

# SSA 方法在气候时间序列分析和预测中的应用

刘健文

周小刚

(空军第七研究所,北京 100085) (北京气象学院,北京 100081)

**摘要** 文章介绍了奇异谱分析(SSA)方法在不同时间尺度气候研究中的应用。表明 SSA 是气候动力重建、统计分析和信息压缩的一种有效工具,并可应用于建立经验长期预报的预报因子集。

**关键词** 奇异谱分析,时空主分量,重建,低频振荡,长期预报

## 1 引言

气候动力学的研究表明,如果考虑较长的时间尺度,则气候是可以由较少的自由度来描述的。但是,由于尺度的相互作用,还难于通过简化运动方程来分离行星波及研究其动力学特性。换句话说,即便大尺度的气候系统是相对低阶的动力系统,也难于从基本的 Navier-Stokes 系统中导出简单的演变方程。因此我们往往只能从观测资料、而不能从方程出发来研究大尺度气候动力学问题。

从观测资料中可大致获得这样一个气候吸引子:吸引子中不同的吸引区域可共存,导致周期性行为。在这些区域中最易发现的是不稳定平衡态和不稳定周期轨迹<sup>[1]</sup>。当相空间轨迹被吸引到周期轨道上时,大气的行为在一定时期经历了一次周期振荡。而当它位于不稳定的固定点上时,大气是准静止的,即所谓天气季节状况(weather regimes)。

奇异谱分析(Singular Spectrum Analysis, 简称 SSA)被认为是一种特别适合于研究周期振荡行为的分析方法<sup>[2]</sup>。它是 EOF 分解的一种特殊应用,分解的空间结构与时间尺度密切相关,可以较好地有噪声的有限长度时间序列中提取信息,目前已应用于多种气候时间序列的分析中。

本文重点在于讨论 SSA 的一些应用结果,包括气候吸引子的重建、年际和季节内低频振

荡的分析以及采用 SSA 的时空主分量为预报因子的经验长期预报方案等。关于 SSA 的数学原理和理论分析,感兴趣的读者可参阅文献[2]、[3]。

## 2 吸引子的重建和近似问题

像许多其它的物理系统一样,气候系统随时间的演变可由一组偏微分方程表示为:

$$\frac{\partial \vec{X}_t}{\partial t} = F(\vec{X}_t) \quad (1)$$

(1)中矢量  $\vec{X}_t$  是  $t$  时刻的气候状态,其离散形式为:

$$\vec{X}_{t+\Delta} = G(\vec{X}_t) \quad (2)$$

这里的  $\Delta$  是时间积分步长。空间离散化使  $\vec{X}_t$  成为一个有限维数的矢量。对气候系统来说,函数  $G$  是包含强迫和耗散的非线性函数。(2)式表示的气候动力系统一般是浑沌的,系统的吸引子有复杂的拓扑结构。计算这一吸引子有两种可能的方法,一是采用近似方程,即建立数值模式,二是直接对观测数据进行分析。

近年来很多工作围绕着第二个方法进行,即对观测时间序列进行非线性动力学重建。根据 Whitney 嵌入定理,如果存在某个连续取有限值的变量,则它仅能适当定义一个具有相同拓扑性质的新状态矢量,即只有一个变量有能力重建吸引子。因此,如果气候是一个低维动力系统的话,只存在一个有限连续观测的量能够进行完备的气候动力学重建。

但是,历史资料中的观测误差总是难以避免的。这种噪声模糊了可能存在的低阶吸引子的细致结构,以致不能恰当估计出系统的维数。同时观测资料的长度也是一个限制问题。因此,实际上存在一个不能描述吸引子细微结构的限度。对于一个有限长度的噪声资料序列来说,只能通过某些数学算法估计出一个宏观的维数。很多人采用过延迟坐标相空间的 Grassberger-Procaccia 算法,但是因为不能进行坐标变换以确定对应的变量,故没有实际意义。

很多研究都以气候维数计算为重点,但得出了不同的数目,这是由于上述的限制造成的。也存在另外一种可能性,即气候系统并非是低阶的。Broomhead 和 King 最早把 SSA 引入这方面的研究中<sup>[4]</sup>。他们通过嵌入一个单一时间序列来进行非线性吸引子的重建,结果表明在一个新的延迟坐标空间中分解的 SSA 能获得比较精确的结果。

