

# 城市轨道交通网络客运强度影响因素研究

付铎<sup>1</sup>, 李瑞敏<sup>2</sup>, 戚扬<sup>1</sup>, 曹瑾鑫<sup>1,3</sup>

(1. 内蒙古大学交通学院, 内蒙古自治区 呼和浩特 010000; 2. 清华大学土木工程系, 北京 100084; 3. 内蒙古科学技术研究院, 内蒙古自治区 呼和浩特 010000)

**摘要:** 城市轨道交通网络客运强度的影响机理较为复杂, 从城市层面分析社会经济与城市轨道交通系统属性对其作用机制, 对于促进城市轨道交通系统的可持续发展具有重要意义。基于2023年中国大陆地区44个开通城市轨道交通城市的实际数据, 从城市轨道交通系统属性、城市轨道交通可达性、社会经济特征以及城市交通系统特征4个方面选取了10个变量, 综合运用主成分分析、极限梯度提升模型和SHAP解释器, 系统探究了城市轨道交通网络客运强度的关键影响因素及其作用机理。结果表明: 各影响因素对网络客运强度的作用存在显著阈值效应。其中, 运营规模、城市交通拥堵指数和换乘车站比例是主要正向影响因素; 平均站间距与网络客运强度呈负相关; 公共汽车线网密度、道路网密度等因素的影响机制较为复杂, 需结合具体城市情况进行分析。

**关键词:** 城市轨道交通; 网络客运强度; 影响因素; XGBoost模型; SHAP分析; 阈值效应

Factors Influencing Passenger Volume Intensity in Urban Rail Transit Networks

Fu Duo<sup>1</sup>, Li Ruimin<sup>2</sup>, Qi Yang<sup>1</sup>, Cao Jinxin<sup>1,3</sup>

(1. School of Transport, Inner Mongolia University, Hohhot Inner Mongolia Autonomous Region 010000, China; 2. School of Civil Engineering, Tsinghua University, Beijing 100084, China; 3. Inner Mongolia Academy of Science and Technology, Hohhot Inner Mongolia Autonomous Region 010000, China)

**Abstract:** The mechanisms influencing passenger volume intensity in urban rail transit networks are highly complex. Analyzing the effects of socioeconomic conditions and urban rail transit system attributes at the city level is significant for promoting the sustainable development of urban rail transit systems. Based on real data from 44 cities in Chinese mainland with operational urban rail transit systems in 2023, this study selects ten variables across four dimensions, including urban rail transit system attributes, rail transit accessibility, socioeconomic characteristics, and urban transportation system features. Principal component analysis, an Extreme Gradient Boosting (XGBoost) model, and the SHAP explainer are used to systematically investigate the key factors affecting passenger volume intensity in urban rail transit networks and their underlying mechanisms. The results indicate that the influencing factors exhibit pronounced threshold effects on network passenger volume intensity. Among them, operational scale, the urban traffic congestion index, and the proportion of transfer stations emerge as the primary positive influencing factors. Average station spacing shows a negative correlation with network passenger volume intensity. The influence mechanisms of factors, such as bus transit network density and road network density, are more complex and require analysis tailored to the specific conditions of individual cities.

**Keywords:** urban rail transit; network passenger volume intensity; influencing factors; XGBoost model; SHAP analysis; threshold effects

收稿日期: 2024-09-12

**基金项目:** 国家重点研发计划项目“交通基础设施复杂网络快速建模与智能仿真”(2021YFB2600502)、国家自然科学基金项目“不确定条件下的建筑固废资源化利用物流网络优化理论与方法研究”(71961024)、内蒙古自治区重点科技攻关计划资助项目“基于多源大数据的交通信息分析技术平台研发及产业化推广”(2019GG287)

**作者简介:** 付铎(1999—), 男, 内蒙古包头人, 硕士, 研究方向为交通运输规划与管理, 电子邮箱 tgfuduo@163.com。

**通信作者:** 曹瑾鑫(1982—), 男, 内蒙古呼和浩特人, 博士, 教授, 研究方向为智慧交通系统, 电子邮箱 imucjx@163.com。

引用格式：付铎，李瑞敏，戚扬，等. 城市轨道交通网络客运强度影响因素研究[J]. 城市交通，2026，24(1)：104-112.

Fu Duo, Li Ruimin, Qi Yang, et al. Factors influencing passenger volume intensity in urban rail transit networks[J]. Urban transport of China, 2026, 24(1): 104-112.

## 0 引言

随着城镇化进程的持续推进，以公共交通为导向的开发(Transit-Oriented Development, TOD)战略已在中国许多城市得到实施<sup>[1]</sup>。城市轨道交通逐渐成为超(特)大城市居民公共交通出行的主要方式。截至2024年底，全国城市轨道交通运营里程10 923.8 km<sup>[2]</sup>。然而在实际运营中，不同城市的轨道交通网络客运强度存在显著差异<sup>[3]</sup>。因此，从城市层面探究客运强度的影响因素，对城市轨道交通系统的规划与运营具有重要意义<sup>[4]</sup>。

在既有研究中，多数学者致力于分析不同城市轨道交通车站进出站客流量差异的成因，并对影响因素进行分类归纳。既有研究选取的解释变量可归纳为以下5类<sup>[1, 5-6]</sup>：1)土地利用特征，例如教育用地、办公用地等；2)交通系统特征，例如路网密度、拥堵程度等；3)社会经济特征，例如人均GDP、收入水平等；4)城市轨道交通系统属性，例如进站口与出站口数量等；5)可达性<sup>[7-9]</sup>。其所采用的因变量主要为日客流量、月客流量以及进站与出站客流量等。例如，有研究表明，城市轨道交通车站出入口数量、周边用地功能的复合程度以及公共汽车线网密度，与城市轨道交通车站客流量呈正相关关系<sup>[10-11]</sup>。

