

基于 Fisher 信息理论的突变检测新方法*

蔡舒平¹⁾²⁾ 戴理^{1)†} 倪溟¹⁾

1) (江苏大学电气信息工程学院, 镇江 212013)

2) (江苏大学, 机械工业设施农业测控技术与装备重点实验室, 镇江 212013)

(2013年1月8日收到; 2013年6月14日收到修改稿)

自然界许多复杂系统的动力学结构因其受外强迫的作用, 常常会发生与人类生活密切相关的突变. 对突变做出准确预测并采取相应对策的有效途径就是检测. Fisher 信息 (FI) 是一种能够敏锐捕捉并刻画系统变量的概率密度分布发生微小变化的量; 而当系统的动力学结构发生突变时, 来自突变前后两种不同动力学结构的数据其概率密度分布肯定会发生不同程度的改变. 基于此, 本文将它用于检测和识别系统动力学结构所发生的突变, 以一种全新的思路来解决突变检测问题. 首先通过理想数据序列展示了使用 FI 来表征各种不同类型理想信号突变的能力. 进而应用于中国气象局国家气象信息中心“西北地区地面气候资料日值数据集”兰州站 1960—2008 年日平均气温资料的实际检测中, 结果表明与已有历史记载相符合, 进一步验证了该方法的有效性和实用性.

关键词: Fisher 信息, 动力学结构, 突变检测, 概率密度分布

PACS: 92.60.Wc

DOI: 10.7498/aps.62.189204

1 引言

由于自然界本身和人类活动的加剧, 许多复杂系统如生态系统、气象系统等常常表现为从一个相对稳定状态翻转到另一个相对稳定状态的现象. 生态系统中一个典型的例子是在一个浅水湖泊中, 由于磷的流入, 生态系统有可能从一个寡营养状态翻转到一个富营养状态而导致藻类过度生长, 进而由于缺乏鱼类生存所需要的氧气而使生物的多样性减少和水质量恶化, 随之而来的是系统的大部分功能和效用丧失. 气候系统中如全球气候在温暖而潮湿的间冰期和凉爽而干燥的冰川期之间快速地翻转等等. 近年来, 对复杂系统状态翻转的研究因其对现代社会的影响而引起了人们广泛的关注. 科学工作者通过深入研究后发现, 所有这些系统的状态翻转具有一个共同的特征, 那就是其内在的动力学结构随着外强迫作用的变化而发生了突变. 因此, 有关动力学结构突变检测的研究近年来受到了普遍重视, 其中气候系统因其本身的复杂性及与人类

日常生活的密切相关性, 使得在这方面的研究显得尤为迫切.

封国林等^[1]从气候突变的检测和研究、气候资料信息的提取和分离、气候系统复杂性的研究、气候系统动力学结构特征及极端气候事件的检测等 5 个方面阐述了近年来在我国气候突变检测技术研究方面所取得的新进展. 传统的检测方法如滑动 t -检验, Cramer 法, Yamamoto 信噪比法以及 Mann-Kendall (M-K) 法等, 在突变检测的早期得到了广泛应用, 然而由于这些传统方法自身固有的一些缺陷使得在实际检测中所得结果均存在不同程度的突变点漂移, 在此情况下, 一些新的检测方法如分割方法 (即 BG 算法) 应运而生, 龚志强等^[2]对此进行了实际检验和应用, 并得到了一些有益的结果. 近年来, 国内学者如封国林、龚志强、何文平、王启光和侯威等人将非线性科学的一些研究成果应用于动力学结构突变检测中, 分别提出了动力学相关因子指数法^[3]、动力学指数分割算法^[4]、滑动去趋势波动算法^[5]、排列熵算法^[6]、近似熵算法^[7]、滑动移除近似熵算法^[8]、重标极差法^[9]等

* 国家自然科学基金 (批准号: 61273154/F030116, 51077066/E070303) 和江苏高校优势学科建设工程 (批准号: 苏政办发 (2011)6 号) 资助的课题.

† 通讯作者. E-mail: spcai@mail.ujs.edu.cn

检测方法. 成海英等^[10]将偏度系数和峰度系数引入时间序列的突变检测中, 提出了基于概率密度分布型变化的突变检测方法. 此外, 还有学者将小波分析、粒子滤波等方法用于突变检测中^[11,12]. 在与气候突变有关的其他研究领域, 龚志强等^[13]把希尔伯特-黄变换 (Hilbert-Huang transform, HHT) 应用于气候资料信息的分离和提取中, 且与小波变换进行了对比分析, 并针对它们各自的缺点提出了可能改进的措施. 与此同时, 封国林等^[14]介绍了一种能判断不同时间序列之间动力学异同性的条件熵算法, 讨论了其物理意义, 并以长三角地区为例, 验证了这一算法的有效性和可靠性. 万仕全等^[15]利用 Q 指数来识别时间序列之间动力结构的异同性, 并利用该方法分析了几种典型代用资料的动力学结构, 得到了一些有益的结果. 而在极端气候事件的检测方面, 封国林等^[16]利用固定阈值法研究了 Lorenz 系统极端事件序列的长程相关性特征, 并在此基础上分析了实际的气象观察资料, 得到了一些有参考价值的结论. 所有这些研究工作极大地丰富了突变检测方法, 拓宽了突变检测的途径. 然而, 由于所研究问题的复杂性, 不可能存在一种万能的方法一劳永逸地解决所有问题, 因此, 人们从未停止过对突变检测新思想、新方法的探寻, 尤其是分析那些已被噪声污染且稀疏的实际野外观察数据方法的研究.

