal communication; published in 2011) to the empirical mode decomposition of signals, establishing a fast filtering method for signal mode decomposition that provided good results. We use the latter method to solve the trend function signals. A detailed description of this method can be found in the references (Zhang 2006, 2007; Xu and Wang, 2011) and a brief introduction is given here.

Discrete Fourier transform is conducted on the wave height sequence $\{x_n\}$, resulting in the sequence $\{x_m\}$. The pre-determined upper and lower limits of the bandpass frequency are $\omega_{m_1}$ and $\omega_{m_2}$, respectively, both of which are converted to $m_1$ and $m_2$ according to $m=N\Delta t\omega_m/(2\pi)$. The sequence $\{W_m\}$ is generated as follows:

$$W_m=\begin{cases} X_0, m=0 \\ 2X_m, m_1\le m\le m_2, m=1, 2,\cdots, N-1. \\ 0, \text{else} \end{cases} \quad (7)$$

When an inverse Fourier transform is performed on the sequence $\{W_m\}$, the real part of the resulting sequence is the signal $\{x_n^{(k)}\}$ of the filtered bandpass.

In Eq.7, take $m_2=N/2$ and $m_1=m_2-1$, $m_2-2$, …, and calculate the filtered signal $\{x_n^{(1)}\}$ for different values of $m_1$. In addition, every calculation result is examined for being an intrinsic mode function (IMF). The procedure is repeated until an $m_1$ value is taken that makes the signal filtered through the bandpass ($\omega_{m_1}$, $\omega_{m_2}$) an IMF while the signal filtered through the bandpass ($\omega_{m_1-1}$, $\omega_{m_2}$) is no longer an IMF. In this way, the first IMF $\{x_n^{(1)}\}$ is decomposed from the wave height sequence $\{x_n\}$.

Similarly, take $m_2=m_1-1$, and repeat the previous procedure, and the second IMF $\{x_n^{(2)}\}$ can be decomposed.

The same procedure is repeated until the $k^{\text{th}}$ IMF is decomposed from the wave height sequence $\{x_n\}$, in which the last signal through the last bandpass (0, $\omega_{k-1}$) is a monotonic function $\{M_n\}$. In this way, the signal $\{x_n\}$ is decomposed into multiple IMFs and one monotonic function, as shown below:

$$\{x_n\}=\{x_n^{(1)}\}+\{x_n^{(2)}\}+...\{x_n^{(k)}\}+\{M_n\}, \quad (8)$$

where $\{x_n^{(1)}\}$, $\{x_n^{(2)}\}$, ... $\{x_n^{(k)}\}$ are all IMFs, and $\{M_n\}$

is the trend function that represents the overall trend of the signal.

### 2.2 An improved MF-DFA method

The improvement focuses on replacing the segmentation procedure of the MF-DFA method for fitting polynomial trends with mode decomposition of trends, while the other steps remain essentially the same.

For a wave height sequence $\{x_n\}$ of a given length $N$, the major steps of the improved MF-DFA method are as follows:

1. Calculate the cumulative deviations $\{y_n\}$ of the wave height sequence:

$$y(i)=\sum_{k=1}^{i}(x_k-\bar{x}),\ i=1,\ 2,\ \cdots,\ N. \tag{9}$$

2. Calculate the mode trend function of the cumulative deviations and segment this cumulative deviation sequence as well as the trend function: the fast filtering method of signal decomposition is used to calculate the mode trend $\{T(n)\}$ of the cumulative sequence $\{T(n)\}$. In order to maintain the remaining part, segmentation is conducted from opposite ends, dividing the sequences $\{y(n)\}$ and $\{T(n)\}$ into $2Ns$ segments of equal length $s$, in which $Ns=[N/s]$ (rounding up the result).

3. Calculate the $q^{\text{th}}$ order fluctuation function: first, calculate the mean of the squared residuals, i.e.,

$$F^2(s,k)=\frac{1}{s}\sum_{i=1}^{s}\{y[(k-1)s+i]-T[(k-1)s+i]\}^2,$$
$$k=1,\ 2,\ \cdots,\ 2Ns. \tag{10}$$

Therefore, the $q^{\text{th}}$ order fluctuation function is

$$\begin{cases} F_q(s)=\left\{\dfrac{1}{2Ns}\displaystyle\sum_{k=1}^{2Ns}[F^2(s,k)]^{q/2}\right\}^{1/q}, & q\neq 0 \\ F_0(s)=\exp\left\{\dfrac{1}{4Ns}\displaystyle\sum_{k=1}^{2Ns}\ln[F^2(s,k)]\right\}, & q=0 \end{cases} \tag{11}$$

4. Calculate the generalized Hurst exponent: the relationship between the fluctuation function $F_q(s)$ and the length $s$ of each segment is as follows:

$$F_q(s)\propto s^{H(q)}. \tag{12}$$

In the expression, $H(q)$ is called the generalized Hurst exponent. For each value of $q$, the corresponding $H(q)$ can be obtained from the linear fitting of the $\ln F_q(s)$ vs $\ln s$ plot.

When the value of the generalized Hurst exponent $H(q)$ is independent of $q$, the wave height sequence is a monofractal series; when the value of $H(q)$ changes with $q$, the sequence is a multifractal series.

## 3 EMPIRICAL ANALYSES

In this section, we analyze the observational data from the Chaolian Island Hydrological Station for the wave height sequence (significant wave heights) using the partition function method, the traditional MF-DFA method, and the improved MF-DFA method.

### 3.1 Data information

The data in this study are the wave height sequence (significant wave heights), measured four times a day in the Chaolian Island Hydrological Station from 1963 to 1989. Data are missing or insufficient in some years. Figure 1 shows the scatter plot of the original wave heights, illustrating that starting from 1963, there are some data anomalies in the first few years. In the subsequent discussion, the self-similarity and scale invariance of wave height through all the measured data are discussed. Based on a shape made of parts similar to the whole in some way, when extrapolating the design wave height, the single extremal data over a certain threshold will be selected to estimate the parameters of the distribution function.

To prevent interference from the abnormal data, the data of the first three years were excluded from the original data. For the sake of the simplification of discussions, this paper selects the wave height over the given threshold. If other thresholds are selected, it does not affect the self-similarity of the wave height data. As the scatter plot illustrates, the observed data in the first three years were belong to anomalies. Therefore, the abnormal data were excluded from the subsequent discussion and the data of the later 23 years were selected. In addition, since the research focuses on larger wave heights, a threshold selection was performed. A total of 13 546 effective wave height data whose threshold values were above average were used for the analysis, as shown in the scatter plot in Fig.2. As the figure illustrates, the overall wave height sequence is statistically similar to its local parts, and therefore it can be further analyzed using fractal theory.

