

自适应多分辨数据滤波器及其 在我国气候诊断中的应用*

尤卫红 林振山 邓自旺

(南京大学大气科学系, 南京 210093)

摘 要 本文以小波分析的多分辨分析理论为基础, 建立了一个自适应的多分辨数据滤波器。该滤波器不仅具有传统方法所不具有的局部化、多层次、多分辨的优点, 而且对于突变点的判断也象 M-K 方法一样的精确。此外, 该滤波器还具有可以同时判别多个突变点而不改变原时间序列的大尺度结构的优点。本文的应用表明: 中国近百年来气温变化主要经历了三个持续近 30 年的冷暖交替, 它们分别是 1919 年以前的偏冷期; 1920 年至 1954 年的偏暖期和 1955 年至 1986 年的偏冷期。对应于这种较大时间尺度的气候演变, 中国的气温变化表现出了十分明显的突变特性, 其冷暖交替的突变点分别发生在 1920 年、1955 年和 1987 年。对应于较小时间尺度的气候演变来说, 中国近百年来气温变化则增加了更多冷暖交替的层次结构和突变点。云南近百年来气温变化与中国的气温变化是基本一致的, 但与北半球和全球的气温变化并不完全一致, 其中最主要的差异是北半球和全球的气温变化在 1955 年至 1978 年是一个极弱的偏冷期, 1979 年发生明显增暖的突变, 而云南和中国的气温变化在 1955 年至 1986 年则是一个极强的偏冷期, 从 1987 年开始才出现明显增暖的突变。

关键词 自适应多分辨 固定尺度窗 气候诊断

1 引言

在对气候或气象数据的分析处理中, 当只使用固定时间尺度的低通滤波器时, 在消去噪音和小尺度波动的同时也平滑了与大尺度振幅有关的激烈变化与突变, 显然这对于分析气候的激烈变化与突变以及研究动力系统波形之间的突然变化等是很不合适的。此外^[1~2], 传统的滤波器以及 Gramer, Yamamoto 和 Mann-Kendall 等方法因不具有任何时空分辨的缺点, 故无法诊断出系统的内在层次结构。因此这就提出了建立一个既能保留大尺度变化的突变特性而又能消去小尺度变化的方差和噪音的自适应多分辨低通滤波器的必要性。

本文将根据小波理论的多分辨分析^[3~5]和自适应的滤波技术^[6~9]来建立一个自适应的多分辨数据滤波器, 即根据一个可移动的窗内数据的多分辨分解和重构来建立滤波器。为了便于与一般的滑动平均的低通滤波器的数据分析比较, 这里的窗宽是以某一尺度固定的, 并通过限制窗内的方差保留量来实现自适应的多分辨分解和重构, 而窗内的每一个数据都是由它在窗内的位置和方差来局部决定的。

本文的第二部分给出了固定尺度窗的分解与重构方法; 第三部分给出了由方差保留

1995-03-06 收到, 1996-02-10 收到再改稿

* 本文为国家“攀登计划”课题和云南省科委应用基础研究基金资助项目

量和移动窗来控制的自适应多分辨滤波输出;第四部分作为该滤波器的应用例子我们分析了中国近百年来气温变化的层次结构和突变特征;第五部分作为该滤波器的进一步应用,我们分析了云南近百年来气温变化的层次结构和突变特征,并与中国、北半球和全球的气温变化情况进行了比较。

2 固定尺度窗的分解与重构

设气象数据的时间序列为:

$$x_i = x(t_i), \quad i = 1, 2, \dots, N, \quad (1)$$

其中 N 为数据总数,数据的采样间隔为 Δt 。

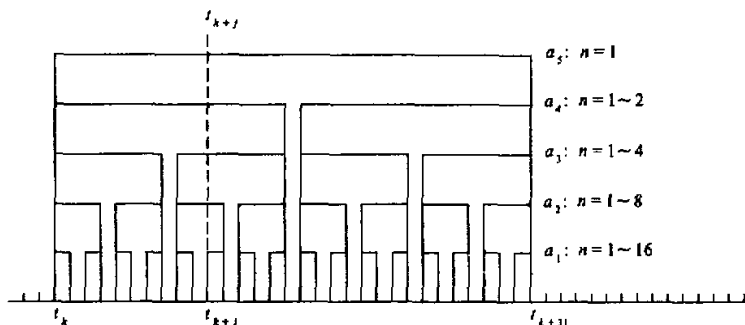


图 1 从 t_k 开始的由 32 个数据确定的固定尺度窗的多分辨分解示意图

见图 1, 设固定的尺度窗是由从 t_k 开始的 $L = 2^M$ 个顺序数据确定的, 其中 M 为大于 1 的整数。显然, 低通信息是由窗内所有数据的平均

$$\bar{x}_k(a_M; 1) = \frac{1}{L} \sum_{i=0}^{L-1} x(t_k + i) \quad (2)$$

来表示的。这里我们已引入通用符号

$$\left. \begin{array}{l} x_k(a_m; n) \\ \bar{x}_k(a_m; n) \end{array} \right\} m = M, M-1, \dots, 1; \quad n = 1, 2, \dots, 2^{M-m} \quad (3)$$

来表示从 t_k 开始的由 2^m 个数据构成的第 n 个不重叠子窗中的 2^m 个数据(图 1)和这 2^m 个数据的算术平均。当 $m = M$ 时, $n = 1$, 这时(3)式分别对应于从 t_k 开始的整个固定窗内的 $L = 2^M$ 个数据和这 L 个数据的平均即(2)式。作为第一步的分解, 可将(2)式改写为

$$\bar{x}_k(a_M; 1) = \frac{1}{2} [\bar{x}_k(a_{M-1}; 1) + \bar{x}_k(a_{M-1}; 2)], \quad (4)$$

显然如果我们取差分

$$\Delta x_k(a_M; 1) = \frac{1}{2} [\bar{x}_k(a_{M-1}; 2) - \bar{x}_k(a_{M-1}; 1)], \quad (5)$$

那么我们有第一步的重构

$$\bar{x}_k(a_{M-1}; 1) = \bar{x}_k(a_M; 1) - \Delta x_k(a_M; 1), \quad (6)$$

$$\bar{x}_k(a_{M-1}; 2) = \bar{x}_k(a_M; 1) + \Delta x_k(a_M; 1). \quad (7)$$

这样的分解与重构可一直进行到最后一步(即第 M 步)。显然这时的重构已变为对固定尺度窗中任一数据的重构,而重构的形式则是由该数据在窗中的位置所决定的。换句话说,为了把窗内不同位置上的原始数据变换为按不同尺度计算的梯度,窗中的数据是根据不同尺度的变换(分解)来重构的,在图1中这对应于通过某数据点 x_{k+j} 作一垂线,与该垂线相交的所有变换(子窗)被用于重构该数据点。因此,具有 $L = 2^M$ 个数据的固定尺度窗中的数据 x_{k+j} 的重构可表示为

$$x_{k+j} = \bar{x}_k + \sum_{m=1}^M (-1)^l \Delta x_k(a_m; n), \quad (8)$$

式中

$$n = 1 + \text{int}\left(\frac{j}{2^m}\right), \quad (9)$$

$$l = 1 + \text{int}\left(\frac{j}{2^{m-1}}\right), \quad (10)$$

$$\Delta x_k(a_m; n) = \frac{1}{2^m} \sum_{i=0}^{2^m-1} (x_{k+2^m(n-1/2)+j-i} - x_{k+2^m(n-1)+j-i}), \quad (11)$$

这里 $\text{int}()$ 表示取整数部分, $\bar{x}_k$ 为由(2)式表示的窗平均。

固定尺度窗内 2^M 个数据的总样本方差 S 也分解为

$$S = \frac{1}{2^M} \sum_{j=0}^{2^M-1} (x_{k+j} - \bar{x}_k)^2 = \sum_{m=1}^M \frac{1}{2^{M-m}} \sum_{n=1}^{2^{M-m}} [\Delta x_k(a_m; n)]^2. \quad (12)$$

(8)式的求和表示了对窗内所有变换(分解)的求和。但作为对数据的滤波分析,我们总是对求和进行截断的,即用从 $m = m_0$ ($m_0 \geq 1$)到 $m = M$ 的求和来代替从 $m = 1$ 到 $m = M$ 的求和。一般来说,这种决定更小尺度被截断的 a_{m_0} 对于窗内不同位置上的数据应该是不相同的。下面我们就来讨论这种小尺度被截断的多分辨滤波问题。

