

陈红,李晨光,刘爽,等.考虑群体差异的地铁出行时空特征及影响因素[J].长安大学学报(自然科学版),2025,45(1):114-124.
CHENG Hong,LI Chen-guang,LIU Shuang,et al.Spatio-temporal characteristics and influencing factors of subway travel considering group differences[J].Journal of Chang'an University (Natural Science Edition),2025,45(1):114-124.

DOI:10.19721/j.cnki.1671-8879.2025.01.010

考虑群体差异的地铁出行时空特征及影响因素

陈红,李晨光,刘爽,刘恩泽,姚振兴

(长安大学运输工程学院,陕西西安710064)

摘要:为分析地铁站点周边建成环境对不同年龄群体出行时空特征的影响,以西安市为例,结合手机信令数据、兴趣点数据和土地利用数据,提取老年群体和非老年群体的轨道交通客流量及4类建成环境变量。对出行时空特征进行深入分析的基础上,使用轻量级梯度提升机(LightGBM)探讨地铁站点周边建成环境对不同群体客流量的特征重要度和非线性影响。研究表明:LightGBM模型在拟合和预测不同年龄群体出行特征方面表现出优异的效果,显著优于传统的梯度提升决策树(GBDT)模型和线性回归(LR)模型;老年群体早晚高峰呈双峰形,老年群体无明显峰值;2类群体地铁出行距离主要集中在7~20 km,活动范围主要位于西安市三环内,但三环外热门景区附近的地铁站点吸引了大量非老年群体;建成环境因素对2类群体地铁出行的影响重要度存在差异,交通相关特征是影响2类群体地铁客流量最重要的特征,对老年群体的影响大于非老年群体,居住人口密度和工作人口密度对老年群体的影响小于非老年群体;医疗中心数量、科教文化数量、公司企业数量、公交站点密度与2类群体客流量呈非线性正相关,且有明显的阈值效应,当医疗中心数量达到58、科教文化数量为48时,促进作用最明显。研究结果对构建老年友好型城市体系具有重要意义,为制定增强老年人出行能力的相关政策提供参考。

关键词:交通工程;城市交通;轻量级梯度提升机;群体差异;建成环境;老年人

中图分类号:U491.17 **文献标志码:**A **文章编号:**1671-8879(2025)01-0114-11

Spatio-temporal characteristics and influencing factors of subway travel considering group differences

CHEN Hong, LI Chen-guang, LIU Shuang, LIU En-ze, YAO Zhen-xing

(School of Transportation Engineering, Chang'an University, Xi'an 710064, Shaanxi, China)

Abstract: To analyze the impact of the built environment around metro stations on spatio-temporal travel characteristics for different groups, Xi'an was conducted as an example, mobile signaling data, points of interest (POI) data and land use data were integrated to extract metro ridership for elderly and non-elderly groups, along with four categories of built environment

收稿日期:2024-09-11

基金项目:国家自然科学基金项目(52002030)

作者简介:陈红(1963-),女,湖南湘潭人,教授,博士研究生导师,E-mail:glch@chd.edu.cn.

通讯作者:姚振兴(1989-),男,浙江金华人,副教授,工学博士,E-mail:yaotraffic@chd.edu.cn.

variables. Based on an in-depth analysis of spatio-temporal travel characteristics, the light gradient boosting machine (LightGBM) was applied to explore the feature importance and nonlinear effects of the built environment around metro stations on ridership for different groups. The results show that the LightGBM model outperforms traditional gradient boosting decision tree (GBDT) models and linear regression (LR) models in fitting and predicting travel characteristics. The elderly group exhibits a bimodal distribution during morning and evening peak hours, whereas no distinct peak is observed for the non-elderly group. The travel distance for both groups predominantly ranges between 7 to 20 km, with activity areas mainly located within the third ring road of Xi'an. However, metro stations near popular attractions outside the third ring road attract a significant number of non-elderly passengers. The importance of built environment factors on metro ridership varies between the two groups. Transportation-related features are the most significant factors influencing both groups, with a greater impact on the elderly group compared to the non-elderly group. Residential population density and employment density have a lesser impact on the elderly group than the non-elderly group. The number of medical centers, educational and cultural facilities, companies, and bus stop density are positively correlated with ridership for both groups, showing clear threshold effects. Specifically, the number of medical centers has the most pronounced effect at 58, while educational and cultural facilities have the greatest impact at 48. These findings are crucial for building an age-friendly urban system and provide valuable insights for formulating policies to enhance the mobility of the elderly. 6 tabs, 8 figs, 26 refs.

Key words: traffic engineering; urban traffic; LightGBM; group difference; built environment; elderly

0 引言

随着城市不断扩张和过度依赖机动化,交通拥堵和空气污染等问题日益突出。地铁出行作为一种解决这些城市问题有效且可持续的选择,受到了越来越多的关注。然而,由于人口老龄化加剧,地铁系统面临新的挑战。老年人口的增加以及他们社会地位的变化导致了其在出行方面的独特需求和特征,与非老年群体存在显著差异。为了提高轨道交通的舒适度和公平性,轨道交通发展需要关注不同群体的出行差异。研究表明,随着年龄增长,老年人更倾向于选择在早高峰和晚高峰后出行,与此同时,城市地铁站的客流量与周边建成环境密切相关,建成环境被认为是影响客流量的关键驱动因素,也是影响地铁出行行为的重要因素^[1-3]。因此,研究不同群体在地铁出行方面的时空特征以及建成环境对其影响极为重要。

目前关于地铁出行时空特征的研究主要侧重于定量分析居民的出行规律,包括出行时长、距离、出行方向、方式选择等方面。Sun等分析了地铁开通后对附近居民出行行为的影响,发现附近居民通勤

出行的平均地铁使用时间为 30.9 min/d,非工作出行的平均地铁使用时间为 16.6 min/d^[4]。何伟等基于深圳地铁智能卡数据对地铁群组进行了模式挖掘,发现深圳市居民在工作日的地铁出行呈现明显的潮汐现象,尤其早高峰时客流量和空间覆盖范围明显大于晚高峰;也有研究发现地铁平均通勤距离为 12.4 km/d,通勤出行方向呈现明显向心性^[5-6]。在方式选择方面,朱佳欣研究发现,通勤者转变出行方式的意愿明显受习惯和线路调整感知的影响,习惯于公共交通工具通勤的个体受更易被轨道线路调整变化影响^[7]。兰招娣等发现上班时间与学校上课时间越宽裕,家庭通勤者将会越倾向于乘坐地铁出行^[8]。除此之外,也有研究者比较了不同群体的出行特征差异。杨红等通过对比老年群体和通勤群体的出行差异,发现老年群体地铁出行频率更高、出行距离更短,出行时长集中在 45 min/d 以内^[9]。熊美成等进一步对比低龄老年群体和高龄老年群体的地铁日常出行和活动模式,发现低龄老年和高龄老年群体的活动空间分布存在明显差异,其中高龄老年群体的活动空间向居住密集区聚集,且出行在较大程度上受建成环境的制约^[10]。

