

土地利用视角下城市轨道交通线路分类 及其客运量影响因素研究

张霄宇¹, 许奇^{1,2*}, 宋丽英², 邓进³

(1. 北京交通大学综合交通运输大数据应用技术交通运输行业重点实验室, 北京 100044;

2. 北京交通大学交通运输学院, 北京 100044; 3. 北京城建交通设计研究院有限公司, 北京 100037)

摘要:理解城市轨道交通客流与土地利用的依赖关系有助于合理配置资源以推动TOD(Transit-oriented development)模式的实施。既有研究多以车站客流及其建成环境为研究对象揭示客流的关键影响因素,从全局视角下分析线路客流对其沿线土地利用依赖关系的研究不够充分。论文基于2019年中国36座主要城市的多源数据构建包含线路功能和建成环境两方面的土地利用以及交通服务等三类指标,并以线路功能和建成环境两类指标为聚类特征划分线路类型,采用XGBoost(eXtreme Gradient Boosting)模型分析不同类型线路客运量的影响因素。案例研究表明:根据线路服务范围及城市发展水平的差异,城市轨道交通线路可分为市区线I型、市区线II型和郊区线。3类线路重要度排序前5的解释变量对客运量的贡献至少达到60%,其中就业岗位密度和中心城区线路覆盖率是3类线路共同的重要土地利用变量。重要度排序高的解释变量对客运量的影响具有阈值效应,且相同类型线路重要变量的阈值趋同,而不同类型线路之间的阈值存在显著差异。因此,TOD客流廊道的构建需因地制宜地统筹职住分布及其配套资源。研究结果有助于理解土地利用影响城市轨道交通线路客流的机理,可为促进TOD客流廊道的发展提供参考。

关键词:城市交通;土地利用;线路客运量;机器学习模型;阈值效应

城市轨道交通客流与土地利用的依赖关系是城市交通研究的经典问题,其目标是促进城市轨道交通与土地利用一体化发展^[1-2]。TOD(transit-oriented development)引导轨道交通与城市高质量融合发展,鼓励居民使用公共交通出行。实践中多数TOD项目仅在车站周边地块层面实施,土地利用对客流的影响研究也多从站点视角下开展。然而,在更大的空间尺度下,TOD还需促进城市用地结构的优化以支撑可持续发展^[3],据此TOD发展仍需统筹协调城市轨道交通车站用地开发以形成“节点+廊道”的城市形态^[4]。因此,分析线路沿线土地利用对其客运量的影响,将有助于合理配置资源以推动TOD模式的成功实施。

城市轨道交通客流与土地利用的研究多基于站点展开。针对研究对象,既有文献关注车站进/出站客流^[5-10]和起讫点(OD)客流^[11-13]等客流与土地利用之间的关系。结果表明,职住密度显著正向影响车站客流^[8-12,14],慢行优先的街区设计也对客流产生积极效果^[8,11,15]。混合用地对客流的影响源于完善的服务设施所产生的吸引力^[7-8,14],但贺鹏等^[9]针对北京地铁的研究发现用地混合度降低正向影响客流,表明土地利用对客流的影响具有异质性。此外,针对OD客流的研究表明旅行时间等出行阻抗亦对客流影响显著^[12-13]。

从研究方法上看,既有研究假定车站客流与土地利用存在线性关系并采用最小普通二乘法估计^[5],

收稿日期:2023-04-13;修订日期:2023-10-23。

基金项目:国家自然科学基金项目(71621001)。[Foundation: National Natural Science Foundation of China, No. 71621001.]

第一作者简介:张霄宇(1998—),男,甘肃张掖人,硕士生,主要研究方向为城市轨道交通客流与土地利用。

E-mail: 21120954@bjtu.edu.cn

*通信作者简介:许奇(1982—),男,云南普洱人,博士,副教授,硕士生导师,主要研究方向为轨道交通发展与土地利用一体化。E-mail: xuqi@bjtu.edu.cn

引用格式:张霄宇,许奇,宋丽英,等.土地利用视角下城市轨道交通线路分类及其客运量影响因素研究[J].地理科学进展,2023,42(12): 2337-2349. [Zhang Xiaoyu, Xu Qi, Song Liying, et al. Classification of urban rail transit lines and determinants of the ridership from the perspective of land use. Progress in Geography, 2023, 42(12): 2337-2349.] DOI: 10.18306/dlkxjz.2023.12.005

该模型未充分考虑客流分布及建成环境的空间特征。针对此问题,空间模型能较好地处理空间效应对估计结果的影响。根据估计参数是否随样本空间位置的变化而改变,这类模型又可分为空间计量模型^[6]和地理加权回归模型^[7,9],彭诗尧等^[16]的研究表明后者的估计效果优于前者。近年来基于机器学习的建模技术进一步放宽线性假设并发现土地利用对客流的影响存在非线性特征且表现为阈值效应^[8-9,14-15,17-18]。此外,土地利用类型及其面积或传统数据已被POI (point of interest) 等多源城市大数据替代,为当前分析土地利用对客流的影响提供新的数据环境^[17,19],进一步推动了机器学习的广泛应用。

建成环境对城市轨道交通车站客流的影响存在非线性的空间效应^[9],但另一方面,线路是车站的集合,其客流是各车站乘降量累加的结果,这种空间尺度的粗粒化处理以及影响效果的平均化处理,既有的分析框架是否适用仍是亟需回答的关键问题,具体如下:首先,从研究对象上看,粗粒化处理导致样本规模降低一个数量级,为此需扩大研究范围,从单个城市扩充为开通城市轨道交通运营的所有城市。研究对象及其规模变化后,不同类型线路客运量的影响因素是否存在差异,以及不同城市相同类型的线路对客运量的影响是否存在普遍规律,仍然是值得研究的问题。其次,从研究方法上看,空间权重矩阵难以刻画不同城市的不同线路之间的关系,因此空间回归模型难以适用于该问题。同时,土地利用对客流的影响存在非线性特征已被许多研究证实,如何构建非参数的机器学习模型是研究线路客运量影响因素的关键问题。最后,从结果解释上看,基于线路的结果与基于车站的结论是否一致,是否有助于理解TOD廊道构建的土地利用因素,具有实践价值。

综上,以36个中国主要城市轨道交通系统为案例,融合城市多源数据刻画线路特征,在基于土地利用因素的线路分类研究的基础上,构建基于机器学习的多维度客流影响因素分析模型,分析土地利用影响不同类型线路客运量的非线性特征和阈值效应,以期为促进线路沿线用地深度融合,推动TOD客流廊道发展提供参考。

1 研究区域、方法及数据

1.1 研究区域与数据来源

本文研究区域为截至2020年中国开通城市轨

道交通运营的36座城市(港澳台地区数据暂缺)。根据第七次人口普查数据,这些城市包括北京等7座超大城市、武汉等14座特大城市和南宁等15座大城市。研究区域及其城市轨道交通规模如图1所示。

研究数据主要包括:

(1) 城市轨道交通系统数据。选取36座城市已运营的城市轨道交通线路164条、车站3304座(换乘站不重复计算),包含地铁、轻轨、单轨和磁悬浮等4种系统制式。

(2) 交通运行情况。城市轨道交通线路客流数据与运行指标采用2019年各运营企业统计数据,主要包括线路日均客运量(万乘次)和高峰小时最小发车间隔。

(3) 土地利用状况。基于高德地图和百度地图等开放平台,获取2019年城市主要建成区建筑轮廓数据与全部139类二级行业POI数据。POI数据包括名称、经纬度坐标与行业分类信息。城市环路、绕城快速路或高速公路等城市道路数据通过开源地图网站OpenStreetMap获取,主要包括名称、编号、道路等级信息。

