

基于时间序列分析对高速公路车流量的研究

摘要：本文通过对附件所给的高速公路车辆轨迹数据合并整理得到了各时段关于各门架、收费站及路段分别的车流量数据。基于整理好的数据，统计得到车流量的时空分布特征、计算并分析了得到各门架的漏记率及收费站与路段的车流量关系。然后通过建立排队论模型，对收费站的车道数量设置进行研究。最后通过对各路段在 6 天内时间序列的分析，建立指数平滑法的简单季节性模型，预测了未来 24 小时的车流量，并给出了各路段的实时限速方案。

针对问题一，本文利用 MATLAB 整理并计算得到了 15 座门架的日均车流量数据。通过分析观察得出：所有门架的日均车流量随时间的变化趋势在很大程度上相似，呈“M”型变化趋势。其中日均最高流量的时段可达近 2000 次/时，而车流量在最低的时段会跌至 50 次/时以下。然后基于日均车流量数据，给出了各门架最为合适的维修时间段。

针对问题二，本文基于数据分析，统计出各门架在 6 日之内的对三种车辆的漏记次数表。结合总记录次数，计算得到各门架的漏记率。分析结果得到结论：所有门架只会对货车出现漏记现象，不会对客车和其他车辆漏记。其中漏记率最高的门架为门架 7，其漏记率超过了 6.5%；而门架 1、2、3、8 的漏记率不到 0.3%，属于漏记率较低的门架。

针对问题三，本文基于对路段流量与收费站流量的统计，通过皮尔逊相关系数来衡量两个变量呈线性相关程度，从而得出路段流量与

收费站流量的关系，然后通过收费站流量以及顺畅通行的条件建立排队论模型，来求出一共需要 19 条 ETC 通行车道，29 条 MTC 车道，15 台便携式收费机从而达到费用最省。

针对问题四，本文根据对路段车流量的时间序列图分析，利用数据含有稳定季节成分、无整体趋势的特性，结合 SPSS 专家建模器建立了指数平滑法的简单季节性模型。计算预测了未来 24 小时内的各路段的车流量，并基于预测数据给出了各路段的实时限速方案。

关键词： 数据分析；车流量；Pearson 相关系数；排队论模型；指数平滑法；简单季节性

一、 问题重述

1.1 问题背景

近年来，随着中国经济社会开发和人民生活水平的提高，高速公路的交通流量增加，高速公路使用者的需求增加，有时会发生停车收费引起的交通堵塞现象。因此，为了克服以往全人工及半自动收费的缺点，持续提高高速公路的交通效率和服务水平，电子收费系统(ETC)诞生了。系统的主要组成部分是建在高速公路主线的横断面上的门架，主要包括汽车视觉识别，天线，边缘计算服务单元等主要设备。ETC依托于在高速公路的特定规模区域中密集且连续配置的门架系统，与高速公路收费站等其他必要的资源组合来实现对车辆的行驶轨道检测。

1.2 问题重述

问题一：通过车辆轨迹信息，描述车流量的时空特征，据此指出维修门架的编号和时间。

问题二：，通过轨迹信息来看门架是否会漏失对过路车辆的记录，若会，则指出各门架漏失记录率以及漏失记录车辆的特征。

问题三：说明收费站流量与路段流量的关系，并确定收费站车道数与应急收费系统的个数，来保证道路通畅的情况下使得其花费最少。

问题四：预测每个路段未来 24 小时的车流量，并利用交通流时空分布不均衡的特点，通过预测模块实时优化未来 24 小时每个路段的限

速值，最大化道路的安全性与流通性。

二、 问题分析

2.1 问题一的分析

问题一首先要求描述车流量的时空分布特征。从空间角度来看，有 15 个 ETC 门架分布在高速公路路段的不同位置；从时间角度来看，一天 24 小时的不同时间段内通过同一门架的车流量也不相同。据此，本文通过分割一天的时间段，求解在 2022 年 2 月 22 日至 2 月 27 日的 6 天时间内，各 ETC 门架的平均每日的车流量随时间变化特征。其次，关于每个门架维修时间段的确定，本文求解得到了各门架日均流量最小的时间段，结合车流量随时间的变化折线图，最终给出每个门架最为合适的维修时间段。

2.2 问题二的分析

问题二要求对漏失记录车辆的门架和相应的车辆类型做出统计，并计算各门架的漏失记录率。本文通过对所给信息进行数据处理，建立六天时间内各门架关于各车辆类型的漏失记录次数表，再结合前文得到的这些天内通过各门架车辆的总流量，可分别计算得到各门架在每种车型下的漏失记录率及各门架的总漏失记录率。据此，我们便可结合数据直观的对漏失记录现象进行分析。

2.3 问题三的分析

针对问题三前半为了确定说明收费站流量与路段流量的关系，首

先要排除两者所在空间分布的干扰。可以通过画图发现路段车流量和收费站车流量的变化趋势与编号、时间的关系。最后我们取所有收费站车流量和路段车流量关于时段或者编号的综合变化趋势，分析二者相关性。针对问题三后半，由于题目中要求汽车顺畅通行，我们将 ETC 和人工通道通过的一天的车流量数据统计出来，以此来建立排队论模型，我们通过研究排队系统的数量指标，通过求解其排队队长和排队机率来描述汽车是否顺畅通行，借此求出来让汽车顺畅通行的最少服务台个数，即各车道的最少数量和应急收费系统的最少数量，来保证费用最少。

2.4 问题四的分析

问题四的第一个子问题要求我们预测出各路段未来 24 小时的车流量。本文通过分析车流量在前 6 天内时间序列图，结合序列特征，借助 SPSS 专家建模器选择最合适的指数平滑模型对时间序列进行分析和预测。第二个子问题要求我们对各路段的实时限速值做出设置。本文基于对车流量与路段限速值关系的假设，结合预测数据能较为简单的求出各路段的实时限速值。

三、 模型假设

1. 假设该路段高速公路上车辆均正常行驶，无逆行等特殊行为
2. 假设每日对门架的维修时间在 4 小时之内，且不间断
3. 假设不同车型对通行的时间没有影响

4. 假设门架维修不影响车辆的通行速度
5. 假设收费后车辆不在此逗留
6. 忽略同一路段的不同车道的限速值差异,即每个路段有同一限速标准
7. 假设车流量与限速值为简单的线性关系

四、 符号说明

符号	符号说明
r_i	门架 i 的漏记率($i = 1, 2, \dots, 15$)
$\bar{\rho}_0$	收费站空闲概率
$\bar{c}(s, \bar{\rho})$	车辆到达时必须等待的概率
\bar{L}_q	收费站排队的平均队长
m	周期长度(小时取 24)
α	水平的平滑参数
γ	季节的平滑参数
h	预测超前期数
\hat{x}_{t+h}	第 h 期的预测值
sl_{ij}	路段 i 在时段 j 限速值
$\rho_{(x,y)}$	皮尔逊相关系数
x	路段平均流量
y	收费站平均流量

五、模型的建立与求解

5.1 模型一的建立与求解

针对问题一的第一个子问题，关于描述车流量的时空分布。本文首先从门架的空间分布特征，分别对各门架进行讨论。再从时间角度分析，将一天 24 小时等间隔地分为 24 个时间段：例如时段 1 表示 0:00—1:00，时段 2 表示 1:00—2:00，以此类推得到一系列时段编号。然后通过对附件中轨迹数据进行整合，基于“收费站/门架编号”和“记录时间”两列数据，进行简化后筛选出了各门架在不同时段的流量表(详见附录)，以门架 1 为例，表 1 展示了门架 1 从 2022 年 2 月 22 日至 2 月 27 日的六天时间内，每天各时段的日均流量，

