

基于大数据的极端暴雨事件下城市道路交通 及人群活动时空响应

易嘉伟^{1,2}, 王楠^{1,2}, 千家乐^{1,2}, 马廷^{1,2}, 杜云艳^{1,2}, 裴韬^{1,2},
周成虎^{1,2}, 涂文娜³, 刘张^{1,2}, 王会蒙^{1,2}

(1. 中国科学院地理科学与资源研究所, 北京 100101; 2. 中国科学院大学, 北京 100049;
3. 华中师范大学, 武汉 430079)

摘要: 随着全球气候变化加剧, 极端降雨增多, 暴雨内涝灾害频发, 严重威胁城市的可持续发展。快速掌握暴雨给城市交通及人群的影响, 有助于提高灾害应急管理水平和事件响应能力。利用实时动态的交通路况信息和手机定位请求数据, 通过一种融合 STL 时序分解技术与极端学生化偏差统计检验的时间序列异常探测方法, 监测和分析暴雨内涝灾害事件中, 城市道路交通和人群活动的时空响应特征, 并以 2018 年 7 月 16 日发生在北京的极端暴雨事件为例开展实证研究。研究结果显示, 在降雨集中的早、晚高峰两个时段(8—9 时、18—19 时), 市区的拥堵道路数量超出往常水平最高可达 150%, 异常检测分析显示拥堵道路数量和交通拥堵指数均达到异常甚至极端水平。人群活动的异常响应分析结果显示, 暴雨事件引起定位请求量异常升高、异常点增多, 且异常点的空间分布与 1 h 前的降雨量分布存在较高相关性。以上结果不仅证明了大数据及异常检测方法对于快速洞察暴雨事件对城市交通及人群影响的有效性, 也为城市暴雨内涝灾害的应急响应与管理提供了新的技术手段。

关键词: 暴雨事件; 城市内涝; 道路交通; 人群活动; 大数据; 异常检测

DOI: 10.11821/dlxb202003005

1 引言

随着全球气候变化的加剧, 极端天气事件日益增多, 城市暴雨内涝灾害愈发频繁^[1-2]。短时集中的降雨一旦超过城市排水能力, 会致使城市道路及社区等低洼地区发生积水, 严重的积水内涝不仅影响人们的正常出行、工作与生活, 还会带来巨大的经济损失, 直接危害人们的生命财产安全^[3-4]。根据住建部 2010 年专项调查, 自 2008 年以来中国有 213 个城市发生过积水内涝^[5]。2017 年, 中国有 104 座县级以上城市发生内涝, 城市洪涝造成 218.72 万人受灾, 直接经济损失 165.68 亿元^[6]。加强暴雨内涝灾害防御能力已成为中国城市可持续发展的重要课题^[7-11]。

快速掌握强降雨事件对城市交通及人群的影响, 是降低灾害损失、提升抗灾救灾能力的重要途径, 也是当前城市灾害研究的热点问题。国内外学者在这方面已经开展了很多研究工作, 其中, 关于降雨对城市道路交通影响的研究主要以模型分析为主^[12-16]。Versini 等^[14]研制了一种道路积水预警系统, 通过分布式水文模型预测不同程度降水可能

收稿日期: 2018-08-31; 修订日期: 2019-10-08

基金项目: 国家重点研发计划(2017YFB0503605); 中国科学院战略性先导科技专项(A类)(XDA19040501) [Foundation: National Key Research and Development Program of China, No.2017YFB0503605; Strategic Priority Research Program of the Chinese Academy of Sciences, No.XDA19040501]

作者简介: 易嘉伟(1988-), 男, 湖南衡阳人, 博士, 助理研究员, 主要从事地理数据挖掘研究。E-mail: yijw@lreis.ac.cn

导致的道路淹没点,实验结果显示该系统能够近100%的准确度预测积水点且误报率不超过30%。Yin等^[16]应用高精度2D水动力模型模拟不同程度降雨对上海中心城区路网的影响与风险评估,发现积涝程度与道路中断的响应关系是成比例但非线性的。通过构建模型可以模拟不同场景下,暴雨灾害对城市居民日常出行的影响,但模拟的精度受参数及实测数据校正的影响较大,模拟结果与实际受灾情况往往存在差异,难以客观地反映灾害的影响,不利于快速的灾害应急响应。另一方面,关于暴雨内涝对城市人群活动影响的研究主要通过灾后调查问卷的方式收集相关的研究数据^[17-18]。Ruin等^[17]通过调查问卷研究极端降雨事件下人类行为响应(或称社会响应)的时空特征,他们在法国瓦尔的实证研究显示对灾情的认知会强烈影响人们应对灾害时的行为选择。Papagiannaki等^[18]通过收集暴雨期间紧急电话量和灾后线上调查问卷,分析希腊阿提卡地区暴雨对人群的影响,他们发现影响强度与局地累积降雨量有关,对于灾害的恐慌与焦虑主要来自行程计划的调整与改变,而提前知晓灾害风险能够减轻这种焦虑。通过调查问卷能够详细了解灾害中人群的行为选择,但问卷收集的时间滞后性及样本有效性上的不足,使得这种方式较难实时全面地感知暴雨灾害对人群活动的影响。

在当前城市大数据技术的发展和推动下,大量实时动态的城市交通路况数据、手机定位数据、社交媒体签到等不仅成为弥补传统观测数据短缺的重要数据源^[19-21],更是促进城市研究乃至地理学研究步入全新发展阶段的重要催化剂^[22-23]。甄峰等^[22]认为大数据为智慧城市研究与规划提供了新的数据思维,丰富了传统城市研究与规划理论及方法。程昌秀等^[23]指出地理大数据与复杂性科学的相互融合支撑将成为新时代地理学的主流科学方法。并且,近年来,很多研究已经开始将大数据的分析手段融入到城市灾害研究中,以突破传统方法的局限^[24-26]。Xiao等^[25]利用twitter数据分析发现纽约市在飓风Sandy事件中的用户发推量与灾损之间存在倒“U”型的非线性关系。Li等^[27]融合了社交媒体数据与遥感数据来监测强降雨给武汉、深圳等城市的影响,克服遥感数据在应急响应中的局限性。

