

基于时间序列的滴滴出行生成预测

王成晨

[上海市政工程设计研究总院(集团)有限公司, 上海市 200092]

摘要: 为了更有效地预测滴滴出行生成,进而使得出行生成预测结果能够更好地运用在交通需求预测的各个方面,在传统交通调查的基础上,针对如何基于时间序列方法实现滴滴出行生成预测进行研究。通过对滴滴打车数据的处理,分析滴滴出行的时间和空间特性,研究滴滴出行的时空变化规律。结合具体案例,提出一种基于 ARMA 模型的滴滴出行生成预测方法,先计算时间序列的自相关性和偏相关性,确定模型阶数,最终得到出行生成预测值。预测结果表明,基于时间序列方法的滴滴出行生成预测方法能够实现滴滴出行时空变化的短期预测,并取得较好的预测精度,能够为交通规划、交通管理等政策的制定提供经验和参考。

关键词: 交通工程;滴滴出行生成预测;ARMA 模型;时间序列

中图分类号: U491

文献标志码: A

文章编号: 1009-7716(2025)08-0034-07

Didi Trip Generation Prediction Based on Time Series

WANG Chengchen

[Shanghai Municipal Engineering Design Institute (Group) Co., Ltd., Shanghai 200092, China]

Abstract: In order to predict the generation of Didi Trips more effectively, the trip generation prediction results can be better applied to all aspects of traffic demand prediction. Based on the traditional traffic survey, how to implement the Didi trip generation prediction based on the time series method is studied. Through the processing of Didi taxi data, the time and space characteristics of Didi Trips are analyzed, and the spatio-temporal variation rules of Didi Trips are studied. Combined with the specific cases, a prediction method of Didi trip generation based on the ARMA model is proposed. The autocorrelation and partial correlation of time series are firstly calculated, and then the model order is determined. Finally, the trip generation prediction value is obtained. The prediction results show that the Didi trip generation prediction method based on time series method can realize the short-term prediction of the spatio-temporal change of the Didi Trips, and the better prediction accuracy is obtained, which can provide experience and reference for the formulation of traffic planning, traffic management and other policies.

Keywords: traffic engineering; Didi trip generation prediction; ARMA model; time series

0 引言

滴滴出行属于居民出行中的一部分,研究滴滴出行生成预测能够进一步挖掘特定人群出行生成规律,并且滴滴数据的应用使得特定人群的出行信息采集方式更加便捷、精准、低成本^[1-2]。滴滴数据不仅包含用户的活动信息,以及用户参与相应活动的位置和时刻,而且具有大样本、低成本的优势,同时能记录较长时间段的连续数据,在交通领域得到了迅速应用^[3]。

目前,使用数据进行出行生成的研究尚不深入,相关的应用也没有广泛展开。本文研究基于时间序

列的滴滴出行预测方法,通过滴滴数据获取出行者的出行信息,并使用 ARMA 模型对滴滴出行生成进行预测。此方法获取的数据量大,且更新快、时效性高、准确性较高^[4]。

1 基于滴滴数据的出行特征分析

1.1 数据预处理

本文采用 M 市 2023 年 12 月 7 日星期四(工作日)全天 24 h 的滴滴打车数据,数据采集记录每个订单的基本信息[包括订单信息、司机信息、乘客信息、起始点位置信息以及订单产生时间(时耗)和金额等,还包括道路拥挤状况、天气信息等]。本文使用相应的数据库管理软件进行数据的存储、管理和操作,原始数据中共有 338 085 条滴滴用户信息。

由于原始滴滴打车数据存在大量冗余、无效数

收稿日期: 2025-01-15

作者简介: 王成晨(1994—),女,硕士,工程师,从事市政道路工程设计工作。

据,对后续研究会产生严重干扰,所以本章先进行数据的基本处理,为后续研究提供高质量的数据来源。

原始滴滴数据往往存在以下几类问题:一是重复冗余数据;二是打车失败数据;三是无效信息数据^[5-6]。针对以上几类问题,本文采用下述方法进行处理:

(1)对重复冗余数据进行删除

在对缺失数据以及错误数据进行删除后,利用编程将记录完全一样的数据删除。

(2)对打车失败数据进行删除

打车失败数据在司机信息数据一栏显示值为 NULL,筛选删除所有 NULL 值。

1.2 滴滴出行时间特征分析

城市交通特征时间主要包括早高峰、晚高峰、平峰以及其他特殊时间段。早高峰和晚高峰是一天中城市交通最为繁忙的时刻,研究这 2 个高峰时刻有利于城市交通部门制定出合理的交通管理措施。工作日和节假日的早晚高峰有所不同,若能够通过滴滴打车数据方便快捷地获取早晚高峰各小区出行生成量,则可以分析了解城市在不同时间周期中的出行发生规律。本文采用 M 市 2023 年 12 月 7 日星期四(工作日)的滴滴打车数据,以此分析得到滴滴出行的一般规律。

本文得到滴滴打车一天的全部数据,经过数据预处理,得到 M 市全天打车出行生成总量的变化图(见图 1)。由图 1 可知,M 市在 2023 年 12 月 7 日打车出行生成总量整体呈现先增加后减少再增加再减少的趋势。在早晨 7:00—9:00 以及傍晚 17:00—19:00,出行具有明显的 2 个峰值;在 10:00—16:00,打车量略有下降,近似于平峰趋势;在 12:00—13:00 的时间段内,打车量有小幅波动;在夜晚 22:00—凌晨 5:00,打车量处于一天中订单量的最低值。综上所述,M 市一天的滴滴打车订单变化具有明显的平峰、早晚高峰趋势。

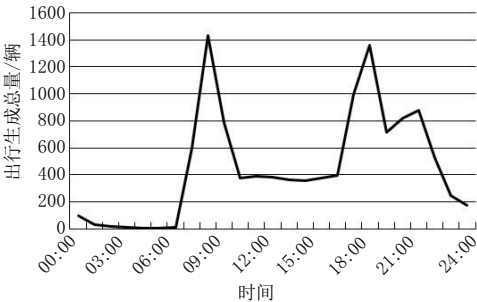


图 1 M 市全天出行生成变化图

1.3 滴滴出行空间特性分析

不同的交通小区有着不同的出行变化规律,为此本研究将 M 市划分为 66 个交通小区,并选择 51 号交通小区(商业区)为研究对象。研究此交通小区全天的出行变化,有助于更深刻地了解商业性质地段的出行变化规律。滴滴出行数据可以很好地记录每次出行时刻的各种信息,比采用传统交通出行调查所获得的出行时刻数据更加准确,也能够更好地反映各交通小区在不同时间段的出行产生量,为后续研究提供更准确的数据分析。

51 号交通小区具有明显的商业性交通区特点,即夜间出行产生量相较于其他交通区高很多。在其他时间段,51 号交通区都保持着较高的出行产生量和出行吸引量,只有夜间的出行变化是区别于居住区、工作区和其他区域的最大特点。工作日,聚会、购物等弹性出行主要集中在晚上,这种集聚性娱乐活动一般会持续到夜间 22:00 左右,结束时会产生大量的出行,所以在这个时间段,51 号交通区出行量显著高于其他小区。以 10 min 为间隔绘制 51 号交通区出行变化,如图 2 所示。

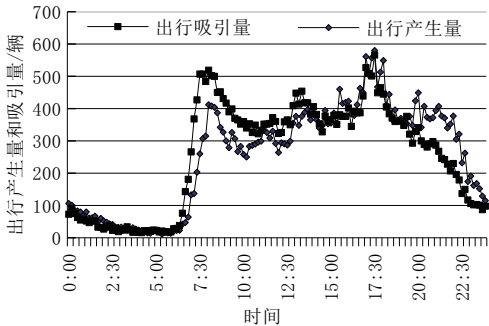


图 2 51 号交通区全天出行变化图

由图 2 可知,51 号交通区也具有明显的早晚高峰,早高峰出行产生量低于出行吸引量,一般是由于商业区也会吸引部分工作人群出行。夜间 20:00 之后,该交通区拥有较高的出行产生量,说明该阶段的娱乐活动相继结束,产生返程出行。

