

滑动移除小波分析法在动力学结构突变检验中的应用

孙东永 张洪波 王义民

Application of moving cut data-wavelet transformation analysis in dynamic structure mutation testing

Sun Dong-Yong Zhang Hong-Bo Wang Yi-Min

引用信息 Citation: *Acta Physica Sinica*, 66, 079201 (2017) DOI: 10.7498/aps.66.079201

在线阅读 View online: <http://dx.doi.org/10.7498/aps.66.079201>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2017/V66/I7>

您可能感兴趣的其他文章

Articles you may be interested in

基于 2013 年 8 月中国中东部地区持续高温的延伸期稳定分量提取及改进研究

Stable components in extended-range forecasting during the continuous high temperature event in August 2013 over the mid-eastern China

物理学报.2015, 64(4): 049203 <http://dx.doi.org/10.7498/aps.64.049203>

基于长程相关性特征的北京气候中心气候系统模式对中国气温的模拟性能评估

Performance evaluation of Chinese air temperature simulated by Beijing Climate Center Climate System Model on the basis of the long-range correlation

物理学报.2014, 63(20): 209201 <http://dx.doi.org/10.7498/aps.63.209201>

中国东北区域盛夏雨季的客观识别

Objective identification study of midsummer rainy season in the northeast China

物理学报.2014, 63(20): 209202 <http://dx.doi.org/10.7498/aps.63.209202>

20 世纪 90 年代末东亚夏季降水年代际变化及其成因初探

Interdecadal variation of precipitation pattern and preliminary studies during the summer of late-1990s in East Asia

物理学报.2014, 63(20): 209204 <http://dx.doi.org/10.7498/aps.63.209204>

2013 年长江下游降水低频分量延伸期预报的扩展复数自回归模型

Extended complex autoregressive model of low-frequency rainfalls over the lower reaches of Yangtze river valley for extended range forecast in 2013

物理学报.2014, 63(19): 199202 <http://dx.doi.org/10.7498/aps.63.199202>

滑动移除小波分析法在动力学结构突变检验中的应用*

孙东永¹⁾ 张洪波^{1)†} 王义民²⁾

1)(长安大学环境科学与工程学院, 旱区地下水文与生态效应教育部重点实验室, 西安 710054)

2)(西安理工大学水利水电学院, 西北旱区生态水利工程国家重点实验室培育基地, 西安 710048)

(2016年11月28日收到; 2016年12月20日收到修改稿)

标度指数计算的即时性与准确性对相关时间序列的动力学结构突变分析至关重要, 然而现有方法在即时性与准确性上一直无法兼顾. 将小波分析方法与滑动移除窗口技术相融合, 提出一种新的动力学结构突变检测方法——滑动移除小波分析法. 通过选取不同的滑动移除窗口, 分别对构建的线性、非线性理想时间序列进行动力学结构突变分析, 结果表明不论是线性时间序列还是非线性时间序列, 滑动移除小波分析能够准确地检测到序列的动力学结构突变点及突变区间, 对于滑动移除窗口长度依赖性较小, 具有很强的稳定性, 而且在计算速度上明显优于滑动移除重标极差和滑动移除方差分析方法, 将在大数据处理中具有一定的优势. 同时分别对线性、非线性理想时间序列添加高斯白噪声, 结果表明滑动移除小波分析具有很强的抗噪能力, 能够准确地检测到加噪后序列的突变点. 对佛坪站日最高温度实测资料的动力学结构突变的准确检测进一步验证了该方法的有效性. 滑动移除小波分析法可为具有相关性的系统动力学结构突变的快速、准确检测提供一种途径.

关键词: 标度指数, 小波分析, 滑动移除小波分析方法, 突变检测

PACS: 92.60.Wc, 92.40.Zg

DOI: 10.7498/aps.66.079201

1 引言

在地球物理学时间序列相关动力学进程的研究过程中, 分形维数和标度指数一直是目前广被认可的方法^[1], 尤其是标度指数法. 如在系统动力学结构突变检测过程中将滑动数据移除和标度指数相结合提出的滑动移除去趋势波动分析(moving cut detrended fluctuation analysis, MC-DFA)、滑动移除重标极差分析(moving cut data-rescaled range analysis, MC-R/S)、滑动移除重标方差分析(moving cut data-rescaled variance analysis, MC-V/S)等^[2-5], 理想时间序列数值试验和实测资料

分析验证结果都表明, 相对于传统统计方法 Mann-Kendall、滑动 t 检验等, 这些方法不仅能够有效地检测系统动力学结构的突变点, 而且能够表征系统动力学结构突变前后的性质变化, 极大地丰富了系统动力学结构突变检测理论与方法体系. 其中, 标度指数的快速、准确的计算是这些系统动力学结构突变方法的核心. 标度指数的计算方法主要包括重标极差分析法(rescaled range analysis, R/S)^[6]、去趋势波动分析(detrended fluctuation analysis, DFA)^[7]、小波分析法(wavelet transformation, WT)^[8-15]、重标方差(rescaled variance analysis, V/S)法等^[16]. R/S分析法是最常用的非

* 国家自然科学基金青年科学基金(批准号: 51409005)、国家自然科学基金重大项目(批准号: 51190093)、国家自然科学基金(批准号: 51379014)和中央高校基本科研业务费专项资金(批准号: 310829161008)资助的课题.

† 通信作者. E-mail: honeber@126.com

参数标度指数计算方法,但在长序列分析中易受短期相关性和周期的影响,结果会出现一定的偏差^[17,18];V/S分析法通过方差代替R/S分析中的极差,具有较强的稳定性,相关实验证明其对于标度指数在临界值0.5附近的估算比R/S更加有效^[19],但其计算效率偏低;由于在各阶趋势成分处理上的优势,DFA方法非常适用于具有各种尺度噪声及趋势的非平稳序列的标度计算^[20,21];这些方法对于中小数据集序列的估计具有一定的有效性,但对于超大数据集序列的分析则需要进行复杂的计算和很高的内存要求,且结果有一定的偏差^[8,22].WT法是在序列尺度和时间域上进行,其多尺度特性与自相似过程的尺度不变性有着自然的联系^[23],可以快速地对数据集序列进行不同尺度的分解,通过分析不同尺度下各小波变幅的标度关系来计算标度指数,具有计算速度快、收敛性好的特点^[24],节省时间和内存;其次,通过改变小波基消失矩的数目,数据集序列的多项式趋势能够被严格地剔除,而相关数值试验模拟证明小波分析还具有很强的抗噪能力^[9,25],因而适用于超大数据集的非平稳序列的分析.

由于气候系统是一个非线性系统,它的观测数据量大,且常常呈现出一些非线性现象,如复杂周期、趋势、突变等,给序列标度指数的快速、准确计算带来一定的困难.通过WT计算标度指数可为解决这一困难提供一条思路.本文参照文献^[3]将滑动移除技术与WT相融合,发展了一种新的动力学结构突变检测方法——滑动移除小波分析法(moving cut data wavelet transformation, MC-WT).该方法与MC-R/S类似,是基于数据的移除对于具有相同动力学属性的相关序列标度指数的估算几乎没有影响的这一特征而提出.为了全方面检验MC-WT方法在动力学结构突变检测中的性能,文中首先通过构造线性和非线性两种理想时间序列,分别检测MC-WT方法的有效性,再以佛坪站日最高温度实测资料对方法进行验证.

