

周积强,黄艳红,姚肖萌,等.基于Python的闪电定位资料分析方法研究及应用[J].气象研究与应用,2020,41(3):49-53.

Zhou Jiqiang, Huang Yanhong, Yao Xiaomeng, et al. Research and application of lightning location data analysis method based on Python[J]. Journal of Meteorological Research and Application, 2020, 41(3): 49-53.

基于Python的闪电定位资料分析方法研究及应用

周积强^{1,4}, 黄艳红^{2,4}, 姚肖萌^{1,4}, 魏巍^{2,4}, 柳佳俊^{1,4}, 杨勇^{1,4}

(1. 宁夏回族自治区气象灾害防御技术中心, 银川 750002; 2. 宁夏回族自治区气象信息中心, 银川 750002;
3. 宁夏回族自治区银川市气象局, 银川 750002; 4. 宁夏回族自治区气象防灾减灾重点实验室, 银川 750002)

摘要: 使用Python自带库和第三方Folium库、pandas库函数,分析了闪电定位资料的处理过程、计算函数、实现途径及数据处理的方法。并使用该方法分析了宁夏2019年雷电活动情况,编制了雷电监测年报。结果显示,使用Python能够快速方便的完成闪电数据的统计、分析和显示,为分析闪电定位监测数据应用提供了一种新方法。

关键词: 闪电;Python;统计;应用

中图分类号: P427.32 **文献标识码:** A **doi:** 10.19849/j.cnki.CN45-1356/P.2020.3.09

OSID:



引言

2006年宁夏建设完成闪电定位监测系统,实现了宁夏全区地闪活动监测,2007年宁夏闪电定位监测系统正式投入运行。闪电定位监测系统对宁夏防雷减灾工作的开展提供了强有力的科技支撑,特别是为雷击灾害调查、雷击灾害风险评估^[1]、雷电防护装置工程设计、防雷减灾科普、雷电活动规律研究^[2]等工作提供了数据保障。随着经济社会的发展,闪电定位资料将会在气象、电力^[3]、能源^[4]、危化、易燃易爆^[5]等领域得到广泛应用。

Python是一种跨平台的计算机程序设计语言,广泛应用于Web和Internet开发、科学计算和统计、人工智能、桌面界面开发、软件开发等。同时其完全免费,众多开源的科学计算库都提供了Python的调用接口。用户可以在任何计算机上免费安装Python及其绝大多数扩展库。通过调用扩展库,可极大的减小用户编程工作量,提高效率。Python已经在气象领域^[6-9]逐步得到应用。鉴于Python在科学计算和统计方面强大的应用能力,本文尝试使用Python软件,

结合闪电定位资料数据特点,使用Folium库和pandas库研究了闪电定位资料的地理信息、时空分布及其强度特征分析的方法,并以分析2019年4—10月宁夏闪电定位资料为例,编制完成了2019年宁夏雷电监测公报并向社会公布,为宁夏防雷减灾工作提供数据支撑。

1 资料来源与方法

1.1 资料格式

目前使用闪电定位资料是中国气象局大气探测中心提供的ADTD数据监测终端自动生成地闪数据记录。当某一地区发生地闪以后,ADTD系统按天生成闪电定位信息记录,存储格式为文本格式,命名规则是日期.txt(例:2019_07_09.txt)。当天没有监测地闪数据时,软件不生成文本文件。闪电定位资料文本文件记录了地闪发生次数编号、发生时间(DateTime格式)、经度、纬度、强度、抖动、误差、定位方式、行政区域信息(省/区、市、县)。数据使用两个空格自动分割。

收稿日期: 2019-12-08

基金项目: 中国气象局旱区特色农业气象灾害监测预警与风险管理重点实验室指令性项目(CAMP-202010)、第二次青藏高原综合科学考察研究项目(2019QZKK0104)

作者简介: 周积强(1986—),男,工程师,主要从事气象灾害防御技术研究。E-mail: 545852413@163.com

1.2 闪电统计特征

闪电统计通常使用数学统计方法对地闪的空间、时间、强度分布进行分析。空间分布是地理位置信息显示、省市县行政区域发生地闪次数和天数。时间分布是闪电的年、月、季、日分布。强度分布指闪电强度的频次分布,地闪电流强度分档,一般分为四档(20kA 以下、20~50kA、50~100kA、100kA 以上),分别计算各档地闪强度频次。

