

基于网络数据的乡村旅游热点识别及成因分析 ——以江苏省为例

荣慧芳^{1,2}, 陶卓民¹

(1. 南京师范大学地理科学学院, 南京 210023; 2. 池州学院资源环境学院, 池州 247000)

摘要: 基于旅游网站数据, 提出一种乡村旅游热点识别方法, 以江苏省为例, 运用趋势面、核密度、热点分析等方法探寻乡村旅游冷热格局及演变特征, 并借助地理探测器揭示其时空演化的影响因素。结果表明: (1) 江苏省乡村旅游热度的时间演化规律明显, 年际变化呈“S”型演变轨迹; 季节变化呈“三峰四谷”特征。(2) 江苏省乡村旅游热度的空间结构在 2009—2017 年间经历了“单核—双核—三核”的演化过程, 但其空间异质性依然显著, 整体表现为“南高北低, 东高西低”的差异特征。(3) 影响因素在不同时段的影响强度各异, 交通条件、服务能力一直是主导因素, 经济水平对乡村旅游发展早期的热度提升有显著影响, 资源禀赋的显著性趋于下降, 生态环境和政府导向的影响力逐渐增强。基于网络数据的乡村旅游热点识别为乡村旅游定量研究提供了新的研究视角, 对乡村旅游资源开发与区域合作具有重要的指导意义。

关键词: 大数据; 热点识别; 演化; 成因; 江苏省

乡村, 是人们回归自然、记忆乡愁的重要载体, 凭借其优美的自然环境、独特的地方文化, 日益受到游客们的青睐。资料显示, 2017 年, 乡村旅游人次达 25 亿, 占据国内旅游总人次的一半以上。尤其是乡村振兴战略提出以来, 乡村旅游作为旅游产业向传统农业延伸的新尝试, 对于优化乡村产业结构, 促进乡村经济复苏, 推动乡村振兴战略的深入实施具有重要的现实意义^[1]。乡村旅游不仅成为地方政府、旅游企业关注的焦点, 也是国内外旅游研究的热点话题。与此同时, 随着互联网的发展与普及, 旅游网站、社交软件等网络平台已成为游客查询出游信息、辅助出游决策、分享出游体验的重要工具, 由此产生的旅游大数据为乡村旅游研究提供了数据来源和方法支撑。因此, 在乡村旅游地位不断提升的新形势下, 基于网络数据对乡村旅游时空演化规律及其驱动机制进行研究, 不仅有助于拓展乡村旅游研究视野, 而且对于理解乡村旅游发展机制, 优化产业布局具有重要的实践价值。

纵观国内外相关文献, 乡村旅游学术研究积累了丰硕的成果。国外研究相对成熟与细化, 研究内容已从乡村旅游概念^[2]、旅游者动机^[3]、发展模式和策略研究^[4]深入到乡村旅游影响^[5]、可持续发展^[6]、利益相关者^[7]等方面。国内研究则主要以国家政策为导向探讨乡村旅游发展中的新问题, 研究内容主要涉及旅游扶贫^[8,9]、旅游地社区参与^[10]、产业融合与生态文明^[11,12]、旅游城镇化^[13,14]、空间分异与影响因素^[15-17]等方面。其中乡村旅游时

收稿日期: 2019-05-24; 修订日期: 2019-08-01

基金项目: 国家自然科学基金项目 (41571139, 51778002); 江苏省研究生科研与实践创新计划项目 (KYCX19_0762)

作者简介: 荣慧芳 (1984-), 女, 河南永城人, 博士研究生, 研究方向为旅游地理与旅游规划。

E-mail: ronghuifang96@126.com

通讯作者: 陶卓民 (1963-), 男, 江苏苏州人, 博士, 教授, 博士生导师, 研究方向为旅游地理、旅游经济与市场。E-mail: taozm@tom.com

空分异研究探寻乡村旅游发展的空间格局、演变过程、作用机制，为乡村旅游产业布局、发展规划提供决策参考，成为旅游地理学者重要的研究领域。梳理文献不难发现，由于乡村旅游景点分布范围广、规模等级低，统计数据的缺失成为乡村旅游定量研究的一大难题，目前乡村旅游时空格局的研究过度关注乡村旅游景点的数量多寡、聚散程度等位置信息的表达，忽视了对乡村旅游资源质量、热度等属性信息的研究。

大数据时代的到来为乡村旅游定量研究提供了更多可能，由此引发的关于旅游大数据的研究成果较多。研究切入点大体归为三个方面：一是运用网络关注度数据进行旅游流时空格局分析。例如，张子昂等^[18]以南京钟山景区为例，基于微博数据对游客活动的时空演变特征进行了实证研究；马丽君等^[19]通过百度指数揭示了北京居民对5A级景区旅游需求的时空格局及成因；孙焯等^[20]以百度指数表征旅游者的网络关注度，预测了三清山景区的一日游客量情况。二是借助地理标记照片研究游客流动轨迹。如秦静等^[21]探索了基于地理标记照片的北京市游客空间分布格局和流动特征。三是基于网络文本的旅游体验与感知研究。如李萍等^[22]通过挖掘网络点评信息研究了北京市典型城市旅游社区的形象感知；刘超等^[23]以黄山为例，基于网络文本分析了游客对山岳型景区的感知体验。综合来看，前人研究多以百度指数或微博数据等单一数据源为研究对象，样本量有限，可能导致一般性和小规模景点无法测度。此外，研究的时间尺度多在一年（含）以内，很难全面客观地反映出旅游流的真实分布状态，而且案例地多选择热点旅游城市或知名旅游景区，而针对乡村旅游的研究尚显不足。

基于以上分析，本文按照乡村旅游“热度测算→冷热点识别→成因解析”的研究思路，以江苏省496家星级乡村旅游景点为研究样本，利用网络数据对其在2009—2017年间的乡村旅游热度进行测度，在此基础上，借助GIS空间分析技术识别乡村旅游的冷热格局及演化过程。继而，运用地理探测器模型揭示时空格局演化背后的影响因素及作用机制。本文旨在拓展乡村旅游定量研究的深度和广度，为乡村旅游时空格局研究提供思路和方法借鉴，并为乡村旅游资源的合理开发及相关政策的完善提供理论指导。

1 研究方法与数据来源

1.1 研究区概况

江苏省地处我国东部沿海地区，社会经济基础雄厚，历史文化悠久，是长三角重要的旅游客源地和目的地。优越的地理区位，发达的经济条件，丰富的旅游资源，为江苏省乡村旅游发展提供了有利条件，使之成为我国乡村旅游发展较早和较成熟的地区之一。截至2017年底，江苏省共有星级乡村旅游景点496家，涵盖了农事体验类、休闲观光类、农业科技类、乡村遗产类、新农村示范类等多种类型。2017年全年乡村旅游接待游客2.87亿人次，营业总收入近900亿元，带动直接从业人数近42万人。以江苏省为研究案例，实证探讨乡村旅游冷热格局演化规律及其影响因素，可为乡村旅游产业布局及可持续发展提供科学指导，也可为其他类似地区的乡村旅游发展提供借鉴和参考。