在研究方法上，早期研究多采用普通最小二乘法(Ordinary Least Squares, OLS)<sup>[12-14]</sup>、地理加权回归(Geographically Weighted Regression, GWR)<sup>[7, 15]</sup>、结构方程模型以及多尺度地理加权回归(Multiscale Geographically Weighted Regression, MGWR)<sup>[16]</sup>等方法，以探索城市轨道交通客运量的影响因素及其作用机制。然而，已有研究指出，这些方法通常仅能刻画变量间的线性关系，对非线性关系的识别能力有限<sup>[13]</sup>。随着机器学习方法的发展，深度学习<sup>[17]</sup>和随机森林等模型<sup>[18]</sup>逐渐被引入交通领域。研究表明，梯度提升回归树在拟合非线性关系方面表现较好，因而被应用于城市轨道交通客流量影响因素的分析中<sup>[8, 11, 19]</sup>；但该方法也存在变量间关系过拟合

的问题，可能影响结果的稳健性。近年来，随着极限梯度提升(eXtreme Gradient Boosting, XGBoost)模型<sup>[10, 20-22]</sup>的提出，学者们发现其不仅能有效识别变量间的非线性影响，也有助于降低过拟合带来的实验误差。

截至目前，针对城市轨道交通网络客运强度影响因素的探讨仍相对不足。鉴于各城市轨道交通网络的客运强度差异显著，且这种差异在很大程度上影响着城市轨道交通系统的运营压力与未来发展方向。为此，本文基于可获取的城市级数据选取变量，采用XGBoost模型与SHAP解释器(SHapley Additive exPlanations)构建分析方法，深入分析不同城市轨道交通网络客运强度差异的原因，量化各变量的影响程度，系统识别不同变量对网络客运强度的影响模式与作用强度，对中国城市轨道交通网络客运强度的差异性进行初步解释。

## 1 数据来源及影响因素变量选取

### 1.1 数据来源

本文以2023年中国大陆地区44个开通城市轨道交通的城市为研究对象，包括42个有地铁(含轻轨)城市、1个有市域快轨城市、1个有跨座式单轨城市，未纳入其他制式的城市轨道交通(如有轨电车)。研究所用城市级数据来源于《2022年城市建设统计年鉴》，城市轨道交通相关数据来源于中华人民共和国交通运输部发布的官方信息。

### 1.2 变量选取

为系统探究各城市轨道交通网络客运强度存在差异的成因，本文在变量选取时综合参考了既有研究与车站客流量、人口密度等相关的变量选取经验，同时兼顾城市轨道交通网络客运强度的潜在影响因素以及数据的可获取性及准确性。最终选取10个变量，并将其归纳为4类：城市轨道交通系统属性、城市轨道交通可达性、社会经济特征以及城市交通系统特征。各变量释义及描述性统计结果见表1，各变量对网络客运强度

的影响模式分析如下。

1) 运营线路数与运营里程。二者共同决定城市轨道交通网络的覆盖与服务范围,较长的运营里程可服务更广泛区域,而较多的线路数量则有助于更好地满足乘客的换乘需求<sup>[7]</sup>。

2) 平均站间距。该变量表示相邻两个城市轨道交通车站之间的平均距离,直接影响乘客步行至车站的意愿和可达性。

3) 换乘车站比例。该变量反映线网中换乘节点的密集程度,较高的换乘车站比例通常意味着网络连通性更强,跨线出行效率更高,从而可能提升城市轨道交通的出行吸引力。

4) 车站密度。指单位建成区面积内的车站数量,反映城市轨道交通服务的空间覆盖水平,直接影响居民出行的便捷程度。

5) 人口密度。该变量表征建成区内人口的集聚程度,较高的人口密度通常伴随着更强的出行需求。既有研究表明,该变量对城市轨道交通客运量具有正向影响<sup>[7,16]</sup>。

6) 人均GDP。作为衡量城市经济发展水平与居民收入的重要指标,该变量在城市轨道交通网络趋于复杂时对客运强度的影响更为显著<sup>[9]</sup>。

7) 公共汽车线网密度。该变量反映建成区内公共汽车线路的密集程度。由于城市

轨道交通车站与公共汽车站在规划布局上密切相关,本文将其作为公共交通接驳特征,以分析其对网络客运强度的影响。既有研究表明,该变量对车站级客运量具有正向影响<sup>[7]</sup>。

8) 道路网密度。该变量既反映道路的通达性,也能体现选择非城市轨道交通出行方式的便捷程度,同时还与城市轨道交通车站的可达性相关。研究表明,该变量与城市轨道交通车站客运量呈正向关系<sup>[8]</sup>。

9) 城市交通拥堵指数。该变量反映城市道路运行状况,较高的拥堵程度会降低私人小汽车出行的可靠性,从而促使部分乘客转向准时性更优的城市轨道交通,有助于提升城市轨道交通网络客运强度。

## 2 研究方法

### 2.1 主成分分析

本文研究发现,城市轨道交通系统的运营线路数与运营里程之间存在较强的相关性。为避免由此引发的多重共线性问题,采用主成分分析方法(Principal Component Analysis, PCA)对这两个变量进行处理。PCA是一种无监督的机器学习算法,主要用于数据降维。通过该方法,可将原有高度相关的变量转化为一组彼此正交的主成分,从

