

# 城市轨道交通站点周边建成环境对地铁客流量的影响研究——以深圳市为例

## Impact of Built Environment of Urban Rail Transit Stations on Metro Passenger Flow: A Case Study of Shenzhen

黄建欣 龚蔚霞 张金林 汤咏诗 吴巍 HUANG Jianxin, GONG Weixia, ZHANG Jinlin, TANG Yongshi, WU Wei

**摘要** 传统线性模型无法显示指标的非线性结果,难以体现各指标的相对重要性和影响阈值。利用深圳轨道交通刷卡数据,构建梯度提升决策树(GBDT)模型,探索建成环境对不同时间段客流量的非线性关系,以期为城市尤其是超大特大城市的交通规划、站点开发、站点周边建成环境优化提供参考。结果表明:不同建成环境对不同时间段客流量的相对重要性不一致,建筑容积率、居住用地比例、商业商务用地比例、距CBD距离对站点客流量均有相对较高且稳健的影响;建成环境与客流量存在明显的非线性关系和阈值;当土地利用混合度在0.60以上和距CBD距离超过17 km时,全天客流量呈现明显的下降趋势;建成环境对不同时间段客流量的影响机制不完全相同;城中村对地铁客流量的促进效果十分显著。

**Abstract** Traditional linear models cannot display the nonlinear results of indicators, making it difficult to reflect the relative importance and impact threshold of each indicator. This study uses Shenzhen rail transit smart card data to construct a gradient boosting decision tree (GBDT) model to explore the non-linear relationship between the built environment and passenger flow in different time periods, hoping to provide references for transportation planning, station development, and optimization of built environment around stations in cities, especially mega cities. The results show that: ① Different built environments have different relative importance to passenger flow in different time periods. Building floor area ratio, proportion of residential land, proportion of commercial land, and distance from CBD all have relatively high and robust effects on site passenger flow. ② There is a significant nonlinear relationship and threshold between the built environment and passenger flow. ③ When the land use mixing degree is above 0.60 and the distance from CBD exceeds 17 km, the daily passenger flow shows a significant downward trend. ④ The impact mechanism of the built environment on the passenger flow in different time periods is not completely the same. ⑤ The promotion effect of urban villages on subway passenger flow is very significant.

**关键词** 城市轨道交通;建成环境;客流量;梯度提升决策树;非线性影响

**Key words** urban rail transit; built environment; passenger flow; gradient boosting decision tree; non-linear influencing

文章编号 1673-8985 (2023) 06-0047-08 中图分类号 TU984 文献标志码 A

DOI 10.11982/j.supr.20230606

### 作者简介

黄建欣

广东省科学院广州地理研究所  
规划师,硕士

龚蔚霞

广东省科学院广州地理研究所主任  
正高级工程师,博士,硕士生导师

张金林

广东省科学院广州地理研究所  
助理工程师

汤咏诗

广东省科学院广州地理研究所  
助理工程师

吴巍 (通信作者)

河北工程大学建筑与艺术学院  
讲师,博士,urban8611@163.com

## 0 引言

在全球气候持续变暖、城市化不断推进和机动车快速发展的背景下,环境和交通矛盾日益凸显,选择公共交通出行逐渐成为控制交通碳排放、缓解交通拥堵和实现低碳出行的关键因素之一<sup>[1]</sup>。城市轨道交通是公共交通网络的骨架,其客流量已广泛应用于分析绿色出行成效、站点选址合理性以及测度站点与周边城市空间的匹配度等方面,但现有研究主要观察全地铁客流量(以下简称“客流量”)的特征,忽略了不同时间

段客流量的变化机制。

建成环境通过影响城市功能分区、人口密度、居民出行行为等方面进而影响轨道交通站点客流量。Cervero等<sup>[2]</sup>提出密度 (density)、多样性 (diversity) 和设计 (design) 3个研究维度, Ewing等<sup>[3]</sup>在此基础上补充公共交通可达性 (distance to transit) 和到目的地距离 (destination accessibility), “5D” 指标体系已广泛应用于建成环境指标测度研究<sup>[4]</sup>。密度是反映居民活动强度的重要指标, 站点合理步行范围内开发强度和人口密度越高, 使用公共交通出行的需求越大<sup>[5]</sup>; 多样性反映站点周边区域不同类型用地均匀度情况, 土地利用通过影响城市空间布局、就业岗位分布和站点周边用地类型等因素从而影响地铁周边居民出行方式的选择, 土地利用混合度是影响客流量的根本原因之一<sup>[6]</sup>。道路交通设计和公共交通可达性反映居民到达轨道交通站点的便捷度、舒适度和安全性, 对居民地铁出行具有显著的积极作用; 路网密度与道路交叉口密度反映街道连通性的高低, 数值越高表示可达性水平越高, 意味着居民可以更便利地获取公共交通服务。目的地距离会影响居民选择出行方式的意愿, 有研究指出站点到城市中心区的距离对客流量有一定的消极作用<sup>[7]</sup>。居住要素方面, 主要以居住建筑面积和居住用地面积为量化指标, 少有直接将城中村建筑面积纳入为模型变量。2018年深圳城中村私宅总建筑面积仅占全市住房面积的34%, 却提供45%以上的城市住房套数<sup>[8]</sup>, 深圳城中村高密度的人口为城市轨道交通提供主要客源。另外, 城中村聚集了城市大量的中低收入群体, 该群体对城市轨道交通出行的需求和依赖性较高, 因此城中村对探究客流量影响机制具有较大的意义。

在有关建成环境与出行方式的已有研究中, 多采用多元回归模型<sup>[9]</sup>、地理加权回归模型 (GWR)<sup>[10]</sup>、聚类分析模型<sup>[11]</sup>等数理模型来分析不同指标与客流量的线性关系。线性回归模型预先假设两者存在线性关系, 可能掩盖建成环境存在的非线性关系和阈值影响结果<sup>[12]</sup>。而通过机器学习算法运行梯度提升决策树 (Gradient Boosting Decision Tree,

GBDT) 模型, 研究建成环境与客流量间的非线性关系, 有利于掌握不同自变量的影响系数、重要性排序和阈值效应。崔叙等<sup>[13]</sup>分析建成环境、经济属性、站域设施数量3大因素与客流量、出行时间的非线性关系, 明确指出容积率对客流量的高敏感值 (容积率为1.5时) 和高稳定值 (容积率 $\geq 3.2$ ); 周扬等<sup>[14]</sup>研究站点周边区域建成环境对居民活动影响的非线性关系, 提出建成环境的适宜建设的指标控制和协同优化建议; Ding等<sup>[15]</sup>认为当土地利用混合度在0.5—0.6区间时, 对客流量有明显的促进作用。可见, 构建非线性模型分析各建成环境要素对客流量的影响机制, 能更具体判断各指标的重要性和阈值效应, 提出更有效的站点空间指标控制与优化建议。