(4) 职住数据。就业岗位数据采信《中国城市轨道交通沿线发展态势年度报告》^[20],其融合通信运营商、移动互联网、国家统计局部门的时空大数据,扩样后测算主要城市的就业岗位数据。居住人口基于第七次人口普查数据对worldpop.org的100 m×100 m栅格数据修正获取。上述数据均采用2019年数据计算。

1.2 解释变量及其计算方法

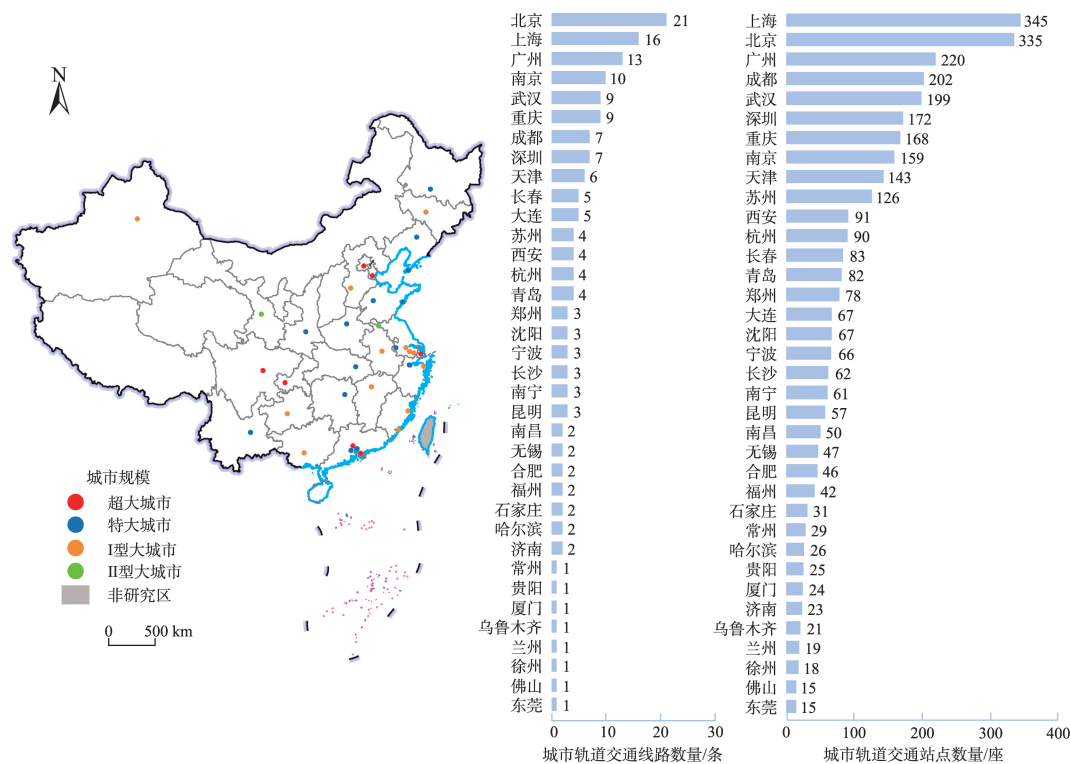
影响城市轨道交通线路客运量的社会经济环境因素繁多且复杂。根据数据可获得性,从土地利用和交通服务等两方面构建16个解释变量以刻画线路客运量的影响因素,如表1所示。其中土地利用又包含线路功能和建成环境两个维度。

(1) 线路功能变量

城市轨道交通沿线土地利用影响其线路功能与定位。城市中心城区是社会经济活动的核心区域,城市轨道交通主要服务上述区域。同时,在城市快速扩张且职住分离背景下,连接中心城区与郊区的轨道交通廊道是缓解向心通勤压力的重要手段。这种对中心城区的强支撑作用是城市轨道交通线路功能和定位的重要体现。鉴于此,采用中心城区线路比例等8个变量刻画线路功能的特征^[21]。

(2) 建成环境变量

建成环境是因人类生产和生活等活动形成的



注：该图基于自然资源部标准地图服务网站下载的审图号为GS(2023)2763号标准地图制作，底图无修改。

图1 研究城市空间分布与其轨道交通线路及站点数量

Fig.1 Spatial distribution of the studied cities and the number of rail transit lines and stations

表1 城市轨道交通线路客运量影响因素及其变量说明

Tab.1 Factors influencing urban rail transit ridership and description of the corresponding variables

因素	解释变量	符号	变量说明
线路功能	中心城区线路比例	L_1	位于中心城区的线路长度与其运营里程之比
	中心城区车站比例	L_2	位于中心城区的车站与线路全部车站之比
	中心城区站间距	L_3	位于中心城区的车站的平均站间距(km)
	非中心城区站间距	L_4	位于非中心城区的车站的平均站间距(km)
	中心城区线路覆盖率	L_5	中心城区的车站影响范围与其面积之比(%)
	非中心城区线路覆盖率	L_6	非中心城区的车站影响范围与其面积之比(%)
	与市中心的距离	L_7	线路所有车站与市中心的平均欧氏距离(km)
	换乘站比例	L_8	线路中换乘站与线路全部车站之比
建成环境	居住人口密度	D_1	线路覆盖区域内常住人口与其面积之比(万人/km ²)
	就业岗位密度	D_2	线路覆盖区域内就业岗位与其面积之比(万个/km ²)
	商业服务设施POI密度	D_3	线路覆盖区域内商业类POI与其面积之比(万个/km ²)
	公共服务设施POI密度	D_4	线路覆盖区域内公共服务类POI与其面积之比(1000个/km ²)
	土地利用混合度	D_5	线路覆盖区域内8类POI土地利用类型的熵值
	经济行业混合度	D_6	线路覆盖区域内20类POI经济行业类型的熵值
	轨道交通车站距离	D_7	车站影响范围内建筑与车站欧氏距离的平均值(km)
交通服务	高峰小时最小发车间隔	T_1	线路高峰小时的最小列车发车间隔(min)

注：中心城区是城市市区范围内政治、经济、社会、文化活动的核心区域，其范围根据最新版的城市规划确定；线路覆盖区域是以车站中心为圆心且半径为500 m的圆叠加而成的区域，圆形重叠部分不重复计算。

人居环境状态。Cervero等^[22]和Ewing等^[23]提出刻画轨道交通与城市融合发展的5D特征。本文采用刻画土地利用的密度(density)、混合用地(diversity)和城市设计(design)等3D特征刻画线路沿线的建成环境。

首先,城市轨道交通沿线聚集的职住人口为其发展提供潜在客流需求,故分别采用职住密度等2个变量刻画该特征;另外,高密度用地开发是线路沿线建成环境的显著特征,故采用商业服务设施和公共服务设施的POI密度等2个变量刻画该特征。其次,混合功能则是另一个显著的建成环境特征,采用信息熵的概念刻画土地利用混合度和经济行业混合度^[16]。当POI根据《城市用地分类与规划建设用地标准》划分为8类时,信息熵刻画了土地利用的多样性;当根据《国民经济行业分类》划分为20类时,则反映了经济行业混合度。最后,轨道交通与城市的融合依赖慢行出行环境的营造,其邻近性是关键特征,故采用线路车站影响范围内建筑体与车站欧氏距离的平均值刻画。

