

文章编号: 0529-6579 (1999) 04-0107-05

广东春雨的长期预报^{*}

周 文, 陈创买, 简裕庚

(中山大学大气科学系, 广州 510275)

摘 要: 利用气候场的主分量逐步回归预报方法, 通过相关分析和逐步回归, 求得春雨场的主分量与高度场、海温场以及环流指数等因子场的主分量之间的联系. 结果表明, 广东春雨场的前 15 个主分量(约占场的总方差的 97.4%)的预报方程, 通过 $\alpha=0.05$ 显著性检验, 其方程复相关系数最小为 0.729, 最大为 0.975. 对广东省长期预报的 19 个代表站的 1990 年春季(2~4 月)降水量作预报的统计结果, 其预报平均绝对误差约 90 mm.

关键词: 气候场; 主分量; 逐步回归; 相关分析; 广东春雨

中图分类号: P 468 **文献标识码:** A

对气候要素场(如降水)的预报, 在日常业务中十分重要. 有不少学者对此进行过研究^[1~4], 但他们往往仅局限于考虑两个场的相关. 而实际上气候变化不仅仅受海温场的影响, 或仅受某高度场的影响. 如何将一个气候场的预报与多个因子场的影响相联系, 是一个值得研究的问题. 文献[5~7]利用主分量回归分析的方法, 分别对美国季降水量、太平洋气压场及雷暴天气进行了研究. 主分量分析的特点就是可将气候场的绝大部分信息集中在前几个分量中, 将对气候场的预报变成对气候场的主分量的预报.

吴明进^[8]认为台湾春雨的年际变化是全球大尺度环流变化的一部分, 而此大尺度环流特征的出现又和全球尺度海面温度的分布息息相关. 海面温度可能成为台湾春雨预报的一领先指标. 广东地区与台湾地区同属低纬度地区, 2、3 月是广东的春播期(早稻), 4 月为壮秧期, 雨量多寡极为重要. 历年广东南部存在着春旱问题^[9], 准确的预报意义十分重要. 本文拟利用气候场的主分量回归预报方法作广东春雨的长期预报.

1 预报模型

气候要素场的主分量是与其特征向量即空间分布型相对应的, 在这种情况下, 所谓预报量与预报因子的相关, 实际上变成气候场的某种分布型与各因子场的一定空间分布型的相关, 它反映了各因子场的分布型对气候场的各部分有不同的影响.

1.1 预报场的标准化处理

设场是一个有 p 个空间点, m 个时间点的标准化变量场:

^{*} 基金项目: 广东省自然科学基金(930966)资助项目

收稿日期: 1998-10-20 作者简介: 周 文, 女, 1969 年生, 讲师.

$$X_{pm} = (x_{ki}) \quad k=1, 2, \dots, p; i=1, 2, \dots, m \quad (1)$$

其中, x_{ki} 为第 k 个空间点第 i 个时间点的标准化变量.

$$x_{ki} = (x_{ki} - \bar{x}_k) / S_k \quad k=1, 2, \dots, p \quad (2)$$

其中, $\bar{x}_k$ 和 S_k 分别为各空间点的平均值和标准差.

1.2 预报场的 PCA 计算

对标准化气候场 X_{pm} 进行 PCA 计算^[10], 得出

$$X_{pm} = V_p Y_m \quad (3)$$

其中, V 为 $(p \times r)$ 阶特征向量, Y 为 $(r \times m)$ 阶主分量, r 为 p 和 m 的最小值. 取前 k (≤ 15) 个主分量进行标准化, 即

$$Y_m = \Lambda^{-\frac{1}{2}} Y_{rm} \quad (4)$$

其中, $\Lambda^{-\frac{1}{2}}$ 为 r 阶对角阵.