基于以上论述, 本文使用深圳市2017年10月的城市轨道交通刷卡数据, 提取不同时间段客流量, 获取建筑、用地、交通、POI (Point of Interest) 等多源数据, 基于梯度提升决策树 (GBDT) 机器学习算法, 从开发强度、土地利用、交通道路设计、公交可达性和距CBD距离等5个方面, 研究站点周边区域内各建成环境要素对全天、早高峰、非高峰、晚高峰和夜间5个时间段客流量的非线性关系, 以期优化

轨道交通站点周边已有建成环境, 缓解城市交通拥堵矛盾, 促进绿色低碳出行, 并为轨道交通站点规划选址与建设提供案例支撑。

## 1 数据与方法

### 1.1 研究区域

本文以深圳市轨道交通站点周边建成环境为研究对象 (见图1)。截至2017年10月, 深圳市共开通运营8条地铁线路共166个站点 (换乘站不重复计算), 其中福邻站和机场北站尚未正式对外运营, 故最终将165个站点纳为研究对象。地铁线路主要经过城市主中心、副中心和主要的生活区与商业区, 站点周边区域建成环境相对成熟完善, 对加强研究成果准确性具有实际意义。深圳城市中心主要位于福田区、罗湖区和南山区3个行政区, 后续测度“距CBD距离”指标将取地铁站点到这3个行政区的平均直线距离。参考已有研究, 本文将城市轨道交通站点周边半径800 m的区域定义为研究范围<sup>[16]</sup>。

### 1.2 数据来源与处理

#### (1) 城市轨道交通站点客流量

本文选取2017年10月16日至10月20日连续1周工作日深圳市165个站点的地铁刷卡数

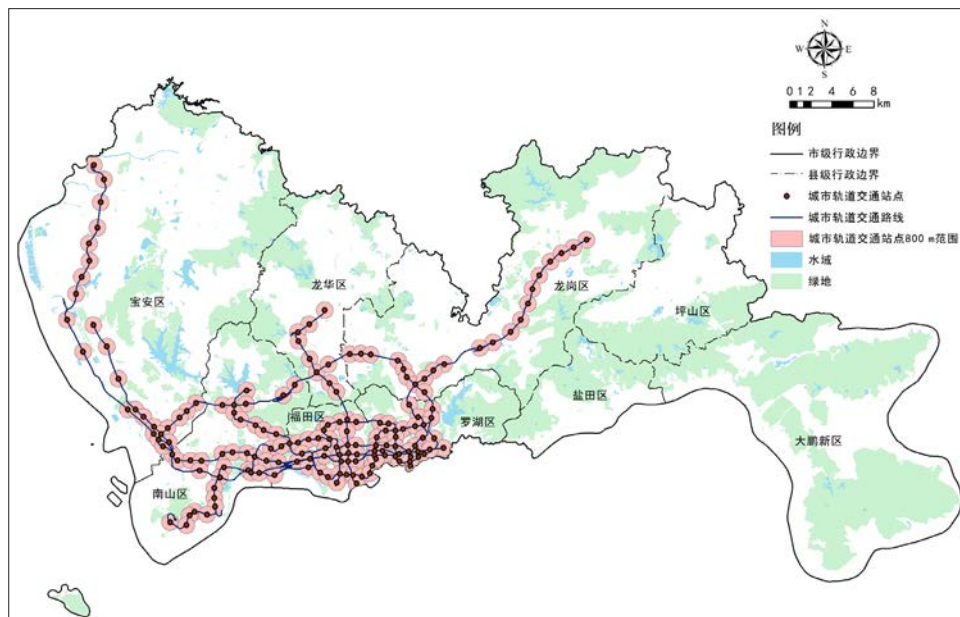


图1 深圳市轨道交通站点及线路分布  
Fig.1 Distribution of Shenzhen rail transit stations and lines

资料来源: 笔者自绘。

据,剔除部分无效数据,根据用卡信息提取出行的OD (Origin-Destination) 数据。取5 d全天客流量平均值得到日均客流量,获得日均OD数据约184万条;以1 h为间隔计算不同时间段的5 d平均客流量,结合客流量大小将全天划分为早高峰时段 (7:00—9:00)、非高峰时段 (11:00—13:00)、晚高峰时段 (17:00—19:00) 和夜间时段 (21:00—23:00)。

(2) 轨道交通站点周边建成环境指标

本文按照建成环境的“5D”指标体系,将自变量分为密度、多样性、道路交通设计、公共交通可达性、到目的地距离5个层面(见表1)。密度主要研究开发强度和城中村建设对客流量的影响;多样性探究土地利用情况的影响机制,其中公共服务用地包含综合服务区和绿地休闲区,主要分析除居住、就业和交通出行外的公共服务功能对客流量的影响;道路交通设计分析步行系统和出行便捷度对客流量的影响,合理高效的交叉口设计保证了居民步行至轨道交通站点的便捷性;到目的地距离主要是测量站点到南山区、福田区和罗湖区3个城市中心区的直线平均距离。

1.3 研究方法

针对城市轨道交通站点800 m范围内建成环境要素对不同时间段客流量的影响机制,本文运用梯度提升决策树模型分析建成环境与客流量的相关性和非线性关系。GBDT模型是由Friedman等<sup>[18]</sup>于2001年提出的一种经典机器学习迭代算法,最初主要用于计算机领域,近年来逐步应用于交通工程<sup>[19]</sup>、城市规划<sup>[20]</sup>、经济地理<sup>[13]61</sup>等研究领域。

相比于传统的线性回归模型,GBDT模型在研究自变量对因变量影响机制上具有较多的优势:(1) GBDT模型不假设自变量和因变量之间具有线性关系,能有效解决自变量的多重共线性问题;(2) 模型数据拟合度较好,能预测每个自变量与因变量的非线性关系并计算出相对重要度,生成偏依赖图 (Partial Dependence Plot, PDP),为分析单个自变量对因变量的影响机制提供支撑;(3) 模型能在训练过程中通过加强机器学习不断调整预

