



Assessing Disaster Impact in Real Time: Data-Driven System Integrating Humans, Hazards, and the Built Environment

Haiyan Hao¹ and Yan Wang, Ph.D., A.M.ASCE²

Abstract: Current disaster impact assessments are conducted within days after an event and usually with limited workforce and resources. Although many recognize the valuable information attainable in this way, the evidence is likely to be disturbed, and collecting information rapidly is critical for timely disaster response. Many proposed approaches use either physical- or social-sensing data, and are not implemented in real time, and therefore cannot assess disaster impacts holistically and timely. It also is challenging to integrate multisource data with disparate data modalities and resolutions, and to accommodate large-volume, high-rate input data for real-time computations. Therefore, we propose an assessing disaster impact in real-time (ADIR) system which harnesses cutting-edge computation and geovisualization platforms. ADIR processes heterogeneous data for humans, hazards, and built environments, and outputs multiscale disaster impacts. The usefulness and effectiveness of ADIR were demonstrated in the city of Houston, which was affected by Hurricane Harvey. ADIR provides a platform for multiagent coordination during community response and recovery, which contributes to more-resilient communities by potentially reducing life and property losses. DOI: [10.1061/\(ASCE\)CP.1943-5487.0000970](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000970). © 2021 American Society of Civil Engineers.

Author keywords: Crisis informatics; Disaster impact assessment; Resilience; Real-time data analytics; Geovisualization.

Introduction

Disaster impacts arise as the interactive consequence of the human, hazard, and the built environment (Zhu et al. 2020). Disastrous events such as hurricanes and earthquakes can destroy residences and roadways, cause injuries and deaths, disrupt utility services such as water and power supplies, and generate much debris, affecting cities' normal operation. In recent decades there has been increasing occurrence and severity of major disasters such as tropical storms and wildfires (Kossin et al. 2020; McWethy et al. 2019). Additionally, urban sprawl has brought large populations and expensive civil infrastructures to disaster-vulnerable areas, which poses increased risks of adverse events to human communities (Coronese et al. 2019; Dottori et al. 2018; Smith 2019). In the US, both government agencies and research communities advocate for resilience strategies to mitigate the increased risk and severe impacts of extreme events (NRC 2012). Disaster resilience, as the key concept, refers to the capability of a community to prepare for, recover from, and adapt to disruptions induced by disasters (NRC 2006). Many studies have emerged to articulate the elements, characteristics, strategies, and measurements of disaster resilience (Bruneau et al. 2003; Cutter et al. 2008; Wang et al. 2020). The research outcomes have been used by urban planners and

emergency managers in project planning and policymaking. These efforts ultimately contribute to communities that are more resilient to the negative impacts of disasters.

Two key measurements of disaster resilience are resourcefulness and rapidity (Bruneau et al. 2003), which measure the ability to identify and mobilize resources (e.g., labor, money, and technology) for problem-solving, and the speed of restoring disrupted systems, respectively. For example, emergency responders have expressed their need for timely knowledge of disaster impacts, such as locations and intensities of disaster hazards, operation status of infrastructure systems, needed resources, and so forth. Timely and accurate knowledge can support effective decisions and actions, which optimizes resource mobilization (i.e., resourcefulness) and expedites disaster recovery processes (i.e., rapidity), and hence enhances disaster resilience. However, existing disaster impact assessments usually are conducted after event occurrence by human assessors or with postdisaster surveys. These conventional approaches require extensive labor and time to complete, which may delay the disaster response and recovery and cause extra losses to disaster-affected communities.

Authoritative agencies and research communities seek approaches to accelerate the acquisition of disaster impact information. Agencies such as FEMA and the National Oceanic and Atmospheric Administration (NOAA) use physical sensing devices (e.g., radar, lidar, in situ sensors, and satellites) to gather data and automate impact assessments. These physical sensing data provide insights regarding locations and intensities of ongoing events and conditions of the built environments, but they do not reflect disaster impacts to humans (e.g., injuries, casualties, and service disruptions). Some researchers used social sensing data (e.g., social media posts) generated by people to mine rapid disaster impact information (Kryvasheyeu et al. 2016; Wang and Taylor 2019; Yao and Wang 2019, 2020; Zhong et al. 2016). The social sensing data record personal experiences and perceived impacts of affected communities, but they are limited by low data quality (Agarwal and Yiliyasi 2010; Goodchild 2007; Zou et al. 2019).

¹Ph.D. Student, Dept. of Urban and Regional Planning and Florida Institute for Built Environment Resilience, College of Design, Construction and Planning, Univ. of Florida, 1480 Inner Rd., Gainesville, FL 32601. Email: hhao@ufl.edu

²Assistant Professor, Dept. of Urban and Regional Planning and Florida Institute for Built Environment Resilience, Univ. of Florida, P.O. Box 115706, Gainesville, FL 32611 (corresponding author). ORCID: <https://orcid.org/0000-0002-3946-9418>. Email: yanw@ufl.edu

Note. This manuscript was submitted on September 29, 2020; approved on February 2, 2021; published online on May 19, 2021. Discussion period open until October 19, 2021; separate discussions must be submitted for individual papers. This paper is part of the *Journal of Computing in Civil Engineering*, © ASCE, ISSN 0887-3801.

Consequently, approaches developed with either physical- or social-sensing data sources do not depict disaster impacts comprehensively. Although some of the data can be acquired during disasters, previous studies mainly analyzed them in offline mode to assess collective impacts, which does not make full use of their real-time nature to provide timely and evolving damage information for prompt decision-making.

The use of real-time data sources does not translate into real-time approaches. The latter should consider the complexity of the timely processing of large data volumes and disparate data formats. A few real-time disaster impact assessment systems have been developed in previous studies, but they mostly rely on single data sources. Emergency agencies use weather reports or in situ environmental sensor measurements to predict affected areas and sequential impacts based on numeric models (Glahn et al. 2009; Schneider and Schauer 2006). The models output predictive results that may suffer from modeling errors. Some research studies developed real-time crisis mapping systems leveraging social media data to improve situational awareness, detect resource needs, and identify impact hot-spots (Chen et al. 2016; Choi and Bae 2015; Kelly et al. 2017). However, these data-driven systems do not fully evaluate the credibility of social media data before use. These limitations call for the integration of data from heterogeneous sources that enables a reliable and holistic understanding of disaster situations.

Therefore, we propose a system, assessing disaster impact in real-time (ADIR), to estimate disaster impacts at multiple spatial scales (e.g., neighborhood or city) with multisource data depicting conditions of humans, hazards, and the built environments in affected communities. The system can automatically process static and streaming data with different data modalities and formats, and outputs multiscale disaster impact information during and after disasters. Specifically, the system outputs point-level eyewitness reports, neighborhood-level integrated disaster impacts, and citywide flood probabilities. We employed a set of data-driven techniques to tackle the challenges associated with integrating heterogeneous data with various data modalities (e.g., number, text, and image), formats (e.g., spatial, nonspatial, temporal, and static), resolution (e.g., point and neighborhood), and semantics (e.g., flooding and property damage). We also employed Apache Spark, Apache Kafka, and ArcGIS to achieve real-time computation and geovisualization. The effectiveness of the system was examined with data collected for Houston during Hurricane Harvey in a simulated real-time scenario. The system provides timely and comprehensive disaster impact information that can be used by different stakeholders to coordinate their response and recovery decisions and operations. The diverse and dynamic disaster impact information improves the public's situational awareness, informs infrastructure managers regarding decisions for restoring disrupted civil infrastructure systems, allows first responders to provide more-effective damage control, and leads to timely resource allocations. These decisions and operations contribute to more-disaster-resilient communities that experience less physical and societal losses and recover from disaster disruptions rapidly.

Related Works

We categorized existing studies on disaster impact assessments according to the data sources that were used for the assessments. Notably, very limited studies intended to achieve real-time damage assessments, so we also reviewed real-time disaster information systems that assist in various tasks.

Estimating Hazard Conditions and Built Environment Impacts with Physical Sensing Data

Many approaches have been proposed to estimate the hazard conditions during disasters, including event types, locations, and magnitude, with physical sensing data such as weather reports and in situ environmental sensor measurements. These approaches use engineering modeling or simulation to predict hazard conditions, such as wind speed and inundation depth. For example, some researchers modeled wind fields based on weather observations for tropic storms (Powell et al. 2010) whereas others simulated inundation depth with wind speed, precipitation, and earthquake warning (Glahn et al. 2009; Musa et al. 2018; Shangquan et al. 2019). These systems can deliver predictions for affected areas during disasters, e.g., every 4–6 h in hurricane advisories (Glahn et al. 2009).

However, the estimation of hazard conditions cannot be translated directly into disaster impacts. Some authoritative agencies developed models to estimate disaster losses with hazard conditions and asset inventories. For example, FEMA developed the Hazus models to estimate losses caused by adverse events including earthquakes, floods, hurricanes, and tsunamis (FEMA 2020; Schneider and Schauer 2006). These models use empirical damage functions and formulate structure damages as responses of exposed hazard conditions and structure types. For example, Hazus estimates earthquake damages with ground motion intensities (Kircher et al. 2006), flood damages with inundation depths (Scawthorn et al. 2006a, b), and hurricane damages with wind speeds (Vickery et al. 2006a, b). Hazus archived hundreds of empirical damage functions for different disaster types. Built environment factors, such as terrain roughness, elevation, asset type, and design code, are used to select the damage function for use. The damage functions determine the damage conditions of assessed structures and the aggregated economic losses are calculated with asset values and determined damage conditions. Similarly, the state of Florida developed the Florida Public Hurricane Loss Model (FPHLM) for hurricane loss estimations, which includes models for wind hazards, structure vulnerability, and insured loss costs (Chen et al. 2009).

These numeric models have been widely used by governmental agencies and local insurance companies to determine the initial disaster damage and economic losses. However, engineering models and simulations suffer from uncertainties, because real-world situations are much more complicated and cannot be characterized fully by numeric inputs. Empirical models that rely on empirical damage functions often are calibrated to the aggregated loss of major disasters, which can be inaccurate for individual structures (Merz et al. 2010; Wing et al. 2020). Sociodemographic factors, such as race, income, and education, are not considered in these models, whereas abundant evidence shows that disasters impact different social groups inappropriately and disaster response decisions should consider the social vulnerabilities of affected communities (Aerts et al. 2018; Eid and El-Adaway 2017). For example, Burton (2010) showed better predictive performance for modeling catastrophe damage when the Social Vulnerability Index (SoVI) is included in the modeling; social vulnerabilities refer to socioeconomic and demographic factors that affect communities' resilience toward disasters (Flanagan et al. 2011). These human-related influences have not been well captured by the physical-sensing data in assessing disaster impacts.

Assessing Disaster Impacts on Human Communities with Social Sensing Data

Social sensing data can record disaster impacts from a population's perspectives. Previous studies developed approaches to mine

disaster impacts on human communities from social sensing data using techniques such as natural language processing and computer vision. For example, Imran et al. (2013) used the naïve Bayesian model to classify tweet messages containing caution and advice, casualties and damage, and donations with subcategories. Zhang et al. (2020) proposed a semiautomatic approach combining keyword search, clustering, and bidirectional long short-term memory (Bi-LSTM) to monitor the public's sentiment toward community disruptions including built environment hazards, casualty, and power outage. Similarly, Pouyanfar et al. (2019) fused the visual and audio features of crowdsourced video clips to predict video contents, e.g., flood/storm, human relief, damage, and victims. These approaches identify predetermined disaster impact information concerning humans, urban services, public facilities, and personal properties. The yielded outcomes reflect the experience and hardship of affected people, and thus can be utilized by disaster responders (Roy et al. 2020; Zhang et al. 2020).

These approaches estimate disaster impacts based on users' observations that are exempted from the uncertainties of numeric models. However, the imbalanced distribution of social-sensing data may aggravate inequitable emergency response and relief operations. The data sparsity may return a few on-topic posts for fine-grained assessments. Additionally, previous work using social media data rarely evaluated data credibility before usage. Some researchers addressed these limitations by integrating developed approaches with human intelligence (Alam et al. 2018; Fan et al. 2021), which made the approaches not fully automatic. Although with these limitations, social sensing data still are investigated widely because they provide free, versatile, and human-centric information for disaster management.

Disaster impact assessment approaches based on either physical or social sensing data may be inadequate to evaluate the impacts on both the built environments and humans. This limitation calls for approaches that integrate heterogeneous data sources depicting different aspects of affected communities. In addition, the aforementioned disaster impact assessments were not tested with real-time computing settings. The complex modeling approaches, large data volumes, and high data throughputs may impede these approaches from delivering assessment results continuously during disasters.