表 1. 门架 1 各时段的日均流量

时段	通过门架次数	时段	通过门架次数
0: 00-1: 00	47.8	12: 00-13: 00	435.8
1: 00-2: 00	38.2	13: 00-14: 00	559
2: 00-3: 00	29.2	14: 00-15: 00	545
3: 00-4: 00	95.5	15: 00-16: 00	474.2
4: 00-5: 00	63.2	16: 00-17: 00	462.3
5: 00-6: 00	144.3	17: 00-18: 00	357.5
6: 00-7: 00	298	18: 00-19: 00	270.5
7: 00-8: 00	402	19: 00-20: 00	204.3
8: 00-9: 00	587.5	20: 00-21: 00	161.7
9: 00-10: 00	532.8	21: 00-22: 00	116.8
10: 00-11: 00	410.7	22: 00-23: 00	98.7
11: 00-12: 00	335.2	23: 00-24: 00	60.2

为了更清晰直观地得到门架在一天内流量随时段变化趋势，我们绘制了各门架在不同时段的流量表(详见附录)，仍以门架 1 为例，图 1 为门架 1 各时段的日均流量图

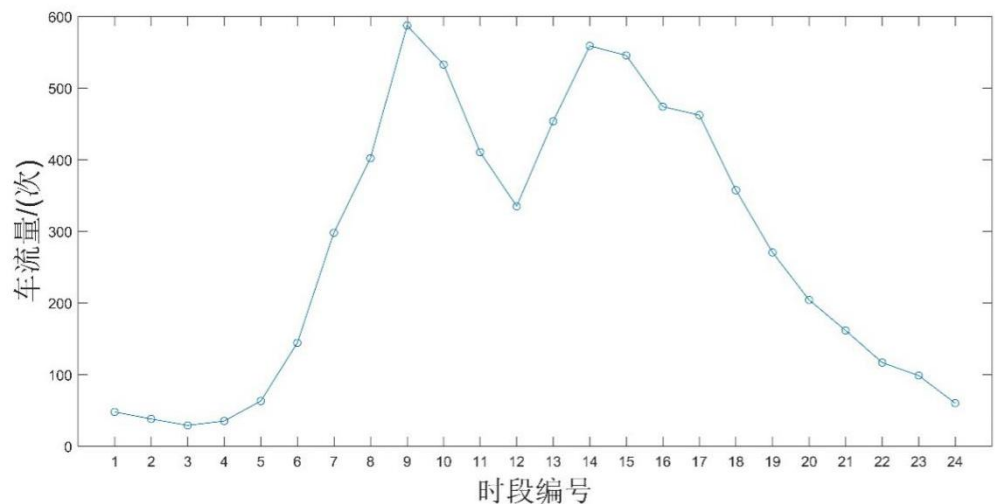


图 1. 门架 1 各时段的日均流量图

从图一中可以得知，门架 1 的日平均车流量从凌晨 0：00 开始略有减少，从凌晨 4 时开始迅速上升，直至上午 9—10 时达到最高峰将近 600 次/时；之后又迅速减少，中午 12 时左右衰减到 300 次/时；接着下午 15 时左右又迅速上升至第二次高峰，最后车流量随时间又逐渐减少，到深夜 24 时下降到 100 次/时以下。我们又将所有门架的车流量变化呈现在同一张图中，如图 2

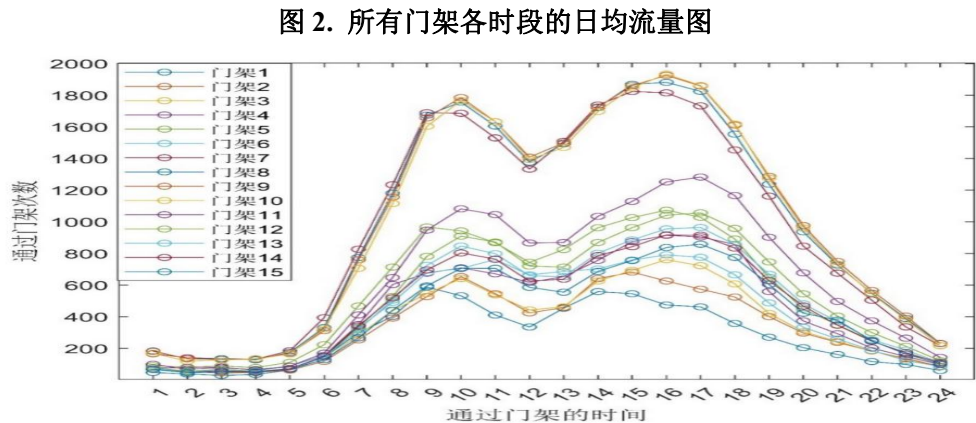


图 2. 所有门架各时段的日均流量图

通过观察得出：所有门架每日内的平均车流量随时间的变化趋势在很大程度上相似，都呈“M”型变化趋势。时间上各门架整体上在上午 10 时及下午 16 时左右达到日最高峰。其中日均最高流量的时段可达近 2000 次/时，而最低流量会低于 50 次/时。空间上，门架 7、8、9、10 在各时段整体上的车流量高于其余门架，门架 1、2、3 车流量整体较低。

针对问题一的第二个子问题，关于对各门架维修时段的选取。通过查阅相关资料得知一般高速公路对门架的维修在 4 个小时以内，故我们简单地选取每个门架日平均流量最小的 4 个时段为维修时段，汇总得到表 2，

表 2. 各门架的最适维修时段表

门架编号	维修时段编号	门架编号	维修时段编号
1	1, 2, 3, 4	9	2, 3, 4, 5
2	2, 3, 4, 5	10	2, 3, 4, 5
3	1, 2, 3, 4	11	2, 3, 4, 5
4	1, 2, 3, 4	12	2, 3, 4, 5
5	1, 2, 3, 4	13	2, 3, 4, 5
6	1, 2, 3, 4	14	1, 2, 3, 4
7	1, 2, 3, 4	15	1, 2, 3, 4
8	2, 3, 4, 5		

5.2 问题二模型的建立和求解

首先我们分析各门架不出现遗漏记录现象的条件：当车辆入站后，根据附件中的高速路段图，入站所记录的收费站点与第一次经过的门

架之间不会存在其它门架；并且从所记录的第一个门架编号开始，车辆轨迹对应的门架编号数据应该按顺序地出现，且间隔为 1。基于这样的漏记判定条件，再结合 MTTLAB 编程处理并统计，通过识别每个车辆类型和其轨迹数据，得到各门架对应各车辆类型的遗漏次数表，见表 3，

表 3. 各门架的记录遗漏次数表

门架编 号	客 车	货 车	其 他	门架编 号	客 车	货 车	其 他
1	0	107	0	9	0	580	0
2	0	90	0	10	0	634	0
3	0	93	0	11	0	571	0
4	0	1165	0	12	0	711	0
5	0	324	0	13	0	875	0
6	0	4373	0	14	0	323	0
7	0	482	0	15	0	361	0
8	0	107	0		0		0