表1 城市轨道交通网络客运强度影响因素变量统计

Tab.1 Statistics on variables of influencing factors for passenger volume intensity in urban rail transit networks

变量类别	变量名称	变量释义	平均值	标准差	最小值	最大值
因变量	网络客运强度/ (万人次·km <sup>-1</sup> ·d <sup>-1</sup> )	年均日客运量与运营里程的比值	0.610	0.330	0.038	1.36
	运营线路数/条	运营线路数量	6.640	5.740	1.000	27.00
城市轨道交通系统属性	运营里程/km	运营线路总长度	227.090	214.010	15.100	825.00
	平均站间距/km	相邻两个车站之间的平均距离	1.560	0.440	0.630	3.25
城市轨道交通可达性	换乘车站比例/ (个·km <sup>-1</sup> )	换乘车站数量与运营里程的比值	0.062	0.039	0	0.15
	车站密度/ (个·km <sup>-2</sup> )	车站数量与建成区面积的比值	0.210	0.098	0.013	0.49
自变量 社会经济特征	人口密度/ (万人·km <sup>-2</sup> )	建成区总人口与建成区面积的比值	1.000	0.260	0.630	1.99
	人均GDP/万元	城市居民的人均地区生产总值	12.070	3.920	5.840	19.83
城市交通系统特征	公共汽车线网密度/ (条·km <sup>-2</sup> )	公共汽车线路数量与建成区面积的比值	0.960	0.580	0.290	3.90
	道路网密度/ (km·km <sup>-2</sup> )	建成区道路网络总长度与建成区面积的比值	6.410	1.290	3.600	9.80
	城市交通拥堵指数	基于浮动车数据计算的城市道路交通拥堵指数	1.610	0.200	1.180	2.09

而消除特征间的强相关性，减少其对后续模型分析造成的干扰。PCA的基本思想是将原始的  $n$  维特征通过线性变换映射到  $k$  维空间 ( $k < n$ )，该  $k$  维空间中的特征即为全新的正交特征，也称为主成分，它们是在原始  $n$  维特征基础上重构得到的低维表征，并能够保留原始数据的主要信息<sup>[23]</sup>。PCA的计算步骤如下：

假设有  $m$  个  $n$  维数据样本，其中每个样本  $X_i$  为一个  $n$  维列向量，构成数据矩阵：

$$X_{n \times m} = (X_1, X_2, \dots, X_m). \quad (1)$$

首先对数据进行去中心化处理，即令

$$X \leftarrow X - \frac{1}{m} \sum_{i=1}^m X_i, \quad (2)$$

计算其协方差矩阵：

$$C = \frac{1}{m} X X^T. \quad (3)$$

对协方差矩阵  $C$  进行特征值分解，得到特征值及对应的特征向量。将特征向量按特征值从大到小排列，取前  $k$  个特征向量组成投影矩阵  $P_{n \times k}$ ，其中  $P$  的每一列代表一个主成分方向。将原始数据投影到  $P$  坐标系下，即可获得降维后的数据表示：

$$Y_{k \times m} = P_{n \times k}^T X_{n \times m}. \quad (4)$$

本文运用PCA将运营线路数与运营里程这两个高度相关的变量整合为一个综合特征，称为“运营规模”。该特征能够保留原始变量的主要信息，同时有效降低共线性对后续建模分析的影响。

## 2.2 XGBoost模型

在分析城市轨道交通车站客流量差异的研究中，部分学者指出传统机器学习模型在处理该类问题时易出现过拟合现象。XGBoost模型由Chen Tianqi等<sup>[20]</sup>于2016年提出，是一种基于梯度提升决策树的改进算法。该模型在控制过拟合的同时，具有较高的预测精度，并能够有效处理缺失值与异常值<sup>[20, 22]</sup>。XGBoost回归模型的目标函数  $T(\varphi)$  通常由训练损失函数与正则化项两部分构成，可表示为：

$$T(\varphi) = L(\varphi) + \Omega(\varphi), \quad (5)$$

式中： $L(\varphi)$  为训练损失函数，用于衡量模型在训练数据上的拟合效果； $\Omega(\varphi)$  为正则化项，用于控制模型的复杂度和过拟合程度。

本文将数据集按8:2划分为训练集与测试集。在模型训练过程中，超参数的选取以追求最大拟合优度为目标，采用网格搜索法进行优化确定，并运用  $K$  折交叉验证评估

模型的泛化能力。XGBoost模型的具体参数配置见表2。

## 2.3 SHAP分析

作为一种基于集成学习的机器学习模型，XGBoost在处理非线性关系时虽能显著提升预测精度，但其模型结构的复杂性也导致了可解释性的降低。在相关研究中，特征重要性分析常被用于模型解释。然而，该方法仅能评估各变量对模型整体预测能力的贡献程度，而无法具体揭示不同特征与模型预测结果之间的内在关联。SHAP是由S. M. Lundberg和S. I. Lee于2017年提出的一种机器学习模型解释方法<sup>[24]</sup>，其理论源于合作博弈论中的Shapley值，通过将模型的预测结果分解为各特征在不同样本上的局部贡献，为机器学习模型提供从全局到局部层面的解释支持。该方法的核心优势在于能够量化各个特征对单一样本预测结果的影响大小与方向。SHAP值的计算公式为<sup>[24]</sup>：

$$\theta_i = \sum_{S \in \mathcal{N} \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f_{S \cup \{i\}}(X_{S \cup \{i\}}) - f_S(X_S)], \quad (6)$$