1.2 数据来源与处理

1.2.1 数据来源说明

本文所用数据涉及三大类：（1）旅游网站数据，主要用于乡村旅游热度测评。在百度指数中输入知名旅游网站名称，近五年整体日均值排名前三的分别是携程、去哪儿、

途牛,另外考虑到大众点评网景点覆盖范围广,样本量丰富,最终选择携程、去哪儿、途牛、大众点评网四家网站作为数据来源网站。(2)统计数据,主要用于影响因子探测,来源于各地统计年鉴和统计公报。(3)矢量数据,用于ArcGIS空间分析及可视化。由国家地球系统数据共享平台——长江三角洲科学数据中心下载,数据年份为2015年。

1.2.2 数据采集与处理

(1)样本选择与可视化。基于样本的典型性、代表性以及数据的可获取性,以江苏省文化与旅游厅官方网站截至2017年12月公布的496家星级乡村旅游景点为研究样本,使用百度地图API实现样本点地理坐标拾取,并运用ArcGIS 10.2软件对所有样本点进行可视化表达。需要指出的是,因江苏省星级景点评定工作开始于2009年,故研究时段确定为2009年1月至2017年12月。

(2)网络数据采集。使用Python程序抓取景点的网络数据,具体采集工作分为三步。第一步,数据抓取。以景点点评数、浏览数、点赞数等热度信息为目标数据,运用PyCharm和Python 3.7编写爬虫,构造get请求获取各网站目标数据的源代码,根据源代码中链接循环进入目标数据具体页面,分析其源代码并截取目标数据详情。第二步,数据清洗。删除研究区范围外同名景点数据、同一ID重复数据和无效数据,存入表格文件中,最终获取有效样本49952条。第三步,数据汇总。分景点、分类型、分年份、分月份,对乡村旅游热度信息进行统计汇总,具体内容如表1所示。

表1 景点网络数据采集内容

Table 1 Collected data content of attraction in website

网站	现实热度信息		潜在热度信息	
携程网	游记数	景点点评数	游记评论数	游记浏览数
去哪儿网	游记数	游记评论数	游记点赞数	游记浏览数
途牛网	游记数	点评数	游记收藏数	游记浏览数
大众点评网	默认点评数	全部点评数	点赞数	回应数

(3)影响因素数据获取。县域是乡村旅游发展的重要载体,乡村旅游是振兴县域经济的重要路径,对其乡村旅游发展驱动机制进行探讨具有重要的现实意义。因此,选择以县域为研究单元解析乡村旅游热度空间差异形成的影响因素,数据主要来源于2009—2017年各县市国民经济与社会发展统计公报、统计年鉴等。需要说明的是,考虑到乡村旅游的发生地域,参考相关研究的处理方法^[24],将各市辖区合并为“市区”、与“县级市”和“县”并行作为县域单位进行研究,以便有效衔接行政区划的调整与旅游宏观政策的制定。

1.3 研究方法

1.3.1 景点热度测评

现有景点热度测评主要针对城市、大型景区等知名旅游目的地,通过百度指数、微博签到等数据评估景点热度,但由于乡村旅游景点具有规模等级小、种类多样的特点,上述测度方法很难全面客观地反映乡村旅游景点的热度信息。基于此,本文在考虑景点等级参数的基础上,选取知名旅游网站为数据源,参考张霞等^[25]的景观热度评价方法,构建乡村旅游景点热度测评模型,测度公式如下:

$$H_i = L_i \sum_{j=1}^4 (w_1 R_{ij} + w_2 P_{ij}) \quad (1)$$

式中： H_i 为第*i*个景点的热度值； L_i 为景点等级； R_{ij} 为第*i*景点在第*j*个网站所呈现的现实热度，用游记数、点评数表达； P_{ij} 为第*i*景点在第*j*个网站所呈现的潜在热度，用点赞数、收藏数表达； w_1 和 w_2 分别为对应的权重，考虑到现实热度的实际影响效应远大于潜在热度，取 w_1 为0.8， w_2 为0.2。

1.3.2 趋势面分析

趋势面分析法通过全局多项式将空间采样点数值用数学函数拟合，将二维空间的采样点数据转换成三维可视化平滑曲线，展示地理要素在空间上的变化趋势。本文利用趋势面表达江苏省乡村旅游热度的空间分异趋势。设 $z_i(x_i, y_i)$ 为第*i*个地理要素的真实观测值， $T_i(x_i, y_i)$ 为趋势面拟合值，则有如下计算公式^[26]：

$$z_i(x_i, y_i) = T_i(x_i, y_i) + \varepsilon_i \quad (2)$$

式中： (x_i, y_i) 为地理坐标； ε_i 为残差，即真实值与拟合值的偏差。

1.3.3 核密度估算

核密度估算法对数据分布不受任何假定条件的限制，是一种从数据样本本身出发研究数据分布特征的方法，它具有表达直观、概念简洁和易于计算的优点，因而成为空间分析中运用最为广泛的非参估计方法。其一般形式是^[27]：

$$f(s) = \sum_{i=1}^n \frac{k}{\pi r^2} \left(\frac{d_{is}}{r} \right) \quad (3)$$

式中： $f(s)$ 是位置*S*的密度； n 为样本数（个）； d_{is} 是点*i*到*S*的距离（m）； k 为 d_{is} 与*r*比值的核函数；*r*为带宽（m）。要注意的是，带宽*r*的选择会对估算结果产生较大影响，需根据现实情况选择不同带宽值进行比较，直至得到与现实情况较吻合的核密度曲面。

1.3.4 热点分析

热点分析（Getis-Ord General G_i^* ），通过计算每个要素的 G_i^* 值，可以识别具有统计显著性的高值和低值空间聚类，进而判断出地域单元属于热点区还是冷点区。本文主要借助热点分析识别乡村旅游冷热区域。具体公式^[28]可表示为：

$$G_i^* = \frac{\sum_{j=1}^n w_{ij} x_j - \bar{X} \sum_{j=1}^n w_{ij}}{S \sqrt{\frac{n \sum_{j=1}^n w_{ij}^2 - \left(\sum_{j=1}^n w_{ij} \right)^2}{n-1}}}, \bar{X} = \frac{\sum_{j=1}^n x_j}{n}, S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2} \quad (4)$$

式中： x_j 是要素*j*的属性值； w_{ij} 为要素*i*和要素*j*的空间权重； n 为要素数（个）； \bar{X} 是平均值； G_i^* 统计是*Z*得分，如果 G_i^* 值高且显著，则表示有一个高值的空间聚类，称之为热点，反之，如果 G_i^* 值低且为负、显著，则表示有一个低值的空间聚类，称之为冷点。

