

文章编号: 2095-2163(2020)07-0273-06

中图分类号: U491

文献标志码: A

# 基于 ARIMA 的短时交通量预测模型

张腾飞, 袁鹏程

(上海理工大学 管理学院, 上海 200093)

**摘要:** 为了对交通控制与诱导系统提供更可靠的数据,有必要对实时、动态的数据进行短时交通量在线预测。本文对短时交通量预测的 ARIMA 建模方法进行了介绍,用 2008 年 10 月 7 日以 3min 为单元的短时交通量数据作为训练数据,将其平滑化处理并标定模型为 ARIMA(5,1,6),采用最大似然估计法估计模型的参数,残差白噪声检验验证其有效性,用静态预测法对 8~10 日三天的数据进行实时动态预测,结果表明 ARIMA 模型在短时交通量预测方面可移植性较强,可以在历史数据较少的情况下准确的捕捉其内在规律,并对大量数据进行静态预测,可以在线、实时、动态预测下一时刻的数据,给交通管理系统提供更可靠的实时数据。

**关键词:** 短时交通量; ARIMA; 静态预测; 实时预测

## Short-term traffic flow forecast based on ARIMA model

ZHANG Tengfei, YUAN Pengcheng

(1 Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

**[Abstract]** In order to provide more reliable data for traffic control and guidance system, it is necessary to carry out short-term traffic volume online prediction for real-time and dynamic data. Firstly, ARIMA modeling method for short-term traffic volume prediction is introduced. The short-term traffic volume data on October 7, 2008 in 3min units is used as training data and smooth it out and calibrated as ARIMA(5,1,6). Then the maximum likelihood estimation method is used to estimate the parameters of the model, which use the white noise test of residual error to verify its effectiveness. Later the static prediction method is used to carry out real-time dynamic prediction on the data on the 8<sup>th</sup>, 9<sup>th</sup> and 10<sup>th</sup>. The results show that ARIMA model has strong portability in short-term traffic volume prediction and ARIMA model can accurately capture its inherent laws and carry out static prediction on a large amount of data under the condition of less historical data. It can also predict the data at the next moment in real time and dynamically, and provide more reliable real-time data for traffic management system.

**[Key words]** Short-term traffic volume; ARIMA; Static prediction; Real-time prediction

## 0 引言

随着社会的发展与人们生活水平的提高,小汽车保有量越来越高,交通变得愈加拥堵,现在道路资源的建设已不能满足快速增长的交通需求,许多研究学者在实践中发现由于缺乏对交通的实时动态管理,交通控制与诱导系统并不能使道路交通系统充分发挥其现有价值<sup>[1]</sup>。交通控制与交通诱导系统的前提就是做好短时交通量预测和交通信息采集等,来提高智能交通系统的运行效率<sup>[2]</sup>。做好交通控制和诱导系统必须要做好短时交通量预测,短时交通量预测是以几分钟为数据单元进行时间序列的分析预测,滚动实时预测未来短时段交通量。近年来,大数据与智能交通结合的愈加紧密,交通数据的采集也愈加便捷,大量的不间断的实时数据,便于

进行实时动态的交通量预测,可以为交通的实时动态管理提供更好的依据<sup>[3]</sup>。道路交通量的变化主要受前段时间交通量的影响以及该路段下游交通量的影响,许多外界的干扰都会对交通量产生一些影响,当然一些突发事件等特定因素都会影响交通状态。日内短时交通量呈现出早高峰以及晚高峰特征,日内早上 6 时至晚 22 时的交通流占比很大,虽然白天的交通量没有多大波动,但是也存在早、晚两个峰值,分别是上午 9 时和晚 18 时左右,所以其变化趋势呈 M 型或马鞍型<sup>[4]</sup>。短时交通量数据是以几分钟为单元的数据(一般不超过 15 分钟),由于统计的时间短,这些数据在时间序列上呈现出不确定性、非平稳性、非线性随机过程,并不能很好的呈现出一定的规律性<sup>[5]</sup>。

