

一种气候预测综合决策的方法 ——递归正权综合决策法*

项静恬

陈国珍

刘海波

赵振国

(中国科学院应用数学研究所, 北京 100080)

(国家气候中心, 北京 100081)

摘要 根据气候预测的特点, 提供了以误差平方和为风险函数, 以正权综合为模式的多途径气候预测决策方案。不仅从理论上论证了几种正权方法的优越性, 还提供了递归技巧, 进一步提高了正权综合的预测精度。对1997年汛期降水预测所做的综合决策表明, 该方法具有较好的综合决策能力。

关键词 气候预测 统计决策 递归正权综合

1 引言

由于气候变化规律及其影响因素的复杂性和多样性, 气候预测需要对多种方法的预测结果进行综合分析。因此, 选择合理的综合决策手段和方案, 对气候预测综合决策客观化和提高预测精度有重要作用。本文根据气候预测的特点, 提出一种递归正权模式的综合决策方案, 不仅从理论上保证了决策结果的优越性, 而且对1997年汛期降水预测所做的综合决策试验表明, 该方法具有较好的综合决策能力。

2 统计决策的思想和关键

统计决策是数理统计学科中一种重要的思想和工具, 其目的是基于所掌握的服从某类统计分布的观测值进行统计分析和判断, 从而对被考察的系统模型或系统输出采取某种决定。这个分析和决定的过程和结果即称为决策。

统计决策的关键, 在于构造风险函数和选择决策方案。前者是衡量决策优劣的尺度, 后者是保证决策成功的手段。然而风险函数是否合理, 决策方案能否成功, 都需具备可靠的理论依据和实际应用的有效验证。

3 气候预测统计决策模型的选择

3.1 权重综合型的气候预测统计决策

气候预测的统计决策, 目的在于将多种方法或多种途径的气候预测结果进行合理综

1997-12-12收到, 1998-10-08收到再改稿

* “九五”国家重中之重项目“96-908”资助

合, 构造出能提高预测效果的综合决策方案。

设 $\{x_t\}$, $t=1, 2, \dots, N$ 为某气候要素的观测结果 (依时间顺序排列), 今用多种途径或方法对未来时刻的该要素进行预测, 结果记为 $\hat{x}_{N+l}(j)$, $l=1, 2, \dots, L$ 为预测步长, $j=1, 2, \dots, J$ 为预测方法或途径, 则气候预测的综合决策结果 $\hat{x}_{N+l}$ 可用如下形式来表示:

$$\hat{x}_{N+l} = \sum_{j=1}^J w_j \hat{x}_{N+l}(j), \quad (1)$$

式中 w_j , $j=1, 2, \dots, J$ 为在综合决策方案中对第 j 个预测方法 (或途径) 所施加的权重。为保证决策方案的无偏性, 需满足约束条件

$$\sum_{j=1}^J w_j = 1. \quad (2)$$

选定式 (1) 和 (2) 的综合模式后, 气候预测统计决策的问题就归结为: 如何依据观测值 $\{x_t\}$ 及拟合值 $\{\hat{x}_t(j)\}$, $t=1, 2, \dots, N$, $j=1, 2, \dots, J$ 来确定权重 $w_1, \dots, w_J$, 使 (1) 式所获得的决策尽可能具有优效性。

3.2 风险函数和决策模型

用 (1) 式进行综合决策时, 气候预测的拟合误差可用下式表示:

$$e_t = x_t - \hat{x}_t = x_t - \sum_{j=1}^J w_j \hat{x}_t(j) = \sum_{j=1}^J w_j [x_t - \hat{x}_t(j)] = \sum_{j=1}^J w_j e_t(j). \quad (3)$$

在统计意义下, 误差平方和的大小是拟合优劣性的合理表征, 因此是理想的决策风险函数, 用 Q 表示风险函数, 则

$$Q = \sum_{t=1}^N e_t^2. \quad (4)$$

气候预测综合决策的最优模型应在约束条件 (2) 式下, 选取 w_j , $j=1, \dots, J$, 使 (4) 式达到最小, 即

$$\begin{cases} \min Q, & Q = \sum_{t=1}^N e_t^2, \\ \text{s.t.} & \sum_{j=1}^J w_j = 1. \end{cases} \quad (5)$$