实际中往往会面临单变量的一维时间序列, 很显然, 系统的非线性突变特征信息就蕴藏在这些序列中, 如何有效地从中提取这类信息成为一个非常关键的问题.

Fisher 信息 (FI) 恰好为我们提供了表征这样一类突变特征信息的可能性. 原因是任何类型的数据和模型本质上都可以转换为信息而不管最初的学科是什么^[17]. 不像系统信息的其他测量方法, Fisher 信息提供了一种通过监测系统变量从而监测系统的状态和状态突变的方法^[18], 这种检测状态和状态突变的能力使得识别发生在系统中的最根本的变化成为可能. 事实上, Frieden^[19,20]使用该理论推导出了许多基本的物理学、热力学、人口遗传学方程. 近年来, 生态学家广泛使用 Fisher 信息理论来研究生态领域的问题, Fath 等^[18]和 Mayer 等^[21]使用 Fisher 信息来作为一种复杂生态系统动态秩序的度量. Mayer 等^[21]和 Karunanithi 等^[22]提议用来作为一种量化的指标来检测和评估生态

系统状态的翻转以及作为一种可持续性标准^[23]. Fisher 信息还被用于研究包含一个多间隔食物链稳定性问题的模型系统^[24-26]以及用于可持续环境管理中的动态模型系统的优化控制问题^[26,27]. Rico-Ramirez 等^[28]把 Fisher 信息应用于医疗化工领域的优化控制问题. 在了解了以上为研究生态系统尝试使用 Fisher 信息理论所取得的成果后, 本文试图用它来解决动力学结构的突变检测问题.

2 方法

2.1 Fisher 信息

统计学家 Ronald Fisher 在 1922 年提出了一种度量不确定性的方法, 即现在称之为费歇信息^[29]. 然而, 由于信息只能从代表行动力系统稳定性的数据中获得, 使得 Fisher 信息尤其被认为是一种对现行动力系统秩序的度量. 许多文献已经展示了使用 Fisher 信息来度量生态系统的秩序的能力^[21,22,30]. 一个变量单次测量的 Fisher 信息 I 计算如下:

$$I = \int \frac{ds}{P(s)} \left(\frac{dP(s)}{ds} \right)^2, \quad (1)$$

这里 $P(s)$ 为概率密度函数 (PDF), s 是一个状态变量.

对于一个稳定的动力系统而言, 系统变量的概率密度分布也较为稳定, 而当动力学结构发生变化后, 系统变量的概率密度分布也会发生不同程度的变化. 由于 Fisher 信息是基于概率密度函数的导数, 因此它反映的是一种局部特性, 这使得它能敏锐捕捉系统变量的概率密度分布发生微小的变化. 鉴于此, 本文用它检测和识别蕴藏在时间序列中的突变特征, 为突变检测提供一种新的途径.

实际计算中, 为了避免因除以较小的 $P(s)$ 值而带来的计算误差, 我们用一个变量的强度来代替 (1) 式中概率密度函数 $P(s)$, 即令 $q^2(s) = P(s)$, 于是 (1) 式变为

$$I = 4 \int ds \left[\frac{dq(s)}{ds} \right]^2. \quad (2)$$

为了使 (2) 式适合于分析和检测实际系统的观察序列, 我们用求和来代替积分以及用差分来代替微分, 即用 $\Delta S = S_i - S_{i+1}$ 和 $\Delta q = q_i - q_{i+1}$ 来分别

代替 ds 和 dq , 从而 Fisher 信息能被近似为

$$I \approx 4 \sum_{i=1}^n \left[\frac{q_i - q_{i+1}}{s_i - s_{i+1}} \right]^2 (s_i - s_{i+1}), \quad (3)$$

(3) 式中, S_i 代表系统的某一个特定状态, 即 S_1 代表状态 1, S_2 代表状态 2 等等. 于是: $S_i - S_{i+1} = 1$, 则计算 Fisher 信息的最终表达式为

$$I \approx 4 \sum [q_i - q_{i+1}]^2, \quad (4)$$

表达式 (4) 将会用于我们以下所有的计算.

2.2 Fisher 信息的计算

检测系统动力学结构突变的第一步是获取描述系统状态随时间变化的观察数据序列集, 然后在数据集上定义一个滑动时间窗口, 窗口宽度取决于可得到的数据量和系统的行为. 按照经验, 窗口宽度至少应该包含 8 个数据点以确保窗口中的每一点不会过度影响整个计算. 窗口的滑动因子应小于窗口宽度, 这样相邻的窗口间就会出现一个重叠, 这样做的好处是能够捕获有可能已经延伸到该窗口边界之外的动力学结构的突变.

即若设观察数据序列集为 $D = \{d(k), k = 1, \dots, N\}$, 其中 N 为序列总长度, 窗宽为 $w \in N$, 滑动因子为 $\delta \in N$, 用数学公式可把滑动窗 W 描述如下:

$$W(m, w, \delta) = \{d(k), k = 1 + m * \delta, \dots, w + m * \delta\}, \quad (5)$$

式中 $m = 1, 2, \dots, M$, M 为窗口个数, $M = (N - w) / \delta$.