(3) 交通服务变量

城市轨道交通系统的服务水平及差异化是影响客流的关键因素。东京等国际都市通过提供不同速度等级列车服务以满足不同乘客的出行需求,极大地提高出行效率^[24]。然而,国内城市轨道交通系统的网络化运营程度较低,仅少部分线路采用快慢车等运营技术。因此,针对交通服务的特征,采用线路的高峰小时最小发车间隔刻画其服务水平。

1.3 研究方法

1.3.1 机器学习模型

土地利用解释变量对城市轨道交通客流的影响存在显著的非线性特征,极度梯度提升(eXtreme Gradient Boosting, XGBoost)是研究上述二者依赖关系的重要方法。同时,该方法对样本空间效应的隐式建模可有效克服空间权重矩阵无法刻画不同线路之间空间特征的局限性。因此,采用XGBoost模型研究线路客运量与其影响因素间关系,即

$$\hat{F}(x) = \sum_{m=1}^M \xi_m f_m(x) \quad (1)$$

式中: $\hat{F}(x)$ 为迭代结束后的模型预测的线路客运量; x 为表1中确定的影响线路客运量的解释变量; ξ_m 为第 m 棵树的权重; $f_m(x)$ 为第 m 棵树的预测结果; M 为模型迭代次数。

1.3.2 机器学习模型的解释工具

XGBoost模型可解释性不及回归模型。针对

可解释性差的黑箱问题,将基于博弈论的SHAP(Shapley Additive exPlanations)值作为XGBoost模型的解释工具,其衡量解释变量对因变量预测结果的边际贡献度,可用于分析解释变量的重要性和非线性影响。

解释变量针对单一样本计算SHAP值,其正负及数值反映解释变量对线路客运量预测结果的正负及影响程度。基于解释变量的各条线路的SHAP值计算全局重要度,该值越大,表明该变量对模型预测结果的影响越显著,计算公式如下:

$$I_i = \frac{1}{N} \sum_{j=1}^N |S_j^i| \quad (2)$$

式中: I_i 为解释变量 i 的全局重要度; S_j^i 为解释变量 i 在样本 j 的SHAP值; N 为样本数。

部分依赖图(partial dependence plot, PDP)是解释机器学习模型预测结果的另一重要工具。通过类似回归分析中边际效益分析的思路,PDP实现对因变量与关键的或期望的解释变量 x_s 之间非线性关系的可视化,以分析解释变量的阈值效应,计算公式如下:

$$f(x_s) = \frac{1}{N} \sum_{j=1}^N f(x_s, x_c^j) \quad (3)$$

式中: $f(x_s)$ 为预测线路客运量与解释变量 x_s 间的非线性函数; $f(x_s, x_c^j)$ 为样本 j 的相应函数; x_c 为除 x_s 外剩余解释变量构成的补集, x_s 仅考虑单一变量; x_c^j 为 x_c 中样本 j 的实际取值; x_s 的取值需要遍历 x_c 中对应的所有取值。

部分依赖图仅刻画非线性关系在所有样本下的平均效应,无法实现对个体差异的关注。因此,使用局部解释工具个体条件期望(individual conditional expectation, ICE)图揭示样本间的异质性,其计算方法与PDP相似,但仅计算个体值而无需计算所有样本的均值。

2 结果分析

2.1 线路分类结果

为分析不同类型线路客运量的影响因素是否存在差异,构建基于K-means++的城市轨道交通线路聚类算法。既有研究多以各类车站客流指标为特征构建其功能识别算法,并以建成环境的差异解释聚类结果的合理性^[25]。据此,本文仍以线路客流量为聚类特征,以降低因城市间社会经济环境等因

素的差异对分类结果的影响,此外,还增加表1中“线路功能”和“建成环境”维度的15个解释变量为聚类特征,重点从土地利用视角识别线路类型。

城市轨道交通线路分类结果为3种类型:市区线I型、II型和郊区线。市区线I型共计52条,其中40条分布在京沪穗深渝蓉等6座超大城市、12条则位于汉宁杭镐等4座头部特大城市;市区线II型共计73条,其中20条分布在京津沪穗渝蓉等6座超大城市、23条位于特大城市、30条位于I、II型大城市;郊区线共计39条,其中20条位于京津沪穗深蓉等超大城市、19条分布在特大城市以及I型大城市苏州。

从线路空间分布看,市区线服务中心城区,其中I型线路构成超大城市和头部特大城市的核心网络,II型线路则构成特大/大城市的核心网络(也包括部分服务超大城市中心城区外围的线路)。郊区线主要服务超大城市和特大城市中心城区和外围组团,通常呈放射状向郊区延伸。图2为京沪穗深等典型超大城市线路分类及其空间分布。

表2为城市轨道交通3种类型线路客运量及其

影响因素的统计特征。由表可知,不同类型线路的土地利用和交通服务的特征存在显著差异。具体而言,由于服务的城市空间及其发展水平的差异,市区线沿线的用地开发和交通服务水平均优于郊区线。此外,虽然同为服务中心城区的线路,但由于城市间发展水平参差不齐,各城市市区线I型和II型的建成环境特征也存在较大差异。例如,前者的居住人口密度和就业岗位密度均明显大于后者。这种不同类型线路沿线发展水平的差异表明:由于不同类型线路的服务区域和建成环境不同,土地利用生成出行需求也相应存在高低,因此为满足线路客流的供给也就存在差异。

2.2 线路客运量的影响因素分析

以表1的16个因子为解释变量,以对数化后的线路日均客运量为因变量,利用Python的XGBoost工具箱实现3类线路的模型求解。模型超参数的最优组合采用网格搜索法与五折交叉验证法计算,拟合效果最佳时,基于3类线路的模型均包含100棵最大深度为3的回归树,其学习率分别为0.23、0.25、0.23,解释能力分别为71%、69%和52%。

