

## 噪声对滑动移除近似熵的影响\*

金红梅<sup>1)</sup> 何文平<sup>2)†</sup> 张文<sup>1)</sup> 冯爱霞<sup>1)</sup> 侯威<sup>2)</sup>

1) (兰州大学大气科学学院, 兰州 730000)

2) (国家气候中心, 北京 100081)

(2011年9月9日收到; 2011年11月8日收到修改稿)

受外强迫和仪器本身的测量误差等因素的影响, 观测数据中经常包含噪声和扰动等一些虚假的信息. 针对这一问题, 本文研究了各种不同噪声对滑动移除近似熵的影响. 研究表明, 滑动移除近似熵的检测结果受尖峰噪声和高斯白噪声的影响较小, 意味着其具有很强的抗噪能力, 这为该方法在实际观测资料中的广泛应用提供了坚实的实验基础.

**关键词:** 滑动移除近似熵, 噪声, 突变检测

**PACS:** 92.60.Wc

## 1 引言

时间序列分析已经有了相当长的历史和成就, 突变分析是其一个重要的研究方向<sup>[1-3]</sup>, 在实际中, 时间序列往往是非平稳的, 这可能会导致分析中得到错误的结论, 譬如“虚假回归”等现象的出现, 即两个本来没有任何因果关系的变量, 却有很高的相关性. 在气象观测资料中, 各种趋势和噪声等造成的非平稳性不可避免, 因此, 研究各种常见的非平稳现象对气候突变分析的影响十分重要. 作者此前通过大量数值试验测试了各种趋势对滑动移除近似熵 (moving cut data-approximate entropy, MC-ApEn) 方法的影响<sup>[4]</sup>, 结果显示, 周期性趋势、线性趋势和非线性趋势对 MC-ApEn 方法的突变检测结果影响较小. 但尚未系统地研究噪声对其突变检测结果的影响.

由于受外强迫和仪器本身测量误差的影响, 气象观测资料中不可避免地存在各种噪声和扰动等非平稳现象<sup>[5-8]</sup>, 使观测值与大气真实状态之间存在着一定的观测误差, 尽管可以对原始数据进行滤

波处理, 但亦不可能完全消除噪声, 这有可能会对气象观测资料的分析结果产生一些不良影响<sup>[9-11]</sup>, 譬如, 尖峰噪声和高斯白噪声的存在会对时间序列长程相关性的分析造成一定影响, 使所获得的时间序列的标度指数较真实信号偏小<sup>[12]</sup>. 本文主要考虑了观测资料中两种比较常见的噪声: 尖峰噪声与高斯白噪声. 产生尖峰噪声的原因有多种, 其中包括电磁干扰以及通信系统的故障和缺陷, 通信系统的电气开关和继电器状态的改变等. 当线路自身产生了随机信号或电磁干扰时, 气象电子通信设备的输出信号可能出现误码、雷达信号失真等现象<sup>[13]</sup>; 由于各种电磁干扰的影响, 气象电子设备会受到不同程度的影响, 因而不可避免影响到气象观测资料的准确性, 高斯白噪声就是其中一种重要的影响因素<sup>[14]</sup>, 它是短波信道中存在的复杂干扰因素之一.

20世纪60年代中期以来, 以 Thom<sup>[15]</sup> 的工作为先导逐步形成了突变理论, 并被广泛地应用于气候、地震、医学等各个领域. 气候突变被称为气候变化的不连续性、气候的跳跃, 是普遍存在于气候系统中的一种重要现象, 而气候系统是一种高度复杂的非线性动力系统, 对其复杂性进行度量则有助

\* 全球变化重大研究计划 (批准号: 2012CB955203), 国家自然科学基金 (批准号: 40930952, 41175067 和 40905034), 公益性行业 (气象) 科研专项 (批准号: GYHY201106016 和 GYHY201106015) 和国家科技支撑计划 (批准号: 2007BAC29B01 和 2009BAC51B04) 资助的课题.

† E-mail: wenping\_he@163.com

于认识气候变化的内在本质<sup>[16]</sup>, 气候突变的检测与归因对认识气候变化的性质及气候预测有重要意义. 传统的气候突变检测方法主要有: 滑动  $-t$  检验、Mann-Kendall 法、Cramer 法、Yamamoto 法、滤波检测法、Fisher 最优分割法和最大概率突变点检测方法等<sup>[17]</sup>. 继传统方法之后, 近几年相继发展了多种新方法<sup>[18-34]</sup>. Pincus 从衡量时间序列复杂性的角度提出了近似熵<sup>[35-37]</sup> (approximate entropy, ApEn) 的概念, 已得到广泛应用<sup>[38-42]</sup>. 王启光等<sup>[43]</sup> 结合滑动技术将 ApEn 用于气候系统的突变检测, 得到了较好的结果. 何文平等<sup>[44]</sup> 基于近似熵发展了滑动移除近似熵, 与滑动近似熵相比, 该方法能够更为准确地检测突变开始的时间, 但是文献中仅仅研究了理想时间序列突变检测的情况.

鉴于噪声在资料分析中潜在的不利影响, 本文在研究各种趋势对 MC-ApEn 影响的基础上<sup>[4]</sup>, 进一步利用非线性时间序列研究了尖峰噪声和高斯白噪声对 MC-ApEn 的影响, 发现 MC-ApEn 具有较强的抗噪能力.