1.3 气候场标准化主分量的回归预报

取占 95% 以上方差的前 n 个主分量逐个建立回归模型, 则第 k 个标准化主分量的回归预报模型为

$$y_{kt} = C_{k0} + \sum_{i=1}^L C_{ki} f_{kit} + e_{kt} \quad (5)$$

其中, C_{ki} 为各回归系数, f_{kit} 为各可能预报因子. L 为该回归方程所含可能预报因子数. e_{kt} 为误差. 为了获得对预报场第 k 个标准化主分量预报方程的 L 个可能预报因子, 必须对由因子场提取的大量因子进行相关筛选计算, 使用

$$\gamma_{kl} = \frac{1}{m} \sum_{i=1}^m (y_{ki} - \bar{y}_k)(f_{li} - \bar{f}_l) / (S_k \cdot S_l) \quad (6)$$

设定显著性水平 α , 若 $\gamma_{kl} > \gamma_\alpha$, 则第 l 个因子与预报量的相关显著, 可作为预选因子参加逐步回归.

对预报场第 k 个主分量进行逐步回归计算, 设有 L_k 个预报因子进入预报方程, 则

$$\hat{y}_{kt} = C_{k0} + \sum_{i=1}^{L_k} C_{ki} f_{kit} \quad k=1, 2, \dots, n; t=1, 2, \dots, m \quad (7)$$

根据前面计算, 对预报场标准化主分量的拟合预报, 可用下式表示

$$Y_{k \times (m+s)} = Y_{k \times m} Y_{k \times s} \quad (8)$$

其中, $Y_{k \times m} = (\hat{y}_{kj}) \quad k=1, 2, \dots, n; j=1, 2, \dots, m \quad (9)$

$$Y_{k \times s} = (\hat{y}_{ki}) \quad k=1, 2, \dots, n; i=1, 2, \dots, m \quad (10)$$

式中, s 为预报长度. 预报模型所作的预报结束, 从第 4 步开始, 所谓的“预报问题”的计算仅仅是对资料处理过程的回算.

1.4 气候场标准化变量的预报

首先对气候场标准化主分量预报的反算,

$$\hat{Y}_{k \times (m+s)} = \Lambda^{1/2} Y_{k \times (m+s)} \quad (11)$$

其中, $\Lambda^{1/2}$ 为对角阵. 然后对预报标准化变量场的回算,

$$\hat{X}_{p \times (m+s)} = V_{p \times r} Y_{r \times (m+s)} = V_{p \times r} (Y_{r \times m} Y_{r \times s}) = \hat{X}_{p \times m} \hat{X}_{p \times s} \quad (12)$$

$$\hat{X}_{p \times m} = (\hat{x}_{kj}) \quad k=1, 2, \dots, p; j=1, 2, \dots, m \quad (13)$$

$$\hat{X}_{p \times s} = \hat{x}_{k(m+i)} \quad k=1, 2, \dots, p; i=1, 2, \dots, s) \quad (14)$$

其中, $Y_{k \times m}$, $Y_{k \times s}$ 前面已给出.

1.5 对原始气候场的回算

$$X_{p(m+s)} = S_p \hat{X}_{p(m+s)} + E_p \quad (15)$$

其中, E_p 和 S_p 分别为原始气候场的均值和标准差序列.

1.6 预报结果的误差估计

第 k 个测站的残差估计为

$$D_k = \left[\frac{1}{m} \sum_{i=1}^m (\hat{x}_{ki} - x_{ki})^2 \right]^{\frac{1}{2}} \quad k=1, 2, \dots, p \quad (16)$$

第 k 个测站的相对误差与总平均相对误差分别为

$$E_{rk} = D_k / E_k \quad (17)$$

$$E_r = \sum_{k=1}^p E_{rk} / p \quad (18)$$

2 广东春雨的预报

取广东 47 个测站每年 2~4 月降水量, 其中 1954~1989 年为统计资料, 试报 1990.

2.1 预报序列阵

主分量 1~15 的累积方差 (%) 依次为 68.8, 76.9, 82.6, 86.5, 88.6, 90.1, 91.4, 92.5, 93.6, 94.5, 95.2, 95.9, 96.5, 97.0, 97.4, 它们共占总方差的 97.4%.

