Annals of the American Association of Geographers

ISSN: 2469-4452 (Print) 2469-4460 (Online) Journal homepage: http://www.tandfonline.com/loi/raag21

Mining Twitter Data for Improved Understanding of Disaster Resilience

Lei Zou, Nina S. N. Lam, Heng Cai & Yi Qiang

To cite this article: Lei Zou, Nina S. N. Lam, Heng Cai & Yi Qiang (2018): Mining Twitter Data for Improved Understanding of Disaster Resilience, Annals of the American Association of Geographers, DOI: 10.1080/24694452.2017.1421897

To link to this article: https://doi.org/10.1080/24694452.2017.1421897

Published online: 14 Mar 2018.

Mining Twitter Data for Improved Understanding of Disaster Resilience

Lei Zou ,* Nina S. N. Lam ,* Heng Cai ,* and Yi Qiang†

*Department of Environmental Sciences, Louisiana State University
†Department of Geography, University of Hawaii–Manoa

Coastal communities faced with multiple hazards have shown uneven responses and behaviors. These responses and behaviors could be better understood by analyzing real-time social media data through categorizing them into the three phases of the emergency management: preparedness, response, and recovery. This study analyzes the spatial–temporal patterns of Twitter activities during Hurricane Sandy, which struck the U.S. Northeast on 29 October 2012. The study area includes 126 counties affected by Hurricane Sandy. The objectives are threefold: (1) to derive a set of common indexes from Twitter data so that they can be used for emergency management and resilience analysis; (2) to examine whether there are significant geographical and social disparities in disaster-related Twitter use; and (3) to test whether Twitter data can improve postdisaster damage estimation. Three corresponding hypotheses were tested. Results show that common indexes derived from Twitter data, including ratio, normalized ratio, and sentiment, could enable comparison across regions and events and should be documented. Social and geographical disparities in Twitter use existed in the Hurricane Sandy event, with higher disaster-related Twitter use communities generally being communities of higher socioeconomic status. Finally, adding Twitter indexes into a damage estimation model improved the adjusted R^2 from 0.46 to 0.56, indicating that social media data could help improve postdisaster damage estimation, but other environmental and socioeconomic variables influencing the capacity to reducing damage might need to be included. The knowledge gained from this study could provide valuable insights into strategies for utilizing social media data to increase resilience to disasters. *Key Words: geographical and social disparities in disaster resilience, Hurricane Sandy, social media, Twitter use.*

面临多重灾害的沿海社区，展现出不一致的回应与行为。这些回应与行为，能够透过将即时社交媒体的数据分类成以下三大紧急管理阶段并进行分析，以获得更佳的理解：准备、回应，以及回復阶段。本研究分析珊蒂飓风在 2012 年十月二十九日侵袭美国东北部期间，推特活动的时空模式。本研究范围包含一百二十六个受到珊蒂飓风影响的郡县。本分析的三大目标如下：(1) 从推特数据衍生出一组普遍的指标，因此能够将其运用于紧急管理和回復力分析；(2) 检视关乎灾害的推特使用中，是否呈现显着的地理与社会不均；以及 (3) 检验推特数据是否能够促进灾后损失评估。本研究验证三大相关假说。分析结果显示，从推特数据衍生出的共同指标，包括比值、标准化比值和情绪，让跨越区域和事件进行比较成为可能，因而应进行纪录。推特使用的社会与地理不均，存在于珊蒂飓风事件之中，其中与灾害相关的推特使用较高的社群，一般而言属于较高社经地位的社群。最后，将推特指标增添至损失评估模型，使得调整过后的 R2 从 0.46 改进至 0.56，显示社交媒体数据有助于改进灾后损失评估，但或许仍须纳入影响降低损失能力的其他环境与社经变因。此一研究所获得的知识，能够对运用社交媒体数据来增进灾害回復力的策略，提供宝贵的洞见。*关键词：灾害回復力中的地理与社会不均，珊蒂飓风，社交媒体，推特使用。*

Las comunidades costeras que se abocan a enfrentar múltiples amenazas muestran respuestas y conductas desiguales. Tales respuestas y conductas podrían entenderse mejor analizando los datos mediáticos sociales de tiempo real por medio de su categorización en tres fases de manejo de la emergencia: preparación, respuesta y recuperación. Este estudio analiza los patrones espacio–temporales de las actividades de Twitter durante el Huracán Sandy, que golpeó el Nordeste de los EE.UU. el 29 de octubre del 2012. El área de estudio comprende 126 condados que fueron afectados por aquel Huracán. Se fijaron tres objetivos: (1) derivar un conjunto de índices comunes desde los datos de Twitter de modo que puedan usarse para el análisis de manejo de emergencia y resiliencia; (2) examinar si existen disparidades geográficas y sociales significativas en el uso de Twitter en relación con el desastre; y (3) examinar si los datos de Twitter pueden mejorar los estimativos de daños calculados después del desastre. Se pusieron a prueba tres hipótesis relacionadas con los objetivos. Los resultados indican que los índices comunes derivados desde los datos de Twitter, incluyendo la ratio, la ratio normalizada y el sentimiento, podrían facilitar la comparación a través de regiones y eventos, lo cual debe documentarse. Las disparidades sociales y geográficas en el uso de Twitter existían durante el evento del Huracán Sandy, con las comunidades de mayor uso de Twitter para cosas relacionadas con el desastre generalmente correspondientes a

Annals of the American Association of Geographers, 0(0) 2018, pp. 1–20 © 2018 by American Association of Geographers
Initial submission, July 2017; revised submission, October 2017; final acceptance, November 2017
Published by Taylor & Francis, LLC.

las comunidades del estatus socioeconómico más alto. Por último, al agregar los índices de Twitter a un modelo de estimación de los daños se mejoró el R2 ajustado de 0.46 a 0.56, indicando que los datos de los medios sociales podrían ayudar a mejorar el estimativo posterior al desastre de los daños causados, aunque otras variables ambientales y socioeconómicas que influyen la capacidad de reducir el daño podrían demandar inclusión. El conocimiento ganado con este estudio podría generar valiosas perspectivas en la estrategia de utilizar datos de los medios sociales para aumentar la resiliencia a los desastres. *Palabras clave: disparidades geográficas y sociales en la resiliencia de desastres, Huracán Sandy, medios sociales, uso de Twitter.*

Disaster resilience broadly describes the ability of a community to bounce back from disaster impacts by generating resilience building activities through the four phases of emergency management: preparedness, response, recovery, and mitigation (Federal Emergency Management Agency [FEMA] 2006). Disaster resilience is a major societal challenge and has been a subject of intense research by many researchers from multiple disciplines, all aiming at building capacity to enhance resilience (National Research Council 2012).

Given the same type of hazard and threat level, the impact of a hazard on a community varies across regions with different socioeconomic and environmental characteristics (Cutter, Boruff, and Shirley 2003, 2008; Adger et al. 2005; Reams, Lam, and Baker 2012; Lam et al. 2015; K. Li et al. 2015; Cai et al. 2016; Lam et al. 2016). These social and geographical disparities of disaster resilience are major obstacles to building long-term resilience for communities. Hence, a better understanding of the patterns and consequences of social and geographical disparities in disaster resilience is necessary for reducing the disparities, which in turn should help reduce damage and increase resilience.

In the past few decades, disaster studies have been relying on traditional sociodemographic data collected at regular time intervals such as data from census and health agencies to quantify disaster resilience. A major shortcoming of the traditional approach is that data describing communities' preparedness, response, and recovery behaviors are generally not available through traditional databases. Social media, as an emerging data source, could provide an innovative approach to observing human behaviors under emergencies in real time. During an emergency, users' posting habits and feelings of the event are expected to be related to the level of threat from the hazard, damage sustained, users' individual characteristics, and local social and environmental conditions (Earle 2010; Kent and Capello 2013; Guan and Chen 2014; Kryvasheyeu et al. 2016). Spatial–temporal surveillance through social media platforms has the potential to reveal the disparities of each community's preparedness, response, and recovery behaviors during hazardous events. Information extracted from social media data could be used to estimate potential damage, reduce disparities, and enhance resilience.

Previous investigations, however, have pointed out the difficulties in deriving scientific results through social media due to its inherent data issues, including false information, lack of validation, and biased demographics of users (Mislove et al. 2011). In addition, social media data, as a type of big data, pose additional methodological and computational challenges. Several research questions need to be addressed to properly apply social media data for disaster resilience studies. First, how can we extract useful information and common indexes to represent the spatiotemporal patterns of social media activities during disasters? Second, are there any geographical and social disparities in social media use and content, and what are the patterns in different phases of disasters? Third, how should we use information extracted from social media to understand and enhance disaster resilience?

In answering these three research questions, this article examines the Twitter reactions to Hurricane Sandy, which hit the northeastern United States on 29 October 2012. The objectives of this study are threefold: (1) to derive a set of common indexes from Twitter data for emergency management and resilience analysis across regions and events; (2) to examine whether geographical and social disparities in Twitter use exist across the preparedness, response, and recovery phases of emergency management; and (3) to examine the utility of Twitter information in improving the postdisaster damage estimation. Three corresponding research hypotheses are tested in this study: (1) communities located near the disaster are expected to have higher levels of Twitter activity before the disaster; (2) given the same level of threat, communities with higher socioeconomic conditions are expected to have higher disaster-related Twitter activities than communities with lower socioeconomic conditions; and (3) communities that have suffered more damage are expected to have higher levels of Twitter activity after the disaster and, in turn, Twitter information can be used for damage estimation. As explained later, the rejection or acceptance of these hypotheses has implications to the resilience of the communities. Findings from this study will shed light on the pros and cons of using social media in disaster research. Knowledge gained from this study will provide valuable insights into strategies of using social media data to reduce disparities in resilience and build long-term resilience to disasters.

The article proceeds as follows. We first provide a brief review of previous investigations on the use of social media data during emergencies and an introduction to Hurricane Sandy. We then detail the data and methodology through which the Twitter data are collected, processed, and analyzed. Following that, we document the results and evaluate the hypotheses and then discuss the methodological uncertainties and limitations of the study and provide suggestions for future research. We conclude with a summary of the findings and their implications.

Background

Social Media and Disaster Resilience

With the advent of the big data era, social media data have provided new opportunities for studying the world (Merchant, Elmer, and Lurie 2011; Tsou and Leitner 2013). Through social media, numerous users can exchange information at any time in any place. With mobile devices enabled with the Global Positioning System (GPS), every human being can act as an intelligent sensor who collects information about the environment and shares feelings on social media in real time and at different locations. With the timestamps and geotags, social media data can be used to uncover the spatiotemporal variation of different processes and offer unique insights into the socioeconomic conditions of human communities.

