ENVIRONMENTAL RESEARCH
INFRASTRUCTURE AND SUSTAINABILITY

PURPOSE-LED PUBLISHING™

Part of the ENVIRONMENTAL RESEARCH SERIES

PAPER • OPEN ACCESS

Analysis of urban flooding in Chicago based on crowdsourced data: drivers and the need for community-based mitigation strategies

To cite this article: Jangho Lee *et al* 2025 *Environ. Res.: Infrastruct. Sustain.* **5** 025008

IOP Publishing Environ. Res.: Infrastruct. Sustain. 5 (2025) 025008 https://doi.org/10.1088/2634-4505/add4eb

ENVIRONMENTAL RESEARCH
INFRASTRUCTURE AND SUSTAINABILITY

RECEIVED
9 January 2025

REVISED
21 April 2025

ACCEPTED FOR PUBLICATION
2 May 2025

PUBLISHED
14 May 2025

PAPER

Analysis of urban flooding in Chicago based on crowdsourced data: drivers and the need for community-based mitigation strategies

Jangho Lee[1,*], **Max Berkelhammer**[1], **Sun Young Park**[2] and **Matthew D Wilson**[3]

[1] Earth and Environmental Sciences, University of Illinois Chicago, Chicago, IL, United States of America
[2] Civil and Environmental Engineering, University of Illinois Urbana-Champaign, Champaign, IL, United States of America
[3] Great Cities Institute, University of Illinois Chicago, Chicago, IL, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: jholee@uic.edu

Keywords: urban, flood, crowdsourcing, precipitation, extreme events, community-based solution

Abstract

This study examines 311 service requests for basement and street flooding in Chicago, focusing on the neighborhoods of Humboldt Park and Chatham. While precipitation is the primary driver citywide, neighborhood-specific factors significantly influence reporting behaviors. In Humboldt Park, flooding service request is driven largely by precipitation, with infrastructure and building conditions contributing to moderate flooding. In Chatham, socioeconomic factors such as homeownership rates and ethnicity, play a critical role, amplifying reporting rates even when physical flood conditions are similar to other areas. Especially, for extreme street flooding service requests, impervious cover, homeownership, and building area have a more pronounced impact in Chatham than in Humboldt Park, reflecting the interplay of physical and social factors. These findings highlight the importance of tailored flood management strategies that integrate infrastructure improvements with community engagement to promote equitable resource allocation and resilience planning.

1. Introduction

The impact of global warming is increasingly evident through compounded risks from rising temperatures, more frequent and intense precipitation, and sea-level rise. Among these factors, flooding poses one of the most severe threats to both human life and infrastructure. Flooding is the leading cause of natural disaster-related deaths globally, contributing to 6.8 million fatalities in the 20th century [1, 2]. In the United States, approximately 100 lives are directly lost to flooding annually and significant economic losses are incurred [3, 4]. Urban areas are particularly vulnerable to flooding, due to their impermeable surfaces—roads, pavements, and buildings—that increase surface runoff during heavy rainfall events [5, 6].

The city of Chicago is highly susceptible to flooding due to heavy storms, dense built areas as well as it being a coastal city (Lake Michigan) [7, 8]. In 2023, a storm in Chicago caused damage to over 60 000 homes, with economic losses exceeding $300 million. Over five years, urban flooding in Chicago has caused more than $773 million in damages though this is likely significantly underestimated due to household basement damage and indirect cost being underreported [9].

Urban flooding in Chicago occurs primarily in the form of basement flooding and street flooding, which are induced by precipitation as opposed to riverbanks and levees being breached. This is partly because of how the sewer system is configured in Chicago. Chicago utilizes a combined sewer system, which collects stormwater runoff and wastewater from homes and businesses into the same pipes rather than using separate systems [10, 11]. The combined sewer system routes water flows to a limited number of large wastewater treatment plants, which can become overwhelmed during extreme weather events, further exacerbating flood risks [12, 13]. During heavy rainfall, this can lead to overflows that release untreated water into nearby rivers or Lake Michigan.

Basement flooding occurs when water penetrates through foundation walls, enters buildings due to the failure of drainage systems, or results from backup of the sewer system into basements. This issue is particularly pronounced in areas with older infrastructure, where outdated systems are unable to handle large volumes of water from storms. Additionally, rising groundwater levels contribute to this issue, especially during periods of heavy rainfall when the soil becomes saturated, creating more pressure against building foundations [14–19]. Especially, Lake Michigan had a historically high-water level in 2020, but it has returned to average conditions since then. Basement flooding can lead to significant damage to personal property, structural integrity issues, and long-term health risks associated with mold and stress/anxiety [20].

Street flooding occurs when stormwater drainage systems are overwhelmed, causing water to accumulate on roads, sidewalks, and public spaces. This flooding is common in areas with high levels of impervious surfaces, like asphalt, which prevent water absorption as well as compacted soils that are common in urban areas [21, 22]. Poorly maintained drainage systems compound this problem, as they fail to remove the excess water quickly enough during storm events [23]. Street flooding not only disrupts transportation by making roads impassable but can also damage public infrastructure such as roads, sidewalks, and sewer systems [24]. Moreover, pooling water poses a public health risk of contaminating public water supplies, particularly if it mixes with untreated sewage or other hazardous substances.

Both basement and street flooding are often worsened by aging infrastructure and socioeconomic disparities [25–27]. In Chicago, neighborhoods with lower socioeconomic status typically contend with poorly maintained sewer systems and drainage infrastructure, which makes these communities particularly vulnerable to the impacts of flooding [28]. This combination of factors highlights the unequal burden of urban flooding, where marginalized communities are disproportionately affected due to a lack of investment in flood mitigation infrastructure [29], raising significant environmental justice concerns.

Despite the need for flood detection and prediction, there are limited resources to detect floods directly and accurately as current models and resources face limitations. Hydrological models, such as National Weather Service (NWS) model [30] or the storm water management model [31, 32] are commonly used to simulate surface runoff, infiltration, and drainage performance based on precipitation data or scenarios. While effective at modeling water flow and assessing stormwater management strategies like catchment or drainage placements, hydrological models do not capture all types of urban flooding, particularly basement flooding. Furthermore, these models often operate at a broad geographic scale and may overlook the fine-scale variability in flood risk that can occur at the neighborhood or even street level. Factors such as the percentage of impervious surfaces, the age and quality of infrastructure, and socioeconomic disparities across neighborhoods can drastically influence flood risks, but these are underrepresented or simplified in traditional hydrological models.

Satellite-based methods, such as multispectral optical imagery and synthetic aperture radar, are another widely used approach for large-scale flood monitoring [33–35]. These methods provide valuable data for monitoring surface water and can quickly map flood extent over large areas. However, this approach has some critical limitations. For example, during heavy rain events—precisely when flood data is most needed—cloud cover can obscure satellite retrievals that are based on surface reflectance. In addition, the long return period of some satellite systems—which can be about 10 to 14 d—limits the continuous monitoring of high frequency transient dynamics associated with flooding. Because of this, satellite-based observations often miss the peak effects of severe flooding events, when they are most needed. Additionally, satellite measurements (as with hydrological models) cannot detect water inside buildings, nor can they estimate the depth of surface water. This means that while they are capable of detecting the presence of standing water, they cannot estimate the severity of the flooding event.

Given the shortcomings of both hydrological models and satellite-based methods, crowdsourced data [18, 36–38], such as citizen-reported 311 service requests [23, 39–41], presents a promising alternative for flood detection and assessment. In cities like Chicago, residents report both basement and street flooding in real-time, offering granular, location-specific information that traditional models and satellite imagery often miss. This community-generated data provides insight not only into where and when flooding occurs but also into how people experience it, capturing the human dimension of flooding in ways that automated sensors or models cannot. While 311 data do not directly measure flood depth or water spread, they serve as a valuable proxy for understanding urban flood dynamics [23, 39–41], especially for events like basement flooding that traditional detection methods fail to capture. By integrating environmental and socioeconomic factors into the analysis of 311 service requests, researchers can identify patterns that might otherwise go unnoticed—such as the vulnerability of older neighborhoods with aging infrastructure to basement flooding or the influence of socioeconomic disparities on the likelihood of encountering and reporting flood incidents. The real-time, localized nature of 311 data highlights areas where traditional models and large-scale monitoring may underestimate risks, offering a more equitable understanding of flood vulnerability.

