

廖微, 姚琪, 廖林, 等, 2021. 基于地理信息系统的空间统计学方法在地震统计中的应用. 中国地震, 37(3): 695~703.

基于地理信息系统的空间统计学方法 在地震统计中的应用

廖微¹⁾ 姚琪^{2,3)} 廖林⁴⁾ 张微⁵⁾ 姜祥华²⁾ 杨文²⁾

1) 北京新能源技术研究所, 北京 102300

2) 中国地震台网中心, 北京 100045

3) 中国地震局地震预测研究所, 北京 100036

4) 中国石油天然气集团有限公司国家特聘专家工作室, 北京 102200

5) 浙江大学软件学院, 浙江宁波 315048

摘要 传统的地震统计学多局限于规则的统计范围,或是需要较多的人工干预,本文引入了基于地理信息系统的空间统计学方法,将其应用于地震空间统计中,能够高效、快速地挖掘含有地理信息的各要素之间的相互联系。本文以川滇地区 2019 年发生的 $M_L \geq 1.0$ 地震数量的空间统计为例,详述了对地震数据进行空间统计的流程。统计结果表明, $0.5^\circ \times 0.5^\circ$ 是最适合该地震数据集空间统计的矩形范围,发生地震次数最多的区域位于四川盆地西南缘的威远-长宁地区,此外龙门山断裂带、龙门山断裂带与鲜水河-小江断裂系交汇处、丽江-小金河断裂沿线以及滇南地区均为地震发生次数较高的区域。此外,还利用该方法计算了 2015—2019 年全国各省地震能量释放情况。结果显示在该时间段内,释放地震能量总量最多的省份为四川省,新疆维吾尔自治区和西藏自治区次之,而台湾地区陆地区域、四川省、西藏自治区等地区具有较高的每万平方公里地震能量释放量和每万人地震能量承受量。并通过上述应用实例,探索了地理信息系统在地震统计领域的初步应用。

关键词: 空间统计学 地震统计 地理信息系统 工作流程 川滇地区

[文章编号] 1001-4683(2021)03-0695-09 [中图分类号] P315 [文献标识码] A

0 引言

在地震监测预报领域,地震的空间统计是非常重要的内容,包括但不限于地震分布、地震频度、地震能量等研究方面,并以此延伸出了 b 值、 c 值等统计参数(国家地震局震害防御司,1992),但是大多仅根据简单规则的形状进行空间统计,譬如矩形、圆形等(吴绍春等,2006;张瑞芳等,2018),而很少能通过复杂的形状进行空间统计,或者是进行地震密度的统计过程较为繁琐,且很难对统计结果进行正确性的评估。在大数据时代,海量不同来源的各种数据需要归并到一起进行深入、快速的数据挖掘(张晁军等,2015;孙静等,2018)。尤其

[收稿日期] 2020-07-22 [修定日期] 2020-10-15

[项目类别] 地震动力学国家重点实验室开放基金项目(LED2019B08)资助

[作者简介] 廖微,男,1980 年生,副研究员,主要研究方向为地理信息系统。E-mail: cp_820@163.com

姚琪,通讯作者,女,1981 年生,副研究员,主要研究方向为地震与数值模拟。E-mail: yqvoxelgeo@163.com

各级地方政府的地震应急响应,需要在极短的时间内完成地灾信息的数据融合、可视域分析、通视分析、滑坡分析、人流扩散分析等工作,为辅助决策提供科学可靠的依据。因此,传统的地震空间统计已经无法满足地震应急的高效、可视化需求。

国外地理信息系统 (Geographic Information System, GIS) 包括 Mapinfo、Envi、ArcGIS、GeoSence、Google Earth 等,国内学者也针对不同的用途开发了各自的地理信息系统软件,譬如地质调查系统的 MapGIS,李胜乐等(2004)研发的 MapSIS,廖微等(2015)开发的应急地理信息系统等。这些地理信息系统具有强大、高性能的大数据挖掘能力,通过高效的分析,能够让数据的价值和潜能得以发掘,并通过可视化进行展示。基于地理信息系统的数据挖掘和大数据分析在地学的各个领域均得到了较好的应用,包括物流运输设计、城市规划、区域地质调查、卫星遥感等方面(杨凡等,2019)。尤其在 2020 年新型冠状病毒肺炎疫情防控期间,GIS 云平台 and WebGIS 功能在疫情统计和实时发布方面起到了较大的作用。目前,地理信息系统在地震监测预报领域尚处于初步应用阶段,主要作用是展示地震发生的地点,或是作为监测数据处理的某一参数(李志杰等,2020),对其强大的空间统计功能尚无大规模应用。

