Coupling sentiment and human mobility in natural disasters: a Twitter-based study of the 2014 South Napa Earthquake

Yan Wang & John E. Taylor

Natural Hazards
Journal of the International Society for the Prevention and Mitigation of Natural Hazards

ISSN 0921-030X
Volume 92
Number 2

Nat Hazards (2018) 92:907-925
DOI 10.1007/s11069-018-3231-1

Springer

Nat Hazards (2018) 92:907–925
https://doi.org/10.1007/s11069-018-3231-1

ORIGINAL PAPER

Coupling sentiment and human mobility in natural disasters: a Twitter-based study of the 2014 South Napa Earthquake

Yan Wang[1] · **John E. Taylor**[2]

Received: 6 October 2017 / Accepted: 19 February 2018 / Published online: 15 March 2018

Abstract Understanding population dynamics during natural disasters is important to build urban resilience in preparation for extreme events. Social media has emerged as an important source for disaster managers to identify dynamic polarity of sentiments over the course of disasters, to understand human mobility patterns, and to enhance decision making and disaster recovery efforts. Although there is a growing body of literature on sentiment and human mobility in disaster contexts, the spatiotemporal characteristics of sentiment and the relationship between sentiment and mobility over time have not been investigated in detail. This study therefore addresses this research gap and proposes a new lens to evaluate population dynamics during disasters by coupling sentiment and mobility. We collected 3.74 million geotagged tweets over 8 weeks to examine individuals' sentiment and mobility before, during and after the M6.0 South Napa, California Earthquake in 2014. Our research results reveal that the average sentiment level decreases with the increasing intensity of the earthquake. We found that similar levels of sentiment tended to cluster in geographical space, and this spatial autocorrelation was significant over areas of different earthquake intensities. Moreover, we investigated the relationship between temporal dynamics of sentiment and mobility. We examined the trend and seasonality of the time series and found cointegration between the series. We included effects of the earthquake and built a segmented regression model to describe the time series finding that day-to-day changes in sentiment can either lead or lag daily changed mobility patterns. This study contributes a new lens to assess the dynamic process of disaster resilience unfolding over large spatial scales.

✉ John E. Taylor
jet@gatech.edu

Yan Wang
wangyan@vt.edu

[1] Charles E. Via, Jr. Department of Civil and Environmental Engineering, Virginia Tech, 121 Patton Hall, Blacksburg, VA 24061, USA

[2] School of Civil and Environmental Engineering, Georgia Institute of Technology, Mason 4140c, Atlanta, GA 30332, USA

Keywords Disaster informatics · Earthquakes · Human mobility · Sentiment · Social media · Twitter

1 Introduction

Natural disasters adversely affect human beings and the built environment. According to the latest report from Munich RE (2017), over the past 7 years from 2010 to 2016, natural loss events (with at least one fatality and/or produced normalized losses larger than 100 thousand, 300 thousand, 1 million, or 3 million U.S. dollars depending on the assigned World Bank income group of the affected country) caused yearly average insurance losses of 55 billion U.S. dollars and overall losses of 174 billion U.S. dollars. During the same time period, there were in total 188 catastrophic natural events, and each of them caused more than 1000 fatalities and/or more than 100 million U.S. dollars in normalized losses (Munich RE 2017). Urban areas, due to their large and concentrated population and complex networked infrastructure systems, are highly susceptible to natural hazards, e.g., flooding, drought, storms, earthquakes, tsunamis and landslides (Godschalk 2003). It is crucial to improve urban resilience to natural hazards. Among the tasks of disaster risk reduction, benchmarks for strengthening urban resilience and adaptation is urgent because it is difficult to manage something that is not measured (UNISDR 2017).

Big data offers the potential to revolutionize our understanding of managing disaster risks in terms of vulnerability assessment, early warning, monitoring and evaluation (Ford et al. 2016). Taking advantage of the increasing use of social media platforms, e.g., Facebook and Twitter (Pew Research Center 2017), researchers have extended our understanding of disaster dynamics from diverse perspectives (Guan and Chen 2014; Tang et al. 2015; Wang and Zhuang 2017; Wang et al. 2016). These platforms can document geographical locations and collective reactions to extreme events in both virtual and physical worlds at a broad scale, which facilitates the development of research and practices in various branches of disaster management.

Twitter, among the most popular social media platforms, provides plentiful opportunities for detecting, tracking, and documenting extreme events. Its open design, wide usage, geo-enabled functionality and limited message lengths are well suited for emergency environments (Kryvasheyeu et al. 2016). Research involving sentiment analysis and human mobility has already taken advantage of the massive crowd-sourced data collected from Twitter. These two research topic areas help disaster managers make bottom-up decisions and play increasingly important roles in disaster relief. Specifically, sentiment analysis of short posts from social media has been shown to be an effective method to identify the dynamic polarity of sentiments over a disaster (Beigi et al. 2016), improve decision making regarding resource assistance, humanitarian efforts and disaster recovery, and obtain particular information (Nagy and Stamberger 2012). Additionally, human mobility, defined as the quantification of an individuals' movement trajectory, provides a basis to understand the perturbed movement patterns during and after disasters, and to predict displacements (Wang and Taylor 2014).

We developed the approach outlined in this study to enable a data-driven understanding of disaster dynamics in urban areas. To examine the spatiotemporal dynamics of urban areas during natural disasters, we: (a) examined the correlation between disaster intensity and collective sentiment, and the spatial association of sentiment in different disaster intensity zones; (b) investigated the temporal characteristics of sentiment and human mobility using an interrupted time series model; and (c) explored the relationship between individuals' sentiment and mobility over the course of a disaster.

The structure of the paper is arranged as follows: Sect. 2 summarizes the current studies on human mobility and sentiment in the disaster context, and develops two sets of hypotheses. Section 3 illustrates the specific disaster case studied, data collection, pre-processing and specific methods for computing sentiment and mobility, and for conducting spatial and temporal analysis. Section 4 explains the research results for each hypothesis. Section 5 discusses our research results and limitations. Finally, Sect. 6 draws conclusions from this study and suggests future studies for broader academic and practical work.

