

上海市不同出行方式通勤效率数据集（2015）研发

岳丽莹¹, 朱宇^{1*}, 李开明^{2,3}

1. 上海大学人口研究所暨亚洲人口研究中心, 上海 200444; 2. 上海大学上海美术学院建筑系, 上海 200444;
3. 自然资源部大都市区国土空间生态修复工程技术创新中心, 上海 200003

摘要: 基于过剩通勤框架, 作者开展不同出行方式通勤效率的比较研究, 其结果有助于加深对城市通勤与土地利用、交通系统之间关系的理解, 为建设绿色低碳城市以及实现城市可持续发展提供数据支持。利用1%人口抽样调查数据, 本文分析了上海市不同学历群体的三种出行方式(非机动、公共交通和小汽车)通勤效率, 得到上海不同出行方式的通勤效率数据集(2015)。数据集内容包括2015年上海市以下数据:(1)各街道通勤量;(2)不同学历群体的通勤出行方式及占比;(3)不同学历群体不同出行方式的通勤效率;(4)不同学历群体不同出行方式的实际通勤流数据。该数据集存储为.xlsx格式, 1个数据文件, 数据量为15.5 MB。

关键词: 通勤效率; 过剩通勤; 职住平衡; 出行方式; 上海; 2015

DOI: <https://doi.org/10.3974/geodp.2024.04.08>

CSTR: <https://cstr.escience.org.cn/CSTR:20146.14.2024.04.08>

数据可用性声明:

本文关联实体数据集已在《全球变化数据仓储电子杂志(中英文)》出版, 可获取:

<https://doi.org/10.3974/geodb.2024.09.05.V1> 或 <https://cstr.escience.org.cn/CSTR:20146.11.2024.09.05.V1>.

1 前言

通勤是城市生活的重要组成部分, 其强度可以看作是城市经济活力的标志^[1,2]。但是, 当其规模和距离达到临界水平时, 城市可能会面临严重的交通拥堵、资源消耗、环境污染等城市管理难题, 深刻影响城市的可持续发展^[3,4]。城市土地利用决定居民的通勤行为, 优化居住-就业空间布局被认为是解决城市通勤问题的根本途径^[5,6]。2014年, 国务院出台的《国家新型城镇化规划(2014–2020年)》将“产城融合”作为破解快速城镇化阶段出现的空城、睡城、鬼城等问题的重要举措。近年来, 上海、北京、深圳、成都等逐渐在城市发展规划中纳入了“职住平衡”理念, 旨在提高城市运行效率。

过剩通勤是评估城市通勤效率的重要研究范式^[7]。然而, 已有研究主要聚焦于出行成本单一视角, 较少关注通勤系统的空间组织方面, 尚未从出行成本和空间组织双视角对通

收稿日期: 2024-08-26; 修订日期: 2024-11-30; 出版日期: 2024-12-24

基金项目: 中华人民共和国教育部(23YJCZH287, 23YJC790169); 自然资源部(CXZX202412)

*通讯作者: 朱宇, 上海大学人口研究所暨亚洲人口研究中心, zhu300@shu.edu.cn

数据引用方式: [1] 岳丽莹, 朱宇, 李开明. 上海市不同出行方式通勤效率数据集(2015)研发[J]. 全球变化数据学报, 2024, 8(4): 414–421. <https://doi.org/10.3974/geodp.2024.04.08>. <https://cstr.escience.org.cn/CSTR:20146.14.2024.04.08>.

[2] 岳丽莹, 朱宇, 李开明. 上海市不同出行方式的通勤效率数据集(2015) [J/DB/OL]. 全球变化数据仓储电子杂志, 2024. <https://doi.org/10.3974/geodb.2024.09.05.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2024.09.05.V1>.