测值与权重赋值,且能灵活处理异常值和缺失值,从而提供预测结果的准确性。

GBDT模型以当前预测值为基准,在不断训练过程中使用决策树模型产生不同层级的弱分类器,新分类器在上一轮分类器的残差基础上再次拟合损失函数并计算新预测值的新残差,迭代形成多棵决策树,最终达到将数据分类或回归的效果,其关键是利用损失函数的负梯度作为提升决策树算法中残差的近似值。具体算法步骤如下。

初始化弱学习器,计算最小化损失函数的常数项:

$$F_0(x)=\operatorname{argmin}_{\delta} \sum_{i=1}^N L(y_i, \delta) \quad (1)$$

式中: $x$ 表示自变量, $y$ 表示不同时间段客流量, $i$ 表示地铁站点样本 ( $i=1,2,\dots,165$ ), $\delta$ 表示每棵回归树中分裂变量、分裂位置和叶子节点的均值, $L$ 表示损失函数。

对于迭代次数 $n$  ( $n=1,2,\dots,N$ ) 有:

$$\delta_{ni} = - \left[ \frac{\mu L[y_i, f(x_i)]}{\mu f(x_i)} \right]_{f(x)=f_{n-1}(x)} \quad (2)$$

式中: $\delta_{ni}$ 表示第 $n$ 轮的第 $i$ 个样本的损失函数的负梯度(即残差), $\mu$ 表示计算参数。

将公式(2)的损失函数得到的残差值 $\delta_{ni}$ 作为样本新计算值,并将数据 $(x_i, \delta_{ni})$ 作为下一棵决策树的训练数据,再次得到一棵新的决策树 $f_n(x)$ 。其对应的叶子节点区域为 $V_{nj}$  ( $j=1,2,\dots,J$ ), $J$ 为决策树的叶子节点数量。

叶子区域 $j$ 的最佳拟合值的计算公式:

$$\delta_{nj} = \operatorname{argmin}_{\delta} \sum_{x_i \in V_{nj}} L(y_i, f_{n-1}(x_i) + \delta) \quad (3)$$

式中: $\delta_{nj}$ 表示对于叶子区域 $J$ 计算得出的最佳拟合值。

更新强学习器:

$$f_n(x) = f_{n-1}(x) + \sum_{j=1}^J \delta_{nj} I(x \in V_{nj}) \quad (4)$$

运行结束后得到最终学习器:

$$f(x) = f_N(x) \quad (5)$$

为避免模型出现过度拟合问题和制约每棵决策树的负梯度学习结果,在模型的每个弱分类器前设置“学习速率” $v$  ( $0 < v \leq 1$ ), $v$ 值越高则学习越快,越容易出现过度拟合现象,得出最终的计算公式:

$$f_n(x) = f_{n-1}(x) + v \sum_{j=1}^J \delta_{nj} I(x \in V_{nj}), 0 < v \leq 1 \quad (6)$$

表1 城市轨道交通站点周边半径800 m研究范围内建成环境指标  
Tab.1 Built environmental indicators within the research range of 800m radius around urban rail transit stations

指标层面	变量名称	指标计算与解析	指标获取方式
密度	建筑容积率	总建筑面积与研究范围面积的比值	2015年深圳市建筑普查数据
	城中村私宅建筑面积占比/%	城中村私宅建筑面积与研究范围面积的比值	
多样性	土地利用混合度	土地利用混合度,采用POI数据熵指数测算,具体计算公式为: $E_{mix} = \frac{-\sum_{i=1}^n P_i \ln(P_i)}{\ln(N)}$ 式中: $P_i$ 为研究范围内第 $i$ 类POI占总POI的比重, $N$ 为POI的类别数。 $E_{mix}$ 数值越大表示土地利用混合度越高,反之越低	高德地图平台获取不同类型POI数据,为2016年数据
	居住用地比例/%	居住用地占研究范围面积的比值	2018年深圳市地块尺度的城市土地利用数据 <sup>[17]</sup>
	商业商务用地比例/%	提供商业、商务办公等就业岗位为主要功能导向区域的用地面积/占研究范围面积的比值	
道路交通设计	路网密度/(km/km <sup>2</sup> )	道路总长度与研究范围面积的比值	Open Street Map 开放地图数据库,为2017年数据
	道路交叉口数量/个	道路交叉口个数	
公共交通可达性	公交线路数量/条	常规公交线路数目	
到目的地距离	距CBD距离/km	每个轨道交通站点到福田、罗湖、南山中心区的直线平均距离	

资料来源:笔者自制。

每棵回归树都乘以“学习速率” $v$ ,目的是实现损失函数的最小化。

## 2 结果分析

### 2.1 站点周边地区建成环境指标的相对重要性

表2显示了建成环境各变量在预测不同时间段客流量相关重要性。总的来说,建筑容积率、居住用地比例、商业商务用地比例、距CBD距离在不同时间段的相对重要性均较高,除了建筑容积率在夜间时段的重要性为8.61%外,这4个变量在5个时间段的重要性均大于10%。可见,无论在哪个时间段,这4个变量对站点客流量都有相对较高且稳健的影响,紧凑的高密度开发、大量的居住人口和就业岗位引发的必要性通勤流是客流量的关键。

(1) 密度方面。建筑容积率的重要性十分显著,站点周边的开发强度是决定站点客流量的最重要因素。值得注意的是,建筑容积率对非高峰时段和夜间时段的重要性有所下滑。这两个时段以非通勤客流量为主,出行目的较多元,客流量变化对居住密度和就业岗位的敏感度相较早晚高峰时段弱。城中村私宅建筑面积占比的相对重要性在8.12%—10.16%之间,居住在城中村的高密度人口为地铁带来大量的客流源。该群体主要依赖公共交通出行,对提高客流量有十分积极的作用。

(2) 多样性方面。土地利用混合度对全天客流量、早高峰时段和晚高峰时段的影响较低,均位于第10位;另外,土地利用混合度对非高峰时段(第6位)和夜间时段(第8位)的重要性有所上升。公共服务用地比例的相对重要性在8.06%—12.03%之间,对不同客流量的相对重要性差异较大;其对夜间客流量的影响性有较为明显的增加。这可能与深圳居民下班后的社交活动、休闲娱乐活动较为丰富有关。

(3) 道路交通设计和公共交通可达性方面。路网密度、道路交叉口数量和公交线路数量的相对重要性弱于其他指标,整体处于3.83%—9.72%区间,仍具有一定的影响作用,提高站点周边步行可达性和公交与地铁接驳