Previous studies have explored the use of crowdsourced data and 311 service requests to analyze urban flooding [23, 39, 42–44]. For instance, Agonafir *et al* [23] utilized a Random Forest model to assess street flooding in New York, examining the connections between precipitation and socioeconomic factors. Kelleher and McPhillips [43] analyzed 311 service request data to investigate the role of topography in flood events, while Agonafir *et al* [39] focused on the impact of infrastructure on street flooding. While these studies offer valuable insights into urban flood patterns, they primarily focus on street flooding and often employ zip-code-level analysis, which can obscure relationships between environmental, infrastructural, and socioeconomic factors that emerge at finer scales.

It is also important to acknowledge the limitations of 311 data as a form of crowdsourced or citizen-generated information. The data can be influenced by the motivations and capabilities of individuals to report or not report the incidents, leading to biases in its representation of actual flooding conditions. Previous research has shown that crowdsourced data may exhibit biases, as a small segment of the population often contributes a disproportionate share of the reports [45, 46]. Studies have further explored the role of demographic differences—such as educational and economic status, as well as cultural variations linked to race and ethnicity—in shaping participation rates and motivations for reporting [47–49]. For example, higher-income residents may be more familiar with municipal reporting systems or more confident that their complaints will elicit a response, leading to higher reporting rates in wealthier neighborhoods. In contrast, lower-income communities may underreport issues due to limited awareness of 311 systems, language barriers, or skepticism about the responsiveness of local government services.

Demographic differences, such as educational attainment, also play a role in shaping participation rates and motivations for reporting. For example, individuals with higher levels of education may be more likely to recognize flooding as a reportable issue or understand how to navigate bureaucratic systems to submit a complaint. Additionally, cultural variations linked to race and ethnicity can influence reporting behaviors. For instance, immigrant communities or communities of color might report flooding at lower rates due to historical distrust of government institutions or fears of unintended consequences, such as scrutiny from municipal authorities. These findings underscore the need to consider such biases when analyzing and interpreting 311 data.

In this study, we analyze both basement and street flooding reports from 311 service requests to provide a localized understanding of flooding impacts in Chicago and its community areas. We perform a census tract-level analysis to examine environmental and socioeconomic factors in greater detail, capturing neighborhood-specific dynamics within the city. By integrating multiple data sources—including precipitation records, canopy cover, building footprints, and socioeconomic data—we explore the complex interactions driving flooding patterns in Chicago's diverse community regions. Focusing on Chicago, where 311 data have not been extensively studied, allows us to offer new insights into the unique challenges of inland urban flooding and its localized impacts. By combining community-reported data with detailed environmental and socioeconomic analyses, this study mitigates potential biases in crowdsourced data and highlights localized flooding drivers. Our findings aim to support targeted mitigation strategies and equitable policies to address the disproportionate impacts of flooding on marginalized communities in Chicago.

2. Data and method of analysis

2.1. 311 requests and precipitation data

We collect 311 service requests data from the Chicago data portal [50]. Each service request includes a request content, the date and time of the request, and the location in latitude-longitude coordinates. For basement and street flooding, we extract complaints related to 'water in basement' and 'water on street' from March to October for the years 2019–2023, as this period (excluding winter) accounts for over 95% of all flooding-related complaints. We then aggregate the data into daily temporal data and spatially at the census tract level. Finally, we normalize the service requests from each census tract to service requests per 1000 population. The spatial distribution and the time series of the service requests are shown in figure 1.

For precipitation measurements, we utilize the Multi-Radar/Multi-Sensor precipitation dataset, which offers a high spatial resolution of 1 km. To align this data with the temporal and spatial scales of our study, we calculate daily average precipitation and aggregate it into daily temporal data. Spatially, the data is aggregated at the census tract level. To achieve this, the original 1 km resolution data is first linearly interpolated to a finer resolution of 50 m. The census tract averages are then derived by spatially averaging the interpolated values. This methodology ensures comprehensive data coverage with no missing values at the census tract level.

The average precipitation pattern, peak precipitation pattern, and time series of daily average precipitation are also shown in figure 1. One notable feature is the low average precipitation values observed along the shoreline in the downtown region (figure 1(c)). This anomaly is likely an artifact caused by radar

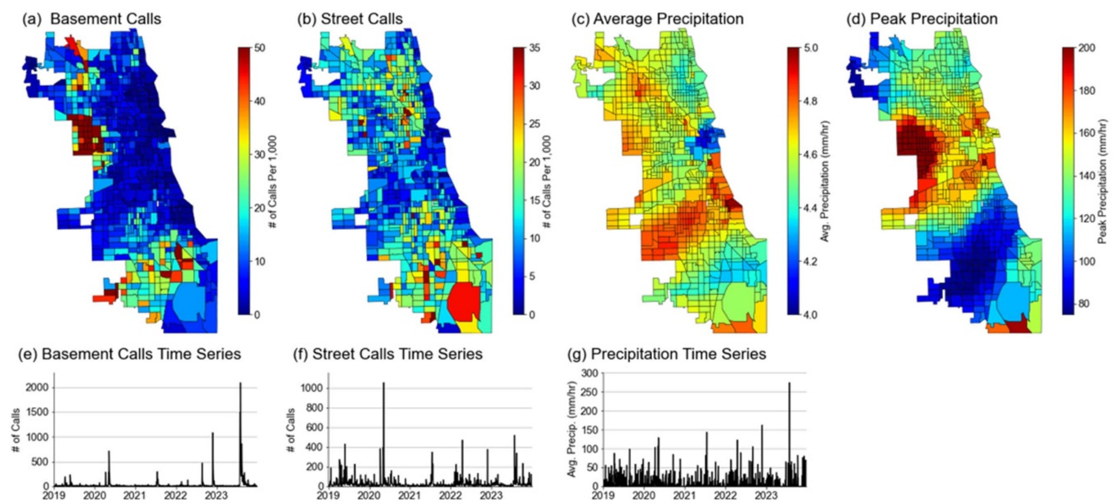

Figure 1. (a) Total number of 311 basement flooding service requests collected between March and October from 2019 to 2023, aggregated at the census tract level. (b) Same as (a), but for street flooding 311 service requests. (c) Average precipitation distribution across Chicago. (d) Peak precipitation values for each census tract during the study period. (e)–(g) Daily time series of total basement flooding service requests, total street flooding service requests, and area-averaged precipitation over the entire study period.

signal interference from skyscrapers. However, this does not affect the overall results of this study, as this region has negligible basement or street flooding-related 311 calls (normalized per 1000 population).

At this scale, street flooding service requests exhibit a strong temporal correlation with precipitation ($R = 0.74$), indicating a direct and predictable relationship between rainfall intensity and surface runoff. In contrast, basement flooding service requests show a more moderate correlation ($R = 0.37$), suggesting that these incidents are influenced by more complex and unevenly distributed factors. With a spatial smoothing (alpha = 0.5), the spatial correlation coefficient between precipitation and basement flooding service requests was 0.1 while the correlation with the street flooding service request was 0.2.

While the 311 reports span 2019–2023, canopy cover and impervious surface data (2021), and socioeconomic data (2020–2022) are assumed to be largely time-invariant. These variables typically change slowly over short time frames, allowing us to combine them into a single analytic framework without significant temporal misalignment. Precipitation, in contrast, is explicitly treated as time-varying.

2.2. Humboldt Park and Chatham

The low correlation coefficient between the spatial patterns of precipitation and flooding-related service requests underscores the uneven distribution of flood vulnerability across Chicago. Even when taking local spatial context into account, adjacent census tracts often exhibit differing responses to similar rainfall patterns. This study highlights the importance of a community-centered approach to analyzing flooding service requests by focusing on two Chicago neighborhoods—Humboldt Park and Chatham. These areas illustrate how the drivers of such requests can vary significantly, influenced by factors like socioeconomic disparities, high frequencies of flooding-related service requests, and elevated precipitation levels. Figure 2(a) provides a map of Chicago with census tract boundaries (defining the study's spatial scale), while figure 2(b) depicts the time series of precipitation and highlights extreme surges in basement and street flooding service requests.

Chicago exhibits a distinct pattern of racial and ethnic segregation—a factor incorporated into our analysis by examining the spatial distributions of White (non-Hispanic or Latino), Hispanic/Latino, and Black (non-Hispanic or Latino) populations (figures 2(c) and (e)), based on data from the 2020 Decennial Census. We use the White (non-Hispanic or Latino) population as a reference group, while considering Hispanic/Latino and Black (non-Hispanic or Latino) communities as minority populations. This approach enables us to investigate how different ethnic groups may influence the reporting or experience of flooding events.

