

基于时间序列分析的货运周转量预测

叶 斌,代晓琴,陈朝海

(成都理工大学信息管理学院,成都 610059)

摘 要:基于重庆市统计局的实际统计数据,通过对实际样本数据进行预处理,确认重庆水运货运周转量序列为平稳非白噪声序列。在此基础上,通过对 1995 - 2009 年重庆水运货运周转量的数据分析,利用时间序列分析方法建立了 ARMA 预测模型,结果显示该模型具有较好的预测效果,对重庆水路货物运输工作及水运的发展决策有一定的参考价值。

关键词:重庆水运;时间序列分析;ARMA 模型;预测

中图分类号:O211

文献标识码:A

时间序列是变量按时间间隔的顺序而形成的随机变量序列,时间序列分析就是估算和研究某一时间序列在长期变动过程中所存在的统计规律性。博克斯和詹金斯的专著《时间序列分析预测与控制》^[1]是时间序列分析发展的里程碑,他们的工作提供了对时间序列进行分析、建模、预测和控制的系统方法,在动态数据的处理分析、预测等方面显示出传统的数理统计静态处理手段无可比拟的优越性。随着时间序列分析方法的不断成熟,已广泛应用于气象、天文、经济、管理、机械、生物、医学、水利、化工和农林等各个部门和研究领域,并取得了很好的效果^[2-4]。重庆水运 1995 - 2009 年的货运周转量可以作为时间序列,通过对相关统计数据建立时间序列模型,再进行参数估计后就可以对重庆市水运的货运周转量进行模拟预测。本文通过对 1995 - 2009 年的重庆水运货运周转量的数据进行分析,建立 ARMA 模型对未来重庆水运的货运周转量进行预测。

1 ARMA 模型与建模方法

1.1 自回归移动平均模型

自回归移动平均模型(ARMA)是平稳时间序列模型,时间序列的平稳性是时间序列建模的重要前提。ARMA 模型是指序列的当前值不仅与自身的过去值有关,而且还与以前进入系统的外部干扰存在一定的依存关系,利用该模型刻画这种动态特征时,模型既包括自身的滞后项,也包括过去的外部干扰,记做 $ARMA(p,$

$q)$ 其模型的表达式为:

$$\begin{cases} X_t = c + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \\ \varepsilon_t \sim WN(0, \sigma^2), \forall s < t, E(X_s \cdot \varepsilon_t) = 0 \end{cases}$$

其中, ϕ_1, \dots, ϕ_p 为自回归系数, $\theta_1, \dots, \theta_q$ 为移动平均系数。

1.2 ARMA 模型的建模方法

平稳时间序列常用的建模方法有 Box - Jenkins^[5]方法、Pandit - Wu 方法以及长阶自回归建模方法等。本文采用 Box - Jenkins 方法进行建模^[6-7]。

(1) 对时间序列进行平稳性分析。检验序列是否为平稳非白噪声序列,若序列是非平稳序列,则需要对序列进行平稳化,主要是根据数据特点有针对性的对数据序列进行对数化、差分或季节差分等预处理。

(2) 模型的识别与建立。对平稳序列进行模式识别,判断序列是属于 AR 模型,还是属于 MA 模型或 ARMA 模型。模型识别后,确定模型的最高阶数并从低阶到高阶对模型进行拟合及检验,以确定最适宜的模型。最后估计模型的位置参数,并检验参数的显著性,以及模型本身的合理性。

(3) 模型评价及预测。通过比较预测值和实际值的差异来评价预测的精确程度。

2 实例建模与预测

2.1 平稳性分析与处理

在建立时间序列的模型之前,先要确定该时间序列的平稳性,才能对模型作出正确的估计。若某时间序列

是非平稳的,通过取对数或差分运算,可以得到平稳性的序列。判断时间序列的平稳性可通过单位根的检验来判断。

单位根检验^[8]目的在于确定一个序列是否经过差分后成为平稳序列,这是建立 ARMA 模型的基础工作。将 1995 - 2009 年重庆水运货运周转量的数据建立时间序列 F1,并做出时间序列 F1 的时序图,如图 1 所示。

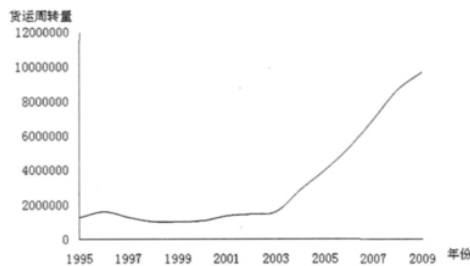


图 1 1995 - 2009 年重庆水运货运周转量 (单位:万吨公里)

运用 Eviews 对时间序列 F1 做单位根检验,如图 2 所示,得到: ADF 的 t 检验量为 -0.268 393,而在置信水平为 1%、5% 和 10% 时 t 检验量分别为 -4.057 910、-3.119 910 和 -2.701 103,其绝对值都比 ADF 的 t 检验量的绝对值要大,所以 F1 是非平稳的。

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.268393	0.9057
Test critical values:		
1% level	-4.057910	
5% level	-3.119910	
10% level	-2.701103	

图 2 F1 序列的 ADF 检验表

通过图 1 可以看出,该时间序列 F1 有非常明显的指数趋势,因此对它进行取对数的运算,可以消除指数趋势的影响。将取对数后的时间序列命名为 F2,即 $F_2 = \log(F_1)$ 。对 F2 做单位根检验后发现,F2 时间序列仍是非平稳的。最终经过对序列 F2 作二阶差分处理后得到 F,继续对 F 做单位根检验,如图 3 所示,发现在置信水平为 1%、5% 和 10% 时 t 检验量的绝对值都比 ADF 的 t 检验量的绝对值要小,说明该序列是平稳序列。