通过空间位置建立各要素之间的统计关系,能够解决地震信息要素数据中隐含的空间不确定因素,用统计方法从凌乱的地理空间数据中发掘地理空间变化规律(刘湘南等,2015)。本文以川滇地区为例,详述了利用地理信息系统软件对该地区 2019 年地震个数进行统计的流程,并展示了 2015—2019 年期间全国各省每万平方公里地震能量释放量和每万人地震能量承受量统计的结果,对地理信息系统空间统计在地震统计学的应用进行了初步展望,以期为各级政府的应急预案配比提供更直观、更符合需求的依据。

1 空间统计流程

利用中国地震台网中心提供的正式目录,以川滇地区(21°~34°N,97°~107°E)2019 年发生的 $M_L \geq 1.0$ 地震为例,对该地区的地震发生次数进行统计,具体流程如图 1 所示。本文基于 ArcGIS 软件,详细阐述基本的地理信息系统空间统计流程。大部分地理信息系统软件和程序都有类似的功能,可以直接实现,或是经过二次开发实现。

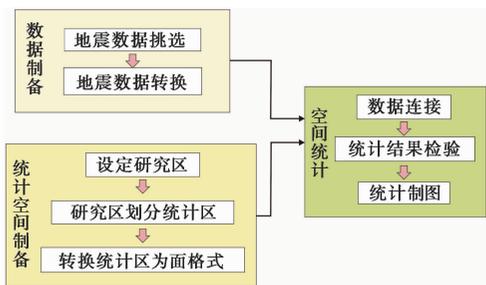


图 1 空间统计流程图

1.1 数据制备

GIS 软件平台界面普遍比较友好,能够接受多种格式的数据,对大部分点格式数据来说,CSV、Excel 或者 txt 格式均能方便地转换为 GIS 接受的格式。点数据格式包括地震的时间、

震级、经度、纬度、深度、地点等空间信息和属性信息。通过数据挑选,正式目录中川滇地区 2019 年 $M_L \geq 1.0$ 地震个数为 69570 个。

有 2 种常用的将点格式数据转换为 GIS 接受的格式的方式,一种是在 ArcGIS 的地图生成程序 ArcMap 中直接加入 txt 或 CSV 文件,然后在其中进行转换;另一种是在 ArcGIS 的数据库管理程序 ArcCatalog 中进行转换。本文推荐将地震文件用 Excel 软件或其他表格软件进行转换,增加一行来指代属性名称,之后再在 ArcCatalog 中进行 ArcGIS 格式的转换。这样有助于指定地震的属性信息,也不容易在指定空间坐标时发生经纬度指向错误的情况。不论是何种格式数据的加载,如果不增加一行属性名称,那么第一行数据将会被认为是属性名称,从而造成数据丢失的情况。此外,由于大部分地震数据为经纬度格式,因此本文采用 WGC-84 坐标系统。

数据制备具体流程如下:①在 ArcCatalog 中,点击地震数据的 Excel 文件右边的“+”,显示 Excel 文件下属的 sheet 1 表格文件;②右键点击此文件,在弹出菜单中选择“Create feature class”中的“From XY Table...”;③按照弹出对话框的提示,选择空间坐标和坐标系统,指定存储位置和文件名称。其中需要注意的是,一般情况下, X 坐标为经度坐标, Y 坐标为纬度坐标。地震数据制备结果如图 2 所示。

