

非线性时间序列的动力结构突变检测的研究^{*}

龚志强¹⁾²⁾ 封国林²⁾³⁾ 董文杰³⁾ 李建平⁴⁾

1) (扬州大学物理科学与技术学院, 扬州 225009)

2) (中国科学院大气物理研究所东亚区域气候-环境重点实验室, 北京 100029)

3) (国家气候中心, 中国气象局气候研究开放实验室, 北京 100081)

4) (中国科学院大气物理研究所大气科学和地球流体力学数值模拟国家重点实验室, 北京 100029)

(2005 年 7 月 4 日收到, 2005 年 12 月 10 日收到修改稿)

基于非线性时间序列分析方法——动力学相关因子指数, 提出一种新的动力结构突变的检测方法——动力学指数分割算法. 通过理想时间序列试验, 验证了该方法检测动力结构突变的有效性, 同时发现相对少量的尖峰噪声对该方法的影响较小, 但连续分布的随机白噪声对其具有一定的影响, 并与传统的滑动 T 检验法和 Yamamoto 法进行比较, 进而讨论它们各自的优缺点.

关键词: 动力学相关因子指数, 动力学指数分割算法, 噪声, 滑动 T 检验, Yamamoto 法

PACC: 9260X, 0545

1. 引言

时间序列分析是一门涉及到几乎一切科学和技术的学问, 并且已经有了相当长的历史和成就. 突变分析是时间序列研究的一个重要方面^[1,2]. 20 世纪 60 年代中期开始, 以 Thom^[3]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、医学等研究领域. 所谓突变就是指系统发生了突然的变化, 是系统对外界条件的光滑变化而做出的突然响应. 突变主要包括高等突变和初等突变, 通常所说的突变一般指初等突变, 如均值突变、频率突变、趋势突变和方差突变等. 目前对突变理论的研究大多集中于这些传统型突变, 研究的方法也较多, 主要有: 滤波检测法、滑动 t , F 检测法、Yamamoto 法等, 但这些方法本质上还是统计和线性的, 对物理过程的描述不太明显^[4-6]. 另外, 有些统计方法本身还存在一定的缺陷, 如滑动 t , F 检测法检测均值突变时, 经常会检测到一些虚假的突变点^[7-9]. 气候系统等许多实际系统的演化和发展可能受一个或多个驱动因子的控制, 往往表现为非线性、非平稳性和复杂性, 其内在的动力结构也可能随着驱动因子的改变而发生

快速的变化, 即其内在的演化方程发生了突变——动力结构突变^[10]. 目前关于这方面研究的理论和方法还相对较少, 迄今为止, 相对于气候动力系统^[11] 还没有一套使人们普遍接受的方法认定多长的时域内、哪些空间点属于同一动力系统^[11]. 由于存在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信息, 尽管可以对原始数据进行滤波处理, 但并不能完全消除噪声, 因此从非线性角度提出一种对噪声不太敏感的动力结构突变的检测方法, 进而研究这些突变形成的物理机制, 对于更好的分析非线性时间序列的本质特征具有重要的现实意义^[12]. 针对上述问题, 本文首先基于非线性时间序列分析方法——动力学相关因子指数, 提出一种新的动力结构突变的检测方法——动力学指数分割算法(下文简称 Q 算法). 通过构造多组理想时间序列对 Q 算法检测动力结构突变的有效性, 及尖峰噪声和白噪声对该方法的影响做了初步的讨论, 就传统的检测方法和 Q 算法进行了比较, 进而分析他们各自的优缺点.

2. 动力学指数分割算法及其物理意义

动力学相关因子指数是基于相空间重构理论的

^{*} 国家重点基础研究发展规划项目(批准号 2006CB400503)和国家自然科学基金(批准号 90411008 和 40325015)共同资助的课题.

