引文格式：余思汗,张楠,陈瑞,等. 基于POI要素空间聚集特征的集镇识别与边界提取[J]. 地理空间信息, 2025, 23(4):100-104. doi:10. 3969/j. issn. 1672-4623.2025.04.022

基于POI要素空间聚集特征的集镇识别与边界提取

余思汗[1]，张　楠[1,2]，陈　瑞[3]，刘　超[1]，杨　顺[1]，孙嘉欣[1]，蔡保祥[3*]
（1. 宁夏回族自治区地震局，宁夏 银川 750001；2. 中国地震局地质研究所，北京 100029；
3. 宁夏回族自治区自然资源信息中心，宁夏 银川 750000）

摘　要：为了高效率精确识别集镇、准确提取边界，为地震应急快速评估和地震应急救援提供精准数据支撑，利用核密度分析、加权叠加分析和空间关联分析，探索集镇的空间分布和边界范围。根据POI数据的覆盖广、聚集度特征明显的特点，针对集镇职能提出居民依赖性合理计算方法，实现对研究区内所有集镇的识别和边界提取。结果表明：利用POI能够识别研究区内的全部集镇，并成功提取集镇边界，在空间重叠度和面积精度方面，具有较好的一致性和准确性，可为集镇识别、边界提取提供新思路。

关键词：集镇；POI数据；空间聚集；核密度分析；加权叠加分析；地震应急
中图分类号：P208　　文献标志码：B　　文章编号：1672-4623（2025）04-0100-05

Market Town Recognition and Boundary Extraction Based on Spatial Clustering Characteristics of POI Elements

YU Sihan[1], ZHANG Nan[1,2], CHEN Rui[3], LIU Chao[1], YANG Shun[1], SUN Jiaxing[1], CAI Baoxiang[3]
(1. Earthquake Agency of Ningxia Hui Autonomous Region, Yinchuan 750001, China;
2. Institute of Geology, China Earthquake Administration, Beijing 100029, China;
3. Natural Resources Information Center of Ningxia Hui Autonomous Region, Yinchuan 750000, China)

Abstract: In order to efficiently and accurately identify market towns and extract their boundaries, to provide accurate data support for rapid earthquake emergency evaluation and rescue, we used kernel density analysis, WSA and spatial correlation analysis to explore the spatial distribution and boundary size of some market towns. Based on the wide coverage and obvious clustering characteristics of POI, we proposed a calculation method for town residents' dependency for the functions of market towns, achieving the recognition and boundary extraction of all market towns in the research area. The results show that using POI can successfully identify and extract the boundaries of all market towns. This method has fair consistency and accuracy in terms of spatial overlap and area accuracy, providing new ideas for town recognition and boundary extraction.

Key words: market town, POI data, spatial aggregation, kernel density analysis, weighted overlay analysis, earthquake emergency

随着中国经济的发展，城镇化进程快速推进，集镇作为城市和乡村之间的纽带，是介于乡村和城市间的过渡型居民点，在居民经济生活中起着重要作用，也迎来了在相当长的一段时间内持续扩展。集镇内部结构的主要特征，是商业街道居于核心的地位，其平面形态则受当地环境以及与相邻村镇联络的道路格局的影响，或作带状伸展，或作块状集聚，并随本身的成长而逐步扩展。随着集镇面积的扩大和常住人口的增加，暴露了无规划的快速自然扩张在自建房屋质量方面的严重问题，在汶川地震、玉树地震、长宁地震以及土耳其地震[1]等地震灾害中，集镇均出现了严重的房屋倒塌和大量人员伤亡现象。

集镇的形态和经济职能兼有乡村和城市2种特点，在发展过程中受种种因素制约，在抗震设防和震灾风险方面存在“城市高风险”兼“农村不设防”的状况，掌握集镇的空间分布和规模大小，不仅对破坏性地震发生后的灾害快速评估结果有影响，还对现场应急救援的重点目标有指导意义[2-5]。因此，在面对当前更高要求的地震应急救援新形势下，掌握详尽的救援目标和精确的评估结果仍是地震应急工作的重点。