SSA 在数学上等价于 EOF 分析在延迟坐标上的表达,因此其状态矢量可分解成由前若干个主分量代表的“信号”分量和由特征值谱末端代表的“噪声”分量之和。对吸引子的近似通过对显著分量的截断来进行,即作为延迟坐标空间的一个子空间。这样的“宏观”近似有着传统 EOF 分析所不具备的动力学意义。

文献[4]指出,当资料中出现白噪声时,特征值谱在某个特定的阶数  $S$  后达到极大,在高阶分量的方向上信号不能从噪声中分辨出。而 Vautard 和 Ghil 的研究表明<sup>[5]</sup>, $S$  定义了包含全部信息的显著方向的数目,它与吸引子的维数没有特别的联系。实际上, $S$  值只是一个上界,任何大于  $S$  的维数估计肯定是不正确的。

Vautard 和 Ghil 还指出,当从 SSA 中计算出明显成对的特征值、而相应的 EOF 几乎是周期性和正交时,则通常对应着信号中的振荡现象。与其它传统的谱分析方法相比,SSA 有很多优点。比如它具有分析过滤器的自适应特点;特别是对窗口长度  $M$  的限定,使得可对振荡的转换进行时间定位。文献[2]表明,当对所有明显成对的特征分量进行重建时,吸引子可以简

单地以准环形轨迹被近似估计;而对其它的分量一般不容易恰当描述,有的仅仅是呈浑沌或倾向变化,对它们的重建只能以一样复杂得多的过程来逼近吸引子。

文献[2]中提出了几种分离振荡和噪声分量的算法,特别指出 SSA 能很好地从时间序列中分离出周期小于  $M$ 、谱宽小于  $1/M$  的振荡。

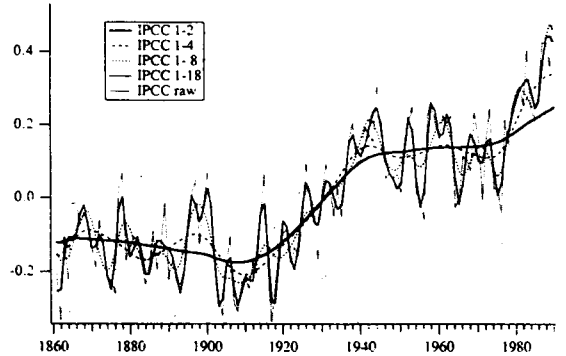


图 1 IPCC 序列以及对不同分量的重建  
(摘自文献[2])

### 3 在低频振荡分析中的应用

SSA 目前主要用于年际和季节尺度的大气低频振荡分析。

Ghil 和 Vautard 用 SSA 分析全球地面气温记录(IPCC 序列)<sup>[2]</sup>。资料是从 1861 到 1990 年的全球年平均地面气温值,共 130 个数据。SSA 计算的窗口长度取为 40 年,以求解从几年到几十年的振荡周期。结果表明(见图 1),前 18 个时空主分量是十分显著的,特别是头 2 个代表了信号的趋势分量,揭示出本世纪气温约有  $0.5^{\circ}\text{C}$  的不均匀增高;分量 1—4 代表着 20—30 年的振荡。与二氧化碳含量的指数式增加不同,增温趋势是分步式的。资料序列中 1990 年是最暖的一年,但从趋势分量看,八十年代并非明显高于此前的年份,因此 1990 年的极值可能只是气候的自然变化。分析所有的主分量后发现,许多振荡分量都是在二十世纪 80 年代后期达到极大,作者由此认为该时期的气温异常是几种振荡的叠加和相互作用。这一结论有待讨论。对这一时间序列计算出的显著周期有 4—

5、10、15、20—30 年。G. Plaut 等在更长的英格兰局地气温时间序列中也发现了类似的周期<sup>[6]</sup>。这证实了气候变化存在着几年到几十年的变化规律。目前研究已表明年际振荡是与厄尔尼诺/南方涛动(ENSO)现象密切相关的。对这一现象 Rasmusson 等用不同资料进行了研究<sup>[7]</sup>,发现其存在两种主要的周期。对于尺度为几十年的振荡,尽管已提出了某些机制,但仍需作进一步的研究。

