

潘 健,毕硕本,沈 香,等. 江苏省雷电强度等级空间分布及其与土地利用类型的关系[J]. 江苏农业科学,2017,45(8):219-223.
doi:10.15889/j.issn.1002-1302.2017.08.061

江苏省雷电强度等级空间分布 及其与土地利用类型的关系

潘 健,毕硕本,沈 香,陆 源,周 浩

(南京信息工程大学地理与遥感学院,江苏南京 210044)

摘要:首先针对雷电强度空间定量分析不足的问题,利用 R 语言对江苏省雷电强度进行聚类分析,得到江苏省 2007—2009 年夏半年雷电强度的 5 个等级,结果发现,占据雷电强度 90% 以上(31 kA 以下)的 1、2 级组内成员最多,而后 3 个等级则呈递减趋势。为了给实际雷电防护工程提供参考,对江苏省各市的雷电强度空间变化范围进行区划,得出各市相应的强度等级分区。其中,淮安市的雷电强度空间变化最突出,连云港市则是 13 个地级市中变化最不显著的。最后探讨了雷电强度等级与 6 种土地利用类型的关系,结果显示,林地对于雷电强度等级的空间分布影响最大,其余 5 种地表类型在雷电强度为 2 级时关联性最大,随着雷电强度等级的增大,雷电强度等级的空间分布与这 5 种土地利用类型的关系逐渐减弱。

关键词:雷电强度;等级;空间分布;江苏省;R 语言;聚类分析;土地利用类型

中图分类号: S161.9;F323.211

文献标志码: A

文章编号: 1002-1302(2017)08-0219-05

雷电学作为气象领域的重要分支,一直深受国内外专家学者的广泛关注。江苏省每年因雷电所引发的直接和间接的经济损失达数亿元,群众的生命财产也受到严重威胁。如何

通过科学合理的防护技术手段将这类损失降到最低,是防雷工作者所急需解决的问题。雷电具有随机性、局域性、分散性、突发性、瞬时性及三维性这些鲜明的特点^[1]。现阶段,随着数据的多元化,在通过人工观测、地基和空基等渠道所获得的大量资料支持下,对于雷电数据的分析越来越深入。Reap 等使用线性逐步回归法和模式输出统计法(MOS)分析闪电资料与套网模式(NGM)预报结果^[2-3],提出其所研究的区域阿拉斯加地区形成雷暴的先决条件是:要求存在大范围的层结不稳定以及由局地风场和湿度提供的辐合^[4]。冯桂力等

收稿日期:2016-01-15

基金项目:国家自然科学基金(编号:41071253,41271410)。

作者简介:潘 健(1990—),男,江苏南京人,硕士,主要从事气象数据挖掘等研究。E-mail:pjhuhusoso@163.com。

通信作者:毕硕本,教授,博士生导师,主要从事气象信息融合、气象史料数据挖掘、三维可视化等研究。E-mail:bishuoben@163.com。

[4] Farahani H J, Izzi G, Oweis T Y. Parameterization and evaluation of the AquaCrop model for full and deficit irrigated cotton [J]. *Agronomy Journal*, 2009, 101(3): 469-476.

[5] Andarzian B, Bannayan M, Steduto P, et al. Validation and testing of the aquacrop model under full and deficit irrigated wheat production in Iran [J]. *Agricultural Water Management*, 2011, 100(1): 1-8.

[6] Jin X, Feng H, Zhu X, et al. Assessment of the AquaCrop model for use in simulation of irrigated winter wheat canopy cover, biomass, and grain yield in the North China Plain [J]. *PLoS One*, 2014, 9(1): e86938.

[7] 杜文勇, 何雄奎, Shamaila Z, 等. 冬小麦生物量和产量的 AquaCrop 模型预测 [J]. *农业机械学报*, 2011, 42(4): 174-178, 183.

[8] 付 驰, 李双双, 李 晶, 等. AquaCrop 作物模型在松嫩平原春麦区的校正和验证 [J]. *灌溉排水学报*, 2012, 31(5): 99-102.

[9] Todorovic M, Albrizio R, Zivotic L, et al. Assessment of AquaCrop, CropSyst, and WOFOST models in the simulation of sunflower growth under different water regimes [J]. *Agronomy Journal*, 2009, 101(3): 509-521.

