第2卷 第4期 应用气象学报 Vol.2,№4
1991年11月 QUARTERLY JOURNAL OF APPLIED METEOROLOGY Nov. 1991

长期预报的相空间相似法

林 振 山*
（北京大学地球物理系）

提 要

通过对分数维 D 的计算，我们用时间序列重建了一个 $d=\mathrm{INT}(l+D)$ 维的相空间。而后，将统计预报的相似理论用于相空间的相点，并假设最近邻相点经演化时间 T 后模相等。根据这一假设，做了一些长期预报试验，其距平符号报准率一般为60—70%，相对误差一般不大于8%。

一、引 言

目前，无论是国内，还是国外，长期天气的统计预报准确率仍然较低，导致长期预报不准确的一个重要因素是长期天气系统的外在随机因素，如随机初始条件、随机参数、随机噪声、物理场的随机起伏等。可是长期以来，我们一直忽略了导致长期预报不准确的另一个重要因素，即系统的内在随机性。所谓系统的内在随机性，指的是某些完全确定性的方程（系统），不需附加任何随机因素，亦可表现出随机行为，即浑沌。浑沌系统的最大特点是系统的演化对初始条件十分敏感。从而导致长期行为的随机性，即不可预报性。研究表明，几乎所有的长期天气系统、气候系统都是浑沌系统。而统计方法却无法消除或滤去内在随机性。

浑沌理论在指出长期预报的不可能的同时，还阐明了相空间短期过程的预报是可能的。因为任一浑沌时间序列都包含了系统简单的确定性关系。有必要指出，经过适当的技术处理，我们所说的长期预报问题，将对应于相空间里的一个短过程。故而，我们认为：将有关的浑沌理论与统计理论结合起来研究长期预报问题，势必同时减小系统的内、外随机性所造成的不准确性。

本文的第二部分阐述了如何重建相空间；第三部分阐述了相空间相似预报法；第四部分则是对该模式的检验。

本文1990年6月4日收到，9月10日收到修改稿。

* 现在南京大学大气科学系工作。

二、相空间和相型

1. 时间序列的处理和拓展

众所周知，任一时间序列都包含着系统参与演化的所有变量的大量信息。而系统的某一态变量随时间的变化，可以用一时间序列来表示。现在的问题是：如何用一时间序列来重建一个 n 维空间，用以描写包含 n 个变量的系统的动态特征。

由于我们事先不知道描述系统的相空间的真实维数，但总可以用时间序列的数据，通过适当的组合方式支起一个 n 维空间 R^n。这种利用数据序列构造成的 n 维空间，通常叫嵌入空间。

设时间序列间隔 $\Delta t=1$（单位），选定一延滞时间 $\tau=h\Delta t(h=1,2,\cdots)$，将原来的时间序列按如下方式排列，构成一个相型：

$$\begin{array}{cccc}
x(t_1) & x(t_2) & \cdots x(t_i) & \cdots x(t_N) \\
x(t_1+\tau) & x(t_2+\tau) & \cdots x(t_i+\tau) & \cdots x(t_N+\tau) \\
x(t_1+2\tau) & x(t_2+2\tau) & \cdots x(t_i+2\tau) & \cdots x(t_N+2\tau) \\
\vdots & \vdots & \vdots & \vdots \\
x(t_1+(n-1)\tau) & x(t_2+(n-1)\tau) & \cdots x(t_i+(n-1)\tau) & \cdots x(t_N+(n-1)\tau) \\
y(t_1) & y(t_2) & \cdots y(t_i) & \cdots y(t_N)
\end{array} \tag{1}$$

其中，$y(t_i)$ 为 n 维相空间里的一个相点，表征该系统在某一瞬时的状态。N 个相点的连线则构成了点在相空间的轨道，它表征了系统状态随时间的演化。而 $y(t_i)$ 的集合，则构成了相空间的一个相型。

2. 分数维 D 的计算

计算分数维 D 的方法很多，而比较适合于用实验资料时间序列来计算 D 的方法是 Grassberger 和 Procaccia 提出的[1]。其主要步骤是：

（1）重建（1）式所示的 n 维相空间。

（2）计算关联函数

$$C(r)=\frac{1}{N^2}\sum_{\substack{i,j \\ i\neq j}}^{N}\theta(r-\|y(t_i)-y(t_j)\|) \tag{2}$$

式中的 $\theta(z)$ 的是 Heaviside 函数，当 $z<0$ 或 $z>0$ 时，它分别取零或 1。r 为临界距离；$\|y(t_i)-y(t_j)\|$ 表示相点 $y(t_i)$ 到 $y(t_j)$ 的 n 维 Euclidean 距离。

