

安庆气温的多重分形消除趋势波动分析与预测

程花花¹, 郑婷婷^{1,2}, 万涛¹, 张琛¹, 章意成¹

(1. 安徽大学 数学科学学院, 安徽 合肥 230039;

2. 安徽大学 智能计算与信号处理教育部重点实验室, 安徽 合肥 230039)

摘 要: 气温变化的研究对人类的生产建设具有重要的指导意义。运用多重分形消除趋势波动分析法(MF-DFA方法)先对安庆1961年至2008年的日平均气温与春夏秋冬四季的平均气温进行分析,然后预测安庆未来四季气温的变化趋势。通过分析发现安庆的日平均气温与春夏秋冬四季的平均气温均存在多重分形现象。并且预测未来四季气温变化的趋势是:春秋两季气温整体呈上升趋势变化不明显,并有可能下降,与过去的气温变化趋势有较弱的相反性;夏冬两季气温变化整体趋势将继续上升,与过去的变化趋势保持一致。

关键词: 气温变化;多重分形;MF-DFA方法;Hurst指数

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2010)12-0065-04

Multifractal and Detrended Fluctuation Analysis and Forecast of Air Temperature in Anqing

CHENG Hua-hua¹, ZHENG Ting-ting^{1,2}, WAN Tao¹, ZHANG Chen¹, ZHANG Yi-cheng¹

(1. Department of Mathematical Sciences, Anhui University, Hefei 230039, China;

2. Ministry of Edu. Key Lab. of Intelligent Computing & Signal
Processing, Anhui Univ., Hefei 230039, China)

Abstract: Studying temperature variation has important guiding significance to the human life. By using multifractal and detrended fluctuation analysis method, it analyzed the daily average temperature and the seasonal average temperature of Anqing from 1961 to 2008, and then forecasted the future tendency of seasonal temperature variation. The results show that both the daily average temperature and the seasonal average temperature have multifractality, they are summarized as follows: the overall trend of temperature variation in spring and autumn will not rise obviously, and is likely dropped, which has weakly contrary to the past. The overall trend of temperature variation in summer and winter will rise obviously, which is consistent with the past.

Key words: temperature variation; multifractal; MF-DFA method; Hurst index

0 引言

目前,全球气候变化已成为国际社会关注的焦点,气候变化引发的各种自然灾害给人类的生存发展带来了极为严重的影响。在近来的一个世纪里,北半球气候明显变暖,我国的气候在北半球气候变化下也明显上升,国内的许多大小城市也开始关注气候变化的问题^[1,2]。安庆市位于我国安徽省西南间,长江北岸。全市行政区域介于北纬29°47′~31°16′与东经115°45′~117°44′之间,总面积15348.53平方公里,占全省总

面积的11.04%。安庆属北亚热带湿润气候区,气候温和,四季分明。

分形(fractal)这个名词是Mandelbort在20世纪70年代为了表征复杂图形和复杂过程首先引入自然科学领域的,它的原意是不规则的、支离破碎的物体^[3]。分形可以分为规则分形和不规则分形。迄今为止,分形还没有一个严格的定义。1982年Mandelbort将分形定义为Hausdorff(豪斯道夫)维数大于拓扑维数的集合。此定义强调维数,而其中的豪斯道夫维数一般不是整数。1986年Mandelbort给出了一个更广泛、更通俗的定义^[4]:分形是局部和整体有某种方式相似的形(A fractal is a shape made of parts similar to the whole in some way)。该定义强调图形中局部和整体之间(包括小的局部和大的局部之间)的自相似性。除了几何图形,还可以研究相空间中的轨迹等图形的自相

收稿日期:2010-04-09;修回日期:2010-07-01

基金项目:安徽省高等学校省级自然科学研究项目(KJ2007B239)