2 统计分析过程 and 关键技术

地闪数据分析前需对多个地闪记录文本文件进行合并,形成一个完整的时间段数据集。然后对数据集中的地闪记录进行格式处理,以便于计算和统计分析。本文中将文本逐条记录合并到 CSV 文件中(逗号分隔文件)。使用 Folium 和 pandas 库相关函数对闪电定位数据进行地理信息显示和统计。

2.1 数据分析流程

因地闪数据为逐天生成文本记录,为便于做整体分析,需将所有逐天记录合并在一起。合并数据方法有很多种,本文选择首先将全部数据合并为一个文本文件,然后将文本文件转为逗号分隔文件(格式为.csv),在转为 csv 文件时,需要文本文件空格分隔符转为逗号分隔符。使用循环和列表的方法将闪电定位数据经纬度和强度数据处理为地图可视化模块(Folium)所要求的显示格式,最后进行地图分布显示。数据处理和分析如图 1 所示。数据处理完毕后使用 Pandas 直接调用逗号分隔文件,即可进行统计和分析。

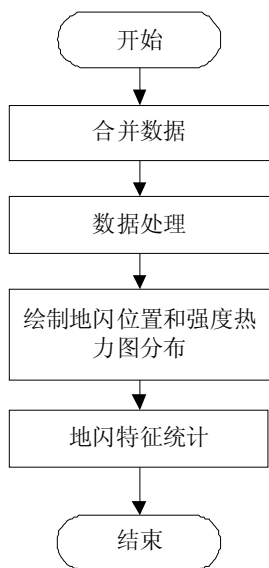


图 1 数据处理和分析流程

2.2 数据处理

资料合并时使用 `os.listdir()` (文件夹名称) 获取该文件夹中全部地闪记录文件名称,并将全部名称写入列表。将地闪各特征参量名称使用列表的方式将名称写入目标 CSV 文件中。当写入标题行时,需要在每次写入一个名称时添加逗号分隔符,最后一个写入时需添加换行符。使用 `open` 函数(“文件路径+文件名称”, “w/a”, 其中文件路径可为相对路径或绝对路径; 文件打开属性中 w 会将所有记录全覆盖, a 为追加形式)采用循环的方式将所有记录读取和写入目标 CSV 文件中。逐行写入地闪数据时先用 `strip()` 函数对记录进行格式处理,然后用 `split()` 将记录转为列表。闪电定位资料的写入和标题行写入方法一致,但需要将写入目标文件的文件打开属性设置为追加,以防止将先写入的标题行覆盖。

2.3 数据调用

使用 pandas 库中 `pd.read_csv()` (文件路径名”, 属性设置) 函数将经过数据处理的文件打开。使用 `info()` 函数可查看打开文件数据的属性,查看属性的目的是为后面调用数据时便于各种格式转换,以适用各相关函数调用的数据格式要求。使用 `[“数据标题”].values` 方法将数据表按列的形式方便调用(该方法调用数据属性为列表),即可通过索引值来直接使用数据。使用 `len()` (读取文件对象) 函数可直接查看打开文件记录条数或者列表长度,为后期数据循环调用提供范围。循环函数中最终范围使用 `range(len(文件对象/列表))` 函数实现动态确定。关键代码: `df = pd.read_csv(out_file, header = None, encoding = "gbk")`, 读取逗号分隔文件语句,其中 `out_file` 为文件的地址,若为相对地址时,需将文件和 Python 程序放在同一个文件夹下。

2.4 地图调用

使用 Folium 库将地闪数据在 Leaflet 地图上进行可视化。因 Folium 库为第三方提供(pandas 库同样),使用 `“pip install folium”` 函数对其进行安装。使用 `folium.Map([维度, 经度], zoom_start=n)` 建立地图对象,使用经纬度设置地图显示中心,以宁夏为例可设置红寺堡(37.4250, 106.0576)为地图中心, `zoom_start` 设置地图显示大小,数值越大区域显示越大。使用 `folium.Circle(参数).add_to(地图对象)` 函数即可将经纬度信息依次添加到地图,其中内部有很多参数根据需要进行设置,如地闪颜色、直径等各种信息,具体设置方法可在其官方网站进行查