[†] 通讯联系人. E-mail: feng_gl@sina.com

时间序列动力结构分析方法,其构造方法和物理意义如下:

对一个长度为 N 的时间序列 $\{x(t), t = 1, 2, \dots, N\}$ 进行嵌入空间上动力学轨线重构^[13],其嵌入向量表达式为

$$X_i = \{x(t_i), x(t_i + \tau), \dots, x(t_i + (m-1)\tau)\}, \quad (1)$$

其中 $\tau = \alpha \Delta t$ 为时间延迟, α 为延迟参数, Δt 为采样时间间隔, m 为嵌入空间维数. 对序列的每个点重构后, 组成了一个 $(N - \alpha(m-1)) \times m$ 维的向量矩阵

$$X = \{X_i, i = 1, 2, \dots, N - \alpha(m-1)\}, \quad (2)$$

它的自关联和定义为^[14]

$$\begin{aligned} C_{xx}(\epsilon) &= P(\|X_i - X_j\| \leq \epsilon) \\ &= \frac{2}{(N - \alpha m)(N - \alpha(m-1))} \\ &\quad \times \sum_{i=1}^{N - \alpha m} \sum_{j=i+1}^{N - \alpha(m-1)} \Theta(\epsilon - \|X_i - X_j\|) \quad (3) \end{aligned}$$

表示在重构空间里 ϵ 距离内找到邻近点 X_i 的概率, $\Theta(h)$ 为 Heaviside 阶跃函数. 在描述混沌信号时, 自关联和具有一定区分潜在动力学结构的能力, 但它还远不能作为识别混沌时间序列间相近性最重要的标准^[14-17].

如何更好地识别混沌时间序列动力异同性呢? 文献[18-20]初步回答了这一问题. 假设 $\{x_i\}$ 和 $\{x_j\}$ 是离散序列上的两点, 当 $|x(i) - x(j)| \leq \epsilon$ 时, $|x(i+1) - x(j+1)| \leq \epsilon$ 的概率 $S_m = C_{xx}^{m+1}(\epsilon) / C_{xx}^m(\epsilon)$ 比自关联和具有更强的预见性, 可用于两个时间序列集动力异同性的识别. 对于两个时间序列 $\{x_i\}, \{y_i\}$, 动力学自相关因子指数 Q_{xy} 定义为^[18-20]

$$Q_{xy} = \lim_{\epsilon \rightarrow 0} \left| \ln \frac{C_{xx}(\epsilon)}{C_{yy}(\epsilon)} \right|, \quad (4)$$

其物理意义是当 Q_{xy} 统计上足够小时, 序列集 $\{x_i\}, \{y_i\}$ 至少具有相近的动力学结构, 否则就不具有相近的动力学特征. 研究表明它能起到直接测量混沌时间序列之间“距离”的作用^[16], 能有效区分不同动力系统, 尤其是它能处理较短的时间序列. 根据已有的研究^[21], m 可取 3—4, α 取 1—4.

基于动力学相关因子指数的分割算法的构造和物理意义介绍如下: 取一宽度为 n 的滑动窗口 W , 分别计算 $x(t)$ 中 n 至 $N - n$ 各点左右两个窗口的动力学指数 $Q_1(i)$ 和 $Q_2(i)$ 以及标准偏差 $s_1(i)$ 和 $s_2(i)$. 计算 Q 指数时, 一般以原序列为参考窗口,

本文将原序列划分为若干个宽度为 n 的窗口 $\{W_i\}$, 分别将其作为参考窗口并计算动力学相关因子指数值, 最后求统计平均, 则 i 点的合并偏差 $s_D(i)$ 为

$$s_D(i) = \frac{1}{n} \times [s_1(i)^2 + s_2(i)^2]^{1/2}, \quad (i = n, n+1, \dots, N-n) \quad (5)$$

我们用统计值 $T(i)$ 来量化表示 i 点左右两个窗口动力学指数的差异, 即

$$T(i) = \lambda \times \left| \frac{Q_1(i) - Q_2(i)}{s_D(i)} \right|, \quad (6)$$

其中 λ 为缩放因子, 一般可取 3—6, 则得到长度为 $N - 2n$ 的检验统计值序列 $T(t)$, T 越大, 表示该点左右两部份动力结构的差异越大. 计算 $T(t)$ 中的最大值 T_{\max} 的统计显著性水平 $P(T_{\max})$:

$$P(T_{\max}) = P(\tau \leq T_{\max}), \quad (7)$$

$P(T_{\max})$ 表示在随机过程中取到 τ 值小于等于 T_{\max} 的概率. $P(T_{\max})$ 可近似表示如下:

$$P(T_{\max}) \approx (1 - I_{(\nu/\nu + T_{\max}^2)}(\delta\nu, \delta))^\eta, \quad (8)$$

