

# 基于小波变换的气温趋势和分形特征分析

毛军军<sup>1,2</sup>, 朱良燕<sup>1,2</sup>, 张学友<sup>1,2</sup>, 吴涛<sup>1,2,3</sup>

(1. 安徽大学数学科学学院, 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, 合肥 230039;

3. 南京大学计算机软件新技术国家重点实验室, 南京 210093)

**摘要:** 将小波变换、多重分形和 R/S 分析法相结合, 选择 db4 小波函数进行小波变换, 利用分解与重构算法, 从分形维数、非周期循环长度、奇异性指数、多重分形谱等方面探讨合肥市近 54 年来气温系统的动力学分形特性和演化规律, 分析合肥市气温趋势的持续性、长期记忆性、统计自相似性和多重分形性。分析结果表明, 合肥市气温变化过程是一个层次分明的过程, 具有多标度结构且 1 月、8 月、11 月的分形维数较高, 气温不规则程度高, 复杂性高。

**关键词:** 小波变换; 分形特征; Hurst 指数; 多重分形

## Analysis of Temperature Trend and Fractal Feature Based on Wavelet Transform

MAO Jun-jun<sup>1,2</sup>, ZHU Liang-yan<sup>1,2</sup>, ZHANG Xue-you<sup>1,2</sup>, WU-tao<sup>1,2,3</sup>

(1. School of Mathematical Sciences, Anhui University, Hefei 230039; 2. Key Laboratory of Intelligent Computing & Signal Processing of Ministry

of Education, Anhui University, Hefei 230039; 3. State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

**【Abstract】** This paper analyses the dynamic characteristics and evolutionary principle of Hefei's temperature over the past fifty-four years, it combine R/S analysis with wavelet transform, choosing db4 wavelet function, deconstruct and reconstruct the original data. It analyses the temperature based on multi-fractal analysis, discusses the stability, long-term memory, statistical self-similarity and multi-fractal of the temperature trend in Hefei from fractal dimension, non-periodic length of circulation, singular index, and multi-fractal spectrum. Analysis results show that the temperature change in Hefei whose level is clear has a multiple-scale structure, the fractal dimension of January, August, November shows that the irregular degrees and the complexities of temperatures are higher than the others.

**【Key words】** wavelet transform; fractal feature; Hurst exponent; multi-fractal

### 1 概述

气候作为人类赖以生存的自然环境的一个重要组成部分, 它的任何变化都会对自然生态系统以及社会经济系统产生影响。复杂的天气过程往往产生纷杂且不规则的信息数据, 而气象数据的奇异性往往蕴涵着丰富的动力系统信息。近年来, 基于配分函数的多重分形形式和 R/S 分析法已广泛地应用于城市气候趋势研究<sup>[1-2]</sup>。它不仅能提供对不同系统的奇异性的一个完整描述, 并能用来分析和测量非线性系统的非周期循环周期。但是由于该技术内在的统计属性, 使其只能提供信号奇异性的全局描述; 并且用这种方法计算奇异性指数时, 都采用了平均或加权平均的平滑技术, 这样就可能掩盖气象数据中的有用信息。

本文将小波变换和 R/S 分析法相结合, 并利用多重分形研究方法, 针对 50 多年的合肥市的平均气温时间序列, 研究气温系统的动力学分形特性和演化规律。

### 2 理论与方法

#### 2.1 小波变换<sup>[3-4]</sup>

小波分析(Wavelets Analysis, WA)是 20 世纪 80 年代中期发展起来的一门新兴的数学理论和方法, 通过小波变换可以将研究对象分解到不同尺度的空间进行分析和处理, 然后根据需要进行相应的重构。

**定义** 如果函数  $\psi(x)$  是一个平方可积函数, 且 Fourier 变

换满足  $C_\psi = \int_{\mathbb{R}} \frac{|\psi(\omega)|^2}{|\omega|} d\omega < \infty$ , 则称  $\psi(x)$  是一个基本小波或母

小波, 将母小波  $\psi(x)$  经伸缩和平移后, 就可以得到一个小波序列。

对于连续的情况, 小波序列为

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad a, b \in \mathbb{R}; a \neq 0 \quad (1)$$