2 MC-WT

2.1 标度指数 γ 的小波估计原理

目前通过小波分析来估计标度指数的方法主要有 wavelet-based analysis, averaged wavelet coefficient, wavelet transform modulus maxima 等,

本文采用文献^[9]提出的 wavelet-based analysis 方法,该方法计算的标度指数在高斯假设条件下是一个无偏估计量,且概念简单、能够对超大数据集的进行直接有效的分析.原理如下:

对于任一时间序列 $x(t)(t = 1, 2, \dots, N, N$ 为序列长度),其能量谱 $\Gamma_X(\omega)$ 满足

$$\Gamma_X(\omega) \propto c_f |\omega|^{1-2\gamma} \quad \omega \rightarrow 0, \quad (1)$$

式中, ω 为能量谱的频率; $c_f = c\pi\Lambda(2\gamma - 1)\sin(\pi - \pi\gamma)$, c 为一正常数; Λ 为 Gamma 函数; γ 为标度指数;

通过小波变换得到小波变换系数 $\{d_{j,k}\}(j = 1, 2, \dots, M; k = 1, 2, \dots, 2^{-j}/N)$, 其中, j 为尺度参数, k 为位置参数, M 为分解尺度;该系数可以度量 $x(t)$ 在时间 $2^j k$, 频率为 $2^{-j}\omega_0$ 处的能量, ω_0 为小波函数 $\psi(t)$ 的参考频率;对给定的尺度 j , $d_{j,k}^2$ 为能量谱估计量,即

$$\Gamma_X(2^{-j}\omega_0) = \left(\sum_{k=1}^{N_j} d_{j,k}^2 \right) / N_j, \quad (2)$$

式中, $N_j = 2^{-j}N$ 为尺度 j 小波变换系数个数;又因

$$\begin{aligned} E[\Gamma_X(2^{-j}\omega_0)] &= \int \Gamma_X(\omega) \cdot 2^j \cdot |\hat{\psi}(2^j\omega)|^2 d\omega, \end{aligned} \quad (3)$$

$E[\Gamma_X(2^{-j}\omega_0)]$ 表示 $\Gamma_X(2^{-j}\omega_0)$ 的数学期望,将(1)式代入(3)式,可得

$$\begin{aligned} E[\Gamma_X(2^{-j}\omega_0)] &= \int c_g |\omega|^{1-2\gamma} 2^j |\hat{\psi}(2^j\omega)|^2 d\omega \\ &= 2^{(2\gamma-1)j} c_g \int |\omega|^{1-2\gamma} |\hat{\psi}(\omega)|^2 d\omega, \end{aligned} \quad (4)$$

式中, c_g 为与 c_f 有关的一常数,对(4)式两边取2为底的对数,

$$\begin{aligned} \log_2 E[\Gamma_X(2^{-j}\omega_0)] &= (2\gamma - 1)j + \log_2 \left[c_g \int |\omega|^{1-2\gamma} |\hat{\psi}(\omega)|^2 d\omega \right] \\ &= (2\gamma - 1)j + b, \end{aligned} \quad (5)$$

式中, $b = \log_2 \left[c_g \int |\omega|^{1-2\gamma} |\hat{\psi}(\omega)|^2 d\omega \right]$, 且可收敛;将(2)式代入(5)式,可得

$$\log_2 \left[\left(\sum_{k=1}^{N_j} d_{j,k}^2 \right) / N_j \right] = (2\gamma - 1)j + b, \quad (6)$$

通过该式 $\log_2 \left[\left(\sum_{k=1}^{N_j} d_{j,k}^2 \right) / N_j \right]$ 与 j 之间的线性回归可得标度指数的估计量 γ .

2.2 标度指数 γ 的小波估计的计算步骤 [26,27]

1) 依据时间序列 $x(t)$ 的长度 N 选取分解尺度 M , 对其进行 Mallat 一维小波分解, 计算小波系数 $d_{j,k} (j = 1, 2, \dots, M; k = 1, 2, \dots, 2^{-j}/N)$.

2) 由小波系数计算中间参量 $\eta_j, s_j (j = 1, 2, \dots, M)$

$$\eta_j = \log_2 \left[\left(\sum_{k=1}^{N_j} d_{j,k}^2 \right) / N_j \right],$$

$$s_j = (N \ln^2 2) / 2^{j+1}.$$

3) 计算标度指数 γ 的小波估计值 $\gamma(j_1, j_2)$, 其中 $1 \leq j_1 \leq j_2 \leq M$,

$$\gamma(j_1, j_2) = \frac{1}{2} \left[\frac{\sum_{j=j_1}^{j_2} j s_j \eta_j - \sum_{j=j_1}^{j_2} j s_j \sum_{j=j_1}^{j_2} s_j \eta_j}{\sum_{j=j_1}^{j_2} s_j \sum_{j=j_1}^{j_2} j^2 s_j - \left(\sum_{j=j_1}^{j_2} j s_j \right)^2} + 1 \right]. \quad (7)$$

2.3 MC-WT 方法

参考文献 [3] 中 MC-R/S 方法, 本文给出 MC-WT 分析方法的具体步骤:

- 1) 依据序列长度 N 选择移除窗口长度 L ;
- 2) 取滑动步长为 L , 从序列 $x(t)$ 的第 $t (t = 1, 2, \dots, N - L + 1)$ 个数据开始连续移除 L 个数据, 形成 $\text{int}(N/L)$ (int 表示取整) 个长度为 $N - L$ 的子序列;

3) 通过小波估计各子序列的标度指数 γ , 可以得到一个长度为 $\text{int}(N/L)$ 的标度指数序列;

4) 对标度指数序列进行方差分析, 根据方差贡献大小确定原序列的突变点或突变区间.

与 MC-R/S 方法相同, 对于无动力学结构突变且具有相关性的时间序列, 任意移除该序列的数据, 对其标度指数计算的影响几乎可以忽略. 因此, 可以通过 1)–4) 检测时间序列不同时间段内数据对于整个序列标度指数贡献的大小来对系统的动力学结构突变进行检测.