在把时间序列转变为一个重叠窗口序列以后, 接下来是将每个滑动窗口划分成如下 I 个区间:

$$W(m, w, \delta) = \bigcup_{i=1}^I Z_i, \quad (6)$$

式中 $\{Z_i = [S_{i-1}, S_i), i = 1, 2, \dots, I\}$, 互不相交.

$$S_0 < S_1 < S_2 < \dots < S_I, \quad (7)$$

$$S_0 = \min[W(m, w, \delta)] \\ = \min[\{d(k), k = 1 + m * \delta, \dots, w + m * \delta\}], \quad (8)$$

$$S_I = \max[W(m, w, \delta)] \\ = \max[\{d(k), k = 1 + m * \delta, \dots, w + m * \delta\}], \quad (9)$$

则数据 $d(k) \in W(m, w, \delta)$ 落入区间 Z_i 的概率 $P(Z_i)$ 等于 $d(k) \in W(m, w, \delta)$ 落入区间 Z_i 的数目与 $W(m, w, \delta)$ 中总数据个数 w 之比.

即使用如下公式计算窗口中相应于 Z_i 区间上的概率密度分布 $P(Z_i)$:

$$P(Z_i) = \frac{\text{落入区间 } Z_i \text{ 的数据个数}}{\text{时间窗口中的总数据个数}}. \quad (10)$$

需要说明的是, 在上述计算中, 窗口中区间个数 I 的取值越大, 区间越窄, 相同条件下, 所计算出的窗口概率密度分布比较平缓, 从而 Fisher 信息值较小, 这使得其对突变的反应不够敏感. 反之, I 取值越小, 区间越宽, 相同条件下, 所计算出的窗口概率密度分布比较陡峭, 从而 Fisher 信息值较大, 而这又容易受到噪声的干扰而引起误判. 一般情况下, I 的取值应根据待分析数据的最大波动范围而定.

接下来, 为每个状态计算 $q(Z_i)$ ($q(Z_i) = \sqrt{p(Z_i)}$) 以及使用 (4) 式为每个时间窗口估计 FI. 值得注意的是计算的 FI 值被分配给每个时间窗口的末端位置.

归纳起来, 计算 FI 的基本步骤如下: 1) 把时间序列数据划分为一连串重叠的时间窗口, 2) 把每个时间窗口划分成长度相等的若干个区间, 清点落入每个区间的数据个数; 3) 计算相应于该窗口的概率密度分布; 4) 对每个时间窗口使用 Fisher 信息实用计算公式从概率密度函数中计算 Fisher 信息.

3 结果

3.1 FI 方法在理想时间序列中的性能测试

在所有的突变测试函数中, 正余玄复合函数因其构造简单, 表现形式多样, 物理意义明确而得到了广泛的应用. 为了便于对照, 本文亦选择它为测试函数来检验新方法的有效性.

$$y(t) = \begin{cases} 2 \sin(0.2t) + 1 & 1 \leq t \leq 1000 \\ 1.5 \sin(0.2t) + 2 \cos(0.5t) - 0.2 & 1000 < t \leq 2000 \end{cases}, \quad (11)$$

(11) 式表明, 上述时间序列来自两种不同的动力学结构, 在区间 $[1, 1000]$ 上, 由单一的正玄函数所组成, 在区间 $(1000, 2000]$ 上, 由正余玄函数复合而成. 很显然, 在 $t = 1000$ 时, 系统从一种形式的动力学结构突变为另一种形式的动力学结构. 图 1(a) 给出了该理想信号序列波形. 为了进一步测试新方法的抗噪能力, 在原理理想信号的基础上, 同时构造如下两种形式的加噪信号. 图 1(b) 为加入信噪比为 30 dB 高斯白噪声后信号波形, 图 1(c) 为随机加入

12 个尖峰噪声后的信号波形, 其大小在 2—5 之间变化.

图 2 为采用 Fisher 信息方法在滑动窗口宽度 $w = 20$, 滑动步长 $\delta = 1$ 情况下, 对 3 种信号进行突变检测所得结果.