Real-Time Data-Driven Systems for Disaster Management

A few real-time data-driven systems were developed to assist in disaster decisions and improving situational awareness. The early systems in this area, such as Ushahidi and Sahana, harness the power of crowdsourcing and integrate information contributed by online volunteers (Careem et al. 2006; Gao et al. 2011). The USGS developed the Did You Feel It (DYFI) system to exploit citizens' reports about earthquake shaking and damages, which generates intensity maps after earthquakes (Wald et al. 2011). The emergence of these tools provides platforms allowing different emergency stakeholders to coordinate disaster relief operations (Gao et al. 2011), but the effectiveness of these tools is affected highly by their user group size. Some researchers developed systems based on real-time social media data to monitor ongoing disasters (Choi and Bae 2015), assist resource dispatching (Chen et al. 2016), detect impact hot-spots (Avvenuti et al. 2018; Tsou et al. 2017), or detect abnormal incidents (Kelly et al. 2017). These systems display results on web-based maps that easily can be viewed and shared among emergency stakeholders, but they only use single data sources, i.e., real-time social media posts.

We found a few systems that integrate data from heterogeneous sources. Huang et al. (2017) presented a cloud-based system that

synthesizes multisource data including social media, remote sensing, Wikipedia, and web data. The system allows users to analyze historical events as well as track real-time events. Nara et al. (2017) proposed a real-time system to support evacuation decision-making during wildfires, which recommends evacuation routes based on real-time population distribution estimation, wildfire parameters, and people's feedback. Wang et al. (2017) proposed a system for real-time wildfire risk management which predicts and visualizes property-level wildfire risks with real-time weather observations and location-specific attributes such as vegetation and elevation.

In summary, a variety of approaches and systems have been proposed to mine disaster impact information, but most of them conduct analyses with archived data that do not yield real-time knowledge. A few real-time disaster information systems rely heavily on social media data, which can be unreliable and suffer from spatial bias. There still is no system that provides comprehensive real-time disaster impact information with multisource data inputs depicting different aspects of affected communities.

System Architecture of ADIR

The proposed ADIR system assesses disaster impacts in real-time with multisource data depicting human, hazard, and the built environment aspects of disaster-affected communities. The system consists of four modules liable to data collection, analysis, storage, and geovisualization (Fig. 1).

Data Collection Module

The data collection module collects data from different sources that characterize the human, hazards, and built environment factors of

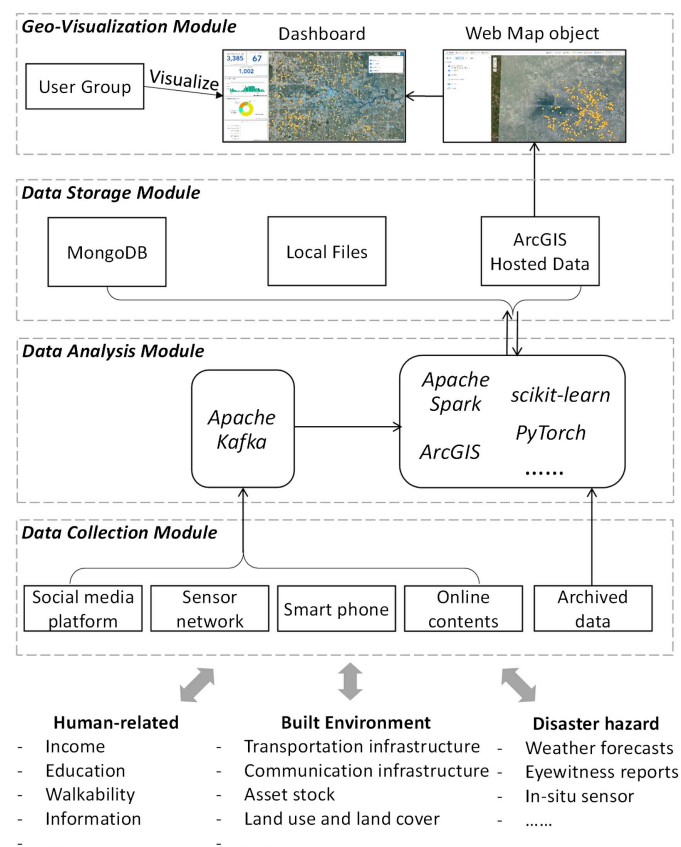


Fig. 1. System framework of ADIR.

Table 1. Multisource data used for disaster impact assessment

Data source	Category	Temporal form	Spatial form
Real-time data			
Streaming tweets	H, D, BE	Streaming	Geographic coordinates
311 service requests	H, D, BE	Streaming	Geographic coordinates
USGS stream gauge readings	D	Updated every 15 min	—
Static data			
USGS DEM	BE	—	10 m
CDC SoVI index	H	—	Census tracts
USGS stream gauge locations	—	—	Geographic coordinates

Note: H = Human; D = Disaster; and BE = Built-environment.

affected communities. The collected data are transmitted to the data analysis module directly for analysis. Both real-time and static data are collected through this module. Some real-time data sources that can be harnessed for disaster impact assessments and the specific data collection approaches include

- Real-time social media data. Disaster victims may post their observations on online platforms (e.g., Twitter) during disasters, but the generated data often are kept proprietary by companies who operate these platforms. Some companies provide open-sourced APIs for application developers to obtain streaming posts with programming (Twitter 2020).
- Real-time infrastructure/environment sensor measurements. In situ and mobile sensors continuously measure the status of monitored environments/infrastructure. Authoritative agencies, such as NOAA and USGS, deploy abundant sensors to monitor meteorological, hydrological, and geological environments. Infrastructure managers in cities also use sensors to monitor the operating status of infrastructure systems. The generated sensor measurements are archived privately in data hubs administrated

by these stakeholders, but some agencies may publish data online or provide subscription services allowing external users to access the real-time sensor measurements.

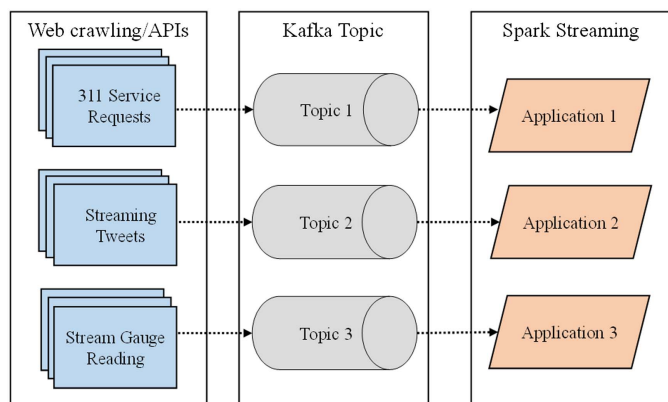
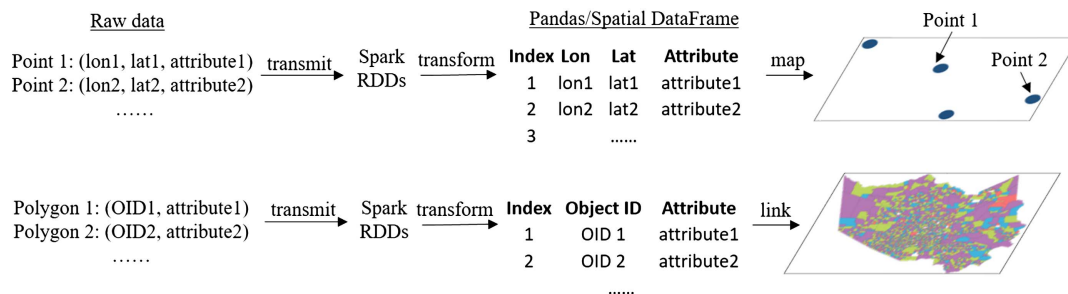
- Real-time hazard data. In addition to the raw sensor measurements, authoritative agencies also deliver data describing the disaster hazards intermittently during the event. For example, National Hurricane Center (NHC) delivers hurricane advisories every 4–6 h during hurricane onsets, which include wind directions, intensities, and wind center locations. Approaches such as web-crawling may download these data automatically.

To implement the proposed ADIR system for hurricane impact assessments, we selected some open-source data that are available for most US cities as the system inputs. Table 1 lists sources, categories, and temporal and spatial formats of input data.

Data Analysis Module

The data analysis module consumes received static and streaming data and conducts analyses. Static data are loaded to the module before analyses. Fig. 2 shows the transmission and processing of heterogeneous real-time data. Specifically, we created different Apache Kafka topics to receive heterogeneous streaming data. Applications written in Apache Spark can access these data by subscribing to corresponding Kafka topics. Both Kafka and Spark are distributed computation platforms capable of parallel computing.

Data streams transmitted to Spark are discretized into micro-batches, or resilient distribution data sets (RDDs). RDDs are read-only data collections that can be partitioned across large computing clusters. Spark regulates two types of functions for RDDs, i.e., transformation and action. A transformation function transforms a RDD into a new RDD, and an action function returns the analysis results. By integrating Spark functions and third-party libraries (e.g., scikit-learn, PyTorch, and ArcGIS), ADIR can achieve a series of complex data transformations and analyses. For example, the system deals with two types of spatial data, i.e., point-level eyewitness reports and neighborhood-level disaster impacts. ADIR uses Spark transformation functions to transform

**Fig. 2.** Transmitting multisourced streaming data.**Fig. 3.** Transforming and mapping spatial data.

received raw data into different data formats (Fig. 3). The raw point data containing numeric coordinates are converted to spatial DataFrames that can be mapped with ArcGIS. Polygons representing neighborhoods have fixed spatial locations. ADIR processes nonspatial attributes describing the changing status of polygons (e.g., flooding conditions) and links these attributes to cloud-stored polygon maps through unique identification numbers. We use action functions to write results to the data storage module.

With the Spark and third-party libraries, we implement three parallel analysis processes (described in the following subsections) in the data analysis module for hurricane impact assessment, which extracts multiscale disaster impact information.

Mining Eyewitness Reports from User-Generated Point Data

ADIR retrieves eyewitness reports describing disaster impacts from 311 service requests and geotagged streaming tweets. Compared with social media posts, 311 service requests generally are more relevant and reliable, but not all requests represent damages caused by disasters. This research mainly focused on six request types associated with hurricane-induced impacts on humans and the built environment: crisis cleanup, storm debris collection, flooding, tree removal, storm damage, and medicine evacuation. A filter is used to limit the mapped 311 service requests to be one of the six types. The raw crawled social media data can be noisy and irrelevant. We use an approach developed by Hao and Wang (2020) to filter out irrelevant reports and extract various disaster damage information from geotagged streaming tweets. The approach uses computer vision and text mining to analyze multimodal social media data (Fig. 4). ADIR checks whether each tweet arriving in the stream includes an image or images, and downloads the image(s) if it does. Then the approach is applied to tweet texts and images to extract disaster impact information, including hazard types (i.e., wind hazard, flood hazard), severities, and damage types (e.g., infrastructure damage, power outage, and house damage). The filtered 311 service requests and tweets identified to include disaster impact information are transformed to point features and saved in a hosted data service in the data storage module.

The eyewitness reports obtained from social media data can be unreliable due to the existence of fake content, photoshopped images, inaccurate spatial and temporal attributes, and less precise models for knowledge extraction, which can undermine the validity

of ADIR. We use a density-based method to check and tag the reliability of eyewitness reports obtained from social media data. An eyewitness report o is tagged to be reliable when the number of reports falling in its temporal-spatial neighborhood $N_{\varepsilon_s, \varepsilon_t}(o)$ exceeds a threshold, ε_n . The neighborhood $N_{\varepsilon_s, \varepsilon_t}(o)$ is determined using Eq. (1), where $\text{dist}_s(p, o)$ and $\text{dist}_t(p, o)$ is Euclidean and temporal distance between report p and o , and ε_s and ε_t are pre-determined spatial and temporal thresholds. The basic assumption is that an eyewitness report is credible if multiple reports are found with close spatial and temporal proximity to that report (Sakaki et al. 2010)

$$N_{\varepsilon_s, \varepsilon_t}(o) = \{p \in D | \text{dist}_s(p, o) \leq \varepsilon_s, \text{dist}_t(p, o) \leq \varepsilon_t\} \quad (1)$$

Computing Citywide Flood Probabilities with Stream Gauge Measurements and Digital Elevation Models

ADIR maps flood probabilities with the USGS digital elevation model (DEM) file and gauge height measurements. It intermittently retrieves the maximum gauge heights of stream gauges located within and near the study region over consecutive time windows (e.g., 3 h). We adopt an inverse distance weighted-based method described by Li et al. (2018) for flood probability mapping. The method generates a flood probability index (FPI) surface for each stream gauge (p) using

$$\text{FPI}_{pi} = \frac{\max(\Delta h_{wp} + h_p - h_i)^a}{(d_{pi})^b} \quad (2)$$

where d_{pi} = Euclidian distance between location i and stream gauge p ; h_p and h_i = elevation of stream gauge p and location i ; Δh_{wp} = increment of gauge heights of stream gauge p at investigated time compared with its normal condition; and a and b are parameters used to adjust impact magnitude of water height and distance. For this study, we tested different combinations of a and b , and set $a = b = 1$.