因此,为了突破传统方法在快速感知灾害影响的局限,检验大数据在灾害研究中的实用价值,本文尝试利用实时更新的道路交通数据和腾讯定位请求数据,通过时序异常探测方法分析暴雨引发的城市交通及人群活动异于平常的时空响应特征,实现暴雨内涝事件对城市影响的量化表达。主要的创新之处在于利用大数据更新快、时效性高、来源多样化的特点,通过异常监测技术,近实时地感知并量化灾害事件下城市交通及人群活动的现时变化,为灾害应急部门提供决策依据。

本文以北京市2018年7月16日的暴雨事件为例开展实证研究。北京是中国的政治、文化中心,拥有超过2000万的常住人口,2012年曾发生过“7·21”特大暴雨事件,造成79人死亡^[28],直接经济损失140多亿元^[29]。2018年7月15日夜间至16日,北京市发生连续降雨,造成市区多处出现积水内涝,北京市气象局将此次降雨命名为“7·16”暴雨事件,列入近8年的极端暴雨事件之一。因此,本文选择此次事件开展实证研究具有代表性。

2 研究资料与方法

2.1 研究区概况

本文以北京市2018年7月16日的暴雨事件为例开展实证研究,研究范围包括北京16个市辖区,总面积16412 km²,常住人口2172.9万^[30],人口和道路集中分布在城六区(东城区、西城区、朝阳区、海淀区、丰台区、石景山区)。北京市位于华北平原西北边缘,背靠太行山余脉和燕山山脉,属于暖温带半湿润大陆性季风气候,全年降雨集中在6—8

月份, 降雨期西北山区易发生暴雨泥石流灾害, 而位于平原区的城六区易发生积水内涝灾害。

2.2 数据资料

(1) 降雨与道路积水数据

降雨数据主要包括由中国气象数据网 (<http://data.cma.cn/>) 提供的气象站逐时降水量数据和气象雷达反演的 1 h 累计降水量数据。其中, 气象站逐时降水量数据包含北京市 18 个台站, 雷达反演数据来自位于大兴区的基站, 空间范围覆盖整个北京市, 台站和基站的分布见图 1。道路积水点位置数据主要通过收集北京交通广播、新闻媒介、网友等在微博及互联网媒体上发布的图文信息, 并对文本中的地名信息进行地理编码而获得。受数据采集方式的限制, 本文收集的积水点数据仅反映降雨事件中的部分积水位置, 并且由于本文描述的模糊性, 积水点坐标主要定位在道路附近或道路交叉口位置, 与实际积水位置存在一定偏离, 因而在本文的分析中, 选择积水点附近的道路交通情况进行统计分析, 以保证分析结果的客观性。

(2) 交通指数

北京市路况数据通过四维交通指数网站 (<http://nitrafficindex.com/>) 抓取获得, 包含 7 月 9—16 日期间 1278 个路段每 15 min 的交通指数数据。该交通指数结合了道路实际速度及道路通行条件, 并加入对交通拥堵的主观感受, 量化地表达了道路的交通运行状况。指数越大, 道路越拥堵。根据四维图新对该指数的划分标准 (表 1), 本文将交通指数在 7 以上的道路定义为拥堵路段。

表 1 道路交通指数分级标准
Tab. 1 Index categorization of road traffic

拥堵指数	0~2	2~4	4~7	7~10	10~18	18及以上
拥堵分级	畅通	基本畅通	轻度拥堵	拥堵	严重拥堵	道路瘫痪

(3) 定位请求数据

反映人群活动及分布的定位请求数据是通过腾讯位置大数据网站 (<https://heat.qq.com/>) 抓取获得。该网站收集了手机用户使用腾讯地图位置服务而公开的位置数据, 为了保护用户隐私, 数据按 1 km×1 km 的网格单元进行了逐小时聚合。在本文中, 数据的时间范围为 7 月 9 日—16 日连续 8 d。该数据仅限于使用腾讯位置服务的用户群体, 且在不同年龄层存在样本偏差, 因而分析结果所反映的人群响应特征亦局限于部分群体。但鉴于腾讯公司的即时通讯软件及其他涉及定位服务调用的软件拥有目前中国最大的用户群体, 使得该数据具有较好的代表性。

2.3 数据处理与研究方法

为了定量分析暴雨内涝给城市交通及人群活动带来的影响, 本文采用了一种融合 STL 时序分解技术^[31]与极端学生化偏差 (Extreme Studentized Deviate, ESD)^[32]统计检验的时间序列异常探测方法^[33], 简称 S-ESD 方法, 通过对比交通指数和定位请求量在暴雨期间监测到的异常表现与以往同期的常态表现, 分析城市交通及人群活动响应暴雨事件的时空演变特征。

本文中, 整个研究区或任意路段、网格的交通指数和定位请求量均可抽象为等间隔时间序列 $T_s = \{Q_i\}$, i 表示时间戳。交通指数数据分析包含 1278 个路段 15 min 为间隔的时间序列数据, 时间范围为 2018 年 7 月 9 日 0 时至 2018 年 7 月 16 日 24 时; 定位请求数据分析包含北京市行政区划内时间序列中值大于 0 的所有公里网格, 时间间隔为 1 h, 时间范围与交通指数一致。S-ESD 方法分两步对时间序列 T_s 进行异常检测:

(1) 首先, 通过 Piecewise Median 改进的 STL 时序分解方法提取时间序列 T_t 的残余项。城市的交通和人群活动都具有明显的昼夜周期特征和长时间的趋势变化, 采用时序分解方法准确剔除时间序列中的趋势与周期, 是降低异常误报率的关键。由于传统的 STL 方法对数据中的异常值极为敏感, S-ESD 方法提出 Piecewise Median 技术, 即划分多个分段, 用分段的中值来估计趋势。

(2) 然后, 对残余值序列进行 ESD 统计检验, 提取具有统计显著性的异常时间点。经典的 ESD 统计检验在给定最大异常点数量 r 参数条件下, 通过样本均值 \bar{x} 和方差 s 计算统计量 R_i :

$$R_i = \frac{\max |x_i - \bar{x}|}{s} \quad (1)$$

每次计算完, 剔除最大化 $|x_i - \bar{x}|$ 的样本值并重复计算, 从而检测有多少个不超过 r 的异常点达到 0.05 的显著性水平。由于样本均值和方差的计算对数据的异常值敏感, S-ESD 方法采用中值和绝对中位差 (MAD) 来计算 R_i , 提高异常检测方法的健壮性。

