

人均 GDP 时间序列模型及预测

官琳琳 门可佩 (南京信息工程大学数理学院, 江苏南京 210044)

摘要 大多数经济时间序列存在惯性或是迟缓性, 通过对这种惯性的分析可以由时间序列的当前值和过去值对未来值进行预测。用 ARIMA 模型可以对天津市人均国内生产总值 1978~2006 年时间序列进行建模和短期外推预测。

关键词 人均 GDP; 非平稳时间序列; ARIMA 模型; 预测

中图分类号 O213 文献标识码 A 文章编号 0517-6611(2009)12-05340-02

The Time Series Mode and Forecast of Per-capita GDP

GUAN Lin-lin et al (College of Math & Physics, Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044)

Abstract The majority of economic time series have a kind of inertia, which is said to be tardiness. It can analyze the kind of inertia in order to forecast future values through the current value and past values. This study mainly discusses how to construct the mode of ARIMA for the economic time series of per capita GDP from 1978 to 2006 in Tianjin and how to forecast the future value in short term.

Key words Per-capita GDP; Non-stationary time series; ARIMA model; Forecast

江泽民同志在十六大报告中提出全面建设小康社会的宏伟目标, 中华民族强国富民的愿望正逐步走向现实。然而 50 多年前的中国, 人均年收入仅 18 美元, 人均工业资本还不足 5 元, 只有美国的 0.13%。新中国经济就是在这样的历史起点上, 在人口增加 2 倍多的情况下, 于 2000 年实现人均 GDP 800 美元的第二步战略目标, 为全面建设小康社会奠定了坚实的基础。人均 GDP 是以某地区一定时期国内生产总值(现价)除以同期平均人口所得出的结果。作为衡量一个国家和地区经济发展水平和综合经济实力的重要指标, 人均 GDP 不仅考虑了经济总量的大小, 而且结合了人口多少的因素, 在国际上它也是划分一个区域经济发展阶段的依据之一。尤其像我国这样的人口大国, 用人均 GDP 指标反映区域经济增长和发展情况会更加准确、深刻, 富有现实意义^[1]。深入分析这一指标对于反映我国经济发展历程、探讨增长规律、研究波动状况, 制定相应的宏观调控政策有着十分重要的意义。我国省区经济是国民经济的重要组成部分, 而省区经济又具有相对独立性。因此, 许多研究文献认为中国省区经济是宏观经济学中相对独立的研究对象和重要的研究层次。天津市是我国省级直辖市之一, 特别是在滨海新区列入国家十一五重点规划项目的情况下, 研究天津市人均 GDP 的统计规律性和变动趋势, 对于地方政策的制定有特别重要的意义。ARIMA 模型是用于一个国家或地区经济和商业预测中比较先进、适用的时间序列模型之一。因此, 笔者试图以天津 1978~2006 年人均 GDP 历史数据为样本, 通过 ARIMA 模型对样本进行统计分析, 以揭示天津市人均 GDP 变化的内在规律性, 并进行后期预测^[2]。

1 时间序列分析法

客观现象都是处在不断发展变化之中, 对现象发展变化的规律, 不仅要从内部结构、相互关联去认识, 而且还应随时间演变的过程去研究, 这就需要运用时间序列分析方法^[3]。时间序列分析是一种广泛应用的数学分析方法, 它主要用于描述和探索现象随时间发展变化的数量规律。时间序列预测是通过对预测目标自身时间序列的处理来研究其变化趋势的。即通过时间序列的历史数据揭示现象随时间变化的

规律, 将这种规律延伸到未来从而对该现象的未来作出预测。现实中的时间序列都是非平稳的, 其变化受许多因素的影响, 有些起着长期的、决定性的作用, 使时间序列的变化呈现某种趋势和一定的规律性, 有些则起着短期的、非决定性的作用, 使时间序列的变化呈现出某种不规则性。时间序列的变化大体可分解为以下几种: 趋势变化。指现象随时间变化朝着一定方向呈现出持续稳定的上升、下降或平稳的趋势; 周期变化(季节变化)。指现象受季节影响, 按一固定周期呈现出的周期波动变化; 循环变动。指现象受不固定的周期影响呈现出的波动变化; 随机变动。指现象受偶然因素的影响而呈现出的不规则波动。

