

试用自然正交函数展开及最优化法做气象产量区域分布预报

张云秋 (沈阳区域气象中心研究所)

本文采用大气环流分型(W、C、E型)出现的日数;副高综合指数(面积+脊线位置+西伸脊点位置);极涡中心强度、平均位置;冬季东亚槽强度、平均位置;西太平洋副高强度指数;经向和纬向环流指数共122个环流因子作为优选普查因子。

辽宁省12个市水稻单产资料抄自省统计局农业处。社会产量的处理采用15年直线滑动平均法模拟历史趋势产量,社会产量与历史趋势产量之差为气象产量,并以此作为自然正交函数展开的要素场。预报年的趋势产量用调和权重法来求算。资料年代为1960—1985年。

一、方法简介

所谓经验正交展开,就是把要素场序列分解成正交的时间函数序列与正交的空间函数序列的乘积。通常对给定的要素场可用下列矩阵形式表示:

$$F = [F_{ij}] \quad i=1,2,\dots,m \\ j=1,2,\dots,n$$

其中, F_{ij} 表示要素在第 j 个测点的第 i 次观测值, n 为场中观测点的个数, m 为总共观测次数。

将要素场 F 按自然正交展开得

$$F = TB$$

式中, T 为 $(m \times n)$ 的时间权重系数矩阵, B 为 $(n \times n)$ 的特征向量矩阵。

首先求出 F 的协方差矩阵 $R = [R_{ij}]$

$$\text{式中, } R_{ij} = \frac{1}{2} \sum_{k=1}^m (F_{kj} - \bar{F}_j)(F_{ki} - \bar{F}_i) \quad (1)$$

(1) 式中

$$\bar{F}_i = \frac{1}{m} \sum_{k=1}^m F_{ki}, \quad \bar{F}_j = \frac{1}{m} \sum_{k=1}^m F_{kj}$$

$$i, j = 1, 2, \dots, n$$

其次用雅可比法求出 R 的特征值 λ 和特征向量 B , 特征向量求出后, 可求出时间权重系数矩阵:

$$T = [T_{ij}]$$

$$\text{式中, } T_{ij} = \sum_{k=1}^n F_{ik} B_{jk} \quad i=1,2,\dots,m \\ j=1,2,\dots,n$$

因用自然正交函数展开收敛比较快, 仅取对应特征值较大的前 H 个特征向量作为自然正交函数就能逼近要素场的趋势, 设 H 个较大特征值对应的时间权重系数序列为: $T_1 = (T_{11}, T_{21} \dots T_{m1})$, $T_2 = (T_{12}, T_{22} \dots T_{m2}) \dots$, $T_H = (T_{1H}, T_{2H} \dots, T_{mH})$, 把 T_1, T_2, \dots, T_H 看做因变量, 把候选因子作为自变量, 用非线性最优化方法计算以下 4 种类型的相关系数:

1. 线性函数类型: $y = x$
2. 幂函数类型: $y = x^a$
3. 指数函数类型: $y = e^{ax}$;
4. 对数函数类型: $y = L_{nx}$ 。

式中, y 为因变量, x 为自变量。

然后比较其 $|R|$ 的大小来确定因子类型, 因子类型确定后, 还要通过 $|R|$ 的显著性检验、样品稳定性检验和参数 a 的稳定性检验, 这样筛选出的因子为最优预报因子。再用逐步回归法建立各预报量的非线性预报方程。

表达式为:

$$\hat{T}_p = b_{0p} + \sum_{i \in A_p} b_{ip} x_i \quad p=1, 2, \dots, H \quad (2)$$

式中, b_{0p} , b_{ip} 为回归系数, x_i 为非线性预报因子, A_p 为引入方程因子序号的集合。

最后利用公式

$$\hat{F}_{m+t, j} = \sum_{p=1}^H \hat{T}_{m+t, p} \cdot B_{pi} \quad j=1, 2, \dots, n, t=1, 2 \quad (3)$$

计算出 n 个点的预报值。

二、计算步骤

根据自然正交函数展开的基本原理以及选最优预报因子的办法, 以本省水稻产量的区域分布预报为例, 取沈阳、大连、鞍山、

抚顺、本溪、丹东、锦州、营口、阜新、辽阳、铁岭、朝阳12个测点, 资料为1960—1985年, 试报1986、1987年。社会产量资料经处理后, 得到气象产量场:

$$F = [F_{ij}] \quad i=1, 2, \dots, 26 \\ j=1, 2, \dots, 12$$

具体数据略。其计算步骤和结果如下:

1. 计算气象产量场 F 的协方差矩阵。利用公式 (1) 得到:

$$R = \begin{pmatrix} 13047.2 & 4858.2 & \dots & 1589.0 \\ 4858.2 & 7753.5 & \dots & 3091.6 \\ \dots & \dots & \dots & \dots \\ 1589.0 & 3091.6 & \dots & 4679.9 \end{pmatrix}$$

2. 用雅可比法求出 R 的特征值 λ 和特征向量 B 。具体值见表1 (λ 值从大到小排列)。

表1 R 矩阵的特征值和特征向量

λ	1	2	3	4	5	6	7	8	9	10	11	12
B_i	79981.6	19533.7	11056.5	2847.9	2767.8	1854.7	1553.7	1060.7	737.6	410.5	284.8	194.5
1	0.4	-0.1	0.4	0.0	0.0	0.0	-0.6	0.1	0.4	0.3	0.2	-0.2
2	0.2	0.2	-0.4	0.4	0.1	-0.2	0.3	0.0	0.3	0.4	-0.3	-0.4
3	0.5	0.7	0.1	0.0	0.3	0.0	0.2	0.0	-0.1	-0.1	-0.1	0.0
4	-0.5	0.7	-0.6	-0.4	-0.1	-0.2	-0.2	-0.4	0.0	-0.4	-0.2	-0.4
5	0.3	0.2	-0.3	-0.5	0.1	0.0	0.0	0.0	-0.2	0.5	-0.1	0.4
6	0.3	0.2	-0.4	0.2	0.0	-0.2	0.0	0.4	-0.1	-0.3	0.6	0.1
7	0.2	0.1	0.4	-0.1	0.0	0.2	0.6	-0.3	-0.1	0.1	0.5	-0.3
8	0.2	0.0	-0.1	0.4	-0.4	0.3	-0.1	-0.6	0.0	-0.1	0.0	0.4
9	0.5	0.2	0.2	0.3	0.2	-0.5	-0.1	-0.1	-0.5	0.0	-0.1	0.2
10	-0.6	0.4	0.1	-0.1	-0.5	0.4	0.0	0.5	-0.3	0.0	-0.3	-0.2
11	-0.7	0.6	0.4	-0.1	0.0	0.0	0.3	0.2	0.6	-0.3	-0.2	0.4
12	0.1	0.3	0.0	0.2	0.7	0.6	0.2	0.0	-0.1	-0.2	0.0	-0.1

3. 计算气象产量场的时间权重系数矩阵。其式为:

$$T = [T_{ij}] \quad i=1, 2, \dots, 26 \\ j=1, 2, \dots, 12$$

具体数字略。根据这种展开收敛速度较快的特

点, 其前三个特征向量之和占全部信息的92.3%, 所以只取前三个特征向量, 即 $H=3$, 相应的时间权重系数如表2。

4. 筛选最优预报因子。以时间权重系数 T_1 、 T_2 、 T_3 为预报量, 与122个大气环流因子进行优选普查, 确定因子类型, 选出通

表2 前三个特征向量的时间权重系数值

T \ H	i												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1	-377.2	-207.3	-228.8	146.6	192.6	140.4	-33.6	289.2	289.2	-503.7	9.7	75.0	-549.8
2	48.5	-62.1	-71.8	-66.8	28.8	35.7	2.3	87.8	43.1	-85.1	34.0	102.2	-239.7
3	112.5	-2.8	24.2	-50.1	-56.8	-117.9	-23.8	29.9	11.0	-17.3	53.7	206.5	-23.7

T \ H	i												
	14	15	16	17	18	19	20	21	22	23	24	25	26
1	148.7	210.0	339.4	-319.0	-42.2	-398.8	-37.1	71.4	-35.6	-17.5	267.1	294.2	-678.0
2	-19.7	45.6	153.6	-58.8	-56.4	595.8	-7.5	7.0	-76.3	-29.2	18.6	-54.3	126.6
3	35.2	9.6	-8.0	-99.7	-100.7	-110.3	-125.7	-56.5	-1.7	58.3	254.3	149.7	264.9