由蒙特卡洛模拟可得到: $\eta = 4.19 \ln N - 11.54$, $\delta = 0.40$, N 是时间序列 $x(t)$ 的长度, $\nu = N - 2$, $I_x(a, b)$ 为不完全 β 函数. 我们设定一个临界值 P_0 , 如果 $P(T_{\max}) \geq P_0$, 则于该点将 $x(t)$ 分割成两段动力结构有一定差异(差异的程度随 P_0 的取值变化)的子序列, 否则不分割^[22]. 对新得到的两个子序列分别重复上述操作, 直至所有的子序列都不可分割. 为确保统计的有效性, 当子序列的长度小于等于 l_0 (最小分割尺度)时不再对其进行分割. 动力学相关因子指数的物理意义表明, 通过上述操作, 我们将原序列分割为若干表征不同动力结构的子序列, 各子序列分别包含了不同层次信息, 分割点即为动力结构突变点. 根据已有的工作^[23-24], l_0 的取值不小于 25, P_0 可取 0.5—0.99 (视具体的分割要求和资料特点而定).

3. 理想时间序列的构建和检测

本节就如何应用 Q 算法识别动力系统发生了改变作初步的研究. 不失一般性, 这里给出简单情形下 Q 算法对理想时间序列动力结构的诊断试验. 我们构建一理想时间序列 $x(t)$ (2000 个点), 因为构建的理想序列为无量纲的数值序列, 因此本文从无量纲的角度进行分析, 其动力学方程组为

$$x(t) = \begin{cases} 2\sin(0.5t) + 1.5\cos(0.2t) + 0.1, & (t < 1000) \\ e^{t/3000} + 2\sin(0.2t), & (1000 \leq t < 1500) \\ \tan(\pi t) + 2\sin(0.2t) + 2, & (1500 \leq t \leq 2000) \end{cases} \quad (9)$$

由(9)式可知,该时间序列包含如下三种子序列:

(1)当 $t < 1000$ 时,子序列由正弦和余弦函数叠加得到;

(2)当 $1000 \leq t < 1500$ 时,子序列由指数函数和正弦函数叠加得到;

(3)当 $1500 \leq t \leq 2000$ 时,子序列由正切和余弦函数叠加得到。

三段子序列分别代表了不同的动力结构, $t_0 = 1000, 1500$ 分别为两个动力结构突变点(图 1(a)). 取窗口宽度 $n = 80, l_0 = 100, P_0 = 0.99$, 对该序列进行检测. 图 1(b)为点 80—1920 的 T 值曲线, 可以看出, 点 80—1920 的 T 值曲线存在两个明显的大值区 920—1080 和 1420—1580, 这两个大值区与原序列结构突变点的位置相对应, 故用 Q 算法进行检测时, 结构突变点一般出现在 T 值较大的区域. $t = 1003$ 时有 $T_{\max 1}$, 此时 $P(T_{\max 1}) > 0.99$, 在该处原序列被分割为两段, 对新得到的子序列继续用 Q 算法进行处理. 图 1(c)为点 1083—1920 的 T 值曲线, 结构突变点 1500 附近同样存在一个大值区, 该区域中 $t = 1485$ 时有 $T_{\max 2}$, 此时 $P(T_{\max 2}) > 0.99$, 故在 1485 处该子序列被二次分割, 对子序列 1—1003 进行检测, 图 1(d)为该序列的 T 值曲线, 其 $P(T_{\max 3}) < 0.99$,

故这段子序列不能继续分割. 如此重复, 我们得到两个分割点分别在 t 等于 1003 和 1485 的位置, 与原序列的两个结构突变点的位置较接近. 参考窗口宽度的取值一般不超过原序列长度的 5%, 故首尾两个窗口内存在动力结构突变的可能性较小, 在试验中不考虑一般不会影响该方法在实际中的应用. 通过上述试验可知, 发生结构突变点的位置一般对应于 Q 算法 T 值的大值区, 同时检测到的突变点与实际突变点相比存在较小的偏差, 故我们定义检测偏差 η , 即

$$\eta = \frac{|t - t_0|}{N} \times 100\%, \quad (10)$$

其中 t_0 为实际突变点, t 为检测到的突变点, N 为原序列长度.