**基金项目:** 国家自然科学基金(7160118);上海理工大学大学生创新项目(XJ2019135,XJ2019144)。

**作者简介:** 张腾飞(1994-),男,硕士研究生,主要研究方向:交通流预测;袁鹏程(1982-),男,博士,副教授,硕士生导师,主要研究方向:交通行为、数据统计与分析。

**通讯作者:** 袁鹏程 Email: danis\_cx@126.com

收稿日期: 2020-01-14

近几十年,许多研究学者在研究短时交通量预测方面取得了很多成就,各学科的研究学者在短时交通量预测方向开发了一系列的短时预测模型,如线性回归模型、人工神经网络模型、时间序列模型、卡尔曼滤波模型、非参数回归模型、小波理论预测模型。因为短时交通量数据具有一定的不确定性、非平稳性及非线性随机过程,所以要选择具备可移植性、准确性、实时性、自适应性和鲁棒性的短时交通量预测方法<sup>[6]</sup>。ARIMA模型是由美国统计学家Box和英国统计学家Jenkins在20世纪70年代提出的时间序列分析模型,即差分自回归移动平均模型(ARIMA),用该模型做的时间序列预测也称博克斯-詹金斯法(Box-Jenkins)。ARIMA是时间序列预测的主要方法之一,主要依靠研究对象的历史时间序列数据,从历史数据自身规律分析该数据并建立合适的模型,对未来数据进行预测。1979年Mohamed S和Ahmed首次在交通量预测领域提出了时间序列模型<sup>[7]</sup>;2001年万建强和文洲在股票指数方面将ARIMA模型与ARCH模型对比,得出ARIMA模型的预测精度受时间序列数据的不确定性的影响<sup>[8]</sup>;2005年石美娟将ARIMA模型应用于投资预测中,得出较好的预测结果,但随着预测步数的增加,会导致预测结果误差较大<sup>[9]</sup>;2006年姚亚夫和曹锋将ARIMA模型应用于短时交通量预测,通过动态预测向前多步预测,发现随着预测步数的增加误差递增<sup>[10]</sup>;2010年芮少权和匡安乐将ARIMA模型应用于长时交通量预测,结果表明ARIMA模型能较好的适应于高速公路月度交通量预测<sup>[11]</sup>;2017年崔伟杰将ARIMA模型应用于安全事故死亡人数的预测,得出ARIMA模型在预测时更适合进行短时预测,其短时预测较长时,预测有着更高的预测精度<sup>[12]</sup>。

通过分析近年来ARIMA模型的应用可以看出,学者们在短时交通量预测领域应用ARIMA模型还不多,ARIMA模型在短时交通量预测方面还可以进一步完善以及深入研究。本文在考虑历史数据较少的基础上,通过建立ARIMA模型进行短时交通量的静态预测,避免多步预测误差递增的问题,采用常州市博爱路2008年10月7日的短时交通量数据(全天480条交通量数据统计,时间间隔为3分钟),对ARIMA模型标定建立预测模型,对8、9、10日的交通量进行静态预测。结果表明,所建立的ARIMA模型可以较好地对短时交通量进行短时预测。

## 1 ARIMA模型

### 1.1 ARIMA理论基础

ARMA模型由自回归AR(Auto Regressive)模型和移动平均MA(Moving Average)模型两部分组成。滞后项为 $p$ 阶的AR模型记为 $AR(p)$ , $AR(p)$ 模型如式(1)所示:

$$x_t = c + a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_p x_{t-p} + \varepsilon_t. \quad (1)$$

其中, $c$ 为常数; $a_1, a_2, \dots, a_p$ 是 $AR(p)$ 模型的自回归系数; $\varepsilon_t$ 为白噪声序列,即随机误差项; $x_t$ 是时间序列 $t$ 时刻的值,即因变量; $x_{t-1} x_{t-2} \dots x_{t-p}$ 是时间序列 $x_t$ 的滞后序列,即自变量。滞后项为 $q$ 阶的MA模型记为 $MA(q)$ , $MA(q)$ 如式(2)所示:

$$x_t = c + \varepsilon_t - b_1 \varepsilon_{t-1} - b_2 \varepsilon_{t-2} - \dots - b_q \varepsilon_{t-q}. \quad (2)$$