过显著性检验和样品及a 值稳定性检验的因子作为最优预报因子。现只把各预报量选出的最优预报因子类型作以统计列于表3，其具体因子和a 值、相关系数从略。

表3 最优预报因子类型统计表

数量 \ 预报量	类型				合计
	1	2	3	4	
T ₁	0	10	8	1	19
T ₂	0	12	7	0	19
T ₃	0	6	4	1	11

从表3 中可以看出，最优预报因子是没有线性类型的，大部分是2，3 两种类型。

5. 建立非线性预报方程。有了T₁、T₂、T₃ 三个预报量，再从各自选出的最优预报因子中，用逐步回归方法选出4—6 个预报因子建立非线性预报方程，其结果如下：

$$T_1 = 187.2 + 255.2[(x_5 - 12.0) \div (-4.0)]^{-4.3} - 582.8e^{-2.9[(x_6 - 34.0)/139.0]} - 332.2e^{-2.3[(x_{10} - 30.0)/36.0]} - 458.7(x_{13}/21.0)^{6.0} + 264.2[(x_{14} - 124.0) \div 131.0]^{4.2} - 439.0[(x_{19} - 529.0)/13.0]^{4.9} \quad (4)$$

式中，x₅ 为前一年7 月极涡中心位置；x₆ 为

前一年1 月纬向环流指数；x₁₀ 为前一年9 月经向环流指数；x₁₃ 为当年6 月E 型环流日数；x₁₄ 为当年4 月东亚槽平均位置；x₁₉ 为当年6 月极涡中心强度。

$$T_2 = 55.4 + 106.5e^{-6.5[(x_3 - 3.0)/22.0]} + 103.6e^{-4.7(x_4/22.0)} + 420.9e^{-3.0[(x_8 - 486.0)/33.0]} - 108.0[(x_{14} - 31.0)/130.0]^{0.1} - 66.3[(x_{18} - 12.0)/(-4.0)]^{1.8} + 25.6[(x_{19} - 430.0)/26.0]^{0.2} \quad (5)$$

式中，x₃ 为前一年1 月C 型环流日数；x₄ 为前一年2 月C 型环流日数；x₈ 为前一年4 月极涡中心强度；x₁₄ 为前一年5 月经向环流指数；x₁₈ 为当年7 月极涡中心位置；x₁₉ 为当年4 月经向环流指数。

$$T_3 = -37.7 - 22.7 \ln(x_1/18.0) + 0.1e^{8.0(x_2/40.0)} - 99.2[(x_4 + 19.0)/30.0]^{1.4} + 94.1e^{-2.9[(x_9 - 50.0)/139.0]} \quad (6)$$

式中，x₁ 为前一年11 月C 环流型日数；x₂ 为前一年4 月西太平洋副高强度指数；x₄ 为前一年2 月极涡中心位置；x₉ 为当年6 月极涡中心位置。

为预报m + t 年的气象产量，先将预报

年的有关因子资料代入(4)、(5)、(6)式, 求出 $\hat{T}_{m+t, p}$, 再将该值代入(3)式, 便可得到展开场内各点的气象产量预报值。

三、方程效果检验及试报

1. 方程检验。将进入方程的有关因子的历史资料代入(4)、(5)、(6)式, 再代入(3)式, 求出各点气象产量的历史拟合值, 以产量误差小于10%为衡量标准, 12个点平均历史拟合率为84.7%, 方程的其他检验结果见表4。

表4 方程效果检验值

项目 \hat{T}_p	复相关系数R	检验值 F	标准差 S	平均绝对误差
1	0.94	26.37	48.43	46.81
2	0.97	48.62	40.41	26.18
3	0.92	21.36	47.61	31.27
平均	0.94	32.12	45.48	34.75

从表4中看出, 各方程的复相关系数R ($R_{0.01, 6} = 0.745$) 和F值 ($F_{0.01, 1} = 7.67$) 均通过显著性检验。

2. 试报。将1986、1987年的有关因子资料代入(4)、(5)、(6)式, 求出 \hat{T} 值, 见表5。

表5 预报年T值

\hat{T} 年代 \ P	1	2	3
1986	6.49	-48.34	-44.65
1987	27.53	-0.49	0.84

再将 \hat{T} 值代入(3)式, 得到辽宁省12个市的水稻气象产量预报值, 列于表6。

表6中的气象产量值与相应的趋势产量之和为1986、1987年的水稻单产预报值, 见表7。

从表7中可以看出, 1986、1987两年试报误差百分率除本溪、丹东、阜新偏高外, 其余

表6 试报年的气象产量值

地点 年代	沈阳	大连	鞍山	抚顺	本溪	丹东	锦州	营口	阜新	辽阳	铁岭	朝阳
1986	-10.43	9.49	-35.06	-63.87	5.67	10.14	21.39	5.76	-30.70	-27.70	-51.41	-13.85
1987	11.40	5.07	13.51	-13.60	7.91	7.82	5.79	5.42	5.83	-16.63	-19.23	2.61

