MSSA-SVD 典型回归模型及其用于 ENSO 预报的试验*

裕国 程正泉
(南京气象学院, 南京, 210044)
程炳岩
(河南省气象局气候中心, 郑州, 450003)

摘 要

文中提出了一种基于多通道奇异谱分析(MSSA)的广义典型混合回归模式。其基本思想是,利用 MSSA-SVD 提取预报因子场和预报变量场的显著耦合振荡信号,对它们的前几个显著典型分布型建立多元线性统计气候预报模式。经对 Nino 海区各季海温距平所进行的短期气候预测试验表明,其预报效果优于其它统计预报方案,从而为探索 ENSO 预测方法提供了一种新的思路。

关键词: 多通道奇异谱分析, 典型混合回归模式, ENSO 预测方法。

1 引 言

众所周知, ENSO 现象乃是目前全球气候系统所能观测到的最强气候信号,它的发生、发展及其演变规律一直是人们关注的焦点,实现对全球 ENSO 气候信号的准确预报,正是当前研究 ENSO 气候型态的一个关键。然而, ENSO 预报既是短期气候预测的研究热点,也是其难点。

目前,统计预报模式仍然是一种花费少、见效快的气候预报途径。在充分考虑物理机制或气候系统某些方面的整体型态基础上,采用适当的统计预测模式,仍不失为行之有效的方案。利用统计学方法对ENSO 作预测的模式很多,例如典型相关分析(CCA)、主振动型分析(POP)、经验正交函数与主分量分析(EOF/PCA)及线性转置分析(LIM)等都是一些较好的方法。采用这些统计预报模式,一般都能获得较优的预报效果。对此,国际上已取得一定的成果,一些国家已陆续投入业务预报;而中国国内近年来关于 ENSO 预测的统计模式也有了一定的发展[1~5]。丁裕国等提出一种典型自回归预测模式用于气象要素场的预报,在此基础上又提出一种推广的典型自回归气象场预测模式(PG-CCA)用于预报 Nino 海区的 SSTA,取得较好的效果[6]。1998 年丁裕国等引用 SSA 基础上的 AR 模型对 ENSO 海区 SSTA 进行预测,也取得了较好的效果[7]。其后,丁裕国等又提出奇异交叉谱分析(SCSA)方法,用于提取 Nino 海区和南方涛动指数(SOI)的耦合振荡信号,并结合回归分析,对 Nino 各海区平均的 SSTA 月际序列作短期气候预测试验,结果证明,预报试验效果较好[8~10]。

尽管20 世纪 80 年代以来,各种海-气耦合模式或统计预报模式已能成功地预报赤道太平洋海域的一些关键区 SSTA,但仍有许多不足之处,例如预报效果的稳定性随季节或年代漂移(春季的预报较差,预报水平有较明显的年代际变化,等等)、预报时效的稳定性和预报时效的延伸长度随不同模式有明显差异(超前6 个月以上的预报仍然非常差),这些都是有待改善的。文中提出一种基于多通道奇异谱分析(MSSA)的广义典型混合回归模式,试图弥补上述缺陷,通过提取预报因子场和预报场的显著耦合振荡信号,结合多元回归统计技术,对各 Nino 海区各季节海温进行了短期气候预测试验,为探索 ENSO 预测方法提供一种新思路。

* 初稿时间: 2000 年 8 月 22 日; 修改稿时间: 2001 年 12 月 24 日。
资助课题: 国家九五重中之重科技项目"我国短期气候预测系统的研究"和(96-908-04-02)课题及河南省计委科技攻关项目。

2 模 型

设 $Y=(y_1, y_2 \cdots y_q)'$ 和 $X=(x_1, x_2 \cdots x_q)'$ 分别表示预报变量场和预报因子场。不失一般性, 约定 X, Y 中各变量已零均值化, 其中 $y_i = y_i(t)$ 和 $x_i = x_i(t)$, $i=1, 2 \cdots q$ 或 p, $t=1, 2 \cdots n$, 参数 q, p, n 分别为预报变量数、预报因子数和序列长度。

首先对 $\boldsymbol{X}$ 种 $\boldsymbol{Y}$ 场分别作多通道奇异谱分析(MSSA), 由此得到各场相应的重建耦合振荡分量序列(RCCS)。所谓多通道奇异谱分析(MSSA)则是奇异谱分析(SSA)的推广[8, 12], 后者仅用于对单一时间序列作时迟 EOF 分解, 以便研究序列内不同振荡模态之间的振荡结构。然而, 单一时间序列并不足以代表整个场的变化, 而现代气候则多着眼于变量场的时空结构特征的研究, 即不仅要靠考虑其时间变化, 而且要考虑其空间型态变化。因此, 类似于 SSA, 将预报变量场和预报因子场的资料序列作如下时迟排列

$$\boldsymbol{X} = \begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_{n-m+1}^1 \\ x_2^1 & x_3^1 & \cdots & x_{n-m+2}^1 \\ \cdots & \cdots & \cdots & \cdots \\ x_m^1 & x_{m+1}^1 & \cdots & x_n^1 \\ x_1^2 & x_2^2 & \cdots & x_{n-m+1}^1 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ x_m^N & x_{m+1}^N & \cdots & x_n^N \end{bmatrix}_{(N\times m)\times(n-m+1)} \tag{1}$$