We calculate the final flood probability for location i , i.e., $P(i)$, as the average value of FPIs associated with that location [Eq. (3)]. Averaging reduces the result difference caused by different numbers of stream gauges used in analyses. Eq. (3) may not return probabilities ranging from 0 to 1. We tested the method and use a step function [Eq. (4)] to map the obtained $P(i)$ into the range

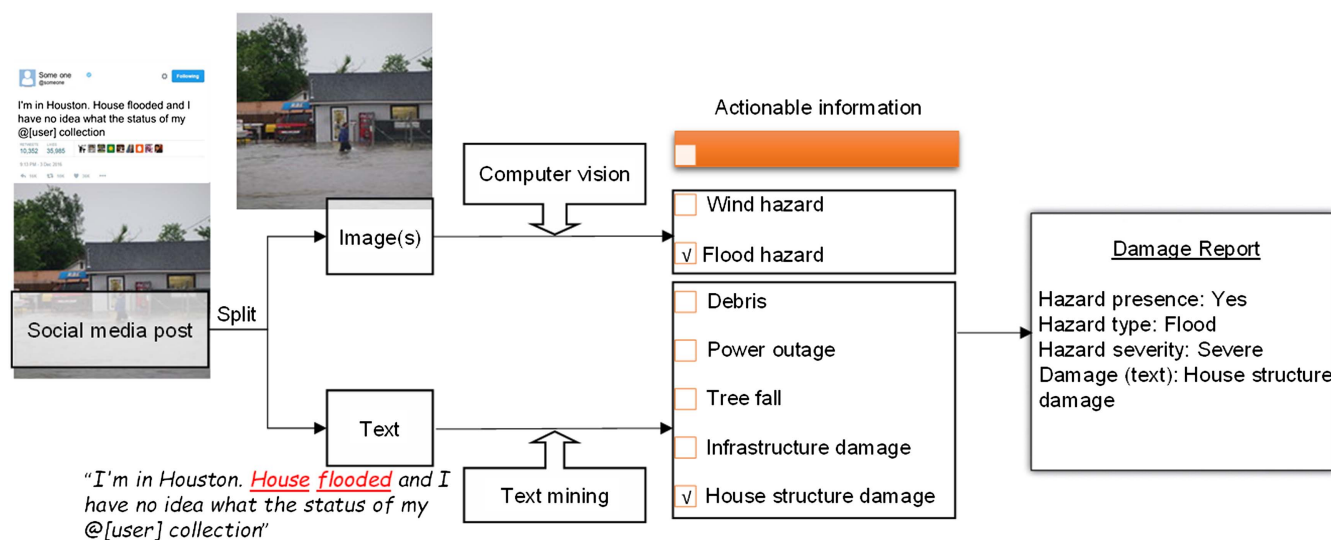


Fig. 4. Multimodal social media analysis. [Image courtesy of Todd Dwyer, under Creative Commons-BY-SA 2.0 (<https://creativecommons.org/licenses/by-sa/2.0/>).]

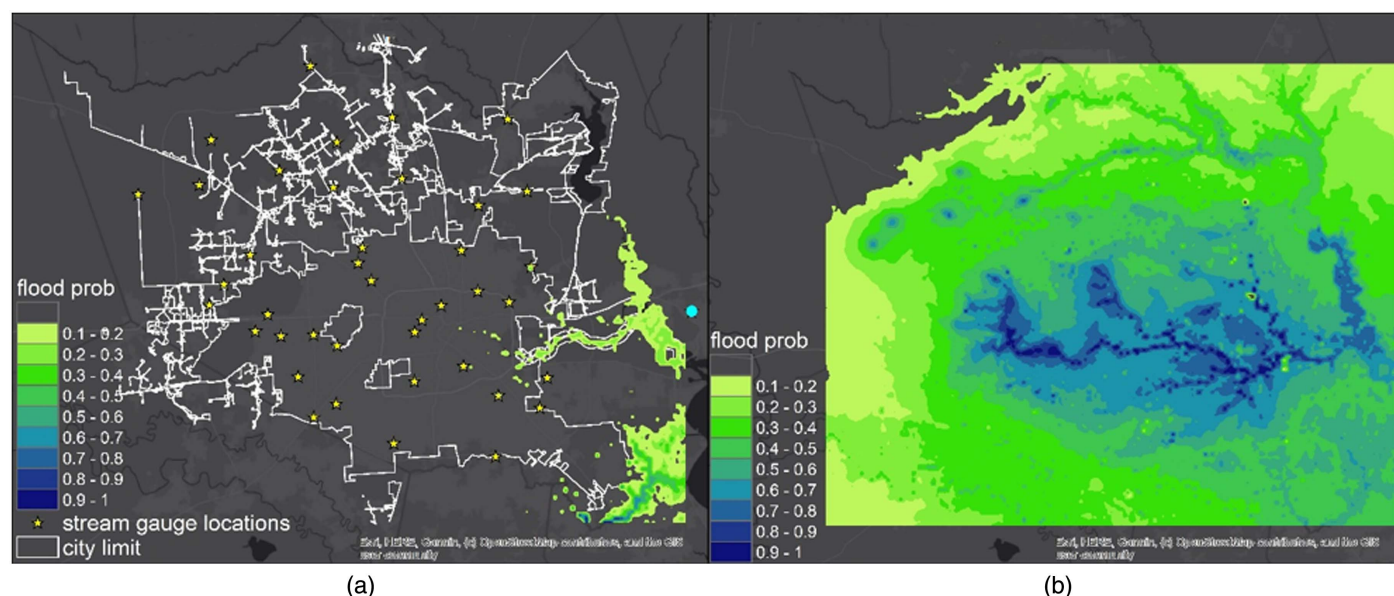


Fig. 5. (a) FPI generated with a single stream gauge measurement; and (b) overall flood probability map. (Base maps by Esri, HERE, © OpenStreetMap contributors, and the GIS user community.)

of 0–1 based on the test results. Fig. 5 demonstrates a FPI surface generated with a single stream gauge measurement and an overall flood probability map. We use ArcGIS's Python module to implement the spatial analyses. The generated flood probability maps are converted to Keyhole Markup Language (KML) files that can be added and updated in web map objects in the geovisualization module

$$P(i) = \frac{1}{n} \sum_{p=1}^n \text{FPI}_{pi} \quad (3)$$

$$P(i) = \min(P(i), 1) \quad (4)$$

Integrating Hazard and Social Vulnerabilities for Neighborhood Impact Assessment

We assess disaster impacts at the neighborhood scale by integrating hazard conditions and social vulnerabilities. For a given time, the flood probability for each neighborhood is calculated with the coordinate of neighborhood centroid and the latest generated flood probability map. Similarly, we also calculate flood probabilities for reliable eyewitness reports with their coordinates and the flood probability map. We use the minimum and quartiles (i.e., 0%, 25%, 50%, and 75%) of the flood probabilities associated with reliable eyewitness reports as thresholds to divide neighborhoods into five hazard-vulnerability levels (Levels 0–4). Thus, a neighborhood with the highest flood vulnerability level (Level 4) has a centroid flood probability higher than the flood probabilities associated with 75% reliable eyewitness reports, which shows evident damage. Because eyewitness reports are used only for setting the thresholds for dividing neighborhoods, the generated flood vulnerability maps are immune to potentially unbalanced spatial distribution of eyewitness reports.

Neighborhoods undergoing more severe hazard conditions may not experience more disaster impacts than other neighborhoods. Social factors, such as population, wealth, connectedness, and education, can modify disaster outcomes (Burton 2010). These factors are represented comprehensively by SoVI. This research used the Centers for Disease Control and Prevention's (CDC) SoVI, which is determined with variables depicting four themes:

socioeconomic status (e.g., income, unemployment, and education), household composition (e.g., age, disability, and single-parent households), minority status (e.g., disability and language), and housing and transportation (i.e., multiunit home, mobile home, and possession of vehicles). The index is normalized into the range 0–1, with which we equally divided neighborhoods into five levels. We use a bivariate map to show the neighborhood hazard vulnerability levels and social vulnerability levels together.

For illustration, Fig. 6 shows the temporal change of raw computed flood probabilities, neighborhood-level hazard vulnerabilities that are curated by reliable eyewitness reports, and a bivariate map showing integrated neighborhood vulnerabilities. The flood probabilities mapped with DEM and stream gauge readings suggest expanding flood zones in western Houston from September 1–3, 2017, but the neighborhood-level hazard vulnerability maps display an inverse trend and predict fewer neighborhoods experiencing flooding on September 3 when taking into account reliable eyewitness reports. The curated trend of neighborhood hazard vulnerabilities also is consistent with the USACE's estimations that flood conditions were relieved after the release of two western reservoirs (USACE 2017).

Data Storage Module

The analysis results from different sections of the data analysis module are stored in different locations, including MongoDB, ArcGIS's hosted data service, and local files. MongoDB was selected to store the analysis results because it is an open-source and cloud-based database suitable for storing big data. MongoDB is a NoSQL (i.e., nonrelational) database that stores data as schema-free documents with key-value structures. In Spark, we write the analysis results continuously to MongoDB. The assessment of disaster impacts may use data across temporal and spatial scales. For example, the mapping of individual eyewitness reports is executed as soon as the report is received, but the mapping of neighborhood impacts refers to reliable eyewitness reports identified within a time window. MongoDB supports queries with temporal, content, or spatial filters, and therefore the archived data can be retrieved according to temporal or spatial attributes for analyses.

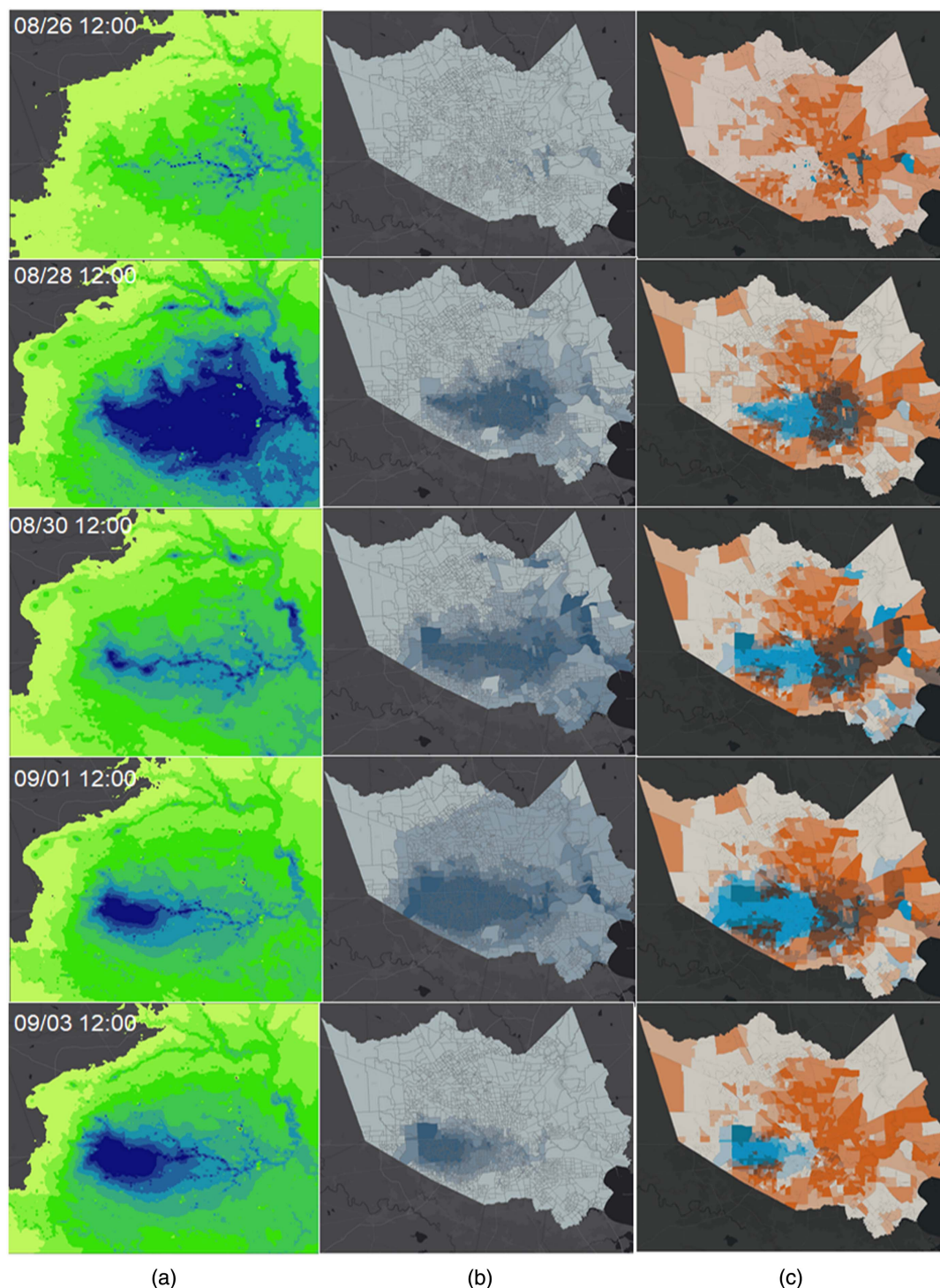


Fig. 6. (a) Mapped flood probabilities; (b) neighborhood hazard vulnerabilities; and (c) neighborhood integrated vulnerabilities. (Base maps by Esri, HERE, © OpenStreetMap contributors, and the GIS user community.)