便利度对客流量有一定的促进作用。

(4) 到目的地距离方面。距CBD距离的相对重要性整体处于较高水平,对全天、早高峰时段和非高峰时段客流量重要性排序均为第4位,对晚高峰时段和夜间时段则分别上升至第2位和第1位。早高峰交通拥堵和上班通勤时间紧迫的情况下,城市轨道交通仍是中远距离通勤者的优选出行方式,而下班返程的时间紧迫性有所下降,居民返程的交通方式和时间段可选择性增加,因而距CBD距离的影响度有所降低。

### 2.2 站点建成环境对不同时间段客流量的非线性影响分析

相比于影响系数不变的线性模型结果,由GBDT模型结果可知,建成环境与不同时间段客流量之间均存在明显的非线性关系,大部分的影响结果存在有效的阈值范围。

“建筑容积率”对5个时间段的客流量均呈现明显的正相关关系(见图2)。全天客流量整体呈现“凸”形特征,影响波动幅度较大,存在2个陡增、1个剧减的临界值和3个平台。当建筑容积率从0.90增长到1.01时,客流量激增2740人次;建筑容积率在1.01—1.90区间变化浮动较小;当容积率由1.90升至2.00时客流量快速增长到全天高峰,与已有研究容积率2.00的阈值效果一致<sup>[21]</sup>;当建筑容积率由2.60

上升至2.78时,客流量明显下降。其他时段第一个激增区间与全天时段具有一致性(建筑容积率在1.00左右),建筑容积率对早高峰时段客流量的敏感性高于其他时段,晚高峰次之。这2个时段在容积率为1.40—1.44区间出现第二个激增段;容积率大于1.00之后,非高峰和夜间时段客流量平缓增加,影响效力逐渐减缓。值得注意的是,建筑容积率在2.30时对早高峰客流量的影响值达到峰值,随后影响力逐渐下降;当容积率至2.75后,影响作用不再产生显著的变化。可见,在一定范围内,开发强度越高,该区域可容纳的人口与就业岗位密度越大,更能诱发居民选择公共交通出行,但当容积率超过一定的舒适度和安全感时,反而对公共交通出行存在消极作用,继而使出行量减少<sup>[22]</sup>。

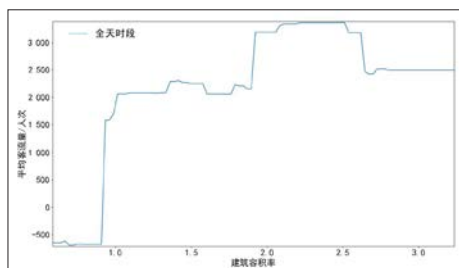
“城中村私宅建筑面积占比”对客流量有着明显的正相关关系(见图2)。在全天客流量中,城中村占比从20.0%升至21.5%时,客流量急速增加,作用效果十分明显;当城中村占比超过44.0%时,全天客流量保持在高稳定位,但对全天客流量增加的作用不再显著。对早高峰时段而言,当城中村占比从0到45.0%时,早高峰客流量逐步平稳增加;城中村占比从45.0%增长至57.0%,早高峰客流量快速增加,随后客流量基本保持不变。城中村占比对

表2 城市轨道交通站点周边建成环境对不同时间段客流量的相对重要性

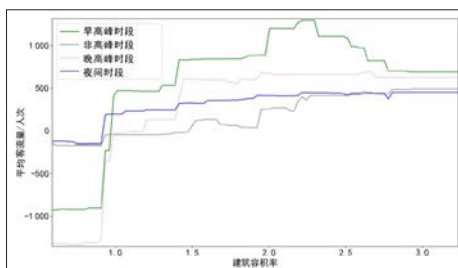
Tab.2 The relative importance of the built environment around urban rail transit stations on passenger flow at different time periods

变量名称	全天		早高峰时段		非高峰时段		晚高峰时段		夜间时段	
	排序	相对重要性/%	排序	相对重要性/%	排序	相对重要性/%	排序	相对重要性/%	排序	相对重要性/%
建筑容积率	1	14.15	1	16.09	3	10.83	1	15.87	6	8.61
城中村私宅建筑面积占比	6	8.89	5	8.41	8	9.02	7	8.12	5	10.16
土地利用混合度	10	6.38	10	4.43	6	9.45	10	5.01	8	7.51
居住用地比例	2	13.79	3	15.35	2	11.96	4	12.26	3	12.05
商业商务用地比例	3	13.04	2	15.47	1	15.56	3	13.02	2	12.23
公共服务用地比例	5	9.57	7	8.06	7	9.19	5	10.90	4	12.03
路网密度	9	6.85	6	8.13	5	9.72	8	6.42	9	5.43
道路交叉口数量	7	8.77	9	6.34	10	6.49	9	5.56	10	3.83
公交线路数量	8	7.57	8	6.95	9	7.77	6	9.35	7	8.53
距CBD距离	4	10.99	4	10.77	4	10.01	2	13.49	1	19.62
训练集		0.75		0.74		0.75		0.73		0.72
测试集		0.87		0.86		0.84		0.82		0.81

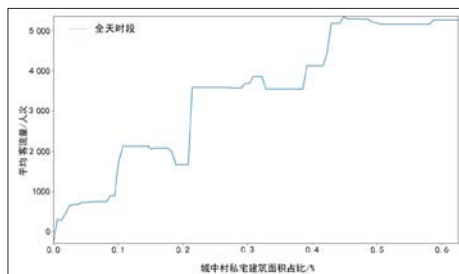
资料来源:笔者自制。



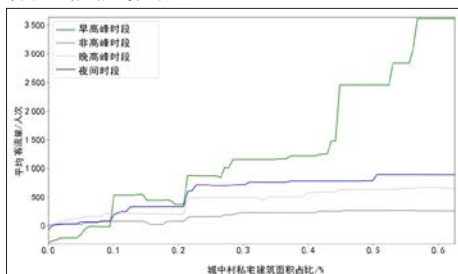
a 建筑容积率对全天客流量的非线性影响



b 建筑容积率对早高峰、非高峰、晚高峰和夜间客流量的非线性影响



c 城中村私宅建筑对全天客流量的非线性影响



d 城中村私宅建筑对早高峰、非高峰、晚高峰和夜间客流量的非线性影响

注：纵坐标以站点平均值客流量为基准值。  
图2 密度要素对不同时间段客流量的非线性影响  
Fig.2 The nonlinear effect of density factors on passenger flow at different time periods

