

EOF-CCA 模型在山东季降水 预测中的应用^①

胡桂芳 张苏平 谢考宪

(山东省气象局, 济南 250031)

提 要

首先介绍了 EOF-CCA 模型的基本原理及建模步骤, 阐明了经验正交函数分解 (EOF) 和典型相关分析 (CCA) 两种统计方法在短期气候预测应用中结合的必然性及优越性, 并指出该模型可由一个气候场或多个气候场去预报另一个气候场。然后以山东 13 站的季降水量场为预报场, 分别以前期整个北半球和东亚区域的 500hPa 月平均位势高度场为因子场, 建立了位势高度场的前 13 个主分量场与季降水场之间的 EOF-CCA 模型, 最后对模型的拟合、预测效果进行了评估, 结果表明: 模型对山东的季节性降水有一定的预报技巧, 可在业务中试验应用。

关键词: EOF-CCA 模型 降水量 季预测

引 言

EOF-CCA 模型是应用自然函数正交分解 (EOF) 和典型相关分析 (CCA) 两种方法建立一种统计预测模型^[1,2]。它是既考虑现象联系, 又考虑成因特征的新型、高效的统计预报方法。美国国家气候预报中心已将 CCA 为基础的统计预报模式应用于长期预报业务, 我国也开始将该模型应用于短期气候预测业务中^[3,4], 山东地理位置特殊, 预报难度较大。本文尝试用该方法对山东的季降水场进行预测试验。

1 资料来源及预处理

本文选用 1951 年 1 月~1999 年 8 月山东分布较为均匀的德州、惠民、临清、济南、淄博、泰安、兖州、菏泽、临沂、潍坊、莱阳、青岛、烟台 13 站的四季降水量场为预报场, 分别以前期北半球 (10~80°N) 576 个格点 (简称 M576) 和东亚 (20~60°N, 60~140°E) 共 81

个格点 (简称 M81) 的 500hPa 月平均位势高度场 (5°×10°) 为预报因子场。

2 EOF-CCA 模型的基本原理及计算步骤

2.1 EOF-CCA 模型的基本原理

EOF 分解在气象统计分析中已被广泛应用, 本文着重阐述典型相关 (CCA) 分析的基本原理。CCA 是在 EOF 基础上发展起来的多变量统计技术。它的基本思想是把气候场中每一个网格点看成一个变量, 对两组变量分别作线性组合构成一对新变量 U_1, V_1 , 使得它们之间有最大的相关系数。再分别作出 U_1, V_1 独立的线性组合 U_2, V_2 , 使得它们之间有次大相关系数。如此进行下去, U_i, V_i 称为典型变量。设因子场 x_1, x_2, \dots, x_p 共 p 个空间点, 预报场 y_1, y_2, \dots, y_q 共 q 个空间点, 资料阵为:

$$X = (x_1, x_2, \dots, x_p)' \quad (1)$$

$$Y = (y_1, y_2, \dots, y_q)' \quad (2)$$

^① 中国气象局“短平快”项目“卫星遥感资料在山东旱涝分析、预测中的应用”资助
国家重中之重科技项目 96-908-05-06-10 研究成果

因子组原 p 个变量的线性组合为一新变量:

$$\begin{aligned} u_1 &= c_{11}x_1 + c_{21}x_2 + \dots + c_{p1}x_p \\ c_1 &= (c_{11}c_{21}\dots c_{p1})' \end{aligned} \quad (3)$$

预报两组原 q 个变量的线性组合为一新变量:

$$\begin{aligned} v_1 &= d_{11}y_1 + c_{21}x_2 + \dots + c_{p1}y_p \\ d_1 &= (d_{11}d_{21}\dots d_{q1})' \end{aligned} \quad (4)$$

典型变量必须具有如下特性:

①平均值为零,方差为1,即:

$$u_1 u_1' / n = c_1' s_{11} c_1 = 1 \quad (5)$$

$$v_1 v_1' / n = d_1' s_{22} d_1 = 1 \quad (6)$$

②它们之间的相关系数(称作典型相关系数)在两组变量的所有的线性组合而成的新变量中最大,即:

$$r_1 = u_1 v_1' / n = c_1' s_{12} d_1 \text{——最大} \quad (7)$$

在式(5)、(6)、(7)的条件下求二次型条件极值,经变换可以导出以下特征值和特征向量方程组的矩阵表达式:

$$(s_{22}^{-1} s_{21} s_{22}^{-1} s_{12} - \lambda_1 I) d_1 = 0 \quad (8)$$

利用求非对称阵的特征值和特征向量的方法求解。求出 λ_1 和系数 d_1 后,利用关系式:

$$c_1 = s_{11}^{-1} s_{12} d_1 / \lambda_1^{-1/2} \quad (9)$$

第一对典型变量的相关系数为:

$$r_1 = \lambda_1^{-1/2}$$

依次可以计算出第二、三、……对典型变量的组合系数及典型相关系数。

典型相关系数是否显著需通过检验,通过显著性检验的典型变量代表了两组变量之间的线性协方差关系的主要信息。利用这种关系建立两组变量的回归关系。再利用典型变量与原两组变量关系建立两组变量的直接的回归方程:

$$Y = S_{22} D A^{1/2} C' X \quad (10)$$

由方程(10),可以由气候场 X 去预报另一气候场 Y 。但前提是两个气候场资料阵各自的方差阵 $S_{11} S_{22}$ 必须有逆存在,而实际资料阵的空间格点数往往大于时间序列长度,即使空间点数小于时间序列长度,因不同格点序列

有很大的相关性,必将导致 $S_{11} S_{22}$ 没有逆存在。因此,在进行 CCA 过程之前,首先对因子场进行 EOF 展开,提取前 m 个主分量,再用 CCA 方法分析前 m 个主分量之间的典型相关,从而建立起 EOF-CCA 预报模型。这样既减少了计算量,又可避免 CCA 过程中可能出现的逆矩阵蜕化现象。

2.2 具体建模步骤

①对资料进行标准化,而后进行 EOF 分解。因本文预报场为小区域降水场,站点数不多 ($m = 13$),所以仅对因子场进行 EOF 分解,提取前 13 个主分量与降水场之间进行 CCA 分析。为了便于计算,对 L 年 576 个格点高度场组成的矩阵 $F_{L,576}$ 的 EOF 分解转换为对其转置矩阵 $F_{576,L}$ 进行计算,这样所得到的特征向量 $C_{P1}, C_{P2}, \dots, C_{PL}$ 与主分量 $V_{P1}, V_{P2}, \dots, V_{P576}$ 与 $F_{L,576}$ 更换了位置。参考文献 [5] 中时空转换,即令:

$$(S_P \times V_{P1})^2 + (S_P \times V_{P2})^2 + \dots + (S_P \times V_{P1})^2 \quad (11)$$

据式(11),求出 SP , 则 $S_P \times V_{P1}, S_P \times V_{P2}, \dots, S_P \times V_{PL}$ 及 $C_{P1}/S_P, C_{P2}/S_P, \dots, C_{PL}/S_P$ 为矩阵 $F_{L,576}$ 的特征向量和时间系数。对矩阵 $F_{L,81}$ 的 EOF 分解采用同样的方法进行时空转换。表 1 为 12 月份北半球 500hPa 高度场 $F_{42,576}$ 的时空转换系数 S_P 值。

表 1 12 月份北半球 500hPa 高度场的时空转换系数 S_P 值 ($\times 10^{-2}$)

P	S_P	P	S_P
1	1.5842	8	1.3927
2	1.77007	9	3.4138
3	2.0722	10	3.8686
4	2.2082	11	4.1094
5	2.29807	12	4.27163
6	2.76468	13	4.52945
7	2.91638		