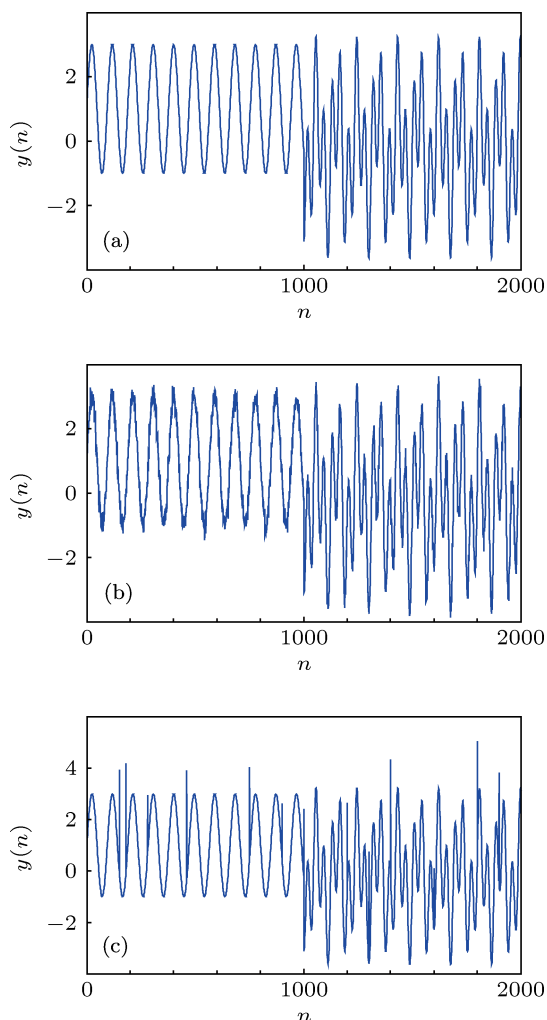


图 1 测试信号 (a) 理想信号; (b) 加白噪声的理想信号; (c) 加尖峰噪声的理想信号

从图 2(a) 所呈现的检测结果可以看出, 在区间 $[1, 1000]$ 上, FI 值表现出疏密有致的有规律波动, 而在区间 $(1000, 2000]$ 上, FI 值表现出疏密不等的无规律变化, 在 $n = 1000$ 处, 出现了一个幅度为 4(最大的 FI 值) 的尖峰. 而由 Fisher 信息的定义可知, Fisher 信息是基于序列概率密度分布的导数, 这意味着在上述两个区间上, 序列的概率密度分布肯定不同, 这种不同意味着在整个区间 $[1, 2000]$ 上, 序列的概率密度分布在 $n = 1000$ 处会出现一次“跃变”. 我们知道, 跃变函数的一阶导数刚好是冲击函数, 这就不难理解为什么在 $n = 1000$ 处会出现一个尖峰冲击. 由此看来, 基于 Fisher 信息的突变检测

方法深刻地揭示了突变发生的机理, 如实地反映了突变发生的整个物理进程. 另一方面, Fisher 信息又是一种对信号稳定程度的一种度量, 从这一点来讲, 突变前序列的稳定性要优于突变后.

从图 2(b) 的测试结果来看, 该方法具有一定的抗噪能力. 和图 2(a) 相比较, 突变前后因其噪声的加入使得在一定程度上会影响序列的概率密度分布, 又因为 Fisher 信息是基于概率密度分布的导数, 反映的是一种局部特性, 因此, 它对任何影响序列概率密度分布发生微小变化的因素较敏感, 其结果是突变前后 FI 值较图 2(a) 出现了一定程度的波动, 尤其是突变前, FI 值呈现出一种毫无规律的小幅尖峰震荡. 但该方法对噪声的这种敏感性仅限于此, 丝毫不会影响突变检测的最终结果, 这一点从图 2(b) 的 FI 分布趋势与图 2(a) 中完全一致就说明了这一点.

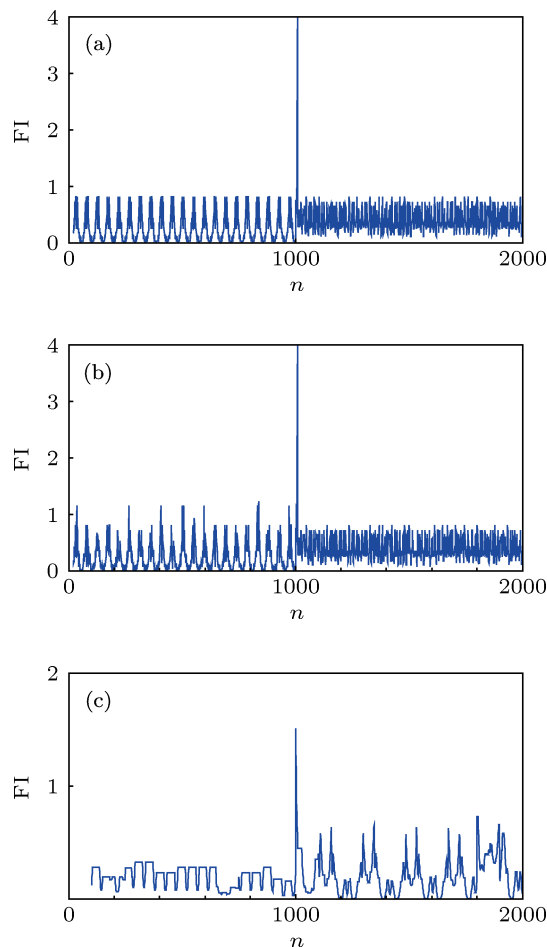


图 2 三种信号的 FI 突变检测 (a) 理想信号; (b) 加白噪声信号; (c) 加尖峰噪声信号

图 2(c) 展示的该突变检测方法抗尖峰噪声的能力. 从图中可以看到, 信号中的尖峰噪声仅仅

会对突变前后 FI 值造成轻微的影响, 并使突变点 $n = 1000$ 处 FI 值的跃变幅度有所减弱, 尽管如此, 从图中仍可以清楚地识别出突变点. 有趣的是尖峰噪声对突变前后 FI 值的影响恰好与白噪声相反, 这是由 FI 值的计算方法所决定的. FI 是以窗口为单位进行计算的, 单个的尖峰噪声对整个窗口的影响被平均掉了, 这样看来, 窗口中的单一信号比复合信号能更有效地抑制尖峰噪声的影响, 这也就是图 2(c) 中所看到的突变前 FI 的变化要比突变后 FI 的变化更平稳的原因.