3.3 最优综合决策的局限性

模型 (5) 式称为最优综合模型, 理论可证明该模型的最优解存在唯一, 且可用解析式表示^[1]:

$$Q_0 = \min Q = (R^T E^{-1} R)^{-1}. \quad (6)$$

相应的最优权重为

$$W_0 = (R^T E^{-1} R)^{-1} E^{-1} R, \quad (7)$$

式中 $W_0 = (w_1, \dots, w_J)^T$ (τ 表示向量转置), $R = (1, \dots, 1)$,

$$E = (e_{ij})_{J \times J}, \quad e_{ij} = \sum_{t=1}^N e_t(i) e_t(j). \quad (8)$$

由于 W_0 能保证综合模型的误差平方和达到最小, 因此该方案的最优性是显然的, 然而计算得到的最优权向量 W_0 有可能出现负值的分量, 这不符合气候预测综合决策的要求, 因此在 (5) 式的决策模型中, 需增加一个边界条件:

$$w_j \geq 0, \quad j = 1, 2, \dots, J. \quad (9)$$

满足 (5) 和 (9) 式的最优决策没有解析表示, 需借助运筹学中的单纯形法通过大量的迭代计算来逼近, 这对于实际应用的数据处理往往带来不便, 因此需要寻求既直观简便又合理有效的综合决策模型和方案。

4 正权综合决策的模型及优性

4.1 几种常用的正权综合决策模型

正权综合决策模型的权向量常有以下形式:

$$(1) \text{ 算术平均加权 } W_1 = \frac{1}{J}(1, 1, \dots, 1)^T = J^{-1}R. \quad (10)$$

$$(2) \text{ 均方倒数加权 } W_2 = \left(\sum_{j=1}^J \sqrt{e_{jj}^{-1}} \right)^{-1} (\sqrt{e_{11}^{-1}}, \sqrt{e_{22}^{-1}}, \dots, \sqrt{e_{JJ}^{-1}})^T. \quad (11)$$

$$(3) \text{ 方差倒数加权 } W_3 = \left(\sum_{j=1}^J e_{jj}^{-1} \right)^{-1} (e_{11}^{-1}, e_{22}^{-1}, \dots, e_{JJ}^{-1})^T. \quad (12)$$

$$(4) \text{ 简单顺序加权 } W_4 = \frac{2}{J(J+1)}(1, 2, \dots, J)^T. \quad (13)$$

$$(5) \text{ 二项系数加权 } W_5 = 2^{-2(J-1)}(C_{2J-1}^0, C_{2J-1}^1, \dots, C_{2J-1}^{J-1})^T. \quad (14)$$

方法 (4) 和 (5) 要求 $e_{11} \geq e_{22} \geq \dots \geq e_{JJ}$, 方法 (5) 中 C_{2J-1}^j 为组合数。

除以上几种外, 常用的还有离异系数法、三点法、层次分析法、德尔菲法等, 也能提供正权综合决策的方案^[1]。

4.2 正权综合决策的优性

可以证明^[1], 正权综合模型的误差平方和 Q 满足以下不等式:

$$\frac{\lambda_{\min}}{J} \leq Q_0 \leq Q \leq \sum w_j e_{jj} \leq \max_{1 \leq j \leq J} \{e_{jj}\}, \quad (15)$$

式中 $\lambda_{\min}$ 为对称正定阵 E [(8)式] 的最小特征根, 式中其他字母含义同前。

(15) 式表明正权综合决策模型有以下的共同优性:

- ① 正权综合决策方案的风险函数有共同上、下界;
- ② 正权综合决策方案可能是次优的 (因 $Q \geq Q_0$);
- ③ 正权综合决策优于参加综合的最差成员 ($Q \leq \max_i \{e_{ii}\}$).