时间序列一般是以上几种变化形式的叠加或组合^[2]。时间序列预测方法分为确定型和随机型 2 大类。确定型时间序列预测方法是用一个确定的时间函数 $y = f(x)$ 来拟合时间序列, 不同的变化采取不同的函数形式来描述, 不同变化的叠加采用不同的函数叠加来描述。传统的时间序列分析方法在经济中的应用, 主要是确定性的时间序列分析方法, 包括指数平滑法、滑动平均法、时间序列的分解等。随着社会的发展, 许多不确定因素在经济生活中的影响越来越大, 必须引起人们的重视。该研究讨论的 ARMA 模型是一类常用的随机时序模型。由博克斯(Box)、詹金斯(Jenkins)创立, 亦称 B-J 方法, 它是一种精度较高的时序短期预测方法, 适用于各种领域的时间序列分析。

时间序列模型不同于经济计量模型的 2 个根本点, 一是时间序列模型的建立不以经济理论为依据, 不考虑被研究变量以外的任何其他变量, 而是依靠被研究变量本身的外推机制描述经济变量的变化, 而经济计量模型的建立以经济理论为基础; 二是在建立经济模型时, 要明确考虑经济时间序列的非平稳问题, 若变量非平稳, 应对该变量进行差分, 即用经济变量的本期值与滞后一期值作相减运算, 差分后得到新的时间序列称做差分序列, 差分序列平稳后再建立经济模型。而在 20 世纪 80 年代以前, 在建立经济计量模型时一直忽视经济变量的非平稳性, 目前已经注意到此问题^[4]。

2 ARMA 模型及数据准备

2.1 ARMA 模型介绍^[3-4] ARMA 模型基本模型有 3 种: 自回归模型 (Autoregressive Model), 简称 AR 模型; 移动平均模

作者简介 官琳琳 (1977-), 女, 山东青岛人, 硕士, 讲师, 从事概率统计及应用研究。

收稿日期 2009-01-19

型(Moving Average Model),简称MA模型;自回归移动平均模型(Autoregressive Moving Average Model),简称ARMA模型。ARMA模型是一种比较成熟的模型,适用于短期预测。

ARMA模型简称B-J模型,B-J模型只适用于平稳时间序列,但在实际问题中,许多时间序列并不近似为平稳时间序列,所以不能直接用均值为常数的平稳过程模型建模。但是可以通过某种处理产生一个平稳的新的时间序列,再用ARMA(p, q)模型建模,这个模型就是ARMA(p, d, q)。d为把非平稳的时间序列转化成平稳时间序列时,对原时间序列进行差分的次数。采用ARMA(p, d, q)模型拟合数据的过程,实质上是先对观测数据进行d次差分处理,然后再拟合ARMA(p, q)模型。

ARMA(p, d, q)的数学形式为: $(B)^d X_t = (B) a_t$ (*) 或 $(B)(1-B)^d X_t = (B) a_t$, 其中 $(B) = 1 - b_1 B - b_2 B^2 - \dots - b_p B^p$, $(B) = 1 - a_1 B - a_2 B^2 - \dots - a_q B^q$ 。称(*)为求和自回归滑动平均模型(Integrated Autoregressive Moving Average Model),简记为ARMA(p, d, q)。其中p、d、q分别为自回归阶数、差分阶数和滑动平均阶数。

2.2 平稳性检查 该研究以天津市人均GDP历史数据(1978~2006)为样本进行分析^[5]。如果对非平稳序列来建立ARMA模型,就会出现虚假回归问题,即尽管基本序列不存在任何关系,也会得到回归模型。因此,要建立ARMA模型,随机序列必须是平稳的,首先对数据进行平稳性检验。

