

# 基于主成分分析法的我国棉花生产 向西北扩张分析\*

朱再清, 王燕玲

(华中农业大学 经济管理学院, 湖北 武汉 430070)

**摘要** 为了探讨我国棉花生产是否具有向西北扩张转移的可能性, 在分析甘肃、陕西、山西 3 省棉花生产的气候及灌溉条件基础上, 运用主成分分析方法对我国 14 个棉花生产省(区)棉花生产的技术经济条件进行了综合评价。结果表明: 甘肃、陕西、山西分别属于最适宜、较适宜、一般的棉花生产区。得出结论: 甘肃省可以发展成为我国西北地区继新疆之后的又一个棉花优势生产区; 陕西省可以进一步稳定为我国棉花主产区。我国棉花主产区存在进一步向西北地区扩张转移的可能性。

**关键词** 棉花主产区; 区域布局; 优势区; 适宜区; 西北地区; 气候条件; 灌溉条件

**中图分类号:** F 323.3   **文献标识码:** A   **文章编号:** 1008-3456(2012)03-0018-05

棉花历来是我国一级战略性物资, 其生产对国计民生、纺织工业以至整个国民经济发展都具有举足轻重的作用, 因此, 我国一直都很重视棉花生产。目前, 我国棉花生产主要分布在黄河流域、长江流域和西北内陆区域。黄河流域棉区生产规模指数波动幅度较大而频繁, 且用地矛盾突出; 长江流域棉区生产规模指数不断下降; 西北内陆棉区生产规模指数不断上升, 且主产省份新疆进一步扩张的空间已十分有限<sup>[1]</sup>。这一现实迫使人们不得不思考我国棉花生产是否可以循着世界棉花发展的一般规律, 进一步向光热条件好但相对干旱<sup>[2-3]</sup>的西北、北部, 尤其是其中的甘肃、陕西、山西 3 省扩张和转移。

国内学者对此作了一些相关性的研究。钟甫宁等、朱启荣等分别从比较利益变化的角度, 研究了我国棉花生产布局变化原因<sup>[1, 3, 4]</sup>。朱启荣认为, 我国棉花生产布局变化受各地区种植业内部的比较效益、农民的非农就业机会、农民的粮食安全保障水平、灌溉条件和自然灾害、技术进步和国家政策的影响<sup>[3]</sup>。朱再清等运用概率优势指数法测算了我国 13 个主要产棉省(区)及全国平均的棉花生产成本、净收益、总产量的一级概率优势值, 结果表明: 在全国 13 个主要产棉省区中, 甘肃、陕西、山西 3 省的棉花生产成本优势排序分别为第 3、第 4、第 7 位; 净收益优势排序分别为第 5、第 12、第 9 位; 产量优势排

序分别为第 1、第 8、第 6 位<sup>[5]</sup>。张晴和刘李峰等认为, 今后中国棉花区域化布局的可能走向为: 长江流域的棉花种植面积有所减少, 西北内陆棉区进一步增加种植面积和生产规模, 而黄河流域棉区则保持平稳<sup>[6-7]</sup>。从这些学者的研究中, 可以看出我国棉花生产向西北、北部, 尤其是其中的甘肃、陕西、山西 3 省扩张和转移的可能性露出了端倪, 但这些学者没有专门从扩张和转移的角度进行实证综合研究, 对“可能性”问题的回答还不够全面和确定, 为此, 本文试图运用主成分分析法对我国 14 个主产省(区)的棉花生产条件进行实证综合分析评价, 重点考察甘肃、陕西、山西 3 省进一步扩大棉花生产的可能性, 从而为我国进行棉花生产区域布局调整、制定棉花发展战略及棉花产业政策提供参考。

## 一、假设的提出及其依据

依据世界棉花发展的一般规律以及甘肃、陕西、山西 3 省的气候与灌溉条件, 本研究提出如下假设: 甘肃省可以发展成为我国西北地区继新疆之后的又一个棉花优势生产区; 陕西省可以进一步稳定为我国棉花主产区。我国棉花主产区存在进一步向西北地区扩张转移的可能性。