[10] Irmak S, Payero J O, Eisenhauer D E, et al. Watermark granular matrix sensor to measure soil matric potential for irrigation

management [R]. Nebraska: University of Nebraska - Lincoln, 2006.

[11] 杨 宁, 孙占祥, 张立桢, 等. 基于改进 AquaCrop 模型的覆膜栽培玉米水分利用过程模拟与验证 [J]. *农业工程学报*, 2015, 31(增刊1): 122-132.

[12] Hsiao T C, Heng L E, Steduto P, et al. AquaCrop - The FAO crop model to simulate yield response to water: III. Parameterization and testing for maize [J]. *Agronomy Journal*, 2009, 101(3): 448-459.

[13] Araya A, Keesstra S D, Stroosnijder L. Simulating yield response to water of Teff (*Eragrostis tef*) with FAO's AquaCrop model [J]. *Field Crops Research*, 2010, 116(1/2): 196-204.

[14] Thompson R B, Gallardo M, Agüera T, et al. Evaluation of the watermark sensor for use with drip irrigated vegetable crops [J]. *Irrigation Science*, 2006, 24(3): 185-202.

[15] 廖 江. 毛管配置模式与灌溉定额对滴灌春小麦土壤水分分布及产量的影响 [D]. 石河子: 石河子大学, 2012.

[16] 余根坚, 黄介生, 高占义. 基于 HYDRUS 模型不同灌水模式下土壤水盐运移模拟 [J]. *水利学报*, 2013, 44(7): 826-834.

[17] 马海燕, 王 昕, 张展羽, 等. 基于 HYDRUS-3D 的微咸水膜孔沟灌水盐分布数值模拟 [J]. *农业机械学报*, 2015, 46(2): 137-145.

利用 1998—2000 年山东地区雷电探测网获取的云对地闪电资料,研究山东地区闪电时空分布特征,指出闪电分布与地形和下垫面性质有关^[5];李霞等利用苏州地区 2002—2004 年闪电定位系统监测资料,分析该地区地闪的月变化、日变化以及强度、闪电密度等特征^[6];冯民学等通过对雷暴日数据和地闪数据的分析,研究了近 43 年江苏省的雷电分布特征,得出了江苏省雷电分布的总趋势并指出江苏省雷电的多发区分布特征^[7];李政在对重庆地区雷电下垫面状况的研究中分析了重庆市雷电分布的时空规律和雷电分布对应下垫面状况,但主要分析了雷电发生频次与下垫面的关系,缺乏对于雷电强度和陡度的系统分析^[8];赵伟等利用浙江省电力和气象部门的相关数据,研究了浙江省雷电的时空分布特征及影响因素,同时得出地闪高密度区的分布,与气候、地形、地貌、地面大型水体和城市热岛效应具有密切的关系^[9];宋晓爽等利用 LS800 闪电定位系统观测地闪的 2009—2011 年资料对上海及周边地区(120.0°~122.5°E、30°~32°N)的地闪活动特征进行了研究并对上海及其周边地区雷电的海陆分布差异进行了探讨^[10]。气象数据和地理数据都有着数据量大、维度高的结构特点,这使得两者毫无争议跻身大数据的潮流之中。在无先验知识的情况下,数据挖掘技术中的聚类分析方法是用于研究这类数据的有效手段^[11]。自聚类算法提出至今,国内外学者结合各自领域的研究需要提出或改进了多种聚类算法,如 K -means、OPTICS、DBSCAN、CURE、CLIQUE、DENCLUE 等聚类算法^[12-18]。聚类分析可以根据样本相似度对数据进行分组,从而发现对象空间的分布特征^[19]。本研究采用的 K -means 算法是硬聚类算法,是典型的基于原型的目标函数聚类方法的代表。R 语言是目前世界上最流行的计算统计软件之一,该语言具有强大的数学分析功能,是适用于各领域的高质量软件扩展包,具有出色的可视化效果及很好的兼容性,支持跨平台运行^[20]。本研究对江苏省闪电定位系统

数据进行降维分析,针对前人对于雷电强度定量分析研究的不足,在 R 语言环境下利用 GIS 功能,通过聚类算法对雷电强度进行等级分类,并对分类结果进行空间分布分析。本研究同时结合前人研究经验,考虑到下垫面因素对雷电强度的可能影响,对雷电强度与土地利用类型进行关系分析。

