

基于网络核密度估计的 河湖人类活动热点区域识别

宋瑞平, 王冬梅, 刘文壮

(江苏省水利科学研究院, 江苏 南京 210017)

摘要:以里下河中部地区湖泊群为研究范围,选择研究区内 2017—2020 年间由遥感监测发现的人类活动变化点数据,基于网络核密度估计的方法识别研究区内人类活动现象的高发区域、高发河段。结果表明,变化点的分布与变化点的类型、周边养殖业发展程度、土地类型有关。此外,网络核密度估计方法在应用于河网内部热点识别时,其结果比平面核密度估计更符合实际情况。

关键词:平面核密度估计;网络核密度估计;遥感监测;里下河腹部地区;湖泊湖荡

中图分类号:X832

文献标识码:B

文章编号:1007-7839(2022)02-0065-04

Hotspots identification of human activities in rivers and lakes based on network kernel density estimation

SONG Ruiping, WANG Dongmei, LIU Wenzhuang

(Jiangsu Hydraulic Research Institute, Nanjing 210017, China)

Abstract:In this study, the lakes in the middle part of the Lixiahe River were selected as the research area, and the data of human activity change points detected by remote sensing monitoring during 2017 and 2020 were selected in the study area. Based on the network kernel density estimation method, the areas and river sections with high incidence of human activity phenomena in the study area were identified. The results showed that the distribution of change point was related to the type of change point, the development degree of surrounding aquaculture and land type. In addition, when the network kernel density estimation method is applied to the hot spot identification in river network, the results are more consistent with the actual situation than the plane kernel density estimation.

Key words: plane kernel density estimation; network kernel density estimation; remote sensing monitoring; belly area of Lixia River; lake

在大尺度空间上,具体的地理实体对象和曾经某些现象发生的地点都表现为点的特征,例如居民点、犯罪现场、违法现象等,这些离散的地理对象或事件点的分布情况对于城市规划、犯罪热点识别、违法热点地段识别有重要作用,对该类空间点分布模式的研究称为空间点模式分析。空间点模式分析大多采取空间聚类法,包括 Getis-Ord (G_i^*) 统计、核密度估计(kernel density estimation, KDE)、最近邻方法、K-means 聚类等,KDE 是一种重要的非

参数检验方法,广泛用于交通中危险路段检测。KDE 是通过将每个位置的空间点个数作为密度来创建空间上点事件的平滑密度表面,其中心密度值最高,并且随着空间点远离中心而下降,通过将空间点分布可视化为平滑连续的表面,借此来识别事件的高发区域。

预防河湖人类活动变化发生的主要关注点之一是确定人类活动频繁的区域,有针对地在该地区巡查。里下河地区历年遥感监测中,人类活动变化

收稿日期:2021-10-20

作者简介:宋瑞平(1997—),男,本科,研究方向为水利信息化、水利空间数据分析。E-mail: 484209521@qq.com

往往发生于圩埂、河流沿岸的陆地之上,而不是分布于河道湖泊内部。如果将河网、圩埂视作网格,可以认为变化点是受该网格约束的事件点。

基于上述分析,本文根据 Okabe 等^[1]提出的网络约束点事件研究思路,使用网络核密度估计方法对河湖遥感监测中的人类活动变化点进行分析,识别人类活动的热点河段,并与平面核密度分析结果进行比较。

1 研究区域和数据

1.1 研究区域

里下河地区位于江苏省中部,全区地势低洼,湖荡相连,人工水网稠密,是全省重要的农业种植和养殖区。研究区域为里下河地区中部的兰亭荡、广洋湖、王庄荡,面积约为 231 km²,区域内网络约 120 km²。河流网络是指区域内主要河流的组成的网络,而圩区网络是圈圩圩埂的线性网络,本文将河流网络和圩区网络组合在一起,并将它们视作一组网络进行分析。

1.2 数据来源

本文选择 2017—2020 年研究区内河湖遥感监测发现的变化点为研究对象,包括疑似彩钢棚,新建房屋等共计 281 处,见表 1。河湖遥感地理数据库以 Excel 形式为主,包含变化点的详细属性,如事件名称、事件类型、发生地点、占用面积、所属河湖等。之后根据事件点经纬度使用 ArcGIS 进行映射,将其转换为 GIS 矢量文件。