Humboldt Park, a community area in Chicago, has been at the center of discussions about the gentrification of its historically Puerto Rican neighborhood (Hispanic and Latino populations), pointing to a dynamic interplay of cultural shifts, economic development, and displacement pressures [51–53]. As shown in figure 2(c), Humboldt Park exhibits a pronounced east-to-west ethnic gradient, as well as higher

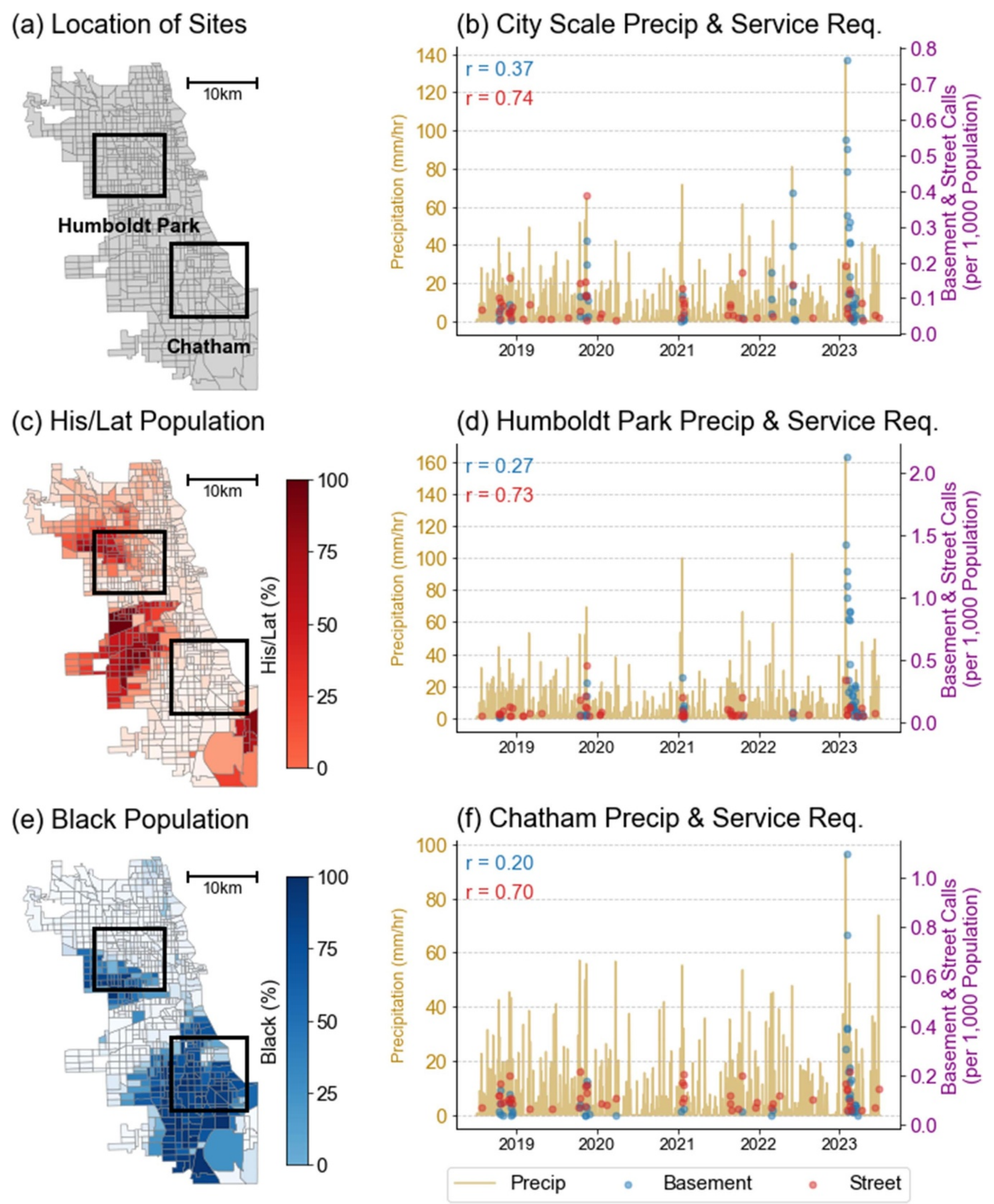

Figure 2. (a) Map of Chicago with census tract boundaries, highlighting the two community areas studied: Humboldt Park and Chatham. (b) Time series of the average precipitation rate (yellow lines) and citywide numbers of basement (blue dots) and street (red dots) flooding service requests per 1000 population. Service requests are plotted only for events within the top 5% of total data. (c) Distribution of Hispanic and Latino populations in Chicago, emphasizing the ethnic characteristics of Humboldt Park. Red shading indicates census tracts dominated by Hispanic and Latino populations. (d) Time series of average precipitation rates and numbers of basement and street flooding service requests, similar to panel (b), but averaged specifically for Humboldt Park. (e) Distribution of Black populations in Chicago, emphasizing the ethnic characteristics of Chatham. (f) Time series of average precipitation rates and numbers of basement and street flooding service requests, similar to panel (b), but averaged specifically for Chatham.

precipitation levels and a greater frequency of 311 service requests when compared to citywide averages (figure 2(d)).

The second community area of focus is Chatham (and its surrounding region) in Chicago. Chatham is predominantly a Black community [54, 55], with a small fraction of White residents in its northeast part (figure 2(e)). Although the overall rate of 311 service requests in Chatham is not as high as in Humboldt Park, it still exceeds the citywide average (figure 2(f)), particularly during severe storm events such as the one recorded in 2023.

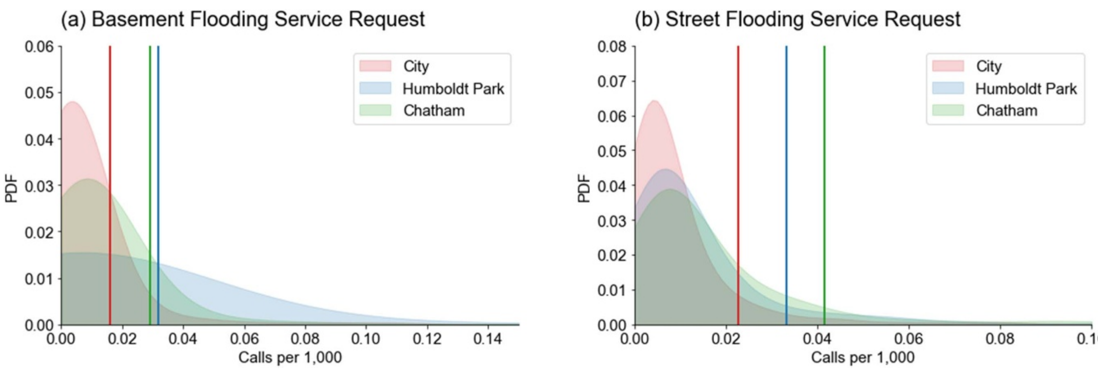

Figure 3. Kernel density estimates (KDE) of basement (panel (a)) and street (panel (b)) 311 service requests (per 1000 population) for the entire city (red) Humboldt Park (blue), and Chatham (green). Vertical lines mark the 90th (extreme) percentile thresholds of non-zero daily service requests in each region.

2.3. Moderate and extreme 311 service requests

Due to the highly skewed distribution of 311 service requests, as illustrated in figures 2(b), (d) and (f), we establish moderate and extreme service request categories based on the 75th and 90th percentile thresholds of non-zero daily service requests in each region (citywide, Humboldt Park, and Chatham). We use these thresholds to isolate the upper-tail of flooding reports, ensuring that more severe or uncommon events are captured separately from lower-frequency or minor incidents. The selection process involved examining the distribution of daily service request counts and identifying percentile cutoffs for each study region; service requests falling between the 75th and 90th percentiles are classified as moderate, while those exceeding the 90th percentile is considered extreme. Figure 3 shows the kernel density estimates (KDE) of basement and street service requests for each region, along with their 90th percentile thresholds.

2.4. Land cover and topography data

For our study region, we source one-meter resolution canopy cover and impervious land cover data from the NOAA coastal change analysis program (C-CAP) for the year 2021. This dataset includes three binary layers: canopy cover, impervious land cover, and water bodies. We calculate the percentage of canopy cover and impervious land cover for each census tracts, representing these as percent canopy cover and impervious cover (figures 4(a) and (b)).