1 资料与数据处理

江苏省架设的 ADTD 雷电探测仪可提供闪电发生的经度、纬度、时间、强度、极性等资料,本研究采用该设备 2007—2009 年的地闪资料,根据雷电发生的显著季节差异,将雷电高发的 4—9 月定义为夏半年,而同年中的 1—3、10—12 月则为受雷电影响极小的冬半年。因为对应年份的地闪时空特征^[21]已经作出了详细的探讨,本试验则侧重结合数据挖掘的算法在更为高效的架构下深化对应年份夏半年雷电强度的定量研究,并对其与相关下垫面因素的关系进行了探索。

1.1 雷电数据分析及预处理

当前对于雷电定位的方式多为 2 站混合、3 站混合、4 站算法、磁向和时差联合法。表 1 是雷电数据的原始记录形式,由于本试验研究的属性主要是纬度、经度、强度和监测的时间等,其余的属性诸如定位方式、误差(雷电电磁波沿复杂地表传播时,会引起闪电定位系统的定位误差,进而影响闪电定位系统的定位精度。表对闪电定位精度的影响主要体现在两个方面,其一,电磁波沿有限电导率起伏地表传播会使高频分量衰减;其二,电磁波传播路径的延长导致传播时间的增加。)等不在本试验重点考虑范围。雷电强度的正负电荷对于损害程度和预防措施来说没有区别,所以对于雷电的强度和陡度取绝对值,经过这一系列的数据预处理后得到表 2。为了从多变量的雷电数据中确定强度的重要程度,本试验先采用降维方法中的因子分析法,通过因子分析法中的相关检验来判定强度对于雷电的重要程度。

表 1 雷电数据的原始记录

监测日期 (年-月-日)	监测时间	北经纬度 (°)	东经经度 (°)	强度 (kA)	陡度 (kA/μs)	误差	定位方式	地域		
								省份	地市	县(市)
2008-08-04	13:05:41	31.320 9	119.620 4	-22.0	-5.7	0	3 站混合	江苏省	无锡市	宜兴市
2007-08-31	00:24:10	33.575 4	118.760 9	17.7	3.0	46.4	4 站算法	江苏省	宿迁市	泗阳县
2007-08-19	15:16:12	32.028 2	118.444 4	32.9	8.4	0	2 站混合	江苏省	南京市	江浦县

表 2 预处理后的雷电数据记录形式

监测日期 (年-月-日)	监测时间	北经纬度 (°)	东经经度 (°)	强度 (kA)	陡度 (kA/μs)
2008-08-04	13:05:41	31.320 9	119.620	22.0	5.7
2007-08-31	00:24:10	33.575 4	118.761	17.7	3.0
2007-08-19	15:16:12	32.028 2	118.444	32.9	8.4

2 雷电强度等级的聚类分析

本试验首先将不同强度的雷电进行聚类等级划分,再依据不同雷电强度聚类等级进行空间分析。

2.1 K -means 聚类算法

K -means 算法是基于距离划分的硬聚类方法,也是适用范围十分广泛的数据挖掘算法之一。对于处理气象类的海量级数据集,这个算法具有相对可伸缩和高效性的优点。 K -means 算法的基本思想是:以空间中 K 个点为中心进行聚类,

对最靠近它们的对象归类。通过迭代的方法,逐次更新各聚类中心的值,直至得到最好的聚类结果。

算法主要流程描述如下:(1)对于要聚类的数据集适当选择 K 个类作为初始中心;(2)通过适当次数的迭代,对任意一个样本,求其到 n 个中心的距离,将该样本归到距离最短的中心所在的类;(3)利用均值等方法更新该类的中心值;(4)对于所有的 n 个聚类中心,如果利用(2)、(3)的迭代法更新后,值保持不变,则迭代结束,否则继续迭代。

2.2 K -means 的 R 语言实现

该方法是在 RStudio 框架下进行 R 语言的编译运行。以 2008 年夏半年数据聚类过程为例,具体编译运行过程如下:(1)数据调用及加载用于分析的软件包(RODBC、stats、ggfortify)。本试验根据聚类对象的数据特点选用基于距离的 K -means 聚类作为具体的聚类函数。利用 R 语言在 RStudio 框架下对雷电的夏半年数据进行 K -means 聚类。该函数在 R

中的基本格式为

Kmeans(*x*, centers, iter. max = 10, nstart = 1, algorithm = c(“Hartigan – Wong”, “Lloyd”, “For – gy”, “Macqueen”)).