在刻画研究区域的总体异常表现上, 本文采用了异常数量 (AQ) 和异常偏离量 (AD) 两个定量指标。AQ 是指任意时段研究区检测到交通指数或定位请求量的异常道路或网格的数量总和, 而 AD 是指所有异常点的实际值与期望值 (趋势和周期分量之和) 偏差的总和。本文将实际值大于期望值的异常称为正异常, 而小于期望值的异常为负异常。在监测 AQ 和 AD 指标随时间变化的表现时, 本文采用统计学中常用的 Tukey's Fences 方法^[34], 当样本值大于 $Q3+1.5(Q3-Q1)$ 或小于 $Q1-1.5(Q3-Q1)$ 时为异常值, 当样本值大于 $Q3+3(Q3-Q1)$ 或小于 $Q1-3(Q3-Q1)$ 时为极端值。

3 结果与分析

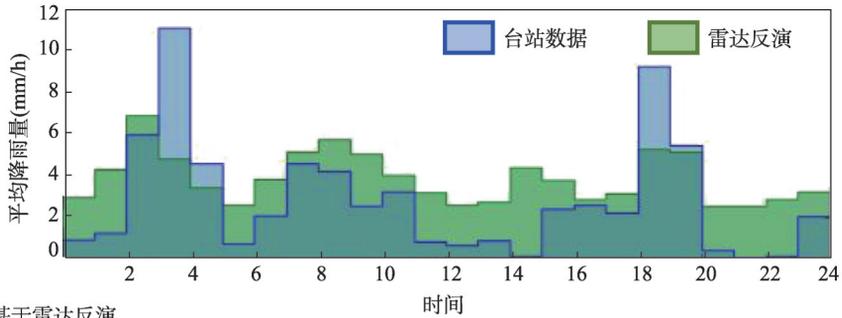
3.1 “7·16”暴雨雨情

根据本文收集的 18 个气象站的降水数据, 从 15 日 20 时—16 日 20 时 (24 h), 北京市平均降雨量 (站点观测值的算数平均) 为 60.3 mm。其中, 城六区的平均降雨量达到 93.8 mm, 超过 50 mm (暴雨) 的台站 11 个, 超过 100 mm (大暴雨) 的台站有 4 个。如图 1 所示, 16 日凌晨 3 时的小时降雨量高达 11 mm, 白天降雨主要集中在早高峰 8 时和晚高峰 18 时两个时段。从雷达反演的小时降雨量分布图上可以看到, 早高峰 8—9 点时段, 降雨带沿西南—东北方向分布, 覆盖海淀、昌平、怀柔、密云、丰台、石景山等地区, 最大降雨区域位于海淀和昌平地区, 小时降雨量超过 38 mm; 晚高峰 18—19 点时段, 降雨区域几乎覆盖北京市平原地区, 降雨主要集中在昌平、海淀、丰台、石景山、门头沟等西部地区, 城区降雨量在 5 mm 以上。根据北京市气象局通报, 此次降雨共持续 58 h, 是 2018 年入汛以来的最强降雨过程, 因降雨集中在早、晚高峰时段, 暴雨引发的道路积水和交通拥堵在诸多媒体和社交网络上都有报道。

3.2 对道路交通的影响

为了探究暴雨事件对城市道路交通的影响, 本文从全市道路总体和道路个体两种尺度, 对拥堵道路数量、拥堵指数、以及异常道路数量进行了分析。首先, 从全市拥堵道路的数量变化上看 (图 2), 7 月 16 日 8—9 时和 18—19 时附近呈现明显峰值特征, 拥堵道路数量为 7 月 9 日 (上周一) 同时段的 2 倍左右, 比一周平均水平高出 14%~150%, 并且异常检测结果显示在这两个时段正异常聚集发生。此外, 全市交通指数总和的分析结果显示, 除了 8—9 和 18—19 两个时段为峰值且正异常集中分布外, 凌晨 3—4 时亦表现为

a. 基于台站观测和雷达反演



b. 基于雷达反演

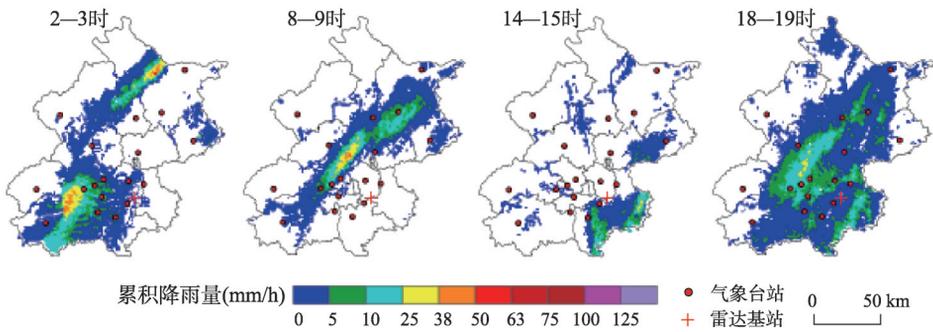


图1 2018年7月16日北京地区小时平均降雨量和累积降雨量空间分布

Fig. 1 1-hour rainfall variations observed by in-situ stations and weather radar (a) and precipitation distribution at different hours observed by weather radar (b) in Beijing on July 16th, 2018