### 1. 世界棉花发展的一般规律

世界上主要产棉国家对棉花布局调整的一般规

收稿日期: 2011-09-11

\* 国家转基因科技重大专项资金资助项目(2008ZX08005-001); 中央高校基本科研业务费专项资金资助项目(2010PY033); 华中农业大学农业经济与社会发展研究项目“转基因水稻商业化生产的风险及对策——基于湖北省的调查分析”(XB0913)。

作者简介: 朱再清(1968-), 女, 教授, 博士; 研究方向: 农业贸易与农业政策。E-mail: zhuzq@mail. hzau. edu. cn

律是:从分散到集中,从湿润地区向干旱地区转移<sup>[2]</sup>。随着灌溉农业的发展,世界棉花生产布局不断向光热条件好、气候干燥的地区集中<sup>[3]</sup>。世界棉花布局调整的一般规律已在我国新疆维吾尔自治区的棉花生产中得到印证。新疆维吾尔自治区推广膜下滴灌节水种植技术,既可提高15%~20%的棉花单产,还可以节约2/3的灌溉用水<sup>[8]</sup>,实现了在干旱地区扩大棉花种植。

## 2. 甘肃、陕西、山西3省的气候与灌溉条件

甘肃、陕西、山西3省均在我国14个棉花主产省份之列,各省棉花面积占全国的比重一直都在1%以上<sup>[1]</sup>,棉花生产有一定的基础。

(1)气候条件。一般来讲,棉花优势区域的基本气候条件是:自然生态指标能满足棉花生长的基本需要,无霜期大于200天,大于10℃的积温大于4200℃,年日照时数大于2000h,非灌溉棉区年降雨量在600~1000mm,区域内各项指标的综合适宜程度在90%以上,棉花比较效益高,在种植业中棉花处于主导地位,以棉花为主的棉田效益比其他大宗作物高10%以上,具有明显的竞争优势,并且在较长时间内难以被其他作物替代<sup>[9]</sup>。

甘肃省棉花种植主要分布于河西走廊北部和西北部。甘肃省气候干燥,气温日较差大,光照充足,太阳辐射强。年均降水量300mm左右,降水各地差异很大,在42~760mm之间,自东南向西北减少,降水主要集中在6—9月。年日照时数为1700~3300h,自东南向西北增多。河西走廊年日照时数为2800~3300h,陇南为1800~2300h,陇中、陇东和甘南为2100~2700h。河西

走廊及甘肃西北部的气候条件适宜于棉花种植<sup>[9]</sup>。

陕西省,秦岭以南为亚热带湿润季风气候,秦岭以北大陆性气候特征明显,无霜期5~6个月,年降水量为800~1000mm,南多北少。秦岭以北的陕西北部地区气温和降水条件均难以达到棉花优势区的条件,而且地形破碎保水保土能力差,基本不适宜于棉花种植。渭南地区无霜期和年降水量基本能满足棉花生长的要求。事实上,20世纪70年代陕西全省棉花面积在农作物中占有一席之地,自20世纪80年代起棉花种植面积陡降,进入20世纪90年代,特别是到了1996年以后,由于棉铃虫的暴发,棉花单产降低、比较效益下降等原因致使棉花面积大幅度减少,到2008年全省棉田面积仅85.1km<sup>2</sup>,主要分布于渭南地区<sup>[10]</sup>。

山西地区无霜期在108~170天,年降水量384~600mm。山西省的棉花主要种植在太行山区。太行山棉区热量条件较差,秋季降温快,棉花的生长期短,而且干旱较频繁,导致棉花的纤维品质较差<sup>[6]</sup>。山西地区的气温和降水条件不适宜发展棉花生产。