### 3.2 The fractal characteristic analysis method

Using a mass exponent $\tau(q)$ and multifractal spectrum $f(\alpha)$, we describe the characteristics of the normalized wave height sequence in intervals of

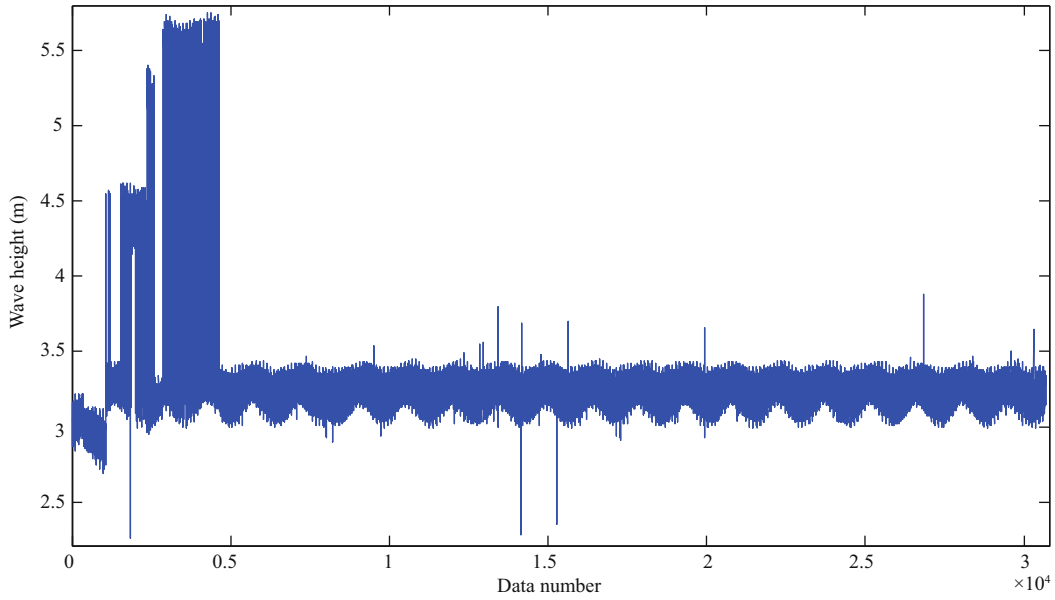


**Fig.1 Scatter plot of all original wave heights**

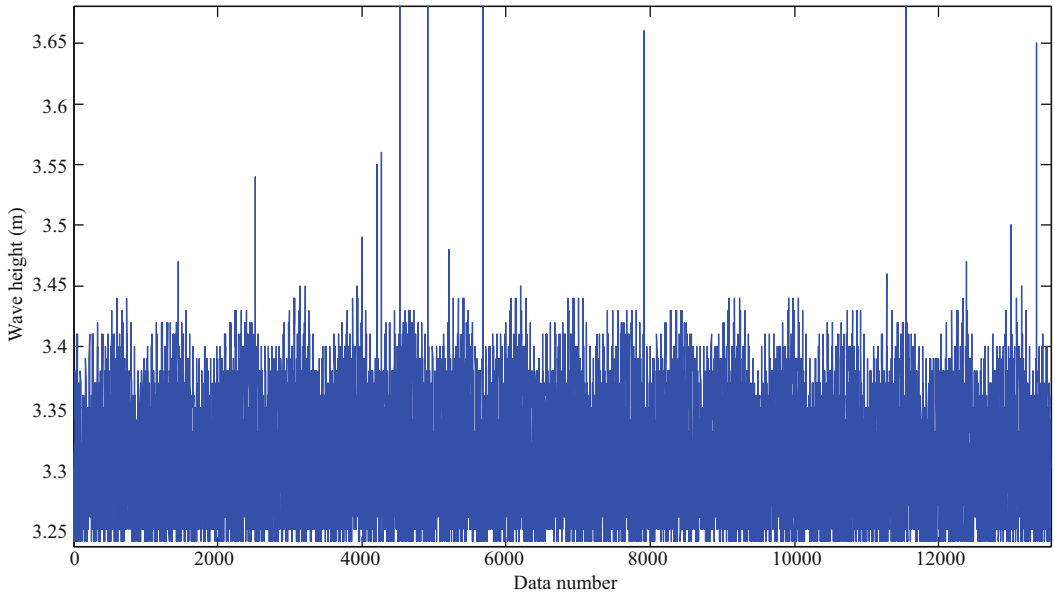


**Fig.2 Scatter plot of over-threshold wave heights**

different scales, to demonstrate the fractal properties of the sequence (Fig.3).

First, the $\ln X_q(s)$ vs $\ln(s)$ relationship was examined. Figure 3a shows the double logarithmic plots of the sequence for different values of $q$, in which a total of 21 integer values ranging from -10 to 10 are taken for $q$. Figure 3a shows only the patterns for even values of $q$. The data points produce a nearly straight line regardless of the value of $q$, suggesting that for any given $q$, the wave height sequence has fractal scaling properties. The figure also shows that for different values of $q$, slopes of the lines in which the data points lie are different from each other, suggesting a multiscale property of the sequence; i.e., the wave height sequence of Chaolian Island has a multifractal characteristic.