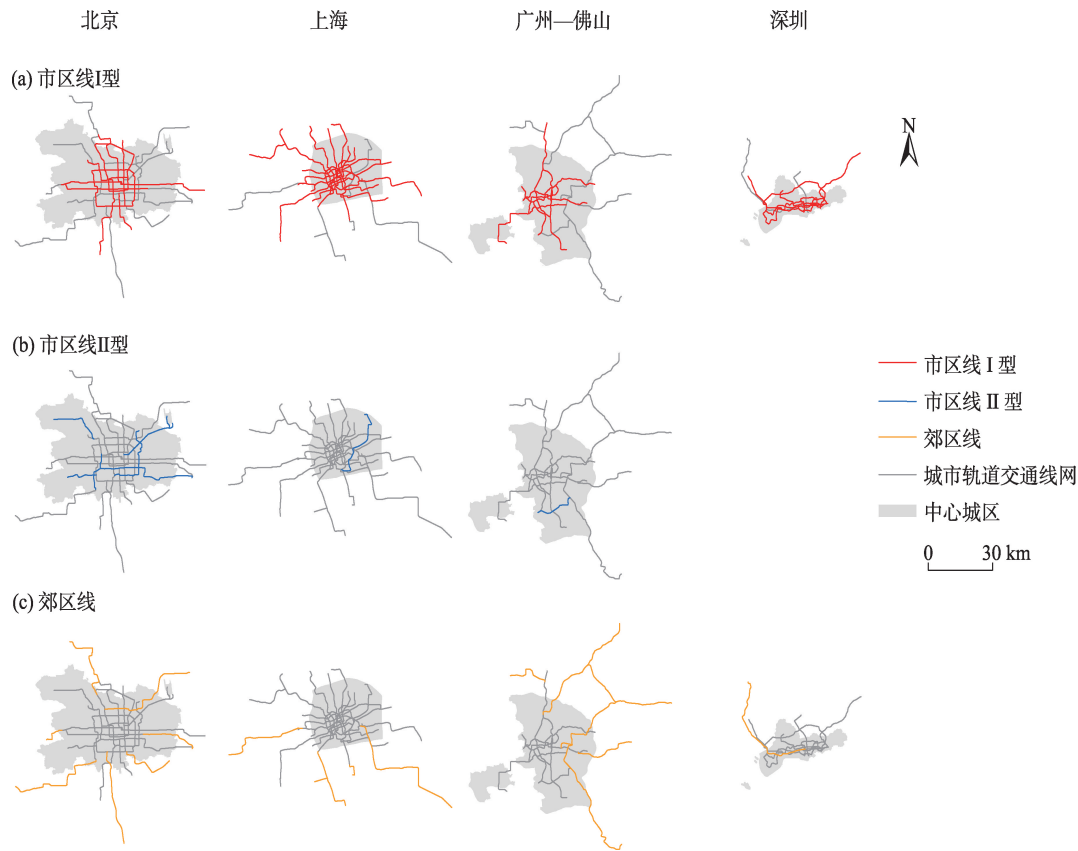


图2 典型超大城市的轨道交通线网及3类线路的空间分布
Fig.2 Spatial distribution of urban rail transit networks and lines in three types for typical supercities

表2 线路客运量解释变量的统计特征

Tab.2 Statistical results of independent variables of urban rail transit ridership at the line level

因素	变量	市区线Ⅰ型		市区线Ⅱ型		郊区线	
		均值	标准差	均值	标准差	均值	标准差
客运量	线路日均客运量/(万乘次/d)	82.275	35.301	24.156	15.198	13.715	13.227
线路功能	中心城区线路比例 L_1	0.790	0.181	0.892	0.161	0.253	0.202
	中心城区车站比例 L_2	0.816	0.158	0.899	0.142	0.293	0.216
	中心城区站间距 L_3 /km	1.344	0.266	1.355	0.317	1.943	1.056
	非中心城区站间距 L_4 /km	1.865	0.685	1.732	1.138	3.595	3.951
	中心城区线路覆盖率 L_5 %	2.512	1.284	3.582	2.341	0.665	0.843
	非中心城区线路覆盖率 L_6 %	0.785	0.619	0.926	0.633	0.595	0.425
	与市中心的距离 L_7 /km	9.243	2.702	8.013	2.950	24.503	10.025
	换乘站比例 L_8	0.297	0.117	0.153	0.114	0.128	0.082
建成环境	居住人口密度 D_1 /(万人/km ²)	2.091	0.627	1.408	0.402	0.704	0.359
	就业岗位密度 D_2 /(万个/km ²)	1.947	1.142	0.912	0.374	0.512	0.341
	商业服务设施POI密度 D_3 /(万个/km ²)	0.382	0.164	0.259	0.118	0.101	0.074
	公共服务设施POI密度 D_4 /(1000个/km ²)	0.685	0.333	0.423	0.153	0.193	0.137
	土地利用混合度 D_5	0.321	0.049	0.320	0.065	0.357	0.069
	经济行业混合度 D_6	0.658	0.025	0.655	0.033	0.658	0.043
	轨道交通车站距离 D_7 /km	0.342	0.006	0.341	0.008	0.341	0.011
交通服务	高峰小时最小发车间隔 T_1 /min	2.742	0.758	5.179	1.646	6.235	3.183

表3 城市轨道交通线路客运量解释变量的全局重要度

Tab.3 Global importance of independent variables on urban rail transit ridership at the line level

解释变量	市区线Ⅰ型		市区线Ⅱ型		郊区线	
	重要度/%	排序	重要度/%	排序	重要度/%	排序
线路功能						
中心城区线路比例 L_1	1.45	14	0	14	3.31	11
中心城区车站比例 L_2	1.01	16	0	15	0.56	15
中心城区站间距 L_3	6.38	6	1.62	9	1.91	13
非中心城区站间距 L_4	3.69	12	0	16	4.56	7
中心城区线路覆盖率 L_5	9.06	5	9.36	4	6.67	5
非中心城区线路覆盖率 L_6	5.58	7	0.67	13	9.73	4
与市中心的距离 L_7	1.11	15	5.19	6	1.14	14
换乘站比例 L_8	9.97	3	9.33	5	6.14	6
建成环境						
居住人口密度 D_1	4.09	10	4.73	7	11.18	3
就业岗位密度 D_2	10.22	2	13.34	2	22.46	1
商业服务设施POI密度 D_3	4.91	9	10.79	3	3.95	9
公共服务设施POI密度 D_4	4.09	11	3.19	8	0	16
土地利用混合度 D_5	5.45	8	1.32	10	3.71	10
经济行业混合度 D_6	1.77	13	0.82	12	4.25	8
轨道交通车站距离 D_7	9.88	4	1.16	11	3.28	12
交通服务						
高峰小时最小发车间隔 T_1	21.34	1	38.48	1	17.15	2

表3为城市轨道交通不同类型线路的解释变量的重要度。易知,影响不同线路的解释变量的重要度差异显著。市区线I型的 T_1 、 D_2 、 L_8 、 D_7 和 L_5 ,累计重要度为60.47%;市区线II型的 T_1 、 D_2 、 D_3 、 L_5 和 L_8 ,累计重要度达到81.30%;郊区线的 D_2 、 T_1 、 D_1 、 L_6 和 L_5 ,累计重要度达到67.19%。后续分析将重点针对重要度排序前5名的解释变量展开。

由表3可以看出:

(1) 就业岗位密度(D_2)、中心城区线路覆盖率(L_5)与高峰小时最小发车间隔(T_1)为3类线路客运量共同的重要解释变量。就业岗位密度是3类线路最重要的土地利用变量;中心城区线路覆盖率重要度在市区线均达到9%以上,即使是郊区线亦超过6%。显然,因车站周边集聚的就业岗位而吸引的通勤客流是城市轨道交通服务的主要对象,而就业聚集的区域通常位于社会经济活动规模庞大的中心城区。

高峰小时最小发车间隔也成为至少排序前2名的重要解释变量。此外,该变量的重要度值在不同线路之间略有差异。三类线路中,郊区线的数值最低,其影响客流的更重要的因素应是出行需求的本源因素,即土地利用;市区线I型低于II型,超大城市的出行需求更加多样化,供需平衡仅是交通服务的一方面,换乘等出行环境的提升也是影响客流的重要因素。

