

基于搜索平均法的气象观测数据的非线性去噪^{*}

侯 威^{1,2)} 廉 毅¹⁾ 封国林^{3,4)†}

1) 吉林省气象科技研究所气候研究开放实验室, 长春 130062)

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) 国家气候中心气候研究开放实验室, 北京 100081)

4) 扬州大学物理科学与技术学院, 扬州 225002)

(2006 年 5 月 22 日收到, 2006 年 6 月 14 日收到修改稿)

介绍了一种新型基于相空间重构和返回图法的非线性降噪方法——搜索平均法, 首先将该方法应用于被高斯白噪声所污染的 Henon 映射时间序列, 说明根据降噪理论所编写的计算程序的正确性; 然后应用于中国气象局公布全国 435 站 1960—2000 年的逐日气温观测时间序列, 并利用非线性预报方法衡量降噪的效果, 验证了该方法的有效性.

关键词: 返回图法, 非线性降噪, 非线性预报

PACC: 9260X

1. 引 言

近年来, 由确定性系统产生的混沌现象得到了广泛的关注, 在很多学科中也得到了广泛的应用. 但是实际中得到的混沌信号总是不可避免地存在误差或者噪声^[1], 对于数值计算的结果会存在截断误差, 而对实际的测量信号则会存在观测误差及仪器误差. 噪声的存在会淹没混沌信号的部分细节, 给分析和研究带来误差. 为了消除噪声对信号分析和结果解释的影响, 可以从两个不同的方面着手. 首先, 可以设计一些抗噪声的算法提取感兴趣的信息, 如计算 Lyapunov 指数鲁棒性的方法^[2,3]、近似熵^[4,5]、复杂度^[6,7]等; 其二, 通过降低噪声的影响, 然后分析信号中包含的有用信息. 本文即属于第二个方面的研究范围. 实际上, 时间序列分析的一个重要内容就是降噪. 降噪就是将实验采集的时间序列分解为干净的信号和随机的噪声波动两部分, 然后去除噪声便得到干净的信号^[8]. 由于产生在低维确定性系统的混沌信号有着跟噪声一样的宽带谱和快速衰减的自相关函数, 因此对于被噪声污染的混沌信号, 线性滤波方法不仅不能将噪声去除, 反而使得原来的混沌信号发生严重扭曲增加了信号的复杂性^[9,10].

现在的许多用于含有噪声的混沌信号的降噪方法^[11-21]都采用了相似的技术路线, 它们都包括三个方面的内容, 即从观测的时间序列中重构原始的吸引子, 选择一类模型估计局部动力学行为并从统计角度拟合出模型参数, 然后修正观测数据, 使得修正后的观测数据更加和模型一致^[17]. 众所周知, 实际系统基本上都是非线性的, 混沌是相当普遍的动力学行为, 因此对于这样的系统, 必须采用非线性降噪方法^[16,21]. 气候系统是一个复杂系统, 有许多证据表明气候信号普遍是混沌性质的, 因此对于气候资料必须采用适合于混沌信号的降噪方法.

本文介绍了一种新型基于相空间重构 (reconstructed phase space) 和返回图法 (recurrence plots) 的非线性降噪方法——搜索平均法 (searching average method, SAM), 同其他非线性降噪方法相比, SAM 具有节省机时、精度较高等优点. 本文将该方法应用于被高斯白噪声所污染的 Henon 映射时间序列以及中国气象局公布的呼玛站 (站点编号 50353; 东经 126.39°, 北纬 51.43°) 1960—2000 年的逐日气温观测时间序列, 说明了本文根据降噪理论所编写的程序的正确性; 并利用非线性预报方法衡量降噪的效果, 验证了该方法的有效性.

^{*} 国家自然科学基金 (批准号 90411008) 和国家重点发展基础研究项目 (批准号 2006CB400503) 资助的课题.

[†] E-mail: feng_gl@sina.com; fenggl@cma.gov.cn

2. 方法介绍

搜索平均法的核心思想是首先确定数据点 x_i 的最佳局部邻域 W , 从被噪声污染的混沌信号中正确地识别出局部流形, 再求出共有 n 个点处于这个最佳邻域内, 然后在局部将原始信号向流形方向上靠拢, 用这个 n 个点的平均值作为点 x_i 的降噪后的修正值, 以此获得降噪效果.

