

# 甘肃省玉米产量的主成分分析\*

韩永翔 葛秉钧

(兰州干旱气象研究所, 兰州 730020)

## 提 要

应用自然正交函数方法将甘肃省玉米产量的趋势产量项( $Y_t$ )和气候产量项( $Y_w$ )进行展开, 展开后的  $Y_t$ 、 $Y_w$  第一特征向量都表明了  $Y_t$ 、 $Y_w$  总趋势的一致性, 其大值中心表明产量变率较大. 通过对各主成分主要时间系数与各站  $Y_t$ 、 $Y_w$  进行相关分布分析, 筛选出全省玉米产量的八个站点, 这些站点基本上代表了本省不同类型的农业气候区域的玉米产量水平和生育状况.

**关键词:** 玉米产量; 甘肃省; 主成分分析.

## 1 引 言

玉米是我省第二位粮食作物, 省内大部分地方均有种植. 由于我省地形复杂、气候多样, 各地玉米产量的变化也不相同. 在进行玉米产量预报模式的研究过程中, 我们曾用回归分析等方法挑选出许多站点来拟合产量. 但发现, 随着时间的变化, 选入模型的站点也相应发生变化. 再加上人为因素, 所选站点不一定具有较好的代表性. 为了克服这些缺陷, 同时也为了从总体上了解全省玉米产量的变化趋势, 我们搜集了全省 55 个县(市、区)22 年的玉米产量资料, 用自然正交函数对产量场进行展开, 在挑选出各主成分代表站点的同时, 还得到了一些满意的结果.

## 2 资料的处理和方法

### 2.1 资料来源

玉米产量资料取自于甘肃省统计局农村处. 由于不少县(市、区)是从 1969 年才开始有完整的玉米产量资料, 所以, 全省 55 个县(市、区)资料的起止年代均为 1969~1990 年计 22 年.

\* 本课题属甘肃省气象局项目, 由甘肃省农委资助.  
1993-07-12 收到, 1993-10-27 收到修改稿.

## 2.2 玉米产量资料的处理

由于玉米产量的形成是由品种、土壤肥力、气候因素以及各种农业技术措施综合影响的结果,因而需对产量资料进行分解,以求出其趋势产量项( $Y_t$ )和气候产量项( $Y_w$ )。我们首先采用直线滑动平均方法,来消除超短波波动的影响。某阶段线性趋势的直线滑动方程为:

$$Y_i(t) = a_i + b_i t$$

其中  $i = 1, 2, \dots, n - k + 1$ ;  $t = i, i + 1, \dots, i + k - 1$ 。  $n$  为样本数,  $k$  为滑动步长, 取其值为 5。 求出各点的  $\bar{Y}_i(t)$ , 组成一组新的产量序列  $Y(t)$ 。 运用调和权重法以不同权重去估计序列样本对趋势产量的影响。 趋势产量的增长量  $W_{t+1} = Y_{t+1} - Y_t$ , 则其平均增长量  $\bar{W} = \sum_{i=1}^{n-1} C_{i+1}^n \cdot W_{i+1}$ 。 式中  $C_{i+1}^n$  为调和权重系数, 它满足  $C_{i+1}^n > 0$ ,  $\sum_{i=1}^{n-1} C_{i+1}^n = 1$  两个条件<sup>[1]</sup>。 计算出趋势产量的值  $Y_{t+1} = Y_t + \bar{W}$ 。 用  $Y = Y_t + Y_w + \Delta Y$  求出气候产量项  $Y_w$ 。 随机误差项  $\Delta Y$  可忽略不计。

## 2.3 用自然正交函数展开 $Y_t$ 和 $Y_w$ 项

$Y_t$  和  $Y_w$  分别用自然正交函数展开<sup>[2]</sup>, 设产量资料矩阵为:

$$mX_n \begin{Bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{Bmatrix}$$

式中  $m$  代表年代, 为 22,  $n$  是参加展开的站点数目, 为 55。 将  $mX_n$  分解为:

$$mX_n = mT_n \cdot A'$$

式中  $mT_n$  是分解后的时间系数,  $A'$  是分解后的特征向量。

由于自然正交函数收敛快, 一般仅前几项就可以很好地反映参加展开产量场的主要信息, 若取项太多, 反而会掩盖对主要问题的分析。 所以, 在本文分析中, 仅取前两项<sup>[3]</sup>。

