

基于遥感和社交媒体数据的城市洪涝灾害监测

张 会, 潘晓宁, 贾 浩, 许德合

(华北水利水电大学 测绘与地理信息学院, 河南 郑州 450046)

摘 要:城市洪涝监测对灾害管理与应急响应至关重要。单一数据源往往存在各自的缺陷,不利于多维度、全方位地分析城市复杂洪涝灾害灾情。结合不受云雨影响且覆盖范围大的 GF-3 SAR 影像与实时性较强的社交媒体数据,建立“7·20”特大暴雨郑州市中心城区洪涝监测方法:采用阈值分割、随机森林对 GF-3 SAR 影像进行多方案的灾前灾中水体提取,获取精度最高的水体提取与洪涝监测结果,分析典型区域的水体提取效果;利用 Python 工具获取城市内涝的社交媒体数据,对其进行处理、可视化与空间分析;结合 GF-3 SAR 影像与社交媒体数据洪涝监测结果探讨二者优势互补性。结果表明:SAR 水体总体提取精度从高到低依次为随机森林、Otsu 阈值法、水体指数法,但是在一些典型区域分析中,随机森林的提取效果低于其他方法;基于 SAR 影像提取的洪涝淹没范围主要集中在三环外的城市边缘地区与大型水体周围,基于社交媒体数据提取的洪涝信息主要集中在城市人口和建筑密集的内环内。

关键词:洪涝灾害;高分三号;阈值分割;随机森林;社交媒体数据

中图分类号:TP79 **文献标志码:**A **doi:** 10.13705/j.issn.1671-6833.2025.01.011

在全球气候变化下,城市极端暴雨内涝灾害频发,已成为制约城市健康稳定发展的重要因素^[1]。由于城市建筑和人口密集、关键基础设施薄弱^[2],城市洪涝灾害事件往往会造成极大的经济损失、人员伤亡与社会影响。对洪涝灾害进行快速监测,并做出有效响应在灾害应急管理中至关重要。合成孔径雷达(SAR)使用长波段电磁波,不受云、雨、雾的影响,具有全天时、全天候获取影像数据的优势^[3],可用于城市洪涝灾害监测。

利用 SAR 影像中水体与非水体地物的后向散射强度的差异可提取洪水淹没信息。SAR 洪水检测方法主要分为阈值法、机器学习、变化检测法和干涉测量法等^[4]。阈值法通过特定阈值提取水体,使用便捷,但受影像直方图双峰分布影响较大。机器学习通过提取先验样本,借助样本的概率统计信息制定分类规则,利用不同分类器模型进行洪涝信息提取,提取效果优于阈值法,但对样本的依赖程度高。变化检测法采用洪水发生前与洪水期间的 SAR 影像生成差值或比值图像,进一步结合阈值分割、分类器法等提取洪涝信息。此方法

可直接去除永久性水体的影响,但斑点噪声的影响、理想参考影像的选取都是其限制因素。干涉测量方法需要另外结合 SAR 影像的相位信息,根据相干性值的高低区分淹没区与非淹没区。该方法在建筑密集区具有一定效果,但需要处理相位信息,时间成本比较高。

目前,基于 SAR 数据的洪涝监测多集中于自然流域的河流和湖泊。谷鑫志等^[5]提出基于 GF-3 影像的阈值分割与马尔可夫随机场(MRF)相结合的自动化水体信息提取方法,在湖南省东北部区域水体提取精度达到 85% 以上;Markert 等^[6]基于 Sentinel-1 影像和谷歌地球引擎(GEE)进行水体提取,发现阈值法中 Edge Otsu 优于 Bmax Otsu,并使用 Edge Otsu 对湄公河下游洪泛平原 2019 年夏季洪水进行提取分析。由于城市地表复杂的后向散射机制与建筑类型和高度、植被面积,以及道路拓扑结构有关,利用 SAR 数据对城市地区的洪水探测具有挑战性。当前只有少量研究探索 SAR 数据在城市地区洪涝淹没监测中的应用。城市地区主要依赖高分辨率 SAR 影像,如 TerraSAR-X, RADARSAT-2;其次是中

收稿日期:2024-06-10;修订日期:2024-06-30

基金项目:国家自然科学基金资助项目(42007423);河南省科技攻关项目(222102320021)

作者简介:张会(1987—),女,河南驻马店人,华北水利水电大学讲师,博士,主要从事遥感与 GIS 应用研究、城市洪涝风险评估的研究,E-mail:zhanghui2019@ncwu.edu.cn。

分辨率 SAR 影像,如 GF-3、Sentinel-1 等。但仅利用 SAR 后向散射强度在城市洪涝监测中获取的信息是有限的,还需要借助相关辅助数据。Li 等^[7]将多时相高分辨率 TerraSAR-X 相干性结合强度信息,Tanguy 等^[8]将 RADARSAT-2 数据与水力数据(汛期)相结合,对城市洪涝进行淹没范围提取,都取得了良好的效果。