2.2 因子场的选择

取北太平洋海温场 5 个区各季节的前 6 个 PCA 分量, 其中 1 区太平洋西部 (180°E 以西, 10°S 以北), 2 区太平洋东北部 (180°E 以东, 15°N 以北), 3 区太平洋赤道区 (180°E 以东, $10^\circ\text{S} \sim 15^\circ\text{N}$), 4 区太平洋北部 (110°E 以东, 15°S 以北), 5 区太平洋热带地区 (110°E 以东, $10^\circ\text{S} \sim 30^\circ\text{N}$); 北半球 500 hPa 高度场 6 个区各季节 PCA 的前 6 个分量, 其中 1 区东北半球北部 ($0 \sim 180^\circ\text{E}$, $45 \sim 85^\circ\text{N}$), 2 区东北半球南部 ($0 \sim 180^\circ\text{E}$, $0 \sim 40^\circ\text{N}$), 3 区西北半球北部 ($180 \sim 360^\circ\text{E}$, $45 \sim 85^\circ\text{N}$), 4 区西北半球南部 ($180 \sim 360^\circ\text{E}$, $0 \sim 40^\circ\text{N}$), 5 区东北半球 ($0 \sim 180^\circ\text{E}$, $0 \sim 85^\circ\text{N}$), 6 区表示西北半球 ($180 \sim 360^\circ\text{E}$, $0 \sim 85^\circ\text{N}$); 以及环流指数等 312 个因子. 时间取 1953~1988 年, 1989 年资料作试报用.

2.3 相关计算

对降水场前 15 个标准化主分量, 用上述因子群逐个进行相关筛选计算, 取显著性水平 $\alpha=0.10$, 结果各个主分量都能找到 20~50 个可能预报因子, 供逐步回归进一步筛选. 单相关系数最小 0.275, 最大 0.613.

2.4 逐步回归计算

给定因子选择的 F -分布临界信度 $F_\alpha=0.95$, 对降水场的前 15 个主分量, 利用上述结果, 逐个进行回归计算, 结果建立了 15 个主分量的预报方程. 各方程的复相关系数如表 1, 通过 $\alpha=0.05$ 显著性检验. 由表 1 可见, 方程复相关系数最低 0.729, 最高 0.975, 且近一半都有 0.90 以上.

表 1 回归方程的复相关系数和因子数目

Tah 1 The multiple correlation coefficients and factors of regression equations

主分量	1	2	3	4	5	6	7	8
复相关系数	0.729	0.925	0.856	0.973	0.970	0.815	0.859	0.912
回归方程因子数	6	10	9	18	16	7	9	10
主分量	9	10	11	12	13	14	15	
复相关系数	0.848	0.970	0.788	0.750	0.914	0.899	0.975	
回归方程因子数	8	16	5	5	12	11	15	

第 1 主分量的回归方程为

$$y_1=-2.7068-0.0207x_1-0.0437x_2+0.0592x_3+0.1534x_4-0.1284x_5+4.8518x_6$$

(19)

其中， x_1 为时间线性函数， x_2 为春季高度场 1 区第 3 主分量， x_3 为秋季高度场 4 区第 2 主分量， x_4 为春季海温场 1 区第 4 主分量， x_5 为冬季海温场 1 区第 3 主分量， x_6 为 3 月份亚洲环流指数。

第 2 主分量方程为

$$y_2=-1.6566+0.0637x_1-0.0841x_2+0.1771x_3-0.0238x_4+0.1529x_5+0.1462x_6+0.1005x_7+3.6995x_8-2.7662x_9+1.9407x_{10}$$

(20)

其中， x_1 为秋季高度场 1 区第 6 主分量， x_2 为秋季高度场 3 区第 6 主分量， x_3 为秋季高度场 4 区第 4 主分量， x_4 为秋季高度场 6 区第 3 主分量， x_5 为冬季高度场 1 区第 5 主分量， x_6 为夏季海温场 1 区第 6 主分量， x_7 为夏季海温 2 区第 6 主分量， x_8 为欧亚环流指数， x_9 为 9 月亚洲环流指数， x_{10} 为 10 月亚洲环流指数。

由上可见，对于全省性多雨或少雨的第 1 主分量的降水分布主要受时间线性趋势项、东北半球北部和西北半球南部高度场、西太平洋海温场以及亚洲环流指数等因子影响为主；而对于广东省南北型（南旱北涝或南涝北旱）的第 2 主分量降水分布主要受东北半球北部与西半球（还包括北部、南部）高度场，西太平洋和东北太平洋海温场以及亚洲、欧亚环流指数等因子影响为主。

