ICONHIC2019
2nd International Conference on Natural Hazards & Infrastructure
23-26 June, 2019, Chania, Greece

Using Social Media to Assess Earthquake Impact on People and Infrastructure: Examples from Earthquakes in 2018

D. Zekkos[1]
University of Michigan, Ann Arbor, USA & ARGO-E LLC

A. Tsavalas-Hardy, G. Mandilaras, K. Tsantilas
Elxis Group, Athens, Greece & ARGO-E LLC, Ann Arbor, MI, USA

ABSTRACT

Social media activity following the occurrence of an earthquake has the potential to provide valuable data on infrastructure condition and for disaster response and reconnaissance purposes. The challenge however is that the data with valuable content (i.e., the signal) is small compared to the earthquake-related social media activity overall (i.e., the noise). We explore Twitter activity related to earthquakes throughout 2018 and also analyze selected earthquakes with the intent to explore social media dynamics and compare the observed patterns to the characteristics of the earthquake event. We find that earthquake-related social media activity, amounting to more than 4 million tweets in 2018, directly relates to earthquake activity. Social media activity associated with a specific event is affected by the earthquake magnitude, and other social factors. However, the most important factor is the earthquake intensity with respect to population density. Even relatively small events (e.g., $4<M_w<5$) in urban areas, had significant social networking response. Damaging earthquakes have a continued social media response compared to events that did not cause significant infrastructure damage. Social media activity was also spatially correlated to earthquake activity. Analyses using machine learning demonstrated that there is significant content related to the impact of the earthquake on infrastructure that can be used to pinpoint expected damage in the immediate aftermath of an earthquake.

Keywords: earthquakes, social media, twitter, infrastructure, impact

INTRODUCTION

Social media networks are changing the way people communicate and exchange ideas, and already have had significant impact on the way information is disseminated and used. Social networking activity following an earthquake has the potential to provide immediate and large amounts of data that can be of value for earthquake response and reconnaissance purposes and for infrastructure condition assessment purposes. In that context, humans can act as in-situ observers that are spatially distributed and are also potentially positioned in the most critical areas. These humans can provide information from the affected areas that may not be available otherwise. Identifying and querying this information can be a challenge as the data that needs to be analyzed to identify the valuable content (i.e., the "signal") is large. In addition, what is considered valuable content largely depends on the perspective and objectives of the investigation. For example, information that may be

[1] Corresponding Author: D. Zekkos, *University of Michigan, zekkos@geoengineer.org*

of interest to someone focused on immediate earthquake response with the objective to save lives in the immediate aftermath of an earthquake, is different from that of someone who is looking to assess infrastructure condition.

A number of studies have been conducted to collect social media data associated with natural disasters (e.g., Kryvasheyeu et al. 2016, Yin et al. 2015, among others). As pointed out by Murthy and Gross (2017), the majority of these studies have focused on first responders and relief organizations. Kaigo (2012) examined the case of the city of Tsukuba in Ibaraki prefecture, where power outage during the 2011 Great East Japan earthquake resulted in use of Twitter for communication and the dissemination of vital information during the disaster. Lachlan et al. (2014) investigated the characteristics of tweets during Hurricane Sandy. The authors collected tweets at specific time points and found that the tweet rate increased during the storm. Government and organizational responses were largely absent. The authors also found that Twitter was used more for emotional release than to provide information. Gao et al. (2011) examined the characteristics and challenges of social media-based crowdsourcing data for disaster relief following the 2010 Haiti earthquake and 2011 Japan earthquake and tsunami. Middleton et al. (2014) found based on work done for Hurricane Sandy and a tornado event in Oklahoma in 2013, that it is feasible to obtain high-precision geoparsing from real-time Twitter data by exploiting large databases of preloaded location information for at-risk areas. Dashti et al. (2014) explored the type of information that can be collected for post-event reconnaissance following the September 2013 floods in Colorado. Sutton et al. (2014) collected all public tweets sent by official government accounts during the Waldo Canyon wildfire and suggested strategies for designing and disseminating messages through networked social media under periods of imminent threat. Murthy and Longwell (2013) studied the social media dynamics during the 2010 Pakistani floods and found differences between users of social media in the western societies and Pakistan. Kim and Hastak (2018) provided insights on the critical role of social media use for emergency information propagation using the 2016 Louisiana floods as examples.