由此可见,辅助信息的使用有助于 SAR 数据对城市地区的洪水监测。常用的辅助信息有土地利用数据、社交媒体数据、水文数据等。在面向城市洪涝的快速提取中,需要考虑辅助数据的处理时间成本。社交媒体数据作为灾害识别的新型数据源,通常以文本或视频形式发布,处理速度快,是城市洪涝监测中极具应用潜力的数据。Li 等^[9]利用卷积神经网络对从微博中抓取的数据进行分类训练,并将模型运用于深圳和武汉的暴雨洪涝事件监测。但该数据受人口分布、经济水平等影响比较大,具有较强的空间异质性。社交媒体数据与遥感数据的结合,可进一步提高洪涝监测的时效性。邢子瑶等^[10]将 SAR 影像与 VGI(volunteered geographic information)数据、微博数据结合进行洪涝范围监测。但是目前在城市洪涝监测研究中,没有充分发挥它们的数据优势。

本文以郑州市“7·20”特大暴雨洪涝为例,利用 GF-3 SAR 影像,辅以社交媒体数据,结合阈值分割法和机器学习方法,进行城市洪涝淹没信息快速提取,探索 SAR 影像与社交媒体数据在城市洪涝灾害监测中的应用潜力。

1 研究区概况与数据处理

1.1 研究区概况

郑州市是国家中心城市,河南省省会。郑州市位于河南省北部,处于黄河中下游和伏牛山脉向黄淮平原过渡交接地带,研究区地形如图 1 所示。2021 年 7 月 20 日,郑州市突发特大暴雨,贾鲁河、双洎河、颍河等 3 条主要河流均出现超水位洪水,市区出现严重内涝。此次暴雨导致城市内涝、河流洪水、山洪滑坡等多灾并发,造成重大人员伤亡和财产损失^[11]。

1.2 数据处理

本文主要采用覆盖郑州中心城区的 GF-3 SAR 影像及 Sentinel-2A 影像,影像信息如表 1 所示。GF-3 SAR 包含 2021 年 7 月 15 日的两景灾前影像和 2021 年 7 月 20 日的两景灾中影像。GF-3 FSII 影像预处理工作主要包括滤波处理、地形校正、影像

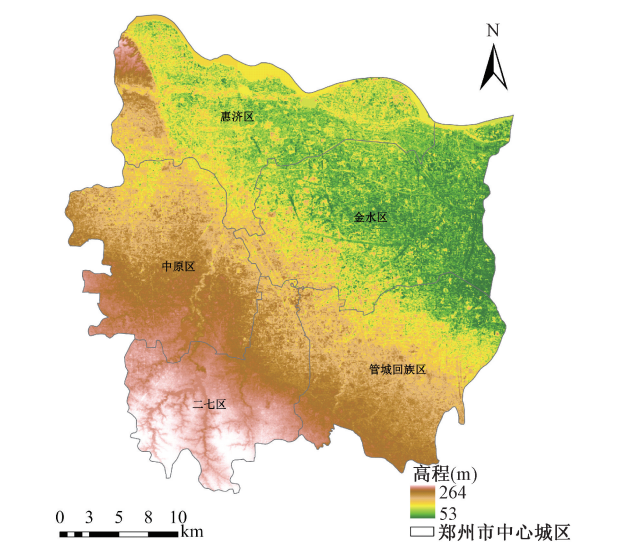


图 1 郑州市中心城区地形图

Figure 1 Topographic map of the central urban area of Zhengzhou

裁剪、影像配准。GF-3 L2 FSII 影像为经过几何校正和辐射校正的产品,不需要再进行相关处理。由于 SAR 图像中存在大量随机分布的相干斑噪声,选择 Frost 滤波器进行滤波处理,Frost 滤波可在减少噪声的同时保持图像的边缘与细节。GF-3 SAR 属于斜距投影,需对影像进行地形校正以减弱几何畸变的影响,本文采用 GMTED2010 DEM 进行地形校正。此外,采用 2021 年 6 月 6 日 Sentinel-2A L2A 级影像,提取空间分辨率为 10 m 的蓝、绿、红、近红外 4 个光谱波段。该级别产品不需要经过辐射定标与大气校正。

表 1 GF-3 与 Sentinel-2A 影像基本信息

Table 1 Basic information of GF-3 and Sentinel-2A images

遥感影像	成像模式	分辨率	成像时间	幅数
GF-3	FSII	10 m	2021 年 7 月 15 日	2
			2021 年 7 月 20 日	2
Sentinel-2A	L2A	10 m	2021 年 6 月 6 日	1

2 研究方法

针对预处理后的 GF-3 SAR 影像,分别采用阈值法和机器学习方法进行郑州市中心城区洪涝信息提取。阈值法包括同极化 HH Otsu、交叉极化 HV Otsu 和 SAR 水体指数法;机器学习方法采用随机森林。对比两种方法的水体提取精度,选择提取精度较高的灾前灾中水体提取结果,通过空间叠加分析获取洪涝淹没空间分布图,并与基于社交媒体数据获取的积水点结合,实现对郑州市中心城区的洪涝灾害监测。