Spatial data (e.g., points and polygons) that are displayed for users' visualization are stored in a hosted data service operated by ArcGIS for Developer ([Esri 2020a](#)). The hosted data service is a cloud-based database that provides access to manage, query, and edit spatial data. Local folders provide extra options to store critical analysis results.

Geovisualization Module

The geovisualization module communicates the analysis results to system users. We implemented the geovisualization layer with

a set of ArcGIS products developed by ESRI, including ArcGIS for Developer, ArcGIS Dashboard, and ArcGIS Online. The spatial data stored in hosted data services can be added directly as feature layers in web map objects of ArcGIS Online. Then we configure web map objects into web-based dashboards and add elements, such as charts, gauges, and indicators, to describe the data. The dashboard offers a comprehensive and engaging view of displayed spatial data and provides key insights for at-a-glance decision-making ([Esri 2020b](#)). When users make a change to the spatial data stored in hosted data services, the change can be shown

immediately in the configured dashboards. In this way, real-time geovisualization is realized.

Application and Evaluation of ADIR in Simulated Real-Time Hurricane

Case and Data Description

We applied the ADIR system to a historical disaster, Hurricane Harvey, that in 2017 affected Houston, a city in southern Texas with a population of more than 2 million. Hurricane Harvey was a Category 4 hurricane that made landfall on coastal Texas on the night of August 25, 2017. The storm brought exceptional rainfall of 76.2–152.4 cm (35–60 in.) and sustained winds of about 177.0 km/h (110 mi/h) to the coastline. Harvey severely impacted southern Texas cities, including Houston. The urban river runoff caused massive flooding that inundated nearly one-third of Houston, disabled almost all major roads in the city, and cut power connections to many households (Blake and Zelinsky 2018). The severe impacts and large affected population make Harvey and Houston an ideal study case to evaluate our proposed system. We collected geotagged tweets in Houston during the study period, i.e., August 25 to September 5, 2017, using a Twitter streaming API (Wang et al. 2020). We downloaded the 311 service requests reported by Houston citizens during the study period from the Houston 311 portal. The portal collects 311 service requests received from multiple resources (e.g., phone calls, online forms, and emails). For hazard data, we used the measurements of 46 stream gauges located in and near Houston which were published online in the National Water Information System (NWIS) by USGS. Fig. 7 shows the location of the 46 stream gauges and the measurement of

a stream gauge during the study period. The measurements were updated every 15 min. For the built environment and social vulnerability data, we used the USGS DEM data with 10-m resolution and the 2016 CDC SoVI for the study region. The SoVI is mapped for each census tract in Harris County. The general data sets are described in Table 1.

Because we investigated a historical event and the real-time data were collected and archived before system implementation, we developed a simulator to simulate these archived data in a streaming manner according to their associated timestamps (Halse et al. 2019). The simulator was developed in Apache Kafka (Fig. 8). It sorted the archived data according to their temporal sequences and then sent the data contents to designated Kafka topics following the same temporal pattern as that in which they were created. We converted the creation time of different data sources to the same time zone, i.e., Coordinated Universal Time (UTC), before running the simulator. When the system is used for an ongoing event, users can connect streaming data inputs to Kafka topics directly.

Analytical and Geovisualization Outcomes from ADIR

With the simulated streaming data, ADIR generates maps for eyewitness reports, flood probabilities, and integrated neighborhood impacts with approaches described in the section “Data Analysis Module.” Then the analysis results are stored and visualized with tools described in the sections “Data Storage Module” and “Geovisualization Module.” The dynamic disaster impact assessments is demonstrated in Video S1 in the Supplemental Materials.

The system maps eyewitness reports as individual points with attributes showing location, time, type of report, and reliability. Images are attached, if they are detected, to show evidence of

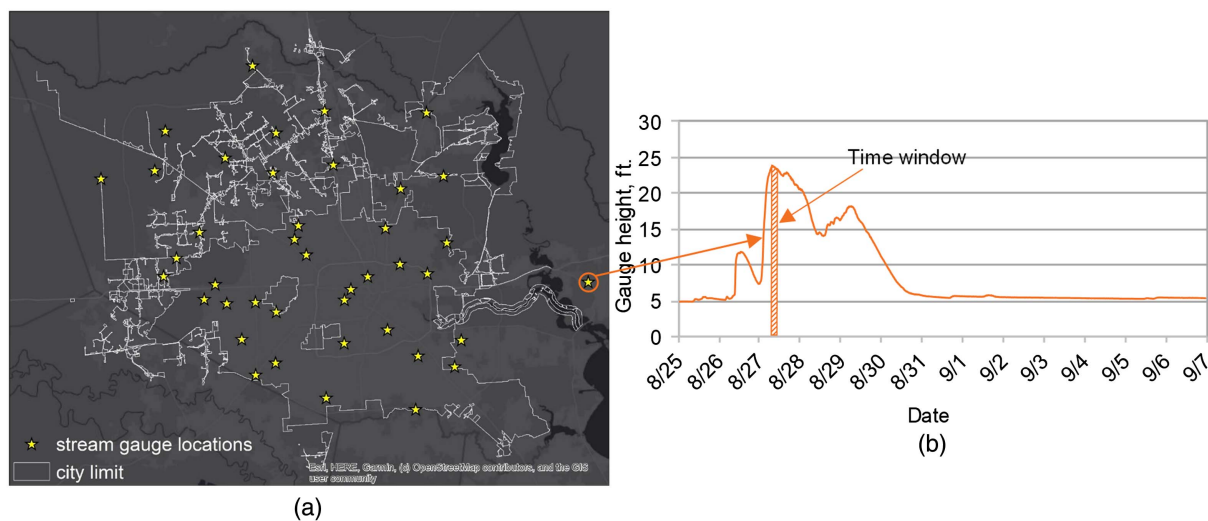


Fig. 7. (a) Location of stream gauge sites (base map by Esri, HERE, © OpenStreetMap contributors, and the GIS user community); and (b) measurements of a single stream gauge.

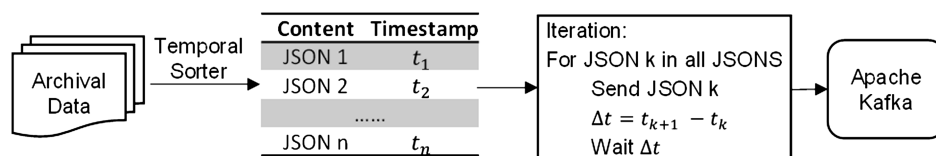


Fig. 8. Simulating real-time streaming data.

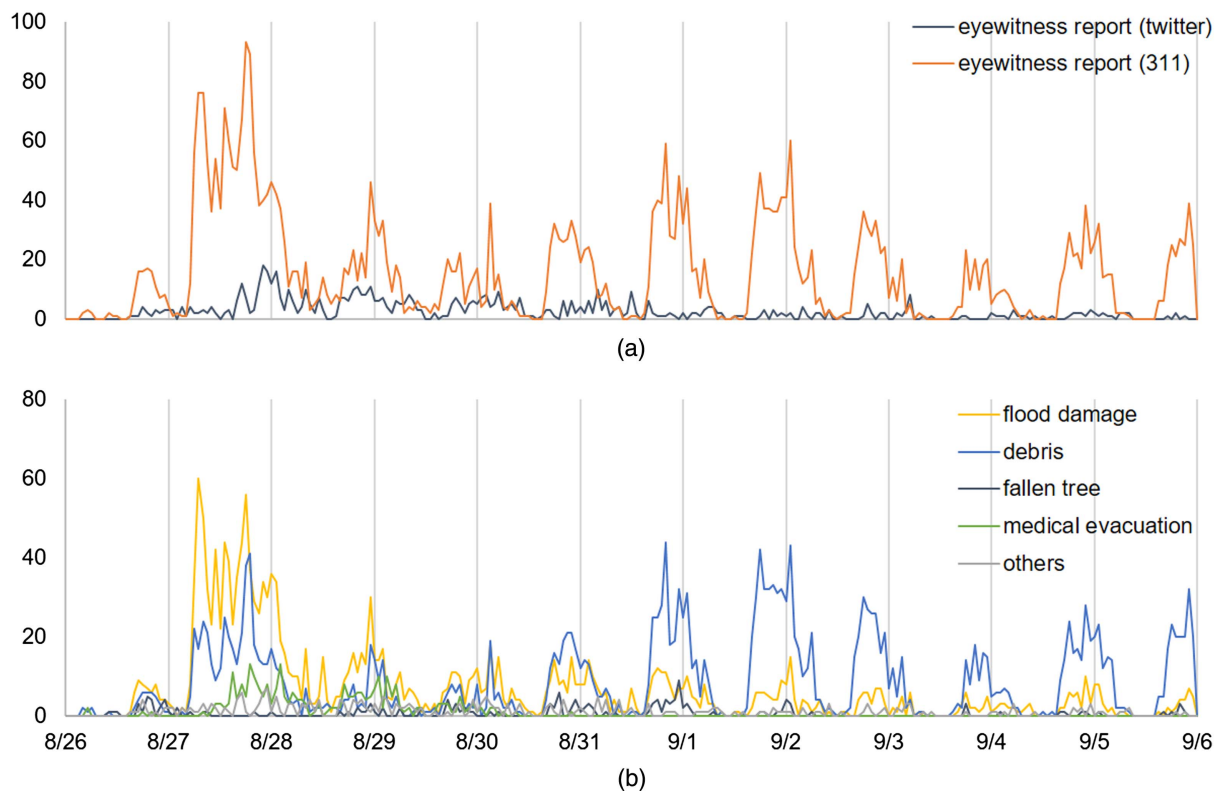


Fig. 9. Temporal distribution of eyewitness reports by (a) source; and (b) type.

flood or wind hazards. ADIR detects nine types of disaster impact information from streaming tweets and 311 service requests: power outage, roadway blockage, residence damage, fallen tree, vehicle damage, debris, flood damage, wind damage, and medical evacuation. Fig. 9 shows the temporal distribution of identified eyewitness reports. The 311 service request data contributed more eyewitness reports than did streaming tweets (Fig. 9). Flood damage and debris were the most reported impact types. The numbers of reports regarding flood damage and medical evacuation peaked on August 27 and then gradually decreased, whereas reports about debris dominated in the later investigation period. We use different symbols to represent different report types. A tweet report is denoted multiple if it contains multiple types of impact information. The tagging of tweet report reliability uses thresholds $\varepsilon_s = 1,000$ m, $\varepsilon_t = 3$ h, and $\varepsilon_n = 3$. The selection of thresholds is based on the intuition that the hurricane situations would not change drastically for locations within 1 km and within a temporal period of 3 h.

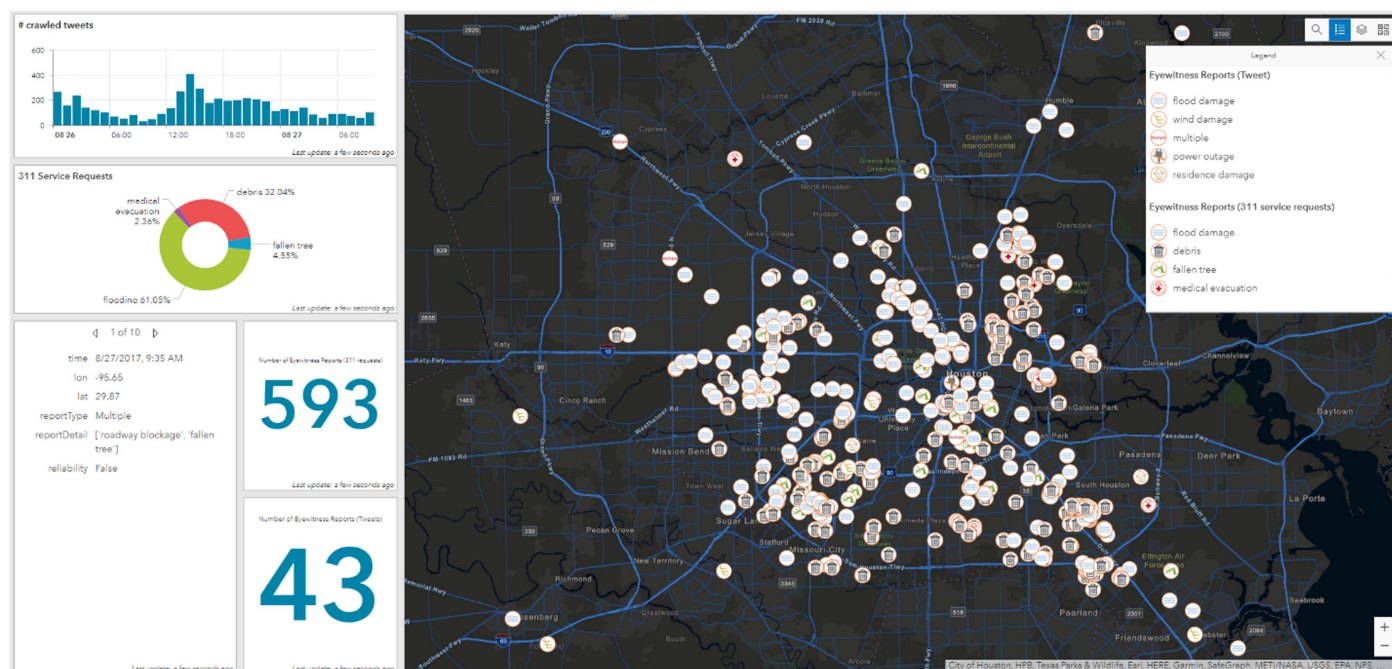
The dashboard sidebar shows the temporal trend of received tweets, report type distribution, report number, and most recently received eyewitness reports. System users quickly can grasp disaster situations by looking at the displayed maps and sidebar charts. They also can zoom in and click individual reports to check their attributes and images. Fig. 10 shows two screenshots showing the mapped eyewitness reports by 9:00 a.m., August 27.