(2)

式中:*X* 为进行聚类分析的数据集,本试验即为处理好的雷电数据集;centers 为预设类别数 *k*;iter. max 为迭代的最大值,且默认值为 10,本试验根据数据的量级将迭代次数设为 1 000; nstart 为选择随机起始中心点的个数,默认值为 1;而参数 algorithm 则提供 4 种算法选择,上述 4 种算法由不同专家学者在不同时期围绕算法的本质提出。本试验选用的“Hartigan – Wong”算法为 Hartigan 和 Wong 等 2 位科学家提出,该算法相较于另外 3 种算法的优势在于较高的运算效率。

(2) 调节聚类优度。对于 *K* – means 聚类算法的关键问题是类别数的确定,在 R 中对于类别参数 center 的取值也是决定聚类效果的一个重要指标。本试验通过讨论预设置的类别数的组间平方和占总平方和的比值来确定最优类别数。通过遍历数据来确定类别数,当类别数小于 5 时,随着类别数的增加,聚类效果越来越好,组间平方和占总平方和的比值快速提高,其值为 89.0%,说明组内差距小,组间差距很大;当类别数超过 5 以后继续增加时,聚类效果提高的非常缓慢。因为约是约值,本试验考虑到较小的类别数对于后续的分析更加方便有效,所以取 *K* = 5。

2008 年夏半年总的样本数为 256 659,5 个簇的中心强度值分别为 18. 216 09、30. 552 05、46. 637 54、74. 920 79、139. 792 07 kA,每个簇的组内成员数量分别为 96 661、99 197、46 142、12 617、2 042。因为是对于强度这个单一变量进行聚类,所以将结果处理为从弱到强的 5 级,这样更方便后续的分析。

2.3 聚类结果与分析

通过分别对 2007、2008、2009 年夏半年雷电强度的聚类,将对应的聚类中心强度值取平均值后,得到表 3 中各夏半年在不同地闪强度等级上的频数分布情况。从表 3 可以看出,占据雷电强度 79.64% 以上(31 kA 以下)的 1、2 级组内成员最多,而后 3 个等级则呈递减趋势。从时间尺度上看,可以反映出 2008 年前 3 个等级强度的发生次数在 3 年中是最低的,这主要是受当年地闪总频次低于另外 2 年的影响。通过整理对应年份的统计年鉴发现,2008 年夏半年江苏省的月平均气温为 23.2℃,低于 2007、2009 年。由于雷电是一种中小尺度的强对流现象,对气温变化较为敏感,所以月平均气温的下降有可能是造成 2008 年雷电总频次下降的原因。但在雷电强度较大的 4、5 级中,3 年发生雷电的次数则几乎不受总地闪频次的影响,呈逐年上升的趋势。这反映出江苏省高强度雷电流的发生次数有逐年上升的趋势。

表 3 3 年各等级雷电频数分布

聚类中心均值 (kA)	等级	各年发生雷电的频数(次)			累计占比 (%)
		2007 年	2008 年	2009 年	
19	1	142 038	96 661	120 300	37.09
31	2	188 375	99 197	124 182	79.64
47	3	46 193	46 142	58 137	95.18
75	4	11 155	12 617	16 596	99.36
138	5	1 340	2 042	2 801	100.00