式中： $\theta_i$  为特征  $i$  的SHAP值； $i$  为当前所考察的特征； $F$  为全部特征的集合； $S$  为  $F$  中除去特征  $i$  后的任意特征子集； $|S|!$  为子集  $S$  中特征数量的阶乘； $f_{S \cup \{i\}}$  为使用子集  $S \cup \{i\}$  训练的模型； $X_{S \cup \{i\}}$  为特征子集  $S$  与当前考察特征  $i$  合并后形成的特征集合所对应的输入特征取值； $f_S$  为仅使用特征子集  $S$  训练的模型； $X_S$  为子集  $S$  对应的特征取值； $[f_{S \cup \{i\}}(X_{S \cup \{i\}}) - f_S(X_S)]$  为两个模型在对应输入下的输出差值。

表2 XGBoost模型参数配置

Tab.2 Parameters configuration of the XGBoost model

参数	释义	取值
树的数量	指定模型中决策树的总数，叠加集成多棵树的预测结果提升模型性能	100.00
学习率	控制每棵树的贡献权重，较低的学习率通常需要更多树来提升模型性能，但有助于提升模型稳定性	0.30
最大树深度	限制单棵树决策的最大深度，用于控制模型复杂度。较浅的树可防止过拟合，但过浅可能导致欠拟合	4.00
随机采样比例	训练每棵树时使用的样本比例，引入随机性以增强模型泛化能力	0.70
列采样率	训练每棵树时使用的特征比例，通过特征采样进一步降低过拟合风险	0.70
叶节点分裂阈值	控制叶节点分裂所需的最小损失降低值，值越大模型越保守，有助于防止过拟合	0.01

### 3 影响因素机制分析

#### 3.1 XGBoost模型拟合结果

XGBoost模型的拟合优度 $R^2$ 为0.86,表明模型能够解释约86%的网络客运强度变异,反映出该模型对城市轨道交通网络客运强度的影响因素具有良好的刻画能力。在误差评估方面,均方误差(Mean Squared Error, MSE)为0.01,平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)为0.28,说明模型预测偏差较小,进一步验证了模型的精度与可靠性。

为验证模型的泛化能力,本文采用5折交叉验证方法。结果显示,模型在不同折子集上的平均预测误差为0.05,标准差为0.02,表明模型在不同数据子集上表现稳定,具有良好的泛化能力。

#### 3.2 影响因素重要性分析

基于XGBoost模型,通过计算各变量在所有样本上的SHAP值的平均绝对值,得到城市轨道交通网络客运强度影响因素的贡献度排序(见图1)。各因素按重要性从高到低依次为:运营规模(由运营线路数和运营里程合成)、城市交通拥堵指数、换乘车站比例、平均站间距、人口密度、公共汽车线网密度、人均GDP、道路网密度和车站密度。

运营规模对城市轨道交通网络客运强度的影响最为显著,体现了系统自身的网络规模效应。尽管运营规模并非决定网络客运强度的唯一因素,但总体趋势表明,较长的运营里程与较多的线路数量共同扩展了服务范围,提升了城市轨道交通的可达性,进而促

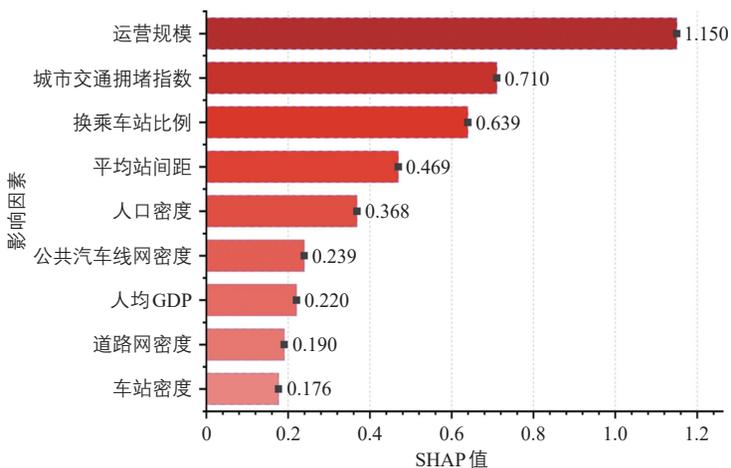


图1 城市轨道交通网络客运强度影响因素的贡献度排序

Fig.1 Ranking of contribution levels of factors influencing passenger volume intensity in urban rail transit networks

进客运强度的增长。城市交通拥堵指数的重要性仅次于运营规模,在北京、上海、广州等高拥堵城市,城市轨道交通以其大容量与高准时性成为部分通勤者与日常出行者的首选交通方式。

影响因素贡献度的排序客观反映了各变量对网络客运强度的相对重要性,对城市轨道交通系统的规划与建设具有参考价值。同时,各变量对网络客运强度的影响机制总体较为复杂,可能存在尚未完全揭示的非线性关系与变量间的交互效应。

#### 3.3 影响因素相关性分析

##### 3.3.1 单一特征分析

相关性检验结果(见图2)显示,除平均站间距与网络客运强度呈负相关外,其余变量均呈正相关。其中,相关性较强的典型影响因素如下:

##### 1) 运营规模。

城市轨道交通系统属性与网络客运强度之间的正相关关系表明,扩大运营规模对提升客运强度具有显著作用。从数据分布来看,在运营规模较小的城市中,提高运营规模对提升网络客运强度的促进作用尤为明显,体现了城市轨道交通的网络规模效应。