本研究选用 51 号交通小区为案例进行重点分析,其他交通小区的出行产生量和吸引量可以采用类似的方法推断其出行变化规律。

2 时间序列

2.1 时间序列的基本概念

时间序列是按照时间的先后顺序排列而成的数据序列,也称作动态数列。这些按时间排序的动态序列一般反映出某类现象的变化规律,包括上升、下

降等^[7]。通过对时间序列进行延伸,可以预测数据短期或长期的变化。

时间序列中的观测数据虽然彼此之间没有特定联系,但从整体上来看又呈现出某种特殊联系。所以,对于时间序列的研究,首先需要观察这类对象是否具有时间序列的特征,如果不存在此类特征,则需要对原始数据作轻微调整,以使其满足时间序列的要求^[8]。在利用SPSS软件建立滴滴出行生成量和道路拥挤程度的模型时,发现所有模型的预测精度都非常低,这说明道路拥挤程度和居民出行生成量之间并不存在一般的线性、二元和指数等关系,所以一般的回归方法并不适用于本文的滴滴出行生成量预测^[9-10]。

在传统的居民出行生成预测中,由于传统交通调查的局限性,数据量少,数据时间可靠度低,所以一般不会分析某个交通区一天内居民出行的时空分布特征。本文利用滴滴数据,可以弥补传统交通调查的这类缺陷,且时间序列分析方法可以填补这方面的空白,这也是本文选择时间序列分析方法进行居民出行预测的原因之一。

时间序列分析方法在数据量大,且数据时间连续性强时,预测效果更好,所以本文选用10 d的滴滴打车数据作为研究对象。将10 d的滴滴全天打车数据按时间排序,发现滴滴打车数据符合时间序列的要求,具有一定的综合性变化规律,每个数据点之间呈现出随机变化性,且每天都包含早晚高峰,具有一定的周期性,故本文采用时间序列的方法来预测居民出行生成。

2.2 AMRA模型

ARMA模型(Auto-Regressive and Moving Average Model)是研究离散时间序列的重要方法,以自回归模型(AR模型)和滑动平均模型(MA模型)为基础,一般用于数据量较大的研究。

ARMA模型又称为自回归滑动平均模型。它与AR模型和MA模型相比,预测效果更好,但是预测过程非常繁琐。

2.2.1 AR模型

AR模型是一种线性预测,是利用已经获取的M个数据去推测未来的数据。AR模型与插值法相似,但插值法一般使用两个数据去推测多个点的数据,而AR模型则是使用多个数据推测多个点。从这一点来看,AR模型的效果更好一点,数据的可靠度也较高^[11-13]。

AR(p)是p阶自回归模型,结构如下:

$$x_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 x_{t-2} \cdots + \beta_p x_{t-p} + \epsilon_t \quad (1)$$

$$\beta_p \neq 0$$

$$E(\epsilon_t) = 0, \text{Var}(\epsilon_t) = \sigma_\epsilon^2, E(\epsilon_t \epsilon_s) = 0, s \neq t$$

$$Ex_s \epsilon_t = 0, s < t$$

2.2.2 MA模型

MA模型又叫滑动平均模型,MA(q)是q阶滑动平均模型,具有如下结构:

$$x_t = \mu + \epsilon_t - \varphi_1 \epsilon_{t-1} - \varphi_2 \epsilon_{t-2} - \cdots - \varphi_q \epsilon_{t-q} \quad (2)$$

$$\varphi_q \neq 0$$

$$E(\epsilon_t) = 0, \text{Var}(\epsilon_t) = \sigma_\epsilon^2, E(\epsilon_t \epsilon_s) = 0, s \neq t$$

2.2.3 ARMA模型

在时间序列中,影响因子会对观测变量造成一定的影响,但是这些变量又存在自身的变化规律,即变量既受外部因子影响,也会依靠自身规律变化。假定影响因子分别为 $x_1, x_2, \dots, x_p$,由回归分析方法可得:

$$Y_t = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \epsilon \quad (3)$$

Y_t 作为观测对象会因自身变化而变化,其内在规律如下:

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \epsilon_t \quad (4)$$

误差项在不同的时期具有一定的内在相关关系,如下式:

$$\epsilon_t = \rho_t + \alpha_1 \rho_{t-1} + \alpha_2 \rho_{t-2} + \cdots + \alpha_q \rho_{t-q} \quad (5)$$

ARMA模型为:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \rho_t + \alpha_1 \rho_{t-1} + \alpha_2 \rho_{t-2} + \cdots + \alpha_q \rho_{t-q} \quad (6)$$

