

# 城市轨道交通客流增长滞后性与预测方法研究

陈小鸿 王翔 李佳

(同济大学道路与交通工程教育部重点实验室 201804, 上海//第一作者 教授)

**摘要** 城市轨道交通客流是确定城市轨道交通建设规模及投资决策的主要依据,然而客流增长速度并不与城市轨道交通建设同步。考虑到城市轨道交通客流增长的滞后性,利用上海城市轨道交通1996—2010年客流与建设数据,基于协整理论建立了向量自回归模型进行城市轨道交通客流预测。研究结果表明:城市轨道交通客流与城市轨道交通网络长度以及连接度存在协整关系;在城市轨道交通建设初期,客流增长对于网络长度具有显著滞后性,随着城市轨道交通系统的完善客流滞后性减弱;2011年上海城市轨道客流预测误差为9.3%。

**关键词** 城市轨道交通; 客流预测; 滞后影响; 协整理论; 向量自回归模型

中图分类号 U 293.13; U 231

## Lagged Effect of Passenger Flow Growth and Forecasting Method in Shanghai Urban Mass Transit

Chen Xiaohong, Wang Xiang, Li Jia

**Abstract** Passenger flow of urban mass transit determines the construction scale and investment. However, the growth of passenger flow is out of sync with rail construction. In view of the lagged effect of passenger flow growth, the annual data of Shanghai urban mass transit during 1996—2010 is used to establish the vector auto regression model on the basis of co-integration theory, so as to forecast the passenger flow. Research results indicate that the passenger flow has a co-integration relationship with the track mileage and network connection degree. At the early phase of urban mass transit construction, the lagged effect on track mile is significant, but it tends to become weaker with the development of urban mass transit, the prediction error of Shanghai urban mass transit passenger flow in 2011 is 9.3%.

**Key words** urban mass transit; passenger flow forecasting; lagged effect; co-integration theory; vector auto regression model

**First-author's address** Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, 201804, Shanghai, China

• 22 •

在我国,对于城市轨道交通客流总量的预测通常采用时间序列数据来预测,如采用逐步回归、多元线性回归等方法<sup>[1]</sup>。然而,在实际应用过程中经常会出现预测失准的情况。其主要原因如下:

(1) 谬误回归。在传统回归分析方法中普遍应用最小二乘法来估计参数。该方法要求随机误差项数学期望为0,方差为定值。然而,多数社会经济时间序列数据具有非平稳性,即随机误差项数学期望与方差随时间变化,不满足统计方法的前提要求,因此,易导致变量间原本不存在因果关系,却得到显著统计关系的错误结论,即谬误回归。

(2) 滞后性考虑不足。上海轨道交通网络的不断建设促使轨道交通客流加快增长,此阶段轨道交通客流对于轨道交通网络建设具有滞后性<sup>[2]</sup>,即当期轨道交通客流变化主要由前期的轨道交通建设造成。传统回归分析方法仅关注变量间当期的影响,而对影响的滞后性考虑不足。

协整理论(co-integration theory)是处理非平稳时间序列的有效方法。若多个非平稳时间序列通过某种线性组合能消除序列之间的随机趋势,则这些序列存在协整关系。协整关系的检验是避免谬误回归的预检验<sup>[3]</sup>。基于协整理论的向量自回归(VAR)模型将内生变量的滞后项作为解释变量,能够体现自变量与因变量的历史变化对于当期因变量的影响。目前,协整理论在社会经济研究中已有广泛应用。文献[4]就货运周转量与工业生产指数、石油价格之间是否存在协整关系进行了分析;文献[5]利用协整理论分析了价格、竞争策略对于机场客流的影响。

上海轨道交通客流呈指数型增长,其时间序列具有非平稳性。协整理论能筛选出对于客流增长具有因果关系的影响因素,避免谬误回归的发生;同时,能体现这些因素的历史变化对于当期客流的影响。本文基于协整理论分析上海轨道交通建设对于轨道交通客流的影响;研究客流增长滞后于轨道交通建设的现象;最后利用VAR模型进行客流预测,