2 Related work

2.1 Sentiment analysis and natural disasters

Sentiment analysis of short posts from social media plays an increasingly important role in disaster relief and urban resilience. It is an effective method to help disaster managers to identify the dynamic polarity of sentiments over the course of a disaster (Beigi et al. 2016), improve decision making regarding resource assistance and requests, humanitarian efforts and disaster recovery, and obtain particular information (Nagy and Stamberger 2012). Vo and Collier (2013) classified and tracked the emotions of affected people using tweets during earthquakes in Japan. Eight types of emotion were selected to annotate tweets, including unconcerned, concerned, calm, unpleasantness, sadness, anxiety, fear and relief. The results revealed that fear and anxiety were the main emotions after an earthquake occurred, while calm and unpleasantness were only detected during severe earthquakes. Cody et al. (2015) explored the collective sentiment of tweets containing the word "climate", and found the connection between climate-change-related topics and a change of happiness. Bai and Yu (2016) proposed an incident monitoring framework in a post-disaster situation based on crowd negative sentiment of Chinese short blogs from Weibo. The framework was applied in the Ya'an earthquake and discovered aftershocks and potential public crises effectively. Although a few critical efforts have been made to classify sentiment during and after disasters, few studies have worked on examining both spatial and temporal dynamics of sentiment over the course of a disaster. A study by Neppalli et al. (2017) found unique spatial tweeting patterns of positive and negative sentiment following Hurricane Sandy: both positive and negative sentiment generally showed increasing clustering tendency to the point of Hurricane Sandy's maximum impact and then dispersed on the following days, while negative sentiment consistently clustered in closer proximity to Hurricane Sandy. It remains unclear what role disaster intensity plays in influencing sentiment, and what spatial patterns of sentiment may be during other types of disasters.

2.2 Human mobility during natural disasters

Human mobility, as a critical quantification basis of human dynamics, has triggered interest from diverse research areas, such as urban planning, traffic congestion, disease diffusion, and natural disasters. Different sources of crowd-sourced geo-referenced data have been utilized including Twitter, mobile phone records, billing records, etc. The analysis of this data has ushered a new era to quantitatively understand urban population dynamics. Recently, human mobility patterns during natural disasters have received considerable research attention. Scholars in the disaster research area have identified scaling laws and evaluated the predictability of human mobility during and after extreme events using mobility patterns from non-perturbed states. Lu et al. (2012) used approximately 1 year of mobile phone data of 1.9 million users and found that population movements following the Haiti earthquake had a high level of predictability, and destinations were correlated with normal-day mobility patterns and social support structure. A study by Wang and Taylor (2014) showed that human mobility was significantly perturbed during Hurricane Sandy but also exhibited high levels of resilience. A more recent study on multiple types of natural disasters revealed a more universal pattern of human mobility, as well as the limitations of urban human mobility resilience under the influence of multiple types of natural disasters (Wang and Taylor 2016). A recent study by Wang et al. (2017) also revealed how human mobility is perturbed by severe winter storms. These studies quantitatively improve our understanding of human mobility dynamics during events; however, the change of mobility pattern associated with sentiment levels over time before, during and after a natural disaster has yet to be investigated.

2.3 Hypothesis development

Although few disaster-oriented studies examine the spatial and temporal characteristics of sentiment and its relationship with mobility over time, these topics have drawn interest of researchers from other fields (e.g., urban studies and computer science). For instance, Bertrand et al. (2013) visualized sentiment in New York City based on 603,954 geotagged Tweets over 2 weeks. They identified that the level of sentiment is connected with location, e.g., it progressively improved with proximity to Times Square. They also found periodic patterns of sentiment at both daily and weekly scales: tweets on weekends tend to be more positive than on weekdays; midnight tweets are the most positive while 9:00 am and noon have the lowest-level-of-sentiment tweets. Lin (2014) examined sentiment segregation in urban communities with 3-months of geotagged tweets in Pittsburg. He explored the sentiment stability of neighborhoods and correlations between their sentiment orientations and the neighborhoods' demographic attributes. The results indicated a significant sentiment segregation effect. Mitchell et al. (2013) investigated the relationship between sentiment and geographic, emotional, demographic and health characteristics with 80 million geotagged tweets in 2011. Frank et al. (2013) characterized sentiment as a function of human mobility using a collection of 37 million geo-located tweets from 180,000 individuals. Research results of the two studies revealed that expressed happiness increased logarithmically with distance from an individuals' center of mass; it also increased logarithmically with the radius of gyration when binning individuals into ten equally sized groups by the radius of gyration.

These studies and our motivation to examine urban population dynamics during disasters inspired us to extend the spatiotemporal analysis of sentiment and mobility over time to the disaster context. We therefore propose two sets of hypotheses below:

Category 1 Spatial characteristics of sentiment

Hypothesis 1a Sentiment level is correlated with earthquake intensity: the higher intensity of disaster polygons that tweets/individuals are in, the lower sentiment levels tweets/individuals have.

Hypothesis 1b Sentiment level is clustered in space: tweets/individuals of similar sentiment levels tend to cluster together.

Category 2 Disasters can disrupt the temporal relationship between sentiment and mobility

Hypothesis 2a There is a significant interuption in time series of sentiment and mobility.

Hypothesis 2b Time series of sentiment and mobility are cointegrated.

Hypothesis 2c Change in sentiment (Δsentiment) and change in mobility (Δmobility) are cross-correlated over time, and Δsentiment is a predictor of Δmobility.

3 Data and methods

3.1 South Napa, California Earthquake

We elected to design our study of sentiment and mobility in natural disasters to focus on a severe earthquake. Geophysical disasters, such as earthquakes, are among the most severe natural disasters in terms of fatalities and damage. The 6.0 magnitude (M6.0) South Napa, California Earthquake was the strongest earthquake in 25 years in the Northern California Bay Area of the United States. The earthquake occurred at 10:20:44 UTC (03:20:44 PDT) on August 24, 2014, north of San Francisco. It reached the Modified Mercalli Intensity (MMI) Scale of VIII (severe) and on the moment magnitude scale a 6.0. MMI is a qualitative measure of the strength of ground shaking at a particular site and the USA employs the MMI scale, which ranges from I (not felt) to X (extreme) (USGS 2017). According to the Earthquake Engineering Research Institute (EERI 2014), the earthquake caused approximately 200 injuries and one fatality, and the total amount of federal aid was 30.8 million USD. Perceived shaking, potential damage and selected cities exposure under different estimated MMI can be found in Table 1 (USGS 2017).

3.2 Data description

The raw data for this study are comprised of geotagged tweets collected from a Twitter Streaming API (Wang and Taylor 2015). We use geotagging as the only filter to collect real-time tweets. As 1.24% of tweets are geotagged (Pavalanathan and Eisenstein 2015) and the streaming API can collect 1% of tweets, our database is representative in terms of geotagged tweets. Additionally, the Twitter geotags are based on GPS Standard Positioning Service, which offers a worst-case pseudo-range accuracy of 7.8 m with 95% confidence (Swier et al. 2015).

We used a spatial bounding box of intensity 2.5 contour to filter geotagged tweets (latitude from 37.382170 to 39.048830, longitude from − 123.561700 to − 121.061700), because an intensity 2.5 contour represents the lowest level of perceived shaking for this disaster case. The study period was set from August 3 to September 27, 2014. Specifically, we consider 3 a.m. (PDT) as the starting time of a day to aggregate 24-h tweets for further daily-based analysis due to the time that the earthquake occurred (3:20 a.m. PDT). In total, we collected 3,737,325 geotagged tweets. As we focused exclusively on tweets in English, the data volume reduced to 3,310,323 tweets. The daily average percentage of English

Table 1 Affected areas and population in different intensity zones

Intensity	II-III	IV	V	VI	VII	VIII	IX	X+
Perceived shaking	Weak	Light	Moderate	Strong	Very strong	Severe	Violent	Extreme
Potential damage	None	None	Very light	Light	Moderate	Mod/Heavy	Heavy	Very heavy
Selected city	Sacramento, Fremont, Stockton	Oakland, San Francisco	–	EI Verano, Sonoma, Temelec	American Canyon, Yountville	Napa	–	–
Exposed population	466 K, 214 K, 292 K	391 K, 805 K	–	4 K, 11 K, 1 K	19 K, 3 K	77 K	–	–

tweets is 88.58% during the study period, which indicates that these tweets are generally representative for the population in the studied area.