(2)灌溉条件。气候较干旱地区是否适宜棉花生产主要取决于灌溉条件。新疆棉区正是节水灌溉的典范。同新疆的情况相比,1999—2008年期间,甘肃、陕西、山西3省的有效灌溉面积并不低,均在1200km<sup>2</sup>之上;新疆的棉花面积2008年已占区内有效灌溉耕地面积的48.11%,而甘肃、陕西、山西3省均占不到10.00%,分别仅为5.79%、6.53%、7.09%(见表1)。说明3省均存在将有效灌溉面积转化为棉花面积的潜力。

表1 选定省(区)的棉花面积占有效灌溉面积的比重及其变化

省份	棉花面积/km <sup>2</sup>			有效灌溉面积/km <sup>2</sup>			棉花占比/%		
	1999	2005	2008	1999	2005	2008	1999	2005	2008
山西	52.6	97.5	89.0	1 092.0	1 088.6	1 254.6	4.82	8.96	7.09
陕西	27.7	70.2	85.0	1 309.2	1 298.8	1 301.4	2.12	5.40	6.53
甘肃	31.0	64.0	72.7	972.6	1 030.4	1 254.7	3.19	6.21	5.79
新疆	995.9	1 160.5	1 718.6	3 065.4	3 204.3	3 572.5	32.49	36.22	48.11

数据来源:根据中国国家统计局《中国统计年鉴》(1999—2008)的数据整理而得。

## 二、实证研究

### 1. 研究方法的选取

具备种植棉花的基本气候、灌溉条件的棉花生产区,同时也是粮食等其他作物的优势产区,是否更适宜于发展棉花生产,还取决于较多的经济技术指标,如要素投入、单产水平、机会成本等等。为此,需要选择一个合适恰当的分析评价方法对全国14个

棉花主产省(区)棉花生产的相关技术经济条件进行综合评价分类,以划分出适宜于发展棉花生产的不同类型区,从而验证本研究所提出的假设。而主成分分析法能使这种评价分类变得简单而精确,所得到的分析评价结果具有客观性和准确性<sup>[11]</sup>,且主成分分析法较多见于空间区划、区域开发适宜性评价等研究成果中<sup>[12]</sup>,故本研究采用之。

主成分分析法是将原本众多具有一定相关性的

指标重新组合成一组新的互相无关的综合指标来代替原来的指标的方法。通常在数学上的处理办法就是将原本  $N$  个指标做线性组合,作为新综合指标。其中,最为经典的办法就是选用  $F_1$  的方差来表达,即  $F_1$  的方差  $[\text{Var}(F_1)]$  越大,表示  $F_1$  包含的信息越多。因此,在所有的新型组合中选取的  $F_1$  应该是方差最大的,所以称  $F_1$  为第 1 个主成分。如果第 1 个主成分不足以代表原来  $N$  个指标的信息,再考虑选取  $F_2$ ,即选第 2 个线性组合,为了有效地反应原来信息, $F_1$  已有信息就不需要再出现在  $F_2$  中,即  $F_1$  和  $F_2$  的协方差为 0  $[\text{Cov}(F_1, F_2) = 0]$ ,则称  $F_2$  为第 2 个主成分,依此类推可以构造出第 3、第 4、……,第  $N$  个主成分。

## 2. 变量选择与数据来源

本文选取 11 个与棉花生产相关的变量,将每个省(区)看作 1 个样本。鉴于每个省(区)的规模大小不同,指标的选择采用了绝对指标和相对指标相结合的办法。11 个变量分别是:棉花播种面积( $X_1$ )、棉花单产( $X_2$ )、有效灌溉面积( $X_3$ )、农用机械总动力( $X_4$ )、农业化肥施用量( $X_5$ )、农村居民纯收入( $X_6$ )、农村劳动力数量( $X_7$ )、农村劳动力中初中及以上文化程度所占比重( $X_8$ )、农业增加值占农林牧渔增加值的比重( $X_9$ )、第一产业所占比重( $X_{10}$ )、耕地面积( $X_{11}$ )。上述各项指标数据均来自 2010 年的《中国农村统计年鉴》。