通过观察表 3，我们得到：所有门架对客车与其它车辆不会出现漏记现象，只会对货车出现漏记。所以各门架在这 6 天内的总漏记次数等于对货车的总漏记次数。据此，我们结合第一问求得的各门架在的 2022 年 2 月 22 日至 2 月 27 日的六天时间内的总记录次数，得到计算门架 i 的漏记率的计算公式

$$r_i = \frac{\text{门架 } i \text{ 的漏记次数}}{\text{门架 } i \text{ 的漏记次数} + \text{门架 } i \text{ 的总记录次数}}$$

* MERGEFORMAT (5.1.1)

由此，我们可以通过简单的计算得到各门架的漏记率，见表 4

表 4. 各门架的漏记率表

门架编号	漏记率	门架编号	漏记率
1	0.27%	9	0.40%
2	0.19%	10	0.44%
3	0.18%	11	0.40%

4	1.77%	12	0.78%
5	0.41%	13	1.15%
6	6.81%	14	0.47%
7	0.34%	15	0.54%
8	0.27%		

为了更直观地展示各门架的漏记次数与漏记率的关系，我们将二者以组合图表的形式汇总在同一张图内，如图 3，

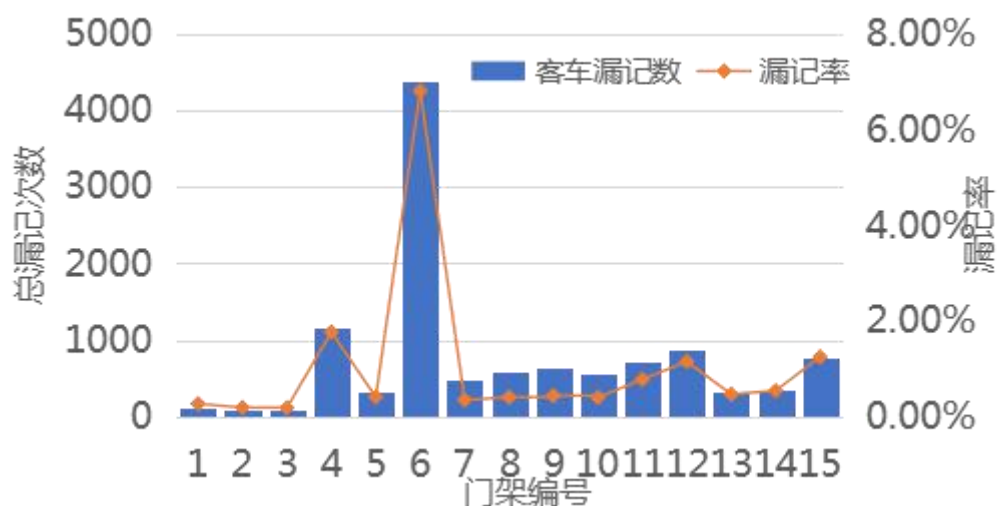


图 3. 各门架总漏记次数和漏记率统计图

从图中我们可以看出，各门架的漏记率与漏记次数高度相关，即漏记次数越高者漏记率越高。其中门架 7 对车辆的漏记率超过了 6.5%，明显高于其余门架；而门架 1、2、3、8 的漏记率不到 0.3%，属于漏记率较低的门架。

5.3 问题三模型的建立和求解

5.3.1 初步分析

为了确定说明收费站流量与路段流量的关系，首先要排除两者所在空间分布的干扰。这里我们绘制了各收费站和各路段分别在不同时段的流量变化图，如图 所示

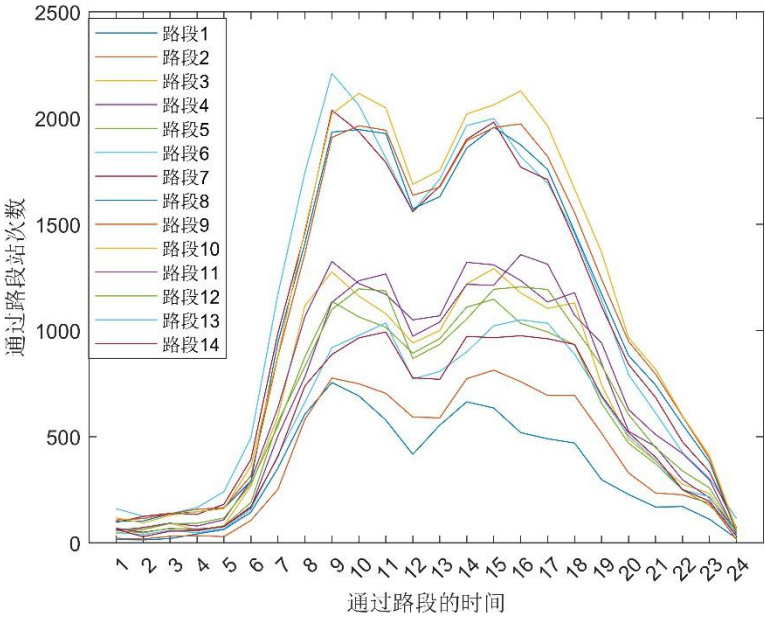


图 4. 不同路段下日均车流量

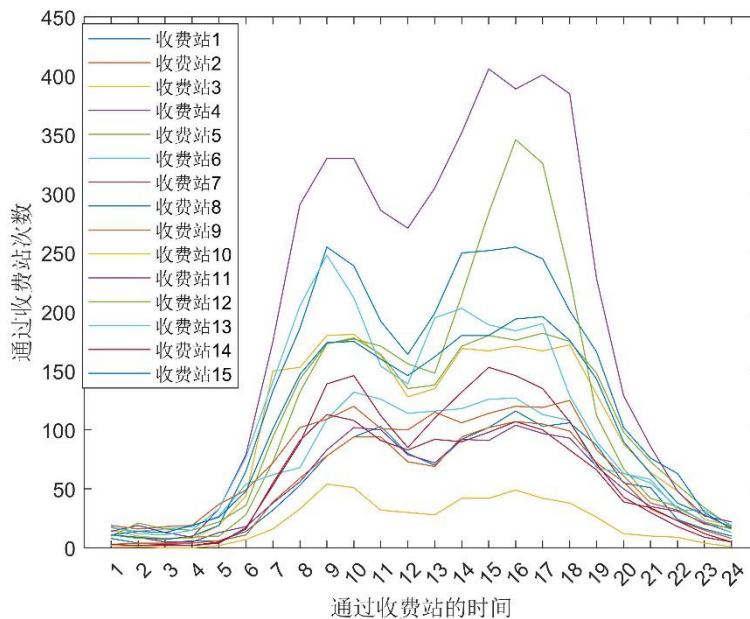


图 5. 不同收费站下日均车流量

从上图可以发现路段车流量和收费站车流量的变化趋势受其编号影响不大，只与其所在时段有关。固我们取所有收费站车流量和路段车流量关于时段的综合变化趋势，分析二者相关性。