In recent years, several studies have attempted to investigate social media activities during disasters. Earle and others explored the capacity of Twitter in reporting an earthquake and estimating its impacts (Earle et al. 2010). Their results demonstrate that Twitter activities could help identify affected areas faster than traditional monitoring methods. A demographic study of online sentiment during Hurricane Irene finds that the level of concern in the days leading up to the hurricane's arrival is dependent on regions and genders (Mandel et al. 2012). Communities that had less hurricane experience before or received more early warning information expressed higher levels of concern on Twitter. Females were more likely to express their concerns on social media than males. Kent and Capello (2013) analyzed the spatial patterns of user-generated contents on Instagram, Twitter, Flickr, and Picasa during the Horsethief Canyon fire of 2012. Their results indicate that communities that were closer to the wildfire location and had more young population, higher population density, and higher situational awareness tended to produce more useful information on social media sites.

Popularity of social media use in disasters has increased its potential as a new data source to understand disaster resilience. Dufty (2012) proposed that social media could help build community resilience to disasters through risk reduction, emergency management, and posthazard development. Patton and others discussed the idea of visualizing community resilience metrics from Twitter by decomposing the time series data into subseries components, standing for different aspects of resilience metrics, and quantitatively plotting each component's temporal trend (Patton, Steed, and Stahl 2013). Instead of passively sensing user-generated information on social media, a study of the 2011 Genoa Floods initiated a Facebook page to organize rescue activities and reconstruction to enhance community resilience (Rizza and Pereira 2014).

Creating a multiscale spatiotemporal analysis framework for social media data and using this framework for disaster studies remains a major challenge in social media and big data research, however (Tsou 2015). Many of the preceding examples have suggested the difficulties in realizing quantitative analysis through social media. They point to the necessity to account for the complexity embedded in the data when using social media data for quantitative analysis (Earle et al. 2010; Shelton et al. 2014), due to the many inherent issues of social media data, such as false information, malicious use, lack of validation, and biased demographic composition of users. Using social media data alone to draw scientific conclusions is still questionable. They will need to be integrated with traditional data to provide valid analysis results (Lindsay 2011; Li, Goodchild, and Xu 2013).

As the amount of volunteered geographic information (VGI) created by social media grows, a framework for analyzing social media responses to natural hazards is needed to help identify common indexes that are generalizable for analysis across space, time, and events. More research is necessary to answer the question of how social media data can be synthesized with other data sources to produce knowledge that can be used to construct concrete plans and strategies to strengthen resilience.

Hurricane Sandy

As the most devastating natural disaster event to hit the United States since Hurricane Katrina in 2005, Hurricane Sandy formed south of Jamaica and was named as a tropical storm on 22 October 2012. It developed into a Category 1 hurricane on 23 October and made landfall near Kingston, Jamaica, on 24 October. Then it grew to a Category 2 hurricane and swept Cuba and Haiti on 25 October (Cable News Network 2012). Early on 29 October 2012, Sandy curved northwest, and then moved ashore near Brigantine, New Jersey. It went through New Jersey, Delaware, and Pennsylvania and ended in Ohio (Figure 1). Sandy was the most destructive hurricane of the 2012 Atlantic hurricane season, sweeping across eight countries. According to the historical records of tropical cyclones from the National Oceanic and Atmospheric Administration (NOAA 2014), Hurricane Sandy caused $71.4 billion in economic damage in the United States, and it is the second costliest hurricane in the U.S. history. The most severe damage took place in New Jersey and New York, caused by strong winds and storm surge; Pennsylvania, Maryland, Delaware, and Connecticut also suffered hurricane impacts at varying levels, such as widespread power outages and paralysis of public transportation services. Due to the immense damage caused by Hurricane Sandy, seven severely affected states received large amounts of assistance from FEMA for postdisaster recovery and reconstruction.

Hurricane Sandy provides a compelling case study of social media activities for several reasons. First, Hurricane Sandy affected numerous people residing in different regions (e.g., United States, Jamaica, and Haiti); therefore, the Twitter activities at multiple geographical scales could be observed and compared. Second, there was considerable media and political attention surrounding Hurricane Sandy, providing abundant data available for analysis. Third, early warning and postdisaster recovery information regarding this event has been published and disseminated through various media sources so that responses on social media could be evaluated over time.

Therefore, analyzing Twitter activity during Hurricane Sandy attracts much interest from researchers and practitioners. Through visualizing the national spatial pattern of Sandy-related tweets during Hurricane Sandy, Shelton et al. (2014) suggested that the more distant the area to the hurricane track, the fewer disaster-related tweets were generated (see also

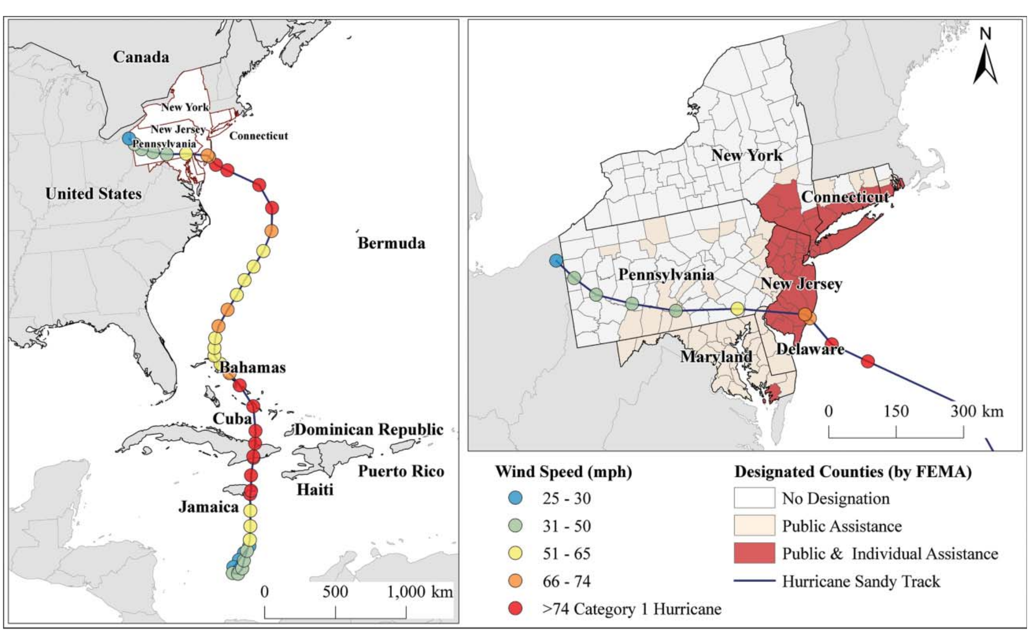

Figure 1. Tracks and affected areas of Hurricane Sandy. (Color figure available online.)

Kryvasheyeu et al. 2015). Wang and Taylor (2014) used Twitter data to analyze the human mobility pattern during Hurricane Sandy. They concluded that human mobility in New York City was significantly affected in the first three days after the landfall of Hurricane Sandy and then returned to steady states afterward. Guan and Chen (2014) and Kryvasheyeu et al. (2016) found a significant relationship between hurricane damage and Twitter activity, suggesting the potential of using social media activities for rapid damage assessment. These pioneered studies on Twitter activities during Hurricane Sandy have provided useful information, but more research on extracting common indexes from Twitter to examine the geographical and social disparities of Twitter use and their potential consequences for community resilience is needed.

Data and Methods

Twitter Data Collection and Processing

Twitter was chosen as the social media data source for our study because it has the advantages of being GPS-enabled, real-time publishing and having a wide audience. Twitter is a service that allows users to send and receive up to 140-character text messages, referred to as tweets, through any Internet-enabled device (Earle et al. 2010). Originally, Twitter was designed as a platform to help people stay in touch by publishing brief updates on their activities. Today Twitter has transformed into a new kind of social network for people to discuss "what is happening" (Kirilenko and Stepchenkova 2014). As one of the most popular social networking sites, with more than 300 million unique monthly visitors in 2012, Twitter makes it possible to obtain large amounts of information for any processes.

In this research, Twitter data were accessed from an online library called Internet Archive (see https://archive.org/details/twitterstream), which provides randomly collected tweets since 2011. The data size in Internet Archive is nearly 1 percent of the full Twitter database, which was equal to about 4 million tweets per day in 2012. In this archive, tweets are encoded in JavaScript Object Notation (JSON) format and represented as a collection of name–value pairs. Each record contains tweet information such as time of tweet, text content, coordinates and places, user profile (i.e., user ID and address), and follower and following status.

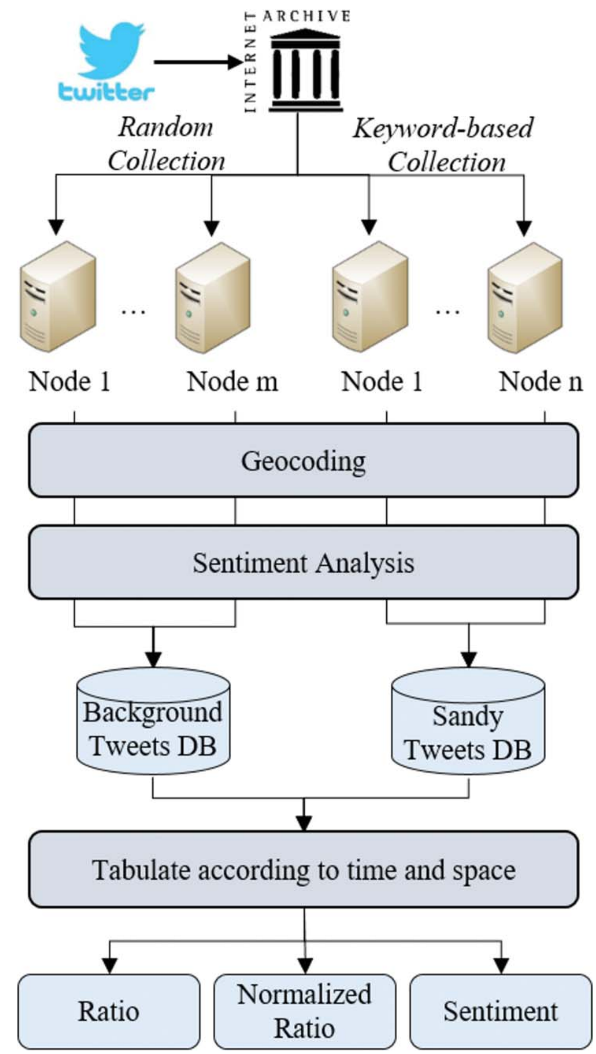

Figure 2. Methodology of Twitter data collection and processing. (Color figure available online.)