3 自适应的多分辨滤波输出

在一个固定的尺度窗内,方差(12)式是按变换的差分来分解的。显然,如果我们规定重构结束时的方差保留为一分数 α ($0 < \alpha \leq 1$),那么我们可以通过 α 来决定重构是否结束。现在我们来对从时间 t_k 开始的整个固定窗进行一步一步的自适应多分辨重构。

首先第一步,考虑具有最大绝对值的差分项 $\Delta x_k(a_{m^*}; n^*)$,即满足

$$|\Delta x_k(a_{m^*}; n^*)| = \max |\Delta x_k(a_m; n)| \quad (13)$$

的变换。如果 $m^* < M$,那么所有包含子窗 $x_k(a_{m^*}; n^*)$ 的更大尺度的子窗变换都应在这第一步中被包含,这是由低通滤波的要求所决定的。

第二步,消去第一步已考虑的变换项(13)式,以其余变换项中具有最大绝对值的项作为第二步的多分辨重构,显然包含该最大绝对值项的所有较大尺度的变换项除第一步

中已被包含的外都应在这第二步中被增加。以后的每一步都以同样的方式来增加某些剩余的变换项来逐渐提高局部的分辨率。

显然, 每一步多分辨重构的结束都使重构的窗内方差总量得到修正。如果修正后的方差总量已大于规定的方差保留量 α , 即

$$S' \geq \alpha S, \quad (14)$$

其中 S' 为修正后的方差总量, S 为(12)式表示的原有方差。那么多分辨重构就宣告结束, 否则继续按前述多分辨重构方法进行下去, 直到满足(14)式为止。

第一步结束时的方差为

$$S' = \sum_{m=m^*}^M \frac{1}{2^{M-m}} [\Delta x_k(a_m; n')]^2, \quad (15)$$

式中

$$n' = 1 + \text{int} \left(\frac{n^* - 1}{2^{m-m^*}} \right), \quad (16)$$

这里的 m^* 和 n^* 由(13)式确定。一般对于某一步结束时的修正方差 S' , 其最简单的计算方法是将该步为止的所有没有被包含的剩余变换令为“0”, 然后再用(12)式对所有的变换进行计算。

当修正方差满足(14)式时, 窗内位置 $k+j$ 上的最小重构尺度的对应指数为函数

$$m_0 = m_0(j, k; \alpha), \quad (17)$$

因此通过把(8)式修改为从 m_0 ($1 \leq m_0 \leq M$) 开始的求和, 就得到了对数据 x_{k+j} 的修正值

$$\tilde{x}_{k+j} = \bar{x}_k + \sum_{m=m_0}^M (-1)^l \Delta x_k(a_m; n), \quad (18)$$

其中 n 和 l 的表达式与(9)和(10)式相同。显然, 如果 $m_0(j, k; \alpha) = 1$, 那么原始数据的重构是精确的(没有被截断); 如果 $m_0(j, k; \alpha) = M$, 那么重构完全被简化为半窗的平均。一般, $\tilde{x}_{k+j}$ 的值是等于 2^{m_0-1} 个连续数据的平均。

对于全部数据记录的自适应多分辨滤波器的输出, 是通过移动这个固定尺度窗来实现的。这时所有可能位置上的重构窗都要考虑, 显然这对应于考虑由数据总数 N 和窗宽 $L = 2^M$ 所确定的不完全重叠窗的最大数目。因此除端点外, 时间序列中的每一个数据都不止一次被重构, 有不只一个输出, 但作为最后的滤波输出结果, 可取它们的平均值

$$\hat{x}_j = \frac{1}{K_j} \sum_{k=k_1}^{k_2} \tilde{x}_{k+j}, \quad (19)$$

式中

$$j = i - k, \quad (20)$$

$$k_1 = \max(i - 2^M, 0), \quad (21)$$

$$k_2 = \min(i - 1, N - 2^M), \quad (22)$$

$$K_j = \min(2^M, i, N - i + 1, N - 2^M + 1). \quad (23)$$

(19)式的求和是在包含数据 x_j 的 K_j 个固定尺度窗上进行的。

到此, 我们即得到了一个自适应多分辨数据滤波器的全部建立过程。作为实际应用

例子，下面我们分别用它来分析中国和云南近百年来气温变化的层次结构和突变特性。

4 在中国近百年气温变化分析中的应用

我们以王绍武等^[10~11]给出的中国近百年的年气温距平值序列作为分析对象。为了使分析不受年际气温变化的振荡影响，我们用该序列的 3 年滑动平均结果作为滤波器的输入数据。