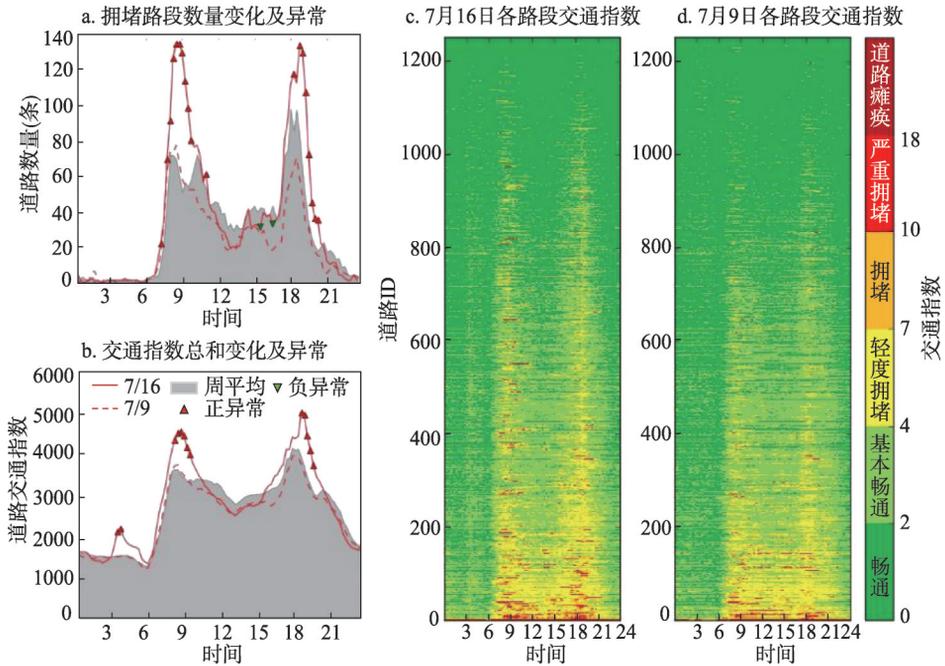


图2 北京市2018年7月16日道路拥堵数量及交通指数变化情况

Fig. 2 Changes in the number of road congestion and traffic index in Beijing on July 16th, 2018

局部峰值并检测出正异常, 这表明因为人的作息规律, 该时段道路车流量较小, 降雨虽未引起道路出现拥堵, 但仍然导致交通指数总量相比周平均增长了约40.9%。对比7月16日与7月9日同时段不同路段的交通指数也可以看到不少道路3时、9时、18时有较明显的差异。

其次, 在道路个体尺度上, 本文对每条路段的交通指数时间序列进行了异常检测, 并对比了7月16日每个时段正、负异常道路的数量和指数与一周平均水平的差异。从图3展示的结果上看, 交通指数异常高的路段数量AQ及这些路段的指数异常偏离量AD在当天的3—4时、8—9时、和18—19时3个时段明显偏离一周平均水平, 根据Tukey's Fences标准已经达到异常甚至极端水平, 而交通指数异常低的AQ和AD值在中午12—13时略高于同时段一周统计的上四分位数。上述结果表明, 受降雨的影响, 城市道路的交通指数出现异常升高, 异常道路的数量以及偏移量明显高于往常水平, 而中午12时一些道路异于平常的通畅可能与降雨抑制人们的出行有关。