**Table 1 Deviation from straight line of mass exponent $\tau(q)$**

| $q$ | -10 | -5 | 0 | 5 | 10 |
|---|---|---|---|---|---|
| $\tau(q)$ | -11.510 | -6.138 | -1.004 | 3.926 | 8.712 |
| Straight line | -11.510 | -6.454 | -1.399 | 3.656 | 8.712 |
| Deviation | 0 | 0.316 | 0.395 | 0.270 | 0 |
| Total sum of squared deviations | | | 1.650 7 | | |

Table 1 shows the differences between the value of

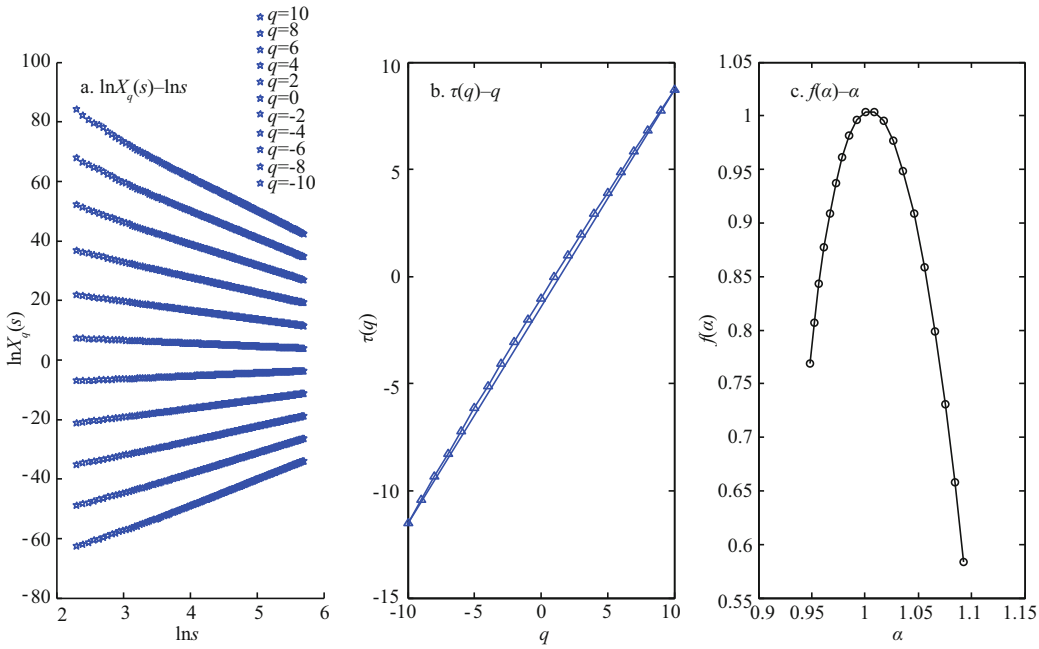


**Fig.3 Fractal characteristic analysis**

$\tau(q)$ and that of the corresponding straight line for different values of $q$. Combined to Fig.3b, it is found that the $\tau(q)$ vs $q$ curve deviates slightly from a straight line, with the total sum of squared deviations reaching 1.650 7, implying insignificant multifractality of the wave height sequence.

When Fig.3c and Table 2 are combined, the multifractal spectrum $f(\alpha)$ mainly appears as a hook-shaped quadratic curve to the left. The Holder singularity exponent $\alpha$ is featured with a scale range of (0.948, 1.093)—the entire scale range is larger than 0.5, while $\Delta\alpha$ is only 0.145, and the maximum value of $f(\alpha)$ is 1.004, suggesting a relatively narrow distribution of the multifractal spectrum curve. This result not only demonstrates the presence of long-term memory behavior in the fractal pattern of this wave height sequence but also depicts its multiscaling property implied by the variability in the scaling exponent versus volatility, both of which further indicate the relatively weak multifractal characteristic of the wave height sequence at Chaolian Island.

### 3.3 A comparison between the analytical results obtained by the conventional MF-DFA method and those by the improved MF-DFA method

A comparison is made between the conventional MF-DFA method and the improved MF-DFA method on the basis of two aspects: trend-fitting effect and final pattern.

The trend-fitting effect obtained by mode decomposition is the overall trend-fitting result for the cumulative deviation sequence. For comparison, third-order polynomials fitted to the cumulative deviation sequence were added. In the conventional MF-DFA method, trend fitting for cumulative deviations is conducted using piecewise polynomials, normally with a first-, second-, or third-order polynomial trend that depends on the volume of data. In this case, the cumulative deviation data were divided into segments, each with 50 data points, and a second-order polynomial was fitted to each segment (Fig.4).

**Table 2 Multifractality spectrum value domain**

| | Maximum | Minimum | Δ |
|---|---|---|---|
| $\alpha$ | 1.093 | 0.948 | 0.145 |
| $f(\alpha)$ | 1.004 | 0.584 | 0.420 |

As Fig.4 shows, the third-order polynomial fitting is too general compared to that obtained by mode decomposition and to that by piecewise polynomial fitting—it cannot accurately represent the local fluctuations of the cumulative deviations of wave heights, and therefore, it provides the worst fit. To further compare the results obtained by mode decomposition with those by piecewise polynomial fitting of the MF-DFA method, the region containing data numbers 4 960–5 160 in Fig.4 is enlarged and shown in Fig.5.

Figure 5 can be analyzed from two perspectives. First, within a complete segment—data points 5 001–