## 2 方法<sup>[44]</sup>

### 2.1 ApEn 方法

ApEn 是一种基于边缘概率分布统计量化时间序列复杂程度的方法, 其算法如下<sup>[35-37,44]</sup>:

1) 对时间序列  $\{u(i), i = 1, 2, \dots, N\}$  进行相空间重构, 重构维数为  $m$ , 据此可构造一组维数为  $m$  的新向量  $X(1), X(2), \dots, X(N-m+1)$ , 其中

$$X(i) = \{u(i), u(i+1), \dots, u(i+m-1)\}, \\ i = 1, 2, \dots, N-m+1; \quad (1)$$

2) 计算任意向量  $X(i)$  与其余向量  $X(j)$  之间的相对欧式距离  $d[X(i), X(j)]$ , 即

$$d[X(i), X(j)] = \max [ |u(i+k) - u(j+k)| ], \\ k = 0, 1, \dots, m-1; \quad (2)$$

3) 设定容许偏差为  $r$ , 统计每个向量  $X(i)$  的  $d[X(i), X(j)]$  小于  $r$  的数目, 求出该数目与向量总数  $N-m+1$  的比  $C_i^m(r)$  为

$$C_i^m(r) = \{d[X(i), X(j)] \\ \leq r \text{ 的个数} / (N-m+1)\}; \quad (3)$$

4) 将  $C_i^m(r)$  取对数, 再求其对所有  $i$  的平均值, 记作  $\phi^m(r)$ , 即

$$\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_i^m(r); \quad (4)$$

5) 维数  $m$  增加 1, 重复 1)–4) 步, 得  $C_i^{m+1}(r)$  和  $\phi^{m+1}(r)$ ;

6) 近似熵的估计值定义为

$$\text{ApEn}(m, r) = \phi^m(r) - \phi^{m+1}(r). \quad (5)$$

显然, ApEn 值与  $m, r$  的取值有关. Pincus 建议取  $m = 2, r = 0.1-0.2 \sigma$  ( $\sigma$  是原始数据的标准偏差). 在本文中, 取参数  $m = 2, r = 0.15 \sigma$ .

近似熵方法反映了时间序列在模式上的自相似程度, 即序列在  $m$  维情况下两点组成的模式间的近似程度, 以及当维数变化时, 产生新模式的可能性大小. ApEn 值愈大, 说明产生新模式的概率愈大, 序列越复杂, 系统可预测性越差. 它给出新模式发生率随维数而增减的情况, 因而反映数据在结构上的复杂性. 同时它并不是企图完全重构吸引子, 因此对各种非线性时间序列有很好的适用性<sup>[43,44]</sup>.

### 2.2 MC-ApEn 方法

MC-ApEn 是在 ApEn 方法的基础上通过滑动移除技术形成的一种突变检测方法, 其具体计算过程如下<sup>[44]</sup>:

1) 选择滑动移除数据的窗口长度  $L$ ;

2) 从待分析时间序列的第  $i(i = 1, 2, \dots, N-L+1, N$  为时间序列中的总记录个数) 个数据开始连续移除  $L$  个数据, 再将剩余  $N-L$  个数据直接连在一起得到一个新的时间序列;

3) 计算新序列的 ApEn 值;

4) 保持移除数据的窗口长度不变, 以步长  $L$  逐步移动窗口, 重复 2), 3) 步操作, 直到原序列结束为止;

5) 通过 1)–4) 步操作可得到一个随着窗口移动的 ApEn 序列;

6) 基于不同动力学性质的数据其复杂性大小不相同, 而具有相同动力学性质的数据的复杂性差异不大这一特点, 结合步骤 5) 中得到的 ApEn 序列判断突变点或突变区间.

### 3 不同噪声对 MC-ApEn 的影响研究

其 Logistic 方程为

$$x_{n+1} = ux_n(1 - x_n), \quad x \in [0, 1]. \quad (6)$$

#### 3.1 理想时间序列的构建

为了模拟气候系统的非线性特征, 类似于文献 [44] 构建了非线性理想时间序列 IS1, 序列总长为 2000, 前 1000 个数据由 Logistic 模型产生, 后 1000 个数据由正态分布的随机数模拟产生.

式中, 初值  $x_0 = 0.8$ , 参数  $u = 3.8$ , IS1 随时间的演化曲线已由图 1 给出, 从该时间序列的构造过程来看, 两段子序列分别具有不同的动力学特征, 当  $t = 1001$ , 序列由确定性模型——Logistic 方程状态转变为随机行为, 即系统在该时刻发生了动力学结构突变.

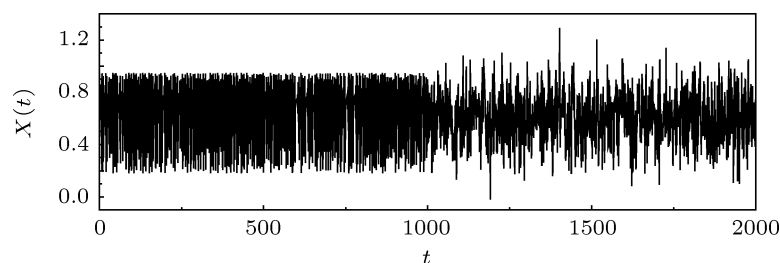


图 1 非线性时间序列 IS1 随时间的演变, 序列总长为 2000, 前 1000 个数据由 Logistic 模型产生, 后 1000 个数据由正态分布的随机数模拟产生

#### 3.2 尖峰噪声对 MC-ApEn 方法的影响

为了测试尖峰噪声对 MC-ApEn 方法的影响, 在理想时间序列 IS1 中分别随机叠加 12, 24 个尖峰

噪声, 分别占原序列长度的 6%和 12%, 尖峰噪声取值大小在 0.9—2.5 之间, 尖峰噪声的数目相对原序列较少, 原序列 IS1 的动力学性质不会发生质的变化 (图 2).