Additionally, the digital elevation model dataset is obtained from the US Geological Survey's 1/3 arc-second (approximately 10 m) map [56, 57] (figure 4(c)). However, for our analysis, the relative elevation of each census tract compared to its surrounding census tracts is more important than absolute elevation, as it can represent the flow of water during flooding events. To capture this, we calculate the relative elevation by subtracting the mean elevation of the surrounding census tracts (figure not shown).

To align these raster datasets with census-tract boundaries, we applied a zonal statistics approach. Specifically, all pixel values within each census tract polygon were averaged, thereby converting high-resolution imagery into tract-level canopy and impervious cover metrics. This step harmonizes the resolution of raster inputs (1 m or 1 km) with the census-tract aggregation of 311 flooding data.

2.5. Socioeconomic and infrastructure data

We also consider home ownership rate and per capita income per census tract, sourced from the 2018–2022 American Community Survey of the U.S. Census Bureau (figures 4(d) and (e)). For the infrastructure analysis, we examine building area (figure 4(f)) and building age (figure 4(g)), both sourced from the Chicago Data Portal's building footprint data [58]. This dataset includes the geometry and construction date of each building. From the geometry, we calculate the percentage of land area covered by buildings within each census tract. Additionally, we compute the average age of buildings in each census tract.

We note a relatively high correlation between White population and income, especially in Humboldt Park ($r = 0.9$). This likely reflects broader demographic and socioeconomic patterns rather than a strictly one-dimensional relationship. Consequently, caution is warranted when interpreting results for these two variables individually, since racial composition and wealth are often intertwined in complex ways shaped by historic and ongoing segregation, housing policies, and economic disparities. Rather than treating either

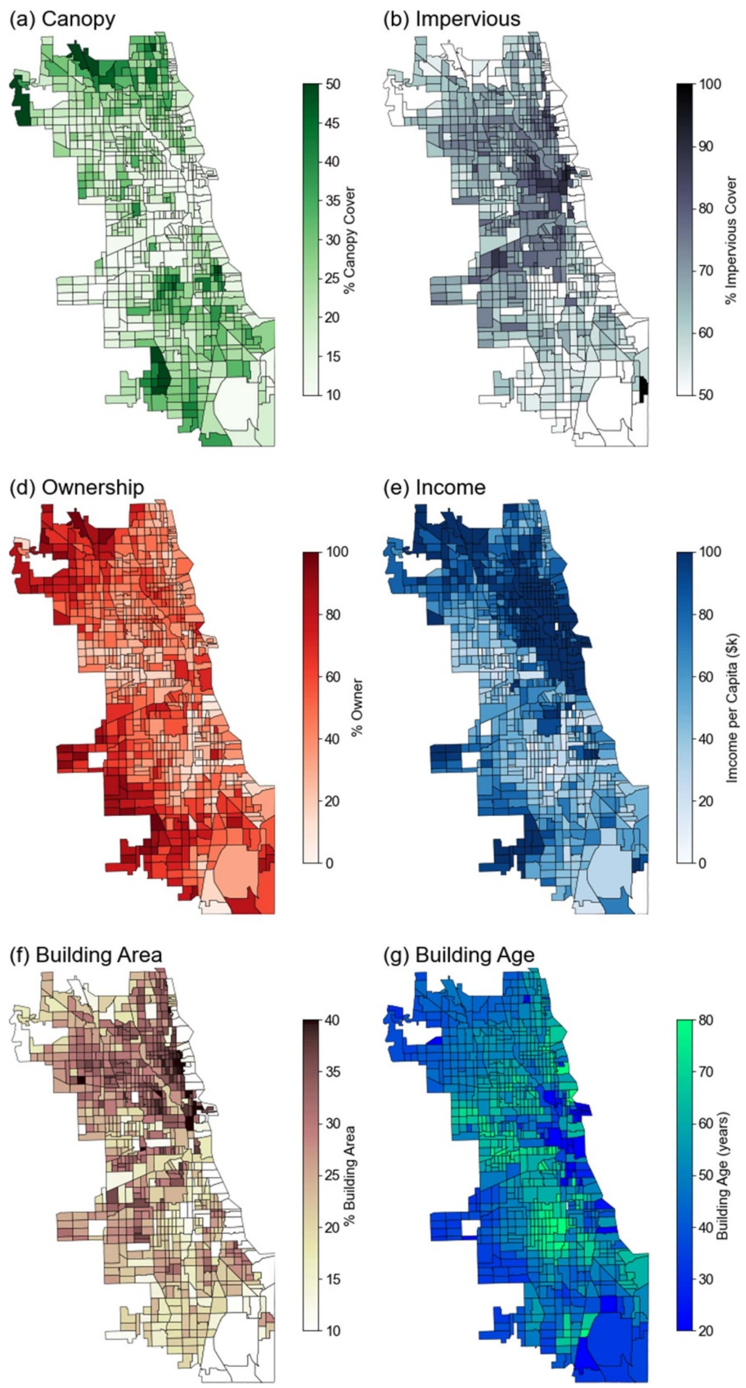

Figure 4. Spatial distribution of key environmental, socioeconomical, and infrastructure variables across Chicago's census tracts. (a) Percent canopy cover. (b) Percent impervious surface. (c) Elevation map of the city. (d) Homeownership rate by census tract. (e) Per capita income (in thousands of dollars). (f) Percent building area, indicating the extent of land covered by buildings. (g) Average building age in years.

variable in isolation, we view them as overlapping proxies for social, economic, and cultural contexts that collectively influence flood vulnerability and reporting behavior.

2.6. Generalized additive models

To assess the impact of predictor variables on basement and street flooding service requests, we utilize generalized additive models (GAM) regression [59]. GAM is particularly useful because it allows for flexible, non-linear relationships between the predictor variables and the response variable. This flexibility is crucial in our study, as many environmental and socioeconomic factors affecting flooding may not have strictly linear relationships with flooding outcomes. GAM can capture these complex interactions by fitting smooth functions to each predictor, enabling more accurate modeling of flood reports dynamics. Additionally, GAM

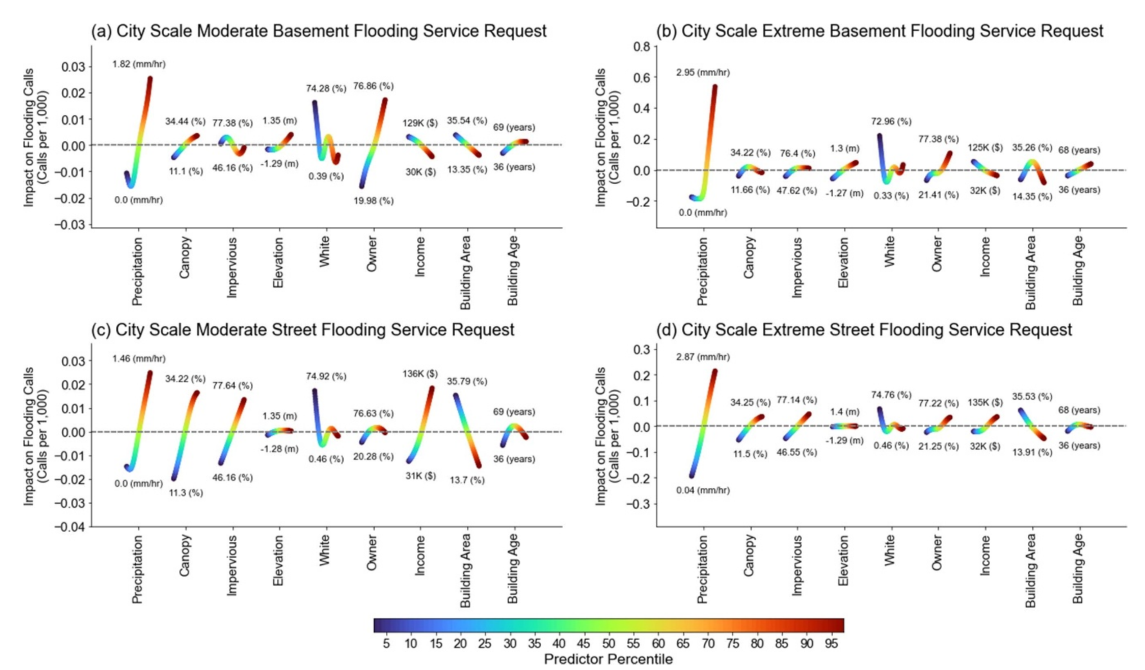

Figure 5. (a) The impact of key environmental and socioeconomic variables on moderate (75th–90th percentile) basement flooding service requests at the city scale. The blue part of each line denotes the lower end of the predictor variable, while the red part shows the higher end. Numbers on the top and the bottom of each plot shows the range of each predictor variables. (b)–(d) same as (a), but for extreme basement flooding service request, moderate street flooding service request, and extreme street flooding service request, respectively.

