

股指波动率尺度行为演化的研究

孙 鹏^{1,2}, 翁同峰¹, 杨会杰¹

(1. 上海理工大学 管理学院, 上海 200093; 2. 上海财经大学 经济学院, 上海 200433)

摘要: 采用相关依赖的平衡扩散熵方法(CBEDE)研究中国股票市场上 4 个重要的股指从 2014 年 1 月 9 日到 2017 年 1 月 5 日的波动率序列的尺度行为及其演化过程。研究发现, 所有 10 个交易日长度的片段都表现出了尺度不变性的特征, 分形行为的演化表现出了丰富的特征。与中国股票市场中的重大事件发生相对应的时间中, 尺度行为具有不同的模式, 或下降随后上升, 或上升, 或下降, 这些特征可用来定量描述金融事件以及金融事件之间的区别。

关键词: 尺度行为演化; 股指波动率; 尺度指数; 扩散熵的相关依赖平衡估计

中图分类号: F 832.5 文献标志码: A

Evolution of Scaling Behavior of Stock Index Volatility

SUN Peng^{1,2}, WENG Tongfeng¹, YANG Huijie¹

(1. Business School, University of Shanghai for Science and Technology, Shanghai 200093, China;

2. School of Economics, Shanghai University of Finance and Economics, Shanghai, 200433, China)

Abstract: The correlation-corrected balanced estimation of diffusion entropy (CBEDE) was used to evaluate the local scale behavior of stock index volatility series and its evolution behavior. The volatility series of four important stock indices in China's stock market from January 9th 2014 to January 5th 2017 were investigated. It is found that all the segments with a 10 trading day length show a scale-invariant behavior. On different time scales, the evolution of scaling behavior shows rich characteristics. In particular, in the duration of important events in China's stock market, the scale behavior has different patterns, or falls then rises, or rises, or falls. These characteristics can be used to quantitatively describe financial events and the differences between them.

Keywords: evolution of scaling behavior; stock index volatility; scale index; correlation-corrected balanced estimation of diffusion entropy

有效市场假说作为现代金融学的基石, 极大地推动了金融理论的发展, 并对金融实践有着深刻的指导意义。人们从海量的金融市场数据中发

现, 金融市场不是完全随机的, 而是具有长程关联(分形)结构, 从而提出了分形市场理论。时间序列的分形结构表明, 复杂系统在不同的时间尺

收稿日期: 2018-01-14

基金项目: 国家自然科学基金资助项目(10975099)

第一作者: 孙 鹏(1991-), 男, 博士研究生. 研究方向: 数量经济学. E-mail: sunpeng4007@163.com

通信作者: 杨会杰(1968-), 男, 教授. 研究方向: 金融系统演化. E-mail: hjyang@usst.edu.cn

度上满足相同的规律，体现为某些统计量的尺度不变性。

1 研究现状和存在的问题

Mantegna 等^[1]首次将尺度不变性方法引入经济系统分析中。一个股指的价格序列 $P = \{P_i, i = 1, 2, \dots, N\}$ ，其波动率序列可定义为

$$X(s) = \{x_m \equiv |\log P_i - \log P_{i-s}|, i = s + 1, s + 2, \dots, N\} \tag{1}$$

其中， x_m 表示的是与 s 时间单位之前的股指价格相比，现在的股指价格波动率。比如，对于 5 min 的股指价格序列，如果选择 $s = 48$ ，就会产生相对于一个交易日之前的股指价格的波动率序列。如果 $X(s)$ 是平稳序列，并且长度 N 趋近于无穷，那么，它的统计特性可由概率密度函数来准确描述。如果序列 $X(s)$ 的概率密度函数 $P(x, s)$ 满足如下形式：

$$P(x, s) \sim \frac{1}{s^\delta} F\left(\frac{x}{s^\delta}\right) \tag{2}$$

则称该序列满足尺度不变性，即在尺度 $x \rightarrow \frac{x}{s^\delta}$ 变换操作下，概率密度函数保持不变。参数 δ 为尺度指数^[2]。

在有效市场中，概率密度函数 $P(x, s)$ 满足高斯分布，即尺度指数 $\delta = 0.5$ 。然而，在实际中通常观察到 $\delta \neq 0.5$ 。如果 $\delta > 0.5$ ，则表明时间序列存在长期记忆性，即当上一个时刻的波动率是上升(下降)的，则下一个时刻上升(下降)的可能性比较大。如果 $\delta < 0.5$ ，则表明时间序列存在反持续性，即波动率高频率地震荡。从而可以利用尺度指数 δ 对股票市场进行预测。问题的关键是如何估计这个尺度指数。