勤效率进行分析。仅从出行成本的视角评估城市通勤效率，不利于建立对城市土地利用与通勤行为关系的全面认知。此外，已有研究对出行方式与居民社会经济属性在通勤行为中的交互作用考虑仍显不足。基于此，本文利用上海市 1%人口抽样调查数据，在过剩通勤框架下，对不同出行方式-学历水平的通勤效率进行分析，量化了出行成本和空间组织视角下通勤效率的非线性关系，得到了《上海市不同出行方式的通勤效率数据集（2015）》。

2 数据集元数据简介

《上海市不同出行方式的通勤效率数据集（2015）》^[8]的名称、作者、地理区域、数据年代、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表 1。

表 1 《上海市不同出行方式的通勤效率数据集（2015）》元数据简表

条 目	描 述
数据集名称	上海市不同出行方式的通勤效率数据集（2015）
数据集短名	CommutingEfficiencyShanghai2015
作者信息	岳丽莹，上海大学人口研究所暨亚洲人口研究中心，liyong128@shu.edu.cn 朱宇，上海大学人口研究所暨亚洲人口研究中心，zhu300@shu.edu.cn 李开明，上海大学上海美术学院建筑系，kaiming1239@shu.edu.cn
地理区域	上海市都市区（不包括崇明区）
数据年代	2015
数据格式	.xlsx
数据量	15.5 MB
数据集组成	各街道通勤量、不同学历群体的通勤出行方式及占比、不同出行方式的通勤效率以及不同出行方式的实际通勤流数据
基金项目	中华人民共和国教育部（23YJCZH287）
数据计算环境	Matlab
出版与共享服务平台	全球变化科学研究数据出版系统 http://www.geodoi.ac.cn
地址	北京市朝阳区大屯路甲 11 号 100101，中国科学院地理科学与资源研究所
数据共享政策	（1）“数据”以最便利的方式通过互联网系统免费向全社会开放，用户免费浏览、免费下载；（2）最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源；（3）增值服务用户或以任何形式散发和传播（包括通过计算机服务器）“数据”的用户需要与《全球变化数据学报（中英文）》编辑部签署书面协议，获得许可；（4）摘取“数据”中的部分记录创作新数据的作者需要遵循 10%引用原则，即从本数据集中摘取的数据记录少于新数据集总记录量的 10%，同时需要对摘取的数据记录标注数据来源 ^[9]
数据和论文检索系统	DOI, CSTR, Crossref, DCI, CSCD, CNKI, SciEngine, WDS, GEOSS, PubScholar, CKRSC

3 数据研发方法

3.1 数据来源

本文利用上海市统计局提供的 2015 年 1%人口抽样调查数据^[10]搭建通勤行为矩阵。原始数据是不等比例抽样的样本数据，均未作总体推算。本文提供的数据均是按照抽样比例推算还原后的结果。关于居民社会经济属性，本研究将学历水平划分为以下三种类型：低学历

(小学及以下)、中学历(初中和高中)及高学历(大学及以上);将出行方式划分为以下三种类型:非机动出行(步行和骑行)、公共交通(公共汽车和轨道交通)及小汽车(摩托车和小轿车)。参照已有研究方法^[11],通勤成本数据由基于 ArcGIS 网络分析的路网距离表示。

上海市包括黄浦区、徐汇区、长宁区、静安区、普陀区、闸北区、虹口区、杨浦区、闵行区、宝山区、嘉定区、浦东新区、金山区、松江区、青浦区、奉贤区、崇明区在内的 17 个区,总面积约 6,340 km²。考虑到地理空间上的连贯性,本研究将上海都市区(不包括崇明区)作为研究范围,并以街道为空间分析单元。因此,数据包含上海 196 个街道的通勤规模、通勤方式、学历水平及通勤流矩阵数据。

3.2 算法原理

在既有城市职住空间结构下,实际通勤与理论最小通勤之间的差距被称为过剩通勤^[12](Excess Commuting),其起源于经济学家汉密尔顿在 1982 年所提出的浪费通勤(Wasteful Commuting)概念^[13]。基于通勤图谱理论(Commuting Spectrum Theory),本文将理论最小通勤作为通勤基准值的下限,将随机通勤作为通勤基准值的上限^[14]。理论最小通勤是就业人口与最邻近就业的平均最短距离,对微观局部空间范围内的居住与就业配置比较敏感,常被用作一种规划工具来反映微观层面的职住平衡^[15]。较低的理论最小通勤表明就业与住房更加混合,二者的空间配置更加平衡,反之亦然。随机通勤是居民不考虑出行距离时(即距离摩擦系数为 0)的通勤模式,反映城市层面的职住不平衡。