Twitter data from 1 October to 30 November 2012 were downloaded from this library, and five attributes were used for subsequent analysis: time when the tweet was created, text content, coordinates, place, and address in the user profile. Time of tweets was used to tabulate the Twitter data into the preparedness, response, and recovery phases. Text contents were used for sentiment analysis, and the latter three elements were used to determine the location of each tweet.

Figure 2 shows the procedure of Twitter data collection and processing. We collected the tweets for three weeks starting from 23 October 2012 after the tropical depression was named Sandy, which was one week before the landfall, to 12 November, which was two weeks after Hurricane Sandy landed in the United States. The first step of processing was to identify whether a tweet was related to Hurricane Sandy. We selected the four most relevant key words based on previous investigations: *hurricane*, *Sandy*, *flood*, and *storm* (Shelton et al. 2014). Although incorporating more key words would result in a larger collection of disaster-related tweets, a shorter list of the most relevant key words is efficient in removing noise messages prior to the event (Kryvasheyeu et al. 2015).

One issue might arise when someone's user name contains one of the key words and was mentioned in the tweet content. For instance, Sandy is a popular user name on Twitter. To avoid the false classification, we defined a @filter to exclude those cases. If the tweet content contains one of the key words but has an @ in front of it, such as @sandy, it means that the key word appearing here is a user name and that this tweet should not be considered to be Hurricane Sandy related. Every tweet containing one of the key words or variations after the filtration was recognized as a disaster-related tweet and stored in the Sandy tweets database. Meanwhile, about 4 million tweets randomly collected for the same period were stored in the background database as a base layer of Twitter uses in different localities.

The second processing step was to determine the location of each tweet. The location of a tweet can be represented in different ways. It can be either an *x*–*y* coordinate pair derived from the built-in GPS in mobile devices or a place name selected by the user from a set of place names suggested by Twitter. In the latter case, the place name is usually recorded as a point of interest, a street, a neighborhood, or a city, which can be converted to a point using a geocoding tool. In the case that a tweet was not explicitly tagged with a location, the address in the user profile was used as the location of the tweet. Previous studies suggest that even though the number of tweets equipped with *x*–*y* coordinate pairs represent only about 1 to 4 percent of total tweets, approximately 55 percent of all tweets could be correctly associated with a city using Google geocoding services (Graham, Hale, and Gaffney 2014).

For cases with no *x*–*y* coordinates but with attached places or addresses in their user profiles, we performed geocoding to associate the address or the place name with geographic coordinates. Geocoding could be accomplished through several online service providers, such as Geocoder.us, Google, MapPoint, OpenStreetMap, ArcGIS, and Yahoo!. A study comparing five commonly used geocoding tools suggests that Google, MapPoint, and Yahoo! provide more accurate points and shorter error distance than other methods (Roongpiboonsopit and Karimi 2010). In this study, we employed the Google geocoding application programming interface (API) to georeference tweets according to the attached places or addresses in the user profiles in cases with no *x*–*y* coordinates.

The last processing step was sentiment analysis. Identifying sentiment expressed by users in an online social networking site can help understand the users' main concerns, panics, and psychological impacts during an emergency (Caragea et al. 2014). Previous studies have demonstrated that sentiment in Twitter reflects spatial–temporal mood variations and is indicative of damage suffered during natural disasters (Kryvasheyeu et al. 2016).

Sentiment analysis evaluates people's mood by assigning sentiment scores or levels from the tweet messages and can be classified into two categories: lexicon-based analysis and classification-based analysis (Agarwal et al. 2011). Classification-based analysis involves building classifiers from manually labeled tweets (negative, neutral, and positive) using machine learning techniques and then classifying instances through a supervised classification (Pang, Lee, and Vaithyanathan 2002). It is effective in identifying major groupings of topics on Twitter, such as classifying the sentiment of the tweets during disasters into four phases of emergency management (Yang et al. 2013). Lexicon-based analysis, on the other hand, assigns scores to words in the tweet content based on a lexicon or dictionary of words and returns a combined sentiment score to represent the synthesis sentiment status of the tweet (Taboaba et al. 2011). Compared to classification-based analysis, lexicon-based analysis provides both the category and extent of the sentiment embedded in one tweet.

This research chose a lexicon-based sentiment analysis tool called VADER (Valence Aware Dictionary for sEntiment Reasoning) to perform the sentiment analysis (Hutto and Gilbert 2014). VADER returns a sentiment score ranging from –1 (*most negative*) to 1 (*most positive*) for each input text. It is specifically attuned to sentiments expressed in social media. VADER has been found to be the best algorithm in evaluating Twitter sentiment when compared to other available sentiment analysis tools (Hutto and Gilbert 2014). In this study, each tweet in the database was assigned a sentiment score according to the tweet content through VADER.

Considering the large amount of data in this study (close to 5 million tweets), parallelized processing is necessary. We employed a cluster of forty desktops to accomplish and accelerate the Twitter data collection and processing (Figure 2). A Python script was developed to parse, classify, geocode, and analyze the sentiment of all tweets. The public Google API provides only 2,500 free geocoding services per day for each node. Through this parallelized method, the cluster of desktops can georeference 100,000 tweets per day. The results were stored as the Sandy and background tweets databases. MongoDB (see https://www.mongodb.com/) was selected as the database system for Twitter data storage and query, because its document model has been demonstrated as a suitable system for social media data management. MongoDB supports geospatial and temporal indexes and has fast speed in data reading and writing (Walther and Kaisser 2013).

Twitter Indexes

The collected Twitter data were analyzed using three indexes—ratio, normalized ratio, and sentiment—to represent three dimensions of Twitter responses to disasters. Ratio index is defined as the number of disaster-related tweets in an area divided by the total number of background tweets in the same area within a period (Equation 1). The Ratio index is also defined as situational awareness (Earle, Bowden, and Guy 2012), risk perception (Mandel et al. 2012), or level of concern (Lachlan, Spence, and Lin 2014) in previous research. Its values range from 0 to 1. A high ratio index means a high level of public concern about the disaster on Twitter, with more disaster-related tweets generated. The ratio index of a community represents the extent of awareness to, or the impact caused by, the event in the community. The ratio index has been suggested for damage estimation and recovery monitoring (Guan and Chen 2014).

$$\text{Ratio} = \frac{\#\text{ Sandy related Tweets}}{\#\text{ Background Tweets}}. \qquad (1)$$

To eliminate the effect of hazard threat level so that we can examine the disparities of Twitter activities under the same threat level for the study area, we developed the second index, the normalized ratio index (NRI). The NRI is calculated as the ratio index divided by the hazard threat level (Equation 2). Methodology to derive the threat level of each county from Hurricane Sandy is discussed later in this article.

$$\text{NRI} = \text{Ratio}/\text{Threat Level}. \qquad (2)$$

The third index, Sentiment, is the average sentiment score of collected tweets in an area within a period (Equation 3). The sentiment scores range from –1 to 1. A tweet with a sentiment score from –1 to 0 is identified as a negative tweet, whereas the sentiment score of a positive tweet is greater than 0. A tweet without any emotional word or symbol detected is assigned a sentiment score of 0 and classified as neutral. The spatial–temporal pattern of average sentiment scores has been applied to investigate individual mood, geography of happiness, and public attitude toward climate change (Dodds et al. 2011; Golder and Macy 2011; Mitchell et al. 2013; Cody et al. 2015). Specifically, the sentiment index was found to be sensitive to major disturbances in the real world (Kryvasheyeu et al. 2016). These three Twitter indicators can be calculated or aggregated by user-defined areas (e.g., county, country, and global) within a period (e.g., daily and hourly).

$$\text{Sentiment} = \text{Mean(sentiment scores)}. \qquad (3)$$

Traditional Data

In addition to social media data, we collected damage data and hurricane-related geographical and socioeconomic data (Table 1). Hurricane-related geographical variables include distance to the hurricane

Table 1. List of geographical, social, and damage variables

Category	Abbreviation	Variable
Geographical	DistTrack	Distances of county centers to the hurricane track
	DistLandfall	Distances of county centers to the landfall location
	DistCoast	Distances of county centers to the coastline
	Threat level	Averaged kernel density by hurricane track and wind speed
Damage	Damage	Per capita economic loss in thousands of dollars
Social (2010)	PctYoung	% population 17 to 29 years old
	Education	% population over 25 with a bachelor's or higher degree
	HIncome	Median household income
	PctUrban	% urban areas within the county
	FemaleHH	% female-headed households
	MobileH	% mobile homes
	NoPhone	% housing units without telephone service available
	Poverty	% population living in poverty
	NoVehicle	% households without a vehicle
	PctEmp	% civilian workforce that is employed

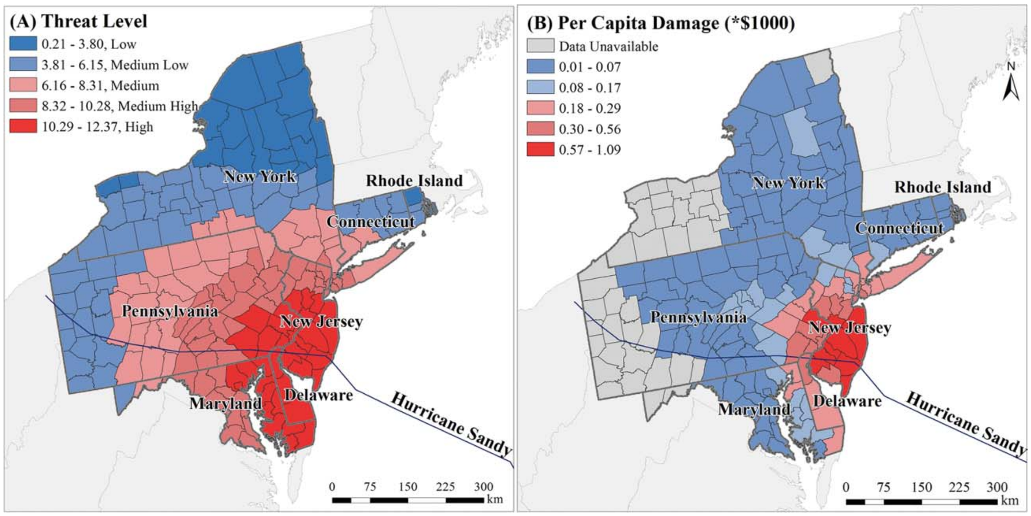

Figure 3. (A) County-level threat level and (B) per capita damage during Hurricane Sandy in affected areas. (Color figure available online.)

track, distance to the landfall location, distance to the coastline, and level of threat from the hurricane. Hurricane tracks with wind speed data were obtained from the NOAA Office for Coastal Management (see https://coast.noaa.gov/hurricanes/). Distances to hurricane track, landfall location, and the coastline were computed as distances between county centers to the corresponding features. The hurricane threat level in a county was calculated as the averaged kernel density based on the location of the hurricane track and its wind speed using ArcGIS (Figure 3A), which has been widely used for hurricane intensity modeling (Knowles and Leitner 2007; Lam et al. 2014; Lam et al. 2015). The grid size is 30 m × 30 m, and the bandwidth is defined as 390 km, which is the extent of Hurricane Sandy in its landfall (NOAA 2012). For each segment of the hurricane track, a kernel density surface was generated through the quadratic kernel function with its wind speed as the center value. The density at each output raster cell is calculated by adding the values from all the kernel surfaces.