作为提取尺度自相似结构的标准方法，基于方差的分析如小波变换模极大值(wavelet transformation modulus maxima, WTMM)^[3-6]和趋势消除波动分析(detrended fluctuation analysis, DFA)^[7-10]，在文献中被广泛使用。而方差分析法具有以下 2 个局限性：

a. 这种方法是依赖于动力学机制的。用 H 来表示采用方差分析法得到的尺度指数的估计值。对于一个分数布朗运动过程，可以观察到 $H = \delta$ 的关系，也就是说尺度不变性可以被正确地探测到并被估计出来。然而对于一个列维过程，存在关系 $\delta = \frac{1}{3-2H}$ ，也就是说尺度不变性可以被定性地探

测出来，但是，定量估计不正确。对于列维过程，随着尺度增加，理论上的方差序列发散，用实际数据估计出来的不满足幂律关系。

b. 这种方法一般都需要很长的时间序列。大量的研究表明，利用 DFA 和 WTMM 去估计尺度指数的时候，时间序列的长度至少要分别达到 2^{13} ^[11]和 2^{14} ^[12]才能得到 δ 的可靠估计。至于重标涨落分析法(R/S)，时间序列长度达到 2^{16} ^[13]，也不能保证能够得到满意的结果。

2002 年，Scafetta 等^[14]提出了扩散熵(diffusion entropy, DE)的概念以解决依赖于动力学机制的问题。这一方法利用概率分布函数计算熵，考察熵和尺度的关系，而不再考察方差与尺度的关系。简单的推导可以得到，熵与尺度的自然对数满足线性关系，其斜率就是尺度指数的无偏估计。无论对于哪种动力学机制产生的时间序列，扩散熵方法都是行之有效的。

然而，大量的研究表明^[15-22]，利用扩散熵方法估计尺度指数需要比较长的时间序列。为了解决短序列的估计问题，Bonachela 等^[23]提出了熵的平衡估计量方法。这种方法的主要思想是使统计误差(fluctuation)及系统误差(bias)同时最小。然而，在推导过程中，明显的限制条件 $\int_{-\infty}^{+\infty} P(x, s) dx = 1$ 通常被忽略。通过添加这个限制条件，Pan 等将熵的平衡估计量方法改进成了新的版本，即相关依赖的平衡扩散熵(CBEDE)^[24-26]。大量的模拟实验表明，CBEDE 方法可以很好地估计短序列的尺度不变性^[25](序列长度只有几百)。本文采用这一方法探究股指波动率序列的尺度不变性行为的演化过程。将中国股市上 4 个重要的股指从 2014 年 1 月 9 日到 2017 年 1 月 5 日的波动率序列分为 10 个交易日为一个小段，一个交易日为一个步长。并且利用 CBEDE 方法估计每个小段的尺度指数，作为相应时间的局部尺度行为的衡量。研究发现，在一定的时间范围内，股指波动率序列表现出尺度不变性特征。有趣的是，与中国股票市场中的重大事件发生相对应的平滑后的尺度指数曲线部分具有十分显著的特征，如或先下降随后上升，或上升，或下降。

2 研究方法

如果股指波动率序列满足尺度不变性，此时

该序列的扩散熵 E 满足

$$E = - \int_{-\infty}^{+\infty} P(x, s) \ln P(x, s) dx = A + \delta \ln s \quad (3)$$

由式(3)可以看出, 扩散熵与窗口 s 的自然对数呈现线性关系, 斜率为不依赖动力学机制的尺度指数的估计。利用扩散熵方法会产生 2 个问题, 即如何从短序列得到扩散熵的可靠估计, 以及如何保证序列的平稳性。

利用滑动平均法使波动率序列 $X(s)$ 平稳^[25], 即让长度为 s 的窗口沿着序列 $X(s)$ 滑动。将窗口覆盖的元素平均值 $v_{m+[\frac{s+1}{2}]} = \frac{1}{s} \sum_{j=m}^{m+s-1} x_j, m = 1, 2, \dots, N-2s+1$, 作为窗口中间元素 $x_{m+[\frac{s+1}{2}]}, m = 1, 2, \dots, N-2s+1$ 的趋势。从最初的序列中去除趋势, 就得到了去趋势的序列

$$Z(s) = \{z_m \equiv x_{m+[\frac{s+1}{2}]} - v_{m+[\frac{s+1}{2}]}, m = 1, 2, \dots, N-2s+1\} \quad (4)$$