在以上测试过程中, 子序列长度即窗口宽度固定为一个特定的值 ($w = 20$), 得出了完全相同的突变检测结果. 由此引出了另外一个问题, 那就是对于任意的子序列长度检测结果是否仍然完全相同

呢? 为此, 需要考察不同子序列长度下该方法对同一信号的突变检测结果.

图 3 给出了子序列长度分别取 10, 50, 100 和 200 时, 对理想信号的突变检测结果.

从图 3 不难看出, 四种不同情况下, FI 的演化规律完全类似, 即以突变点为界, FI 划分为两个明显不同的演化阶段, 在两个不同阶段的衔接处即为突变点, 此处的 FI 值表现为一个急促向上的跳变, 跳变幅度及突变前后 FI 波动幅度随着子序列长度的增加有所减弱, 这是因为子序列取得越大, 子序列中的单个数据点对由整个子序列计算出来的 FI 值的影响愈小, FI 值愈稳定, 但这丝毫不会影响最终的检测结果. 不同子序列长度下检测结果的一致性, 表明了该方法的检测结果不依赖于子序列长度.

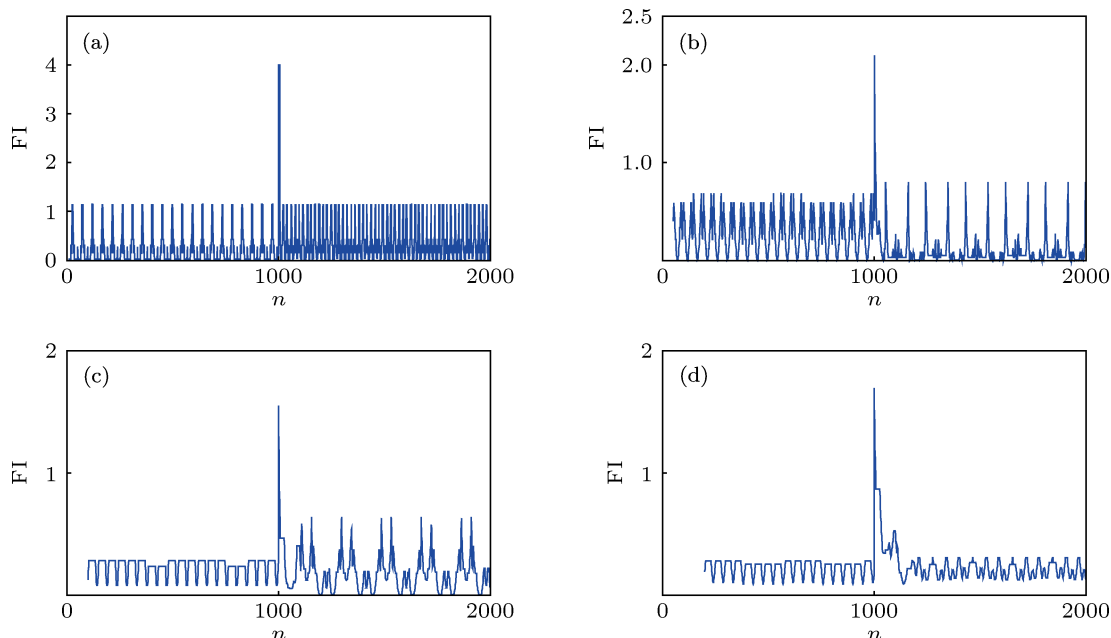


图 3 理想序列 FI 突变检测 (a) $w = 10$; (b) $w = 50$; (c) $w = 100$; (d) $w = 200$

一种方法之所以被提出必有其不同于传统方法的独特之处. 为此, 有必要把该方法与一些传统方法进行比较, 以检验其有效性. 仍以上述线性理想序列为例, 分别基于滑动 t 检验和 Yamamoto 法对该序列进行检测, 其中子序列长度取 200, 滑动步长取 1, 图 4 所示为这两种方法的检测结果.

在图 4(b) 和 (c) 中, 对应突变点 $t = 1000$, 滑动 t 检验和 Yamamoto 法均给出了一个明显的突变区域 (800—1200 之间), 在突变点附近, 无论 t 值还是

信噪比均出现了最大值, 且超过了 0.01 显著性水平. 然而, 若想进一步准确定位突变点则存在一定的困难. 而由 FI 法在此种情况下所给出的检测结果中 (见图 3(d)), 对突变点的检测能够做到准确定位.

进一步考察当子序列的长度取不同值时, 滑动 t 检验和 Yamamoto 法的检测结果是否仍然保持一致. 图 5 给出了当子序列长度取 20, 滑动步长取 1 时, 滑动 t 检验和 Yamamoto 法对此的检测结果.