5.3.2 散点图分析

为了更加直观方便的观察收费站车流量和路段车流量之间的关系，我们对各时段下的车流量数据进行可视化，画出散点图，如图所示

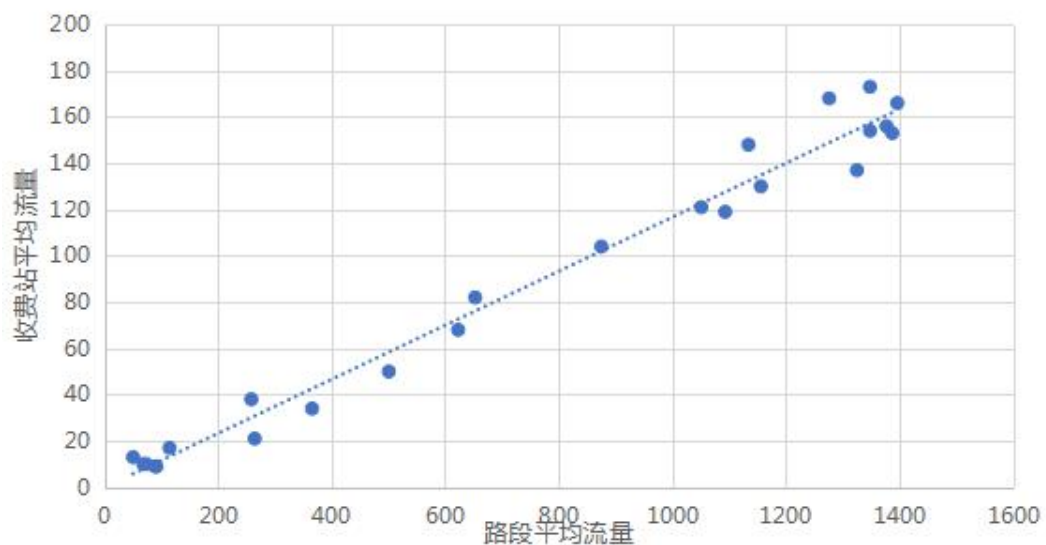


图 6. 路段平均流量与收费站平均流量散点图

通过观察散点图我们可以看出收费站车流量和路段车流量之间大致上成线性关系，猜测路段车流量与收费站车流量之间具有良好的线性关系，我们考虑用 Pearson 相关系数进一步探讨两者之间的关系。

5.3.3 相关分析

Pearson 相关系数是用来衡量两个变量呈线性相关程度的指标。当两个变量都大致呈正态分布且线性关系时，经常选用 Pearson 相关系数刻画两者的相关程度，公式如下：

$$\rho_{(x,y)} = cov(x,y)/(\sigma_x \sigma_y) = E[(x - x_i)(y - y_i)]/(\sigma_x \sigma_y)$$

* MERGEFORMAT (5.1.2)

其中 x 代表路段平均流量， y 代表收费站平均流量。相关系数 $\rho_{(x,y)}$ 越接近于 1，两者相关性越高。当 $\rho_{(x,y)}=1$ 时 x 与 y 之间完全相关；当 $\rho_{(x,y)} > 0.8$ 时， x ， y 之间称为高度相关；当 $\rho_{(x,y)} < 0.3$ 时， x ， y 之间称为低度相关，其余情况称为中度相关。在计算出相关系数后我们进行假设检验，我们设定原假设所得相关系数不可信，然后通过 SPSS 软件计算检验值。当检验值小于 0.01 时我们有 99% 的把握拒绝原假设，当检验值小于 0.05 时我们有 95% 的把握拒绝原假设，此时称相关性显著。我们用 SPSS 软件将数据进行了相关性分析与假设检验，分析结果如表 5 所示：

表 5. 相关系数及假设检验结果表

		路段流量	收费站流量
路段流量	皮尔逊相关性	1	.990**
	显著性（双尾）		.000

个案数		24	24
收费站流量	皮尔逊相关性	.990**	1
	显著性（双尾）	.000	
	个案数	24	24

**. 在 0.01 级别（双尾），相关性显著。

由表 5 可以看出，路段流量与收费站流量其相关系数都十分接近于 1，且假设检验的相关性十分显著。由此可知路段流量与收费站流量呈现出极强的线性相关性。

5.3.4 建立多服务台等待制模型

我们设车道每小时平均到达率为 λ (辆/h)，收费站收费系统平均服务数量为 μ (辆/h)，交通强度为 ρ ，每辆车通过的时间为 t_1 秒。则由排队论模型公式可得：

车道通行能力：

$$w_1 = \frac{3600}{t_1} \quad \backslash * \text{ MERGEFORMAT } (5.1.3)$$

收费站平均车辆数：

$$L_s = \frac{\rho}{1 - \rho} \quad \backslash * \text{ MERGEFORMAT } (5.1.4)$$

每辆车在收费站系统消耗的时间：

$$W_s = \frac{L_s}{\lambda} \quad \backslash * \text{ MERGEFORMAT } (5.1.5)$$

收费站系统中排队等待的车辆数：

$$L_q = L_s - \rho \quad \backslash * \text{ MERGEFORMAT } (5.1.6)$$

每辆车排队等待接收服务的时间：

$$W_q = \frac{L_q}{\lambda} \quad \backslash * \text{ MERGEFORMAT (5.1.7)}$$

收费站的忙期，即服务台工作的时间长度，可以用来体现收费站系统的利用率：

$$\rho = \frac{\lambda}{\mu} \quad \backslash * \text{ MERGEFORMAT (5.1.8)}$$

在单通道的情况下，若 $\rho > 1$ 将导致排队的队列无限远，则考虑运用服务台等待制模型 M/M/s/ ∞ ，服务台数目 s 的选取考虑排队系统在平稳状态下的概率分布情况即 $\bar{\rho}_s = \frac{\bar{\lambda}}{\mu s} < 1$ 的情况。

收费站空闲概率：

$$\bar{\rho}_0 = \left[\sum_{n=0}^{s-1} \frac{\bar{\rho}^n}{n!} + \frac{\bar{\rho}^s}{s! (1 - \bar{\rho}_s)} \right]^{-1}$$

$\backslash * \text{ MERGEFORMAT (5.1.9)}$

车辆到达时必须等待的概率：

$$\bar{c}(s, \bar{\rho}) = \frac{\bar{\rho}^s}{s! (1 - \bar{\rho}_s)} \times \bar{\rho}_0$$

$$\backslash * \text{ MERGEFORMAT (5.1.10)}$$

基于上式可得，收费站排队的平均队长为：

$$\bar{L}_q = \frac{\bar{\rho}^s \times \bar{\rho}_s \times \bar{\rho}_0}{s! \left(1 - \frac{\bar{\rho}}{s}\right)^2} \quad \backslash * \text{ MERGEFORMAT (5.1.11)}$$

5.3.5 模型求解

首先我们统计出六天每个车站每个时间段内通过 ETC 的车辆数求

其平均值得到表 6，这里我们只展示收费站 11 一天的车流量（续表见附录）

表 6. ETC 一天的车流量

时间段/收费站	收费站 11
0: 00-1: 00	7
1: 00-2: 00	5
2: 00-3: 00	2
3: 00-4: 00	3
4: 00-5: 00	17
5: 00-6: 00	43
6: 00-7: 00	104
7: 00-8: 00	168
8: 00-9: 00	174
9: 00-10: 00	172
10: 00-11: 00	158
11: 00-12: 00	141
12: 00-13: 00	151
13: 00-14: 00	186
14: 00-15: 00	223
15: 00-16: 00	216
16: 00-17: 00	220
17: 00-18: :00	214
18: 00-19: 00	120
19: 00-20: 00	68
20: 00-21: 00	44
21: 00-22: 00	23

22: 00-23: 00	11
23: 00-24: 00	10

由表 6 可得，不同时段的车流量不同，从早六点到晚六点车辆较多，其他时间段较少。由文献我们可以得知，一辆装有 ETC 的车通过收费口的时间约为 3 到 5 秒，我们假定取其中间值 4 秒，则由式 1 得出其 ETC 车道通行能力为 900(辆/h)，例如我们选取收费站 11 作为计算：