利用 CBEDE 方法得到扩散熵的可靠估计。首先将 $Z(s) = \{z_m, m = 1, 2, \dots, N-2s+1\}$ 的分布范围分成 $M(s)$ 段, 然后计算每段内元素出现的个数, 用向量 $\mathbf{n} = \{n_1, n_2, \dots, n_{M(s)}\}$ 表示。如果每段的发生概率为 $p = \{p_1, p_1, \dots, p_{M(s)}\}$, 那么, n 的概率分布函数 $Q(n, p)$ 可以用如下的多项式分布函数来描述:

$$\begin{cases} Q(n, p) = \frac{(N-2s+1)!}{n_1! n_2! \dots n_{M(s)}!} p_1^{n_1} p_2^{n_2} \dots p_{M(s)}^{n_{M(s)}} \\ \text{s.t. } p_1 + p_2 + \dots + p_n = 1 \\ n_1 + n_2 + \dots + n_{M(s)} = N-2s+1 \end{cases} \quad (5)$$

CBEDE 方法的主要思想是要平衡统计误差和系统误差, 从而使两者都被接受。如果扩散熵的估计量为 $\hat{S}$, 那么, 统计误差和系统误差为

$$\begin{cases} \Delta_{\text{fluc}}^2(n, p) \equiv \langle (\hat{S} - S)^2 \rangle \\ \Delta_{\text{bias}}^2(n, p) \equiv (\langle \hat{S} \rangle - S)^2 \end{cases} \quad (6)$$

其中, $\langle \cdot \rangle$ 是统计平均, 而且 $S \equiv \sum_{m=1}^{M(s)} -p_m \ln p_m$ 是扩散熵的确切估计。从而总体误差为

$$\begin{aligned} \Delta^2(n, s) &= \int \sum_{m=1}^{M(s)} p_m = 1 dp [\Delta_{\text{bias}}^2(n, p) + \Delta_{\text{fluc}}^2(n, p)] Q(n, p) = \\ &= \int \sum_{m=1}^{M(s)} p_m = 1 dp \left\{ \sum_{m=1}^{M(s)} n_m = N-2s+1 Q(n, p) \hat{S}^2(n) + \hat{S}^2(p) \right\} - \\ &= \int \sum_{m=1}^{M(s)} p_m = 1 dp \left[2S(p) \sum_{m=1}^{M(s)} n_m = N-2s+1 Q(n, p) \hat{S}^2(n) \right] \quad (7) \end{aligned}$$

估计量 $\hat{S}$ 使平均总体误差最小, 即

$$\frac{\partial \Delta^2(n, s)}{\partial n} = 0 \quad (8)$$

经过一系列的计算, 得到

$$\hat{S}(n, s) = \frac{1}{N-2s+1+M(s)} \cdot \sum_{m=1}^{M(s)} [n_m + 1] \sum_{k=n_m+2}^{N-2s+1+M(s)} \frac{1}{k} \quad (9)$$

将式(9)称为扩散熵的相关依赖平衡估计 (CBEDE)。

综上所述, CBEDE 算法步骤如下^[27-28]:

第 1 步 通过原始的股指价格序列 P 来计算波动率序列 $X(s)$;

第 2 步 消除波动率序列 $X(s)$ 的趋势, 进一步得到去趋势序列 $Z(s)$;

第 3 步 将去趋势序列 $Z(s)$ 分割成许多段。每段尺度选择为波动率序列 $X(48)$ 的标准差 (std) 的分形维数, $\frac{1}{\varepsilon} \text{std}[X(48)]$, 适当选取窗口长度 ε 能够保证尺度指数独立于 ε , 这里选择 $\varepsilon = 2$, 可以计算出每段中出现的 $Z(s)$ 元素的个数;

第 4 步 用式(9)去估计相关依赖的平衡扩散熵, 用 $S(s)$ 表示;

第 5 步 重复上述 4 个步骤, 选取 ε 正比于 $Z(s)$ 的标准差, 得到扩散熵的估计值 $S_0(s)$;

