

基于百度指数的长江三角洲虚拟旅游流 流动特征和影响因素分析

杜家祺¹ 徐菁² 靳诚^{1, 31}

(1. 南京师范大学 地理科学学院, 江苏 南京 210023,

2. 南京晓庄学院 旅游与社会管理学院, 江苏 南京 211171;

3. 江苏省地理信息资源开发与利用协同创新中心, 江苏 南京 210023)

【摘要】: 旅游流是游客在空间区域内的集散过程所带来的一系列经济社会发展效应, 是旅游系统的神经中枢和纽带。传统旅游流研究多集中于现实旅游流, 对虚拟旅游流涉及较少, 数据获取多依靠调查问卷与景区提供, 具有局限性。文章基于百度指数, 在保证数据容量与准确度的基础上以长江三角洲为研究对象, 构建 O-D 虚拟旅游流矩阵探究区域虚拟旅游流流动特征, 并利用地理加权回归 (GWR) 模型从流入市域、流出市域两方面, 分析不同市域的虚拟旅游流影响因素。结果表明: (1) 长江三角洲虚拟旅游流空间分布不均, 流动性较强, 形成以合肥、南京、苏州、杭州、上海为核心节点的放射状网络。(2) A 级景区数量、住宿和餐饮业从业人员以及公路里程可以较好的解释流入虚拟旅游流市域分布。其中, A 级景区对流入虚拟旅游流影响最为显著, 住宿和餐饮业的发展对长江三角洲西北部虚拟旅游流的流入具有推动力, 而公路里程对整个长江三角洲流入虚拟旅游流影响最小。(3) 私人汽车拥有量、社会消费品零售总额、互联网宽带用户接入量对流出虚拟旅游流有较强的影响。社会消费品零售总额代表的经济因素是市域虚拟旅游流流出的决定因素, 经济越发达的区域, 人们的出游意愿越强烈, 也更愿意利用网络抓取旅游信息流追求高质量旅游出行行为。

【关键词】: 虚拟旅游流 影响因素 百度指数 GWR 模型

【中图分类号】: F301.1 **【文献标识码】:** A **【文章编号】:** 1004-8227(2021)02-0290-12

旅游流是游客在空间区域内的集散过程以及其引发的一系列社会经济发展效应, 是旅游系统的神经中枢和纽带。旅游流研究主要指从空间角度考察客源地 (Origin) 与目的地 (Destination) 的空间相互作用, 又称为 O-D 对研究^[1, 2]。客源地与目的地区域差异性的客观存在, 使得不同区域之间的空间相互作用强度产生明显差异, 其重要表现就是旅游流的规模、性质、空间尺度等流动特征的差异性以及区域旅游流影响因素的多样性^[3]。虚拟旅游流指依托于互联网检索数据的旅游流量在网络空间流动的过程映射, 由具有流向、流量、空间差异等特征的流动性检索数据组成, 是一种非现实性旅游流流动。国内学者对现实旅游流的研究较多, 而对网络空间虚拟旅游流的研究才刚刚起步^[4-6]。随着我国旅游业的不断发展, 人均出游率从 1992 年的 0.25% 跃升至 2015 年的 8.15%^[7], 随着出游率不断增加, 旅游流给旅游目的地带来了巨大效益的同时也带来了一系列问题。一方面, 旅游目的

¹作者简介: 杜家祺(1995~), 男, 硕士研究生, 主要研究方向为区域旅游流. E-mail: dujiazhen1@163.com
徐菁, E-mail: xujing2431@163.com

地的人均旅游收入的增加提升了目的地居民生活水平和幸福感,促进了就业问题的解决,有利于历史文化遗产的保护与开发,对区域经济发展起到了积极作用;另一方面,过多的游客进入会占用目的地基础设施,引发景区的物价上涨,破坏目的地生态环境。由于不同空间中旅游流的流动特征和影响因素的差异性以及旅游目的地的独特性,旅游流的运行机理具有明显的空间异质性,旅游流的负面影响因素的消除也需要考虑区域差异性具体分析。而虚拟旅游流作为现实旅游流产生的晴雨表,具有先兆性,能够预测旅游流的流动趋势和行为规律。通过对旅游者的出游规律等流动特征的预测,可完善旅游目的地基础设施,优化游憩休息场点,并结合旅游流影响因素差异性保证旅游业态的合理开发。因此,研究虚拟旅游流流动特征和影响因素,可反映市场的动态规律以及趋势,对旅游目的地游憩场所设置,公共设施规划,缓解目的地之间的恶性竞争以及旅游业的持续发展具有重要意义。

在旅游流流动特征研究方面,国内学者们大多依托模型构建旅游流网络结构对其特征进行分析^[8~10],主要从入境旅游视角探究旅游流流向、密度、轨迹演变,进一步通过对各节点网络特征进行中心性分析,利用不同时间尺度下的数据进行纵向对比,得出流动性特征^[11~15];国外学者对旅游流流动特征研究包括利用时间序列模型、结构方程模型等模型研究旅游流流动状态和流动意向^[16,17],以及旅游流的流动演变过程,例如不同边界划分下的旅游流变迁^[18],航空体制约束性影响下的旅游流流动规律^[19],或是基于交通数据对旅游流特征进行临界条件下分析^[20],但主要也是从入境旅游角度进行讨论。而在旅游流影响因素方面,国内地理学者主要从旅游流形成机制探索其影响因素,依托于空间相互作用理论,需求理论,竞争力理论等,从旅游者的内在驱动力、旅游需求的外在动力等多方面进行探讨^[21],包括:旅游资源禀赋,区域经济发展水平,基础设施建设,区域交通^[22~24];国外研究者同样构筑旅游流影响因素分析模型,探究目的地可达性^[25]、旅游基础设施^[26]、对外旅游政策^[27]、距离因素^[28]、国际关系^[29]、游览时间^[30]以及旅游目的地环境^[31]等因素对旅游流影响。然而,传统的旅游流数据获取,大多利用调查问卷或是旅游景区提供的人数(客流量)作为入境旅游流一手数据^[8,10,14]。前者存在样本数量有限性以及在此基础上的分析很难全面反映旅游者行为特征的问题;后者则主要表现旅游者的空间分布特征而非流动性特征,较难体现点对点之间的旅游流流入、流出的流动特征。同时,以往旅游流多从现实角度进行分析,缺少对于网络空间中虚拟旅游流的关注。目前国内外普遍将虚拟旅游流看作是一种网络营销工具,利用网络空间导向功能,结合网络评价,网站的信息有效性与娱乐性剖析旅游吸引力在网络空间中的实现,注重旅游业和日常生活去分化下的互联网这一工具对于游客目的地选择可能性的提升。尽管对虚拟旅游流与现实旅游流的变现关系做了一定的研究,但并未重视虚拟旅游流的网络空间映射过程,忽视了虚拟旅游流的流动性特征,具有很大的局限^[4~6,32~35]。伴随着第三次互联网革命的发生,网络用户数量激增,虚拟旅游流的流动规模也大大提高,进一步增强了其对于现实旅游流的预测的准确性和全面性。在 Web2.0 时代^[36],各种新技术层出不穷,为旅游流的研究提供了海量的数据,也为虚拟旅游流的研究提供了巨大的样本容量。因此,准确获取网络世界中不同空间尺度上区域点对点的虚拟旅游流流动数据,保证数据容量的宽度和准确度,并在此基础上从出入境两方面分析虚拟旅游流流动特征和影响因素是旅游流研究需要进一步探讨的问题。