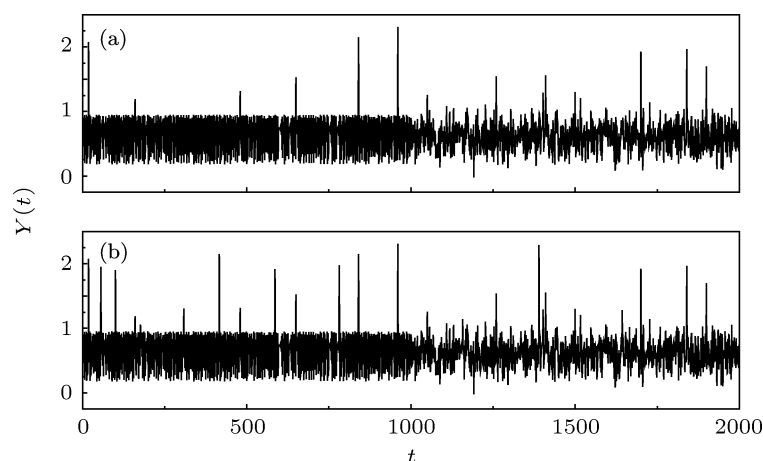


图 2 含尖峰噪声的时间序列, 尖峰噪声大小在 0.9—2.5 之间变化 (a) 尖峰噪声个数  $n = 12$ ; (b)  $n = 24$

图 3 给出了图 2 所示的理想时间序列的 MC-ApEn 检测结果, 当滑动子序列长度为 20 时, 尽管通过 MC-ApEn 得到的 ApEn 值大小略有差异, 但突变开始时间完全一致, ApEn 随时间的演变趋势

十分相似, 均在  $t < 1000$  时, 通过滑动移除计算得到的 ApEn 值明显大于  $t > 1000$  时的情形, 这表明在  $t = 1000$  前后, 时间序列的复杂性发生了明显变化, 意味着其动力学结构在此处发生了转变, 这与

理想序列 IS1 实际的动力学结构改变情况完全一致: 在该序列中, 前 1000 个数据是由确定性动力学方程 Logistic 映射产生, 其复杂性要小于后 1000 个随机数. 根据 ApEn 的物理意义, ApEn 值越大, 则表示序列复杂性越高, 若移除等长度的数据, 移除复杂性相对较大的随机数后所获取的 ApEn 值更小一些, 即部分随机数的移除导致剩余数据的复杂性减小, 与理想序列的动力学性质完全相符. 从 ApEn 的变化趋势可以看出, 序列从一种稳定状态突然转变到了另一种稳定状态, 据此可以容易地判断出突变的开始时间.

为了考察子序列长度对于 MC-ApEn 检测结果的影响, 本文对子序列长度取任意值时含尖峰噪声的 IS1 进行了测试. 作为一个分析试验的例子, 图 3(c) 和 (d) 给出了子序列长度  $L = 2$  时的 MC-ApEn 检测结果, 发现尖峰噪声对于子序列长度较小时的影响仍然可以忽略, 而且对于子序列长度更长时的情形也做了大量试验, 得到了同样的结果. 不仅如此, 数值试验过程中还对时间序列随机添加了不同大小和不同数目的尖峰噪声, 通过取不同尺度的滑动子序列发现, MC-ApEn 均能准确地检测到突变点. 为了避免单一个例对研究结果的可靠性说服力不足的缺陷, 本文对类似于 IS1 的众多理想时间序列进行了检测, 与 IS1 时得到的检测结果

完全一致. 在数值试验过程中通过取不同长度的原序列和滑动子序列进行数值试验的过程中发现, 当原序列长度一定, 滑动子序列越长, MC-ApEn 方法检测得到的突变点的位置越清晰. 这可能主要在于样本量的增加改善了 ApEn 计算结果的稳定性. 这表明相对少量的尖峰噪声对 MC-ApEn 方法突变检测结果不会产生显著的影响, MC-ApEn 方法可以很好地适用于具有尖峰噪声的时间序列的突变检测.

信号含噪比例的大小对分析信号内禀特性非常重要, 为了探究尖峰噪声数目占原序列长度的比例对 MC-ApEn 突变检测结果的影响, 首先保持尖峰噪声取值范围不变, 增大 IS1 中所添加尖峰噪声的比例 (含不同比例尖峰噪声的 IS1 随时间的演变图略), 通过多次数值试验发现 MC-ApEn 方法突变检测结果受尖峰噪声比例影响较小. 图 4 给出了对 IS1 添加尖峰噪声比例分别为 5%, 10%, 15%, 20% 时 MC-ApEn 的突变检测情况, 从图中可以看出, 不论 IS1 含尖峰噪声的比例多大, 其检测结果呈现出惊人的一致性, 均可以非常明显地检测到突变点的位置. 数值试验过程中还对 IS1 添加了更大比例的尖峰噪声, 但试验结果表明尖峰噪声占原序列长度的比例对 MC-ApEn 突变检测结果影响仍然较小 (图略).