1.3.5 地理探测器

地理探测器模型通过探测两个变量在空间分布上的一致性，判断两变量之间存在的因果关系。由于地理探测器无需考虑数据标准化及正态分布等前提条件，可有效避免传统数学计量模型的局限性。具体模型^[29]如下：

$$P_{D,U} = 1 - \frac{1}{n\sigma^2} \sum_{i=1}^m n_i \sigma_i^2 \quad (5)$$

式中： $P_{D,U}$ 为要素 D 对乡村旅游的影响力； n 为研究区样本数（个）， $n=55$ ； σ^2 和 σ_i^2 分别为所有区域与次级区域的方差； m 为次级区域数（个）。 $\sigma^2 \neq 0$ ，模型成立。 $P_{D,U}$ 的取值区间为 $[0, 1]$ ， $P_{D,U}$ 的值越大，说明要素 D 对乡村旅游热度的影响力就越强，反之则越弱。

2 结果分析

2.1 时间演化

2.1.1 时序变化特征

运用乡村旅游景点热度测评模型计算2009—2017年江苏省乡村旅游景点热度值。2009—2017年间江苏省乡村旅游热度值呈“S”型变化轨迹（图1），大体可划分为缓慢发展、快速发展与稳定发展三个阶段。江苏省乡村旅游起步较早但前期产品层次较低，发展速度较慢。2012年开始的村庄整治计划为乡村旅游发展提供了契机，2013—2015年间江苏省乡村旅游进入快速上升期。尤其2015年底印发的《江苏省乡村旅游发展三年行动计划（2016—2018）》为乡村旅游产业发展指明了方向，江苏省休闲农业与乡村旅游现已进入稳定发展阶段。

2.1.2 季节变化规律

从江苏省乡村旅游热度的月统计图（图2）可以看出，乡村旅游热度的季节变化总体上呈“三峰四谷”特征，4月、8月和10月构成“三峰”；1月、6月、9月和12月形成“四谷”。总体看来，江苏省乡村旅游旺季在4月、8月和10月，淡季为1月、6月、9月和12月。4月前后正值春季花开时节，是乡村旅游观光游的最佳选择时段；10月是秋季采摘的黄金时期，再加上清明、国庆闲暇时间的增加，对乡村旅游出游起到很大的促进

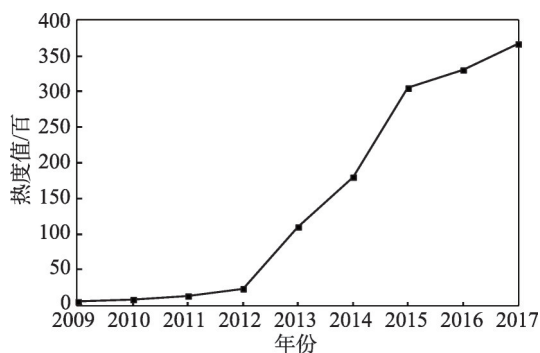


图1 2009—2017年江苏省乡村旅游热度的年度变化
Fig. 1 Annual count of rural tourism popularity in Jiangsu province from 2009 to 2017

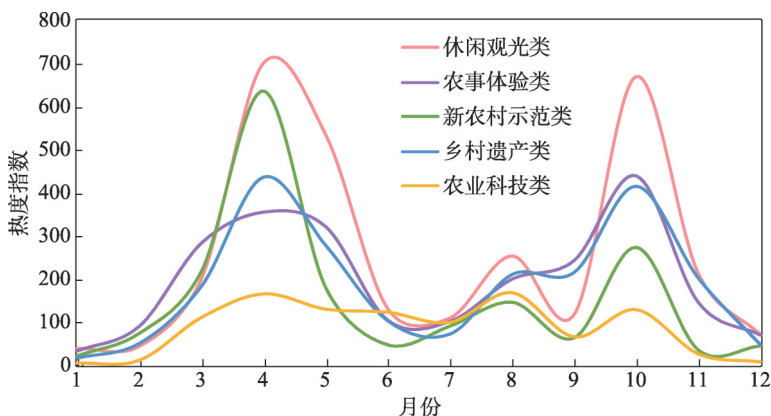


图2 2009—2017年江苏省乡村旅游热度的月统计图

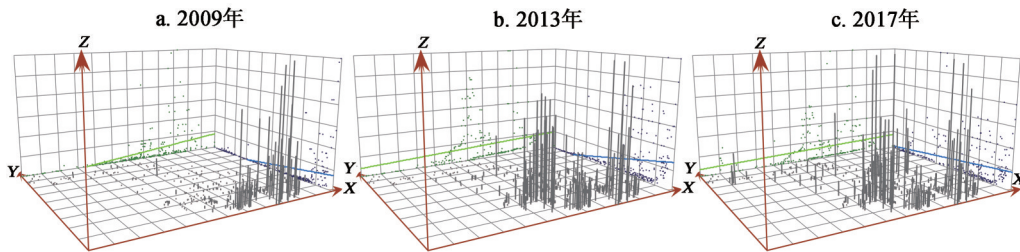
Fig. 2 Monthly count of rural tourism popularity in Jiangsu province from 2009 to 2017

作用。8月正值暑期，亲子游是这一时段的乡村旅游主体。值得关注的是，不同类型的乡村旅游景点受季节影响不同。休闲观光类景点受季节影响最大，主峰最为突出，农业科技类景点的热度值受季节影响较小，主峰特征不是很明显。反映了乡村旅游热度与乡村旅游景点的资源类型和客源主体存在很强的正相关关系。

2.2 空间演化

2.2.1 空间趋势演化

明确乡村旅游热度空间趋势变化有助于整体把握乡村旅游冷热格局演化的过程与机制。趋势面拟合结果显示（图3），研究时段内，江苏省乡村旅游热度的空间趋势线基本保持“南高北低，东高西低”的布局态势，表现出明显的空间指向性特征，南部和东部为乡村旅游的高热度区域。同时注意到，不同方向上的趋势线变化存在差异，南北方向上的趋势面过渡较为陡峭，而东西方向上的趋势面过渡趋于平缓，意味着乡村旅游热度在南北方向上的分异特征更为明显和强烈。研究发现，乡村旅游热度作为地区乡村旅游吸引力与竞争力的重要体现，其空间表现趋势与区域经济发展水平、旅游资源丰度等的空间差异以及旅游支撑系统的空间配置高度一致。



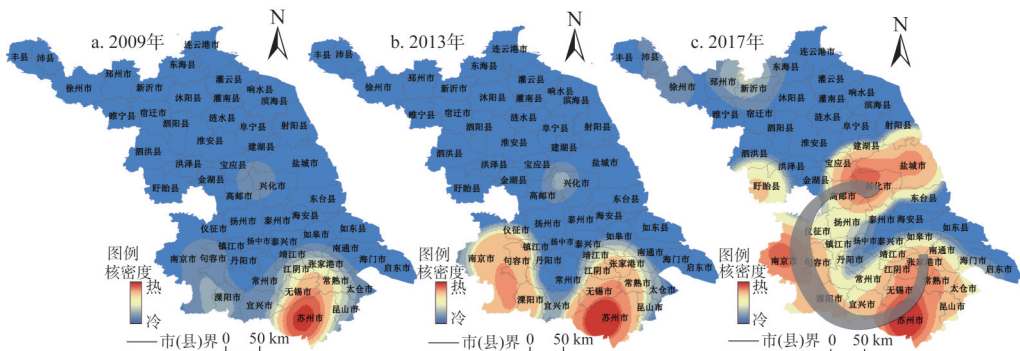
注：绿线表示东西方向，蓝线表示南北方向。

图3 2009—2017年江苏省乡村旅游热度空间趋势面拟合

Fig. 3 Trend analysis of rural tourism popularity in Jiangsu province from 2009 to 2017