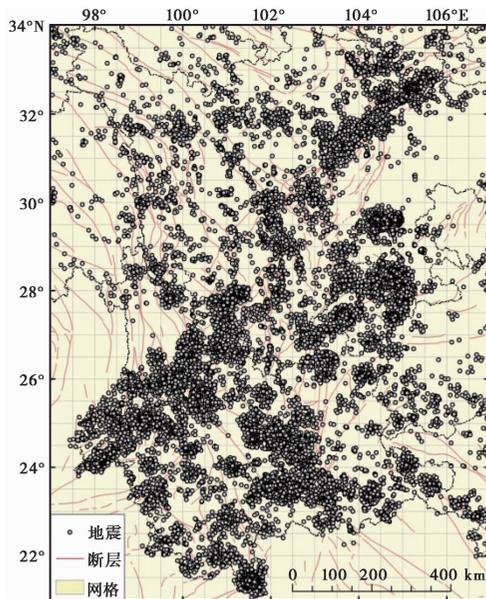


图 2 2019 年 $M_L \geq 1.0$ 地震分布和用于统计的 $0.5^\circ \times 0.5^\circ$ 网格

1.2 统计空间制备

ArcGIS 的统计空间可以为典型的矩形,或者不规则的几何形态,但必须是面要素。为了与典型的地震统计结果进行对比,川滇地区采用 $0.5^\circ \times 0.5^\circ$ 的矩形进行计算(图 2)。

统计空间制备具体流程如下:①在 ArcCatalog 中,右键单击文件夹,选择“New”中的“shapefile...”,按照弹出的对话框分别选择新建文件的名称、类型(Polygon)、空间坐标系统(最好选择与地震文件相同的坐标系统);②在 ArcMap 中,将新建统计空间的 Polygon 文件加载,点击工具栏中“Editor”中的“Start Editing”,新建面文件,并对每个点的经纬度根据研

究区的范围进行指定,然后保存改动并停止编辑;③由于川滇地区的统计区域为矩形,本文用“Fishnet”工具对其进行网格划分,打开 ArcToolbox,依次选择“Data Management Tools”、“Sampling”、“Create Fishnet”,在弹出的对话框中,按照要求输入要建的网格名称,选择研究区范围,输入网格大小 $0.5^{\circ} \times 0.5^{\circ}$;④将生成的网格转化为面,在 ArcToolbox 中依次选择“Data Management Tools”、“Features”、“Feature to Polygon”,选择文件进行转换。

1.3 空间统计

本文所用的 ArcGIS 空间统计是通过文件之间的“Join”来实现的。ArcGIS 用来实现文件关联的方式很多,如 Join、Relate、QueryLayer、ArcSDE 视图、关系类等。其中“Join”的作用是将一个要素类数据与一个普通表进行关联,一般常规是使用“Join”将另一个表的属性标注显示到面文件上,既可以实现“一对一”对应,也可以实现“一对多”、“多对一”、“多对多”等对应方式,但是对于输入表中的每条记录,输出表中必须有且只能有一条记录。使用“Join”进行连接。利用“Join”方式进行关联,不仅可以实现表与表之间某属性的连接,也可以实现利用空间位置进行表与表之间的连接。本文通过统计网格的空间位置与地震震中位置进行连接,统计每个网格中的地震个数。

空间统计具体流程如下:①在 ArcMap 中加入地震数据文件和统计网格文件;②右键点击统计网格文件,选择“Properties”,打开属性框;③选择“Joins&Relates”页面,点击“Add...”,在第一个“What do you want to join to this layer”的问句下选择“Join data from another layer based on spatial location”,然后选择需要连接的地震文件,选择需要统计的功能,例如平均值、最大值、最小值、标准差、方差等,由于需要统计地震的总个数,因此选择“Sum”;④连接生成后,会独立生成一个连接的文件,右键选择“Open attribute table”,就可以看到每一个网格后均增加了若干属性,其中“Count”一栏显示了每个网格中包含的地震个数。统计结果显示,川滇地区 $0.5^{\circ} \times 0.5^{\circ}$ 的网格内,2019 年 $M_L \geq 1.0$ 地震最少发生了 0 次,最多发生了 16831 次,平均每个网格发生 139 次,发生最多次地震的区域位于四川盆地西南缘的威远-长宁地区,此外汶川余震区、三岔口地区、丽江-小金河断裂沿线以及川滇菱形块体南缘均为地震频发的区域(图 3)。

1.4 统计检验

由于地震分布具有不均匀性,某些丛集的地震在不同的划分方案中被划分到不同的网格,将会导致统计网格的大小对空间统计的结果具有决定性作用。因此,地震空间统计分析必须进行统计检验,只有统计结果达到平稳,找到统计个数与统计空间之间的平衡,才能认为统计结果是可信的。