近年来,建成环境特征对地铁客流量的影响受到了广泛关注。Chen 等研究了纽约市地铁乘客的日变化规律,划分出了 5 类主要模式,并指出车站乘客量日变化规律与当地建成环境密切相关^[11]。影响轨道交通客流量的建成环境要素主要包括开发密度、混合土地利用、目的地可达性、到站点的距离及街道设计等,其中开发密度以及用地混合度对地铁客流量的影响极为显著^[12]。此外,公共交通接驳及轨道线网覆盖率等也对车站客流量有显著影响,具有良好的街道网络衔接也有助于增加地铁客流量^[13-14]。现有研究多使用统计方法来探究地铁出行的影响因素,Ding 等使用多元分析模型分析了不同时间段北京地铁客流量与建成环境间的统计关系^[15]。Yu 等使用多项 Logit 模型,通过捕捉轨道站点的相似性变化来研究建成环境对站点出行量的异质性关系^[16]。Pang 等采用多项逻辑回归分析来探索乘客模式与建成环境间的关联,发现多样性、密度和区位因素与大多数车站功能类型显著相关,而目的地可达性仅与就业型车站呈正相关。然而,建成环境和交通客流量间的关系不是简单的线性关系,有研究发现建成环境要素对地铁客流量的影响存在阈值,并呈复杂的非线性效应^[17]。Ding 等应用梯度增强决策树来研究建成环境变量对车站上车人数的非线性影响时发现,建成环境变量对地铁客流量具有阈值效应^[14]。Liu 等采用梯度提升决策树(GBDT)模型,通过探讨始发地和目的地建成环境对重庆地铁客运量的非线性影响,发现到市中心的距离、公司数量和建筑量对地铁乘客量有显著的积极影响,且始发地的建成环境对地铁客流量的影响大于目的地的建成环境^[18]。Champagne 等研究发现,新建地铁站尤其在开通之后对新建商业建筑的影响主要集中在地铁站附近 750 m 范围内,在 0~250 m 是消极影响,而 250~750 m 是积极影响^[19]。Wu 等在对比 2005 年和 2009 年北京地铁基础设施建设情况对通勤行为影响时发现,地铁站的可达性水平与地铁使用呈正相关,同时还认为地铁站数量变化与地铁使用间的关联不显著,可能此时地铁站数量还未达到某阈值^[20]。

然而,既有地铁出行行为研究往往忽略了年龄差异所带来的影响因素。Kim 等研究发现老年人的出行时间预算(TTB)主要受收入、汽车拥有量等经济条件的影响,而非老年人的出行时间预算则受就业类型等活动条件的影响^[21]。Cheng 等利用 2013 年南京旅游调查数据发现,居住区位对出行行

为的影响在老年人和年青人之间存在显著差异。对老年人来说,生活环境变化对出行频率和出行时间的影响更为重要;对年青人来说,自我选择效应影响要大于对老年人的影响。研究结果表明,在地铁出行行为与建成环境的研究中,应充分考虑年龄差异的影响^[22]。

本文使用西安市的手机信令数据、兴趣点(POD)数据以及土地利用数据。首先分析不同年龄群体的出行特征,随后深入探究不同年龄群体与建成环境间的非线性关系。研究有助于预测不同年龄的出行需求,同时为改善交通服务环境提供依据。

1 数据准备

1.1 研究区域

西安市至 2022 年底,常住人口达 1 299.59 万。西安统计年鉴显示,市内老年人口比例逐年增长,人口老龄化趋势明显。根据最新的第七次中国人口普查,60 岁及以上人口为 2 075 318 人,占总人口的 16.02%,相较于第六次普查,增长了 3.48%。到 2021 年 7 月,西安市已开设 8 条地铁线(1、2、3、4、5、6、9 号和 14 号),共计 146 个站点,如图 1 所示。

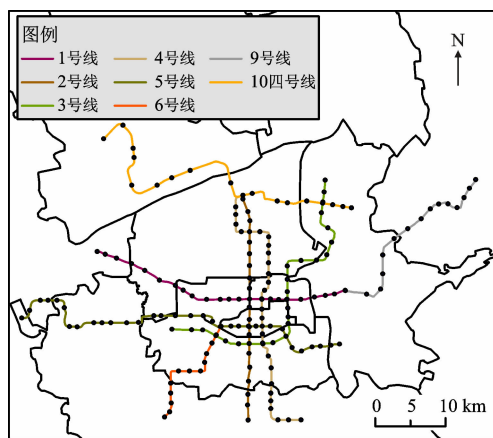


图 1 研究区域

Fig. 1 Research area

1.2 手机信令出行数据

手机信令出行数据由“智慧足迹”公司提供,原始数据来自中国联通公司。为保证数据安全,平台将用户 ID 设置为违规字段(用户无法查询),对年龄进行分段编号。对于用户端,平台仅供导出集计数据,比如各个站点的出发量和到达量。

数据结构如表 1 所示,其中,weight 标签为智慧足迹公司提供,通过比联通用户和全量人口的时空分布状况,使用后台算法识别得出,为后续出行

表 1 手机信令数据结构
Tab. 1 Structure of mobile signal data

字段	内容	样例
Uid	匿名者 ID	*** 92
stime	经过地铁站的时间	2021-7-1 **: **: **
station	地铁站名称	三元桥站
age	年龄段编号	6
weight	将联通用户数外推到全量人数的比例因子	3. 285
flag	经过类型	出站

量扩样的比例因子。stime 标签可以提取出地铁的出行日期,age 标签可以识别出不同年龄群体,flag 标签可以识别出进站与出站站点,用于提取站点出发量。

1.3 客流数据与建成环境变量构造

本文定义 60 岁以下为非老年群体,60 岁及以

上为老年群体,使用 2021-07-01~2021-07-07 共 1 周的手机信令数据,通过剔除数据异常值,进行集计等处理,最终保留的字段包括出发站点、终点站点、OD 出行量。其中,老年群体日均出行量为 5. 71 万人次,非老年人群体日均出行量为 251. 48 万人次,这与 2021 年西安地铁年日均客流 281. 14 万人次基本一致。根据出发站点进行集计,进一步得到老年群体和非老年群体在 1 周内的站点进站人数。

同时,为了探究建成环境与不同年龄群体地铁出行行为的关系,考虑到 15 min 步行距离阈值,本文以地铁站点周边 800 m 缓冲区作为研究区域,从社会经济特征、交通相关特征、土地利用特征和地铁系统特征 4 个方面选取了 13 个变量,变量描述性统计如表 2 所示。