为进一步验证 Q 算法的有效性, 本文对 n 取不同数值进行试验, 表 1 给出了其中三组的试验结果(其余试验结果与此类似). 由表 1 可知, Q 算法检测得到的突变点与实际的结构突变点之间可能存在一定的偏差, 即 Q 算法能够检测到原序列可能在某一区域内发生了突变. 同时检测偏差随 n 取值的不同可能发生变化, 反复试验发现, $n = 80$ 时检测偏差最小, 点 1000 和 1500 处的两个结构突变点的检测偏

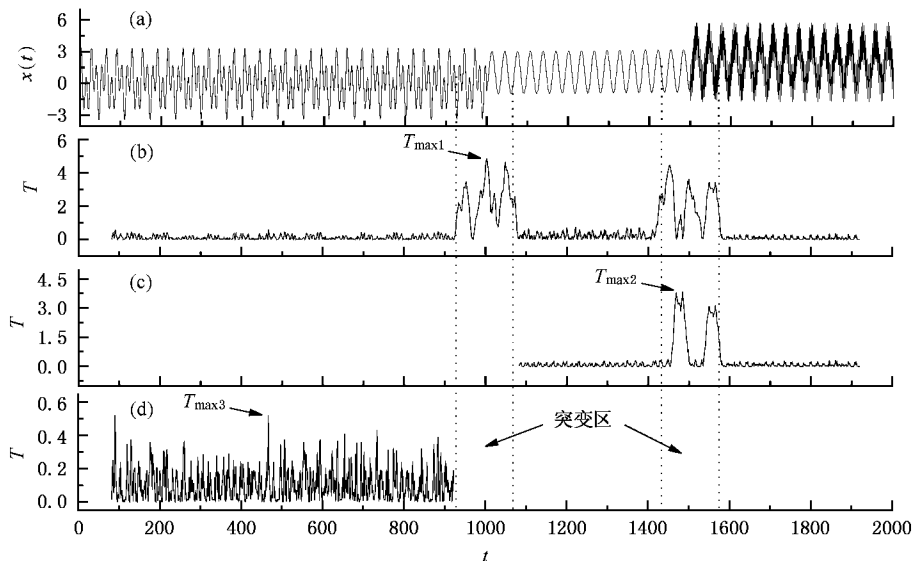


图 1 理想时间序列及其 Q 算法检测结果 (a)为理想序列 (b)(c)(d)分别为点 80—1920, 点 1083—1920 和点 80—823 的 T 值曲线

差分别为0.15%和0.75%，故应用 Q 算法进行检测时存在一个最佳窗口宽度 ω ， $n = \omega$ 时检测效果最好（ ω 可能随实际序列的变化而变化），我们将理想时

间序列的动力学方程组(9)式变为(11)式，用 Q 算法对新序列进行检测，检测结果与前者类似。上述试验表明 Q 算法具有一定的普适性。

$$x(t) = \begin{cases} 2\sin(0.5)t + 1.5\cos(0.2t) + 0.1, & (t < 1000) \\ (t/1000)^2 - t/1000 + 2\cos(0.1t), & (1000 \leq t < 1500) \\ \tan(\pi t) + 2\sin(0.2t) + 2, & (1500 \leq t \leq 2000) \end{cases} \quad (11)$$

表1 不同窗口宽度的检测结果

窗度宽度		$n = 60$		$n = 80$		$n = 100$			
t		1023	1486	1003	1485	937	1039	1486	1570
检测结果	Δt	23	14	3	15	63	39	14	70
	η	1.15%	0.70%	0.15%	0.75%	3.15%	1.95%	0.70%	3.50%

4. 噪声和扰动对 Q 算法的影响

由于存在外强迫和仪器本身的测量误差等因素的影响，实际观测数据大都包含尖峰噪声、白噪声和扰动等的信息，即使对序列进行预处理，一般情况下很难完全滤除这些“坏数据”^[25, 26]，下面就噪声对 Q 算法的影响作初步的探讨。

在(9)式构造的理想时间序列中随机加入15个尖峰噪声，噪声大小在3—8之间变化(图2(a))。因为尖峰噪声的数目相对原序列的长度较小，故不

会改变原序列的结构特点。基于 Q 算法对加噪序列进行检测， $n = 80$ ， $l_0 = 100$ ， $P_0 = 0.99$ 。由图2(b)可以看出，检测到的两个突变区域分别为920—1080和1420—1580，大小较不加噪声时变化不大，同时最后检测到的突变点为1002和1469，检测偏差分别为0.10%和1.55%，相对于不加噪声时变化较小。由此可见相对少量的尖峰噪声对 Q 算法的检测结果影响较小。

在(9)式构造的理想序列中的多个时段叠加随机白噪声 $\epsilon(t)$ ， $-1 \leq \epsilon(t) \leq 1$ ，且每段白噪声宽度为20—30不等，各段内噪声连续分布，各段的宽度