大数据时代的到来、地理时空大数据方面的研究兴起[6-10]，为集镇识别、边界提取提供了新方法途径。兴趣点（point of interest，POI）作为地理时空大数据的一种，具有空间定位特征，且拥有数据量大、实效性强、覆盖广、精度高、更新快、易获取、成本低等特点[11]，被广泛应用于城市边界提取[12-17]、城市功能区

项目来源：中国地震局地震应急青年重点任务（CEAEDEM20240212）；宁夏自然科学基金资助项目（2023AAC03807、2022AAC03701）；基于遥感和人工智能的承灾体信息获取技术创新团队项目（CX2023-4）。
第一作者简介：余思汗（1992—）工程师，主要从事地震应急技术、遥感与GIS应用研究工作，E-mail：251400967@qq.com。
通信作者：蔡保祥（1975—），正高级工程师，主要从事大地测量与遥感应用工作，E-mail：973111372@qq.com。

识别[18-22]、城市空间结构[23-25]、人口活动研究[26-27]、时空演变[28-29]等方面研究。但是利用POI数据在集镇识别、边界提取方面的研究鲜见。宁夏平罗县作为一个以农业经济发展为主的农业县，集镇对当地乡村的生产、生活服务具有重要意义，将平罗县作为集镇识别、边界提取的研究对象具有很好的代表性。本文利用核密度分析、空间关联分析、统计学、加权叠加分析、自然间断点分级法（Jenks）等手段，根据POI要素的空间集聚特征，对POI数据进行分析，通过识别集镇、提取边界，以期为基于POI数据的集镇识别、边界提取提供新的研究思路和方法，为破坏性地震灾害应急救援和快速评估提供可靠的数据支撑。

1　研究区与数据

1.1　研究区域概况

本文研究区为宁夏回族自治区石嘴山市平罗县（图1），地处宁夏北部，辖13个乡镇144个行政村，总人口31.6万人（截至2022年）。从地质地貌来看，平罗县具有贺兰山山地、贺兰山东麓洪积扇冲积平原、黄河冲积平原和鄂尔多斯台地4种类型的地形地貌，处于著名的银川断陷盆地之北部，新构造运动强烈，存在贺兰山东麓断层和正谊关断层2条裸露断层，也存在芦花台隐伏断层、银川隐伏断层北段以及正谊关断层隐伏段3条隐伏断层，地震活动频繁，历史上曾经发生过1739年银川—平罗8级大震，有发生大地震的背景。

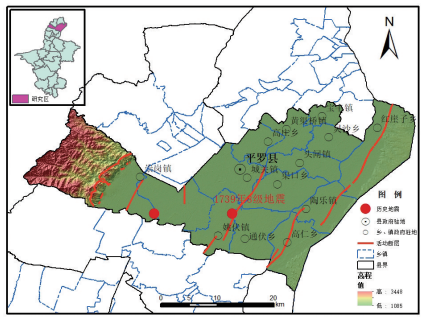


图1　研究区概况图

审图号：宁S(2022)第001号

1.2　数据准备

本文所用POI数据来自宁夏地理空间基础信息共享库（2021版）。数据库内的POI数据全部放置在同一图层内，因此采用研究区空间位置选择的方法，获得研究区内POI数据共计9 060条，坐标系为CGCS2000，为方便后续的空间计算，将所有的要素定义为投影坐标系。值得注意的是，该POI的每一类数据都有其特定地理国情代码，因此接下来的分类工作将会十分便捷。根据集镇在其服务范围内的居民教育、生产生活、医疗、娱乐等职能，本文将POI数据分为公共服务、生活居住、休闲娱乐、商业金融和农业服务5类，表1为数据分类结果。