Ghil 和 Mo 最早用 SSA 研究大气季节内变化<sup>[8]</sup>,Plaut 和 Vautard 后来扩展到多通道的 SSA<sup>[9]</sup>,所用资料为经过过滤的 700hPa 高度场。研究发现了一些明显的季节内振荡,如众所周知的热带 30—50 天低频振荡,并且发现温带大气也有相同周期的振荡行为,但机制不一样,主要出现在北半球的冬季。用 SSA 和最大熵谱方法(MEM)结合,精确揭示出北半球大气中存在 70 天、50 天(太平洋地区)、35 天(大西洋地区)的振荡周期。认为行星尺度和天气尺度的非线性相互作用可能促使大尺度大气呈现许多可能的状态,因而季节内大气只能用较大数目的自由度来描述。

通过对振动位相和振幅作合成分析,SSA 可用于研究振荡行为与其它现象的相互作用。文献<sup>[9]</sup>指出,大西洋的 35 天波型是北半球 70 天波型的锁相,这表明 35 天波型是 70 天波型的一个谐波。而大西洋振荡与太平洋振荡在位相和振幅上没有关系。还提出振荡行为的振幅大大影响着天气季节转换发生的条件概率。如阻塞现象就常常是以 35 天振荡的一种特殊位相出现的。认为是这些振荡为某些天气季节转换的突发提供了有利的环境条件。

上述结论对气候模拟和长期(短期气候)预报有重要意义。由此推论,一个好的数值或统计气候模式必须能正确模拟出大气的季节内振荡,进而才能模拟好天气季节变化特征;而天气季节转换不是随机的,它们受振荡行为的影响。鉴于振荡行为本身是大气的可预报现象,因此它们可以成为长期大气可预报性的一个贡献源。

#### 4 长期预报试验

通常认为对天气系统的决定性预报不可能超过 2 周。就目前而言,逐日预报可用时效是 1 周左右。但这并不意味着超出这一时限就没有任何预报指示信息。事实上,模拟研究已经表明,在逐日可预报性界限外,大尺度预报场还是有一定相关性。关于长期可预报性的作用源主要在两个方面。第一个方面是海洋,海洋状况对气候变化和气候异常的形成有重要作用。大气常常以某些持续性现象响应着海洋的异常,特别是在热带地区,在中纬由于起主要作用的是大气内部变化,故响应不那么强。可预报性的第二个方面完全来自大气内部,即天气季节转换和振荡的交替。但由于存在初始场以及如何正确反映下垫面作用等问题,目前的各种环流模式还难以取得满意的结果。统计模式还在广泛采用。SSA 应用于年际和季节内长期预测时取得了较好的结果。

采用经验统计模式时的一个主要问题是如何选择预报因子。考虑到资料长度的限制,模型中不宜包含太多的参数,否则交叉独立检验的技巧分会大大低于某一个优化检验时段的技巧分。因而,预报因子集必须既小且又包含相关的预报信息。SSA 正是用于预报因子信息压缩,即提取可预报分量上。

文献<sup>[2]</sup>采用 SSA 对 IPCC 序列进行了大量的组合试验,并通过自回归(AR)模型外推出显著的时空主分量。进而作预报主分量的重建,并与观测值进行了详尽比较。结果表明,直至提前 5 年作预报,SSA 的结果仍优于持续性和气候预测的结果。Keppenne 和 Ghil 采用类似的预报方法对 ENSO 现象进行时效 3 年的预报,结果优于由多个 GCM 作出的预报<sup>[10]</sup>。

把 SSA 引入预报的优点在于其显著主分量对应着某些大气低频振荡,它们有较好的可预报性,而其它分量是难以用线性模式来作预报的。如果建立 AR 模型时不排除非显著分量,则预报技巧会低得多。

对于大气季节内变化的预报,由于存在较

大的自由度,因此初始资料向预报因子压缩的投影必然会损失一些有价值的信息,但这是保证预报技巧分水平的代价。此外,中纬度的月或季节预报还存在瞬变斜压强迫的强烈反馈作用问题。瞬变起着随机强迫作用,是不可预报的,可导致计算的大尺度大气呈现不同状态。因此,由大尺度初始场信息是不能确切预报其未来演变的,季节内的长期预报属于概率预报问题。