Additionally, for detailed analysis, five different intensity polygons were generated based on the contours of macroseismic intensity including 7.0, 6.0, 5.0, 4.0, and 3.0 intensity polygons. For instance, the 7.0 polygon refers to the polygon surrounded by the 6.5 intensity line, and the 6.0 polygon refers to the polygon surrounded by 5.5–6.5 intensity polylines. The GIS files in this study were obtained from USGS (2014). Allocations of the Twitter geolocations in the different intensity polygons on August 24, 2014, are illustrated in Fig. 1. Data volumes of English geotagged tweets in distinct intensity polygons are: 3008 (Intensity 7.0 polygon), 2280 (Intensity 6.0 polygon), 2655 (Intensity 5.0 polygon), 34,336 (Intensity 4.0 polygon), 40,173 (Intensity 3.0 polygon), and 1282 (Intensity 2.5 polygon).

3.3 Sentiment analysis

Twitter allows its users to share short 140-character messages. The texts can include words, URLs, mentions, emotions, abbreviations, etc. We cleaned the text by removing URL links and user mentions (@). We did not delete negations and kept as much context as possible for more accurate sentiment analysis. We adopted an unsupervised lexicon-based method to measure the sentiment. The method is based on an affective word list AFINN to assign sentiment scores to words in tweets (Nielsen 2011). The latest version of the word list includes 2477 words. The valence of a word ranges from − 5 (very negative) to + 5 (very positive) as an integer. The sum of valence without normalization of words represents the combined sentiment strength for a tweet. A Python Package "afinn" was used to compute the sentiment scores.

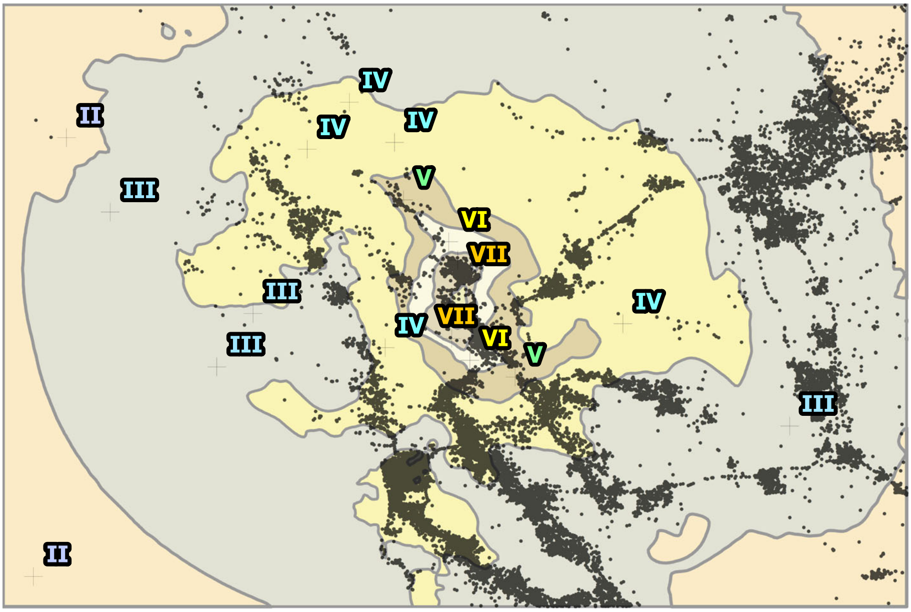


Fig. 1 Filtered Twitter geolocations in different intensity polygons on August 24, 2014

Comparing with methods that classify tweets to nominal categories (e.g., "negative, neutral, and positive"), numerical scores for sentiment contain more information about levels of sentiment and are more suitable for statistical analysis in this study. Additionally, AFINN is a Twitter-based sentiment lexicon including Internet slangs and obscene words. It has been tested in different types of tweets corpora and performs at a consistently satisfactory level of accuracy for both two classes (positive and negative) and three classes (positive, negative and neutral), compared with other unsupervised methods for sentence-level sentiment analysis (Ribeiro et al. 2015). Moreover, the AFINN word list has shown its advantages in analyzing tweets for disaster and crisis sentiment detection (e.g., Nagy and Stamberger 2012; Walther and Kaisser 2013). We therefore selected the AFINN lexicon to evaluate the sentiment polarity of our collected tweets.

3.4 Radii of gyration

Radii of gyration (r_g), a measurement of object movement from physics, has been widely used to quantify the size of trajectory of individuals since the study of Gonzalez et al. (2008). To achieve a more nuanced understanding of the perturbation of human mobility patterns, we computed the daily r_g of each distinct Twitter user over 8 weeks to identify the change of daily radii of gyration over time. The authors adopted the formula in Eq. 1 (Wang and Taylor 2016) to calculate the r_g of each distinct individual in the data set.

$$r_g = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left[2r \times \sin^{-1}\left(\sqrt{\sin^2\left(\frac{\phi_k - \phi_c}{2}\right) + \cos\phi_k\phi_c \sin^2\left(\frac{\varphi_k - \varphi_c}{2}\right)}\right)\right]} \quad (1)$$

where n is the total frequency of visited locations of a Twitter user, k is each visited location by the user during a 24 h period, c is the center location of the user's trajectories, ϕ is the latitude, and φ is the longitude.

3.5 Spatial autocorrelation

We employed Moran's I (Moran 1950) and Geary's C (Geary 1954) to measure the spatial autocorrelation of sentiment in earthquake-affected areas. As Moran's I is a more global measurement and sensitive to extreme values, and Geary's C is more sensitive to differences in small neighborhoods, we adopted both statistics.

Moran's I (Moran 1950) can measure how sentiment level of a location is similar to others surrounding it. Its value ranges from − 1 (perfect clustering of dissimilar values) to 1 (perfect clustering of similar values), and 0 indicates no autocorrelation (perfect randomness).

$$I = \frac{n}{\left(\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\right)} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \quad (2)$$

where $\bar{x}$ is the mean of the x variable, w_{ij} are the elements of the weight matrix.