因此利用搜索平均非线性降噪方法进行去噪时, 必须确定一个局部邻域尺度. 对局部邻域尺度的估计实质上就是要对含有噪声的混沌信号的噪声水平进行正确地估计^[1,2]. 如果局部邻域尺度太小, 由于噪声的存在, 就会使得在局部邻域内的点的分布几乎是均匀的, 从而无法正确地识别原始的流形; 如果局部邻域尺度太大, 又会造成把吸引子不同的分支误认为局部邻域的后果, 同样不能正确地识别原始的流形^[12]. 由此可见, 合理地估计局部邻域的大小是十分重要的, 对提高降噪的质量有着重要的影响. 本文将利用返回图法确定最佳局部邻域尺度, 其具体过程如下^[12,18].

2.1. 局部邻域大小的估计

首先构造返回图. 返回图实际上就是一个距离矩阵, 是一个在复杂的时间序列中寻找隐含相关性的有力工具, 它的一个优点是可以将时间序列自治性等方面的性质可视化. 要构造返回图, 第一步要解决的就是对给定的时间序列 $\{x_1, x_2, \dots, x_N\}$ 进行相空间重构, 通常采用延时坐标法. 这一步的物理含义在于通过相空间重构, 使得重构的动力系统与产生时间序列的原始动力系统在拓扑上等价 (Takens 定理), 这样, 两者的动力学不变量 (比如 Lyapunov 指数, 广义维数等) 是相同的. 因此, 可以通过重构的动力系统获知原始动力系统的有关信息. 而相空间重构的关键在于确定相空间嵌入维数和延时时, 假设嵌入维数为 m , 延时时为 τ , 则延时重构矢量为 $X_n = (x_n, x_{n+\tau}, \dots, x_{n+(m-1)\tau})$, $n = N - (m-1)\tau$. 关于 m 和 τ 的确定, 可参见文献 [22]. 第二步便可以通过计算重构矢量之间的距离构造出返回图, 其方阵为

$$r_{ij} = \Theta\{W - |X_i - X_j|\}, \quad (1)$$

这里 W 是预设的容限水平, $\Theta\{\}$ 是 Heaviside 阶梯函数, 即如果 $x > 0$, 那么 $\Theta\{x\} = 1$, 否则 $\Theta\{x\} = 0$. 对

所有矢量对 X_i 和 X_j , 如果它们之间的距离小于 W , 则矩阵的 (i, j) 元便为 1, 也就是在矩阵 (i, j) 这一点处显示出这一点. 怎样阅读此图见文献 [19, 20].

然后利用返回图分别计算两个标量 $N_p(W)$ 和 $N_\perp(W)$, 它们的含义将在后面介绍. 由返回图, 确定平行于主对角线方向的直方图 $h_i = \sum_{k-j=i} r_{jk}$ 是返回图上的点. 计算出直方图 h_i 后, 确定一个阈值 h_{th} , 比如取所有 h_i 的平均值加上它们的标准偏差的三倍作为阈值, 通过判断满足条件 $h_{i-1} < h_{th}$ 和 $h_i > h_{th}$ 的 h_i 的个数, 即 h_i 大于阈值 h_{th} 的尖峰的个数 $N_p(W)$. 从 $N_p(W)$ 的计算过程可知, 容限水平 W , 即局部邻域尺度的选取对返回图的构造至关重要. 如果 W 取得太小, 噪声就会严重地影响返回图的表现形式, 使得在它上面的点较为均匀随机地分布着. 这样, h_i 就彼此很接近, 使得 $N_p(W)$ 减小. 如果 W 取得太大, 相距很远的点都误认为很靠近, 返回图上的点便很均匀地整齐地分布着, 同样会使得 $N_p(W)$ 减小. 只有适当大小的 W 才能尽量地减小噪声的影响, 突出测量信号中确定性的成分, 使得 h_i 在一定范围内剧烈波动, 这样才能使得 $N_p(W)$ 变大. 因此, 最佳的 W 是要使得 $N_p(W)$ 最大. 这就与搜索平均非线性降噪算法所需确定的局部邻域尺度联系起来了. 为了更容易地确定这个局部邻域尺度, 需再引入一个被称为 W 邻域内平均邻域点数 $N_\perp(W) = \sum_{i,j} r_{ij} / N$ 的量是必要的. 该量实际上是垂直于主对角线方向但不计主对角线上的点的直方图的平均数. 如果局部邻域 W 选择得当, 就会使得含有噪声的信号中的确定性部分掩盖住噪声的行为, 从而使平行于主对角线方向而且具有明显的线条外形的线条 (由返回图上的离散点组成) 的个数, 即 $N_p(W)$ 与平均邻域点数 $N_\perp(W)$ 很接近. 定义如下的指标:

$$g(W) = \frac{|N_\perp(W) - N_p(W)|}{N_\perp(W)}. \quad (2)$$

由此式可以看出, 最佳的局部邻域大小 W 就是要使得 $g(W)$ 最小. 至此也可以明白引入 $N_\perp(W)$ 的物理意义, 实质上就是放大 $g(W)$ 的局部动力学行为, 并对 $g(W)$ 进行正规化处理, 以便容易确定局部邻域尺度 W . 同时要指出的是, 在计算上述各个量时, 需要排除返回图的边缘效应^[12]. 现在分析 $g(W)$ 随局部邻域尺度 W 的变化行为. 对于一个小

的 W 在返回图上除主对角线外,几乎没有点分布在其他地方,因此 $N_{\perp}(W)$ 接近于零. 而此时 $N_p(W)$ 为一个有限值,这样 $g(W)$ 便很大. 对于一个大的 W ,几乎所有的点都出现在返回图上, $N_{\perp}(W)$ 便接近于时间序列长度 N (必须排除返回图的边缘效应,否则不成立). 此时平行于主对角线方向的直方图没有孤立的尖峰,因此 $g(W) \approx 1$. 因此,最佳的局部邻域尺度 W 为 $g(W)$ 曲线最小值处的 W 值. 确定了最佳的局部邻域尺度 W 后,有利于提高降噪效果.

2.2. 搜索平均非线性降噪算法

现在对搜索平均非线性降噪算法的原理和算法进行简单介绍. 该方法的基本思想是从被噪声污染的混沌信号中正确地识别出局部流形,再在局部将原始信号向流形方向上靠拢,对含有噪声的数据进行一个修正,对修正后的数据再进行同样的降噪处理,一般经过几次迭代便可以获得较为干净的数据.

对于时间序列 $\{x_1, x_2, \dots, x_N\}$ 中的每一个 x_i , 首先确定正整数 k 和 l , 得到重构向量 $X_i = (x_{i-k}, \dots, x_{i-l})$, X_i 包含了数据点 x_i 过去和未来的演化信息. 关于 k 和 l 的确定可以参考文献 [23, 24]. 找出 x_i 的所有最近邻点 x_j, x_j 满足

$$\sup\{|x_{j-k} - x_{i-k}|, \dots, |x_{j+l} - x_{i+l}|\} \equiv \|X_j - X_i\|_{\sup} < W. \tag{3}$$

也就是找出从过去的 k 次迭代到未来的 l 次迭代这一个时间段内,时间序列 $\{x_1, x_2, \dots, x_N\}$ 演化轨迹上 x_i 的所有最近邻点. 假定共找到 u_i^W 个,用这 u_i^W 个最近邻点的平均值代替原 x_i 值,作为 x_i 降噪后的修正值,即

$$x_i^{\text{corr}} = \frac{1}{u_i^W} \sum_{u_i} x_j. \tag{4}$$

在这里只对延迟窗口 X_i 的中心点 x_i 进行修正,因为在 X_i 中只有 x_i 处于最佳可控状态,即只有 x_i 的值在稳定或不稳定流形中是固定的. 有关这一替代算法的详细论述可参考文献 [21].

3. 对被高斯白噪声污染的 Henon 映射时间序列降噪

Henon 映射是具有两个参数的平面映射族,即
$$x_1 = 1 + y - ax^2,$$

$$y_1 = bx. \tag{5}$$

对 Henon 映射取初值为 $a = 1.4, b = 0.3, x_1 = 0.2, y_1 = 0.12$,取落在吸引子上的 20000 个点作为标量时间序列,向其中加入高斯白噪声(信噪比为 25),形成人为的含有噪声混沌信号,对此信号估计最佳局部邻域尺度为 $W = 0.26$,如图 1 所示. 然后进行搜索平均非线性降噪,干净的数据和被高斯白噪声所污染的数据以及降噪后的数据时间历程如图 2 所示. 从图 3 可以明显地看出所加的噪声完全湮没原始的细致结构,经过一次降噪即可获得较好的结果.