Vautard 等用从 SSA 得到的时空主分量作为经验预报因子,作 700hPa 高度场月平均预报的尝试<sup>[11]</sup>。计算资料取 1951—1993 年大西洋范围 113 个格点的高度场分析值,并经过滑动平均和距平化等处理。预报因子集由前 15 个时空主分量组成。预报量为同一范围的格点值,划分为 3 级:偏低、正常、偏高。用相似法建立预报模型。采用交叉检验法把 42 年序列划分为 7 个 6 年时段,对冬季月份(12 月至 3 月)进行了预报检验。图 2 是基于 SSA 的预报和持续预报的得分情况,评分采取 3 个级别预报的平均准确率。图中显见,SSA 预报在 1951—1957 年以外的所有检验时段都优于持续预报,而两者均明显优于随机预报(1/3 机会)。

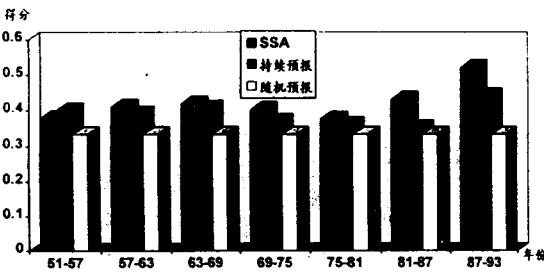


图 2 700hPa 高度场月平均预报准确率 (根据文献[11]整理)

刘健文和 Vautard 用类似方法,以北半球 700hPa 高度场的主分量为预报因子,对 1036 个气候站的月平均地面气温进行预报试验,并对 SSA 和传统的 EOF 方法进行了对比。统计评分采用概率空间线性误差 (LEPS) 值来衡量,资料序列为 1952—1988 年,划分成 12 个 3 年交叉检验时段。图 3 是冬季(1、2、3 月)预报技巧分的空间分布。图 4 是基于 SSA 和

EOF 的预报技巧对比,前者在大部分预报检验时段占优势,表明时空主分量较之空间主分量更起到可预报分量的作用。

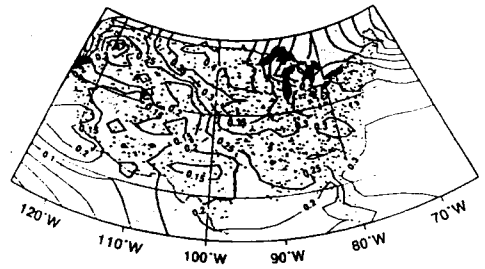


图 3 美国大陆 1036 个站地面气温 SSA 预报 LEPS 得分的空间分布(黑点为站点位置)

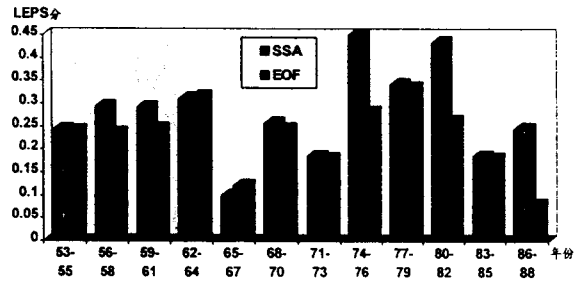


图 4 两种方法在 12 个交叉检验期的预报技巧对比

为了尽可能利用气候系统的信息,可尝试把动力方程与上述经验统计预报方案组合起来,例如把 SSA 计算的某些振荡分量作为 GCM 的限定条件。比如,我们考虑在下述洛仑兹系统(x, y, z)中加入一个低频振子(u, v):

$$\begin{cases} \frac{dx}{dt} = -\sigma(x - y) + au \\ \frac{dy}{dt} = \gamma x - y - xz + av \\ \frac{dz}{dt} = xy - bz \\ \frac{du}{dt} = \Omega v - Ku - \alpha x \\ \frac{dv}{dt} = -\Omega u - Kv - \alpha y \end{cases} \quad (3)$$

假设系统的当前状态为(x<sub>0</sub>, y<sub>0</sub>, z<sub>0</sub>, u<sub>0</sub>, v<sub>0</sub>)。但当前观测是有噪声的,故实际初始状态应为(x<sub>0</sub> + δx<sub>0</sub>, y<sub>0</sub> + δy<sub>0</sub>, z<sub>0</sub> + δz<sub>0</sub>, ...)。用 ŷ(t)表示变量 u 的初猜值,它可用经验预报方法得出。图 5 给出