Geary's C statistic (Geary 1954) is based on the deviations in responses of each observation with one another. It ranges from 0 (perfect positive autocorrelation) to a positive value (high negative autocorrelation). If the value is less than 1, it indicates positive spatial autocorrelation.

$$C = \frac{n-1}{2\left(\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\right)} \frac{\sum_i \sum_j w_{ij}(x_i - x_j)^2}{\sum_i (x_i - \bar{x})^2} \tag{3}$$

3.6 Time series analysis

3.6.1 Interrupted time series

We employed a segmented regression model (Wagner et al. 2002) to describe the interrupted time series of sentiment and mobility before, during and after the earthquake. This regression model is powerful in assessing the intervention effects in interrupted time series over time. We regarded the earthquake as a change point to divide the time series into two portions. Two parameters were used to define each segment of the time series: level and trend. The level refers to the value of the series at the start of a certain interval, while the trend is the rate of change during a portion.

The segmented regression model for sentiment (Eq. 4) and mobility (Eq. 5) are listed below:

$$S_t = \beta_0 + \beta_1 \times \text{time}_t + \beta_2 \times \text{disaster}_t + \beta_3 \times \text{time\_after\_disaster}_t + e_t \tag{4}$$

$$M_t = \beta_0 + \beta_1 \times \text{time}_t + \beta_2 \times \text{disaster}_t + \beta_3 \times \text{time\_after\_disaster}_t + e_t \tag{5}$$

S_t and M_t: daily average value of individual's adjusted sentiment and radius of gyration respectively (removed seasonality); time_t: a continuous variable indicating time in days at time t from the start of the observation period; $disaster_t$: an indicator for time t occurring pre-earthquake ($\text{disaster}_t = 0$) or post-earthquake ($\text{disaster}_t = 1$), which was implemented at day 22 in the series; $\text{time\_after\_disaster}_t$: a continuous variable counting the number of days after the disaster at time t; β_0 estimates the baseline level of the outcome, mean value of adjusted sentiment per individual per day, at time zero; β_1 estimates the change in the mean value of adjusted sentiment that compute with each day before the disaster occurs (i.e., the baseline trend); β_2 estimates the level change immediately after the disaster, that is, from the end of the preceding segment; β_3 estimates the change in the trend after the earthquake, compared with the daily trend before the disaster; the sum of β_1 and β_3 is the post-disaster slope.

3.6.2 Cointegration of time series

The relationship of cointegration reveals the *co-movement* of two time series in the long term. It can be illustrated by the simplest possible regression equation (Granger 1981) in our research scenario (Eq. 6):

$$M_t = \alpha + \beta S_t + \varepsilon_t \tag{6}$$

where M_t is the dependent variable, S_t the single exogenous regressor, and $\{\varepsilon_t\}$ a white-noise, mean-zero sequence. The definition of cointegration can be illustrated as "two non-stationary time series are cointegrated if some linear combination of them is a stationary series" (Metcalfe and Cowpertwait 2009, pp 217).

3.6.3 Cross-correlation of time series

We also employed cross-correlation analysis (see Eq. 7) to identify lags of the daily changed sentiment that might be useful predictors of daily changed mobility.

$$M_{t+h} = b + a\sum_{j=0}^{k-1} S_{t-j} + \varepsilon_t \tag{7}$$

where M_t is the time series of radius of gyration, S_t is the time series of sentiment. ε_t is the Gaussian noise, and h is the lag hyperparameter. A negative value for h is a correlation between sentiment at a time before t and the mobility at t, which means S_t leads M_t; while positive h means S_t lags M_t.

4 Results

4.1 Earthquake intensity and sentiment

We focused our analysis on the first 24 h after the earthquake. We specifically classified the collected tweets into six intensity polygons and conducted both tweet-based analysis and individual-based analysis. Firstly, we grouped the tweets into six bins of distinct intensities. The average *sentiment of tweets* in each bin is taken as the sentiment strength of the bin. The relationship between sentiment of tweets and intensity is shown in Fig. 2a. Sentiment decreases linearly with the intensity level. We also placed individuals into six intensity groups based on their center of mass (average location). For individuals who tweeted more than once during the 24 h, we calculated different statistics of sentiment of their tweets, including sum, average, median, maximum and minimum. Sentiment score of each group is the statistic of individuals' sentiment. The correlation between different statistics of individual's sentiment and earthquake intensity are plotted in Fig. 2b–f.

According to the results of the linear regressions, earthquake intensity can linearly explain the sum sentiment of individual's tweets better than other statistics ($R^2 = 0.944$). The average sentiment of tweets also has a significant correlation with the intensity. However, extreme sentiment of individual's daily tweets (i.e., max and min) cannot be explained by the earthquake intensity very well ($p > 0.1$).

4.2 Spatial association of sentiment

We performed spatial correlation analysis of sentiment scores across six disaster zones of distinct intensity and the whole area with Geary's C and Moran's I. The results of Geary's C and Moran's I are statistically significant with all p values less than 0.01, and all but two p values less than 0.001. Moran's I at different spatial scales is positive and values of Geary's C are less than 1, which indicates spatial dependencies for similar levels of sentiment in all earthquake-affected areas of different intensities. Note, however, that the values of Moran's I for the whole area and all the intensity zones are small, suggesting a weak but significant tendency of similar levels of sentiment to cluster in different areas. The autocorrelation analysis results can be found in Table 2.