资料来源：笔者自绘。

非高峰时段、晚高峰时段和夜间时段客流量的影响阈值在20.0%左右，随后增量维持在一个稳定的区间。可见，深圳城中村高居住密度的空间属性和居住者选择公共交通出行的社会属性是提高客流量的重要因素，20.0%和45.0%是城中村占比的重要阈值。

“土地利用混合度”在0.45—0.60的区间，对全天客流量有高稳定性和高促进作用的影响（见图3），当混合度在0.52左右的正向影响效果最为显著。土地利用混合度在0.60以上，影响呈现负相关趋势。“土地利用混合度”对早高峰、非高峰、晚高峰和夜间时段客流量同样存在转为负相关的阈值，但客流量整体波动幅度较全天客流量小，土地利用混合度与客流量双高的站点主要位于城市中心区，如深大、桃园、罗湖等站点。对站点周边区域规划开发时，应注意控制各类型用地的数量和比例，避免站点周边开发功能过于单一。

“居住用地比例”对不同时间段客流量的影响基本呈现正相关（见图3），对早高峰时段的影响机制相对复杂。这进一步说明居住人口密度对客流量的正向促进作用。从全天客流量来看，

居住用地占比在0—42%区间，该指标对客流量的影响作用较为缓慢，对客流量的增加效果较为微弱，除了占比为35%的特殊值外。当居住用地比例从42%上升至45%时，客流量暴增；当比例从52%上升到54%时，全天客流量快速增加，但整体增量少于42%；占比超过54%时，则不再产生显著影响。同样的，居住用地比例对其他时段客流量影响的阈值同为42%，随后居住用地比例对非高峰、晚高峰和夜间时段的影响缓慢增加直至不再显著，而早高峰时段出现明显的“凹”，居住用地在47%—54%区间，对早高峰客流量的影响有所下滑。

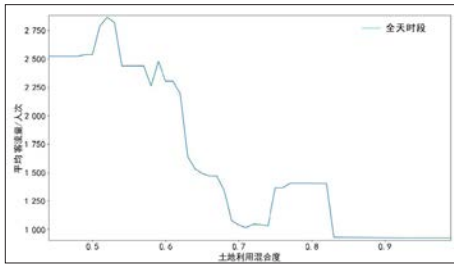
“商业商务用地比例”对不同时间段客流量影响呈现逐级上升到平稳的“阶梯”形态（见图3）。商业商务用地比例对全天客流量的阈值为33%，在0增长至33%时，全天客流量逐步缓慢上升（增加1万多人次）；当比例约为33%时，全天客流量急速增长（增加6 000多人次），随后影响不再显著。这意味着边际效应随着自变量数值的增加而逐渐降低，最后处于高稳定的状态。早高峰、非高峰、晚高峰和夜间客流量高稳定值对应的商业服务用地比例

阈值分别是31.5%、30%、37%和35.5%。总的来说，当商业商务用地占比在31.5%—37%区间，客流量影响最为显著，且相对居住用地和公共服务用地占比，商业商务占比对不同时间段客流量增量作用更为显著，可能因为商业商务用地能更好满足居民日常通勤、消费、社交等不同层面的需求。

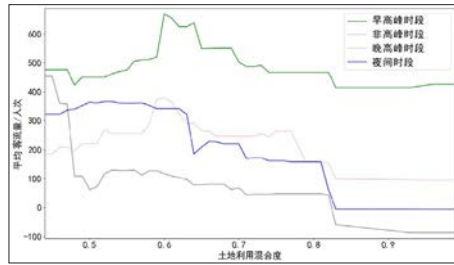
“公共服务用地比例”对不同时间段客流量的影响存在较大的差异（见图3）。其对全天、非高峰和晚高峰客流量有明显的非线性关系，但对早高峰和夜间客流量的非线性关系不显著。这与居民早高峰以通勤客流为主，非通勤目的优先选择错峰出行和休闲娱乐活动主要集中在晚高峰前后等原因有关。公共服务用地比例对全天客流量的高敏感值在6%和20%，主要集中在行政办公、教育和大型公共场所的地铁站点。

与存在明显正相关关系的非线性效果不同，“路网密度”对不同时间段客流量的非线性结果波动较大（见图4）。路网密度超过9 km/km<sup>2</sup>，全天客流量激增3 000多人次，高于路网密度对成都市地铁客流量的提升阈值（7.5 km/km<sup>2</sup>）<sup>[13]68</sup>；路网密度在13.0 km/km<sup>2</sup>左右时，全天客流量达到峰值。这说明路网密度高的区域一般街区尺度较小、步行可达性较高、与地铁站的接驳换乘越便利高效，有利于提高地铁站点客流量。

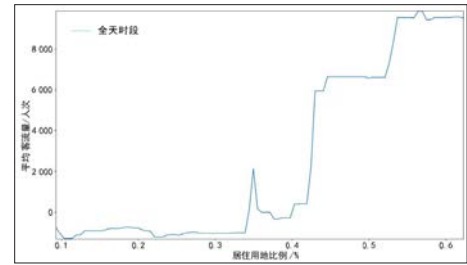
“道路交叉口数量”对不同时间段客流量都有显著的积极影响（见图4）。在全天客流量中，道路交叉口小于2 425个时，对客流量几乎没有影响；但当提升到3 400个时，客流量提升1 000多人次；当道路交叉口数量超过3 800个时，对客流量的影响不再显著。道路交叉口数量对早高峰的正向促进作用优于非高峰、晚高峰和夜间时段，其中夜间时段的敏感性最低。同样的，“公交线路数量”对全天客流量呈明显的正向影响，对其他时间段的非线性影响较弱，公交线路数量达170条时，全天客流量增量达到最高值，公共交通线路与地铁存在既互补又竞争的关系，应合理评估站点周边公交站点与公交线路的数量以及到轨道交通站点的换乘距离。