3 时间序列

3.1 数据录入

本论文选取具有代表性的51号交通小区进行研究。此交通区每天的滴滴打车数据最多,并且根据上述分析可知,此小区具有早晚高峰的滴滴出行时空规律,符合一般时间序列分析的特点。此外,51号交通小区在研究时间段的出行变化幅度较大,便于模型的建立。

本论文选取51号交通区12月4日—12月13日10 d的数据量,并选取早高峰至晚高峰这段时间作为研究时间区间,即7:30—19:00。此段时间为出行的正常时间,影响因素较其他时间段少。

将51号交通区10 d的出行产生来量和吸引量分别按时间序列进行排序,得到的时间序列如图3所示。从10 d的出行产生量和出行吸引量时间序列图可以看出,这些序列具有一定的周期性,并且曲线

变化趋势相仿。

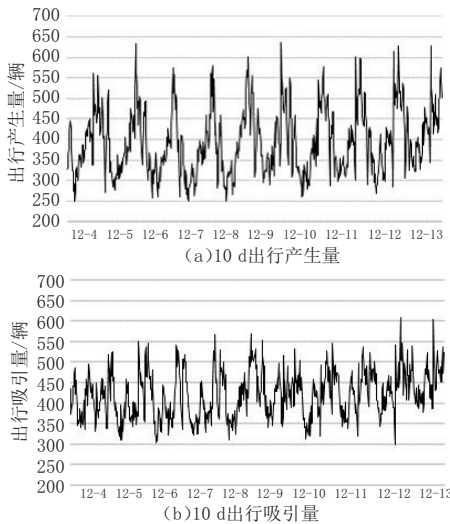


图3 51号交通小区出行时间序列图

3.2 自相关偏相关分析

本论文选取具有代表性的51号交通区进行研究。此交通区每天的滴滴打车数据最多,并且根据上述分析可知,此小区具有早晚高峰的滴滴出行时空规律,符合一般时间序列分析的特点。

3.2.1 AR(p)的自相关函数

已知AR(p)模型为公式(1),则

(1)自协方差函数

$$\begin{aligned} r_k &= E(x_t \times x_{t+k}) \\ &= E[x_t(\phi_1 x_{t+k-1} + \phi_2 x_{t+k-2} + \cdots + \phi_p x_{t+k-p} + \epsilon_{t+k})] \\ &= E(x_t \phi_1 x_{t+k-1}) + E(x_t \phi_2 x_{t+k-2}) + \cdots + E(x_t \phi_p x_{t+k-p}) \\ &= \phi_1 r_{k-1} + \phi_2 r_{k-2} + \cdots + \phi_p r_{k-p} \end{aligned} \quad (7)$$

其中E是数学期望。

(2)自相关函数

$$\begin{aligned} \rho_k &= \frac{r_k}{r_0} = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p} \\ \rho_k &= \rho_{-k}, \quad \rho_0 = 1 \end{aligned} \quad (8)$$

Yule-Walker 方程:

$$\begin{cases} \rho_1 = \phi_1 + \phi_2 \rho_1 + \cdots + \phi_p \rho_{p-1} \\ \rho_2 = \phi_1 \rho_1 + \phi_2 + \cdots + \phi_p \rho_{p-2} \\ \dots\dots\dots \\ \rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \cdots + \phi_p \end{cases} \quad (9)$$

3.2.2 MA(q)的自相关函数

已知MA(q)序列,则

$$X_t = \sum_{j=0}^q \theta_j Z_{t-j}, \quad \{Z_t\} \rightarrow N(0, \sigma^2) \quad (10)$$

自相关函数为

$$\rho_k = \begin{cases} 1, k = 0 \\ \frac{\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \cdots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2}, k = 1, 2, \dots, q \\ 0, k > q \end{cases} \quad (11)$$

K大于q时为零,称 ρ_k 具有截尾性。

自协方差函数为

$$r(k) = E(X_t X_{t+k}) = \begin{cases} \sigma^2 \sum_{j=0}^q \theta_j \theta_{j+k}, |k| \leq q \\ 0, |k| > q \end{cases} \quad (12)$$