基于此,本文以泛长江三角洲(江苏、浙江、上海、安徽)为研究区域,利用百度指数构建旅游出发地-目的地流量矩阵,探究区域虚拟旅游流流动特征。并运用 GWR 软件对虚拟旅游流影响因素进行显著性检验,进一步分析了影响区域虚拟旅游流流动的社会经济影响因素,以期研究区域政府部门和旅游科研机构制定更为科学的措施,为区域旅游空间布局优化以及旅游地精准营销提供依据。

1 研究区域和数据说明

1.1 研究区域

长江三角洲范围覆盖三省一市,包括江苏省、浙江省、安徽省以及上海市共 41 个市级行政单位,经济发达,人均 GDP 位于全国前列。截至 2018 年,长江三角洲 GDP 总量为 213764.46 亿元,约占 2018 年全国 GDP 的 1/5;网络覆盖率高;江苏、安徽、浙江三省高速公路里程达 13967km²,高速铁路里程达 3734km²,上海全市轨道交通运营线路长度达 704.91km²,交通通达性较好;地形以平原为主,同时具有盆地和低山丘陵等地貌,共有 A 级以上景区 1895 个,旅游资源禀赋性较好;旅游总收入 34760.15 亿元,旅游业较为发达。因此,无论从经济因素和旅游资源,还是城市间的时空收敛效应来看,长江三角洲对于反映虚拟旅游流

流动特征及其影响因素，都具有典型性。

1.2 研究方法

1.2.1 节点流量

在 O-D 虚拟旅游流矩阵中，每个节点与其他节点存在着旅游流的集聚和扩散的关系，每一个节点既是客源地又是目的地。由于不同节点之间的旅游流的规模存在差异，因此造成了节点的集散能力的差异性，也就是节点在矩阵中的重要性不同。对于节点的重要性可以利用流入量、流出量以及总流量进行描述，计算公式如下：

$$\text{流入量: } C_{ini} = \sum_{j=1}^n l_{ji} \quad (1)$$

$$\text{流出量: } C_{outi} = \sum_{j=1}^n l_{ij} \quad (2)$$

$$\text{总流量: } C_i = C_{ini} + C_{outi} \quad (3)$$

式中：C_{ini} 是节点 i 的流入量；C_{outi} 是节点 i 的流出量；l_{ji} 为从节点 j 流向节点 i 的流量；l_{ij} 为节点 i 流向节点 j 的流量；C_i 为节点 i 的总流量。根据节点的流入量，流出量计算节点流出率，具体公式如下：

$$\text{流出率: } Z_i = (C_{outi} - C_{ini}) / C_i \quad (4)$$

式中：Z_i 为节点 i 的流出率；C_{outi} 为节点 i 的流出量；C_{ini} 为节点 i 的流入量；C_i 为节点 i 的总流量；Z_i>0 时，i 节点为净流出，数值越大，流出率越大；Z_i<0 时，i 节点为净流入，数值越小，流入率越大。

1.2.2 GWR 模型

地理加权回归(Geographically Weighted Regression,GWR)模型利用非参数局部加权回归技术，通过引入对不同区域的影响进行评估，反映参数在不同空间尺度内的非平稳性以及变量间的关系随空间变化的特征，相较于传统的线性回归而言更符合现实^[37~39]。本文利用 GWR 模型对长三角地区虚拟旅游流的影响因素进行空间回归分析，模型如下：

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{k,i} + \varepsilon_i \quad (5)$$

式中：(u_i, v_i) 是第 i 个单元的空间地理坐标；β_k(u_i, v_i) 为独立变量 x_k 在 i 单元的系数值；y_i 为第 i 单元上的因变量；x_{k,i} 为 i 单元上的第 k 个解释变量；ε_i 为随机误差。

1.2.3 数据说明

本文构建 O-D 虚拟旅游流矩阵数据来源于百度指数。百度作为全球最大的中文搜索引擎，具有海量的数据，百度指数正是基于此的数据分享平台，包括搜索指数和媒体指数。搜索指数是以网民在百度引擎上的搜索量为统计对象，科学分析并计算出网民检索关键词在百度网页中的搜索频次加权和^[40~44]。当某一地区百度指数以另一地区为关键词时，两地之间就出现了信息的

流动, 这种流动对于旅游流具有一定预测性, 可以作为现实旅游流的晴雨表反馈潜在旅游流的数量, 与现实旅游流之间的流量存在相关关系, 是网络空间中旅游流流动的映射。因此, 百度指数作为具有流动性以及具有变现能力的居民自检索数据, 可以代表虚拟旅游流进行数据分析, 具有较高的可信度和合理性。本文选取数据为 2018 年全年长江三角洲三省一市百度搜索真实检索量。由于中国目前大多社会经济要素多以县、市为范围进行划分, 本文依照长江三角洲最新县市行政区划进行统计, 通过百度指数平台, 以 41 个市级行政单位为基础, 利用软件以“城市名+景点”, “城市名+旅游”为关键字, 依据不同市域 IP 定位, 提取 2018 年检索数据共 1681 条, 构建 41X41 二维有向流量矩阵。由于主要考察不同市级行政区域之间的旅游流的流动性特征及影响因素, 故将市域自检索数据进行剔除, 获得有效数据 1640 条。本文涉及的相关社会经济数据, 来源于 2018 年长江三角洲各省市统计公报^[45~48]。

2 研究结果与分析

2.1 流动特征分析

由于本文主要通过构建 O-D 虚拟旅游流矩阵对长江三角洲虚拟旅游流进行探讨, 而旅游流矩阵则是由旅游节点和节点间的旅游流路径构成, 故本文主要从虚拟旅游流流动路径以及旅游节点规模两方面进行分析。为便于分析, 本文通过 ArcGIS 软件, 将长江三角洲旅游流流动路径以及市域流量、流出率空间分布进行可视化处理, 结果如图 1 所示。