In this study, we investigate social media activity related to earthquakes for an entire year. We explore social media dynamics throughout the year and immediately after earthquakes that occurred in 2018 with the intent to understand the type of social media content that is generated following an earthquake. We explore social media trends observed following major earthquakes and compare them to the characteristics of the earthquake events. We identify challenges associated with taking full advantage of the data. Finally, we use machine learning approaches to characterize the impact of earthquakes on infrastructure based on the social networking data.

METHODOLOGY

During 2018, we have investigated social media activity associated with earthquakes using the Twitter social network. We developed a platform that allows the collection, querying and analysis of all tweets associated with earthquakes. We used Twitter's streaming API to collect in real time tweets containing the word "earthquake" written in either lowercase or uppercase characters or having the "#earthquake" hashtag. Tweets are provided as JSON objects that include the tweet content (including media attached to the tweet) as well as related information, such as retweets, replies, and a timestamp for each tweet. These are then stored in a MySQL database for retrieval. Subsequent analysis of the tweets for specific earthquake events was conducted using Python.

Initial analyses were conducted heuristically for two earthquake events in January. Through this analysis, valuable lessons were learned about the type of content collected. Machine learning techniques were subsequently implemented to analyze and classify larger volumes of tweets for events that occurred later in the year. The machine learning techniques used to classify tweets were also tested against heuristic classifications. The methodology involved the following steps: First, blacklists of known users were created based on their Twitter identity. This allowed removal of spam from the analyzed data. Text-based classification was then executed for the remaining tweets using developed vocabularies and a scoring associated with specific words. The developed classes are not mutually exclusive, so a tweet may belong to one or multiple classes. The following classes were created:

(a) "Automated" class: This class included automated tweets from known machines or sensors transmitting data through a Twitter account. Many of these tweets are commonly associated with notifications of earthquake occurrence;
(b) "Impact" class: This class is the primary goal of this study and includes information on the impact of the earthquake on the affected area;

(c) "Felt Intensity" class: This class includes tweets posted by people who experienced the earthquake and shared their experience;
(d) "Supporting message" class: This class includes tweets with a supportive or positive message towards people affected by an earthquake and tweets associated with aid (e.g., donations, support, Red Cross activities);
(e) "Funny" tweets: This class includes messages with funny or entertaining content;
(f) "Undetermined" class: These are tweets that do not belong to the above classes. They may still include some useful information, but were more difficult to classify in classes. Additional work is being conducted to further improve our methodology to assign these tweets to appropriate classes.

As described subsequently, additional analyses were also conducted in the "Impact" class, which is the focus of this study with the aim to get more quantitative information on the affected infrastructure and natural hazards involved.

VOLUME OF SOCIAL MEDIA ACTIVITY ASSOCIATED WITH EARTHQUAKES

Fig. 1 shows a plot of daily tweet activity related to earthquakes for 2018. In total, 4,380,121 earthquake-related tweets were published in 2018, resulting in an average of 12,112 tweets/day. As shown in Fig. 1, major earthquake events are identifiable as distinct "spikes" in tweet activity. Note also that for a reason that is not yet well understood we were not able to collect all tweets for a period from May 17th to July 4th. During that approximately two-month "downtime" period, the number of tweets that we were able to collect was limited to about 1,000 per day.

According to the USGS earthquake website (https://earthquake.usgs.gov/), the first 5 months of the year (January 1st to May 17th) 42 major ($M_w \geq 6$) earthquakes occurred. From July 4th to the end of the year, i.e., the last 6 months, there were 90, i.e., nearly twice the earthquake activity. Only two major earthquakes occurred during the "downtime" period. The higher earthquake activity in the second part of the year is also reflected in the Twitter activity of Fig. 1 where the second part of the year is clearly significantly more active than the first, providing a first indication that Twitter activity is correlated to earthquake activity. Table 1 also lists the major Twitter activity spikes in Fig. 1. All occur, immediately following major earthquakes, or in three cases only, smaller earthquakes in urban areas in California (1, 4, and 8). It is interesting to note that these smaller earthquakes, despite not being damaging to infrastructure, resulted in a significant reaction by people who felt them.