4.3 常用正权综合决策方案的比较

由 (15) 式可对各种常用的正权综合方案算出相应的风险函数 Q 的上界 U , 并以此为依据比较各正权综合方案的优性, 具体推导如下。

①算术平均加权:

$$Q_1 = W_1^T E W_1 = J^{-2} R^T E R < \frac{1}{J} \sum_j e_{jj} = U_1. \quad (16)$$

②均方倒数加权:

$$Q_2 = W_2^T E W_2 \leq \sum_{j=1}^J w_j^{(2)} e_{jj} = \sum_{j=1}^J \sqrt{e_{jj}} / \sum_{j=1}^J \sqrt{e_{jj}^{-1}} = U_2. \quad (17)$$

③方差倒数加权:

$$Q_3 = W_3^T E W_3 \leq \sum_{j=1}^J w_j^{(3)} e_{jj} = J / \sum_{j=1}^J e_{jj}^{-1} = U_3. \quad (18)$$

④简单顺序加权:

$$Q_4 = W_4^T E W_4 \leq \sum_{j=1}^J w_j^{(4)} e_{jj} = \frac{2}{J(J+1)} \sum_{j=1}^J j e_{jj} = U_4. \quad (19)$$

⑤二项系数加权: (要求 $e_{11} \geq e_{22} \geq \dots \geq e_{JJ}$)

$$Q_5 = W_5^T E W_5 \leq \sum_{j=1}^J w_j^{(5)} e_{jj} = 2^{-2(J-1)} \sum_{j=1}^J C_{2J-1}^{j-1} e_{jj} = U_5. \quad (20)$$

可以证明^[1], (16) ~ (20) 式中的风险函数上界满足以下两个不等式:

$$U_3 \leq U_2 \leq U_1; \quad U_5 \leq U_4 \leq U_1. \quad (21)$$

用 (21) 式为依据, 可对上述五种常用的正权综合方案进行排序。方差倒数加权的决策方案是五种当中最为简便有效的。然而即使是五种当中排序最后的算术平均权重, 也有可能是最优的 (当且仅当 $\sum_{j=1}^J e_{jj}$ 为常数, $i = 1, 2, \dots, J$ ^[1])。由此可见, 上述正权综合决策方案都是直观可行且合理优效的。

5 正权综合决策方案的改进——递归正权综合决策

运用递归技巧, 能进一步降低正权综合决策的风险, 并明显提高模型的预测精度, 实现递归技巧的程序步骤如下:

①将 J 个初始方法 (或途径) 获得的 $\hat{x}_t(j)$, $j=1, 2, \dots, J$, $t=1, 2, \dots, N$ 记为初始方案, 计算这 J 种方案的误差平方和 $e_{jj} = \sum_{t=1}^N [x_t - \hat{x}_t(j)]^2$, $j=1, 2, \dots, J$ 。

②将 J 个方案的误差平方和重新按降序排列, 记为 $e_{(1)} \geq e_{(2)} \geq \dots \geq e_{(J)}$, 则

$$e_{(1)} = \max_{1 \leq j \leq J} \{e_{(j)}\}, \quad e_{(J)} = \min_{1 \leq j \leq J} \{e_{(j)}\}.$$

③给定迭代精度 ε , 若 $\Delta = e_{(1)} - e_{(J)} < \varepsilon$, 则终止迭代。计算方案 (J) 相应于初始方法 (或途径) 的权重, 此即所求的综合决策方案, 否则作下一步。

④计算以上 J 个方案的某种正权综合 [记为方案 $(*)$] 及其误差平方和 $e_{(*)}$, 记录方案 $(*)$ 关于初始方法 (或途径) 的权系数向量 $W_{(*)}$, 用 $e_{(*)}$ 取代 $e_{(1)}$, 方案 $(*)$ 取代方案 (1), 回到步骤②重新作下轮迭代。