由图1可见,1978-2006年整个时期,人均GDP呈现出指数增长趋势,具有明显的非平稳性。

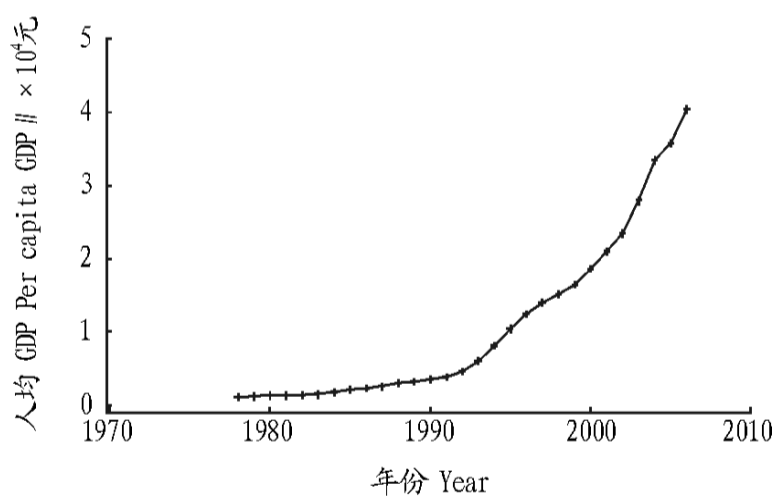


图1 1978~2006年天津市人均GDP增长时间序列

Fig.1 Time Series of per capita GDP increasing of Tianjin City

2.3 数据平稳化与零均值化过程 ARMA模型都是在平稳时间序列基础上建立的,任何齐次非平稳时间序列只要通过适当阶数的差分运算就可以实现平稳,就可以对差分后的序列用ARMA模型拟合,对于原始序列即可用ARMA模型拟合了。如果用非平稳序列来建立模型,就会出现虚假回归问题,即尽管基本序列不存在任何关系,也会得到回归模型。当随机变量不平稳时,统计量的拒绝域远远超过了检验的正常值,由按照一般的检验方法得出的接受假设很可能是错的。因此,要建立ARMA模型,随机序列必须是平稳的。对于含有指数趋势的时间序列,可以通过取对数将指数趋势转化为线性趋势,然后再进行差分以消除线性趋势。从理论上而言,足够多次的差分运算可以充分地提取序列中的非平稳确定性信息。但应当注意的是,差分运算的阶数并不是越多越好。因为差分运算是一种对信息的提取、加工过程,每次

差分都会有信息的损失,所以在实际应用中差分运算的阶数要适当,应当避免过度差分,直至成为平稳序列。由图2和图3可知,经过转化与差分运算,1978~2006年天津市人均GDP增长时间序列已基本平稳。

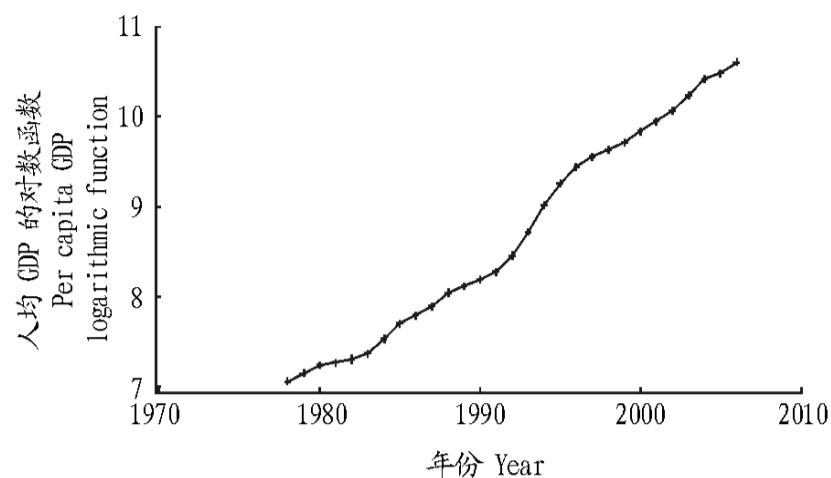


图2 通过取对数将指数趋势转化为线性趋势的天津市人均GDP增长时间序列

Fig.2 The linearity trend time series of per capita GDP increasing of Tianjin City by taking a logarithm