其中,  $a$  为伸缩因子;  $b$  为平移因子。

对于离散的情况, 小波序列为

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi\left(2^{-\frac{j}{2}} t - k\right) \quad j, k \in \mathbb{Z} \quad (2)$$

由多分辨分析理论, 被处理的信号  $A_{j-1}$  可分解为低一级逼近(低频)信号  $A_j$  和细节(高频)信号  $D_j$  之和。运用 Mallat 算法可以将信号一层层进行分解, 每一层分解的结果是将上次分解得到的逼近(低频)信号再分解成逼近(低频)和细节(高

**基金项目:** 国家自然科学基金资助项目(60675031); 安徽高等学校省级自然科学基金资助项目(KJ2008B093); 中国博士后科学基金资助项目(20070411028); 安徽大学人才队伍建设经费基金资助项目

**作者简介:** 毛军军(1973 - ), 女, 副教授、博士, 主研方向: 人工智能, 信息优化; 朱良燕、张学友, 硕士研究生; 吴涛, 教授、博士

**收稿日期:** 2009-09-30 **E-mail:** maojunjun@ahu.edu.cn

频)2 部分。Mallat 将多分辨分析和小波变换结合在一起,提出了离散信号系数分解与重建的 Mallat 算法:

$$cA_{j+1}^k = \sum_n \bar{h}_{k-2n} cA_j^k$$

$$cD_{j+1}^k = \sum_n \bar{g}_{k-2n} cA_j^k \quad (3)$$

式(3)称为信号的分解算法。

$$cA_j^k = \sum_n cA_{j+1}^k \bar{h}_{k-2n} + \sum_n cD_{j+1}^k \bar{g}_{k-2n} \quad (4)$$

式(4)称为信号的重建算法。

小波的选择不是唯一的,但它的选择应满足定义域紧支撑条件和小波容许条件,小波函数的消失阶也很重要。必须尽可能选择一个最小的消失阶次数而又能保证其检测出信号的最大奇异性指数。经过计算比较,db4 小波函数不仅可以满足算法的条件,而且可以节省内存和计算时间。

## 2.2 R/S 分析法<sup>[5]</sup>

R/S(Rescaled Range Analysis)分析法又称为重标极差法。其计算结果如下:假设  $P$  是长度为  $N$  的时间序列。将这  $N$  个数据划分为  $m$  个子区间,每个子区间的长度为  $n$ 。每个子区间记  $I_a$ , 其中,  $a=1,2,\dots,m$ 。  $I_a$  中的元素为  $N_k$ , 其中,  $k=1,2,\dots,n$ 。子区间  $I_a$  的平均值  $M_a$  计算如下:

$$M_a = \frac{1}{n} \sum_{k=1}^n N_{k,a} \quad (5)$$

每个子区间  $I_a$  相对平均值的累计离差  $X_{k,a}$  为

$$X_{k,a} = \sum_{i=1}^k (N_{i,a} - M_a), k=1,2,\dots,n \quad (6)$$

定义极差  $RI_a$  为子区间中最大值和最小值间的差值:

$$RI_a = \max(X_{k,a}) - \min(X_{k,a}), 1 \leq k \leq n \quad (7)$$

子区间  $I_a$  的标准差记为  $SI_a$ , 计算方法如下:

$$SI_a = \sqrt{\frac{1}{n} \sum_{k=1}^n (N_{k,a} - M_a)^2} \quad (8)$$

由每个子区间的  $RI_a$  和  $SI_a$ , 计算重标极差  $FI_a = RI_a / SI_a$ , 由于划分了  $m$  个连续子区间,因此必须计算 1 个平均的重标极差  $F_n$ , 计算方法如下:

$$F_n = \frac{1}{m} \sum_{a=1}^m \left( \frac{RI_a}{SI_a} \right) \quad (9)$$

通过增加  $n$  的数值,  $m$  的值减少,直到  $n=(m-1)/2$ 。当然,  $n$  要能整除  $(m-1)$ 。

Hurst 指数  $H$  关系式如下:

$$F_n = c \times n^H \quad (10)$$

对式(10)端取对数可得:

$$\lg(F_n) = \lg(c) + H \times \lg(n) \quad (11)$$

通过对式(11)用最小二乘法拟合求得  $H$  估算值。

$H$  指数的取值一般在 0~1 之间。 $H$  的大小可以揭示许多重要的非线性动力学特性和演化规律。

本文在上述  $H$  指数算法的基础上,加入了最大非周期循环长度的寻找,在算出  $H$  指数的同时也算出了非循环周期的长度。具体算法如下:

算出的平均重标极差  $F_n$  有  $q$  个,记为  $F_n(1), F_n(2), \dots, F_n(q)$ , 令

$$X(j) = |F_n(j+1) - F_n(j)|, j=1,2,\dots,q-1 \quad (12)$$

因此,可以得到  $X(j)$  的最大值  $\max_{1 \leq j \leq q-1} X(j)$ , 记为  $X(r)$ , 其中的  $r$  即为非周期循环的长度,也就是说系统平均经过  $r$  天后失去对初始条件的记忆。

## 2.3 多重分形谱分析

设时间序列为  $\{x_i\}$ ,  $i=1,2,\dots,T$ , 按大小  $\varepsilon$  ( $\varepsilon < 1$ ) 均匀等分序列,份数为  $N$ , 令

$$P_j(\varepsilon) = I_j(\varepsilon) / \sum I_j(\varepsilon), j=1,2,\dots,N \quad (13)$$

其中,  $I_j(\varepsilon)$  为大小为  $\varepsilon$  时第  $j$  份内各数之和。

定义配分函数:

$$\chi_q(\varepsilon) = \sum P_j^q(\varepsilon) = \varepsilon^{\tau(q)} \quad (14)$$

从  $\tau(q)$  中可以计算出多重分形谱  $f(\alpha)$ , 其计算公式如下:

$$\frac{d}{dq} [\tau(q)] = \alpha(q) \quad (15)$$

$$\tau(q) = q\alpha(q) - f(\alpha) \quad (16)$$

多重分形的奇异性指数  $\alpha$  描述每一次分割形成的质量分布概率子集的大小。分形谱的宽度  $\Delta\alpha = \alpha_{\max} - \alpha_{\min}$  表征最大、最小概率间的差别,即概率变化的不均匀性,表明多重分形的强弱变化程度。相应地最大、最小概率子集分形维数的差别  $\Delta f = f(\alpha_{\min}) - f(\alpha_{\max})$  反映高低位出现频率的变化。

## 3 实例分析

### 3.1 气温的时序变化趋势

本文选取安徽地区的合肥气象站自 1953 年~2007 年气象资料,使用数据来源于中国气象资料共享网。气象要素为平均气温,季节划分为春季(3 月~5 月)、夏季(6 月~8 月)、秋季(9 月~11 月)、冬季(12 月~次年 2 月)。

图 1 为合肥市 1953 年~2007 年春、夏、秋、冬气温时序变化趋势。

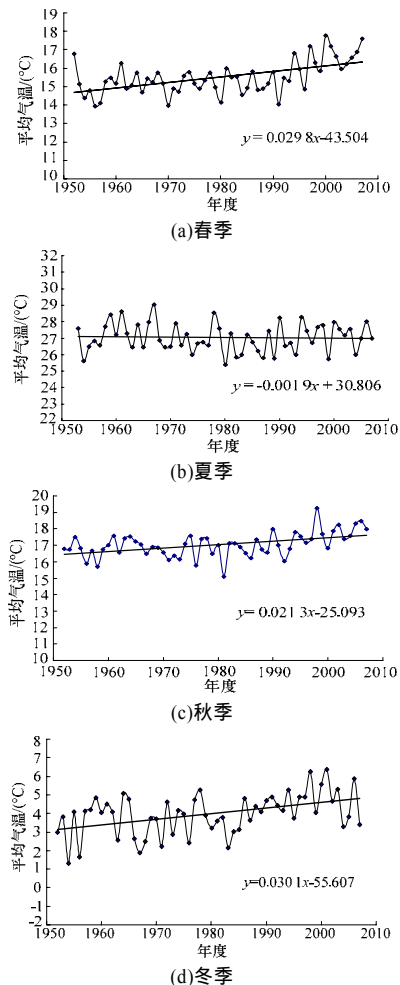


图 1 合肥市 1953 年~2007 年 4 季气温时序变化趋势

从图中可以看出，春、秋、冬 3 个季度气温有明显的上升趋势。夏季气温不具有上升趋势，反而有稍微的下降趋势。春、夏、秋、冬 4 个季度均具有周期性震荡。震荡周期、幅度均不同，以冬季震荡最明显。