关于通勤基准值的计算,本文应用 Matlab 软件,利用线性规划最优化模型(Transportation Problem in Linear Programming)计算理论最小通勤(C_{min})及其通勤矩阵。然后,将理论最小通勤矩阵作为双约束空间相互作用模型(Doubly Constrained Spatial Interaction Model)的数据基础,基于 Newton-Raphson 迭代算法估计理论最小通勤模式的距离摩擦系数(β),并得到理论最小通勤熵(H_{min})。根据给定的就业人口和就业岗位量、实际通勤流矩阵和通勤距离矩阵,可以计算实际平均通勤距离(C_{obs})和实际通勤模式的熵值(H_{obs})。此外,将距离衰减系数(β)设为 0 时所计算的结果即为随机通勤模式的平均通勤距离(C_{ran})和熵值(H_{ran})。基于表 2 中的过剩通勤指标,测度不同出行方式的通勤效率。

4 数据结果

4.1 数据集组成

数据集共包括 2015 年上海市以下四方面数据:(1)各街道通勤量;(2)不同学历群体的通勤出行方式及占比;(3)不同学历群体不同出行方式的通勤效率;(4)不同学历群体不同出行方式的实际通勤流数据。

4.2 数据结果

经推算统计,2015 年上海市常住就业人口共 1,247.47 万人,在上海都市区(除崇明区外)就业且居住的为 1,204.23 万人。其中,50.22%的就业人口居住地与就业地不在同一街道,49.78%的就业人口在同一街道居住且就业。大规模的跨区通勤流主要分布在中心城区

和郊区五大新城。图 1 展示了就业人口和就业岗位规模的分布情况（即，起讫点的通勤规模），可以看出就业人口分布呈现去中心化格局，就业岗位呈现集中与分散并存的现象。

表 2 本研究采用的通勤基准值与过剩通勤指标对比表

算法指标	测度指标	公式
通勤基准值	实际通勤距离（ C_{obs} ）	$C_{obs} = \frac{1}{W} \sum_i \sum_j T_{ij} c_{ij}$
	理论最小通勤距离（ C_{min} ）	$C_{min} = \min \left(\sum_i \sum_j T_{ij} c_{ij} \right)$
	随机通勤距离（ C_{ran} ）	$C_{ran} = \frac{1}{W^2} \sum_i \sum_j O_i D_j c_{ij}$
	实际通勤熵（ H_{obs} ）	$H_{obs} = - \sum_i \lambda_i^{obs} O_i - \sum_j \mu_j^{obs} D_j + \beta^{obs} C_{obs}$
	理论最小通勤熵（ H_{min} ）	$H_{min} = - \sum_i \lambda_i^{min} O_i - \sum_j \mu_j^{min} D_j + \beta^{min} C_{min}$
过剩通勤指标	随机通勤熵（ H_{ran} ）	$H_{ran} = - \sum_i \lambda_i^{ran} O_i - \sum_j \mu_j^{ran} D_j + \beta^{ran} C_{ran}$
	过剩通勤率（ EC ）	$EC = \frac{C_{obs} - C_{min}}{C_{obs}} \times 100$
	标准化过剩通勤率（ NEC ）	$NEC = \frac{C_{obs} - C_{min}}{C_{ran} - C_{min}} \times 100$
	标准化过剩通勤熵（ NEH ）	$NEH = \frac{H_{obs} - H_{min}}{H_{ran} - H_{min}} \times 100$

注： T_{ij} 为居住在街道 i 到街道 j 工作的就业人口规模， O_i 、 D_j 分别为街道 i 的就业人口和就业岗位规模， C 是平均通勤距离， c_{ij} 为街道 i 与 j 之间的距离， λ_i 、 μ_j 为拉格朗日系数， A_i 和 B_j 为平衡因子^[16]。