## 3. 主成分分析

(1) 计算主成分贡献率和累计贡献率。从样本相关矩阵出发,对 2009 年全国 14 个省(区)横截面资料进行主成分分析,从而计算出各指标组合所构成的  $N$  个主成分的特征值和贡献率(见表 2)。

表 2 各主成分的特征值、贡献率和累计贡献率

主成分	特征值	贡献率/%	累计贡献率/%
$F_1$	6.150	55.911	55.911
$F_2$	2.314	21.033	76.944
$F_3$	0.852	7.748	84.692
$F_4$	0.752	6.840	91.531
$F_5$	0.426	3.871	95.402
$F_6$	0.264	2.397	97.800
$F_7$	0.127	1.150	98.950
$F_8$	0.048	0.435	99.386
$F_9$	0.042	0.381	99.767
$F_{10}$	0.020	0.183	99.950
$F_{11}$	0.006	0.050	100.000

(2) 提取主成分并计算各变量在主成分上的载荷。主成分个数提取的原则是:提取各主成分对应的特征值大于 1 的前  $M$  个主成分。根据这一原则,本文所研究的内容只需提取表 2 中的前 2 个主成分

$F_1$  和  $F_2$  即可。经主成分旋转分析后,得出了各变量分别在前 2 个主成分上的载荷(见表 3)。

表 3 各变量在主成分上的载荷

变量	主成分	
	$F_1$	$F_2$
棉花播种面积( $X_1$ )	0.442	0.600
棉花单位面积产量( $X_2$ )	-0.684	0.542
有效灌溉面积( $X_3$ )	0.913	0.257
农用机械总动力( $X_4$ )	0.935	-0.030
农用化肥施用量( $X_5$ )	0.959	0.020
农村居民纯收入( $X_6$ )	-0.083	-0.726
农村劳动力数量( $X_7$ )	0.943	-0.027
农村劳动力中初中以上文化程度所占比例( $X_8$ )	0.514	-0.590
农业增加值占农林牧渔增加值的比重( $X_9$ )	0.018	0.769
第一产业所占比重( $X_{10}$ )	0.945	-0.078
耕地面积( $X_{11}$ )	0.895	0.349

表 3 显示,棉花单位面积产量、有效灌溉面积、农用机械总动力、农用化肥施用量、农村劳动力数量、第一产业所占比重和耕地面积等变量指标在第 1 主成分上有较高载荷,说明第 1 主成分基本反映了这些指标的信息;棉花播种面积、农村居民纯收入、农村劳动力中初中以上文化程度所占比例、农业增加值占农林牧副渔增加值的比重和第一产业增加值占地区生产总值比重等变量指标在第 2 主成分上有较高载荷,说明第 2 主成分基本反映了这些指标的信息。提取前 2 个主成分  $F_1$  和  $F_2$  可以基本反映全部(11 个)变量指标的信息,因此决定用这 2 个新变量来替代原来的 11 个变量。

(3) 计算各变量在主成分上的得分系数并建立主成分得分函数。根据以上数据很难从某个主成分看出 14 个省(区)各自发展棉花生产的优劣势,而表 2 显示的前 2 个主成分  $F_1$  和  $F_2$  的累计贡献率已达到 76.944%,因此,可以利用前 2 个主成分  $F_1$  和  $F_2$  作为评价的综合指标,计算出各变量在该 2 个主成分上的得分系数。本文采用回归算法计算出了各变量在 2 个主成分  $F_1$  和  $F_2$  上的得分系数(见表 4)。