provides interpretability, allowing us to visualize the individual effects of each predictor variable on the likelihood of flooding service requests while controlling for other factors. We set up our gam model as equation (1).

$$311\ \mathrm{Requests}_{b,t} = \mathrm{GAM}\left(\mathrm{s}\left(\mathrm{precipitation}_{b,t}\right) + \mathrm{s}\left(\mathrm{white}_b\right) + \mathrm{s}\left(\mathrm{canopy}_b\right) + \mathrm{s}\left(\mathrm{impervious}_b\right) + \mathrm{s}\left(\mathrm{elevation}_b\right) + \mathrm{s}\left(\mathrm{ownership}_b\right) + \mathrm{s}\left(\mathrm{income}_b\right) + \mathrm{s}\left(\mathrm{building\ area}_b\right) + \mathrm{s}\left(\mathrm{building\ age}_b\right)\right). \quad (1)$$

The GAM model is designed to estimate the number of 311 flooding-related service requests per 1000 population (either basement or street flooding) in census tract b at time t based on several predictor variables. The model incorporates a smooth function $s()$ which consists of 8 cubic splines, each with its own smoothing parameter calculated using restricted maximum likelihood (REML). We have tested multiple number of splines, but with the smoothing parameter calculated with REML, the resulting smooth curve was nearly identical whichever number of splines used.

The predictors include both time-varying and time-invariant variables. Precipitation is the only time-varying variable in the model, reflecting daily changes in rainfall across block groups, while the remaining variables, such as canopy cover, impervious surfaces, population demographics, and infrastructure, are time-invariant. By this, the GAM model allows us to effectively analyze the complex and potentially non-linear relationships between flooding-related 311 service requests and a variety of environmental, socioeconomic, and infrastructural factors.

3. Results

3.1. City-wide analysis

Analyzing the moderate basement flooding service request in city scale (figure 5(a)), precipitation emerges as the primary environmental driver. Furthermore, the rate of homeownership shows a positive association, suggesting that homeowners may be more motivated to report basement flooding—due to stronger incentives to protect property investments or a greater familiarity with municipal reporting channels. The subtle negative correlation with White (non-Hispanic or Latino) population, although not straightforward, may reflect differences in neighborhood reporting behaviors, infrastructure quality, or resource allocation. For the extreme basement flooding service request (figure 5(b)), precipitation is the dominant driver, showing strong influence citywide.

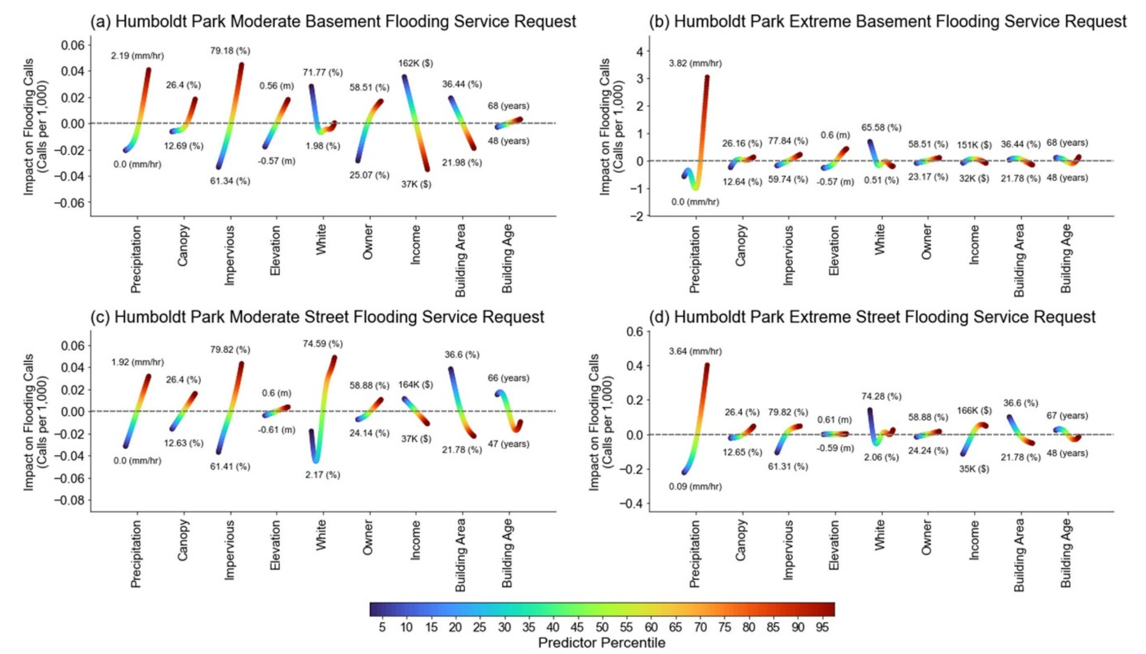

Figure 6. Same as figure 5, but for Humboldt Park.

We also analyze the factors impacting moderate street flooding service request (figure 5(c)). At the city scale, precipitation remains a key factor, but canopy cover and impervious surfaces gain prominence. Dense tree canopies may direct rainfall and introduce leaf debris into gutters, potentially obstructing drains, while expansive impervious surfaces prevent infiltration and increase ponding. By contrast, areas with relatively lower canopy and impervious cover have more grassy or open spaces available to absorb runoff, reducing standing water and the likelihood of residents filing 311 service requests. Furthermore, it is worth noting that much of the urban tree canopy is managed, often situated near paved surfaces or tree pits with limited infiltration potential. In such cases, the presence of canopy does not necessarily imply substantial pervious area. Socioeconomic and demographic variables—such as percent White (non-Hispanic or Latino) population and income—further modulate number of street flooding service requests, potentially influencing how communities perceive, manage, and report moderate street flooding. Larger building footprints appear to be associated with fewer service requests, possibly reflecting neighborhoods with improved stormwater controls or proactive property management. For the extreme street flooding service request (figure 5(d)), the primary driver is precipitation, similar to the extreme basement flooding service request.

3.2. Local analysis in Humboldt Park

Examining the moderate basement flooding service request in Humboldt Park (figure 6(a)), the influence of multiple environmental and socioeconomic variables—ranging from precipitation to demographic and infrastructural indicators—suggests a complex interplay in which no single factor dominates. In such a diverse setting, moderate basement flooding may arise from the cumulative effect of several different factors, from slight imperfections in building maintenance to subtle socioeconomic factors that influence whether residents choose to report flooding. The driver of the extreme basement flooding service request mirrors the pattern from the city-wide scale, which is dominated by the precipitation pattern.

Moderate street flooding service request presents a multifaceted context (figure 6(c)), where precipitation, land cover, and socioeconomic variables all impacts the number of service requests. Moderate street flooding service request in this area likely results from an interplay of different influences, including how local runoff is managed, how well drains are maintained, and how readily communities engage with 311 services. Economic capacity, historic experiences, and trust in municipal responsiveness may collectively shape reporting behaviors, making moderate flooding in Humboldt Park a product of overlapping environmental and social dimensions. Similar with the extreme basement flooding request, precipitation is the principal driver for the extreme street flooding request (figure 6(d)).