作者简介:程花花(1986-),女,硕士研究生,主要研究领域为人工智能;郑婷婷,博士,副教授,硕士生导师,主要研究领域为人工智能、分形。

似性,这种自相似性应该有若干层次,其尺度的变化应该尽可能大。

随着分形理论不断发展,人们已经将它广泛地应用于物理、化学、生物、经济等方面。2002 年 Kantelhardt 首次提出多重分形消除趋势波动分析法^[5,6] (multifractal and detrended fluctuation analysis, MF-DFA),跟其它的分形方法相比,该方法的优点在于它能够发现非平稳时间序列中的长范围相关性,更适合分析非平稳时间序列的多重分形特征,同时能够避免对相关性的错误判断^[7,8]。在近几年,多重分形方法在分析气候数据、天气预测等领域已得到成功的应用,2004 年江田汗等人对全球和我国的过去近百年气温变化趋势进行了分形分析^[9,10]。文中主要利用多重分形消除趋势波动分析法对时间序列气温数据进行分析,并结合过去近 50 年气温变化趋势,对安庆未来的气温变化做出相关预测。

1 数据和方法

1.1 资料来源

根据国家气象信息中心气象数据室(中国气象科学资料共享服务网, <http://cdc.cma.gov.cn/>)提供的气象数据,分别整理出安庆 1961~2008 年的日平均气温以及春夏秋冬四季的平均气温,其中,春季(3,4,5 月)、夏季(6,7,8 月)、秋季(9,10,11 月)、冬季(12,次年 1,2 月)^[11]。

1.2 分析方法

多重分形消除趋势波动分析法是验证一个非平稳时间序列是否具有多重分形性的有效方法^[12],对于一个给定长度的时间序列 $\{x_i\}$, $i = 1, 2, \dots, N$, MF-DFA 方法介绍如下。

一般 MF-DFA 方法可分为以下五个步骤:

(1) 求序列对于 $\{x_i\}$ 均值的累积离差序列 $\{y_k\}_{k=1}^N$:

$$y_k = \sum_{i=1}^k (x_i - \bar{x}), k = 1, 2, \dots, N \quad (1)$$

$$\text{其中 } \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

(2) 分割 $\{y_k\}_{k=1}^N$ 为等长小段。

对任意正整数 s , 记 $N_s = [N/s]$, 这里 $[x]$ 表示 x 的整数部分。将序列 $\{y_k\}_{k=1}^N$ 自首向尾分割成长为 s 的 N_s 个互不重叠的小段, 由于 N 未必是 s 的整数倍, 为了使末尾的数据不致遗漏, 再将序列自尾向首重复这一过程, 于是得到 $2N_s$ 个小段:

$$\{y_{(\lambda-1)s+j}\}_{j=1}^s, \lambda = 1, 2, \dots, N_s \quad (2)$$

$$\{y_{N-(\lambda-N_s)s+j}\}_{j=1}^s, \lambda = N_s + 1, N_s + 2, \dots, 2N_s \quad (3)$$

(3) 消除每小段序列的趋势。

通过最小二乘法拟合每一小段上的局部趋势, 记 $P_\lambda(j) = a_0 + a_1j + \dots + a_jj^k$ 是第 λ 小段上的局部趋势, $\lambda = 1, 2, \dots, 2N_s, k = 1, 2, \dots$ 。消除每一小段的趋势, 得残差序列为:

当 $\lambda = 1, 2, \dots, N_s$ 时,

$$D_\lambda(j) = y_{(\lambda-1)s+j} - P_\lambda(j), j = 1, 2, \dots, s \quad (4)$$

当 $\lambda = N_s + 1, N_s + 2, \dots, 2N_s$ 时,

$$D_\lambda(j) = y_{N-(\lambda-N_s)s+j} - P_\lambda(j), j = 1, 2, \dots, s \quad (5)$$

(4) 对任意不为 0 的实数 q , 计算序列的 q 阶波动函数。

$$F_q(s) = \left\{ \frac{1}{2N_s} \sum_{\lambda=1}^{2N_s} [F^2(s, \lambda)]^{q/2} \right\}^{1/q} \quad (6)$$

$$\text{其中 } F^2(s, \lambda) = \frac{1}{s} \sum_{j=1}^s D_\lambda(j)^2, \lambda = 1, 2, \dots, 2N_s$$