上式中, 矩阵元素 $\boldsymbol{x}_i^j$, $(i=1, 2 \cdots n; j=1, 2 \cdots N)$ 其下标 i, j 分别为样本序号(时序)和地理空间站点或变量序号, m 为最大时迟步数(又称嵌套维数)。对此矩阵作 EOF 展开, 其分量形式为

$$\boldsymbol{x}_{li+j} = \sum_{k=1}^{N\times m} a_i^k \boldsymbol{E}_{lj}^k \tag{2}$$

$$a_i^k = \boldsymbol{X}_i \cdot \boldsymbol{E}^k = \sum_{l=1}^{N} \sum_{j=1}^{m} \boldsymbol{x}_{li+j} \boldsymbol{E}_{lj}^k \tag{3}$$

这里, 特征向量 $\boldsymbol{E}^k$ 是 $N\times m$ 维的, 分量 $\boldsymbol{E}_{lj}{}^k$ 是第 k 个特征向量在 l 通道滞后 j 的分量, 它既反映了空间型(随 l 变化), 也反映了时间演变(随 j 变化), 是 N 维向量(空间型)的 m 个不同时间滞后的序列, $\boldsymbol{E}^k$ 又称为空间-时间域的 EOF(即 ST-EOF)。而其时间系数 $a_i{}^k$ 则是第 i 个状态 X_i 在 $\boldsymbol{E}^k$ 上的投影, $a_i{}^k (1 \leqslant i \leqslant N-M+1)$ 序列称为第 k 个空间时间主成分(简称 ST-PC)。

显然, SSA 和 PCA 都是 MSSA 的特例, MSSA 中 $N=1$ 就是 SSA, $m=1$ 时就是常规的 PCA。

类似于 SSA, 这里, 当仅使用第 k 个特征成分来重建耦合振荡分量序列(RCCS)时, 公式为

$$\boldsymbol{X}_{li}^k = \begin{cases} \dfrac{1}{m} \displaystyle\sum_{i=1}^{m} a_{ij}^k \boldsymbol{E}_{ij}^k \\ \quad m \leqslant i \leqslant n-m+1 \\ \dfrac{1}{i} \displaystyle\sum_{i=1}^{i} a_{ij}^k \boldsymbol{E}_{ij}^k \\ \quad 1 \leqslant i \leqslant m-1 \\ \dfrac{1}{n-i-1} \displaystyle\sum_{j=i-n+m}^{m} a_{ij}^k \boldsymbol{E}_{ij}^k \\ \quad n-m+2 \leqslant i \leqslant n \end{cases} \tag{4}$$

上式中, $\boldsymbol{X}_{li}^k$ 表示场中第 l 个变量在第 i 个时次的第 k 个重建分量。显然, 在此基础上, 可以选取感兴趣的某些特征成分的子集 A 来重建耦合振荡分量序列,

$$\boldsymbol{X}_{lt}^A = \sum_{k\in A} \boldsymbol{X}_{lt}^k \tag{5}$$

为了提取预报变量场和预报因子场的显著耦合振荡信号, 可分别对 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 做 MSSA 后重建上述耦合振荡分量序列(RCCS), 即分别提取其前 k_1 和 k_2 个耦合振荡分量序列作为其主要信号, 由此构成新的变量场 $\boldsymbol{X}^*$ 和 $\boldsymbol{Y}^*$, 则它们必然包含了原预报因子场和预报变量场的主要耦合振荡信息。换言之, 新变量场 $\boldsymbol{X}^*$ 和 $\boldsymbol{Y}^*$ 可认为是从原预报因子场 $\boldsymbol{X}$ 和预报变量场 $\boldsymbol{Y}$ 经 MSSA 过滤后的主要信号场。若对新变量场 $\boldsymbol{X}^*$ 和 $\boldsymbol{Y}^*$ 作奇异值分解(SVD)[13, 14], 就可有

$$\boldsymbol{U} = \boldsymbol{C}'\boldsymbol{X}^* \quad \boldsymbol{V} = \boldsymbol{D}'\boldsymbol{Y}^* \tag{6}$$

在 $\boldsymbol{C}$, $\boldsymbol{D}$ 的列向量为正交化向量的条件下, 即若

$$\boldsymbol{C}\boldsymbol{C}' = \boldsymbol{I} \quad \boldsymbol{D}\boldsymbol{D}' = \boldsymbol{I} \tag{7}$$

使得 $\boldsymbol{U}$, $\boldsymbol{V}$ 之间有极大化协方差

$$\mathrm{cov}(\boldsymbol{U}, \boldsymbol{V}) = \boldsymbol{C}'\boldsymbol{X}\boldsymbol{Y}'\boldsymbol{D} = \boldsymbol{D}'\boldsymbol{Y}\boldsymbol{X}'\boldsymbol{C} = \begin{Bmatrix} \Sigma & 0 \\ 0 & 0 \end{Bmatrix} = \max \tag{8}$$