The flood probability map provides ADIR users with a general sense of flood locations and intensities, in which darker colors represent areas with higher flood probabilities. Fig. 11 shows screenshots of the displayed flood probability maps at 10:00 a.m. and 4:00 p.m. on August 27. Central Houston and areas along the Buffalo Bayou River were flooded, and the flooding areas sprawled over the 6 h (Fig. 11). The integrated neighborhood impact map displays neighborhood hazard vulnerabilities and

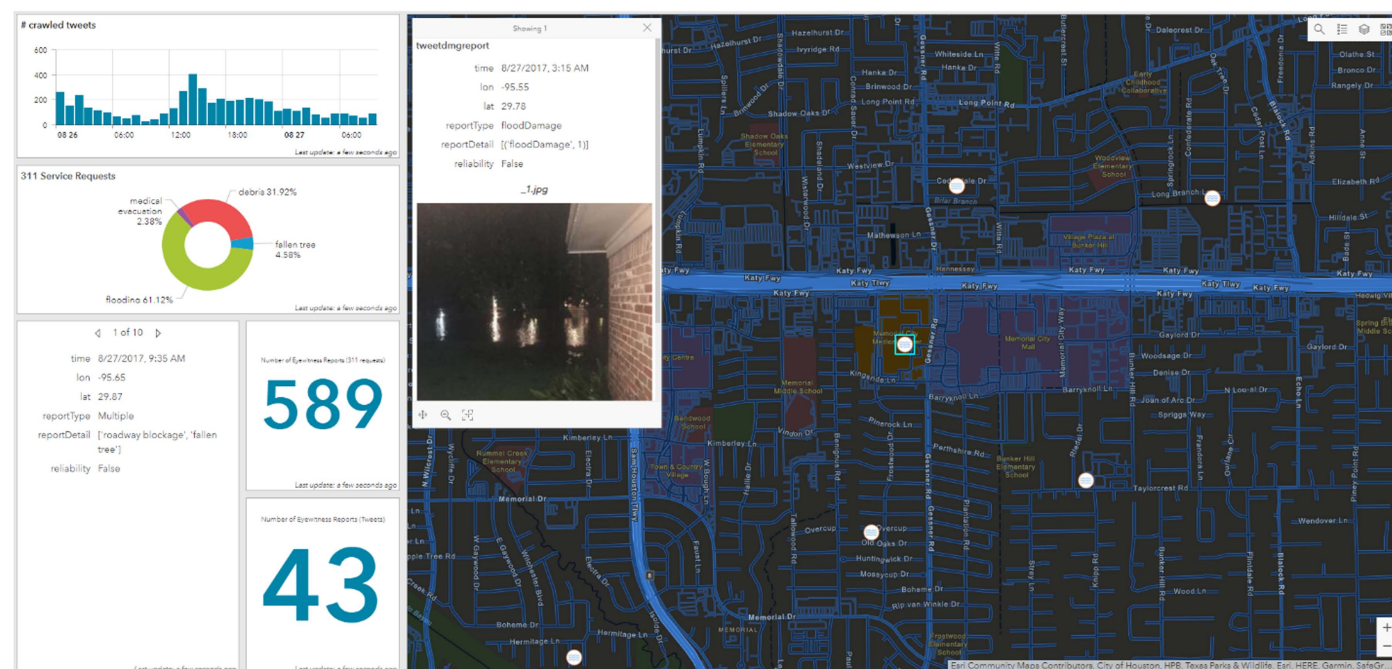
social vulnerabilities with red and blue colors, respectively. Fig. 12 shows the integrated neighborhood impact maps. Stakeholders can refer to the map to plan disaster response activities. For example, disaster managers may want to dispatch resources and rescue teams to neighborhoods displayed in deep purple first, which suggests the exposure of both severe flooding and high social vulnerabilities (Fig. 12). Disaster managers may check crowdsourced reports (Fig. 9) collected within or near these neighborhoods to acquire a general sense of disaster situations before dispatching. These three maps are displayed on a single dashboard in which users can switch or combine the feature layers for display.

Performance Evaluation of ADIR

We evaluated the system performance regarding the (1) assessment accuracy and (2) timeliness of the processing. First, we examined the accuracy of impact assessments using an existing damage assessment benchmark, which annotated buildings located in areas affected by Hurricane Harvey with five damage levels (i.e., none, affected, minimum, major, and destroyed) (Choe et al. 2018). The annotation referred to the FEMA's postdisaster estimation and a crowdsourced project. We randomly inspected some spatial windows containing damaged buildings (Fig. 13). We found that affected or damaged buildings mostly were located in areas with sustained flooding probabilities higher than 0.7 during the study period. We then treated the building levels with two independent variables, i.e., average flood probabilities and the number of inundated days, which are calculated with the flood probability maps generated from August 2 to September 4, 2017. The basic assumption is that a building is more likely to experience severer damage if it is associated with higher flood probabilities and/or longer inundation periods. The correlation tests had positive correlations, with coefficients 0.23 (p -value < 0.0001) and 0.28



(a)



(b)

Fig. 10. Mapped eyewitness reports: (a) overview; and (b) zoomed-in view. (Base maps by City of Houston, HPB, Texas Parks & Wildlife, Esri, HERE, Garmin, SafeGraph, METI/NASA, USGS, EPA, NPS, USDA.)

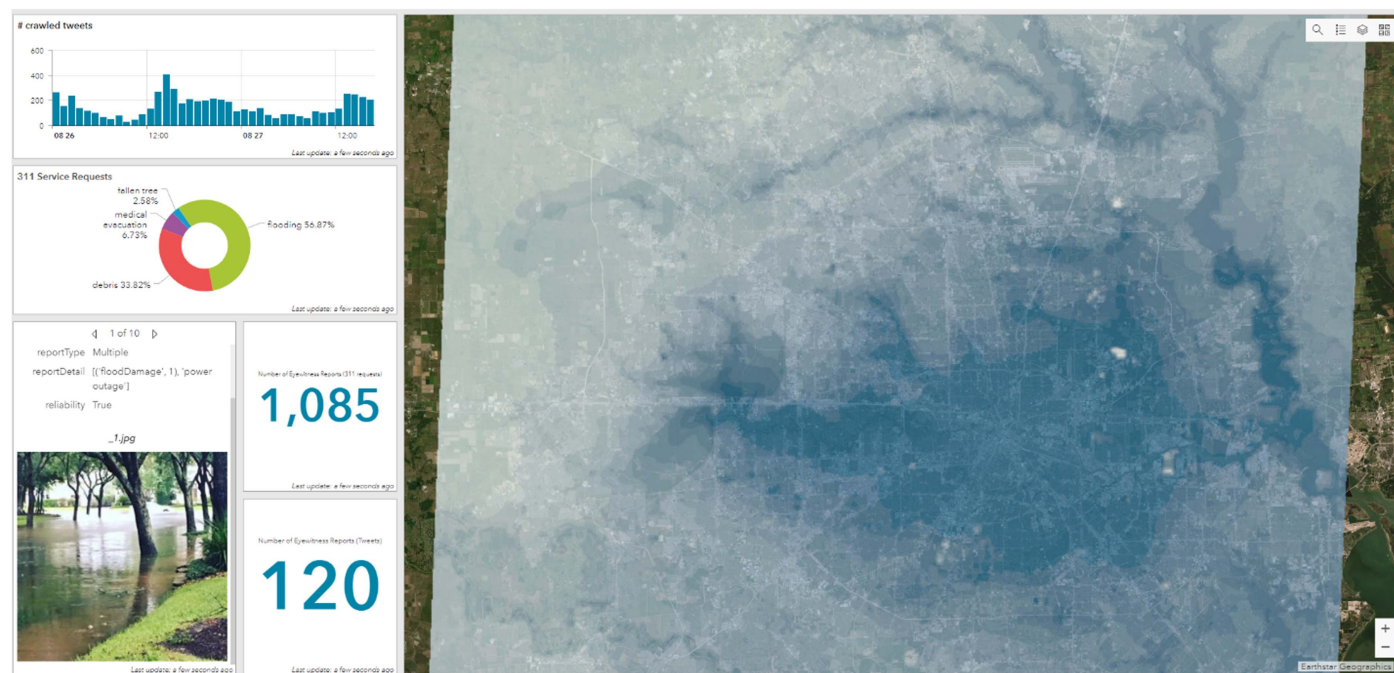
(p -value < 0.0001) for the two independent variables (Table 2). The building damage levels also are determined by other factors such as building qualities. The correlation tests showed that our system generally can predict areas that are more likely to experience property damages.

The real-time setting can deliver results throughout events and can show changing disaster situations for AIDR users. We checked whether the dynamic mapping results delivered by AIDR corresponded to the real-world changes. However, we found only a video created by the USACE, Galveston District showing the

changing inundation conditions for western Houston (USACE 2017). The USACE's video screenshots are shown in Fig. 14(a) and the ADIR flood maps (with flood probabilities equal or larger than 0.7) are shown in Fig. 14(b) with the same temporal and spatial scales. Flood maps from both sources showed scattered flooding in central Houston on August 26 and the massive flooding on August 28, which almost inundated the whole city. Flood waters receded from central Houston on August 30. The USACE's video shows that the flood water accumulated in two western reservoirs, i.e., Addicks and Barker Reservoirs, after



(a)



(b)

Fig. 11. Flood probability map for August 27, 2017 at (a) 10:00 a.m.; and (b) 4:00 p.m. (Base maps by Earthstar Geographics.)

August 30, which relieved the flooding conditions in other places. However, the flood probability maps created with our approach did not capture the discrepancies caused by the human-operated flood control devices. They predicted that western Houston gradually was inundated with flood water accumulated there. However, we consider that our system can show the general flooding conditions and locations over time. However, future improvements should consider flood-control infrastructures.

We then examined the timeliness of the processing performance for the three data streams, i.e., streaming tweets, 311 service requests, and stream gauge readings. In particular, we monitored their

input rates (i.e., the numbers of records in the present data batch), processing time (i.e., the time spent to process the present data batch), and scheduling delay (i.e., the time delay need to wait to process the present data batch). We tested the system rigorously by accelerating the simulator by 10 times, meaning that the data volume the system received in 1 s corresponded to the 10-s data volume collected in the real world. We did this considering that the data volume generated by sensors and citizens increases rapidly, and the future application of ADIR may incorporate more real-time data sources. We set the batch interval to 30 s and ran the experiment for 5 day's data, i.e., from 12:00 p.m. August 26 to 12:00 p.m.