表1　研究区POI数据分类

数据分类	POI内容	数量/个	占比/%
公共服务	教育、医疗、政府单位、公厕、交通设施等	1 825	20.14
生活居住	住宅小区、公寓、宿舍、购物、餐饮等	5 858	64.66
休闲娱乐	公园广场、娱乐场所、运动场馆、电影院等	259	2.86
商业金融	银行、公司企业、ATM机、保险、金融等	961	10.61
农业服务	批发市场、良种、农资、兽医、农业合作社等	157	1.73
合计		9 060	100

2　研究方法

2.1　技术线路

本文研究技术线路主要是分3个方面。首先对兴趣点与研究区空间位置选择，得到研究区内的POI数据，依据研究对象的职能，将数据进行分类；再利用核密度分析得到POI数据的聚集态势，通过分析集镇对各类数据的依赖性，确定权重值，进行加权叠加分析；最后将自然间断点分级法结果和等值线数据相结合，确定集镇边界等值线的核密度值。技术线路如图2所示。

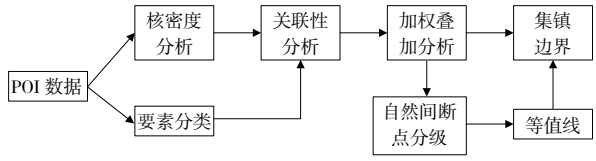


图2　研究区集镇识别与边界提取技术线路

2.2　核密度估计方法

核密度分析法是探析研究区域内POI数据聚集区的重要方法，能够连续模拟空间点要素分布在周围邻域的密度。其原理是基于地理学第一定律，距离越近的点联系度越高。该方法就是设定每个点要素上方均覆盖着一个曲面，在点所在位置处表面值最高，随着与点的距离的增大表面值逐渐减小，在与点的距离等于搜索半径的位置处表面值为零，计算公式如下：

$$P_i = \frac{1}{n\pi h^2} \times \sum_{j=1}^{n} K_j \left(1 - \frac{D_{ij}^2}{h^2}\right)^2 \qquad (1)$$

式中，P_i 为空间内的任意一点 i 的估计密度值；n 为带宽内的POI的总数；h 为规定的带宽；K_j 为研究对象 j 点的权重。前面所提到的搜索半径即带宽 h，也称为距离衰减阈值，该值的设定直接影响到核密度分析结果，为避免人为主观误判等因素，因此，采用了Silverman经验法则计算确定阈值。该法则能够根据数据自身的空间分布情况，计算所有点的平均中心之间的距离和这些距离的中值 D_m，确定标准距离 SD，

得到 h 的最优阈值。计算公式如下：

$$h=0.9\times\min\left(SD,\sqrt{\frac{1}{\ln(2)}}\times D_m\right)\times n^{-0.2} \quad (2)$$

式中，SD 为标准距离；D_m 为中值距离；方程的min部分表示将使用两选项中计算得出值较小者，确定了带宽 h 后，便可以代入式（1）进行计算估计密度值。

2.3 集镇识别方法

加权叠加是常用的叠加分析方法之一来解决多准则问题。在加权叠加分析中，每个条件的重要性不相同，可根据分析对象与区域的关联度情况，赋予每类数据不同的优先等级，且每一个优先等级值是具有相对性的。本文将POI数据分为5个大类（公共服务、生活居住、休闲娱乐、商业金融和农业服务），每一类数据和集镇的关联度都是不同的，这取决于集镇居民对各类要素功能的依赖程度，根据重要性的大小，确定每类数据的影响权重，赋值相应的优先等级，最终进行加权叠加分析，将分析结果用于集镇识别。依据集镇供应乡村所需的生产资料和生活资料，以及农产品贸易和自身需求方面的考虑，来确定5类POI要素的权重等级。按照5类数据核密度分析结果，进行加权叠加分析（图3所示），初步得到研究区内集镇的空间分布情况。