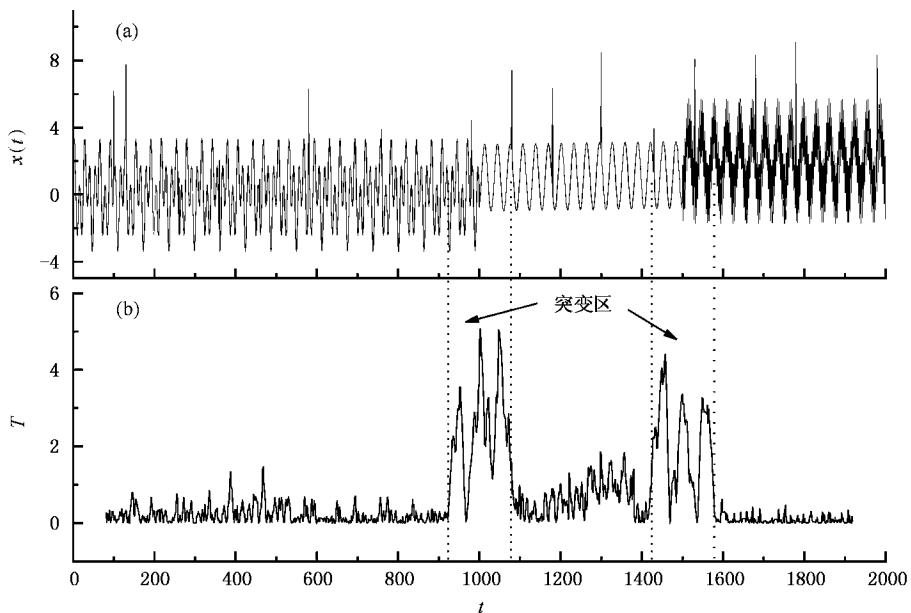


图2 加尖峰噪声序列及其 Q 算法检测结果 (a) 叠加尖峰噪声的理想序列 (b) 加噪序列的 T 值曲线

(数据点个数)之和为 100, 均值 $\epsilon \approx 0$ (图 3(a)). 随机白噪声的总宽度为原序列长度的 5%, 故原序列的结构特点也不会因为加入噪声而发生质的变化. 基于 Q 算法对加噪序列进行检测, $n = 80$, $l_0 = 100$, $P_0 = 0.99$. 由图 3(b)可以看出, 加噪声后的 T 值曲线较不加噪声时, 其波动反常的区域增多、增宽, 但仔细比较后发现结构突变点附近的 930—1070 和 1420—1580 两个区域内 T 值及其波动的幅度最大. 同时最后检测得到的突变点分别为 1003 和 1485, 相

应的检测偏差为 0.15% 和 0.75%, 这和不加噪声时的结果相同. 显然, 受白噪声的影响 T 值曲线的变化较大, 多个区域出现类突变现象, 故连续分布的随机白噪声对 Q 算法检测突变区域的影响较大, 但就最后检测到的结构突变点而言, 受白噪声的影响较小, 可以借助最后检测到的突变点排除那些虚假的类突变区. 将上述白噪声离散为多个更小时段并叠加到理想时间序列中, 检测结果与加尖峰噪的情况类似(图略).