首先我们求出收费站 11 一天平均车流量 $\bar{\lambda} = \frac{2480}{24} \approx 103.3333$ 。

假设具有 s 个 ETC 服务平台，我们需要满足 $\bar{\rho}_s < 1$ ，由 $\bar{\rho} = \frac{\bar{\lambda}}{\mu} \approx 0.1148$ ，

而 $\bar{\rho}_s = \frac{\bar{\rho}}{s}$ ，即得出 $s > \bar{\rho}$ ，并且 s 必然为正整数，则 s 取 1，即开辟一条 ETC 通道，通过式 (5.1.9) 计算可得，收费窗口空闲概率：

$$\bar{\rho}_0 = \left[\frac{(0.1148)^0}{0!} + \frac{(0.1148)^1}{1! \times (1 - 0.1148)} \right]^{-1} = 0.8852$$

通过式 (5.1.10) 计算可得，车辆到达收费站等待的概率：

$$c(1, 0.1148) = \frac{(0.1148)^1 \times (0.8852)}{1! (1 - 0.1148)} = 0.1148$$

通过式 (5.1.11) 计算可得，收费站排队的平均队伍长度：

$$\bar{L}_q = \frac{(0.8852) \times (0.1148)^1 \times (0.1148)}{1! \times \left(1 - \frac{0.1148}{1}\right)^2} = 0.0149$$

由上述计算可知，每辆车到达收费站 11 的 ETC 通道时有 11.48% 的概率需要排队等待，由于我们要求需要顺畅通行，我们通过 matlab 求解出当所有收费站 ETC 通道个数为 1 时车辆到达收费站等待的概率

求其平均值为 4%，所以我们尽量将排队概率降到 4%及以下。

此时我们对收费站 ETC 通道进行增加，当 s 增加到两个时，即两条 ETC 通道，通过式 (5.1.9) 计算可得，此时收费窗口空闲概率：

$$\bar{\rho}_0 = \left[\frac{(0.1148)^0}{0!} + \frac{0.1148^1}{1!} + \frac{0.1148^2}{2! \times (1-0.0574)} \right]^{-1} = 0.8914$$

通过式 (5.1.10) 计算可得，车辆到达等待的概率：

$$\bar{c}(2, 0.1148) = \frac{0.1148^2 \times 0.8914}{2! \times (1-0.0574)} = 0.0063$$

通过式 (5.1.11) 计算可得，车辆排队的平均队长：

$$\bar{L}_q = \frac{0.0574 \times 0.1148^2 \times 0.8914}{2! \times \left(1 - \frac{0.1148}{2}\right)^2} = 0.0003$$

此时增加一个车道后，车辆等待概率为 0.6%，概率极低，即可顺畅通行。

其次我们统计出六天每个车站每个时间段内通过 MTC 的车辆数求其平均值得到表 2，详表见其附录。通过文献可知 MTC 处理一辆车的时间为 16 到 20 秒，我们依然取其中间值 18 秒，同理求出 MTC 车道通行能力为 200 (辆/h)，处理方式同上述收费站 11。通过对表一表二数据统计得知车辆通过 MTC 的数量约为 ETC 数量的两倍，利用 matlab 求解当 ETC 通道数量为 1 时所有收费站 ETC 通道的等待概率大部分概率均在 10%以上，但由于增加车道数会让排队概率大幅降低，并且由于 MTC 通过时间长于 ETC，故我们应当将汽车等待概率降到更低，由于 MTC 等待时间约为 ETC 的四倍，故我们设置将等待概率调到 1%及以下，最后利用 matlab 求解其 ETC、MTC 的车道数。

应急收费系统（便携式收费机或移动电站收费车），便携式收费机相当于一个便携式移动收费亭，可根据现场车流情况，实时调整其位置，单车道两车同时收费，车道“一变二”，有效提高车辆通过收费站的效率。由于应急收费系统只能用于 MTC 车道，所以根据原来结果上减少一条车道而采用此收费系统来代替，利用 Matlab 求解其结果见表格 7，

表 7. 各车道及应急收费系统数量

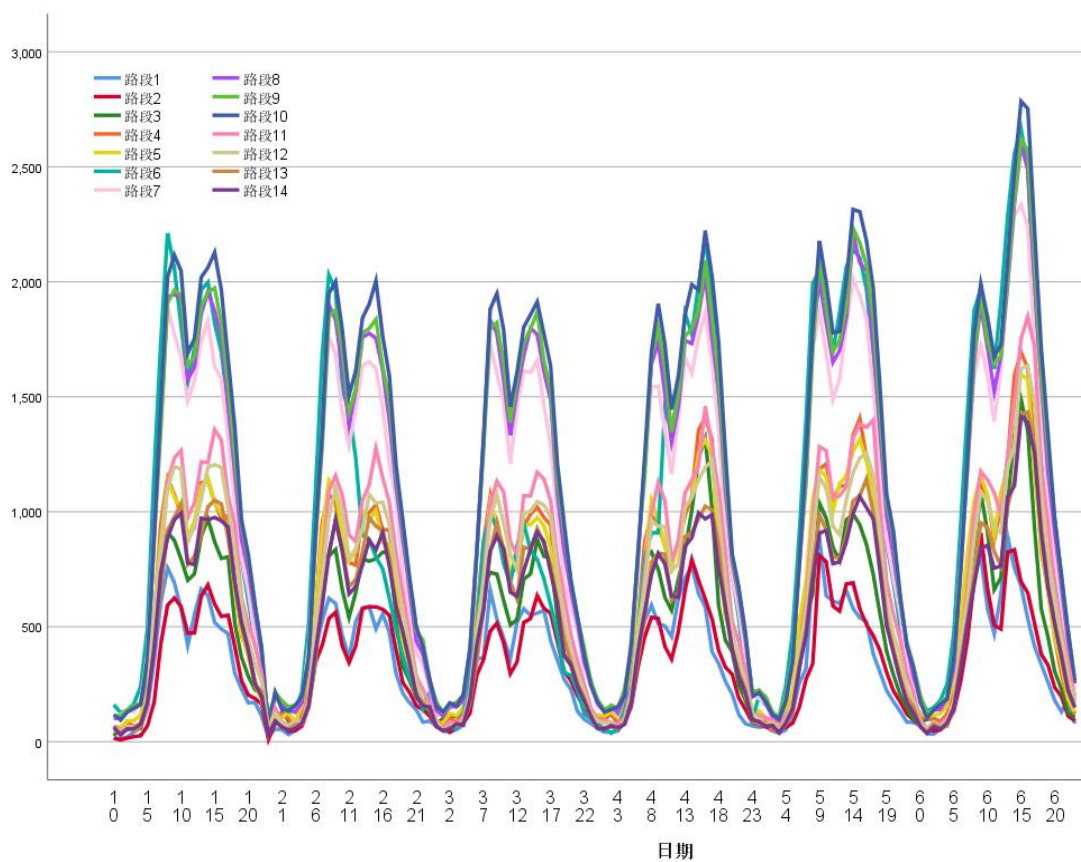
收费站编号	ETC 通道个数	MTC 通道个数	便携式收费机
1	1	1+2	1
2	1	1+2	1
3	1	1+2	1
4	1	1+2	1
5	1	1+2	1
6	1	1+2	1
7	1	1+2	1
8	2	1+2	1
9	1	1+2	1
10	1	1+1	1
11	2	1+2	1
12	2	1+2	1
13	2	1+2	1
14	1	1+2	1
15	1	1+2	1