2.1.1 流动路径分析

长江三角洲三省一市共 41 个市域范围内存在路径 1640 条, 利用 ArcGIS 软件对所有流动路径进行可视化表达(图 1a)。为更好分析虚拟旅游流路径分布, 发现区域差异, 现将虚拟旅游流路径值在 200 以上的路径称为高流量路径, 虚拟旅游流路径值在 100 以上 200 以下的路径称为次流量路径, 虚拟旅游流路径值在 100 以下的路径称为低流量路径。高流量路径主要存在于江苏南部、浙江、安徽东南部以及上海之间, 路径值大于 400 以上的流动路径以南京、苏州、杭州、上海为节点形成放射状网络, 由苏南中部、浙江中部与北部以及上海组成; 路径值在 200~400 之间的流动路径以合肥、南京、上海、杭州此类省会城市为核心向安徽东南部、江苏南部以及浙江全省进行辐散; 杭州、上海在整个高流量路径中, 吸收大量旅游流同时又辐散出大量旅游流, 具有强大的旅游流集散效应。次流量路径中, 以南京、杭州、上海为核心的江苏中南部, 浙江北部以及上海地区虚拟旅游流路径密集程度最高, 其次为浙江南部地区, 安徽省则以其省会城市合肥为中心向外进行辐散, 其密度最低。低流量路径覆盖整个长江三角洲三省一市, 但长江三角洲北部流量最小。可以看出, 长江三角洲中部的苏南和浙北地区是中、高流量路径分布最多的区域, 而此部分区域也是整个长江三角洲经济最为发达、旅游业最为繁荣的市域单元集中所在。低流量路径值虽覆盖整个长江三角洲三省一市, 但北部经济相对相差, 区域低流量路径分布最多, 分析区域经济发展水平和旅游资源禀赋情况可能对虚拟旅游流分布造成一定影响。

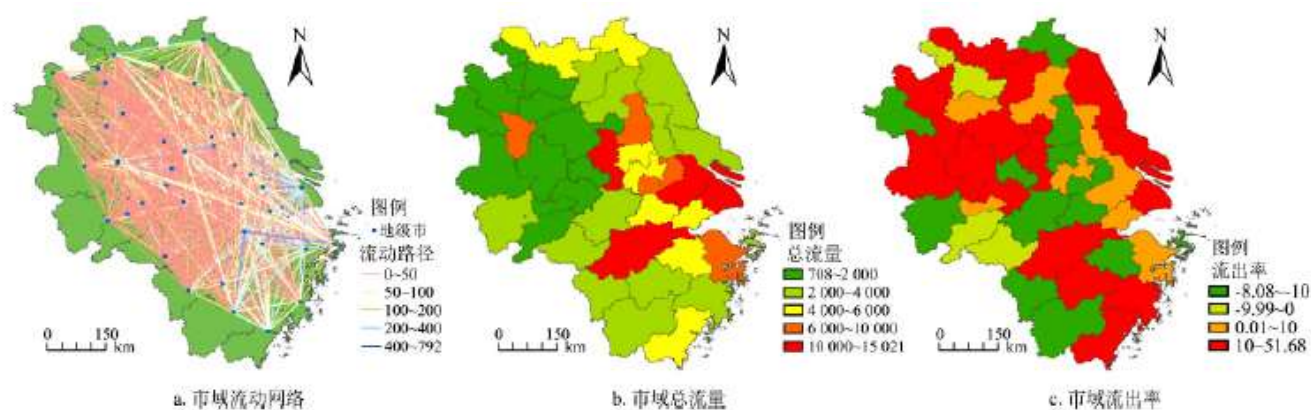


图 1 长江三角洲市域之间流动特征

进一步分析市域之间不同流量数据及流量占比,高流量路径中,路径值在 300 以上的路径数有 41 条,占全部流量的 18.78%,在网络中优势较大;次流量路径有 205 条,占全部流量的 32.15%,所占比例也较大;低流量路径中,路径值小于 30 的路径数达 1124 条,但仅占全部流量的 7.4%。可以看出,整个长江三角洲虚拟旅游流流动主要集中于苏州南部、浙江、安徽东南部以及上海几处区域,大部分市域区间虚拟旅游流流动较少。

表 1 市域间流量分布

流量规模	高流量路径 (>200)	次流量路径 (100~200)	低流量路径 (<100)
路径数(条)	104	205	1331
流量占比(%)	35.58	32.82	31.6

2.1.2 流量特征分析

统计各市域的虚拟旅游流流入量与流出量,从而计算市域虚拟旅游流总流量(图 1b)。从虚拟旅游流总流量计算结果可知,整体上呈现东高西低的空间分布,为了进一步分析市域之间虚拟旅游流的空间差异,对其进行重分类,将其划分为 5 个层级。第一个层级为总流量为 10000 以上的市域,包括 4 个市域。虚拟旅游流流量最高的是上海,为 15021,占全部虚拟旅游流总量的 8.36%,其次是杭州,虚拟旅游流流量为 13657;苏州和南京分列 3、4 位,分别为 13222 和 12496。除安徽省外,其余两省一市均以省会城市为旅游流节点,这些市域具有丰富的旅游资源和交通区位优势,且经济较为发达,在吸纳了大量旅游流流入的同时也有大量的居民选择出游,因此流量巨大。江苏省更是拥有除省会城市外的另一个旅游流节点苏州,依托上海形成了巨大的旅游集聚和扩散效应。第二个层级为总流量在 6000~10000 的市域,包括 4 个市域,分别为合肥、扬州、无锡、宁波,其中无锡因位于苏州和南京两个第一层级旅游节点之间,受其集聚和扩散效应影响,虚拟旅游流流量达到了 9523,而安徽则利用其省会城市的优势排名第二,虚拟旅游流流量为 7573,此部分市域相对于第一层级而言,经济发展水平,旅游资源丰富度均较低,且部分市域失去省会城市的交通集散中心作用,流量稍有下降。第三层级为总流量在 4000~6000 的市域,包含 8 个市域单元,主要分布在江苏中南部和浙江中北部,其中镇江、常州、湖州、嘉兴、绍兴因靠近第一、二层级旅游节点而形成,享受到了高级旅游节点的旅游流辐散,徐州、连云港以及温州则因靠近省界而获得了一定的虚拟旅游流流量。此层级市域具有一定的旅游资源和交通区位优势,并因此获得了一定的旅游流流量。第四层级包括 13 个市域,主要分布于江苏北部,浙江南部以及安徽东南部。第五层级包括 12 个市域,全部位于安徽省境内。第四,第五层级的市域共 25 个,经济发展水平、旅游资源以及交通区位优势均不如前 3 个层级,且总流量仅占全部流量的 30.55%,在网络中占比较低,所起作用不强。