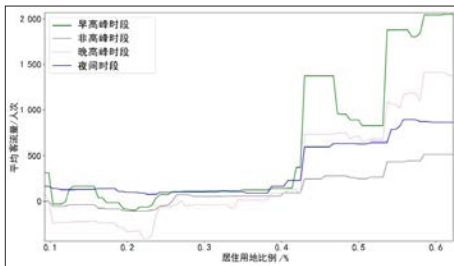
a 土地利用混合度对全天客流量的非线性影响



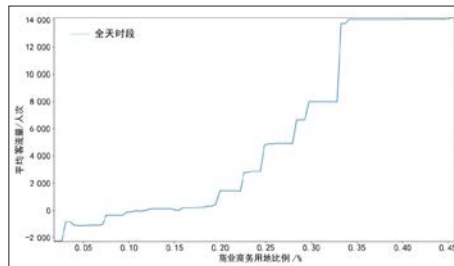
b 土地利用混合度对早高峰、非高峰、晚高峰和夜间客流量的非线性影响



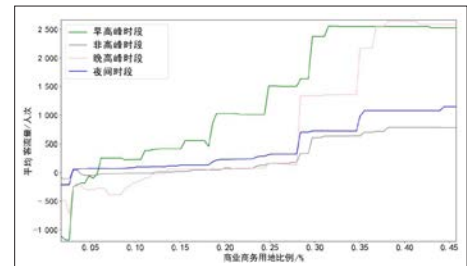
c 居住用地比例对全天客流量的非线性影响



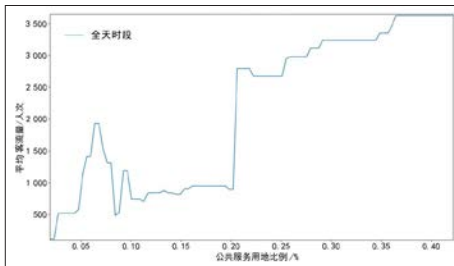
d 居住用地比例对早高峰、非高峰、晚高峰和夜间客流量的非线性影响



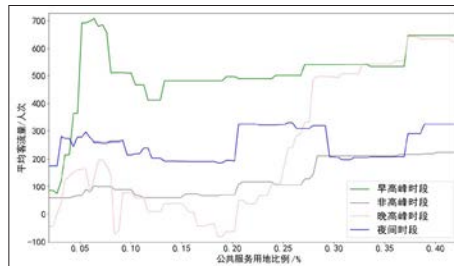
e 商业商务用地比例对全天客流量的非线性影响



f 商业商务用地比例对早高峰、非高峰、晚高峰和夜间客流量的非线性影响



g 公共服务用地比例对全天客流量的非线性影响



h 公共服务用地比例对早高峰、非高峰、晚高峰和夜间客流量的非线性影响

注:纵坐标以站点平均值客流量为基准值。

图3 多样性要素对不同时间段客流量的非线性影响

Fig.3 The nonlinear effect of diversity factors on passenger flow at different time periods

资料来源:笔者自绘。

距CBD距离与不同时间段客流量呈先逐步增加后稳定再降的非线性关系(倒U型)(见图4)。在全天客流量中,距CBD距离从0 km增加至12 km,客流量增加4 000多人次;在12.0—16.7 km区间,客流量显著性较小;当距离超过17 km后,客流量下降明显。由本文研究数据的进出站时间点可以计算得出,深圳地铁平均出行时长约为33.6 min/次,与17 km阈值结果具有一致性。城市中心区完善的服务设施配套吸引大量居民移动至城市中心区获取就业、教育、医疗等服务,但当通勤距离超过一定忍耐值时,地铁对出行者的吸引效用有所下降。另外,距CBD距离对早高峰时段客流量的敏感度高于非高峰、晚高峰和夜间时段,说明大部分居民会因

担心早高峰交通拥堵影响工作通勤,从而优先选择搭乘地铁出行,另外3个时段选择出租车、公交车等其他出行方式的可能性增加。

### 3 结论与结语

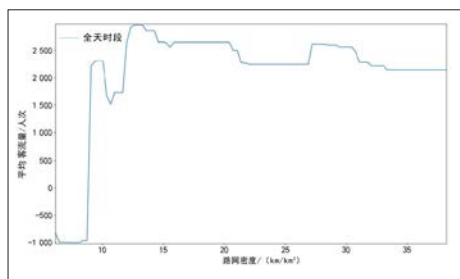
#### 3.1 结论

本文利用深圳地铁刷卡数据,构建梯度提升决策树GBDT模型,研究建成环境要素对不同时段客流量的影响。建成环境与客流量不连续、非单调、存在阈值的非线性研究结果,能更好地帮助决策者和城市规划师了解建筑环境特征对不同时间段客流量的影响机制。具体结论如下:

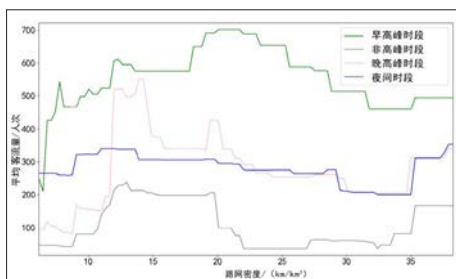
第一,各建成环境要素在不同程度上对

客流量存在明显的影响。整体而言,建筑容积率、居住用地比例、商业商务用地比例、距CBD距离的重要性最为突出。从全天客流量来看,建成环境对客流量的相对重要性从高到低依次是:建筑容积率、居住用地比例、商业商务用地比例、距CBD距离、公共服务用地比例、城中村私宅建筑面积占比、道路交叉口数量、公交线路数量、路网密度和土地利用混合度。对早高峰、非高峰、晚高峰和夜间时段,相对重要性排序会有所调整。这是由于不同时间段居民通勤、办公、休闲娱乐、社交等活动目的有所不同,各变量对轨道交通的需求会有所变化。如土地利用混合度对早高峰客流量的重要性为4.43% (第10位),非高峰时段为9.45% (第6位)。这很可能是因为早晚高峰以工作通勤为主,选择城市轨道交通的目的性较强;而非高峰时段和夜间时段以综合服务和休闲娱乐为主,土地利用混合度越高,站点周边功能越丰富,对居民吸引力更高。此研究结果能为完善站点周边建成环境提供干预顺序参考。

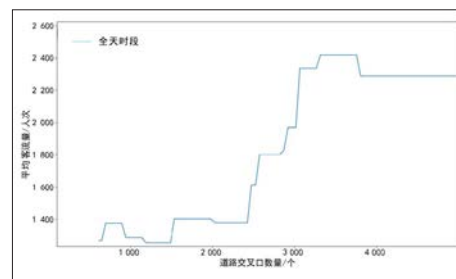
第二,建成环境与客流量存在明显的非线性关系和阈值。非线性关系主要分为以下2种:一是部分指标增长到一定范围后,对客流量的正相关影响达到高稳定值的平台区间,如建筑容积率、城中村私宅建筑面积占比、路网密度、