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.573085	0.0048
Test critical values:		
1% level	-4.121990	
5% level	-3.144920	
10% level	-2.713751	

图 3 F 序列的 ADF 检验表

2.2 模型的识别与建立

通过单位根检验之后,我们得到序列 F 是平稳序列。通过 Eviews 软件计算时间序列 F 的自相关系数和

偏自相关系数,利用自相关系数图和偏自相关系数图进行模型的识别和定阶,如图 4 所示。从图 4 可以初步判断该序列的自相关图六阶截尾,而对于序列的偏自相关图不能马上作出判断。

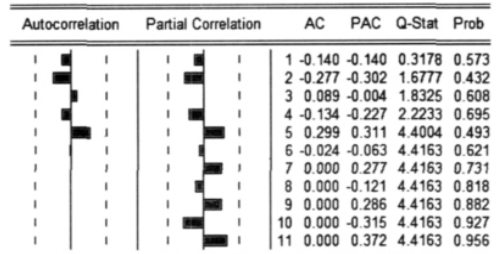


图 4 F 序列的自相关与偏自相关图

针对上述情况,尝试几种不同的模型拟合,例如 MA(6), ARMA(1,6), ARMA(2,6), ARMA(3,6) 等。经过多次试验尝试,最终选择 ARMA(5,6) 模型,而且由于该模型中自回归和移动平均部分的系数只有 AR(5) 和 MA(6) 的系数是显著的,所以我们把其他阶的系数全部放弃,最终的参数估计结果如图 5 所示,调整后的拟合优度达到了 0.986 878,可见该模型相当可靠。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(5)	-0.200981	0.054114	-3.714030	0.0099
MA(6)	0.985071	2.37E-06	415845.2	0.0000
R-squared	0.988753	Mean dependent var	-0.016527	
Adjusted R-squared	0.986878	S.D. dependent var	0.226233	
S.E. of regression	0.025915	Akaike info criterion	-4.255686	
Sum squared resid	0.004029	Schwarz criterion	-4.235826	
Log likelihood	19.02274	Durbin-Watson stat	1.163077	

图 5 ARMA 模型参数估计与效果指标表

综合上述分析过程,针对原序列(F1): 1995 - 2009 年重庆水运货运周转量的数据序列,建立了 ARMA(5,6) 模型进行拟合,模型的表达式为:

$$F_t = -0.200981F_{t-5} + \varepsilon_t + 0.985071\varepsilon_{t-6}, \varepsilon_t \sim WN(0, \sigma^2)$$

2.3 模型的预测与分析

根据建立的模型可以得到重庆水运货运周转量的变化规律,同时也可以利用该模型对重庆水运的货运周转量进行短期预测,得到 2010 年的重庆水运货运周转量预测值(表 1)。

表 1 2005 - 2009 年的实际值与 2010 年的测值对照表(单位:万吨公里)

年份	货运周转量
2005	4004600
2006	5331900
2007	6998600
2008	8655800
2009	9684000
2010	11568928

由于该模型是经过二阶差分并且通过有限个数据的拟合得到,因此该模型只适合进行短期预测,预测得到重庆水运 2010 年货运周转量将增加到 115 689 28 万吨公里。从预测结果看出,重庆水运货运周转量有望在今后几年保持平稳高速增长,对重庆水运部门的可持续发展决策有一定的参考价值。

参 考 文 献:

- [1] 博克斯,詹金斯. 时间序列分析预测与控制 [M]. 北京: 中国统计出版社, 1997.
- [2] 胡永红,吴志峰,李定强. 基于 ARIMA 模型的区域水生态足迹时间序列分析 [J]. 生态环境, 2006, 15(1): 94-98.
- [3] Romilly P. Time series modeling of global mean temperature for managerial decision-making [J]. Journal of Environment Management 2005, 76: 61-70.
- [4] Nickerson M, Madsen B C. Nonlinear regression and ARIMA models for pre-cipitation chemistry in East Central Florida from 1978 to 1997 [J]. Environment Pollution 2005, 135: 371-379.
- [5] 王晓鹏,曹广超,丁生喜. 基于 Box-Jenkins 方法的青南高原降水量时间序列分析模型与预测 [J]. 数理统计与管理 2008, 27(4): 565-570.
- [6] 王黎明,王连,杨楠. 应用时间序列分析 [M]. 上海: 复旦大学出版社, 2009.
- [7] 朱宗元. 我国保险赔付的时间序列分析—建模与预测 [J]. 数理统计与管理, 2010, 29(4): 698-704.
- [8] 张晓峒. EViews 使用指南与案例 [M]. 北京: 人民邮电出版社, 2007.

Transportation Freight Quantity Forecasting Based on Time Series Analysis

YE Bin, DAI Xiao-qin, CHEN Zhao-hai

(College of Information Management, Chengdu University of Technology, Chengdu 610059, China)

Abstract: Based on the practical datum of statistics bureau in chongqing, the datum are pre-processed and the chongqing water transportation freight quantity series is considered to be steady non-white noise series. On this basis, through the 1995-2009 chongqing water transportation freight quantity of data analysis, we use time series analysis method to establish the ARMA forecasting model. The results show that the model has better prediction results, which can be used to have certain reference value for waterway transport of goods work and water transport development decision-making.

Key words: chongqing water carriage; time serie analysis; ARMA model; forecast