

文章编号: 2095-2163(2021)04-0135-04

中图分类号: U293.5

文献标志码: A

# 基于 ARIMA 模型的城市轨道交通客流预测及研究

倪 杰, 于 莉, 靳笑楠

(上海工程技术大学 城市轨道交通学院, 上海 201620)

**摘 要:** 在城市轨道交通新线开通运营后期, 客流趋于稳定的情况下, 利用改进的时间序列模型 ARIMA (Autoregressive Integrated Moving Average Model), 对新线开通后期的站点进出站客流进行预测。为增加预测精度, 利用小波分析方法对客流进行去噪, 同时利用差分的方法消除数据的趋势性和季节性。将预测进站客流值与实际客流量对比可知, 平均误差约为 3.708%, 预测效果较好, 具有很强的应用性。

**关键词:** 城市轨道交通; ARIMA 模型; 小波去噪; 客流预测

## Forecast and research of urban rail transit passenger flow based on ARIMA model

NI Jie, YU Li, JIN Xiaonan

(College of Urban Rail Transit, Shanghai University Of Engineering Science, Shanghai 201620, China)

**[Abstract]** When the passenger flow of new urban rail transit line tends to be stable in the later period of operation, the improved time series model ARIMA (autoregressive integrated moving average model) is used to predict the passenger flow in and out of the new line. In order to increase the prediction accuracy, the wavelet analysis method is used to de noise the passenger flow, and the difference method is used to eliminate the trend and seasonality of the data. Compared with the actual passenger flow, the average error is about 3.708%.

**[Key words]** Urban rail transit; ARIMA model; wavelet denoising; passenger flow prediction

## 0 引 言

近年来,随着城市轨道交通行业的高速发展,由于新线开通运营后期,站点周边公交配套设施逐渐完备,乘客对于新线已经较为熟悉,且原有的客流出行习惯也会由于新线所提供的便利性而得到改变。此时站点的进出站客流波动趋于稳定,因此可以利用历史数据进行客流预测。

目前,众多学者针对不同场景提出不同的预测方法。郭文雅等<sup>[1]</sup> 比对武汉轨道交通 3 号线开通前后,线网及各线路客流变化情况,定量分析了新线开通对线网及各线路客流的影响,提出了提高线网客流的几项建议和措施。杨德明等<sup>[2]</sup> 对深圳地铁新开通线路进行观察,结合网络后台客流数据分析,找出网络中存在较大问题的车站,掌握新线开通客流变化规律。朱霞等<sup>[3]</sup> 从上海轨道交通运营线路的客流变化入手,探讨网络化进程对客流变化的影响,分析网络化进程中客流变化的原因。郑清杰<sup>[4]</sup> 在分析历史新线接入后客流变化情况,基于土地利用性质对城市轨道交通车站进行 K 聚类,基于聚类结果及新线车站的评估,

进行新线车站进出站量的预测。

考虑到不同重要度等级下的站点客流特性有较大差异<sup>[5]</sup>, 时间序列模型的优势在于能够较好的挖掘出数值变化背后的规律性。因此,本文采用时间序列模型进行站点进出站客流的预测,并利用小波去噪法,针对客流信号中不规则波动的噪点进行去除,提高预测精度。

## 1 客流预测时序模型建立

时间序列模型能够对一个时间序列进行时频两方面的分析。时域分析认为当前信号组成的时间序列是由过去的状态以及相关的外部因素(噪声、干扰等)组成的,即将来的数据能够由之前的时序数据预测而出;频域分析则认为时间序列可以通过不同周期振幅下的正弦波相加而得,即将正弦波通过平移拉伸产生一系列不同周期及振幅的波系来对一个信号进行表示<sup>[6]</sup>。