3 雷电强度等级空间分布分析

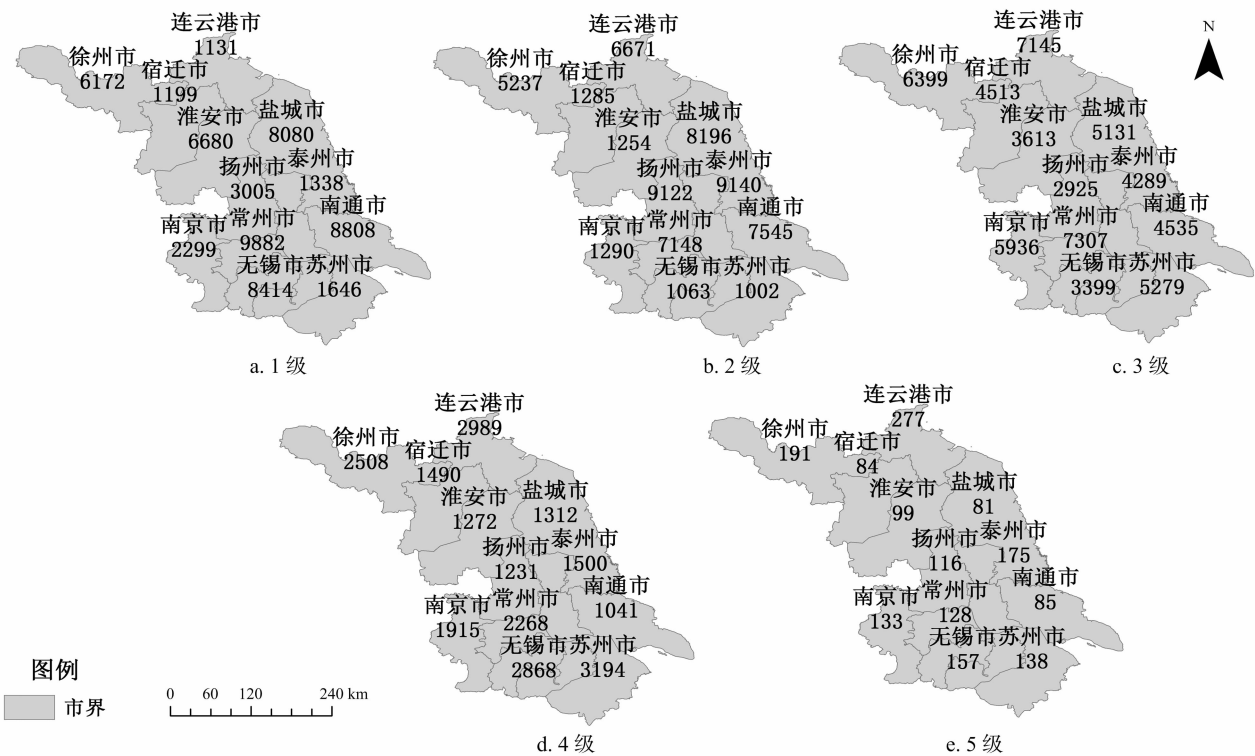
将在 R 语言聚类后的数据导入 GIS 中,利用 GIS 对 5 个等级雷电簇组内成员的空间分布进行分析,结果如图 1 所示,全省 13 个地级市在各雷电等级强度上存在着明显的波动。在雷电强度为 1 级即聚类中心雷电强度值约为 19 kA 时,淮安、南京、扬州、盐城和镇江为频次最高的前 5 位城市,频次总和占总值的 56.38%;在雷电强度等级为 2 级即聚类中心雷电强度值约为 31 kA 时,淮安、盐城、苏州、徐州和南京分列前 5 位,占总值的 51.40%;在雷电强度为 3 级即聚类中心雷电强度值约为 47 kA 时,南京、盐城、淮安、扬州和苏州为前 5 位的城市,频次总和占总值的 50.01%;在雷电强度等级为 4 级即聚类中心雷电强度值约为 75 kA 时,苏州、盐城、南通、淮安和徐州为排名前 5 位的城市,频次总和占总值的 53.89%;最后一类即当聚类中心雷电强度值约为 138 kA 时,盐城、淮安、宿迁、徐州和南通为占据前 5 位的城市,且频次总和占总值的 52.71%。从上述统计分析情况可以看出,各等级全省排名前 5 位的城市的频次总和均超过了该等级频次总和的一半以上。这说明各等级排名前 5 位的城市最具有代表性。