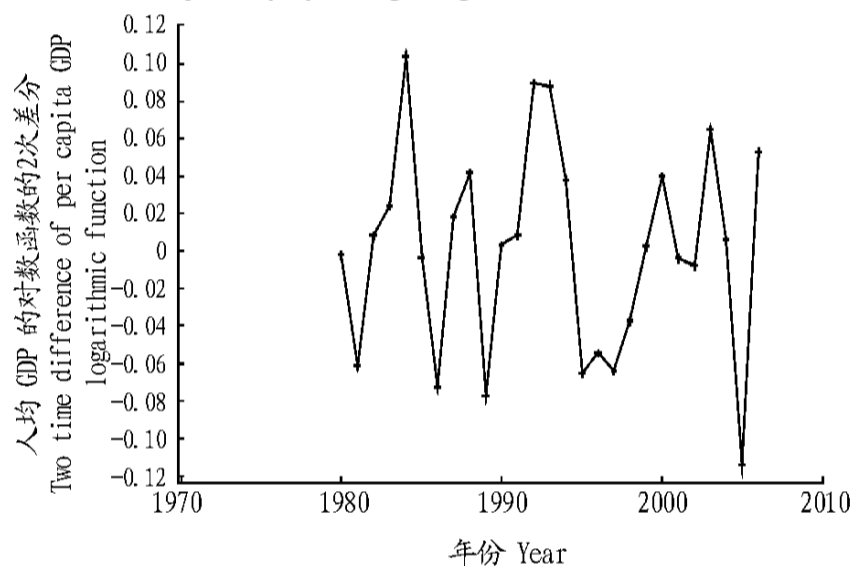


图3 2次差分后的天津市人均GDP增长时间序列

Fig.3 Time series of per capita GDP increasing of Tianjin City after two time difference

3 时间序列模型的建立

3.1 模型建立 ARMA模型的识别与定阶可通过样本的自相关与偏自相关函数的观察获得,如:AR(p)模型自相关函数拖尾,偏自相关函数p步截尾;MA(q)模型自相关函数q步截尾,偏自相关函数拖尾;而ARMA模型的自相关函数与偏自相关函数都具有拖尾性^[6-7]。通过筛选,取对数后的序列适合ARMA(0,2,2)模型,即

$$(1 - B^2)\ln(s_t) - 0.203161 = (1 + 1.9496B + B^2) \epsilon_t$$

3.2 模型预测 由于该时间序列模型经过取对数再作二阶差分才平稳,且模型由有限个数据拟合而成,所获得的模型反映的是短期变化关系,而不是长期变化关系,因此只适合进行短期预测。通过计算得出2007、2008年天津市人均GDP预测的结果。利用此模型对天津市的人均GDP进行预测,最终的预测结果2007年为41588.17元,2008年为45034.79元。从预测结果来看,天津市的人均GDP在2007、2008年将保持平稳高速增长。

4 小结

研究表明,时间序列预测法是一种重要的预测方法,其模型比较简单,对资料的要求比较单一,只需变量本身的历

也相对较少,从而产量也最低。在满足有效积温的条件下,应尽早播种,赶上地温回升较快的时期,确保蓖麻出苗好,发育正常,成熟充分。该试验结果表明,在神东矿区采煤塌陷区适当早播优于晚播。