3 数值实验

3.1 MC-WT 在线性时间序列突变中的检测

为了对 MC-WT 方法的性能进行全面了解, 首先进行线性序列的动力学突变检测试验. 理想序列 IS0 采用如下方程构建 [28]:

$$y(t) = \begin{cases} 2 \sin(0.2t) + 1 & (1 \leq t \leq 1000) \\ 1.5 \sin(0.2t) + 2 \cos(0.5t) - 0.2 & (1000 < t \leq 2000). \end{cases} \quad (8)$$

由 (8) 式可知, 序列 $y(t)$ 在 $t = 1000$ 处发生了动力学结构突变, 序列由初始的正弦函数方程变为正余弦函数控制的方程, 如图 1 所示. 选取滑动移除窗口 $L = 2$, 采用 Mallat 离散小波变换算法计算各子序列标度指数 γ , 其中滤波器组选用 sym8, 根据序列长度 $N = 2000$, 取 $M = 9, j_1 = 1, j_2 = M$. 图 2(a) 给出了在滑动移除窗口 $L = 2$ 情况下理想序列 IS0 的 MC-WT 检测结果, 容易看到, 在 $t = 1001$ 处, 标度指数 γ 发生了一次显著的均值突变, 突变前后呈现明显的两种动力学结构特征, 准确地刻画了原系统的动力学结构变化, 因此可通过标度指数的变化确定原序列的动力学突变. 图 2(b)–(d) 分别给出了序列在滑动移除窗口 $L = 5, 10, 50$ 情况下的检测结果, 可以看到, 不论滑动移除窗口 L 如何变化, 序列均在 $t = 1001$ 处发生了突变, 说明 MC-WT 方法在对线性序列的突变检测中受移除窗口长度的影响较小, 能够准确地检测系统的动力学突变. 相关研究表明, 在信号处理领域, 由于电子设备或通信系统内部缺陷 (如电路电流突变、元件静电感应、磁感应等) 和外部电磁干扰 (如太阳辐射电磁波、信号发射基站信号等), 信号从输入端开始不可避免地叠加了不同程度的噪声, 使得信道中的模拟信号受到干扰, 输出信号可能出现失真、误码等情况, 因此在进行数据分析时必须考虑强噪声对检测结果造成的影响 [29,30]. 为了测试噪声对 MC-WT 方法检测结果的影响程度, 分别对理想时间序列 IS0 依次添加信噪比 (signal-noise ratio, SNR) 为 20, 25, 30 dB 的高斯白噪声, 图 3 可以看到, 在 $SNR = 20, 25, 30$ dB 情况下加噪后的理想时间序列 (滑动移除窗口 $L = 5$) 标度指数 γ 均在 $t = 1001$ 处发生了突变, 说明 MC-WT 具有较强

的抗噪能力, 其他滑动步长结果类似. 以上分析表明, 对于线性序列的动力学结构突变MC-WT方法有着很强的检测能力, 然而在自然界中, 系统的演化呈现出复杂、动态的非线性特征, MC-WT方法的适用性如何, 需要进一步检测.

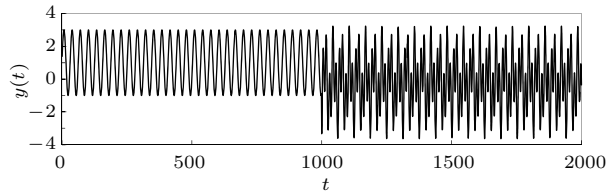


图1 理想时间序列 IS0
Fig. 1. The ideal time series IS0.

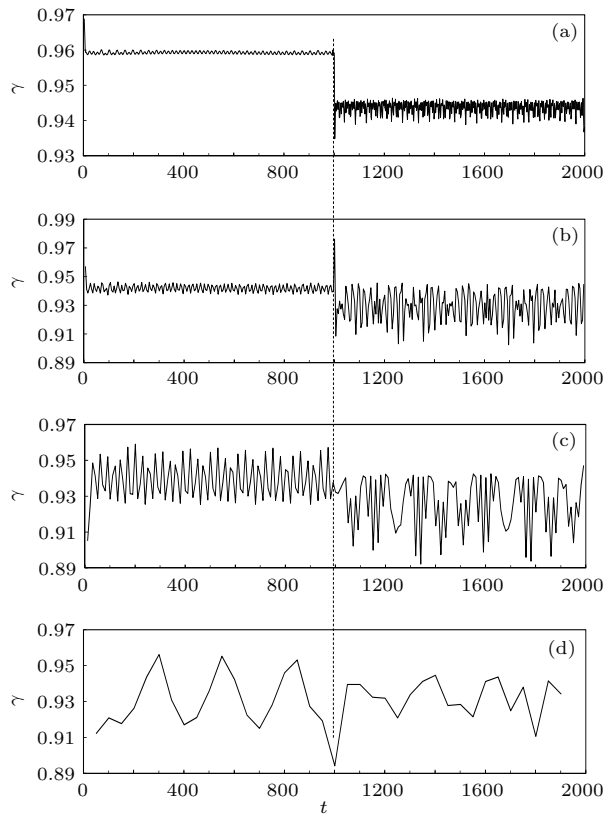


图2 理想序列 IS0 的 MC-WT 检测结果 (a) $L = 2$; (b) $L = 5$; (c) $L = 10$; (d) $L = 50$

Fig. 2. The MC-WT detection result of ideal time series IS0: (a) $L = 2$; (b) $L = 5$; (c) $L = 10$; (d) $L = 50$.

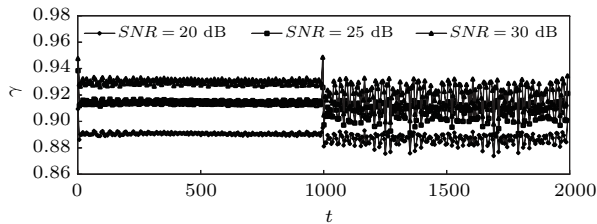


图3 加噪后理想序列 IS0 的 MC-WT 突变检测 ($L = 5$)
Fig. 3. The MC-WT mutation detection of ideal time series IS0 after adding noise ($L = 5$).

3.2 MC-WT 在非线性时间序列中的动力学结构突变检测

采用文献[29]中构造的理想时间序列 IS1(图4), 序列前 1000 个数据由 Logistic 映射产生, 后 1000 个数据由满足正态分布的随机数组成, 序列在 $t = 1001$ 处发生了突变, 由一种非线性状态转变为另一种随机状态. Logistic 映射方程如(9)式, 其中初值 $x_0 = 0.8$, 参数 $\mu = 3.8$.

$$x_{n+1} = \mu x_n(1 - x_n), \quad x \in [0, 1]. \quad (9)$$

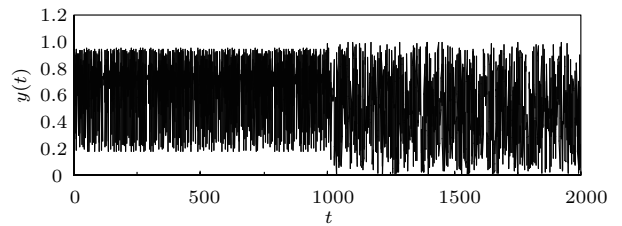


图4 理想时间序列 IS1
Fig. 4. The ideal time series IS1.