迭代精度 ε 的大小可依据实际需要和计算能力来决定 (例如 0.001), 本递归对任何一种满足 (2) 式的正权都适用, 这是由于 (15) 式保证了每轮递归中正权误差平方和必有不等式 $0 \leq e_{(i+1)} \leq e_{(i)}$, 递归程序必定收敛, 且精度优于非递归情形。因此递归正权综合决策可以进一步降低决策风险。

6 对 1997 年汛期降水预测的递归正权综合决策试验

1997 年 4 月初在国家气候中心和水利信息中心联合召开的全国汛期旱涝趋势预测会商会上, 有气象、水文、科研院校等 8 大单位制作全国大范围降水趋势预测。8 个单位对当年汛期降水趋势预测的意见分歧较大, 出现了两种大致相反的预测结果。图 1、图 2 分别是第一种 (3 个单位) 和第二种 (5 个单位) 预报夏季降水出现正距平概率综合图。从图 1 中可看出, 第一种预报意见夏季 (6~8 月) 主要雨带在黄河至长江之间, 江南大部地区少雨; 第二种预报意见 (图 2) 夏季主要多雨区在长江以南, 黄河至长江之间少雨。由此, 我们对 8 个单位的预报结果进行了递归正权综合决策试验。根据 1994~1996 年各个单位的预报水平, 计算了各个单位 3 年来每年的预报误差方差, 并以误差平方和为预报的风险函数, 计算了综合后的误差方差, 递归后的综合误差方差等, 计算结果见表 1。从表 1 中可见, 综合后的误差方差接近 8 个单位中最小误差方差, 递归后误差方差较递归前的误差方差又有明显降低, 其值低于 8 个单位中最小误差方差。可见以方差倒数为权重的正权综合决策具有稳健性, 递归后预测精度有明显提

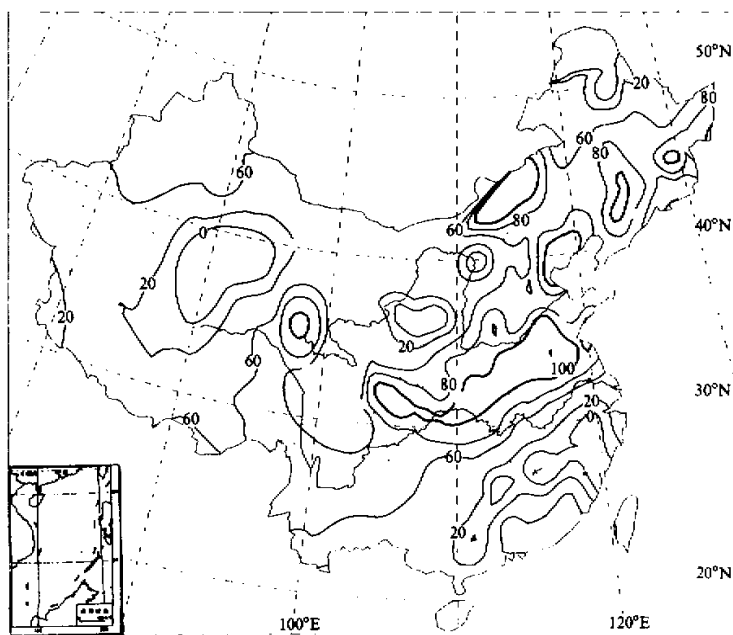


图 1 1997 年夏季 (6~8 月) 降水正距平概率 (第一种预报意见综合)