其中, $c$ 为常数; $b_1, b_2, \dots, b_q$ 是 $MA(q)$ 模型的自回归系数; $\varepsilon_t$ 为白噪声序列,即随机误差项; $x_t$ 是时间序列 $t$ 时刻的值,即因变量。

AR模型可以刻画一个时间序列对过去自身的记忆,MA模型可以刻画一个时间序列对过去冲击的记忆。如果一个时间序列的变化规律中不仅包含了对过去状态的记忆,还包含了对过去冲击的记忆,则此时间序列的变化规律可以被ARMA模型捕获,滞后项分别为 $(p, q)$ 的ARMA模型,可记为 $ARMA(p, q)$ , $ARMA(p, q)$ 模型如式(3)所示:

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_p x_{t-p} + \varepsilon_t - b_1 \varepsilon_{t-1} - \dots - b_q \varepsilon_{t-q}, \quad (3)$$

$$E(\varepsilon_t) = 0, \quad (4)$$

$$Var(\varepsilon_t) = \sigma_t^2, \quad (5)$$

$$E(\varepsilon_t \varepsilon_s) = 0, s \neq t, \quad (6)$$

$$E x_s \varepsilon_t = 0, \forall s < t. \quad (7)$$

其中, $x_t$ 是 $t$ 时刻样本序列值; $\varepsilon_t$ 是服从高斯分布的白噪声序列; $a_1, a_2 \dots a_p$ 是自回归系数; $b_1, b_2 \dots b_q$ 是移动平均系数;式(4)~(7)是式(3)中变量约束条件。在ARMA模型中若 $q$ 为0,则该模型 $ARMA(p, 0)$ 退化为 $AR(p)$ 。若 $p$ 为0,则该模型 $ARMA(0, q)$ 退化为 $MA(q)$ 。

若一个时间序列是非平稳的,则不能直接用ARMA模型进行拟合(容易出现伪回归),此时必须对此时间序列进行平稳化处理。差分方法是平稳化时间序列最常用的一种方法,若一个非平稳时间序列经过 $d$ 阶差分后平稳,则依据此时间序列构建的 $ARMA(p, q)$ 模型称为 $ARMA(p, d, q)$ 模型。

### 1.2 ARIMA模型的构建

ARIMA标准建模步骤:

(1)白噪声检验。白噪声序列规律性难以捕捉,不适合应用于ARMA模型中进行预测,非白噪声序列有很好的规律性,可以更好的建立模型,白噪声检验可以通过Q统计量检验以及自相关函数和偏自相关函数判断是否为白噪声序列。

(2)序列的平稳性检验以及平稳化处理。序列的平稳性通常用单位根检验和序列相关图法判断,如果序列是非平稳时间序列数据,则需要进行 $d$ 差分变换,使得序列满足建立时间序列模型平稳可逆的要求;如果序列为平稳性序列,则可以构建模型。

(3)模型识别与定阶。对平稳时间序列或者差分变换后的平稳序列识别,通过自相关系数图以及偏自相关系数图的截尾性和拖尾性识别,采用AR、MA、ARMA的其中一个模型<sup>[13]</sup>。模型定阶可以参考自相关函数图与偏自相关函数图,初步确定模型的 $p, d, q$ 的值。 $d$ 是非平稳时间序列经过 $d$ 次差分转化为平稳时间序列。若时间序列的自相关函数拖尾,而偏自相关函数截尾,则选用AR( $p$ );若时间序列的偏自相关函数拖尾,而自相关函数 $q$ 步后截尾,则选用MA( $q$ );若时间序列的自相关函数 $q$ 步后拖尾,而偏自相关函数 $p$ 步后截尾,则选用ARMA( $p, q$ ),由此可以初步判断模型的阶数<sup>[14]</sup>。选择多组( $p, q$ )值对模型的不同阶数进行多次试验,最后通过AIC、SC准则,确定最优( $p, q$ )的阶数<sup>[15]</sup>。