并解释误差增大的原因。

## 1 协整分析方法

### 1.1 平稳性检验

平稳性检验是协整检验的前提,只有同阶单整序列才可能具备协整关系。本文中讨论的平稳性为弱平稳,即时间序列数据的期望( $E$ )、方差( $V_{ar}$ )和协方差( $C_{ov}$ )不随时间变化。若时间序列 $\{Y_t\}$ 对于所有时间 $t$ 满足 $E(Y_t) = \mu$ ,  $V_{ar}(Y_t) = \sigma^2$ 为不变常数,  $C_{ov}(Y_t, Y_{t-j}) = c$ 不依赖时间 $t$ 而仅依赖滞后期 $j$ 时,则时间序列 $\{Y_t\}$ 平稳,为零阶单整序列。

具有单位根的时间序列为非平稳序列。通常选择增广迪基—富勒(ADF)检验法对变量进行单位根检验,滞后阶数由 AIC (Akaike Information Criterion) 和 SC (Schwartz Criterion) 最小准则确定。原假设  $H_0$  表示序列具有单位根,非平稳;备选假设  $H_1$  表示序列没有单位根,平稳。

ADF 检验估计由如下回归构成:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^m \alpha_i \Delta Y_{t-i} + \varepsilon_t \quad (1)$$

式中:

$\varepsilon_t$ ——白噪声误差项;

$\Delta Y_t$ ——一阶差分序列,即  $\Delta Y_t = Y_t - Y_{t-1}$ ;

$m$ ——滞后阶数;

$\alpha, \beta, \delta$ ——估计参数,若  $\delta$  显著为零,则时间序列平稳。

### 1.2 协整检验

若多个非平稳序列具有协整关系,则可通过这些序列的某种线性组合形成平稳序列。基于回归系数的 Johansen 极大似然法是目前经济学研究中常用的多变量协整关系的分析方法,其能得到所有变量之间协整关系的个数与协整方程的估计结果。Johansen 检验是基于向量自回归(VAR)模型的检验方法,原假设  $H_0$  表示有 0 个协整关系;备选假设  $H_1$  表示有  $M$  个协整关系。检验方法通常采用关于系数矩阵协整似然比( $R_{LM}$ )检验法:

$$R_{LM} = -n \sum_{i=M+1}^{N'} \log(1 - \lambda_i) \quad (2)$$

式中:

$M$ ——协整关系个数;

$\lambda_i$ ——按由大到小排列的第  $i$  个特征值;

$n$ ——样本容量;

$N'$ ——变量个数。

当  $R_{LM}$  大于临界值时,则存在  $M$  个协整关系。

### 1.3 向量自回归模型

VAR 模型反映多个变量之间的动态互动关系。在 VAR 模型中,因变量由其自身的滞后项与自变量的滞后项来解释。由非平稳时间序列建立 VAR 模型,首先需通过协整检验以保证因变量与自变量之间具备因果关系,否则 VAR 模型无逻辑意义;此外,VAR 模型需进行稳定性检验,当 VAR 模型所有特征根的绝对值小于 1 时,模型稳定收敛。

滞后期为 2 的双变量 VAR 模型形式如下:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-2} \end{bmatrix} + \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{bmatrix} \quad (3)$$

式中:

$y_t, x_t$ ——具有协整关系的非平稳序列;

$\varepsilon_{yt}, \varepsilon_{xt}$ ——不相关的白噪声扰动项;

$C_1, C_2$ ——常数项;

$b, g$ ——各滞后项系数。

## 2 城市轨道交通数据处理

### 2.1 数据说明

本文使用上海市 1996—2010 年城市轨道交通客流数据和网络建设发展数据,来源包括《上海统计年鉴》、《上海综合交通年报》与《上海第四次综合交通调查总报告》。

文献[6]将城市轨道交通客流的影响因素分为社会经济、城市道路公交系统与城市轨道交通系统 3 大类。其中,城市轨道交通系统的影响最为显著。可达性好、可靠性强的城市轨道交通网络将促进客流的快速增长<sup>[7-8]</sup>。本文重点分析城市轨道交通客流相对于网络规模与结构变化的滞后性。选择变量如下。