询使用。使用 `add_child(folium.LatLngPopup())` 函数可提供地图交互操作,该函数目的是显示出鼠标点击地方的经纬度信息,该方法可直接查看地闪位置准确经纬度。使用 `HeatMap(热力图数据).add_to(地图对象)` 绘制雷电强度热力图。热力图数据的属性为列表,列表中的单个数据是由经纬度和雷电流强度数据组成的列表。关键代码:`map = folium.Map([37.4250,106.0576], zoom_start=7)`,方括号为地图中心经纬度,`zoom_start` 为地图初始化放大倍数,本文设置为 7。`folium.Circle(radius= 0.1,location = ([lng [i],lat [i]]),popup= None ,tooltip= None, control_scale=True,color='#FF0000')`.`add_to(map)`,在地图显示地闪为的位置信息。`map.add_child(folium.LatLngPopup())`,为地图提供交互信息显示,点击地图时显示点击点为经纬度信息。`HeatMap(heatmap_data).add_to(map_Dheat)` 绘制地闪强度分布,其中 `heatmap_data` 为地闪强度数据,其格式为 `[[维度 1,经度 1,强度 1],[维度 2,经度 2,强度 2].....,[维度 n,经度 n,强度 n]]`。

2.5 统计方法

地闪天数参照雷暴日定义,为某一区域只要发生一次地闪即记当天为一个地闪天。根据闪电定位资料特点,使用 `counts()` 函数可直接统计出全区、各市、县、区域雷电发生次数。使用条件判断方法将符合要求的数据用 `append()` 函数装入列表,使用 `len`

`()` 和 `nunique()` 函数可直接统计出个全区、各市、县、区域在地闪天数,以及月份、24h、各强度分档闪电次数。使用 `math` 和 `python` 自带的统计函数,可方便的统计出闪电强度最大值、最小值、平均数等基本统计特征。其中 `counts` 函数为统计数据中每条纪录的数量,`nunique()` 函数为统计数据中去除重复记录的数量。

3 应用个例分析

使用宁夏 2019 年 4—10 月的闪电定位监测资料,通过 `python` 软件数据格式化、文件处理、地图显示、数据统计等方法分析闪电空间、时间、强度等分布信息,编制完成了 2019 年宁夏雷电监测公报。

3.1 空间分布

宁夏全区有 107d 发生雷云对地闪击(以下简称地闪)。全区地闪总数为 3180 次。图 2 为全区地闪空间分布情况。从全区地闪位置分布来看,全区各市县均有地闪发生,地闪次数由南向北逐渐减小。

各市地闪天数分别为:银川市 38d、石嘴山市 29d、吴忠市 78d、固原市 56d、中卫市 41d,其中吴忠市地闪天数最多,石嘴山市地闪天数最少。各市地闪次数分别为:银川市 245 次、石嘴山市 180 次、吴忠市 1198 次、固原市 953 次、中卫市 604 次,其中吴忠市地闪次数最大,石嘴山市地闪次数最小。图 3 为