2.1 阈值分割法

2.1.1 Otsu 阈值法

Otsu 阈值法不需要任何先验知识,且不需要参数设置,只根据图像灰度直方图进行分割阈值选取。Otsu 阈值法根据图像的灰度特性将图像分成背景和前景两部分,计算两者的类间方差,当类间方差最大时,完成对图像背景和前景的分割。该算法可求出水体和非水体两类地物类间方差最大值^[12],从而得到分割水体和非水体的全局阈值。Otsu 算法假设水体的像元值为 $[p_1, p_2, \dots, p_n]$,非水体的像元值为 $[p_{n+1}, p_{n+2}, \dots, p_m]$, p_n 表示分割水体与非水体的阈值,其类间方差为

$$\sigma = W_1(M_1 - M)^2 + W_2(M_2 - M)^2; \quad (1)$$

$$W_1 + W_2 = 1; \quad (2)$$

$$W_1M_1 + W_2M_2 = 1. \quad (3)$$

式中: W_1 、 W_2 分别为水体和非水体像素占整张图像的比例; M_1 、 M_2 分别为水体和非水体像元的平均值; M 为影像的像元平均值。通过不断迭代计算得到类间方差 σ 最大时的阈值 p_n 。

本文利用 Otsu 阈值法分别提取灾前、灾中的 GF-3 HH 和 GH-3 HV 影像的水体与非水体分割阈值。

2.1.2 水体指数法

贾诗超等^[13]提出的基于 Sentinel-1 双极化数据的 SDWI 水体信息提取方法的公式为

$$K_{SDWI} = \ln 10HH \cdot HV - a. \quad (4)$$

式中: K_{SDWI} 为水体指数; HH 和 HV 极化影像相乘,乘以 10 并取其对数值得到合成影像,以此扩大水体与非水体地物之间的差异,便于它们的分割; $a = 8$ 为水体与非水体直方图的波谷值;当 K_{SDWI} 大于 0 时为水体,小于 0 时为非水体。

由于不同 SAR 影像水体与非水体直方图的波谷值可能存在差异,本文在郑州市通过波段计算的合成影像中选取常庄水库区域作为 SDWI 模型改进的试验区,如图 2 所示,由合成影像直方图可知双峰波谷在 8.7 左右。根据 Otsu 阈值法原理,当选取 8.7 作为 SDWI 模型的参数时,水体提取的效果较好。借鉴 SDWI 水体信息提取方法,改进后的 GF-3 水体指数 GFWI 公式为

$$K_{GFWI} = \ln 10HH \cdot HV - 8.7. \quad (5)$$

2.2 随机森林

随机森林通过构造不同的训练集增加分类模型间的差异从而提高组合分类模型的外推预测能力^[14]。随机森林基本思想为首先利用 Bootstrap 抽样从原始训练集抽取 n 个样本;其次对 n 个样本分

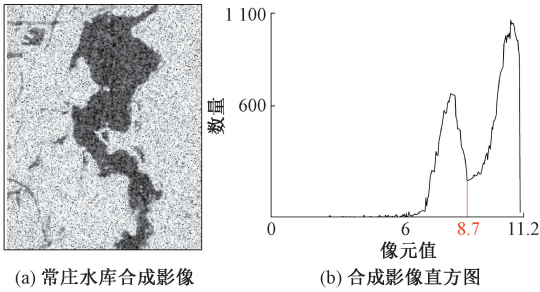


图 2 常庄水库合成影像与合成影像直方图
Figure 2 Composite image and its histogram of Changzhuang Reservoir

别建立 n 个决策树模型得到 n 种分类结果;最后根据 n 种分类结果对每个记录进行投票表决,决定其最终分类。

样本是机器学习分类的关键输入之一,样本的数量、准确性和纯度都会影响最终的分类结果^[15]。本文利用 ENVI 5.6 的 Layer Stacking 工具将水体指数影像、HH 和 HV 极化影像进行波段合成,得到 7 月 15 日和 7 月 20 日的三波段合成影像。选取水体和非水体(建筑、道路)3 类样本执行随机森林分类,以提取水体信息。

2.3 水体提取精度验证

本文参照沈兰芝等^[16]的水体提取精度验证方法。选取研究区 2021 年 6 月 6 日的 Sentinel-2A 影像,对 GF-3 SAR 影像水体信息提取结果进行精度验证。在 Sentinel-2A 影像上研究区范围内随机选择 200 个检验样本,如图 3 所示,其中水体样本点 96 个。在 Sentinel-2A 影像通过目视解译识别样本点的水体/非水体属性,作为水体提取结果精度评价的参考依据。