(4)参数估计与检验。模型参数估计可以选择最大似然估计法、矩估计法、最小二乘法。用拟合优度 $R^2$ 分析模型的拟合效果,以估计最佳参数。模型的有效性检验是验证残差是否为白噪声序列,若残差为白噪声序列,则表明原序列中所有相关信息基本上已经被模型提取出来,否则需要进一步对模型进行优化,可以通过相关图法和残差Q检验等方法对残差进行白噪声检验。若模型有效,则对模型的参数进行t统计量的显著性检验。

(5)模型预测。模型预测分静态预测和动态预测。静态预测指每一步都采用序列的真实值向前一步预测;动态预测除了第一步是利用序列的真实值向前一步预测,之后每一步都采用预测值向前预测。

## 2 模型构建

### 2.1 数据

本文采用的数据为常州市博爱路的真实交通量数据,数据由自动检测器获得,采用2008年10月7日至2008年10月10日连续4天的短时交通量数据,数据统计时间间隔为3 min,即3 min内通过道路某断面的交通量 $Q$ 。将该数据的第一天( $Q_7$ )作

为训练数据,后三天( $Q_8, Q_9, Q_{10}$ )作为预测数据。用 $Q_7$ 对ARIMA模型标定,建立预测模型,并在此基础上预测 $Q_8, Q_9, Q_{10}$ ,将预测值与真实值对比分析。

在预测之前,首先对样本数据进行描述性统计分析。从交通量时间规律性角度上分析,从图1~图4可以看出,从时间上看,交通量数据走势呈现早高峰以及晚高峰特性,这与日内短时交通量特性相一致。日内短时交通量在早7时之前流量较小,7时之后数据骤变,全天交通流量集中在早7时~晚6时之间,曲线的走势显示出短时交通流的不确定性、突变性、以及非线性特征。基本统计特征见表1。

表1 基本统计特征

Tab. 1 Basic statistical characteristics

	样本均值	最大值	最小值	标准差
7日	34.45	78	0	22.00
8日	34.01	78	0	21.63
9日	34.59	77	0	22.12
10日	35.49	88	0	22.23