Hurricane damage data were retrieved from the FEMA Modeling Task Force (MOTF). MOTF is a group of experts in hazard loss modeling and impact assessment, including hurricanes, flooding, and storm surges. In MOTF's project on Hurricane Sandy impact analysis, county-level economic losses measured in thousands of dollars are provided using the Hazus-Multi-Hazard (Hazus-MH) application (FEMA 2012). The advantage of this data source is that Hazus-MH incorporates as many factors as possible, thus providing a comprehensive damage assessment at the county level. Previous research has demonstrated that this data source is valid for understanding the damage caused by Hurricane Sandy (Guan and Chen 2014; Kryvasheyeu et al. 2016). A total of 154 counties in seven states have economic damage data. We divided the total damage in each county by its population to compute the per capita damage caused by Hurricane Sandy (Figure 3B).

A total of ten socioeconomic variables collected for the year 2010 (before Hurricane Sandy) were accessed from the U.S. Census (see https://www.census.gov/) and the National Historical Geographic Information System (see https://www.nhgis.org/; Table 1). The selection of these variables was based on two reasons. First, previous studies show that urban communities with more residents who are between eighteen and twenty-nine years old, wealthy, and well-educated tend to have more Twitter use (Kent and Capello 2013; Li, Goodchild, and Xu 2013; Sloan et al. 2015). The selected ten variables include some of the variables suggested in these previous studies. Second, these ten variables were found to be important in representing the resilience capacity of communities in the resilience literature (Cai et al. 2016; Lam et al. 2016). These variables represent the socioeconomic condition of a community; thus, they can be used to test the hypothesis that social and geographical disparities in Twitter use during disasters exist.

Methods of Analysis

Twitter indexes (ratio, NRI, and sentiment) were derived at multiple spatiotemporal scales. First, global ratio and sentiment indexes were calculated hourly and daily to provide a general trend of public awareness and sentiment toward this event. Country-level daily ratio indexes were computed to reveal how Twitter use differs from affected regions to nonaffected regions. Second, a U.S. county-level map of ratio index for the whole period was also created to show the spatial patterns of public awareness toward Hurricane Sandy during the three weeks in the United States.

Third, the detailed local analysis focuses on the 191 severely affected counties in the states of New York, New Jersey, Pennsylvania, Maryland, Delaware, Connecticut, and Rhode Island, which were declared to receive financial assistance from FEMA. Because county-level damage information is available for these 191 counties, Twitter activities and damage under diverse threat levels could be examined and compared. If a county has very few background tweets, the small number problem might occur when computing the ratio indexes. Therefore, we used a threshold of fifty background tweets to select counties for the detailed local analysis. A total of 126 out of 191 counties were selected based on the threshold.

Ratio, NRI, and sentiment indexes for each selected county were tabulated into preparedness, response, and recovery phases and for the whole three-week period. We applied FEMA's emergency management framework in this study, which includes four phases: preparedness, response, recovery, and mitigation (Figure 4; FEMA 2006). Preparedness is the short-term actions taken in the days and weeks before the disaster to reduce impacts, such as moving property or persons out of harm's way and preparing emergency supplies and shelters. Actions in the response phase involve the first wave of core emergency services, such as dispatching first responders into disaster-affected areas to do search and rescue, firefighting, or shelter victims. In the recovery phase, affected communities are restored to their previous states following the disaster, including repairing or rebuilding property and reemployment. Mitigation indicates long-term actions for preventing, reducing, and eliminating risks to people and property from future hazards. The four phases of emergency management have been employed in many studies to represent the cycle of disaster resilience (FEMA 2006; Yang et al. 2013).

Figure 4. The four phases of emergency management. (Color figure available online.)

Because this investigation is an analysis of short-term Twitter activities, only the first three phases—preparedness, response, and recovery—were used. By tabulating Twitter activities into the three phases, we can better understand what the factors are in affecting public discussion on the disaster in different phases; in turn, this information could guide stakeholders in better using social media platforms to reduce the social and geographical disparities and enhance disaster resilience. Hurricane Sandy made landfall on the United States on 29 October 2012 and dissipated on 31 October 2012, so we defined a week before the landfall of Sandy as the preparedness phase (23–28 October), three days during the landfall as the response phase (29–31 October), and approximately two weeks after as the recovery phase (1–12 November).

Fourth, we tested whether geographical and social disparities in Twitter use exist across different phases. To test the first hypothesis that communities located near the disaster are expected to have higher levels of Twitter activity, a county's proximity to the disaster was represented by four geographical factors, including distance to the hurricane track, distance to the coastline, distance to the landfall location, and the level of threat from Hurricane Sandy. Twitter activities were represented by the ratio indexes at different phases. Pearson correlation coefficients between each pair of Twitter index and geographical factor were computed and compared. To test the second hypothesis that given the same level of threat, communities with higher socioeconomic conditions are expected to have higher levels of disaster-related Twitter activity, NRI was used to represent the Twitter activity adjusted by the threat level. Four NRIs for the preparedness, response, and recovery phases, and for the whole period were derived and denoted $NRI_{preparedness}$, $NRI_{response}$, $NRI_{recovery}$, and NRI_{Total}, respectively. For each NRI, we conducted a stepwise regression analysis to identify the socioeconomic characteristics that are significant in affecting the NRI. The independent variables are the ten socioeconomic variables described in Table 1.

The fifth analysis is to test the third hypothesis that communities that suffered more damage have higher levels of Twitter activity and, in turn, Twitter information can be used for damage estimation. A stepwise linear regression was conducted with the normalized per capita damage as the dependent variable, and the independent variables were the eight Twitter indexes (ratio and sentiment indexes in phases of preparedness, response, recovery, and the whole period) and the level of threat from the hurricane.

Results

During the data collection time period (23 October–12 November 2012), a total of 298,615 out of 80 million (0.37 percent) tweets were identified as disaster-related tweets. Among them, 167,959 (58.1 percent) tweets were successfully geocoded to a location (this number includes 3,472, or 1.2 percent of tweets that have latitude and longitude coordinates). Meanwhile, 4,890,050 background tweets were randomly selected from the data library, and 1,858,543 (38.0 percent) of them were geocoded as a base layer to normalize the disaster-related tweets. It took twenty-one days to finish the geocoding process. Because the selected background tweets are one sixteenth of the whole database, all calculated ratio indexes were divided by sixteen to derive the true ratio indexes. Results from the five analyses are summarized as follows.

The Global Spatial–Temporal Pattern

The first objective of this study is to derive common indexes from Twitter data so that they can be used for emergency management and resilience analysis across

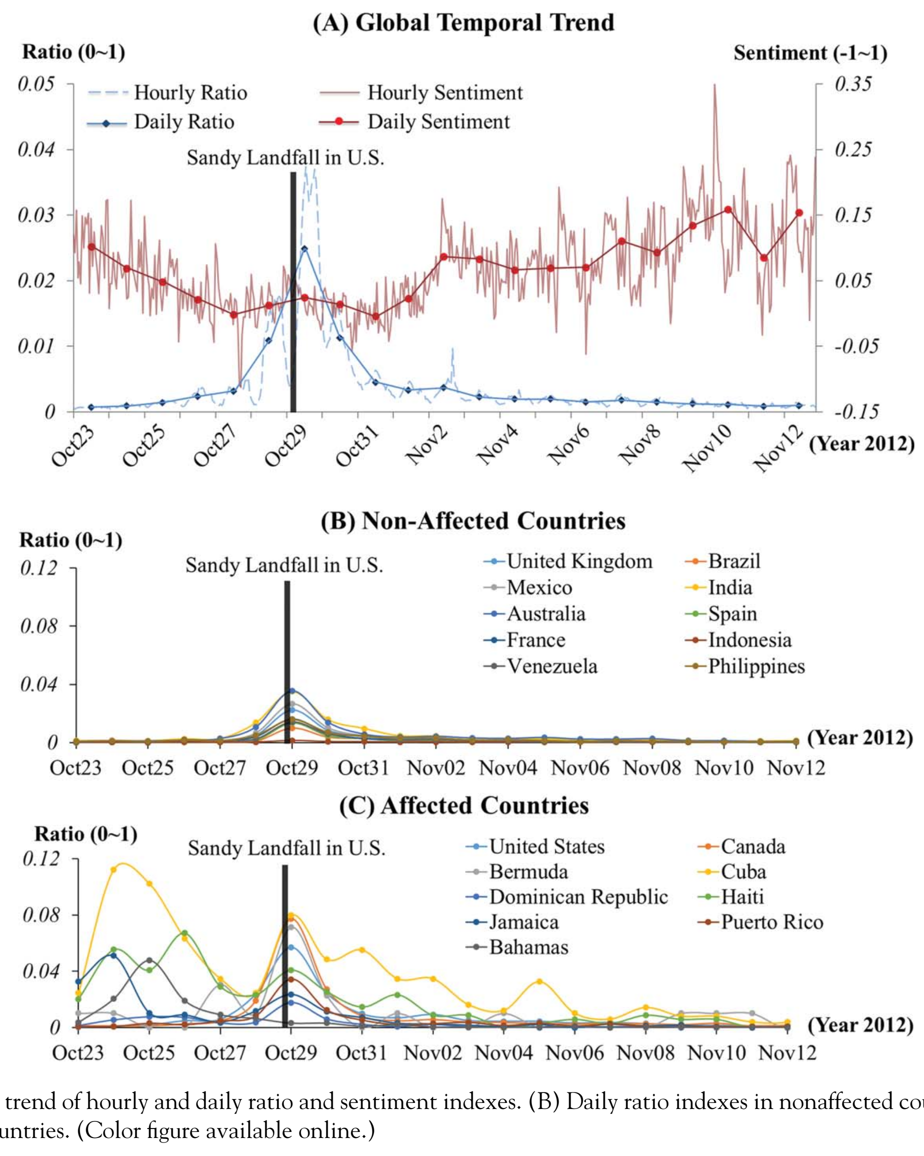

Figure 5. (A) Global trend of hourly and daily ratio and sentiment indexes. (B) Daily ratio indexes in nonaffected countries. (C) Daily ratio indexes in affected countries. (Color figure available online.)