由上述表格 7 可知我们共需要建造 19 条 ETC 通行车道,29 条 MTC 车道，且每个收费站配备一台便携式收费机即 15 台便携式收费机。

5.4 问题四模型的建立和求解

5.4.1 模型准备

首先，为了更准确的找出适合符合实际情况的车流量预测模型，我们基于问题三中对路段车流量的统计方法，进行更详细的处理，得到各路段在 6 天内各时段的车流量。图 是利用 SPSS 软件给出的所有



路段在 2022 年 2 月 22 日至 2 月 27 日这 6 天之内连续的车流量变化时间序列图，

图 7. 各路段车流量的时间序列图

从图中我们大致可以得到：各路段在 6 日之内的变化趋势十分相似，下面我们以路段 1 为例，建立指数平滑法模型来预测未来 24 时内的

车流量。

由于我们汇总统计出的 2022 年 2 月 22 日至 2 月 27 日的车流量数据无缺失值，而且属于“季度”数据(有 6 个周期)，对于路段 1，我们可以做出其时间序列图，如图 8

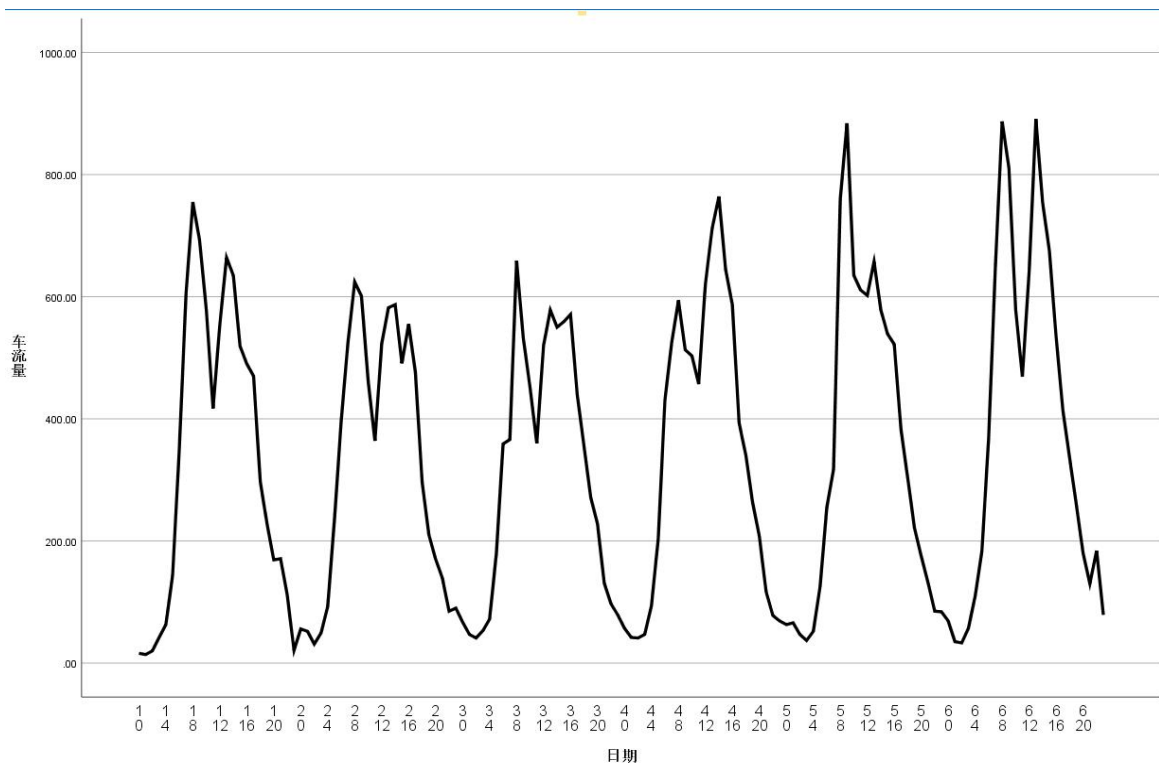


图 8. 路段 1 在 6 天内的车流量时间序列图

时间序列图中可以看出，车流量有明显的“季节性波动”，即每个周期内车流量的变化趋势相同，都成“M”型变化趋势。并且整体数据呈稳定的波动，即无明显的上升或下降的趋势。为了更客观的描述车流量的季节变化规律，我们对其数据进行加法时间序列分解得到表 8，

表 8. 路段 1 的时间序列分解表

周期	季节因子	周期	季节因子	周期	季节因子
1	-272.760	9	364.765	17	213.573
2	-288.068	10	328.090	18	100.411

3	-299.310	11	185.757	19	-13.627
4	-290.456	12	111.411	20	-93.481
5	-255.689	13	233.598	21	-143.014
6	-154.639	14	307.890	22	-195.860
7	22.432	15	292.148	23	-243.160
8	136.540	16	219.432	24	-265.981

从各时段的季节因子可得到，季节因子大于 0 的时段为高于日均车流量水平；反之，小于 0 的时段为低于日均车流量水平。例如，在时段 10 内，平均车流量高于每日平均水平约 328 次，在时段 2 内，平均车流量低于每日平均水平约 288 次。

SPSS 软件的专家建模器会查找每个相依序列的最佳拟合模型。对于指定的自变量，专家建模器考虑指数平滑法模型或 ARIMA 模型中的内容选择那些与该相依序列具有统计显著关系的模型。这里我们借助专家建模器得到了最适合所给数据的模型为指数平滑法模型中的**简单季节性**(Simple seasonal)。此模型适用于没有趋势并且季节性影响随时间变动保持稳定的序列，这也符合我们上文对车流量时间序列特征的分析结果。简单季节性模型的具体形式如式

$$\begin{cases} l_t = \alpha(x_t - s_{t-m}) + (1 - \alpha)l_{t-1} \\ s_t = \gamma(x_t - l_{t-1}) + (1 - \gamma)s_{t-m} \\ \hat{x}_{t+h} = l_t + s_{t+h-m(k+1)}, k = \left\lceil \frac{h-1}{m} \right\rceil \end{cases}$$

* MERGEFORMAT (5.1.12)

其中第一式为水平平滑方程，第二式为季节平滑方程，第三式则为预

测方程。而 m 表示周期长度(这里取 24)， α 表示水平的平滑参数， γ 表示季节的平滑参数， h 表示预测超前期数， \hat{x}_{t+h} 表示第 h 期的预测值。利用 SPSS 软件，我们得到了参数的估计值，见表 9

表9. 指数平滑法模型参数

模型		估算	标准误差	t		显著性
路段1-模型_1	不转换	Alpha（水平）	.800	.088	9.127	.000
		Gamma（季节）	5.974E-6	.259	2.306E-5	1.000