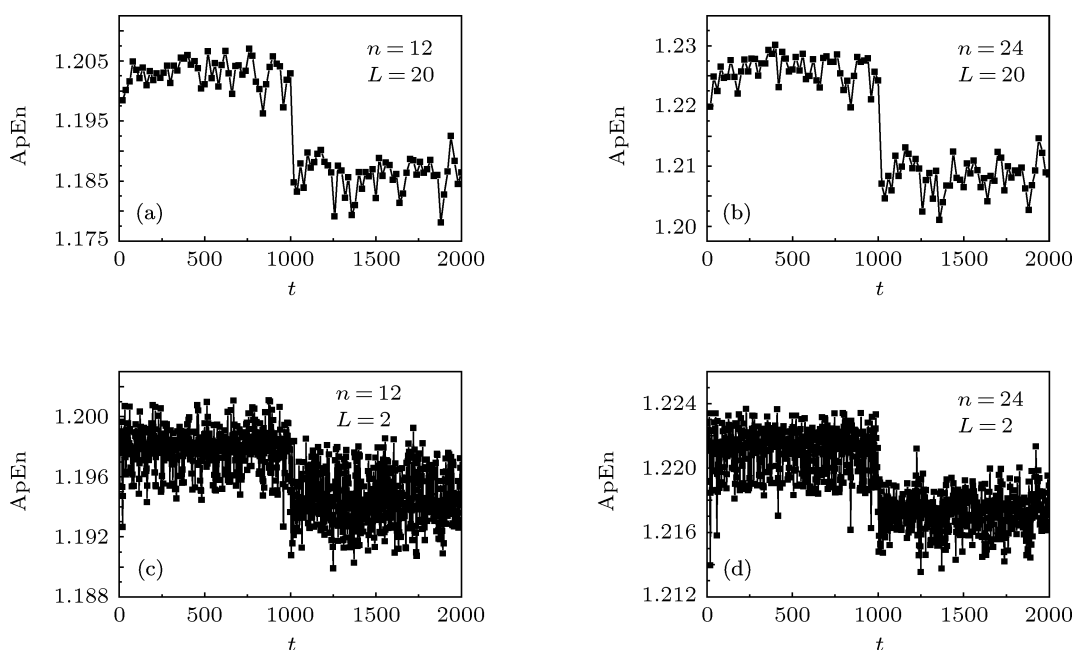


图 3 MC-ApEn 对含不同尖峰噪声的理想序列 IS1 的突变检测 (a) 滑动子序列长度  $L = 20$ , 尖峰噪声个数  $n = 12$ ; (b) 同 (a), 但  $n = 24$ ; (c) 同 (a), 但  $L = 2$ ; (d) 同 (b), 但  $L = 2$

此外,为了测试尖峰噪声的大小对 MC-ApEn 突变检测结果的影响,在时间序列 IS1 中添加了个数占原序列长度为 10% 的尖峰噪声,所含尖峰噪声的最大值约为原序列中尖峰噪声最大值的 5—6 倍(图 5(a)). 图 5(b), (c), (d) 分别展示了取不同

滑动子序列长度时 MC-ApEn 的突变检测情况,发现 ApEn 值随时间的演变趋势均以  $t = 1000$  为界呈现着两种不同的状态,可以清楚地看到突变的开始时间,说明 MC-ApEn 方法突变检测结果受尖峰噪声振幅的影响较小.

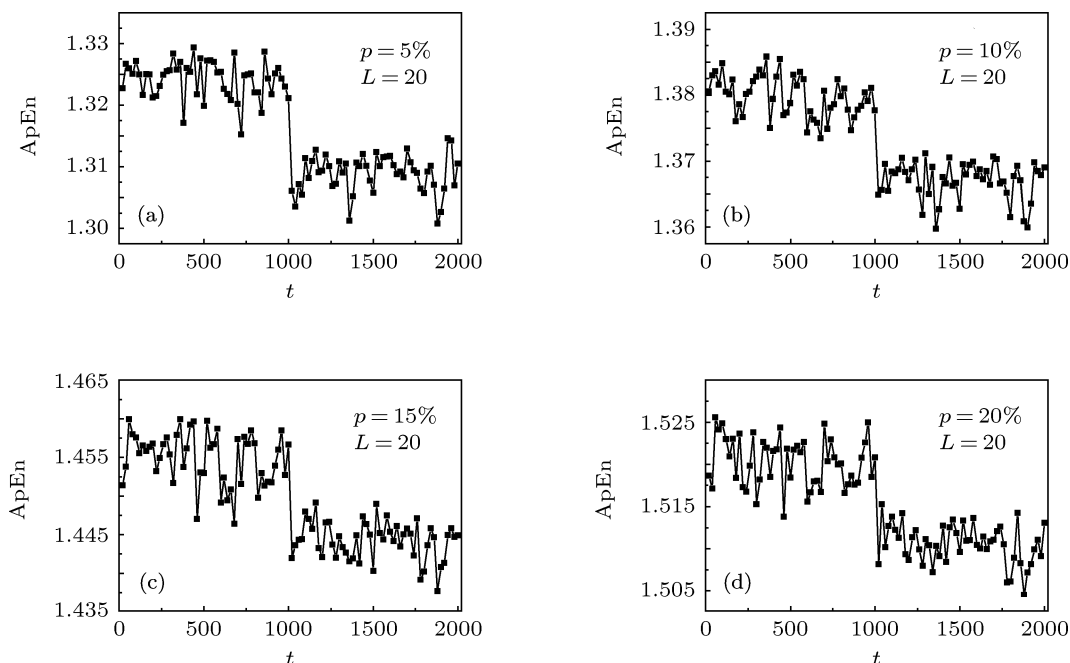


图 4 MC-ApEn 对含不同比例尖峰噪声的非线性时间序列 IS1 的突变检测,尖峰噪声大小在 0.9—2.5 之间变化,滑动子序列长度  $L = 20$  (a) 尖峰噪声个数占原序列长度的比例  $P = 5\%$ ; (b)  $P = 10\%$ ; (c)  $P = 15\%$ ; (d)  $P = 20\%$