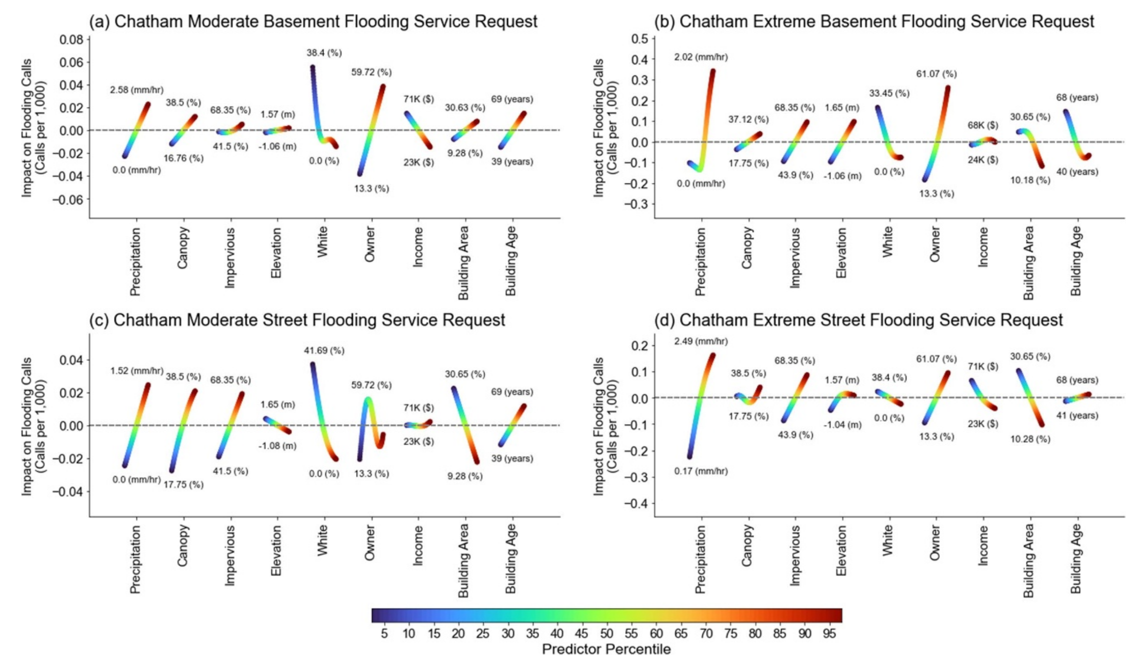

Figure 7. Same as figure 5, but for Chatham.

3.3. Local analysis in Chatham

Looking at the moderate basement flooding service request in Chatham, Chatham more closely mirrors the citywide pattern, but with a heightened emphasis on demographic factors such as percent White (non-Hispanic or Latino) population and homeownership rate (figure 7(a)). This pattern implies that in some neighborhoods, the decision to report flooding to 311 may be just as critical to understanding local flood dynamics as the physical factors causing the event itself. More importantly, Chatham presents an important characteristic in drivers of extreme basement service requests, by having impact of homeownership rate comparable to that of precipitation (figure 7(b)). While extreme rainfall sets the background condition, the community's reporting behavior appears crucial in translating physical floods into reported incidents. High homeownership may correlate with a greater tendency to seek immediate assistance, document damages, or trust that municipal intervention could mitigate losses. These factors can inflate the volume of 311 service requests.

In terms of moderate street flooding service requests, drivers in Chatham are similar to citywide trends with additional emphasis on demographic and socioeconomic characteristics (figure 7(c)). While physical conditions such as canopy and impervious cover still influence flood occurrences, reporting patterns here appear closely tied to community composition and resource management. Factors like income and demographic makeup may determine the propensity to report moderate flooding, either by fostering greater engagement with city services or by motivating more active infrastructure maintenance. The driver of extreme street flooding service in Chatham shows another important distinction (figure 7(d)). Impervious cover, homeownership rate, and building area collectively are comparable to the predictive power of precipitation. High impervious coverage amplifies surface runoff, intensifying the effects of extreme precipitation. Meanwhile, a high homeownership rate may encourage more frequent reporting, enhancing the visibility of flooding events. Areas with greater building coverage may experience different drainage dynamics, as large structures can alter local runoff patterns. Depending on neighborhood-level characteristics, building footprints might influence how rainwater flows, either by directing it away from streets or leaving certain areas more prone to water accumulation.

4. Discussions

4.1. Drivers of flood service request in Humboldt Park and Chatham

This study demonstrates that precipitation is a major and consistent driver of 311 service requests for both basement and street flooding across Chicago, yet the influence of local conditions—both physical and social—varies substantially among neighborhoods. City-wide analyses show that heavy rainfall can cause surges in reported flooding incidents. Nonetheless, factors such as land cover (impervious surfaces, canopy) and socioeconomic attributes (race, income, homeownership) significantly modulate how flooding is both experienced and reported, particularly when storms are less intense. These secondary influences help explain why some neighborhoods generate more 311 calls than others.

In comparing the neighborhoods of Humboldt Park and Chatham, we found that the drivers of 311 service request calls can be significantly different between the community areas. In Humboldt Park, precipitation remains the most influential predictor of 311 calls, especially during extreme storms. Moderate basement flooding in Humboldt Park arises from an interplay of rainfall, building age, and drainage infrastructures. When rainfall intensifies, the localized complexity often becomes secondary to the sheer volume of water inundating streets and basements. This leads to spikes in both basement and street flooding reports, mirroring broader city-wide trends where rainfall intensity correlates strongly with the volume of flood-related calls. It is important to interpret these results as associations rather than definitive causal links. For instance, higher homeownership may correlate with more frequent 311 calls due to greater financial stakes in property and stronger civic engagement, but it could also reflect confounding factors such as historically better infrastructure or organized neighborhood advocacy.

Although moderate street flooding in Humboldt Park also depends on rainfall, the motivation or need to report such events can be shaped by how well drains are maintained, the degree of trust in municipal agencies, and the cultural or socioeconomic contexts of different census tracts within the community. Infrastructure maintenance lapses—such as clogged drains—can aggravate minor flood events enough to prompt calls, but in severe storms, those local drivers also become overshadowed by the city-wide breakdown of stormwater systems under heavy rainfall.

Chatham, in contrast, shows a different interplay between precipitation and socioeconomic variables, especially for the extreme cases. In this area, the volume of basement-flooding reports is shaped almost equally by rainfall and social factors. High homeownership rates, for instance, can incentivize prompt reporting if residents believe filing a 311 request might lead to assistance or mitigation. In neighborhoods where residents feel confident that reporting triggers a municipal response, or have the resources to follow up on damages, the number of calls can outpace what one might expect solely from the physical severity of flooding. This phenomenon may also produce higher reported flood frequencies relative to neighborhoods facing comparable rainfall but exhibiting lower trust in government or fewer economic resources.

Street flooding in Chatham reflects a complex interplay between physical and social factors. On one hand, impervious surface coverage and larger building footprints exacerbate runoff during intense storms, mirroring citywide patterns. However, social dimensions—such as the proportion of White residents in a census tract—also play a role, influencing how communities engage with municipal services and utilize homeowner resources. These factors can contribute to higher reporting rates, even when physical flood conditions may not differ substantially from other areas. For the drivers of extreme street flooding service request, the combined influence of impervious cover, homeownership rates, and building area is notably more pronounced in Chatham compared to Humboldt Park. This highlights the ways in which physical and social factors converge to shape flooding experiences and reporting behaviors. Such differences underscore that crowdsourced data such as 311 service requests reflect not only the severity of physical flood events but also the broader social and infrastructural contexts in which these events occur.

Taken together, Humboldt Park and Chatham highlight the multifaceted dynamics underlying flood-related 311 service requests. Although precipitation stands out as the principal catalyst, the local context—ranging from drainage infrastructure conditions to socioeconomic attributes—can magnify or mitigate reporting behaviors. In Humboldt Park, moderate flooding emerges from a nuanced set of interacting factors, but extreme storms almost invariably trigger large increases in calls. In Chatham, meanwhile, demographic composition and economic capacity can be nearly as decisive as rainfall in explaining extreme basement and street flooding reports. These differences underscore why flood management efforts should not treat all neighborhoods as uniform. Where systemic upgrades to infrastructure might yield the most benefit in some areas, complementary investments in outreach, reporting education, and trust-building may prove equally crucial elsewhere.

4.2. Implications of the study

Such targeted approaches also matter for equitable resource allocation. Where socioeconomic factors strongly influence 311 service requests, high-call neighborhoods may receive more attention, even if physical flooding conditions elsewhere are just as severe or worse but underreported. By cross-referencing 311 data with alternative sources—such as insurance claims, hydraulic modeling, or satellite observations—decision-makers can gain a more holistic understanding of actual flood impacts across the city. This is increasingly pressing as climate change intensifies precipitation events and places further strain on Chicago's infrastructure. Areas already susceptible to heavy rainfall, like Humboldt Park, may see more frequent street and basement flooding in the future, while in communities like Chatham, where socioeconomic factors drive robust reporting behaviors, the gap between reported and unreported floods in other neighborhoods might widen.