根据虚拟旅游流流入量、流出量以及总流量来计算市域流出率,计算结果如图 1c 所示。从市域流出率结果来看,长江三角洲虚拟旅游流流动性较强,41 个市域中,只有 11 个能保持在(-10%,10%)的相对均衡区间中,大部分市域具有巨大的流动性,呈现出明显的净流入和净流出特征。流出率大于 10%的市域多集中于安徽中部、江苏东部、上海以及浙江中部,共 17 个市域,其中江苏和安徽各有 6 个市域在此区域内。流出率在 40%以上的市域有 4 个,流出率最高的市域为阜阳,达到了 51.68%,其次为淮北 50.48%,宿州 47.87%,合肥 45.75%,旅游流净流出十分巨大。大部分流出市域都集中分布于安徽,结合旅游流量分析,安徽省尽管旅游流总流量较低,但是居民出游欲望较高,旅游流流出量大,同时因区域内旅游资源较少,经济发展水平较低,基础设施不如长江三角洲其他市域,所以旅游流多为向外流出。净流出较低的为六安、南京、杭州,平均为 11%,但南京和杭州由于较大的人口基数,较低的净流出率并不能抹杀其重要的旅游客源地的地位。同时由于南京和杭州的旅游资源丰富,知名度较高,吸引了大量的旅游人数,也削弱了旅游流净流出可能。流出率小于 10%的市域有 13 个,集中分布于安徽南部、江苏中部、浙江

南部。净流入最高的市域为黄山，其流出率为-80.18%，其次为丽水-62.32%，舟山-61.46%，净流入特征十分明显。此类市域有着充足的旅游吸引力但是因自身人口因素，流入人口远远大于流出口导致流出率为负。净流入较低的是淮南，其流出率为-12.69%，基本可看作相对均衡。

2.2 影响因素分析

对于市域尺度研究单元而言，由于多以社会经济要素进行划分，同时社会经济要素也具有易获取性的特征，故本文主要从市域虚拟旅游流流入和流出两方面探讨虚拟旅游流与社会经济要素之间的相互关系。在参考前人研究的基础以及上文的旅游流流动特征分析结果^[2,3,49]，依据旅游流推拉理论，需求理论为基础，结合旅游流驱动力模型，从目的地拉力，客源地推力，距离阻力^[21]三个方面选取旅游流影响因素。目的地拉力包括吸引力和支撑力两部分，吸引力是刺激旅游者产生出游欲望的首要动力，支撑力则为被吸引而来的游客提供旅游质量保证，是目的地拉力中的后续动力。客源地推力是基于旅游需求角度出发，包括内在推力和外在推力，内在推力指旅游流的内在驱动机制，主要是旅游者自身需求而形成，外在推力则是包括经济发展水平、可自由支配收入和时间等外在指标，主要指保证旅游者因内在推力激发的旅游需求实现的可能。距离阻力则主要是从空间角度探讨阻止游客出游的因素。流入市域选取 A 级旅游景区、星级饭店、住宿和餐饮业从业人数、公路里程、GDP、医疗卫生机构、空气优良率 7 个要素作为解释变量，其中 A 级旅游景区、星级饭店、空气优良率作为目的地拉力中的吸引力，GDP、住宿和餐饮业从业人员、医疗卫生机构作为支撑力，公路里程作为距离阻力。流出市域选取人口、GDP、住户存款数、私人汽车拥有量、城镇职工基本养老保险人数、社会消费品零售总额、互联网宽带用户接入数作为解释变量，其中人口、GDP、住户存款数、社会消费品零售总额作为外在推力，城镇职工基本养老保险人数、互联网宽带用户接入量作为内在推力，私人汽车拥有量作为距离阻力。这些解释变量均来自 2018 年长江三角洲各省统计公报^[45-48]。

为进一步分析这些解释变量对区域虚拟旅游流的影响程度以及在空间上表现出的影响特征，本文引入 GWR 模型分析解释变量之间的空间相互关系。为避免出现多重共线性问题，首先利用 spss 软件对所有解释变量进行标准化处理，然后利用线性回归方法对标准化后的解释变量进行共线性检验，将方差膨胀因子(VIF)大于 10 的解释变量剔除。最终流入市域选取 A 级旅游景区、公路里程、住宿和餐饮业从业人数作为最终解释变量，流出市域选取互联网宽带用户接入量、私人汽车拥有量、社会消费品零售总额作为最终解释变量，分别构建回归模型如下：

$$IN(i) = \alpha_0(i) + \alpha_1(i) VIEW(i) + \alpha_2 SERVICE(i) + \alpha_3 ROAD(i) + \varepsilon(i) \quad (6)$$

式中：IN(i)为市域流入总流量；VIEW(i)为各市域 A 级旅游景区；SERVICR(i)为各市域住宿和餐饮业从业人员；ROAD(i)为公路里程数。

$$OUT(i) = \beta_0(i) + \beta_1 NETWORK(i) + \beta_2 CAR(i) + \beta_3 SALE(i) + \varepsilon(i) \quad (7)$$

式中：OUT(i)为市域流出总流量；NETWORK(i)为各市域互联网宽带用户接入量；CAR(i)为各市域私人汽车拥有量、SALE(i)为各市域社会消费品零售总额。

为更好体现 GWR 模型相较于传统的线性回归模型更符合现实情况，依次对虚拟旅游流市域流出和流入解释变量分别进行 OLS 回归分析与 GWR 回归分析。虚拟旅游流市域流入 OLS 回归结果 R^2 为 0.6987，调整后 R^2 为 0.5739，AICc 值为 107.7291。而 GWR 模型回归结果 R^2 为 0.8786，调整后 R^2 为 0.8497，AICc 值为 95.0111。虚拟旅游流市域流出 OLS 回归结果 R^2 为 0.8198，调整后 R^2 为

0.7941, AICc 值为 62.4598。而 GWR 模型回归结果 R^2 为 0.9707, 调整后 R^2 为 0.9398, AICc 值为 38.0803。无论是虚拟旅游流市域流入还是市域流出, 从 R^2 看 GWR 模型回归结果远优于传统的 OLS 回归结果, 解释变量拟合度较高。若 AICc 值相差大于 3, 表明 GWR 模型拟合性能优于 OLS 回归模型^[50], 而本文虚拟旅游流市域流入中 AICc 值相差为 12.6169, 大约为 4 倍, 而虚拟旅游流市域流出 AICc 值相差 24.3495, 大约为 8 倍, 同样证明了 GWR 模型拟合结果优于 OLS 拟合。

2.2.1 虚拟旅游流流入市域影响因素分析

对 α_1 、 α_2 和 α_3 以及 t 值进行可视化处理, 可以看出其空间分布特征存在明显差异, 对市域间虚拟旅游流影响存在正、负相关, 各市域解释变量系数参数估计结果不同, 符合现实情况。