图5为 IS1 在不同滑动移除窗口 L 下的 MC-WT 检测结果, 滤波器组选用 sym8, 取分解尺度 $M = 9$ ($j_1 = 1, j_2 = M$). 从图5(a)—(d) 可以看到, 不论是滑动移除窗口 $L = 10$, 还是 $L = 20, 25, 50$, 其标度指数 γ 的演变趋势非常类似, 均在 $t = 1001$ 处发生了突变, 突变前后呈现两种状态, 表现为由 Logistic 映射所产生的数据的标度指数序列变化幅度相对平稳, 而由随机数据生成的标度指数序列其变化幅度相对较大, 表明数据的移除对于随机序列的影响较大. 同时随着移除窗口 L 的增大, 其序列动力学结构的突变更明显, 这说明 MC-WT 对于非线性时间序列的动力学结构突变同样有着良好的检测能力, 且对移除窗口 L 的长度依赖性较小. 作为比较, 图6给出了 IS1 的滑动 t 检验 ($n_1 = 10, n_2 = 10, n_1, n_2$ 分别为基准点前后子序列的长度) 和 Mann-Kendall 的检测结果, 由图6(a) 可以看到, 曲线呈现两个明显的阶段, 约在 $t = 1000$ 左右发生了动力学结构突变, 但很难准确定位突变点的位置; 图6(b) 中 UF 和 UB 线在置信区间 ($\alpha = 0.05$) 内 $t = 1000$ 左右均发生了改变, 但 UF 和 UB 线并没有相交, 依据 Mann-Kendall 定义判断此处并没有发生突变, 与实际情况不符.

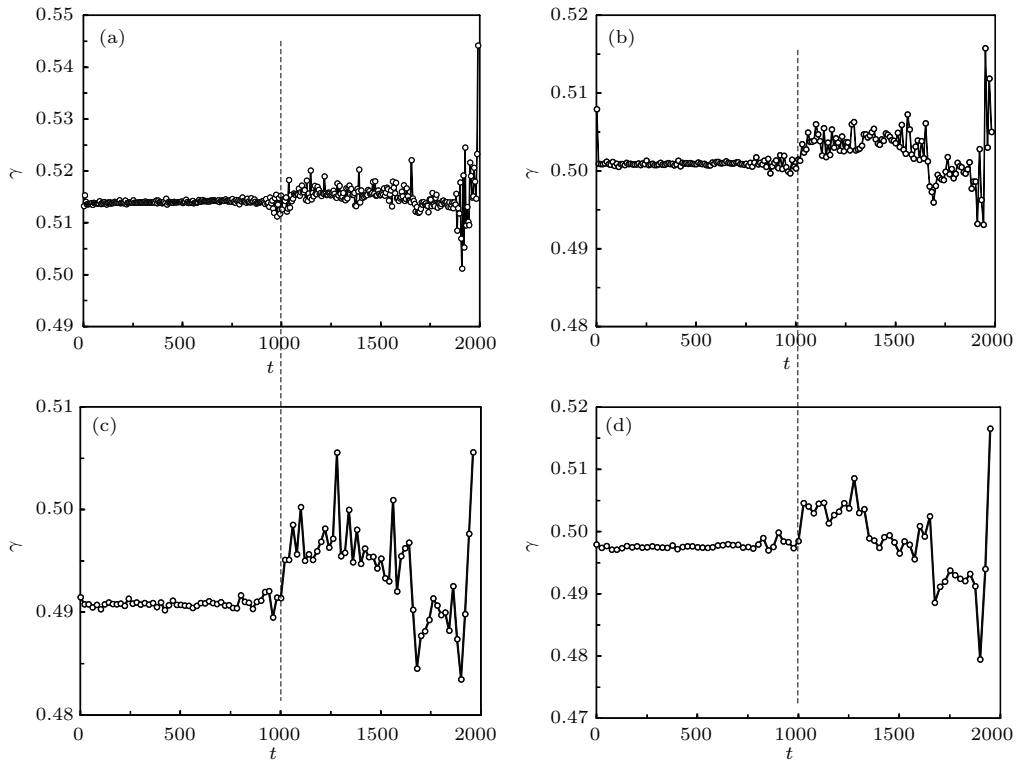


图5 理想时间序列 IS1 的 MC-WT 检测结果 (a) $L = 10$; (b) $L = 20$; (c) $L = 25$; (d) $L = 50$

Fig. 5. The MC-WT detection result of ideal time series IS1: (a) $L = 10$; (b) $L = 20$; (c) $L = 25$; (d) $L = 50$.

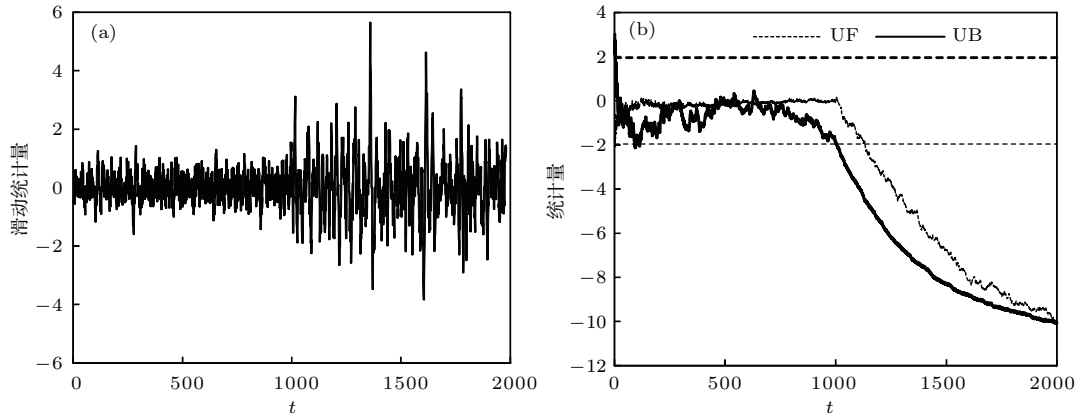


图6 理想时间序列 IS1 检测结果 (a) 滑动 t 检验; (b) Mann-Kendall

Fig. 6. The detection result of ideal time series IS1: (a) Moving t -test; (b) Mann-Kendall.

3.3 MC-WT 在非线性时间序列区间动力学结构突变中的检测

以上所考虑的是单点突变的情况, 即系统突然由一种状态过渡到另一种状态, 而实际情况中还可能发生区间突变的情况, 即系统在演变过程中某一段时间发生了动力学结构突变之后又恢复到原来的状态. 依据文献 [3] 构造理想时间序列 IS2, 即在 Logistic 映射产生一条长度为 1000 的理想演化序

列中, 预想时间序列在区间 [301, 330] 由确定性方程转变为随机状态. 故 IS2 在区间 [301, 330] 发生了一次动力学结构突变 (图 7). Logistic 映射方程见 (9) 式.