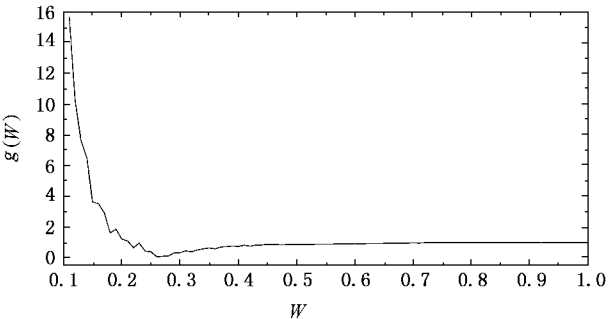


图 1 指标 $g(W)$ 随局部邻域尺度 W 的变化曲线

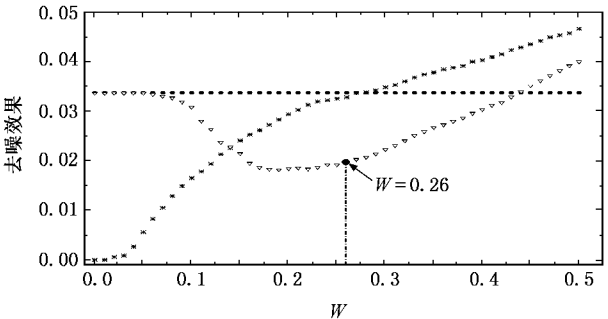


图 2 一次去噪效果“均方根误差”的绝对值随局部邻域尺度 W 的变化曲线(●代表原始噪声的均方根误差,*代表修正值的均方根误差,▽代表去除噪声后信号的均方根误差)

为了验证本文估计局部邻域大小方法的有效性,将用本文方法和文献 [21] 中提出的方法分别计算得到的局部邻域尺度进行了对比. 采用文献 [21] 中的方法,分别计算了原始噪声、一次去噪修正量、剩余噪声的均方根误差(root mean square, RMS),取一次去噪修正量的均方根误差增加缓慢时的 W 值为局部邻域尺度,如图 2 所示,当 $W = 0.25$ 左右时,一次去噪修正量的均方根误差增加缓慢,而用本文方法计算得到的最佳局部邻域为 $W = 0.26$,说明了本文方法的正确性和快捷性.

为了衡量降噪效果,用一个以分贝为单位的增益进行度量,它定义为

$$f = \left[\frac{\sum (y_k - x_k)^2}{\sum (\hat{y}_k - x_k)^2} \right]^{1/2}, \quad (6)$$

式中 x_k 为干净的信号, y_k 是人为的噪声的混沌信

号 \hat{y}_k 是降噪后的信号. 降噪前的增益为 17.79, 经过一次降噪后的增益为 21.58, 从而说明降噪后增益有明显地提高. 从此算例可说明, 在本文中根据搜索平均非线性降噪理论所编写的程序的正确性. 下面将此降噪方法应用于实际气候资料进行降噪.