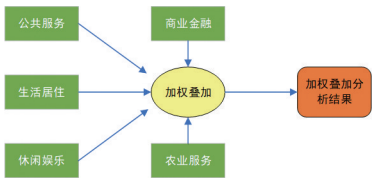


图3 加权叠加分析原理

2.4 集镇边界提取方法

自然间断点分级法是数据特定的分类方法，基于数据中固有的自然分组，对分类间隔加以识别，进行最恰当的分组，使得各个类之间的差异最大化。集镇中心向农村过渡过程中POI会逐渐稀少，聚集程度也随之下降，使用自然间断分级法能够在数据值差异相对较大的地方设置其边界。再利用核密度生成等值线的方法来提取集镇边界，等值线的边界处核密度值应遵循自然间断分级的边界，基于统计学原理来确定核密度值，所确定的等值线曲线即为集镇边界。

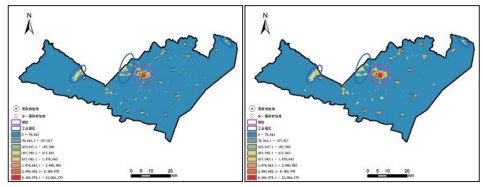
3 结果与分析

3.1 集镇识别结果分析

集镇识别需要不同职能POI的差异服务来确定。本文将POI分为5类，通过集镇对每一类服务的依赖性确定权重，从产业聚集度来讲，依赖性大小和POI类型在集镇范围内存量占比是正相关的，只考虑集镇的依赖性的情况下，将权重分为5个等级，即1~5级。利用遥感影像解译出5个已知集镇的边界范围，结合实际调查结果确定解译结果的正确性，并分别统计这5个集镇内的不同类型的POI，采用区位熵的计算方法，得到集镇与每类POI依赖程度，区位熵计算公式如下：

$$E_i=\frac{\left(\frac{t_i}{\sum_{i=1}^{m}t_i}\right)}{\left(\frac{T_i}{\sum_{i=1}^{m}T_i}\right)} \quad (3)$$

式中，E_i 为 i 类POI数据的区位熵值；t_i 为已知集镇内 i 类POI数据的数量；T_i 为研究区域内 i 类POI数据的总数；m 为POI数据分类的类别总数，在本文中 m 等于5。计算出要素区域内的空间分布特征，结果是农业服务类指数最大，休闲娱乐类的指数最小，按照区位熵值的大小，分别对应5个权重等级，再利用加权叠加分析得出空间分布结果。图4a为POI数据经核密度分析后的结果，图4b为经区位熵计算后赋值权重的加权叠加分析结果，可以明显发现在同一尺度下的渲染结果，城区和工业园区和密度值变化不大，甚至减小，而集镇的核密度全部有较大幅度增加，说明采用区位熵结算结果赋值权重等级是一种有效的方法，再利用加权叠加分析能够很好地识别出集镇的空间分布情况。通过乡镇政府驻地的定位点也能够说明识别的精度，另外，有部分集镇是没有在乡镇政府驻地的，经过进一步的调查发现，这些地方一部分属于曾经的乡镇政府驻地，后经改革合并入其他乡镇，从而取消了乡镇建制，但是集镇的功能却被完整地保留了下来；另一部分由农贸集市发展到集镇，兼顾了集镇的基本职能，因此也作为集镇被识别。



a 核密度分析结果 b 加权叠加分析结果

图4 研究区内POI聚集分析结果

审图号：宁S(2022)第001号

3.2 集镇边界提取结果分析

利用加权叠加后的核密度结果生成等值线，确定合适的核密度值等值线作为集镇边界。通过自然间断分级法将核密度结果分成8类，一共是9个密度间隔值，去掉2个极值，即密度值为0和23 884 170的值，剩下的值根据统计学原理取中位数417 043作