regions and events. Figure 5 depicts the temporal variations of ratio and sentiment indexes at the global scale, which indicates the general trend of public awareness and feelings toward Hurricane Sandy on Twitter through time. Global ratio and sentiment indexes were aggregated hourly and daily and represented as blue and red lines, respectively. At the global scale, the ratio indexes ranged from 0.000 to 0.037, with an average value of 0.0024. The highest hourly ratio index value was observed at 1:00 p.m. on 29 October 2012, after Hurricane Sandy made landfall in New Jersey. During 23 to 28 October, prior to its landfall on the East Coast, very few tweets were discussing the hurricane, even though it had already caused huge impacts in the Caribbean region. During 29 to 31 October when Hurricane Sandy approached, landed, and devastated the East Coast of the United States, it aroused intensive discussion on Twitter. Two weeks after the landfall, the intensity of discussion dissipated gradually and returned to the same level as 23 October.

At the global scale, affected countries included the United States, Canada, Bermuda, Cuba, the Dominican Republic, Haiti, Jamaica, Puerto Rico, and the Bahamas, whereas the nonaffected regions were represented by the top ten countries with the most disaster-related tweets. In the non affected regions (Figure 5B), the temporal variations of public awareness on Hurricane Sandy were similar to the global trend (Figure 5A). The highest daily ratio was less than 0.036 on 29 October, indicating that less than 4 percent of tweets were talking about Sandy in nonaffected countries on that day. Public discussion on Sandy in nonaffected regions lasted a short time, no more than four days (from 28 October to 1 November 2012). On the contrary, the temporal patterns of the affected regions (Figure 5C) were quite different. Affected regions show higher daily ratio indexes, with a range of 0.000 to 0.112, and the discussion lasted up to twenty days. In Cuba (yellow line in Figure 5C), for example, the highest daily ratio was on 24 October when Hurricane Sandy landed in Cuba, and the second discussion peak was reached on 29 October when Sandy landed in the United States. The highest ratio value was 0.112, indicating that 11.2 percent of tweets were talking about Hurricane Sandy on the day Sandy landed in Cuba, and the discussion lasted for more than two weeks, from 23 October to 12 November.

The difference in the ratio indexes between affected and nonaffected regions is expected. It confirms that affected countries have higher levels of Twitter activity and longer discussion than nonaffected countries. Twitter users' awareness in nonaffected countries could be media driven, which means that people discuss the event because there are mass media reporting it on social media platforms. In addition, tweets from governmental and nongovernmental agencies might also generate new discussions. On the contrary, residents in the affected region discuss the event because not only were there many reports on traditional and social media but the event might also affect their daily lives.

Table 2. Top five saddest and happiest Sandy-related tweets of all tweets by sentiment scores

Rank	Tweet content	Scores
Top 5 saddest tweets		
1	I want to cook eggs so bad right now. Stupid electric stove. Stupid storm. Stupid metabolism. Stupid life. Stupid existence. Stupid eggs.	−0.9766
2	The world is coming to an end! :(An earthquake just happened in Hawaii :(hurricane sandy, is coming to New York!! :(:(:(:(:(:(	−0.9646
3	I want Hurricane Sandy to be bad enough that we don't get school on Monday, but not so bad that people die	−0.9590
4	Why are my lights flickering!! Bad snow bad storm bad bad bad mother nature	−0.9589
5	Terrible scenes New York 16 killed. in other news (somewhere ??) same sandy wreaks havoc in Caribbean 67 killed	−0.9542
Top 5 happiest tweets		
1	TGIF Tweeters . . . Have a fabulous day!! Love laugh smile n inspire!! Enjoy your weekend :-) Prayers to all those touched by "Sandy"	0.9819
2	TY mah gorgeous BFF! @MeditDancer Have a gr8 weekend, we R headed 2 a Halloween party and R hoping 4 the BEST with #Hurricane #Sandy	0.9806
3	@AmongTheHidden_ HAPPY BIRTHDAY LOVE! You're awesome!!! I hope you have a great birthday and the hurricane does not blow down your house!	0.9739
4	I LOVE weather!!! Storms are the best!! People need to respect mother nature more! And enjoy her magnificent power!!	0.9712
5	Anyone who has Hurricane Sandy in their path hope you and your love ones near and far will be OK!!! Be safe	0.9684

In terms of sentiment scores, public sentiment toward the event was also changing through time. The sentiment scores varied from –0.11 (negative) to 0.35 (positive), with an average value of 0.04. Sentiment scores were the lowest right after Hurricane Sandy made landfall, implying that Hurricane Sandy negatively affected human lives during that time period. To further exemplify how Twitter users used this platform during disasters and how those tweets associated with the sentiment scores the top five saddest and happiest Sandy-related tweets by sentiment scores during the three-week period were retrieved and are listed in Table 2. The lowest and highest sentiment scores of all tweets were –0.9766 and 0.9819, respectively. According to the tweet contents, most of the saddest Sandy-related tweets mentioned the damage caused by Hurricane Sandy, such as people who lost their lives, unstable power, and destroyed homes for children. A closer look at the top five happiest Sandy-related tweets shows that most of the "happy" tweets were praying for friends and families, and most of them were not affected by this event.

The Spatial–Temporal Pattern in the United States

The spatial pattern of county-level ratio indexes in the United States during the three weeks are shown in Figure 6. To avoid the small number problem, only counties with more than twenty background tweets are shown in the map, which leads to 1,767 out of 3,233 counties that meet the criterion. The average ratio index was 0.0074. High-ratio counties were found to concentrate along the East Coast of the United States, especially in the severely affected area. The Pearson correlation coefficient between ratio indexes and distances of county centers to the hurricane track was –0.37, which was significant at the $p < 0.05$ level, indicating that counties closer to the hurricane track had higher levels of Twitter activity at the national scale.

Figure 6 also shows that a few counties in nonaffected states had very high ratio values. A total of fourteen counties with a ratio index greater than 2.62 percent (>2.5 standard deviations) were found in nonaffected states, with most of them being urban counties. As always, there are anomalies from the general trend, which could be due to multiple reasons, including the small number problem, geocoding errors, and Twitter user composition. For example, Aransas County in Texas, which had a high ratio value, is part of the Corpus Christi metropolitan area and has experienced frequent coastal hazards. Therefore, Twitter users living in Aransas County might be more concerned about hurricane events than users in other counties.

The spatial–temporal pattern is consistent with previous studies that used data with geotags (Shelton et al. 2014). The use of geocoding, such as in this study, has greatly increased the amount of data available for analysis. Although inaccurate georeferencing might occur due to imprecise or inaccurate address

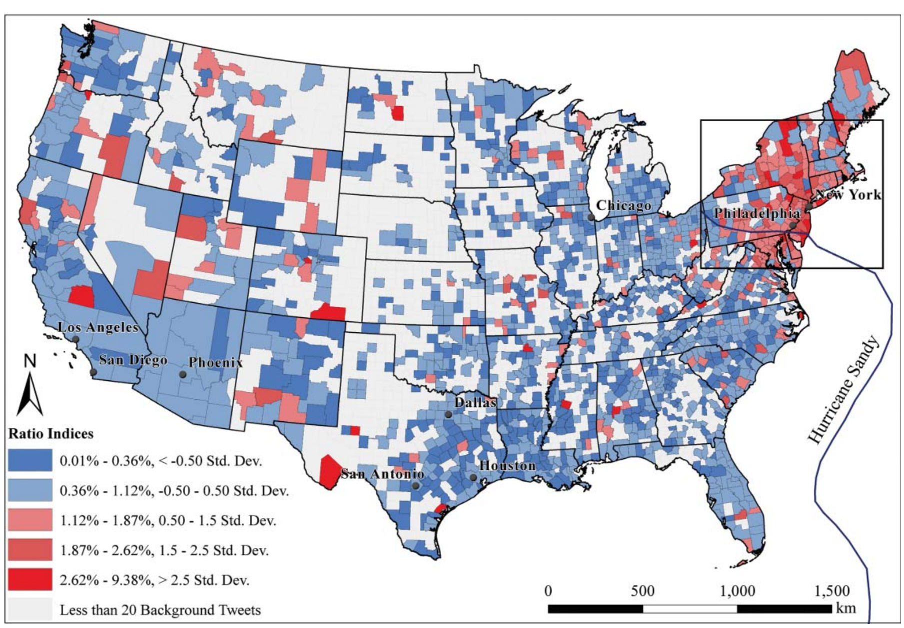

Figure 6. County-level ratio indexes of Hurricane Sandy in the United States. (Color figure available online.)

Table 3. Summary statistics of Twitter indexes in different phases

Twitter indexes	Phase	Minimum	Maximum	M
Ratio	Whole	0.0016	0.0439	0.0185
	Preparedness	0.0008	0.0625	0.0199
	Response	0.0075	0.1750	0.0623
	Recovery	0.0000	0.0275	0.0070
Normalized ratio	Whole	0.0003	0.0059	0.0024
	Preparedness	0.0002	0.0069	0.0025
	Response	0.0014	0.0272	0.0083
	Recovery	0.0000	0.0033	0.0009
Sentiment	Whole	−0.1492	0.3063	0.0398
	Preparedness	−0.8877	0.8627	0.0363
	Response	−0.2187	0.4709	0.0265
	Recovery	−0.4588	0.5386	0.0906

information in the user profiles, normalizing disaster-related tweeting frequencies by background tweets (i.e., the ratio index) should still show the relative awareness of the disastrous events.