根据表 4,可以得到下面的前 2 个主成分  $F_1$  和  $F_2$  的得分函数:

$$F_1 = 0.068X_1 - 0.115X_2 + 0.147X_3 + 0.152X_4 + 0.156X_5 - 0.008X_6 + 0.154X_7 + 0.088X_8 - 0.003X_9 + 0.154X_{10} + 0.143X_{11}$$

$$F_2 = 0.256X_1 + 0.239X_2 + 0.104X_3 - 0.020X_4 + 0.002X_5 - 0.313X_6 - 0.019X_7 - 0.259X_8 + 0.332X_9 - 0.041X_{10} + 0.144X_{11}$$

表4 各变量在主成分上的得分系数

变量	主成分	
	$F_1$	$F_2$
棉花播种面积( $X_1$ )	0.068	0.256
棉花单位面积产量( $X_2$ )	-0.115	0.239
有效灌溉面积( $X_3$ )	0.147	0.104
农用机械总动力( $X_4$ )	0.152	-0.020
农用化肥施用量( $X_5$ )	0.156	0.002
农村居民纯收入( $X_6$ )	-0.008	-0.313
农村劳动力数量( $X_7$ )	0.154	-0.019
农村劳动力中初中以上文化程度所占比例( $X_8$ )	0.088	-0.259
农业增加值占农林牧渔增加值的比重( $X_9$ )	-0.003	0.332
第一产业所占比重( $X_{10}$ )	0.154	-0.041
耕地面积( $X_{11}$ )	0.143	0.144

(4)计算各省(区)棉花生产条件综合评价及排序。根据上述主成分分析逐步得出的数据对14个省(区)的棉花生产条件(11个变量指标)进行综合评价,得出的综合评价值用 $F$ 表示。其方法是:先根据前2个主成分 $F_1$ 和 $F_2$ 的方差贡献率确定权重,因为,其在较大程度上反映了原变量的大部分信息,其累计贡献率为76.944%,因此,可用其方差贡献率作为综合评价的权重;再根据前2个主成分 $F_1$ 和 $F_2$ 的得分函数,计算出14个省(区)各自的 $F_1$ 值和 $F_2$ 值;最后用前2个主成分各自的方差贡献率所确定的权重分别乘以14个省(区)各自的 $F_1$ 值和 $F_2$ 值,将其积相加便得出14个省(区)各自的综合评价值 $F$ 。于是,14个省(区)棉花生产条件各自的综合评价值计算公式如下:

$$F=0.55911F_1+0.76944F_2$$

由综合评价得分值的大小确定某一省(区)棉花生产条件的综合水平。然后根据 $F$ 值的大小进行排序、分类。计算结果如下(见表5):

表5 14个主产省(区)棉花生产条件综合评价及排序

省(区)	$F_1$ 值	$F_2$ 值	$F$ 值	排序
天津	-1.324 42	-1.585 12	-1.960 15	14
河北	1.109 37	-0.040 61	0.589 01	5
山西	-0.554 33	-0.368 38	-0.593 38	11
江苏	0.480 82	-0.653 26	-0.233 81	10
浙江	-0.861 06	-1.067 88	-1.303 10	13
安徽	0.416 31	-0.105 50	0.151 59	6
江西	-0.726 43	-0.391 61	-0.707 47	12
山东	1.672 81	0.045 33	0.970 16	3
河南	1.842 37	0.097 36	1.105 00	2
湖北	0.197 08	-0.378 53	-0.181 07	9
湖南	-0.019 57	0.097 91	0.064 39	7
陕西	-0.593 66	0.364 09	-0.051 78	8
甘肃	-1.050 31	1.559 15	0.612 43	4
新疆	-0.588 96	2.427 04	1.538 17	1

(5)各省(区)棉花生产条件综合评价结果。根据 $F$ 值的大小可以将我国14个主要产棉省(区)分为3类:第一类棉花生产区,综合评价值在0.5以上,包括新疆、河南、山东、甘肃、河北5省,称为最适宜的棉花生产区;第二类棉花生产区,综合评价值在-0.5~0.5之间,包括安徽、湖南、陕西、湖北、江苏5省,称为较适宜的棉花生产区;第三类棉花生产区,综合评价值小于-0.5,包括山西、江西、浙江、天津4省(区),称为一般的棉花生产区。