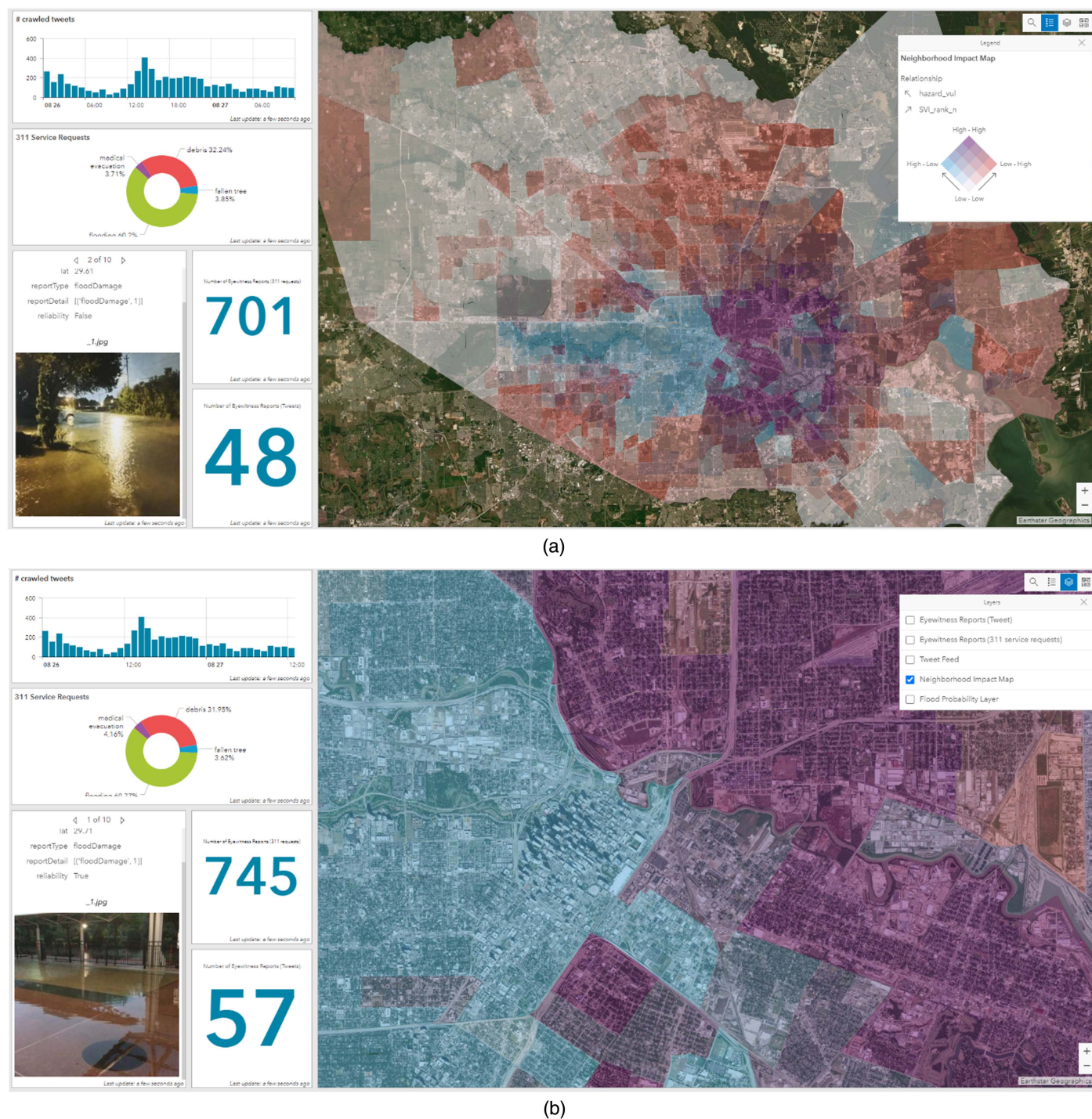


Fig. 12. Integrated neighborhood impact map: (a) overview; and (b) zoomed-in view. (Base maps by Earthstar Geographics.)

August 31, 2017, on a local computer with an Intel Core i7 processor and 16 GB memory.

Fig. 15 shows the performance metrics for the processing of the three data sources. The input rates of streaming tweets and 311 service requests had periodic patterns in which more reports were received in the daytime than at nighttime. Stream gauge measurements were routinely updated every 15 min. The average input rates were about 20 and 5 for streaming tweets and 311 service requests, and the average processing times were 25 and 5 s, respectively. The longer processing time of streaming tweets was caused by the higher input rate and potential image processing, which

included checking and downloading images, analyzing image contents, and uploading images to the cloud storage. The processing time for stream gauge data was about 1 s, which was much faster than that for 311 service requests; 311 service requests needed to be transformed and mapped as spatial points, whereas stream gauges had fixed locations. Both 311 service requests and stream gauge data had low scheduling delays close to zero, suggesting that these data were analyzed immediately after they were received. However, there was a significant scheduling delay for streaming tweets from August 27 to 30. The delay accumulated when the processing time exceeded the batch interval (i.e., 30 s). More tweets were received

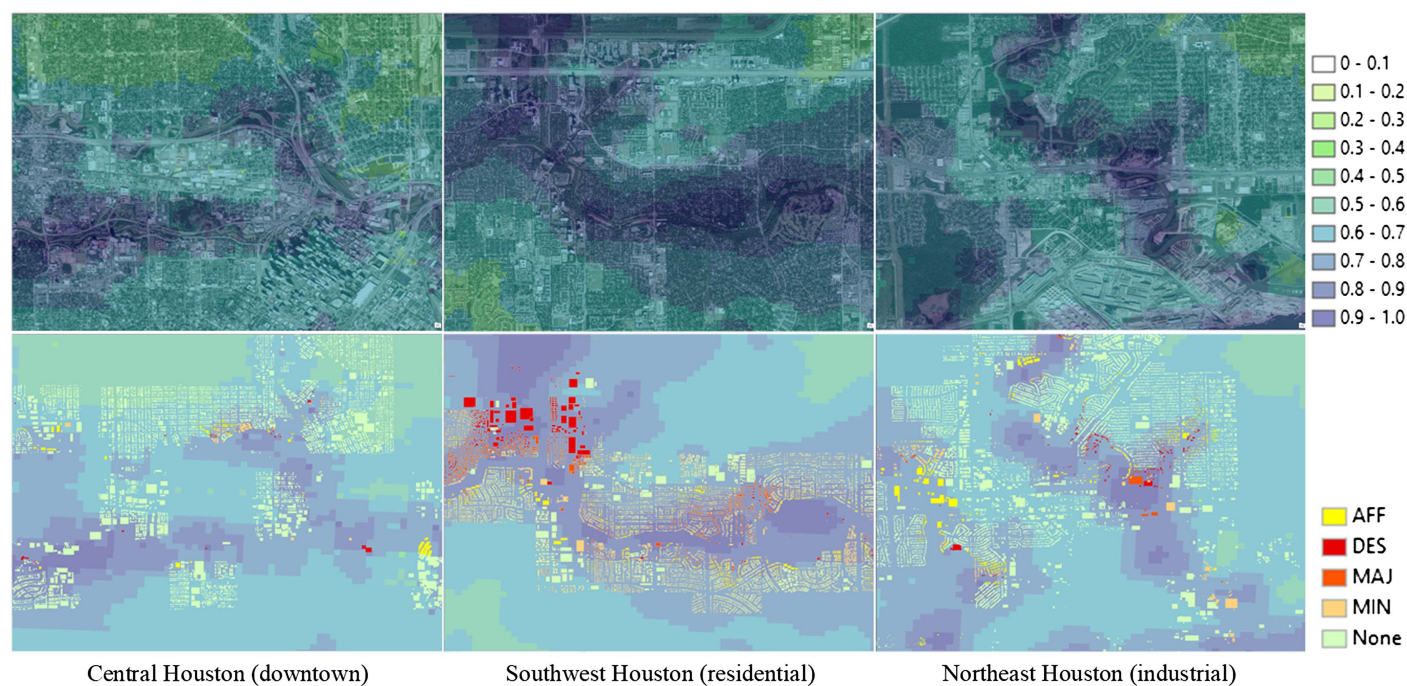


Fig. 13. Mapped flood probabilities and building damage assessment results. (Base maps by Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community.)

Table 2. Pearson correlation tests for building damage levels, and average flood probabilities and number of inundated days

Variable	Correlation coefficient	95% confidence limits		p-value
Average flood probabilities	0.22559	0.221265	0.229909	<0.0001
Number of inundated days	0.27528	0.271067	0.279483	<0.0001

on these days, and the received tweets were more likely to include texts and images describing disaster situations. A tweet maximally needed to wait 16 min for processing on August 30, 2017.

Discussion

We proposed and evaluated an end-to-end data-driven system, ADIR, for assessing disaster impacts in real time. Previous disaster impact assessments mainly focused on single data modalities and sources (Glahn et al. 2009; Roy et al. 2020; Shangguan et al. 2019; Zhang et al. 2020), which may not track holistically disaster impacts to both the built environments and population. This research contributes to the computational civil engineering with (1) domain-specific parameter tuning, system implementation, and extensive experiments with real-time multi-sourced and multimodal data streams processing; (2) an automated approach assessing fine-grained disaster impacts with integrated considerations of social vulnerability, physical built environment, and natural hazard factors; and (3) the integration human-cyber-physical system for more reliable and informative disaster situation awareness.

Some existing work has developed approaches and systems for disaster impact assessments based on social sensing data, but without evaluating the data credibility. A few researchers proposed integrating human intelligence with artificial intelligence approaches to improve system reliability (Alam et al. 2018; Fan et al. 2021), which makes the high-throughput social media posts difficult to synthesize. ADIR addresses this gap by automatically tagging

the reliability of social media posts with a density-based approach. This data reliability assessment makes the system more self-sustainable and credible for use.

Additionally, acquiring and processing streaming data in real-time can be challenging, especially when the data are from heterogeneous sources and have multiple formats. Therefore, very few previous studies implemented their developed disaster impact assessment approaches or tools within real-time settings, and the extant data-driven assessments mostly use simple analyses such as hot-spot detection to assess the spatial scale of disaster impacts and frequency analyses to assess the temporal scale of disaster impacts (Chen et al. 2016; Choi and Bae 2015; Kelly et al. 2017). ADIR addresses this gap by harnessing a series of advanced computation and geovisualization platforms, including Spark, Kafka, and ArcGIS, to realize real-time data analytics and mapping. We also use a set of third-party libraries and APIs to achieve detailed analyses, such as multimodal social media analysis to extract actionable information from individual tweet posts and spatial analysis to map citywide flood probabilities. The incorporation of these tools makes ADIR powerful enough to include more detailed analyses and flexible enough to be adapted to other scenarios.

Lastly, we applied ADIR with data collected for a historical hurricane case and running in a simulated real-time manner. The evaluation showed that ADIR quickly can pinpoint areas with the most-damaged buildings and detects and displays changing disaster situations during and immediately after disasters. The identified eyewitness reports can provide nuanced information about the disturbances to infrastructure (e.g., power outage and roadway

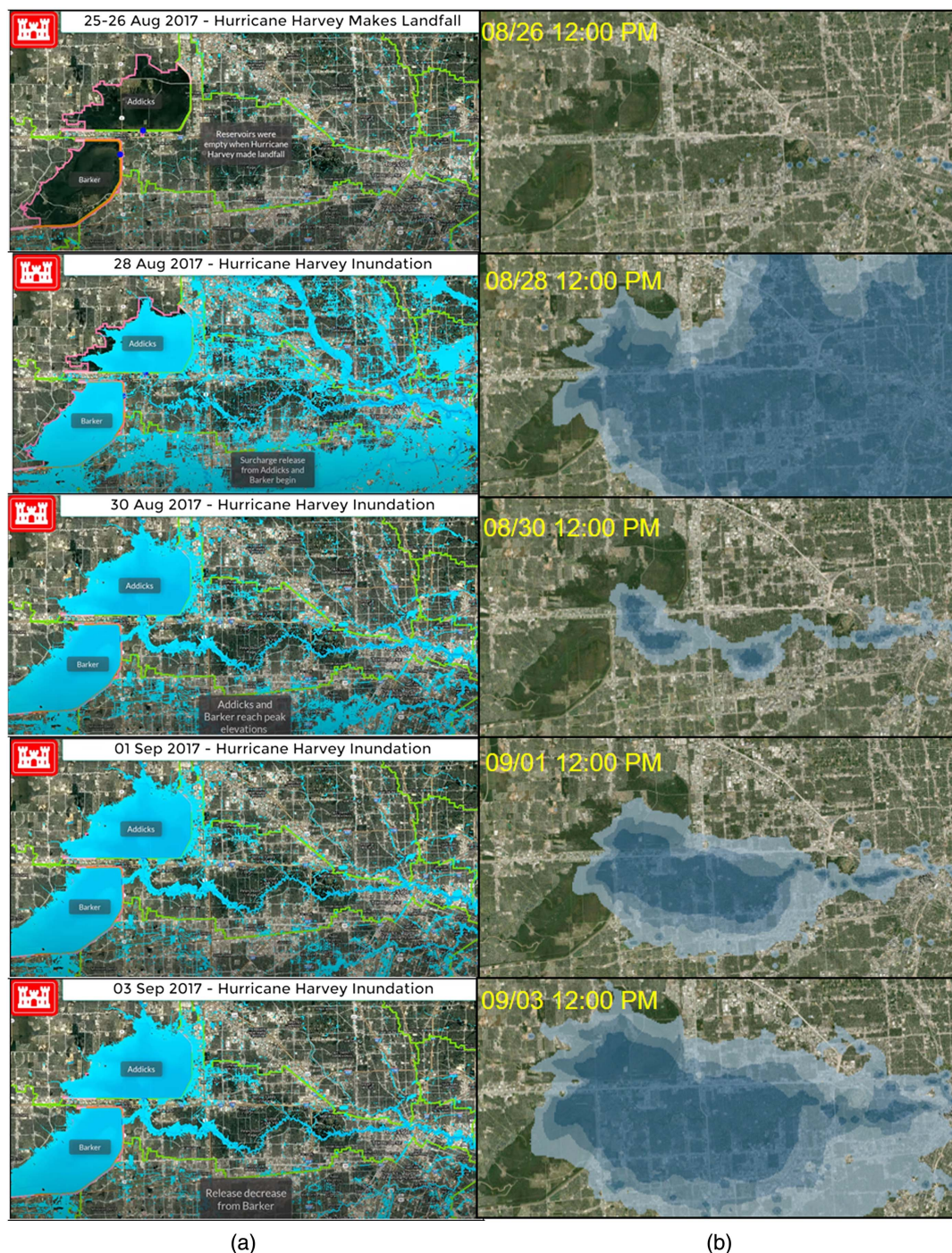


Fig. 14. Temporal change of inundation mapped by (a) USACE; and (b) flood probability map. (Base maps by Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community.)

blockage), situations of urban floods, debris/garbage removal, and the impacts on citizens' daily lives at fine spatial scales. Different stakeholders (e.g., utility companies, city managers, and emergency responders) can refer to the report details to coordinate their disaster relief activities and bring affected communities back to normal operation quickly. Although ADIR was applied with a hurricane case, it has the versatility to be applied in other adverse events (e.g., earthquakes, tornados, severe winter storms, and flooding) by using new data sources and reconfiguring the analysis processes. For example, when applied during and immediately after an earthquake, system users can replace the hazard vulnerability map (i.e., flood probabilities, in this study) with earthquake intensity

maps generated with seismometer measurements or with crowd-sourced reports (Wald et al. 2011). Analysis modules for social sensing data also should be changed to new computer vision models and text mining approaches that are trained for seismic damage detections.