2.5 预报结果及其误差

选取广东省 19 个有代表性的测站作 1990 年 2~4 月的降水总量预报，其结果表明大部分台站预报误差小于 10%，平均绝对误差约 90.0 mm（表 2）。

表 2 1990 年广东 19 个站春雨的预报值和观测值

Tah 2 The predicted and observed results of spring-period precipitation of 19 weather stations in 1990

项目	南雄	韶关	连县	连平	梅州	怀集	清远	河源	肇庆	广州
预测值/mm	582.8	576.1	627.5	602.1	492.6	535.8	532.7	541.8	224.5	378.9
实际值/mm	627.8	547.6	653.4	614.7	616.7	560.7	647.4	678.0	294.6	403.0
误差/mm	-45.0	28.5	-25.9	-12.6	-123.5	-24.9	-114.7	-136.2	-70.1	-24.1
项目	惠州	深圳	汕尾	汕头	信宜	台山	阳江	湛江	徐闻	
预测值/mm	336.1	401.5	341.7	465.8	395.9	376.8	411.4	371.5	176.2	
实际值/mm	342.8	453.9	473.4	463.6	586.4	423.7	594.1	301.5	159.7	
误差/mm	-6.7	-52.4	-131.7	2.2	-190.5	-46.9	-182.7	70.0	16.5	

参考文献:

- [1] 徐瑞珍, 张先恭. 经验正交函数在两个气象场相关分析中的应用 [J]. 气象学报, 1982, 40 (1): 117 ~ 122.
- [2] 谢炯光, 秦冰冰, 王静渊. 奇异值分解法在季降水预报中的应用 [J]. 气象学报, 1997, 55 (1): 117 ~ 123.
- [3] 张尧庭, 方开泰. 多元统计分析引论 [M]. 北京: 科学出版社, 1983. 302 ~ 315.
- [4] [英] M 肯德尔. 多元分析 [M]. 中国科学院计算中心概率统计组译. 北京: 科学出版社, 1983. 76 ~ 89.
- [5] HARNACK R P, LANZANTE J R. Spacification of United States seasonal precipitation [J]. Mon Wea Rev, 1985, 113: 319 ~ 325.
- [6] BROCCOLI A J, HARNACK R P. Predictability of monthly North Pacific sea level pressure from monthly sea surface temperature for the period 1933 ~ 1976 [J]. Mon Wea Rev, 1981, 109: 2107 ~ 2117.
- [7] MICHAELS P J, CLUCCI S J. Modification of Mos-driven thunderstorm probabilities over complex terrain with continentalscale upper air data. Pripints 10th conf. on weather forecasting and analysis [J]. Amer Met Soc, 1984, 65 (1): 160 ~ 164.
- [8] 吴明进. 台湾春雨之长期预报 [J]. 大气科学, 1992, 20 (3): 199 ~ 216.
- [9] 宋丽莉, 张昌昭. 广东省水稻旱灾的评价及防旱抗旱对策 [J]. 热带地理, 1996, 16 (3): 191 ~ 195.
- [10] 黄嘉佑. 气象统计分析与预报方法 [M]. 北京: 气象出版社, 1990. 152 ~ 195.

Long-range Forecasting of the Spring-period Precipitation in Guangdong Area

ZHOU Wen^{*}, CHEN Chuang-mai, JIAN Yu-geng

Abstract: Long-term forecast of spring-period precipitation in Guangdong was conducted with stepwise regression analysis of principal component on climatic field. The relationship between the principal component of spring-period precipitation field and the principal component of various factor fields, namely, geopotential height field, sea surface temperature field and circulation index was revealed. It shows that the prediction equations of the first 15 principal components of spring-period precipitation field in Guangdong have passed the $\alpha = 0.05$ significance test. The multiple correlation coefficients range from 0.729 to 0.975. The mean absolute error of predictional spring-period precipitation in 1990 (19 representative stations in Guangdong) is about 90 mm.

Keywords: climatic field; principal component; stepwise regression; correlation analysis; spring-period precipitation of Guangdong

^{*} Department of Atmospheric Sciences, Zhongshan University, Guangzhou 510275, China.