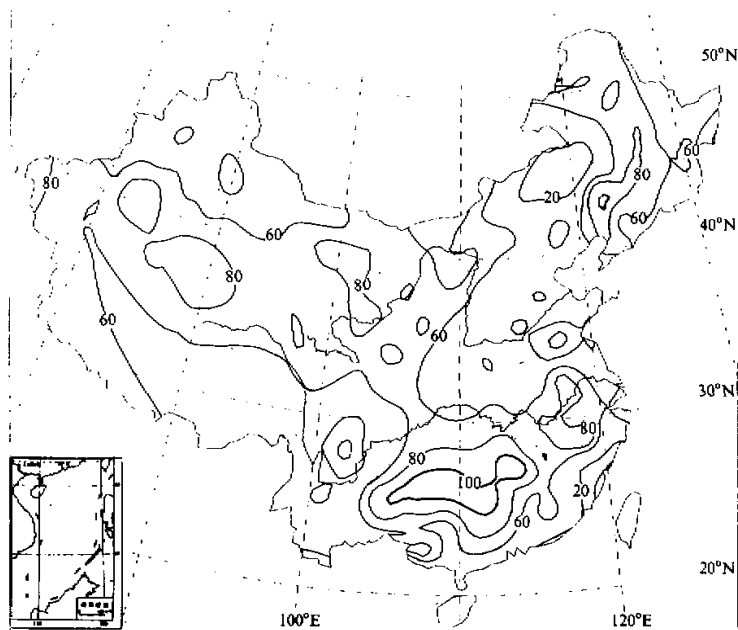


图 2 1997 年夏季 (6~8 月) 降水正距平概率 (第二种预报意见综合)

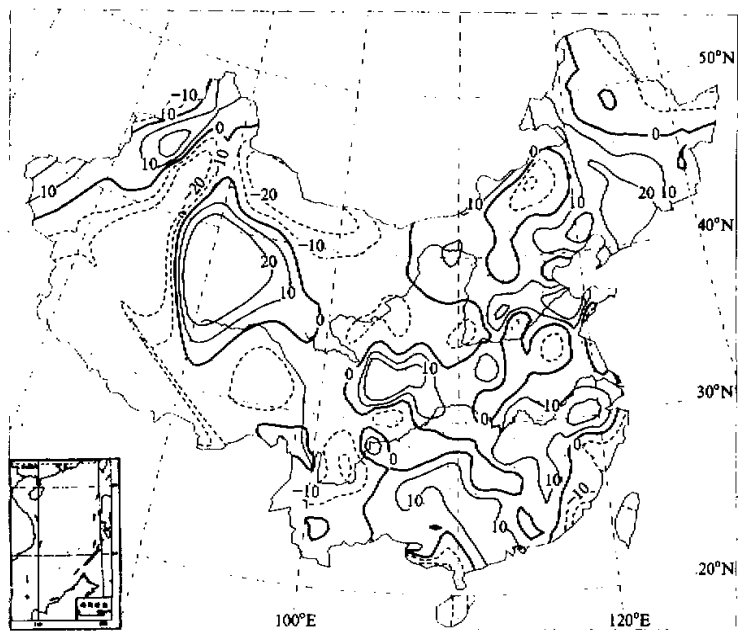


图 3 1997 年夏季 (6~8 月) 递归正权综合预报 (距平百分率)

高。根据 3 年来各个单位对全国 160 个站汛期降水预报权重（即 3 年平均权重）、对 1997 年 8 月汛期预测结果进行了综合决策试验。综合后的预报结果较好地反映了各个单位在近几年的预测表现和水平。从图 1、2、3 中可以看出，综合后，长江以南大部地区降水偏多，反映了第二种预报意见的主要信息，黄河至长江间既反映了第一种预报意见，也反映了第二种预报意见的某些信息，多雨范围较第一种预报意见小。东北及河套地区由于两种预报意见均报多，综合后仍报偏多。综合后的预报与实况（图 4）比较接近。