式中 $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2 \cdots \sigma_p)$, 其中 σ_i 为 $\boldsymbol{X}^*$ 和 $\boldsymbol{Y}^*$ 的交叉协方差矩阵的奇异值, 这里, 称 $\boldsymbol{C}$ 的列向量为交叉协方差矩阵的左奇异向量, 称 $\boldsymbol{D}$ 的列向量为交叉协方差矩阵的右奇异向量, 而时间系数 $\boldsymbol{U}$, $\boldsymbol{V}$ 分别表示相应的空间分布型 $\boldsymbol{C}$, $\boldsymbol{D}$ 随时间的演变。由此可见, 奇异向量的线性组合近似地描述了原资料场的主要耦合振荡。即由式(6)~(8), 可有

$$\boldsymbol{X}^* = \boldsymbol{C}\boldsymbol{U} \tag{9}$$

$$Y^* = DV \tag{10}$$

利用左、右奇异向量的正交性,类似于典型相关分析(CCA)方法,可建立两个场之间的回归关系(预报关系),即

$$Y^* = H' X^* \tag{11}$$

式中 $H' = DVU'(UU')^{-1}C'$; X^* 为预报因子场, Y^* 为预报变量场。很显然,这里的系数矩阵 H' 与原始场 X 与 Y 的 MSSA 的嵌套维数 m、提取重建分量的个数 k_1 与 k_2 以及 SVD 中提取主要模态的个数 k_3 有关。因此,参数(m, k_1, k_2, k_3)的选择直接影响模式(11)的预测估计精度。

3　预报试验方案的设计

上述 X 和 Y 可表示:(1)不同气象变量场的各有限格点的同期观测序列;(2)同一变量场具有不同时间后延的各有限格点的观测序列;(3)某几个气象变量场各自的混合场序列,等等。

文中考虑的预报变量场是厄尔尼诺各海区(记为 Nino 区,下同)的各季节平均值。因此,预报变量场是由 Nino 区海温的某一季节内的每一个月的海温值组成,最后的预报结果由季节内各月预报值的平均值构成。而将前期不同间隔时期的相应 SST 作为其预报因子场,即预报因子场和预报变量场并非两个同期空间场,而是同一变量的两个不同时滞值所构成的场,即(2)的情况。这里,将每个月的 SST 作为一个变量,它等价于空间场的一个格点变量。对某一季节而言(含 3 个月),代表 3 个变量,相当于 3 个空间格点。换言之,假定以 SST 作为预报对象,前期同一区域的 SST 作为预报因子,如再加入其它预报因子,也作同样处理。即将每个月的某一因子作为变量,若以连续 m 个季的 n 个因子作为预报因子场,则相当于有 $3\times m\times n$ 个变量,等价于空间场的 $3\times m\times n$ 个格点。基于上述考虑,其预报试验方案可设计如下:

(1)选取赤道东太平洋 Nino1+2 区、Nino3 区和 Nino4 区某个季节的各月平均 SSTA 构成预报变量场;选取相应海区的前期 SSTA 构成预报因子场或再加入相应时期(如季度)的其它预报因子(如表 1)。

表 1　所用变量的资料情况

预报因子	简记	经纬度	资料长度
Nino1+2 区 SST	Nino1+2	0~10°S, 90~80°W	1951 年 1 月~1997 年 2 月
Nino3 区 SST	Nino3	5°N~5°S, 150~90°W	1951 年 1 月~1997 年 2 月
Nino4 区 SST	Nino4	5°N~5°S, 160°E~150°W	1951 年 1 月~1997 年 2 月
南方涛动指数	SOI		1900 年 1 月~1999 年 2 月
全球月平均气温	GLT		1856 年 1 月~1997 年 12 月

(2)取前期一个季节的要素场作为预报因子场的基本序列,分别按提前 1, 2, 3, 4 个季拟定预报方案,即起报时刻与预报时刻分别差 1, 2, 3, 4 个季度。例如,如果要预报某一年冬季海温,可以分别选取当年秋季(提前 1 季)、当年夏季(提前 2 季)、当年春季(提前 3 季)、前一年冬季(提前 4 季)的要素场作为预报因子场。

(4)设计两种预报试验方案:①以原始资料(全部)序列(1951~1996 年共 46 a)为基础资料对有关参数作敏感性试验,并在建立 MSSA-SVD 预报模型的基础上作独立样本的预报试验,用于考察预报效果的一方面依据;②以原始资料序列的前 40 a(即 1951~1990 年共 40 a)为建模资料序列,对 1991~1997 年的各 Nino 区海温进行实际预报试验。

4　模式参数对预报结果的敏感性

MSSA-SVD 预报模式方程为 $Y^* = H' X^*$,如前所知,其系数矩阵 H' 与预报因子场 X 和预报变量场 Y 的 MSSA 嵌套维数、提取重建分量的个数 k_1 与 k_2 以及 SVD 中提取主要模态的个数 k_3 有关。因此,(m, k_1, k_2, k_3)参数的选择直接影响模式式(11)的预测估计精度。现分别讨论预报模式对各个参数的敏感性。