2.2.2 空间结构演化

趋势面分析直观地表达了乡村旅游热度的空间分布趋势，为进一步探究其空间结构及演化特征，利用ArcGIS 10.2对江苏省乡村旅游热度进行核密度制图（图4）。结果显示：



注：c图中灰色代表“C”字形空间结构。

图4 2009—2017年江苏省乡村旅游热度核密度分布

Fig. 4 Kernel density of rural tourism popularity in Jiangsu province from 2009 to 2017

(1) 江苏省乡村旅游热度空间结构特征突出, 分布层级明显, 具体表现为苏南区连片高密度分布, 苏中区局部集聚, 苏北区零星散落布局, 形成了以苏州、南京、盐城为核心的“C”字形空间结构, 基本规定了江苏省乡村旅游热的整体脉络和格局。乡村旅游空间结构特征是自然本地和资源禀赋综合差异的反映, 苏北地势平坦, 具备农业生产的基础条件, 但由于接待服务设施配套薄弱, 生态环境较差, 客源吸引能力有限; 苏南经济基础雄厚, 设施配置完善, 生态环境优越, 具备乡村旅游发展的优势区位。

(2) 近十年来, 江苏省乡村旅游热度空间结构呈现出“单核—双核—三核”的演化过程。2009年, 乡村旅游热点区集中于苏南地区, 其中苏州的乡村旅游热度最为突出; 2013年, 热点区进一步向西北扩散至南京; 2017年, 苏中地区的盐城市周边也逐渐成为乡村旅游热点区, 现已形成以苏州、南京、盐城为核心的三大集聚区。总体来看, 江苏省乡村旅游表现出从南向北逐步扩散的演变趋势, 但高热度值仍显著集聚在苏南地区。由此不难看出, 江苏省乡村旅游热度存在着“近朱者赤、近墨者黑”的现象, 区域之间的相互影响对于乡村旅游热度的扩散和空间结构的演化具有重要作用, 区域间应打破传统的地带区划, 开展跨区域旅游开发与合作。

2.2.3 县域格局演化

趋势面分析与核密度估计一定层面上揭示了乡村旅游热度的总体分布规律, 但无法精确判别各研究单元的集聚状态。因此, 本文以县域为评价单元, 利用ArcGIS软件的热点分析方法精确识别乡村旅游热度的高值集聚区和低值集聚区, 即热点区和冷点区。首先计算出各县域的 G_i^* 得分, 继而采用 Jenks 的“自然间断点分级法”将三个典型年份的 G_i^* 得分值从高到低分成四类, 并考虑到结果的可比性, 参照曹瑞芬等^[30]的处理方法, 对各年份冷热区划分的区间值进行了统一, 进而生成乡村旅游冷热格局演化图(图5)。结果分析如下:

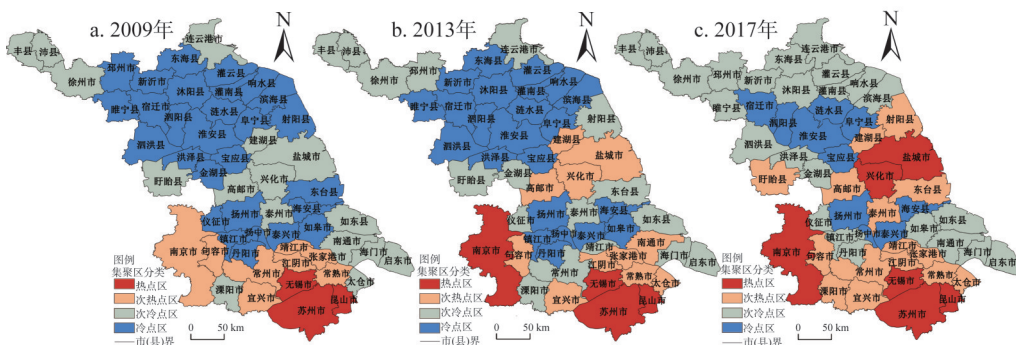


图5 2009—2017年江苏省乡村旅游冷热格局演化

Fig. 5 Hotspot distribution evolution of rural tourism popularity in Jiangsu province from 2009 to 2017

(1) 冷热点格局演化。数量变化上, 随着乡村旅游热潮的到来, 2009—2017年江苏省乡村旅游热点地区不断增多, 热点区与次热点区由2009年的11个增加到2017年的21个, 所占全省县域数量的比例由20%上升至38.2%, 冷点区与次冷点区数量明显减少, 尤其是冷点区由28个急剧降低到10个, 数量占比由50.9%降为18.2%。空间分布上, 2009年热点区集中分布在苏南地区, 形成明显的乡村旅游热度高值集聚区, 向北逐渐演替为冷点区, 冷热格局的梯度变化较为明显; 到2013年, 高值集聚区开始向西

北方向推移,南京市由次热点区升级为热点区,苏中地区由冷点区跃迁为新的次热点区;2017年,热点区进一步向东北方向扩散,苏中地区升级为热点区,苏北地区也由冷点区升级为次冷点区。总体来看,热点区有自南向北扩散的趋势,但苏南地区仍然占据江苏省乡村旅游的半壁江山。

(2) 县域热度变化。盐城、兴化和东台、射阳变化程度最大,分别由次冷点区和冷点区跃升为热点区和次热点区。苏州、无锡、昆山和句容、宜兴、江阴、张家港、常熟一直是热点区和次热点区,扬州、扬中、泰兴、海安、宝应、淮安、泗阳、宿迁、涟水、阜宁和如东、海门、启东、丰县、沛县、徐州分别为冷点区和次冷点区,乡村旅游热度变化不大。其他县域中,除常州、靖江、南通在个别年份出现了逐级减弱外,其他24个县域均表现为逐级增强。可以发现,经济社会发展水平高的县市乡村旅游热度较高,对周边地区的扩散作用也较大。苏锡常等相对发达地区之间的贸易、旅游、信息、交通等联系更加频繁与密集,带来乡村旅游产品开发、体验宣传、管理服务等多合作与发展机会,乡村旅游热度日益增强,并形成区域间乡村旅游集聚区。相对来说,苏北地区经济发展水平较弱,再加上气候条件、生态环境等自然因素的制约,在乡村旅游竞争中无法形成对等的地位与优势。