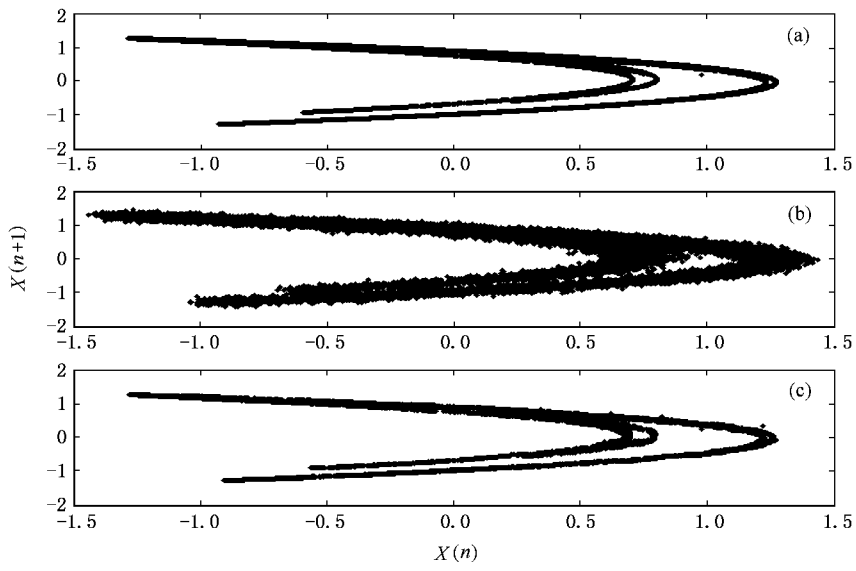


图 3 干净的和被高斯白噪声污染的时间历程以及降噪后的时间历程 (a) 没有加入噪声时的信号 (b) 加入噪声后的信号 (c) 去除噪声后的信号

4. 对实际气候资料序列的降噪

气候系统是一个复杂系统,有许多证据表明气候信号普遍都是混沌性质的,因此对于气候资料必须采用适合于混沌信号的降噪方法. 利用中国气象局公布的全国 720 个常规观测站的观测数据,选取其中包含 1960—2000 年逐日气温观测数据并且无缺测值的站点数据,共 435 个站点的观测数据做为本文的研究资料. 做为示例,我们首先给出了的呼玛站(站点编号 50353;东经 126.39°,北纬 51.43°)温度序列用本文方法进行去噪的结果(图 4,图 5).

同样,首先确定局部邻域尺度,然后进行降噪,连续两次,每降噪一次,就计算一次局部邻域尺度. 图 4 实际温度资料降噪前的时间历程以及降噪后的时间历程,可见降噪后时间序列变得更加光滑,由于去除了噪声成分,因此在数值上去噪后要小于去噪前,但去噪后的数据保持了原始信号的非线性特征,去噪前数据的 Lyapunov 指数为 0.9514384,去噪后的 Lyapunov 指数为 0.9559014,可见去噪处理并没有使得原始信号发生变形.

由于在实验条件下干净的信号是不存在的,不能像前一节那样计算增益. 为了衡量降噪的效果,本文提出对实际温度资料进行非线性预报分析评价降噪效果^[25-26]. 混沌信号的一个特点是它具有短期可预报,而长期不可预报的性质. 这里以正规化预报误差(normalized prediction error, NPE)来量化实际数据去噪前和去噪后的可预报性. 文献 26 给出了关于 NPE 的计算过程和详细说明. 如果 NPE 接近于 0,则说明可预报性很好,如果 NPE 接近于 1,则说明可预报性很差,或不可预报. 此时用序列的平均值便是最好的预报. 很容易知道噪声会使得可预报性能降低,即使得 NPE 值升高. 对实际温度资料在降噪前后的预报如图 5 所示,当预报步数截至为 6 步时,可预报性随着预报步数的增加逐渐减小,6 步之后的可预报性改变不大,逐步趋于收敛. 这同实际预报的理想期限是一致的. 无论是去噪前还是去噪后的数据都有短期可预报,长期不可预报性. 即当预报步数较短时,可预报性较好;当预报步数逐渐增加,可预报性也随之降低,逐渐收敛于 $NPE = 1$,这符合数据本身的混沌性质. 然而在相同的预报步数条件下,降噪后实际温度资料的 NPE 值比降噪前的 NPE