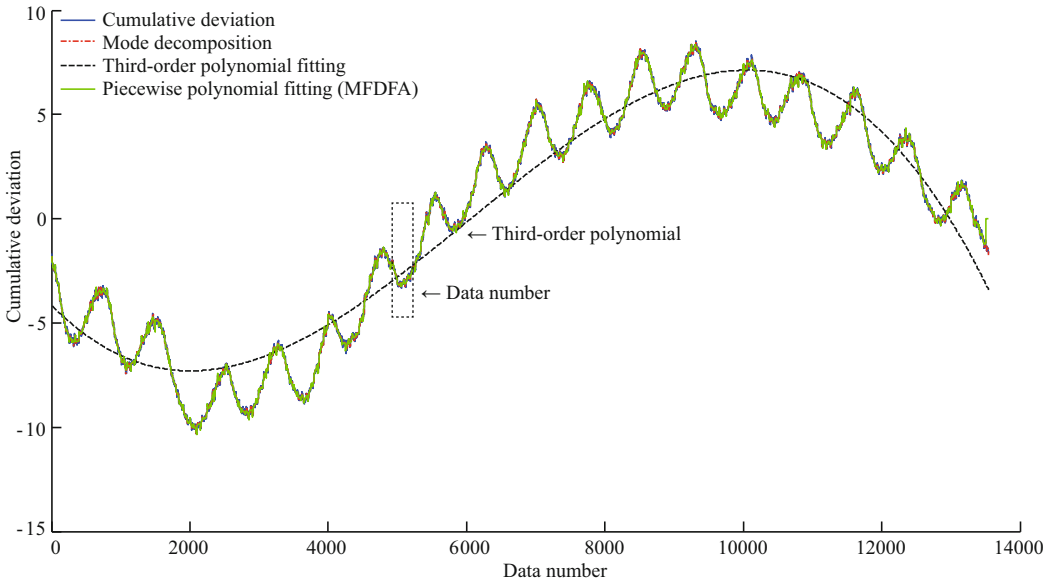


**Fig.4 Trend fitting of the cumulative deviations**

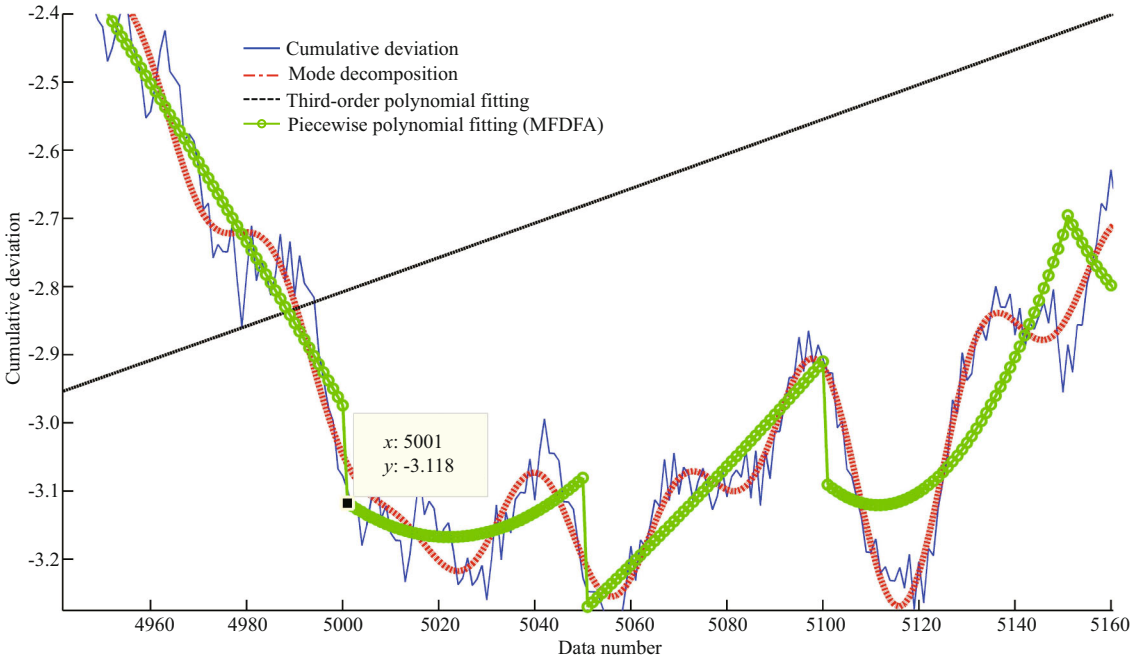


**Fig.5 Plot of the local trends**

5 050, for example—the trend obtained by mode decomposition is smoother than the MF-DFA piecewise polynomial trend, and it works better to depict the projections and declines of the fluctuations, implying a better fitting effect; second, at the connection points (i.e., data points 5 001 and 5 050), there is a large jump present in the curve of the MF-DFA method, which means that the fitting is discontinuous at connection points of adjacent segments, which generates new pseudo-fluctuation errors that interfere with the subsequent calculations. However, such a discontinuity does not occur in the mode decomposition method, and therefore no pseudo-fluctuation error is generated.

Next, we examine the detrending results for the wave height cumulative deviations (Fig.6). Figure 6a shows the normalized cumulative deviations, and Figure 6b shows the cumulative deviations after the overall third-order polynomial function is detrended. It reveals that the goal of detrending is not achieved. Figure 6c and d shows the cumulative deviation results obtained by detrending the piecewise

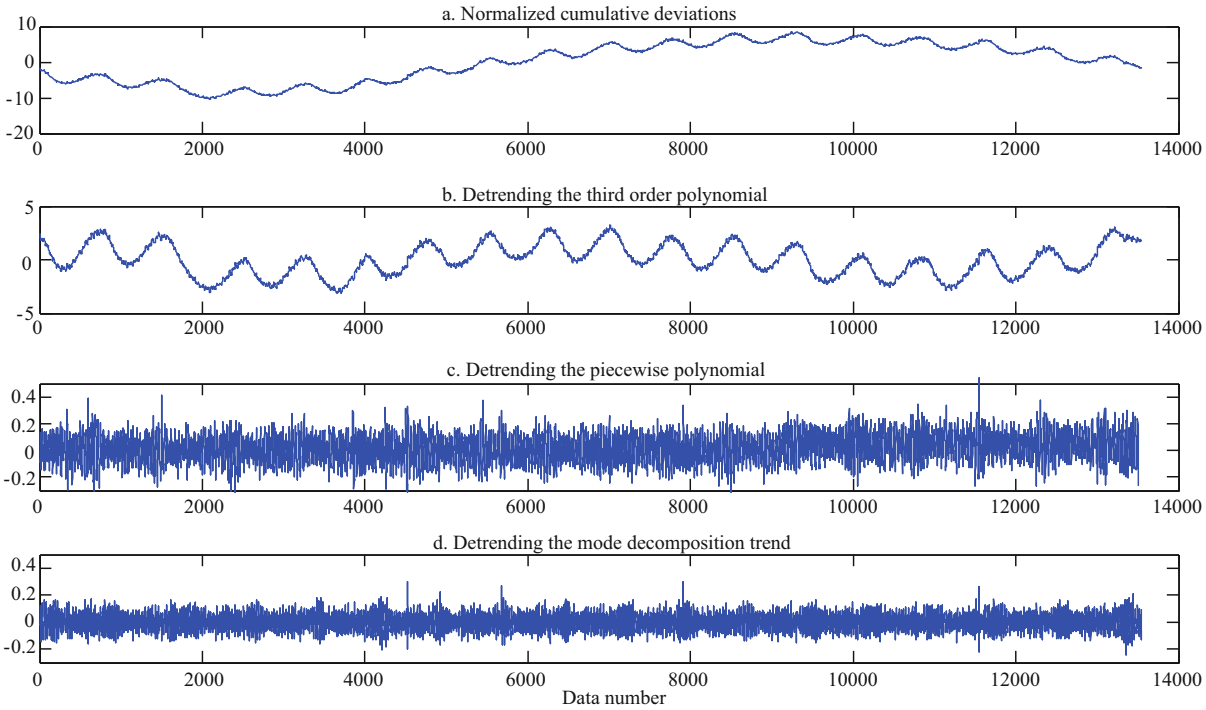


**Fig.6 Plot of detrended cumulative deviations**

**Table 3 Comparison of fluctuation of detrended cumulative deviations**

| Cumulative deviations | Mean | Variance |
|---|---|---|
| Detrended piecewise polynomial | -0.004 8 | 0.015 4 |
| Detrended mode decomposition | 2.505 5e-17 | 0.003 0 |

polynomial as well as mode decomposition, respectively, both of which reach the goal of detrending. In addition, as the figure illustrates, because the trend obtained by mode decomposition implies a better fitting effect at the connection points , when the long-term measured data and the short-term measured data are nested each other, the advantage of the fitting, based on the signal mode decomposition method, will more evident with the increase of short-term nested data nested. The means and variances of the cumulative deviations in the latter two groups are calculated for further analysis (Table 3).