In addition, climate change is poised to exacerbate existing inequities. Neighborhoods facing a combination of old infrastructure, lower incomes, or weaker trust in city services will likely struggle to adapt, potentially underreporting serious flood events and losing out on timely assistance or long-term planning support. In the same context, communities that readily report even moderate incidents could appear to bear a disproportionate burden, thus attracting more municipal resources at the expense of systematically underrepresented areas. Policymakers must therefore interpret 311 data by considering underlying social and infrastructural variations.

Ultimately, effective flood resilience in a city as diverse as Chicago demands a detailed, place-based approach that addresses both physical risk factors and the social realities of reporting. Infrastructure expansions like storm sewers or permeable pavements remain vital in areas prone to intense runoff, but they must be paired with community outreach, education on flood risks, and reliable reporting mechanisms that do not disadvantage less-engaged populations. Where homeownership rates or trust in government shape the volume of 311 requests, tailored campaigns to build confidence in municipal services can promote more accurate and equitable data collection. By recognizing that flooding is both a physical and social phenomenon, urban planners, engineers, and community leaders can design interventions that tackle not only the immediate challenges of stormwater management but also the broader societal context in which flood events are experienced, reported, and addressed.

While this study focuses on Chicago—a city with unique attributes such as a combined sewer system and distinctive socio-spatial patterns—many of its challenges are similar in those of older industrialized cities worldwide, as well as rapidly urbanizing regions coping with inadequate stormwater infrastructure [60–64]. Our findings illustrate how combining community-driven 311 data with socioeconomic, infrastructural, and environmental variables can capture localized flooding dynamics that formal hydrologic models or satellite methods may miss. Ultimately, leveraging crowdsourced data with fine-scale environmental and socioeconomic indicators, supports more targeted and equitable flood-resilience measures in urban centers globally.

4.3. Limitations and future directions

Several limitations highlight the need for continued investigation. Although 311 requests offer useful real-time data, they reflect community engagement patterns rather than a direct measure of flood magnitude. Underreporting can mask significant events in neighborhoods with lower trust in government or fewer resources for navigating municipal systems. Future studies might triangulate 311 data with claims, field surveys, or satellite observations to develop more robust risk profiles. Nevertheless, these analyses advance our understanding of how hydrological conditions intersect with socioeconomic factors to shape flood reporting. The distinctions between Humboldt Park and Chatham emphasize that stormwater intensity remains critical, yet local structural vulnerabilities and social dynamics considerably modulate reporting. With climate change intensifying rainfall extremes, a multifaceted approach—encompassing infrastructure upgrades, community engagement, and inclusive data interpretation—will be essential for equitable and effective urban flood resilience.

Data availability statement

311 Service requests data for Chicago can be accessed at Chicago data portal (https://data.cityofchicago.org/widgets/v6vf-nfxy?mobile_redirect=true). Census tract level socioeconomic data can be retrieved from Census data portal (https://data.census.gov/). Building footprints data can be downloaded from Chicago Data Portal (https://data.cityofchicago.org/Buildings/Building-Footprints-current-/hz9b-7nh8). Digital Elevation Map (DEM) used in this study can be accessed at USGS data collection (https://data.usgs.gov/datacatalog/data/USGS:3a81321b-c153-416f-98b7-cc8e5f0e17c3). Canopy and Impervious cover data can be retrieved from C-CAP project data archive from NOAA (www.fisheries.noaa.gov/inport/item/47841).

All data that support the findings of this study are included within the article (and any supplementary information files).

Acknowledgments

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research's Urban Integrated Field Laboratories CROCUS project research activity, under Award Number DE-SC0023226.

Conflict of interest

The authors declare no competing interests.

ORCID iDs

Jangho Lee https://orcid.org/0000-0002-8942-1092
Max Berkelhammer https://orcid.org/0000-0002-8924-716X
Sun Young Park https://orcid.org/0009-0000-4018-1254
Matthew D Wilson https://orcid.org/0000-0002-6569-4950

References

[1] Doocy S, Daniels A, Murray S and Kirsch T D 2013 The human impact of floods: a historical review of events 1980–2009 and systematic literature review *PLoS Curr.* **5** ecurrents-dis
[2] Hamidifar H and Nones M 2021 Global to regional overview of floods fatality: the 1951–2020 period *Nat. Hazards Earth Syst. Sci. Discuss.* **2021** 1–22
[3] Ashley S T and Ashley W S 2008 Flood fatalities in the United States *J. Appl. Meteorol. Climatol.* **47** 805–18
[4] Han Z and Sharif H O 2021 Analysis of flood fatalities in the United States, 1959–2019 *Water* **13** 1871
[5] National Academies of Sciences, Engineering, and Medicine, Division on Earth and Life Studies, Committee on Urban Flooding in the United States, Water Science and Technology Board, Policy and Global Affairs *et al* 2019 *Framing the Challenge of Urban Flooding in the United States* (National Academies Press)
[6] Sohn W, Kim J-H, Li M-H, Brown R D and Jaber F H 2020 How does increasing impervious surfaces affect urban flooding in response to climate variability? *Ecol. Indic.* **118** 106774
[7] Nanía L S, Leon A S and García M H 2015 Hydrologic-hydraulic model for simulating dual drainage and flooding in urban areas: application to a catchment in the metropolitan area of Chicago *J. Hydrol. Eng.* **20** 04014071
[8] Hejazi M I and Markus M 2009 Impacts of urbanization and climate variability on floods in Northeastern Illinois *J. Hydrol. Eng.* **14** 606–16
[9] Festing H, Copp C, Sprague H, Wolf D, Shorofsky B and Nichols K 2014 *The Prevalence and Cost of Urban Flooding: A Case Study of Cook County, IL* (Center for Neighborhood Technology)
[10] Quijano J C, Zhu Z, Morales V, Landry B J and Garcia M H 2017 Three-dimensional model to capture the fate and transport of combined sewer overflow discharges: a case study in the Chicago Area Waterway System *Sci. Total Environ.* **576** 362–73
[11] Tarr J A 1979 The separate vs. combined sewer problem: a case study in urban technology design choice *J. Urban Hist.* **5** 308–39
[12] Brose D A, O'Brien P, Moran-Andrews J, Grabowy J S, Kunetz T E and Podczerwinski E W 2020 Long-term biosolids planning with an operational mega reservoir for combined sewer overflow-impacted stormwater capture *Water Environ. Res.* **92** 5–10
[13] Oberg N, Schmidt A R, Landry B J, Leon A S, Waratuke A R, Mier J M and García M H 2017 Improved understanding of combined sewer systems using the Illinois conveyance analysis program (ICAP) *Urban Water J.* **14** 811–9
[14] Sandink D 2016 Urban flooding and ground-related homes in Canada: an overview *J. Flood Risk Manage.* **9** 208–23
[15] Kaur B, Binns A, Sandink D, Gharabaghi B and McBean E eds 2021 Reducing the risk of basement flooding through building-and lot-scale flood mitigation approaches: performance of foundation drainage systems *Canadian Society of Civil Engineering Annual Conf.* (Springer)
[16] Steis Thorsby J, Miller C J and Treemore-Spears L 2020 The role of green stormwater infrastructure in flood mitigation (Detroit, MI USA)–case study *Urban Water J.* **17** 838–46
[17] Liu W, Chen W and Peng C 2014 Assessing the effectiveness of green infrastructures on urban flooding reduction: a community scale study *Ecol. Modelling* **291** 6–14
[18] Wang R-Q, Mao H, Wang Y, Rae C and Shaw W 2018 Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data *Comput. Geosci.* **111** 139–47
[19] Zellner M, Massey D, Minor E and Gonzalez-Meler M 2016 Exploring the effects of green infrastructure placement on neighborhood-level flooding via spatially explicit simulations *Comput. Environ. Urban Syst.* **59** 116–28
[20] Lane K, Charles-Guzman K, Wheeler K, Abid Z, Graber N and Matte T 2013 Health effects of coastal storms and flooding in urban areas: a review and vulnerability assessment *J. Environ. Public Health* **2013** 913064
[21] Qin Y 2020 Urban flooding mitigation techniques: a systematic review and future studies *Water* **12** 3579
[22] Mignot E, Li X and Dewals B 2019 Experimental modelling of urban flooding: a review *J. Hydrol.* **568** 334–42
[23] Agonafir C, Lakhankar T, Khanbilvardi R, Krakauer N, Radell D and Devineni N 2022 A machine learning approach to evaluate the spatial variability of New York City's 311 street flooding complaints *Comput. Environ. Urban Syst.* **97** 101854
[24] Kermanshah A, Derrible S and Berkelhammer M 2017 Using climate models to estimate urban vulnerability to flash floods *J. Appl. Meteorol. Climatol.* **56** 2637–50
[25] Elliott J R, Loughran K and Brown P L 2023 Divergent residential pathways from flood-prone areas: how neighborhood inequalities are shaping urban climate adaptation *Soc. Prob.* **70** 869–92
[26] Fielding J and Burningham K 2005 Environmental inequality and flood hazard *Local Environ.* **10** 379–95
[27] Sanders B F, Brady D, Schubert J E, Martin E-M H, Davis S J and Mach K J 2024 Quantifying social inequalities in flood risk *Multidiscip. J. Civ. Eng.* **2** 04024004
[28] Sinha D 2023 *Urban Stormwater Flooding: Inequal Stress and Associated Vulnerabilities in Chicago, IL, USA* (Indiana State University)
[29] Tate E, Rahman M A, Emrich C T and Sampson C C 2021 Flood exposure and social vulnerability in the United States *Nat. Hazards* **106** 435–57
[30] Adams T III 2016 Flood forecasting in the United States NOAA/national weather service *Flood Forecasting* (CRC Press) pp 249–310
[31] Huber W C, Rossman L A and Dickinson R E 2005 EPA storm water management model, SWMM5 *Watershed Models* **338** 359
[32] Niazi M, Nietch C, Maghrebi M, Jackson N, Bennett B R, Tryby M and Massoudieh A 2017 Storm water management model: performance review and gap analysis *J. Sustain. Water Built Environ.* **3** 04017002
[33] Matgen P, Hostache R, Schumann G, Pfister L, Hoffmann L and Savenije H 2011 Towards an automated SAR-based flood monitoring system: lessons learned from two case studies *Phys. Chem. Earth A/B/C* **36** 241–52