为等值线间距，并生成等值线。由集镇中心向边缘地带过渡，核密度值会逐渐递减，因此生成的等值线半径也会越来越大，且基本上处于均匀递增的过程，当到达某一值时，等值线范围会突然打破原有的均匀扩大，半径也随着突增，这个时候的保持均匀扩大的最后等值线便是集镇的边界，且核密度值唯一（图5所示），确定集镇边界的核密度值时，将小于或大于此值的等值线全部删除后，保留的即为集镇边界线。

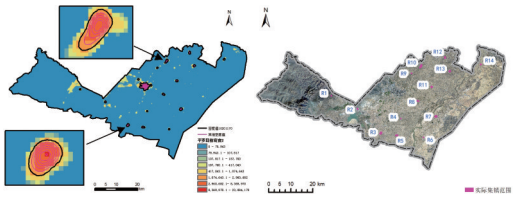


图5　等值线集镇边界　图6　遥感影像提取实际集镇范围

审图号：宁S(2022)第001号

为了检验提取集镇边界的精度，采用与集镇实际边界比对的方法。利用遥感影像目视解译，提取出集镇的实际边界（图6所示），然后再将核密度提取的等值线边界转为面要素，将解译的实际边界与等值线边界做相交，得到二者的交集部分，计算核密度边界与实际边界重叠度，检验空间位置的精度，另外，计算面积误差比，确定核密度提取边界范围的准确度。计算公式如下：

$$O_i = \frac{A_i}{\left(B_i + C_i - A_i\right)} \tag{4}$$

$$D_i = \frac{\left(B_i - C_i\right)}{C_i} \tag{5}$$

式中，O_i 为 i 处集镇的重叠度 D_i 为 i 处的面积误差比；A_i 为 i 处重叠部分的面积；B_i 为实际边界面积；C_i 为提取边界面积。根据表2结果来看，重叠率最高为87.28%，最低仅为51.76%，绝大多数的重叠度都处于70%~80%之间，平均重叠度为73.81%，说明根据核密度提取的集镇边界和实际边界在空间位置上具有较好的一致性；面积误差最大的为31.56%，最小的为−7.15%，绝大多数的误差比都为负值，说明核密度提取集镇边界范围基本上都存在高估的情况，总体来说提取边界范围和实际范围误差比较小。

表2　集镇边界的重叠度和误差

地点	核密度提取面积/㎡	集镇实际面积/㎡	重叠度/%	面积误差/%
R1	336 659	291 775	78.29	−13.33
R2	790 691	1 040 251	77.15	31.56
R3	1 031 016	862 634	73.60	−16.33
R4	246 286	228 686	82.14	−7.15
R5	572 076	475 824	64.47	−16.83
R6	289 049	257 375	71.38	−10.96
R7	1 577 330	1 739 776	87.28	10.30
R8	647 190	549 891	73.81	−15.03
R9	355 835	301 589	71.29	−15.24
R10	1 351 254	1 148 516	72.84	−15.00
R11	674 111	537 818	72.82	−20.22
R12	796 690	726 847	84.16	−8.77
R13	733 246	585 667	72.32	−20.13
R14	475 133	354 821	51.76	−25.32
平均			73.81	

4　结　论

本文基于POI数据，利用核密度分析和加权叠加分析方法，将POI数据分为5类，通过计算区位熵来确定集镇与每类要素的关联度，成功识别了研究区内所有集镇，并使用确定等值线集镇边界密度值的方法，提取集镇边界，得到了研究区内所有集镇的空间分布和边界范围结果。经过对结果的分析，提取的集镇边界范围与实际范围在空间位置上具有较好的一致性，在边界范围大小上也比较接近。

POI数据在集镇识别和边界提取上的应用，能够以低成本、高效率提取大范围内的集镇边界，为POI数据在集镇边界研究带来新视角。同时根据POI数据易获取、更新快等特点，能够动态掌握集镇时空变化，为破坏性地震发生后的灾害应急快速评估和应急救援提供最新可靠的数据，从而为政府应急处置决策与救援力量部署提供科学支撑，最大程度地减轻破坏性地震带来的灾害损失。