表 2 变量描述性统计
Tab. 2 Descriptive statistics of variables

变量类型	变量名称	变量解释	最大值	最小值	均值	标准差
因变量	老年群体出行量	老年人在 1 周内的进站数	675. 20	5. 79	260. 91	171. 06
	非老年群体出行量	非老年人在 1 周内的进站数	543 508. 22	4 074. 50	129 439. 70	119 065. 06
社会经济特征	居住人口数量	站点影响范围内的居住人口数	76 913. 93	241. 74	25 221. 71	19 693. 41
	工作人口数量	站点影响范围内的工作人口数	50 130. 49	195. 49	13 674. 05	11 141. 62
交通相关特征	路网密度	站点影响范围内的路网长度/km	35. 36	2. 67	15. 33	7. 05
	公交站点密度	站点影响范围内的公交站点数量	50. 00	0. 00	9. 38	6. 13
土地利用特征	生活服务数量	站点影响范围内的生活服务数量	628. 00	1. 00	160. 77	139. 97
	医疗中心数量	站点影响范围内的医疗中心个数	242. 00	0. 00	49. 39	43. 88
	餐饮美食数量	站点影响范围内的餐饮美食数量	1 012. 00	0. 00	224. 07	204. 15
	科教文化数量	站点影响范围内的科教文化数量	228. 00	0. 00	61. 10	53. 33
	公司企业数量	站点影响范围内的公司企业数量	635. 00	2. 00	91. 05	99. 80
	休闲娱乐数量	站点影响范围内的休闲娱乐数量	105. 00	0. 00	14. 77	20. 39
	用地混合度	站点影响范围内的 POI 类型的熵指数	2. 69	1. 49	2. 32	0. 21
地铁系统特征	是否为换乘站	站点是否为换乘站(是 1,否 0)	1. 00	0. 00	0. 09	0. 28
	是否为终点站	站点是否为终点站(是 1,否 0)	1. 00	0. 00	0. 08	0. 27

社会经济特征主要由居住人口数和工作人口数描述。数据来源于“智慧足迹”平台,该平台将 21:00~08:00 停留时间最长的位置定义为居住位置,如果用户在该位置的月度驻留次数超过 10 次,标记用户为该位置的居住用户。同理,将工作日 09:00~15:00 停留时间最长且不是居住位置定义为工作位置,如果月度驻留次数超过 10 次,标记为工作用户。之后,分别对轨道车站影响范围内的居住用户以及工作用户进行集计,得到居住与工作人口数。

交通相关特征主要由路网密度和公交站点密度来表示,在 OpenStreetMap(OSM)平台获取西安市

的路网,利用 Python 的 geopandas 数据库对路网进行处理得到路网密度,选择高德地图作为公交站点的数据来源,通过 API 接口获取公交站点的位 置,统计轨道站点影响范围内的站点数量。

土地利用特征包括 6 类 POI 兴趣点数量(医疗中心、科教文化、公司企业、休闲娱乐、餐饮美食、生活服务)以及用地混合度。其中,POI 兴趣点均是通过高德开发平台 API 接口获取并统计缓冲区内数量。用地混合度表示土地的利用多样性和均衡性,通过 Shannon’s 熵指数进行计算^[23]

$$H_k = - \sum_{h=1}^6 q_{kh} \ln(q_{kh})$$

式中: H_k 为第 k 个缓冲区的 Shannon’s 熵指数; q_{kh}

为第 k 个缓冲区内第 h 种土地类型所占比率。

地铁系统特征包括站点是否为换乘站以及是否为终点站,通过调用 Python 的 networkx 数据库计算每个站点的节点度,节点度值为 1 是终点站,节点度值大于 2 是换乘站。

2 研究方法

2.1 梯度提升决策树

梯度提升决策树是一种集成方法,由决策树学习和梯度提升 2 种技术组成。该方法以决策树作为基本模型,决策树是根据数据进行预测的弱学习器。第 m 次迭代预测值如式(2)所示

$$f_m(x) = b_m I(x; R_m) \quad (2)$$

式中: $f_m(x)$ 为第 m 次迭代中样本 x 的预测值; b_m 为 R_m 的常数值; R_m 为第 m 次迭代决策树叶子节点所对应的区域; $I(x; R_m)$ 表示 x 是否属于 R_m ,若是,则 $I(x; R_m) = 1$,否则为 0。

梯度提升技术将许多基本决策树组合成模型 $f(x)$,其目标是 minimized 损失函数 $L(y, f(x))$ 的值。具体来说,通过顺序合并新树来纠正先前树产生的错误,并通过最小化负梯度后的损失函数来拟合附加树。因此, $f_m(x)$ 更新如式(3), b_m 更新如式(4),损失函数使用均方误差计算,如式(5),即

$$f_m(x) = f_{m-1}(x) + \xi b_m I(x, R_m) \quad (3)$$

$$b_m = \arg \min_b \left\{ \sum_{i=1}^N L[y_i, f_{m-1}(x_i) + b I(x; R_m)] \right\} \quad (4)$$

$$L[y, f(x)] = \frac{1}{N} \sum_{i=1}^N (f(x_i) - y_i)^2 \quad (5)$$

式中: ξ 为学习率; N 为样本数; $f(x_i)$ 为第 i 个样本的预测值; y_i 为第 i 个样本的真实值; x_i 为第 i 个样本; b 为在当前模型迭代中决策树的某个叶节点的预测值。

2.2 LightGBM

LightGBM 是一种高效的 GBDT 算法实现,相较于传统 GBDT 方法,它具有更高的速度和准确率。通过将连续特征离散化并采用区间统计和梯度累加的方法,LightGBM 大幅提高了存储、计算效率,同时也增强了模型的稳定性和鲁棒性。此外,该算法采用叶子生长策略,优先选择增益最大的叶子节点进行扩展,并限制最大深度以防止过拟合。通过将互斥的稀疏特征捆绑在一起,并通过贪心搜索和互斥特征合并算法降低特征维度,进一步减少了特征处理的时间复杂度和内存需求。本文采用