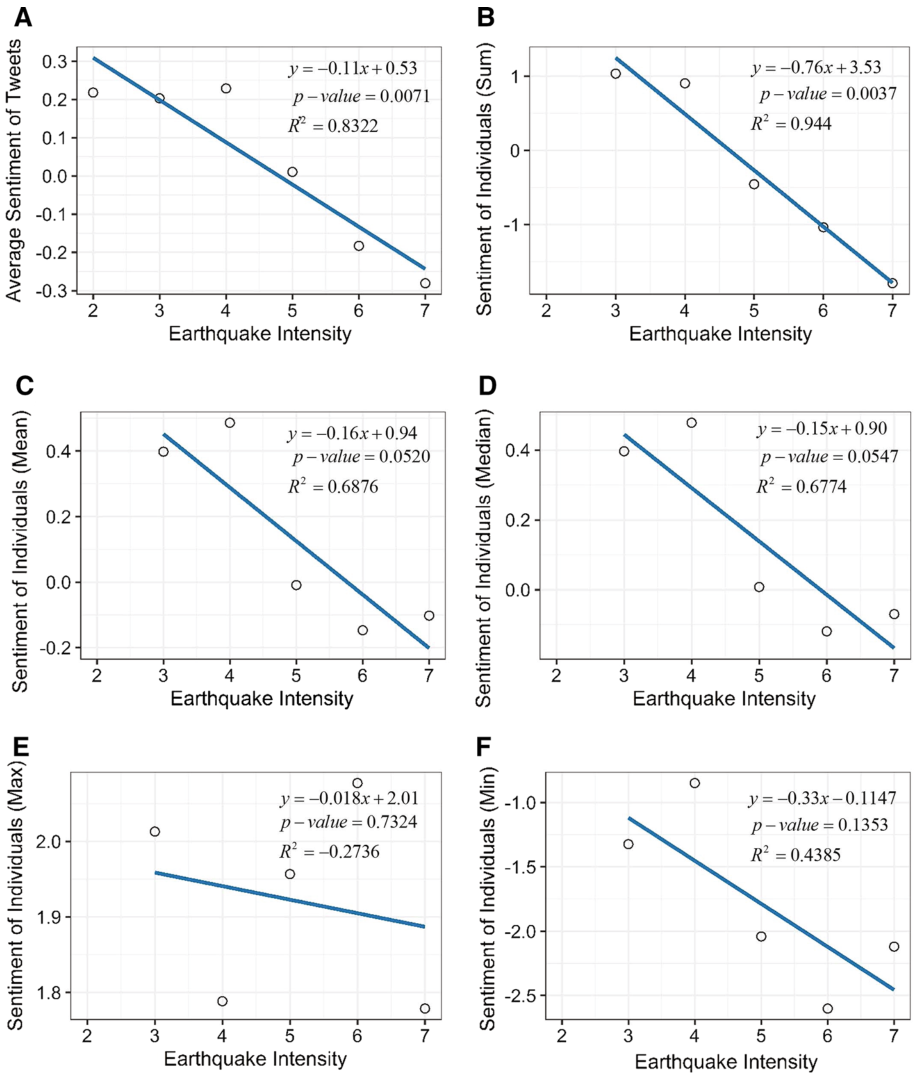


Fig. 2 **a** Average sentiment of tweets was a function of earthquake intensity, with tweets grouped into six different intensity polygons. **b–f** Average sentiment of individuals was a function of earthquake intensity, with individuals grouped into six different intensity polygons

4.3 Temporal analysis of sentiment and mobility

4.3.1 Decomposition of time series

We normalized daily sentiment and radius of gyration by dividing by the number of individuals (Eqs. 8 and 9). Only individuals with at least two distinct locations in a single day were included into the analysis. Examining Twitter postings from at least two distinct locations removes static tweets from bots and organizations.

Table 2 Spatial autocorrelation of sentiment during 24 h after the earthquake

	Moran's *I*			Geary's *C*		
	Statistics	*p* value	SD	Statistics	*p* value	SD
Whole area	0.0692	< 2.2e−16***	30.98	0.9258	< 2.2e−16***	15.637
7 intensity zone	0.0653	1.311e−08***	5.565	0.9425	0.0007042***	3.1929
6 intensity zone	0.0623	1.819e−06***	4.6311	0.9521	0.002832**	2.7666
5 intensity zone	0.0651	9.001e−08***	5.2189	0.9236	6.299e−05***	3.8342
4 intensity zone	0.0668	< 2.2e−16***	19.19	0.9291	< 2.2e−16***	9.0739
3 intensity zone	0.0689	< 2.2e−16***	21.389	0.9239	< 2.2e−16***	10.812
2 intensity zone	0.0573	0.0007405***	3.1784	0.9356	0.004594**	2.605

**p* value < 0.05, ***p* value < 0.01, ****p* value < 0.001

$$\text{Normalized\_sentiment} = \frac{\sum \text{Daily\_average\_sentiment\_of\_individuals}}{\text{Number\_of\_individuals}} \tag{8}$$

$$\text{Normalized\_mobility} = \frac{\sum \text{Radius\_of\_gyration}}{\text{Number\_of\_individuals}} \tag{9}$$

The changes of sentiment and mobility over the study period are plotted in Figs. 3 and 4, respectively. We decomposed the time series based on a moving average method to investigate the trends and seasonal effects (assumed weekly). The additive decomposition within each plot includes the observed time series, trend, seasonal effect, and random variables with mean zero (irregular plot). The unit of time scale is a week. According to the trend plot in the decomposition plots, there is a decreasing trend before the earthquake occurred, and an increasing trend post-disaster for both series.

4.3.2 Stationarity and cointegration

We adjusted the time series by removing the seasonality prior to further analysis. This is necessary to avoid the impact of the intrinsic autocorrelation of the time series and to avoid the false identification of a relationship between sentiment and mobility. We conducted the Augmented Dickey–Fuller (ADF) Test (Said and Dickey 1984) with the null hypothesis being that a unit root is present in a time series sample, and an alternative hypothesis of stationarity. We did not find support to reject the hypothesis in both time series of sentiment S_t (p value = 0.8089, Dickey–Fuller = − 1.4189) and time series of radius of gyration M_t (p value < 0.1442, Dickey–Fuller = − 3.0673). This means we did not find evidence to support the stationarity for both M_t and S_t. Therefore, there is no clear tendency for time series of mobility and sentiment to return to or fluctuate around a constant value or a linear trend.

We then tested if the two adjusted non-stationary time series are cointegrated with the Phillips–Ouliaris Cointegration Test (Phillips and Ouliaris 1990). We found support to reject the null hypothesis that the two series are not cointegrated (Phillips–Ouliaris demeaned = -32.458, p value $\leq$ 0.01). This cointegration means that there exists a linear combination of the two variables that is stationary. In another words, sentiment and mobility share a trend together over time: specifically, a change in one will be permanent

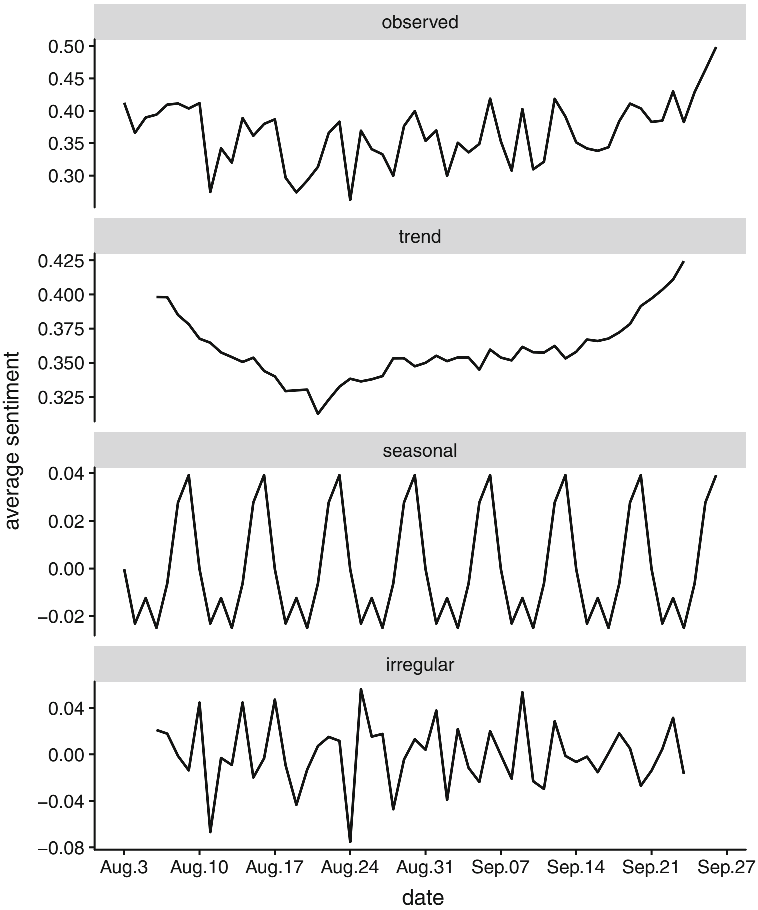


Fig. 3 Additive decomposition of time series of average sentiment

only if both change; an interruption to only one will be meaningless in the long run because it will be pulled back to the long-term path determined by the other one.

