

引用格式: 邬群勇, 张良盼, 吴祖飞. 顾及空间异质性的出租载客与公交客流回归分析[J]. 地球信息科学学报, 2019, 21(3): 337-345. [Wu Q Y, Zhang L P, Wu Z F. Regression analysis of taxi pick-up and bus passenger flow considering the spatial heterogeneity[J]. Journal of Geo-information Science, 2019, 21(3): 337-345.] DOI:10.12082/dqxxkx.2019.180380

顾及空间异质性的出租载客与公交客流回归分析

邬群勇^{1,2*}, 张良盼^{1,2}, 吴祖飞^{1,2}

1. 福州大学 卫星空间信息技术国家地方联合工程研究中心, 福州 350000;
2. 空间数据挖掘与信息共享教育部重点实验室, 福州 350000

Regression Analysis of Taxi Pick-up and Bus Passenger Flow Considering the Spatial Heterogeneity

WU Qunyong^{1,2*}, ZHANG Liangpan^{1,2}, WU Zufei^{1,2}

1. National & Local Joint Engineering Research Center of satellite-spatial Information Technology, Fuzhou University, Fuzhou 350000, China;
2. Key Laboratory of Spatial Data Mining & Information Sharing of MOE, Fuzhou 350000, China

Abstract: Taxi traffic has always been regarded as a supplement to public transportation. However, this may be in part due to previous studies focusing on independent research of taxi and bus passenger flow. Research around the relationship between taxi and bus passenger flow has not yet been thoroughly investigated. Taxi passenger hotspots not only provide real-time understanding of urban traffic hotspots, but also guide taxi drivers and enable taxi companies to make effective dispatches. Taxi passenger hotspots tend to occur in areas where demand for transportation is high and in areas of intensive crowding. Bus passengers' IC card data can reflect real-time traffic demand within the city. This study used Xiamen Island taxi GPS trajectory data and public transportation system data, along with the kernel density estimation method and geographic weighted regression (GWR) model to analyze the OD (Origin-Destination) passenger flow in both morning and evening peak travel times. Results showed a significant spatial heterogeneity in the kernel density value of the taxi passenger O. However, within the same area, the impact of bus passenger O and bus passenger D on the taxi passenger O was found to be opposite; in various regions, the negative impact of bus passenger O on the taxi passenger O in areas with complex urban functional types, bus passenger O had a positive impact on the taxi O at a single function area, while the bus passenger D was just the opposite. Compared to the ordinary least squares (OLS) model, GWR provided a much better fit (with the goodness of fit increasing from 0.13 and 0.11 to 0.59 and 0.53 in the morning and evening peak hours, respectively). Results of this study could provide the basis to forecast the number of taxi passengers.

Key words: passenger OD; spatial heterogeneity; geographically weighted regression; kernel density estimation; urban traffic

*Corresponding author: WU Qunyong, E-mail: qywu@fzu.edu.cn

收稿日期: 2018-08-16; 修回日期: 2018-12-19.

基金项目: 国家自然科学基金项目(41471333); 中央引导地方科技发展专项项目(2017L3012). [**Foundation items:** National Natural Science Foundation of China, No.41471333; The Central Guided Local Development of Science and Technology Project, No.2017L3012.]

作者简介: 邬群勇(1973-), 男, 山东诸城人, 博士, 研究员, 研究方向为时空大数据分析、地理信息服务。E-mail: qywu@fzu.edu.cn

摘要:出租车一直以来被看作公共交通的补充,但是以往研究多侧重于出租客流与公交客流的独立研究,对于二者的关联关系分析没有足够得到关注。预测出租车载客热点区域不仅能够实时的了解城市交通热点区域,还能够很好地指引出租车司机,帮助出租车司机快速寻客。出租车载客热点常发生在人流密集并且交通出行需求较高的区域,公交乘客IC卡数据能够实时的反映城市中的交通需求。因此,本文使用厦门岛出租车GPS轨迹数据与公共交通运输系统运营数据,利用核密度估计法和地理加权回归模型分析了早晚高峰时段出租车载客与公交上下车(OD)客流之间的时空分布关系。研究发现,出租乘客O点的核密度值在空间上存在分布不均衡性,聚集特征明显。在同一区域,公交乘客O点和公交乘客D点对出租乘客O点所产生的影响刚好相反;在不同区域,城市功能类型复杂的地区公交乘客O点对出租乘客O点产生负的影响,在城市功能类型单一的地区公交乘客O点对出租乘客O点产生正的影响,公交乘客D点则刚好相反。与普通线性回归模型相比,地理加权回归模型的拟合效果显著提高,早晚高峰拟合优度分别从0.13和0.11提升到了0.59和0.53。研究结果可为出租车载客数量的预测提供相关依据。

关键词:乘客OD;空间异质性;地理加权回归;核密度估计;城市交通

1 引言

出租车是公共交通系统的重要补充,在满足居民出行需求中发挥重要作用。因此,公交站点乘客IC卡数据能够从侧面反映出当地的出租乘客数量。理解出租车、公交车客流的空间分布特征和出租车客流与公交车客流之间的关系,一方面可以利用公交车客流变化规律指引出租车司机快速寻客,从而增加出租车司机的收入;另一方面也能利用出租车客流空间分布规律发现不合理的公交站点和公交线路,为公交站点和公交线路的规划提供依据。