## 3 甘肃省玉米产量场的空间分布特征

由于我省各地玉米产量水平差异很大, 为了便于从空间场上分析玉米产量的主要特征, 我们对  $Y_t$  和  $Y_w$  分别加以讨论。

### 3.1 玉米趋势产量 $Y_t$ 的空间分布

图 1 是分解后玉米产量  $Y_t$  的第一特征向量, 方差贡献率为 79.9%。 其特点是全省均以正值出现, 大值中心主要分布在以下四个区域: (1) 临洮、会宁以西的广大地区; (2) 甘谷、武山一带; (3) 灰县、成县盆地; (4) 陇东南部的平凉、宁县一带。 最大值在张掖、临泽、高台等地。  $Y_t$  第一特征向量全省均以正值出现的事实, 反映出省内玉米趋势产量的变化具有较好的一致性。 大值中心突出了上述地方是趋势产量变率大的区域, 这与生产实际是

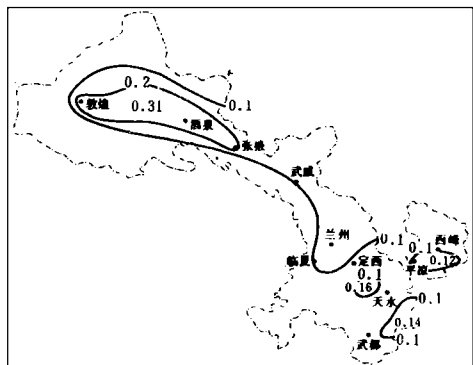


图 1  $Y_1$  第一特征向量

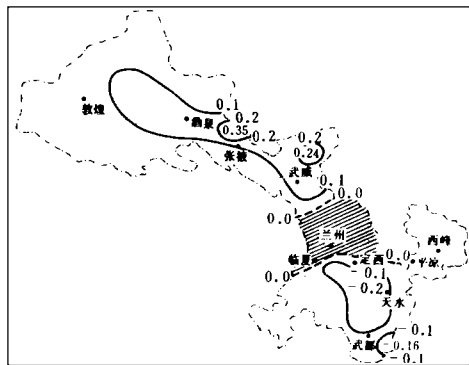


图 2  $Y_2$  第二特征向量(斜线为过渡带)

相吻合的. 分析其原因, 仍在于上述四个区域, 尤其是临洮、会宁以西的广大地区多为灌溉条件较好的水浇地. 加之品种的更换, 农业技术的改进(如地膜玉米面积的迅速扩大, 弥补了热量条件的不足, 带田玉米种植和规范化栽培技术的推广应用, 提高了玉米对光、热、水资源的利用率) 和农业投入逐年增加(化肥用量的增加、病虫害的及时防治) 等, 致使玉米产量呈较快的增长趋势.

图 2 是玉米趋势产量  $Y_t$  的第二特征向量, 方差贡献率 10.1%. 其分布特点是: 河西五地(市) 为正值区, 中心位于临泽、高台附近, 陇东、陇南、天水及中部地区的东部均为负值区. 中心有两个: 一个在广河、临洮, 另一个在西和、礼县到甘谷、武山一带, 而在白银经兰州到临夏一线为一条东北—西南向的有正、负值分布的过渡区. 这种分布格局, 表明了我省玉米趋势产量( $Y_t$ ) 的第二特征向量在过渡区的东西两侧具有反相关的变化趋势.