具体从路网交通指数分布(图4)上看, 8—9时段的严重拥堵路段数较多, 主要分布在回龙观、西二旗、中关村、东坝、广安门、十里河、玉泉营等区域, 并且回龙观、

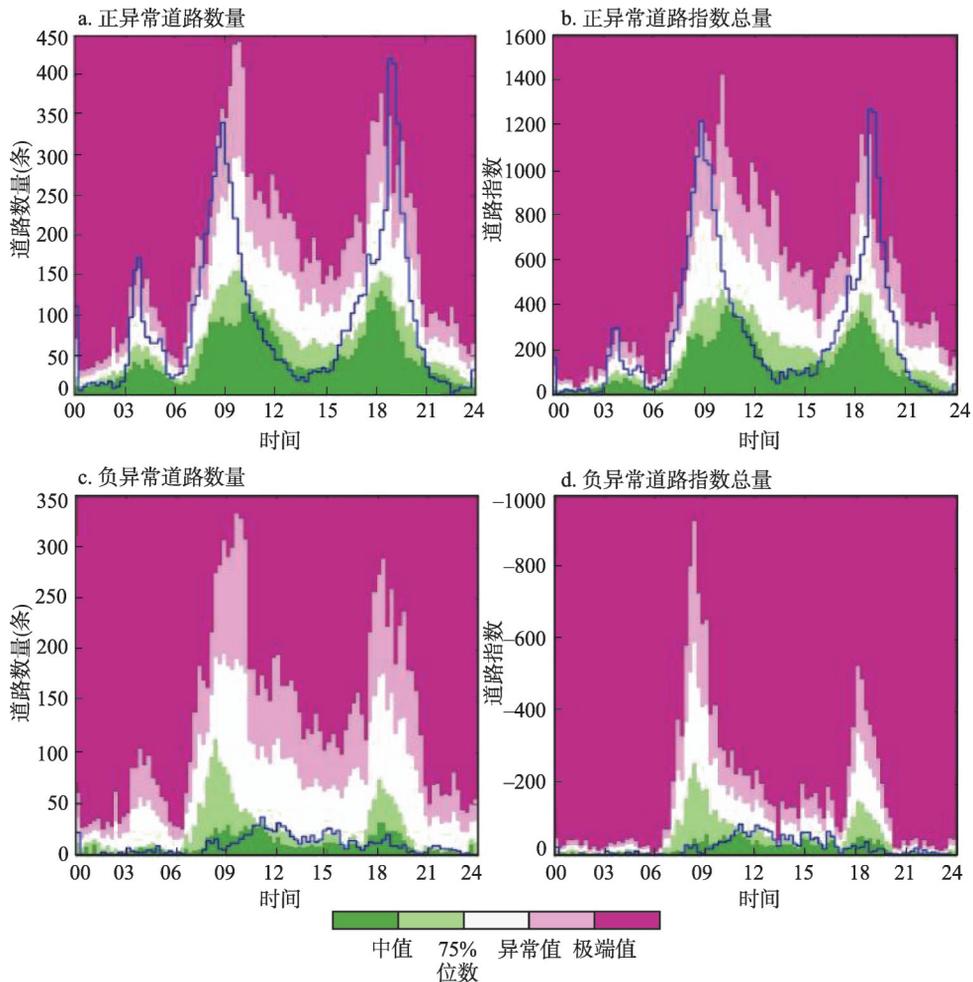


图3 道路异常检测分析结果

Fig. 3 Anomaly detection of congested roads and traffic index

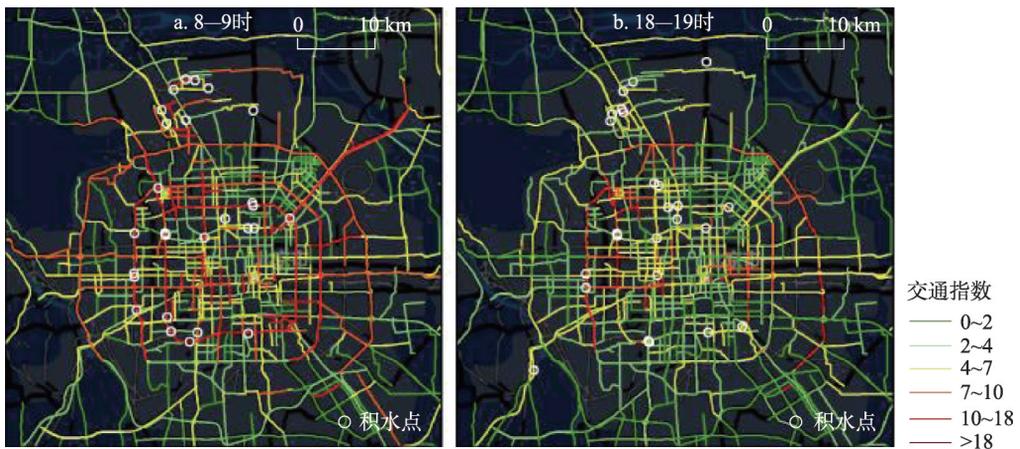


图4 早、晚高峰时段北京城区路网交通指数与积水点分布情况

Fig. 4 Road traffic and inundation distribution during morning and evening rush hours

西二旗、丰益桥等地点附近存在不同程度的道路积水，如回龙观育知东路、上地东路等。18—19时段的严重拥堵路段主要集中在万泉河桥、联想桥、长椿街等地点附近，尽管上地、回龙观、六道口等附近仍有不少道路存在积水，但对交通的影响已小于早高峰时段。如表2所示，相比早高峰时段，晚高峰时段的降雨导致的积水点数量略有减少，积水点附近道路交通指数及异常偏离量指数均已下降，导致这种变化的潜在因素是多方面的，首先，道路交通的有效疏导能够逐步缓解暴雨内涝给城市

交通带来的影响，其次，预警信息及新闻媒体的舆情传播能够提高公众在暴雨事件中的警惕性，从而改变人们在高峰时段的出行行为，避开因积水而拥堵甚至中断的道路。

3.3 对人群活动的影响

为了探究“7·16”暴雨事件对城市人群活动的影响，本文对7月9日—16日的定位请求时间序列数据进行了逐网格的异常探测分析。分析结果中，绝大多数异常网格为正异常，即实际值高于期望水平，而负异常网格数量较少，且与暴雨事件无明显响应特征，因而下文仅对正异常的分析结果进行阐述和讨论。

首先，从正异常点数量上看（图5），7月16日8—9时、16—17时、19—20时三个时段的异常网格数量较多，高出一周统计的同时段Q3水平，而夜间22—24时段低于平均（中值）水平；从异常偏离量上看，前面3个时段同样明显高于平均（中值）水平，而22—24时段较往常明显偏低。由此可见，8—9时、16—17时、以及19—20时3个时段的降雨引起了不少地区人们使用定位请求服务（如打车、导航等）的频次异常增高，而预警信息的发布以及鼓励市民错峰出行、提前放学下班等应对措施的执行是导致晚间22—24时段定位请求行为明显低于平常的潜在原因。

其次，通过核密度方法对正异常点的空间分布（图6）进行可视化分析，可以看到8—9时段，定位请求异常的网格与此时段的降雨分布相似，呈西南—东北方向分布，顺

表2 早晚高峰时段道路积水点及
周边道路拥堵情况

Tab. 2 Statistics of road congestion near inundation spots during morning and evening rush hours

时段	积水点数量 ^①	邻近道路交通指数 ^②	邻近道路交通指数异常偏离量 ^②
8—9时	28	9.91	3.20
18—19时	25	6.99	2.10

注：①^{*}积水点数据主要来自北京交通广播、新闻媒介、网友等在微博及互联网媒体上发布的图文信息整理并地理编码而来，可能存在未统计到的积水路段；②^{**}表中的交通指数及异常偏离量指数均为积水点周边邻近道路统计均值。