以往研究多侧重于公交客流与出租客流的独立研究。大量研究利用IC卡数据和公交车GPS轨迹数据识别乘客的上车站点,并使用出行链方法推算乘客的下车站点^[1-3]。为提高计算效率,邬群勇等^[4]以IC卡数据、车辆GPS数据、线路站点数据和车辆基本信息数据为数据源,基于分布式存储、并行计算技术进行公交乘客OD(Origin-Destination)并行推算。Ma等^[5]利用IC卡和公交车GPS数据计算网络运行速度、公交客流量和车辆到站时间间隔等指标来评价公交系统的性能;Wang等^[6]构建以最小化乘客候车时间为目标的优化模型,从海量的公交数据中提取基于时间的交通路况和出行需求,基于真实的数据求解优化策略,结果能够有效降低乘客平均候车时间;邹智杰^[7]综合利用出租车和公交数据,通过出租载客热点发现不合理公交站点与公交线路,利用遗传算法对公交系统调度优化。Guo等^[8]使用SNN(Shared Nearest Neighbors)距离的K-Means聚类算法从海量出租车GPS轨迹数据中提取显著的热区,从出行轨迹中提取流向类簇,以深圳市的出租车数据为例分析居民移动的时空模式;Shen等^[9]利用改进的DBSCAN聚类算法

(Density-Based Spatial Clustering of Applications with Noise)从出租车轨迹数据中提取出行热区,根据居民出行的OD点的相似性对出行轨迹进行聚类,理解居民的出行需求和移动规律,尝试为空载出租车推荐寻找客源的最佳路线;孙飞等^[10]用出租车空载率来评价出租车运营效率并且得到了出租车优质客源的时空分布;Qi等^[11]探讨了出租车OD时空动态与城市社会功能区的关系,借此关系对社会功能区进行识别和提取;Liu等^[12]考虑了公民高度重复的活动链和任意两种具体土地利用类型之间的相互作用模式,对研究单元随机分配功能区类别,利用最大期望聚类(Expectation-Maximization, EM)算法对研究单元进行聚类,来理解城市的功能区结构。

在城市中,出租乘客轨迹数据和公交乘客刷卡数据能够实时反映城市的交通需求。吴建生等^[13]利用出租车GPS轨迹数据计算了不同时间段OD点全局Moran's I指数,证明了上客点与下客点的空间分布呈现显著为正的空间自相关,之后建立了出租车上下客事件的地理加权回归模型,分析并探讨了小城市出租车上下客时空动态与各类城市设施的时空关系和不同城市设施对上下客事件的影响的分布规律及其驱动机制。但是目前对于出租乘客与公交乘客出行的关联关系分析的研究少之又少。

出租乘客和公交乘客数据存在空间异质性,地理加权回归是一种考虑空间权重的局部回归方法,可以很好地解决空间异质性的问题^[14-20]。本文以厦门岛为研究区域,以公共交通运输系统运营数据和出租车轨迹数据为主要数据源,利用核密度估计和GWR回归模型分析了早晚高峰时间段出租车载客与公交OD(O点为起点,D点为目标地)客流之间的时空分布差异和关联关系,旨在为出租车载客数量与载客热点的预测提供相关依据。

2 研究区概况、数据源与研究方法

2.1 研究区概况

研究区为厦门岛,由湖里区和思明区(不含鼓浪屿)两个行政区,面积约157 km²,2015年常住人口大约200万,约占厦门市总人口的52%。研究区域土地利用图来源于厦门市规划委员会(2015年),如图1所示,岛内主要用地类型有居住用地、商业服务用地和工业用地。岛内用地类型的分布较复杂,居住用地主要分布在思明区北部和湖里区南部,除东部沿岸有较多集中规划的商业服务用地外,其它商业服务用地无规律地分布在居住地聚集区中。

2.2 数据源

研究数据为2015年6月13日至19日厦门市公共交通运输系统运营数据和出租车GPS轨迹数据,来源于厦门卫星定位应用股份有限公司。公共交通运输系统运营数据包括公交IC卡数据、公交车GPS轨迹数据、线路站点数据。公交车GPS轨迹数据包括车辆属性信息(设备编号)、线路属性信息(站点编号、线路编号)、车辆状态属性信息(时间、定位标识、经度、纬度、车次类型、进出站标识),根据进出站标识信息可将轨迹点划分为行驶轨迹点、进站轨迹点和出站轨迹点。出租车GPS轨迹数据的属性分为2类,分别为车辆属性(车辆ID)和车辆状态属性(定位时间、经度、纬度、车速、载客状态),采集间隔30 s。

2.3 研究方法

研究流程图如图2所示,主要分为乘客出行OD提取和空间异质性与回归分析2个部分。

2.3.1 乘客出行OD提取

出租车GPS轨迹数据中记录车辆的载客状态,通过载客状态的变化推算乘客的OD点,当载客状态从空载变为载客,载客的轨迹点即为乘客的上车点;当载客状态由载客变为空载时,空载轨迹点即为乘客的下车站点。乘客OD数据中车辆编号为出租车的唯一标识,上下车时间为车辆载客发生变化的轨迹点对应的GPS时间,上下车点为车辆载客状态发生变化的轨迹点经纬度坐标。

对于公交乘客出行OD提取,本文利用文献[4]的方法,将IC卡数据和车辆GPS数据存储到HBase数据库,线路站点数据存储到HDFS。运用MapReduce并行计算框架通过IC卡刷卡时间和车辆GPS轨迹时间匹配计算得到乘客上车站点,然后融合连续出行链方法和历史相似出行行为规律推算出乘客下车站点和下车时间,从而得到较为全面、准确的公交乘客出行OD记录。