②利用 CCA 方法建立因子场主分量 A 与降水场 Y 的回归方程:

$$Y_{m \times 1} = R_{m \times P} \cdot A_{P \times 1}$$

③当给定因子场某一时刻的资料时,由第一步得到的特征向量 V_P ,求出该时刻的主分量:

$$A = V_{P \times Q} \cdot X_{Q \times 1}$$

④根据 CCA 建立的回归方程,可由 X 的主分量场求出降水场:

$$Y_{m \times 1} = R_{m \times p} \cdot A_{p \times 1}$$

3 预测试验

本文把有限的资料序列分作两部分,前一部分(1951~1992 年)用于建模,后一部分(1993~1999 年)用作预报验证。预报因子场选用前期 8 个月至 2 个月。

EOF 的计算结果表明:500hPa 高度场的收敛速度快,前几个主分量场浓缩了整个气候场的主要信息,本文选用前 13 个主分量作为预报因子场,其方差贡献占总方差的 75%(M576)或 90%(M81)以上。

采用 M576 及 M81 两种方案对主分量

场和季降水量场进行典型相关分析,结果表明:山东季降水场与前期 500hPa 月平均高度场关系密切,第一典型相关系数大部分接近或超过 0.9,典型相关系数大,以其为因子场建模的历史拟合效果往往也比较显著,但预报能力却不一定很好。选取拟合、试报效果均佳的前期月平均高度场作为因子场建立预报模型。M576 方案的春季与上年 12 月,夏季与当年 4 月,秋季与当年 7 月,冬季与当年 4 月拟合及试报效果最佳。M81 方案的冬季与同年 5 月,春季也与上年 12 月,夏秋季均与当年 2 月拟合、试报效果最好。预报和拟合的效果检验标准采用中国气象局下发的“短期气候预测质量评定暂行办法”中预测评分和技巧评分进行评定,春、夏、秋三季的预测评分采用六级评分制进行评定,冬季采用四级评分制进行评定,各季预测评分和技巧评分的得分情况见表 2、表 3。

表 2 1993~1999 年四季的预测评分/%

		1993	1994	1995	1996	1997	1998	1999	预测平均	多年区域平均	拟合
春季	M ₅₇₆	48.1	75.8	93.3	95.0	87.5	7.7	42.9	64.3	70.2	81.1
	M ₈₁	48.1	92.3	82.4	95.2	63.0	7.7	69.2	65.4		83.1
夏季	M ₅₇₆	53.8	92.9	85.7	76.5	93.9	53.8	55.6	73.2	73.8	82.5
	M ₈₁	55.6	57.1	93.5	72.4	92.6	86.2	79.3	76.7		87.1
秋季	M ₅₇₆	81.3	63.0	71.4	40.7	46.7	89.5		65.4	66.3	75.6
	M ₈₁	75.0	69.2	85.7	85.7	61.5	85.0		77.0		78.5
冬季	M ₅₇₆	85.2	69.2	93.5	51.7	100.0	100.0		83.3	81.7	80.0
	M ₈₁	63.0	92.3	30.8	61.5	38.5	72.4		59.7		79.4

表 3 1993~1999 年四季的技巧评分/%

		1993	1994	1995	1996	1997	1998	1999	预测平均	拟合
春季	M ₅₇₆	-32.8	24.1	43.1	81.0	43.1	-127.7	-51.8	-3.0	27.3
	M ₈₁	-51.8	24.1	43.1	81.0	-13.8	-127.7	5.1	-5.7	34.3
夏季	M ₅₇₆	-10.6	55.8	55.8	44.7	88.9	0.5	33.6	38.4	51.8
	M ₈₁	22.6	11.5	66.8	44.7	88.9	66.8	66.8	52.6	56.6
秋季	M ₅₇₆	48.5	22.8	22.8	-15.8	-15.8	74.3		39.9	22.8
	M ₈₁	35.7	22.8	61.4	61.4	-2.9	61.4		39.9	46.1
冬季	M ₅₇₆	74.5	-1.9	87.3	10.9	100.0	100.0		61.8	44.8
	M ₈₁	61.8	36.3	23.6	36.3	74.5	87.3		53.3	45.1

