

时间序列分析模型在山东省粮食总产量预测中的应用

张晓杰¹, 张希良²

(1. 鲁东大学地理与资源管理学院, 烟台 264025; 2. 莒县四中, 日照 276500)

摘 要:对比传统时间序列分析模型(线性回归、二次滑动平均、一次平滑、二次指数平滑和三次指数平滑等)与 ARIMA 模型在山东省粮食总产量中的拟合精度, 并应用 ARIMA(2, 1, 12)模型预测了未来 3 年内山东省粮食总产量。结果表明, 在山东省粮食总产量拟合中, ARIMA(2, 1, 12)模型得到的粮食总产量拟合值与观测值的相对误差处于 $\pm 10\%$ 和 $\pm 5\%$ 范围内的分别为 73.333% 和 53.333%, 回归方程的决定系数为 0.959, 优于传统时间序列分析模型; 利用 ARIMA(2, 1, 12)模型预测未来 3 年内山东省粮食总产量, 粮食总产量有逐年上升的趋势, 且增长率逐年上升。

关键词:传统时间序列分析模型; ARIMA 模型; 拟合精度; 预测

中图分类号: F307

文献标识码: A

文章编号: 1005-3409(2007)03-0309-03

Application of Time Series Analysis Model on Total Corn Yield of Shandong Province

ZHANG Xiao-jie¹, ZHANG Xi-liang²

(1. Geography and Resources Management College of Ludong University, Yantai, Shandong 264025, China;

2. No. 4 High Middle School of Juxian, Rizhao 276500, China)

Abstract: The classical time series analysis model (linear regression, two step moving average, one step smoothing, two step EXSMOOTH, three step EXSMOOTH, etc.) and the ARIMA model were compared to predict total corn yield, and ARIMA(2, 1, 12) model was applied to predict the total corn yield in future 3 years. Results showed that the ARIMA(2, 1, 12) model was better than the classical time series analysis model in total corn yield of Shandong Province. The determine coefficient of regressive equation was 0.959 and the relative error between fitted value and measured value among 73.333% and 53.333% were $\pm 10\%$ and $\pm 5\%$, respectively. The total corn yield and the increasing ratio ascended year by year in future 3 years with predicting model of ARIMA(2, 1, 12) applied.

Key words: classical time series analysis model; ARIMA; fitted precision; prediction

传统时间序列分析有线性回归法、滑动平均法和指数平滑法等模型, 它是假定时间序列资料存在某个确定的模式, 随机变量 相对来说并不显著。由于一些社会经济现象的时间序列资料并不具备这种确定性趋势特征, 因此, 传统时间序列分析模型在实际应用中受到一定程度的限制^[1]; 20 世纪 70 年代 Box 和 Jenkins 提出了以随机理论为基础的 ARIMA(自回归差分移动平均)模型, 它把事物在某一固定时刻的状态视为一个随机过程, 利用随机过程去分析描述事物的发展趋势^[2], 因此, 该模型更符合社会经济现象的客观事实。由于 ARIMA 模型在时间序列分析中具有以上所述的优势, 其在经济、交通、卫生、矿业等领域得到了广泛应用^[3-7], 但在粮食产量预测中的应用较少涉及。本文应用 ARIMA 模型预测山东省粮食总产量, 并与传统时间序列分析方法相比较, 分析其在粮食总产量预测中的应用优势, 为

有关部门制定合理的经济政策提供科学的理论依据。

1 山东省粮食总产量概况

1949~2005 年山东省粮食总产量变化趋势如图 1、图 2。由图 1、图 2 可知, 1949~2005 年山东省粮食总产量大体可以分为 5 个阶段, 即 1949~1956 年的慢速增长期, 1957~1960 年的快速下降期, 1961~1999 年的稳定增长期, 2000~2002 年的慢速下降期和 2003 年以后的再度增长期。其中在慢速增长期的末期, 山东省粮食总产量达到 1956 年的 1 372.5 万 t, 之后由于国家经济政策的失误, 粮食产量迅速下降, 1960 年下降到 829.5 万 t, 粮食总产量增长率相对于 1959 年为 -20.9247%, 达到建国以来的最低值。由于国家及时采取了措施, 致使山东省粮食总产量从 1961 年开始进入稳定增长期, 平均增长率为 4.644%, 1996 年为 4.332.7