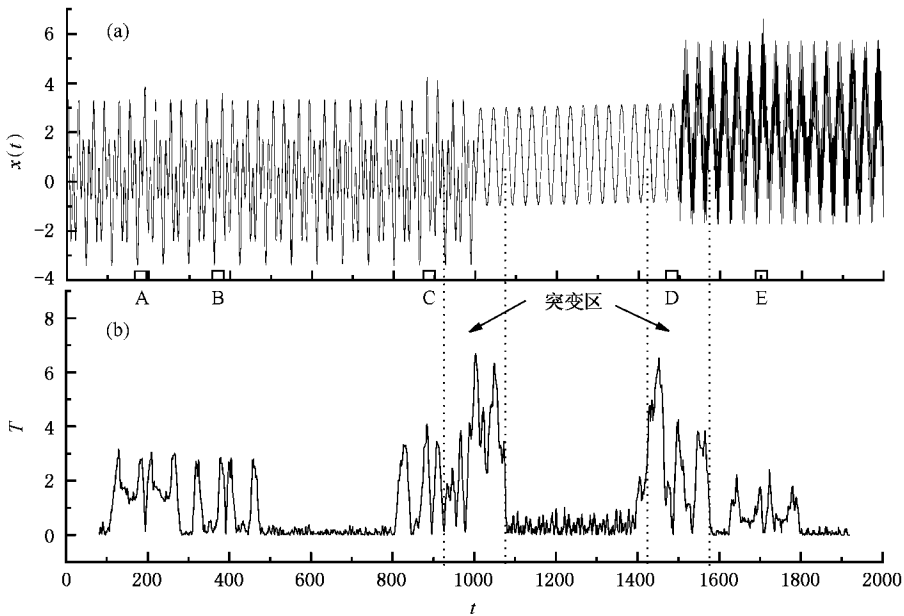


图 3 加随机白噪声序列及其 Q 算法检测结果 (a) 叠加随机白噪声的理想序列, 其中 A, B, C, D, E 五处分别叠加了宽度不等随机白噪声 (b) 加噪序列的 T 值曲线

5. Q 算法和传统分割算法的比较

滑动 T 检验、Yamamoto 法等传统分割算法大多基于时间序列是平稳过程的思想, 一般也只是从外部特征(均值和趋势等)的角度进行检测, 在实际检测中大都存在一定的缺陷^[27, 28]. Q 算法则是建立在对原时间序列进行相空间重构的基础上, 从动力结构变化的角度进行检测, 检测得到的突变区域和突变点的物理意义较明确, 即对应原序列发生了动力结构突变. 可通过比较滑动 T 检验和 Q 算法的检测结果, 对 Q 算法的有效性作进一步的分析.

在 (9) 式构造的理想序列中叠加一全局的线性外强迫 $\varphi(t)$ (图 4(a)),

$$\varphi(t) = t/500 \quad t = 1, 2, \dots, 2000 \quad (12)$$

分别基于滑动 T 检验、Yamamoto 法和 Q 算法对该

序列进行检测. 由图 4(b)可以看出, 基于滑动 T 检验进行检测, 滑动窗宽度取为 80, 显著性水平取为 0.01, 共检测到 8 个突变点, 分别为 5, 24, 26, 527, 779, 1004, 1257 和 1506. 这些突变点虽然包括了结构突变点, 由于它们都是基于各段序列的不同均值得到, 故很难区别其物理意义, 即很难将那些虚假的突变点剔除而影响对序列整体性质的分析.

由图 4(c)可以看出, 基于 Yamamoto 法进行检测, 窗口宽度取 80, 连续设置基准点. 对应信噪比 (SNR) 曲线中存在的两个反常区域, 分别近似为 920—1080 和 1420—1580. 其中第一个区域的反常较明显, 但第二个反常区域较平稳不易辨别. 当显著性水平确定为 0.01, SNR 为 0.37, 此时只有点 1003 符合突变条件, 显著性水平降低为 0.08 时, 对应信噪比 (SNR) 为 0.28 时, 点 1003 和 1508 符合突变条件, 检测偏差分别为 0.15% 和 0.40%. 可见, 显著性水

平的主观确定,可能在一定程度上影响 Yamamoto 法的检测结果.同时 Yamamoto 法本质也是基于均值的思想进行检测,故对于这两个突变点,就其物理意义而言,仅反映了原序列均值的突变,而未能很好地体现其动力结构的变化.

图 4(d)为 Q 算法的检测结果, $n = 80$, $l_0 = 100$, $P_0 = 0.99$,可以看出,虽然叠加了全局外强迫,仍能

检测得到两个明显的突变区域,且较不加全局外强迫时没有发生太大的变化,分别为 920—1080 和 1420—1580.而最终检测到的动力结构突变点为 1047 和 1452,检测偏差分别为 2.35% 和 2.40%.虽然 Q 算法检测得到的突变存在一定的偏差,但就整体而言,相对于传统的突变检测方法在检测动力结构突变方面具有一定的优越性.