在这里我们取 $\alpha=0.8$ ，并分别取 $M=5$ 和 4。显然，这对应于分别选择一个 32 年尺度窗和 16 年尺度窗的重构，并以窗内的方差保留量为 0.8 作为小尺度被截断的控制条件。在图 2 和图 3 中我们分别给出了对这两个滤波器重复使用 3 次的滤波输出结果（粗实线）。

从图 2 中可以看出，中国近百年来气温变化主要经历了三个持续近 30 年的冷暖交替。它们分别是 1919 年以前的偏冷期，1920~1954 年的偏暖期和 1955~1986 年的偏冷期。对应于这种持续近 30 年的冷暖交替，中国的大尺度气候变化表现出了十分明显的突变特性，其冷暖交替的突变点分别出现在 1920、1955 和 1987 年。对应于较小时间尺度的变化来说（图 3），中国近百年来气温变化则表现出了更多相对冷暖交替的

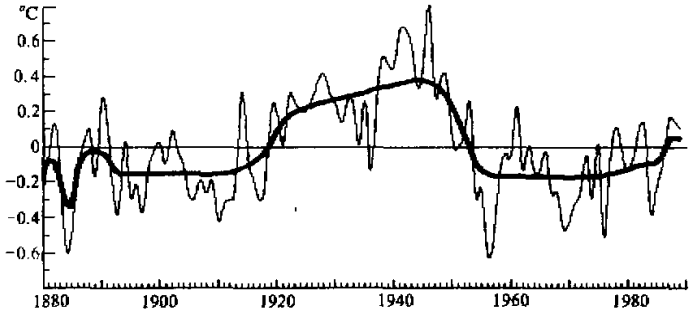


图 2 中国近百年气温变化对应于 32 年固定尺度窗的多分辨滤波输出曲线
 $M=5, \alpha=0.8$ 。图中细线为原始序列曲线

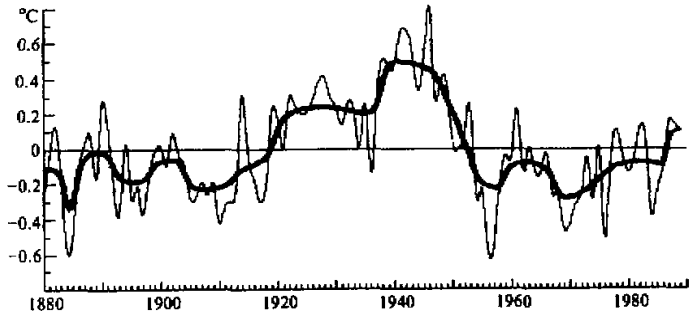


图 3 中国近百年气温变化对应于 16 年固定尺度窗的多分辨滤波输出曲线
 $M=4, \alpha=0.8$ 。图中细线为原始序列曲线

层次演变结构和突变特性。

5 在云南近百年气温变化分析中的应用

我们以昆明的年平均气温距平值序列作为云南的代表, 资料从 1901 年开始到 1995 年。与上节相似, 我们也用该气温距平值序列的 3 年滑动平均结果作为自适应多分辨率数据滤波器的输入数据。这里我们取 $M = 5$, $L = 2^M = 32$, $\alpha = 0.8$ 。在图 4a 中我们给出了这个滤波器对云南年平均气温距平值的 3 年滑动平均序列重复使用 3 次的滤波输出结果。在图 4b、4c 和 4d 中我们也分别给出了这个滤波器对中国、北半球和全球年平均