2.3 冷热格局演化的成因解析

2.3.1 指标选择

乡村旅游作为发生在乡村地区的一种新型旅游形式,具有传统旅游项目的共同特性,其旅游空间格局演化受到传统旅游吸引物空间布局的影响,但乡村旅游有着自身的鲜明特征,如乡村旅游的主要客源来自城市,出游方式以自驾为主。因此,本文参考相关研究^[31,32],在县域数据可获得基础上,充分考虑乡村旅游的客源特征,以旅游引力模型为基础,选取吸引力、驱动力、保障力、调控力四项基础要素,遴选经济发展水平、交通便利程度、接待服务能力、旅游资源禀赋、生态环境条件、政府政策导向六项指标作为探测因子,具体指标及因子的选取依据和数据来源如表2所示,并构建乡村旅游冷热格局演化影响机制分析框架(图6)。

2.3.2 因子探测

以各县域范围内乡村旅游景点热度值作为因变量,以上述六项探测因子作为自变量。由于地理探测器善于分析类别变量,故首先借助K-means分类算法对各因子变量进行离散化处理,然后运用GeoDetector软件分别计算出2009年、2013年和2017年各因子的 $P_{D,U}$ 值,结果如表3所示。

(1) 经济发展水平。区域经济发展水平对乡村旅游热度的影响值从2009年的0.612上升到2013年的0.671,但在2017年没有通过显著性检验。表明乡村旅游出游行为选择存在收入“拐点”,当居民收入达到一定的“拐点”后,经济变量的影响将不再显著。在居民收入有限时,旅游决策对经济收入具有高度依赖性,收入水平越高的地区,居民参与旅游的可能性就越大,但随着社会的发展和收入水平的提高,人们的旅游观念也随之发生改变,旅游者出游决策更加灵活和独立,不再仅仅依赖于经济收入。江苏省乡村旅游南热北冷的空间格局以及北向扩散的演变趋势,正是体现了以上规律。

(2) 交通便利程度。表3显示,交通便利程度在3个典型年份的 $P_{D,U}$ 值均大于0.7,且有逐渐增强的趋势,说明交通条件对乡村旅游热度的影响显著且深远。交通网络的重

表2 乡村旅游冷热格局演化的影响因素分析指标体系

Table 2 Index system of factors influencing rural tourism popularity

表征因子	因子解释	数据来源
人均可支配收入	人均可支配收入是决定居民收入和消费的基础, 可从供需两个方面影响乡村旅游发展, 用于表征经济发展水平对乡村旅游的影响	各县市统计年鉴
公路里程数	公路交通是乡村旅游的主要交通方式, 直接影响乡村旅游景点的可进入性, 本文选用公路里程数来衡量交通条件对乡村旅游的驱动作用	各县市统计年鉴
星级饭店数	星级饭店、旅行社是旅游服务接待能力的重要表征因子 ^[33] , 但考虑到乡村旅游以自助游为主, 旅行社影响作用不大, 故选用星级饭店数这一因子来表征	各县市国民经济与社会发展统计公报、统计年鉴
星级景点数	乡村旅游星级景点是乡村旅游发展的核心要素, 代表了区域旅游资源的富集程度, 是吸引游客的前提和基础。因此, 星级景点数成为学者衡量旅游资源禀赋的常用表征因子 ^[34]	江苏省文化和旅游厅官方网站公布的星级景点评定文件
空气质量优良率	空气质量已成为影响旅游需求和偏好的一大因素, 考虑数据的可获取性, 本文选取空气质量优良率来表征生态环境条件对乡村旅游热度的影响 ^[35]	数据来源于各县市环境质量公报及统计年鉴, 如无直接统计, 可由空气优良天数除以年天数(365)计算得到
旅游总收入占GDP的比例	政府部门往往通过资金投入发挥其政策导向作用, 旅游投资比例是较为理想的表征因子, 但大多县市并未将这一指标纳入官方统计, 为此, 学者多选取旅游总收入占GDP的比例来替代(该指标可反映出旅游业在县域经济中的地位, 是政府政策导向作用的综合体现) ^[36]	数据来源于各县市统计年鉴和国民经济与社会发展统计公报。若无旅游总收入直接统计值, 则由旅游外汇收入乘以当年汇率再与国内旅游收入求和得到

注: 个别县市的空气质量优良率数据缺失, 采取插值的方式补充完善。

大变化和交通技术的革新均会促使乡村旅游空间结构的演变, 尤其是高速公路网的完善对于以自驾游为主的乡村旅游发展意义重大。2009年江苏省全面建成了“四纵四横四联”高速公路网, 使得客源地与目的地、目的地与目的地之间的联系强度得以提升, 促使苏南地区形成乡村旅游集聚带, 且进一步影响着苏中及苏北地区, 反映了交通条件的改善对乡村旅游热度提升的重要作用。乡村旅游资源大多分布在山区、丘陵地带, 交通是发展乡村旅游的关键, 应以乡村振兴为契机, 构建乡村道路网, 破解“最后一公里”及停车难问题,

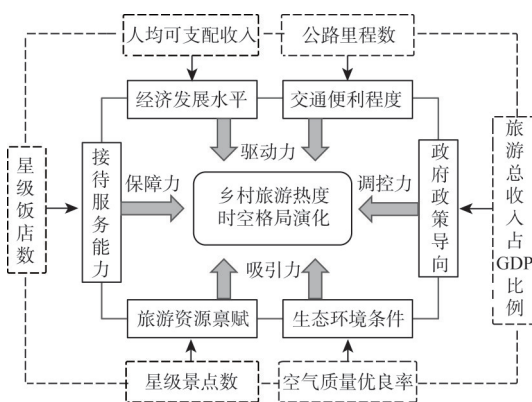


图6 乡村旅游冷热格局演化的影响机制分析框架
Fig. 6 Analytical framework of influencing mechanism to rural tourism popularity

表3 2009—2017江苏省乡村旅游热度影响因子 $P_{D,U}$ 值Table 3 Determinants $P_{D,U}$ value of rural tourism popularity in Jiangsu province from 2009 to 2017

年份	经济发展水平	交通便利程度	旅游资源禀赋	生态环境条件	接待服务能力	政府政策导向
2009	0.612***	0.738***	0.508**	0.402	0.892***	0.421
2013	0.671***	0.781***	0.623***	0.512**	0.827***	0.522**
2017	0.418	0.825***	0.566**	0.536**	0.804***	0.633***

注: ***和**分别表示影响因子在0.01和0.05置信水平上显著。

满足游客更加高效便捷的旅游需求。

(3) 接待服务能力。2009—2017年间,接待服务能力的影响力排名一直位于前三位,且呈现逐年增强的趋势,说明区域接待服务能力是保障乡村旅游客源的重要条件。以星级酒店为代表的旅游服务设施质量和服务水平决定着游客的感知体验,尤其是在体验经济与网络经济的共同作用下,旅游者通过旅游网站、个人社交平台等渠道分享旅游感知和体验,影响着潜在游客的选择倾向,进而影响实际客流量。苏南地区旅游服务设施分布密度高,等级及服务质量也明显优于其他地区,使得苏南地区始终是江苏省乡村旅游的高热度区。苏北地区乡村旅游接待设施配套不够完善,游客感知体验较差,致使其市场占有率较低,吸引力与竞争力不高。