4.1　嵌套维数对模式的影响

在 MSSA 分析中适当选取嵌套维数 m 是一个关键。一般说来,m 愈大,谱分辨率愈灵敏。试验表明,m 一般不应超过原序列长度的 $\frac{1}{3}$。为此文中取 $m\leqslant 15$,作参数的敏感性试验,分析 m 取何值时

能使模式的预报相关达到最大。试验结果表明, *MSSA* 的嵌套维数受资料长度限制而不宜取得过大, 其最佳取值随不同的海区、不同的季节及不同的预报时效均有所差异, 但总的来说当 m 取 3 至 6 时, 预报都能取得较高的相关。表 2 是以 Nino3 区为例的试验结果。对各 Nino 区各季节的海温做相同的预报试验, 也得到类似结论。

表 2 Nino 3 区海温预测的最优参数组合

预报季	1 季	2 季	3季	4 季
春	(4 1 1 2)	(3 1 1 2)	(3 3 3 4)	(4 1 1 1)
夏	(6 6 6 7)	(3 1 1 1)	(3 1 1 1)	(3 3 3 5)
秋	(6 6 6 2)	(3 1 1 2)	(4 2 2 2)	(4 4 4 2)
冬	(3 3 3 5)	(3 3 3 5)	(3 1 1 2)	(6 3 3 2)

注: 表中括号内的 4 个数字依次是 m, k_1, k_2 和 k_3

4.2 重建分量数 k_1, k_2 对模式的影响

独立样本预报试验表明, 在建模过程中, 选取预报因子场和预报变量场的 MSSA 重建耦合振荡分量序列的个数 k_1, k_2 与 MSSA 的嵌套维数有关, 一般情况下, 其个数大约等于嵌套维数, 能取得较好的预报效果。但有的海区季节预测 k_1, k_2 的选取有些差异。MSSA 与 SSA 的差别在于每个空间时间主成分(SF-PC) 涉及到所有通道的谱性质, 所以由 MSSA 的特征成分与时间主成分重建的序列 RCCS 提取的信号包含了所有通道的耦合振荡信息。在建模过程中, 预报因子场和预报对象场中 RCCS 个数的选取对预报模式的影响是显然的, 过少会使提取的信息不够, 预报精度降低; 过多则可能使噪声分量增大, 信噪比减小, 导致预测精度下降。由于文中试验的预报因子场和预报对象场均用海温资料, 故 k_1 和 k_2 的选取差别不大, 试验也证明了这一点。所以在后文的预报中均取 $k_1 = k_2$。

4.3 SVD 模态数 k_3 对模式的影响

此外, SVD 模态数的选取对预报模式同样有影响。SVD 中模态个数的选取也不宜过大, 过大会使历史信息中的噪声分量增大, 预测信噪比减小, 导致预报精度降低。具体个数的选取与前几个模态的累积平方协方差贡献有关。一般说来, 取多少个奇异值(k_3) 所对应的典型分布型, 可通过各模态对两个场的交叉协方差的贡献来确定, 即通过对奇异值的显著性检验(或确定累积平方协方差贡献) 来定。在固定参数 m, k_1, k_2 的情况下, 从预报效果随 k_3 的变化可见, 模式对参数 k_3 也较敏感。当 k_3 取一定的值后能使预报相关较高, 随着 k_3 的增大, 预报相关反而逐渐降低。这也许是因为 k_3 的增大使得历史资料中的噪声分量增加, 预测信噪比减小, 导致预测精度降低的缘故(图略)。通过试验, 发现当 k_3 取 1~ 6 时预报就能达到较高的相关。综上所述, 模式的 4 个参数(m, k_1, k_2, k_3) 对模式的预报效果均有影响, 但影响的形式有所不同。在实际建模过程中需要用试验来寻找最优的参数组合。表 3 列出了固定参数 $m^* = 4$, $k_1^* = k_2^* = 4$, $k_3^* = 2$ 时 Nino3 区海温的预报相关和调整参数取得的最优预报相关, 由表可见, 调整参数后的最高预报相关比固定参数的预报相关在整体上要高得多, 当然, 各区各季增加的情况有所不同。

表 3 固定参数和调整参数的预报比较(Nino3 区)

预报季	1 季		2 季		3季		4 季		平均	
春	0.32	0.52	0.43	0.54	0.06	0.36	-0.14	0.27	0.17	0.42
夏	0.64	0.77	0.04	0.30	-0.05	0.11	-0.01	0.02	0.15	0.30
秋	0.86	0.86	0.44	0.72	-0.10	0.21	-0.00	0.00	0.30	0.45
冬	0.87	0.89	0.71	0.77	0.32	0.76	0.06	0.07	0.49	0.62
平均	0.67	0.76	0.40	0.58	0.06	0.36	-0.02	0.09	0.28	0.45

