

# 新疆棉花产量的分析与预测

谭斌 (新疆石河子大学商学院, 新疆五家渠 831300)

**摘要** 通过对新疆历年棉花产量数据的分析, 在比较和选择的基础上, 建立适当的分析预测模型, 对新疆棉花在今后一段时期的产量进行预测。

**关键词** 新疆棉花产量; 模型; 预测

中图分类号 S11+4 文献标识码 A 文章编号 0517-6611(2007)19-05693-02

目前新疆是我国最大的商品棉生产基地。棉花产业已经成为新疆经济的重要支柱产业, 在新疆国民经济中占有举足轻重的战略地位。近年来, 新疆棉花产量基本保持在占全国棉花总产量的1/3和占世界棉花总产量的8%。新疆棉花生产不仅对国内棉花供求起着重要影响, 而且对国际棉花供求也有一定的影响。因此, 分析与预测新疆棉花产量, 不仅可以了解新疆棉花产量的走势, 而且有助于系统安排我国棉花的消费、出口、库存等相关事宜, 从而稳定棉花市场的供求。笔者通过对新疆历年(1950~2005年)棉花产量数据的分析, 在比较和选择的基础上, 建立适当分析预测模型, 对新疆棉花在今后一段时期的产量进行预测。以EViews作为主要的分析工具。

## 1 对原始数据的分析

从图1可以看出, 1980年以前新疆棉花产量增加缓慢, 但1980年以后特别是1990年以后, 新疆棉花产量开始大幅度增长。这主要是由于“八五”时期国家将新疆建设为“优质商品棉生产基地”, “九五”时期国家明确提出将新疆棉区建成国家级年产150万t优质棉生产基地, 新疆维吾尔自治区提出“一黑一白”发展战略。

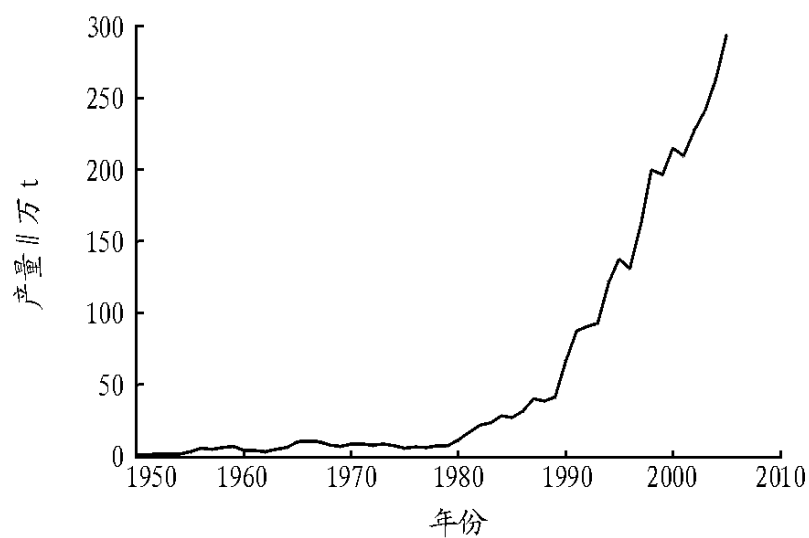


图1 新疆历年棉花产量时序图

## 2 模型的选择

**2.1 复合曲线模型** 采用确定性方法进行建模分析。根据新疆历年棉花产量的走势, 选择复合曲线。

$$P = 0.125 \times 0.096^t \quad (1)$$

式中,  $P$  为产量,  $t$  为时间。新疆历年棉花产量时序曲线与拟合的复合曲线见图2。

通过式(1)对原始数据进行拟合。具体拟合后模式的表达式为:

$$P = 0.125 \times 0.096^t \quad (2)$$

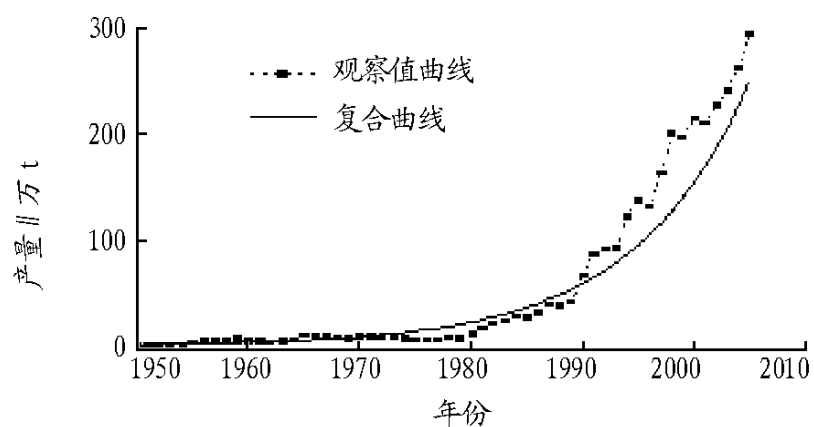


图2 新疆历年棉花产量时序曲线与拟合的复合曲线

尽管方程2)的调整判定系数高达0.92, 但是DW检验结果表明该方程存在明显序列相关。进一步对该模型残差的检验结果见表1。

表1 对模型残差的检验结果

滞后期	自相关系数	偏自相关系数	Q统计量	相伴概率
1	0.858	0.858	43.456	0.000
2	0.690	-0.172	72.119	0.000
3	0.556	0.037	91.060	0.000
4	0.401	-0.192	101.080	0.000
5	0.250	-0.060	105.050	0.000
6	0.183	0.196	107.230	0.000
7	0.165	0.074	109.030	0.000
8	0.098	-0.213	109.680	0.000
9	0.060	0.062	109.930	0.000
10	0.009	-0.196	109.940	0.000
11	-0.059	0.015	110.190	0.000
12	-0.133	-0.054	111.490	0.000

从表1可以看出, Q统计量进一步提示该模型的残差为非白噪声。如果采用该模型进行预测, 那么一定会造成很大的偏差。

**2.2 ARIMA模型** 采用非确定性时间序列分析方法ARIMA进行建模预测。从图1可以看出, 该序列非平稳, 并且有指数上升趋势, 故对原序列取对数后再次绘制时序图, 见图3。

ADF检验结果表明, 该序列在显著性0.01和0.05水平上非平稳, 对取对数后的序列进行一阶差分, 并绘制时序见图4。再一次进行ADF检验, 结果表明显著性0.01和0.05水平上序列平稳。对平稳序列计算EACF(延伸自相关函数), 提示可尝试拟合ARIMA(5, 1, 5)。剔除不显著的系数后, 拟合结果如下:

$$(1 - B)(1 + 0.44B^3 + 0.48B^5) \lg Pt = (1 - 0.52B^2 - 0.47B^5) \quad (3)$$

式(3)的系数的T检验值分别为3.1、-3.1、-3.7、-3.2、2.9。

残差检验结果见表2。

作者简介 谭斌(1966-), 男, 湖南长沙人, 副教授, 从事统计学、时间序列分析等的教学与科研工作。

收稿日期 2007-05-22

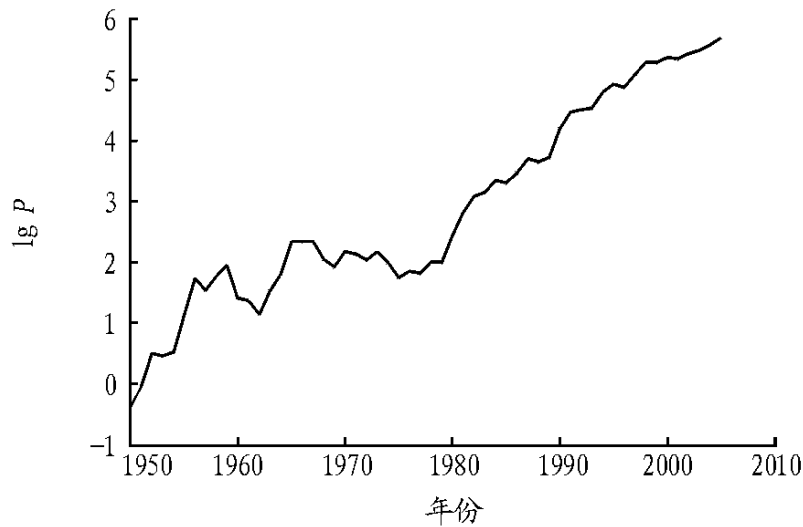


图3 新疆棉花产量取对数后时序图

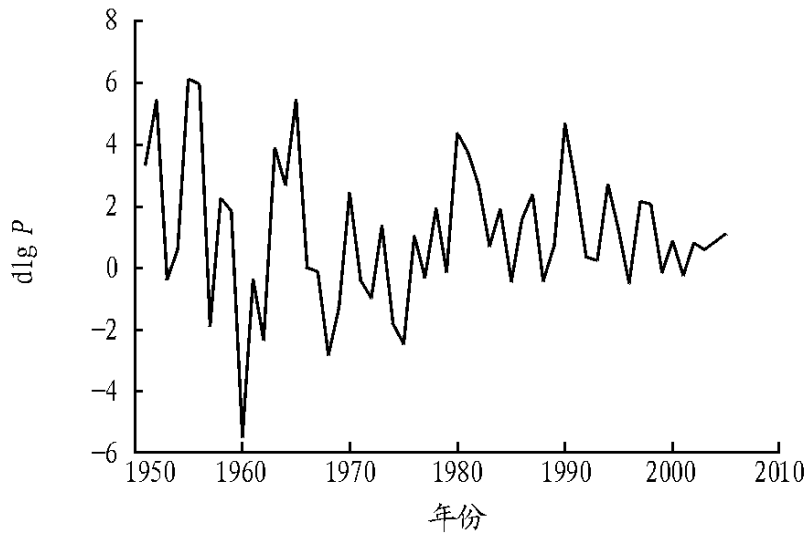


图4 新疆棉花产量取对数后再进行一阶差分后时序图

Q 统计量提示,残差为白噪声。

表2 残差检验结果

残差自相关图	残差偏自相关图	滞后期	自相关系数	偏自相关系数	Q 统计量	相伴概率
		1	0.018	0.018	0.016 3	
		2	0.064	0.064	0.237 3	
		3	-0.060	0.063	0.438 4	
		4	-0.028	-0.030	0.483 1	
		5	-0.090	-0.082	0.949 7	0.330
		6	0.039	0.043	1.038 7	0.595
		7	0.080	0.087	1.424 6	0.700
		8	0.007	-0.026	1.427 3	0.839
		9	0.011	-0.000	1.434 4	0.921
		10	-0.056	-0.052	1.638 4	0.950
		11	0.003	0.014	1.638 9	0.977
		12	-0.096	-0.078	2.265 9	0.972

2.3 残差自回归模型 为了进一步提高拟合效果,最后考虑建立残差自回归模型。最终的模型表达式为:

$$P = 0.1^t + \hat{\epsilon}_t \quad (4)$$

$$\hat{\epsilon}_t = 1.07 \hat{\epsilon}_{t-1} - 0.22 \hat{\epsilon}_{t-2} \quad (5)$$

上式的系数的 T 检验值分别为 26.7、7.2 和 -1.6。DW 检验结果表明上式已不存在序列相关,调整判定系数为 0.98。

进一步对模型进行残差检验,结果见表 3。

表3 残差检验结果

残差自相关图	残差偏自相关图	滞后期	自相关系数	偏自相关系数	Q 统计量	相伴概率
		1	0.017	-0.017	0.016 7	
		2	0.002	-0.002	0.017 0	
		3	0.114	0.114	0.781 9	0.377
		4	0.062	0.067	1.017 7	0.601
		5	-0.221	0.222	4.025 0	0.259
		6	0.053	0.035	4.204 0	0.379
		7	0.156	0.158	5.767 7	0.329
		8	0.025	0.078	5.809 6	0.445
		9	0.118	0.138	6.746 2	0.456
		10	-0.059	-0.164	6.983 0	0.538
		11	0.038	0.017	7.083 8	0.628
		12	-0.130	-0.101	8.299 5	0.600

Q 统计量提示残差为白噪声,说明该模型拟合效果比较理想。

2.4 三种模型效果比较 分别用“复合曲线模型”、“ARIMA 模型”和“残差回归模型”对新疆历年棉花产量进行拟合后,得到相应的残差方差,分别为 368.38、117.90 和 70.38。模型拟合效果表明,残差自回归模型精度最高(残差方差最小),故选择残差自回归模型作为最终的预报模型。

### 3 对今后一段时期新疆棉花产量的预测结果

选择残差自回归模型预测新疆今后一段时期棉花产量,预计 2006~2009 年新疆棉花产量大致为 315.16 万、335.40 万、358.20 万、384.47 万 t。

### 参考文献

[1] 新疆统计局.新疆 50 年 M.北京:中国统计出版社,2005.  
 [2] 新疆统计局.新疆统计年鉴—2006[Z].北京:中国统计出版社,2006.  
 [3] 新疆生产建设兵团统计年鉴编辑委员会.新疆兵团统计年鉴—2006 [Z].北京:中国统计出版社,2006.  
 [4] 易丹辉.数据分析与 Eviews 应用 M.北京:中国统计出版社,2002.