As shown in Table 3, both means are close to zero. Since both methods need to solve the fluctuation function after detrending the cumulative deviations, the mean value does not have a significant effect on the succeeding calculations; the major impact is the volatility of the cumulative deviations. Variance reflects data dispersion, i.e., the magnitude of volatility. The variance of the first method is five times that of the latter, showing that the dispersion of the detrended cumulative deviations obtained by mode decomposition is smaller, with smaller fluctuations, and therefore that this trend tends to stabilize more easily.

In the previous discussions, the advantages and disadvantages of the MF-DFA method and its improved version in terms of fitting effect have been discussed; it is demonstrated that the latter can achieve a better fitting effect while avoiding some drawbacks of the original method.

In the following discussion, we further analyze the advantages and disadvantages of these two methods by comparing the final graphs. For both methods, a total of 21 integer values from -10 to 10 were set for the parameter $q$.

Figures 7 and 8, respectively, show the $\ln F_q(s)$ vs $\ln s$ double logarithmic plots obtained with the MF-DFA method and with the improved MF-DFA method. Figures 7a and 8a show the $\ln F_q(s)$ vs $\ln s$ double logarithmic plots, and Figs.7b and 8b show the corresponding linear curve fitting results. When comparing the two figures, we see that the double logarithmic plots obtained with both methods are of good linearity, indicating that the wave height sequence is long-range power-law correlated and thus that it has a multifractal characteristic, which is consistent with the results obtained by the partition function method. In addition, both of the wave height fluctuation functions obtained tend to converge at some level, the difference being that the former tends

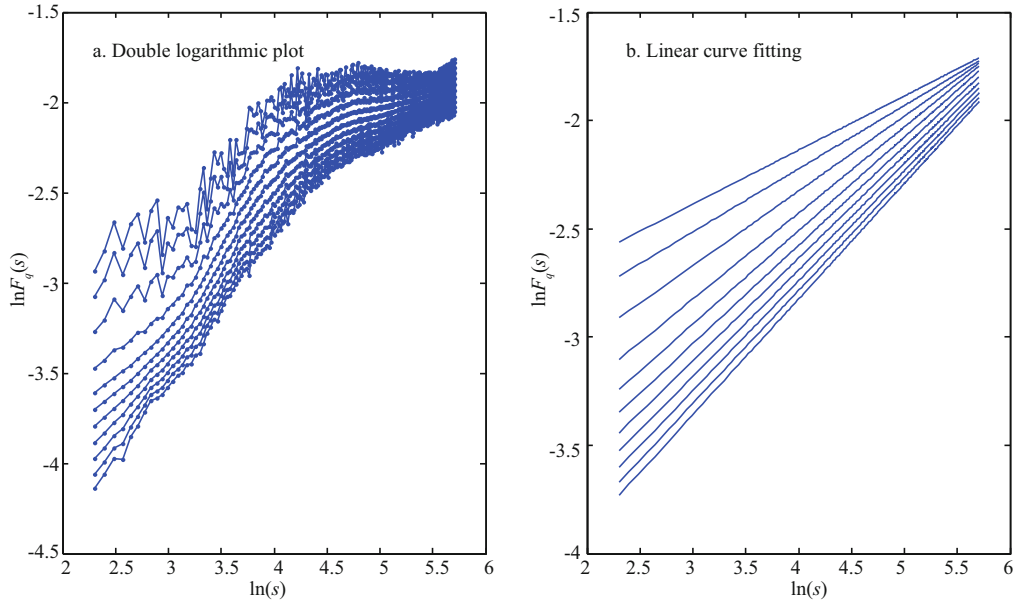


**Fig.7 MF-DFA double logarithmic plots**

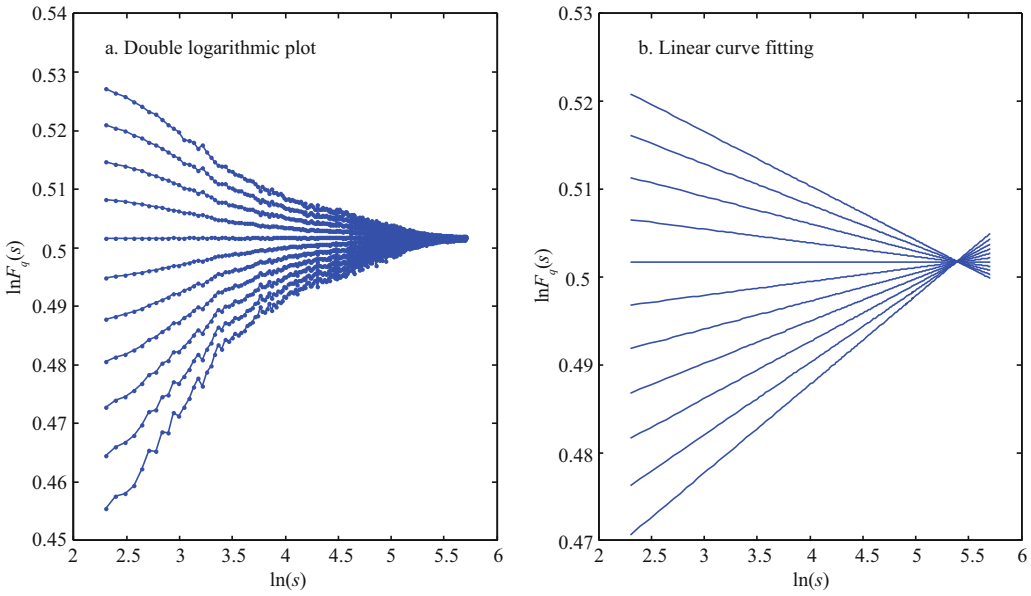


**Fig.8 Double logarithmic plots of the improved MF-DFA**

to converge slowly whereas the latter has a faster convergence speed; its fluctuation becomes smaller and smaller and therefore tends to stabilize more easily. This is inconsistent with the results shown in Table 3, which further corroborates the good performance of the improved method.