2.3.2 空间异质性与回归分析

分析早晚高峰出租乘客OD的核密度空间分布规律,通过剖面分析空间异质性。结合公交乘客OD和出租载客核密度估计值创建公交站点空间数据表,计算全局Moran's *I*指数,分析出租载客的空间聚集程度,通过地理加权回归分析公交站点OD

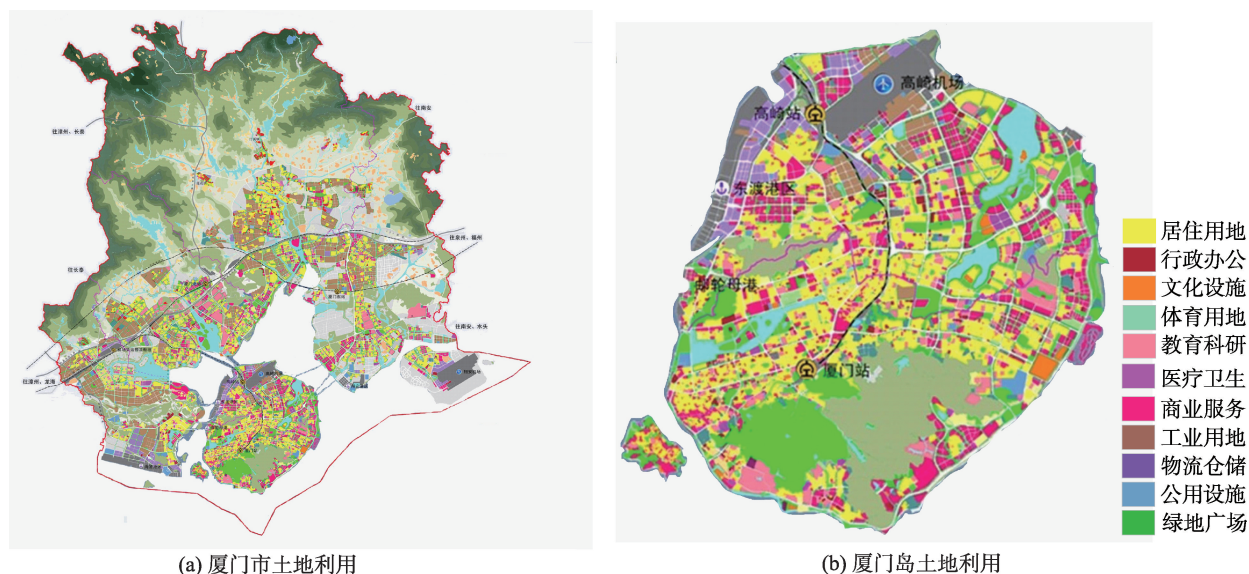


图1 厦门市及厦门岛2015年土地利用情况

Fig. 1 Land-use in Xiamen City and Xiamen Island in 2015

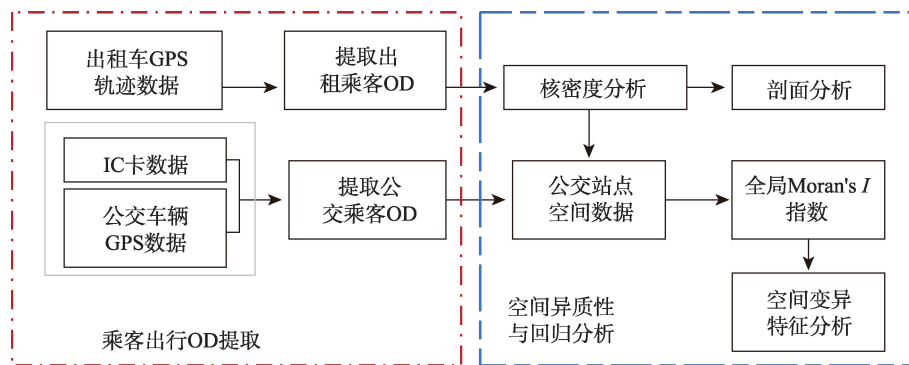


图2 出租载客与公交客流空间异质性回归分析研究流程

Fig. 2 The study flow of regression analysis of spatial heterogeneity of Taxi and bus passenger flow

数量对出租载客影响空间变异特征。在本研究中将公交站点处出租乘客O点核密度值作为因变量,公交站点处公交乘客O和D点数量作为解释变量进行地理加权回归拟合,模型表示如式(1)所示。

$$y_{i,t} = \beta_{0,t}(u_i, v_i) + \sum_{k=1}^p \alpha_{k,t}(u_i, v_i) x_{k,i,t} + \sum_{k=1}^p \beta_{k,t}(u_i, v_i) z_{k,i,t} + \varepsilon_{i,t} \quad (1)$$

式中: $y_{i,t}$ 为在 t 时间段公交站点 i 的出租乘客O点核密度值; $\beta_{0,t}(u_i, v_i)$ 为截距常量; (u_i, v_i) 为公交站点 i 的地理坐标, $x_{k,i,t}$ 为在 t 时间段公交站点 i 处公交乘客O点数量; $z_{k,i,t}$ 为在 t 时间段公交站点 i 处公交乘客D点数量, $\alpha_{k,t}(u_i, v_i)$ 为在 t 时间段公交站点 i 处O点的回归系数; $\beta_{k,t}(u_i, v_i)$ 为在 t 时间段公交站点 i 处D点的回归系数。用公交站点 i 邻近数据建立独立的局部最小二乘回归模型,使得接近公交站 i 的公交站具有高的权重。本文的权重由高