4.3.3 Interrupted time series analysis

The parameter estimate from the linear segmented regression model of effects of the earthquake on the mean sentiment of population can be found in Table 3. The fitted model is demonstrated by Eq. 10. The results indicate that just before the study period, the daily average sentiment in the study period was 0.413. Before the earthquake, there was a significant day-to-day change in the value (p value for baseline trend = 6.65e−05). Just after the earthquake, the day-to-day change in sentiment increased by 0.006 statistically (p value for trend change = 2.60e−05), but there was no significant change in the sentiment level. We eliminated the non-significant term and the most parsimonious model includes only intercept, baseline trend and trend change in the daily average sentiment of individuals.

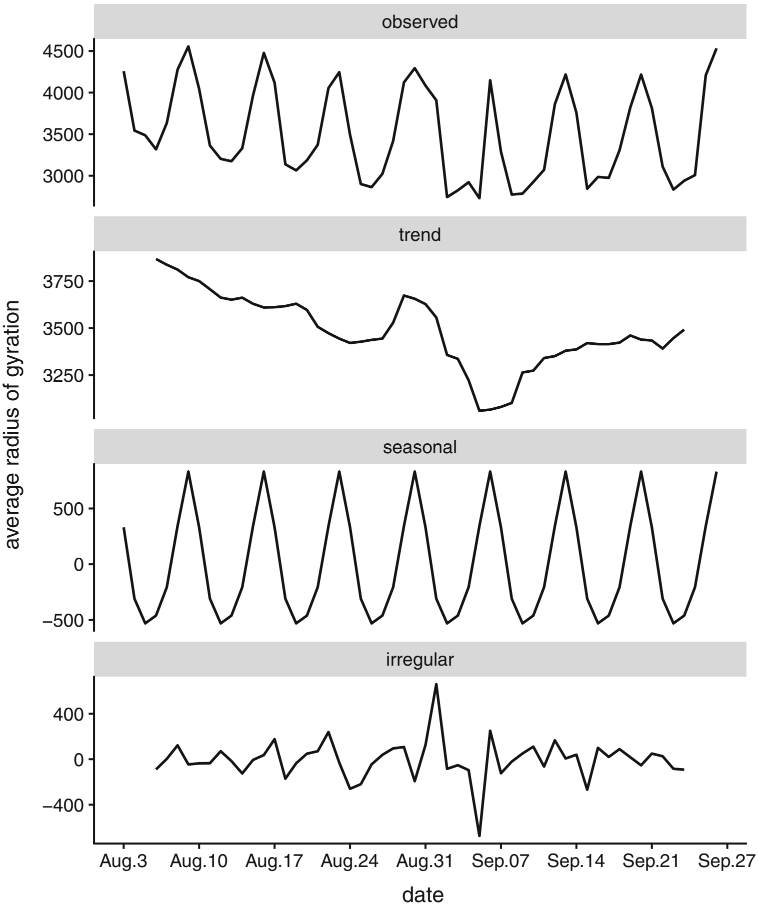


Fig. 4 Additive decomposition of time series of average radius of gyration (mobility)

$$S_t = 0.413 - 0.005 \times \text{time}_t + 0.023 \times \text{disaster}_t + 0.006 \times \text{time\_after\_disaster}_t \quad (10)$$

The parameter estimate from the linear segmented regression model of effects of earthquake on the mean value of the radius of gyration of individuals can be found in Table 4. The fitted regression models are expressed in Eq. 11. The results indicate that just before the study period, the daily average radius of gyration in the study period was 3920.155 m. Before the earthquake, there was a significant day-to-day decreasing trend in the value (p value for baseline trend = 0.0351). After the earthquake, the fitting results show that the day-to-day change in sentiment increases by 17.363 statistically and the level decreases by 110.076, but not significantly for both parameters. We eliminated the non-significant term and the most parsimonious model includes only the intercept and the baseline trend.

$$M_t = 3920.155 - 19.549 \times \text{time}_t - 110.076 \times \text{disaster}_t + 17.363 \times \text{time\_after\_disaster}_t \quad (11)$$

Table 3 Segmented regression model for interrupted time series of sentiment

	Coefficient	Standard error	*t*-statistic	*p* value
a. Full segmented regression model				
Intercept β_0	0.413039	0.013548	30.488	$< 2e{-}16^{***}$
Baseline trend β_1	− 0.004743	0.001079	− 4.396	$6.65e{-}05^{***}$
Level change after EQ β_2	0.022834	0.017154	1.331	0.19
Trend change after EQ β_3	0.006028	0.001286	4.686	$2.60e{-}05^{***}$
Adjusted *R*-squared: 0.2958; *p* value: 0.0002885				
b. Most parsimonious segmented regression model				
Intercept	0.4064146	0.0127058	31.987	$< 2e{-}16^{***}$
Baseline trend	− 0.0038399	0.0008458	− 4.540	$4.04e{-}05^{***}$
Trend change after EQ	0.0056762	0.0012695	4.471	$5.06e{-}05^{***}$
Adjusted *R*-squared: 0.284; *p* value: 0.0001728				

**p* value < 0.05, ***p* value < 0.01, ****p* value < 0.001

Table 4 Segmented regression model for interrupted time series of mobility

	Coefficient	Standard error	*t*-statistic	*p* value
a. Full segmented regression model				
Intercept β_0	3920.155	112.962	34.703	$< 2e{-}16^{***}$
Baseline trend β_1	− 19.549	8.996	− 2.173	0.0351^{*}
Level change after EQ β_2	− 110.076	143.032	− 0.770	0.4456
Trend change after EQ β_3	17.363	10.726	1.619	0.1125
Adjusted *R*-squared: 0.3311, *p* value: 9.459e−05				
b. Most parsimonious segmented regression model				
Intercept	3813.927	73.871	51.63	$< 2e{-}16^{***}$
Baseline trend	− 12.061	2.572	− 4.69	$2.38e{-}05^{***}$
Adjusted *R*-squared: 0.3043, *p* value: 2.383e−05				

**p* value < 0.05, ***p* value < 0.01, ****p* value < 0.001

4.3.4 Cross-correlation

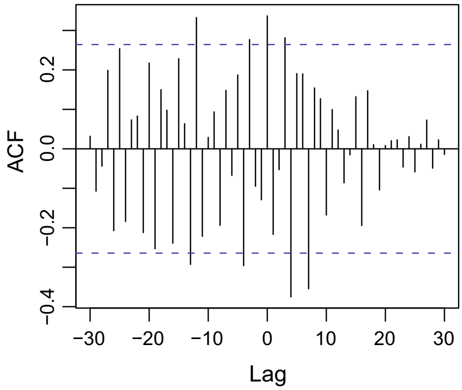
We further computed the first order difference of M_t(Eq. 12) and S_t (Eq. 13), ΔM_t and ΔS_t, which also denote the daily change in radius of gyration and sentiment. Both ΔM_t and ΔS_t are stationary after ADF test with p value < 0.01.

$$\Delta M_t = M_t - M_{t-1} \tag{12}$$

$$\Delta S_t = S_t - S_{t-1} \tag{13}$$

To explore the relationship between ΔM_t and ΔS_t over the time, we employed cross-correlation analysis to examine the relationship between the two time series. Cross-correlogram for ΔM_t and ΔS_t can be found in Fig. 5.