Figures 9 and 10 show the curves of the generalized Hurst exponent $H(q)$ calculated using the two methods. When comparing the two figures, we see that the generalized Hurst exponent $H(q)$ decreases as the partition order $q$ increases. According to the theory, this suggests that the value of $H(q)$ actually changes as $q$ changes and that the wave height sequence has a multifractal characteristic.

Meanwhile, it is roughly shown that the $H(q)$ curve of the improved method has a more marked inflection point compared with the original method—the inflection point occurs at $q$=-3 and decays much faster.

As Table 4 shows, for the same set of $q$ values, the fluctuation $\Delta H(q)$ of $H(q)$ calculated with the original method is only 0.284 3, while that of the improved method reaches 0.694 3, which is almost 2.5 times that of the original method. This further confirms that the improved MF-DFA method has a faster decay rate as well as better stability.

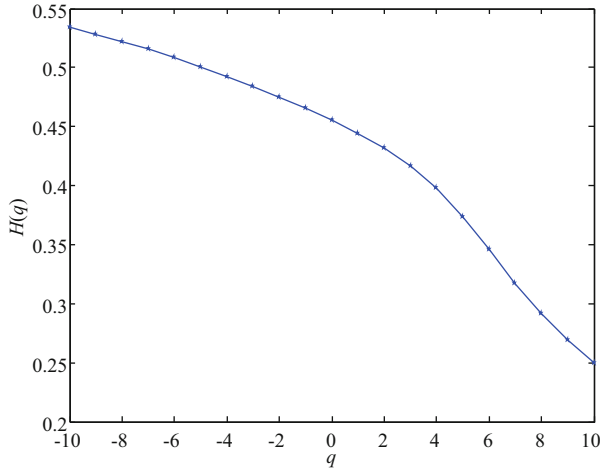


**Fig.9 Generalized Hurst exponent of the MF-DFA method**

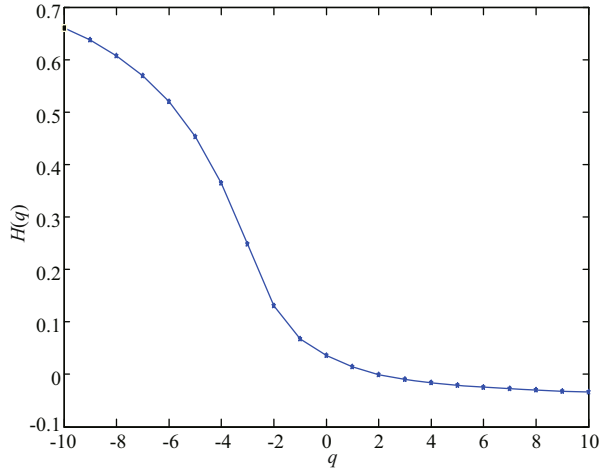


**Fig.10 Generalized Hurst exponent of the improved MF-DFA method**

**Table 4 Comparison between MF-DFA and its improved generalized Hurst exponent**

| $q$ | $H(q)$ of the conventional MF-DFA | $H(q)$ of the improved MF-DFA |
|---|---|---|
| -10 | 0.534 0 | 0.660 6 |
| -9 | 0.528 2 | 0.637 0 |
| -8 | 0.521 9 | 0.607 3 |
| -7 | 0.515 2 | 0.569 2 |
| -6 | 0.508 1 | 0.519 4 |
| -5 | 0.500 5 | 0.453 2 |
| -4 | 0.492 5 | 0.364 2 |
| -3 | 0.484 0 | 0.248 0 |
| -2 | 0.475 0 | 0.130 6 |
| -1 | 0.465 5 | 0.068 1 |
| 0 | 0.455 3 | 0.035 9 |
| 1 | 0.444 3 | 0.014 3 |
| 2 | 0.431 8 | -0.000 2 |
| 3 | 0.416 8 | -0.009 8 |
| 4 | 0.397 6 | -0.016 3 |
| 5 | 0.373 2 | -0.021 0 |
| 6 | 0.345 4 | -0.024 5 |
| 7 | 0.317 4 | -0.027 4 |
| 8 | 0.291 7 | -0.029 8 |
| 9 | 0.269 1 | -0.031 8 |
| 10 | 0.249 7 | -0.033 7 |
| $\Delta H(q)$ | 0.284 3 | 0.694 3 |

## 4 CONCLUSION

The multifractal spectrum $f(\alpha)$ of Chaolian Island appears as a hook-shaped quadratic curve to the left and its Holder singularity exponents $\alpha$ obtained are all larger than 0.5. However, the entire scale range is relatively small: $\Delta\alpha$ is only 0.145, and the maximum value of $f(\alpha)$ is 1.004. The multifractal spectrum has a relatively narrow distribution, showing weak multifractality in the wave fluctuations of Chaolian Island.

In this paper, we have applied fractal theory to the analysis of the fluctuation characteristics of waves. A partition function-based multifractal analysis of the observational wave height data of Chaolian Island was conducted, and it was observed that the multifractal spectrum $f(\alpha)$ appears as a hook-shaped quadratic curve to the left. In addition, the Holder singularity exponents $\alpha$ obtained are all larger than 0.5. However, the entire scale range is relatively small: $\Delta\alpha$ is only 0.145, and the maximum value of $f(\alpha)$ is 1.004. The multifractal spectrum has a relatively narrow distribution, showing weak multifractality in the wave fluctuations of Chaolian Island.

In addition, in this paper, the signal mode decomposition method is proposed to solve the trend function against the existing problems of the piecewise polynomial method used during the detrended fluctuation analysis of the MF-DFA method. An improved MF-DFA method is established based on mode decomposition, which is applied to the analysis of the wave height observational data from Chaolian Island. A comparison with the conventional MF-DFA method shows that:

(1) Compared to the original method, the improved method achieves finer trend fitting, enabling it to better depict the projections and declines of the

fluctuations, thus achieving a better fitting effect. In addition, the use of the improved method avoids the problems of the original one wherein the fitted trend is discontinuous at connection points of adjacent intervals; thus, it does not generate a new pseudo-fluctuation error, which is beneficial to the succeeding calculations of fluctuation functions.