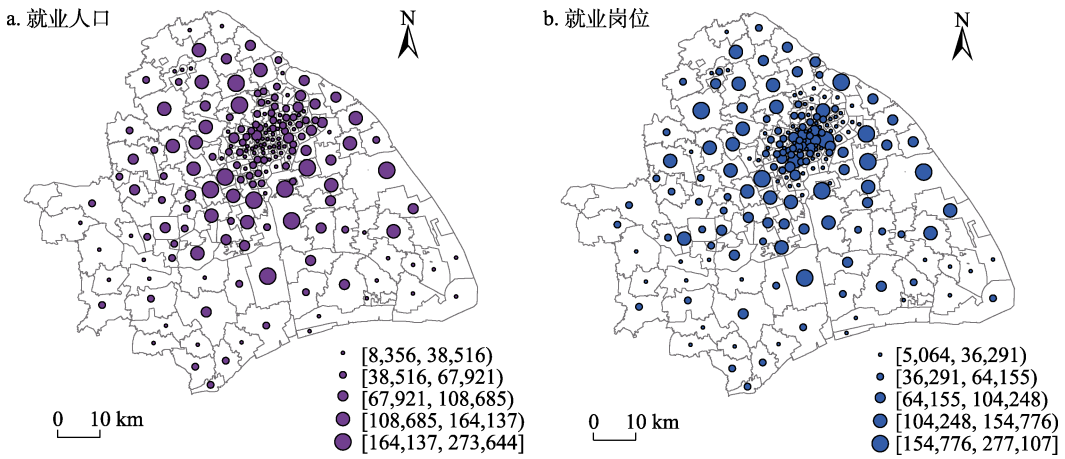


图 1 上海市就业人口和就业岗位规模分布图

从出行方式来看，非机动、小汽车和公共交通的出行占比分别为 55.73%、19.08%和 25.19%。这表明，与西方发达国家相比，中国城市的通勤出行仍以公共交通和非机动出行为主，小汽车出行占比较低。从不同学历群体的出行方式来看（图 2），几乎所有的低学历

就业者（92.34%）选择非机动出行方式。与之相比，高学历群体的非机动出行占比较低，仅 22.85%，而小汽车和公共交通出行的占比分别高达 30.46%和 46.69%。对于中学历群体，小汽车和公共交通出行占比分别为 13.25%和 13.42%。

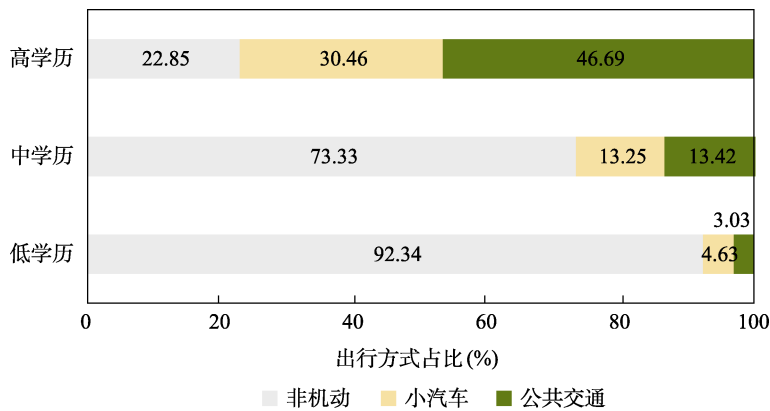


图 2 上海市不同学历群体的出行方式占比对比图

表 3 和图 3 为不同出行方式-学历水平的通勤效率结果。从理论最小通勤(C_{min})来看，低学历群体的通勤距离最长(4.17–6.56 km),高学历通勤者的出行距离最短(3.02–5.06 km),中学历通勤者的出行距离居中(3.94–5.26 km)。随机通勤(C_{ran})亦呈现一致的趋势，低学历就业者的通勤距离比高学历就业者长约 10 km。这表明，无论是微观还是宏观层面，低学历就业者的就业可达性最低，职住平衡程度最差。尽管中国城市不存在美国的种族歧视和种族隔离问题，但是居住地与就业地的过度分离，会给低学历、低技能等弱势群体带来巨大的时间成本和经济成本。已有研究表明职住分离会影响弱势群体的信息搜索能力、流动能力，制约其在城市住房市场和劳动力市场中的机会获得，进而可能引发更大的城市社会问题和空间问题^[17]。