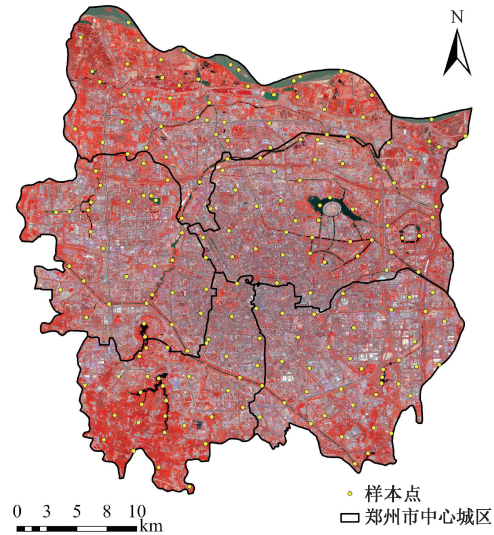


图 3 样本点分布图
Figure 3 Distribution map of sample points

比较样点在 GF-3 SAR 影像水体提取结果和 Sentinel-2A 参考影像的水体/非水体属性,若一致则为正确检验样本 (CS),不一致则为错误检验样本 (FS)。利用 CS 与 FS 计算总体分类精度 OA:

$$OA = \frac{CS}{CS + FS} \times 100\%。 \quad (6)$$

2.4 基于社交媒体数据的积水点提取

本文使用的社交媒体数据来源于新浪微博。作为国内用户数量众多的社交媒体平台,新浪微博包含了丰富的语义内容特征和关系特征,可支持发布文本、图片、视频等,并可带有地理位置标签,是灾害研究中社交媒体数据的常用数据源^[17]。

本文使用基于 Python 开发的网络爬虫程序进行社交媒体数据的获取。首先在浏览器中登陆微博,获取微博后台中的初始 URL (uniform resource locator) 并输入程序。在程序中设置需获取的关键词(使用了“暴雨”、“洪水”、“内涝”、“洪涝”等关键词),搜索的时间(2021 年 7 月 20 至 26 日)以及搜索的地点(河南省郑州市)。最终生成的数据表中包含有对应网址、用户名、时间、博文内容、点赞数、转发数与评论数等信息。共获取 22 860 条数据,数据多来自居民发布的微博博文。对获取的数据进行数据清理、模糊融合匹配、去重与人工筛选等操作,得到基于微博发布的积水点数据。另外,基于用户名为郑州交警的微博账号在 2021 年 7 月 20 日至 26 日发布的关于郑州市区积水点的微博博文,获取微博文本中的积水点位置信息。本文选择将居民与交警微博账号发布的内涝位置信息相结合,获取尽可能全面且准确的灾情信息。交警部门可能关注三环内的重点区域,此部分信息比较可靠。居民关注的空间范围比交警部门关注的范围大,但是可信度可能比交警数据低一些。将微博爬虫数据和郑州交警微博博文积水点数据进行去重,综合得到社交媒体积水点数据,并利用百度地图定位工具获取地理坐标,最后在 ArcGIS 中进行可视化,结果如图 4 所示。

3 结果分析与讨论

3.1 基于 SAR 的水体监测结果分析

3.1.1 精度评价

利用 Otsu 阈值法、水体指数法和随机森林分别对 2021 年 7 月 15 日和 7 月 20 日郑州市中心城区 GF-3 SAR 影像进行水体提取,水体提取结果如图 5 所示,水体提取精度如表 2 所示。可以看出,Otsu 阈值法中交叉极化 HV 影像提取效果优于同极化 HH 影像。SAR 极化方式的不同会导致接收器接收到

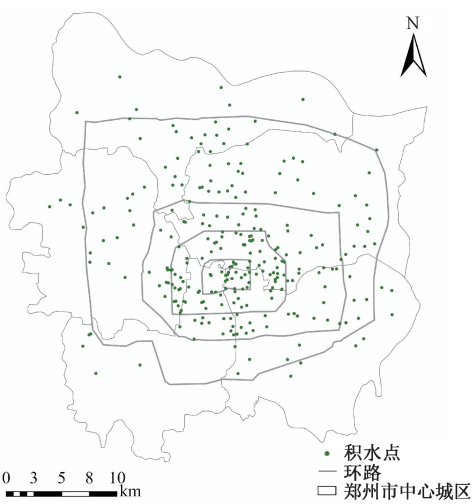


图 4 积水点可视化图

Figure 4 Visualization of waterlogging points

回波信号的不同,交叉极化 HV 相比同极化 HH 可以显著地增加雷达图像信息。水体指数法在灾前的提取精度与同极化 HH 提取精度相似,但低于交叉极化 HV 的提取精度。水体指数法在灾中的提取精度低于交叉极化 HV 与同极化 HH 的提取精度。随机森林在灾前灾中的水体提取中精度都最高,分别高达 88.54%与 86.46%。相比阈值分割的单特征,本文中随机森林输入了同极化 HH、交叉极化 HV 及水体指数等多个特征,可利用多个决策树进行预测。且随机样本的输入及随机性决策树的构建可以避免过拟合现象的发生,因此其分类精度较高。此外,在 3 种方法中,灾前影像的水体提取精度均高于灾中影像。