(1) 城市轨道交通网络规模方面有:

$L$ ——网络长度,所有线路长度总和(共线段仅计算 1 次);

$N$ ——站点数,网络所有站点数量;

$A$ ——弧段数,网络所有相邻站点之间的连接弧段数;

$C_r$ ——网络闭环数,网络中闭环的数量,  $C_r = A - N + 1$ 。

(2) 城市轨道交通网络结构方面有:

$S$ ——平均站点间距,反映网络建设更注重高区域覆盖(短间距)还是高运营速度(长间距),  $S =$

$L/A$ 。

$R$ ——网络成环率,网络实际环数除以最大可能环数(最大可能环为 $2N-5$ ),成环率越高则乘客可选路径越多 $R=(A-N+1)/(2N-5)$ 。

$P$ ——网络复杂度, $P=A/N$ ,最小值为0.5,此时网络仅有2个站点和1条连接弧段;随站点和弧段的增加, $P$ 逐渐逼近1;当闭合网络中存在交叉连接时 $P>1$ 。

$C$ ——网络连接度,网络实际弧段除以最大可能弧段(最大可能弧段为 $3(N-2)$ );网络连接度较差,则乘客在出行过程中易发生绕行,无效出行比例增加; $C=A/(3N-6)$ 。

## 2.2 数据处理与检验

(1) 城市轨道交通线路开通运营时间修正。上海部分轨道交通线路正式开通运营时间位于年末,如3号线“上海南站—江湾镇”段与“江湾镇—江杨北路”段分别在2000年、2006年的12月开通。线路开通后短期内客流培育不完全,因此不能以开通当年的建设情况进行分析。本文将年末11、12月开通运营的线路计入下一年。

(2) 对数处理。上海轨道交通客流( $V$ )具有指数增长趋势,为降低时间序列数据中存在的异方差影响,将各项指标取对数(在字母名称前加 $\ln$ 表示)。

## 3 城市轨道交通客流协整分析

### 3.1 平稳性检验

利用EViews软件,采用ADF检验法判断序列平稳性。当 $t$ 检验值的绝对值大于5%临界值时,拒绝原假设。各变量原序列与一阶差分序列检验结果见表1。

表1 ADF法单位根检验结果

变量	原序列			一阶差分序列		
	检验形式	$t$ 检验值	5%临界值	检验形式	$t$ 检验值	5%临界值
$\ln V$	(C, T, I)	-3.69	-3.83	(C, D, I)	-3.57	-3.14
$\ln L$	(C, T, I)	-2.83	-3.93	(C, D, I)	-4.96	-3.14
$\ln N$	(C, T, I)	-2.26	-3.86	(C, D, I)	-3.79	-3.12
$\ln A$	(C, T, I)	-2.27	-3.88	(C, D, I)	-3.72	-3.12
$\ln C_r$	(C, T, I)	-0.86	-4.11	(C, D, I)	-2.31	-3.32
$\ln S$	(C, T, I)	-2.64	-3.79	(C, D, I)	-5.16	-3.12
$\ln R$	(C, T, I)	-0.73	-4.11	(C, D, I)	-2.14	-3.32
$\ln P$	(C, T, I)	-4.07	-3.93			
$\ln C$	(C, D, I)	-1.53	-3.10	(C, D, I)	-3.55	-3.12

注:检验形式的3项分别表示是否有截距项、是否有趋势项及滞后阶数;C表示有截距项,I表示有趋势项

上海轨道交通客流( $\ln V$ )原序列非平稳而一阶差分序列平稳,为一阶单整序列。只有同阶单整序列才可能存在协整关系。根据单位根检验结果,城市轨道交通网络里程( $\ln L$ )、站点数( $\ln N$ )、弧段数( $\ln A$ )、站间距( $\ln S$ )、连接度( $\ln C$ )为一阶单整序列,可能与客流( $\ln V$ )存在协整关系。