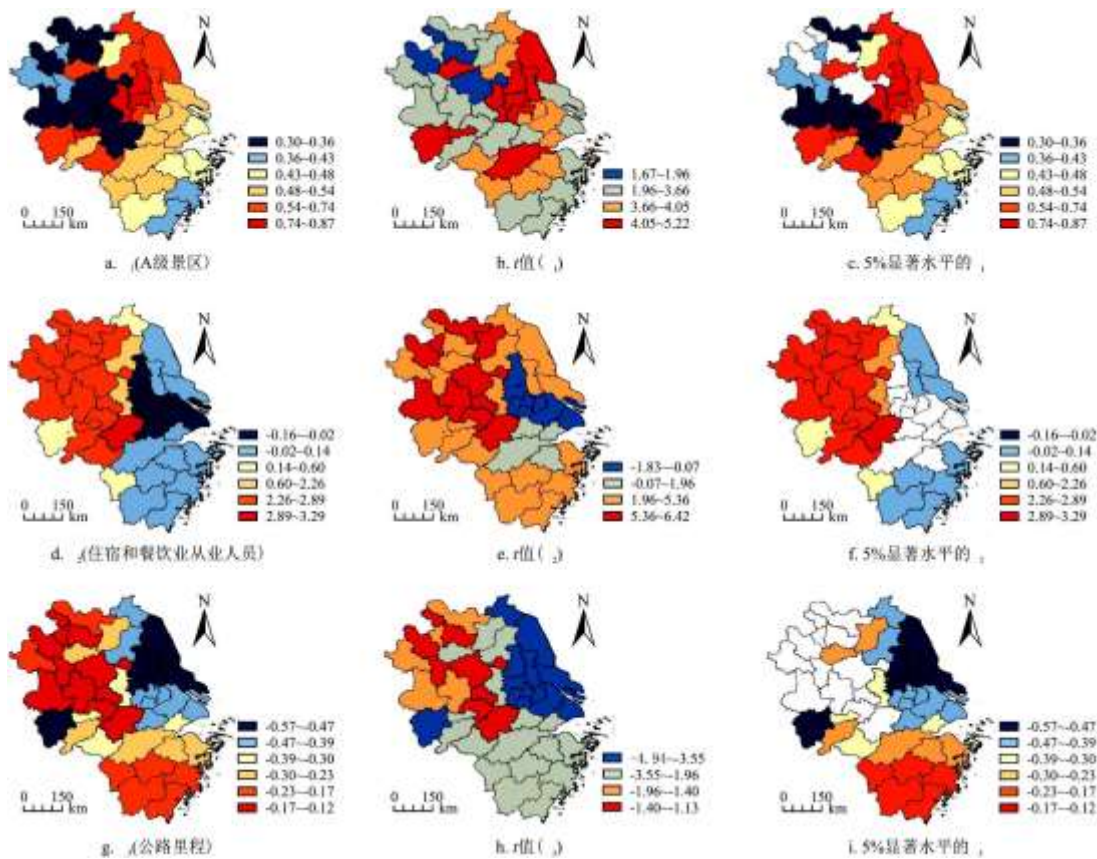


图 2 流入市域地理加权回归结果

(1) α_1 系数总体上呈现出东北高、西北低趋势, 市域 A 级旅游景点对整个长江三角洲三省一市流入虚拟旅游流都成正相关, 对流入虚拟旅游流流动有推动作用, 可以看出人们在进行网络检索时, 仍然优先考虑的是旅游景点所代表的旅游资源。从显著性上看, 有 38 个市域可以通过 5% 水平的显著性检验, 只有安徽的安庆、亳州、阜阳未通过 5% 水平的显著性检验。分析结果表明 A 级旅游景点数量对整个长江三角洲流入虚拟旅游流有显著的正向影响, 尤其是对于江苏南部、浙江北部的市域而言, 以南京、扬州、苏州、杭州组成了四核心的流入虚拟旅游流流量网络, 同时以其自身的高质量旅游形象与旅游资源, 加之互为旅游集散中心, 吸引了大量的流入虚拟旅游流, 进一步增强了 A 级旅游景点对此部分市域的影响。未能通过 5% 水平的显著性检验的安徽 3 市则是由于旅游资源开发水平仍有待提高, 导致 A 级旅游景点对其影响并不显著。

(2) α_2 系数值与 α_1 相反, 高数值区域集中于西部地区, 而东部地区则为低数值区域, 整体上为正相关, 江苏部分市域成负

相关,负相关系数绝对值远小于正相关系数绝对值。从显著性上看,有10个市域未能通过5%水平的显著性检验,江苏中、南部,上海以及浙江北部显著性不明显,而住宿和餐饮业从业人员这一解释变量对长江三角洲西部地区有着强烈的正相关作用。长江三角洲西部地区经济发展水平以及旅游业发展水平与东部、南部相差较大,积极发展住宿和餐饮业这一作为旅游资源的重要影响因素对整个长江三角洲西部地区流入虚拟旅游流有较强的引导效果,在一定程度上消除了与东部地区差距过大所带来的马太效应。长江三角洲东部、南部的江苏中、南部、上海以及浙江北部,由于经济发达、旅游业繁荣,住宿和餐饮业发展程度较高,游客的食宿选择便捷多样,旅游者出游时不会对此方面产生担忧,也并非首要考虑短板,因此住宿和餐饮业的进一步发展,对流入虚拟旅游流的吸引作用有限,呈现不相关或负相关。

(3) α_3 系数值与 α_2 类似,高数值区域同样集中于长江三角洲西部和南部地区,低数值区集中于东北部,公路里程这一解释变量对整个长江三角洲均产生负相关影响。从显著性上看,共有29个市域通过了5%水平的显著性检验,是三个解释变量中通过检验数量最少的变量,可以看出在高铁、飞机以及航运一系列完善交通通达性工具出现的现在,尤其是区域通达性较好的长江三角洲地区,公路里程这一解释变量对流入虚拟旅游流的影响远小于前两个变量。未通过5%水平显著性检验的市域大部分位于安徽省境内,这是由于安徽省与江苏、上海、浙江相比较,旅游业发展水平存在滞后性,境内公路里程的逐渐增加并不能对旅游业产生直接的带动作用,需要配合其他配套旅游设施的不断升级进而刺激虚拟旅游流的流入。而通过的29个市域集中于江苏、上海以及浙江,且呈现较弱的负相关,说明在交通网络较为完善的江苏、浙江以及市内交通系统发达的上海,游客的出行选择并不只依靠公路里程,高铁、飞机等交通方式也十分常见,公路里程的选择首要性被削弱,其承担游客出游的功能性被替代,所以公路里程的增加对扩大虚拟旅游流流量并未能产生有效引导,呈现微弱的负相关。