The Spatial–Temporal Pattern in the Northeast

Zooming into the severely affected area and tabulating the tweets into three phases provides more information on how public awareness changed through the three phases. A total of 126 out of 191 counties in the seven severely affected states had more than fifty background tweets in the period and were selected for the detailed analysis. The statistics of ratio, sentiment, and NRI indexes for the affected counties are summarized in Table 3.

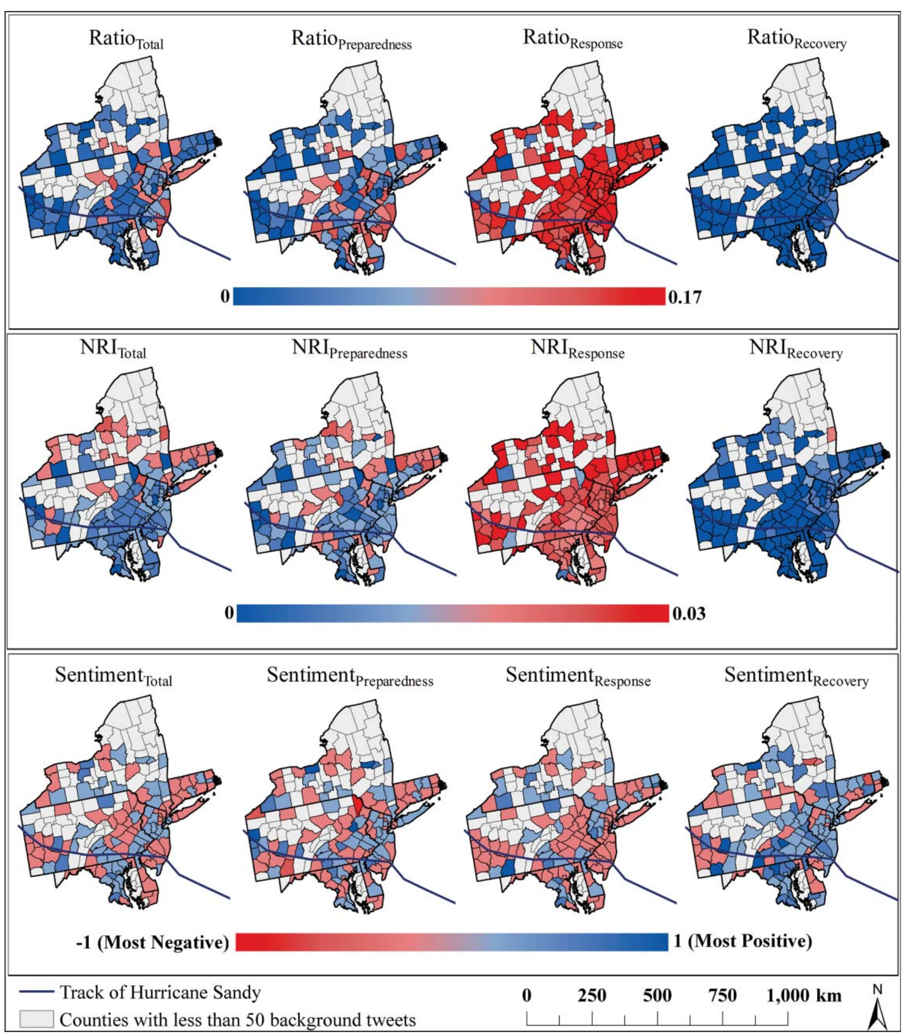

Figure 7. County-level ratio, normalized ratio, and sentiment indexes at different phases in the affected area. (Color figure available online.)

The spatial patterns of county-level ratio, NRI, and sentiment indexes at different phases in the severely affected zone are shown in Figure 7. During the three-week period, the average ratio was 0.0185, which is much higher than the national average (0.0074). High total ratios were concentrated in counties along the coastline where hurricane-induced wind and storm surge were stronger. The highest ratio indexes were in the response phase, and the lowest were in the recovery phase. The averaged NRI during the three-week period was 0.0024. The highest and lowest NRIs were also at the response and recovery phases, with average values of 0.0083 and 0.0009, respectively. The average sentiment score for the entire three-week pattern was 0.0398, with the lowest score being at the response phase (0.0265) and the highest score at the recovery phase (0.9060).

The highest sentiment score found at the recovery phase could be interpreted as follows. During the preparedness phase, people could be very anxious because of the unknown potential damage that could be caused by the hazard. During the response phase, most people in the study area might be affected by the hazard. Therefore, Twitter users are more likely to express negative sentiment on Twitter in the preparedness and response phases. In the recovery phase, however, Twitter users are more likely to post positive tweets to show the recovery process and support people who suffered from the disaster. Also, many people who suffered severe damage might be busy rebuilding homes and are unlikely to tweet in the recovery phase. This results in fewer negative tweets and overall higher sentiment score at the recovery phase.

The top five saddest and happiest Sandy-related tweets by sentiment scores in the northeastern United States at the recovery phase were retrieved and are listed in Table 4. The lowest and highest sentiment scores were –0.9487 and 0.9588, respectively. Most of the saddest Sandy-related tweets complained of the damage caused by Hurricane Sandy, whereas the happiest Sandy-related tweets involved praying for friends and families, checking on their safety, or updating the recovery process.

Geographical and Social Disparities

To test the first hypothesis that communities located near the disaster are expected to have higher levels of Twitter activity, four geographical factors including level of threat, distance to the hurricane track, distance to the landfall location, and distance to the coastline were used to correlate with the Twitter activities. The results are summarized in Table 5. In the local analysis, all the geographical variables except distance to the hurricane track were found to be significantly correlated with

Table 4. Top five saddest and happiest Sandy-related tweets by sentiment scores in the severely affected zone at the recovery phase

Rank	Tweet content	Scores
Top 5 saddest tweets		
1	@SoapZoneCoggie Last yr's storms were bad, but nothing like this. So many still in the dark, no heat, limited gas & food. Such devastation :(	−0.9487
2	i hate the cold. i hate rain. i hate snow. i hate hurricanes. go away.	−0.9413
3	4 Montauk classrooms badly damaged in #HurricaneSandy, but destruction in perspective as community mourns Dee Wright.	−0.9280
4	can this insane weather just leave my poor dad alone?? hurricane then snow! awful :(	−0.9267
5	@MarkRoseHfx not mine, a friend who lost her apartment and studio. Her kitty is sick, else I'd take it in. : (stupid storm.	−0.9264
Top 5 happiest tweets		
1	@LIBN We at #blueribbonpet hope you and all your pets are safe after super storm Sandy!!! Best of luck!	0.9588
2	@IslandgirlOK Hello! Yes, thank you, I didn't even lose power through the storm, so I'm very fortunate. TY for asking. :)	0.9503
3	The wonderful people @Food52 lend a hand for hurricane relief in the best way possible – cooking comfort foods. http://t.co/9SwxSRmr	0.9493
4	@marbrenda123 I have missed you as well. Been on&off since #Sandy! Happy Hump Day! Hope your safe&warm. Sending smiles and hugs UR way!	0.9483
5	One thing this storm reminded me is to celebrate & appreciate LIFE. Love LIFE! Give LIFE! Live LIFE! God is faithful. #GreatExploits	0.9475

Table 5. Pearson correlations between Twitter ratio indexes and geographical factors

Twitter indexes	Threat level	DistTrack	DistLandfall	DistCoast
$\text{Ratio}_{\text{Total}}$	0.445*	–0.098	–0.471*	–0.372*
$\text{Ratio}_{\text{Preparedness}}$	0.596*	–0.234*	–0.587*	–0.522*
$\text{Ratio}_{\text{Response}}$	0.278*	–0.018	–0.289*	–0.186*
$\text{Ratio}_{\text{Recovery}}$	0.353*	–0.013	–0.454*	–0.381*

*Significant at $p < 0.05$.

Twitter use in all three phases and in total. The $\text{Ratio}_{\text{Total}}$ index was positively correlated with the level of threat but negatively correlated with distances to the landfall location and the coastline. The correlations between Twitter activities and geographical variables were the highest in the preparedness phase but lowest in the response phase. The changing correlations between ratio indexes and geographical factors during the three phases reveals the changing roles of geographical factors affecting Twitter use through time. In the preparedness phase, people in communities near the landfall location and coastline were at high risk of potential impacts by strong wind and storm surge. Therefore, these communities are expected to be more aware of the hurricane threat. The results support the hypothesis that significant geographical disparities existed in Twitter use during disasters, especially in the preparedness phase.

To address the research question regarding whether social disparities affected Twitter activities under the same level of threat and to test the second hypothesis that counties with higher socioeconomic status have higher disaster-related Twitter activities, stepwise regression analyses between the NRIs and the ten socioeconomic variables at the three phases and for the entire period were conducted. This resulted in four regression models with adjusted R^2 of 0.29, 0.19, 0.41, and 0.33 for the total, preparedness, response, and recovery phases, respectively (Table 6). All models are significant at the $p < 0.05$ level.

From Table 5, the first $\text{NRI}_{\text{Total}}$ model provides a general trend of the relationship between socioeconomic conditions and disaster-related Twitter activities, although the model has the second lowest R^2 (0.29). The variables that correlate negatively with $\text{NRI}_{\text{Total}}$ were percentage urban area within the county and percentage mobile homes, whereas percentage of the population between seventeen and twenty-nine years old and percentage of civilian workforce that is employed were positively correlated with $\text{NRI}_{\text{Total}}$. This result agrees with those found in the literature except for the urban area variable. The social media literature suggests that urban, young, more educated people often tweet more (Li, Goodchild, and Xu 2013; Sloan et al. 2015), but in this study, urban area is negatively associated with disaster-related Twitter use. This discrepancy will need to be further evaluated in future studies.

A closer look at the three models at different phases reveals how socioeconomic conditions affect Twitter use through time. In the preparedness phase, the model yielded the lowest R^2 (0.19), with young and income positively associated and urban area negatively associated with Twitter use. The social disparities are most significant in the response phase when the third model reached the highest adjusted R^2 of 0.41. In this model, education was added as a new positive contributor. In the recovery phase (the fourth model), percentage of female-headed households, a well-known poverty indicator, and percentage of mobile homes were negatively associated with disaster-related Twitter use, whereas percentage of households without a car was the lone variable positively associated with disaster-related Twitter use. Overall, these regression results suggest that the hypothesis that higher disaster-related levels of Twitter activity are associated with communities with higher socioeconomic conditions is true.