### 3.2 协整检验

利用Johansen检验变量是否与城市轨道交通客流量存在协整关系,当LR检验值的绝对值大于5%临界值时拒绝原假设。检验结果见表2。

表2 Johansen检验结果

变量	方程形式	H0	H1	$R_{LM}$ 检验值	5%临界值
$\ln L$	仅有截距项	$M=0$	$M=1$	16.28	15.49
		$M \leq 1$	$M=2$	0.07	3.84
$\ln N$	线性趋势	$M=0$	$M=1$	24.12	25.87
		$M \leq 1$	$M=2$	7.91	12.52
$\ln A$	线性趋势	$M=0$	$M=1$	23.84	25.87
		$M \leq 1$	$M=2$	7.84	12.52
$\ln S$	线性趋势	$M=0$	$M=1$	25.09	25.87
		$M \leq 1$	$M=2$	11.36	12.52
$\ln C$	线性趋势	$M=0$	$M=1$	41.90	25.87
		$M \leq 1$	$M=2$	9.78	12.52

上海城市轨道交通客流( $\ln V$ )与网络长度( $\ln L$ )、网络连接度( $\ln C$ )分别存在1个协整关系。协整方程系数标准化后,结果见表3。

表3 协整方程表达式

变量	协整方程	对数似然比
$\ln L$	$\ln V = 0.924 \ln L$	13.669
$\ln C$	$\ln V = 42.65 \ln C + 0.613 t$	56.728

1996—2010年,上海城市轨道交通网络长度每提高1%,客流量提高0.924%;连接度每提高1%,客流量提高42.65%,同时,年份每增加1年(从1997年开始),客流量再提高0.613%。网络连接度的改善比长度的增长对于吸引客流具有更好的效果。在城市轨道交通网络规划阶段需注重网络结构设计,保证网络连通性,以提高对客流的吸引力。

## 4 城市轨道交通客流增长滞后性

### 4.1 城市轨道交通发展情况

1996—2010年上海城市轨道交通客流与网络长度、连接度的变化关系如图1、图2所示。

上海城市轨道交通网络建设进入快速发展期,

长度年均增长 24.3%; 同时客流规模不断膨胀, 年均增长 26.3% (见图 1)。上海轨道交通网络连接度经历了先减小后增加的过程(见图 2) 2001 年与 2010 年网络连接度大幅提高, 主要是由于: 2000 年 12 月的 3 号线开通使得上海轨道交通网络拥有半个环形线路; 2009 年末正式开通运营的 9 号线二期以及 2010 年上半年的 7 号线一期与 10 号线在已成环的 4 号线内交织成网。环线与环内网络的形成提高了网络连接度。

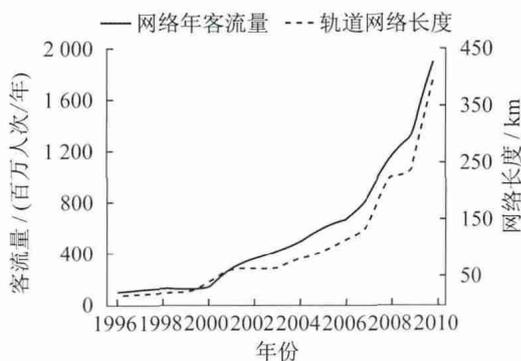


图1 城市轨道交通客流量与网络长度的关系

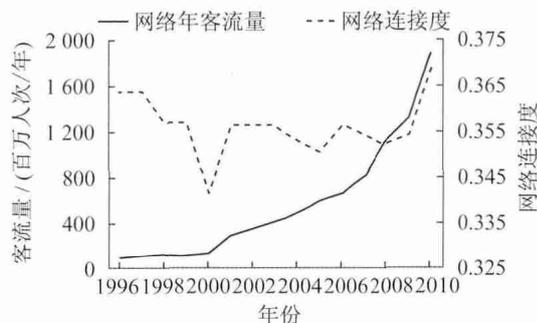


图2 城市轨道交通客流量与网络连接度的关系

#### 4.2 城市轨道交通客流增长滞后性

在上海城市轨道交通网络建设初期, 客流增长对于网络建设变化的反应在时间上存在滞后性, 如图3所示。上海城市轨道交通网长度在2000年(当年6月2号线一期开通)与2004年(2003年11月5号线开通)出现快速增长的现象, 而客流增速加快则在2001年与2005年完成, 比网络建设延后1年。因此, 在建设初期, 轨道交通运营单位应做好面对客流缓慢增长的准备, 采取适当措施控制运营成本。然而, 滞后性随网络的发展而减弱, 如2008年及以后(2007年12月6、8、9号线开通), 上海轨道交通网络长度增长率与客流增长率同步。