由表 2、3 可见,EOF-CCA 模型对山东季节性降水预测表现出一定的预报技巧,大多数年份的预测评分在 60 分以上,技巧评分

为正值。但预报效果存在着明显的年际差异和季节差异。不同的年份预报效果差异很大,春季预报最不稳定,冬季的预报效果最为理

想。预报、拟合的水平与因子选取的区域有关, M81 选取的东亚区域的环流作预报比 M576 利用整个半球的环流作预报效果明显(冬季除外), 这主要是因东亚地区的环流直接影响山东降水, 整个半球的环流是通过环流间相互联系对山东降水产生影响。选择拟合、预报效果均优的建模方案作为最终预测方案, 春、夏、秋采用 M81 方案, 冬季采用 M576 方案。除春季外, 其他季节逐站预报的预测评分在 75 分以上, 均高于历史区域平均的预测评分。技巧评分在 40 分以上, 远高于气候概率预报。从 13 站 1993~1999 年降水实况与预报值分布图看(图略), 预报效果存在着明显的空间和季节差异。以距平符号相同或符号不同数量上只差一级为预报准确, 春季对鲁西南、半岛地区预报效果较好, 准确率为 5/7 或 6/7, 对鲁西北、鲁中地区预报效

果较差。秋季对半岛北部地区预报效果较差, 准确率仅为 3/6。其它地区预报效果较好, 准确率为 5/6。夏季对鲁西北地区预报效果良好, 准确率为 7/7, 对鲁西南、鲁中山区预报效果不太理想。冬季降水预报效果最理想, 13 个站中有 9 个站的预报准确率为 5/6 或 6/6。

季区域平均降水量与前期 500hPa 月平均环流之间的对应, 反映了降水与大气环流之间明显的韵律关系, 这种韵律关系主要表现在半年左右。最佳月份一般出现在预报季前期 5~7 个月, 即表现出明显的隔季相关现象。当降水量集中于季节中间时段时, 季平均降水量与月平均环流联系比较好, 可以作出比较好的预报, 如 1994、1996 年春季, 1995、1996 年秋季, 图 1 为 1996 年春季降水距平百分率的实况值和预报值的分布情况。

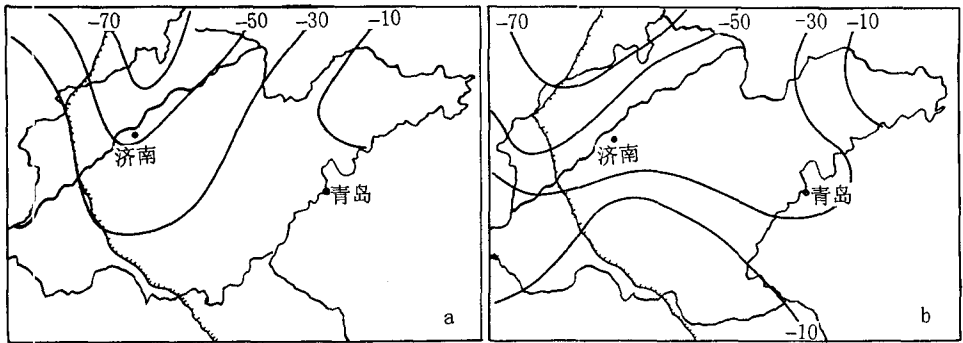


图 1 1996 年春季降水实况值(a)和预报值(b)