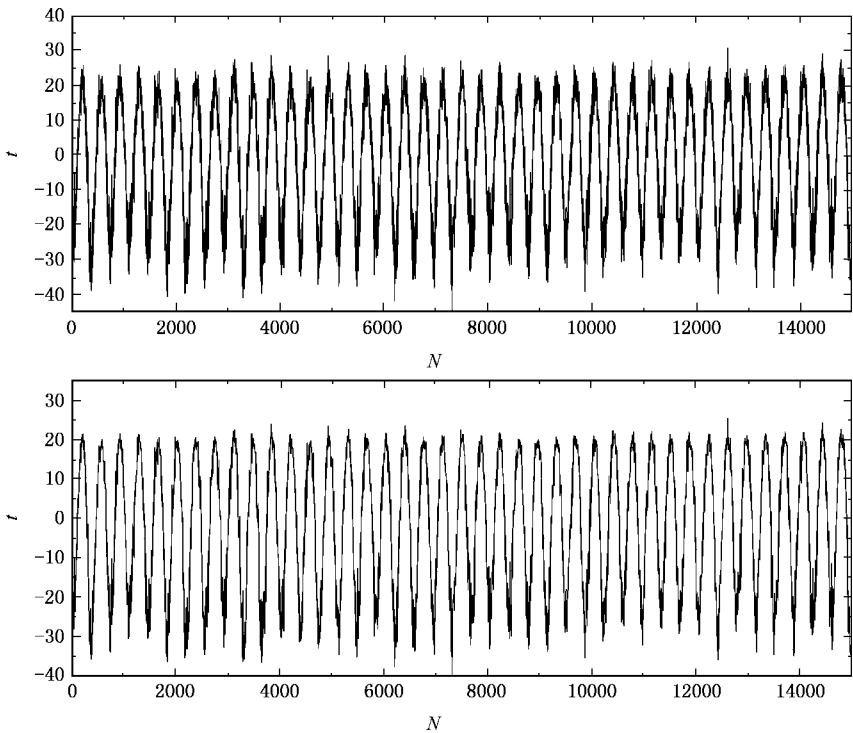


图 4 实际温度资料降噪前的时间历程以及降噪后的时间历程 (a) 去除噪声前的信号 (b) 去除噪声后的信号

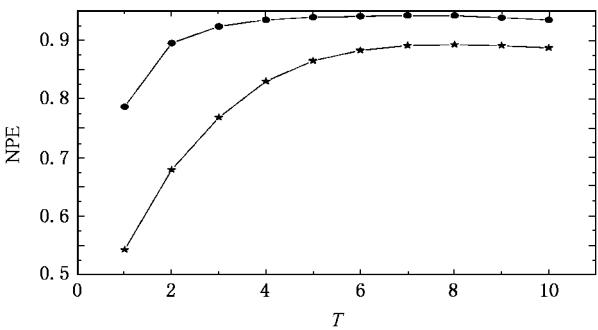


图 5 实际温度资料在降噪前后的正规化预报误差 NPE 随预报步数 T 的变化情况 (●代表去除噪声前的正规化预报误差值, *代表去除噪声后的正规化预报误差值)

值要低许多,这说明降噪处理降低了数据中的噪声成分,增强了实际温度资料的可预报性。

图 6 显示了全国 435 个常规气象站 1960—2000 年逐日气温观测时间序列去噪前 1—2 步正规化预报误差减去去噪后 1—2 步正规化预报误差的差值,即经过去噪处理之后各站点可预报性的增加值。可以看出,在用本文的方法去除噪声之后,当预报步数相同时,所有站点资料的正规化预报误差均有明显减小,可预报性有了较大的提高。当预报步数为 1—2 步时,可预报性的增加由南向北呈递增趋势。

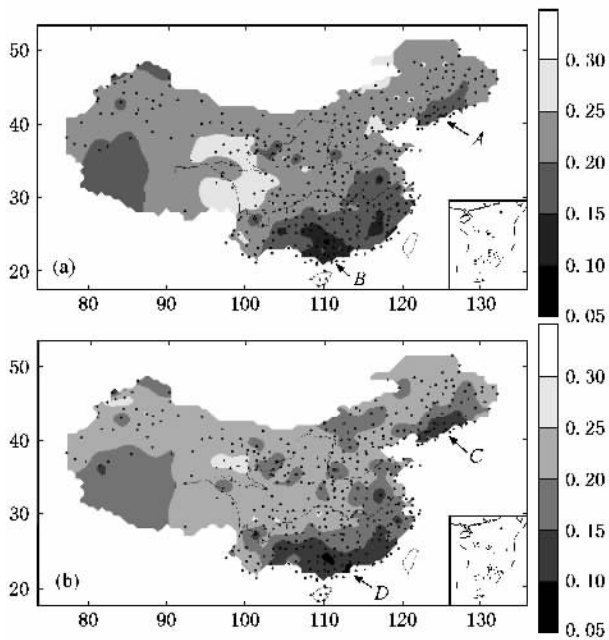


图 6 全国 435 站实际温度资料降噪后可预报性增加值 (a) 去噪前后 1 步正规化预报误差的差值 (b) 去噪前后 2 步正规化预报误差的差值

中国位于亚欧大陆东南部,不仅受东亚季风的显著影响,而且还受到印度季风的作用,是世界上季

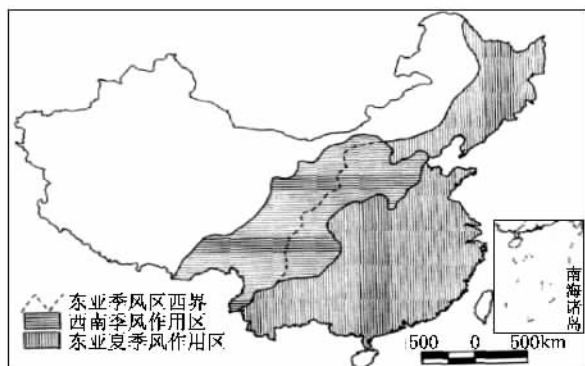


图 7 东亚夏季风和印度西南季风各自影响区域^[31]