当 $q = 0$ 时, 波动函数不能由(6)式直接计算, 而应由(7)式计算:

$$F_0(s) = \exp \left\{ \frac{1}{4m} \sum_{\lambda=1}^{2m} \ln [F^2(s, \lambda)] \right\} \quad (7)$$

(5) 利用最小二乘法拟合 $\log F_q(s)$ 与 $\log s$ 的散点图, 得到拟合直线:

$$\log F_q(s) = \log C + h(q) \log s \quad (8)$$

其中标度指数 $h(q)$ 即为广义赫斯特(Hurst)指数。这里当 $h(q)$ 为 q 的函数时, 就称序列 x_i 为符合多重分形特征。

对应于不同大小的赫斯特指数($0 \leq h(q) \leq 1$), 存在以下几种情况:

(1) $h(q) = 0.5$, 即气温各个数据之间是完全独立的, 相互间没有依赖性, 是符合完全随机游走的。

(2) $0 \leq h(q) < 0.5$, 说明序列具有短程相关性。即将来气温的变化与过去相反, 过去整体增加的趋势预示着未来将整体减少, 过去整体减少的趋势预示着未来将整体增加。 $h(q)$ 值越接近于 0, 短程相关性越强。

(3) $0.5 \leq h(q) < 1$, 说明序列具有长程相关性。即将来气温的变化趋势与过去一致, 过去整体增加的趋势预示着未来将整体增加, 过去整体减少的趋势预示着未来将整体减少。 $h(q)$ 值越接近于 1, 长程相关性就越强。可以看出, $h(q)$ 指数能很好地揭示时间序列中的变化趋势, 并且可以由 Hurst 指数值的大小来判断趋势成分的短程相关性或长程相关性的强弱。

2 结果分析

文中所指的安庆日平均气温是取 1961 年 1 月到 2008 年 11 月中每月的 1 号、11 号、21 号、月底最后一

天的日平均气温,共 2300 个数据,其中 s 取 10 ~ 575(通常要求 $s \geq \max\{k+2, 10\}$)。春季平均气温是取 1961 年到 2008 年的每年 3 月 1 日至 5 月 31 日的日平均气温,共 4416 个数据。夏季平均气温是取 1961 年到 2008 年的每年 6 月 1 日到 8 月 31 日的日平均气温,共 4416 个数据。秋季平均气温是取 1961 年到 2008 年的每年 9 月 1 日到 11 月 30 日的日平均气温,共 4368 个数据。冬季平均气温是取 1961 年到 2008 年的 12 月 1 日至次年 2 月 28(29)日的日平均气温,共 4242 个数据。其中计算春夏秋冬四季平均气温时 s 皆取 16~1000。另外,根据安庆 1961~2008 年的年平均气温以及春夏秋冬四季平均气温整理出气温随年份的变化曲线,再用直线拟合得到的斜率即是下文中所说的气温变化的速率(图略)。经分析得知:安庆的总体气候将继续变暖,这与全球气候变暖有着直接的关系。其中,安庆的年平均气温以 $0.36^{\circ}\text{C}/10\text{year}$ 的速率上升,春夏秋冬四季平均气温的上升速率分别是: $0.49^{\circ}\text{C}/10\text{year}$ 、 $0.18^{\circ}\text{C}/10\text{year}$ 、 $0.33^{\circ}\text{C}/10\text{year}$ 、 $0.42^{\circ}\text{C}/10\text{year}$,相比较之下,夏季的气温变暖趋势较平缓些,春秋冬三季气温上升的趋势比较明显。

在 MF-DFA 方法中的 k 表示对 s 个点进行拟合时用的是 k 阶多项式拟合的($k=1, 2, 3, 4$),并且称此时的 MF-DFA 为 k 阶 MF-DFA 方法。由 MF-DFA 中的步骤(3)可知, k 阶的 MF-DFA 可以消去累积误差中的 k 阶趋势成分以及原始序列中的 $k-1$ 阶趋势成分。通过公式可求得安庆日平均气温及其春夏秋冬