This research has a few limitations that can be improved in future studies. First, we use the static SoVI indexes derived from census data to account for the social vulnerabilities. The census data may not represent the dynamic population distributions during disasters considering that people may evacuate or take shelters. Our work can benefit by integrating real-time human mobility data (Wang and Taylor 2016). Second, we use three streaming data

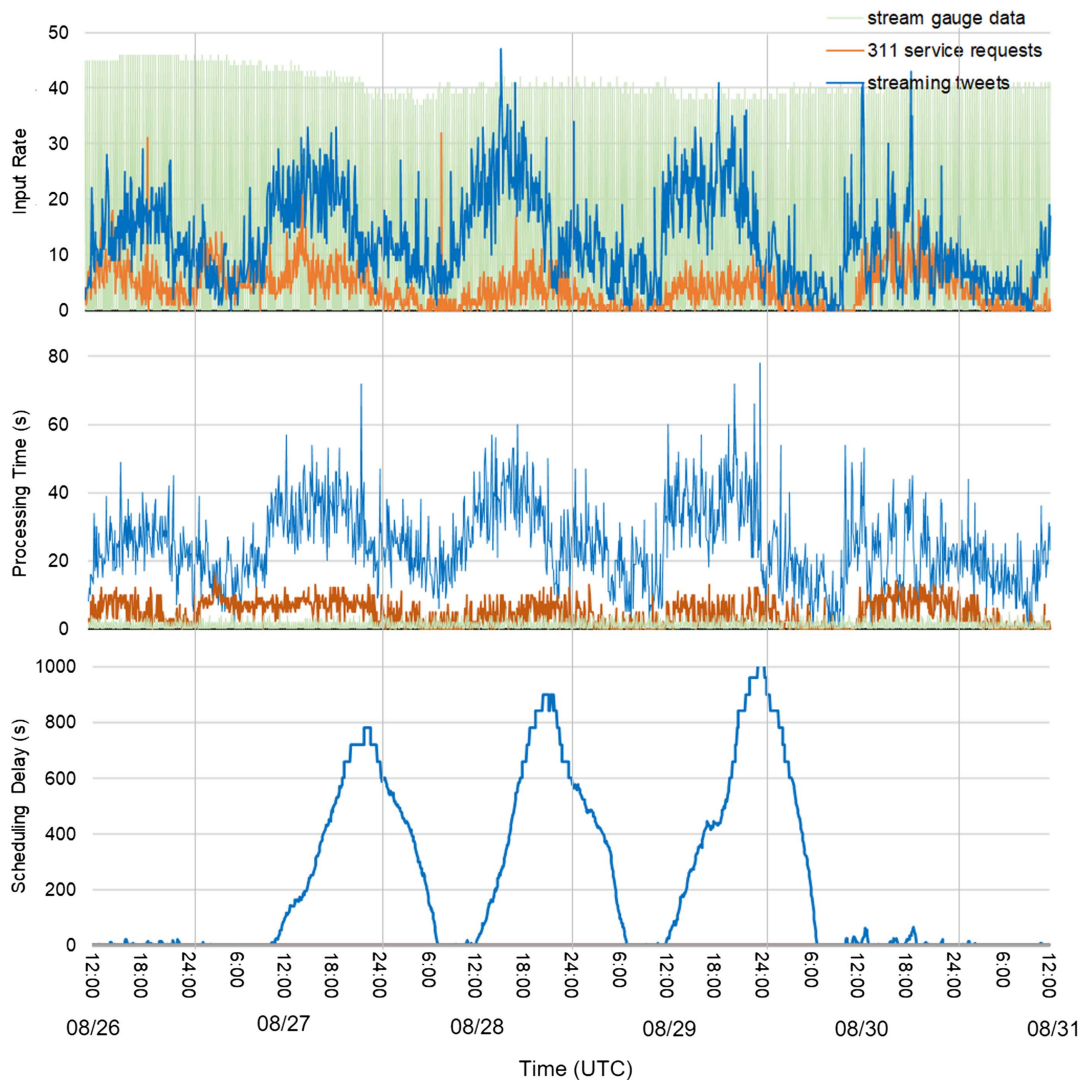


Fig. 15. Performance metrics for the streaming processing of simulated tweets, 311 service requests, and stream gauge data.

sources that generally are available for all US cities. Additionally, the flood mapping does not consider the ground flood control facilities such as reservoirs and dams. Future work may improve flood predictions by taking into consideration flood control devices. Our system easily can be extended to include other real-time data sources. For example, city departments may use infrastructure sensors to monitor the operating status of critical infrastructure systems. Hospitals and shelters also may report their real-time capacities. Our system can be more useful when including more data representing ongoing disaster situations. Third, because our system uses distributed computing platforms (i.e., Kafka and Spark) for data transmissions and analyses, a cloud infrastructure may be considered in future work to further accelerate the system and accommodate more data sources. Forth, our system focuses on assessing disaster impacts on humans and the built environments from a perspective of assisting prompt decision-making. Other types of disaster impacts, e.g., economic impacts and psychological impacts, which can be evaluated in a less time-critical manner, were not considered in this study. In addition, our system also may benefit from external validation. Video S1 in the Supplemental Materials demonstrates system outputs that can be evaluated by practitioners of disaster management and response. We will continue improving the system with their advice in future work.

Conclusion

We developed a system, ADIR, to assess the impact of disasters on humans and the built environments in real-time and evaluated its effectiveness and applicability with a historical hurricane event. The evaluation showed that ADIR is capable of delivering timely, accurate, and comprehensive disaster impact information at multiple scales and across disastrous events. The system provides a platform for different stakeholders, including infrastructure managers, service companies, public agencies, and citizens, to share disaster impact information and gain situational awareness. The diverse and dynamic information then can be used by emergency responders to adjust their decisions and operations, such as waste management, infrastructure fixation, and service restoration. The system also may encourage the wider community to participate in disaster relief assistance, for example, through contributing accurate eyewitness data and taking part in voluntary activities. These operations and activities accelerate the restoration of community disruptions and reduce physical and societal losses from disastrous events, and thus enhance disaster resilience concerning resourcefulness and rapidity. This research also is innovative in integrating real-time computing, big data, and geospatial analytics to address an urban problem. Consequently, the emergence of computer-aided

tools such as ADIR relieves the stress posed to the growing population and civil infrastructures in disaster-vulnerable locations such as coastal lines and flood plains.

Data Availability Statement

All the data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

This material is based upon work supported by the early-career faculty start-up fund, graduate research assistantships at the University of Florida, National Science Foundation, under Grant Nos. 2028012 and 1951816. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation and University of Florida.

Supplemental Materials

Video S1 is available online in the ASCE Library (www.ascelibrary.org).

References

- Aerts, J. C. J. H., W. J. Botzen, K. C. Clarke, S. L. Cutter, J. W. Hall, B. Merz, E. Michel-Kerjan, J. Mysiak, S. Surminski, and H. Kunreuther. 2018. "Integrating human behaviour dynamics into flood disaster risk assessment." *Nat. Clim. Change* 8 (3): 193–199. <https://doi.org/10.1038/s41558-018-0085-1>.
- Agarwal, N., and Y. Yiliyasi. 2010. "Information quality challenges in social media." In *Proc., 15th Int. Conf. on Information Quality 2010*. Cambridge, MA: M.I.T. Information Quality Program.
- Alam, F., F. Ofli, and M. Imran. 2018. "Processing social media images by combining human and machine computing during crises." *Int. J. Hum.-Comput. Interact.* 34 (4): 311–327. <https://doi.org/10.1080/10447318.2018.1427831>.
- Avvenuti, M., S. Cresci, F. Del Vigna, T. Fagni, and M. Tesconi. 2018. "CrisMap: A big data crisis mapping system based on damage detection and geoparsing." *Inf. Syst. Front.* 20 (5): 993–1011. <https://doi.org/10.1007/s10796-018-9833-z>.
- Blake, E. S., and D. A. Zelinsky. 2018. "Hurricane Harvey." Accessed September 29, 2020. https://www.nhc.noaa.gov/data/tcr/AL092017_Harvey.pdf.
- Bruneau, M., S. E. Chang, R. T. Eguchi, G. C. Lee, T. D. O'Rourke, A. M. Reinhorn, M. Shinozuka, K. Tierney, W. A. Wallace, and D. von Winterfeldt. 2003. "A framework to quantitatively assess and enhance the seismic resilience of communities." *Earthquake Spectra* 19 (4): 733–752. <https://doi.org/10.1193/1.1623497>.
- Burton, C. G. 2010. "Social vulnerability and hurricane impact modeling." *Nat. Hazards Rev.* 11 (2): 58–68. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2010\)11:2\(58\)](https://doi.org/10.1061/(ASCE)1527-6988(2010)11:2(58)).
- Careem, M., C. De Silva, R. De Silva, L. Raschid, and S. Weerawarana. 2006. "Sahana: Overview of a disaster management system." In *Proc., 2nd Int. Conf. on Information and Automation, ICIA 2006*, 361–366. New York: IEEE.
- Chen, S.-C., M. Chen, N. Zhao, S. Hamid, K. Chatterjee, and M. Armella. 2009. "Florida public hurricane loss model: Research in multi-disciplinary system integration assisting government policy making." *Gov. Inf. Q.* 26 (2): 285–294. <https://doi.org/10.1016/j.giq.2008.12.004>.
- Chen, X., G. Elmes, X. Ye, and J. Chang. 2016. "Implementing a real-time Twitter-based system for resource dispatch in disaster management." *GeoJournal* 81 (6): 863–873. <https://doi.org/10.1007/s10708-016-9745-8>.
- Choe, Y., V. Staneva, T. Schneider, A. Escay, C. Haberland, and S. Chen. 2018. "Benchmark dataset for automatic damaged building detection from post-hurricane remotely sensed imagery." Accessed September 28, 2020. <https://iee-dataport.org/open-access/benchmark-dataset-automatic-damaged-building-detection-post-hurricane-remotely-sensed>.
- Choi, S., and B. Bae. 2015. "The real-time monitoring system of social big data for disaster management." In Vol. 330 of *Computer science and its applications. Lecture notes in electrical engineering*, edited by J. Park, I. Stojmenovic, H. Jeong, and G. Yi, 809–815. Berlin: Springer. https://doi.org/10.1007/978-3-662-45402-2_115.
- Coronese, M., F. Lamperti, K. Keller, F. Chiaromonte, and A. Roventini. 2019. "Evidence for sharp increase in the economic damages of extreme natural disasters." *Proc. Natl. Acad. Sci. U.S.A.* 116 (43): 21450–21455. <https://doi.org/10.1073/pnas.1907826116>.
- Cutter, S. L., L. Barnes, M. Berry, C. Burton, E. Evans, E. Tate, and J. Webb. 2008. "A place-based model for understanding community resilience to natural disasters." *Global Environ. Change* 18 (4): 598–606. <https://doi.org/10.1016/j.gloenvcha.2008.07.013>.
- Dottori, F., et al. 2018. "Increased human and economic losses from river flooding with anthropogenic warming." *Nat. Clim. Change* 8 (9): 781–786. <https://doi.org/10.1038/s41558-018-0257-z>.
- Eid, M. S., and I. H. El-Adaway. 2017. "Integrating the social vulnerability of host communities and the objective functions of associated stakeholders during disaster recovery processes using agent-based modeling." *J. Comput. Civ. Eng.* 31 (5): 04017030. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000680](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000680).
- Esri. 2020a. "Hosted data." Accessed July 30, 2020. <https://developers.arcgis.com/features/hosted-data/>.
- Esri. 2020b. "What is a dashboard." Accessed July 30, 2020. <https://doc.arcgis.com/en/dashboards/get-started/what-is-a-dashboard.htm>.
- Fan, C., C. Zhang, A. Yahja, and A. Mostafavi. 2021. "Disaster City Digital Twin: A vision for integrating artificial and human intelligence for disaster management." *Int. J. Inf. Manage.* 56 (Feb): 102049. <https://doi.org/10.1016/j.ijinfomgt.2019.102049>.
- FEMA. 2020. "Hazus." Accessed July 28, 2020. <https://www.fema.gov/hazus>.
- Flanagan, B. E., E. W. Gregory, E. J. Hallisey, J. L. Heitgerd, and B. Lewis. 2011. "A social vulnerability index for disaster management." *J. Homeland Secur. Emergency Manage.* 8 (1): 1–22. <https://doi.org/10.2202/1547-7355.1792>.
- Gao, H., G. Barbier, and R. Goolsby. 2011. "Harnessing the crowdsourcing power of social media for disaster relief." *IEEE Intell. Syst.* 26 (3): 10–14. <https://doi.org/10.1109/MIS.2011.52>.
- Glahn, B., A. Taylor, N. Kurkowski, and W. A. Shaffer. 2009. "The role of the SLOSH model in National Weather Service storm surge forecasting." Accessed September 29, 2020. http://www.nws.noaa.govmdl/pubs/Documents/Papers/Role_of_SLOSH_Model_August2009.pdf.
- Goodchild, M. F. 2007. "Citizens as sensors: the world of volunteered geography." *GeoJournal* 69 (4): 211–221. <https://doi.org/10.1007/s10708-007-9111-y>.
- Halse, S. E., R. Grace, J. Kropczynski, and A. Tapia. 2019. "Simulating real-time Twitter data from historical datasets." In *Proc., 16th Int. ISCRAM Conf.*, 780–787. Valencia, Spain: Information Systems for Crisis Response And Management.
- Hao, H., and Y. Wang. 2020. "Leveraging multimodal social media data for rapid disaster damage assessment." *Int. J. Disaster Risk Reduct.* 51 (Dec): 101760. <https://doi.org/10.1016/j.ijdrr.2020.101760>.
- Huang, Q., G. Cervone, and G. Zhang. 2017. "A cloud-enabled automatic disaster analysis system of multi-sourced data streams: An example synthesizing social media, remote sensing and Wikipedia data." *Comput. Environ. Urban Syst.* 66 (Nov): 23–37. <https://doi.org/10.1016/j.compenvurbsys.2017.06.004>.
- Imran, M., S. Elbassuoni, C. Castillo, F. Diaz, and P. Meier. 2013. "Extracting information nuggets from disaster-related messages in social media." In *Proc., 10th Int. ISCRAM Conf.*, 791–801. Baden-Baden, Germany: Information Systems for Crisis Response and Management.
- Kelly, S., X. Zhang, and K. Ahmad. 2017. "Mining multimodal information on social media for increased situational awareness." In *Proc., 14th Int.*