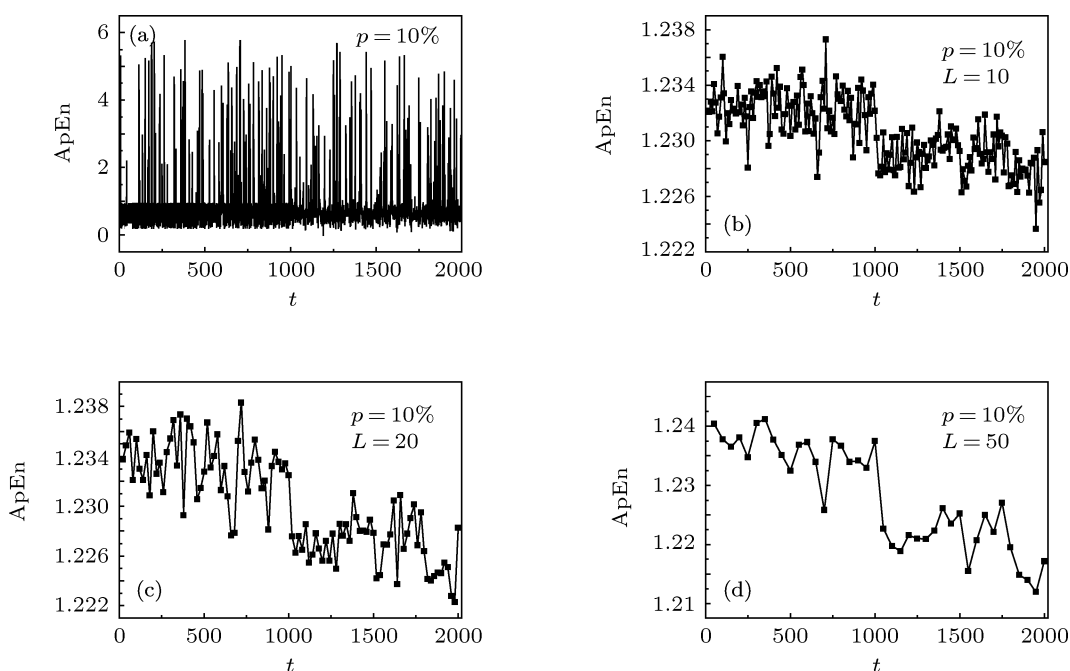


图 5 含尖峰噪声的非线性时间序列 IS1 及其 MC-ApEn 突变检测 (a) 含个数占原序列长度 10% 的尖峰噪声的非线性时间序列 IS1; (b) 滑动子序列长度  $L = 10$ ; (c)  $L = 20$ ; (d)  $L = 50$

### 3.3 高斯白噪声对 MC-ApEn 方法的影响

图 6 为了测试高斯白噪声对 MC-ApEn 方法突变检测的影响程度, 对理想时间序列 IS1 依次添加了信噪比 (signal-to-noise ratio, SNR) 从 5 dB 到 100 dB 不等的高斯白噪声. 图 6 分别给出了 IS1

添加 22 dB, 23 dB, 25 dB, 30 dB 的高斯白噪声后的时间序列演变图. 图 7 展示了滑动子序列长度  $L = 50$  时, 相应含噪序列的 MC-ApEn 检测结果, 从中可以看到, 随着 SNR 逐渐增大, MC-ApEn 检测到的突变点变得更加清晰, 即便对于强度较大的高斯白噪声 (SNR = 23 dB), MC-ApEn 也能够较