四季平均气温的广义赫斯特指数 $h(q)$ 的估计值(其结果见表 1)以及 $h(q)$ 与 q 的关系图(图略)。

根据上述公式求得表和图的结果可知:

(1)对于固定的 k 时,安庆的日平均气温和四季日平均气温均随着 q 增加而减少,即 $h(q)$ 为 q 的减函数, $h(q)$ 是依赖于 q 的函数,所以说安庆的日平均气温和四季平均气温均具有多重分形的特征,若用简单的单一分形来描述是不合适的。

(2)对于给定的 q 值,春夏秋冬四季的平均气温的赫斯特指数 $h(q)$ 均随着 k 的增加而增加,而对于给定的 k 值时,春夏秋冬四季的日平均气温的赫斯特指数 $h(q)$ 均随着 q 的增加而减少。并且,对于不同的 k 与 q 值下,夏秋冬三季的赫斯特指数都要大于春季的赫斯特指数。其中,夏季最大,冬季第二,秋季第三,春季最小,这表明,春季的日平均气温的波动程度要比其它三季要小,相比较而言,夏季的日平均气温的波动程度是最大的。

(3)对春夏秋冬四季的平均气温,用 1、2、3、4 阶的 MF-DFA 方法分析,发现四季均用 2 阶的分析比较好。并且在表 1 中可以看到 $q=1$ 的情况下,四季的不同赫斯特指数。通过对表和图的分析,发现用 $k=2, q=1$ 时的赫斯特指数资料,来对安庆春夏秋冬四季平均气温的相关性程度进行预测较合适,预测结果如下:在 $k=2, q=1$ 的 MF-DFA 方法下,春季赫斯特指数为 0.4781,即春季平均气温具有很弱的短程相关性,将来的气温变化整体趋势与过去的相反;夏季的赫斯