图 8 给出了 IS2 序列在不同滑动移除窗口 L 下 MC-WT 方法的检测结果, 滤波器组选用 sym8, 取分解尺度 $M = 9$ ($j_1 = 1, j_2 = M$). 可以看到, 滑动移除窗口长度 $L = 5, 10, 15, 30$, 在区间 [301, 331] 内, 其标度指数 γ 的变化明显大于其他区域,

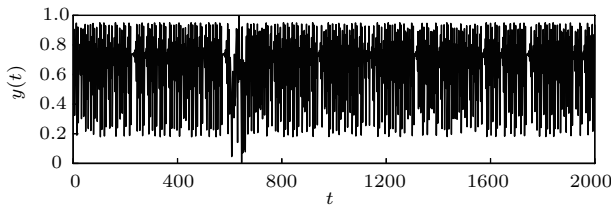


图7 理想时间序列 IS2
Fig. 7. The ideal time series IS2.

说明数据的移除对于该区间标度指数估计影响较大, 这与 IS2 构造的突变区间基本符合, 表明 MC-WT 对序列的区间突变有着良好的检测能力, 且对移除窗口长度依赖较小. 同时进一步证明文献 [3] 提到的具有相同动力学性质的数据对于序列标度指数计算的贡献度大致相同, 而具有不同动力学属性的数据对计算整个序列标度指数的贡献存在着显著的差异. 为了进一步验证突变区间的准确性, 采用文献 [3] 提出的方差分析方法来定量区分不同动力学特性对于标度指数估算的贡献, 即定义方差阈值为三倍平均标准方差, 超过该值即认为系统发生了突变. 图 9 分别为滑动移除窗口 $L = 5, 10, 15, 30$ 时的方差贡献图, 可以看到, 除了在区间 [301, 331] 内标度指数计算的方差贡献超过了三倍方差阈值, 在其他区域内方差贡献基本接近于 0 值, 可以判定序列在区间 [301, 331] 发生了突变, 与

MC-R/S 和 MC-V/S 分析结果一致 [31], 说明 MC-WT 方法具有良好动力学突变检测能力. 同时也注意到, 在滑动移除窗口 $L = 5$ 结尾附近和 $L = 10$ 开始端, 有个别方差贡献也超过了方差阈值, 可能与算法本身有关, 实验结果表明加大移除窗口的长度 L 可以消除该影响, 如图 9(c) 和图 9(d). 其次, 为了分析 MC-WT 的运行效率, 表 1 给出了不同滑动移除窗口下 MC-WT, MC-R/S 和 MC-V/S 在同一电脑下 (Inter Core (TM) i7-4510, 2.4 GHz, 4 GB, Win7) Matlab 2014 b 平台的运行时间, MC-WT 花费时间大约为 MC-R/S 的 1/6 和 MC-V/S 的 1/22, 因而 MC-WT 在处理大数据时将有明显的优势. 最后, 为了测试高斯白噪声对 MC-WT 方法检测结果的影响程度, 分别对理想时间序列 IS2 依次添加 SNR 为 15, 20, 25, 30 dB 的高斯白噪声, 图 10 分别给出了在滑动步长 $L = 10$ 情况下, 加噪后 IS2 序列的 MC-WT 动力学突变检测方差贡献图, 可以看到方差贡献的突变区间基本与真实区间一致 (除 SNR = 15 dB 情况下个别点超出阈值), 且随着 SNR 的逐渐增大突变区间愈加清晰, 没有出现虚假的突变区间, 说明 MC-WT 具有很好的抗噪能力.

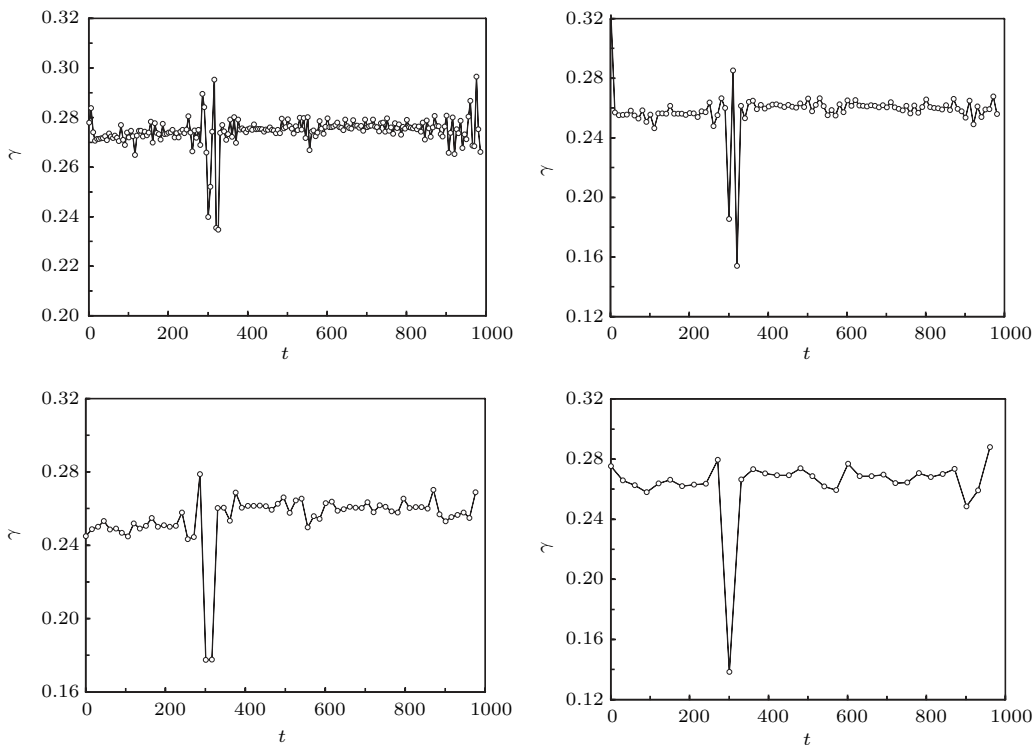


图8 理想时间序列 IS2 的 MC-WT 检测结果 (a) $L = 5$; (b) $L = 10$; (c) $L = 15$; (d) $L = 30$.
Fig. 8. The MC-WT detection result of ideal time series IS2: (a) $L = 5$; (b) $L = 10$; (c) $L = 15$; (d) $L = 30$.