第 6 步 用 $s = 1, 2, \dots$ 进行迭代, 得到 $S(s) - S_0(s)$ 与 $\ln s$ 的关系曲线, 可以估计出尺度行为。

3 数 据

以中国股票市场上 4 个重要的股指的波动率序列为研究对象, 4 个股指分别为: 上证综合指数 (SSE Composite Index)、沪深 300 指数 (CSI 300 Index)、中证小盘 500 指数 (CSI SmallCap 500 Index) 和上证 50 指数 (SSE 50 Index)。时间跨度为 2014 年 1 月 9 日到 2017 年 1 月 5 日, 共 3 年。

图 1 为上证综指的价格和波动率变化图, T 为时间 (5 min), P 为价格, V 为波动率。上证综指的价格分布很宽, 其范围为 [1 991.25, 5 166.35], 它是一个具有多峰的复杂形状。表 1 展示了 2014—2016 年中国股市的 3 次重要事件。

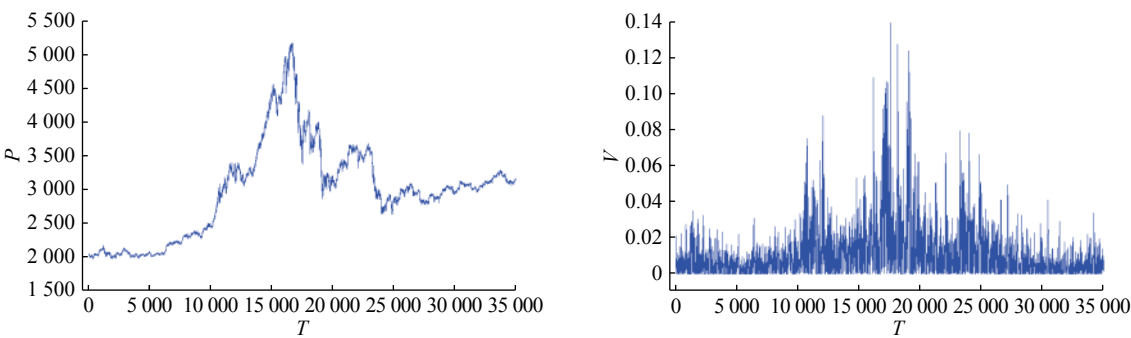


图 1 上证综指的价格和波动率变化图

Fig.1 Figure of price and volatility changes of SSE Composite Index

表 1 2014—2016 年中国股市重要事件

Tab.1 Famous events in China's stock market, 2014—2016

时间	事件名称	事件描述
2014-11-24—2015-06-12	疯牛	2014 年 11 月, 中国央行时隔两年的首次降息引燃了 A 股疯牛, 市场对央行的宽松预期成为牛市不断上涨的重要引擎。在强有力的政策以及政策暗示之下, 金融和宏观经济管理部门对股市打开了闸门, 各路“国家队”资金和机构资金纷纷进入股市, 并带动社会资金大量进入, 共同造就了牛市的疯狂繁荣。
2015-06-15—2015-07-8	股灾 1.0	仅 17 个交易日, 上证综指、创业板指最大跌幅分别达 34.9%, 42.9%, 跌幅超过 30% 和 50% 的公司分别有 2 139 家和 1 390 家, 占比为 77% 和 50%, 调整之快、幅度之大均属历史少见。
2015-08-18—2015-08-26	股灾 2.0	上涨综指由 4 006 点下跌至 2 850 点, 跌幅 29%, 创业板指由 2 721 点下跌至 1 843 点, 跌幅达 32%。

4 结 果

让覆盖 10 个交易日的窗口沿着波动率序列滑动, 滑动步长为一个交易日。然后计算每个片段的 $S(s)$ 和 $S_0(s)$, 可以得到 $S(s)-S_0(s)$ 与 $\ln s$ 的关系曲线, 图 2 为上证综指的 8 个典型片段 (编号为 seg01, seg02, ..., seg08) 的结果。研究发现, 每一条曲线在大的尺度范围内都是直线, 相应的片段表现出了尺度不变性的特征, 即中国股票市场不是