(3) 种子播深的确定。种子播深主要影响温度和水分。

播深较深有利于保持充足水分,但会造成温度低,不利于种子萌发。播深较浅,地温较高,但水分条件稍差。从试验结果看,在神东矿区采煤塌陷区保持充足的水分比温度更重要。

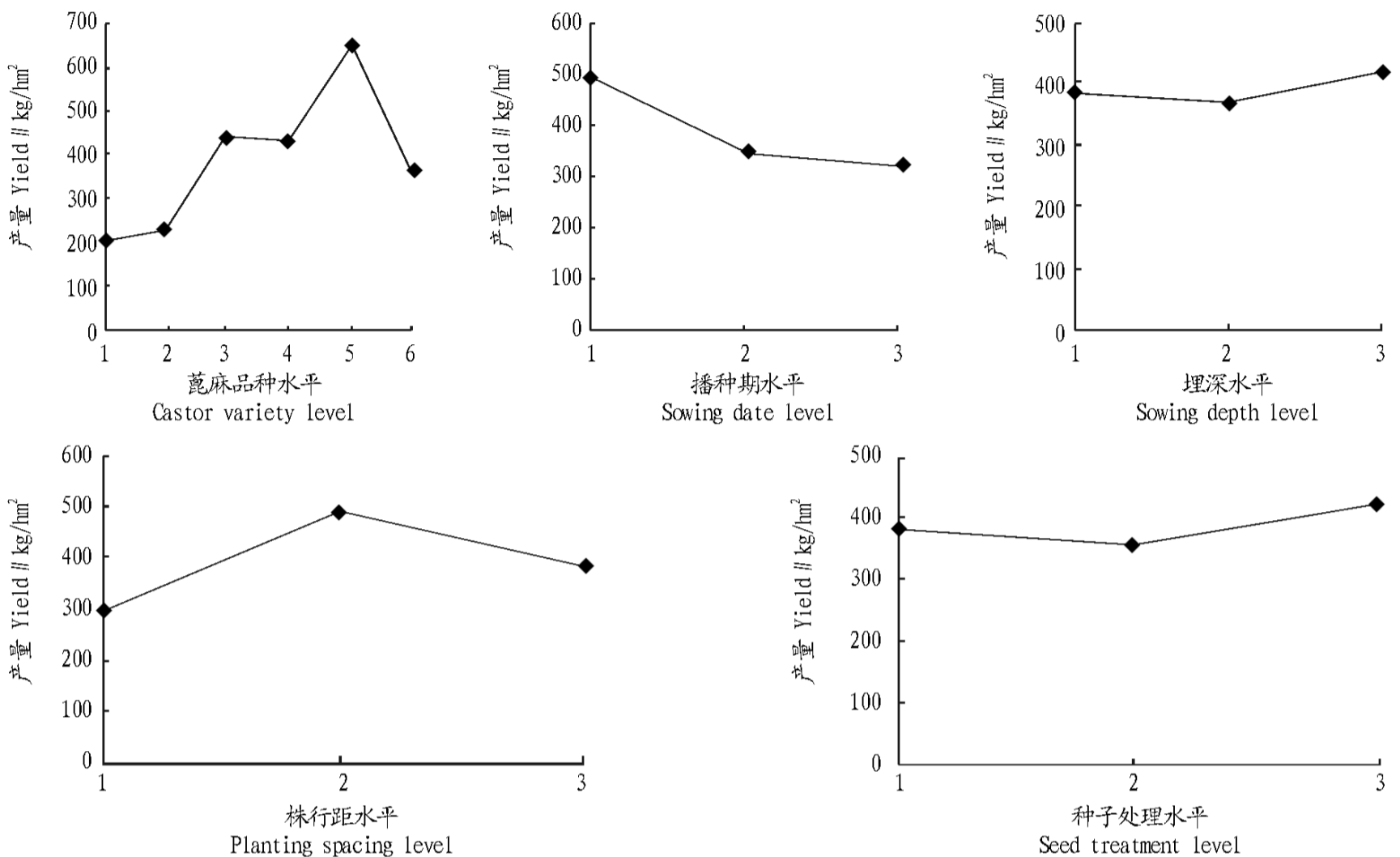


图1 各水平因素趋势

Fig.1 Factor trends of all levels

各因素不同水平对产量的影响: 法国杂交种 CSR24.181 较其他5 个品种更适宜当地环境,比排名第2 的无刺青产量高出45.8%; 种子埋深8 cm 较5.3 cm 更为适合蓖麻生长,增产9.8%和13.9%; 适当早播可以延长生长期,提高产量,5月2日播种的蓖麻产量明显高于5月9日和5月16日播种的产量,分别提高产量44.6%和55.0%; 株行距60 cm×60 cm 最为合适,较其他2 种株行距提高产量67.0%和28.7%; 冷水浸种较其他2 种种子处理方法可提高产量12.2%和19.4%。单从提高产量方面考虑,由试验结果可知,法国杂交种CSR24.181 在5月初经过冷水浸种,播深8 cm,种植密度60 cm×60 cm 的情况下能够取得较好的产量。

参考文献

[1] 尹秀玲,徐兴友,孟宪东,等.蓖麻的开发利用[J].生物学杂志,1998

(2):35.

- [2] 吕相义,李军,毕万新.国外蓖麻生产的现状、问题及对策[J].中国农技推广,2003(6):25-26.
- [3] TONGGONA P, MCHENA S C, MARGAI K. Response of castor (*Ricinus communis* L.) cultivars to different plant spacings in contrasting agro-ecological zones in Zimbabwe[J]. Zimbabwe Journal of Agricultural Research, 1991, 27(1):1-9.
- [4] 马德富,张春华,包红霞,等.蓖麻不同种植密度对群体生育动态的影响[J].中国油料作物学报,1998,20(1):57-58.