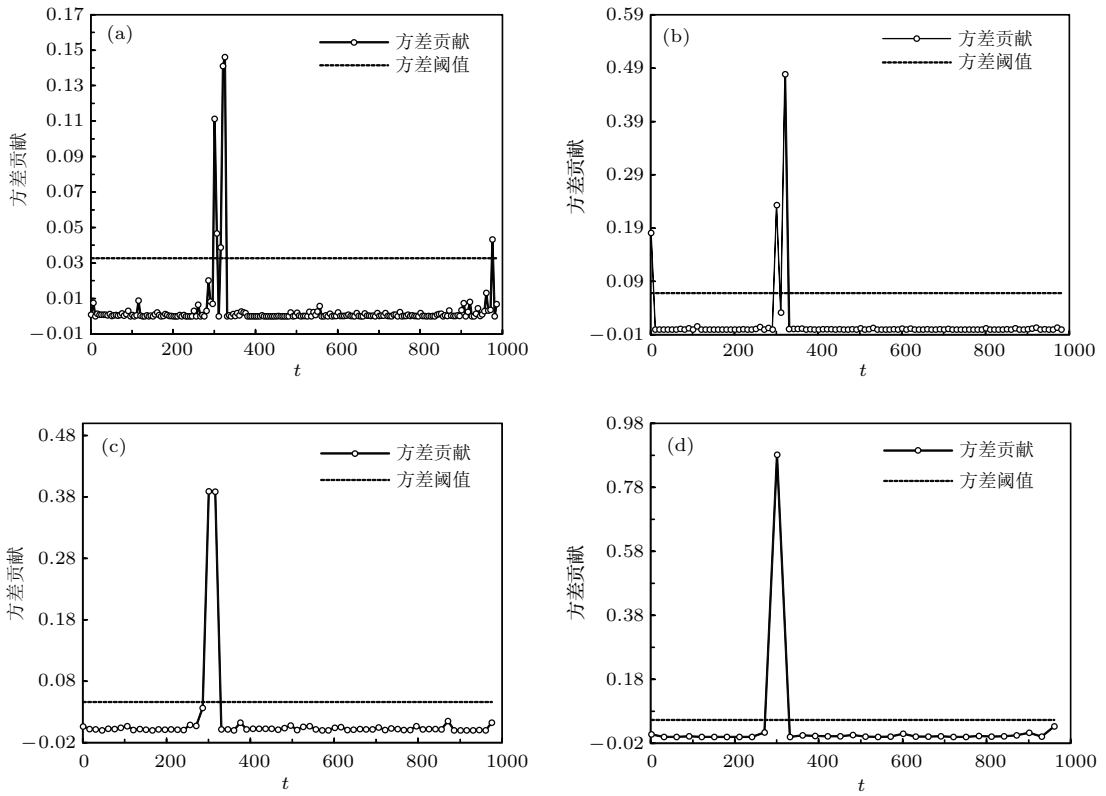


图9 理想时间序列 IS2 的 MC-WT 方差贡献 (a) $L = 5$; (b) $L = 10$; (c) $L = 15$; (d) $L = 30$

Fig. 9. The variance contribution of MC-WT detection result for IS2: (a) $L = 5$; (b) $L = 10$; (c) $L = 15$; (d) $L = 30$.

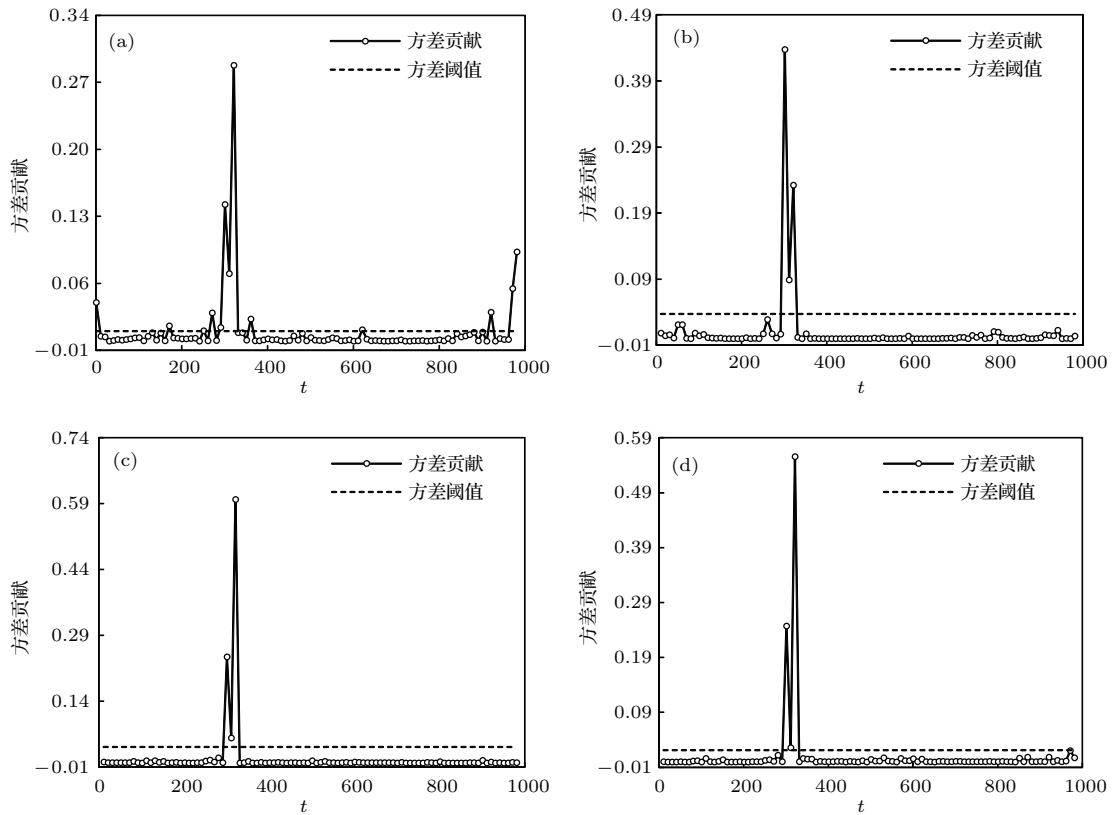


图10 加噪后 IS2 序列 MC-WT 方差贡献 (a) $SNR = 15$ dB; (b) $SNR = 20$ dB; (c) $SNR = 25$ dB; (d) $SNR = 30$ dB

Fig. 10. The variance contribution of MC-WT detection result for the IS2 after adding noise: (a) $SNR = 15$ dB; (b) $SNR = 20$ dB; (c) $SNR = 25$ dB; (d) $SNR = 30$ dB.

表1 不同移除窗口下 MC-WT, MC-R/S 和 MC-V/S 运行时间 (单位: s)

Table 1. The run time of MC-WT, MC-R/S, and MC-V/S under different remove windows (unit: s).

方法	窗口长度				
	$L = 2$	$L = 5$	$L = 10$	$L = 15$	$L = 30$
MC-WT	2.56	1.052	0.533	0.375	0.207
MC-R/S	17.068	6.821	3.364	2.230	1.131
MC-V/S	61.076	24.955	12.127	7.986	3.961

4 MC-WT 在实测资料中的应用

前文分析了 MC-WT 在理想时间序列动力学突变中的应用, 而实测资料则呈现出更加复杂的非线性动态特性. 鉴于此, 本文拟以实测温度资料测试 MC-WT 在突变检测中的性能. 实测温度资料采用渭河流域佛坪站 1960.1.1—2012.7.31 (共 19207 个数据) 逐日最高温度数据, 资料来源于中国气象数据网 (<http://data.cma.cn/>), 质量得到了控制. 图 11 给出了佛坪站日最高温度的 MC-WT 检测结