川滇地区 2019 年 $M_L \geq 1.0$ 以上地震最大网格内统计个数、平均统计个数与网格大小之间的关系如图 4 所示。网格内地震统计个数的平均值曲线基本较为连续,随着网格的增大而不断增大。这是由于平均值是对所有网格内地震个数的平均,弱化了网格内个数的差异。然而地震统计个数的最大值曲线并不平直,统计网格从 0.1° 增至 0.4° ,网格中所统计到地震个数的最大值随之增大,然而在网格边长从 0.5° 增至 1.5° 区间,最大个数曲线趋于平缓,且在网格大小为 0.6° 和 1.3° 时出现了相对低值。事实上,在本文的统计时间段内,川滇交界东侧的长宁-威远地区出现了显著的小震丛集现象,统计网格的边界是否横跨小震丛集区,对最大个数的影响较大。因此最大个数曲线的不平整是由于小震的丛集造成的。由于统计的目标是地震个数在区域的分布情况,因此我们认为最大个数曲线在 0.5° 网格时出现了拐点,

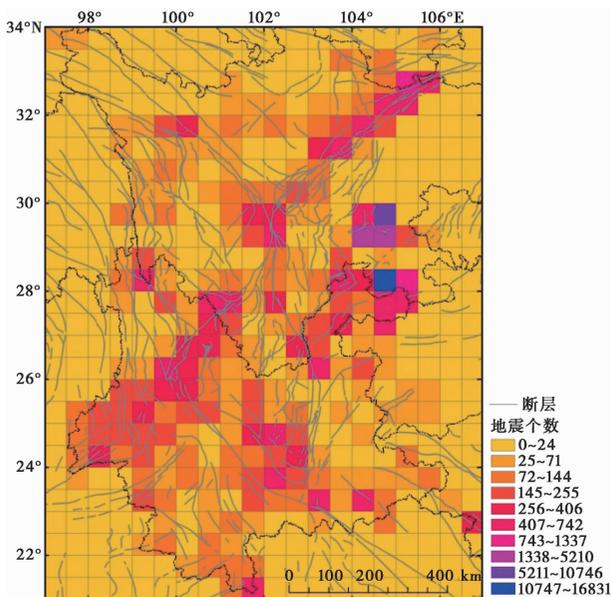


图 3 0.5°×0.5°网格统计地震个数的结果

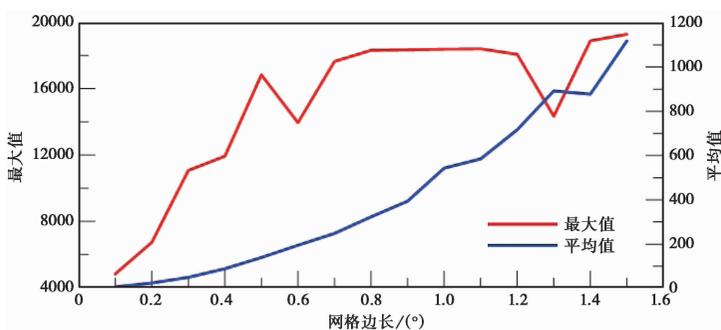


图 4 网格内地震统计个数最大值、平均值与网格大小的关系

选用边长为 0.5°的正方形网格进行网格统计最能够反应地震个数在空间上的分布(图 3)。

2 2015—2019 年全国各省年均地震能量释放密度

社会财政、政策、决策计划一般 3 年或 5 年为一个周期,并大多以省为单位进行统筹安排,因此,我们计算了 2015—2019 年全国各省的地震情况。由于地震不同震级释放的能量不一样,不同省的面积大小也不同,如果机械地按照地震个数来统计,则无法凸显出受地震灾害严重的区域,而如果忽略统计区域的面积,则会造成受灾面积显示过量的问题。因此,我们统计了 2015—2019 年期间,全国各省市 $M_s \geq 1.0$ 每万平方公里地震能量释放释放。计算公式如下

$$\lg E = 4.8 + 1.5M \tag{1}$$

$$\rho_E = \frac{\sum E}{A} \tag{2}$$

式中, E 为地震能量,单位为J; M 为 M_s 震级; A 为统计区面积,单位为 km^2 ; ρ_E 为统计区的每万平方公里地震能量释放。式(1)通过地震的震级可计算每个地震所释放的能量,这里采用经验公式(国家地震局震害防御司,1992)。公式(2)则根据地震的能量和次数,计算了统计区内的面积单位能量释放密度。2015年1月1日至2019年12月31日全国共有 $M_s \geq 1.0$ 地震8304个,其中最大地震为2017年8月8日九寨沟 $M_s 7.1$ 地震,此外中国大陆有9个6级以上地震,台湾地区陆地区域有1个6级以上地震。