2.2.2 虚拟旅游流流出市域影响因素分析

对 β_1 、 β_2 和 β_3 以及 t 值进行可视化处理,进一步分析私人汽车拥有量、社会消费品零售总额、互联网宽带用户接入量对长江三角洲流出虚拟旅游流影响的空间差异。

(1) β_1 系数值具有南北较高、中间低的空间分布特征,除少数市域外,大部分市域的私人汽车拥有量与流出虚拟旅游流成正相关,安徽省为整个长江三角洲 β_1 数值最高的省份,而江苏省数值则最低。从显著性上看,共有11个市域未能通过5%水平的显著性检验,大部分集中于江苏省,从苏南的南京一直向上延伸至苏北的连云港。结果表明,苏北地区经济与苏南和浙江相比存在一定差距,私人汽车拥有量的数量也相对较少,故部分市域未能通过显著性检验。作为游客集散中心的扬州、南京、杭州,交通基础设施完善,航空、铁路运输便利,出游方式多样化且便捷性较高,大大削弱了居民的汽车出游选择可能性。同样作为旅游集散中心的苏州是进入上海的主要通道,凭借自身旅游资源优势同时依托上海形成了整体效应,旅游者可选择苏州-上海之间的双向自驾游,体验高质量的游憩活动,强化了驾车出游的可能性。而安徽经济发展水平相对较差,居民私人汽车拥有量低于长江三角洲其余两省一市,但私人汽车拥有量表现出显著的正相关,表明居民私人汽车拥有量的微小增加会对此部分市域居民的出游欲望产生强烈的刺激,极大增加居民的出游可能,因而私人汽车拥有量对此部分市域流出虚拟旅游流的影响显著。

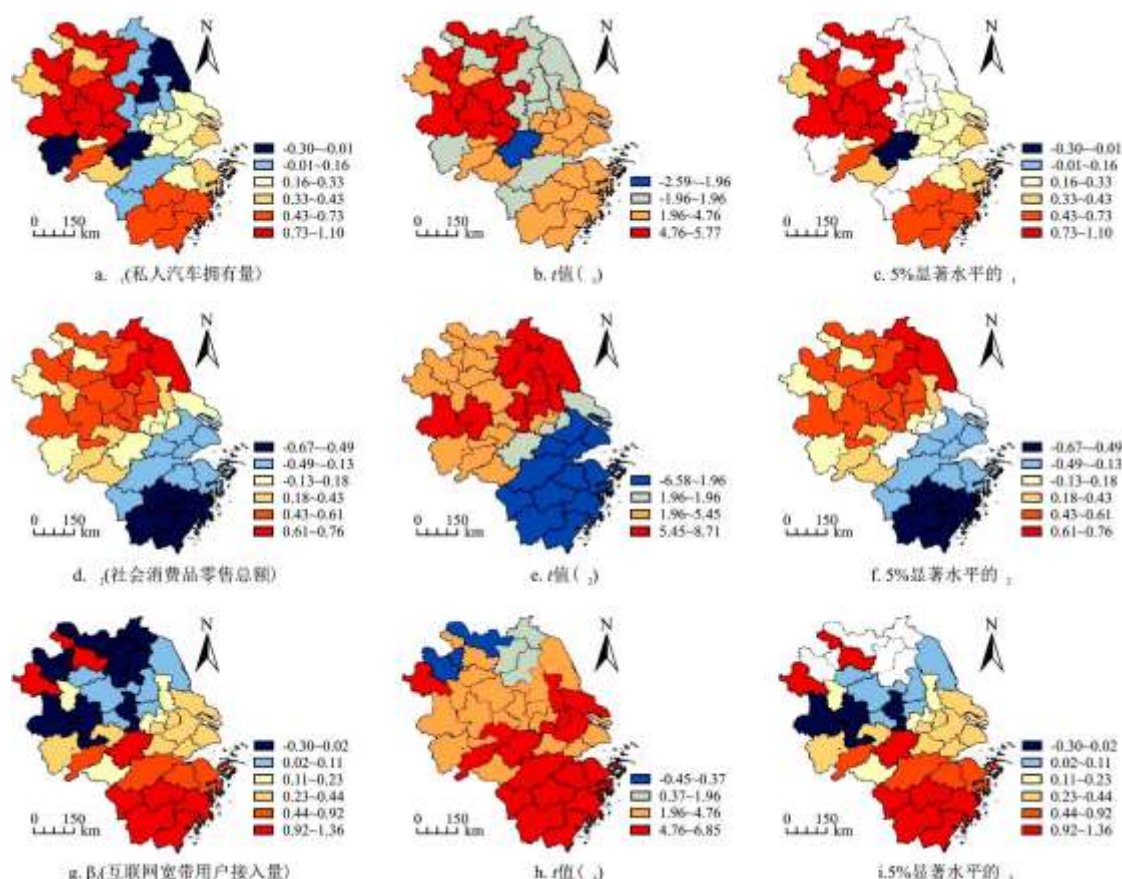


图3 流出市域地理加权回归结果

(2) β_2 系数值由北向南逐渐降低, 社会消费品零售总额对长江三角洲北部和中部为正相关, 对南部则为负相关。从显著性来看, 有 38 个市域通过显著性检验, 除长江三角洲中部的黄山、无锡、南通 3 市外, 其余市域显著性较好, 社会消费品零售总额对整个长江三角洲影响极为显著。可以看出, 长江三角洲北部的安徽省经济稳步发展, 社会消费品零售总额逐渐增加, 居民对于高层次的追求不断增强, 居民经济能力的增强也使得旅游这一放松身心的游憩活动成为了备选项, 人们有了一定的出游意愿和将出游意愿转化为实际出游的经济能力, 对于虚拟旅游流流出具有较强的推动作用。而江苏省内经济相对落后的苏北地区则同样因为经济的快速发展催生了人们的出行欲望, 带来了大量的虚拟旅游流流出。苏南地区, 浙江和上海则是由于经济较为发达, 除旅游外, 居民还可选择其他大量的社会活动和消费行为以达到放松身心, 愉悦自身的需要, 社会消费品零售总额的增加被其他消费所占据。同时, 旅游已被作为一种定期的游憩需要, 一种固定的开支而存在, 并不会有明显增加。部分具有经济实力的居民还会进行出国旅游, 而非在国内进行游玩, 也在一定程度上弱化了社会消费品零售总额对虚拟旅游流增加的有效刺激, 社会消费品零售总额与虚拟旅游流之间的转化率较低, 所以此部分市域社会消费品零售总额的增加并不能对虚拟旅游流的流出带来较大的刺激作用, 只有微弱的负相关或不相关。