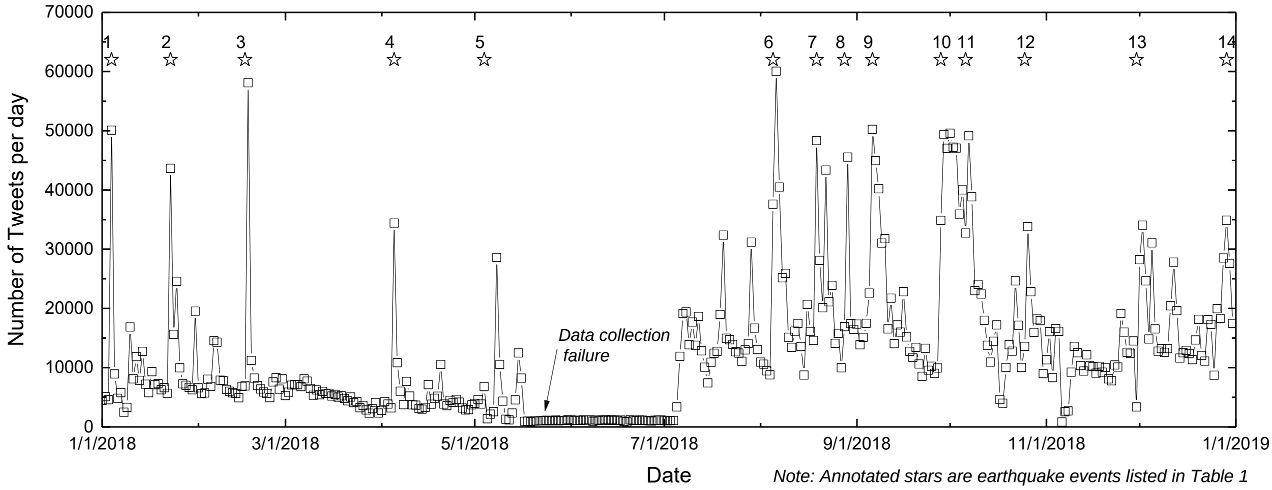

Figure 1. Number of tweets per day for the year 2018.

Table 1. Earthquakes resulting in greatest number of tweets in 2018

ID	Date	Location	Country	Magnitude
1	1/4/2018	Berkeley	USA	4.4

2	1/23/2018	Gulf of Alaska	USA	7.9
3	2/16/2018	Oaxaca	Mexico	7.2
4	4/5/2018	South California (offshore)	USA	5.3
5	5/4/2018	Hawaii	USA	6.9
6	8/5/2018	Lombok	Indonesia	6.9
7	8/19/2018	Lombok aftershock	Indonesia	6.3
8	8/28/2018	La Verne, South California & New Caledonia	USA &West Pacific	4.4 & 7.1
9	9/6/2018	Hokkaido	Japan	6.7
10	9/28/2018	Sulawesi & aftershocks & tsunami	Indonesia	7.5
11	10/6/2018	Haiti	Haiti	5.9
12	10/25/2018	Zakynthos	Greece	6.8
13	11/30/2018	Anchorage, Alaska[1]	USA	7
14	12/29/2018	Mindanao	Philippines	7

[1]*Data was lost for a limited time (~ 2hrs) shortly after the earthquake; number of tweets shown in Fig. 1 is lower than actual.*