利用全国各省和直辖市边界的面数据作为空间统计范围,计算所得各省份5年地震释放的总能量、面积能量密度,如图5所示。显然,四川省释放地震的总能量最多,新疆维吾尔自治区和西藏自治区次之,表明这几个省具有较高的地震活动背景,这与地震活动水平是相当的,主要与地震个数和震级有关。在统计的时间窗和空间范围内,四川省的绝大部分地震能量来自于2017年九寨沟 $M_s 7.1$ 地震和2019年长宁 $M_s 6.0$ 地震,新疆维吾尔自治区在2015至2019年发生3个6级以上地震,其中2015年皮山 $M_s 6.5$ 地震和2017年精河 $M_s 6.6$ 地震均为 $M_s 6.5$ 以上的中强地震。西藏自治区在统计的时间段内则发生了2017年米林 $M_s 6.9$ 地震和2019年墨脱 $M_s 6.3$ 地震。

将地震能量与各省的面积相除后,得到了每万平方公里地震能量释放量。结果显示,四川省仍然为每万平方公里地震能量释放量最高的省份,台湾地区陆地区域也变得更为显著。台湾地区陆地区域的地震总能量在统计的时间段内并不突出,仅发生2016年高雄 $M_s 6.7$ 地震,由于岛区面积狭小,每万平方公里地震能量释放量则显得较为突出。另一个较为显著的特点是统计时间段内青海省的每万平方公里地震能量释放量小于甘肃省。在统计时间段内,虽然青海省境内发生了2个6级以上地震,而甘肃省境内未发生6级以上地震,但两省之间显著的面积差异,导致了青海省的每万平方公里地震能量释放量相对较低。

将地震释放能量总量与各省份常驻人口数据相除后,可以展示各省份每万人地震能量承受量,这对于地震应急预案的设计具有重大意义。西藏自治区由于人口稀少而地震较多,其每万人地震能量承受量较大。四川省作为全国地震能量释放集中省份,由于人口众多,每万人地震能量承受量较高,但仍小于西藏自治区。新疆维吾尔自治区和青海省在统计时间段内均发生多次6级以上中强震,且人口密度较小,因此每万人地震能量承受量与这一时间段地震活动最强的四川省处于同一量级。

总体来说,无论是地震释放能量总量、每万平方公里地震能量释放量,还是每万人地震能量承受量,均为地震活动的统计结果,其结果主要还是依赖于地震活动的强弱。但是不同的空间统计方法和空间统计结果,能够挖掘出更有社会意义的信息。

3 讨论

快速发展的地理信息系统为地震数据的深入挖掘提供了更多的可能,其不仅能够快速实现对传统的矩形、圆形、蜂窝型等统计形状的数据统计,也能够根据社会需求调整统计区,实现对国界、省界、研究区等不规则统计区的地震破坏的统计,以及实现对多来源地理信息数据的共同统计,如人口、交通、水文、地貌、滑坡等,均可以和地质、地震等数据进行融合统计,并实现实时更新,为政府制定政策、基建选址、防震减灾等工作提供更快速、更实时、更综合、更有效的依据。