LightGBM 模型探究建成环境变量与不同群体地铁出行量之间的关系,算法框架如图 2 所示。

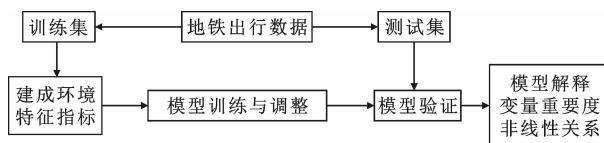


图 2 LightGBM 框架

Fig. 2 Framework of LightGBM

本文将数据集随机分为训练数据集和验证数据集(比例为 7 : 3),其中,训练数据集用于模型拟合。通过网格搜索算法调整正则化系数 $l1$ (λ_{l1})、正则化系数 $l2$ (λ_{l2})、叶子节点最少样本数 ($\min_data_in_leaf$)、叶子节点数量 (νm_leaves) 和特征采样比例 ($feature_fraction$) 等几个超参数,找到最佳参数组合。使用验证数据集检验 LightGBM 的预测能力,通过判定系数 R^2 、均方根误差 R_{MSE} 、平均绝对误差 M_{AE} 等统计指标对模型的性能进行评价。通过特征重要性和部分依赖图对模型进行解释。其中,特征重要性是指该特征在模型使用时对均方误差减少的贡献度,计算见式(6),即

$$F_{x_i}^1 = \frac{1}{T} \sum_{t=1}^T (M_{SE_{x_i}}^t - M_{SE_{x_i}}^t) \quad (6)$$

式中: $F_{x_i}^1$ 为变量 x_i 的重要性; T 为随机森林模型中决策树的个数; $M_{SE_{x_i}}^t$ 为决策树 t 包含变量 x_i 之前的均方误差; $M_{SE_{x_i}}^t$ 为决策树 t 包含变量 x_i 之后的均方误差。

影响因素对预测结果的边际效应见式(7),即

$$F(h_j) = \frac{1}{N} \sum_{i=1}^N f(h_j, h_{i./j}) \quad (7)$$

式中: $F(h_j)$ 为第 j 个影响因素子集 h_j 对预测结果的平均贡献; $f(h_j, h_{i./j})$ 为 h_j 与其他影响因素子集 $h_{i./j}$ 的预测结果。

3 结果分析

3.1 不同年龄群体时空特征分析

图 3 为不同群体出行时空特征,图 3(a)显示了老年群体与非老年群体出行时刻的差异,非老年群体的出行早高峰和晚高峰集中在 07:30~08:30、17:00~18:30,而老年群体没有明显的早晚高峰出行,同时,老年群体出行量最大的 2 个时刻为 12:00 和 21:00。可以推断出,大部分老年人已经退休,考虑安全与舒适度等因素,老年人更愿意错峰出行。

图 3(b)揭示了 2 类群体的出行距离差异,75%

的老年群体和非老年群体地铁出行距离都集中在 20 km 内,约 20%的地铁出行距离在 7 km 以下,结合 2 类群体出行距离的差异可以发现,老年群体平均出行距离低于非老年群体。这表明随着年龄的增长,居民会缩小活动范围。

图 3(c)、图 3(d)分别揭示了老年群体和非老年群体的出行 OD 分布状况,发现老年群体和非老年群体的地铁活动范围主要集中在西安市三环内(包括莲湖区、碑林区、雁塔区、新城区),这些区域是西安市的主要市区。此外,非老年群体在西安市三环外的华清池、秦陵西、韦曲南站出行量较多,主要原因可能是这些站点在华清池、兵马俑、终南山等热门景点的附近。

3.2 LightGBM 模型结果

3.2.1 模型确定及精度评价

通过网格搜索算法,针对 2 类群体分别确定了 LightGBM 模型的最优参数组合(表 3),该参数组合通过控制正则化系数、叶子节点数量和特征采样比例等参数,防止模型过拟合,同时也显著增强了模型的预测精度。此外,采用五折交叉验证进一步评估网格搜索优化的 LightGBM 模型鲁棒性,通过计算获得每组测试集的 R^2 、 R_{MSE} 和 M_{AE} ,结果如表 4 所示。由表 4 可知,在 5 个子集中,对于老年群体, R^2 为 0.51~0.68, R_{MSE} 为 69.54~93.05, M_{AE} 为 65.53~98.79;对于非老年群体, R^2 为 0.49~0.60, R_{MSE} 为 60 369.51~88 495.36, M_{AE} 为 47 437.80~62 418.56。结果表明 LightGBM 对 2 类群体均表现出较好的鲁棒性。

将 LightGBM 的结果与残性回归(LR)和 GBDT 的结果进行比较。以 M_{AE} 、 R_{MSE} 和 R^2 为指标评价不同模型的效果,如表 5 所示。由表 5 可知,相比于 GBDT 模型,老年群体和非老年群体的 R^2 分别提高了 6.15%和 3.09%, R_{MSE} 分别降低了 6.18%和 15.19%, M_{AE} 分别降低了 13.15%和 11.31%。相比于 LR 模型,老年群体和非老年群体的 R^2 分别提高了 10.77%和 20%, R_{MSE} 分别降低了 9.67%和 9.26,老年群体的 M_{AE} 降低了 13.18%,而非老年人群体的 M_{AE} 与 LR 模型持平。总体而言,LightGBM 模型在精度和拟合性能方面均优于 GBDT 和 LR 模型。