风影响显著的国家之一。在文献^[31]中认为东亚季风和印度季风的分界线以 100°E 为界较为适宜,在此以西主要反映印度季风的特征,在此以东主要反映东亚季风的特征。图 7 为中国季风区范围^[31],它大致位于大兴安岭—阴山—贺兰山—日月山—巴颜喀拉山一线以东。东亚季风和印度季风既有各自的活动区域,又有共同重叠的作用空间,西南地区 and 西北地区东部主要受印度季风的影响,华北中南部地区是印度季风和东亚季风共同作用最大的区域;中国东部、南部主要受东亚季风的影响^[27-30, 31]。

对比图 6 和图 7,可以看出二者非常相似。中国南部尤其是华南地区南部(箭头 B 所指)和东北地区南部(箭头 A 所指)增加较小,这两个地区主要受东亚季风的影响,气候系统比较单一,尤其华南南部地区是东亚季风最先影响我国的地区,观测数据中的气候信号较强,随机性成分较少,因此去噪处理对数据的修正较小,可预报性增加不大;中国广大中部地区和华北地区增加较大,而这些地区也正是同时受到东亚季风和印度季风的共同作用,加之不同季风系统之间的非线性相互作用,因此这些地区气候系统比较复杂,包含了相对较强扰动成分,因此观测数据中的随机成分也相对较强,去噪处理的修正较大,可预报性增加也相对较大。另外,由于近代以来中国南部地区和东北地区社会经济较发达,气象观测站点较多,观测设施和观测水平也较其他地区相对较高;广大的西北和西南地区地处欧亚大陆腹地,

主要受西风带气候、高原季风影响,而内蒙古地区绝大部分位于温带,东亚季风仅影响东南部边缘的狭长地带,主要为温带大陆性气候,这些地区一方面由于复杂的气候系统,另一方面由于经济欠发达,加之地理位置和自然环境的限制,观测站点分布稀疏,观测设施和水平也不及发达地区,在一定程度上影响了数据的准确性,这也造成了去噪处理后可预报性增加较大。当预报步数增加为 2 时,如图 6(b)所示,可预报性的提高情况也十分类似于图 6(a),其中可预报性增加较小既去噪处理对数据的修正较小的地区也同样在中国南部和东北的南部(如箭头 C、D 所指)。进一步增加预报步数,全国各地区可预报性的增加情况类似于图 6(图略)。

对比图 6(a)和(b),可以看出当预报步数为 2 时,相同地区可预报性的增加值小于预报步数为 1 时的情况。说明随着预报步数的增加,正规化预报误差有增加的趋势,愈来愈接近于不可预报,也正表明了气候系统的混沌性质。

以上分析表明,本文所用的搜索平均非线性降噪方法对实际温度资料的处理也同样有着较好的降噪效果,可以显著提高数据的可预报性,这对于提高数值预报的时限和精度有着重要意义。

5. 结 论

本文介绍了一种新型基于相空间重构和返回图法的非线性降噪方法——搜索平均法(searching average method, SAM),用返回图法估计了最佳的局部邻域尺度,不再试探局部邻域尺度的大小,或者仅仅凭借肉眼观测估计,这就显著地提高了降噪效果。同其他非线性降噪方法相比, SAM 具有节省机时、精度较高等优点。本文首先将该方法应用于被高斯白噪声所污染的 Henon 映射时间序列;然后应用于中国气象局公布全国 435 站 1960—2000 年的逐日气温观测时间序列,并利用非线性预报方法衡量降噪的效果,验证了该方法的有效性,为实际温度资料的进一步处理创造了条件。

[1] Schreiber T 1993 *Phys. Rev. E* **48** R13

[2] Tanaka N ,Okamoto H , Natio M 1999 *Chaos* **9** 436

[3] Kantz H 1994 *Phys. Lett. A* **185** 77

[4] Yang S Q ,Zhang X H ,Zhao C A 2000 *Acta Phys. Sin.* **49** 636 (in Chinese) [杨绍清、章新华、赵长安 2000 物理学报 **49** 636]

[5] Pincus S M ,Gladstone I M , Ehrenkranz R A 1991 *Journal of Clinical Monitoring* **7** 335

[6] Hou W , Feng G L ,Gao X Q , Chou J F 2005 *Acta Phys. Sin.* **54** 2441 (in Chinese) [侯 威、封国林、高新全、丑纪范 2005 物理学报 **54** 2441]

[7] Hou W , Feng G L ,Dong W J *Acta Phys. Sin.* 2005 **54** 3940 (in Chinese) [侯 威、封国林、董文杰 2005 物理学报 **54** 3940]