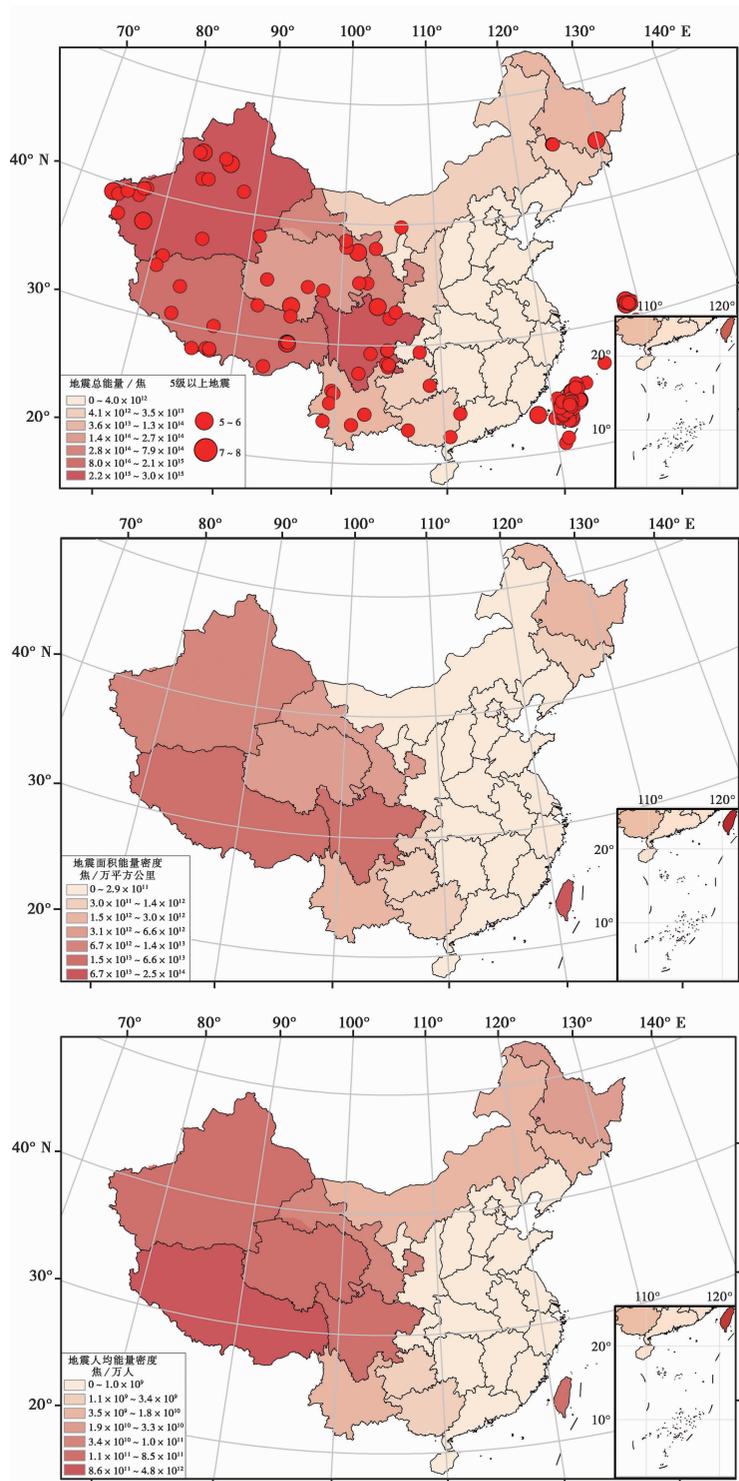


图 5 2015—2019 年地震能量总量(a)、每万平方公里地震能量释放量(b)、
每万人地震能量承受量(c)

从本文的统计过程和统计结果来看,无论是传统的规则网格,还是复杂多变的空间范围,利用 ArcGIS 的空间统计功能均能快速、高效地完成统计,并且能够对数据和统计结果进行二次加工,获得更好、更广泛的应用。不论应用何种统计方式,在数据制备和空间制备完成的基础上,通过在数据的地理信息与统计范围之间建立联系,就可以快速获得统计范围内的平均值、最大值、最小值、标准差、方差等各种常见统计结果,并能够利用 ArcGIS 的数据展示功能,实现多图层叠加和快速出图,能够在地震震后应急、灾害评估、震后政府预算编制、资源评估等工作中发挥更大的作用。

地震数据是典型的不确定性地理空间数据,这种不确定性不仅来自地震震中信息的误差,更多源自于地震的空间现象、发展过程和发展特征无法被准确确定。随着人们将专家系统、人工神经网络、模糊逻辑、遗传算法等人工智能技术(刘湘南等,2008)运用到地理信息系统中,智能化空间技术将极大地改进现今地震监测预报现状。随着技术的进一步发展,结合具体的地学知识和地理信息,以地理信息系统智能化空间分析和数据挖掘为手段,有望获得更精确的、能够反应实际地震发生规律的分析结果。

由于本文主要目标是介绍方法和展示结果,因此仅展示了利用规则网格和公开发行的地理数据作为空间统计范围的结果。从深入研究地震发生的时空规律的角度来说,应该使用更有地质意义的空间统计范围来进行统计,如板块构造边界,或地震带等。在必要的时候,需要对空间统计范围进行进一步的细化,并且在重点区域开展包括但不限于时间序列的空间统计。这些工作均可在 ArcGIS 或者其他具有空间统计功能的地理信息系统软件上实现。