表 1 研究区内变化点类型

年份	彩钢房	砖混房屋	堆放物	电力设施	公墓	桥梁	合计
2017 年	29	4	7	0	0	2	42
2018 年	11	4	9	0	1	0	25
2019 年	69	6	14	31	0	0	120
2020 年	54	8	12	17	3	0	94

2 违法热点地区识别方法

2.1 网络 KDE

KDE 是一种用于估计概率密度函数的非参数方法,通过将每个事件点的数据和带宽当作特定核函数的参数,对估计函数归一化生成平滑的密度表面,可视化后能够体现变化点的分布模式,在交通犯罪预防、经济领域有广泛的应用。KDE 在根据路

径分析的原理不同分为平面 KDE 和网络 KDE,网络 KDE 在平面 KDE 的基础上有以下 2 个方面改进^[2-3]。

(1) 空间点分布受到路网、河网等网络约束,使用网络 KDE 的结果误差较小。图 1 中 a 和 b 中的点在空间中分布完全一致,以网络距离进行事件搜索时有两起事件作为核函数的参数被计算,而基于欧式距离搜索时有 4 起事件在搜索范围内,平面 KDE 会造成网络事件点的过度聚集,影响最终结果。

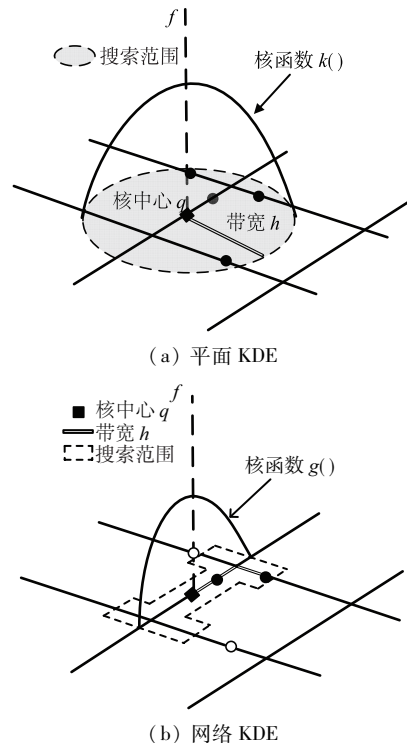


图 1 平面 KDE 与网络 KDE 估计原理对比

(2) 平面 KDE 估计中密度值相同的节点可能包含不同数量的网格段,网络 KDE 估计考虑了要素在节点处密度计算的特殊情况,节点处的核函数根据节点数量的不同做出调整。

实际的河网、圩网中,人类活动变化事件的发生和分布受到河网、圩网空间格局的影响,具有非均质性的特征,故本文使用网络 KDE 来估计人类活动的高密度地区。

河网、圩网中基本河段的网络 KDE 函数为

$$K(q) = \begin{cases} \sum_{i=1}^n \frac{1}{h} k\left[\frac{d_s(q, p_i)}{h}\right] & d_s(q, p_i) < h \\ 0, & d_s(q, p_i) \geq h \end{cases} \quad (1)$$

式中: $K(q)$ 为点 q 处的网络核密度值; n 为事件点个数; h 为搜索半径(或搜索带宽); p_i 为点事件, $i = 1, 2, 3, \dots, n$; $d_s(q, p_i)$ 为点事件 p_i 到核中心 q 的距离; $k()$ 为核函数。

实际的河网、圩网中,河段网络核密度经过节点处后会出现衰减。针对不同节点数的节点处的衰减,Okabe 等使用了“节点等分不连续”内核函数。

图 2 中,点 q 为核心中心,为交叉口节点, v_{ik} 内核函数基于节点处的节点数进行分割,当经过下个节点时,再次进行分割,此过程一直持续到内核中心到达边界点。交叉口区域的网络 KDE 函数为

$$K(q)=\begin{cases}\frac{\sum_{i=0}^n\frac{1}{h}k\left[\frac{d_s(q,p_i)}{h}\right]}{(n_{i1}-1)(n_{i2}-1)\cdots(n_{ik}-1)}&d_s(q,p_i)<h\\0,&d_s(q,p_i)\geq h\end{cases}\quad(2)$$