5.4.2 残差检验

接下来我们需要判断残差是否为白噪声，即该时间序列是否能被模型识别完全，图9分别是残差的ACF和PACF图，表10为残差的Q检验表，

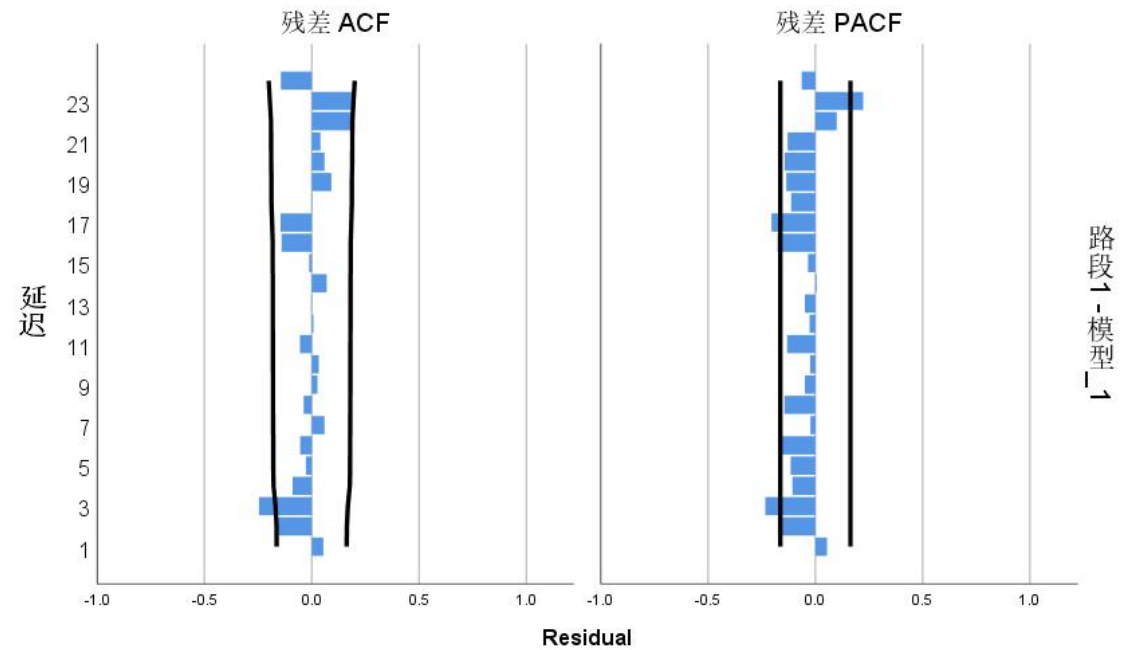


图 9. 残差 ACF、PACF 图

表10. 残差Q检验结果表

模型	预测变量数	模型拟合度统计	杨-博克斯 Q(18)			离群值数
		平稳 R 方	统计	DF	显著性	

路段1-模型_1	0	.546	24.526	16	.079	0
----------	---	------	--------	----	------	---

从残差的ACF和PACF图形中可以看出,几乎所有的滞后阶数的自相关系数和偏自相关系数均和0没有显著差异;另外从表10可以看出,对残差进行Q检验得到的p值为0.079,即我们在95%的置信区间内无法拒绝原假设,认为残差就是白噪声序列,因此简单季节模型能够很好的识别本例中的车流量数据。

5.4.3 模型求解

基于上述讨论,我们在95%的置信水平下可以给出路段1未来24小时的车流量预测值及置信区间(具体见表)。另外我们绘制了路段1关于车流量的实测、拟合和预测值的组合图,见图10,

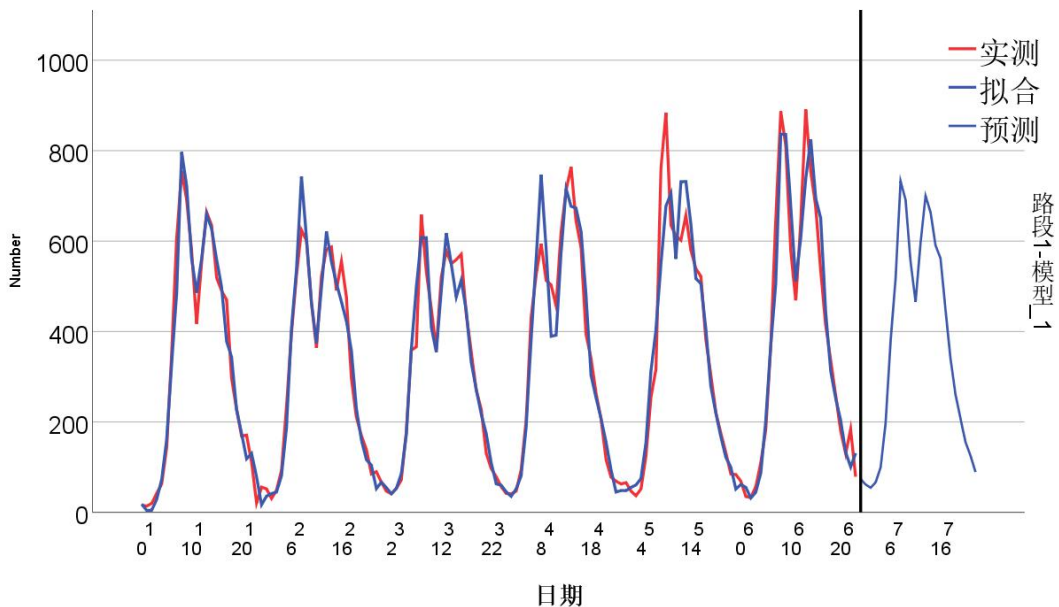


图 10. 路段 1 车流量拟合及预测效果图

直观上来看,真实数据和拟合数据的走势相同,拟合度较高,这说明简单季节模型对原始数据的拟合效果较好;另外,预测所得的第7天

的数据既保留了原始序列的季节效应,也同时保证了整体稳定的趋势,这说明简单季节模型能很好的对该路段的车流量数据进行预测。客观的看,可以根据具体的指标来评价模型。如本模型的平稳的 R 方为 0.546,较于其他模型中表现最好;而正态化 BIC 值为 8.078,相对较小,表示模型在复杂度方面也较为优良。

类似的,我们可以得到所有路段的在 2 月 27 日的车流量预测值和置信区间,由于数据量较大,表 11 仅展示了两个路段的预测值(详细预测数据的图表见附录)

表11. 两个路段的车流量预测结果表

时段编号	路段 1			路段 2		
	置信下界	预测值	置信上界	置信下界	预测值	置信上界
1	-35	74	182	-37	72	182
2	-77	62	201	-94	61	216
3	-109	55	218	-138	52	242
4	-119	67	252	-162	58	277
5	-105	100	304	-172	73	319
6	-25	197	420	-142	127	395
7	141	380	618	-26	264	554
8	263	517	771	91	402	712
9	464	732	1001	210	539	868
10	409	691	973	320	667	1013
11	259	554	849	193	557	921
12	158	465	773	55	435	815
13	277	596	915	74	469	865
14	369	700	1031	246	657	1067

15	322	664	1006	269	694	1119
16	237	590	943	201	640	1078
17	199	562	926	135	587	1039
18	74	448	822	70	535	1000
19	-43	340	724	-68	410	888
20	-132	261	654	-185	305	796
21	-195	207	610	-256	246	749
22	-256	156	568	-324	191	705
23	-295	126	547	-382	144	670
24	-340	89	519	-446	91	628

至此，本文完成了车流量的预测工作。下面我们会基于预测数据，讨论未来 24 小时各个路段的限速值，以达到最大限度提高道路网安全畅通的目的。这里，我们认为车流量与路段限速值有着负的线性关系。结合一般高速公路的限速值的上下限，我们根据预测数据便可推算出各路段在各时段的实时限速值。