Table 6. Stepwise linear regression models of normalized ratio indexes in different phases

Phases	Model	Adjusted R^2
Whole	$\text{NRI}_{\text{Total}} = 0.407 * \text{PctYoung} + 0.28 * \text{PctEmp} - 0.312 * \text{PctUrban} - 0.385 * \text{MobileH}$	0.29
Preparedness	$\text{NRI}_{\text{preparedness}} = 0.494 * \text{PctYoung} + 0.231\text{HIncome} - 0.255\text{PctUrban}$	0.19
Response	$\text{NRI}_{\text{Response}} = 0.432 * \text{PctYoung} + 0.434 * \text{PctEmp} + 0.162 * \text{Education} - 0.511 * \text{PctUrban} - 0.359 * \text{MobileH}$	0.41
Recovery	$\text{NRI}_{\text{Recovery}} = 0.363 * \text{NoVehicle} - 0.468 * \text{MobileH} - 0.313\text{FemaleHH}$	0.33

Note: All models are significant at the $p < 0.05$ level.

Table 7. Stepwise regression model for damage estimation

Variable	Coefficients β	Standardized β	Significance	M
Level of threat	0.363	0.736	0.000	8.159
$Ratio_{Recovery}$	68.583	0.322	0.000	0.007
$Sentiment_{Recovery}$	–0.732	–0.100	0.017	0.091
Dependent variable	$\log_{10}$(damage per capita)			
R^2	0.572		Adjusted R^2	0.563

Note: Model with threat level to hurricane alone: R^2 = 0.463, adjusted R^2 = 0.461.

Postdisaster Damage Estimation

The fifth analysis is to test the third hypothesis that communities that suffered more damage have higher levels of Twitter activity and, in turn, Twitter information can be used for damage estimation. Results of the stepwise regression between normalized per capita damage and the eight Twitter indexes (ratio and sentiment indexes in phases of preparedness, response, recovery, and the whole period) and hurricane threat level are summarized in Table 7. Only three variables, including level of threat from the hurricane and ratio and sentiment indexes in the recovery phase, were selected as significant indicators at the $p < 0.05$ level. Level of threat from hurricane and $Ratio_{Recovery}$ were positively correlated with damage, whereas $Sentiment_{Recovery}$ was negatively correlated with damage. The regression results suggest that the third hypothesis can be confirmed. By incorporating the Twitter indexes, the model can explain 56 percent of the damage variation. We also conducted a linear regression with the level of threat from the hurricane alone as the independent variable to explain damage variation. The regression model yielded an adjusted R^2 of 0.46, indicating that incorporating social media indexes in the model had increased the damage assessment model accuracy about 10 percent.

Integrating social media with traditional data for damage estimation has the potential to complement traditional damage estimation approaches, such as satellite monitoring or postdisaster surveying. Data availability, image resolution, and high expenses have often limited the application of satellite monitoring. Postdisaster surveying might be the most accurate method to identify the worst damaged areas, but this type of physical site inspection is time consuming and requires a large amount of labor forces and resources. Using social media data for real-time damage detection can help emergency managers not only quickly detect the hardest hit areas but also derive information on human response that is not detectable by physical monitoring, thus offering a timely and useful approach for emergency management (Hotz 2016).

Discussion

This study analyzed Twitter activity during Hurricane Sandy, and three associated hypotheses were tested and confirmed. Despite the success of the study, there is room for improvement for future studies. First and foremost is the issue of algorithm uncertainty in big data analysis that could introduce errors in research findings (Kwan 2016). To transfer the large amount of noisy raw Twitter data into practical indexes, this research employed several algorithms for data cleaning, parsing, geocoding, and text mining. Implementing different algorithms to process data could introduce uncertainties or errors, thus affecting the research findings. In the following, we outline these issues and make suggestions for future research.

First, this study used a short list of four key words (hurricane, Sandy, flood, and storm) to extract disaster-related tweets. This algorithm might overlook some disaster-related tweets related to recovery, thus potentially leading to underestimation of the ratio index at the recovery phase. For instance, Twitter users might use key words like *power outage* or *blackout* instead of *hurricane* or *Sandy* to show the specific damage they encountered in the recovery phase. In this study, however, these four key words have been demonstrated as significant in collecting Twitter data during Hurricane Sandy (Shelton et al. 2014; Kryvasheyeu et al. 2015). Previous studies have also demonstrated that adding more key words will collect many more irrelevant tweets before the disaster, while only slightly underrating the ratio index in the recovery phase. To determine the best set of key words, one possible solution is testing the sensitivity of each key word or combination of key words through a manually labeled training data set (Guan, Chen, and Work 2016).

Second, tweets without geotags were georeferenced according to addresses in user profiles. When users sign up or update their address information in their user profiles, some users fill in only broad addresses, such as United States or a state or county name, instead of writing their specific addresses. This practice results in some tweets geocoded to the center of the United States, a state, or a county. To understand the potential uncertainty introduced by the geocoding errors, it is necessary to investigate the proportion of each spatial scale in Twitter users' addresses. In addition, more research on mining accurate location information from the historical tweet contents needs to be addressed.

Third, the use of Google Geocoding API might affect the efficiency of the framework. The rate limitation of the public Google Geocoding API restricts the use of large data sets. This limitation could be solved by purchasing Google's premium plans or implementing a local geocoding server using open source data, such as OpenStreetMap planet data.

Fourth, analysis in areas with few Twitter users could cause the small number problem, as the reliability of ratio and sentiment indexes is highly dependent on the amount of data collected. A possible solution to alleviate the small number problem is to increase the data collection size, such as acquiring the full data set from Twitter agencies.

Finally, county-level analysis might be too coarse, because the county scale might overlook the socioeconomic inequalities within each county. A more accurate analysis conducted at the ZIP code level should help in testing the relationships, yield better understanding, and assist emergency management agencies to transfer the findings into practical tools in the future. This type of finer scale analysis would also rely on a full data set in which more Twitter data would be available.

Conclusion

This study analyzed the Twitter response to Hurricane Sandy between 23 October and 12 November 2012 at global, country, and county levels. The three research objectives were (1) to derive a set of common indexes from Twitter data so that they can be used to evaluate and compare disaster events across regions for emergency management and resilience analysis; (2) to examine whether there are significant geographical and social disparities in Twitter use; and (3) to test whether Twitter data can improve post-disaster damage estimation. Three corresponding hypotheses were tested and confirmed. Results show that common indexes derived from Twitter data, including ratio, NRI, and sentiment, could enable comparison across regions and events and should be documented. Social and geographical disparities in Twitter use existed in the Hurricane Sandy event. Higher disaster-related Twitter-use communities generally were communities of higher socioeconomic status, especially in the response phase. Finally, adding Twitter indexes in the recovery phase into a damage estimation model improved the adjusted R^2 from 0.46 to 0.56, indicating that social media data in conjunction with other traditional data could help improve postdisaster damage estimation.

There are important implications of the findings. For the first objective, the computed Twitter indexes provide baseline information on Twitter activity in Hurricane Sandy, which can be used to compare with those of similar disaster events across space and through time. Findings from the second objective, that geographical and social disparities in Twitter use existed in the three phases of emergency management, raise the question of whether and in what ways such disparities in Twitter use would affect disaster recovery and resilience. Also, future detailed studies could look into those counties that deviate from the hypothesis and explore why and how they deviate to gain further understanding of the disasters. Findings from the third objective suggest the potential of adding Twitter indexes in damage estimation. Rapid damage estimation information could help decision makers in resource allocation and management under emergencies. Finally, the methodological framework developed in this study, such as the use of geocoding, common indexes, and sentiment analysis, could guide researchers to properly use Twitter data to study disaster resilience. With a relatively small and imperfect data set, this study shows that the framework can be used to derive useful information that leads to improved understanding of disaster resilience.

Funding

This article is based on work supported by two research grants from the U.S. National Science Foundation: one under the Dynamics of Coupled

Natural Human Systems (CNH) Program (Award No. 1212112) and the other under the Interdisciplinary Behavioral and Social Science Research (IBSS) Program (Award No. 1620451). 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 funding agencies.

ORCID

Lei Zou http://orcid.org/0000-0001-6206-3558
Nina S. N. Lam http://orcid.org/0000-0002-5344-9368
Heng Cai http://orcid.org/0000-0001-6068-5150

References

Adger, W. N., T. P. Hughes, C. Folke, S. R. Carpenter, and J. Rockström. 2005. Social-ecological resilience to coastal disasters. *Science* 309 (5737):1036.

Agarwal, A., B. Xie, I. Vovsha, O. Rambow, and R. Passonneau. 2011. Sentiment analysis of Twitter data. In *Proceedings of the workshop on languages in social media*, ed. M. Nagarajan and M. Gamon, 30–38. Stroudsburg, PA: Association for Computational Linguistics.

Cable News Network. 2012. Hurricane Sandy fast facts. Accessed April 17 2017. http://www.cnn.com/2013/07/13/world/americas/hurricane-sandy-fast-facts/index.html.

Cai, H., N. S. N. Lam, L. Zou, Y. Qiang, and K. Li. 2016. Assessing community resilience to coastal hazards in the lower Mississippi River basin. *Water* 8 (2):46.

Caragea, C., A. C. Squicciarini, S. Stehle, K. Neppalli, and A. H. Tapia. 2014. Mapping moods: Geo-mapped sentiment analysis during hurricane sandy. In *Proceedings of 11th conference on information systems for crisis response and management*, ed. S. R. Hiltz, M. S. Pfaff, L. Plotnick, and P. C. Shih, 642–51. University Park, PA: The Pennsylvania State University.

Cody, E. M., A. J. Reagan, L. Mitchell, P. S. Dodds, and C. M. Danforth. 2015. Climate change sentiment on Twitter: An unsolicited public opinion poll. *PLoS ONE* 10 (8):e0136092.

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. *Local Evidence on Vulnerabilities and Adaptations to Global Environmental Change* 18 (4):598–606.

Cutter, S. L., B. J. Boruff, and W. L. Shirley. 2003. Social vulnerability to environmental hazards. *Social Science Quarterly* 84 (2):242–61.

Dodds, P. S., K. D. Harris, I. M. Kloumann, C. A. Bliss, and C. M. Danforth. 2011. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS ONE* 6 (12):e26752.

Dufty, N. 2012. Using social media to build community disaster resilience. *Australian Journal of Emergency Management* 27 (1):40.

Earle, P. S., D. C. Bowden, and M. Guy. 2012. Twitter earthquake detection: Earthquake monitoring in a social world. *Annals of Geophysics* 54 (6):708–15.