(4) 旅游资源禀赋。表3显示,虽然旅游资源禀赋的探测值在各研究年份均大于0.5,但其影响程度呈现先增强后减弱的趋势。资源禀赋水平在乡村旅游发展早期起到决定性作用,乡村物质或非物质旅游资源富集且优越的地区乡村旅游吸引力强,往往会首先形成区域性旅游中心。然而,随着旅游层次的提升和新型旅游产品的出现,旅游资源禀赋的影响程度逐渐弱化。例如,扬州市虽然在乡村旅游资源禀赋上占有优势,但并未形成明显的乡村旅游热点区;而盐城市在资源条件并不占优势的情况下,充分挖掘当地丰富的农业资源优势,开发了荷兰花海、恒北新村等一批乡村旅游特色产品,成为苏中地区重要的乡村旅游目的地,乡村旅游热度与日俱增。这也佐证了前人仅以景点的多寡与聚散来描述乡村旅游热度差异的片面之处。

(5) 生态环境水平。因子探测结果显示,生态环境影响的显著性在2013年开始显现,且其影响力表现出不断增强的变化趋势。“绿水青山就是金山银山”,“绿水青山可以创造金山银山”,优良的生态环境质量是乡村旅游的重要引力源,更是乡村旅游可持续发展的根基所在。随着大健康时代的到来,良好的生态环境已经成为乡村旅游目的的选择的决定因素之一,而较差的生态环境将不可避免地导致潜在旅游者流失,对乡村旅游市场产生负面效应,这对乡村旅游发展的相关趋向以及目标决策具有一定的参考价值。

(6) 政府政策导向。政府政策导向通过政策引导、资金支持鼓励乡村旅游发展,是乡村旅游时空演化的外在调控力。从各年份的 $P_{D,U}$ 值来看,政府导向的影响力已经由不显著变为显著,成为区域乡村旅游发展的重要驱动因素。例如,兴化市近年来通过转换发展思路,利用良好的生态环境和垛田特色增加投资,逐步从因工业污染严重的不锈钢产业基地发展为环境优美、百姓致富的乡村旅游目的地,成为苏中板块的高热点区。下一步应充分发挥政府的政策导向作用,通过合理布局、要素优化来实现乡村旅游吸引力的提升,积极引导乡村旅游投资由南向北扩散,为江苏北部地区的乡村旅游发展注入新动力,开拓新思路。

2.3.3 成因解析

乡村旅游热度时空演化过程是各种驱动力共同作用的结果,经济水平和交通条件是乡村旅游热度空间演变的直接驱动力,资源禀赋和生态环境是重要的吸引力要素,接待能力和政府导向则分别起着内在保障和外在调控的作用。前文分析可知,各要素在不同时期的影响强度存在差异。交通条件和接待能力在3个研究时段均产生重要影响,其他因子的影响程度在不同时段存在一定差异。2009年,经济发展水平作为主要影响因子,对乡村旅游热度提升起到了重要的推动作用,直接影响江苏省乡村旅游冷热格局的形

成; 2013年, 资源禀赋和政府导向的作用开始呈现; 2017年, 经过近十年的发展, 江苏省乡村旅游进入稳定发展阶段, 经济发展的驱动作用逐步减弱, 区域生态环境质量的影响力日益凸显。随着居民生活水平的提高, 乡村旅游需求由单纯追求感官刺激向更加重视特色化、品质化转变, 交通条件、接待能力、生态环境必将在相当长时间内成为乡村旅游目的地选择的主导因素。

3 结论与讨论

面向公众服务的旅游网站数据包含旅游时间、地点、感知等属性信息, 具有数据丰富、实时性强的特点, 为乡村旅游定量研究提供了数据源。同时 Python 等数据挖掘技术的进步为数据获取提供了技术支持。本文以江苏省为研究对象, 探讨基于旅游大数据的乡村旅游热度测评方法。在此基础上, 从时间与空间两个维度对江苏省乡村旅游冷热格局及演化特征进行分析, 并借助地理探测器对格局演化成因进行解析。研究结果表明:

(1) 时间上, 江苏省乡村旅游热度的年际变化特征及季节变化规律明显。年际变化呈“S”型演化轨迹; 季节变化呈“三峰四谷”特征, 淡旺季分明。总体来看, 不同类型的乡村旅游景点受季节影响不同, 而休闲观光类景点受季节影响最大。

(2) 空间上, 江苏省乡村旅游热度区域差异显著, 整体表现为“南高北低, 东高西低”的空间特征, 冷热格局的梯度变化较为明显。江苏省乡村旅游热度的空间结构经历了“单核—双核—三核”的演化过程, 目前呈现出“西进北拓”的扩散趋势。

(3) 乡村旅游热度时空演化与经济水平、资源禀赋、交通条件、接待能力、生态环境、政策导向等因素密切相关, 但不同时段各因素影响强度各异。其中, 交通条件、接待能力一直是主导因素; 经济水平对乡村旅游发展早中期的热度提升有显著影响; 资源禀赋的影响力呈下降趋势, 生态环境和政策导向的正向影响力不断增强。

基于网络数据的乡村旅游热点识别尝试为乡村旅游定量研究提供新的研究视角, 实用性方面有利于明晰乡村旅游热度时空演化特征, 深化区域乡村旅游发展认知, 对乡村旅游区域合作、规划与决策具有理论和实践指导意义。本文还存在一些不足, 一方面, 由于现阶段旅游大数据还存在着人群异质性、代表性不足等问题, 尽管研究结果较好地切合了区域乡村旅游的发展现状, 但测评模型的科学性和普适性仍需在后续研究中进一步通过实验、模拟等手段进行验证。另一方面, 受到统计年鉴数据类型、数据内容及地域间数据差异的限制, 乡村旅游热度驱动机制研究仅基于县域尺度进行了分析, 如能针对不同类型景点对其热度影响因素进行研究, 可能对乡村旅游发展与规划的指导价值和意义更大。为此, 本文在测评模型构建与验证、方法创新以及影响机制研究方面尚有较大改进空间, 今后仍需继续努力与探索, 以期做出更具价值的学术研究成果。

参考文献(References):

- [1] 刘彦随. 中国新时代城乡融合与乡村振兴. 地理学报, 2018, 73(4): 637-650. [LIU Y S. Research on the urban-rural integration and rural revitalization in the New Era in China. *Acta Geographica Sinica*, 2018, 73(4): 637-650.]
- [2] SHARPLEY R, ROBERTS L. Rural tourism-10 years on. *International Journal of Tourism Research*, 2004, 6(3): 119-124.
- [3] LEPP A. Tourism and dependency: An analysis of Bigodi village, Uganda. *Tourism Management*, 2008, 29(6): 1206-1214.
- [4] PARK D B, YOON Y S. Segmentation by motivation in rural tourism: A Korean case study. *Tourism Management*, 2009, 30(1): 99-108.