本文研究内容为城市轨道交通客流数据,该序列受多种因素的影响,呈现出时序的趋势及特征。因此在对客流信号进行分析时主要还是关注于时间

**作者简介:** 倪 杰(1996-),男,硕士研究生,主要研究方向:城市轨道交通运营与管理;于 莉(1994-),女,硕士研究生,主要研究方向:城市轨道交通运营与管理。

收稿日期: 2020-06-23

哈尔滨工业大学主办 ◆ 系统开发与应用

域,以达到预测的目的。而频率域的分析,则集中于对客流信号的去噪,以及不同频率成分的分析。由于本文所采用的数据为地铁日进出站客流量,且数据时间跨度较大,故而季节性与趋势性并存,因此采用  $ARIMA(p, d, q)$  ( $P, D, Q$ )<sup>S</sup> 进行预测模型的建立。ARIMA 方法的模型运用主要包括:模型平稳性分析、模型阶数判定以及模型估计。利用季节时间序列模型即可预测出各等级下的站点新线开通后期的进出站客流。

为了验证方法的可行性,在站点客流预测前,优先针对线网全日总客流进行预测尝试。图1即为利用2006年1月~11月的全日客流数据,预测2016年12月的预测值与实际值之间的对比图。

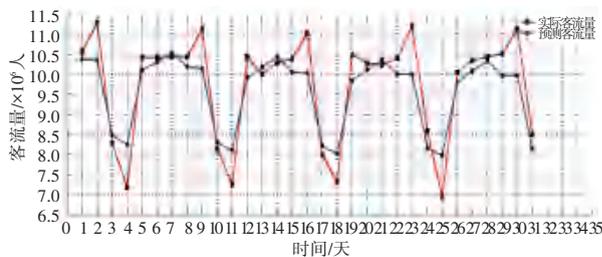


图1 全日总客流时间序列预测值与实际值对比图

Fig. 1 Comparison of predicted and actual values of total daily passenger flow

由此可见,总体预测精度较高,能够较好地保留客流变化的趋势,证明了预测方法的可用性。

## 2 预测实例研究

时间序列模型即为利用站点历史客流数据的变化对新线开通后期的站点客流量进行预测。由于不同线路所在线网中位置走向的差异,则在进行实际预测时,需要对不同类别线路的站点分别进行建模,以提高预测精度。本文选取上海地铁9号线杨高中路站作为实例,进行换乘站点进站客流预测模型的建立与演示。

### 2.1 数据处理与分析

预测采用2015年1月1日~2018年6月24日的上海地铁9号线杨高中路站的全日进站客流数据,利用Matlab绘制该站进站客流的总体趋势如图2所示。

从图中清晰可见客流变化趋势,春节期间,客流量降至全年的低谷,各大节假日,客流量也会有较大的降低。此外还有少量数据存在着较大幅度的波动,若将这些异常波动点纳入时间序列模型中,则会降低预测的精度及模型的拟合效果,因此需要利用小波分析方法进行去噪。

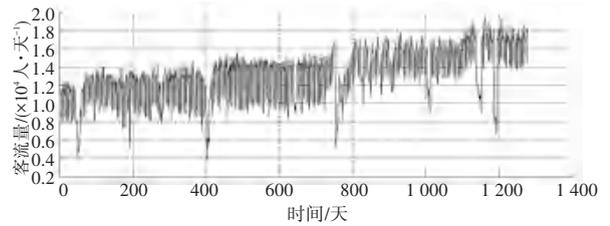


图2 杨高中路站进站客流总体趋势图

Fig. 2 The overall trend of passenger flow in Middle Yanggao Road Station

(1)借助Matlab对客流数据进行db9小波基下的3层分解,编程求取阈值(约为3745.3)。将分解后的低频信号进行软阈值去噪,对去噪后的信号进行重构,可得新的客流趋势如图3所示。

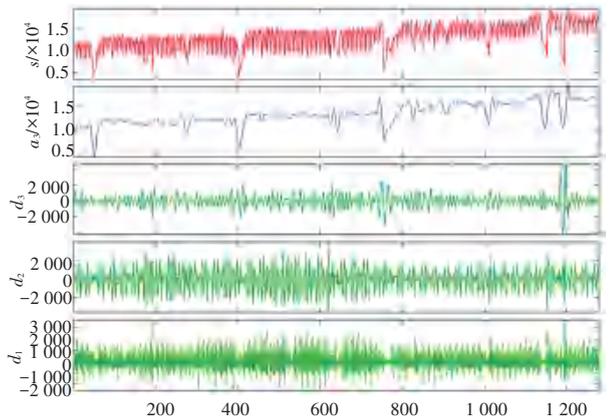


图3 三层分解后的客流高低频信号

Fig. 3 High and low frequency signals of Three-layer decomposition of passenger flow