表 1 不同 k 与 q 值下,安庆春夏秋冬四季里平均气温的赫斯特指数 $h(q)$ 的估计值

q	春季				夏季				秋季				冬季			
	$K=1$	$K=2$	$K=3$	$K=4$	$K=1$	$K=2$	$K=3$	$K=4$	$K=1$	$K=2$	$K=3$	$K=4$	$K=1$	$K=2$	$K=3$	$K=4$
-10	0.5297	0.7553	0.8351	0.9269	0.6603	0.7300	0.8434	0.9320	0.5474	0.7768	0.8583	0.9479	0.6492	0.7453	0.7898	0.8365
-9	0.5202	0.7437	0.8235	0.9153	0.6518	0.7229	0.8354	0.9234	0.5388	0.7664	0.8475	0.9367	0.6432	0.7374	0.7824	0.8294
-8	0.5093	0.7301	0.8100	0.9019	0.6417	0.7146	0.8260	0.9134	0.5285	0.7542	0.8348	0.9236	0.6365	0.7280	0.7738	0.8212
-7	0.4967	0.7142	0.7942	0.8864	0.6295	0.7049	0.8150	0.9016	0.5163	0.7398	0.8200	0.9085	0.6290	0.7170	0.7640	0.8120
-6	0.4825	0.6953	0.7759	0.8687	0.6150	0.6936	0.8023	0.8880	0.5017	0.7226	0.8026	0.8908	0.6207	0.7040	0.7529	0.8017
-5	0.4664	0.6729	0.7547	0.8485	0.5980	0.6809	0.7879	0.8726	0.4845	0.7021	0.7820	0.8702	0.6114	0.6889	0.7404	0.7904
-4	0.4488	0.6464	0.7306	0.8262	0.5786	0.6670	0.7722	0.8557	0.4648	0.6775	0.7579	0.8465	0.6012	0.6719	0.7266	0.7783
-3	0.4301	0.6157	0.7043	0.8023	0.5578	0.6527	0.7558	0.8381	0.4433	0.6476	0.7299	0.8200	0.5901	0.6533	0.7119	0.7656
-2	0.4107	0.5812	0.6771	0.7781	0.5367	0.6387	0.7396	0.8206	0.4207	0.6111	0.6980	0.7910	0.5776	0.6334	0.6965	0.7525
-1	0.3907	0.5446	0.6502	0.7544	0.5166	0.6255	0.7244	0.8039	0.3968	0.5676	0.6633	0.7603	0.5632	0.6128	0.6808	0.7393
0	0.3698	0.5093	0.6247	0.7319	0.4983	0.6133	0.7103	0.7880	0.3707	0.5207	0.6275	0.7289	0.5461	0.5917	0.6648	0.7261
1	0.3483	0.4781	0.6011	0.7111	0.4822	0.6021	0.6975	0.7728	0.3426	0.4770	0.5934	0.6985	0.5264	0.5706	0.6489	0.7131
2	0.3277	0.452	0.5799	0.6922	0.4684	0.5916	0.6858	0.7584	0.3155	0.4415	0.5630	0.6706	0.5049	0.5499	0.6332	0.7005
3	0.3099	0.4308	0.5612	0.6752	0.4566	0.5818	0.6749	0.7448	0.2930	0.4147	0.5373	0.6462	0.4832	0.5302	0.6182	0.6883
4	0.2953	0.4136	0.5449	0.6601	0.4467	0.5726	0.6649	0.7319	0.2756	0.3947	0.5161	0.6253	0.4624	0.5117	0.6039	0.6768
5	0.2838	0.3996	0.5307	0.6468	0.4382	0.5640	0.6555	0.7197	0.2622	0.3794	0.4986	0.6076	0.4436	0.4947	0.5906	0.6660
6	0.2746	0.3878	0.5184	0.6350	0.4309	0.5560	0.6468	0.7085	0.2516	0.3672	0.4840	0.5926	0.4271	0.4795	0.5785	0.6560
7	0.2671	0.3779	0.5077	0.6245	0.4246	0.5486	0.6389	0.6983	0.2431	0.3572	0.4717	0.5798	0.4129	0.4659	0.5676	0.6470
8	0.2609	0.3693	0.4983	0.6152	0.4190	0.5420	0.6317	0.6890	0.2360	0.3487	0.4611	0.5687	0.4007	0.4540	0.5579	0.6390
9	0.2557	0.3618	0.4900	0.6069	0.4140	0.5359	0.6253	0.6807	0.2300	0.3414	0.4520	0.5591	0.3904	0.4436	0.5494	0.6318
10	0.2511	0.3552	0.4826	0.5995	0.4095	0.5305	0.6196	0.6734	0.2248	0.3350	0.4440	0.5506	0.3815	0.4345	0.5419	0.6256

特指数为 0.6021,即夏季平均气温具有较弱的长程相关性,将来气温变化整体趋势与过去相同;秋季的赫斯特指数为 0.4770,即秋季平均气温具有很弱的短程相关性,将来气温变化整体趋势也与过去相反;冬季的赫斯特指数为 0.5706,即冬季平均气温具有比较弱的长程相关性,将来气温变化整体趋势与过去相同。

(4)结合前面分析得出的安庆过去几十年气温变化的速率得出的结果,即春夏秋冬四季平均气温的整体趋势均随着年份的增加而升高,春秋两季的平均气温呈现上升趋势较明显,夏季上升趋势稍平缓些。再利用结论(3)中较合适的赫斯特指数,预测安庆未来四季的气温变化趋势如下:春秋两季气温整体呈上升趋势变化不明显,并有可能下降,与过去的气温变化趋势有较弱的相反性;夏冬两季气温变化的整体趋势将继续上升,与过去的变化趋势保持一致。

3 结束语

综合上述分析得到如下结论:安庆的日平均气温和春夏秋冬四季的平均气温都具有多重分形特征,不适合用单一分形来描述;并且安庆的年平均气温整体呈上升趋势,这符合全球气候变暖的趋势;