### 3.2 玉米气候产量 $Y_w$ 的空间分布

图 3 是  $Y_w$  的第一主分量场, 图中有两个正的高值中心, 一个在酒泉、张掖、民勤一线, 一个在华家岭以东的广大地区. 其大值区域主要集中在通渭、陇西县及天水、平凉、庆阳三地区以及临夏东部、陇南西北部一带. 这一分布态势不仅仅反映了上述地区气候产量变率较大的特点, 同时还反映出全省玉米气候产量变化的总趋向较为一致. 其余主分量场从略.

气候产量( $Y_w$ ) 的收敛速度是比较慢的, 第一主成分方差贡献仅占总方差的 15.9%. 前 12 个主成分累积方差贡献值才超过 80%(表 1). 相比之下, 趋势产量( $Y_t$ ) 的收敛速度就要快得多, 其第一个主成分就占总方差的 79.9%. 这说明  $Y_t$  的一致性比较好, 收敛快, 而  $Y_w$  因受气候条件的影响, 局地性强, 收敛速度则较慢.

上述分析结果为玉米产量预报的制作提供了一个思路: 在具体工作中首先考虑趋势

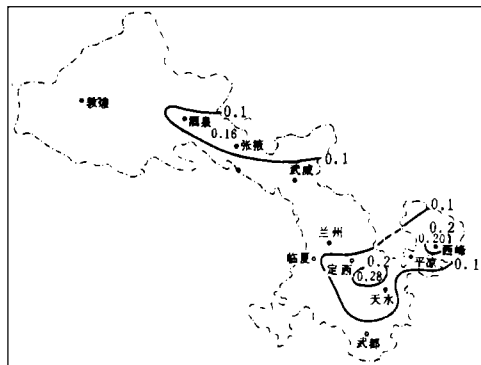


图 3  $Y_w$  第一特征向量

产量的总体变化趋势的一致性,并注意到东部与西部  $Y_1$  变化具有相反的趋势的事实;这样基本上就把握了玉米产量的大体丰、欠趋势;最后通过对  $Y_w$  的详细分析和预测,就有可能较准确的做出玉米产量预报。

表 1  $Y_1$  各主成分占总方差百分比(%)

主成分	1	2	3	4	5	6	7	8	9	10	11	12
个别方差	15.9	10.0	8.3	7.0	6.4	5.5	5.4	5.0	4.9	4.2	3.9	3.6
累积方差	15.9	25.9	34.2	41.2	47.6	53.1	58.5	63.5	68.4	72.6	76.5	80.1

#### 4 玉米载荷量的空间分布

为了获得玉米产量的代表站点,我们分别用  $Y_1$  和  $Y_w$  展开后的主要时间系数与各站点的  $Y_1$  和  $Y_w$  进行相关分析,求出其相关系数,称为载荷量,以反映各站  $Y_1$  和  $Y_w$  在该典型场中所起作用的大小。

图 4 是  $Y_1$  第一主分量的载荷量,正值中心主要分布在河西、兰州到临夏一线、陇东、灰成盆地;负值中心在舟曲至礼县一带。这种分布与我省热量状况由于受地形地势的影响,使其纬向分布不明显,而形成随地形分割成条块相间的不规则热量分布区是一致的<sup>[4]</sup>。我们预选出高台、临夏、兰州、西峰、平凉、灰县、舟曲七个站点,它们基本上代表了河西、中部、陇东、陇南的几个主要农业气候类型区,其载荷量均在 0.94 以上。 $Y_1$  的第二主成分载荷量大值中心与图 4 几乎一样,但载荷量相差较大。河西、兰州、临夏等大部分地区均在 0.90 以上,平凉、庆阳的大值仅 0.72、灰县 0.76、舟曲是 -0.95,但信度均达到 0.001,从而进一步证明了上述七个站点具有较好的代表性。

$Y_w$  的第一主成分载荷量分布如图 5 所示,大值中心主要分布在平凉、庆阳、定西三地区 and 天水市北部、陇南地区西北部及酒泉经张掖至民勤一线。在这些大值区域中选择平凉市、西峰市、礼县、陇西、高台为代表站,其信度都通过了 0.001 的检验。