3.2.2 特征变量重要性

影响 2 类群体轨道出行量的特征变量重要度如表 6 所示。由表 6 可知,影响老年群体和非老年群体地铁客流的建成环境变量重要度存在差异。社会

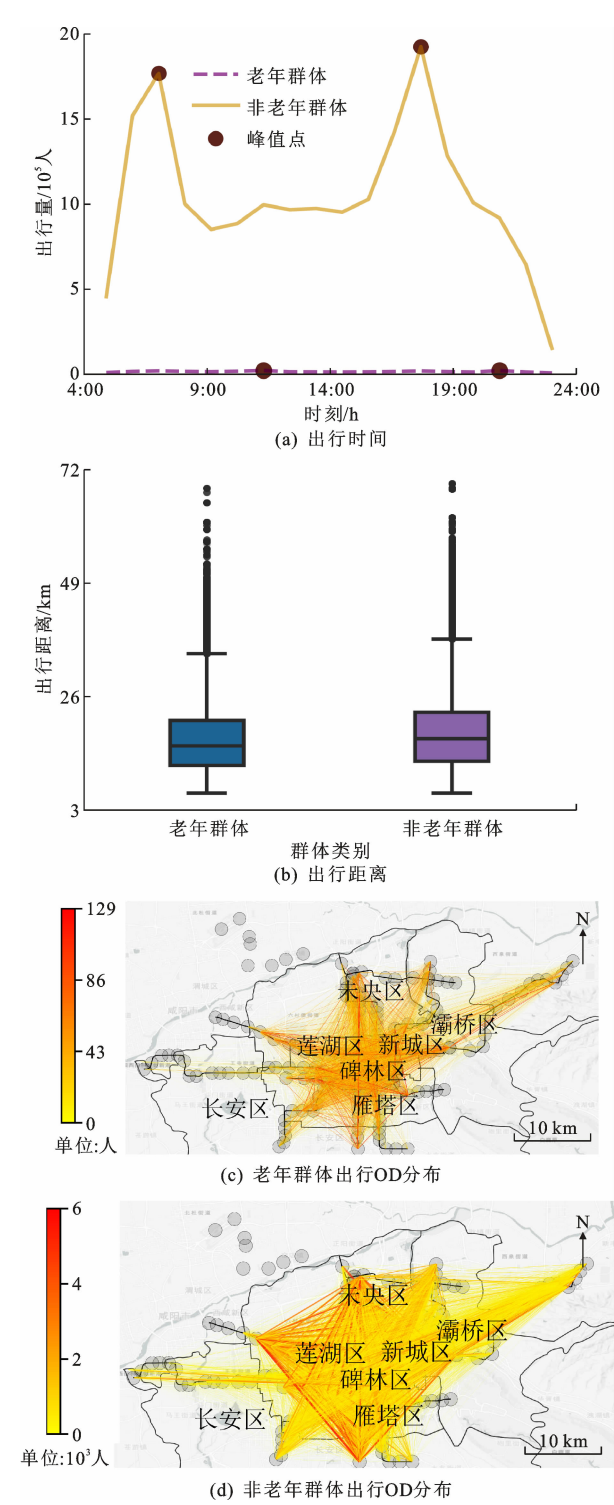


图 3 不同群体出行时空特征

Fig. 3 Spatio-temporal travel characteristics of different groups
经济特征、交通相关特征、土地利用特征以及地铁系统特征对老年群体出行量的平均贡献为 2.94%、19.64%、6.45%、0.03%,对非老年群体出行量的平均贡献为 5.94%、12.24%、9.09%、0.01%。很明显,土地利用特征对非老年群体的影响大于老年群体,这主要是因为老年群体可能更倾向于在家中进

表 3 LightGBM 模型的调优参数

Tab. 3 Tuning parameters of LightGBM model

超参数	范围	参数值	
		老年群体	非老年群体
lambda_l1	{0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0}	0.0	0.9
lambda_l2	{0.5,0.6,0.7,0.8,0.9,1.0}	0.9	0.5
min_data_in_leaf	{10,20,30,40,50,60}	20	30
num_leaves	{10,20,30,40,50,60,70,80,90,100}	20	20
feature_fraction	{0.6,0.7,0.8,0.9}	0.7	0.8

表 4 五折交叉验证结果

Tab. 4 Results of five-fold cross-validation

验证次数	老年群体			非老年群体		
	R_{MSE}	M_{AE}	R^2	R_{MSE}	M_{AE}	R^2
0	93.05	85.54	0.64	61 001.18	48 786.65	0.59
1	69.54	65.53	0.68	62 563.82	50 431.79	0.49
2	89.37	98.79	0.60	60 369.51	47 437.80	0.60
3	83.50	71.32	0.55	86 765.52	62 418.56	0.53
4	91.10	87.91	0.51	88 495.36	61 047.40	0.50

表 5 不同学习模型的比较

Tab. 5 Comparison of different learning models

学习模型	老年群体			非老年群体		
	R_{MSE}	M_{AE}	R^2	R_{MSE}	M_{AE}	R^2
LightGBM	107.60	81.48	0.65	86 219.99	59 239.55	0.55
GBDT	114.25	92.20	0.61	99 324.58	65 944.05	0.38
LR	118.01	92.22	0.58	94 209.87	58 960.02	0.44

行休闲活动,而非外出娱乐或购物。此外,交通相关特征对 2 类群体的平均贡献均超过 10%,说明地铁站周边的交通相关变量对地铁客流量有显著

表 6 轨道客流的建成环境变量重要度排序

Tab. 6 Ranking of importance of built environment variables for subway ridership

建成环境	解释变量	老年群体			非老年群体		
		重要度/%	排序	均值/%	重要度/%	排序	均值/%
社会经济	工作人口数量	4.76	8	2.94	4.88	7	5.94
	居住人口数量	1.12	10		6.99	6	
交通相关	路网密度	29.16	1	19.64	20.28	2	12.24
	公交站点密度	10.12	5		4.20	8	
土地利用	科教文化数量	16.08	2	6.45	7.69	5	9.09
	医疗中心数量	14.29	3		9.09	4	
	公司企业数量	11.32	4		30.05	1	
	餐饮美食数量	5.95	6		9.79	3	
	休闲娱乐数量	5.36	7		0.02	11	
	生活服务数量	1.20	9		4.19	9	
	用地混合度	0.58	11		2.80	10	
地铁系统	是否为换乘站	0.06	12	0.03	0.02	12	0.01
	是否为终点站	0.00	13		0.00	13	

影响。

路网密度对老年群体和非老年群体的重要度分别达到了 29.16%、20.28%,这说明路网密度和连通性直接影响了乘客到达地铁站点的便利程度,进一步决定了地铁客流量。其次,居住人口和工作人口对老年群体的影响小于非老年群体,这主要是因为非老年群体可能更频繁的在居住地和工作地之间通勤,而老年群体大多已退休,地铁站周边的居住人口和工作人口数会对非老年群体的出行需求有更显著影响。

在剩余因素中,公司企业数量和医疗中心数量的重要度高于用地混合度和休闲娱乐数量,这说明相比于休闲娱乐,人们可能更看重地铁带来的工作和医疗服务的便利性。此外,地铁系统因素(是否为换乘站、终点站)对地铁客流量的重要可以忽略不计。这与 Shao 等^[24]在深圳的研究结果不同,后者表明地铁网络因素对地铁客流量有显著影响,这种差异可能归因于西安和深圳 2 个城市不同的结构特征。

3.2.3 非线性影响

提取影响老年群体出行量最重要的 5 种因素(路网密度、科教文化数量、医疗中心数量、公司企业数量以及公交站点密度)部分依赖图,如图 4~图 8 所示。由图 4~图 8 可知,这 5 个建成环境变量对 2 类群体的地铁出行量均有明显的阈值效用。