客流增长相对于连接度的滞后现象不明显(见图4)。2001年、2005年、2008年与2010年客流增

长率均随当年连接度增长率提高而提高。客流对于连接度的滞后性不显著, 说明连接度改善对于快速吸引客流具有较好的效果, 与协整方程分析结果一致。

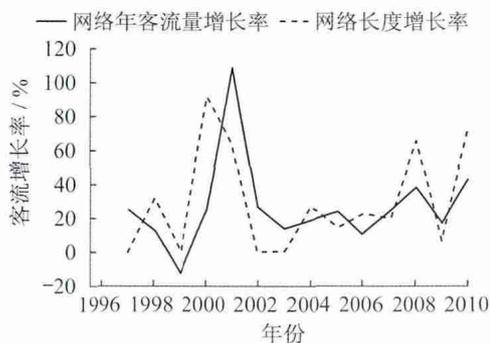


图3 城市轨道交通客流与网络长度增长率的关系

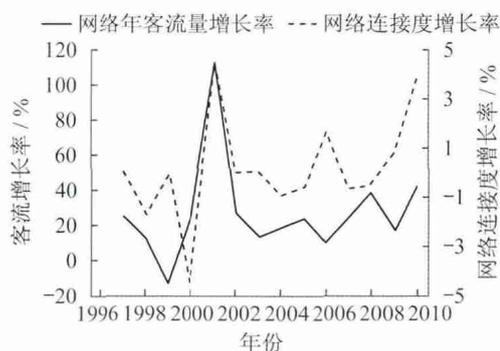


图4 城市轨道交通客流与网络连接度增长率的关系

### 5 城市轨道交通客流量向量自回归模型

#### 5.1 VAR 模型

根据协整检验结果, 客流分别与网络长度以及连接度存在协整关系, 故可以建立 VAR 模型, 滞后阶数为 2。VAR 模型特征根检验结果见表 4。

表4 VAR 模型特征根检验结果

VAR 模型	特征根模	稳定性
长度与客流	0.983 $\rho$ .746 $\rho$ .746 $\rho$ .376	稳定
连接度与客流	1.057 $\rho$ .057 $\rho$ .337 $\rho$ .337	不稳定

长度与客流的 VAR 模型特征根均小于 1, 模型稳定收敛; 而连接度与客流的 VAR 模型部分特征根大于 1, 模型不稳定收敛, 说明客流关于连接度的滞后性不显著。故仅能建立网络长度与客流 VAR 模型如下( $\ln V(-1)$ 、 $\ln V(-2)$  与  $\ln L(-1)$ 、 $\ln L(-2)$  分别表示滞后 1 年、2 年的客流量与网络长度):

$$\begin{bmatrix} \ln V \\ \ln L \end{bmatrix} = \begin{bmatrix} 6.535 \\ -0.410 \end{bmatrix} + \begin{bmatrix} 0.157 & -0.107 \\ -0.414 & 0.546 \end{bmatrix} \begin{bmatrix} \ln V(-1) \\ \ln V(-2) \end{bmatrix} + \begin{bmatrix} 0.517 & 0.350 \\ 0.873 & -0.016 \end{bmatrix} \begin{bmatrix} \ln L(-1) \\ \ln L(-2) \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (4)$$

表5为城市轨道交通客流VAR模型检验统计量。

表5 城市轨道交通客流VAR模型检验统计量

	lnV	lnL
调整R方	0.959	0.892
残差平方和	0.271	0.851
标准误差	0.184	0.326
F统计量	71.365	25.904
对数似然比	6.694	-0.731
因变量均值	10.718	4.533
因变量标准差	0.911	0.995