### 3.2 小波变换与 Hurst 指数分析

利用 Matlab 软件中的小波工具箱分别进行 3 层 Daubechies 小波分解，然后保留低频部分，将高频部分全置为 0 进行重构。编程环境为 Matlab7.0，小波函数选择为 db4。实验结果见表 1。

表 1 原数据与重构数据分形特征统计

| 时间<br>标度 | 原数据      |      | 小波变换重构  |          |      |
|----------|----------|------|---------|----------|------|
|          | Hurst 指数 | 周期/年 | 分形维数    | Hurst 指数 | 周期/年 |
| 1 月      | 0.484 9  | 16   | 2.137 1 | 0.676 8  | 15   |
| 2 月      | 0.505 0  | 19   | 1.960 2 | 0.685 7  | 17   |
| 3 月      | 0.555 8  | 21   | 1.877 0 | 0.682 0  | 14   |
| 4 月      | 0.498 8  | 21   | 1.981 7 | 0.668 6  | 22   |
| 5 月      | 0.539 8  | 18   | 1.964 8 | 0.684 4  | 23   |
| 6 月      | 0.519 3  | 11   | 1.932 5 | 0.688 7  | 11   |
| 7 月      | 0.517 8  | 15   | 1.882 3 | 0.692 8  | 11   |
| 8 月      | 0.504 5  | 13   | 2.014 7 | 0.685 1  | 10   |
| 9 月      | 0.520 6  | 19   | 1.789 5 | 0.676 5  | 22   |
| 10 月     | 0.514 9  | 20   | 1.991 4 | 0.683 1  | 18   |
| 11 月     | 0.404 3  | 24   | 2.056 9 | 0.669 1  | 23   |
| 12 月     | 0.533 0  | 21   | 1.876 1 | 0.678 1  | 16   |
| 春        | 0.513 1  | 16   | 1.852 3 | 0.682 2  | 19   |
| 夏        | 0.516 9  | 20   | 1.923 5 | 0.687 1  | 11   |
| 秋        | 0.507 0  | 19   | 1.994 7 | 0.683 2  | 15   |
| 冬        | 0.517 1  | 23   | 1.893 5 | 0.682 0  | 18   |
| 年        | 0.605 5  | 21   | 1.777 4 | 0.680 3  | 11   |

从实验结果可看出：

(1)分形维数代表非规则性，无序、弯曲或填充空间的能力。一般地，维数越大，弯曲越大，不规则程度越大，复杂性越高。可看出合肥市在 1 月、8 月、11 月的分形维数较大。说明这 3 个月在近 50 年中，气温不规则程度高，复杂性高。

(2)经过小波变换重构后的 Hurst 指数明显高于原始数据，这说明文中引入小波重构，确实消除了数据中的噪声，提高了信息数据的质量，提取了有效信息，得到了更加符合实际生活的指数。

(3)从总体来看，Hurst 指数都接近 0.7，大于 0.5，具有持续性，即过去的整体变暖趋势在未来整体仍然将变暖，与实际相符。20 世纪 80 年代以来，气候变化已成为全球关注的热点问题，联合国政府间气候变化专门委员会第 4 次评估报告指出：21 世纪末气温升高幅度为 1.1°C~6.4°C。