由此可见,经过小波去噪后的客流信号,相对较为光滑且将客流波动变化的特征尖峰很好地进行了保留。去噪后的进站客流总体趋势如图4所示。

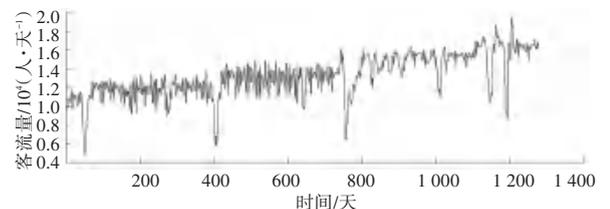


图4 去噪后的进站客流总体趋势图

Fig. 4 Overall trend of inbound passenger flow after denoising

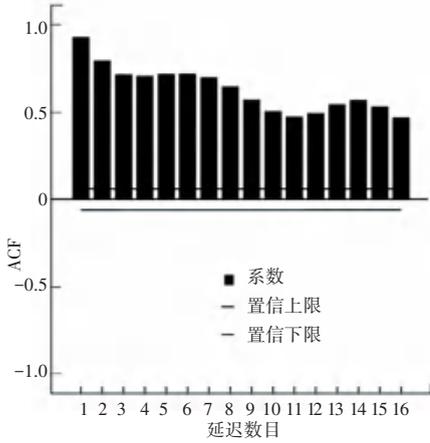
(2)利用去噪后的客流数据进行时间序列模型的构建,首先绘制自相关图和偏自相关图,如图5所示。

显而易见,两图皆未呈现拖尾或是截尾趋势,因此序列非平稳。观察偏自相关图中,数据在前三阶显著突出,但随着阶数的增长,数据在零水平线上下波动未显著趋于零。因此可知,序列趋势性与季节性并存。

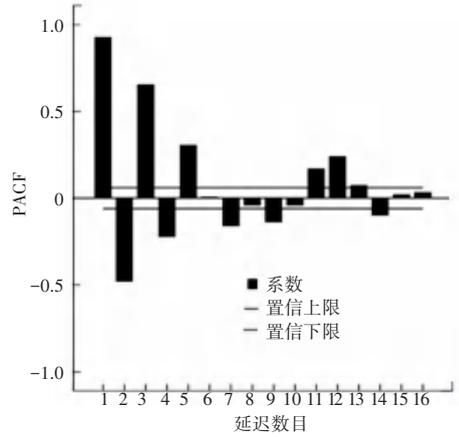
(3)为了消除数据的趋势性和季节性,拟采用一阶差分和一阶季节差分对数据进行差分。经增加

阶数进行差分以及对数变换等尝试,效果并不明显。因此,为了不过分丢失原始数据的特性,最终选择一阶差分。差分结果如图 6 所示。

经差分后数据的自相关图、偏自相关图,如图 7 所示。数据的趋势性得到了一定的消除,但仍有少许季节性波动。



(a) 自相关  
(a) Autocorrelation



(b) 偏自相关  
(b) Partial autocorrelation

图 5 去噪后客流数据的自相关图和偏自相关图

Fig. 5 Autocorrelation and partial autocorrelation of passenger flow data after denoising