对比发现10日的时间序列数据的标准差较大,数据波动较大,数据较为离散。

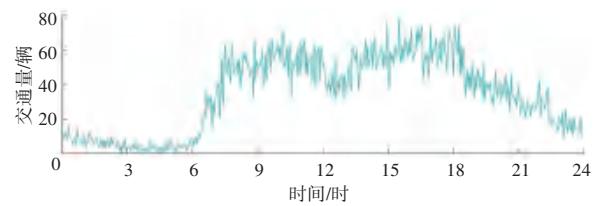


图1 7日交通量

Fig. 1 Traffic volume on the 7<sup>th</sup>

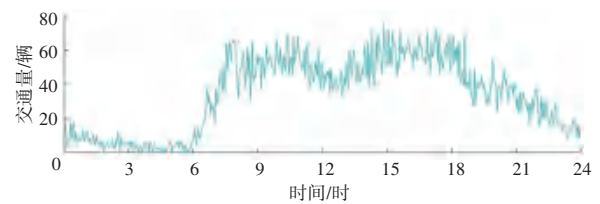


图2 8日交通量

Fig. 2 Traffic volume on the 8<sup>th</sup>

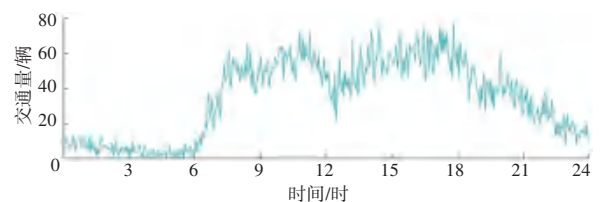


图3 9日交通量

Fig. 3 Traffic volume on the 9<sup>th</sup>

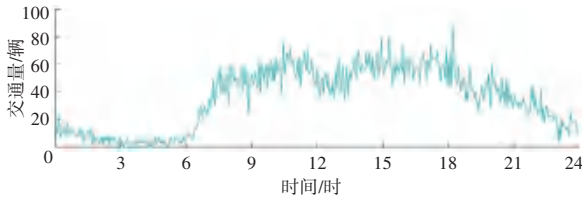


图4 10日交通量

Fig. 4 Traffic volume on the 10<sup>th</sup>

2.2 序列平稳性检验以及平稳化处理

平稳的序列在图像上一般显示出围绕其均值上下波动的特性,而非平稳序列一般在不同的时刻有不同的趋势,不同时刻有不同的均值。从图1可以看出7日交通量数据随着时间的变化有一定的规律性,且符合交通量的时间规律性特征,可以初步判断该数据为非平稳数据。为了更精准的检测该数据的

平稳性,本文采用单位根检验时间序列的平稳性。单位根检验分别测试含趋势项和截距项、仅有截距项、无趋势项和截距项,结果见表2,三种情况下时间序列  $Q$  的  $t$  统计量值在10%显著水平下均接受原假设存在单位根,所以原序列不平稳,进而对原序列进行平稳化处理,消除其非平稳特征。首先,对其进行一阶差分处理;其次,检验平稳性,差分后的序列为  $dQ$ 。结果表明差分后的序列 ADF 检验的  $t$  统计量值小于1%显著水平下的值,所以拒绝原假设,即一阶差分后的序列表现出平稳性,  $dQ$  如图5所示,符合 ARMA 模型对时间序列数据平稳可逆的要求。因此在 ARIMA ( $p, d, q$ ) 中首先确定  $d=1, p$  和  $q$  的阶数通过相关图以及 AIC 和 SC 准则确定。

表2 ADF 单位根检验结果

Tab. 2 ADF unit root test results

	Trend and intercept		Intecept		None	
	t-Statistic	Prob	t-Statistic	Prob	t-Statistic	Prob
$Q$	-2.151 6	0.516 0	-2.240 1	0.192 2	-1.123 3	0.238 1
$dQ$	-27.688 7	0				
Test	-3.962 8	1%level	-3.433 6	1%level	-2.566 1	1%level
Critical	-3.412 2	5%level	-2.862 8	5%level	-1.941 0	5%level
values	-3.128 0	10%level	-2.567 5	10%level	-1.616 6	10%level

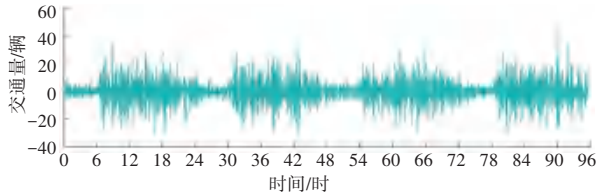


图5 交通量一阶差分图

Fig. 5 Traffic volume first-order difference diagram

2.3 模型阶数的确定

本文采用相关图法以及 AIC 和 SC 准则确定模型的阶数。首先通过自相关图和偏自相关图初步判断模型的阶数,由图6可以看出自相关系数在5阶以后基本上都落在两倍标准差范围之内,可以判定5阶以后拖尾,偏自相关系数也在5阶以后基本上

都落在两倍标准差之内,为5阶以后拖尾。现初步确定 ARIMA ( $p, 1, q$ ) 的  $p$  不大于5,而  $q$  不大于5。模型最终定阶,根据 AIC 和 SC 准则以及拟合优度确定,经过多次反复对 ( $p, q$ ) 不同阶数组合的验证,得出结果见表3,仅显示部分阶数组合的测试结果。