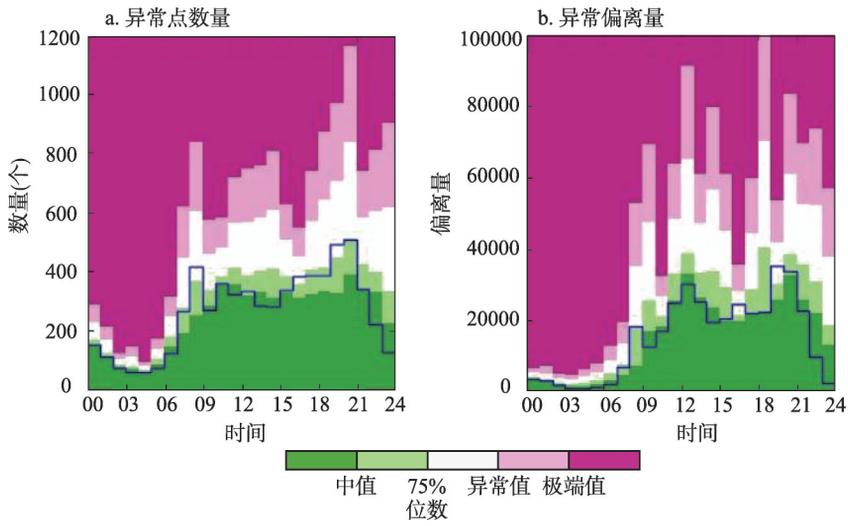


图5 定位请求量异常网格点检测分析结果

Fig. 5 The number of anomalous grids and the deviations of location requests

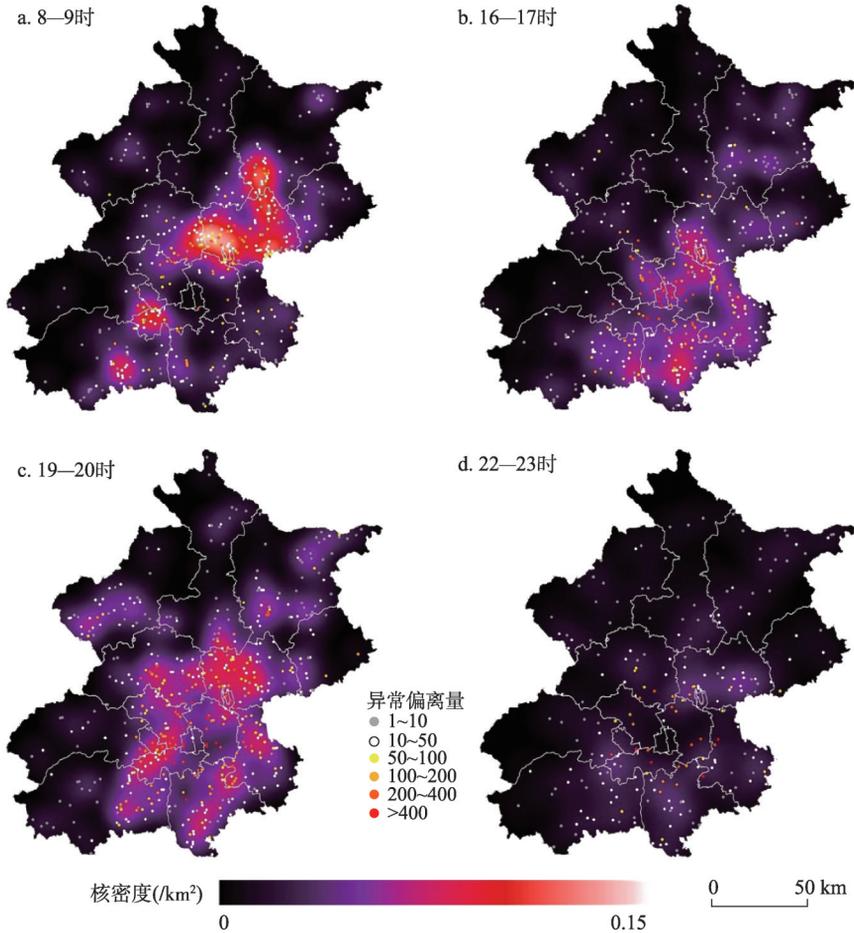


图6 基于核密度方法的不同时段异常网格点空间分布

Fig. 6 Spatial distribution of anomalous grids at different times based on kernel density analysis

义、密云的人口密集区异常分布较为集中, 数量较多; 16—17时段, 北京南部大兴区异常点分布较为密集; 19—20时段, 京西丰台、石景山、海淀、昌平四区, 以及顺义区的异常点分布较为密集。

为了探究异常点的分布特征与降雨分布的空间响应关系, 本文进一步将8—9时、16—17时、19—20时3个时段的异常点核密度值与对应时段及前后一段时间雷达反演的降雨量分布数据进行Spearman秩相关^[35]分析。结果显示(图7), 上述时段的异常点分布与1 h前降雨分布的相关系数在降雨前后2 h范围内达到峰值, 相关系数在0.4~0.6之间, 且在0.05的置信水平上显著相关。上述结果表明, 暴雨不仅会引起人群活动的异常响应, 并且异常点的空间分布与降雨量的空间分布相关, 但有1 h的时间滞后, 即当前时段的强降雨会在1 h后引起降雨地区的人群活动异常。

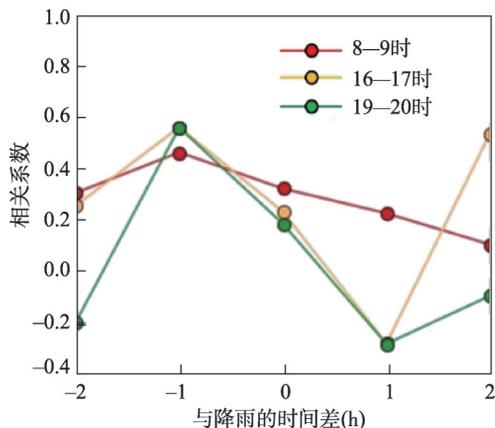


图7 不同时段异常点核密度值与前后时段降雨量的相关性分析结果

Fig. 7 Correlation between the kernel density of anomalous grids and precipitation intensity with different time lags

4 总结与展望

暴雨内涝是近年来中国诸多城市面临的重要气象灾害之一, 及时掌握暴雨事件对城市道路交通及人群的影响, 有助于提高城市防灾减灾能力, 降低灾害的经济损失与社会影响。本文以北京“7·16”暴雨事件为例, 采用了一种融合STL时序分解技术与极端学生化偏差统计检验的时间序列异常探测方法, 利用实时动态的路况信息与定位请求数据, 对比降雨期间监测到的异常表现与以往同期的常态表现, 分析暴雨内涝对城市道路交通及人群活动的影响。

分析结果显示, 7月16日北京市降雨主要集中在3—4时、8—9时、和18—19时3个时段, 其中, 3—4时因凌晨人们的出行等活动很少, 因而对城市的道路及人群几无影响; 而在8—9时和18—19时两个通勤高峰时段, 降雨引发了城市多条道路异于平常的拥堵, 拥堵道路数量高出一周平均水平14%到150%不等, 交通指数异常高的道路数量在相同时段一周统计结果中也表现为异常值甚至极端值。同时, 通过对定位请求数据的监测与分析, 本研究发现城市的人群活动在早晚高峰时段的降雨也呈现相应的时空响应变化。在“7·16”暴雨事件中, 降雨导致定位请求异常升高, 异常网格点数量高于一周平均水平, 其空间分布与降雨量分布相关, 但存在1 h的滞后。