注: 表中左列为固定参数的预报相关, 右列为调整参数的最优预报相关。

5 各 Nino 海区独立样本预报试验

独立样本预测试验表明, 各 Nino 区用 MSSA-SVD 为基础所建立的预报模式, 一般都优于其它统计预测方法(如 SSA, CCA 或 SVD 及 PC-CCA), 从预报效果看, 也与其它方法有类似之处, 即不但各 Nino 区预报效果有所不同[12~15], 而且各 Nino 区不同季节预报效果也有显著差异。现分述如下:

就本方法与其它方法的比较而言, 以 MSSA-SVD 与单独用 SVD 方法的预报对比为例, 一般以 MSSA-SVD 方法求得的预报相关高于 SVD 方法, 不过对各季和不同的预报时效来说, 情况略有不同。

提前 1 季的预报, 春季的相关均低于 SVD 方法。因为, 一般来说典型的 ENSO 事件往往在春季建立, 往往春季海温距平在时间上具有准随机的不可连续性, 导致其准周期振荡较弱, 故用 MSSA 提取

的春季振荡信号较差。MSSA-SVD 方法求得的预报相关低于单纯 SVD 方法可能与此有关。而 MSSA-SVD 方法与 SVD 方法的最大差异在于, 提前 2 季以上的预报, 前者比后者的预报相关高得多, 一般相差达 0.1~ 0.3, 甚至更大。可见对预报时效较长的预报来说, MSSA-SVD 方法要优于 SVD 方法。其次, 各 Nino 区海温预报的差异表明, Nino4 区预报效果最好, Nino1+ 2 区预报效果最差, 这与文献[9, 10] 结论一致。其原因可能是, 这些预报结果与 ENSO 暖事件随时间在空间上的发展演变特点有关。典型 ENSO 暖事件是冬季和第 2 年春季首先在南美沿岸(Nino1+ 2 区) 开始出现, 之后由东向西发展, 大约在第 2 年的夏季和早秋延伸到赤道中太平洋(Nino3 区、Nino4 区)。由于典型的 ENSO 事件一般在 Nino1+ 2 区建立, 使得 Nino1+ 2 区的海温异常在时间演变上具有准随机的不可连续性, 导致了 Nino1+ 2 区的海温预报相关技巧较低, 而当海温异常传播到 Nino3 区、Nino4 区后 ENSO 事件已经发展起来, 因此预报准确性要高于 Nino1+ 2 区, 尤其是 Nino4 区预报。预报事件的持续当然要易于预报事件的先兆, 这也许能解释为什么 Nino1+ 2 区的海温预报效果最差而 Nino4 区的海温预报效果最好。就预报效果的季节差异来看, 在 3 个海区中都以春季效果最差, 而预报效果最好的季节在各区有所不同: Nino1+ 2 区的秋季预报相关最高, 而 Nino3 和 Nino4 区则是冬季预报效果最好。从提前 1, 2, 3 和 4 季的预报来看, 各区均较好。但其中 Nino1+ 2 区和 Nino3 区提前 1 季的预报效果较好, 而提前 2, 3 和 4 季的预报 Nino1+ 2 区比较差, Nino3 区相对较好。图 1a~ d 分别是 Nino3 区提前 1, 2, 3, 4 季的 SSTA 预报与实况比较。

由此可见, 热带海气耦合系统的可预报性依赖于季节, 文献[5] 曾指出, 以北半球冬季作初始场的预报, 其评分低于以春季作初始场的预报, 这表明存在着一个北半球春季的"可预报性障碍"。短期 ENSO 预报技巧的季节性差异在其他一些统计或动力预报模式的研究中也有发现, 被认为是固有的季节变化对 ENSO 自身循环的影响, 已被一些观测研究所证实, 本文所试验的结果基本上与这些研究结果吻合, 限于篇幅, 不作详述[1, 2, 4, 5, 11]。

6　增加预报因子的预测试验

ENSO 是全球海气相互作用下的强信号, 研究 ENSO 预测有必要引进海气系统的特征量[15~ 17]。显然, 当预报因子场增加其它因子时, 新构成的预报因子场其周期振荡可能会发生变化。为此, 在原始预报因子场中分别加入全球气温(简记 GLT) 、南方涛动指数(SOI) 、或同时加入 GLT 和 SOI 构成 3 种新的预报因子场。以 Nino3 区预报为例的试验结果表明, 绝大部分的预报相关比单纯考虑前期 SSTA 场作预报因子场有显著提高, 其中以提前 3 季和提前 4 季的预报提高更显著。从季节效果来看, 春季仍然最差, 效果最好的是秋、冬两季。对提前 1~ 2 季的预报来说, 预报相关一般都有一定的增加, 而对提前 3~ 4 季的预报来说, 增加得更为显著(冬季预测的增加略低)。