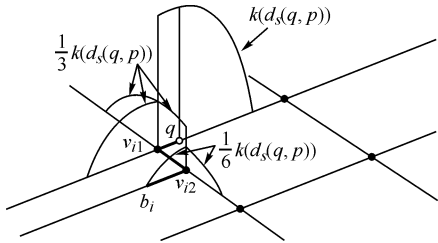


图 2 节点等分函数示意

3 实证分析

3.1 网络核密度估计

为了网络 KDE 能顺利进行,研究区内河流网络和圩区路网均矢量化后在 arcmap 中合并为一组河网。网络 KDE 主要由核函数类型、搜索带宽、单位网络长度 3 个因素决定,而其中对结果影响最大的是搜索带宽。通过比较不同的搜索半径和单位网络长度的结果,确认带宽 500m、单位网络长度 20 m 的条件下密度图的平滑度最好。

采用 SANET 软件,对研究区域内 2017—2020 年间的变化点进行网络 KDE,并在 ESRI ArcGIS 中可视化网络核密度计算结果。在带宽 500 m,单位网络距离 20 m 的条件下,对 2017—2020 年 4 年的核密度估计结果如图 3 所示。

研究区内遥感变化点的核密度分析结果显示,河段的密度值基本代表该河段人类活动程度。2017 年人类活动的热点区域主要分布在向阳河、杨家河河段,沙黄河河段。2018 年人类活动的热点区域主要分布在沙黄河、杨家河河段,芦范河段。2019 年人类活动的热点区域主要分布在杨家河河段,芦范河段,广洋湖圩(w77)周边。2020 年类活动的热点区域主要分布在向阳河、杨家河河段,杨家河河段,广洋湖圩(w77)周边。

2017—2020 年,向阳河(兰亭荡段)和杨家河

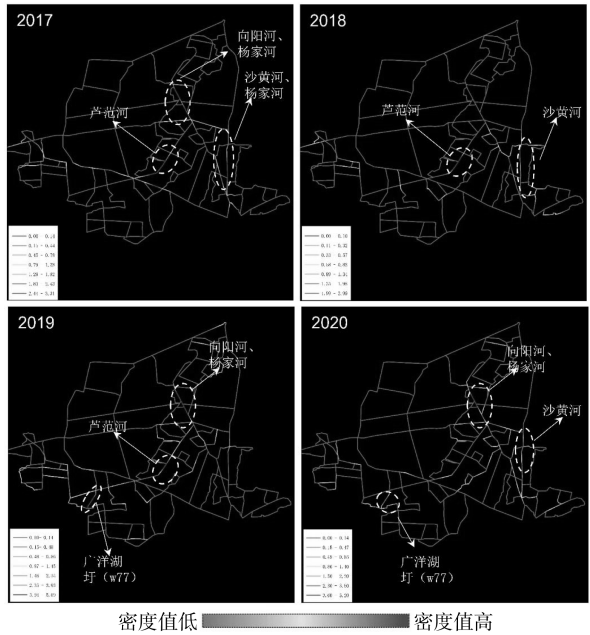


图 3 2017—2020 年研究区变化点网络核密度估计结果(兰亭荡段)周边进行渔光互补光伏发电项目建设,人类活动频繁,自 2019 年起广洋湖圩(w77)也在进行渔光互补光伏发电项目建设。沙黄河、芦范河周边渔业养殖发达,渔民常会新建鱼塘看护彩钢房,这是这些地区人类活动频繁的内因。

表 2 研究区内主要热点河段核密度值

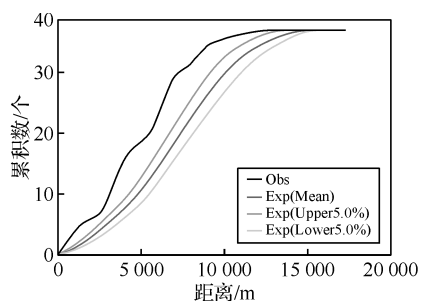
河段名称	平均核密度值				合计
	2017 年	2018 年	2019 年	2020 年	
芦范河(广洋湖段)	2.65	2.98	4.70		10.33
向阳河(兰亭荡段)	1.75		2.78	2.60	7.13
杨家河(兰亭荡段)	2.86		5.81	5.25	13.92
沙黄河(王庄荡段)	2.83	1.89		2.70	7.42
广洋湖圩(w77)			5.15	2.90	8.05

3.2 多距离空间聚类分析

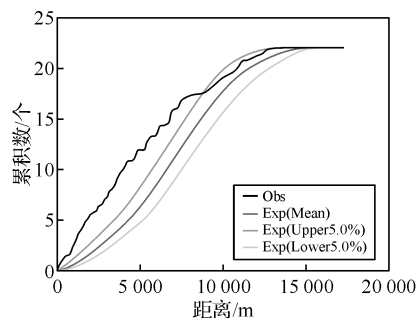
网络 KDE 的问题在于它是一种先验方法,不具有统计学意义。通过多距离空间聚类分析可以对一定距离范围内的空间自相关性进行汇总,如果某个特定计算距离的平均相邻要素数高于整个研究区域内要素的平均密度,则该距离的空间聚类具有统计显著性。对于位于网络中的空间点应当使用网络多距离空间聚类分析^[4-5]。