参考文献

[1] Chen W K, Rao G, Kang D J, et al. Early Report of the Source Characteristics, Ground Motions, and Casualty Estimates of the 2023 Mw 7.8 and 7.5 Turkey Earthquakes[J]. Journal of Earth Science, 2023, 34(2): 297-303

[2] 聂高众, 安基文, 邓砚. 地震应急灾情服务进展[J]. 地震地质, 2012, 34(4): 782-791

[3] 余思汗, 雷启云, 王银, 等. 高精度静态人口空间分布研究: 以银川市西夏区为例[J]. 震灾防御技术, 2020, 15(4): 757-766

[4] 陈文凯, 周中红, 张灿, 等. 新一代区域地震灾害快速评估系统设计与实现以甘肃省为例[J]. 地震工程学报, 2020, 42(6): 1683-1692

[5] Yu S H, Xie X F, Du P, et al. A Method for Rapidly Determining the Seismic Performance of Buildings Based on Remote-sensing Imagery and Its Application[J]. Advances in Civil Engineering, 2022, Article ID 5760913, 15 pages

[6] 余思汗, 陈文凯, 张楠, 等. 基于时空大数据的城市地震应急避难场所布局研究: 以银川市西夏区为例[J]. 华北地震科学, 2022, 40(3): 8-15

[7] Liu Y, Liu X, Gao S, et al. Social Sensing: A New Approach to Understanding Our Socioeconomic Environments[J]. Annals of the Association of American Geographers, 2015, 105(20): 512-530

[8] Long Y, Liu X J. Automated Identification and Characterization of Parcels with OpenStreetMap and Points of Interest[J]. Environment and Planning B: Planning and Design, 2016,43(10):341-360

[9] Liu K, Yin L, Lu F, et al. Visualizing and Exploring POI Configurations of Urban Regions on POI-type Semantic Space[J]. Cities, 2020,99(12):102 610

[10] Zhai W, Bai X Y, Shy Y, et al. Beyond Word2Vec: An Approach for Urban Functional Region Extraction and Identification by Combining Place2vec and POIs[J]. Computers, Environment and Urban Systems,2019,74(8):1-12

[11] 陈德权,兰泽英.基于POI数据的城市中心体系识别与边界提取:以长沙市中心六区为例[J].现代城市研究,2020(4):82-89

[12] 王方民,骆畅,杨朝现,等.基于兴趣点密度与城市扩张曲线的城市建成区边界识别[J].西南大学学报(自然科学版),2021,43(12):115-126

[13] 刘金花,张家玮,贾琨.基于POI数据的中心城区边界识别与空间格局优化:以高唐县为例[J].城市发展研究,2021,28(6):74-83

[14] 张慧馨,张丽君,秦耀辰,等.基于POI大数据的开封市城市功能扩张与演替特征分析[J].地球信息科学学报,2023,25(3):560-572

[15] 方斌,张宪哲,杨柳.基于POI数据的城市边界变化提取研究:以山西运城市城区为例[J].现代测绘,2017,40(5):20-22

[16] 李娜,吴凯萍.基于POI数据的城市功能区识别与分布特征研究[J].遥感技术与应用,2022,37(6):1 482-1 491

[17] Tannier C, Thomas I. Definingan Characterizing Urban Boundaries: A Fractal Analysis of Theoretical Cities and Belgian Citie[J]. Computers, Environment and Urban Systems,2013,41(9):234-248

[18] 苟爱萍,张振,王江波.基于POI视角的上海城市功能区演化特征及驱动因素[J].资源开发与市场,2022,38(9):1 052-1 062

[19] 胡晓鸣,黎小龙,蔚芳.基于POI的城市功能区及其混合度识别研究:以重庆市核心城区为例[J].西南大学学报(自然科学版),2021,43(1):164-173