比较仅增加 GLT 和仅增加 SOI 的预报效果, 可以发现, 增加 SOI 因子时的预报效果略高于增加 GLT 的效果。实际上, 增加 GLT 时, 提前 1 季和提前 2 季的各个季节预测与单纯考虑前期 SST 场变化不大, 有的略微增加, 有的反而略低, 而增加 SOI 后的预报效果就显著得多, 一般都能增加 0.05 左右。若同时增加 GLT 和 SOI 因子, 其预报效果的稳定性反而较差, 有些季节虽提高很多(如提前 1 季的秋季预测, 相关高达 0.934) 但是许多的季节预测并无显著增加, 而有的甚至下降许多。因此, 对提前 1~ 2 季的 Nino3 区海温预测来说, 引入 SOI 来提高该区的预报技巧有最佳的效果[16, 17]。而对某些季节来说, 同时考虑 SOI 和 GLT(还有其它的各种因子) 也有一定的意义(图表略)。

7　实际预测试验

进一步以 1951~ 1990 年资料为样本序列, 分别对各海区建立 MSSA-SVD 预报模型, 对未来 6 a(1991 年春季~ 1996 年冬季) 的标准化 SSTA 作实际预测试验。选定 1951~ 1990 年的某一海区某季标准化 SSTA 为预报场, 以提前 k 季的前期标准化 SSTA 场为预报因子场, 求出预报模式的回归系数矩阵, 然后将 1990~ 1995 年的提前 k 季的前期标准化 SSTA 资料代入预报方程, 作提前 k 季的标准化 SSTA 预报。以冬季为预报因子为例, 将 1990 年冬季标准化 SSTA 代入预报方程, 分别提前 1 季预报次年春季、提前 2 季预报次年夏季、提前 3 季预报次年秋季和提前 4 季预报次年冬季共 4 个季度的预报值。仿照这种方法, 分别用春、夏、秋和冬 4 个季度作为预报因子作 1991~ 1996 年的预报, 其预报相

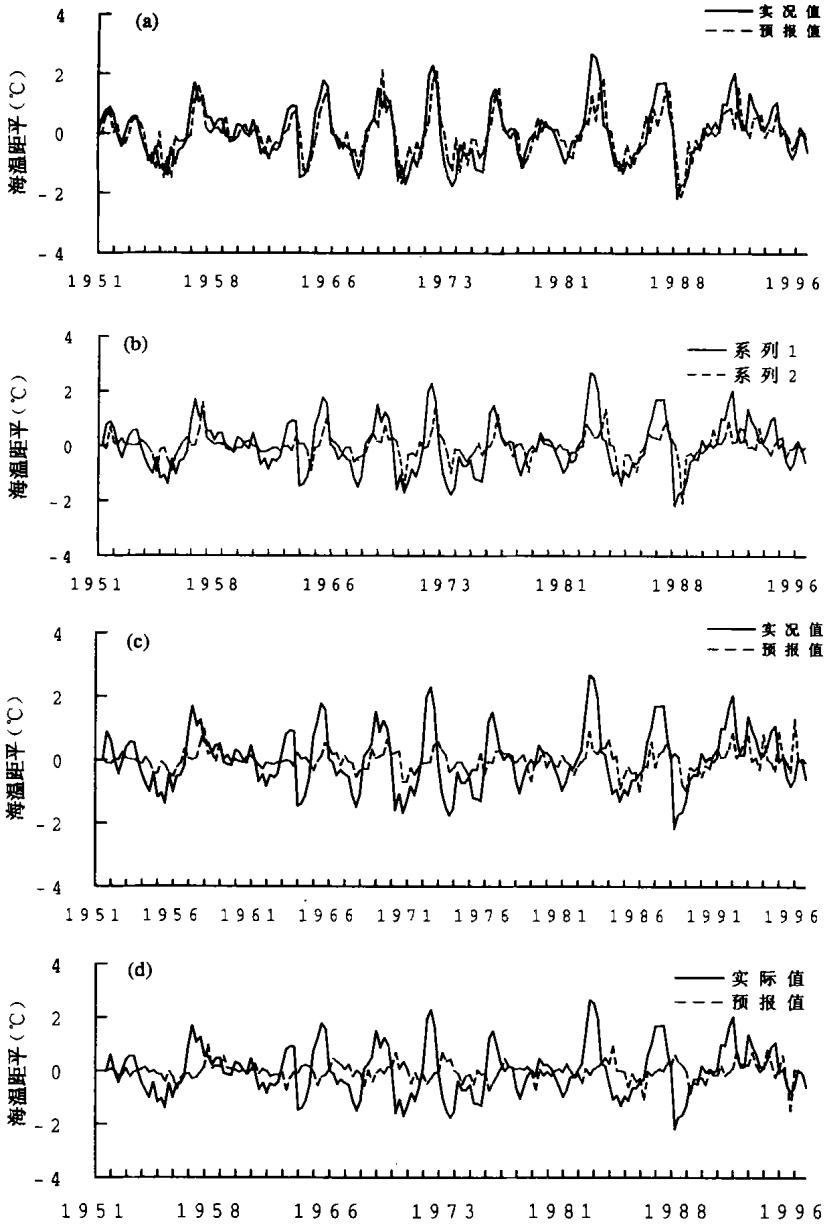

图 1 Nino3 区提前 1~ 4 季的预报与实测距平序列对比

(a, b, c, d 分别是提前 1, 2, 3, 4 季对 Nino3 区海温的独立样本预报试验)