Earle, P., M. Guy, R. Buckmaster, C. Ostrum, S. Horvath, and A. Vaughan. 2010. OMG earthquake! Can Twitter improve earthquake response? *Seismological Research Letters* 81 (2):246–51.

Federal Emergency Management Agency (FEMA). 2006. *Principles of emergency management, independent study*. Washington, DC: FEMA.

———. 2012. *Hurricane Sandy impact analysis*. Accessed April 17, 2017. https://www.arcgis.com/home/item.html?id=307dd522499blk14;d-4a44a33d7296a5da5ea0

Golder, S. A., and M. W. Macy. 2011. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science* 333 (6051):1878.

Graham, M., S. A. Hale, and D. Gaffney. 2014. Where in the world are you? Geolocation and language identification in Twitter. *The Professional Geographer* 66 (4):568–78.

Guan, X., and C. Chen. 2014. Using social media data to understand and assess disasters. *Natural Hazards* 74 (2):837–50.

Guan, X., C. Chen, and D. Work. 2016. Tracking the evolution of infrastructure systems and mass responses using publically available data. *PLoS ONE* 11 (12): e0167267.

Hotz, R. L. 2016. Twitter storms can help gauge damage of real storms and disasters, study says. *Wall Street Journal*, March 11. Accessed May 17, 2017. http://www.wsj.com/articles/twitter-storms-can-help-gauge-damage-of-real-storms-and-disasters-study-says-1457722801.

Hutto, C. J., and E. Gilbert. 2014. VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international AAAI conference on weblogs and social media*, ed. E. Adar and P. Resnick, 216–25. Palo Alto, CA: The Association for the Advancement of Artificial Intelligence Press.

Kent, J. D., and H. T. Capello. 2013. Spatial patterns and demographic indicators of effective social media content during the Horsethief Canyon fire of 2012. *Cartography and Geographic Information Science* 40 (2):78–89.

Kirilenko, A. P., and S. O. Stepchenkova. 2014. Public microblogging on climate change: One year of Twitter worldwide. *Global Environmental Change* 26:171–82.

Knowles, J. T., and M. Leitner. 2007. Visual representations of the spatial relationship between Bermuda High strengths and hurricane tracks. *Cartographic Perspectives* 56:37–51.

Kryvasheyeu, Y., H. Chen, E. Moro, P. V. Hentenryck, and M. Cebrian. 2015. Performance of social network sensors during Hurricane Sandy. *PLoS ONE* 10 (2): e0117288.

Kryvasheyeu, Y., H. Chen, N. Obradovich, E. Moro, P. V. Hentenryck, J. Fowler, and M. Cebrian. 2016. Rapid assessment of disaster damage using social media activity. *Science Advances* 2 (3):e1500779.

Kwan, M. P. 2016. Algorithmic geographies: Big data, algorithmic uncertainty, and the production of geographic knowledge. *Annals of the American Association of Geographers* 106 (2):274–82.

Lachlan, K. A., P. R. Spence, and X. Lin. 2014. Expressions of risk awareness and concern through Twitter: On the utility of using the medium as an indication of audience needs. *Computers in Human Behavior* 35:554–59.

Lam, N. S. N., H. Arenas, P. L. Brito, and K.B. Liu. 2014. Assessment of vulnerability and adaptive capacity to coastal hazards in the Caribbean region. *Journal of Coastal Research* 70 (Suppl. 1):473–78.

Lam, N. S. N., Y. Qiang, H. Arenas, P. Brito, and K. Liu. 2015. Mapping and assessing coastal resilience in the Caribbean region. *Cartography and Geographic Information Science* 42 (4):315–22.

Lam, N. S. N., M. Reams, K. Li, C. Li, and L. P. Mata. 2016. Measuring community resilience to coastal hazards along the northern Gulf of Mexico. *Natural Hazards Review* 17 (1):04015013.

Li, K., N. S. N. Lam, Y. Qiang, L. Zou, and H. Cai. 2015. A cyberinfrastructure for community resilience assessment and visualization. *Cartography and Geographic Information Science* 42 (Suppl. 1):34–39.

Li, L., M. F. Goodchild, and B. Xu. 2013. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science* 40 (2):61–77.

Lindsay, B. R. 2011. *Social media and disasters: Current uses, future options, and policy considerations*. Washington, DC: Congressional Research Service.

Mandel, B., A. Culotta, J. Boulahanis, D. Stark, B. Lewis, and J. Rodrigue. 2012. A demographic analysis of online sentiment during Hurricane Irene. In *Proceedings of the second workshop on language in social media*, ed. M. Nagarajan, S. Owsley Sood, and M. Gamon, 27–36. Stroudsburg, PA: Association for Computational Linguistics.

Merchant, R. M., S. Elmer, and N. Lurie. 2011. Integrating social media into emergency-preparedness efforts. *New England Journal of Medicine* 365 (4):289–91.

Mislove, A., S. Lehmann, Y. Y. Ahn, J. P. Onnela, and J. N. Rosenquist. 2011. Understanding the demographics of Twitter users. In *Proceedings of the fifth international AAAI conference on weblogs and social media*, ed. N. Nicolov and J. G. Shanahan, 554–57. Spain, Barcelona: The Association for the Advancement of Artificial Intelligence Press.

Mitchell, L., M. R. Frank, K. D. Harris, P. S. Dodds, and C. M. Danforth. 2013. The geography of happiness: Connecting Twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS ONE* 8 (5):e64417.

National Oceanic and Atmospheric Administration (NOAA). 2012. *Hurricane SANDY*. National Hurricane Center, NOAA. Accessed September 22, 2017. http://www.nhc.noaa.gov/archive/2012/al18/al182012.public.029.shtml?.

———. 2014. *The thirty costliest mainland United States tropical cyclones 1900–2013*. Hurricane Research Division, NOAA. Accessed May 22, 2017. http://www.aoml.noaa.gov/hrd/tcfaq/costliesttable.html.

National Research Council. 2012. *Disaster resilience: A national imperative*. Washington, DC: National Academies Press.

Pang, B., L. Lee, and S. Vaithyanathan. 2002. Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on empirical methods in natural language processing*, ed. J. Hajic and Y. Matsumoto, Vol. 10, 79–86. Stroudsburg, PA: Association for Computational Linguistics.

Patton, R. M., C. A. Steed, and C. G. Stahl. 2013. Visualizing community resilience metrics from Twitter data. In *2013 International Conference on Weblogs and Social Media (ICWSM) Workshop*, ed. D. Archambault, E. Kandogan, and M. Harrigan, 6–9. Palo Alto, CA: The Association for the Advancement of Artificial Intelligence (AAAI) Press.

Reams, M. A., N. S. N. Lam, and A. Baker. 2012. Measuring capacity for resilience among coastal counties of the US northern Gulf of Mexico region. *American Journal of Climate Change* 1 (4):194–204.

Rizza, C., and A. G. Pereira. 2014. Building a resilient community through social network: Ethical considerations about the 2011 Genoa floods. In *Proceedings of the 11th international information systems for crisis response and management (ISCRAM) conference*, ed. S. R. Hiltz, M. S. Pfaff, L. Plotnick, and P. C. Shih, 289–93. University Park, PA: The Pennsylvania State University.

Roongpiboonsopit, D., and H. A. Karimi. 2010. Comparative evaluation and analysis of online geocoding services. *International Journal of Geographical Information Science* 24 (7):1081–1100.

Shelton, T., A. Poorthuis, M. Graham, and M. Zook. 2014. Mapping the data shadows of Hurricane Sandy: Uncovering the sociospatial dimensions of "big data." *Geoforum* 52:167–79.

Sloan, L., J. Morgan, P. Burnap, and M. Williams. 2015. Who tweets? Deriving the demographic characteristics of age, occupation and social class from Twitter user meta-data. *PLoS ONE* 10 (3): e0115545.

Taboada, M., J. Brooke, M. Tofiloski, K. Voll, and M. Stede. 2011. Lexicon-based methods for sentiment analysis. *Computational Linguistics* 37 (2):267–307.

Tsou, M. H. 2015. Research challenges and opportunities in mapping social media and big data. *Cartography and Geographic Information Science* 42 (Supp1.):70–74.

Tsou, M.-H., and M. Leitner. 2013. Visualization of social media: Seeing a mirage or a message? *Cartography and Geographic Information Science* 40 (2):55–60.

Walther, M., and M. Kaisser. 2013. Geo-spatial event detection in the twitter stream. In *Advances in information retrieval*, ed. P. Serdyukov, P. Braslavski, S. O. Kuznetsov, J. Kamps, S. Rüger, E. Agichtein, I. Segalovich, and E. Yilmaz, 356–67. Berlin: Springer.

Wang, Q., and J. E. Taylor. 2014. Quantifying human mobility perturbation and resilience in hurricane sandy. *PLoS ONE* 9 (11):e112608.

Yang, S., H. Chung, X. Lin, S. Lee, L. Chen, A. Wood, A. L. Kavanaugh, S. D. Sheetz, D. J. Shoemaker, and E. A. Fox. 2013. PhaseVis: What, when, where, and who in visualizing the four phases of emergency management through the lens of social media. In *10th international conference on information systems for crisis response and management*, ed. T. Comes, F. Fiedrich, S. Fortier, J. Gelderman, and T. Müller, 912–17. Karlsruhe, Germany: Karlsruhe Institute of Technology.

LEI ZOU is a Postdoctoral Research Associate in the Department of Environmental Sciences at Louisiana State University, 2281 Energy, Coast and Environment (ECE) Building, Baton Rouge, LA 70803. E-mail: lzou4@lsu.edu. His research interests include GIScience, location-based social media data mining, coupled natural–human system modeling, and disaster resilience.

NINA S. N. LAM is Professor and E. L. Abraham Distinguished Professor of Louisiana Environmental Studies in the Department of Environmental Sciences at Louisiana State University, Baton Rouge, LA 70803. E-mail: nlam@lsu.edu. Her research interests include GIScience, remote sensing, environmental health, resilience, and sustainability.

HENG CAI is a Postdoctoral Research Associate in the Department of Environmental Sciences at Louisiana State University, Baton Rouge, LA 70803. E-mail: hcai1@lsu.edu. Her research interests include geospatial modeling of coupled natural–human systems and the assessment of community resilience to hazards.

YI QIANG is Assistant Professor in the Department of Geography at University of Hawaii–Manoa, Honolulu, HI 96822. E-mail: yiqiang@hawaii.edu. His research interests include GIScience in general, geocomputation, visual analytics, and hazard vulnerability in particular.