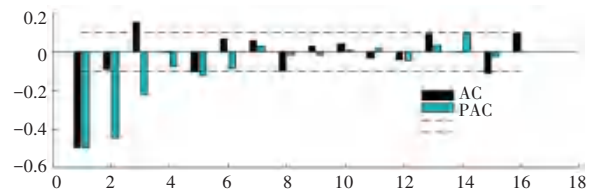


图6 自相关和偏自相关图

Fig. 6 Autocorrelation and partial autocorrelation graphs

表3 部分 ARMA 模型的不同阶数组合测试

Tab. 3 Combination Test of Different Orders for Some ARMA Models

	P value							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
AR(1)	0	0	0	0	0	0	0	0
AR(2)	0	0	0	0	0.520 7	0	0	0.320 8
AR(3)	0	0	0	0	0	0	0	0
AR(4)	0	0	0	0	0	0	0	0
AR(5)	0	0	0	0	0.371 6	0	0	0.047 5
MA(1)	0.942 9	0.901 9	0.021 0	0	-	0.900 3	0.101 8	-
MA(2)	0.367 6	0.604 5	0.774 3	-	0.003 9	-	0.304 7	0
MA(3)	0	0	0	0	0.001 2	0	0	-
MA(4)	0.517 3	0.651 8	-	-	0.072 8	0.519 8	-	0.069 0
MA(5)	0.260 5	-	-	-	0.003 9	0.408 0	0.210 7	0
MA(6)	0	0.000 9	0	0	0.932 6	0.002 0	0	-
AIC	6.772 4	6.770 0	6.767 0	6.762 9	6.796 2	6.769 7	6.768 9	6.792 0
SC	6.869 0	6.858 4	6.846 0	6.833 1	6.884 0	6.857 5	6.856 7	6.871 0
$R^2$	0.452 9	0.452 8	0.453 7	0.454 8	0.438 6	0.453 3	0.453 7	0.439 8

通过建立  $(p, q)$  不同阶数的模型进行多次测试, 对比分析其 AIC、SC 以及拟合优度  $R^2$ , 由表 3 可以看出 Model 4 的 ARIMA (5, 1, 6) 的拟合优度  $R^2$  相对较大, 而 AIC 和 SC 最小, 并且所有参数均通过显著性水平测试, 所以最终确定模型阶数  $p = 5, q = 6, d = 1$ , 即 ARIMA (5, 1, 6), 用最大似然估计法进行参数估计, 得到 ARIMA (5, 1, 6) 模型, 如式 (8):

$$d(Q_t) = -0.9063d(Q_{t-1}) - 0.7548d(Q_{t-2}) + 0.8576d(Q_{t-3}) + 0.7683d(Q_{t-4}) + 0.6699d(Q_{t-5}) + \hat{\varepsilon}_t + 0.0924\hat{\varepsilon}_{t-1} - 1.3372\hat{\varepsilon}_{t-3} + 0.3994\hat{\varepsilon}_{t-6} \quad (8)$$

将一阶差分序列还原后的预测模型, 如式 (9):

$$Q_t = 0.0937Q_{t-1} + 0.1515Q_{t-2} + 1.6124Q_{t-3} - 0.0893Q_{t-4} - 0.0984Q_{t-5} - 0.6699Q_{t-6} + \hat{\varepsilon}_t + 0.0924\hat{\varepsilon}_{t-1} - 1.3372\hat{\varepsilon}_{t-3} + 0.3994\hat{\varepsilon}_{t-6} \quad (9)$$

其中,  $\hat{\varepsilon}_t$  为  $\varepsilon$  的估计值。

表 4 残差序列相关检验

Tab. 4 Residual series correlation test

滞后期数	1	2	3	4	5	6	7	8
AC	-0.025	0.026	-0.006	0.038	-0.004	0.007	0.016	-0.043
PAC	-0.025	0.025	-0.005	0.037	-0.002	0.005	0.017	-0.044
Q-Stat	0.288	0.604	0.622	1.298	1.307	1.334	1.465	2.348
Prob	0.592	0.739	0.891	0.862	0.934	0.97	0.983	0.968

### 2.5 模型预测

模型的预测分为静态预测和动态预测。本文主要用标定好的模型 ARIMA (5, 1, 6) 对后三天的 1 440 条交通量数据进行静态预测。得到实际值和预测值趋势图, 如图 8~图 10 所示, 模型预测精度见表 5。模型预测精度表,