雷电主要通过直击雷和雷电感应现象威胁人类的生命和财产安全。本试验将雷电按强度等级进行聚类,并根据聚类结果进行空间分析,其意义在于对某地区进行雷电风险评估,对建筑或服务设备防雷定级时,可以考虑周边的环境分量。在实际的工程实施中,是否须要提高防护等级、提高到多少,则须要参考当地的雷电强度变化。因此,本试验通过对雷电强度的聚类,得到反映雷电强度空间分布的 5 类雷电簇,以此作为防雷工程实施时的参考依据,以期在采用最有效的等级防护措施的同时将成本降到最低。通过对聚类后的结果进行 3 年 5 类强度的均值计算,得到全省市 1 级的雷电强度基本变化情况,并按照其均值对 13 个地市进行等级区划,结果如图 2 所示,1 级区代表该等级内的地级市雷电强度的变化最高,情况最复杂,在施工时应综合多方因素针对不同等级的设施进行最大化的防护,淮安市属于这一等级范畴,其数值远超其他城市,为 9 019 次,往后各等级区划所代表的雷电强度复杂度递减;盐城、南京属于 2 级,均值分别为 7 239、7 053 次;徐州、扬州和苏州属于 3 级,均值范围在 5 000 ~ 7 000 次之间;宿迁、泰州、南通、镇江、常州和无锡为 4 级,均值范围在 4 000 ~ 6 000 次之间;5 级区均值最小,为 3 136 次,该等级仅连云港市。依照上述等级区划,在江苏省不同城市进行雷电防护施工时,可参考其所处的雷电强度变率等级区,进行相应的防雷措施的调整,以加大对施工工程有效保护的作用。

4 雷电强度等级及与土地利用类型关系的分析

雷电与下垫面因素之间的关系挖掘是研究雷电空间分布的一个重要组成部分。MODIS 地表反射率产品(MOD90A1)提供经过大气校正的地表反射率数据。本试验对 MOD90A1 中的 1 ~ 7 波段从可见光到近红外、短波红外丰富的光谱信息进行监督分类,将江苏省土地利用类型划分成林地、草地、湿地、耕地、建设用地和其他六大类,并与前文中各雷电等级频数进行分析。

从图3可以看出,江苏省4.14%的土地为林地,草地占



图中各城市下方对应的数值代表发生相应强度雷电等级的雷电次数

图1 13市雷电强度等级分布

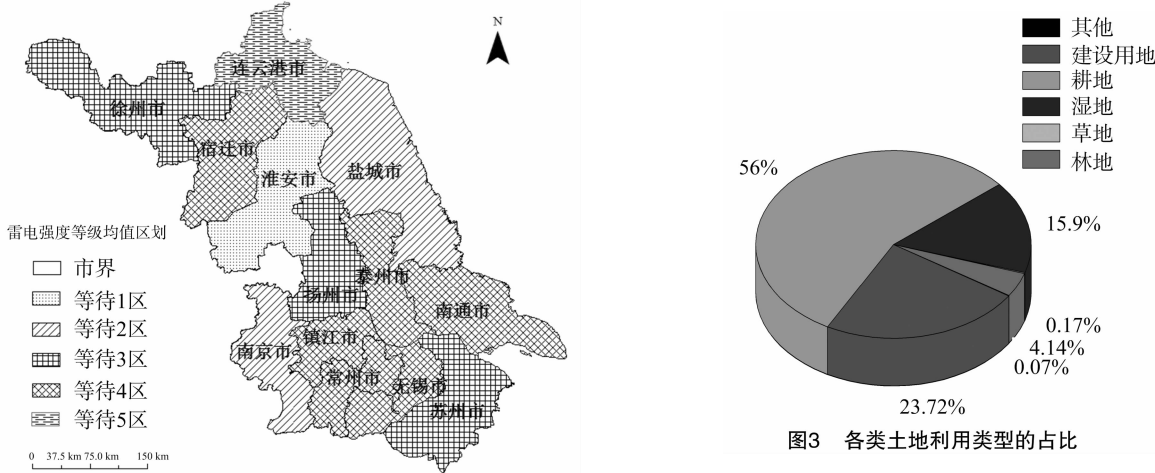


图2 雷电强度变率区划

图3 各类土地利用类型的占比

0.17%, 湿地占 15.9%, 耕地占 56%, 建设用地占土地总量的 23.72%, 其他占 0.07%。图 4 给出了反演后的江苏土地利用类型, 该图能够直观地反映出各土地利用类型的空间分布状况。