图 4 为路网密度与 2 类群体地铁客流量关系对比。由图 4 可知:路网密度与地铁客流量,均呈凹形(非线性关系),这与现有研究结果一致^[25]。具体来

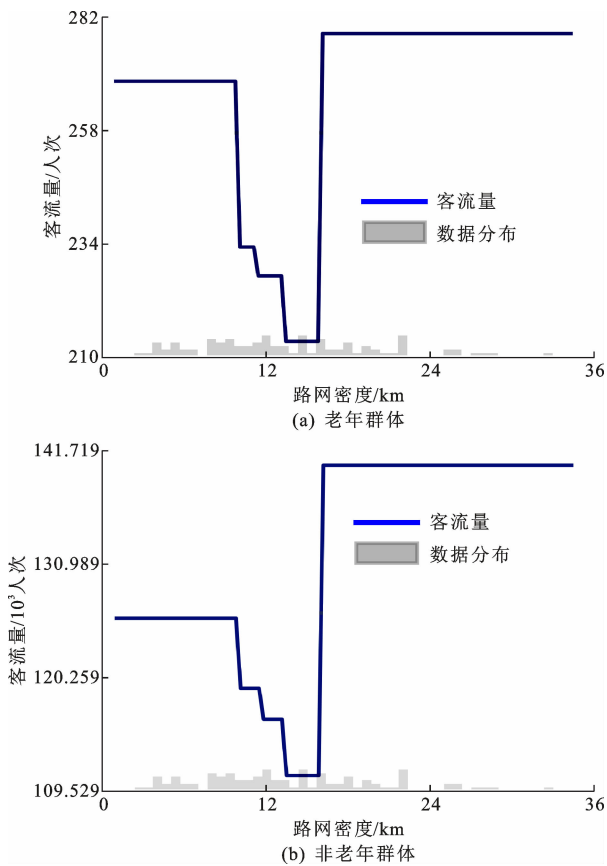


图 4 路网密度与客流量关系

说,路网密度在10 km以下对客流量基本没有影响;而当路网密度从10 km增加15 km左右时,老年群体和非老年群体的客流量分别下降了62、14 000人次;15~17 km基本保持不变,之后2类群体的客流量分别在17.5 km处迅速提升至280、14 000人次。这表明站点影响范围内17.5 km的路网密度对地铁出行量具有显著提升作用,主要原因可能是较高的路网密度直接影响了步行便利性和地铁接驳便利性,这有助于增加公众使用轨道交通的频次。

图5为轨道站点缓冲区内医疗中心数量与2类群体地铁出行量关系对比。由图5可知:对于老年群体,当医疗中心数量从0增加到55左右时,对出行量的影响较小;而超过55之后,客流量从224人上升至325人;之后医疗中心数增加客流量都保持不变。对于非老年群体,医疗中心数量从55增加到58时,客流量增加了35 000人,之后客流量保持不变。研究结果表明,增加医疗中心数量对2类群体出行的促进作用比较类似,且医疗中心数量在55~58时,促进作用最明显。

图6为科教文化数量与2类群体地铁出行量关系对比。由图6可知:科教文化数量对轨道客流量

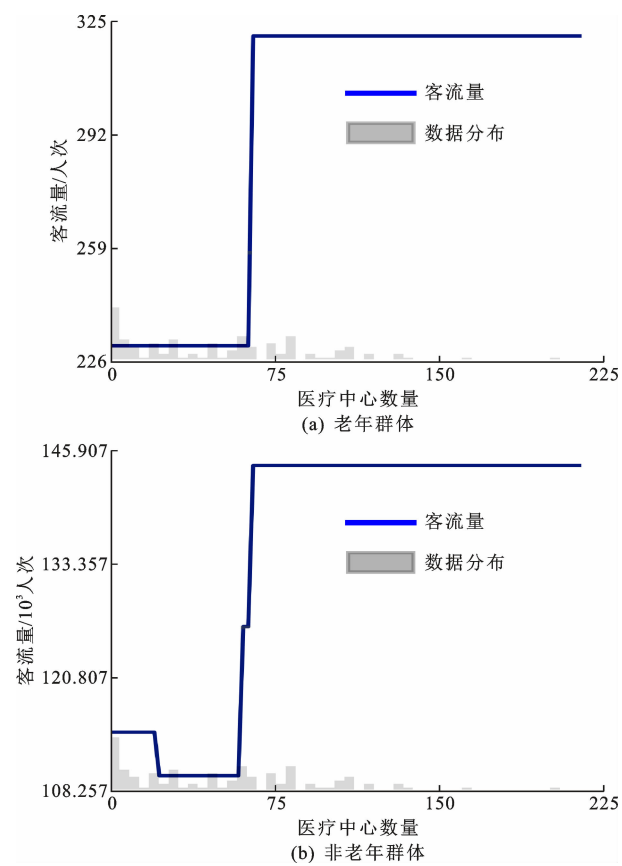


图 5 医疗中心数量与客流量关系

具有明显的阈值作用,且2类群体的变化趋势很相似。当科教文化数量在25~48时,对2类群体地铁出行量有显著促进作用,分别可提升92、24 000人次的客流,达到峰值后,客流量保持不变。科教文化数量较多的站区,以学生与教师群体的通学、通勤为主,老年人可能会接送小孩,因此科教文化数量对客流量的影响有显著正向作用。

图7为公交站点密度与2类群体地铁出行量关系对比。由图7可知:公交站点对2类群体均有显著的正向影响,这与Ding等^[14]的研究结果相似。2个群体均呈现保持不变(0~8)、急剧上升(9~12)、保持不变(13~60)的形态,当轨道站点影响范围内的公交站点数达12个时,老年群体的出行量增加36人,非老年群体增加10 800人。公交车站会对地铁客流产生积极的影响,主要原因可能是公交站点会提升城市轨道交通接驳的便利性,有利于站点区域以及站区以外的人群使用地铁。

图8为公司企业数量与2类群体地铁出行量关系对比。由图8可知:缓冲区内公司企业数量对2类群体客流量的影响整体呈非线性正相关作用,进一步证实了通勤对地铁客流量有显著影响^[26]。具

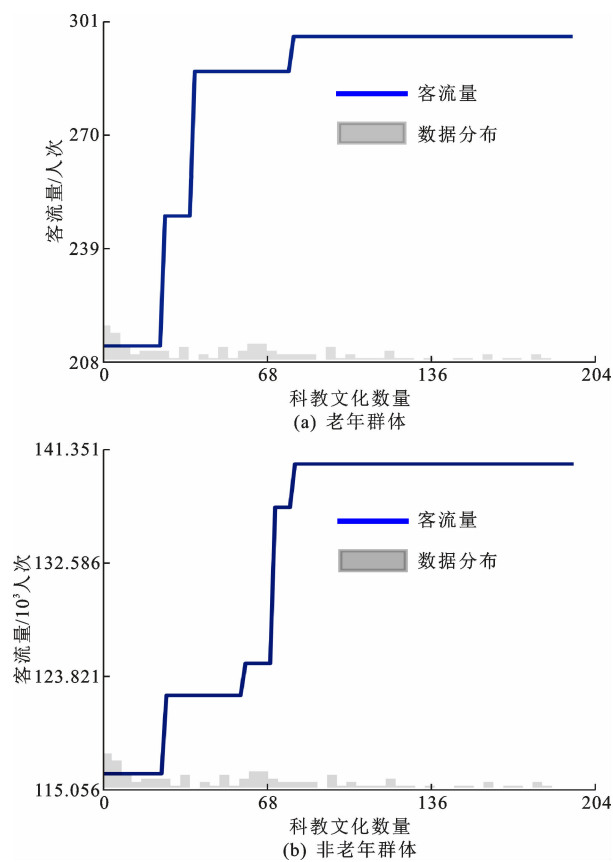


图 6 科教文化数量与客流量关系

Fig. 6 Relationship between number of cultural and educational facilities and ridership