有效市场。

图 3 为上证综合指数、沪深 300 指数、中证小盘 500 指数和上证 50 指数 4 只股指的尺度行为演化过程, T 为时间 (5 min), δ 为尺度指数, 灰色圆圈为股指的尺度指数, 蓝色曲线为傅里叶平滑后的曲线。每一个十个交易日片段的尺度指数在它的第一天被表示。例如, τ 日的尺度指数对应的是开始于 τ 日的十个交易日片段。可以发现, 尺度指数在 0.44~1.02 很宽的范围 内变化。总体来说, δ 的大部分值显著大于整个序列的尺度值。随着时间的推移, δ 出现了一些高峰和低谷, 即各个股指的演化过程在不同尺度上表现出了丰富的特征。

如图 3 所示, 很有意思的是找到了表 2 所展示的中国股市的重要事件和股指尺度行为之间的可能的关系。为了显示事件的影响, 本文考虑了以事件发生的当天为中心的 10 个交易日长度的尺度指数曲线, 可以找到与事件对应的曲线部分具有十分明显的特征。比如, 在疯牛阶段, 上证综指、沪深 300 指数、中证 500 指数和上证 50 指数的尺度指数曲线都是先下降随后又上升。股灾 1.0 的时候, 4 只股指的尺度指数曲线都是上升的。而股灾 2.0 的时候, 4 只股指的尺度指数曲线都是下降的。中国股市在疯牛阶段之前, 长期处于低迷中, 一直徘徊在 2 000 点左右, 2014 年 11 月 21 日晚间的降息政策让市场为之一振, 使投

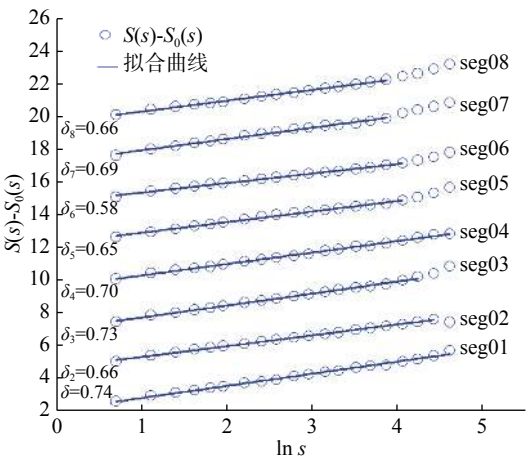


图 2 上证综指的 8 个典型片段的 $S(s)-S_0(s)$ 与 $\ln s$ 的关系

Fig.2 Relations of $S(s)-S_0(s)$ versus $\ln s$ in eight typical segments of SSE Composite Index