## 5.2 客流预测

VAR模型预测分为静态预测和动态预测。静态预测各期均使用实际值,适合近期预测,准确度较高;动态预测只有初值使用实际值,其后各期采用递推计算,适合远期趋势预测。样本数据范围为1996—2010年,利用式(4)对2011年上海城市轨道交通客流量进行预测,将lnV转换为实际值V,结果如图5所示。

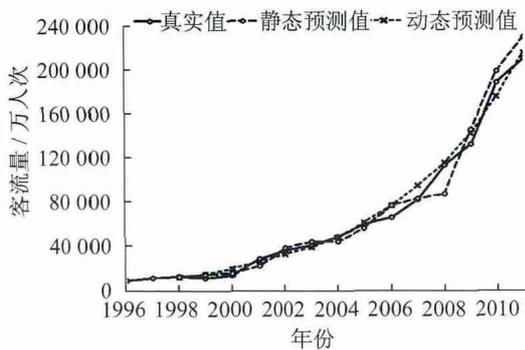


图5 上海城市轨道交通客流真实值与VAR模型预测结果比较

2011年上海城市轨道交通客流量静态预测结果为22.98亿人次,与实际值21.01亿人次的误差为9.4%。在图5中,动态预测值的曲线平滑,能体现客流快速增长的趋势。

城市轨道交通客流预测误差如图6所示。在图6中,各期静态预测与动态预测的误差都在0值上下波动,模型稳定。2007年后,静态预测与动态预

测的各期误差开始增加,主要原因是城市轨道交通发展导致客流滞后性减弱,VAR模型滞后项影响降低,预测误差增大。

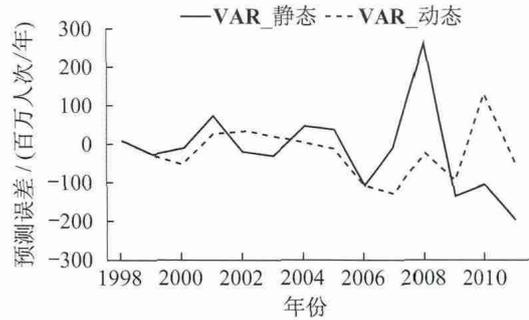


图6 城市轨道交通客流预测误差

## 5.3 方法验证

为判断上述城市轨道交通客流预测方法是否具有推广型,使用北京轨道交通客流数据进行验证。北京轨道交通于1981年正式对公众开放,现选取1981—2010年的客流数据与网络建设情况进行分析。按照2.2节的方法进行数据处理,经过单位根检验与协整检验,北京轨道交通客流与网络里程为一阶单整序列,具备一个协整关系。VAR模型所有特征根绝对值小于1,模型(见式5)稳定。

$$\begin{bmatrix} \ln V \\ \ln L \end{bmatrix} = \begin{bmatrix} 1.352 \\ 0.206 \end{bmatrix} + \begin{bmatrix} 0.649 & 0.133 \\ -0.236 & 0.168 \end{bmatrix} \begin{bmatrix} \ln V(-1) \\ \ln V(-2) \end{bmatrix} + \begin{bmatrix} 0.271 & -0.006 \\ 1.117 & 0.038 \end{bmatrix} \begin{bmatrix} \ln L(-1) \\ \ln L(-2) \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (5)$$

表6为北京城市轨道交通客流VAR模型检验统计量。

表6 北京城市轨道交通客流VAR模型检验统计量

	lnV	lnL
调整R方	0.957	0.949
残差平方和	0.496	0.601
标准误差	0.147	0.162
F统计量	151.073	106.055
对数似然比	16.743	14.048
因变量均值	10.656	4.128
因变量标准差	0.708	0.658

2011年北京轨道交通客流量静态预测结果为23.14亿人次,与实际值21.93亿人次相差6.0%。动态预测结果能反映北京轨道交通客流快速增长的变化趋势(见图7)。上述预测结果符合实际情况,可见本文提出的方法可适用于上海以外的城市,具有推广意义。