3.2.3 ARMA模型的自相关函数

$$\rho_k = \begin{cases} \frac{(1 + \theta_1 \phi_1)(\theta_1 + \phi_1)}{1 + \theta_1^2 + 2\theta_1 \phi_1}, k = 1 \\ \phi_1 \rho_{k-1}, k \geq 2 \end{cases} \quad (13)$$

由模型可得到出行产生量和出行吸引量的自相关和偏相关检验,如图4所示。

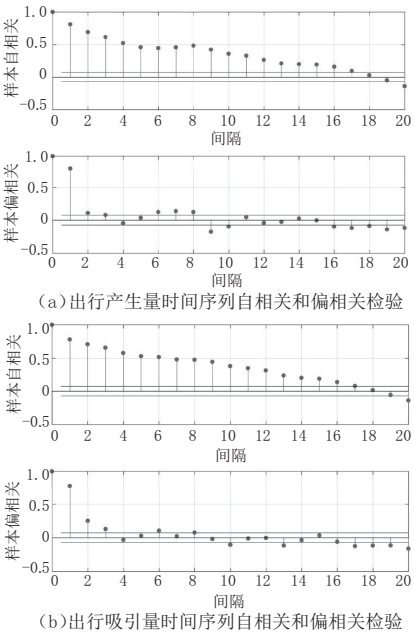


图4 出行时间序列自相关和偏相关检验图

辨识时间序列所属模型种类是通过自相关和偏相关系数,而自相关和偏相关系数描述的都是当前观测数据与之前观测数据的相关性。一般情况下,观测数据会在零值附近来回摆动。当样本的自相关和偏相关系数在前k阶段出现明显大于2倍的标准差,而后面大部分系数都在2倍的标准差范围之内时,称为k阶截尾,否则为拖尾。ARMA模型判定基本原则见表1。

表1 ARMA模型判定基本原则

p	q	模型
拖尾	q阶截尾	AR(P)模型
P阶截尾	拖尾	MA(q)模型
拖尾	拖尾	ARMA(p,q)模型

3.3 差分归零化处理

自回归移动平均 ARMA(p, q)模型为

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q \epsilon_{t-q} \quad (14)$$

线性差分方程为

$$\begin{aligned} Y_t - \beta_1 Y_{t-1} - \cdots - \beta_p Y_{t-p} \\ = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \end{aligned} \quad (15)$$

其中, $\{Z_t\} \sim \text{WN}(0, \sigma^2)$, p 阶自回归多项式为

$$\beta(z) = 1 - \beta_1 z - \cdots - \beta_p z^p \quad (16)$$

与 q 阶滑动平均多项式为

$$\theta(z) = 1 - \theta_1 z - \cdots - \theta_q z^q \quad (17)$$

$\{Y_t, t \in Z\}$ 为 p 阶自回归、 q 阶滑动平均序列, 简称 ARMA(p, q) 序列, 满足 ARMA(p, q) 模型。

当自相关系数和偏相关系数都呈现出不截尾的情况, 并且末尾收敛于零值时, 则一般情况下该序列可能为 ARIMA(p, d, q) 模型, 要想拟合模型, 首先必须利用差分运算将其转换成 ARMA(p, q) 模型, 并使其成为平稳序列。对出行产生量和出行吸引力时间序列也需要进行归零化处理, 减少多余因素的影响程度, 得到平稳的时间序列, 有利于后续的出行生成预测。其结果如图 5 所示。