* 收稿日期: 2006-09-08

基金项目: 鲁东大学大学生科技创新基金资助

作者简介: 张晓杰(1983)女, 山东潍坊人, 本科生, 主要从事自然地理学习和研究。

万 t,达到历史最高纪录。1999 年山东省粮食总产量为 4 269.0 万 t,之后由于耕地面积的减少,粮食总产量也进入慢速下降期,平均增长率为 - 8. 218 %。随后由于科技投资力度的加大,粮食单产量迅速增加,表现为在耕地面积减少的条件下总产量反而增加,即山东省粮食总产量进入 2003 年以后的再度增长期。

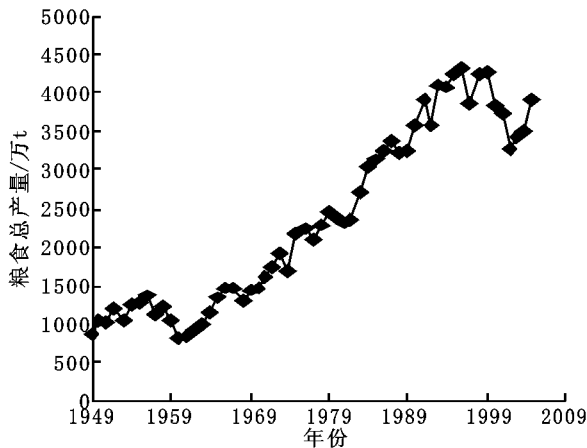


图 1 1949~2005 年山东省粮食总产量

2 传统时间序列分析拟合粮食总产量

分别利用传统的线性回归、二次滑动平均、一次平滑、二次指数平滑和三次指数平滑等时间序列分析模型拟合山东省粮食总产量,各时间序列模型的拟合结果如图 3、表 1。

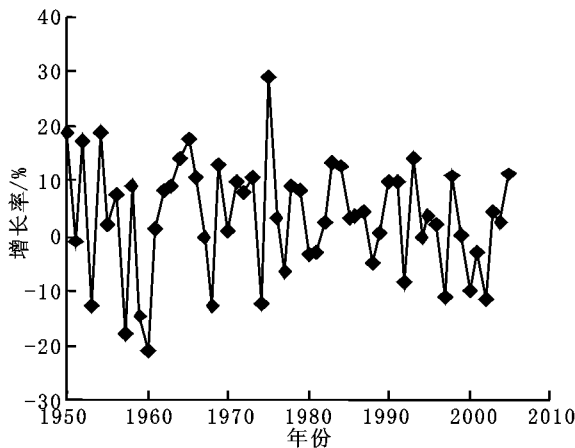


图 2 1950~2005 年山东省粮食总产量增长率变化

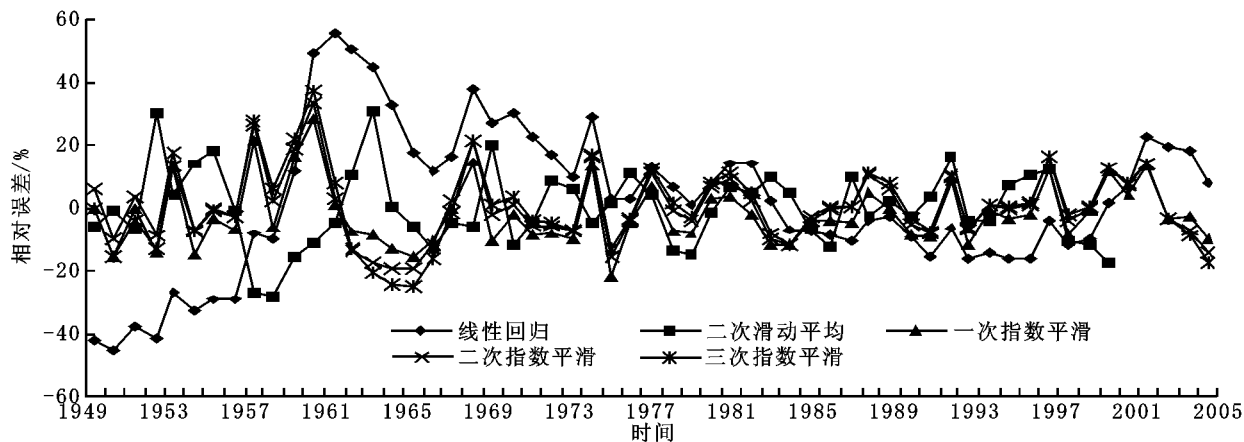


图 3 传统时间序列模型拟合结果

表 1 相对误差范围百分比

相对误差范围	线性回归 / %	二次滑动平均 / %	一次平滑 / %	二次指数平滑 / %	三次指数平滑 / %
±10 %	33.333	57.692	61.404	66.667	61.404
±5 %	14.035	36.538	38.596	36.842	40.351