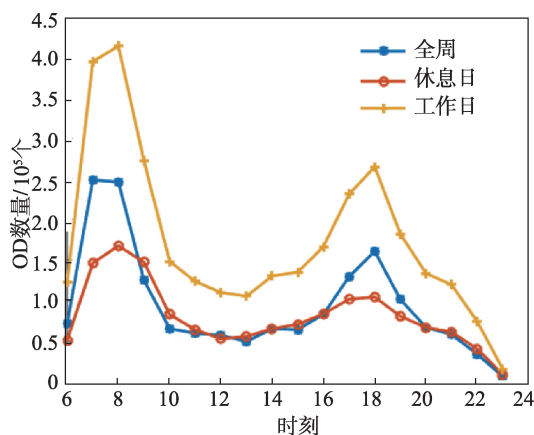
斯核函数确定,带宽的选择使用AIC_c准则。

3 结果与分析

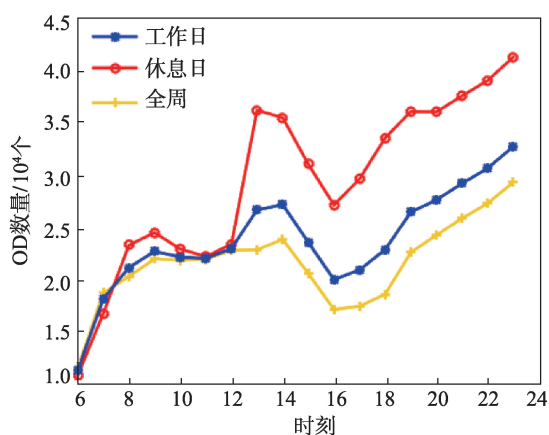
3.1 出租与公交客流时间分布差异分析

经数据预处理提取出租和公交乘客OD点数据,将一周七天、工作日(周一至周五)和周末(周六和周日)两天的OD数量分时段分别取平均,平均一天中各时间段OD数量的变化如图3所示。本文将6:00–9:00定为早高峰时间段,17:00–20:00定为晚高峰时间段。

图3(a)显示公交车各时段客流量总体呈现双峰分布,工作日和周末早晚高峰客流量差异突出。工作日的早晚高峰时段客流量占全天的46.2%,高于双休日的35.6%。相比工作日,双休日的早晚高峰表现得更为平缓。图3(b)显示出周末比工作日的出租车活动强度大。在工作日,从



(a) 公交车OD时间分布



(b) 出租车OD时间分布

图3 厦门公共交通出行时间分布

Fig. 3 Public transportation travel time in Xiamen

早晨 6:00 活动强度开始增强进入早高峰出行期；10:00 左右活动强度开始稳定持续到下午 15:00 左右；由于出租车司机交换班，16:00 出租车活动强度进入相对低谷期；从 17:00 活动强度又开始逐渐增大进入晚高峰出行期，活动强度持续增大。在周末休息日，在 13:00 之前 OD 时间分布和工作日大致吻合，13:00 之后周末活动强度骤然上升，明显高于工作日。无论公交出行还是出租出行工作日早晚高峰客流呈现较强的规律性，因此本文主要研究工作日期间早晚高峰客流。

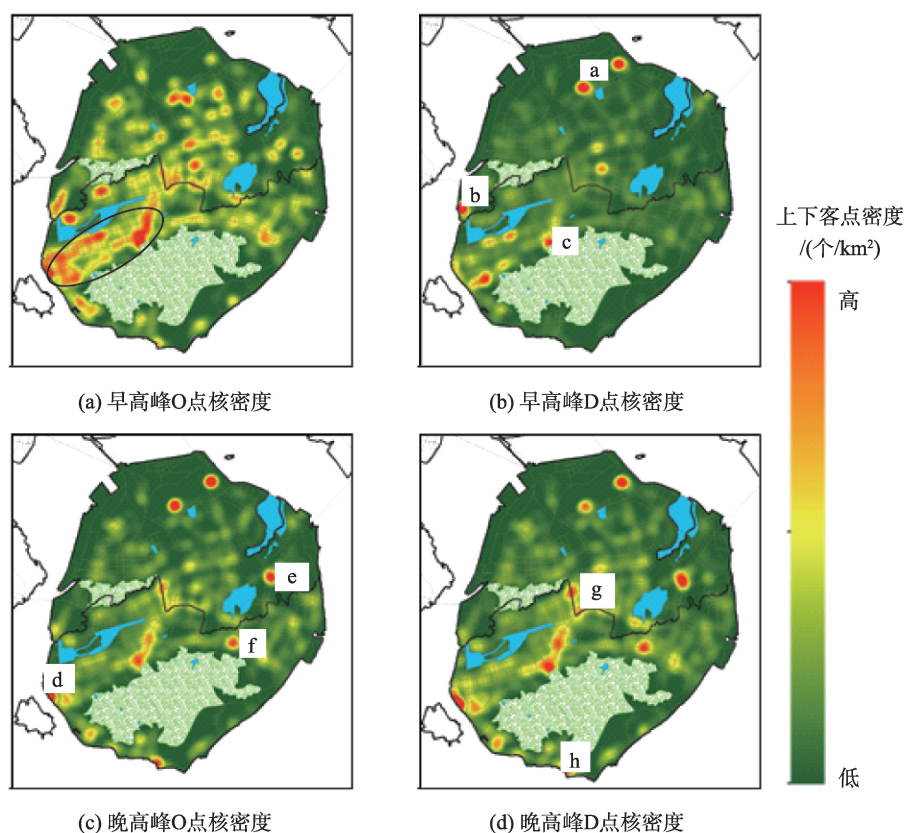
3.2 出租乘客 OD 核密度分析

核密度估计是地理空间数据挖掘的重要方法和工具，可表达城市空间点的分布模式、分布密度等特征。它是一种非参数估计的方法，对原始数据分布不基于任何数学前提假定，弥补了参数估计法中需要人们依靠经验来假定参数的缺陷^[21-22]。本文，首先基于提取的出租乘客 OD 位置数据，利用核密度估计的方法生成厦门岛内出租乘客 OD 客流连续表面，从而分析出租乘客 OD 客流的时空分布；然