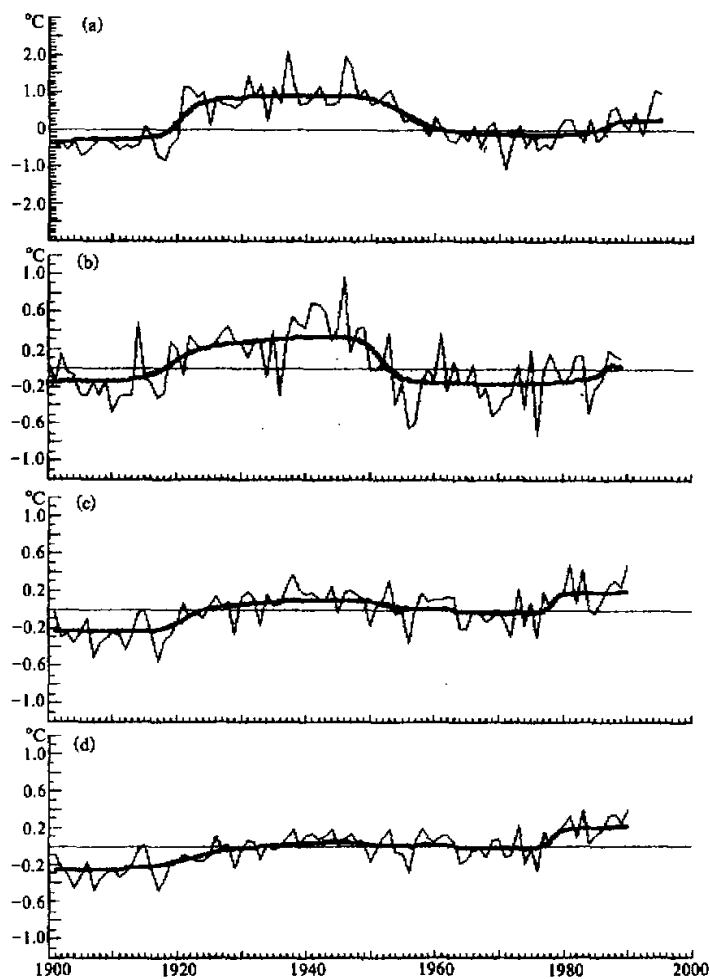


图 4 云南、中国、北半球和全球近百年年平均气温变化的层次结构和突变特征比较
 $M = 5$, $\alpha = 0.8$ 。图中细线为原始序列曲线

(a) 云南, (b) 中国, (c) 北半球, (d) 全球

气温距平值序列^[12,13]作同样处理的滤波输出结果。

从图 4a 中我们可以分析出,云南近百年来气温变化主要经历了四个较大时间尺度的层次演变,它们分别是 1919 年以前的偏冷期;1920 至 1954 年的偏暖期;1955 至 1986 年的偏冷期和 1987 年以后的偏暖期。对应于这种较大时间尺度的层次演变,云南的气温变化也表现出了十分明显的突变特性,其冷暖交替的突变点分别发生在 1920、1955 和 1987 年。

从图 4a 与图 4b、4c 和 4d 的比较我们可以得出,云南近百年来气温变化与中国的气温变化是基本一致的,但与北半球和全球的气温变化并不完全一致,其中最主要的差异是北半球和全球的气温变化在 1955 至 1978 年是一个极弱的偏冷期,1979 年发生明显增暖的突变,而云南和中国的气温变化在 1955 至 1986 年则是一个极强的偏冷期,从 1987 年开始才出现明显增暖的突变。

6 结论

从图 2~图 4 可以看出,本文建立的自适应多分辨数据滤波器具有以下优点:

(1) 同时具有局部化、多层次、多分辨的特点。这是目前所有统计分析技术所不具有的。

(2) 既有接近于 Mann-Kendall 方法对于突变点的精确判断,还可以同时判别多个突变点,并保留有原时间序列的多层次动态结构的优点。

(3) 该滤波器可用于多时间尺度、系统不同层次结构及主突变点和次突变点的分析和诊断。

从图 2~图 4 中我们还得到以下结论:

(1) 中国近百年来气温变化对于较大时间尺度的气候演变来说,主要经历了三个持续近 30 年的冷暖交替。它们分别是 1919 年以前的偏冷期,1920 至 1954 年的偏暖期和 1955 至 1986 年的偏冷期。对应于这种较大时间尺度的气候演变,中国的气温变化表现出了十分明显的突变特性,其冷暖交替的突变点分别发生在 1920、1955 和 1987 年。