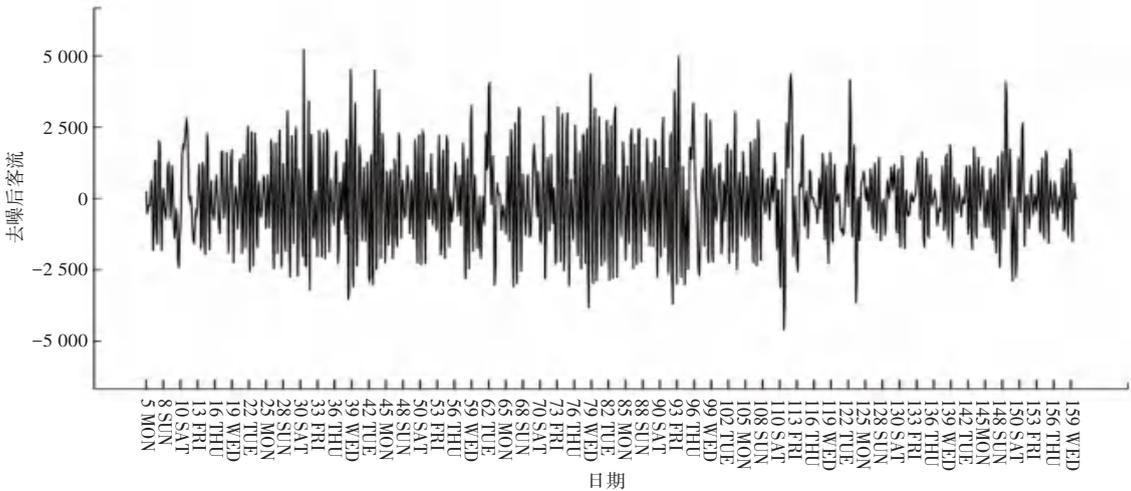
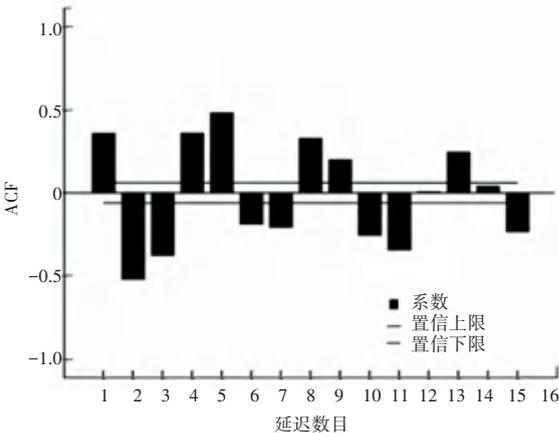
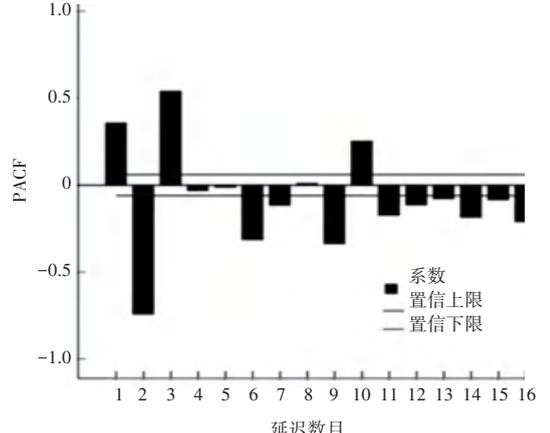


图 6 客流信号的一阶差分图

Fig. 6 First order difference in the passenger flow



(a) 自相关  
(a) Autocorrelation



(b) 偏自相关  
(b) Partial autocorrelation

图 7 差分后数据的自相关图和偏自相关图

Fig. 7 Autocorrelation plot and partial autocorrelation plot of differential data

### 2.2 模型构建及预测

考虑到数据既存在着趋势性又存在季节性,在构建季节时间序列模型时,根据图7的拖尾及截尾情况,尝试建立  $ARIMA(5,1,3)(1,0,1)^s$  模型,利用SPSS进行模型参数的拟合,实际拟合结果见表1。