表7 预报产量与实际产量比较

单位(公斤/亩)

地点 年代 \ 产量	沈阳	大连	鞍山	抚顺	本溪	丹东	锦州	营口	阜新	辽阳	铁岭	朝阳	
1986	实产	429	356	405	280	173	384	392	567	242	403	349	198
	预产	424.1	360.4	434.2	337.3	324.1	413.5	413.8	545.1	189.1	431.5	377.5	219.8
1987	实产	454	308	429	365	286	285	386	556	290	393	365	240
	预产	456.1	363.9	462.4	366.9	327.8	419.2	424.1	566.4	219.9	444.5	408.9	231.9
平均误差%	0.8	9.7	7.5	10.0	50.9	27.3	7.7	2.9	23.1	10.1	10.0	7.2	

各点均在10%左右,基本符合目前产量预报规定的误差标准,效果良好。关于本溪、阜新、丹东试报平均误差偏高的问题,正在深入研究。

四、方法讨论

利用自然正交函数展开产量场,由于场内各点可任意选取,便于应用于各行政区域的产量预报(因现有产量资料都以行政区划为准)。这种展开收敛速度较快,很容易将大量资料信息浓缩集中,利用较少的几个主分量就能反映出产量场的主要特征,从而减少了不必要的计算步骤。利用最优化方法筛选预报因子,避免漏掉一些与预报量非线性

(上接28页)

微机及遥感技术在农业气象中应用研究。

3. 扬长避短,加强农气科研队伍建设

与国内其他省区相比,我省得天独厚的长处是有一支由160多人组成的农业气象科技队伍。据不完全统计,我省气象部门有农气专业大专以上学历的人员88人,获中级职务的49人,这是开展农气科研工作的雄厚人力基础。在机构方面,除沈阳区域气象中心研究所外,还有4个市级气科所和1个农业气象试验站。我们一定要利用我省农气科技人员多、台站多、资料系统性好、微机利用率较高的优势,围绕气象业务和农业生产中的关键气象问题,做到省、市、县密切配合,通力协作,积极开展农业气象科学研究和成果推广应用工作。要加强农气科技队伍建设,通过请进来,派出去等办法培训农气科技人员,提高农气科技人员的技术素质,合理安排科研机构各类科技人员比例,不断为出成果、出人才创造条件。要加强横向联系,吸收其他学科的新方法、新经验,发展农业气象科学。注意学习国外的先进技术与经验,争取同一些国家及地区的同行建立联系,通过引进消化吸收国外的先进经验,促进我省农业气象科学研究的发展。

相关好的因子,从而建立的时间权重系数的非线性预报方程提高了展开精度和预报的准确率。由于自然正交展开的空间函数是不变的,所以在历史回代及预报过程中对特殊年份的预报准确率较低,有待进一步研究改进。

参 考 文 献

- 【1】于系民等,用车贝雪夫多项式预报气象产量地理分布的试验,气象学报,第45卷,第1期,1987年。
- 【2】冯耀煌,优化的非线性预报方法及其应用,现代若干动力统计分析预报方法研讨会材料,1988年。
- 【3】吉书琴等,辽宁省粮食作物产量气象预报方法研究论文集,1987年。
- 【4】屠其璞等,气象应用概率统计学,气象出版社,1984年。

本溪地区暴雨档案及检索

系统通过鉴定

本刊讯 由本溪市气象局研制的本溪地区暴雨档案及检索系统经省市有关专家鉴定,认为在同类研究中,达到了国内先进水平。

该系统采用新颖的dBASE III数据库管理语言及BASICA高级语言编制,对本溪地区历史上出现的暴雨过程的天气形势进行数字化处理,实现了检索最佳相似暴雨天气过程的目的,为制作暴雨预报提供了主要参考依据,比人工翻阅图表找相似提高时效百余倍。(魏军)