[20] 李欣.基于POI要素空间聚集特征的城市多中心结构识别:以郑州市为例[J].北京大学学报(自然科学版),2020,56(4):692-702

[21] Wang Y, Huang H, Yang G, et al. Ecosystem Service Function Supply-Demand Evaluation of Urban Functional Green Space Based on Multi-source Data Fusion[J]. Remote Sensing,2023,15(4):118

[22] 李莹,涂志德,刘艳芳,等.基于时间序列相似度的城市功能区识别研究[J].地理空间信息,2021,19(1):22-29

[23] 薛冰,肖骁,李京忠,等.基于兴趣点(POI)大数据的东北城市空间结构分析[J].地理科学,2020,40(5):691-700

[24] 于丙辰,刘玉轩,陈刚.基于夜光遥感与POI数据空间耦合关系的南海港口城市空间结构研究[J].地球信息科学学报,2018,20(6):854-861

[25] 刘彤,周伟,曹银贵.沈阳市城市功能区分布与人口活动研究[J].地球信息科学学报,2018,20(7):988-995

[26] 淳锦,张新长,黄健锋,等.基于POI数据的人口分布格网化方法研究[J].地理与地理信息科学,2018,34(4):89-95

[27] Li J, Long Y, Dang A. Live-work-playcenters of Chinese Cities: Identification and Temporal Evolution with Emerging Data[J]. Computers, Environment and Urban Systems, 2018(71):58-66

[28] Jiang S, Alives A, Rodrigues F, et al. Miningpoint-of-interest Data From Social Networks for Urban land Use Classification and Disaggregation[J]. Computers, Environment and Urban Systems,2015(71):36-46

[29] Atsuyuki O, Toshiaki S, Kokichi S H. A Kernel Density Estimation Method for Networks, Its Computational Method and a GIS-based Tool[J]. International Journal of Geographical Information Science,2009,23(1):7-32

(上接第82页)

石力学与工程学报,2007(3):433-454

[3] 李益敏,邓选伦,杨一铭,等.基于GIS的怒江州建设用地生态适宜性评价[J].地理空间信息,2023,21(9):27-30

[4] Corominas J, Van W C, Frattini P, et al. Recommendations for the Quantitative Analysis of Landslide Risk[J]. Bull Eng Geol Environ,2014,73(3):209-263

[5] Pourghasemi H R, Rahmati O. Prediction of the Landslide Susceptibility:Which Algorithm, Which Precision? [J]. Catena,2018,162(10):177-192

[6] Li L, Lan H, Guo C, et al. A Modified Frequency Ratio Method for Landslide Susceptibility Assessment[J]. Landslides,2017,14(6):727-741

[7] 李郎平,兰恒星,郭长宝,等.基于改进频率比法的川藏铁路沿线及邻区地质灾害易发性分区评价[J].现代地质,2017,31(5):911-929

[8] 仉义星,兰恒星,李郎平,等.综合统计模型和物理模型的地质灾害精细评估:以福建省龙山社区为例[J].工程地质学报,2019,27(3):608-622

[9] Zhang Y X, Lan A N H X, Li L P, et al. Optimizing the Frequency Ratio Method for Landslide Susceptibility Assessment:A Case Study of the Caiyuan Basin in the Southeast Mountainous Area of China[J]. Journal of Mountain Science,2020,17(2):340-357

[10] Li L, Lan H. Bivariate Landslide Susceptibility Analysis: Clarification Optimization Open Software and Preliminary Comparison[J]. Remote Sensing,2023,15(5):1 418

[11] 田乃满,兰恒星,伍宇明,等.人工神经网络和决策树模型在滑坡易发性分析中的性能对比[J].地球信息科学学报,2020,22(12):2 304-2 316

[12] Gong P, Li X, Wang J, et al. Annua Maps of Global Artificial Impervious Area (GAIA) Between 1985 and 2018[J]. Remote Sensing of Environment,2020,236(17):111 510