Fig. 2 shows Twitter activity related to earthquakes for January 2018. On average, for the days when seismic activity is low, Twitter activity includes about 2500-5000 tweets per day. The January 4th 2018 Berkeley event and the January 23rd 2018 Gulf of Alaska event stand out as distinct spikes in Twitter activity. The January 4th 2018 event in Berkeley was a small earthquake (M_w=4.4) with an epicenter essentially on an urban area. Although the magnitude of the earthquake was small and did not cause any damage, the shaking was felt by a large population resulting in a Twitter activity that was about 10 to 20 times greater than the tweet activity before the earthquake. The January 23rd earthquake event was a major event (M_w=7.9) that occurred in the Gulf of Alaska. The earthquake epicenter was approximately 350 miles southwest from Anchorage, the main high density population area. The closest populated area, Kodiak island, was still about 175 miles away. As a result, social media activity for this earthquake, despite its much greater magnitude, was lower than the Berkeley one. Social media activity was still significant, partly due to a Tsunami Watch issued for California, Oregon, and Washington by NOAA's National Tsunami Warning Center. The tsunami ended up being small (less than 30 cm in maximum wave height and in many places much smaller). Following the arrival of the tsunami on the coastline, Twitter activity dropped significantly, since there was no other damage or consequences from that earthquake.

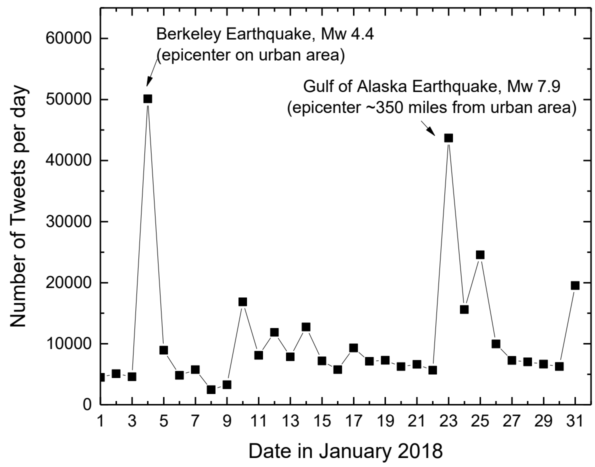

Figure 2. Daily twitter activity associated with earthquakes in January 2018.

Interestingly, on November 30th 2018, another major earthquake event, with a magnitude of 7.0, took place in Anchorage. This earthquake caused significant damage particularly in terms of roadways, pipelines and other infrastructure. Social media activity associated with this event appears to be comparable to the January Gulf of Alaska event. This is caused by interruptions in the data streaming that occurred shortly after the earthquake and at the time of peak activity, significantly reducing the collected tweets for the second event. The November 30th 2018 event shows continued social media activity as more information about damage to infrastructure and the consequences to people was becoming available. Overall, the activity associated with this event was higher than the January 23rd Alaska event, although the affected population was practically the same.

In general, it appears that Twitter activity associated with specific earthquake events follows a trend that can be described by the relationship shown in Fig. 3a. Fig. 3a is a conceptual figure where the x-axis represents time (expressed in minutes, hours or days), and the y-axis is the number of tweets posted during that time. Immediately (i.e., 0-1 hrs) following the earthquake, a large increase in the number of tweets is observed. The number of tweets reaches a maximum within the first hour, but in some cases a few hours, and then is followed by a gradual decline in the number of tweets with time, going back to "background" tweet activity. Aftershocks, particularly major aftershocks will disturb the trend of Fig. 3a, since they cause a new rise in Twitter activity. Example earthquake events from 2018 are shown in Figure 3b for the first 96 hrs (4 days) after the earthquake. All of them generally fit the conceptual Fig. 3a, but the peak and rate of decrease in activity varies.

Peak activity associated with an earthquake event is influenced by a number of factors. First, larger magnitude earthquakes cause higher peak twitter activity compared to smaller magnitude earthquakes. Larger Magnitude events (M_w 7+) that occurred in 2018 were followed by about 40,000-50,000 tweets during the first day. The numerous smaller events (M_w 4.0-) that occur globally, represent the background activity of about 2,500-5000 tweets. Intermediate magnitude events have tweets in the order of 10,000-40,000. However, other factors affect peak activity. Probably the most critical factor is the spatial distribution of seismic intensity compared to densely populated areas. Densely populated areas hit by smaller earthquakes, such as events 1, 4 and 8 of Table 1 result in a significant increase in social media activity as people feel the need to communicate that they experienced an earthquake. In addition, other social factors such as the availability of internet access, mobile phones, how much Twitter is used by the affected population, as well as the local language used by the population influences Twitter results. In this study, Twitter activity in the English language is analyzed. This activity represents only a portion of the total activity in areas where English is not the main language of communication.