- [5] 田福东,李金芹,张春华.密度和肥料对蓖麻光合性能及产量的影响[J].吉林农业科学,2000,25(1):29-31.
- [6] 刘秦.蓖麻栽培技术要点[J].中国种业,2004(1):51.
- [7] 任余艳.神东矿区采煤塌陷对旱作农田生产力影响的试验研究[D].呼和浩特:内蒙古农业大学,2007.
- [8] 中国农科院油料作物研究所.中国特油作物品种资源目录[M].北京:中国环境科学出版社,1997:1-80.
- [9] 盖钧镒.试验统计方法[M].北京:中国农业出版社,2002:172-175.
- [10] 兰巨生.作物遗传统计法[M].石家庄:河北人民出版社,1982:73-88.

(上接第5341页)

史数据,在实际应用中有着广泛的适用性。在应用中,应该根据所要解决的问题及问题的特点等方面因素来综合考虑并选择相对最优的模型。天津市是我国环渤海地区的物流中心,滨海新区已被列入国家“十一五”发展规划,在此情况下研究如何不断调整经济结构,改善经济运行质量,如何提高自主创新能力,发展节约能源资源和保护生态环境的制造业和高新技术产业,对于实现天津经济在数年内的快速增长,并且带动环渤海地区的经济发展以及国家统筹区域经济

都具有重要意义。

参考文献

- [1] 陈泽中,李锋,杨启智.中国人均GDP(1952~2002)时间序列分析[J].统计与决策,2005(2):61-62.
- [2] 弗朗西斯·X·迪博尔德.经济预测[M].北京:中信出版社,2003.
- [3] 王振龙.时间序列分析[M].北京:中国统计出版社,2000.
- [4] 张晓峒.计量经济分析[M].北京:经济科学出版社,2000.
- [5] 国家统计局.中国统计年鉴-2007[M].北京:中国统计出版社,2007.
- [6] GEORGE E P BOX, GWLYMMJENKINS, GREGORY C REINSEL. 时间序列分析预测与控制[M].北京:中国统计出版社,1997.
- [7] 岳朝龙.SAS系统与经济统计分析[M].合肥:中国科学技术大学出版社,2003.