- [5] KOMPPULA R. The role of individual entrepreneurs in the development of competitiveness for a rural tourism destination: Case study. *Tourism Management*, 2014, 40(1): 361-371.
- [6] HWANG J H, LEE S W. The effect of the rural tourism policy on non-farm income in South Korea. *Tourism Management*, 2015, 46(1): 501-513.
- [7] CAMPÓN-CERRO A M, HERNÁNDEZ-MOGOLLÓN J M, ALVES H. Sustainable improvement of competitiveness in rural tourism destinations: The quest for tourist loyalty in Spain. *Journal of Destination Marketing & Management*, 2017, 6(3): 252-266.
- [8] 汪侠, 甄峰, 沈丽珍, 等. 基于贫困居民视角的旅游扶贫满意度评价. *地理研究*, 2017, 36(12): 2355-2368. [WANG X, ZHEN F, SHEN L Z, et al. Evaluation of poor residents' satisfaction with tourism poverty alleviation. *Geographical Research*, 2017, 36(12): 2355-2368.]
- [9] 杨德进, 白长虹, 牛会聪. 民族地区负责任旅游扶贫开发模式与实现路径. *人文地理*, 2016, 31(4): 119-126. [YANG D J, BAI C H, NIU H C. Development models and realization paths of responsible pro-poor tourism in ethnic minority areas. *Human Geography*, 2016, 31(4): 119-126.]
- [10] 黄越, 赵振斌. 旅游社区居民感知景观变化及空间结构: 以丽江市束河古镇为例. *自然资源学报*, 2018, 33(6): 1029-1042. [HUANG Y, ZHAO Z B. Residents' perception and spatial structure of landscape change in tourism communities: A case study of Shuhe Old Town in Lijiang. *Journal of Natural Resources*, 2018, 33(6): 1029-1042.]
- [11] 方世敏, 王海艳. 基于系统论的农业与旅游产业融合: 一种粘性的观点. *经济地理*, 2018, 38(12): 211-218. [FANG S M, WANG H Y. Agriculture and tourism industry convergence based on system theory: A viewpoint of viscosity. *Economic Geography*, 2018, 38(12): 211-218.]
- [12] 许黎, 曹诗图, 柳德才. 乡村旅游开发与生态文明建设融合发展探讨. *地理与地理信息科学*, 2017, 33(6): 106-111, 124. [XU L, CAO S T, LIU D C. Discussion on the integration development of rural tourism development and ecological civilization construction. *Geography and Geo-Information Science*, 2017, 33(6): 106-111, 124.]
- [13] 陆林, 韩娅, 黄剑锋, 等. 基于扎根理论的杭州市梅家坞旅游城市化特征与机制. *自然资源学报*, 2017, 32(11): 1905-1918. [LU L, HAN Y, HUANG J F, et al. Characteristics and mechanism of tourism-driven urbanization based on grounded theory: A case study of Hangzhou Meijiawu village. *Journal of Natural Resources*, 2017, 32(11): 1905-1918.]
- [14] 黄震方, 陆林, 苏勤, 等. 新型城镇化背景下的乡村旅游发展: 理论反思与困境突破. *地理研究*, 2015, 34(8): 1409-1421. [HUANG Z F, LU L, SU Q, et al. Research and development of rural tourism under the background of new urbanization: Theoretical reflection and breakthrough of predicament. *Geographical Research*, 2015, 34(8): 1409-1421.]
- [15] 郑治伟. 北京市乡村旅游空间分布及发展评价. *中国农业资源与区划*, 2018, 39(5): 224-230. [ZHENG Z W. Spatial characteristics and development evaluation of rural tourism in Beijing. *Chinese Journal of Agricultural Resources and Regional Planning*, 2018, 39(5): 224-230.]
- [16] 王建英, 黄远水, 叶新才. 可达性及游憩压力视角下乡村旅游景区(点)空间结构优化. *地理与地理信息科学*, 2016, 32(5): 110-114. [WANG J Y, HUANG Y S, YE X C. Layout optimization of rural tourism scenic spots based on accessibility and recreational pressure. *Geography and Geo-Information Science*, 2016, 32(5): 110-114.]
- [17] 许贤棠, 刘大均, 胡静, 等. 国家级乡村旅游地的空间分布特征及影响因素: 以全国休闲农业与乡村旅游示范点为例. *经济地理*, 2015, 35(9): 182-188. [XU X T, LIU D J, HU J, et al. Spatial distribution characteristics and influence factors of rural tourism destination in China: The case of national level leisure agriculture and rural tourism demonstration sites. *Economic Geography*, 2015, 35(9): 182-188.]
- [18] 张子昂, 黄震方, 靳诚, 等. 基于微博签到数据的景区旅游活动时空行为特征研究: 以南京钟山风景名胜區为例. *地理与地理信息科学*, 2015, 31(4): 121-126. [ZHANG Z A, HUANG Z F, JIN C, et al. Research on spatial-temporal characteristics of scenic tourist activity based on Sina microblog: A case study of Nanjing Zhongshan mountain national park. *Geography and Geo-Information Science*, 2015, 31(4): 121-126.]
- [19] 马丽君, 郭留留. 基于网络关注度的北京市居民对5A级景区旅游需求时空特征分析. *干旱区资源与环境*, 2017, 31(10): 203-208. [MA L J, GUO L L. Spatial and temporal distribution characteristics of Beijing residents' demand for 5A class scenic spots. *Journal of Arid Land Resources and Environment*, 2017, 31(10): 203-208.]
- [20] 孙焯, 张宏磊, 刘培学, 等. 基于旅游者网络关注度的旅游景区日游客量预测研究: 以不同客户端百度指数为例. *人文地理*, 2017, 32(3): 152-160. [SUN Y, ZHANG H L, LIU P X, et al. Forecast of tourism flow volume of tourist attrac-