了一次计算实例,其中 5a 是  $u(t)$  和  $\hat{u}(t)$  的演变曲线,5b 是变量  $x$  的几种演变。如果  $\hat{u}(t)$  提供的信息没被考虑,则时间  $t$  的预报结果就不是真正的演变曲线 ( $x(t), \dots$ ), 而是错误轨迹 ( $x_p(t), \dots$ )。考虑  $\hat{u}(t)$  的一种办法是寻找系统的轨迹 ( $x_s(t), \dots$ ), 使对应的  $u_s(t)$  曲线最接近  $\hat{u}(t)$ 。这一轨迹可以用传统的非线性最优化方法来计算。图 5 清楚表明,初猜值的信息可改进预测的效果。这个例子说明,从统计和动力两方面综合考虑,建立混合预报模型是有实际意义的。

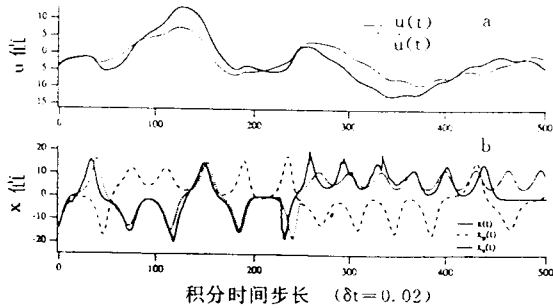


图 5 混合模型预报试验实例:(a)变量  $u$  演变的真值(实线)和初猜值(点线);(b)变量  $x$  演变的真值(点线)、不采用初猜的预报值(长划线)、混合模型预报值(实线)

## 5 结语

本文介绍了 SSA 方法在气候分析和预测方面的一些应用。已有结果表明,SSA 是一种有效的统计分析工具,特别是可以定位周期振荡,过滤气候噪声。它在数学处理上与扩展 EOF 方法相近,但在空间类型和时间类型的解释上不同。SSA 可以处理大于空间自由度的时间自由度问题,其时空主分量可提供所研究物理系统的动力学信息。系统的自然振荡可以在位相正交的典型场中表现为偶极子的形式,易于识别。

从 SSA 得到的主分量的规则行为使之较完全信号更易于预报。因此,在一定意义上,SSA 不仅是有效压缩了研究对象的时空信息,

而且是把最可预报的分量聚集到若干个时间序列、即头几个时空主分量上。以此作为经验预报因子有可能为长期预报提供好的基础。文中给出的在年际和季节内预报两方面的应用说明了这一点。另外,以某种合适的方式,把从较长历史记录中获得的经验统计结果和动力学方程结合起来,有望显著改进长期预报的效果。

## 参考文献

- 1 Ghil M, S Childress. Topics in Geophysical Fluid Dynamics; Atmospheric Dynamics, Dynamo Theory and Climate Dynamics. Springer, New York, 1987
- 2 Vautard R, P Yiou and M Ghil. Singular spectrum analysis: A toolkit for short, noisy chaotic signals. Physica D, 1992 (58): 95-126
- 3 周小刚, 刘健文. 时空扩展 EOF 分析及其初步应用. 北京气象学院学报, 1996, 第 1 期
- 4 Broomhead D S, G P King. Extracting qualitative dynamics from experimental data. Physica D, 1986(20): 217-236
- 5 Vautard R, Ghil M. Singular spectrum analysis in nonlinear dynamics with applications to paleoclimatic time series. Physica D, 1989(35): 395-424
- 6 Plaut G, Ghil M, Vautard R. Interannual and Interdecadal Variability in 335 years of central England temperature. Science, 1995(268): 710-713
- 7 Rasmusson E M, Wang X L, Ropelewski C F. The biennial component of ENSO variability. J Mar Syst, 1990(1): 71-96
- 8 Ghil M, Mo K C. Intraseasonal oscillations in the global atmosphere, Part I: Northern hemisphere and tropics. J Atmos Sci, 1991(48): 752-779
- 9 Plaut G, Vautard R. Spells of low-frequency oscillations and weather regimes in the Northern hemisphere. J Atmos Sci, 1994(51): 210-235
- 10 Keppenne C L, Ghil M. Adaptive filtering and prediction of the Southern Oscillation index. J Geophys Res, 1992, 97D, 20441-20252
- 11 Vautard R. Patterns in time; SSA and MSSA, In: Statistical Analysis of Climate Variability, proceedings of the E. C. Autumn school, Isola d'Elba, Italy, 1995