由图 3 和表 1 可知,山东省粮食总产量传统时间序列拟合模型中,线性回归的拟合效果最差,二次滑动平均次之,一次平滑、二次指数平滑和三次指数平滑较好,其中一次平滑和三次指数平滑的拟合精度较接近。所以在山东省粮食总产量拟合和预测中,一次平滑、二次指数平滑和三次指数平滑模型较理想。

3 ARIMA 模型拟合粮食总产量

建立 ARIMA 模型要求时间序列数据符合正态、零均值和平稳的要求。由图 1 可知,1949~2005 年山东省粮食总产量总体呈波动状态,因此,首先需要对粮食总产量时间序列数据进行适当的处理,以满足建模的要求。对原始时间序列数据进行自然对数转换和一阶差分后,Daniel 平稳性检验统计量 $p=0.211$,样本分组的正态性检验统计量 $p=0.278$,

满足建模的正态、平稳性的要求,可以利用新的时间序列分析方法进行分析和建立预测模型。经过反复调试,确定 ARIMA 模型的自回归阶次 $p=2$,滑动平均阶次 $q=12$,利用所建立的 ARIMA(2,1,12) 模型对山东省粮食总产量进行拟合,对残差序列进行检验, $Q=18.281$,查自由度 $df=30$, $=0.01$ 时, $\chi^2=50.892$; $=0.05$ 时, $\chi^2=43.773$, $Q<\chi^2_{0.05}$,故可以认为在极显著水平下残差序列为白噪声。模型的其他检验指标如表 2。

表 2 ARIMA(2,1,12) 模型检验指标

检验指标	剩余方差	最小信息量	拟合度	相对系数
数 W 值	0.568	- 312.92	62.137	0.788

山东省粮食总产量拟合值与观测值的相对误差分布及数值关系如图 4、图 5。

由图 4 可知,利用 ARIMA(2,1,12) 模型拟合山东省粮食总产量,拟合值与观测值的相对误差中,处于 $\pm 10\%$ 和 $\pm 5\%$ 范围内的分别为 73.333 %和 53.333 %,对比表 1 可知模型的拟合精度得到了很大的提高,拟合值与观测值的回归方程决定系数为 0.959(图 5),因此 ARIMA 模型较传统时间序列分析模型在山东省粮食总产量精确拟合中更具优势。

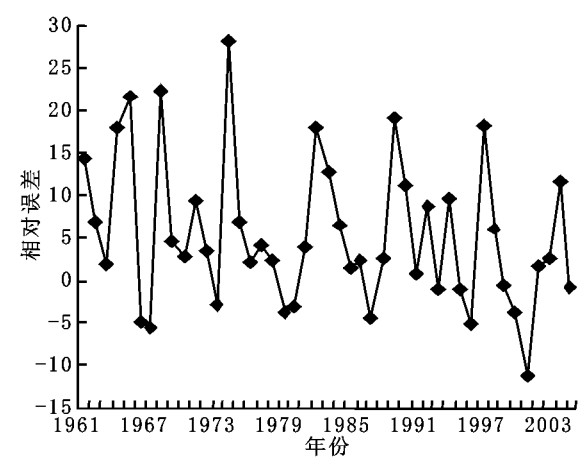


图 4 ARIMA(2,1,12)模型拟合结果

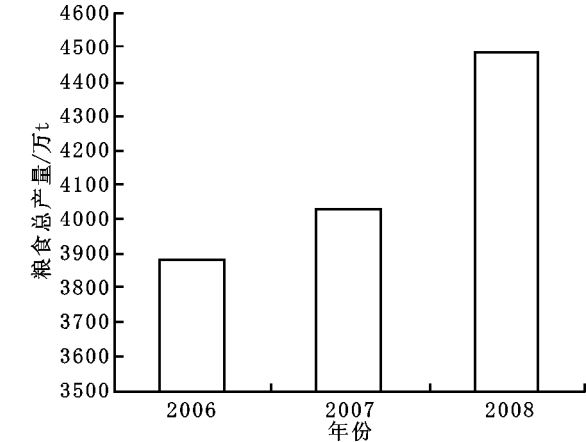


图 6 山东省粮食总产量预测值

4 粮食总产量预测

利用 ARIMA(2,1,12)模型预测未来三年内山东省粮

参考文献:

[1] 石美娟. ARIMA 模型在上海市全社会固定资产投资预测中的应用[J]. 数理统计与管理, 2005, 24(1): 69 - 74.