3.1.2 典型区域水体监测结果对比

在研究区范围内选取 5 个典型区域,用于对比 4 种方法在水体提取细节方面的差异,如图 6 所示,其中区域 1 为尖岗水库区域,区域 2 为刘江黄河桥南侧区域,区域 3 为潮河区域,区域 4 为索须河和贾鲁河交汇区域,区域 5 为国际会展中心区域。在区域 1 中,3 种方法获取的洪涝淹没范围基本一致,但 Otsu 阈值法和随机森林的水体提取细节优于水体指数法。区域 2 分布着大量的水塘,对比发现,水体指数法提取的水塘周围与城市道路淹没面积均小于另外两种方法,且区域 2 中存在着部分道路错分为水体的现象,这是由于道路与水体在 SAR 影像上均表现为深色,造成了水体误提取。区域 3 为分布着内涝点的居民区,Otsu 阈值法与随机森林提取的淹没区域明显大于水体指数法的提取结果。由此可以看出,水体指数法在细小水体提取方面存在着不足。结合精度评价结果,Otsu 阈值法和随机森林提取效果较好,随机森林由于算法的优越性,提取效果优于

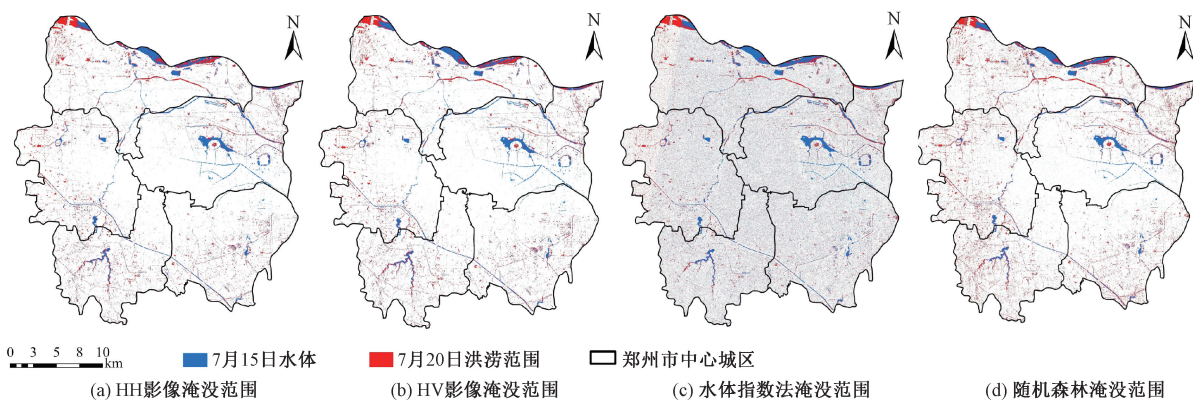


图 5 水体提取效果图
Figure 5 Maps of water extraction

表 2 精度评价表

Table 2 Precision evaluation table

水体提取方法		影像	水体样点数量	正确检验样点数量	总体分类精度/%
阈值分割	Otsu 阈值法	7. 15 HH	96	82	85. 42
		7. 15 HV	96	83	86. 46
		7. 20 HH	96	80	83. 33
		7. 20 HV	96	81	84. 38
	水体指数法	7. 15 GFWI	96	82	85. 42
		7. 20 GFWI	96	79	82. 29
随机森林		7. 15 HH+HV+GFWI	96	85	88. 54
		7. 20 HH+HV+GFWI	96	83	86. 46