体而言,对于老年群体,公司企业数量超过 100 时,对客流的促进作用最显著,对于非老年群体,公司企业数量大于 60 时,有显著的促进作用。

4 结 语

(1)LightGBM 模型在老年群体和非老年群体的数据拟合与预测性能上均表现出色,相较于 GB-DT 和 LR 模型具有显著优势。

(2)非老年群体在轨道交通客流的时序变化规律呈现双峰形,即呈现明显的早晚高峰潮汐特征,而老年群体则呈现无明显峰值的特征。此外,2 类群体在地铁出行距离方面主要集中在 7~20 km,活动范围集中在西安市三环内。值得注意的是,在三环外的站点如华清池、秦陵西、韦曲南站,吸引了大量非老年群体的出行,这与这些站点靠近兵马俑、华清池等热门景点有很大关联。

(3)就客流量影响因素的重要度排序而言,对于 2 类群体,其影响程度排列为交通相关特征、土地利用特征、社会经济特征和地铁系统特征。在所有建成环境变量中,路网密度对 2 类群体的客流量影响

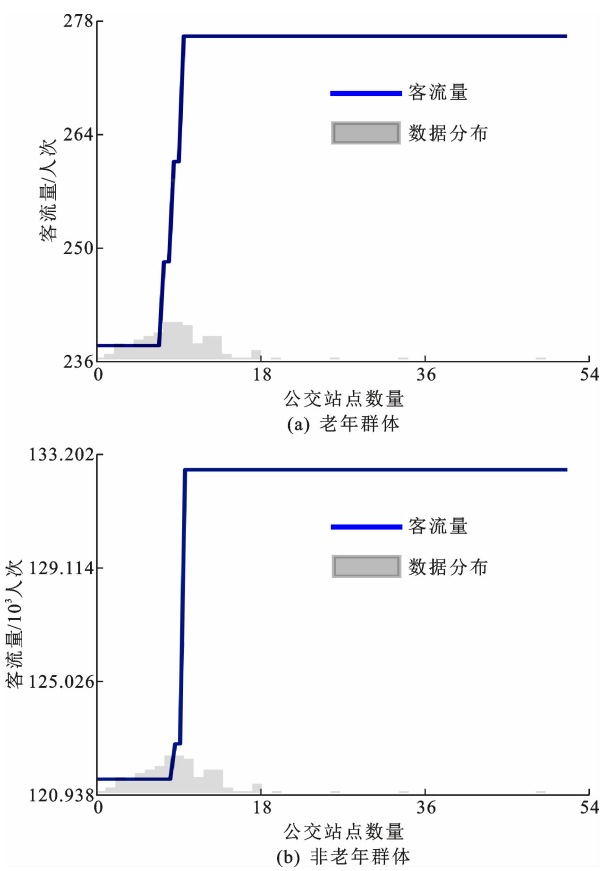


图 7 公交站点密度与客流量关系

Fig. 7 Relationship between density of bus stations and ridership 显著,贡献率均超过 10%。这表明,高可达性的站区建设对提高地铁出行需求有积极影响。

(4)在影响客流量的因素中,医疗中心数量、科教文化数量、公司企业数量和公交站点密度对 2 类群体的客流量呈非线性正相关作用,而路网密度呈现凹形(非线性作用)。其中,在轨道站点影响范围内,医疗中心数量达到 55~58 时,对 2 类群体的地铁出行量都产生了明显促进作用;科教文化数量的影响阈值为 48,之后其数量的增加不再对客流量产生影响。

(5)本文仅使用了 1 周的手机信令数据,因此在后续的研究中,需要集成更长时间范围的数据,以进一步验证不同群体出行在不同时间段的差异。

参考文献:

References:

[1] SZETO W Y, YANG L C, WONG R C P, et al. Spatio-temporal travel characteristics of the elderly in an ageing society [J]. Travel Behaviour and Society, 2017, 9: 10-20.
[2] DING C, CHEN P, JIAO J F. Non-linear effects of the

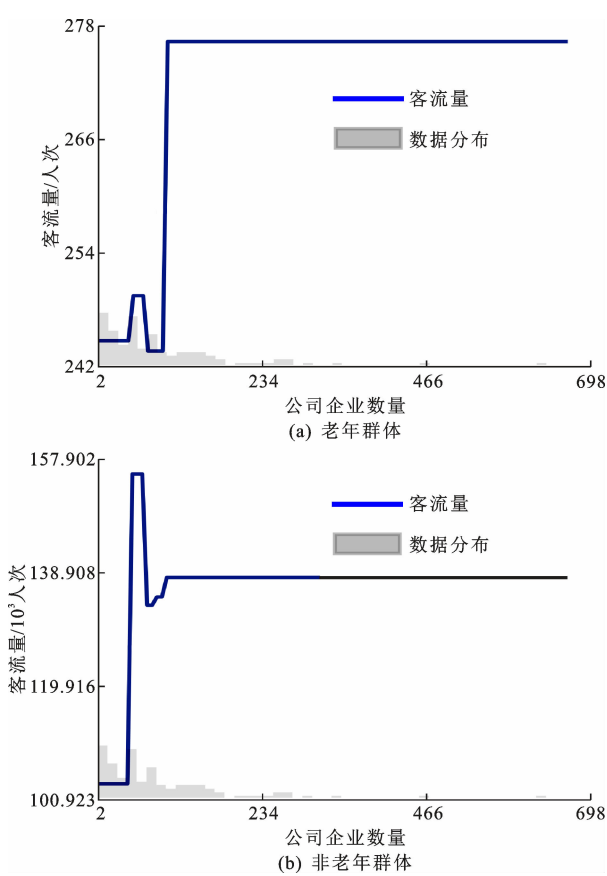


图 8 公司企业数量与客流量关系

Fig. 8 Relationship between number of enterprises and ridership

built environment on automobile-involved pedestrian crash frequency: A machine learning approach[J]. Accident Analysis & Prevention, 2018, 112: 116-126.

[3] 李康康,杨东峰. 城市建成环境如何影响老年人体力活动:模型构建与大连实证[J]. 人文地理, 2021, 36 (5): 111-120.

LI Kang-kang, YANG Dong-feng. How does the urban built environment affect the physical activities of the elderly: Model construction and Dalian demonstration[J]. Human Geography, 2021, 36(5): 111-120.