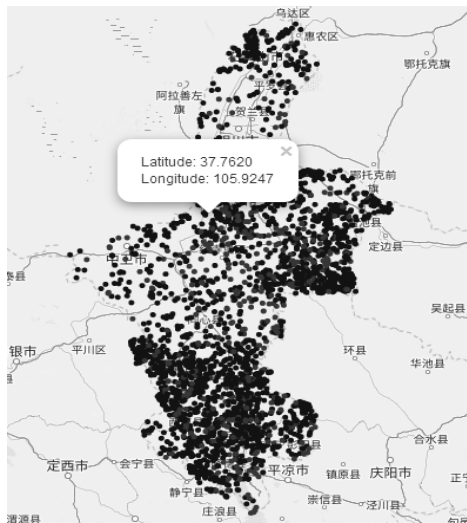


图 2 宁夏地闪空间分布

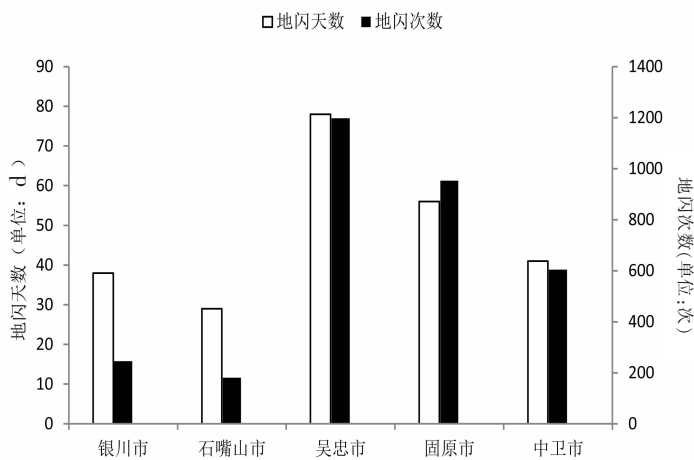


图 3 宁夏各市地闪天数和次数分布

宁夏各市地闪天数和次数分布情况。

宁夏各县、区地闪次数分布如图 4 所示。其中发生地闪天数最多为同心县,地闪天数为 50d;最少为

永宁县,地闪天数为 5d。发生地闪次数最大为盐池县,地闪次数为 682 次;最小为泾源县,地闪次数为 16 次。

3.2 时间分布

宁夏地闪发生次数4月份最大,地闪次数为1452次;10月份最小,地闪次数为33次。其中2019年4月

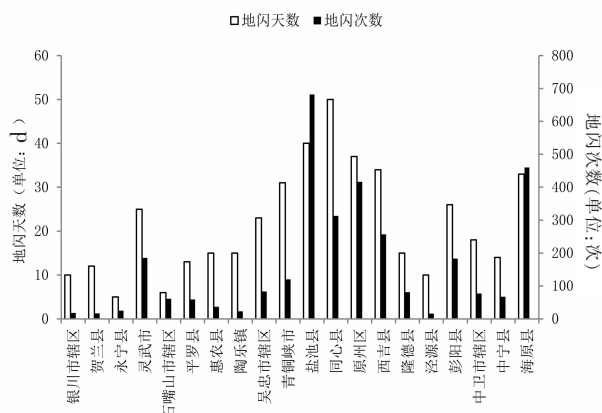


图4 宁夏各县、区地闪天数和次数分布

地闪次数按照24h逐时进行对应累加,发现地闪在各时间段均会发生。其中22点到次日凌晨2点地闪发生频次最高。4—12点,地闪发生频次较低。13—21点地闪发生频次高于4—12点。图6为地闪次数24h逐时分布情况。

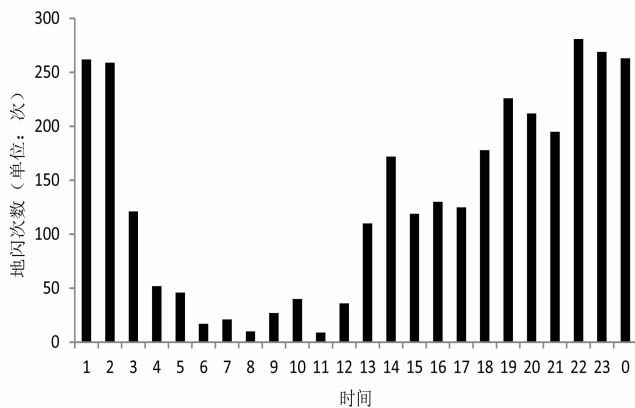


图6 宁夏地闪次数24h逐时分布

3.3 强度分布

3.3.1 地闪强度热力分布

当热力图显示区域为全区,热力信息显示不太清晰,几乎看不出位置分布区别,但Folium输出地图可以按照需要进行放大到局部,这样既可清晰的看出强度分布特点。虽然Folium能够快速的显示雷电强度热力分布,以及根据需要对地区进行放大,但该热力分布不提供层次清晰的强度分档色标显示。