(2) 郊区线的重要土地利用变量与市区线不同。郊区线的就业岗位密度和中心城区线路覆盖率的重要度排序与市区线相似,但其居住人口密度(D_1)和非中心城区线路覆盖率(L_6)的重要度则均明显高于市区线,这源于郊区线主要联系中心城区的就业聚集区域和外围居住组团,服务由职住分离而产生的通勤客流。此外,换乘站比例(L_8)对郊区线的影响弱于市区线,主要是因为郊区线多为放射线,相较于市区线的换乘便捷,其更强调组团间的直达。综上,重要土地利用变量在郊区线与市区线间的差异源自郊区和中心城区用地开发的差距,也由郊区线的廊道功能及其服务范围决定。

(3) 市区线I和II型均服务中心城区,但重要土地利用变量亦存在差异。一方面,商业服务设施POI密度(D_3)对市区线I型的影响低于市区线II型。后者通常是特大/大城市的核心线路,串联中心城区最重要的商圈,良好的轨道交通可达性可生成更多的出行需求;而超大城市和头部特大城市的中

心城区发展更加充分,联通商圈仅是其市区线的功能之一。另一方面,市区线I型的轨道交通车站距离(D_7)重要度远高于市区线II型,这主要源自不同城市及其轨道交通线网规模的差异。各类城市与其轨道交通一体化发展都面临“首末一公里”问题,但特大/大城市的线网密度较低,市区线II型是公共交通的主动脉,乘客对末端出行效率不敏感;而超大城市和头部特大城市中心城区的线网密度更高,乘客获取出行服务更便捷。

综上,土地利用变量重要度在市区线和郊区线之间的差异源自3种类型线路服务范围及其沿线用地开发特征的不同,而市区线I型和II型的不同则由城市社会经济发展水平的差异而导致。上述差异及其影响程度可进一步由解释变量的SHAP值刻画,如图3所示。显然,解释变量SHAP值分布与其重要度相关,且对线路客运量的影响存在复杂关系。

2.3 重要土地利用变量的阈值效应

2.3.1 不同类型线路的阈值效应

(1) 市区线I型

城市轨道交通市区线I型的客运量与其重要土地利用变量间的非线性关系如图4所示。

由图可知,就业岗位密度和换乘站比例对线路客运量的影响呈先增加后降低的趋势,前者的阈值效应发生在1.3万个/ km^2 和3.0万个/ km^2 处附近,后者为0.25和0.45,这种双阈值特征表明资源配置存在的合理区间。线路沿线适度的就业聚集和合理换乘产生的网络化效应均能吸引更多的客流,但就业的过度聚集将引发拥挤效应,换乘站过多将破坏线网的合理拓扑结构,则对客流的影响呈负效应。

此外,轨道交通车站距离对客运量的影响呈负相关,在350 m附近具有显著的阈值效应;与之相反,中心城区线路覆盖率的影响则呈正相关,在1.0%处附近具有阈值效应。显然,城市轨道交通服务若覆盖更广的中心城区并改善慢行环境将有效提升客运量。

(2) 市区线II型

城市轨道交通市区线II型的客运量与其重要土地利用变量间的非线性关系如图5所示。与市区线I型不同,市区线II型客运量与4个土地利用变量均呈正相关,对应的阈值效应分别出现在0.4万个/ km^2 、0.05万个/ km^2 、0.4%和0.02处附近。这种市区线I型和II型的差异源自城市社会经济发展水平及城市轨道交通线网规模的差异。

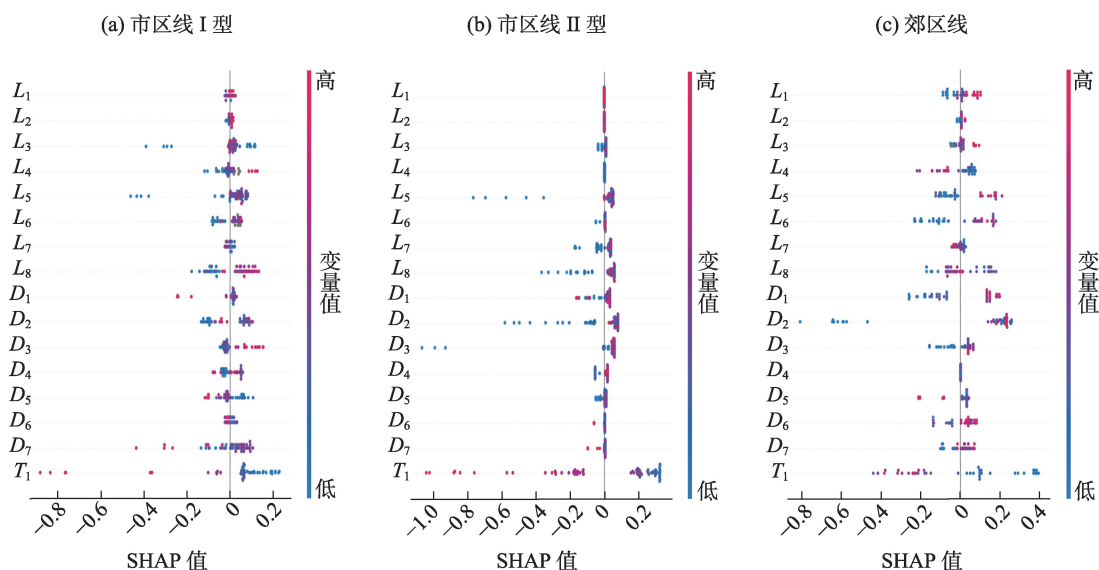
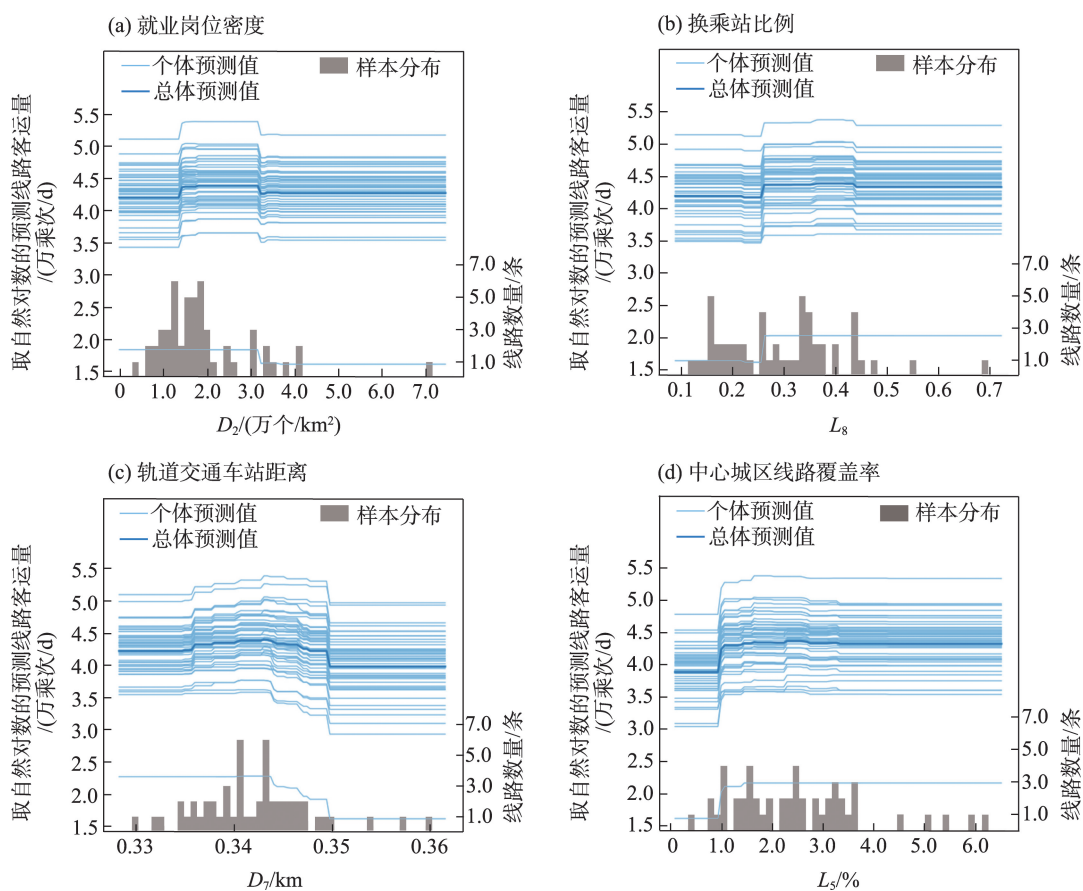