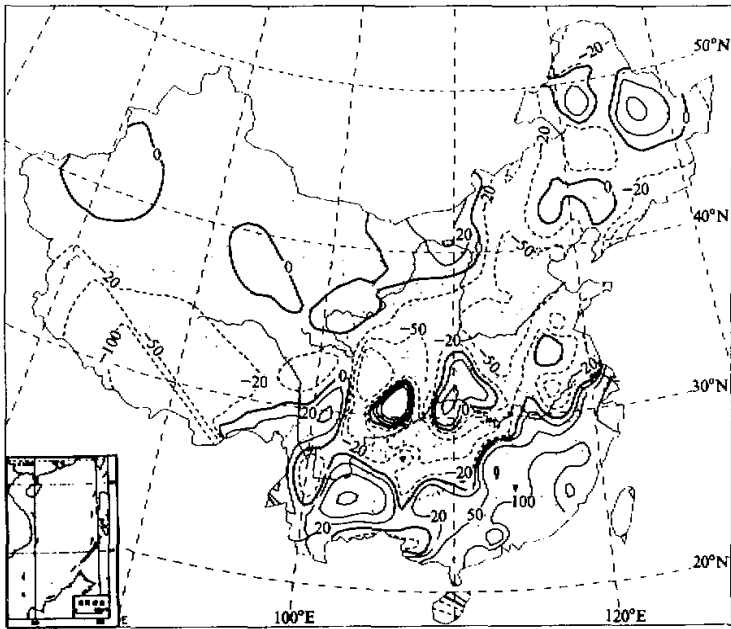


图 4 1997 年夏季（6~8 月）降水实况（距平百分率）

表 1 方差倒数加权的递归正权综合决策

年份	单位 方差权重	1	2	3	4	5	6	7	8	递归前 综合误差 方 案	递归 次数	递 归 综合误差 方 差
1994	误差方差	10.966	14.605	14.596	13.390	12.316	11.622	14.193	11.038	11.906	18	10.898
	递归综合权重	0.458	0.007	0.014	0.037	0.054	0.106	0.025	0.299			
	预报评分	0.86	0.65	0.64	0.69	0.73	0.79	0.70	0.81			
1995	误差方差	7.962	8.575	8.992	8.842	8.887	8.924	8.278	7.934	7.698	10	7.606
	递归综合权重	0.194	0.138	0.025	0.108	0.093	0.061	0.166	0.214			
	预报评分	0.84	0.72	0.71	0.77	0.81	0.70	0.72	0.76			
1996	误差方差	14.659	16.598	13.949	16.409	15.775	15.863	15.394	15.087	14.566	19	13.993
	递归综合权重	0.126	0.015	0.526	0.031	0.067	0.048	0.083	0.103			
	预报评分	0.67	0.62	0.81	0.66	0.64	0.71	0.63	0.70			
1994~1996 年 综合权重平均		0.260	0.053	0.188	0.059	0.071	0.072	0.091	0.205			

任何数理统计方法都不能从根本上解决短期气候预测问题, 预报准确率的提高关键还在于好的预报因子及改进预报方法本身。研究综合决策方法的目的不仅在于它较人工综合客观化、定量化, 而且要求通过综合决策后能提高预报准确率。试验表明, 本综合决策方法对预报精度的提高有一定作用, 对趋势预报的准确性提高并不十分明显, 反映了本综合决策方法的决策能力还需要进一步改进提高。

另外将 8 个单位 3 年预测的综合权重与国家气候中心预测室原评分方法^[2]所评出的分数相比较 (见表 1), 可以看出, 综合权重较评分更能使各个单位预报评定拉开档次, 并可运用预报综合权重客观地对各个单位预报进行综合。

参 考 文 献

- 1 项静恬, 史久恩, 1997, 非线性系统中数据处理的统计方法, 北京: 科学出版社。
- 2 赵振国, 1996, 我国汛期旱涝趋势预测进展, 气候预测研究, 北京: 气象出版社, 84~85。

The Synthetic Decision Method for Climate Forecasts — The Synthetic Decision Method of Recurrent Positive Weights

Xiang Jingtian

(*Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing 100080*)

Chen Guozhen, Liu Haibo and Zhao Zhenguo

(*National Climate Center, Beijing 100081*)

Abstract The synthetic decision plan of manifold climate forecasts is presented. The risk function of decision making is the sum of error squares, and its model is the synthesis of positive weights. This plan not only discusses the advantages of some positive weight methods in theory, but also provides the recurrent skill and enhances the forecasting precision of the synthesis of positive weights. This method has been proved to have quite good synthetic decision abilities through the synthetic decision making in the forecasts of the precipitation in the 1997's flood season.

Key words climate forecast statistical decision making synthesis of recurrent positive weights