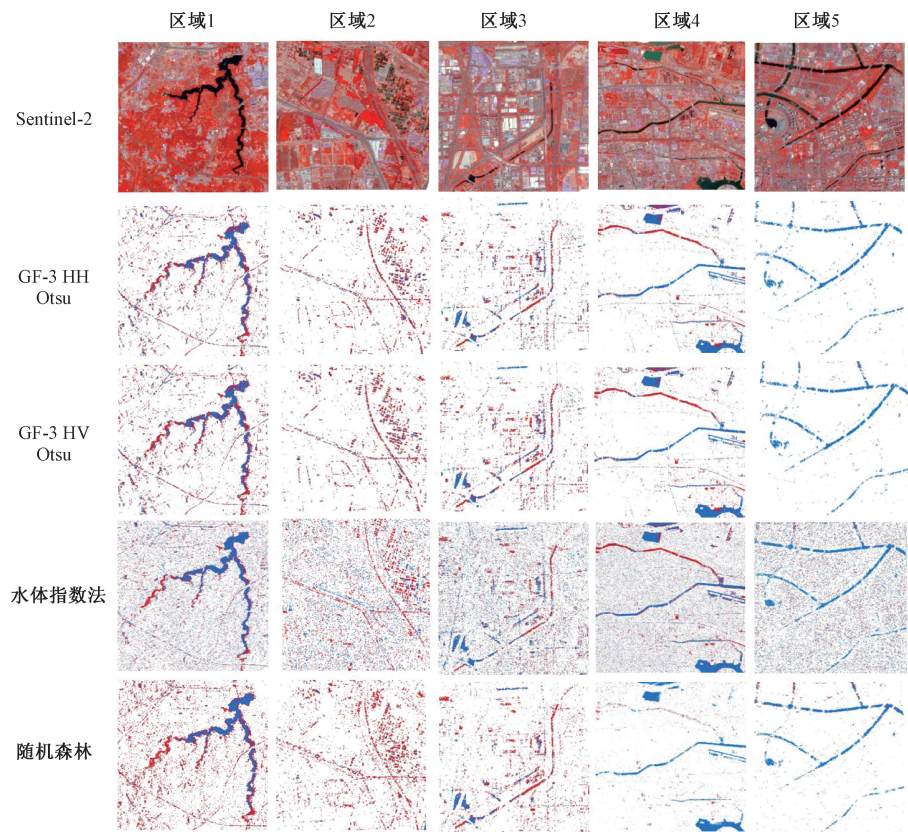


图 6 典型区域水体结果对比图
Figure 6 Comparison of flood monitoring results in typical

Otsu 阈值法。但是由于 Otsu 阈值法的阈值并非最优阈值,随机森林存在着一定程度的过拟合现象,3 个区域均存在着部分洪涝范围的误提取。在区域 4 中,Otsu 阈值法和水体指数法的提取效果一致,河流提取连续、间断少。但是由于 SAR 影像噪声的影响,随机森林样本选取受到限制,该区域河流周边淹没范围提取效果不理想。在区域 5 这 3 种方法的应用效果均不理想,无法对该区域洪涝淹没范围进行有效提取。Otsu 阈值法在该区域基本失效,水体指数法和随机森林可提取到少部分水体,其中后者提取的淹没范围大于前者,这与 7 月 20 日研究区遭遇突破历史极值的强降雨有关。强降雨对 SAR 信号的衰减作用^[18]和高层建筑物对 SAR 信号的散射吸收作用共同造成了该区域 SAR 影像地表信息缺失。

3.2 SAR 与社交媒体综合监测结果分析

本文将基于社交媒体数据的积水点与通过随机森林提取的 SAR 影像洪涝范围进行空间叠加,得到郑州市中心城区的洪涝监测图如图 7 所示。从图 7 中可以明显看出,基于 SAR 影像提取的洪涝淹没范围集中在城市周边地区;基于社交媒体数据提取的积水点在三环内比较密集,在三环外的分布比较分散。

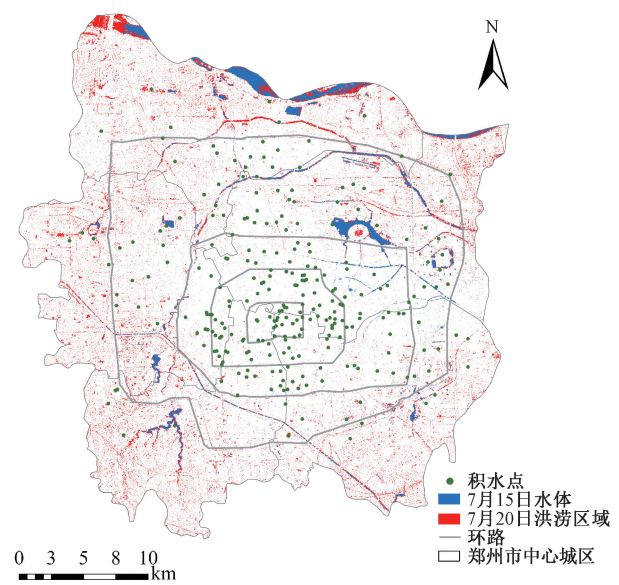


图 7 郑州市中心城区洪涝监测图

Figure 7 Flood monitoring map of Zhengzhou central urban area

对 SAR 遥感监测的洪涝淹没范围进行统计分析可知,相比于 7 月 15 日,7 月 20 日郑州市中心城区洪涝淹面积为 51.49 km²,在三环外,洪涝淹面积为 50.81 km²,占总体洪涝淹范围的 98.7%;三环内提取的洪涝淹面积为 0.68 km²,仅占总体的 1.3%。由此可知,基于 GF-3 影像提取的洪涝淹

没范围绝大部分位于三环外,特别是大型水体如水库、河流周围的淹没范围比较明显。而在发生大范围内涝的三环以内区域,基于 SAR 影像提取的洪涝淹没范围较小且较为分散。这表明基于 SAR 影像的洪涝灾害监测方法适用于城市周边地区大型水体的监测,在城市高密度建成区的洪涝监测效果有限。

通过微博平台共得到 225 个积水点。积水点主要集中在城市人口与建筑密集的核心区域,特别是三环以内,共分布 172 个积水点,占积水点总数的 76.4%。在三环外,内涝点比较分散,共分布 53 个积水点,占积水点总数的 23.6%。三环内存在大量老旧管网,当暴雨来临时,管道内雨水不能及时向河流排泄,降低整体管网行洪能力^[19],导致积水点的集中出现。如图 8 所示,对积水点进行核密度空间分析,可以看出积水点呈现多核空间分布。核密度高值区域相对比较集中,主要处在二环以内,核密度最高值区域出现在一环内,次一级高值区域位于二环东北部以及西南方向。从空间分布上看,积水点主要分布在城市中东部区域,城市西部区域分布较少。这一现象与 7 月 20 日当天暴雨中心从郑州市中心城区西部向中东部移动有关^[1]。