本文通过高时空分辨率的雷达数据、交通指数、定位请求等大数据及时序异常探测方法, 监测到城市道路交通及人群活动在暴雨事件下的响应特征, 利用该特征能够近实时地掌握暴雨灾害过程中的雨情变化对城市不同空间位置的影响。因此, 在遇到类似暴雨事件时, 本文建议结合以上大数据资源及其分析手段, 通过实时交通指数的变化监测掌握城市交通的运行状况与恢复情况, 指导交管部门开展交通疏导工作; 利用手机定位请求数据快速了解受影响人群的空间分布与时空变化, 合理部署应急救援力量, 提高对暴雨内涝灾害的应急响应能力。后续的研究工作还将进一步结合大数据方法与气象预报

模型、交通模型、灾情评估等技术,实现暴雨内涝灾害的情景模拟与提前预警,降低灾害对城市带来的影响,有效提高城市抗灾减灾能力。

参考文献(References)

- [1] Wang Yingchun, Zheng Dawei, Li Qingchun. Urban Meteorological Disaster. Beijing: China Meteorological Press, 2009. [王迎春, 郑大玮, 李青春. 城市气象灾害. 北京: 气象出版社, 2009.]
- [2] Jha A, Lamond J, Bloch R, et al. Five feet high and rising cities and flooding in the 21st century//Policy Research Working Paper 5648. Washington: The World Bank, 2011.
- [3] Hu Haibo. Research progress of surging urban flood risks. Progress in Geography, 2016, 35(9): 1075-1086. [扈海波. 城市暴雨积涝灾害风险突增效应研究进展. 地理科学进展, 2016, 35(9): 1075-1086.]
- [4] Ye Chaofan, Zhang Yichi, Cheng Weiming, et al. Analysis on urban pluvial floods and causes during rapid development in Beijing. China Flood & Drought Management, 2018, 28(2): 19-25. [叶超凡, 张一驰, 程维明, 等. 北京市区快速城市化进程中的内涝现状及成因分析. 中国防汛抗旱, 2018, 28(2): 19-25.]
- [5] Wang Wei, Chen Renze, Liu Yi, et al. Why big cities experienced waterlogging frequently. People's Daily, 2012-07-24. [王伟, 陈仁泽, 刘毅, 等. 大城市为何频频内涝. 人民日报, 2012-07-24.]
- [6] State Flood Control and Drought Relief Headquarters, Ministry of Water Resources. China Water and Drought Disaster Bulletin 2017. Beijing: Sinomap Press, 2017. [国家防汛抗旱总指挥部, 中华人民共和国水利部. 中国水旱灾害公报 2017. 北京: 中国地图出版社, 2017.]
- [7] Xie Yiyang, Han Suqin, You Lihong, et al. Risk analysis of urban rainfall waterlogging in Tianjin City. Scientia Meteorologica Sinica, 2004, 24(3): 342-349. [解以扬, 韩素芹, 由立宏, 等. 天津市暴雨内涝灾害风险分析. 气象科学, 2004, 24(3): 342-349.]
- [8] Jing Yinna, Yin Zhane, Yin Jie, et al. GIS-based analysis on rainstorm waterlogging hazards in Pudong New Area in Shanghai. Journal of Catastrophology, 2010, 25(2): 58-63. [景垠娜, 尹占娥, 殷杰, 等. 基于GIS的上海浦东新区暴雨内涝灾害危险性分析. 灾害学, 2010, 25(2): 58-63.]
- [9] Gu Xiaotian, Li Ning, Zhou Yang, et al. Thinking on urban waterlogging disaster defense initiated by "7·21" extraordinary rainstorm in Beijing. Journal of Natural Disasters, 2013, 22(2): 1-6. [顾孝天, 李宁, 周扬, 等. 北京“7·21”暴雨引发的城市内涝灾害防御思考. 自然灾害学报, 2013, 22(2): 1-6.]
- [10] Zhang Dongdong, Yan Denghua, Wang Yicheng, et al. Research progress on risk assessment and integrated strategies for urban pluvial flooding. Journal of Catastrophology, 2014, 29(1): 144-149. [张冬冬, 严登华, 王义成, 等. 城市内涝灾害风险评估及综合应对研究进展. 灾害学, 2014, 29(1): 144-149.]
- [11] Yin Jie, Xu Shiyuan, Jing Yameng, et al. Evaluating the impact of fluvial flooding on emergency responses accessibility for a mega-city's public services: A case study of emergency medical service. Acta Geographica Sinica, 2018, 73(9): 1737-1747. [殷杰, 许世远, 经雅梦, 等. 基于洪涝情景模拟的城市公共服务灾害应急响应空间可达性评价: 以医疗急救为例. 地理学报, 2018, 73(9): 1737-1747.]
- [12] Sun Chao, Zhong Shaobo, Deng Yu. Scenario deduction based emergency rescue plan assessment and decision optimization of urban rainstorm water-logging: A case study of Beijing. Acta Geographica Sinica, 2017, 72(5): 804-816. [孙超, 钟少波, 邓羽. 基于暴雨内涝灾害情景推演的北京市应急救援方案评估与决策优化. 地理学报, 2017, 72(5): 804-816.]
- [13] Yin Zhane, Xu Shiyuan, Yin Jie, et al. Small-scale based scenario modeling and disaster risk assessment of urban rainstorm water-logging. Acta Geographica Sinica, 2010, 65(5): 553-562. [尹占娥, 许世远, 殷杰, 等. 基于小尺度的城市暴雨内涝灾害情景模拟与风险评估. 地理学报, 2010, 65(5): 553-562.]
- [14] Versini P, Gaume E, Andrieu H. Application of a distributed hydrological model to the design of a road inundation warning system for flash flood prone areas. Natural Hazards and Earth System Sciences, 2010, 10: 805-817.
- [15] Yang S, Yin G, Shi X, et al. Modeling the adverse impact of rainstorms on a regional transport network. International Journal of Disaster Risk Science, 2016, 7(1): 77-87.
- [16] Yin J, Yu D, Yin Z, et al. Evaluating the impact and risk of pluvial flash flood on intra-urban road network: A case study in the city center of Shanghai, China. Journal of Hydrology, 2016, 537: 138-145.
- [17] Ruin I, Lutoff C, Boudevillain B, et al. Social and hydrological responses to extreme precipitations: An interdisciplinary strategy for postflood investigation. Weather, Climate, and Society, 2014, 6: 135-153.
- [18] Papagiannaki K, Kotroni V, Lagouvardos K, et al. Urban area response to flash flood-triggering rainfall, featuring human behavioral factors: The case of 22 October 2015 in Attica, Greece. Weather, Climate, and Society, 2017, 9: 621-638.