实际通勤距离(C_{obs})的取值范围在 4.35–15.91 km 之间。在小汽车出行中，低学历就

表 3 不同出行方式的过剩通勤指标统计表

出行方式	学历	C_{min} (km)	C_{obs} (km)	C_{ran} (km)	H_{min}	H_{obs}	H_{ran}	EC (%)	NEC (%)	NEH (%)
非机动	低学历	4.17	4.72	39.71	4.98	5.62	9.41	11.65	1.55	14.50
	中学历	3.94	4.91	36.78	5.23	6.25	9.84	19.87	2.90	22.10
	高学历	3.02	4.35	28.22	5.43	6.98	10.06	30.49	5.26	33.34
小汽车	低学历	5.33	8.70	38.95	5.01	6.63	8.98	38.71	10.01	40.95
	中学历	4.59	10.45	37.47	5.35	7.84	9.71	56.10	17.83	57.07
	高学历	4.57	12.87	29.10	5.51	8.77	9.77	64.51	33.85	76.58
公共交通	低学历	6.56	15.91	34.90	5.30	8.46	9.41	58.77	32.98	77.03
	中学历	5.26	14.20	25.53	5.61	9.27	9.89	62.98	44.11	85.44
	高学历	5.06	13.66	20.04	5.48	9.23	9.57	62.94	57.42	91.78