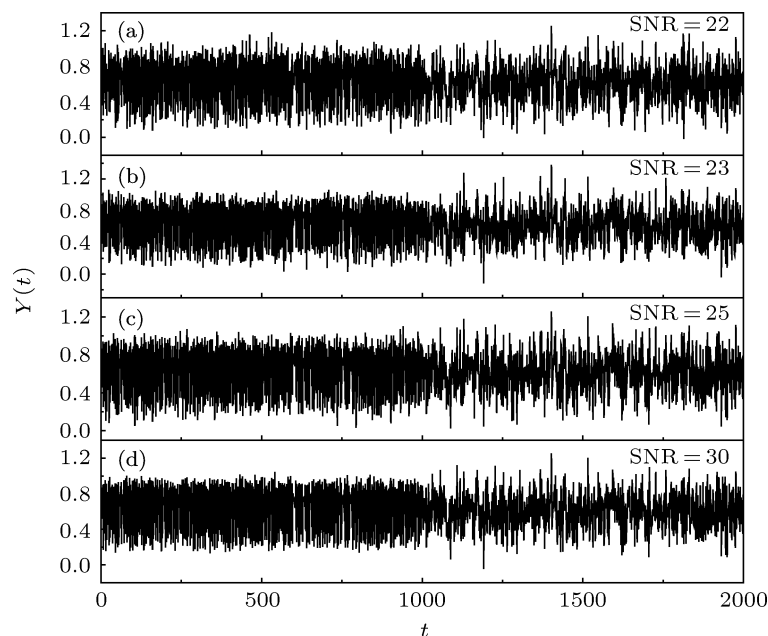


图 6 含高斯白噪声的理想时间序列 (a) 信噪比 SNR = 22; (b) SNR = 23; (c) SNR = 25; (d) SNR = 30

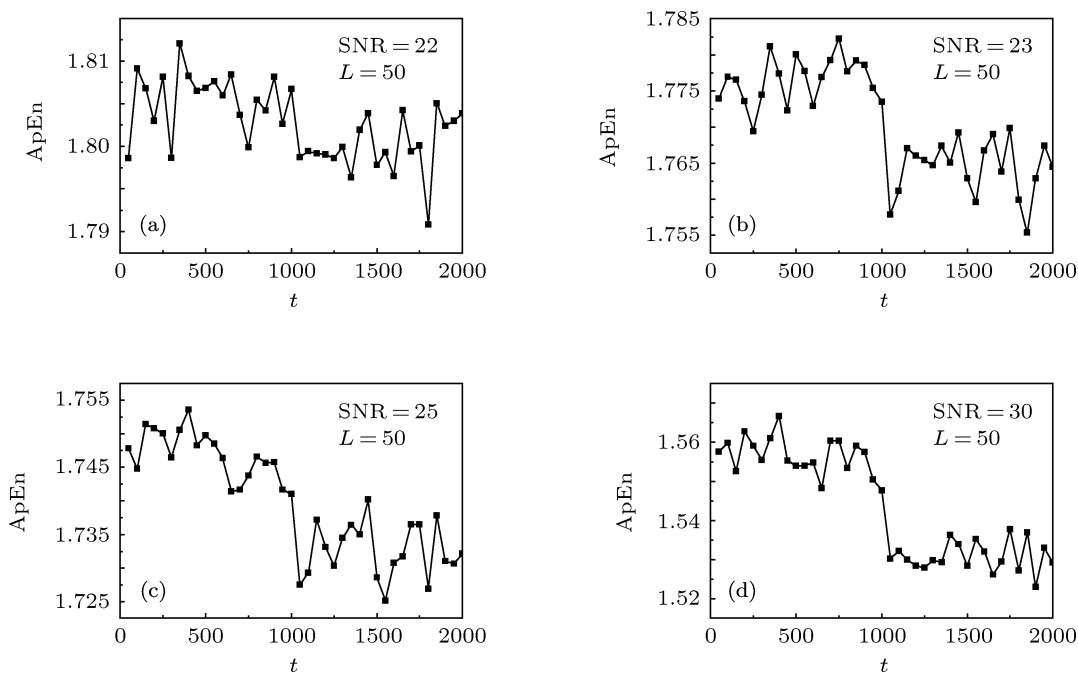


图 7 MC-ApEn 对含不同高斯白噪声的理想序列 IS1 的突变检测, 滑动子序列长度  $L = 50$  (a) 信噪比 SNR = 22; (b) SNR = 23; (c) SNR = 25; (d) SNR = 30

好地识别原始序列中的突变信息. 通过多组数值试验发现,  $\text{SNR} = 22 \text{ dB}$  是可以检测到突变点的  $\text{SNR}$  临界值, 当  $\text{SNR} < 22 \text{ dB}$ ,  $\text{MC-ApEn}$  检测到的突变点位置的清晰程度逐渐递减; 当  $\text{SNR} > 22 \text{ dB}$ , 其检测到的突变位置的清晰程度逐渐递增. 如图 7(b) 所示, 当  $\text{SNR} = 23 \text{ dB}$  时,  $\text{ApEn}$  值随时间的演变趋势以  $t = 1000$  为界呈现着两种趋势, 可以清楚地看到突变的开始时间.

图 8 给出了保持原理想时间序列和所添加高斯白噪声信噪比大小不变, 减小滑动子序列长度到  $L = 20$  时  $\text{MC-ApEn}$  方法的突变检测情况.

当  $\text{SNR} = 22 \text{ dB}$  时,  $\text{ApEn}$  值随时间的演变趋势从一种稳定状态跳跃式地转变到了另一种稳定状态, 在  $t = 1000$  这一时刻, 即时间序列动力学结构发生改变的时刻, 可以清楚地看到突变的发生 (图 8(a)). 为了验证试验结果的可靠性, 数值试验中还取了不同长度的滑动子序列, 发现其  $\text{MC-ApEn}$  突变检测结果基本一致 (图略), 当  $\text{SNR}$  达到  $22 \text{ dB}$  时就可以清楚地检测到突变点的位置,  $\text{SNR}$  值越大, 检测到的突变点越清晰, 对  $\text{MC-ApEn}$  的影响越小. 可检测到突变点的  $\text{SNR}$  临界值的大小可能与样本量的大小有关系.