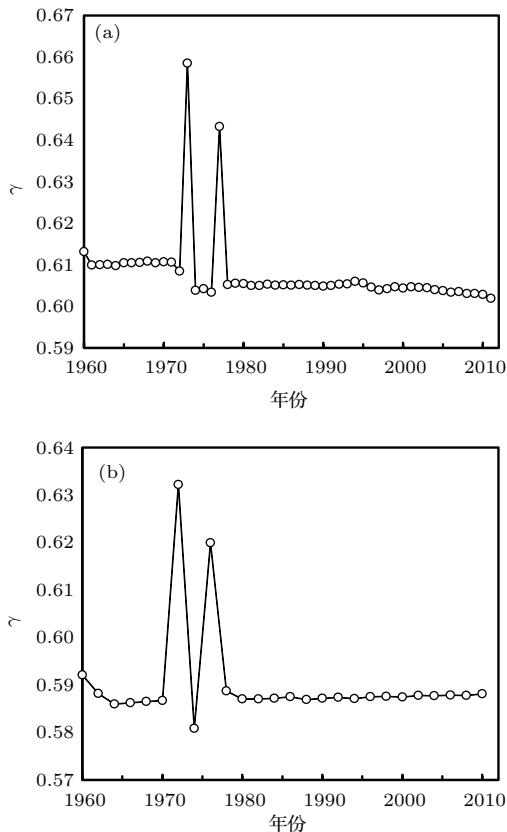


图 11 佛坪站逐日最高温度序列 MC-WT 检测结果 (a) $L = 365$ d; (b) $L = 730$ d

Fig. 11. The MC-WT detection results of daily maximum temperature sequence in Foping station: (a) $L = 365$ d; (b) $L = 730$ d.

果, 可以看到, 与理想试验结果类似, 不论滑动移除窗口 $L = 365$ d ($d = 1$ 日) 或 $L = 730$ d, 逐日最高温度的标度指数 γ 序列出现了一个基本相同的突变区间: 1972 (1973)—1978 年, 在 1978 年以后系统的动力学结构发生了突变, 由一种状态进入到另一种状态, 系统的标度指数降低, 随机性加大, 这与 20 世纪 70 年代末期全球的气候突变相一致 [32–36]. 作为对比, 图 12 给出了在滑动移除窗口 $L = 365$ d 的情况下 MC-WT 和 MC-R/S 方法突变检测的方差贡献 (MC-V/S 方差贡献超出方差阈值, 故剔除). 可以看到两种方法所得到的突变区间完全一致, 而 MC-WT 方法所花费的时间是 MC-R/S 方法的 1/25 左右, 说明在进行大数据分析中, MC-WT 具有更高的效率.

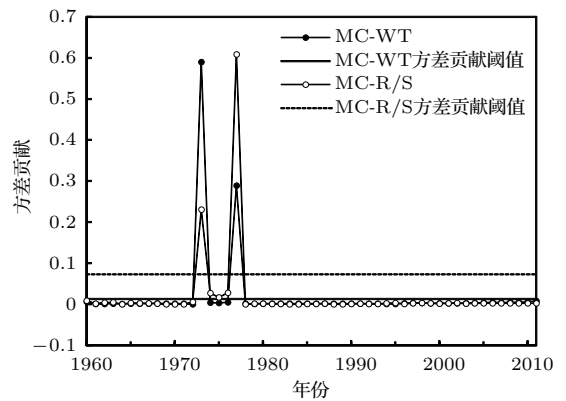


图 12 MC-WT 和 MC-R/S 突变检测方差贡献图 ($L = 365$ d)

Fig. 12. The variance contribution of MC-WT and MC-R/S detection result ($L = 365$ d).

表 2 佛坪站极端温度不同移除窗口下 MC-WT, MC-R/S 运行时间 (单位: s)

Table 2. The run time of MC-WT, MC-R/S under different remove windows in Foping station (unit: s).

方法	窗口长度	
	$L = 365$ d	$L = 730$ d
MC-WT	6.152	2.708
MC-R/S	154.018	74.232

5 结 论

本文通过融合小波标度指数与数据移除技术, 提出一种新的动力学结构突变检测方法—MC-WT. 理想时间序列的试验结果表明, MC-WT 的检测结果对滑动移除窗口的长度依赖小, 对噪声具有

一定的抗干扰能力, 不仅能对线性序列的动力学结构突变实现准确检测, 且对非线性序列的动力学结构变点、突变区间同样具有很好的检测能力. 实测资料的突变检测结果进一步印证了以上结论, 并证明其在更复杂的实测序列上仍能获得较好的检测效果. 与MC-R/S, MC-V/S相关时间序列动力学结构突变分析方法相比, MC-WT检测不仅具有相当的精确度, 且检测速度优势明显, 在大量数据分析中具有一定的优势, 可为相关时间序列的动力学结构突变分析提供一条新的途径. 同时研究中也注意到, 在某些情况下MC-WT在检测开始时会出现1—2个虚假的突变点, 这可能与小波分解算法的选取有关, 可以通过对比不同滑动窗口下检测结果予以剔除; 其次, 对于强噪声对信号序列的影响, 文中只考虑了高斯白噪声的情况, 实际情况中各种噪声(如尖峰噪声)对信号序列的影响不同^[37,38], 因此, 下一步将展开相关研究.