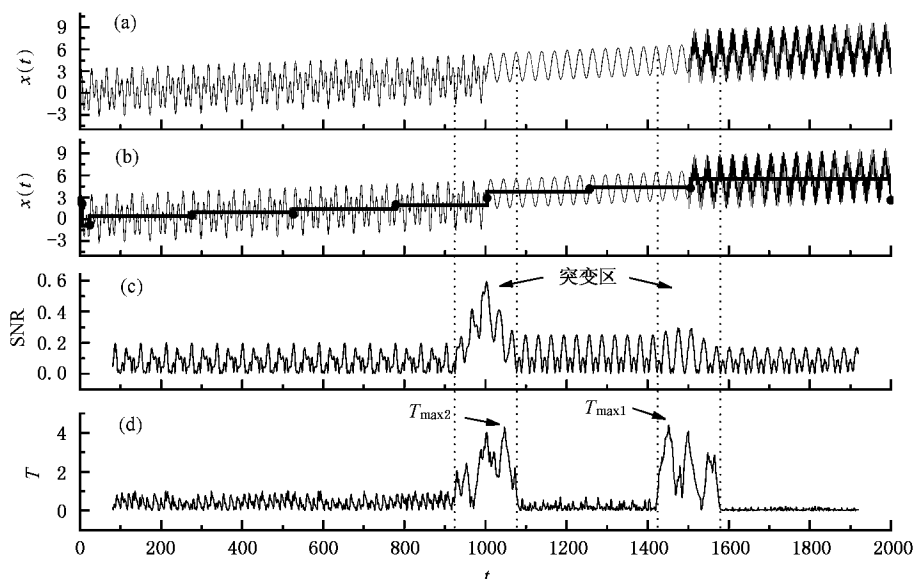


图 4 加全局线性外强迫序列及其滑动 T 检验、Yamamoto 法和 Q 算法检测结果 (a) 叠加线性全局外强迫的理想序列 (b) (c) (d) 分别为滑动 T 检验、Yamamoto 法和 Q 算法的检测结果 (b) 中黑点为均值突变点,粗线为均值曲线,细线为原序列)

6. 结论和讨论

1) 本文在非线性时间序列分析方法——动力学相关因子指数的基础上提出了一种新的检测动力结构突变的方法——动力学指数分割算法,同时构造理想时间序列验证了该方法能够有效检测序列在某一区域内动力结构发生了变化.

2) 相对少量的尖峰噪声对 Q 算法的检测结果影响较小,连续分布的随机白噪声对 Q 算法检测突变区域的影响较大,但就最后检测得到的结构突变点而言,受白噪声的影响较小.

3) 比较滑动 T 检验、Yamamoto 法和 Q 算法检测时间序列结构突变点时,相对于一些传统的检测方法而言, Q 算法具有虚假突变点少,检测到的突变区域较明显,物理意义较明确等特点.

Q 算法本身也存在一定的缺陷,如只能在一定程度上检测出发生动力结构突变的区域,最终检测得到的突变点一般也存在一定的检测偏差,同时,在具体应用中还涉及如何选择相空间重构的最优维数,如何选择最优窗口宽度等问题,这些问题的有效解决将会更好地实现 Q 算法在实际资料检测中的广泛应用.

[1] Shi N 2005 *Chin. Phys.* **14** 844

[2] Dai X G, Wang P, Chou J F 2003 *Chin. Sci. Bull.* **23** 2483

[3] Ling F H 1984 *Adv. Mech.* **14** 389 (in Chinese) [凌复华 1984 力学进展 **14** 389]