关分别为 0. 574, 0. 529, 0. 135 和 0. 361。图 2 是 Nino4区的预报值和实际值曲线。由图可见, 以春季为因子场的预报序列与实测距平序列趋势非常相近, 1991~ 1995 年连续发生的 3 次 El Niño 事件在预报序列中均能较好的反映出来, 只是春季的预报值比实测值要小。以夏秋冬 3 季为预报因子的预报序列与实测距平序列差异较大, 主要反映在位相上有差异, 由于春季的预报值与实际值相差较大, 而且提前4 个季的预报较差, 这些原因导致了预报相关较低。而用每个季节提前 1 个季节预报下个季节的值, 效果相当好。图 2 是该区提前 1 个季节的预报值与实际值的比较, 显然, 夏、秋、冬 3 个季节的预报值与实际值差别不大, 趋势也大致相同, 但春季的预报值一般小于实际值, 这使得提前 1 个季的预报相关不高(r = 0. 55) 。

8 结论与评述

(1) 本文提出一种多通道奇异谱分析与奇异值分解相结合(即 MSSA-SVD) 的方法, 提取预报因子场和预报变量场的显著耦合振荡信号, 并在此基础

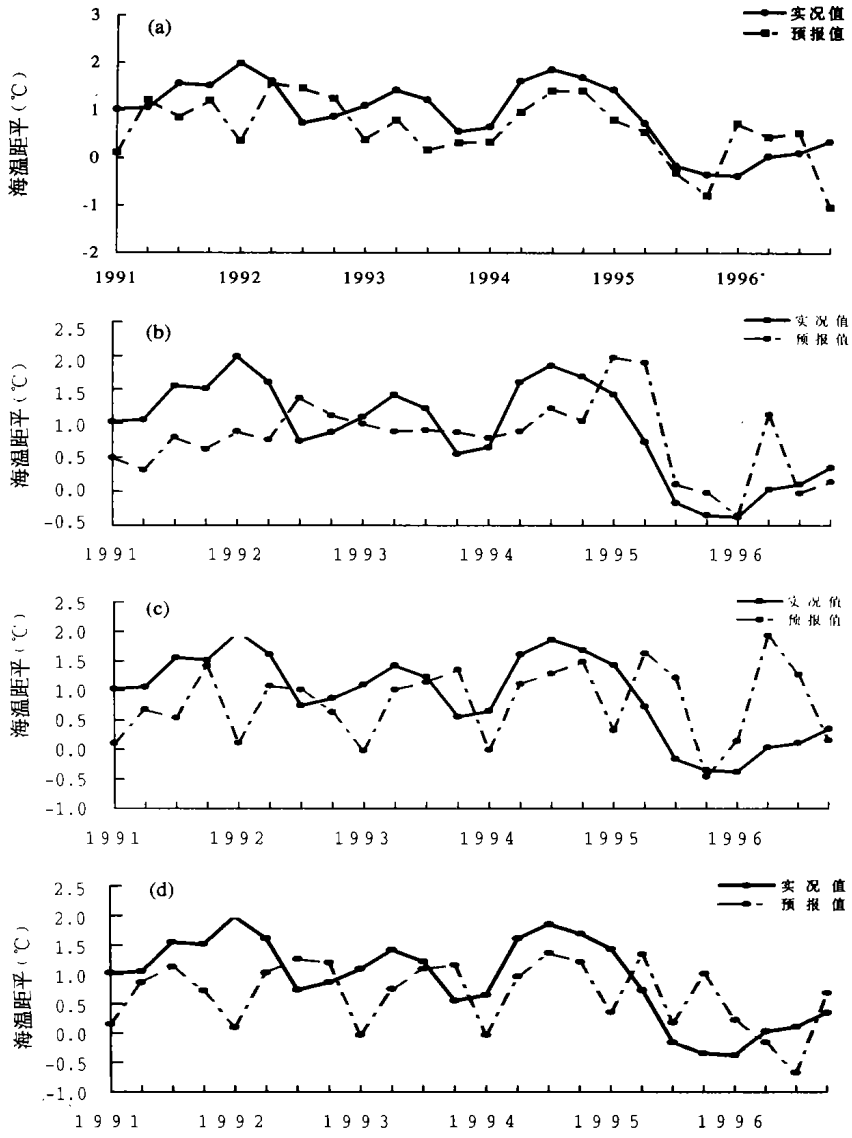

图 2　以各季海温为预报因子的 Nino4 区海温距平预报

(a, b, c, d 分别是以各季海温为预报因子场作 1991~ 1996 年 Nino4 区海温预报)

上建立一种广义典型混合回归模式用于 ENSO 短期气候预测。对 Nino 海区各季 SSTA 的短期气候预测试验表明, 其预报效果在总体上要优于其它统计预报方案。从理论上说, 基于 MSSA-SVD 的典型混合回归集中了 CCA 或 SVD 及作者以前提出的 PG-CCA 典型混合回归的优点, 因而其效果优于其它方法。