预测到安庆未来四季气温的变化趋势是:春秋两季气温整体呈上升趋势变化不明显,并有可能下降,与过去的气温变化趋势有较弱的相反性;夏冬两季气温变化整体趋势将继续上升,与过去的变化趋势保持一致。

参考文献:

- [1] 林学春,于淑秋.近 40 年我国气候变化趋势[J].气象,1990,16(10):16-21.
- [2] 陈隆勋,朱文琴,王文,等.中国近 45 年来气候变化的研究[J].气象学报,1998,56(3):257-271.
- [3] 沙震,阮火军.分形与拟合[M].杭州:浙江大学出版社,2005:1-90.
- [4] 孙霞,吴自勤,黄昀.分形原理及应用[M].合肥:中国科学技术大学出版社,2003:23-88.
- [5] Kantelhardt J W, Zschiegner S A, Koscielny - Bunde Eva, et al. Multifractal Drended Fluction Analysis of Nonstationary Time Series[J]. Physica A, 2002, 316(1-4):87-114.
- [6] Kantelhardt J W, Koscielny - Bude Eva, Braun P, et al. Long-term Persistence and Multifractality of Precipitation and River Runoff Records[J]. Journal of Geophysical Research, 2006, 322(1-4):120-137.
- [7] 苑莹,庄新田.国际汇率的多重分形消除趋势波动分析[J].管理科学,2007(8):80-82.
- [8] 施锡铨,艾克风.股票市场风险的多重分形分析[J].统计研究,2004(9):33-35.
- [9] 江田汗.我国近百年气温的非趋势波动分析[J].气象科学,2004(6):200-204.
- [10] 江田汗.全球气温变化的多重分形谱[J].热带气象学报,2004(12):673-678.
- [11] 朱良燕,毛军军.合肥市降水变化趋势分形特征分析与预测[J].计算机技术与发展,2009,19(9):17-19.
- [12] 卢方元.中国股市收益率的多重分形分析[J].系统工程理论与实践,2004(6):50-52.
- [2] Agrawal R, Gunopulos D, Leymann F. Mining process models from workflow logs[C]//Proc. of the 6th Inter. Conf. on Extending Database Technology (EDBT). Valencia, Spain: Expanded version available as IBM Research Report, 1998.
- [3] Stohr E A, Leon Zhao J. Workflow automation: overview and research issues[J]. Information Systems Frontiers, 2001, 3(3):281-296.
- [4] 曾庆田.过程挖掘的研究现状与问题综述[J].系统仿真学报,2007,19(16):275-280.
- [5] 袁崇义. Petri 网原理与应用[M]. 北京:电子工业出版社,2005.
- [6] 李小白,陈攸跻,张塞红.工作流挖掘技术研究综述[J].电脑知识与技术,2009,5(6):1283-1286.
- [7] Herbst J, Karagiannis D. Workflow mining with In WoLvE[J]. Computers in Industry, 2004(53):245-264.
- [8] Schinam G. Mining exact models of concurrent workflows[J]. Computers in Industry, 2004(53):265-281.
- [9] Wainer J, Kim K, Ellis C A. A workflow mining method through model rewriting[C]//CRIWG 2005, Lecture Notes in Computer Science. [s.l.]:[s.n.], 2005:184-191.
- [10] Jonathan E, Alexander L. Discovering models of software processes from event-based data[J]. ACM Transactions on Software Engineering and Methodology, 1998, 7(3):215-249.
- [11] 李嘉菲,刘大有,杨博.过程挖掘中一种能发现重复任务的扩展 α 算法[J].计算机学报,2007,30(8):1437-1439.
- [12] Dumas M, Van der Aalst W. Process-Aware Information Systems[M]. 王建民,译.北京:清华大学出版社,2009:174-180.
- [13] 潘海兰.一种建模的新技术:过程挖掘[J].上海第二工业大学学报,2006,23(2):127-132.
- [14] Weijters A, Van der Aalst W M P. Rediscovering Workflow Models from Event-Based Data Using Little Thumb[J]. Integrated Computer-Aided Engineering, 2003, 10(2):151-162.

(上接第 64 页)