（3）对于 r 的某个适当范围，吸引子的维数 D 与关联函数 $C(r)$ 应满足对数线性关系：

$$D(n)=\ln C(r)/\ln r \tag{3}$$

（4）增高嵌入空间的嵌入维数 n，重复（1）—（3）三个步骤，直到相应的维数估计值 $D(n)$ 不随 n 的增长发生有意义的变化（保持在给定的误差范围内）为止，即

$$D = D(n_\infty) \simeq D(n_\infty + 1) \tag{4}$$

(4)式的 D 就是我们所要求的分数维。

3. 用于预报的 d 维相空间和相型

如果从时间序列$\{x_t\}$($t=1,2,\cdots N$)里，求得分数维 D 为非整数，则$\{x_t\}$为一浑沌时间序列，其所表征的系统的行为是浑沌的。对于浑沌系统，我们可以嵌入一个 $d=\text{INT}(l+D)$ 维的相空间，$l=1,2,\cdots$。

将时间序列$\{x_t\}$分为两段：x_1 至 x_m 和 x_{m+1}至 x_N，后者很短，仅供检验用，前者按如下排列(注意与(1)的不同)，构成相型：

$$\begin{array}{cccc}
x(t_m) & x(t_{m-1}) & \cdots x(t_i) & \cdots x(t_1+(d-1)\tau) \\
x(t_m-\tau) & x(t_{m-1}-\tau) & \cdots x(t_i-\tau) & \cdots x(t_1+(d-2)\tau) \\
\vdots & \vdots & \vdots & \vdots \\
x(t_m-(d-1)\tau) & x(t_{m-1}-(d-1)\tau) & \cdots x(t_i-(d-1)\tau) & \cdots x(t_1) \\
Y(t_m) & Y(t_{m-1}) & \cdots Y(t_i) & \cdots Y(t_1+(d-1)\tau)
\end{array} \tag{5}$$

这里的 τ 为延滞时间，$Y(t_i)$为相点(或态)，$Y(t_i)$在 d 个基矢上有 d 个分量：$x(t_i)$、$x(t_i-\tau)\cdots x(t_i-(d-1)\tau)$。为方便起见，我们把它们分别叫做 $Y(t_i)$的第一分量、第二分量、…第 d 分量。而第一分量的时间标志 t_i 和 $Y(t_i)$的时间标志一致。

三、相空间相似预报法

1. 基本假设

由于提前预报的时间 T 一般与重建相空间里所用的延滞时间 τ 及时间间隔 Δt 同量级，它们相对于时间序列的长度是很小的，故假设：给定时间序列的结构，在预报区域内不变。

2. 最近邻点

用$\|\ \|$表示 d 维 Euclidean 模，则$\|Y(t_i)-Y(t_j)\|$表示在 d 维相空间里态 $Y(t_i)$到态 $Y(t_j)$的距离。

设参考态为 $Y(t_m)$，其最近邻态为 $Y_{nbt}(t_i)$，则 $Y_{nbt}(t_i)$与 $Y(t_m)$的关系为：

$$Y_{nbt}(t_i) = \min[\|Y(t_m) - Y(t_j)\|] \qquad j = 1,\cdots(m-d+1) \tag{6}$$

3. 相空间相似预报法

设参考态 $Y(t_m)$及其最近邻态 $Y_{nbt}(t_i)$，经预报时间 T 后，分别演化为 $Y(t_m+T)$和 $Y_{nbt}(t_i+T)$，而 $Y(t_m+T)$的第一分量为 $x(t_m+T)$。显然，只要我们提前预报的时间 $T\leqslant\tau$，则 $Y(t_m+T)$中，只有 $x(t_m+T)$这一分量是未知的，而其余的$(d-1)$个分量，都是已知的。故 $x(t_m+T)$即为我们的预报对象。

相似预报是常用的一种近似方法。即利用前期资料选相似，利用相似年的演变作预报。将这一原理与浑沌理论结合起来，我们提出以下的相空间相似预报模式：

$$\| Y(t_m + T) \| = \| Y_{nbt}(t_i + T) \| \tag{7}$$

四、预报及其检验

严绍瑾、彭永清等利用广州1873—1980年月平均气温时间序列资料，求得分数维 $D=2.3$[2]。所以，我们分别取 $d=3,4,5$；$T=\tau=1,2,3,4$；用广州近40年月平均气温时间序列资料进行了一些抽样计算，结果见表1。表1中 $d=3$ 的距平符号报准率平均为62.5%，相对误差平均为4.1%；$d=4$ 和5时，用1961—1970年时间序列资料的距平符号平均报准率为75%，相对误差平均为1.8%；用1951—1988年时间序列资料的距平符号报准率平均为66.7%，相对误差平均为4.9%。