(2) 独立样本预报试验表明, 不但各 Nino 区预报效果有所不同, 各 Nino 区不同季节预报效果也有显著差异。其中以 Nino 4 区效果最好, Nino1+ 2 区效果最差。秋冬季预报较好(除某些年份的预报振幅小于实况值外, 预报与实况曲线基本重合。夏季预报曲线与实况趋势也大致相同, 但在整体上, 预报振幅比实况小); 而春季的预报效果最差。从提前 1, 2, 3 季和 4 季节的各区各季独立预报试验结果来看, 提前 1 个季节的预报效果最好(相关在 0. 80 以上), El Niño 事件和 La Nina 事件几乎都与实际一致。仅只有个别 ENSO 年份预报的异常值要小于实际值, 而提前 2 季的预报曲线与实况也比较一致, 相关大多在 0. 61 左右。但提前 3 季以上的预报只在趋势上较一致, 其预报相关为 0. 53 左右。

(3) 关于增加因子的试验表明, 若增加 GLT 和 SOI 或同时增加二者时, 各 Nino 海区海温的预报准确率一般都有显著提高, 但对不同的海区、不同的季节和不同的预报时效, 预报相关的变化情况有一定

差异。

(4) 模式对 4 个参数(m, k_1, k_2, k_3) 的敏感性较高, 它们对模式的预报效果均有影响, 但影响的形式有所不同。在实际建模过程中需要用试验来寻找最优的参数组合, 方能有最高的预报效果。

参考文献

1 Penland C, Mgorian T. Prediction of Nino3 sea-surface temperatures using linear inverse modelling. J Climate, 1993, 6: 1067~ 1076

2 赵宗慈. ENSO 现象及其成因的模拟与预测研究进展. 气象科技. 1995, 141(1) : 4~ 12

3 Xu J S, Storch H. Principal oscillation patterns-prediction of state of ENSO. J Climate, 1990, 3: 1316~ 1429

4 Barnston A G, ropelewski C F. Prediction of ENSO episodes using CCA. J Climate, 1992, 5: 1316~ 1345

5 Ding Yuguo, Jiang Zhihong. Study on canonical autoregression prediction of meteorological elements fields. Acta Meteor Sinica, 1996, 10(1) : 41~ 51

6 丁裕国, 江志红, 朱艳峰. Nino 海区 SSTA 短期气候预测模型试验. 热带气象学报, 1998, 14(4) : 289~ 296

7 丁裕国, 江志红, 施能. 奇异交叉谱分析及其在气候诊断中的应用. 大气科学, 1999, 23(1) : 91~ 99

8 江志红, 丁裕国, 周琴芳. 用于 ENSO 预测的一种广义典型混合回归模式及其预报试验. 热带气象学报, 1999, 15(4) : 322~ 329

9 余锦华, 丁裕国. Nino 区 SST 与 SOI 的耦合振荡信号及其预测试验. 南京气象学院学报, 1999, 22(4) : 637~ 643

10 余锦华. 厄尔尼诺海区海温季节预测比较. 南京气象学院学报, 1999, 22(3) : 374~ 380

11 Plaut G, Vautard R. Spells of low-frequency osillation and weather regimes in the Northern Hemisphere. J Atmos Sci, 1994, 51: 210~ 236

12 丁裕国, 江志红. SVD 方法在气象场诊断分析中的普适性. 气象学报, 1996, 54(3) : 365~ 372

13 Wallace J M, et al. SVD of wintertime sea surface temperature and 500 mb height anomalies. J Climate, 1992, 5: 567~ 576

14 江志红, 丁裕国. 奇异谱分析的广义性及其应用特色. 气象学报, 1998, 56(6) : 732~ 745

15 Pan Y H, Oort. Global climate variations connected with sea surface temperature anomalies in eastern equatorial Pacific Ocean for the 1958~ 1973 period. Mon Wea Rev, 1983, 111: 1244~ 1258

16 Rasmusson E M, Carpenter T H. Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/ El Nino. Mon Wea Rev, 1982, 110: 354~ 384

A PREDICTION EXPERIMENT BY USING THE GENERALIZED CANONICAL MIXED REGRESSION MODEL BASED ON MSSA-SVD FOR ENSO

Ding Yuguo　Cheng Zhengquan

(Nanjing Institute of Meteorology, Nanjing 210044)

Cheng Bingyan

(Henan Climatic Center, Zhengzhou 450003)

Abstract

A generalized Canonical Mixed Regression Model based on MSSA-SVD is presened to prediction of ENSO. The MSSA is a Multichannel Singular Spectrum Analysis and the SVD is Singular Value Decomposation. The basisc idea of the method is that (1) the prominent coupled oscillation signals are segregated between the forecasted fields and the forecastor fields by using MSSA-SVD method; (2) the generalized Canonical Mixed Regression Model is constructed according to the first several prominent coupled oscillation patterns for ENSO prediction. The results of statistical forecast test based on the MSSA-SVD method show that the predictional model possesses more advantages and better effects than other statistical prediction methods.

Key words: Multichannel singular spectrum analysis, Canonical mixed regression model, ENSO prediction.