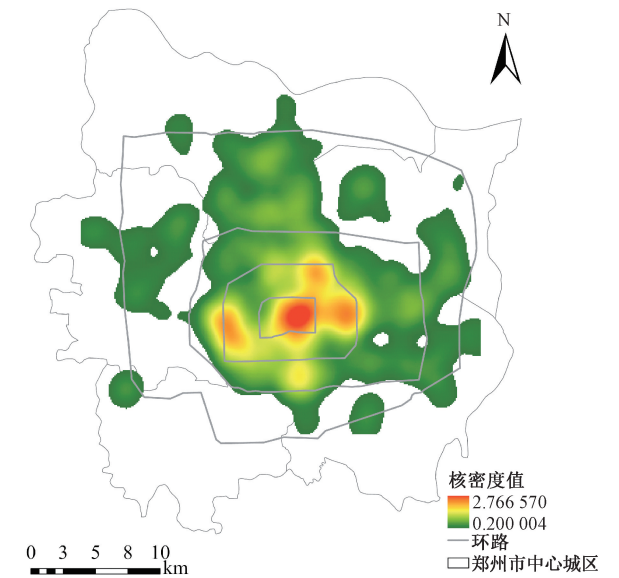


图 8 积水点核密度值图

Figure 8 Kernel density map of waterlogging points

由上述结果可知,社交媒体数据与 SAR 影像在城市洪涝淹没监测范围方面形成了明显的优势互补。遥感卫星监测范围大,能够更加真实客观地表征地表信息,一直是灾害监测和应急管理领域的常用手段。但受遥感卫星重返周期的客观限制,在突然且迅速的极端降雨或洪涝灾害发生时,可用的遥感数据源是其应用的限制性因素之一。SAR 影像相对于光学影像,虽然不易受天气条件影响,但城市

核心区域复杂的地物反射机制会对 SAR 影像成像造成影响,使得部分地物信息缺失,进而影响该地区的洪涝灾害监测。而基于社交媒体数据的洪涝灾害监测方法依托网络的便捷性,使用低成本、易获取的社交媒体数据,可快速便捷地提取城市内部的洪涝灾害信息,提高洪涝灾情判断的速度。因此在城市洪涝监测中,遥感数据可以与社交媒体数据相互补充,共同完成城市区域洪涝灾害监测。

4 结论

本文使用基于遥感数据和社交媒体数据相结合的洪涝监测方法,将提取精度最优的随机森林所提取的洪涝淹没范围与社交媒体数据提取的洪涝信息相结合,实现“7·20”郑州市中心城区洪涝监测。主要结论如下。

(1) SAR 水体总体提取精度从高到低依次为随机森林、Otsu 阈值法、水体指数法,但是在一些典型区域分析中,随机森林的提取效果低于其他方法。

(2) 基于 SAR 影像提取的洪涝淹没范围主要集中在三环外的城市边缘地区与大型水体周围,而基于社交媒体数据提取的洪涝信息主要集中在城市人口和建筑密集的三环内,在三环外的分布比较分散。

本文研究成果结合多源数据优势,获取更为全面的监测结果,为城市洪涝灾害治理与应急管理提供了更科学有效的参考。

参考文献:

[1] 刘家宏,裴羽佳,梅超,等. 郑州“7·20”特大暴雨内涝成因及灾害防控[J]. 郑州大学学报(工学版), 2023, 44(2): 38-45.
LIU J H, PEI Y J, MEI C, et al. Waterlogging cause and disaster prevention and control of “7·20” torrential rain in Zhengzhou[J]. Journal of Zhengzhou University (Engineering Science), 2023, 44(2): 38-45.

[2] SHI J, CUI L L, TIAN Z. Spatial and temporal distribution and trend in flood and drought disasters in East China[J]. Environmental Research, 2020, 185: 109406.

[3] LIANG J Y, LIU D S. A local thresholding approach to flood water delineation using Sentinel-1 SAR imagery [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 159: 53-62.

[4] 高寒新,陈波,孙洪泉,等. SAR 卫星影像洪水检测研究进展及展望[J]. 地球信息科学学报, 2023, 25(10): 1933-1953.
GAO H X, CHEN B, SUN H Q, et al. Research progress and prospect of flood detection based on SAR satellite images[J]. Journal of Geo-Information Science,

2023, 25(10): 1933-1953.

[5] 谷鑫志,曾庆伟,湛华,等. 高分三号影像水体信息提取[J]. 遥感学报, 2019, 23(3): 555-565.
GU X Z, ZENG Q W, SHEN H, et al. Study on water information extraction using domestic GF-3 image[J]. Journal of Remote Sensing, 2019, 23(3): 555-565.

[6] MARKERT K N, MARKERT A M, MAYER T, et al. Comparing sentinel-1 surface water mapping algorithms and radiometric terrain correction processing in Southeast Asia utilizing Google Earth Engine [J]. Remote Sensing, 2020, 12(15): 2469.

[7] LI Y, MARTINIS S, WIELAND M. Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence[J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2019, 152: 178-191.

[8] TANGUY M, CHOKMANI K, BERNIER M, et al. River flood mapping in urban areas combining RADAR-SAT-2 data and flood return period data [J]. Remote Sensing of Environment, 2017, 198: 442-459.

[9] LI J, HE Z, PLAZA J, et al. Social media: new perspectives to improve remote sensing for emergency response[J]. Proceedings of the IEEE, 2017, 105(10): 1900-1912.