根据棉花生产条件综合评价结果,甘肃、陕西、山西分别属于最适宜的棉花生产区、较适宜的棉花生产区、一般的棉花生产区。这一结果客观而准确,有力地支持了本研究所提出的假设。

### 三、结 论

根据采用的主成分分析法所考察的甘肃、陕西、山西3省棉花生产的综合评价值在全国14个省(区)中的排位,即第4、第8、第11位,再结合对这3省棉花生产的气候及灌溉条件的分析,可得出如下结论:甘肃省的西北部、河西走廊北部的较大区域具有发展棉花生产的良好条件,可以发展成为我国西北地区继新疆之后的又一个棉花优势生产区;陕西省的棉花生产的综合评价值在全国居中间位置,可以进一步稳定为棉花主产区。我国棉花主产区存在进一步向西北地区扩张转移的可能性。

### 参 考 文 献

- [1] 钟甫宁,胡雪梅.中国棉花生产区域格局及影响因素研究[J].农业技术经济,2008,(1)4-9.
- [2] 张海亮,吴楚材,陈家其.新疆棉花生产发展的方向[J].科技导报,1997(8):49-52.
- [3] 朱启荣.中国棉花主产区生产布局分析[J].中国农村经济,2009(4):31-38.
- [4] 朱启荣,田新建,张旭青.基于比较利益的中国棉花种植空间布局变化的成因分析——对湖北、山东和新疆三省(区)的实证分析[J].农业现代化研究,2004,25(5):340-343.
- [5] 朱再清,张献龙.我国转基因抗虫棉推广与生产优势区域变化实证分析[J].华中农业大学学报:社会科学版,2010(2):12-17.
- [6] 张晴.中国棉花主产区生产条件及发展对策[J].中国棉花,2007(7):8-10.
- [7] 刘李峰,张晴,张照新.中国三大棉区比较优势分析及发展战略探讨[J].麦类文摘:种业导报,2007(1):9-12.
- [8] 农业部农业研究中心.新疆有能力撑起中国棉花半壁江山[EB/OL].(2006-05-08)[2011-07-26].<http://www.texindex.com.cn/articles/2006-5-8/59577.html>.

- [9] 马淑萍. 我国棉花发展思路及对策[J]. 中国农业信息, 2003(1):7-8.
- [10] 鲍亚宁, 刘宝康, 李侠. 关于陕西省棉花生产的思考[J]. 陕西农业科学, 2008(6):77-78.
- [11] 云虹. 主成分分析法在承诺评价模型中的应用[J]. 管理学报, 2006, 3(5):538-542.
- [12] 李德一, 张树文. 基于主成分方法的黑龙江省土地开发适宜性评价[J]. 农业系统科学与综合研究, 2010, 26(1):75-79.

## An Analysis on Expansion of China's Cotton Production Area to Northwest Based on Main Ingredients Analytical Method

ZHU Zai-qing, WANG Yan-ling

(College of Economics and Management, Huazhong Agricultural University, Wuhan, Hubei, 430070)

**Abstract** In order to probe into the possibility of China's cotton production expansion to the northwest regions, based on analyzing climate and irrigation condition of cotton production in Gansu, Shaanxi and Shanxi provinces, this paper uses the main ingredients analytical method to comprehensively evaluate the cotton production technological and economical conditions in 14 cotton production provinces. The result shows that Gansu province, Shaanxi province and Shanxi province belong to the best dominant, the better dominant and ordinary cotton production areas respectively. Hence, a conclusion is drawn that Gansu province can be consolidated to become another dominant cotton production district besides Xinjiang Uygur Autonomous Region, Shaanxi province can be stabilized to become the main cotton production area, and Shanxi province is not suitable to expand cotton production any longer.

**Key words** main cotton production area; area layout; advantage area; proper area; northwest region; climate condition; irrigation condition

(责任编辑:刘少雷)