各雷电强度等级的频次统计量与各土地利用类型的面积计算进行密度计算后, 对照各雷电强度等级, 由图 5 可知, 单位面积上各雷电强度等级发生的概率最大的为林地, 其余 5 个土地利用类型的对应密度分别为 1.37、1.23、1.03、0.40、0.02 d/km², 而林地仅占江苏省土地面积的 4.41%。这说明雷电强度与林地的关联性最大。其余 5 种土地利用类型与各雷电强度等级的分布趋势大体一致, 值得注意的是该 5 类并

非同林地曲线一样呈衰减状分布, 而是在雷电强度为 2 级, 即中心雷电强度为 31 kA 时, 出现拐点并达到峰值, 而后才成衰减趋势。峰值大小依次为 0.88、1.13、1.05、1.14、0.93、1.23 d/km²。造成该处拐点原因可能有 2 个: 一是云层在江苏境内当带电离子数量达到最佳的泄放量级时, 即本试验 2 级雷电流强度达到 31 kA 时的云层电荷结构, 最有利于地闪的产生; 二是在聚类过程中对第 2 类的聚类宽度范围略大于其余几类, 所以造成该雷电簇组内成员数量增多。以上 2 点认识丰富并完善了文献[9]的相关结论。林地除对雷电的频次有影响外, 也对雷电的强度分布存在显著的影响。此外, 第 2 个认识则反映出了地域差异性, 即在强度 2 级时, 对除林地外的 5 种土地利用类型的影响远大于其他强度, 这也反映出雷电强度受土地利用类型影响。

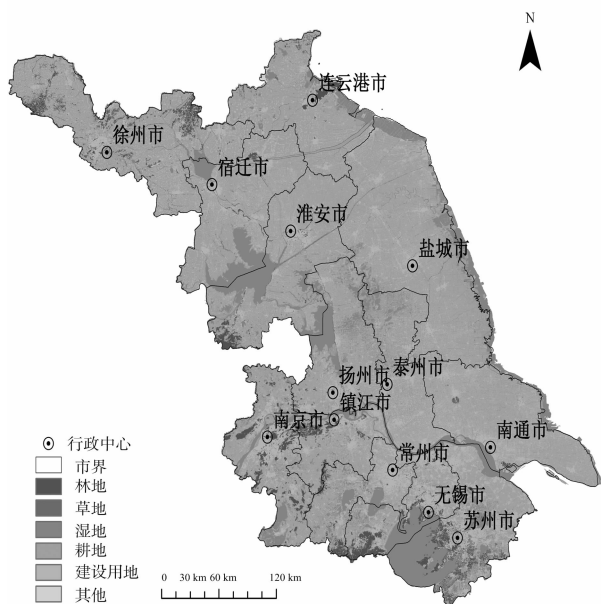


图4 江苏省土地利用类型分布

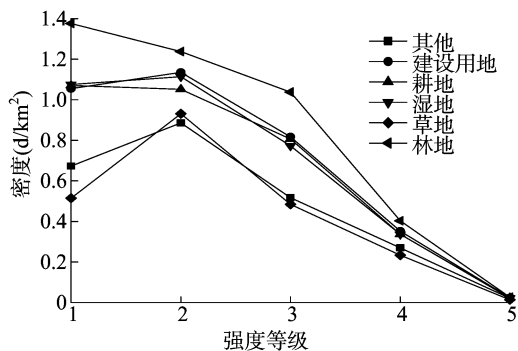


图5 不同土地类型的雷电强度密度分布

5 结束语

本试验针对雷电强度的空间定量分析不足问题,考虑到雷电强度在雷电防护工程中的重要参考价值。将 R 语言架构下的数据挖掘技术与 GIS 技术相结合,通过聚类算法将江苏省夏半年的雷电进行基于强度等级的聚类。在得到相应的等级区划后还与土地利用类型的关系进行分析,结果显示,在 2007、2008、2009 年夏半年,占据雷电强度 90% 以上 (31 kA 以下) 的 1、2 级组内成员最多,而后 3 个等级则呈递减趋势。在雷电强度较大的 4、5 级中,3 年发生雷电的次数则几乎没有受到总地闪频次的影响,呈逐年上升趋势,这反映出江苏省高强度雷电流的发生次数有逐年上升的趋势。雷电强度复杂程度的等级划分如下:淮安市属于 1 级范畴;盐城、南京属于 2 级;徐州、扬州和苏州属于 3 级;宿迁、泰州、南通、镇江、常州和无锡为 4 级;5 级区为连云港。林地除对雷电的频次有影响外,对雷电的强度分布也存在着显著的影响。此外,第 2 个认识则反映出了地域差异性,即在强度等级为本试验的第 2 类时,对除林地外的 5 种土地利用类型的影响远大于其他强度,这也反映出雷电强度受土地利用类型的影响。