- tion based on degree of tourist attention of travel network: A case study of Baidu index of different clients. *Human Geography*, 2017, 32(3): 152-160.]
- [21] 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. *地理学报*, 2018, 73(8): 1556-1570. [QIN J, LI L P, TANG M D, et al. Exploring the spatial characteristics of Beijing inbound tourist flow based on geotagged photos. *Acta Geographica Sinica*, 2018, 73(8): 1556-1570.]
- [22] 李萍, 陈田, 王甫园, 等. 基于文本挖掘的城市旅游社区形象感知研究: 以北京市为例. *地理研究*, 2017, 36(6): 1106-1122. [LI P, CHEN T, WANG F Y, et al. Urban tourism community image perception and differentiation based on online comments: A case study of Beijing. *Geographical Research*, 2017, 36(6): 1106-1122.]
- [23] 刘超, 胡梦晴, 林文敏, 等. 山岳型景区旅游形象感知研究: 基于2014—2016年黄山网络游记分析. *山地学报*, 2017, 35(4): 566-571. [LIU C, HU M Q, LIN W M, et al. Tourism destination image perception of mountain resorts: A comprehensive survey of the Huangshan web travel notes for 2014-2016. *Mountain Research*, 2017, 35(4): 566-571.]
- [24] 张子昂, 黄震方, 曹芳东, 等. 浙江省县域入境旅游时空跃迁特征及驱动机制. *地理研究*, 2016, 35(6): 1177-1192. [ZHANG Z A, HUANG Z F, CAO F D, et al. The space-time transition characteristics and its driving mechanism of county scale inbound tourism in Zhejiang province. *Geographical Research*, 2016, 35(6): 1177-1192.]
- [25] 张霞, 赵玮丹, 江文萍, 等. 基于网络热度的道路景观评价与最美路径推荐. *同济大学学报: 自然科学版*, 2016, 44(3): 383-388. [ZHANG X, ZHAO W D, JIANG W P, et al. Land scape evaluation and most beautiful route planning from big website data. *Journal of Tongji University: Natural Science*, 2016, 44(3): 383-388.]
- [26] 徐建华. *现代地理学中的数学方法(第2版)*. 北京: 高等教育出版社, 2002: 98-105. [XU J H. *Mathematical Methods in Contemporary Geography (Edition 2)*. Beijing: Higher Education Press, 2002: 98-105.]
- [27] WEST M. Kernel density estimation and marginalization consistency. *Biometrika*, 1991, 78(2): 421-425.
- [28] GETIS A, ORD J K. The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 1992, 24(3): 189-206.
- [29] 刘彦随, 李进涛. 中国县域农村贫困化分异机制的地理探测与优化决策. *地理学报*, 2017, 72(1): 161-173. [LIU Y S, LI J T. Geographic detection and optimizing decision of the differentiation mechanism of rural poverty in China. *Acta Geographica Sinica*, 2017, 72(1): 161-173.]
- [30] 曹瑞芬, 张安录. 土地发展非均衡度测算及其敏感性分析: 基于农地非农发展受限视角. *自然资源学报*, 2017, 32(5): 715-726. [CAO R F, ZHANG A L. Measurement and sensitivity analysis of unbalance degree of land development: From the perspective of farmland development restriction. *Journal of Natural Resources*, 2017, 32(5): 715-726.]
- [31] 程哲, 蔡建明, 崔莉, 等. 乡村转型发展产业驱动机制: 以盘锦乡村旅游为例. *农业现代化研究*, 2016, 37(1): 143-150. [CHENG Z, CAI J M, CUI L, et al. Facilitating rural transformation development by tourism industry: A case study of Panjin, Liaoning province. *Research of Agricultural Modernization*, 2016, 37(1): 143-150.]
- [32] 周杨, 何军红, 荣浩. 我国乡村旅游中的游客满意度评估及影响因素分析. *经济管理*, 2016, 38(7): 156-166. [ZHOU Y, HE J H, RONG H. Satisfaction evaluation of tourist and influence factors analysis in rural tourism. *Economic Management*, 2016, 38(7): 156-166.]
- [33] 张红贤, 游细斌, 白伟杉, 等. 目的地旅游吸引力测算及相关因素分析. *经济地理*, 2018, 38(7): 199-208. [ZHANG H X, YOU X B, BAI W S, et al. Measurement of destination tourism attractiveness and influence factors. *Economic Geography*, 2018, 38(7): 199-208.]
- [34] 于秋阳, 颜鑫. 区域旅游经济的时空分异及收敛性研究: 以江苏省为例. *华东经济管理*, 2019, 33(1): 11-18. [YU Q Y, YAN X. Spatial-temporal differentiation and convergence of regional tourism economy: A case of Jiangsu province. *East China Economic Management*, 2019, 33(1): 11-18.]
- [35] 李涛, 刘家明, 刘锐, 等. 基于“生产—生活—生态”适宜性的休闲农业旅游开发. *经济地理*, 2016, 36(12): 169-176. [LI T, LIU J M, LIU R, et al. The development of leisure agriculture based on suitability about "production-living-ecology". *Economic Geography*, 2016, 36(12): 169-176.]
- [36] 赵磊, 方成, 吴向明. 旅游发展、空间溢出与经济增长: 来自中国的经验证据. *旅游学刊*, 2014, 29(5): 16-30. [ZHAO L, FANG C, WU X M. Tourism development, spatial spillover and economic growth: An empirical evidence from China. *Tourism Tribune*, 2014, 29(5): 16-30.]

Hotspot identification and cause analysis of rural tourism based on website data: Take Jiangsu province as an example

RONG Hui-fang^{1,2}, TAO Zhuo-min¹

(1. School of Geography Science, Nanjing Normal University, Nanjing 210023, China;

2. Resource Environment College, Chizhou University, Chizhou 247000, Anhui, China)

Abstract: Rural tourism has important practical significance for optimizing the rural industrial structure, and recovering the rural economy, especially for the implementation of the rural revitalization strategy. Thus, rural tourism is not only the focus of local government and tourism enterprises, but also a hot topic in domestic and international tourism research. At the same time, with the development and popularization of the Internet, travel websites, social software and other online platforms have become important tools to obtain travel information, make travel decisions, and share travel experiences. Tourism big data provides data sources and methodological support for rural tourism research. Based on data of tourism network, this paper puts forward a method for identifying rural tourism hotspots. Taking Jiangsu province as an example, this paper uses the methods of trend surface, nuclear density estimation and hot spot analysis to explore the cold and hot patterns of rural tourism, and reveals the influencing factors of the evolution with the help of geographic detectors. The results show that: (1) The annual and seasonal changes of rural tourism heat are obvious. The annual change presents an "S" shaped evolution track, seasonal variation is characterized by "three peaks and four valleys", but the degree of seasonal influence on different types of rural tourist attractions is slightly different. (2) The spatial structure of rural tourism in Jiangsu province experienced the evolution of "mononuclear-dual-nuclear-trinuclear" in 2009-2017, but its heterogeneity is still significant, basically maintaining the overall characteristics of "high in the south and east regions, while low in the north and west regions". The hot spots are concentrated in southern Jiangsu and gradually evolve into cold spots in the north. The evolution of spatial structure shows a trend of "expanding from the west to the north". (3) There are obvious strength differences and time variations among the influencing factors. Transportation convenience and reception capacity have always been the main influencing factors. The economic development has a significant positive impact on the early development of rural tourism, and the influence of tourism resource tends to decline. The positive influence of ecological environment and government orientation on rural tourism is increasing. Hotspot identification based on network data provides a new perspective for quantitative research of rural tourism. In terms of practicability, it is helpful to clarify the evolution characteristics of the cold and hot patterns of rural tourism so as to provide important guiding significance for rural tourism resource development and regional cooperation.

Keywords: big data; hotspot identification; evolution; cause; Jiangsu province