Fig. 5 Cross-correlogram for ΔM_t and ΔS_t



Based on the cross-correlation analysis, the most dominant correlation is -0.554, when lag (h) is -1 or 1. The negative value of lag indicates that ΔS_t leads ΔM_t, and ΔS_t is a predictor of ΔM_t, while a positive value of lag means that changed ΔS_t lags ΔM_t, and ΔM_t can be a predictor of ΔS_t. Therefore, we conclude that daily changed sentiment can either lead or lag daily changed mobility.

5 Discussion

Massive amounts of geocoded data on human–environment interactions have increased the potential for researchers to quantitatively assess dynamic processes over disaster-affected areas. Specifically, geo-referenced tweets allowed us to examine large-scale sentiment and human mobility patterns over the course of a severe earthquake and form "bottom-up" perspectives to understand urban population dynamics during an extreme event perturbation. Our study examined these dynamics before, during and after the earthquake by analyzing the spatial and temporal characteristics, and the relationship between sentiment and mobility. The proposed framework couples collective sentiment and mobility to evaluate the diversity of human–environment interactions and to inform efforts to improve disaster resilience.

Prior research demonstrates some spatial characteristics of sentiment correlate with nominal types of sentiment, and this is used to characterize locations (Bertrand et al. 2013; Lin 2014). Although one study (Neppalli et al. 2017) has evaluated the clustering pattern of positive and negative sentiment during Hurricane Sandy, the spatial patterns of sentiment has not been explored fully in the disaster context in terms of disaster intensity/magnitude and its spatial scale, nor considered other types of natural disaster. Also, the spatial characteristics of sentiment level in a numerical form have yet been examined. Our linear regression analysis of sentiment and earthquake intensity quantitatively reveals the negative correlation between disaster emotion and severity: the higher the earthquake intensity is, the lower the level of the collective sentiment. Different statistics of individuals' daily sentiment have different correlations with intensity level, of which, the average values of mean and sum of individuals' sentiment during the 24 h after earthquake show high correlation with the earthquake intensity. These statistically significant results provide support to accept Hypothesis 1a. With more available earthquake datasets that include both geographical locations and semantic contents, in future research, we will strive to find the best fitting model to explore the relationship between sentiment level and earthquake intensity level, i.e., comparing results from the linear model and nonlinear models, in order to achieve more general conclusions.

In addition, by employing Geary's C and Moran's I, we found evidence that sentiments of similar levels tend to cluster in geographical space, though the spatial autocorrelation is weak, and this spatial association has been found across disaster-affected areas of distinct intensities (Hypothesis 1b). These findings extend the former studies that classified sentiment in earthquakes (Vo and Collier 2013) by including the spatial factor, e.g., the scale of earthquake intensity and the clustering pattern.

Furthermore, employing the interrupted time series model, we descriptively demonstrate the temporal dynamics of human mobility and sentiment with the interruption of the earthquake. Results of statistical tests for the parameters in the model reveal that the earthquake interrupted the time series of sentiment significantly by changing the trend of the time series, while both level and trend of time series of mobility have not been perturbed significantly by the earthquake. As both fitted models for sentiment and mobility are statistically significant, we found evidence to support Hypothesis 2a regarding the interruption of the earthquake, although with deviations for level and trend changes for sentiment and mobility.

We further investigated the relationship between sentiment and mobility over time (Hypothesis 2b) and found that the time series are cointegrated, which indicates their co-evolution over time. Moreover, we expected that Δsentiment and Δmobility were cross-correlated over time, and Δsentiment was a predictor of Δmobility (Hypothesis 2c). Our analysis results found support for the hypothesis that Δsentiment can lead the change of Δmobility; however, we also found that Δsentiment can lag the change of Δmobility. This first effort to examine disaster dynamics over time by coupling sentiment and mobility contributes a new, expanded and quantitative understanding of these dynamics. These findings also extend former studies (Frank et al. 2013) regarding the relationship between mobility pattern and sentiment to the disaster context.

However, there exist some limitations in this study in terms of data characteristics, geographical scales, and sentiment methods. We exclusively analyzed English tweets due to the unequal development of methods for analyzing sentiment in other languages and the dominant role of English language in studied area. The research results should be generalized to the diverse-language-speaking population with caution because the demographic structure of users posting the English tweets is unknown. Fortunately, as the English tweets occupy nearly 90% of collected geo-referenced daily tweets in our sample, the results are able, by and large, to reveal the urban population dynamics in the disaster-affected area.

Our studied area is the spatial bounding box of a 2.5 intensity line and includes uneven distribution of urban areas in different intensity polygons. Diverse geographical scales of earthquakes and other natural disasters may lead to dissimilar effects on sentiment levels and human mobility patterns. We plan to address these differences among multiple types of disasters and disasters of distinct geographical scales in future studies to achieve a more nuanced understanding of the spatiotemporal dynamics of resilience. As this paper focuses more on the collective influence of a disaster at a large scale, our analysis is adequate in terms of examining the proposed hypotheses.

Additionally, with the development of methods for sentiment analysis in specific domains, especially in the context of natural disasters and extreme events, further studies can reveal more practical information in terms of contents and sentiment levels for targeted disaster topics. We collectively analyzed the sentiment of geotagged tweets in a disaster-affected area to achieve a broad understanding of the influence of an earthquake in different intensity zones. Our next efforts will also focus on developing disaster-specific lexicon generated from social media to classify tweets to more specific emotion types based on disaster psychology.

6 Conclusion

This study expands our understanding of disaster resilience and urban dynamics with crowd-sourced data from a social media platform. It examined hypotheses of spatial characteristics of sentiment before, during and after a severe earthquake. The results uncovered a significant negative correlation between sentiment levels and earthquake intensity levels and demonstrated that sentiment tends to cluster in space in distinct earthquake intensity zones. Moreover, the study investigated the temporal relationship of sentiment and human mobility, including the dynamic effects of the earthquake over time. The time series of radius of gyration and sentiment exhibited co-movement over time, and Δsentiment can either lead or lag the change of Δmobility in the disaster context. We hope to extend the proposed research framework on other types of disasters to generalize findings of the relationship between disaster magnitude and sentiment levels and correlation between sentiment and mobility, and to utilize the framework to evaluate the dynamic process of disaster resilience at different spatial scales. With more specific "small data", e.g., government strategies and disaster characteristics, the research findings based on "big data" can provide "bottom-up" knowledge to facilitate disaster informatics and management in terms of monitoring and evaluation in the built environment.