26日地闪次数为738次、2019年4月27日为548次,2天发生地闪次数占4—10月总地闪次数的40%。图5为宁夏4—10月份地闪次数逐月分布情况。

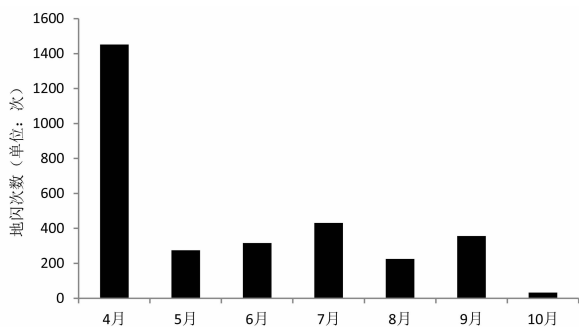


图5 宁夏4—10月份地闪次数逐月分布

3.3.2 地闪强度密度分布

宁夏地闪强度分布为:20kA以下地闪次数为2673次,占比84%;20~50kA地闪次数为212次,占比7%;50~100kA地闪次数为205次,占比6%;100kA以上地闪次数为90次,占比为3%。地闪电流强度100kA以下地闪频次分布为:91%的地闪强度在50kA以下。其中发生频次最高的地闪强度主要集中在10kA以下;地闪强度为10~20kA的发生频次低于10kA以下;强度20kA以上的地闪次数快速减小。

4 结论与讨论

本文使用Python自带库函数和第三方Folium库、pandas库函数,分析了闪电定位资料的处理过程和用到的函数,及详细的使用方法和数据处理过程中需重点关注的技巧。使用该方法分析了宁夏2019年雷电活动情况,编制完成了雷电监测年报并向社会进行公布。

Python作为一种开源免费的计算机程序设计语言,基于功能强大、语言简洁、易上手的特点,能够快速实现闪电定位资料的数据处理、数据调用、地理信息显示、各特征参数统计和分析,可为闪电定位资料统计和分析提供一种新的方法。使用Python分析闪电定位数据有多种方式可选择,本文采用的方法对数据处理和分析思路比较明确,但过程较为繁琐。虽然实现了分析目的,但程序代码较多,仍需继续优化。

参考文献:

- [1] 扈海波, 李京校, 潘进军. 北京地区雷电灾害风险评估方法与应用[J]. 气象, 2012, 38(8): 1004–1011.
- [2] 王学良, 张科杰, 黄小彦, 等. 湖北省雷暴日数与云地闪电密度关系研究[J]. 气象, 2012, 38(6): 728–732.
- [3] 高金阁, 李京校, 张仲, 等. 架空传输线雷电感应电压计算公式分析与修订[J]. 气象科学, 2017, 37(6): 845–850.
- [4] 卢志红, 张世谨, 朱华松. 风力发电机组雷电防护的分析[J]. 气象研究与应用, 2017, 38(2): 93–96.
- [5] 沈海波, 陈贻亮. 梧州市雷电灾害变化规律与易燃易爆场所雷电防护简析[J]. 气象研究与应用, 2017, 38(4): 74–76+87.
- [6] 连高欣, 王琪, 沈岳峰. Python 语言在 CIMISS 中的应用[J]. 气象科技, 2018, 46(4): 699–706.
- [7] 刘志, 郝克俊. 基于 ArcGIS Python 的安全射界图自动化制作[J]. 气象科技, 2016, 44(5): 816–821.
- [8] 吕劲文, 申华羽, 涂小萍, 等. 宁波雷电时空分布特征及潜势预报方法探讨[J]. 气象研究与应用, 2019, 40(2): 30–34.
- [9] 刘平英, 周清倩, 胡颖, 等. 近 12 年云南省云地闪活动变化及雷电灾害时空分布特征[J]. 气象研究与应用, 2018, 39(3): 86–91.

Research and application of lightning location data analysis method based on Python

Zhou Jiqiang^{1,4}, Huang Yanhong^{2,4}, Yao Xiaomeng^{1,4}, Wei Wei^{2,4}, Liu Jiajun^{1,4}, Yang Yong^{1,4}

(1. Ningxia Meteorological Disaster Prevention Technology Center, Yinchuan 750002;

2. Ningxia Meteorological Information Center, Yinchuan 750002;

3. Yinchuan Meteorological Bureau, Yinchuan 750002;

4. Ningxia Key Laboratory of Meteorological Disaster Prevention and Mitigation, Yinchuan 750002)

Abstract: Using Python library, third-party Folium library and pandas library function, the processing process, calculation function, realization way and data processing method of lightning location data are analyzed. Besides, the lightning activities in Ningxia in 2019 were analyzed and the annual report of lightning monitoring was compiled. The results show that Python can quickly and easily complete the lightning data statistics, analysis and display, which provides a new method for analyzing lightning location monitoring data.

Key words: lightning; Python; statistics; application