后基于公交站点要素提取栅格像元值，并将这些值记录到公交站点要素类的属性表，从而估计厦门岛内每个公交站点出租上车点核密度值，进而分析公交乘客 OD 与出租上车量的关系。

核密度计算结果和搜索半径的大小有很大关系，随着搜索半径的增大核密度会趋于平滑。本文利用厦门岛内早晚高峰出租乘客 OD 位置点分别选取了不同搜索半径进行实验，最终选择搜索半径为 500 m，像元大小为 50 m×50 m。一方面在搜索半径为 500 m 时核密度在光滑的同时能够清晰地发现密度中心从而发现交通热区，另一方面 500 m 的搜索半径对于乘客候车和寻车都属于可接受范围，最终结果如图 4 所示。

厦门岛内出租乘客 OD 点从整体上具有西密东疏和中心密度较高的特点，在不同时间段出租乘客 O 点和 D 点核密度图差异较大。如图 4(a) 和 4(b) 早高峰 OD 点核密度图所示，在厦门岛内早高峰时间段出租车 O 点呈片状分布，上车 O 点高密度区域多集中在厦禾路和湖滨南路，在火车站商圈局部呈



注：a 厦门高崎国际机场；b 东渡码头；c 厦门火车站；d 中山路步行街；e 万达广场；f 瑞景商业广场；g SM 城市广场；h 胡里山炮台。

图 4 厦门岛出租乘客出行 OD 核密度

Fig. 4 OD kernel density map of taxi passengers on Xiamen Island

带状分布。早高峰时间段出租乘客D点核密度图和早高峰O点核密度有明显差异,出租乘客D点呈现多中心的空间分布特征,其主要分布在厦门高崎国际机场(a)、东渡码头(b)和厦门火车站(c)等区域,从而说明了早高峰乘客出行具有相同的目的性。如图4(c)和4(d)晚高峰OD点核密度图所示,晚高峰乘客OD点高密度区域较为相似,主要分布在高崎机场和火车站商圈,另外还有中山路步行街(d)、万达广场(e)、瑞景商业广场(f)、SM城市广场(g)和胡里山炮台(h)等商业广场和景点区域。综合对比四幅图,厦门早高峰O点核密度图较为不同,主要因为早高峰乘客上车位置多为居民小区和宾馆区域,从早高峰出租乘客O点可以反映厦门岛内人口主要居住区分布。

剖面图是通过有关的图形根据一定剖切方向所展现的内部构造图例。为了更加清晰地分析厦门岛内局部出租乘客OD早晚高峰的空间分布特征,分别从岛内东西方向、南北方向、西南东北方向和西北东南方向对其核密度图做了剖面分析,剖面图如图5所示。剖面图横轴代表距离,纵轴代表核

密度(个/ km^2),剖面原点分别为厦门岛的正北、正西、西南和西北。从厦门岛早晚高峰剖面图4个方向出租乘客OD核密度值的变化,可以分析岛内出租乘客OD数量的空间分布的内部结构和变化。

从出租乘客OD核密度剖面线可以看出:①早晚高峰乘客出行OD在相同方向空间分布非常相似,并且存在多个峰值,在不同方向差异性十分明显。在南北方向厦门岛内出租乘客OD核密度峰值多集中在厦门岛中间区域,从北至南核密度值先增大后减小,南部区域多为山地并且分布区域较广,所以核密度值大片区域为0;在东西方向出租乘客OD核密度值呈现中间高两边低的分布规律,早晚高峰OD在东西方向密度分布非常相似;在西南东北方向,出租OD密度西南高东北低,从西南至东北方向呈现递减阶梯分布,在数量上早晚高峰差距较大;在西北东南方向,早晚高峰出租乘客OD核密度值差距较大,其分布也呈现中间高两边低的分布。②厦门岛内出租乘客OD核密度值在空间上分布不均匀,存在不均衡性。乘客OD密度值高低和空间位置有很大的关联关系。若分析出租乘客与公交