(2) The cumulative deviations of the wave height data have a smaller dispersion after the detrending procedure by the improved method and thus have smaller fluctuations—the variance obtained is only 1/5 that of the original method, suggesting that the improved method can better meet the requirements of detrending fluctuations.

(3) A comparison between the double logarithmic plots $\ln F_q(s)$ vs $\ln s$ obtained with the two methods shows that both methods can achieve good linearity, indicating that the wave height sequence is long-range power-law correlated and thus has a multifractal characteristic, which is consistent with the results obtained by the partition function method. However, the wave height fluctuation function calculated with the improved method has a more obvious convergence tendency, with a faster convergence speed, and thus tends to stabilize more easily.

(4) The generalized Hurst exponent $H(q)$ curve calculated with the improved method has a more marked inflection point, which occurs at $q$=-3. Under the same set of $q$ values, the fluctuation range $\Delta H(q)$ of $H(q)$ calculated with the original method is only 0.284 3, while that of the improved method reaches 0.694 3, nearly 2.5 times that of the original method, implying that the generalized Hurst exponent has a faster decay rate, thus indicating better stability of the improved method.

In this paper, we introduced fractal to explore the fluctuation characteristics of sea waves. The introduction of the method can better analyses the internal statistical characteristics of waves. By discussing the statistical self-similaritities of the multifractal and the probability distribution models known (Pearson-III, Gumbel Distribution) ,we can not only calculate the design wave height and the design water level, but also get more details about the internal statistical characteristics of waves.The research work in this paper facilitates to accurately classify the disaster grades of storm surges. By discussing the statistical characteristics of dense discrete data, it provides feasible theoretical support for extrapolating the design wave height with discrete extremum data.

## 5 DATA AVAILABILITY STATEMENT

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

**References**

Barrs A, Chen B Y. 2018. How emerging technologies could transform infrastructure. http://www.governing.com/commentary/col-hyperlane-emerging-technologies-transform-infrastructure.html. Accessed on 2018-03-06.

Bhimani J, Yang Z Y, Mi N F et al. 2018. Docker container scheduler for I/O intensive applications running on NVMe SSDs. *IEEE Transactions on Multi-Scale Computing Systems*, https://doi.org/10.1109/TMSCS.2018.2801281.

Cai W, Chan J, Garmire D. 2011a. 3-axes MEMS hall-effect sensor. *In*: Proceedings of 2011 IEEE Sensors Applications Symposium. IEEE, San Antonio, TX, USA. p.141-144.

Cai W, Gouveia L L. 2013. Modeling and simulation of Maximum power point tracker in Ptolemy. *Journal of Clean Energy Technologies*, **1**(1): 6-9.

Cai W, Li C, Luan S W. 2016. SOI RF switch for wireless sensor network. *Advances in Engineering*: *an International Journal*, **1**(2): 1-9.

Cai W, Zhou X R, Cui X L. 2011b. Optimization of a GPU implementation of multi-dimensional RF pulse design algorithm. *In*: Proceedings of the 2011 5th International Conference on Bioinformatics and Biomedical Engineering. IEEE, Wuhan, China.

Cao Y Z, Tian N N, Bahr D et al. 2016b. The influence of cellulose nanocrystals on the microstructure of cement paste. *Cement and Concrete Composites*, **74**: 164-173.

Cao Y Z, Zavattieri P, Youngblood J et al. 2016a. The relationship between cellulose nanocrystal dispersion and strength. *Construction and Building Materials*, **119**: 71-79.

Chen B Y, Escalera S, Guyon I et al. 2016a. Overcoming calibration problems in pattern labeling with pairwise ratings: application to personality traits. *In*: Hua G, Jégou H eds. European Conference on Computer Vision (ECCV 2016) Workshops. Springer, Cham. p.419-432, https://doi.org/10.1007/978-3-319-49409-8_33

Chen B Y, Liu G L, Wang L P. 2017a. Predicting joint return period under ocean extremes based on a maximum entropy compound distribution model. *International Journal of Energy and Environmental Science*, **2**(6): 117-126.

Chen B Y, Liu G L, Zhang J F. 2016b. Method for calculate design wave height the shows three elements of typhoon: CN107103173A. 2017-08-29. (in Chinese)

Chen B Y, Wang B Y. 2017. Location selection of logistics center in e-commerce network environments. *American Journal of Neural Networks and Applications*, **3**(4): 40-48, https://doi.org/10.11648/j.ajnna.20170304.11.

Chen B Y, Yang Z Y, Huang S Y et al. 2017b. Cyber-physical system enabled nearby traffic flow modelling for

autonomous vehicles. *In*: Proceedings of the 36th IEEE International Performance Computing and Communications Conference. IEEE, San Diego, California, USA. p.1-6.

Chen G D, Wang Z C. 2012. A signal decomposition theorem with Hilbert transform and its application to narrowband time series with closely spaced frequency components. *Mechanical Systems and Signal Processing*, **28**: 258-279.

Chen G D, Warren J, Evans J. 2008. Automatically generated consumer health metadata using semantic spaces. *In*: Proceedings of the Second Australasian Workshop on Health Data and Knowledge Management. ACM, Wollongong, NSW, Australia. p.9-15.

Ding M, Fan G L. 2015. Multilayer joint gait-pose manifolds for human gait motion modeling. *IEEE Transactions on Cybernetics*, **45**(11): 2 413-2 424.

Ding M, Fan G L. 2016. Articulated and generalized gaussian kernel correlation for human pose estimation. *IEEE Transactions on Image Processing*, **25**(2): 776-789.

Escalante H J, Ponce-lópez V, Wan J et al. 2016. ChaLearn joint contest on multimedia challenges beyond visual analysis: an overview. *In*: Proceedings of the 23rd International Conference on Pattern Recognition (ICPR). IEEE, Cancun, Mexico. p.67-73.

Hu S L J, Yang W L, Li H J. 2013. Signal decomposition and reconstruction using complex exponential models. *Mechanical Systems and Signal Processing*, **40**(2): 421-438.