[2] G P E Box, G M Jenkins. Time Series Analysis: Forecasting and Control [M]. San Francisco: San Francisco Press, 1978.

[3] 高雷,张蕾. 基于市场周转率的股票市场收益预测[J]. 经济经纬, 2004, (6): 127 - 128.

[4] 李勇,等. 基于乘积 ARIMA 模型的产品不确定性需求预测[J]. 系统工程与电子技术, 2005, 27(1): 60 - 62.

[5] 韦丽琴,等. ARIMA 模型在交通事故预测中的应用[J]. 包头医学院学报, 2004, 20(4): 287 - 288.

[6] 史其信,郑为中,等. 道路网短期交通流预测方法比较[J]. 交通运输工程学报, 2004, 4(4): 68 - 71.

[7] 张莹,等. 露天矿一类时间参数的 ARIMA 模型预测[J]. 矿冶, 2004, 13(4): 80 - 82.

(上接第 308 页)

参考文献:

[1] 山仑,徐萌. 节水农业及其生理生态基础[J]. 应用生态学报, 1991, 2 (1): 70 - 76.

[2] 李吉跃. 油松侧柏苗抗旱特性初探[J]. 北京林业大学学报, 1988, 10(2): 23 - 30.

[3] 汤章城. 名词解释[J]. 植物生理学报, 1984, (3): 71 - 72.

[4] 朱万泽. 台湾桉木种源对水分胁迫的光合响应及其抗旱性[J]. 水土保持学报, 2004, 18(4): 170 - 173.

[5] 孙景生. 冬小麦水分生理特征与适宜土壤水分指标研究[A]. 作物栽培生理研究[M]. 北京: 中国农业出版社, 1998. 121 - 124.

[6] 刘庚山,郭安红. 不同覆盖对夏玉米叶片光合和水分利用效率日变化的影响[J]. 水土保持学报, 2004, 18 (2): 152 - 156.

[7] 山仑. 植物生理学及国土整治[J]. 植物生态学通讯, 1994, 30(3): 218 - 227.

[8] 王天铎,等. 小麦对水的利用效率的实验研究 - 单叶与群体测定结果的对比分析[A]. 中国科学院禹城综合实验站年报 (1988 - 1990) [M]. 北京: 气象出版社, 1991. 4 - 13.

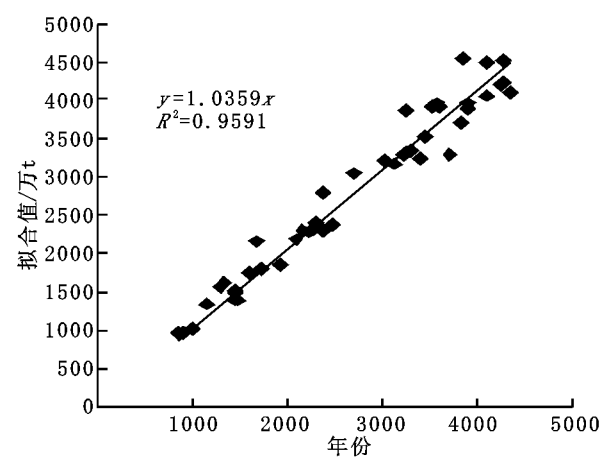


图 5 粮食总产量拟合值与观测值关系

食总产量,预测结果如图 6。由图 6 可知,未来三年内山东省粮食总产量有逐年上升的趋势,且增长率逐年上升,2007 年相对于 2006 年的增长率为 3.770%,而 2008 年总产量值将达到 4 485.618 万 t,相对于 2007 年的增长率为 11.262%。因此,有关部门应该根据粮食产量变化的事实制定合理的经济政策,以保证山东省经济的快速发展。

5 结论

分别利用传统时间序列分析模型(线性回归、二次滑动平均、一次平滑、二次指数平滑和三次指数平滑等)和 ARIMA(2,1,12)模型拟合了山东省粮食总产量,结果表明,在传统时间序列模型的拟合结果中,一次平滑、二次指数平滑和三次指数平滑模型较理想;相对于传统时间序列模型,ARIMA(2,1,12)模型得到的山东省粮食总产量拟合值与观测值的相对误差处于 $\pm 10\%$ 和 $\pm 5\%$ 范围内的分别为 73.333% 和 53.333%,拟合精度显著提高。利用 ARIMA(2,1,12)模型预测未来三年内山东省粮食总产量,结果表明,粮食总产量有逐年上升的趋势,且上升率逐年提高。