- [4] Yang M X , Yao T D 1999 *Exploration of Nature* **18** 30 (in Chinese)
[杨梅学、姚檀栋 1999 大自然探索 **18** 30]
- [5] Cui J X , Zhou S Z 2001 *J. Hebei Normal Univ.* (Natural Science Ed.) **25** 264 (in Chinese) [崔建新、周尚哲 2001 河北师范大学学报(自然科学版) **25** 264]
- [6] Feng G L , Gong Z Q , Dong W J *et al* 2005 *Acta Phys. Sin.* **54** 3947 (in Chinese) [封国林、龚志强、董文杰等 2005 物理学报 **54** 3947]
- [7] Shi Y F , Zhang P Y. 1996 *The variation trend and its effect of the climate in China and sea level change historical climatic change in China.* (Jinan : Shandong Science and Technology Press) 383-393 (in Chinese) [施雅风、张丕远 1996 中国气候与海面变化及其趋势和影响中国历史气候变化 (济南 : 山东科学技术出版社) 383-393]
- [8] Yang W F , Li Z Y , Li X M 1997 *Quarterly J. Appl. Meteorology* **8** 119
- [9] Yan Z W , Li Z Y , Wang X C 1993 *Sci. Atmospheric Sin.* **17** 663
- [10] Wan S Q , Feng G L , Dong W J *et al* 2005 *Acta. Phys. Sin.* **54** 5487 (in Chinese) [万仕全、封国林、董文杰等 2005 物理学报 **54** 5487]
- [11] Chen B M , Ji L R , Yang P C *et al* 2003 *Chin. Sci. Bull.* **48** 513 (in Chinese) [陈伯民、纪立人、杨培才等 2003 科学通报 **48** 513]
- [12] Tan Z , Li A G 2003 *J. Xi 'an Jiaotong Univ.* **37** 338 (in Chinese)
[覃 征、李爱国 2003 西安交通大学学报 **37** 338]
- [13] Sauer T 1994 *Phys. Rev. Lett.* **72** 3811
- [14] Grassberger P 1983 *Phys. D* **9** 189
- [15] Kantz H 1994 *Phys. Rev. E* **49** 5091
- [16] Provenzale A , Smith L A , Vio R , Murante G 1992 *Phys. D* **58** 31
- [17] Grassberger P 1990 *Phys. Lett. A* **148** 63
- [18] Savit R , Green M 1991 *Phys. D* **50** 95
- [19] Manuca R , Savit R 1996 *Phys. D* **99** 134
- [20] Li C G , Pei L Q 2003 *Acta. Phys. Sin.* **52** 2114 (in Chinese)
[李春贵、裴留庆 2003 物理学报 **52** 2114]
- [21] Liu S D , Zhen Z G 1993 *Acta Meteor. Sin.* **51** 333 (in Chinese)
[刘式达、郑祖光 1993 气象学报 **51** 333]
- [22] Bernaola-Galvan P , Ivanov P Ch , Amaral L A N , Stanley H E 2001 *Phys. Rev. Lett.* **16** 168105
- [23] Fukuda K , Stankey H E , Amaral L A N 2004 *Phys. Rev. E* **69** 021108
- [24] Feng G L , Dong W J 2004 *Chin. Phys.* **13** 413
- [25] Zhang D E , Liu C Z , Jiang J M 1997 *Quater. Sci.* **1** 1
- [26] Hou W , Feng G L , Gao X Q *et al* 2005 *Acta. Phys. Sin.* **54** 2441 (in Chinese) [侯 威、封国林、高新全等 2005 物理学报 **54** 2441]
- [27] Liang Y K , Zhang D E 2004 *Meteor. Sci. and Tech.* **32** 137 (in Chinese) [梁叶宽、张德二 2004 气象科技 **32** 137]
- [28] Qu S X , He D R 1997 *Acta Phys. Sin.* **46** 1307 (in Chinese) [屈世显、何大韧 1997 物理学报 **46** 1307]

The research of dynamic structure abrupt change of nonlinear time series *

Gong Zhi-Qiang^{1,2)} Feng Guo-Lin^{2,3)†} Dong Wen-Jie³⁾ Li Jian-Ping⁴⁾

1) *Department of Physics, Yangzhou University, Yangzhou 225009, China*

2) *Key Laboratory of Regional Climate-Environment for Temperate East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China*

3) *Laboratory for Climate Studies of China Meteorological Administration, National Climate Center, Beijing 100081, China*

4) *State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Beijing 100029, China*

(Received 4 July 2005 ; revised manuscript received 10 December 2005)

Abstract

For a long time in the past, researches of time series were often based on their external characters and used linear and statistical methods. However, most actual systems are nonlinear, nonstationary and complicated, which increased the difficulties in treating them. The research of abrupt change is one of most important research aspects of nonlinear time series, for which the traditional method based on the external characters of data and using linear process lacks enough physical foundation, and has obvious limitations. How to find out the essence of complicated systems from time series, in other words, to check the abrupt change in dynamical structure of actual data series is a really important problem pending solution. In the present paper, we present a new method—the dynamical correlation exponent segmentation algorithm for checking dynamical abrupt change based on the dynamical lag correlation exponent. The validity of this method is verified by constructing an ideal time series and put it to test. It was found that a few noise spikes have little influence, but continuously distributed white noise has some influence to this new method. Comparison with conventional t-test and Yamamoto method was made to show the relative merits of the methods.

Keywords : dynamical lags correlation exponent, dynamical correlation exponent segmentation algorithm, noise, student's t-test, Yamamoto method

PACC : 9260X, 0545

* Project supported jointly by the State Key Development Program for Basic Research (Grant No. 2006CB400503) and the National Natural Science Foundation of China (Grant Nos. 90411008 and 40325015).

† Corresponding author. E-mail feng_gl@sina.com