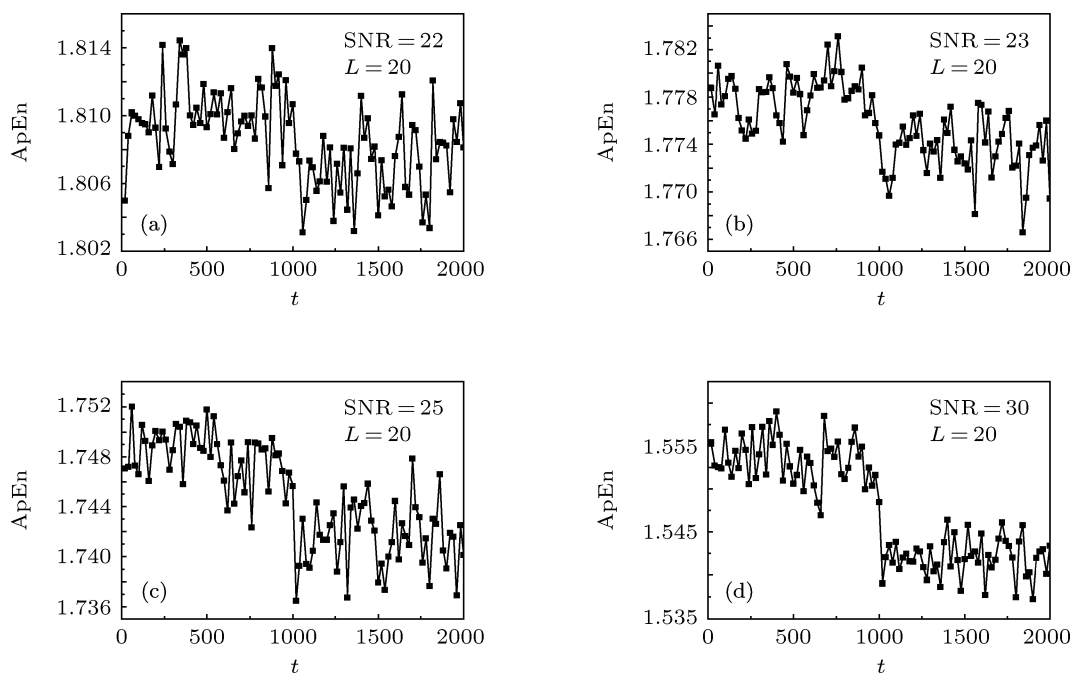


图 8  $\text{MC-ApEn}$  对含不同高斯白噪声的理想序列 IS1 的突变检测, 滑动子序列长度  $L = 20$  (a) 信噪比  $\text{SNR} = 22$ ; (b)  $\text{SNR} = 23$ ; (c)  $\text{SNR} = 25$ ; (d)  $\text{SNR} = 30$

## 4 结果与讨论

本文研究了尖峰噪声、高斯白噪声对  $\text{MC-ApEn}$  突变检测的影响, 发现时间序列中所含尖峰噪声的数目占原序列长度的比例和尖峰噪声的大小对  $\text{MC-ApEn}$  的突变检测结果影响较小, 表明  $\text{MC-ApEn}$  方法对尖峰噪声具有较强的抗噪能力. 当对长度为 2000 的非线性理想时间序列 IS1

加入高斯白噪声, 对于不同长度的滑动子序列, 可检测到突变点的  $\text{SNR}$  临界值均约为  $\text{SNR} = 22 \text{ dB}$ , 这可能与样本量的大小有关系, 这表明  $\text{MC-ApEn}$  的检测结果对高斯白噪声也具有较强的抗干扰能力. 结合本文和此前的研究结果<sup>[4]</sup>可知, 噪声、趋势等时间序列的非平稳性对  $\text{MC-ApEn}$  影响较小, 这为该方法在实际观测资料中的应用提供了坚实的实验基础.