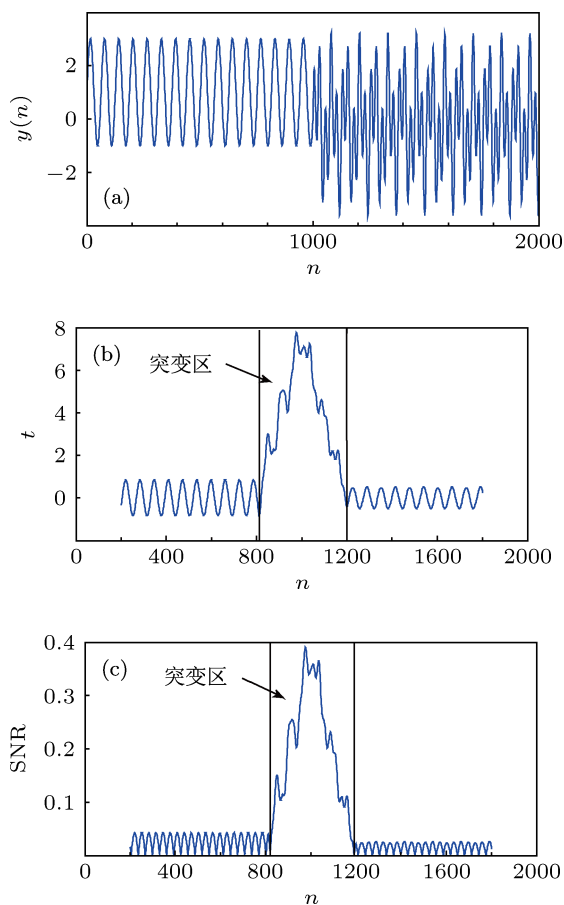


图4 子序列长度取 200 时检测结果 (a) 理想序列; (b) 滑动 t 检验; (c) Yamamoto 法

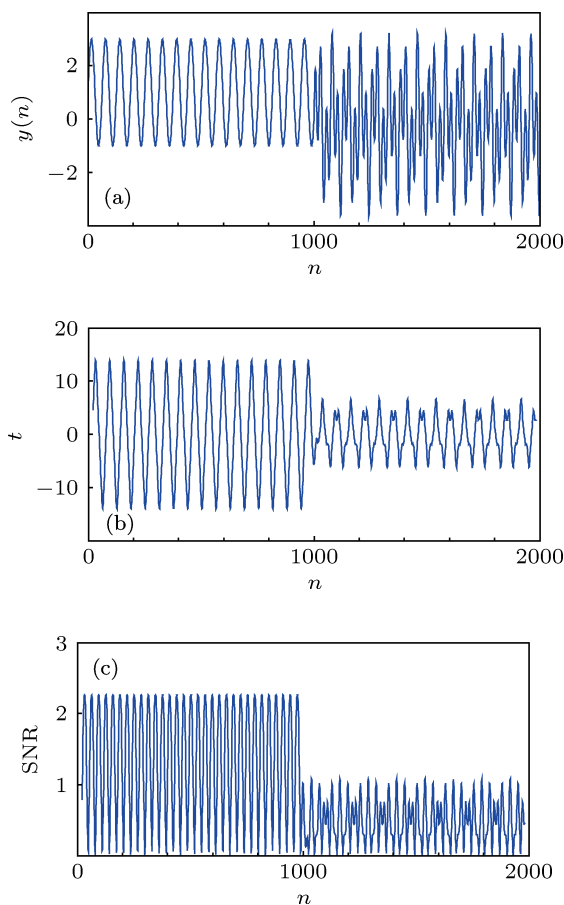


图5 子序列长度取 20 时检测结果 (a) 理想序列; (b) 滑动 t 检验; (c) Yamamoto 法

从图 5 中可以看出, 尽管滑动 t 检验和 Yamamoto 法在此情况下仍能对突变点做出一定的反应, 但要对检测结果做出合理的解释已变得不可能; 而 FI 法在此种情况下对突变点的检测结果 (见图 2(a)) 未发生任何改变. 改变子序列的长度 (如取 50, 100) 会得出类似的结果 (图略). 同时对在本质上也是基于均值思想的传统检测方法——Cramer 法也进行了类似的实验, 发现存在同样的问题. 而另一种传统检测方法——M-K 法常常被用来检测序列中的趋势突变, 若时间序列中不存在明显的趋势突变, 则对此序列中动力学结构突变的检测就无从谈起. 综合对比以上检测实验结果和分析可知, FI 法无论在对动力学结构突变检测的准确性还是稳定性方面较传统方法均具有明显的优势.

3.2 FI 方法在实际气象观察资料中的应用

气候系统具有典型的时空多尺度、结构多层次、本质非线性特征, 不同层次之间关系及其相互作用十分复杂^[31]. 自从 Lorenz^[32,33] 和 Charney

等^[34] 从理论上论证了气候突变存在的可能性, 有关气候突变的研究得以广泛开展. 按照控制气候系统的动力学方程是否发生变化将气候突变分为动力学结构突变和气候状态变量在统计意义上发生了显著变化的突变, 如均值突变、趋势突变等. 传统的突变检测方法如滑动 t - 检验, Cramer 法, Yamamoto 法以及 M-K 法等均是基于状态变量在统计意义上的显著变化来判断突变的发生, 其检测结果均具有多尺度特征. 然而系统的动力学结构突变发生的时间与特定的时间尺度没有必然的关联, 因此基于系统状态变量在统计意义上的显著变化来判断突变的传统检测方法不能够对系统的动力学结构突变进行有效识别^[5]. 近年来, 一些新的突变检测方法如近似熵 (ApEn) 算法, Q 算法等用于检测气候系统的动力学结构突变均取得了一定的效果. 然而随着人们对复杂气候系统愈来愈多的了解和掌握, 一些原先未曾认识到的问题逐渐暴露了出来, 人们迫切需要探寻新的突变检测方法来应对不断涌现的新情况、新问题. 因此, 有关气候系统的

动力学结构突变检测问题一直处于持续的研究和探索中.