[8] Matassini L ,Kantz H ,Holyst J , Hegger R 2002 *Phys. Rev. E* **65** 021102

[9] Kantz H , Schreiber T 1997 *Nonlineartimeseriesanalysis* (Cambridge Univ. Press)

[10] Schreiber T , Grassberger P 1991 *Phys. Lett. A* **160** 411

[11] Grassberger P ,Hegger R ,Kantz H ,Schaffrath C , Schreiber T 1993 *Chaos* **3** 127

[12] Rowlands G , Sprott J C 1992 *Physica D* **58** 251

[13] Kostelich E J , Schreiber T 1993 *Phys. Rev. E* **48** 1752

[14] Hammel S M 1990 *Phys. Lett. A* **14** 8421

[15] Kantz H ,Schreiber T , Hoffmann I 1993 *Phys. Rev. E* **48** 1529

[16] Farmer J D , Sidorowich J J 1991 *Physica D* **47** 373

[17] Sauer T 1992 *Physica D* **58** 193

[18] Xie Y ,Xu J X ,Kang Y M *et al* 2003 *Acta Phys. Sin.* **52** 1121 (in Chinese) [谢 勇、徐健学、康艳梅等 2003 物理学报 **52** 1121]

[19] Zbilut J P , Webber Jr C L 1992 *Phys. Lett. A* **171** 199

[20] Webber Jr C L , Zbilut J P 1994 *J. Appl. Phys.* **76** 965

[21] Thomas S 1993 *Phys. Rev. E* **47** 2401

[22] Sugihara G , May R M 1990 *Nature* **344** 734

[23] Racicot D M , Longtin A 1997 *Physica D* **104** 18

[24] Xie Y , Xu J X , Yang H J *et al* 2002 *Acta Phys. Sin.* **51** 205 (in Chinese) [谢 勇、徐建学、杨红军等 2002 物理学报 **51** 205]

[25] Sauer T , Yorke J , Casdagli M 1991 *J. Stat. Phys.* **65** 579

[26] Sauer T 1994 *Physical Review Letters* **72** 3811

[27] Dai X G , Wang P , Chou J F 2003 *Chinese Science Bulletin.* **48** 2483 (in Chinese) [戴新刚、汪 萍、丑纪范 2003 科学通报 **48** 2483]

[28] Wang P ,Dai X G 2005 *Acta Phys. Sin.* **54** 4961 (in Chinese) [汪 萍、戴新刚 2005 物理学报 **54** 4961]

[29] Wang Y ,Shi N ,Gu J Q *et al* 2006 *Chinese Journal of Atmospheric Sciences* **30** 162 (in Chinese) [王 颖、施 能、顾骏强等 2006 大气科学 **30** 162]

[30] Shi Neng ,Cao H X 1996 *Journal of Chinese Journal of Atmospheric Sciences* **20** 337 (in Chinese) [施 能、曹鸿兴 1996 大气科学 **20** 337]

[31] Yang J P , Ding Y J , Chen R S *et al* 2005 *Scientia Geographica Sinica* **25** 441 (in Chinese) [杨建平、丁永建、陈仁升等 2005 地理学报 **25** 441]

Nonlinear noise reduction for the observation data of climatology based on the searching average method^{*}

Hou Wei^{1 2)} Lian Yi¹⁾ Feng Guo-Lin^{3 4)†}

¹ *Institute of Meteorological Science of Jilin Province, Changchun 130062, China*

² *Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China*

³ *Key Laboratory of National Meteorological Bureau Climate Research Center, Beijing 100081, China*

⁴ *Department of Physics, Yangzhou University, Yangzhou 225002, China*

(Received 22 May 2006 ; revised manuscript received 14 June 2006)

Abstract

The searching average nonlinear noise reduction method, which is based on local linear fit to the nonlinear dynamics, is introduced to reduce the noise in the observation data of climatology. Recurrence plots are used to estimate the size of local neighbors. The noise reduction is improved markedly. In order to show the validity of the program in noise reduction, it is first applied to a noise time series of Henon map contaminated by Gaussian white noise. And then, this noise reduction scheme is applied separately to the observation data of meteorology. The analyses of a nonlinear prediction demonstrate the efficiency of the method for noise reduction.

Keywords : recurrence plots, nonlinear noise reduction, nonlinear prediction

PACC : 9260X

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

[†] E-mail : feng_gl@sina.com ; fenggl@cma.gov.cn