由图 1 可见, 实况降水量距平百分率的负值中心位于鲁西北的德州地区, 自北向南, 自西向东降水距平百分率呈增加趋势, 预报值的中心及分布趋势与实况基本一致, 同号率为 11/13, 预测评分为 95.2, 气候一级异常(降水距平百分率绝对值 $\geq 50\%$)的预报评分为 50%, 二级异常(降水距平百分率绝对值 $\geq 20\%$)的预报评分为 90%。

当降雨量集中在季节初或季节末时, 季平均降雨量与月平均环流之间的对应关系不

好, 这种情况下对降水的预报效果往往不理想, 甚至很差。1994、1996 年的冬季, 1993、1998 年春季都是这种情况, 1998 年最典型。1998 年 5 月份全省平均降雨量 90mm, 占季降雨总量的 50%以上, 而且又主要集中于 5 月下旬(5 月下旬的全省平均雨量为 50mm)。因此这一年的预报完全失败。

EOF-CCA 模型对 1999 年山东 13 站四季降水量的预报结果为: 春季鲁西北、鲁中地区偏多, 其他地区偏少; 夏季大部分地区降水

偏少 10%~30%;秋季降水均为正距平,大部地区偏多 20%~40%;冬季大部地区降水为负距平,降水偏少 10%~50%。

综上所述,本文建立的 EOF-CCA 模型充分利用了整个气候场的信息,尤其选用了直接影响山东降水的东亚区域 500hPa 月平均高度场为因子场,可能反映了东亚环流与山东降水的内在联系,是一种较为理想的预报工具。

5 结论

① EOF-CCA 模型是一种既考虑现象联系,又考虑成因特征的新型、高效的统计预报方法。该模型可由一个气候场或多个气候场去预报另一个气候场。

② 山东季降水场与前期 500hPa 月平均位势高度场关系密切,尤其与东亚区域的月平均环流场关系更佳,表现出明显的半年左右韵律关系。

③ 以 500hPa 月平均高度场为因子场的

EOF-CCA 模型对我省季节性降水预报效果较好,除冬季外,东亚区域的 500hPa 月平均高度场是山东季降水重要的预报因子。预报效果存在着较大的空间差异和季节差异,以冬季预报效果最好,春季最不稳定,对夏、秋、冬季的降水趋势具有较好的预报能力,平均最大技巧评分为 61.8。

参考文献

- 1 黄嘉佑. 气象统计分析与预报方法. 北京:气象出版社, 1990:227~253.
- 2 陈佑民,王绍武. 气候可预报性的诊断研究. 气候预测研究. 北京:气象出版社,1996:139~158.
- 3 毛恒青,李小泉. 典型相关分析(CCA)对我国冬季气温的短期气候预测实验. 应用气象学报,1994,(4):386~391.
- 4 朱洁华,徐家良. EOF-CCA 模型在区域降水预测中的试验研究. 大气科学研究及应用,1998,(1):115~119.
- 5 杨洪昌等. 北半球 500 百帕候平均环流的季节演变. 山东中长期天气预报研究文集. 北京:气象出版社,1990:41~42.
- 6 中国气象局. 短期气候预测质量评定暂行办法. 1999,7.

The Application of EOF-CCA Model to Seasonal Rainfall Prediction in Shandong Province

Hu Guifang Zhang Suping Xie Kaoxian

(Shandong Meteorological Bureau, Jinan 250031)

Abstract

The principle of EOF-CCA model and its steps of development were briefly introduced. The inevitability and superiority were illustrated when the two methods EOF and CCA were used in the combination form of EOF-CCA in the prediction of short-range climatic change. It was also pointed out that with this model another climatic field based on one or more climatic fields could be forecasted. Seasonal precipitation of 13 stations in Shandong province was the forecasted field and previous monthly average geopotential height fields at 500hPa over northern hemisphere and east Asia were used as factor fields. Thus the EOF-CCA model was set up. Finally the model's historical fit and prediction effect were assessed, which shows that the model has rather good ability in forecasting seasonal precipitation and can be used in the operation.

Key Words: EOF-CCA model Shandong province seasonal precipitation