气候系统中的两大气象要素温度和降水是两个最能体现气候动力学结构发生突变的量,因此,本文采用两大气象要素之一的温度来检验本文所提方法在检测气候系统动力学结构突变中的有效性.文中所用数据来自中国气象局国家气象信息中心“西北地区地面气候资料日值数据集”兰州站1960—2008年日平均气温资料.图6给出了该观察站1960—2008年逐日平均气温的FI检测结果.

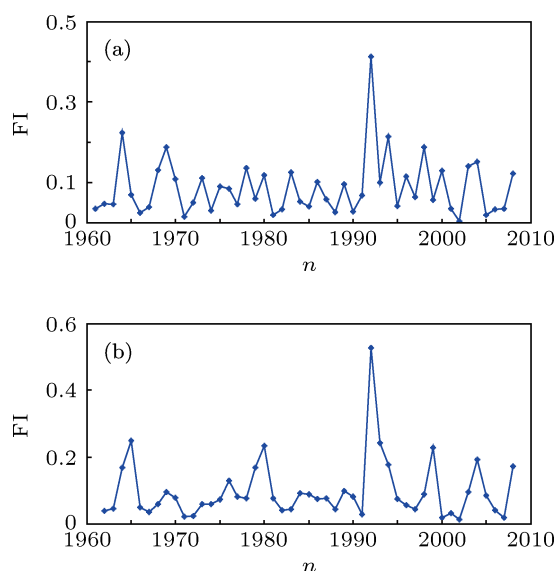


图6 兰州站逐日平均温度序列FI变化情况 (a)子序列长度取365 d(1年);(b)子序列长度取730天(2年)

从图中可以看出,20世纪90年代初该观察站日均温度发生了一次明显的突变,这与已有该站所在地区历史上日均温度突变年份的研究结果相一致^[35,36].另一方面,从图6中子序列长度取1年和2年时,所检测到的突变点相一致来看,进一步表明该方法所得检测结果不依赖于子序列的长度.

4 讨论与结论

本文引入常被用来检测生态系统稳定性和生

态状态翻转的指标——Fisher信息来识别和检测系统动力学结构突变,是因为本文紧扣Fisher信息是基于概率密度分布的导数这一基本原则.基于不同动力学性质的数据其概率密度分布大小不同,而具有相同动力学性质的数据其概率密度分布差异不大.因此,当系统的动力学结构发生突变时,基于突变前后两种不同动力学结构的数据其概率密度分布肯定会发生不同程度的改变.这种细微的改变恰好通过Fisher信息加以定格和放大.因此,理论上Fisher信息是对动力学结构突变的一种很好的测量.

本文首先将它用于分析理想时间序列,结果表明该方法能够准确定位突变点,并且具有良好的抗噪性能,同时,检测结果不依赖于子序列长度的选择.紧接着,比较了新方法与传统检测方法的检测效果,彰显了新方法的优越性.最后,应用于“西北地区地面气候资料日值数据集”兰州站1960—2008年(49年)日平均气温资料,发现检测结果与已有该站所在地区历史上日均温度突变年份的研究结果完全符合,进一步展示了该方法的有效性和实用性,表明该方法在动力学结构突变检测中的应用前景和实用价值.

不像在研究气候时空分布时那样需要做时空插值,由于受观察资料的限制,边界效应是一个不容忽视的问题.而本方法是按照所给时段内的观察数据,采用滑动窗技术逐点计算FI值,并把计算结果分配给窗口的末端,这就意味着对突变点的真正检测是从观察数据的第 L 点开始的(L 为窗口宽度即子序列长度),对前 $L-1$ 点上的突变检测无法进行,这是基于滑动窗技术的此类检测方法的一个共性问题(包括滑动 t 检验, Yamamoto法和Crammer法等基于子序列均值的传统检测方法).

本方法可用于检测任何时间尺度的观察数据序列,观察数据的时间尺度仅仅影响检测的精度即分辨率.

- [1] Feng G L, Gong Z Q, Zhi R 2008 *Acta Metrol. Sin.* **66** 892 (in Chinese) [封国林, 龚志强, 支蓉 2008 气象学报 **66** 892]
- [2] 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]
- [3] Feng G L, Gong Z Q, Zhi R 2006 *Nonlinear Theories and Methods on Spatial Temporal Distribution of the Observational Data* (Beijing: China Meteorological Press) pp1-100 (in Chinese) [封国林, 龚志强,

支蓉 2006 观察数据时空分布的非线性理论和方法(北京: 中国气象出版社)第1—100页]

- [4] Gong Z Q, Feng G L, Dong W J 2006 *Acta Phys. Sin.* **55** 3180 (in Chinese) [龚志强, 封国林, 董文杰 2006 物理学报 **55** 3180]
- [5] 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]