值得注意的是,南京、天津、郑州、青岛等城市的城市轨道交通运营里程均超过300 km,但其网络客运强度却相对较低。如表3所示,这4个城市在运营线路数和运营里程的排名均靠前(约前10名),而网络客运强度的排名却处于18~28名。以南京市和青岛市为例,二者均表现出换乘车站比例较低、平均站间距较大的共同特征。这表明,网络客运强度不仅取决于运营规模,还受到平均站间距、换乘便利性等多重因素的共同影响。

##### 2) 城市交通拥堵指数。

城市交通拥堵指数与网络客运强度的正相关关系表明,随着经济发展和居民对出行时效性要求的提高,较高的拥堵指数反映出小汽车出行效率下降、道路通行困难。此时,城市轨道交通系统凭借快速与准时的特性,能够有效规避道路交通拥堵问题。这一现象在人口密集、通勤压力突出的超(特)大城市(如北京、上海、广州、深圳等)中尤为明显,当地居民选择出行方式时更倾向于城市轨道交通,从而推动其网络客运强度的提升。

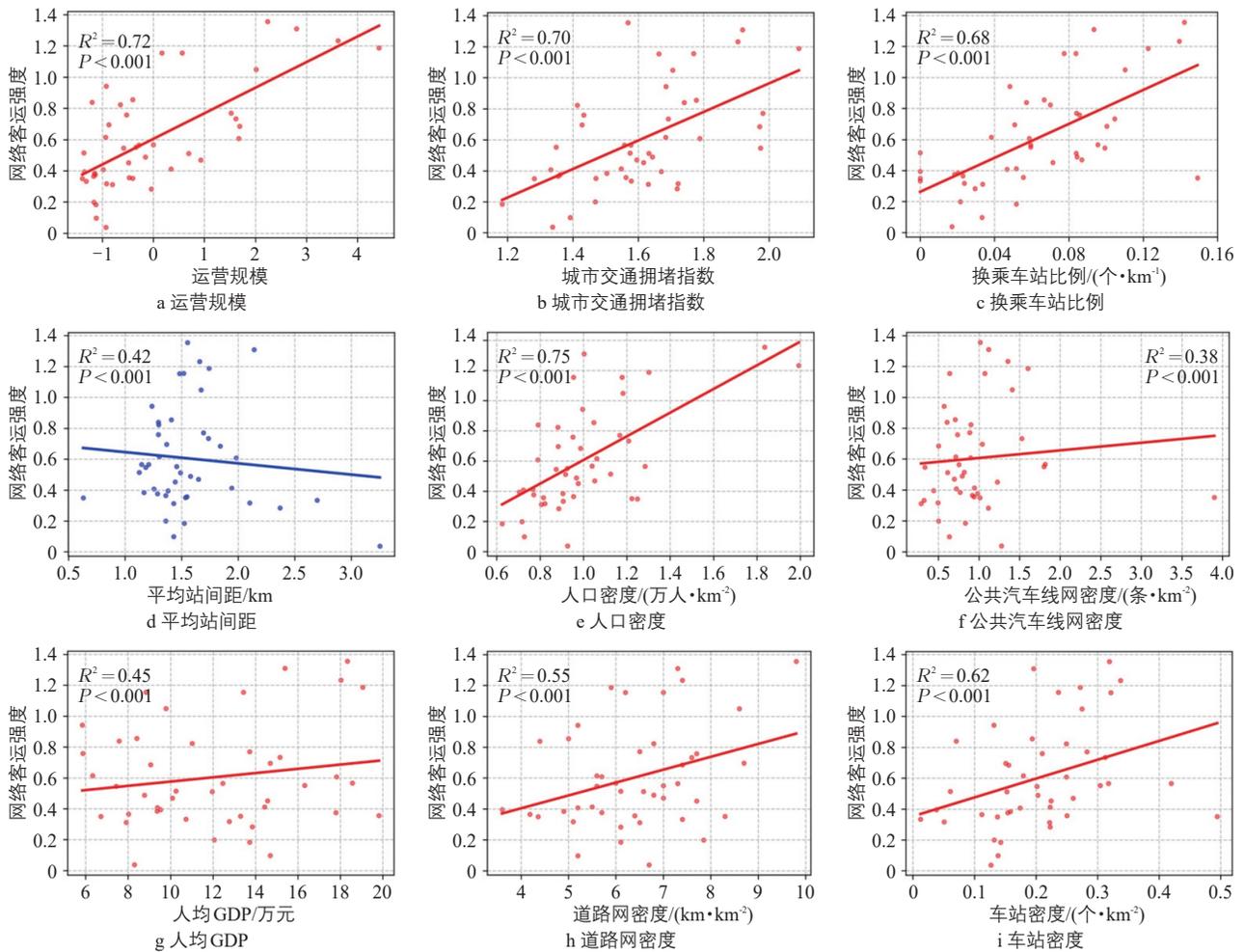


图2 城市轨道交通网络客运强度影响因素相关性检验结果

Fig.2 Results of correlation tests on factors influencing passenger volume intensity in urban rail transit networks

### 3) 换乘车站比例。

换乘车站比例与网络客运强度的正相关性表明，在城市轨道交通线网规模一定的条件下，提高换乘车站比例有助于增强网络连通性，对提升网络客运强度具有关键作用。从总体趋势看，线网规模较小的城市往往换乘车站的比例较低，这限制了网络客运强度的提升，进而从连通性层面体现了轨道交通的网络效应。