The duration of social media activity associated with an earthquake event appears to be largely a function of the consequences of that earthquake to society. Twitter activity following earthquakes (small or large) without significant consequences to people and infrastructure, such as the January Berkeley (#1 in Table 1) or Gulf of Alaska event (#2 in Table 1) drops sharply following its peak, as shown in Fig. 2 and 3b. Earthquakes with more significant consequences on infrastructure such as the Hokkaido and Lombok event have continued social media activity, as shown in Fig. 3b, as more information is becoming available about the event and people still experience the consequences of the event.

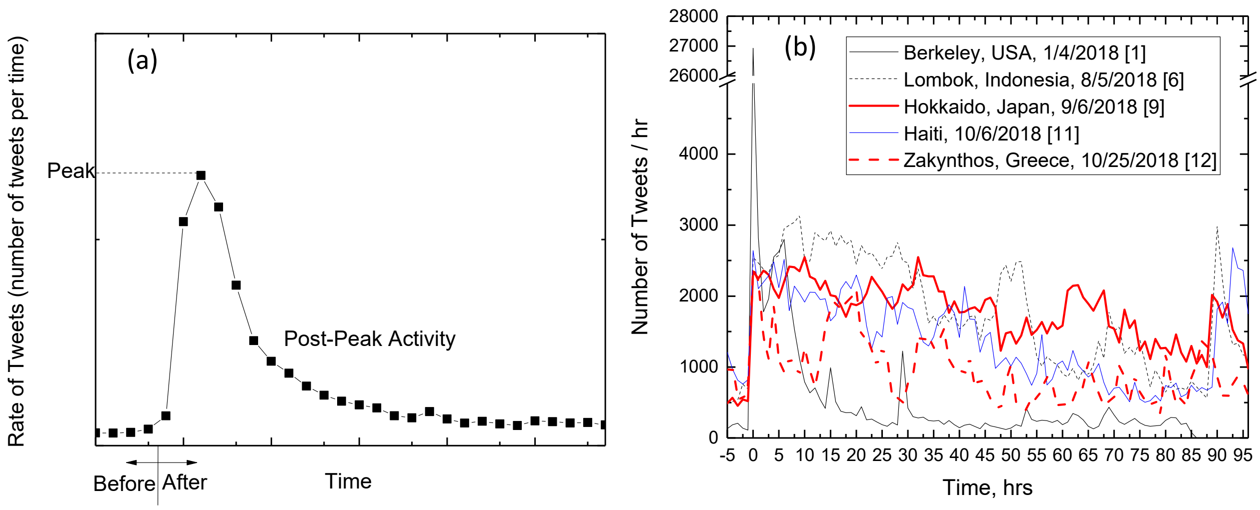

Figure 3. (a) Concept figure illustrating social media activity before and after an earthquake; (b) Examples of social media activity following specific 2018 earthquakes (t=0 hrs is the time of earthquake occurrence).

CHARACTERISTICS OF EARTHQUAKE-RELATED SOCIAL MEDIA ACTIVITY

As described earlier, on any single day, there is significant activity on Twitter related to earthquakes. For 2018, in the most "quiet" of days, when no significant seismic activity occurs, there are still 2,500-5,000 tweets related to earthquakes. When there is increased seismic activity, the daily number of tweets increases significantly and then attenuates in the manner described earlier. Part of our investigation aimed at gaining an understanding of the tweet content and some results are presented subsequently.

Contributors to Social Media Earthquake-related Activity

In general, there are three main contributors to earthquake-related social media activity:

(a) Automatic tweets generated by machines, sensors, or specific platforms that post about earthquake activity, or provide warnings or other notifications. Among those, the most prolific ones are the Twitter accounts @everyEarthquake and @QuakesToday as well as numerous other ones that are regional and tweet the epicenter and characteristics of earthquakes occurring around the world. Other platforms may retweet or create a separate tweet associated with an earthquake. Local agencies or foreign earthquake monitoring agencies may do the same also.