- [19] Goodchild M F, Glennon J A. Crowdsourcing geographic information for disaster response: A research frontier. *International Journal of Digital Earth*, 2010, 3(3): 231-241.
- [20] Granell C, Ostermann F O. Beyond data collection: Objectives and methods of research using VGI and geo-social media for disaster management. *Computers, Environment and Urban Systems*, 2016, 59: 231-243.
- [21] Liu Baowang, Shi Enming. Spatial pattern of population daily flow among cities based on ICT: A case study of "Baidu Migration". *Acta Geographica Sinica*, 2016, 71(10): 1667-1679. [刘望保, 石恩名. 基于ICT的中国城市间人口日常流动空间格局: 以百度迁徙为例. *地理学报*, 2016, 71(10): 1667-1679.]
- [22] Zhen Feng, Qin Xiao. The application of big data in smart city research and planning. *Urban Planning International*, 2014, 29(6): 44-50. [甄峰, 秦箫. 大数据在智慧城市研究与规划中的应用. *国际城市规划*, 2014, 29(6): 44-50.]
- [23] Cheng Changxiu, Shi Peijun, Song Changqing, et al. Geographic big-data: A new opportunity for geography complexity study. *Acta Geographica Sinica*, 2018, 73(8): 1397-1406. [程昌秀, 史培军, 宋长青, 等. 地理大数据为地理复杂性研究提供新机遇. *地理学报*, 2018, 73(8): 1397-1406.]
- [24] Barberia L, Amaro J, Aran M, et al. The role of different factors related to social impact of heavy rain events: Considerations about the intensity thresholds in densely populated areas. *Natural Hazards and Earth System Sciences*, 2014, 14: 1843-1852.
- [25] Xiao Y, Huang Q, Wu K. Understanding social media data for disaster management. *Natural Hazards*, 2015, 79(3): 1663-1679.
- [26] Wu Xianhua, Xiao Yang, Li Lianshui, et al. Review and prospect of the emergency management of urban rainstorm waterlogging based on big data fusion. *Chinese Science Bulletin*, 2017, 62(9): 920-927. [吴先华, 肖扬, 李廉水, 等. 大数据融合的城市暴雨内涝灾害应急管理评述. *科学通报*, 2017, 62(9): 920-927.]
- [27] Li J, He Z, Plaza J, et al. Social media: New perspectives to improve remote sensing for emergency response. *Proceedings of the IEEE*, 2017, 105(10): 1900-1912.
- [28] State Flood Control and Drought Relief Headquarters, Ministry of Water Resources. *China Water and Drought Disaster Bulletin 2012*. Beijing: China Water & Power Press, 2012. [国家防汛抗旱总指挥部, 中华人民共和国水利部. *中国水旱灾害公报 2012*. 北京: 中国水利水电出版社, 2012.]
- [29] Yang Jinhuai, Ding Yueyuan, Liu Dagen. Revelation and strategy of natural disasters such as Beijing "7·21" severe rainstorm. *China Water Resources*, 2012, 17: 29-31. [杨进怀, 丁跃元, 刘大根. 北京"7·21"特大暴雨自然灾害的启示与对策思考. *中国水利*, 2012, 17: 29-31.]
- [30] Beijing Municipal Bureau of Statistics, Survey Office of the National Bureau of Statistics in Beijing. *Beijing Statistical Yearbook 2017*. Beijing: China Statistics Press, 2018. [北京市统计局, 国家统计局北京调查总队. *北京统计年鉴-2017*. 北京: 中国统计出版社, 2018.]
- [31] Cleveland R, Cleveland W, MaRae J, et al. STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 1990, 6(1): 3-73.
- [32] Rosner B. On the detection of many outliers. *Technometrics*, 1975, 17(2): 221-227.
- [33] Vallis O, Hochenbaum J, Kejarival A. A novel technique for long-term anomaly detection in the cloud//*Proceedings of the 6th USENIX Conference on Hot Topics in Cloud Computing*, 2014: 15-21.
- [34] Tukey J W. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley, 1977.
- [35] Hollander M, Wolfe D A. *Nonparametric Statistical Methods*. New York: Wiley, 1973.

Spatio-temporal responses of urban road traffic and human activities in an extreme rainfall event using big data

YI Jiawei^{1,2}, WANG Nan^{1,2}, QIAN Jiale^{1,2}, MA Ting^{1,2}, DU Yunyan^{1,2}, PEI Tao^{1,2},
ZHOU Chenghu^{1,2}, TU Wenna³, LIU Zhang^{1,2}, WANG Huimeng^{1,2}

(1. Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Central China Normal University, Wuhan 430079, China)

Abstract: As global climate change intensifies, extreme rainfalls and floods become more frequent and pose a serious threat to urban sustainable development. Fast assessment of the rainfall disaster impact upon urban traffic and population plays an important role in improving disaster emergency management and incident response capabilities. This study adopts a time series anomaly detection method to discover and quantify the impact of rainfall-triggered flood on road traffic and human activities using real-time traffic condition information and mobile phone location request data. The anomaly detection method combines the STL time series decomposition technique and the extreme student deviation statistics to identify the response characteristics of traffic data and location requests during the event. The extreme rainfall event that occurred in Beijing on July 16, 2018 is used as a case study to examine the method effectiveness. The results show that the precipitation peaked in the morning and evening rush hours, during which the number of congested roads exceeded the average level by up to 150%. The anomaly detection analysis indicates that the number of congested roads and the traffic congestion index reached the outlier level. The anomaly analysis of human activity responses shows that the heavy rainfall event also caused an abnormal increase in the number of location requests, and the spatial distribution of the anomalous grids was highly correlated with the rainfall distribution one hour before. The above results not only prove the effectiveness of the big data and the anomaly detection method in understanding the impact of heavy rainfall events on urban traffic and population, but also provide new means for urban emergency response and management against rainfall disasters.

Keywords: heavy rainfall; urban flood; road traffic; human activity; big data; anomaly detection