(3) β_3 系数值与 β_2 相反, 由南向北逐渐降低, 高数值市域主要集中在江苏南部, 上海以及浙江, 而低数值市域多集中于江苏北部以及安徽。从显著性看, 有 6 个市域未通过 5% 水平的显著性检验, 多集中于安徽北部和江苏北部。浙江和上海显著性较好, 35 个通过 5% 水平的显著性检验的市域多位于此区域。可以看出, 经济发展较差的安徽和苏北地区互联网宽带用户接入量对流出虚拟旅游流影响程度不高, 网络用户数量的增加并不能代表流出虚拟旅游流的增加, 依旧需要区域经济的发展进而促进居民的出游需求提高, 而经济水平较高的江苏南部、上海、浙江, 互联网的普及与发展对于流出虚拟旅游流有巨大的推动和引导作用, 尤其是浙江, 表现得最为明显。证明在 Web2.0 时代, 尽管区域经济发展程度存在差异性, 人们更愿意通过网络进行旅游检索从而决定出游与否, 经济越发达区域, 网络旅游检索数越易变现为实际出游人数。总体而言, 互联网宽带用户接入量对长

江三角洲流出虚拟旅游流的影响程度不如社会消费品零售总额但优于私人汽车拥有量。

3 结论与讨论

3.1 结论

基于百度指数平台,借助于软件爬取 2018 年长江三角洲虚拟旅游流检索数据,构建 O-D 虚拟旅游流网络,探讨长江三角洲虚拟旅游流流动路径与流动特征,并进一步从市域虚拟旅游流流入和流出两个角度讨论虚拟旅游流流量和 A 级旅游景区、公路里程、社会消费品零售总额等社会经济因素之间的相互关系,主要结论归纳如下:

(1)长江三角洲三省一市共存在虚拟旅游流路径 1640 条,高流量路径主要存在于江苏南部、浙江、安徽东南部以及上海之间,形成合肥、南京、苏州、杭州、上海的五核结构网络;次流量路径围绕五核结构形成二级网络,长江三角洲北部旅游流流动路径值最小。

(2)市域总流量呈东高西低的分布趋势,流量最高的市域为上海、杭州、苏州、南京,而低流量市域全部位于长江三角洲北部安徽省境内;虚拟旅游流市域流动较大,41 个市域中有 31 个都呈现净流入或净流出特征,净流入率最高的市域为黄山,流出率达到了-80.18%,净流出率最高的为阜阳,其流出率为 50.68%。

(3)从 GWR 结果来看,对于流入虚拟旅游流而言,A 级景区数量、住宿和餐饮业从业人数以及公路里程可以较好的解释流入虚拟旅游流市域分布。其中 A 级景区对流入虚拟旅游流影响最为显著,在长江三角洲东部和南部表现较为强烈;住宿和餐饮业的发展对长江三角洲西北部虚拟旅游流的流入具有推动力;公里里程对整个长江三角洲流入虚拟旅游流影响最小,尤其是安徽省,影响趋近于无。

(4)根据 GWR 模型分析可以看出,私人汽车拥有量、社会消费品零售总额以及互联网宽带用户接入量对流出虚拟旅游流有较强的影响。私人汽车拥有量的不断增加可吸引安徽省游客出游率不断增加,社会消费品零售总额对长江三角洲影响最为显著,长江三角洲虚拟旅游流的流出量完全受制于区域经济发展水平,在长江三角洲北部表现最为明显。互联网宽带用户接入量则在经济发达地区对虚拟旅游流的流出有着良性的导向作用,经济越发达的区域,人们越倾向于利用网络抓取旅游信息,追求高质量旅游出行。

3.2 讨论

尽管以百度指数为基础,利用 GWR 软件对长江三角洲虚拟旅游流的流动特征以及影响旅游流流动的解释变量的空间关系探讨具有一定的意义,但仍有以下问题需要注意:(1)本文中 GWR 模型主要是对虚拟旅游流流入流出的面域进行了探讨,对虚拟旅游流流动特征的研究需要进一步深化分析。(2)百度指数可以避免传统的抽样调查数据的随机性,但是由于网络空间信息的前兆性^[44,51],其与现实世界旅游流存在一定差异,需要在后续研究中,分析如何消除此类差异并保证其正向引导性。(3)本文从虚拟旅游流的流入和流出两个视角分别对定义为客源地(Origin)与目的地(Destination)的市域进行了影响因素的分析,但是缺乏对流入与流出两方面角度结合的分析。对于某一市域而言,在 O-D 矩阵中既是旅游客源地又是旅游目的地,如何判断虚拟旅游流流入与流出对自身的双向影响,何种因素对既是旅游客源地又是旅游目的地的市域影响最为显著,也需要更深的研究。上述问题的解决有待于数据的进一步挖掘和研究的进一步深化,这也是下一步研究的方向。

参考文献:

[1]秦静,李郎平,唐鸣镝,等.基于地理标记照片的北京市入境旅游流空间特征[J].地理学报,2018,73(8):1556-1570.

-
- [2]汪德根, 陈田, 陆林, 等. 区域旅游流空间结构的高铁效应及机理——以中国京沪高铁为例[J]. 地理学报, 2015, 70(2): 214-233.
- [3]王永明, 王美霞, 吴殿廷, 等. 基于 ZINB 模型的中国省域间入境旅游流影响因素[J]. 经济地理, 2018, 38(11): 234-240.
- [4]涂玮, 黄震方, 方叶林. 基于网络团购的虚拟旅游流空间差异及动力机制研究[J]. 地域研究与开发, 2013, 32(4): 84-89.
- [5]石进. 长三角八个城市国内虚拟旅游流季节性波动研究[D]. 上海师范大学, 2013.
- [6]殷晶. 基于社会网络理论的沪宁杭虚拟旅游流对比分析[D]. 上海师范大学, 2012.
- [7]蒋依依, 温晓金, 刘焱序, 等. 2001-2015 年中国出境旅游流位序规模演化特征[J]. 地理学报, 2018, 73(12): 2468-2480.
- [8]王金莹, 吴晋峰, 唐澜. 基于 SNA 的中国入境欧洲旅游流网络性质和结构特征研究[J]. 人文地理, 2013, 28(6): 147-153.
- [9]马耀峰, 林志慧, 刘宪锋, 等. 中国主要城市入境旅游网络结构演变分析[J]. 地理科学, 2014, 34(1): 25-31.
- [10]刘宏盈, 韦丽柳, 张娟. 基于旅游线路的区域旅游流网络结构特征研究[J]. 人文地理, 2012, 27(4): 131-136.
- [11]刘法建, 张捷, 陈冬冬. 中国入境旅游流网络结构特征及动因研究[J]. 地理学报, 2010, 65(8): 1013-1024.
- [12]李亚娟, 曾红春, 李超然, 等. 山地民族地区旅游流网络结构及流动方式研究——以恩施州为例[J]. 长江流域资源与环境, 2019, 28(7): 1612-1622.
- [13]姚梦汝, 陈焱明, 周桢津, 等. 中国—东盟旅游流网络结构特征与重心轨迹演变[J]. 经济地理, 2018, 38(7): 181-189.
- [14]王金莹, 吴晋峰, 唐澜, 等. 亚洲入境旅游流地理分布及网络结构特征分析[J]. 资源科学, 2013, 35(8): 1701-1709.
- [15]蔚海燕, 戴泽钊, 许鑫. 上海迪士尼对上海旅游流网络的影响研究——基于驴妈妈游客数字足迹的视角[J]. 旅游学刊, 2018, 33(4): 33-45.
- [16]JIN X, et al. Impact of crisis events on Chinese outbound tourist flow: A framework for post-events growth[J]. Tourism Management, 2019, 74: 334-344.
- [17]FRASH R E, JR, J E BLOSE. Serious leisure as a predictor of travel intentions and flow in motorcycle tourism[J]. Tourism Recreation Research, 2019, 44(4): 516-531.
- [18]GAO J, et al. Tourism border-making: A political economy of China's border tourism[J]. Annals of Tourism Research, 2019, 76: 1-13.
- [19]ZHANG Y, C FINDLAY. Air transport policy and its impacts on passenger traffic and tourist flows[J]. Journal of Air Transport Management, 2014, 34: 42-48.