- [1] Shi N 2009 *Statistics forecast of Meteorology* (Beijing: Metrological press) p270–287 (in Chinese) [施能 2009 气象统计预报 (北京: 气象出版社) 第 270–287 页]
- [2] Shi N 2005 *Chin. Phys.* **14** 844
- [3] Dai X G, Wang P, Chou J F 2003 *Chin. Sci. Bull.* **23** 2483
- [4] Jin H M, He W P, Hou W, Zhang D Q 2012 *Acta Phys. Sin.* **61** 069201 (in Chinese) [金红梅, 何文平, 侯威, 章大全 2012 物理学报 **61** 069201]
- [5] Li L J, Wang B, Zhou T J 2007 *Chin. Sci. Bull.* **52** 1820 (in Chinese) [李立娟, 王斌, 周天军 2007 科学通报 **52** 1820]
- [6] Jame M F 2004 *Sti. Nasa. Gov.* 213249
- [7] Hoffmann P, Jaccobi C 2008 *Wiss. Mitteil. Inst. f. Meteor. Univ. Leipzig* **42** 115
- [8] Gong Z Q, Feng G L, Wan S Q, Li J P 2006 *Acta Phys. Sin.* **55** 477 (in Chinese) [龚志强, 封国林, 万仕全, 李建平 2006 物理学报 **55** 477]
- [9] He W P, Feng G L, Dong W J, Li J P 2006 *Acta Phys. Sin.* **55** 969 (in Chinese) [何文平, 封国林, 董文杰, 李建平 2006 物理学报 **55** 969]
- [10] Du J 2002 *J. Appl. Meteor. Sci.* **13** 16 (in Chinese) [杜均 2002 应用气象学报 **13** 16]
- [11] He W P 2008 *Ph. D. Dissertation* (Lanzhou: Lanzhou University) p45 (in Chinese) [何文平 2008 博士学位论文 (兰州: 兰州大学) 第 45 页]
- [12] He W P, Wu Q, Cheng H Y, Zhang W 2011 *Acta Phys. Sin.* **60** 029203 (in Chinese) [何文平, 吴琼, 成海英, 张文 2011 物理学报 **60** 029203]
- [13] Zuo H J 2001 *Meteor. J. Hubei* **4** 29 (in Chinese) [左惠娟 2001 湖北气象 **4** 29]
- [14] Wang C T, Li X G, Wang A H 2010 *Transaction of Beijing Institute of Technology* **30** 474 (in Chinese) [汪春霆, 李修国, 王爱华 2010 北京理工大学学报 **30** 474]
- [15] Thom R 1992 *Structural stability and Morphogenesis* (Sichuan: Sichuan Education Press) p54–59 (in Chinese) [雷内·托姆 1992 结构稳定性与形态发生学 (四川: 四川教育出版社) 第 54–59 页]
- [16] Li S C, Liu F Y, Zhao Z Q 2008 *Clim. Environ. Res.* **13** 31 (in Chinese) [李双成, 刘逢媛, 赵志强 2008 气候与环境研究 **13** 31]
- [17] Feng G L, Dong W J, Gong Z Q, Hou W, Wan S Q, Zhi R 2006 *Nonlinear Theories and Methods on Spatial-Temporal Distribution of the Observational data* (Beijing: the weather publishing company) p5–8 (in Chinese) [封国林, 董文杰, 龚志强, 侯威, 万仕全, 支蓉 2006 观测数据非线性时空分布理论和方法 (北京: 气象出版社) 第 5–8 页]
- [18] Babak A S, Krishnaprasad P S 2004 *EURASIP J. Appl. Signal Processing* **15** 2295
- [19] Abhisek U, Rastko Z 2006 *Electric Power Systems Research* **76** 815
- [20] Liu S D, Chen J, Liu S K 1999 *J. Appl. Meteor. Sci. (Suppl.)* **10** 10 (in Chinese) [刘式达, 陈炯, 刘式适 1999 应用气象学报 (增刊) **10** 10]
- [21] Liu S D, Rong P P, Chen J 2000 *Acta Meteor. Sin.* **58** 111 (in Chinese) [刘式达, 荣平平, 陈炯 2000 气象学报 **58** 111]
- [22] Marko P, Mika R, Mika R 2005 *Control Engineering Laboratory* **26** 1
- [23] Elsner B J, Niu X, Jagger T 2004 *J. Climate* **4** 2652
- [24] Solow A R, Beet A R 2005 *Fish. Oceanogr* **14** 236
- [25] Herron M M, Langway C C 1985 *Greenland Ice Core: Geophys. Geochem. Environ.* **33** 77
- [26] Audrey L M, Christopher W P, Heriberto C 2006 *Ecological Modelling* **95** 72
- [27] István M 2011 *Theor Appl Climatol* **401** 4
- [28] Brian D F, Heriberto C, Christopher W P 2003 *J. Theoretical Biology* **222** 517
- [29] Wan S Q, Feng G L, Dong W J, Li J P 2005 *Acta Phys. Sin.* **54** 5487 (in Chinese) [万仕全, 封国林, 董文杰, 李建平 2005 物理学报 **54** 5487]
- [30] Feng G L, Gong Z Q, Dong W J, Li J P 2005 *Acta Phys. Sin.* **54** 5494 (in Chinese) [封国林, 龚志强, 董文杰, 李建平 2005 物理学报 **54** 5494]
- [31] Feng G L, Hou W, Dong W J 2006 *Acta Phys. Sin.* **55** 962 (in Chinese) [封国林, 侯威, 董文杰 2006 物理学报 **55** 962]
- [32] Zhang W, Gao X Q, Dong W J, Li J P 2006 *Acta Phys. Sin.* **55** 2657 (in Chinese) [张文, 高新全, 董文杰, 李建平 2006 物理学报 **55** 2657]
- [33] Gong Z Q, Feng G L, Dong W J, Li J P 2006 *Acta Phys. Sin.* **55** 3180 (in Chinese) [龚志强, 封国林, 董文杰, 李建平 2006 物理学报 **55** 3180]
- [34] Wan S Q, Feng G L, Dong W J, Li J P, Gao X Q, He W P 2005 *Chin. Phys. Soc. and IOP Publishing Ltd.* **14** 628
- [35] Pinus S M 1991 *Proc. Natl. Acad. Sci. USA* **88** 2297
- [36] Pincus S M, Goldberger A L 1994 *Am. J. Physiol.* **266** H1643
- [37] Pinus S M 1995 *Chaos* **5** 110
- [38] Lei G X, Zeng Y C, Li L 2007 *Technical Acoustics* **26** 111 (in Chinese) [雷国雄, 曾以成, 李凌 2007 声学技术 **26** 111]
- [39] Li Q, Wang T Y, Leng Y G, He G Y, He H L 2007 *Acta Phys. Sin.* **56** 6803 (in Chinese) [李强, 王太勇, 冷永刚, 何改云, 何慧龙 2007 物理学报 **56** 6803]
- [40] Li X F, Yu W L 2008 *Acta Phys. Sin.* **57** 2587 (in Chinese) [刘小峰, 俞文莉 2008 物理学报 **57** 2587]
- [41] Xie Y, Xu J X, Yang H J, Hu S J 2002 *Acta Phys. Sin.* **51** 205 (in Chinese) [谢勇, 徐健学, 杨红军, 胡三觉 2002 物理学报 **51** 205]
- [42] Cao B, Lv X Q, Zeng M, Wang Z M, Huang S S 2006 *Acta Phys. Sin.* **55** 1696 (in Chinese) [曹彪, 吕小青, 曾敏, 王振民, 黄石生 2006 物理学报 **55** 1696]
- [43] 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]
- [44] 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]



# Effect of noises on moving cut data-approximate entropy\*

Jin Hong-Mei<sup>1)</sup> He Wen-Ping<sup>2)†</sup> Zhang Wen<sup>1)</sup> Feng Ai-Xia<sup>1)</sup> Hou Wei<sup>2)</sup>

1) (*Department of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China*)

2) (*National Climate Center, China Meteorological Administration, Beijing 100081, China*)

(Received 9 September 2011; revised manuscript received 21 October 2011)

## Abstract

Affected by some factors such as external forcing and the measurement errors of instrument itself, observational data often contain noises, disturbances and some other false information. To solve this problem, the effects of different noises on moving cut data-approximate entropy (MC-ApEn) are investigated in this paper. The results indicate that MC-ApEn is little affected by random spikes and Gaussian white noise, which means that the MC-ApEn method has strong anti-noise ability. The results provide an essential experimental basis for the wide applications of the present method to observational data.

**Keywords:** moving cut data-approximate entropy, noises, abrupt change detection

**PACS:** 92.60.Wc

---

\* Project supported by the National Basic Research Program of China (Grant No. 2012CB955203), the National Natural Science Foundation of China (Grant Nos. 40930952, 41175067 and 40905034), the Special Scientific Research Fund of Public Welfare Profession of China (Grant Nos. GYHY201106016 and GYHY201106015), and the State Key Program of Science and Technology of China (Grant Nos. 2007BAC29B01, 2009BAC51B04).

† E-mail: wenping\_he @163.com.