参考文献

- [1] Rehman S, Siddiqi A H 2009 *Chaos Soliton. Fract.* **40** 1081
- [2] He W P, Feng G L, Wu Q, He T, Wan S Q, Chou J F 2012 *Int. J. Climatol.* **32** 1604
- [3] He W P, Wu Q, Zhang W, Wang Q G, Zhang Y 2009 *Acta Phys. Sin.* **58** 2862 (in Chinese) [何文平, 吴琼, 张文, 王启光, 张勇 2009 物理学报 **58** 2862]
- [4] He W P, Deng B S, Wu Q, Zhang W, Cheng H Y 2010 *Acta Phys. Sin.* **59** 8264 (in Chinese) [何文平, 邓北胜, 吴琼, 张文, 成海英 2010 物理学报 **59** 8264]
- [5] Sun D Y, Zhang H B, Huang Q 2014 *Acta Phys. Sin.* **63** 209203 (in Chinese) [孙东永, 张洪波, 黄强 2014 物理学报 **63** 209203]
- [6] Hurst H E 1951 *Trans. Am. Soc. Civ. Eng.* **116** 770
- [7] Peng C K, Buldyrev S V, Havlin S, Simons M, Stanley H E, Goldberger A L 1994 *Phys. Rev. E* **49** 1685
- [8] Simonsen I, Hansen A, Nes O M 1998 *Phys. Rev. E* **58** 2779
- [9] Veitch D, Abry P C 1999 *IEEE Trans. Inf. Theory* **45** 878
- [10] Jones C L, Lonergan G T, Mainwaring D E 1996 *J. Phys. A: Math. Gen.* **29** 2509
- [11] Hu K, Ivanov P C, Chen Z, Carpena P, Stanley H E 2001 *Phys. Rev. E* **64** 011114
- [12] Gloter A, Hoffmann M 2007 *Ann. Stat.* **35** 1947
- [13] Manimaran P, Panigrahi P K, Parikh J C 2005 *Phys. Rev. E* **72** 046120
- [14] Ciftlikli C, Gezer A 2010 *Turk. J. Elec. Eng. & Comp. Sci.* **18** 117
- [15] Wu L, Ding Y M 2015 *Int. J. Wavelets Multiresolut Inf. Process.* **13** 1550044
- [16] Giraitis L, Kokoszka P, Leipus R, Teyssiere G 2003 *J. Econom.* **112** 265
- [17] Clausel M, Roueff F, Taqqu M S, Tudor C 2014 *Esaim Probab. Stat.* **18** 42
- [18] Taqqu M S, Teverovsky V 1997 *Comm. Stat. Stoch. Model* **13** 723
- [19] Cajueiro D O, Tabak B M 2005 *Math. Comp. Sim.* **70** 172
- [20] Kantelhardt J W, Koscielny-Bunde E, Rego H H A, Havlin S, Bunde A 2001 *Physica A* **295** 441
- [21] Matos J A O, Gama S M A, Ruskin H J, Sharkasi A A, Crane M 2008 *Physica A* **387** 3910
- [22] Mielniczuk J, Wojdylo P 2007 *Comput. Stat. Data Anal.* **51** 4510
- [23] Zhao Y Z, Wu L W 2014 *Comput. Eng. Appl.* **50** 154 (in Chinese) [赵彦仲, 吴立文 2014 计算机工程与应用 **50** 154]
- [24] Dang T D, Molnar S 1999 *Period. Polytech, Electr. Eng.* **43** 227
- [25] Giordano S, Miduri S, Pagano M, Russo F, Tartarelli S 1997 *Proceedings of 13th International Conference on Digital Signal Processing* Santorini, Greece, July 2–4, 1997 p479
- [26] Li X B, Ding J, Li H Q 1999 *Adv. Water Sci.* **10** 144 (in Chinese) [李贤彬, 丁晶, 李后强 1999 水科学进展 **10** 144]
- [27] Li X B, Ding J, Li H Q 1998 *J. Hydraul. Eng.* **8** 21 (in Chinese) [李贤彬, 丁晶, 李后强 1998 水利学报 **8** 21]
- [28] Wang Q G, Zhang Z P 2008 *Acta Phys. Sin.* **57** 1976 (in Chinese) [王启光, 张增平 2008 物理学报 **57** 1976]
- [29] He W P, He T, Cheng H Y, Zhang W, Wu Q 2011 *Acta Phys. Sin.* **60** 049202 (in Chinese) [何文平, 何涛, 成海英, 张文, 吴琼 2011 物理学报 **60** 049202]
- [30] Jin H M, He W P, Zhang W, Feng A X, Hou W 2012 *Acta Phys. Sin.* **61** 129202 (in Chinese) [金红梅, 何文平, 张文, 冯爱霞, 侯威 2012 物理学报 **61** 129202]
- [31] He W P, Liu Q Q, Jiang Y D, Lu Y 2015 *Chin. Phys. B* **24** 049205
- [32] Powell A M, Xu J J 2011 *Theor. Appl. Climatol.* **104** 443
- [33] Feng G L, Gong Z Q, Zhi R 2008 *Acta Meteor. Sin.* **66** 892 (in Chinese) [封国林, 龚志强, 支蓉 2008 气象学报 **66** 892]
- [34] Shi N, Chen J Q, Tu Q P 1995 *Acta Meteor. Sin.* **53** 431 (in Chinese) [施能, 陈家其, 屠其璞 1995 气象学报 **53** 431]
- [35] Tong J L, Wu H, Hou W, He W P, Zhou J 2014 *Chin. Phys. B* **23** 049201
- [36] Wu H, Hou W, Yan P C, Zhang Z S, Wang K 2015 *Chin. Phys. B* **24** 089201
- [37] Zhang M L, Qu H, Xie X R, Kurths J 2017 *Neurocomputing* **219** 333
- [38] Wan L, Zhang Y, Lin J, Jiang C D, Lin T T 2016 *Chin. J. Geophys.* **59** 2290 (in Chinese) [万玲, 张扬, 林君, 蒋川东, 林婷婷 2016 地球物理学报 **59** 2290]

Application of moving cut data-wavelet transformation analysis in dynamic structure mutation testing*

Sun Dong-Yong¹⁾ Zhang Hong-Bo^{1)†} Wang Yi-Min²⁾

1) (Key Laboratory of Subsurface Hydrology and Ecological Effect in Arid Region of Ministry of Education, School of Environmental Science and Engineering, Chang'an University, Xi'an 710054, China)

2) (Key Laboratory of Northwest Water Resources and Environment Ecology of MOE, Institute of Water Resources and Hydro-electric Engineering, Xi'an University of Technology, Xi'an 710048, China)

(Received 28 November 2016; revised manuscript received 20 December 2016)

Abstract

The scaling exponent is an effective nonlinear dynamic index, which can be used to detect the dynamic structure mutations of the correlation time series by the moving cut a fixed window technology. The immediacy and accuracy of scaling exponent is very important for detecting the series change points, however, some of the existing scale index calculation methods (such as rescaled range analysis and rescaled variance analysis) take none of these into account. Wavelet transform analysis can quickly decompose the sequence on different scales, and then the scaling index can be calculated by analyzing the scaling relation of wavelet coefficients on different scales, which has the characteristics of fast calculation speed and good convergence and memory saving. By moving cut window technology, in the present paper we put forward a new method, i. e., the moving cut data-wavelet transformation for detecting a series of dynamic structure mutations. The principle is that the removal of the data has little effect on the estimation of the scaling exponents of the correlation time series with the same dynamical properties. In order to test the performance of the method, first of all, the dynamic structure mutation analyses of linear ideal time series and nonlinear ideal time series are carried out by selecting different moving cut fixed windows. The test results show that the method can quickly and accurately detect the dynamic structure change points and intervals both in linear time series and nonlinear time series, besides, its calculation speed is obviously better than the moving cut data-rescaled range analysis and the moving cut data-rescaled variance analysis. It has strong stability, and depends less on the moving cut window length, which will have some advantages in the large data processing. At the same time, in order to detect the influence of noise on the method, the linear and nonlinear ideal time series are added to the white Gaussian noise ($SNR = 20, 25, 30$ dB), respectively, and the results show that the method has a strong anti-noise ability with different moving cut window lengths, can still quickly and accurately detect the mutation point or interval in different noise additions. Finally, the method is used to detect the dynamic structure mutation of measured daily maximum temperature data of Foping station in Wei basin, the experimental results indicate that the mutation interval is consistent with the abrupt change in 1970's on a global scale, which further verifies the validity of the method.

Keywords: scaling exponent, wavelet analysis, moving cut data-wavelet transformation analysis, mutations detection

PACS: 92.60.Wc, 92.40.Zg

DOI: 10.7498/aps.66.079201

* Project supported by the Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 51409005), the Major Program of the National Natural Science Foundation of China (Grant No. 51190093), the National Natural Science Foundation of China (Grant No. 51379014) and the Fundamental Research Funds for the Central Universities of Ministry of Education of China (Grant No. 310829161008).

† Corresponding author. E-mail: honeyber@126.com