[10] 邢子瑶,董芯蕊,胥菁莉,等. 融合 VGI 和遥感等多源数据的洪涝范围提取与模拟方法[J]. 地球信息科学学报, 2023, 25(9): 1869-1881.
XING Z Y, DONG X R, ZAN X L, et al. flood inundation mapping and estimation using VGI, remote sensing images and other multi-source data[J]. Journal of Geo-Information Science, 2023, 25(9): 1869-1881.

[11] 国务院灾害调查组. 河南郑州“7·20”特大暴雨灾害调查报告[R/OL]. (2021-01-21) [2024-06-04]. <https://www.mem.gov.cn/gk/sgcc/tbzdsgdcbg/202201/P020220121639049697767.pdf>.
The Disaster Investigation Group of the State Council. Disaster investigation report of Zhengzhou “7·20” heavy rain [R/OL]. (2021-01-21) [2024-06-04]. <https://www.mem.gov.cn/gk/sgcc/tbzdsgdcbg/202201/P020220121639049697767.pdf>.

[12] HUANG W L, DEVRIES B, HUANG C Q, et al. Automated extraction of surface water extent from sentinel-1 data[J]. Remote Sensing, 2018, 10: 797.

[13] 贾诗超,薛东剑,李成绕,等. 基于 Sentinel-1 数据的水体信息提取方法研究[J]. 人民长江, 2019, 50(2): 213-217.
JIA S C, XUE D J, LI C R, et al. Study on new method for water area information extraction based on Sentinel-1 data[J]. Yangtze River, 2019, 50(2): 213-217.

[14] 方匡南, 吴见彬, 朱建平, 等. 随机森林方法研究综述[J]. 统计与信息论坛, 2011, 26(3): 32-38.
FANG K N, WU J B, ZHU J P, et al. A review of technologies on random forests[J]. Statistics & Information Forum, 2011, 26(3): 32-38.

[15] TANG H L, LU S L, BAIG M H A, et al. Large-scale surface water mapping based on landsat and sentinel-1 images[J]. Water, 2022,14(9): 1454.

[16] 沈兰芝, 谌华, 范风云, 等. 基于 Q-OTSU 算法高分三号水体快速提取方法研究: 以河南鹤壁洪涝灾害监测为例[J]. 自然灾害学报, 2022, 31(4): 134-143.
SHEN L Z, SHEN H, FAN F Y, et al. Research on rapid extraction method of water based on Q-OTSU algorithm using domestic GF-3 image: a case of flood disaster monitoring in Hebi, Henan Province[J]. Journal of Natural Disasters, 2022, 31(4): 134-143.

[17] WANG Z Y, YE X Y. Social media analytics for natural disaster management[J]. International Journal of Geographical Information Science, 2018, 32(1): 49-72.

[18] 余水. 强降雨对合成孔径雷达风速反演的影响[D]. 杭州: 浙江大学, 2017.
YU S. Effects of large precipitation to the wind retrieval from synthetic aperture radar[D]. Hangzhou: Zhejiang University, 2017.

[19] 张金萍, 张朝阳, 左其亭. 极端暴雨下城市内涝模拟与应急响应能力评估[J]. 郑州大学学报(工学版), 2023, 44(2): 30-37.
ZHANG J P, ZHANG Z Y, ZUO Q T. Urban waterlogging simulation and emergency response capacity evaluation with extreme rainstorms[J]. Journal of Zhengzhou University (Engineering Science), 2023, 44(2): 30-37.

Urban Flood Disaster Monitoring Based on Remote Sensing and Social Media Data

ZHANG Hui, PAN Xiaoning, JIA Hao, XU Dehe

(College of Surveying and Geo-informatics, North China University of Water Resources and Electric Power, Zhengzhou 450046, China)

Abstract: Urban flood monitoring is crucial for disaster management and emergency response. Single data sources often have their own shortcomings, which are not conducive to the multi-dimensional and all-round analysis of complex urban flooding disasters. The GF-3 SAR images were combined, which were not affected by clouds and rain and had a large coverage, with real-time social media data to establish a flood monitoring method for the downtown area of Zhengzhou City during the "7·20" heavy rainstorm. Threshold segmentation and random forest were used to extract water bodies from GF-3 SAR images before and during the flood, the most accurate water body extraction results and flood monitoring results were obtained, and the effect of water body extraction in typical areas was analyzed. Using Python tools, the obtained social media data about urban waterlogging were processed, visualized and spatially analyzed. GF-3 SAR images and social media data flood monitoring results were combined to explore the complementary advantages of the two. The results showed that the overall extraction accuracy of SAR water bodies was in the order of random forest, Otsu threshold method, and water body index method. However, in some typical regional analyses, the extraction effect of random forest was lower than that of the other methods. The flood inundation range extracted based on SAR images was mainly concentrated in the urban fringe areas outside the Third Ring Road and around large water bodies. The flood information extracted based on social media data was mainly concentrated in the Third Ring Road, where the urban population was densely populated with buildings.

Keywords: flood disaster; GF-3; threshold segmentation; random forest; social media data