图3 线路客运量解释变量的SHAP值

Fig.3 Shapley Additive exPlanations (SHAP) value of independent variables of urban rail transit ridership at the line level



注:各条细线对应单个样本,粗线为该线路下所有样本的平均效应,即PDP;样本分布柱状图的高度表示变量数值位于对应区间内的线路数量,与回归分析中估计参数的显著性作用相似。图5~6类似。

图4 基于重要土地利用变量的市区线I型客运量的ICE图

Fig.4 ICE plots of ridership of downtown line type I based on important land use variables

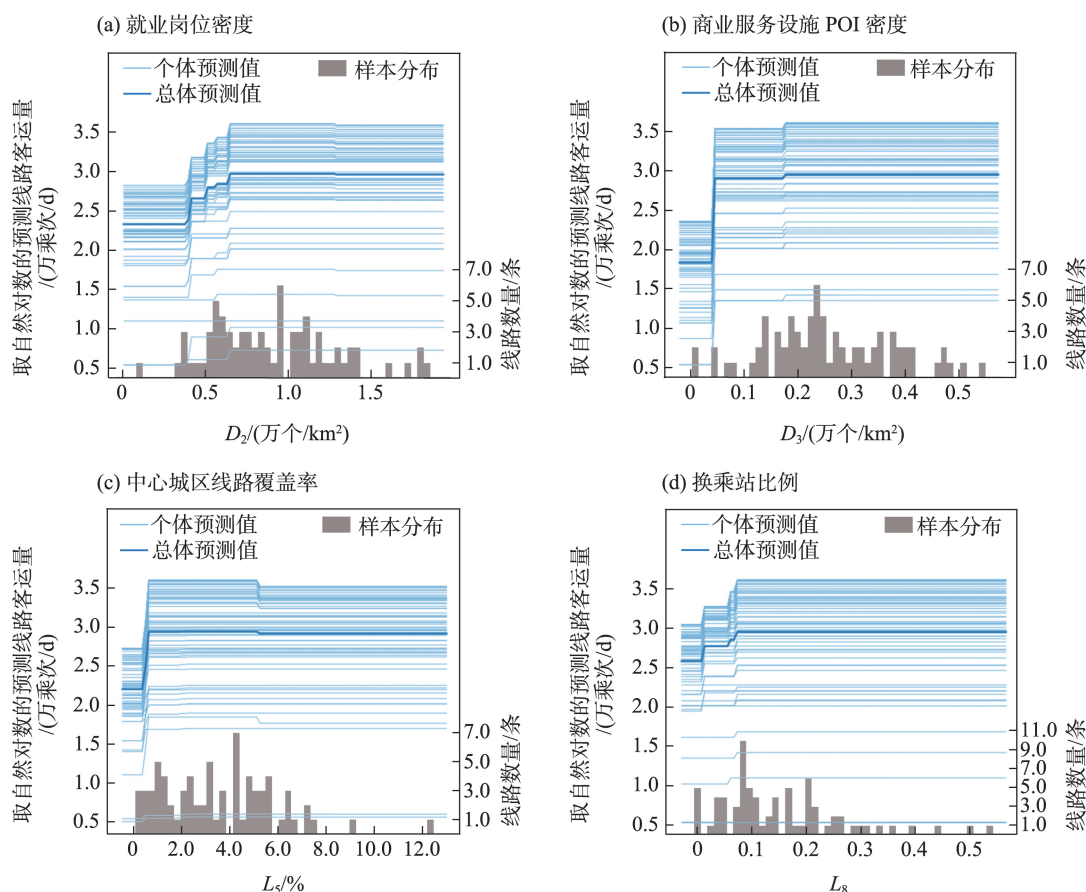


图5 基于重要土地利用变量的市区线Ⅱ型客运量的ICE图

Fig.5 ICE plots of ridership of downtown line type II based on important land use variables

(3) 郊区线

城市轨道交通郊区线客运量与其重要土地利用变量间的非线性关系如图6所示。郊区线客运量与4个土地利用变量均呈正相关,对应的阈值效应分别出现在0.3万个/ km^2 、0.5万人/ km^2 、0.4%和0.7%。郊区线与市区线的阈值高低不同源自城市轨道交通线路所服务的城市区域不同,以及这些区域用地开发水平的差异。

2.3.2 不同类型线路阈值的对比分析

根据图4~6,从总体预测线看,重要土地利用变量与城市轨道交通线路客运量均存在非线性关系并表现为阈值效应;从各样本的个体预测线看,相同类型的线路均呈现相似的趋势且与总体预测线基本一致。这表明尽管各城市社会经济环境各不相同,但线路沿线相似的土地利用特征仍是影响客流的关键因素,且产生显著影响的阈值趋同。因此,同类型线路普遍存在的阈值效应是基于单个城市研究的补充,而不同类型线路阈值的差异有助于更好地理解土地利用因素。

从城市发展水平看,城市轨道交通3类线路的就业岗位密度阈值分别为1.3万个/ km^2 、0.4万个/ km^2 和0.3万个/ km^2 ;对于就业聚集的中心城区,这3类线路的中心城区线路覆盖率阈值分别为1.0%、0.4%和0.7%。显然,市区线Ⅰ型显著高于其他2类线路。与其他类型线路分布的城市及其服务的区域相比,超大城市与头部特大城市中心城区的社会经济活动水平更高,线路沿线聚集的就业岗位也更多,激活客流增长的阈值也更大。中心城区线路覆盖率的影响机理也类似,但郊区线的阈值高于市区线Ⅱ型,其原因是二者服务范围及其沿线就业岗位分布的差异。

另外,从城市轨道交通线网布局看,市区线Ⅰ型换乘站比例的阈值达到0.25,而市区线Ⅱ型的阈值仅为0.02,究其原因不同城市的线网规模大小而导致的换乘便捷性差异^[26]。具体而言,全部市区线Ⅱ型的车站数均不足50座,且多为特大/大城市的首条线路。这些线路的换乘站数量“从无到有”或“从有到多”均能显著改善出行效率,进而显著提升客运量。而超大城市与头部特大城市的轨道交通