### 4) 平均站间距。

平均站间距与网络客运强度的负相关关系表明，车站间距增大会降低车站 500 m 或 800 m 步行可达范围的覆盖比例，导致更多乘客面临“最后一公里”接驳不便，从而削弱城市轨道交通对客流的吸引力。

### 5) 公共汽车线网密度。

公共汽车线网密度与网络客运强度的正相关性主要源于两个机制：一方面，较高的公共汽车线网密度通常反映城市总体客运需求较大，从而间接为城市轨道交通提供了广泛的客流基础；另一方面，公共汽车线网密

表3 典型城市轨道交通网络客运强度影响因素排名

Tab.3 Ranking of key influencing factors for passenger volume intensity in typical urban rail transit networks

城市	网络客运强度排名	影响因素排名				
		运营线路数	运营里程	城市交通拥堵指数	换乘车站比例	平均站间距
南京	18	6	9	7	23	6
天津	24	11	12	20	15	21
郑州	26	10	10	23	11	14
青岛	28	14	11	30	27	7

度较高有助于改善与城市轨道交通的接驳条件，通过提升“最后一公里”可达性，进一步促进城市轨道交通网络客运强度的提高。

### 3.3.2 阈值效应分析

图3展示了网络客运强度各影响因素SHAP值的散点图分布，每个散点代表一个城市。分析结果显示，当城市交通拥堵指数较高时，其对网络客运强度表现为显著的正向影响，表明此时乘客更倾向于选择城市轨

道交通出行，从而提升网络客运强度；而当拥堵指数较低时，其影响转为负向，说明乘客可能更偏好其他方式出行，导致网络客运强度相应降低。这一现象表明，部分影响因素存在明显的阈值效应。

图4通过SHAP值在各影响因素中的分布进一步验证了阈值效应的存在。具体来

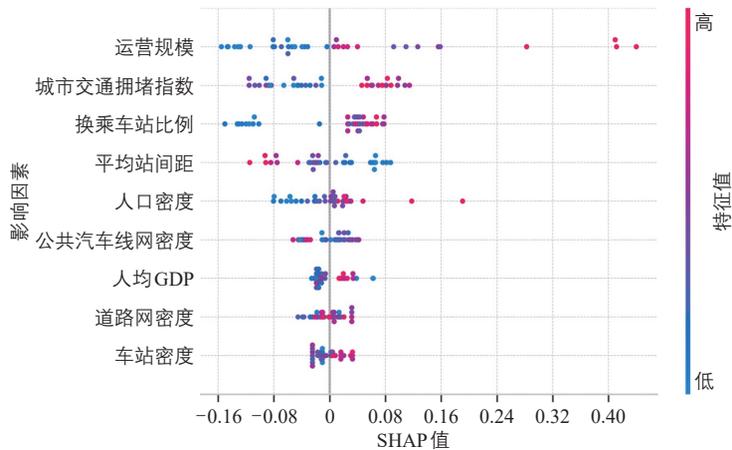


图3 城市轨道交通网络客运强度影响因素SHAP值

Fig.3 SHAP values of factors influencing passenger volume intensity in urban rail transit networks

看：1)城市交通拥堵指数的阈值为1.65，超过此值后其对网络客运强度产生显著正向影响，表明严重拥堵会促使乘客转向城市轨道交通出行；2)换乘车站比例达到0.04个·km<sup>-1</sup>时转为正向影响，反映了网络连通性对提升出行效率的关键作用；3)平均站间距的阈值为1.4 km，超过此值会因“最后一公里”可达性下降而降低网络客运强度；4)当建成区人口密度高于0.9万人·km<sup>-2</sup>时，对网络客运强度有显著促进作用，这体现了客流基础对系统利用率的支撑作用；5)车站密度通过提高空间覆盖率直接提升了服务便捷性；6)公共汽车线网密度呈现双重影响，多数情况下通过接驳服务产生协同效应，少数情形下因竞争关系表现为负相关；7)人均GDP的影响相对有限，其与机动化水平的复杂交互可能在一定程度上抑制城市轨道交通需求，与既有研究结论一致<sup>[25-26]</sup>。