根据《道路交通安全法实施条例》，高速公路限速值为 60—120km/h，在这里我们对所有路段的最低限速值都设置为 60km/h，根据预测数据只计算最高限速值。

故我们得到求解路段 i 在时段 j 限速值 sl_{ij} 与车流量 λ_{ij} 关系公式，

$$\frac{sl_{ij} - 120}{\lambda_{ij} - \min\{\lambda\}} = \frac{sl_{ij} - 60}{\lambda_{ij} - \max\{\lambda\}}$$

* MERGEFORMAT (5.1.13)

其中 $\max\{\lambda\}$ 、 $\min\{\lambda\}$ 表示已知数据中车流量的最大、小值。代入预测所得数据可求得结果。由于数据量较大，这里只展示路段 7 的限速表

(所有路段限速表见附录)

表 12. 路段 7 各时段的实时最高限速值表

时段	限速值(km/h)	时段	限速值(km/h)
0: 00-1: 00	116.4	12: 00-13: 00	86.9
1: 00-2: 00	116.7	13: 00-14: 00	81.9
2: 00-3: 00	116.7	14: 00-15: 00	80.1
3: 00-4: 00	116.8	15: 00-16: 00	80.2
4: 00-5: 00	115.6	16: 00-17: 00	82.0
5: 00-6: 00	111.1	17: 00-18: 00	88.1
6: 00-7: 00	101.7	18: 00-19: 00	94.4
7: 00-8: 00	92.8	19: 00-20: 00	101.3
8: 00-9: 00	82.9	20: 00-21: 00	105.0
9: 00-10: 00	82.9	21: 00-22: 00	108.7
10: 00-11: 00	86.4	22: 00-23: 00	112.4
11: 00-12: 00	90.7	23: 00-24: 00	115.2

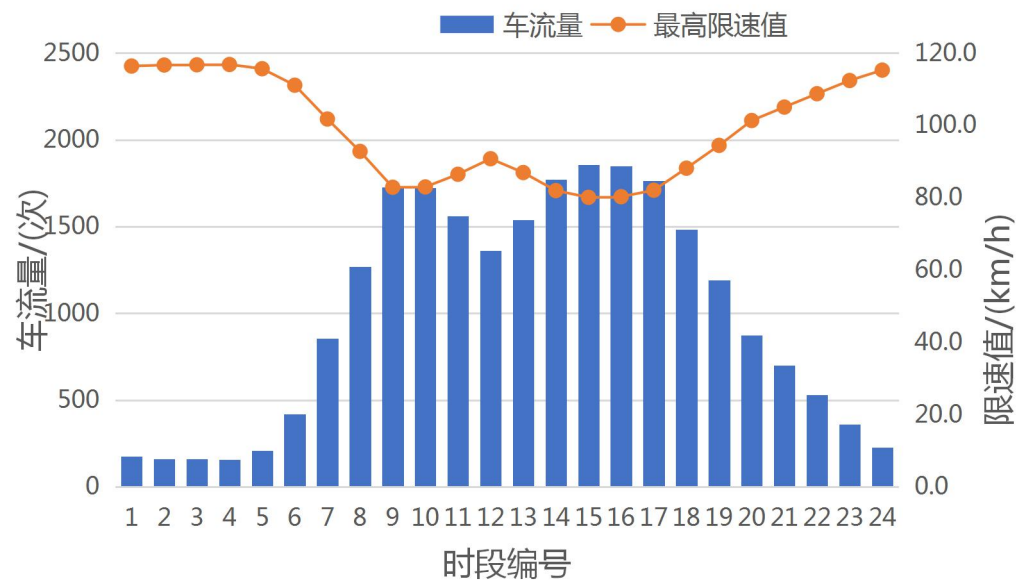


图 11. 路段 7 实时车流量与限速值关系图

从图中可以得到：一天内的车流量高峰时段对应的实时限速值较低，

最低时车辆时速最高行驶速度限制 80km/h 左右;而车流较少的时段对应的限速值设置较为宽松, 接近实际的限速规定 120km/h, 这也符合现实的行驶规律。综上, 我们便完成了对问题四的全部求解。

六、 模型的评价与推广

6.1 模型优点

(1). 程序对大量数据的处理效率高, 对相关信息的统计和筛选准确率高, 运行速度快;

(2). 各模型都通过了的评价指标的检验, 误差较小, 表示模型的可信度较高;

(3). 本文建立的排队论模型能与实际紧密联系, 结合实际情况对所提出的问题进行求解, 使模型更贴近实际, 通用性、推广更强;

(4). 本文的时间序列分析模型结合了序列特征而构建, 使得模型的可靠性高、移植性强, 对各路段的车流量数据都有较好的匹配性;

6.2 模型缺点

(1). 本文对维修时段和关于车流量与限速关系的假设较强, 舍去了一些影响小的因素, 虽然很大程度上简化了模型, 但降低了其准确度;

(2). 由于所给的数据存在一定的局限性, 例如关于补录出入站的数据, 可能对相关车流量的统计有所影响, 造成统计值与实际值偏小的现象;

6.3 模型改进与推广

(1). 改进:；问题四中关于限速值与车流量的关系模型，可以考虑采用更精确的非线性模型来提高对限速值设置的合理性和准确性。

(2). 推广:本文所构建的排队论模型有很好的通用性和实际意义，结合统计的数据，可以利用模型解决现实社会中的其他排队问题或站点数设置问题；时间序列分析模型有很强的移植性，通过选择合适的指数平滑法模型或 ARIMA 模型可以解决类似的，例如某地铁站的人流量的时间序列分析或预测问题等。

参考文献:

- [1] 冯尧, 袁小青, 敬梦娇. 高速公路收费站 ETC 车道配置策略研究[J]. 内江科技. 2021(10):57-58+84
- [2] 王涛. 车速离散性对高速公路通行效率的影响研究[C]. 东南大学, 2016.
- [3] 武文瑜. 基于时空特征的短期车流量预测研究[C]. 西安理工大学, 2021.
- [4] 张方方. 设置 ETC 车道的高速公路收费站车道规模确定方法探讨[J]. 上海公路. 2013(02):58-61+14

附录

代码及数据文件	
data1.mat	门架及时间预处理数据
data2.mat	门架及车辆特征预处理数据
data33.mat	收费站及时间预处理数据
data4.mat	门架及日期预处理数据
door.mat	各门架日均车流量数据
forecast.mat	时间序列预测数据
code1.m	计算各门架日均车流量门架
code2.m	计算各门架总漏记次数和漏记率
code3_E.m	计算 ETC 车道数
code3_M.m	计算 MTC 车道数及各收费站总流量
code4.m	计算各路段车流量时间序列，各路段日均流量及限速值设置
图表附录	
表 1. 各门架在各时段的平均流量分布表	
图 1. 各门架车流量的日均时间变化趋势图	
表 2. 各收费站 ETC 日均车流量	
表 3. 各收费站 MTC 日均车流量	
表 4. 各路段的时间序列分解表	
表 5. 各路段指数平滑法模型参数表(从路段 2 开始)	
图 2. 各路段残差的 ACF、PACF 图	
表 6. 各路段残差 Q 检验结果表	
表 7. 各路段的车流量预测结果表(从路段 3 开始)	
图 3. 各路段车流量拟合及预测效果图	
表 8. 各路段在各时段的实时最高限速值表	