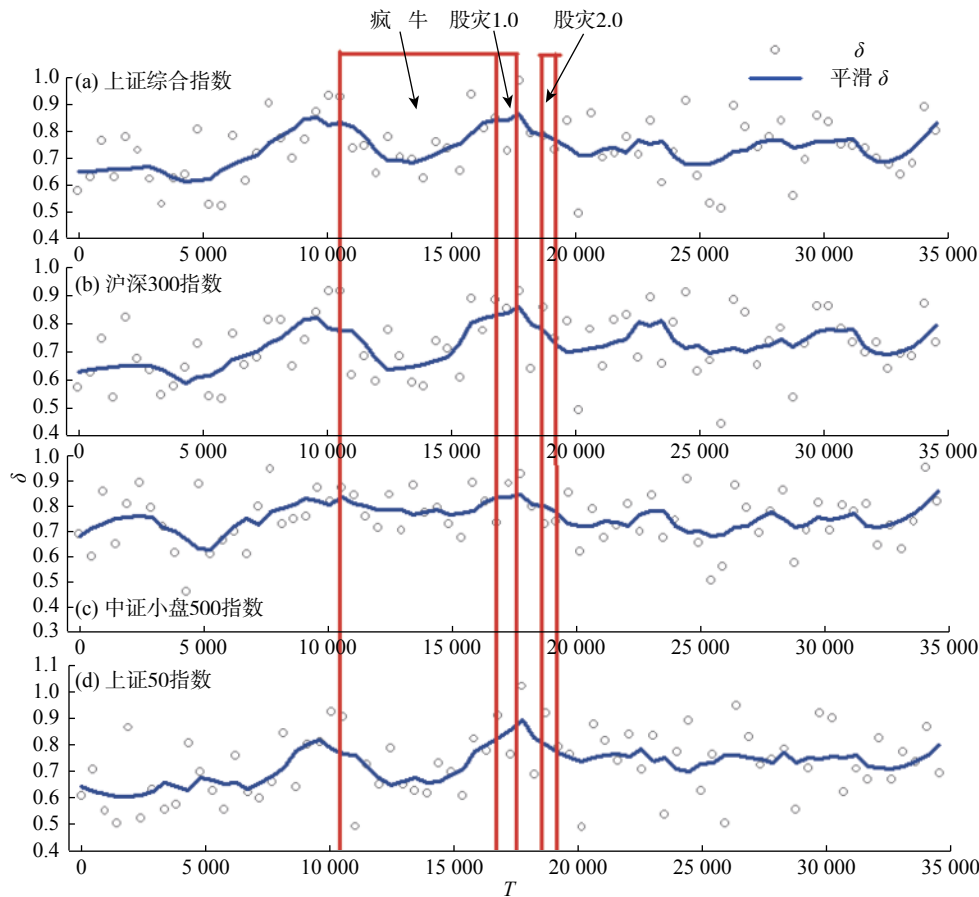


图 3 4 只股指的尺度行为的演化

Fig.3 Evolution of the scale behaviors of four stock indexes

资者相信股市即将上涨，所以， δ 处于高位，中国股市的预测性很强。但是，有些投资者认为基本面无法支撑牛市，因为，GDP 增速仍在下滑，领先经济指标如采购经理指数和贸易进出口等数据并不令人满意，市场中出现了分歧，所以， δ 开始下降。然而，进入 2015 年，在“改革牛”、“国家牛”等口号的鼓吹下，投资者的信心暴涨，从而推动大量资金涌入，继续推高股指，不断上涨的指数又刺激更多资金跟进，“一致性”的牛市预期持续强化，所以， δ 一路上升，中国股市的预测性进一步加强。如图 3 所示，在疯牛阶段，上证综指、沪深 300 和上证 50 的尺度行为的演化曲线的凹度明显，而中证 500 的尺度行为曲线凹度不明显，本文认为，在评价中国股市的时候，应选取上证综指、沪深 300 和上证 50 作为参照物。

5 结 论

金融市场的尺度行为因其在复杂系统理论中的可能应用而备受关注。然而，所有评价时间序

列尺度行为的标准工具都是基于概率论的思想，而概率论需要大量数据的支持。随着时间的推移，尺度行为会逐渐演化。本文利用相关依赖的平衡扩散熵方法来研究中国股票市场上 4 个重要的股指的波动率序列的尺度行为及其演化过程。研究发现，股指波动率序列的每个部分都存在尺度不变性，它们在演化过程中的不同时间尺度上具有丰富的特征。如果提取与中国股票市场上重要事件相对应的股指尺度行为曲线，则可以发现，曲线具有十分显著的特征，如先下降随后上升，或上升，或下降。这些特征可用来定量描述金融事件以及金融事件之间的区别，希望对进一步研究金融系统的演化行为有一定的帮助。

参考文献：

[1] MANTEGNA R N, STANLEY H E. Scaling behaviour in the dynamics of an economic index[J]. Nature, 1995, 376(6535): 46–49.
[2] STANLEY H E, AMARAL L A N, GOPIKRISHNAN P, et al. Scale invariance and universality: organizing