- [6] Hou W, Feng G L, Dong W J, Li J P 2006 *Acta Phys. Sin.* **55** 2663 (in Chinese) [侯威, 封国林, 董文杰, 李建平 2006 物理学报 **55** 2663]
- [7] 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]
- [8] 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]
- [9] 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]
- [10] Cheng H Y, He W P, He T, Wu Q 2012 *Acta Phys. Sin.* **61** 039201 (in Chinese) [成海英, 何文平, 何涛, 吴琼 2012 物理学报 **61** 039201]
- [11] Babak A S, Krishnaprasad P S 2004 *EURASIP J. Appl. Signal Processing* **15** 2295
- [12] Abhisek U, Rastko Z 2006 *Electric Power System Research* **76** 815
- [13] Gong Z Q, Zou M W, Gao X Q, Dong W J 2005 *Acta Phys. Sin.* **54** 3947 (in Chinese) [龚志强, 邹明玮, 高新全, 董文杰 2005 物理学报 **54** 3947]
- [14] 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]
- [15] 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]
- [16] Feng G L, Wang Q G, Hou W, Gong Z Q, Zhi R 2009 *Acta Phys. Sin.* **58** 2853 (in Chinese) [封国林, 王启光, 侯威, 龚志强, 支蓉 2009 物理学报 **58** 2853]
- [17] Cabezas H, Fath B D 2002 *Fluid Phase Equilibria* **194–197** 3
- [18] Fath B, Cabezas H, Pawlowski C W 2003 *J. Theor. Biol.* **222–224** 517
- [19] Frieden B R 1998 *Physics from Fisher Information: A Unification* (Cambridge, UK: Cambridge University Press)
- [20] Frieden B R 2001 *Probability, Statistical Optics and Data Testing* (3th Ed.) (New York: Springer-Verlag)
- [21] Mayer A L, Pawlowski C W, Cabezas H 2006 *Ecol. Model.* **195** 72
- [22] Karunanithi A T, Cabezas H, Frieden B R, Pawlowski C W 2008 *Ecol. Soc.* **13** 22
- [23] Fath B D, Cabezas H 2004 *Ecol. Model.* **174** 25
- [24] Cabezas H, Pawlowski C W, Mayer A L, Hoagland N T 2005 *J. Clean. Product.* **13** 455
- [25] Cabezas H, Pawlowski C W, Mayer A L, Hoagland N T 2005 *Resources, Conservation and Recycling* **44** 279
- [26] Shastri Y, Diwekar U, Cabezas H 2008 *Environ. Sci. Technol.* **42** 5322
- [27] Shastri Y, Diwekar U, Cabezas H, Williamson J 2008 *Environ. Sci. Technol.* **42** 6710
- [28] Rico-Ramirez V R, Qunitana-Hernandez P A, Ortiz-Cruz J A, Hernandez-Castro S 2008 *Computer-Aided Chem. Engin.* **25** 1155
- [29] Fisher R A 1922 *Philosoph. Trans. Roy. Soc. London* **222** 309
- [30] Eason T, Cabezas H 2012 *J. Environ. Manage.* **94** 41
- [31] Li Y Q 2001 *Plateau Meteorol.* **20** 88 (in Chinese) [李跃清 2001 高原气象 **20** 88]
- [32] Lorenz E N 1963 *J. Atmos.* **20** 130
- [33] Lorenz E N 1976 *Quat. Res.* **6** 495
- [34] Charney J G, Devore J G 1979 *J. Atmos. Sci.* **36** 1205
- [35] Wei F Y, Cao H X 1995 *Sci. Atmos. Sin.* **19** 140 (in Chinese) [魏凤英, 曹鸿兴 1995 大气科学 **19** 140]
- [36] Zhao F F, Xu Z X 2006 *Sci. Atmos. Sin.* **64** 246 (in Chinese) [赵芳芳, 徐宗学 2006 气象学报 **64** 246]

A new approach to abrupt change detection based on Fisher information*

Cai Shu-Ping¹⁾²⁾ Dai Li^{1)†} Ni Hao¹⁾

1) (College of Electrical Information and Engineering, Jiangsu University, Zhenjiang 212013, China)

2) (Key Laboratory of Facility Agriculture Measurement and Control Technology and Equipment of Machinery Industry,

Jiangsu University, Zhenjiang 212013, China)

(Received 8 January 2013; revised manuscript received 14 June 2013)

Abstract

The dynamics structure of many complex system there often happens a suddenly change due to the effect of outside force in nature. This abrupt change is closely related to human life. In order to make an accurate prediction of abrupt change and take corresponding measures, abrupt change should be detected effectively. Fisher information (FI) can keenly catch and characterize a small change in probability density distribution (PDD) of a system variable. While there is a change in dynamic structure of a system, the PDD of the system variable will have some changes correspondingly. In view of this, in this paper we describe in detail that FI is used to detect and recognize the abrupt change of system dynamics structure, and it is a new method to solve the detection of abrupt change in dynamic structure of a system. First of all, this method is used to present the ability to detect abrupt change lying in ideal datasets of different types of ideal signals. Next, this method is used to analyze daily datasets of ground climate from national meteorological information center of China meteorological administration between 1960 and 2008 in Lanzhou meteorological observation station in Northwest area. The results show that the daily datasets are consistent with historical records, which further verifies that the method is effective and practical.

Keywords: Fisher information, dynamics structure, abrupt change detection, probability density distribution

PACS: 92.60.Wc

DOI: 10.7498/aps.62.189204

* Project supported by the National Natural Science Foundation of China (Grant Nos. 61273154/F030116, 51077066/E070303) and the Priority Academic Program Development of Jiangsu Higher Education Institutions, China (Grant No. (2011) No.6).

† Corresponding author. E-mail: spcai@mail.ujs.edu.cn