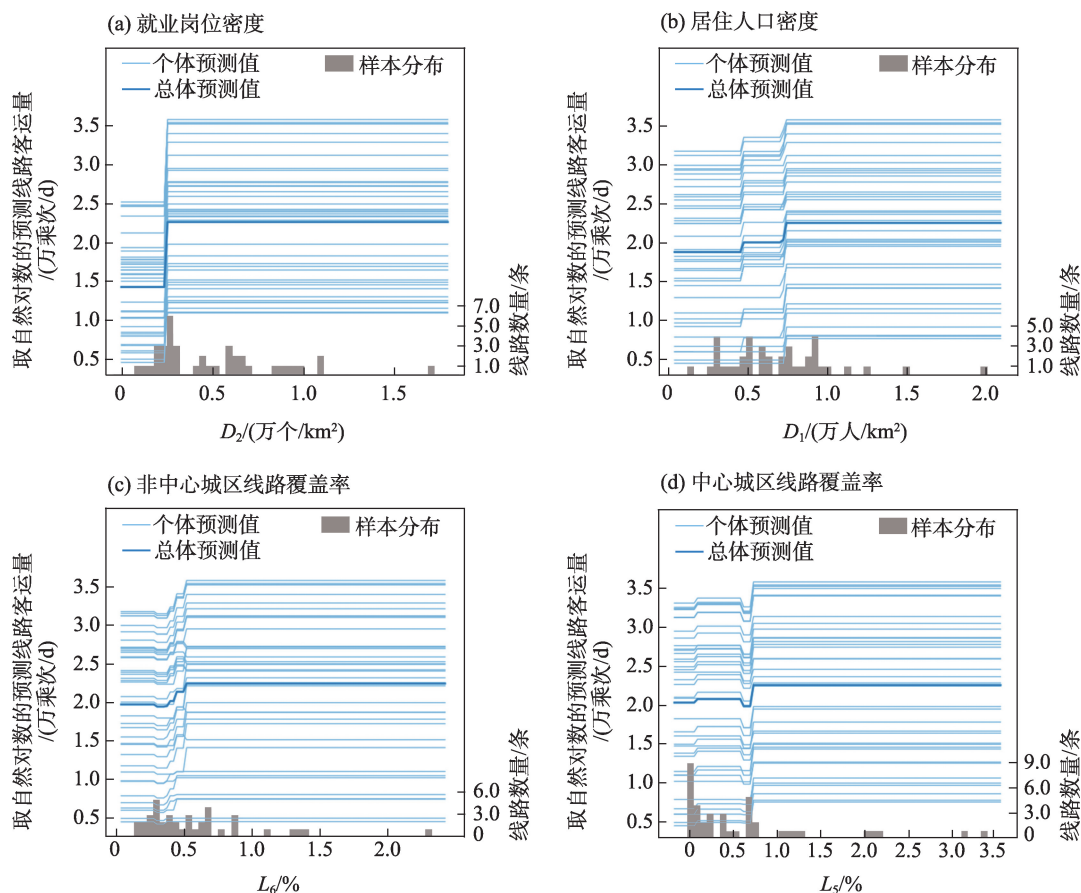


图6 基于重要土地利用变量的郊区线客运量的ICE图

Fig.6 ICE plots of ridership of suburb line based on important land use variables

则已进入网络化运营时代,较高的阈值表明线路间较多的换乘机会将进一步产生网络化效应进而吸引更多的客流。

3 结论与讨论

3.1 主要结论

本文以中国36座主要城市为案例,采用机器学习模型分析不同类型线路客运量的主要影响因素及其非线性特征,主要结论如下:

(1) 线路功能和建成环境等土地利用因素可以更细致地刻画城市轨道交通线路服务范围及其沿线用地开发的特征,以二者为指标的聚类分析方法将线路划分为市区线I型、II型和郊区线等3类,分类效果较好。市区线主要服务中心城区,其I型和II型分布的城市发展水平差异显著;郊区线联通中心城区与外围组团,沿线用地的开发程度不及中心城区。

(2) 基于XGBoost模型的研究表明:3类线路重要度排序前5的解释变量对客运量的贡献至少达到

60%,且土地利用变量的个数均为4个。就业岗位密度与中心城区线路覆盖率为3类线路共同的重要变量,而郊区线亦受居住人口密度与非中心城区线路覆盖率等因素的影响,这表明通勤是城市轨道交通服务的主要对象,客流廊道的构建需统筹职住分布。此外,超大/特大城市需要考虑线网换乘组织及效率对出行的影响,充分发挥网络化效应。

(3) 重要度排序高的解释变量对客运量的影响具有阈值效应。基于ICE图的分析表明,相同类型线路重要变量的阈值趋同,其对客运量的非线性影响具有一致的普遍规律,表明尽管不同城市发展各异,但线路沿线相似的土地利用是影响客运量的关键因素。不同类型线路之间的阈值存在显著差异,究其原因,一方面是线路服务范围及其沿线用地开发特征的不同,另一方面更是各城市社会经济发展水平存在差异,因此,客流廊道的构建需因地制宜。

3.2 讨论

首先,职住分布及服务配套设施等建成环境是决定城市轨道交通客流效果的关键因素,这与既有

站点层面^[6,8-11]的研究结论一致,表明城市轨道交通应主要服务中心城区,不同规模的城市均应提高线路沿线的开发强度,以提供支撑城市轨道交通健康发展的潜在客流需求;线路重要土地利用变量的阈值效应亦表明不同城市不同线路车站开发资源配置均宜规定差异化的下限值,以有效激活用地开发的规模效应以改善客流效果。此外,同一线路不同车站的功能及建成环境存在差异进而需统筹开发。

其次,实现线路的统筹开发需基于全局视角而非站点周边局部空间。在全局视角下,TOD客流廊道的建设需要统筹线路沿线开发以提升客运量并最终实现轨道交通与城市用地一体化发展。中心城区与郊区间土地利用存在差异,各城市间的发展水平也存在不同。在轨道交通支撑城市空间优化的过程中,郊区TOD线路应统筹土地利用“量”的增加与职住分布“质”的优化。因而在推进郊区线沿线用地开发、提升线路服务效能过程中,覆盖更多外围向通勤区域并实现与中心城区就业岗位规模的匹配应成为郊区线的实践目标。

最后,不同于对郊区发展的规划引导,城市轨道交通建设早期发展往往追随中心城区出行需求。在存量主导的城市更新背景下,继续提升城市轨道交通吸引力的关键在于通过既有线优化推动其沿线存量空间更新进而促进轨道交通与城市的深度融合。

既有线优化既需要从宏观线网层面考虑,也需要解决微观的“最后一公里”问题。既有研究多从复杂网络中心性的角度研究车站重要度对客流的影响^[8,10,19],可能导致过度关注枢纽的建设而缺乏对线网系统性的规划。因此,已进入网络化运营时代的超大和特大城市需考虑线网形态与城市空间结构的匹配问题,合理的换乘网络可吸引更多出行者使用城市轨道交通。这种网络化效应对大城市的效果更显著,表现为极小的换乘站比例阈值。此外,末端出行环境的改善对各类城市的市区线的客流效果均具有显著影响,这也表明全出行链的竞争力提升是轨道交通与城市融合的重要方向。

参考文献(References)