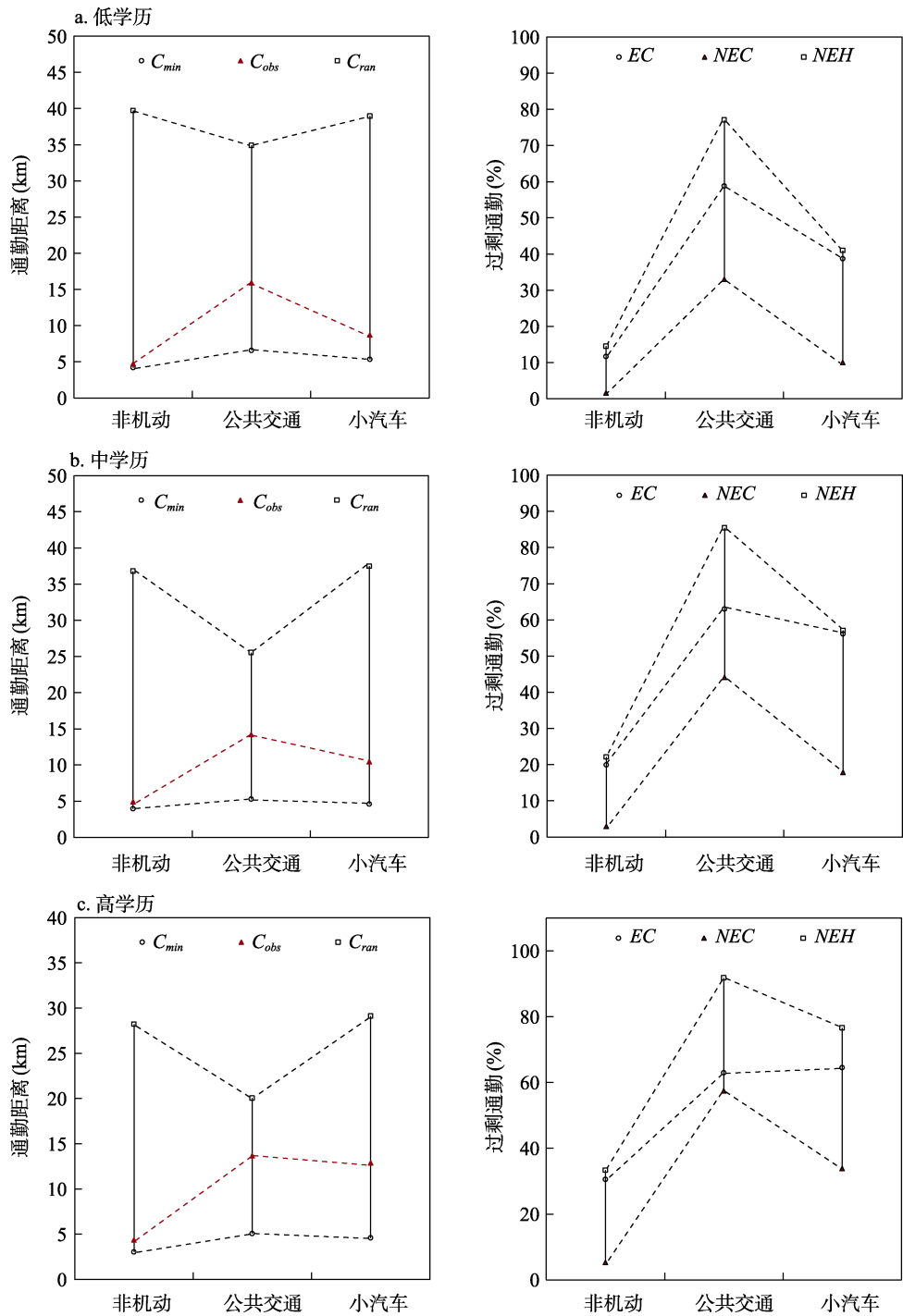


图 3 上海市不同学历群体间通勤效率的模式差异对比图

业者的通勤距离（8.70 km）远短于高学历就业者（12.87 km）。孙斌栋等^[18]在上海发现了两种类型的小汽车通勤：一种是长距离通勤，另一种是短距离通勤。他们认为，短通勤可能源于中国新富阶层的习惯和偏好，把开车作为一种炫耀性消费^[19]。在公共交通出行中，

低学历就业者比高学历就业者多通勤出行 14.14%。这是因为上海的居住-就业空间分布决定了低学历通勤者平均需要比高学历通勤者出行距离多 22.87%。另一个可能的解释是,低学历通勤者愿意为了低住房成本而忍受长距离通勤。高租金和高房价使得低学历通勤者无法在就业地附近找到负担得起的房子,他们被迫住在远离就业地的郊区。已有研究表明,住房压力在中国城市职住关系和居民主观幸福感中一直起着重要作用^[20]。

通勤熵值可用于反映通勤空间组织的有序性,通勤熵值越高,通勤空间组织越有序。其中,非机动出行的通勤熵值最低(5.62–6.98),公共交通出行的通勤熵值最高(8.46–9.27)。从表3可以看出,实际通勤熵值和通勤距离之间存在曲线关系。以低学历和高学历公共交通出行为例,二者的实际平均通勤距离(C_{obs})分别为15.91 km和13.66 km,实际通勤熵值(H_{obs})分别为8.46和9.23。这表明,相较于高学历群体,选择公共交通出行方式的低学历就业者承受着更长的通勤距离,但是通勤模式的空间结构更为有序。同样地,低学历与高学历非机动通勤者的实际平均通勤距离(C_{obs})分别为4.72 km和4.35 km,实际熵值(H_{obs})则分别为5.62和6.98。也就是说,低学历-非机动通勤者有着较长的平均通勤距离和较为有序的通勤空间组织,而高学历-非机动通勤者有着较短的通勤距离和无序的空间结构。

在平均出行距离视角下,低学历-非机动出行的过剩通勤最低($EC=11.65\%$)、通勤效率最高;高学历-小汽车出行的过剩通勤最高($EC=64.51\%$)、通勤效率最低。但当考虑通勤基准上限时,高学历-公共交通出行的通勤效率最低($NEC=57.42\%$),这与公共交通尤其是轨道交通的中心化布局有关。在空间组织视角下,除低学历-公共交通和高学历-小汽车出行外,标准化过剩熵(NEH)与标准化过剩通勤率的排序几乎相同。此外,长通勤并不必然是低效率的(如低学历-公共交通),而短通勤也并不一定是高效率的(如高学历-公共交通)。这是因为通勤效率是由实际通勤在理论最小通勤与随机通勤范围内的相对位置决定的。

5 讨论和总结

本文利用上海市1%人口抽样调查数据,计算了不同出行方式的通勤基准值和过剩通勤指标,量化了平均出行距离和通勤空间组织的非线性关系,为通勤效率相关研究提供了思路与参考。研究发现:(1)从出行方式来看,非机动、小汽车和公共交通的通勤出行占比分别为55.73%、19.08%和25.19%。与西方发达国家相比,中国城市的机动化出行占比较低。(2)不同出行方式的通勤效率存在明显的差异。无论是通勤距离视角还是通勤空间组织视角,通勤效率均表现为非机动>小汽车>公共交通。这表明,公共交通出行的通勤效率最低,职住优化空间最大,当前的城市土地利用不利于公共交通出行。(3)出行方式与居民社会经济属性在通勤效率中存在交互作用。在出行方式-学历双维度的交叉分析中,长通勤(如低学历-公共交通)并不必然是低效、无序的,短通勤(如高学历-公共交通)也并不必然是高效、有序的。(4)距离视角和空间结构视角下的通勤效率呈曲线关系。通勤空间组织可以反映城市通勤的有序性和结构性,且与交通拥堵与交通需求管理密切相关。对于具有相同平均通勤距离的两个城市,通勤空间组织较为无序的城市通常存在更多混乱而随机的交叉通勤,更容易造成严重的交通拥堵。因此,有必要从平均通勤距离和通勤空