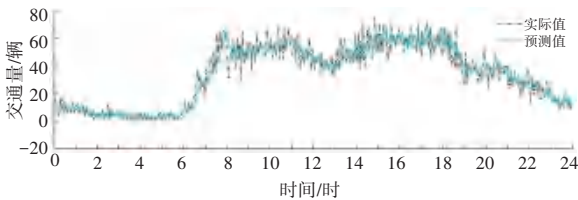


图 8 8 日交通量预测值

Fig. 8 Forecast value of traffic volume on 8<sup>th</sup>

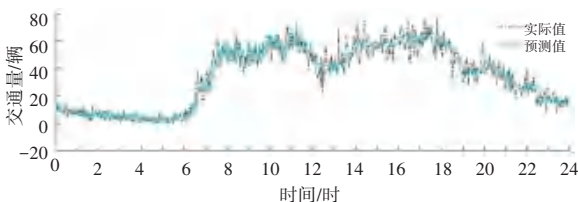


图 9 9 日预测值

Fig. 9 Forecast value of traffic volume on 9<sup>th</sup>

### 2.4 模型的检验

建立模型以后, 需要对残差进行白噪声检验。如果残差满足白噪声序列, 则模型拟合有效。因此, 有必要测试残差序列的平稳性和序列相关性。具体结果如图 7 所示。

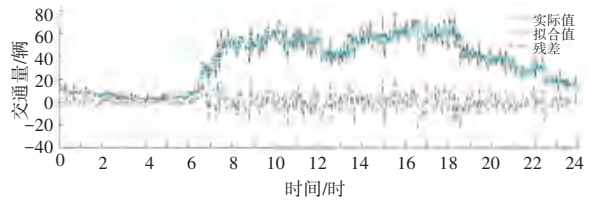


图 7 7 日交通量拟合图

Fig. 7 Traffic volume fitting diagram on 7<sup>th</sup>

从图 7 可以看出, 拟合值与实际值的变化基本一致, 残差也围绕零均值附近波动, 可以初步判断模型的残差序列是平稳的。为了判断更加准确, 继续对残差序列进行序列相关的 Q 检验, 结果见表 4。显示残差序列在滞后 8 期仍然接受原假设不存在序列相关, 因此残差满足白噪声序列。

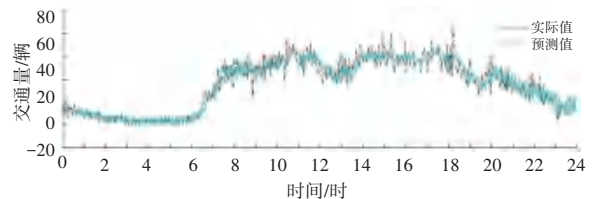


图 10 10 日预测值

Fig. 10 Forecast value of traffic volume on 10<sup>th</sup>

从实际值和预测值的对比图可以看出, 该模型能够更好地预测短时交通量数据, 从预测效果来看, 预测值一直伴随真实值左右, 由于 7-18 时交通流量不确定性较大, 在该段的预测效果相对较差, 总体来说预测效果不错, 表明该模型能够在一定程度上预测未来时刻的交通量, 有着较好的可移植性。

交通量预测需要有一定的指标来衡量, 以下几个指标可以从不同方面反应模型的预测效果<sup>[16]</sup>。设原序列  $y_t$ , 预测结果  $\hat{y}_t$ :

(1) 平均绝对百分比误差 (MAPE):

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (10)$$

(2) 平均绝对误差 (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (11)$$

(3) 均方根误差 (RMSE):

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (12)$$

从表5可以看出8日和9日的MAE仅为5.73和5.78, RMSE为7.61和7.47,表明模型的拟合效果很好,但是平均绝对百分比误差MAPE为32%和34.5%,表明模型的预测精度不够。而10日的MAE为6.34, RMSE为8.22, MAPE也相对较大为38.03%。本研究采用了静态预测法,已经避免了多步预测误差递增的问题,对比表1的基本统计特征值可以发现10日的原时间序列数据标准差较大,不确定性较强,数据较为离散,而8日数据的标准差相对较小,预测精度较好。因为短时交通流数据呈现出不确定性、非平稳性、非线性随机过程,所以离散程度大的数据的预测效果较差。由于模型自身也有一定的局限性,不能很好的反映短时交通流的不确定性影响,也就无法对短时交通流数据的不确定性以及非线性数据做出反应。