- ISCRAM Conf., 613–622. Albi, France: Information Systems for Crisis Response and Management.
- Kircher, C. A., R. V. Whitman, and W. T. Holmes. 2006. “HAZUS earthquake loss estimation methods.” *Nat. Hazards Rev.* 7 (2): 45–59. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2006\)7:2\(45\)](https://doi.org/10.1061/(ASCE)1527-6988(2006)7:2(45)).
- Kossin, J. P., K. R. Knapp, T. L. Olander, and C. S. Velden. 2020. “Global increase in major tropical cyclone exceedance probability over the past four decades.” *Proc. Natl. Acad. Sci. U.S.A.* 117 (22): 11975–11980. <https://doi.org/10.1073/pnas.1920849117>.
- Kryvasheyev, Y., H. Chen, N. Obradovich, E. Moro, P. Van Hentenryck, J. Fowler, and M. Cebrian. 2016. “Rapid assessment of disaster damage using social media activity.” *Sci. Adv.* 2 (3): e1500779. <https://doi.org/10.1126/sciadv.1500779>.
- Li, Z., C. Wang, C. T. Emrich, and D. Guo. 2018. “A novel approach to leveraging social media for rapid flood mapping: A case study of the 2015 South Carolina floods.” *Cartography Geog. Inf. Sci.* 45 (2): 97–110. <https://doi.org/10.1080/15230406.2016.1271356>.
- McWethy, D. B., et al. 2019. “Rethinking resilience to wildfire.” *Nat. Sustainability* 2 (9): 797–804. <https://doi.org/10.1038/s41893-019-0353-8>.
- Merz, B., H. Kreibich, R. Schwarze, and A. Thieken. 2010. “Review article ‘Assessment of economic flood damage’.” *Nat. Hazards Earth Syst. Sci.* 10 (8): 1697–1724. <https://doi.org/10.5194/nhess-10-1697-2010>.
- Musa, A., O. Watanabe, H. Matsuoka, H. Hokari, T. Inoue, Y. Murashima, Y. Ohta, R. Hino, S. Koshimura, and H. Kobayashi. 2018. “Real-time tsunami inundation forecast system for tsunami disaster prevention and mitigation.” *J. Supercomputing* 74 (7): 3093–3113. <https://doi.org/10.1007/s11227-018-2363-0>.
- Nara, A., X. Yang, S. Ghanipoor Machiani, and M.-H. Tsou. 2017. “An integrated evacuation decision support system framework with social perception analysis and dynamic population estimation.” *Int. J. Disaster Risk Reduct.* 25 (Oct): 190–201. <https://doi.org/10.1016/j.ijdrr.2017.09.020>.
- NRC (National Research Council). 2006. *Community disaster resilience*. Washington, DC: National Academies Press.
- NRC (National Research Council). 2012. *Disaster resilience: A national imperative*. Washington, DC: National Academies Press.
- Pouyanfar, S., Y. Tao, H. Tian, S.-C. Chen, and M.-L. Shyu. 2019. “Multi-modal deep learning based on multiple correspondence analysis for disaster management.” *World Wide Web* 22 (5): 1893–1911. <https://doi.org/10.1007/s11280-018-0636-4>.
- Powell, M. D., et al. 2010. “Reconstruction of Hurricane Katrina’s wind fields for storm surge and wave hindcasting.” *Ocean Eng.* 37 (1): 26–36. <https://doi.org/10.1016/j.oceaneng.2009.08.014>.
- Roy, K. C., S. Hasan, and P. Mozumder. 2020. “A multilabel classification approach to identify hurricane-induced infrastructure disruptions using social media data.” *Comput.-Aided Civ. Infrastruct. Eng.* 35 (12): 1387–1402. <https://doi.org/10.1111/mice.12573>.
- Sakaki, T., M. Okazaki, and Y. Matsuo. 2010. “Earthquake shakes Twitter users: Real-time event detection by social sensors.” In *Proc., 19th Int. Conf. on World Wide Web, WWW ’10*, 851–860. New York: Association for Computing Machinery.
- Scawthorn, C., P. Flores, N. Blais, and H. Seligson. 2006a. “HAZUS-MH flood loss estimation methodology. II. Damage and loss assessment.” *Nat. Hazards Rev.* 7 (2): 72–81. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2006\)7:2\(72\)](https://doi.org/10.1061/(ASCE)1527-6988(2006)7:2(72)).
- Scawthorn, C., N. Blais, H. Seligson, E. Tate, E. Mifflin, W. Thomas, J. Murphy, and C. Jones. 2006b. “HAZUS-MH flood loss estimation methodology. I: Overview and flood hazard characterization.” *Nat. Hazards Rev.* 7 (2): 60–71. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2006\)7:2\(60\)](https://doi.org/10.1061/(ASCE)1527-6988(2006)7:2(60)).
- Schneider, P. J., and B. A. Schauer. 2006. “HAZUS—Its development and its future.” *Nat. Hazards Rev.* 7 (2): 40–44. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2006\)7:2\(40\)](https://doi.org/10.1061/(ASCE)1527-6988(2006)7:2(40)).
- Shangguan, B., P. Yue, Z. Yan, and D. Tapete. 2019. “A stream computing approach for live environmental models using a spatial data infrastructure with a waterlogging model case study.” *Environ. Modell. Software* 119 (Sep): 182–196. <https://doi.org/10.1016/j.envsoft.2019.06.009>.
- Smith, A. B. 2019. “2018’s billion dollar disasters in context.” Accessed July 30, 2020. <https://www.climate.gov/news-features/blogs/beyond-data/2018s-billion-dollar-disasters-context>.
- Tsou, M.-H., C.-T. Jung, C. Allen, J.-A. Yang, S. Y. Han, B. H. Spitzberg, and J. Dozier. 2017. “Building a real-time geo-targeted event observation (Geo) viewer for disaster management and situation awareness.” In *Lecture notes in geoinformation and cartography*, 85–98. New York: Springer.
- Twitter. 2020. “Consuming streaming data.” Accessed July 30, 2020. <https://developer.twitter.com/en/docs/tutorials/consuming-streaming-data>.
- USACE. 2017. “USACE announced plans for Addicks and Barker Releases at local press conference.” Accessed August 26, 2020. <https://www.swg.usace.army.mil/Media/News-Releases/Article/1299979/usace-announced-plans-for-addicks-and-barker-releases-at-local-press-conference/>.
- Vickery, P. J., J. Lin, P. F. Skerlj, L. A. Twisdale, Jr., and K. Huang. 2006a. “HAZUS-MH hurricane model methodology. I: Hurricane hazard, terrain, and wind load modeling.” *Nat. Hazards Rev.* 7 (2): 82–93. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2006\)7:2\(82\)](https://doi.org/10.1061/(ASCE)1527-6988(2006)7:2(82)).
- Vickery, P. J., P. F. Skerlj, J. Lin, L. A. Twisdale, Jr., M. A. Young, and F. M. Lavelle. 2006b. “HAZUS-MH hurricane model methodology. II: Damage and loss estimation.” *Nat. Hazards Rev.* 7 (2): 94–103. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2006\)7:2\(94\)](https://doi.org/10.1061/(ASCE)1527-6988(2006)7:2(94)).
- Wald, D. J., V. Quitoriano, B. Worden, M. Hopper, and J. W. Dewey. 2011. “USGS ‘Did You Feel It?’ internet-based macroseismic intensity maps.” *Ann. Geophys.* 54 (6): 688–707. <https://doi.org/10.4401/ag-5354>.
- Wang, Q., and J. E. Taylor. 2016. “Process map for urban-human mobility and civil infrastructure data collection using geosocial networking platforms.” *J. Comput. Civ. Eng.* 30 (2): 04015004. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000469](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000469).
- Wang, Y., and J. E. Taylor. 2019. “DUET: Data-driven approach based on latent Dirichlet allocation topic modeling.” *J. Comput. Civil Eng.* 33 (3): 04019023. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000819](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000819).
- Wang, Y., J. E. Taylor, and M. J. Garvin. 2020. “Measuring resilience of human-spatial systems to disasters: Framework combining spatial-network analysis and Fisher information.” *J. Manage. Eng.* 36 (4): 04020019. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000782](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000782).
- Wang, Z., H. T. Vo, M. Salehi, L. I. Rusu, C. Reeves, and A. Phan. 2017. “A large-scale spatio-temporal data analytics system for wildfire risk management.” In *Proc., GeoRich 2017—4th Int. ACM Workshop on Managing and Mining Enriched Geo-Spatial Data, in Conjunction with SIGMOD 2017*, 19–24. New York: Association for Computing Machinery.
- Wing, O. E. J., N. Pinter, P. D. Bates, and C. Kousky. 2020. “New insights into US flood vulnerability revealed from flood insurance big data.” *Nat. Commun.* 11 (1): 1–10. <https://doi.org/10.1038/s41467-020-15264-2>.
- Yao, F., and Y. Wang. 2019. “Tracking urban geo-topics based on dynamic topic model.” *Comput. Environ. Urban Syst.* 79 (Jan): 101419. <https://doi.org/10.1016/j.compenurbysys.2019.101419>.
- Yao, F., and Y. Wang. 2020. “Towards resilient and smart cities: A real-time urban analytical and geo-visual system for social media streaming data.” *Sustainable Cities Soc.* 63 (Dec): 102448. <https://doi.org/10.1016/j.scs.2020.102448>.
- Zhang, C., W. Yao, Y. Yang, R. Huang, and A. Mostafavi. 2020. “Semi-automated social media analytics for sensing societal impacts due to community disruptions during disasters.” *Comput.-Aided Civ. Infrastruct. Eng.* 35 (12): 1331–1348. <https://doi.org/10.1111/mice.12576>.
- Zhong, X., M. Duckham, D. Chong, and K. Tolhurst. 2016. “Real-time estimation of wildfire perimeters from curated crowdsourcing.” *Sci. Rep.* 6 (1): 1–10. <https://doi.org/10.1038/srep24206>.
- Zhu, R., J. Lin, B. Becerik-Gerber, and N. Li. 2020. “Human-building-emergency interactions and their impact on emergency response performance: A review of the state of the art.” *Saf. Sci.* 127 (Jul): 104691. <https://doi.org/10.1016/j.ssci.2020.104691>.
- Zou, L., N. S. N. Lam, S. Shams, H. Cai, M. A. Meyer, S. Yang, K. Lee, S.-J. Park, and M. A. Reams. 2019. “Social and geographical disparities in Twitter use during Hurricane Harvey.” *Int. J. Digital Earth* 12 (11): 1300–1318. <https://doi.org/10.1080/17538947.2018.1545878>.