(4)不同标度下，非周期循环长度基本相同，这也体现了系统的分形自相似特性。

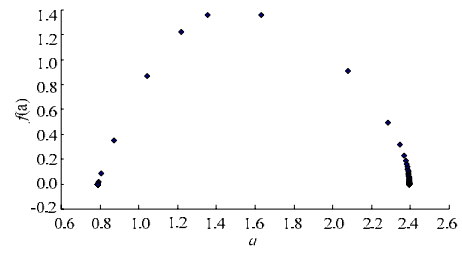
(5)从春夏秋冬 4 季来看，夏季的指数最高，说明未来夏季的变凉趋势维持，春季、秋季和冬季都为 0.68 左右，说明未来 3 季仍然变暖。即夏季越来越凉，春秋冬季越来越暖。

### 3.3 多重分形谱分析

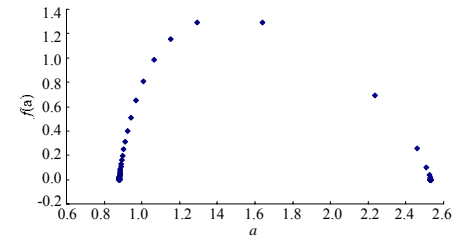
利用多重分形分析法将合肥市自 1953 年~2007 年的 54 年间时间标度为天、月、年的平均气温时间序列作为研究样本。

从以上实验结果可以看出：

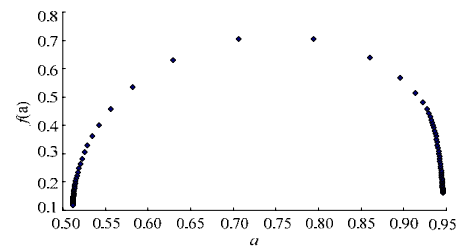
(1)如图 2 所示( $q=-50:50$ )不同的时间标度下奇异谱曲线具有相似的形状，基本呈单峰形状，说明多重分形的形状不随时间标度的改变而改变，也就说明了合肥市气温变化过程确实是一个层次分明的过程，具有多标度结构。



(a)时间标度为天



(b)时间标度为月



(c)时间标度为年

图 2 3 种不同时间标度下  $f(\alpha)-\alpha$  曲线

(2)根据表 2，奇异谱宽度( $\Delta(\alpha)$ )随着标度的减小而缩短，说明标度的细化，相应的分形维的离差( $\Delta f$ )增加，即，多重分形的强弱程度在增强。根据表 3，也可得到，当时间标度不变时， $q$  值增大，奇异谱的宽度和分形维的离差均增大，多重分形也随之增强。

表 2 分形谱宽度( $\Delta\alpha$ )和分形离差( $\Delta f$ )随时间标度的变化

| 时间标度           | 天       | 月       | 季度      | 年       |
|----------------|---------|---------|---------|---------|
| $\Delta\alpha$ | 1.226 9 | 1.223 9 | 0.962 7 | 0.434 5 |
| $\Delta f$     | 1.356 4 | 1.292 4 | 1.184 6 | 0.587 3 |

表 3 分形谱宽度( $\Delta\alpha$ )和分形离差( $\Delta f$ )随  $q$  值的变化

| 时间标度           | 年         |           |           |
|----------------|-----------|-----------|-----------|
| $\Delta\alpha$ | 0.427 4   | 0.431 9   | 0.434 5   |
| $\Delta f$     | 0.533 6   | 0.560 5   | 0.58 73   |
| $q$            | [-20, 20] | [-30, 30] | [-50, 50] |

(3)通过对合肥市平均气温的时间序列多重分形的研究，它所反映出的特点可以说明：多重分形的特点可以真实地反映气温以及气候的复杂性。虽然它不能确切地预测未来的气候如何，但它的确也是对气候的一种更符合实际的描述。

## 4 结束语

本文将小波变换和重标极差法相结合，选择 db4 小波函数进行小波变换，利用分解与重构算法，并结合多重分形的研究方法，针对合肥市近 54 年来气温系统的动力学分形特性和演化规律，并从分形维数、非周期循环长度、奇异性指数等方面探讨合肥市气温趋势的持续性、长期记忆性、统计自相似性和多重分形性，得出以下结论：(1)小波重构确实消除了数据中的噪声，提高了信息数据的质量；(2)1 月、8 月、11 月的分形维数较高。说明这 3 个月在近 54 年中，气温不规则程度高，复杂性高。(3)过去的整体变暖趋势在未来整体

(下转第 24 页)