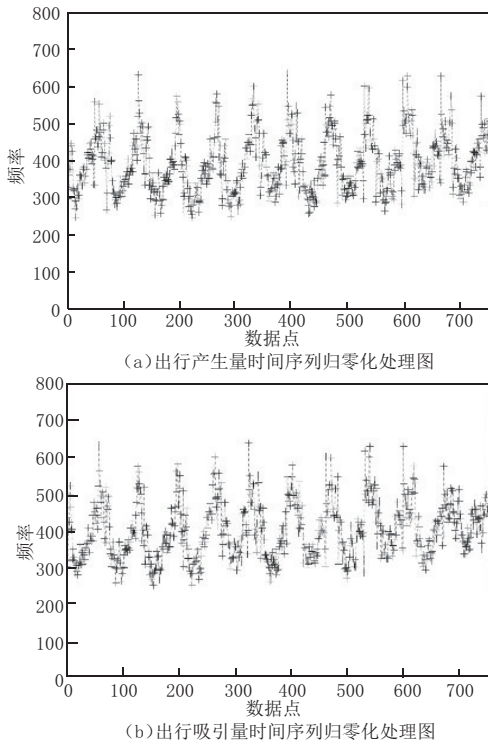


图5 出行时间序列归零化处理图

3.4 确定模型阶数

AIC 准则又称最小信息准则, 用来确定 ARMA 模型阶数和参数的最佳估计。此准则的主要目的是判断预测时间序列的发展与哪一个随机发展过程最为

相近。当观测数据非常多时, 观测数据的自相关函数才能够与原时间序列的自相关函数吻合。首先由从低到高的顺序确定模型阶数, 再计算 AIC 值, 最小 AIC 值的阶数满足最优模型。

模型参数最大似然估计:

$$\text{AIC} = (n - d) \lg \sigma^2 + 2(p + q + 2) \quad (18)$$

模型参数最小二乘估计:

$$\text{AIC} = (n - d) \lg \sigma^2 + (p + q + 1) \lg n \quad (19)$$

根据模型最优解计算得到 ARMA(3, 4) 模型满足出行产生量时间序列, ARMA(9, 8) 模型满足出行吸引力时间序列, 所以确定的出行产生量的模型阶数分别为 3 和 4, 出行吸引力的模型阶数分别为 9 和 8。

3.5 出行生成预测

由 ARMA 模型可得到 1 月 14 日的预测值, 预测的出行产生量和出行吸引力分布如图 6 所示。

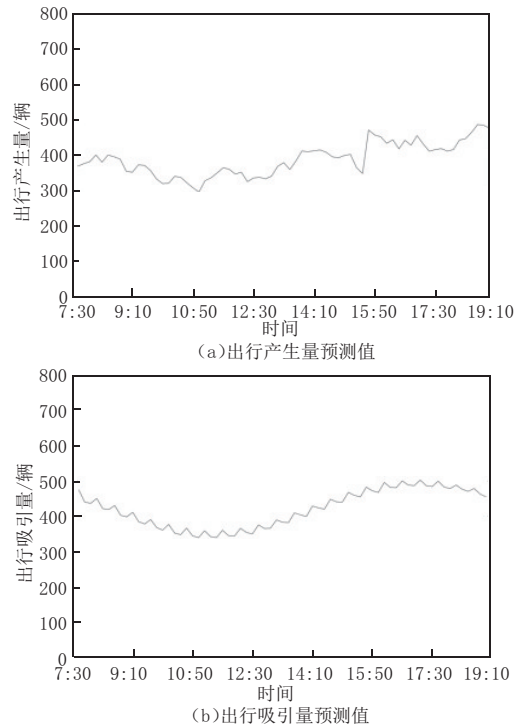


图6 出行生成预测图

4 预测结果分析

4.1 预测结果

由 ARMA 模型可以得到 51 号交通区 1 月 14 日的出行预测时间序列, 并且由实际的 1 月 15 日滴滴出行生成数据, 可得到出行产生量对比图(见图 7)和出行吸引力对比图(见图 8)。