表1　用广州月平均气温资料抽样计算结果

时间序列	d	τ	T	t_m	实测 $x(t_m+T)$ (℃)	预报结果 (℃)	实际距平 (℃)	预报距平 (℃)	距平符号报对与否	相对误差 (%)
1961年1月——1970年12月 $\Delta t=1$ 个月	3	1	1	1969年7月	28.4 1969年8月	27.8	+0.2	−0.4	错	2.1
	3	2	2	1969年7月	28.0 1969年9月	27.6	+1.0	+0.6	对	1.4
	3	3	3	1969年7月	23.7 1969年10月	23.3	−0.1	−0.5	对	1.6
	3	4	4	1969年7月	17.8 1969年11月	19.8	−2.0	0	/	11.2
1961年1月——1970年12月 $\Delta t=1$ 个月	4	1	1	1969年7月	28.4 1969年8月	27.9	+0.2	−0.3	错	1.8
	4	2	2	1969年7月	28.0 1969年9月	27.5	+1.0	+0.5	对	1.7
	5	1	1	1969年7月	28.4 1969年8月	29.3	+0.2	+1.2	对	3.1
	5	2	2	1969年7月	28.0 1969年9月	28.1	+1.0	+1.1	对	0.4
1951年1月——1988年12月 $\Delta t=1$ 个月	5	1	1	1985年2月	14.8	17.8	−3.0	+0.1	错	20.3
	5	2	2	1985年2月	20.6	20.1	−1.3	−1.0	对	1.4
	5	1	1	1985年6月	28.0	28.2	−0.5	−0.3	对	0.7
	5	2	2	1985年6月	28.8	28.9	+0.6	+0.7	对	0.3
	5	1	1	1985年7月	28.8	28.5	+0.6	+0.3	对	1.0
	5	2	2	1985年7月	26.4	28.2	−0.6	+1.2	错	6.8

五、结　　语

从上面的一系列试验可以看出，利用相空间相似法制作长期预报，是有意义的，是对传统长期预报理论的补充。由于在重建相空间时，我们可以取 $d=\mathrm{INT}(1+D)$、$\mathrm{INT}(2+D)$…维和不同的 τ 值，从而，有不同的预报方案。通过对模式的检验、调试，我们总可以找到一组较合适的(d,τ)值，从而做出预报。大量的试验表明：

1. 在求分数维 D 时，所需的时间序列长度约为 10^{D} 个，但在预报时，所需的时间序列长度为 180—500 个。

2. 在求分数维 D 时，其嵌入相空间的维数 n 要充分大于 D 的估计值，但在做预报时，嵌入相空间的维数 d 只要略大于分数维 D 即可，即取 $l=1,2,3$ 为宜。

3. $T=\tau$ 时，即提前预报的时间 T 等于延滞时间时，预报效果较好。这意味着，利用相空间相似法超前预报 $T=1,2,3$(月)的差别只在于分别取 $\tau=1,2,3$(月)而已，而计算量和预报精度都是一样的。

4. 通过调试，相空间相似预报法的距平符号报准率一般为 60—70%，相对误差一般不大于 8%。

鸣谢：刘式达先生对本文提出了不少有益的意见，谨此致谢！

参 考 文 献

[1] Grassberger, P. and I. Procaccia, Characterization of strange attractors, *Phys. Rev. Lett.*, **50**, 346—349, 1983.

[2] 彭永清、严绍瑾等，一维气候时间序列拓展及其相空间中浑沌吸引子维数的确定，全国第二届大气统计物理学术会议文集，1989 年。

AN ANALOG METHOD OF PHASE SPACE FOR LONG-RANGE FORECAST

Lin Zhenshan

(*Department of Geophysics, Peking University*)

Abstract

By way of the calculation of fractal dimension D, a phase space with d-dimensions ($d=\mathrm{INT}(l+D)$) is established using a time series. Then, the analog theory of statistical forecast is applied to the phase points of phase space, and it is hypothesized that the mode of state at the nearest phase points after evolution time T is equal to each other. A lot of trial studies based on the hypothesis show that the forecast accuracy of anomaly sign can reach 60—70 per cent, and the relative error is generally less than 8 per cent.