[34] Kussul N, Shelestov A and Skakun S (eds) 2011 Flood Monitoring from SAR Data *Use of Satellite and In-situ Data to Improve Sustainability* (Springer)
[35] Refice A, Capolongo D, Pasquariello G, DaAddabbo A, Bovenga F, Nutricato R, Lovergine F P and Pietranera L 2014 SAR and InSAR for flood monitoring: examples with COSMO-SkyMed data *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **7** 2711–22
[36] Dede M, Widiawaty M A, Pramulatsih G P, Ismail A, Ati A and Murtianto H 2019 Integration of participatory mapping, crowdsourcing and geographic information system in flood disaster management (case study Ciledug Lor, Cirebon) *J. Inf. Technol. Util.* **2** 44–47
[37] Helmrich A M *et al* 2021 Opportunities for crowdsourcing in urban flood monitoring *Environ. Modelling Softw.* **143** 105124
[38] Sadler J, Goodall J, Morsy M and Spencer K 2018 Modeling urban coastal flood severity from crowd-sourced flood reports using Poisson regression and Random Forest *J. Hydrol.* **559** 43–55
[39] Agonafir C, Pabon A R, Lakhankar T, Khanbilvardi R and Devineni N 2022 Understanding New York City street flooding through 311 complaints *J. Hydrol.* **605** 127300
[40] Eugene A, Alpert N, Lieberman-Cribbin W and Taioli E 2022 Using NYC 311 call center data to assess short-and long-term needs following Hurricane Sandy *Disaster Med. Public Health Prep.* **16** 1447–51
[41] Schellong A and Langenberg T eds 2007 Managing citizen relationships in disasters: hurricane wilma, 311 and miami-dade county *2007 40th Annual Hawaii Int. Conf. on System Sciences (HICSS'07)*(IEEE)
[42] Chen Z, Wang R, Hong Y, Zhou Y, He Y and Gronewold A D 2024 The influence of road network topology on street flooding in New York City—a social media data approach *J. Hydrol.* **638** 131471
[43] Kelleher C and McPhillips L 2020 Exploring the application of topographic indices in urban areas as indicators of pluvial flooding locations *Hydrol. Process.* **34** 780–94
[44] Smith B and Rodriguez S 2017 Spatial analysis of high-resolution radar rainfall and citizen-reported flash flood data in ultra-urban New York City *Water* **9** 736
[45] Basiri A, Haklay M, Foody G and Mooney P 2019 *Crowdsourced Geospatial Data Quality: Challenges and Future Directions* (Taylor & Francis) pp 1588–93
[46] Comber A, Mooney P, Purves R S, Rocchini D, Walz A and Matisziw T C 2016 Crowdsourcing: it matters who the crowd are. The impacts of between group variations in recording land cover *PLoS One* **11** e0158329
[47] Dixon B, Johns R and Fernandez A 2021 The role of crowdsourced data, participatory decision-making and mapping of flood related events *Appl. Geogr.* **128** 102393
[48] Barbier G, Zafarani R, Gao H, Fung G and Liu H 2012 Maximizing benefits from crowdsourced data *Comput. Math. Organ. Theory* **18** 257–79
[49] Zhao Y and Zhu Q 2014 Evaluation on crowdsourcing research: current status and future direction *Inf. Syst. Front.* **16** 417–34
[50] City of Chicago 2024 311 service requests (available at: https://data.cityofchicago.org/widgets/v6vf-nfxy?mobile_redirect=true)
[51] Wilson D and Grammenos D 2005 Gentrification, discourse, and the body: chicago's Humboldt Park *Environ. Plan. D* **23** 295–312
[52] Rinaldo R 2002 Space of resistance: the Puerto Rican cultural center and Humboldt Park *Cult. Crit.* **50** 135–74 (available at: www.jstor.org/stable/1354691?casa_token=L0m1UDoKsIgAAAAA%3ARMoIfen5HiOvAic5kvQdKKjyuOGKQWTuX4akAbt3OG1qpfdHvJt8WRdmHJRiVod2xxJsc9ZebeITN9HbggUFbVT8vuL8MqiSXbOXpJ9Tt5Rk3JuTaiaE)
[53] Lee J, Berkelhammer M, Wilson M D, Love N and Cintron R 2024 Urban land surface temperature downscaling in Chicago: addressing ethnic inequality and gentrification *Remote Sens.* **16** 1639
[54] Cooley W 2018 *Moving Up, Moving Out: The Rise of the Black Middle Class in Chicago* (Northern Illinois University Press)
[55] Toussaint-Comeau M and Newberger R 2017 Reinvesting in the Greater Chatham neighborhoods in Chicago: new data and insights from practitioners and policymakers *ProfitWise News and Views*
[56] Maune D F, Kopp S and Zerdas C 2007 Digital elevation model technologies and applications *The DEM Users Manual*
[57] Gesch D B, Oimoen M J and Evans G A 2014 *Accuracy Assessment of the US Geological Survey National Elevation Dataset, and Comparison with Other Large-area Elevation Datasets: SRTM and ASTER* (US Department of the Interior, US Geological Survey Sioux Falls)
[58] City of Chicago 2024 Building footprints (available at: https://data.cityofchicago.org/Buildings/Building-Footprints-current-/hz9b-7nh8)
[59] Hastie T J 2017 *Generalized Additive Models. Statistical Models in S* (Routledge) pp 249–307
[60] Su X, Liu T, Beheshti M and Prigiobbe V 2020 Relationship between infiltration, sewer rehabilitation, and groundwater flooding in coastal urban areas *Environ. Sci. Pollut. Res.* **27** 14288–98
[61] Pallathadka A, Sauer J, Chang H and Grimm N B 2022 Urban flood risk and green infrastructure: who is exposed to risk and who benefits from investment? A case study of three US Cities *Landsc. Urban Plan.* **223** 104417
[62] Shatkin G 2019 Futures of crisis, futures of urban political theory: flooding in Asian coastal megacities *Int. J. Urban Reg. Res.* **43** 207–26
[63] Chan F, Chuah C J, Ziegler A, Dąbrowski M and Varis O 2018 Towards resilient flood risk management for Asian coastal cities: lessons learned from Hong Kong and Singapore *J. Clean. Prod.* **187** 576–89
[64] Skougaard Kaspersen P, Høegh Ravn N, Arnbjerg-Nielsen K, Madsen H and Drews M 2017 Comparison of the impacts of urban development and climate change on exposing European cities to pluvial flooding *Hydrol. Earth Syst. Sci.* **21** 4131–47