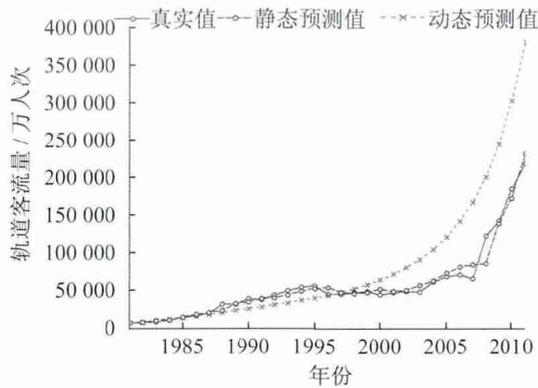


图 7 北京轨道交通客流真实值与预测结果比较

## 6 结语

本文针对上海城市轨道交通客流增长特征,以协整理论为基础,分析了城市轨道交通客流关于网络规模与结构的增长滞后性,并利用向量自回归模型进行客流预测,主要结论如下:

(1) 上海城市轨道交通客流与网络长度以及连接度存在协整关系。网络长度与连接度的提高均能促进客流的增长,其中连接度的改善更能吸引客流。在城市轨道交通网络规划过程中应注重网络的连通性。

(2) 城市轨道交通客流增长对于网络长度存在显著的滞后性。在城市轨道交通建设初期,新开通的线路并不能马上吸引大量客流,需关注该阶段运营成本,但是这种滞后性会随着城市轨道交通的完善而逐渐减弱。

(3) 向量自回归(VAR)模型能体现城市轨道交通建设对于客流增长的滞后影响,并能进行客流

预测。2011年上海与北京轨道交通客流静态预测误差分别为9.4%和6.0%,动态预测能体现客流增长趋势。预测方法具有推广性。然而,由于滞后影响的减弱,针对城市轨道交通不同建设时期,建立滞后后期变化的客流预测模型是下一阶段的重点工作。

## 参考文献

- [1] 孟迎春. 我国城市轨道交通规模研究[D]. 北京: 北京交通大学, 2009.
- [2] 吴娇蓉, 陈小鸿. 轨道网络形成期轨道交通客流培育与聚集策略研究[J]. 武汉理工大学学报: 交通科学与工程版, 2009(2): 211.
- [3] Granger C W J. Developments in the study of co-integrated economic variables [J]. Oxford Bulletin of Economics and Statistics, 1986(48): 226.
- [4] Ramanatha R. The long-run behaviour of transport performance in India: a co-integration approach [J]. Transportation Research Part A 2001(35): 309.
- [5] Pitfield D E. The southwest effect: a timeseries analysis on passengers carried by selected routes and a market share comparison [J]. Journal of Air Transport Management, 2008(14): 113.
- [6] 陈必壮, 沈云樟, 王忠强, 等. 上海市轨道交通客流预测理论方法研究[J]. 城市轨道交通, 2009, 1(1): 13.
- [7] 朱霞, 毕艳祥. 上海轨道交通网络化对客流的影响研究[J]. 城市轨道交通研究, 2010, 13(3): 8.
- [8] 刘颖杜. 广州市轨道交通网络客流特征及成长规律[J]. 交通运输系统工程与信息, 2012(1): 205.

(收稿日期: 2013-08-08)

(上接第 21 页)

- [8] Zhang J, Xu X, Hong L, et al. Networked analysis of Shanghai subway network in China [J]. Physica A, 2011, 390: 4562.
- [9] Freeman L C. Centrality in social networks: conceptual classification [J]. Social Networks, 1978, 1(1): 215.
- [10] Tu Y. Centrality Characteristics Analysis of Urban Rail Network [C] // 2013 IEEE International Conference on Intelligent Rail

Transportation. Beijing: IEEE, 2013.

- [11] 丁小兵. 复杂网络理论及其在上海城市轨道交通网络可靠性分析评价中的应用[J]. 城市轨道交通研究, 2012(11): 50.

(收稿日期: 2013-04-08)

欢迎订阅《城市轨道交通研究》

服务热线 021—51030704