[4] SUN G B, ZHAO J T, WEBSTER C, et al. New metro system and active travel: A natural experiment[J]. Environment International, 2020, 138: 105605.

[5] 何 伟,张 彤,黄 靖. 基于地铁的乘客出行群组模式发现与可视化[J]. 测绘科学, 2021, 46 (8): 156-164.

HE Wei, ZHANG Tong, HUANG Jing. Discovery and visualization of passenger travel group patterns based on subway systems[J]. Science of Surveying and Mapping, 2021, 46(8): 156-164.

[6] 许志榕. 上海市职住关系和通勤特征分析研究:基于轨道交通客流数据视角[J]. 上海城市规划, 2016(2): 114-121.

XU Zhi-rong. Study on job-housing relationship and characteristic of commuting in Shanghai: Based on the perspective of rail transit passenger flow data [J]. Shanghai Urban Planning Review, 2016(2): 114-121.

[7] 朱佳欣. 城市轨道线路调整对居民通勤出行方式选择的影响研究[D]. 北京: 北京交通大学, 2021.

ZHU Jia-xin. Study on the influence of urban rail line adjustment on residents' choice of commuting modes [D]. Beijing: Beijing Jiaotong University, 2021.

[8] 兰招娣,田丽君. 活动视角下基于家庭出行时通勤者的出行方式选择研究[J]. 技术与市场, 2022, 29(7): 15-20.

LAN Zhao-di, TIAN Li-jun. Study on commuters' travel mode choice based on family travel from the perspective of activity [J]. Technology and Market, 2022, 29(7): 15-20.

[9] 杨 红,申犁帆,胡议文,等. 老年人地铁出行时空特征及与建成环境非线性关系:以武汉市为例[J]. 地理科学进展, 2023, 42(3): 491-504.

YANG Hong, SHEN Li-fan, HU Yi-wen, et al. Spatial and temporal characteristics of elderly people's metro travel behavior and its non-linear relationship with the built environment: A case study of Wuhan City [J]. Progress in Geography, 2023, 42(3): 491-504.

[10] 熊美成,黄 洁,王姣娥,等. 活动空间视角下的老年群体地铁出行时空特征:以昆明市为例[J]. 热带地理, 2022, 42(12): 2052-2062.

XIONG Mei-cheng, HUANG Jie, WANG Jiao-e, et al. Spatiotemporal characteristics of subway travel by the elderly based on activity space: A case study of Kunming [J]. Tropical Geography, 2022, 42 (12): 2052-2062.

[11] CHEN C, CHEN J, BARRY J. Diurnal pattern of transit ridership: A case study of the New York City subway system[J]. Journal of Transport Geography, 2009, 17(3): 176-186.

[12] JUN M J, CHOI K, JEONG J E, et al. Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul [J]. Journal of Transport Geography, 2015, 48: 30-40.

[13] 李 想,晏启鹏,骆 晨. 空间异质性建成环境对地铁与公交换乘客流的影响[J]. 交通运输系统工程与信息, 2023, 23(2): 100-110.

LI Xiang, YAN Qi-peng, LUO Chen. Impact of built environment on flow of transfer passengers between subway and bus considering spatial heterogeneity[J]. Journal of Transportation Systems Engineering and Information Technology, 2023, 23(2): 100-110.

- [14] DING C, CAO X Y, LIU C. How does the station-area built environment influence metrorail ridership? Using gradient boosting decision trees to identify non-linear thresholds [J]. *Journal of Transport Geography*, 2019, 77: 70-78.
- [15] DING C, LIU T T, YAO B Z, et al. Understanding the time-dependent effect of built environment attributes on station-level metro ridership uncertainty in Beijing: A big data analytic approach [J]. *Tunnelling and Underground Space Technology*, 2023, 137: 105148.
- [16] YU Z D, ZHU X L, LIU X T. Characterizing metro stations via urban function: Thematic evidence from transit-oriented development (TOD) in Hong Kong [J]. *Journal of Transport Geography*, 2022, 99: 103299.
- [17] PANG L, JIANG Y X, WANG J J, et al. Research of metro stations with varying patterns of ridership and their relationship with built environment, on the example of Tianjin, China [J]. *Sustainability*, 2023, 15 (12): 9533.
- [18] LIU B, XU Y F, GUO S Z, et al. Examining the non-linear impacts of origin-destination built environment on metro ridership at station-to-station level [J]. *ISPRS International Journal of Geo-Information*, 2023, 12(2): 59.
- [19] CHAMPAGNE M P, DUBÉ J, BARLA P. Build it and they will come: How does a new public transit station influence building construction? [J]. *Journal of Transport Geography*, 2022, 100: 103320.
- [20] WU W J, HONG J. Does public transit improvement affect commuting behavior in Beijing, China? A spatial multilevel approach [J]. *Transportation Research Part D*, 2017, 52: 471-479.
- [21] KIM T, CHOO S, SHIN Y, et al. Identifying differences of travel time budgets between the elderly and the nonelderly groups using PSL structural equation models: A case study for Seoul metropolitan area [J]. *International Journal of Sustainable Transportation*, 2016, 10(5): 455-465.
- [22] CHENG L, DE VOS J, SHI K B, et al. Do residential location effects on travel behavior differ between the elderly and younger adults? [J]. *Transportation Research Part D*, 2019, 73: 367-380.
- [23] 许心越, 孔庆雪, 李建民, 等. 建成环境对轨道交通客流的时空异质性影响分析 [J]. *交通运输系统工程与信息*, 2023, 23(4): 194-202, 281.
- XU Xin-yue, KONG Qing-xue, LI Jian-min, et al. Analysis of spatio-temporal heterogeneity impact of built environment on rail transit passenger flow [J]. *Journal of Transportation Systems Engineering and Information Technology*, 2023, 23(4): 194-202, 281.
- [24] SHAO Q F, ZHANG W J, CAO X Y, et al. Threshold and moderating effects of land use on metro ridership in Shenzhen: Implications for TOD planning [J]. *Journal of Transport Geography*, 2020, 89: 102878.
- [25] 汪雨菲, 杨皓森, 喻冰洁, 等. 站域建成环境与地铁客流量的非线性关系和协同效应——可解释机器学习分析 [J]. *都市快轨交通*, 2024, 37(2): 1-7.
- WANG Yu-fei, YANG Hao-sen, YU Bing-jie, et al. Nonlinear and synergistic effects of station-area built environments on metro ridership: A shapley additive explanations (SHAP) analysis [J]. *Urban Rapid Rail Transit*, 2024, 37(2): 1-7.
- [26] LI L C, ZHONG L J, RAN B, et al. Analysis of the relationship between metro ridership and built environment: A machine learning method considering combinational features [J]. *Tunnelling and Underground Space Technology Incorporating Trenchless Technology Research*, 2024, 144: 105564.