基于上述阈值分析，建议各影响因素处于阈值以上的城市重点优化轨道交通的网络结构与换乘效率，以提升网络客运强度；而

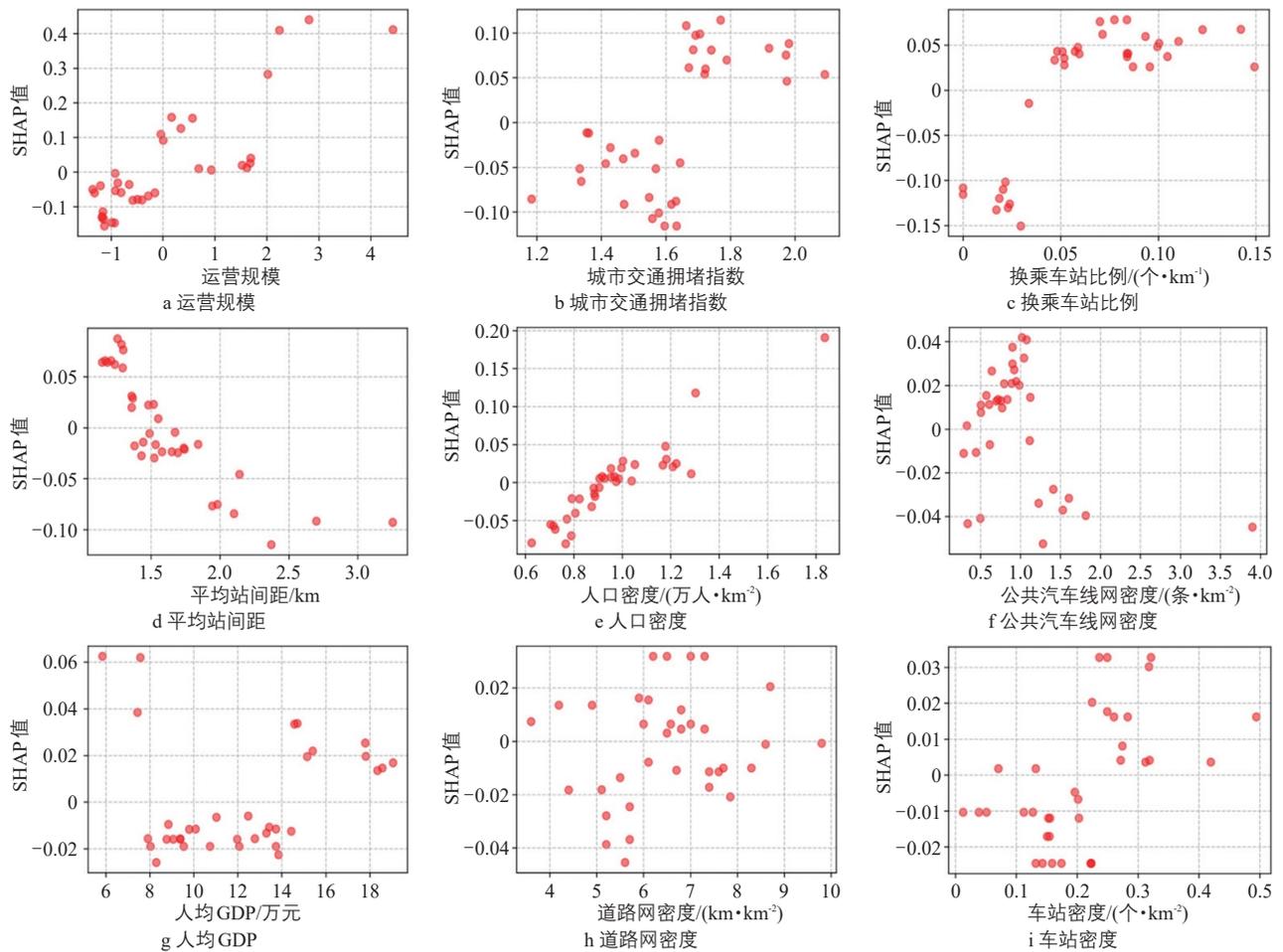


图4 城市轨道交通网络客运强度影响因素与SHAP值散点图

Fig.4 Scatter plot of influencing factors and SHAP values for passenger volume intensity in urban rail transit networks

影响因素处于阈值以下的城市，则应综合研判各因素现状，通过规划引导使关键影响因素逐步接近或超过阈值水平，从而提升城市轨道交通系统的吸引力与使用效率。在未来的城市轨道交通规划中，应统筹考虑人口分布、平均站间距、换乘车站比例等社会经济与系统属性因素的阈值综合影响，科学论证城市轨道交通建设的必要性与发展路径。

## 4 结论

本文以2023年中国大陆地区44个开通城市轨道交通的城市为研究对象，采用XGBoost模型和SHAP分析方法系统探讨了城市轨道交通网络客运强度的影响因素及其作用关系。研究结果表明：

1) 平均站间距与城市轨道交通网络客运强度呈负相关；而运营规模、城市交通拥堵指数等系统属性类影响因素均对其表现出正向影响。

2) 各影响因素对网络客运强度存在明显的阈值效应。这反映出当前中国不同城市在社会经济和交通设施建设上存在不均衡，这一不均衡性直接导致多数城市的轨道交通系统尚未充分发挥其通行能力。

3) 在部分超(特)大城市(如北京、上海、深圳和广州等)中，除平均站间距外，提升其他影响因素均可有效提升网络客运强度。

总体而言，运营规模、城市交通拥堵指数、换乘车站比例和平均站间距是影响城市轨道交通网络客运强度的主要因素，而公共汽车线网密度、道路网密度等因素的作用较为复杂，需结合具体城市情况进行分析。

本文研究结果可为处于不同发展阶段的城市轨道交通规划提供参考。建议在未来的规划和建设中，重点关注关键影响因素，并着力使其达到或超过本研究识别出的阈值水平。

本研究也存在局限性，主要是受数据所限，未能考虑非机动车保有率、平均通勤距离等潜在影响因素。未来研究可引入更多隐含变量，以降低遗漏变量导致估计偏误的风险。随着数据采集技术的不断完善与多源数据的融合，可进一步采用大数据驱动的方法，持续识别影响城市轨道交通网络客运强度的关键因素，从而为中国城市轨道交通系统的可持续发展提供更科学的支撑。

参考文献：

References:

- [1] Yu Zidong, Zhu Xiaolin, Liu Xintao. Characterizing metro stations via urban function: thematic evidence from Transit-Oriented Development (TOD) in Hong Kong[J]. *Journal of transport geography*, 2022, 99: 103299.
- [2] 中华人民共和国交通运输部. 2024年交通运输行业发展统计公报[R/OL]. (2025-06-12) [2025-07-05]. [https://xxgk.mot.gov.cn/2020/jigou/zhghs/202506/t20250610\\_4170228.html](https://xxgk.mot.gov.cn/2020/jigou/zhghs/202506/t20250610_4170228.html).
- [3] 王玉萍, 马超群. 城市轨道交通客运量影响因素与成长规律[J]. *长安大学学报(自然科学版)*, 2013, 33(3): 69-75.  
Wang Yuping, Ma Chaoqun. Influencing factors and development trends of urban rail transit passenger flow[J]. *Journal of Chang'an University (natural science edition)*, 2013, 33(3): 69-75.
- [4] 刘成帅, 陈鹏. 北京市城市轨道交通出行特征及其影响因素分析[J]. *科学技术创新*, 2024(3): 92-95.  
Liu Chengshuai, Chen Peng. Analysis of subway travel characteristics and influencing factors in Beijing[J]. *Scientific and technological innovation*, 2024(3): 92-95.
- [5] 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.
- [6] Fu Xin, Zhao Xiaoxuan, Li Cengceng, et al. Exploration of the spatiotemporal heterogeneity of metro ridership prompted by built environment: a multi-source fusion perspective[J]. *IET intelligent transport systems*, 2022, 16(11): 1455-1470.
- [7] Gao Fan, Yang Linchuan, Han Chunyang, et al. A network-distance-based geographically weighted regression model to examine spatio-temporal effects of station-level built environments on metro ridership[J]. *Journal of transport geography*, 2022, 105: 103472.
- [8] Li Linchao, Zhong Liangjian, Ran Bin, 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*, 2024, 144: 105564.
- [9] Xin Mengwei, Shalaby A. Investigation of the

- interaction between urban rail ridership and network topology characteristics using temporal lagged and reciprocal effects: a case study of Chengdu, China[J]. *Transportation research part A: policy and practice*, 2024, 179: 103922.
- [10] Wang Zhenbao, Li Shihao, Li Yongjin, et al. Investigating the nonlinear effect of built environment factors on metro station-level ridership under optimal pedestrian catchment areas via the machine learning method[J]. *Applied sciences*, 2023, 13(22): 12210.
- [11] Du Qiang, Zhou Yuqing, Huang Youdan, et al. Spatiotemporal exploration of the non-linear impacts of accessibility on metro ridership[J]. *Journal of transport geography*, 2022, 102: 103380.
- [12] Sohn K, Shim H. Factors generating boardings at Metro stations in the Seoul metropolitan area[J]. *Cities*, 2010, 27(5): 358–368.
- [13] Zhao Jinbao, Deng Wei, Song Yan, et al. What influences metro station ridership in China? insights from Nanjing[J]. *Cities*, 2013, 35: 114–124.
- [14] Zhao Jinbao, Deng Wei, Song Yan, 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.
- [15] Yang Hongtai, Xu Taorang, Chen Dexin, et al. Direct modeling of subway ridership at the station level: a study based on mixed geographically weighted regression[J]. *Canadian journal of civil engineering*, 2020, 47(5): 534–545.
- [16] Wang Zhenbao, Song Jiarui, Zhang Yuchen, et al. Spatial heterogeneity analysis for influencing factors of outbound ridership of subway stations considering the optimal scale range of "7D" built environments[J]. *Sustainability*, 2022, 14(23): 16314.
- [17] Su Shiliang, Wang Zhuolun, Li Bozhao, et al. Deciphering the influence of TOD on metro ridership: an integrated approach of extended node-place model and interpretable machine learning with planning implications[J]. *Journal of transport geography*, 2022, 104: 103455.
- [18] Yang Linchuan, Yu Bingjie, Liang Yuan, et al. Time-varying and non-linear associations between metro ridership and the built environment[J]. *Tunnelling and underground space technology*, 2023, 132: 104931.
- [19] Liu Mengyang, Liu Yuxuan, Ye Yu. Nonlinear effects of built environment features on metro ridership: an integrated exploration with machine learning considering spatial heterogeneity[J]. *Sustainable cities and society*, 2023, 95: 104613.
- [20] Chen Tianqi, Guestrin C. XGBoost: a scalable tree boosting system[C]// Association for Computing Machinery. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, CA, USA: Association for Computing Machinery, 2016: 785–794.
- [21] Liu Hao, Chen Xi, Liu Xiaoxiao. Factors influencing secondary school students' reading literacy: an analysis based on XGBoost and SHAP methods[J]. *Frontiers in psychology*, 2022, 13: 948612.
- [22] Kim S, Lee S. Nonlinear relationships and interaction effects of an urban environment on crime incidence: application of urban big data and an interpretable machine learning method[J]. *Sustainable cities and society*, 2023, 91: 104419.
- [23] Zhang Kenan, Cao Jianhong, Yin Haining, et al. Microclimate diversity drives grape quality difference at high-altitude: observation using PCA analysis and Structural Equation Modeling (SEM) [J]. *Food research international*, 2024, 191: 114644.
- [24] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[C]// Association for Computing Machinery. *Proceedings of the 31st international conference on neural information processing systems*. Long Beach, USA: Association for Computing Machinery, 2017: 4768–4777.
- [25] Ingvardson J B, Nielsen O A. How urban density, network topology and socio-economy influence public transport ridership: empirical evidence from 48 European metropolitan areas[J]. *Journal of transport geography*, 2018, 72: 50–63.
- [26] Raza A, Zhong Ming, Akuh R, et al. Public transport equity with the concept of time-dependent accessibility using Geostatistics methods, Lorenz curves, and Gini coefficients[J]. *Case studies on transport policy*, 2023, 11: 100956.