由图 7 可知, 出行产生量的预测值和实际值的总体趋势相同。10~48 数据点(9:00—15:20)之间的趋

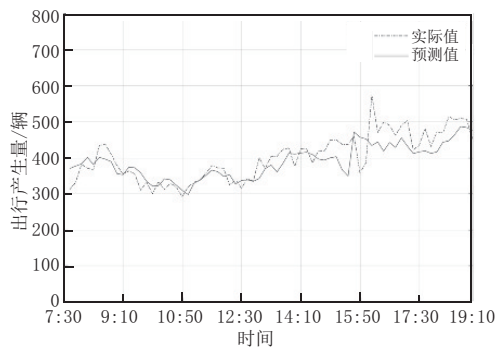


图7 出行产生量对比图

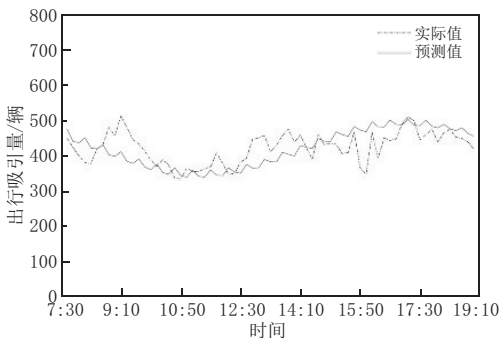


图8 出行吸引量对比图

势基本相同,9:00—15:20处于出行平峰时间,出行产生量的总体变化程度小,预测准确度较高;0~10数据点(7:30—9:00)之间的变化趋势有略微差别;后29个数据点(15:20—19:00)的出行变化趋势相差很大,15:20—19:00为晚高峰时间段,出行量的变化值较大,因此该阶段预测值与实际值的误差较大。预测值在48~50数据点(15:20—15:40)之间出现了出行波动,在实际值中时间后移至15:40—16:00。总体而言,出行产生量预测值的时间变化图同样包括早晚高峰:早高峰从7:30开始,9:00结束;晚高峰从15:40左右开始,19:00左右结束。

由图8可知,出行吸引量的预测值和实际值的总体趋势相同。20~45数据点(10:40—14:50)之间的趋势基本相同,10:40—14:50处于出行平峰时间,出行吸引量的总体变化程度小,预测准确度较高;0~20数据点(7:30—10:40)之间的变化趋势有略微差别;后24个数据点(14:50—19:00)的出行变化趋势相差很大,17:00—19:00为晚高峰时间段,出行量的变化值较大,因此该阶段预测值与实际值的误差较大。总体而言,滴滴出行产生量预测值的时间变化同样包括早晚高峰:早高峰从7:30开始,9:00结束;晚高峰从15:40左右开始,19:00左右结束。

51号交通小区为商业区,所以预测的出行早晚高峰时间与传统的居住区和工作区不同。

4.2 误差分析

几类误差计算公式见表2。

表2 误差计算公式

准确度(误差)		精密度(偏差)	
绝对误差 $\delta = x - \mu$		平均偏差 $\bar{d} = \frac{\sum_{i=1}^n x_i - \bar{x} }{n}$	标准偏差 $S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$
相对误差 $\frac{\delta}{\mu} = \frac{x - \mu}{\mu} \times 100\%$		相对平均偏差 $\frac{\bar{d}}{\bar{x}} \times 100\%$	相对标准偏差 $RSD = \frac{S}{\bar{x}} \times 100\%$

绝对误差可以反映预测值误差的实际情况,相对误差能够反映测量的可信程度。平均偏差代表一组测量值中任意数值的偏差,相对平均偏差是考虑到分析结果的误差。标准偏差可以度量数据分布的分散程度,相对标准偏差表示标准偏差占平均偏差的百分率。

由表3和表4可知:出行产生量的绝对误差为12.79%,相对误差为3.29%,平均偏差为30.82%,相对平均偏差为7.68%,标准偏差为40.72%,相对标准偏差为10.15%;出行吸引量的绝对误差为1.76%,相对误差为0.42%,平均偏差为36.43%,相对平均偏差为8.64%,标准偏差为46.18%,相对标准偏差为10.95%。

表3 出行产生量误差计算表

准确度(误差)		精密度(偏差)	
绝对误差	12.79%	平均偏差	30.82%
相对误差	3.29%	标准偏差	40.72%
		相对平均偏差	7.68%
		相对标准偏差	10.15%

表4 出行吸引量误差计算表

准确度(误差)		精密度(偏差)	
绝对误差	1.76%	平均偏差	36.43%
相对误差	0.42%	标准偏差	46.18%
		相对平均偏差	8.64%
		相对标准偏差	10.95%

出行产生量和吸引量的相对误差都小于4%。相对误差的大小是预测准确度的体现,因本文的相对误差都在预测误差允许范围之内,所以基于LBS的出行生成预测是可行的。

结合图7和图8得出:使用时间序列预测滴滴出行生成的优点在于,通过滴滴出行生成的历史数据来预测未来滴滴出行生成变化,基于历史时间序列建立模型,考虑事物发展的随机性。此方法根据观测数据的历史变化规律建立模型,不需要对事物发展的因果关系作深入研究,这一点不同于一般的回归分析方法。因其不需要影响预测变量的因果关