表5 预测精度表

Tab. 5 Forecast precision table

	MAE	RMSE	MAPE
8日	5.73	7.61	32.00
9日	5.78	7.47	34.50
10日	6.34	8.22	38.08

### 3 结束语

本文采用ARIMA(5,1,6)模型对短时交通量进行预测,发现ARIMA模型在进行短时交通量预测时有着较好的实时性、稳定性和可移植性,采用第一天的Q7的480条数据标定模型并应用于后三天(Q8、Q9、Q10)的1440条数据的预测,采用静态预测避免了多步预测误差递增的问题,表明ARIMA模型在短时交通量预测方面有着较好的可移植性,虽然历史数据较少,但是ARIMA可以很好的捕捉序列的内在规律,并对大量的数据进行在线预测,得出模型的预测精度随着数据的离散程度增大而增大。随着智能交通的发展和5G时代的到来,交通量数据的获取也越来越便捷,实时动态的数据会给更短时的交

通量在线预测提供可靠的基础,用ARIMA模型可以完成在线、实时、连续的短时交通量预测,从而为交通控制与诱导系统提供实时可靠的数据。本文的不足在于历史数据较少,仅用一天480条交通量数据对模型进行标定,由于短时交通量数据有着一定的不确定性和非线性特征,导致预测精度不是很高,如何用更多的历史数据对模型进行训练,提高模型的预测精度;如何预处理短时交通量数据以及优化模型结构;如何更好的将ARIMA模型和实时、动态的数据结合,进行在线预测是下一步的研究内容。

### 参考文献

- [1] 李旭,周彤梅. 基于动态交通分配的交通诱导与控制协同研究[J]. 中国人民公安大学学报(自然科学版), 2017, 23(4): 79-81.
- [2] 高慧,赵建玉,贾磊. 短时交通流预测方法综述[J]. 济南大学学报(自然科学版), 2008, 22(1): 88-94.
- [3] Isaacson M A, Gail Shoval, Noam Wahl, Hans-Werner Oswald, Frank. Compliance and data quality in GPS-based studies [J]. Transportation 2016, 43(1): 25-36.
- [4] 翁小雄 谭,姚树申,黄征. 城市交叉口交通流特征与短时预测模型[J]. 交通运输工程学报, 2006.03.
- [5] 许伦辉,唐德华,邹娜,等. 基于非线性时间序列分析的短时交通流特性分析[J]. 重庆交通大学学报(自然科学版), 2010, 29(01): 110-113.
- [6] 袁健 范. 交通流短时预测研究进展[J]. 城市交通, 2012, 10(06): 73-79.
- [7] Mohamed S. Ahmed A R C. Analysis of Freeway Traffic Time-series Data by Using Box-Jenkins Technique [J]. Transportation Research Record, 1979, 722.
- [8] 万建强,文洲. ARIMA模型与ARCH模型在香港股指预测方面的应用比较[J]. 数理统计与管理, 2001(06): 1-4.
- [9] 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(01): 69-74.
- [10] 姚亚夫 曹. 基于ARIMA的交通流量短时预测[J]. 交通科技与经济, 2006(3): 105-107.
- [11] 芮少权,匡安乐. 高速公路月度交通量ARIMA预测模型[J]. 长安大学学报(自然科学版), 2010, 30(4): 82-85+91.
- [12] 崔伟杰. 以ARIMA模型预测全国生产安全事故[J]. 安全, 2017, 38(6): 27-28+31.
- [13] 韦凌霄. 基于RVM和ARIMA的短时交通流量预测方法研究[D]. 长安大学, 2016.
- [14] 陈强. 高级计量经济学及Stata应用[M]. 183. 高等教育出版社, 2014.
- [15] 沈根祥. 计量经济学[M]. 上海财经大学出版社, 2013(11): 158-185.
- [16] 张利,李星毅,施化吉. 基于ARIMA模型的短时交通流量预测算法研究[J]. 郑州轻工业学院学报(自然科学版), 2008, 23(4): 89-92.