-
- [20]CAMPISI T,et al.The importance of assessing the level of service in confined infrastructures:Some considerations of the Old Ottoman Pedestrian Bridge of Mostar[J].Applied Sciences-Basel,2019,9(8).
- [21]杨兴柱, 顾朝林, 王群. 旅游流驱动力系统分析[J]. 地理研究, 2011, 30(1):23-36.
- [22]刘军胜, 马耀峰. 基于发生学与系统论的旅游流与目的地供需耦合成长演化与驱动机制研究——以西安市为例[J]. 地理研究, 2017, 36(8):1583-1600.
- [23]史春云, 张捷, 尤海梅, 等. 四川省旅游区域核心—边缘空间格局演变[J]. 地理学报, 2007(6):631-639.
- [24]吴晨. 网络视角下目的地城市旅游流空间结构与空间自相关效应研究[D]. 南京大学, 2017.
- [25]YANG Y,et al.Modeling sequential tourist flows:Where si the next destination?[J]. Annals of Tourism Research, 2013, 43:297-320.
- [26]VITOVA P,et al.Determinants of tourism flows in Small Island Developing States (SIDS) [J]. Island Studies Journal, 2009, 14(2):3-22.
- [27]XU L,et al.Modelling international tourism flows to China:A panel data analysis with the gravity model[J]. Tourism Economics, 2019, 25(7):1047-1069.
- [28]GHALIA T,et al.Institutional quality,political risk and tourism[J].Tourism Management Perspectives, 2019, 32.
- [29]SEYFI S,C M HALL.Sanctions and tourism:effects,complexities and research[J].Tourism Geographies, 2019.
- [30]PETREVSKA B,et al.Prevaling motives for creating Ottoman heritage packaged tours in Macedonia[J]. Journal of Tourism And Cultural Change, 2019.
- [31]CALDERON-VARGAS F,et al.Sustainable tourism and renewable energy:Binomial for local development in cocachimba, amazonas, peru[J]. Sustainability, 2019, 11(18).
- [32]FROLOVA E,et al.Mediatization of tourism:New trends and tourist attraction in the Online Space[J]. Mediaobrazovanie-Media Education, 2019(2):249-257.
- [33]CHOI J,et al.Outcomes of destination marketing organization website navigation:the role of telepresence [J]. Journal of Travel & Tourism Marketing, 2016, 33(1):46-62.
- [34]CHAO D,et al.Experimental study on tourist satisfaction using participatory simulation in a virtualen vironment[J]. Springerplus, 2013, 2.
- [35]GALE T.Urban beaches,virtual worlds and ‘The End of Tourism’ [J]. Mobilities, 2009, 4(1):119-138.

-
- [36]程昌秀,史培军,宋长青,等.地理大数据为地理复杂性研究提供新机遇[J].地理学报,2018,73(8):1397-1406.
- [37]庞瑞秋,腾飞,魏冶.基于地理加权回归的吉林省人口城镇化动力机制分析[J].地理科学,2014,34(10):1210-1217.
- [38]赵先超,袁超,张茜茜,等.基于 GWR 模型的长株潭地区县级乡村重构?时空格局及城镇驱动力研究[J].农业现代化研究,2019,40(3):498-506.
- [39]阚博颖,濮励杰,徐彩瑶,等.基于 GWR 模型的南京主城区住宅地价空间异质性驱动因素研究[J].经济地理,2019,39(3):100-107.
- [40]马丽君,孙根年,黄芸玛,等.城市国内客流量与游客网络关注度时空相关分析[J].经济地理,2011(4).
- [41]汪秋菊,黄明,刘宇.城市旅游客流量—网络关注度空间分布特征与耦合分析[J].地理与地理信息科学,2015,31(5):102-106.
- [42]虞洋,宋周莺,史坤博.基于百度指数的中国省域间信息联系网络格局及其动力机制[J].经济地理,2019,39(9):147-155.
- [43]阮文奇,张舒宁,李勇泉,等中国赴泰旅游需求时空分异及其影响因素[J].旅游学刊,2019,34(5):79-92.
- [44]李山,邱荣旭,陈玲.基于百度指数的旅游景区络空间关注度:时间分布及其前兆效应[J].地理与地理信息科学,2008(6):102-107.
- [45]2018 年江苏省国民经济和社会发展统计公报[J].统计科学与实践,2019(3):22-27,60.
- [46]2018 年浙江省国民经济和社会发展统计公报[J].统计科学与实践,2019(3):3-10.
- [47]安徽省统计局国家统计局安徽调查总队.安徽省 2018 年国民经济和社会发展统计公报~([1])[N].安徽日报,2019-02-28(006).
- [48]2018 年上海市国民经济和社会发展统计公报[J].统计科学与实践,2019(3):11-21.
- [49]黄泰.长三角城市群旅游流潜力格局演变及其影响因素[J].资源科学,2016,38(2):364-376.
- [50]靳诚,徐菁,黄震方,等.基于高速公路联网收费数据的江苏省交通流动特征与影响因素[J].地理学报,2018,73(2):248-260.
- [51]刘军胜,马耀峰,李振亨.1997~2010 年中部六省入境旅游流集散时空动态分析[J].地理科学,2013,33(4):450-456.