Hu X M, Song X F. 2003. Multifractal analysis of both Shenzhen and Shanghai stock market. *Quantitative & Technical Economics*, (8): 124-127. (in Chinese with English abstract)

Jiang Z Q, Zhou W X. 2008. Multifractal analysis of Chinese stock volatilities based on the partition function approach. *Physica A*: *Statistical Mechanics and its Applications*, **387**(19-20): 4 881-4 888, https://doi.org/10.1016/j.physa.2008.04.028.

Kantelhardt J W, Zschiegner S A, Koscielny-Bunde E et al. 2002. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A*: *Statistical Mechanics and Its Applications*, **316**(1-4): 87-114, https://doi.org/10.1016/S0378-4371(02)01383-3.

Koag M-C, Kou Y, Ouzon-Shubeita H et al. 2014. Transition-state destabilization reveals how human DNA polymerase β proceeds across the chemically unstable lesion N7-methylguanine. *Nucleic Acids Research*, **42**(13): 8 755-8 766.

Kou Y, Koag M-C, Lee S. 2015. N7 methylation alters hydrogen-bonding patterns of guanine in duplex DNA. *Journal of the American Chemical Society*, **137**(44): 14 067-14 070.

Li Z, Burgueño R. 2010. Using soft computing to analyze inspection results for bridge evaluation and management. *Journal of Bridge Engineering*, **15**(4): 430-438.

Li Z, Zhang Z X. 2005. Response of Shanghai running-metro line 2 to the construction of adjacent tunnels. *Chinese Journal of Rock Mechanics and Engineering*, **24**(1): 5 125-5 129. (in Chinese with English abstract)

Liu G L, Zheng Z J, Wang L P et al. 2015. Power-type wave absorbing device and using method thereof: CN105113452B. 2015-12-02. (in Chinese)

Lu F Y. 2004. The multifractal analysis on stock market returns in China. *System Engineering-Theory & Practice*, **24**(6): 50-54. (in Chinese with English abstract)

Ponce-López V, Chen B Y, Oliu M et al. 2016. ChaLearn LAP 2016: first round challenge on first impressions—dataset and results. *In*: Hua G, Jégou H eds. Computer Vision—ECCV 2016 Workshops. Springer, Cham. Springer International Publishing, Switzerland. p.400-418. https://doi.org/10.1007/978-3-319-49409-8_32.

Shen W, Li D S, Zhang S F et al. 2017. Analysis of wave motion in one-dimensional structures through fast-Fourier-transform-based wavelet finite element method. *Journal of Sound and Vibration*, **400**(2017): 369-386.

Wang L P, Chen B Y, Chen C et al. 2016. Application of linear mean-square estimation in ocean engineering, *China Ocean Engineering*, **30**(1): 149-160.

Wang L P, Chen B Y, Zhang J F et al. 2013b. A new model for calculating the design wave height in typhoon-affected sea areas. *Nat*. *Hazards*, **67**(2): 129-143.

Wang L P, Huang G L, Chen Z S et al. 2014. Risk analysis and assessment of overtopping concerning sea dikes in the case of storm surge. *China Ocean Engineering*, **28**(4): 479-488.

Wang L P, Liu G L, Chen B Y et al. 2015. Typhoon based on the principle of maximum entropy waters affect the design wave height calculation method: CN102063564B. 2010-12-20. (in Chinese)

Wang L P, Liu G L, Chen B Y et al. 2013a. Method for calculating combined return period of ocean extreme value considering typhoon influence: CN102063527B, 2011-05-18. (in Chinese)

Wang L P, Xu X, Liu G L et al. 2017. A new method to estimate wave height of specified return period. *Chinese Journal of Oceanology and Limnology*, **35**(5): 1 002-1 009.

Wang S Y, Tian H. 2006. Empirical research on the multifractal behavior in China security market. *Journal of Beijing Institute of Technology* (*Social Sciences Edition*), **8**(2): 71-73. (in Chinese with English abstract)

Xu D L, Wang L P. 2011. Analysis of Ocean Random Data — Principles, Methods and Applications. Higher Education Press, Beijing. 231p. (in Chinese)

Xu J C, Huang L P, Yang L. 2017a. Magnetic transforms of modulus type applied at region of lower latitudes in SE China. *Journal of Applied Geophysics*, **139**: 188-194.

Xu J C, Ren Q W, Shen Z Z. 2017b. Sensitivity analysis of the influencing factors of slope stability based on LS-SVM. *Geomechanics and Engineering*, **13**(3): 447-458.

Xu J. 2011. A study on the multifractal characteristics of the stock market in China and their relationship with risks. *Journal of Jiamusi University* (*Natural Science Edition*), **29**(1): 158-160. (in Chinese)

Yang Z Y, Wang Y F, Bhimani J et al. 2018. EAD: elasticity aware deduplication manager for datacenters with multi-

tier storage systems. *Cluster Computing*, https://doi.org/10.1007/s10586-018-2141-z.

Zhang K Y, Kleit A N. 2016. Mining rate optimization considering the stockpiling: A theoretical economics and real option model. *Resources Policy*, **47**: 87-94.

Zhang K Y, Olawoyin R, Nieto A et al. 2017a. Risk of commodity price, production cost and time to build in resource economics. *Environment*, *Development and Sustainability*. https://doi.org/10.1007/s10668-017-0003-0.

Zhang L Z. 2006. Fast filtering decomposing signal into intrinsic mode function and its applications in marine data analysis. Ocean University of China, Beijing, China. (in Chinese with English abstract)

Zhang L Z. 2007. Decomposing method of orthogonal intrinsic mode function. *Journal of Vibration and Shock*, **26**(5): 27-32. (in Chinese with English abstract)

Zhang S F, Shen W, Li D S et al. 2018. Nondestructive ultrasonic testing in rod structure with a novel numerical laplace based wavelet finite element method. *Latin American Journal of Solids and Structures*, **15**(7): 1-17.

Zhang X, Ding M, Fan G L. 2017b. Video-based human walking estimation using joint gait and pose manifolds. *IEEE Transactions on Circuits and Systems for Video Technology*, **27**(7): 1 540-1 554.

Zhe Z, Ou J P, Li D S et al. 2017. A thermography-based method for fatigue behavior evaluation of coupling beam damper. *Fracture and Structural Integrity*, **11**(40): 149-161.

Zhuang X Y. 2012. Empirical study of the volatility of Shanghai composite index based on R-MFDFA method. *Journal of Huaibei Normal University* (*Natural Science*), **33**(2): 21-26. (in Chinese with English abstract)