间组织两个视角开展通勤效率研究。

本数据集反映了上海居民采用不同出行方式的职住关系及其通勤效率特征，为优化城市居住-就业空间分布及推动低碳出行提供了数据支撑。未来的研究可以基于本数据集，开展上海通勤效率的演化特征分析及通勤效率的城际差异比较研究。

作者分工：朱宇和岳丽莹对数据集的开发做了总体设计；岳丽莹和李开明采集和处理了数据；岳丽莹设计了模型和算法；岳丽莹、朱宇和李开明撰写了数据论文等。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献

- [1] Niedzielski, M. A. A spatially disaggregated approach to commuting efficiency [J]. *Urban Studies*, 2006, 43(13): 2485–2502.
- [2] 刘贤腾, 陈雪明, 周江评. 就业-居住空间关系及通勤效率——过剩通勤的评估潜力[J]. *城市交通*, 2017, 16(2): 10–18.
- [3] Horner, M. W. Extensions to the concept of excess commuting [J]. *Environment and Planning A: Economy and Space*, 2002, 34(3): 543–566.
- [4] 刘望保, 闫小培, 方远平等. 广州市过剩通勤的相关特征及其形成机制[J]. *地理学报*, 2008, 63(10): 1085–1096.
- [5] 朱玮, 梁雪媚, 桂朝等. 上海职住优化效应的代际差异[J]. *地理学报*, 2020, 75(10): 2192–2205.
- [6] 韩会然, 杨成凤, 宋金平. 公共交通与私家车出行的通勤效率差异及影响因素——以北京都市区为例[J]. *地理研究*, 2017, 36(2): 253–266.
- [7] Yue, L. Y., O’Kelly, M. E. Commuting performance in Shanghai: efficiency and orderliness [J]. *Travel Behaviour and Society*, 2023, 31: 223–231.
- [8] 岳丽莹, 朱宇, 李开明. 上海市不同出行方式的通勤效率数据集（2015）[J/DB/OL]. 全球变化数据仓储电子杂志, 2024. <https://doi.org/10.3974/geodb.2024.09.05.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2024.09.05.V1>.
- [9] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策[OL]. <https://doi.org/10.3974/dp.policy.2014.05> (2017 年更新).
- [10] 上海市统计局. 2015 年上海市 1%人口抽样调查资料[M]. 北京: 中国统计出版社, 2017.
- [11] Frost, M., Linneker, B., Spence, N. Excess or wasteful commuting in a selection of British cities [J]. *Transportation Research Part A: Policy and Practice*, 1998, 32(7): 529–538.
- [12] Zhang, H., Xu, S., Liu, X., et al. Near “real-time” estimation of excess commuting from open-source data: evidence from China’s megacities [J]. *Journal of Transport Geography*, 2021, 91(4): 1–14.
- [13] White, M. J. Urban commuting journeys are not wasteful [J]. *Journal of Political Economy*, 1988, 96(5): 1097–1110.
- [14] Yang, J. W., Ferreira, J. Choices versus choice sets: a commuting spectrum method for representing job-housing possibilities [J]. *Environment and Planning B: Planning and Design*, 2008, 35(2): 364–378.
- [15] Horner, M. W., Murray, A. T. Excess commuting and the modifiable areal unit problem [J]. *Urban Studies*, 2002, 39(1): 131–139.
- [16] O’Kelly, M. E., Niedzielski, M. A. Are long commute distances inefficient and disorderly? [J]. *Environment and Planning A: Economy and Space*, 2009, 41(11): 2741–2759.
- [17] 刘志林, 王茂军. 北京市职住空间错位对居民通勤行为的影响分析——基于就业可达性与通勤时间的讨论[J]. *地理学报*, 2011, 66(4): 457–467.
- [18] 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.
- [19] 李琬, 但波, 孙斌栋等. 轨道交通对出行方式选择的影响研究——基于上海市 80 后微观调查样本的实证分析[J]. *地理研究*, 2017, 36(5): 945–956.
- [20] Zhan, D. S., Kwan, M. P., Zhang, W. Z., et al. The impact of housing pressure on subjective well-being in urban China [J]. *Habitat International*, 2022, 127(3): 1–10.