4 结论

本文详细介绍了利用地理信息系统的空间统计学方法来进行地震数据空间统计的流程,并以公开发行的数据作为对象,统计了川滇地区 2019 年发生的 $M_L \geq 1.0$ 地震个数,并做了统计检验。结果表明,2019 年发生地震次数最多的区域位于四川盆地西南缘的威远-长宁地区,此外龙门山断裂带、龙门山断裂带与鲜水河-小江断裂系交汇处、丽江-小金河断裂沿线以及滇南地区均为地震发生次数较高的区域。此外,本文还展示了以不规则的行政边界为空间统计范围的例子,统计了 2015—2019 年全国各省地震能量释放情况,包括每万平方公里地震能量释放量和每万人地震能量承受量。统计结果显示了地震释放能量、统计区域面积、统计区其他相关属性(本文仅展示了人口数量的影响)三者联合计算下的统计结果,展示了地震空间统计更多的应用和发展方向。

参考文献

- 国家地震局震害防御司,1992. 地震工作手册. 北京:地震出版社.
- 李志杰,陈智勇,2020. 基于空间相关的地震前兆数据库的信息提取与数据更新方法. 地震工程学报,42(1):123-128.
- 李胜乐,陆远忠,车时,2004. MapInfo 地理信息系统二次开发实例. 北京:电子工业出版社.
- 廖微,王少华,彭向军,等,2015. 应急指挥地理信息系统的设计与实现. 电子设计工程,23(3):54-56.
- 刘湘南,黄方,王平,等,2008. GIS 空间分析原理与方法. 北京:科学出版社.
- 孙静,李亚龙,万杰,等,2018. 大数据在地震科研中的应用探讨. 电脑知识与技术,14(35):3-5.
- 吴绍春,胡华山,2006. 一个地震数据挖掘网格及其作业调度. 上海大学学报(自然科学版),12(5):529-533.

杨凡,陈雅慧,李华玥,2019. 利用 GIS 研究延怀盆地地震与断裂关系. 地震地磁观测与研究,40(4):16~22.

张晁军,陈会忠,李卫东,等,2015. 大数据时代对地震监测预报问题的思考. 地球物理学进展,30(4):1561~1568.

张瑞芳,黄泽纯,洪安东,等,2018. 利用样方分析小微地震空间分布模式. 测绘与空间地理信息,41(6):21~24.

Application of Spatial-Based Statistics on Geographic Information System in Seismic Statistics

Liao Wei¹⁾ Yao Qi^{2,3)} Liao Lin⁴⁾ Zhang Wei⁵⁾ Jiang Xianghua²⁾ Yang Wen²⁾

1) Beijing Institute of New Energy Technology, Beijing 102300, China

2) China Earthquake Networks Center, Beijing 100045, China

3) Institute of Earthquake Forecasting, CEA, Beijing 100036, China

4) National Special Merited Experts Institute, CNPC, Beijing 102200, China

5) School of Software, Zhejiang University, Ningbo 315048, Zhejiang, China

Abstract Traditional seismic statistics are usually limited to regular statistical scope, or require a lot of manual intervention somehow. To solve this problem, in this paper we introduce the spatial statistics method based on geographic information system, and apply it to seismic spatial statistics, which can efficiently and rapidly excavate the interrelation among various elements containing geographic information. Take a case-study to demonstrate the basic and detailed technological process of spatial statistics, we analyzed the seismic data ($M_L \geq 1.0$) of the Sichuan and Yunnan region. The results show that $0.5^\circ \times 0.5^\circ$ is the most suitable scale size for this case-study, and the region with most earthquake events in 2019 is the Weiyuan-Changning region on the southwestern margin of the Sichuan Basin. In addition, the Longmenshan region, junction area of the Longmenshan fault zone and Xianshuihe-Xiaojiang fault zones, Lijiang-Xiaojinhe region, and the southern margin of the Yunnan Province have relative high rate of earthquake occurrence in 2019. Besides, we applied the method to statistically study the earthquake energy released of every province during 2015—2019. We found that the Sichuan Province has released the largest amount of seismic energy, Xinjiang and Xizang Autonomous Region has gone through the second largest amount energy. Besides the Sichuan Province, the land area of Taiwan, Sichuan Province and Xizang Autonomous Region have experienced very high seismic energy per 10^4 square kilometers by area or per 10^4 people by population. Through above two case-studies we show the preliminary application of GIS in the earthquake statistical analysis.

Keywords: Spatial statistics; Seismic statistics; Geographic information system; Workflow; Sichuan and Yunan region