对 2017—2020 年间的 4 年变化点进行网络多距离空间聚类分析后得出图 4 结果。

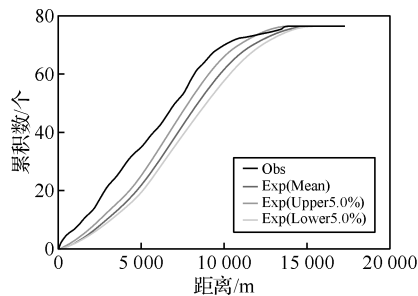
预期结果(随机分布模式)以红线表示,观测结果(实验数据分布模式)以蓝线表示,较高置信区间范围以绿线表示,较低置信区间范围则以粉线表



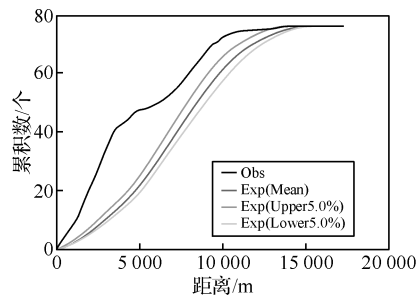
(a) 2017 年分析结果



(b) 2018 年分析结果



(c) 2019 年分析结果



(d) 2020 年分析结果

图 4 2017—2020 年研究区网络空间聚类分析结果

示。观测线在预期线之上表明数据集在该距离内表现为聚类。观测线在预期线之下表明数据集在该距离内表现为离散,观测结果与预期结果的差值表示该数据集的聚类/离散程度。从图 4 得出,观测结果在非常大的范围内均大于较高置信区间范围,表明 2017—2020 年间研究区内人类活动呈多区域聚集模式,而 2017 年和 2020 年的聚类程度更为明显,这两年也是向阳河(兰亭荡段)、杨家河(兰亭

荡段)新建光伏发电设施比较频繁的 2 年,与网络 KDE 相比符合预期,为网络 KDE 的结果提供了统计学参考价值。

4 结 论

本文以里下河地区 2017—2020 年遥感监测变化点为基础,采用基于河网、圩网距离的网络 KDE 方法分析了研究区内人类活动变化,并采用多距离空间聚类分析法对数据空间分布模式进行验证,识别出人类活动热点河段。具体结论如下:在里下河地区,变化点主要集中在建设开发强度大,养殖业发达的河段;与平面核密度方法相比,网络核密度估计的原理和结果都更符合事件点在网络上分布的实际情况;多距离空间聚类分析对网络核密度估计的结果有统计学上的参考价值。两者结合可以使结果更具合理性。

本文从时空尺度上研究人类活动热点,结论能有效指导地方河湖巡查单位巡查人员调度、任务分配,通过对热点河段加大巡查力度,减少人类活动较少区域的巡查,可以提高地方河湖巡查人员巡查效率。但是本文所选时间跨度较小,仅选择了年度间的变化点数据,不足以分析各河段热点变化趋势,无法对未来人类活动演变做出预测,未来将考虑选择时间跨度更大的数据,建立研究区热点演变分析模型。

参考文献:

- [1] OKABE A. Spatial analysis along networks: statistical and computational methods[M]. New Jersey: John Wiley & Sons, 2012.
- [2] KHALID S, SHOAI B F, QIAN T. Network constrained spatio-temporal hotspot mapping of crimes in faisalabad[J]. Applied Spatial Analysis and Policy, 2018(11): 599-622.
- [3] XIE Z, YAN J. Detecting traffic accident clusters with network kernel density estimation and local spatial statistics: an integrated approach[J]. Journal of Transport Geography, 2013, 31(7): 64-71.
- [4] 钱天陆,苑振宇,倪建华,等. 基于平面与网络核密度估计的南京市主城区 ATM 机分布热点探测[J]. 科技通报, 2018, 34(1): 105-110.
- [5] 龙雪琴,周萌,赵欢,等. 基于网络核密度的网约车上下客热点识别[J]. 交通运输系统工程与信息, 2021, 21(3): 86-100.