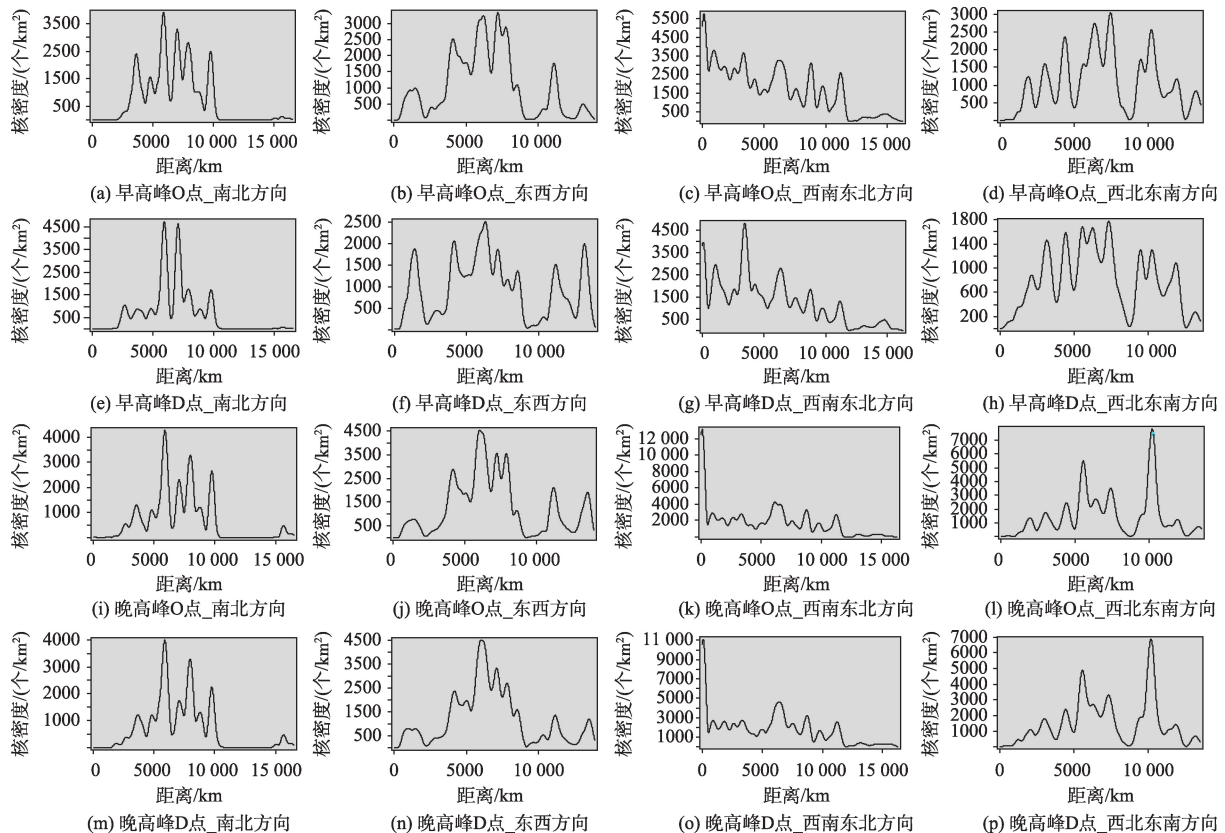


图5 厦门岛出租乘客出行OD核密度剖面图

Fig. 5 OD kernel density profile map of taxi passengers on Xiamen Island

客流的数量关系,应考虑空间权重的影响,普通最小二乘法则不能满足。

3.3 公交站点OD客流与出租载客客流空间异质特征分析

为了分析研究区公交站点处出租乘客O点分布集聚特征,计算全局Moran's I 指数。在早晚高峰出租乘客O点核密度的Moran's I 分别为0.86和0.76,表明早晚高峰出租乘客O点核密度空间分布呈现显著为正的空间自相关。建立地理加权回归模型进一步分析其空间异质性。

分别从早高峰和晚高峰2个时间段提取每个公交站点处的出租乘客O点的核密度估计值,然后基于GWR模型对厦门岛内公交乘客OD和出租乘客O核密度估计值进行拟合,拟合后公交乘客O和公交乘客D的回归系数空间分布如图6所示。

回归系数常可以用来证明自变量与因变量之间的相关关系^[16]。在本文中,公交乘客O和公交乘客D的回归系数可表示早晚高峰公交乘客上下车数量对出租乘客上车数量影响的强度,其空间变化

规律为:

早高峰,如图6(a)可以看出,公交乘客O的回归系数大部分为正值,对出租乘客O为正的,说明早高峰厦门岛大部分区域公交上车乘客多的区域出租上车的核密度值也高;图6(b)显示出早高峰公交乘客D点回归系数大部分区域为负值,对出租乘客O为负的影响。早高峰时刻公交乘客O的回归系数为正值区域多分布在厦门岛东部和北部,回归系数为负值的区域分布在厦门岛的西南部,然而公交乘客D的回归系数正负值在空间分布上则刚好相反。对比上文出租乘客OD核密度分布,在出租乘客OD核密度密集区域回归系数正负值较为复杂,而核密度值较低的东部和北部回归系数正负值则较为简单,这与城市繁华程度和功能区域类型有很大的关系。对于晚高峰,如图6(c)和图6(d),公交乘客O和公交乘客D的回归系数正负相间分布,对出租乘客上车数量核密度值的影响有很强的地域性特点。同一区域公交乘客O和公交乘客D的回归系数正负值相反。