- principles in complex systems[J]. *Physica A: Statistical Mechanics and its Applications*, 2000, 281(1/4): 60–68.
- [3] MUZY J F, BACRY E, ARNEODO A. Wavelets and multifractal formalism for singular signals: application to turbulence data[J]. *Physical Review Letters*, 1991, 67(25): 3515–3518.
 - [4] MUZY J F, BACRY E, ARNEODO A. Multifractal formalism for fractal signals: the structure-function approach versus the wavelet-transform modulus-maxima method[J]. *Physical Review E*, 1993, 47(2): 875–884.
 - [5] XIE W J, JIANG Z Q, GU G F, et al. Joint multifractal analysis based on the partition function approach: analytical analysis, numerical simulation and empirical application[J]. *New Journal of Physics*, 2015, 17(10): 1030120.
 - [6] JIANG Z Q, GAO X L, ZHOU W X, et al. Multifractal cross wavelet analysis[J]. *Fractals*, 2017, 25(6): 1750054.
 - [7] PENG C K, BULDYREV S V, HAVLIN S, et al. Mosaic organization of DNA nucleotides[J]. *Physical Review E*, 1994, 49(2): 1685–1689.
 - [8] KANTELHARDT J W, ZSCHIEGNER S A, KOSCIELNY-BUNDE E, et al. Multifractal detrended fluctuation analysis of nonstationary time series[J]. *Physica A: Statistical Mechanics and its Applications*, 2002, 316(1/4): 87–114.
 - [9] PODOBNIK B, STANLEY H E. Detrended cross-correlation analysis: a new method for analyzing two nonstationary time series[J]. *Physical Review Letters*, 2008, 100(8): 084102.
 - [10] QIAN X Y, GU G F, ZHOU W X. Modified detrended fluctuation analysis based on empirical mode decomposition for the characterization of anti-persistent processes[J]. *Physica A: Statistical Mechanics and its Applications*, 2011, 390(23/24): 4388–4395.
 - [11] OLIVER R, BALLESTER J L. Is there memory in solar activity?[J]. *Physical Review E*, 1998, 58(5): 5650–5654.
 - [12] EKE A, HERMÁN P, BASSINGTHWAIGHTE J, et al. Physiological time series: distinguishing fractal noises from motions[J]. *Pflügers Archiv*, 2000, 439(4): 403olINK.
 - [13] EKE A, HERMAN P, KOCSIS L, et al. Fractal characterization of complexity in temporal physiological signals[J]. *Physiological Measurement*, 2002, 23(1): 1–38.
 - [14] SCAFETTA N, GRIGOLINI P. Scaling detection in time series: diffusion entropy analysis[J]. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 2002, 66(3): 036130.
 - [15] YANG H J, ZHAO F C, ZHANG W, et al. Diffusion entropy approach to complexity for a Hodgkin-Huxley neuron[J]. *Physica A: Statistical Mechanics and its Applications*, 2005, 347: 704–710.
 - [16] SCAFETTA N, WEST B J. Multiscaling comparative analysis of time series and geophysical phenomena[J]. *West Complexity*, 2005, 10(4): 51–56.
 - [17] CAI S M, ZHOU P L, YANG H J, et al. Diffusion entropy analysis on the scaling behavior of financial markets[J]. *Physica A: Statistical Mechanics and its Applications*, 2006, 367: 337–344.
 - [18] CAI S M, ZHOU P L, YANG H J, et al. Diffusion entropy analysis on the stride interval fluctuation of human gait[J]. *Physica A: Statistical Mechanics and its Applications*, 2007, 375(2): 687–692.
 - [19] TSAI C Y, SHIEH C F. A study of the time distribution of inter-cluster earthquakes in Taiwan[J]. *Physica A: Statistical Mechanics and its Applications*, 2008, 387(22): 5561–5566.
 - [20] JIANG Z B, YANG H J, WANG J B. Complexities of human promoter sequences[J]. *Physica A: Statistical Mechanics and its Applications*, 2009, 388(7): 1299–1302.
 - [21] SCAFETTA N, WEST B J. Interpretations of climate-change data[J]. *Physics Today*, 2009, 62(11): 8–10.
 - [22] HUANG J J, SHANG P J. Multiscale multifractal diffusion entropy analysis of financial time series[J]. *Physica A: Statistical Mechanics and its Applications*, 2015, 420: 221–228.
 - [23] BONACHELA J A, HINRICHSSEN H, MUNOZ M A. Entropy estimates of small data sets[J]. *Journal of Physics A: Mathematical and Theoretical*, 2008, 41(20): 202001.
 - [24] PAN X, HOU L, STEPHEN M, et al. Long-term memories in online users' selecting activities[J]. *Physics Letters A*, 2014, 378(35): 2591–2596.
 - [25] YAN Y, GU C G, XIAO Q, et al. Evolution of scaling behaviors embedded in sentence series from *A Story of the Stone*[J]. *PLoS ONE*, 2017, 12(2): e0171776.
 - [26] ZHANG J, ZHANG W Q, YANG H J. In search of coding and non-coding regions of DNA sequences based on balanced estimation of diffusion entropy[J]. *Journal of Biological Physics*, 2016, 42(1): 99–106.
 - [27] 赵丽丽, 唐镇, 王建勇, 等. 基于复杂网络理论的时间序列分析[J]. *上海理工大学学报*, 2011, 33(1): 47–52.
 - [28] 严捷冰, 杨会杰. 基于扩散熵的平衡估计的儿童行走序列研究[J]. *上海理工大学学报*, 2014, 36(2): 147–153.