系,当事物发展的动态性较强,关于时间的数据量多,且关于影响预测变量的决定性因素信息很少时,采用时间序列分析模型可以达到很好的效果。本文符合时间序列分析的要求,数据量大,影响出行生成的因素少,且因果关系不强,所以使用时间序列法建立出行生成模型比一般的回归模型等具有一定的优势。

由于时间序列分析法以时间为中心,重点考虑时间在预测中的作用,不考虑外界具体因素的影响,将所有的影响因素都归为时间的影响,所以存在一定的预测误差。当外部环境变化大时,可能会存在较大的偏差。本论文采用滴滴数据进行研究,可以适当减少时间序列分析方法的长期误差,由于该技术能够随时获取出行数据,所以可以将出行生成的长期预测转化为多个时间段的短期预测(短期一般为半年)。这样既可以减少时间序列的预测误差,也可以随时校正滴滴出行生成量,从而更好地分析滴滴出行特征。这一优势是传统交通调查难以实现的。

5 结 论

本研究定量和定性地分析了滴滴出行的时间、空间出行特征,挖掘滴滴出行的一般规律,并建立了ARMA出行预测模型。对滴滴出行进行预测分析,得出以下结论:

(1)使用时空分析方法研究滴滴出行特征,能够定量和定性地得到滴滴出行变化规律。

(2)通过对滴滴出行生成预测结果进行分析,表明时间序列方法能够满足滴滴出行生成预测的需求。以典型交通区为例,对其进行空间特征研究,并

与实际滴滴出行结果进行对比分析,证明该技术具有较高的可行性。

(3)本文提出一套基于时间序列预测滴滴出行生成的方法,能够实现滴滴出行时空变化的短期预测,并且取得一定的预测精度,能够为交通规划、交通管理等政策的制定提供经验和参考。

参考文献:

- [1] 张钧豪. 基于时空预训练增强型Transformer的交通流量预测研究[D]. 成都: 西南大学, 2024.
- [2] 杨帆, 徐光霁. 大数据环境下城市出行需求建模方法研究[J]. 市政技术, 2022, 40(3): 52-57.
- [3] 鲍凡. 基于时间序列模型的停车需求预测分析[D]. 西安: 西安电子科技大学, 2020.
- [4] 梁柯. 基于ARMA模型的城市轨道交通客流特征分析与预测[D]. 上海: 上海工程技术大学, 2020.
- [5] 胡浩, 闫伟, 李泓明. 基于组合预测方法的城市道路短时交通流预测[J]. 工业工程与管理, 2019, 24(3): 107-115.
- [6] 赵再先, 刘晓玲. 大数据时代“四阶段”交通模型体系技术升级方法[J]. 交通与运输, 2019, 32(增刊1): 71-74.
- [7] 卢志义, 聂惟聪, 陈丽珍. 基于ARMA模型的城市轨道交通客流量预测[J]. 河南科学, 2018, 36(5): 646-651.
- [8] 杨高飞, 徐睿, 秦鸣, 等. 基于ARMA和卡尔曼滤波的短时交通预测[J]. 郑州大学学报(工学版), 2017, 38(2): 36-40.
- [9] 刘勇. 新城区交通生成预测方法研究[J]. 低碳世界, 2016(13): 198-200.
- [10] 冯明兵. 基于活动的出行需求预测模型的研究[D]. 重庆: 重庆交通大学, 2015.
- [11] 李永晟, 夏超. 基于用地混合熵的城市交通生成预测模型[J]. 天津城建大学学报, 2014(5): 328-331.
- [12] 胡郁葱, 陈海伟, 欧阳剑, 等. 新城土地利用空间分布模拟与交通需求预测[J]. 华南理工大学学报(自然科学版), 2014(12): 27-34.
- [13] 董春娇, 邵春福, 诸葛承祥, 等. 基于时空特性的交通自由流短时预测状态空间模型[J]. 土木工程学报, 2013, 46(8): 111-118.

《城市道桥与防洪》杂志

是您合作的伙伴,为您提供平台,携手共同发展!

欢迎新老读者订阅期刊 欢迎新老客户刊登广告

官方网址: <http://www.csdqyfh.com> 电话: 021-55008850 联系邮箱: roadbridgeflood@163.com