由于厦门岛内经济发展程度不同,公交乘客上

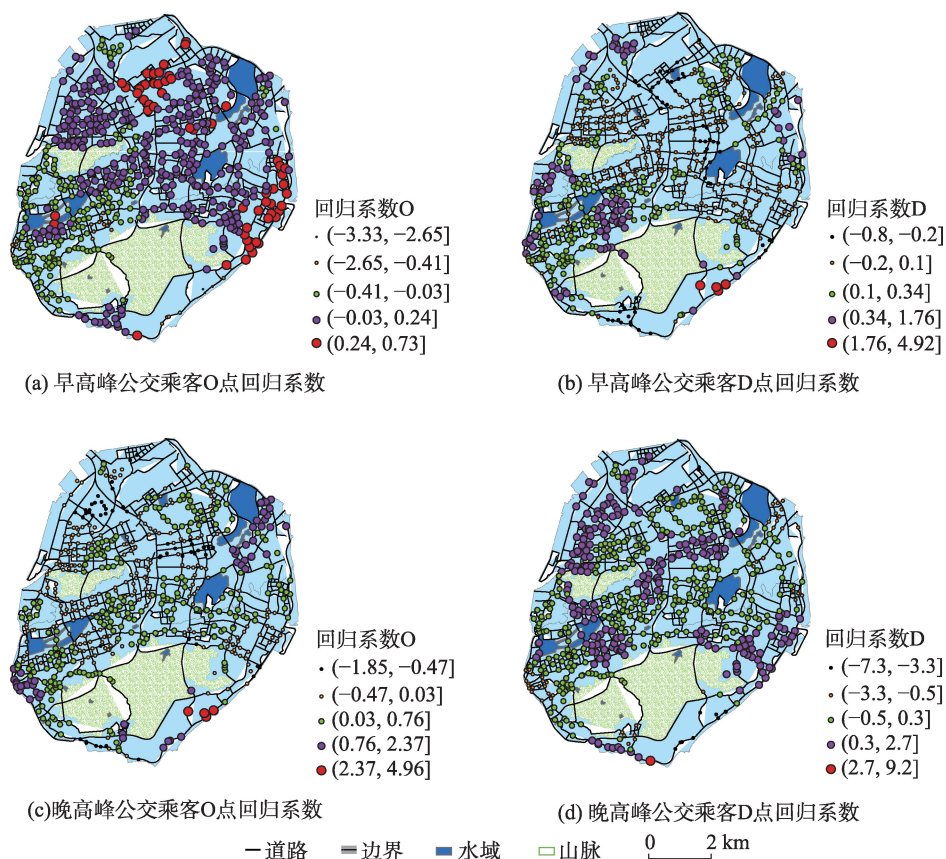


图6 早晚高峰公交乘客OD的回归系数

Fig. 6 Regression coefficient of bus passenger OD at morning and evening peak travel times

车数量、下车数量与出租乘客上车核密度值的关系复杂多变。公交乘客上车数量和公交乘客下车数量对出租乘客上车核密度值产生影响相反,在城市中,并非所有地点公交乘客上车数量与出租乘客上车核密度值成正相关,在功能区复杂的区域呈现不相关或负相关。

3.4 模型对比

基于GWR模型和OLS模型拟合评价指标 R^2 和调整 R^2 如表1所示。从早晚高峰2个时间段评价指标分析,早晚高峰GWR模型的拟合结果调整后 R^2 分别达到了0.59和0.53,基于OLS模型的拟合结果调整后 R^2 分别为0.134和0.109。基于GWR模型具有最大的调整 R^2 ,结果表明基于GWR模型的公交乘客OD与出租乘客O点核密度值的拟合结果优于OLS模型,反映了出租车载客数量与公交客流在空间上存在异质性。

表1 GWR模型和OLS模型拟合指标比较
Tab. 1 Comparison of fitting indicators between GWR model and OLS model

指标	GWR早高峰	GWR晚高峰	OLS早高峰	OLS晚高峰
R^2	0.71	0.68	0.14	0.11
调整 R^2	0.59	0.53	0.13	0.11

4 结论与讨论

本文结合2015年一周工作日早晚高峰公共交通运营数据和出租车GPS轨迹数据,通过核密度估计方法分析了厦门岛内出租乘客OD的时空分布,然后利用地理加权回归(GWR)模型对公交站点公交乘客OD与出租乘客核密度估计值进行了回归分析。主要研究结论如下:

(1)出租乘客OD早晚高峰核密度图能够较好的发现城市交通热点区域和乘客OD空间分布规律。通过核密度分析发现,出租乘客OD存在很高的空间异质性问题。出租乘客OD核密度值较高的区域多分布在厦门岛的西部和中部,尤其在在厦禾路和湖滨南路分布密度较高。

(2)通过GWR模型拟合后公交乘客O和D的系数空间分布分析发现,出租乘客O的核密度值和公交乘客OD数量的关系存在明显区域差异。不同地点公交乘客OD客流对出租乘客O的核密度值产生的影响不同。

(3)在工作日早晚高峰,GWR模型的拟合结果

调整后 R^2 为0.59和0.53,而OLS模型的拟合结果调整后 R^2 仅有0.13和0.11。GWR模型较OLS模型调整 R^2 高,表明GWR模型拟合优度显著大于OLS模型。GWR模型拟合效果更好,可以很好地解决空间异质性问题。

本研究的结论可以为出租载客热点区域的预测和出租车公司的调度提供相关依据。以公交车乘客上下客数量预测出租车载客热点,从而满足出租车快速寻客的需求,进而解决现存城市中打车困难的难题。在出租载客与公交客流强相关性区域,结合机器学习算法,可以为公交调度优化和出租寻客算法提供数据支撑。本研究只分析了出租乘客O点与公交乘客OD的回归关系,其他关系有待进一步研究。

参考文献(References):

[1] Wang W, Attanucci J, Wilson N. Bus passenger origin-destination estimation and related analyses using automated data collection systems[J]. Journal of Public Transportation, 2011,14(4):131-150.

[2] Munizaga M, Palma C. Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile[J]. Transportation Research Part C, 2012,24(9):9-18.

[3] Munizaga M, Devillaine F, Navarrete C, et al. Validating travel behavior estimated from smart card data[J]. Transportation Research Part C, 2014,44(4):70-79.