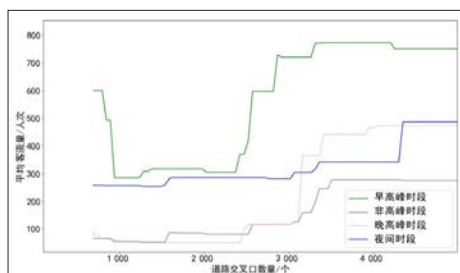
a 路网密度对全天客流量的非线性影响



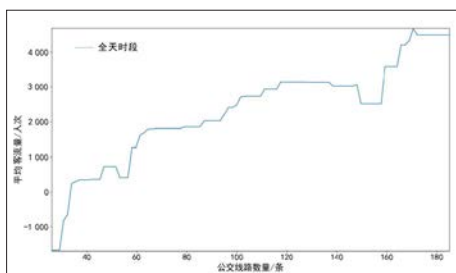
b 路网密度对早高峰、非高峰、晚高峰和夜间客流量的非线性影响



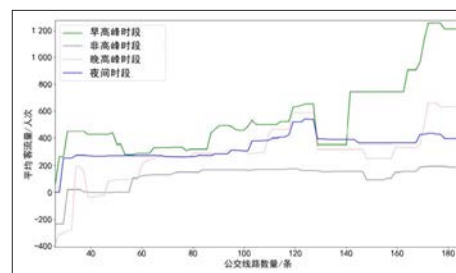
c 道路交叉口数量对全天客流量的非线性影响



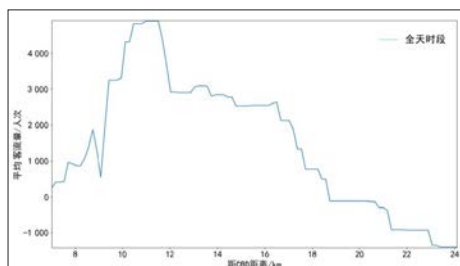
d 道路交叉口数量对早高峰、非高峰、晚高峰和夜间客流量的非线性影响



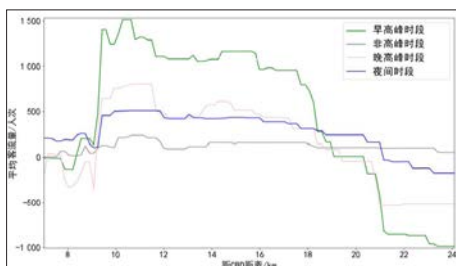
e 公交线路数量对全天客流量的非线性影响



f 公交线路数量对早高峰、非高峰、晚高峰和夜间客流量的非线性影响



g 距CBD距离对全天客流量的非线性影响



h 距CBD距离对早高峰、非高峰、晚高峰和夜间客流量的非线性影响

注: 纵坐标以站点平均值客流量为基准值。

图4 道路交通设计、公交线路可达性、到目的地距离对不同时间段客流量的非线性影响

Fig.4 The nonlinear effects of road design, destination accessibility, and distance to transit on passenger flow at different time periods

资料来源: 笔者自绘。

道路交叉口、公交线路数量等指标,说明并非无限提高该指标必然不断提升客流量,当自变量增加到一定阈值后,不再对客流量产生显著影响;如商业商务用地比例超过33%,其对客流量的影响作用保持不变。二是部分指标对客流量有一定正相关性,但到了一定阈值后影响效果转为负相关,如土地利用混合度在0.60以上和距离CBD超过17 km等情况,全天客流量呈现明显的下降趋势。同时,提高站点周边开发强度、完善步行系统、改善道路交叉口连通性、提升公交接驳效率和控制站点周边住宅户型面积与户数等措施,能有效增加客流量。掌握建成环境对客流量的阈值效应和非线性关系,为确定

各变量的优化范围和地铁网络布局规划提供更加准确的数据支撑。

第三,建成环境对不同时间段客流量的影响机制不完全相同。从相对重要性排序(见表2)和非线性影响(见图2-图4)结果来看,建成环境对全天、早高峰、非高峰、晚高峰和夜间客流量影响程度与非线性机制不同。这可以为城市轨道交通管理中出行需求优化和城市规划中建设环境平衡提供参考,而不是简单地只以每日总客流量作为唯一的研究对象。

第四,城中村对客流量具有十分显著的促进作用。深圳城中村空间属性和社会属性与普通住宅存在明显差异,其中20%和45%是城

中村占比的重要阈值。对比本文已有数据,在城中村和普通住宅的用地面积接近的条件下,城中村对客流量的促进作用更优,城中村人均居住面积远低于普通住宅,两者的人口密度差距较大。现阶段控规层面较少考虑对站点周边住宅户型面积、居住户数等“人本性”指标的控制<sup>[23]</sup>。以香蜜地铁站为例,轨道交通站点周边各居住小区的容积率均符合深圳法定图则要求,但由于大部分小区以大户型为主,因而该站点居住人口密度较低,且住户以高收入群体为主,较少选择公共交通出行,使得该地铁站客流量较少。城市轨道交通规划应优先考虑与保障性住房、地铁上盖居住设计等以小户型为主的住区规划相结合,城市规划师应考虑在控制性详细规划中增加站点周边区域的户型面积与居住套数等控制指标。

### 3.2 结语

基于建成环境与不同时间段客流量的非线性关系,可以准确分析各变量对地铁站点的影响机制,能为城市轨道交通规划建设 and 客流量预测提供技术支撑,对促进公共交通出行有十分积极的作用。但本文尚未将人口数据、就业数据、票价等社会经济属性等指标纳入,模

型对客流量的解析度有待提升,未来可考虑廉租房、保障性住房等居住属性对客流量的影响机制。技术方面通过多源数据获得建成环境对客流量的非线性关系和阈值效应,分析不同变量对不同时间段的最佳阈值,但在机器学习中,每一个变量数值的变化会影响整个模型结果的变化,所以并非将变量调整至本文提及的最佳阈值必定会得到最高的客流量。未来应结合其他数理模型和实地调研等方面的研究,才能更加全面地反映客流量的影响机制,进而更好地制定轨道交通发展政策。■