- [1] Hu N, Legara E F, Lee K K, et al. Impacts of land use and amenities on public transport use, urban planning and design [J]. Land Use Policy, 2016, 57: 356-367.
- [2] Zhao P J, Yang H Z, Kong L, et al. Disintegration of metro and land development in transition China: A dynamic analysis in Beijing [J]. Transportation Research Part A: Policy and Practice, 2018, 116: 290-307.
- [3] Ibraeva A, de Almeida Correia, Silva C, et al. Transit-oriented development: A review of research achievements and challenges [J]. Transportation Research Part A: Policy and Practice, 2020, 132: 110-130.
- [4] 陈威, 焦洪赞, 王涛. 廊道式TOD土地利用遗传优化模型 [J]. 武汉大学学报(工学版), 2020, 53(11): 995-1005. [Chen Wei, Jiao Hongzan, Wang Tao. Corridor TOD land use genetic algorithm optimization model. Engineering Journal of Wuhan University, 2020, 53(11): 995-1005.]
- [5] Kuby M, Barranda A, Upchurch C. Factors influencing light-rail station boardings in the United States [J]. Transportation Research Part A: Policy and Practice, 2004, 38: 223-247.
- [6] Sung H, Choi K, Lee S, et al. Exploring the impacts of land use by service coverage and station-level accessibility on rail transit ridership [J]. Journal of Transport Geography, 2014, 36: 134-140.
- [7] 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.
- [8] 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. doi: 10.1016/j.jtrangeo.2020.102878.
- [9] 贺鹏, 李雯茜, 李妍, 等. 建成环境对轨道交通客流非线性影响的空间特征 [J]. 交通运输系统工程与信息, 2023, 23(3): 187-194. [He Peng, Li Wenxi, Li Yan, et al. Spatial patterns of nonlinear effects of built environment on Beijing subway ridership. Journal of Transportation Systems Engineering and Information Technology, 2023, 23(3): 187-194.]
- [10] Sohn K, Shim H. Factors generating boardings at metro stations in the Seoul metropolitan area [J]. Cities, 2010, 27(5): 358-368.
- [11] Choi J, Lee Y J, Kim T, et al. An analysis of metro ridership at the station-to-station level in Seoul [J]. Transportation, 2012, 39(3): 705-722.
- [12] Zhao J B, Deng W, Song Y, et al. Analysis of metro ridership at station level and station-to-station level in Nanjing: An approach based on direct demand models [J]. Transportation, 2014, 41(1): 133-155.
- [13] Iseki H, Liu C, Knaap G. The determinants of travel demand between rail stations: A direct transit demand mod-

- el using multilevel analysis for the Washington D.C. metrorail system [J]. *Transportation Research Part A: Policy and Practice*, 2018, 116: 635-649.
- [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] Gan Z X, Yang M, Feng T, et al. Examining the relationship between built environment and metro ridership at station-to-station level [J]. *Transportation Research Part D: Transport and Environment*, 2020, 82: 102332. doi: 10.1016/j.trd.2020.102332.
- [16] 彭诗尧, 陈绍宽, 许奇, 等. 基于POI的土地利用与轨道交通客流的空间特征 [J]. *地理学报*, 2021, 76(2): 459-470. [Peng Shiyao, Chen Shaokuan, Xu Qi, et al. Spatial characteristics of land use based on POI and urban rail transit passenger flow. *Acta Geographica Sinica*, 2021, 76(2): 459-470.]
- [17] 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. *Economic Geography*, 2021, 41(7): 61-72.]
- [18] 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. *Progress in Geography*, 2023, 42(3): 491-504.]
- [19] An D D, Tong X, Liu K, et al. Understanding the impact of built environment on metro ridership using open source in Shanghai [J]. *Cities*, 2019, 93: 177-187.
- [20] 中国国际工程咨询有限公司, 北京城建交通设计研究院有限公司. 2021年中国城市轨道交通沿线发展态势年度报告 [R/OL]. 2022-12-08 [2023-04-01]. <https://mp.weixin.qq.com/s/sJ5ZEoqJOoul40Qn-inROg>. [China International Engineering Consulting Corporation, Beijing Urban Construction Transport Planning & Design Institute Co. Limited. Annual report on the development trend of urban rail transit in China, 2021. 2022-12-08 [2023-04-01]. <https://mp.weixin.qq.com/s/sJ5ZEoqJOoul40Qn-inROg>.]
- [21] 国家市场监督管理总局, 国家标准化管理委员会. 城市轨道交通运营指标体系: GB/T 38374—2019 [S]. 北京: 中国标准出版社, 2019. [State Administration for Market Regulation, Standardization Administration of the People's Republic of China. Operational index system of urban rail transit: GB/T 38374—2019. Beijing, China: Standards Press of China, 2019.]
- [22] 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.
- [23] 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.
- [24] 贺鹏, 毛保华, 李妍, 等. 城市群交通与土地利用互动策略研究 [M]. 北京: 中国铁道出版社, 2023. [He Peng, Mao Baohua, Li Yan, et al. Research on interactive strategies of transportation and land use in urban agglomeration. Beijing, China: China Railway Publishing House, 2023.]
- [25] 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. *Journal of Geo-information Science*, 2021, 23(7): 1246-1258.]
- [26] 李明高, 毛保华, 蒋玉琨, 等. 城市轨道交通网络换乘便捷性研究 [J]. *中国铁道科学*, 2015, 36(3): 113-118. [Li Minggao, Mao Baohua, Jiang Yukun, et al. Transfer convenience of urban rail transit network. *China Railway Science*, 2015, 36(3): 113-118.]

Classification of urban rail transit lines and determinants of the ridership from the perspective of land use

ZHANG Xiaoyu¹, XU Qi^{1,2*}, SONG Liying², DENG Jin³

- (1. Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, Beijing Jiaotong University, Beijing 100044, China;
2. School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China;
3. Beijing Urban Construction Transport Planning & Design Institute Co. Limited, Beijing 100037, China)

Abstract: The comprehension of the interdependence between urban rail transit ridership and land use is beneficial for rational resource allocation to facilitate the implementation of transit-oriented development (TOD). Existing studies have primarily focused on ridership and the built environment of stations to identify key influencing factors, while there remains a lack of research from a global perspective regarding the dependence of ridership on land use along urban rail transit lines. To this end, this study used multi-source urban geographic data from 36 major cities in China's mainland in 2019 to establish line function and built environment factors that depict land use, and classify line types using the two as clustering features, while also establishing transportation service factors. XGBoost (eXtreme Gradient Boosting) was employed to analyze the influencing factors of ridership across different types of lines. Taking seven supercities such as Beijing, 14 megacities such as Wuhan, and 15 big cities such as Nanning and their urban rail transit systems as research objects, the results show that urban rail transit lines can be categorized into downtown line type I, downtown line type II, and suburb line, based on the different service scopes of the lines and variations in urban development levels. The cumulative contribution of the explanatory variables in the top five, which are of utmost importance to ridership, accounts for at least 60% across all three types of lines. Among these, both employment density and the coverage level of line on the downtown area emerged as significant land use explanatory variables shared by all three types of lines. Moreover, it was observed that important explanatory variables exhibited a threshold effect on ridership, with converging thresholds among important variables within the same type of line but notable difference between different types. Therefore, the construction of TOD corridor should align the distribution of employment and residential areas, as well as supporting resources, in accordance with local conditions. The research findings can help understanding the mechanism of land use affecting urban rail transit ridership and provide a reference for promoting the development of corridor TOD.

Keywords: urban traffic; land use; line-level ridership; machine learning modeling; threshold effects