由模型拟合结果可知,平稳的R方和R方都达到0.858,Ljung-Box Q(18)(杨博克斯统计量)的显著性为0.713,说明序列已经达到平稳,消除了自相

关性,而且正态化贝叶斯统计值也达到13.924,可见模型总体拟合效果良好,可用性较高。

利用所建模型对一周内的进站客流进行预测(2018年6月25~7月1日数据),预测出的客流量见表2,客流预测值的拟合效果如图8所示。

预测进站客流值与实际客流量作对比结果见表3。求得平均误差约为3.708%,可见预测效果较好,能够达到一定的运用程度。

表1 模型拟合效果评价表

Tab. 1 Model fitting effect evaluation

模型	预测变量	模型拟合统计量			Ljung-Box Q(18)		
		平稳的R方	R方	BIC	统计量	DF	Sig.
客流-模型	0	0.858	0.858	13.924	5.408	8	0.713

表2 模型预测结果

Tab. 2 Model prediction result

模型		星期一	星期二	星期三	星期四	星期五	星期六	星期日
客流-模型	预测	14 710	17 807	17 021	16 914	19 248	17 547	17 253
	UCL	16 191	19 338	19 533	19 868	20 673	20 948	18 514
	LCL	12 228	14 276	13 940	13 959	15 420	14 548	12 991

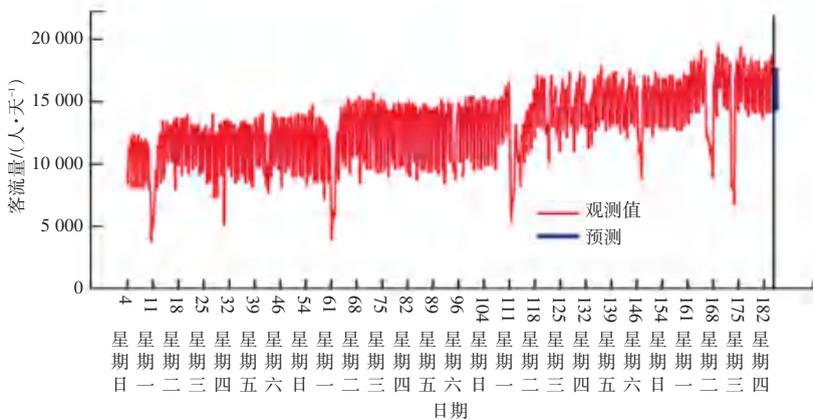


图8 客流预测值拟合图

Fig. 8 Passenger flow prediction value fitting map

表3 预测结果比较

Tab. 3 Comparison of prediction results

实际客流	预测客流	客流差值	客流误差/%
14 415	14 710	-295	-2.05
16 880	17 807	-927	-5.49
16 863	17 021	-158	-0.94
17 247	16 914	333	1.93
18 291	19 248	-957	-5.23
16 851	17 547	-696	-4.13
16 248	17 253	-1 005	-6.19

(约为3.708%),证明该方法具有一定的实用性和工程应用性,可以应用于新开线路后期的客流预测研究及工程应用。

### 参考文献

- [1] 郭文雅,王泽清. 武汉轨道交通循礼门站实时换乘客流研究[J]. 武汉工程职业技术学院学报, 2016, 28(2): 26-28,46.
- [2] 杨德明,杨丽,杨良,等. 深圳地铁三期新线开通客流调查分析[J]. 现代城市轨道交通, 2017, 1(5):44-48.
- [3] 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): 8-12,16.
- [4] 郑清杰. 城市轨道交通新线接入客流预测及网络客流分布研究[D]. 北京: 北京交通大学,2016.
- [5] 程涛,周峰,郦海通,等. 西安地铁2号线南段运营初期客流预测[J]. 都市快轨交通, 2015, 28(5): 45-49.
- [6] 白丽. 城市轨道交通常态与非常态短期客流预测方法研究[J]. 交通运输系统工程与信息, 2017, 17(1): 127-135.

### 3 结束语

本文以2015年~2018年上海地铁客流数据为基础,以轨道交通市区段新线开通后的进出站客流量为预测对象,预测城市轨道交通新线开通后期的客运量。在进行时序建模时需要区段的站点分别进行建模,对杨高中路站进行了基于小波去噪的时间序列模型预测的实例分析,预测平均误差较小