## 参考文献 References

- [1] 张赫,张建勋,王睿,等.小城市建成环境对居民出行交通碳排放的影响机理[J].城市问题,2020(7):4-10.  
ZHANG He, ZHANG Jianxun, WANG Rui, et al. Built environment factors influencing CO<sub>2</sub> emissions from residential trips in small Chinese cities[J]. Urban Problems, 2020(7): 4-10.
- [2] CERVERO R, KOCKELMAN K. Travel demand and the 3Ds: density, diversity, and design[J]. Transportation Research Part D: Transport and Environment, 1997, 2(3): 199-219.
- [3] EWING R, CERVERO R. Travel and the built environment: a synthesis[J]. Transportation Research Record: Journal of the Transportation Research Board, 2001, 1780(1): 87-114.
- [4] 张文佳,鲁大铭.影响时空行为的建成环境测度与实证研究综述[J].城市发展研究,2019,26(12):9-16.  
ZHANG Wenjia, LU Daming. Measuring built environment for spatiotemporal behavior studies: a review[J]. Urban Development Studies, 2019, 26(12): 9-16.
- [5] 孙斌栋,但波.上海城市建成环境对居民通勤方式选择的影响[J].地理学报,2015,70(10):1664-1674.  
SUN Bindong, DAN Bo. Impact of urban built environment on residential choice of commuting mode in Shanghai[J]. Acta Geographica Sinica, 2015, 70(10): 1664-1674.
- [6] 盛来芳,宋彦.城市土地利用对轨道交通运营的影响——以纽约地铁为例[J].城市交通,2012,10(2):33-39.  
SHENG Laifang, SONG Yan. Impact of urban land use development on the rail transit operation: a case study of New York City subway[J]. Urban Transport of China, 2012, 10(2): 33-39.
- [7] ZHAO J, DENG W, SONG Y, et al. What influences metro station ridership in China? Insights from Nanjing[J]. Cities, 2013, 35: 114-124.
- [8] 仝德,高静,龚咏喜.城中村对深圳市职住空间融合的影响——基于手机信令数据的研究[J].北京大学学报(自然科学版),2020,56(6):1091-1101.  
TONG De, GAO Jing, GONG Yongxi. Impact of urban village on job-housing balance in Shenzhen: a study using mobile phone signaling data[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2020, 56(6): 1091-1101.
- [9] CHAVA J, PETER N, REENA T. Gentrification of station areas and its impact on transit ridership[J]. Case Studies on Transport, 2018, 6(1): 1-10.
- [10] 任鹏,彭建东,杨红,等.武汉市轨道交通站点周边地区职住平衡与建成环境的关系研究[J].地球信息科学学报,2021,23(7):1231-1245.  
REN Peng, PENG Jiandong, YANG Hong, et al. Relationship between jobs-housing balance and built environment in areas around urban rail transit stations of Wuhan[J]. Journal of Geo-information Science, 2021, 23(7): 1231-1245.
- [11] 李清嘉,彭建东,杨红.武汉市不同站域建成环境与轨道交通站点客流特征关系分析[J].地球信息科学学报,2021,23(7):1246-1258.  
LI Qingjia, PENG Jiandong, YANG Hong. Research on relationship analysis between passenger flow characteristics of rail transit stations and built environment of different station areas in Wuhan[J]. Journal of Geo-information Science, 2021, 23(7): 1246-1258.
- [12] 仝照民,安睿,刘耀林.建成环境对居民通勤方式选择的影响——以武汉市城中村为例[J].地理科学进展,2021,40(12):2048-2060.  
TONG Zhaomin, AN Rui, LIU Yaolin. Impact of the built environment on residents' commuting mode choices: a case study of urban village in Wuhan City[J]. Progress in Geography, 2021, 40(12): 2048-2060.
- [13] 崔叙,喻冰洁,杨林川,等.城市轨道交通出行的时空特征及影响因素非线性机制——基于梯度提升决策树的成都实证[J].经济地理,2021,41(7):61-72.  
CUI Xu, YU Bingjie, YANG Linchuan, et al. Spatio-temporal characteristics and non-linear influencing factors of urban rail transit: the case of Chengdu using the gradient boosting decision tree[J]. Economic Geography, 2021, 41(7): 61-72.
- [14] 周扬,邵天元,钱才云.南京市城市轨道交通站点周边地区建成环境对居民活动的影响——基于梯度提升决策树与SHAP解释模型的分析[J].科学技术与工程,2023,23(17):7509-7519.  
ZHOU Yang, SHAO Tianyuan, QIAN Caiyun. Influence of built environment on residents' activities in Nanjing urban rail transit station areas: an analysis based on gradient boosting decision tree and SHAP interpretation model[J]. Science Technology and Engineering, 2023, 23(17): 7509-7519.
- [15] DING C, CAO X, 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.
- [16] 申犁帆,王焱,张纯,等.轨道站点合理步行可达范围建成环境与轨道通勤的关系研究——以北京市44个轨道站点为例[J].地理学报,2018,73(12):2423-2439.  
SHEN Lifan, WANG Ye, ZHANG Chun, et al. Relationship between built environment of rational pedestrian catchment areas and URT commuting ridership: evidence from 44 URT stations in Beijing[J]. Acta Geographica Sinica, 2018, 73(12): 2423-2439.
- [17] GONG P, CHEN B, LI X, et al. Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018[J]. Science Bulletin, 2020, 65(3): 182-187.
- [18] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine[J]. The Annals of Statistics, 2001, 29(5): 1189-1232.
- [19] 丁聪,倪少权,吕红霞.基于梯度提升的城市轨道交通客流量预测分析[J].城市轨道交通研究,2018,21(9):60-63.  
DING Cong, NI Shaoquan, LYU Hongxia. Forecast and analysis of urban rail transit passenger flow based on gradient boosting[J]. Urban Mass Transit, 2018, 21(9): 60-63.
- [20] 夏陈红,翟国方.基于GBDT机器学习算法的国土空间演变动力机制研究——以长江三角洲为例[J].城市发展研究,2022,29(11):12-19.  
XIA Chenhong, ZHAI Guofang. Research on dynamic mechanism of territorial space evolution in Yangtze River Delta based on GBDT machine learning algorithm[J]. Urban Development Studies, 2022, 29(11): 12-19.
- [21] 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 Lifan, HU Yiwen, 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.
- [22] SUN B D, ERMAGUN A, DAN B. Built environmental impacts on commuting mode choice and distance: evidence from Shanghai[J]. Transportation Research Part D: Transport and Environment, 2017, 52: 441-453.
- [23] 本刊编辑部.“空间治理体系下的控制性详细规划改革与创新”学术笔谈会[J].城市规划学刊,2019(3):1-10.  
The Editorial Department. Academic symposium on "regulatory detailed planning reform and innovation under the space governance system"[J]. Urban Planning Forum, 2019(3): 1-10.