综合以上分析,玉米产量预报的初选站点是高台、临夏、兰州、西峰、平凉、灰县、礼县

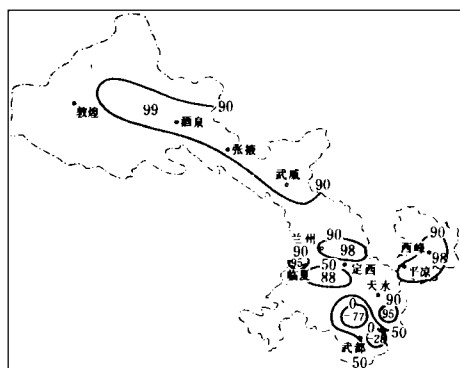


图 4  $Y_1$  第一时间系数与各站  $Y_1$  相关系数分布图(%)

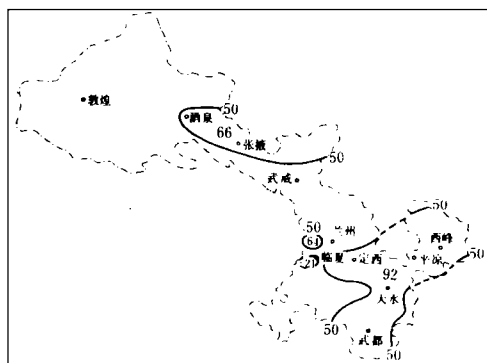


图 5  $Y_w$  第一时间系数与各站  $Y_w$  相关系数分布图(%)

和陇西八个站点.其中高台代表河西灌溉农业区,西峰、平凉代表陇东雨养农业区,兰州、陇西代表中部干旱区,临夏代表二阴山区,灰县、礼县代表北亚热带向暖温带的过渡区.

## 5 结 语

应用自然正交函数方法将甘肃省玉米产量的趋势产量 $Y_t$ 展开,展开后的 $Y_t$ 第一特征向量表明了全省趋势产量具有较好的一致性,大值中心反映了其变率较大的事实.第二特征向量表明趋势产量在其总趋势一致的前提下,以白银经兰州到临夏一线过渡区为界,东部与西部的趋势产量的变化呈反相关趋向.

用自然正交函数将玉米气候产量 $Y_w$ 展开后的第一特征向量表明全省 $Y_w$ 总趋势的一致性,并反映出中部、陇东、陇南西北部以及河西中部是气候产量变率较大的地区.由于受气候条件的影响,局地性很强, $Y_w$ 的收敛速度较慢.

通过对玉米产量载荷量的分布分析,初选八个站点基本上代表了本省不同类型的农业气候区域的玉米产量水平和生育状况,而且代表性较好.

## 参 考 文 献

- 1 王觀棠等.农业气象预报概论.北京:农业出版社,1991.443~446.
- 2 张尧庭,方开泰.多元统计分析引论.北京:科学出版社,1982.322.
- 3 徐国昌,朱炳璠,王先芬.表征青藏高原地面冷热状况的一种方法.夏半年青藏高原对我国天气的影响.北京:科学出版社,1987.
- 4 缪培俊,余优森,葛秉钧,黄关寿.从甘肃的农业气候特点谈农业布局问题.甘肃农业科技,1981,第一期.

## THE PRINCIPAL COMPONENT ANALYSIS FOR CORN YIELD IN GANSU PROVINCE

Han Yongxiang Ge Bingjiun

(Lanzhou Arid-Meteorological Institute, Lanzhou 730020)

### Abstract

The tendency corn yield term ( $Y_t$ ) and climatic corn yield term ( $Y_w$ ) of Gansu province are expanded by means of EOF. Their first characteristic vectors show that general tendency is consistent with the tendency yields of Gansu. The centers of large values is the large variabilities. By means of correlation analysis between  $Y_t$  time coefficients and  $Y_t$  for each station, and the same method to  $Y_w$ , eight stations for predicting corn yield are chosen, which represent basically yield levels and growing conditions under various kinds of agroclimatic regions of Gansu province.

**Key words:** Corn yield; Gansu province; Principal component analysis.