(2) 对应于较小时间尺度的气候演变来说,中国近百年来气温变化则增加了更多相对冷暖交替的层次演变结构和突变点。

(3) 云南近百年来气温变化与中国的气温变化是基本一致的,但与北半球和全球的气温变化并不完全一致,其中最主要的差异是北半球和全球的气温变化在 1955 至 1978 年是一个极弱的偏冷期,1979 年发生明显增暖的突变,而云南和中国的气温变化在 1955 至 1986 年则是一个极强的偏冷期,从 1987 年开始才出现明显增暖的突变。

参 考 文 献

- 1 符淙斌、王强, 1992, 气候突变的定义和检测方法, 大学科学, 16(4), 482~493.
- 2 林振山, 1993, 长期预报的相空间理论和模式, 北京: 气象出版社, 96~106.
- 3 刘贵忠、邸双亮, 1992, 小波分析及其应用, 西安: 西安电子科技大学出版社.
- 4 宋国乡、甘小冰, 1993, 数值泛函及小波分析初步, 郑州: 河南科学技术出版社.
- 5 Mallat, S. G., 1989, The theory of multiresolution signal decomposition: The wavelet representation, *IEEE*

- Trans. Pattern Anal. Machine Intell.*, **11**(7), 674~693.
- 6 Keppenne, C. L. and Ghil, M., 1992, Adaptive filtering and prediction of the SOI, *J. Geophys. Res.*, **97**, 20449~20454.
 - 7 Mallat, S. G. and Hwang, W. L., 1992, Singularity detection and processing with wavelets, *IEEE Trans. Inform Theory*, **38**, 617~643.
 - 8 Penland, C., Chil, M. and Weickmann, K. M., 1991, Adaptive filtering and maximum entropy spectra with application to changes in atmospheric angular momentum, *J. Geophys. Res.*, **96**, 22650~22671.
 - 9 Howell, J. F. and Mahrt, L., 1994, An adaptive multiresolution data filter: Applications to turbulence and climatic time series, *J. Atmos. Sci.*, **51**(14), 2165~2178.
 - 10 Yi Yuhong(衣育红) and Wang Shaowu(王绍武), 1992, Abrupt warming of global climate in 1980s, *Chinese Science Bulletin*, **37**, 528~531.
 - 11 王绍武, 1994, 近百年气候变化与变率的诊断研究, *气象学报*, **52**(3), 261~273.
 - 12 Jones, P. D., Raper, S. C. B., Bradley, R. S. et al., 1986, Northern Hemisphere surface air temperature variations, *J. Climate Appl. Meteor.*, **25**, 161~179.
 - 13 Jones, P. D., Raper, S. C. B. and Wigley, T. M. L., 1986, Southern Hemisphere surface air temperature variations, *J. Climate Appl. Meteor.*, **25**, 1215~1230.

An Adaptive Multiresolution Data Filter and Its Application in Climatic Diagnosis in China

You Weihong, Lin Zhenshan and Deng Ziwang

(Department of Atmospheric Sciences Nanjing University, Nanjing 210093)

Abstract In this paper, an adaptive multiresolution data filter has been established by using multiresolute theory of wavelet analysis. This filter not only has advantages such as locality, multiscale, multiresolution, which the traditional methods lack, but also is as accurate to determining catastrophe points as the M-K method. In addition, this filter can determine the several catastrophe points at the same time without smoothing the large-scale structure. The applications results show that the air temperature variations during the last 100 years in China have passed three main cold and warm vicissitudes of lasting 30 years, which include the cold period before the 1919, the warm period from 1920 to 1954 and the cold period from 1955 to 1986. Corresponding to the larger-scale climatic changes, the air temperature variations in China obviously show the catastrophe features whose jumping points occurred in 1920, in 1955 and in 1987. Corresponding to the smaller-scale climatic changes, the air temperature variations in China increase more the hierarchies and jumping points with the cold and warm vicissitudes. Though the temperature variations of Yunnan Province are similar to those of China, they are different from those of the Northern Hemisphere and Globe. Their main differences are that there is a weaker cold period in temperature variations from 1955 to 1978 and a jumping point of warming in 1979 in the Northern Hemisphere and Globe, but there is a stronger cold period from 1955 to 1986 and a jumping point of warming in 1987 in China.

Key words adaptive multiresolution definite scale window climatic diagnosis