Acknowledgements This study is supported by the National Science Foundation under Grant No. 1760645 and a Virginia Tech BioBuild IGEP grant. 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 or Virginia Tech.

References

Bai H, Yu G (2016) A Weibo-based approach to disaster informatics: incidents monitor in post-disaster situation via Weibo text negative sentiment analysis. Nat Hazards 83:1177–1196

Beigi G, Hu X, Maciejewski R, Liu H (2016) An overview of sentiment analysis in social media and its applications in disaster relief. In: Sentiment analysis and ontology engineering. Springer, pp 313–340

Bertrand KZ, Bialik M, Virdee K, Gros A, Bar-Yam Y (2013) Sentiment in New York City: a high resolution spatial and temporal view. arXiv preprint arXiv:13085010

Cody EM, Reagan AJ, Mitchell L, Dodds PS, Danforth CM (2015) Climate change sentiment on twitter: an unsolicited public opinion poll. PLoS ONE 10:e0136092

EERI (2014) EERI Special Earthquake Report: M 6.0 South Napa Earthquake of August 24, 2014. USA Earthquake Clearinghouse

Ford JD et al (2016) Opinion: big data has big potential for applications to climate change adaptation. Proc Natl Acad Sci 113:10729–10732

Frank MR, Mitchell L, Dodds PS, Danforth CM (2013) Happiness and the patterns of life: a study of geolocated tweets. Sci Rep 3:2625. https://doi.org/10.1038/srep02625

Geary RC (1954) The contiguity ratio and statistical mapping. Inc Stat 5:115–146

Godschalk DR (2003) Urban hazard mitigation: creating resilient cities. Nat Hazards Rev 4:136–143

Gonzalez MC, Hidalgo CA, Barabasi A-L (2008) Understanding individual human mobility patterns. Nature 453:779–782

Granger CW (1981) Some properties of time series data and their use in econometric model specification. J Econom 16:121–130

Guan X, Chen C (2014) Using social media data to understand and assess disasters. Nat Hazards 74:837–850

Kryvasheyeu Y, Chen H, Obradovich N, Moro E, Van Hentenryck P, Fowler J, Cebrian M (2016) Rapid assessment of disaster damage using social media activity. Sci Adv 2:e1500779

Lin Y-R (2014) Assessing sentiment segregation in urban communities. In: Proceedings of the 2014 international conference on social computing. ACM, p 9

Lu X, Bengtsson L, Holme P (2012) Predictability of population displacement after the 2010 Haiti earthquake. Proc Natl Acad Sci 109:11576–11581

Metcalfe AV, Cowpertwait PSP (2009) Introductory time series with R. Springer, New York. https://doi.org/10.1007/978-0-387-88698-5

Mitchell L, Frank MR, Harris KD, Dodds PS, Danforth CM (2013) The geography of happiness: connecting twitter sentiment and expression, demographics, and objective characteristics of place. PLoS ONE 8:e64417

Moran PA (1950) Notes on continuous stochastic phenomena. Biometrika 37:17–23

Munich RE (2017) NatCatSERVICE: number of catastrophic natural loss events worldwide 2010–2016. http://natcatservice.munichre.com/. Accessed 17 Aug 2017

Nagy A, Stamberger J (2012) Crowd sentiment detection during disasters and crises. In: Proceedings of the 9th international ISCRAM Conference, pp 1–9

Neppalli VK, Caragea C, Squicciarini A, Tapia A, Stehle S (2017) Sentiment analysis during Hurricane Sandy in emergency response. Int J Disaster Risk Reduct 21:213–222

Nielsen FÅ (2011) A new ANEW: evaluation of a word list for sentiment analysis in microblogs. arXiv preprint arXiv:11032903

Pavalanathan U, Eisenstein J (2015) Confounds and consequences in geotagged Twitter data. arXiv preprint arXiv:150602275

Pew Research Center (2017) Social media fact sheet. http://www.pewinternet.org/fact-sheet/social-media/. Accessed 11 Apr 2017

Phillips PC, Ouliaris S (1990) Asymptotic properties of residual based tests for cointegration. Econom J Econom Soc 58:165–193

Ribeiro FN, Araújo M, Gonçalves P, Benevenuto F, Gonçalves MA (2015) SentiBench-a benchmark comparison of state-of-the-practice sentiment analysis methods. arXiv preprint arXiv:151201818

Said SE, Dickey DA (1984) Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 71:599–607

Swier N, Komarniczky B, Clapperton B. (2015) Using geolocated Twitter traces to infer residence and mobility. Office for National Statistics GSS Methodology Series, 41

Tang Z, Zhang L, Xu F, Vo H (2015) Examining the role of social media in California's drought risk management in 2014. Nat Hazards 79:171–193

UNISDR (2017) Disaster statistics. https://www.unisdr.org/we/inform/disaster-statistics. Accessed 2 May 2017

USGS (2014) M 6.0—6 km NW of American Canyon, California. https://earthquake.usgs.gov/earthquakes/eventpage/nc72282711#executive. Accessed 11 Apr 2017

USGS (2017) M 6.0—6 km NW of American Canyon, California. https://earthquake.usgs.gov/earthquakes/eventpage/nc72282711#shakemap. Accessed 9 Apr 2017

Vo B-KH, Collier N (2013) Twitter emotion analysis in earthquake situations. Int J Comput Linguist Appl 4:159–173

Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D (2002) Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther 27:299–309

Walther M, Kaisser M (2013) Geo-spatial event detection in the twitter stream. In: European Conference on Information Retrieval. Springer, pp 356–367

Wang Q, Taylor JE (2014) Quantifying human mobility perturbation and resilience in Hurricane Sandy. PLoS ONE 9:e112608

Wang Q, Taylor JE (2015) Process map for urban-human mobility and civil infrastructure data collection using geosocial networking platforms. J Comput Civ Eng 30:04015004

Wang Q, Taylor JE (2016) Patterns and limitations of urban human mobility resilience under the influence of multiple types of natural disaster. PLoS ONE 11:e0147299

Wang B, Zhuang J (2017) Crisis information distribution on Twitter: a content analysis of tweets during Hurricane Sandy. Nat Hazards 89:1–21

Wang Z, Ye X, Tsou M-H (2016) Spatial, temporal, and content analysis of Twitter for wildfire hazards. Nat Hazards 83:523–540

Wang Y, Wang Q, Taylor JE (2017) Aggregated responses of human mobility to severe winter storms: an empirical study. PLoS ONE 12(12):e0188734. https://doi.org/10.1371/journal.pone.0188734