[4] 邹群勇,苏克云,邹智杰.基于MapReduce的海量公交乘客OD并行推算方法[J].地球信息科学学报,2018,20(5):647-655. [Wu Q Y, Su K Y, Zou Z J. A mapreduce-based method for parallel calculation of bus passengers origin and destination from massive transit data[J]. Journal of Geo-information Science, 2018,20(5):647-655.]

[5] Ma X L, Wang Y H. Development of A data-driven platform for transit performance measures using smart card data and GPS data[J]. Journal of Transportation Engineering, 2014,140(12):04014063.

[6] Wang Y, Zhang D, Hu L, et al. A data-driven and optimal bus scheduling model with time-dependent traffic and demand[J]. IEEE Transactions on Intelligent Transportation Systems, 2017,18(9):2443-2452.

[7] 邹智杰.数据驱动的公交调度分析与优化研究——以厦门市为例[D].福州:福州大学,2018. [Zou Z J. Data-driven bus scheduling analysis and optimization: A case study of Xia Men city[D]. Fu Zhou: Fuzhou University, 2018.]

[8] Guo D, Zhu X, Jin H, et al. Discovering spatial patterns in Origin-Destination mobility data[J]. Transactions in GIS,

- 2012,16(3):411-429.
- [9] Shen Y, Zhao L, Fan J. Analysis and visualization for hot spot based route recommendation using short-dated taxi GPS traces[J]. *Information*, 2015,6(2):134-151.
- [10] 孙飞,张霞,唐炉亮,等.基于GPS轨迹大数据的优质客源时空分布研究[J].*地球信息科学学报*,2015,17(3):329-335. [Sun F, Zhang X, Tang L, et al. Temporal and spatial distribution of high efficiency passengers based on GPS trajectory big data[J]. *Journal of Geo-information Science*, 2015,17(3):329-335.]
- [11] Qi G, Li X, Li S, et al. Measuring social functions of city regions from large-scale taxi behaviors[C]//*Proceedings of the 9th IEEE International Conference on Pervasive Computing and Communications*, WIP. Seattle, WA, USA, 2011:384-388.
- [12] Liu X, Kang C, Gong L, et al. Incorporating spatial interaction patterns in classifying and understanding urban land use[J]. *International Journal of Geographical Information Science*, 2016,30(2):334-350.
- [13] 吴健生,李博,黄秀兰.小城市居民出行行为时空动态及驱动机制研究[J].*地球信息科学学报*,2017,19(2):176-184. [Wu J S, Li B, Huang X. Spatio-temporal dynamics and driving mechanisms of resident trip in small cities[J]. *Journal of Geo-information Science*, 2017,19(2):176-184.]
- [14] Zhao Z, Gao J, Wang Y, et al. Exploring spatially variable relationships between NDVI and climatic factors in a transition zone using geographically weighted regression[J]. *Theoretical & Applied Climatology*, 2015,120(3-4):507-519.
- [15] 姜磊,周海峰,柏玲.外商直接投资对空气污染影响的空间异质性分析——以中国150个城市空气质量指数(AQI)为例[J].*地理科学*,2018,38(3):351-360. [Jiang L, Zhou H, Bai L. Spatial heterogeneity analysis of impacts of foreign direct investment on air pollution: Empirical Evidence from 150 Cities in China Based on AQI. *Scientia Geographica Sinica*, 2018,38(3):351-360.]
- [16] 韩雅,朱文博,李双成.基于GWR模型的中国NDVI与气候因子的相关分析. *北京大学学报(自然科学版)*,2016,52(6):1125-1133. [Han Y, Zhu W, Li S. Modelling relationship between NDVI and climatic factors in China using geographically weighted regression. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2016,52(6):1125-1133.]
- [17] 王海宾,侯瑞萍,郑冬梅,等.基于地理加权回归模型的亚热带地区乔木林生物量估算[J].*农业机械学报*,2018,49(6):184-190. [Wang H, Hou R, Zheng D, et al. Biomass estimation of arbor forest in subtropical region based on geographically weighted regression model[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2018,49(6):184-190.]
- [18] Fotheringham A S, Brunson C, Charlton M. Geographically weighted regression: The analysis of spatially varying relationships[J]. *American Journal of Agricultural Economics*, 2004,86(2):554-556.
- [19] Wang Q, Ni J, Tenhunen J. Application of a geographically-weighted regression analysis to estimate net primary production of Chinese forest ecosystems[J]. *Global Ecology and Biogeography*, 2005,14:379-393.
- [20] 袁玉芸,瓦哈甫·哈力克,关靖云,等.基于GWR模型的于田绿洲土壤表层盐分空间分异及其影响因子[J].*应用生态学报*,2016,27(10):3273-3282. [Yuan Y, H W, Guan J, et al. Spatial differentiation and impact factors of Yu Tian Oasis's soil surface salt based on GWR model[J]. *Chinese Journal of Applied Ecology*, 2016,27(10):3273-3282.]
- [21] 汤国安,杨昕. ArcGIS地理信息系统空间分析实验教程[M].北京:科学出版社,2016. [Tang G A, Yang X. ArcGIS geographic information system spatial analysis experiment tutorial[M]. Beijing: Science Press, 2016.]
- [22] Chu H J. Integration of fuzzy cluster analysis and kernel density estimation for tracking typhoon trajectories in the Taiwan region[J]. *Expert Systems with Applications*, 2012,39(10):9451-9457.