基于上述研究的成果,本试验后续将对雷电强度等级的空间尺度进一步细化,得到 13 个地级市各县(市)的强度等

级划分,以进一步提高在实际工程中的参考价值;此外,考虑将其他下垫面因素与雷电强度进行分析,以拓展研究内容。

参考文献:

- [1] 陈渭民. 雷电学原理[M]. 北京:气象出版社,2003:112-149.
- [2] Reap R M. Climatological characteristics and objective prediction of thunderstorms over Alaska[J]. Weather & Forecasting,1991,6(3): 309-319.
- [3] Reap R M, Foster D S. Automated 12~36 h probability forecasts of thunderstorms and severe local storms[J]. Journal of Applied Meteorology,1979,18(10):1304-1315.
- [4] 许小峰. 国外雷电监测和预报研究[M]. 北京:气象出版社,2003:274-278.
- [5] 冯桂力,陈文选,刘诗军,等. 山东地区闪电的特征分析[J]. 应用气象学报,2002,13(3):347-355.
- [6] 李霞,汪茂森,巩晴霞,等. 苏州地区雷电分布规律分析[J]. 气象科学,2006,26(4):442-448.
- [7] 冯民学,焦雪,韦海容,等. 江苏省雷电分布特征分析[J]. 气象科学,2009,29(2):246-251.
- [8] 李政. 重庆地区雷电动活动规律及下垫面状况分析[D]. 南京:南京信息工程大学,2011.
- [9] 赵伟,童杭伟,张俊,等. 浙江省雷电时空分布特征及影响因素分析[J]. 电网技术,2013,37(5):1425-1431.
- [10] 宋晓爽,郑栋,张义军,等. 上海及周边地区地闪活动特征及海陆差异[J]. 气象科技,2014,42(1):164-172.
- [11] Hand D, Mannila H. Principles of data mining[M]. Beijing:China Machine Press,2003:38-55.
- [12] Rodriguez A, Laio A. Clustering by fast search and find of density peaks[J]. Science,2014,344(6191):1492-1496.
- [13] Han J, Kamber M, Tung A. Spatial clustering methods in data mining: a geographic data mining and knowledge discovery[M]. London:Taylor and Francis Group,2001:117-211.
- [14] Birant D, Kut A. ST-DBSCAN: an algorithm for clustering spatial-temporal data[J]. Data & Knowledge Engineering,2007,60(1):208-221.
- [15] Haralick R, Harpaz R. Linear manifold clustering in high dimensional spaces by stochastic search[J]. Pattern Recognition,2007,40(10):2672-2684.
- [16] Dang E K F, Luk R W P, Ho K S, et al. A new measure of clustering effectiveness: algorithms and experimental studies[J]. Journal of the American Society for Information Science & Technology,2008,59(3):390-40.
- [17] Kryszkiewicz M, Lasek P. TI-DBSCAN: clustering with DBSCAN by means of the triangle inequality[C]. Rough Sets and Current Trends in Computing Proceedings,2010:60-69.
- [18] Zelnik-Manor L, Perona P. Self-tuning spectral clustering[J]. Advances in Neural Information Processing Systems,2004,16:1601-1608.
- [19] 侯荣涛,朱斌,冯民学,等. 基于 DBSCAN 聚类算法的闪电临近预报模型[J]. 计算机应用,2012,32(3):847-851.
- [20] Fraley C, Raftery A, Gneiting T, et al. Probabilistic weather forecasting in R[J]. R Journal,2011,3(1):55-63.
- [21] 焦雪,冯民学,钟颖颖. 2006—2009 年江苏省地闪特征分析及应用[J]. 气象科学,2011,31(2):205-210.