(b) Individuals or groups who are either interested in or experienced earthquakes. These may represent scientists and engineers with an interest in earthquakes, but also policy makers, or other groups that could be considered stakeholders. Individuals who experienced an earthquake also tweet about it and describe what they felt or what they saw. This information is probably of the most value for the purposes of post-earthquake reconnaissance and infrastructure assessment. Additionally, other individuals may participate, primarily to provide support and encouragement or communicate their own feelings. The latter becomes common in major events. These individuals are also more likely to retweet or reply to previous tweets.

(c) Spammers and irrelevant topics represent a third group. Spammers may be individuals or machines who aim to promote something irrelevant to the topic of earthquake, but take advantage of the up-rise in earthquake-related social media activity to promote products or services. Such products may include commercial products (phones, sunglasses), or services (insurance). In addition, some activity is irrelevant, using earthquakes in a different context (e.g., pokemons creating earthquakes). This group is generally small and is most commonly identified by the handler's account.

Tweet Characteristics

In general, there are three types of tweets: original, retweets or replies. In 2018, we collected a total of 4,380,121 tweets that consisted of 2,107,562 original tweets, 2,073,368 retweets and 199,191 replies. The replies are far fewer than original tweets and retweets, whereas original tweets and retweets each represent about 50% of the activity. However, as shown in Table 2, when examining specific events, it appears that this observation is not valid. Specifically, for the August 5th 2018 M_w 6.9 earthquake in Indonesia (#5 in Table 1), and the September 5th 2018 M_w 6.6 Hokkaido earthquake (#9 in Table 1) that we examined more closely, original tweets were about a fourth (~20-25%) of total tweet activity, and retweets represented about three quarters (~70-75%) of the activity. Replies to tweets still represented a small part of total tweet activity (<5%). It is important to note that the Twitter Streaming API limits the ability to collect data beyond 1% of total Twitter activity at any given time. In cases where tweet activity is higher than the allowed threshold, the collected tweets represent a smaller part of the total population. Although Twitter does not notify us when the threshold was reached, it is typically relatively easy to notice in the data, because subsequently collected retweets, refer to original tweets that were not previously collected. Overall, this did not seem to be a critical issue during the study period, but it may be an important issue in the case of a major US earthquake impacting a densely populated area.

Geographic Distribution of Earthquake-Related Tweet Activity

Tweet metadata may include the exact location of the Twitter user when the tweet is posted. For a number of reasons, probably the most common being privacy, the vast majority of tweets are set so that location is not included in the tweet. Specifically, out of 2,107,562 tweets, about 78,000 were geotagged, but the majority of these tweets were automated. Further filtering to remove the automated tweets indicated that only 2,700 had locations and were likely to have been posted by individuals, representing 0.13% of the total. This is a very small portion of the tweets, which is unfortunate, because having the location of the tweet is particularly valuable for infrastructure assessment and reconnaissance purposes. Still, due to the large number of tweets involved, it was worthwhile to explore the spatial characteristics of these tweets. Fig. 4a illustrates the location of M_w 4.5+ earthquakes in 2018 from the USGS website. As expected, the earthquakes line up nicely with the tectonic plate boundaries. Fig. 4b illustrates the location of all tweets in 2018 and Fig. 4c the location of tweets that we believe were posted by humans only. It is clear, that the tweets are spatially coincident with earthquake activity and have a higher concentration in areas of high population density.

Table 2. Earthquake and social media characteristics of two events

Location	Hokkaido, Japan	Lombok, Indonesia
Date and Time	September 5th 2018, 18:07:58 UTC	August 5th 2018, 11:46:37 UTC
Magnitude	6.6	6.9
Effects on people	41 killed, 680 injured[1]	563 killed, 7,000+ injured, 431,436 displaced[2]
Total Number of Tweets Analyzed	117,810	92,185
Original Tweets	30,487 (26%)	23,211 (25%)
Retweets	83,686 (71%)	67,159 (73%)
Replies	3,637 (3%)	1,815 (2%)
% of tweets in English language	66%	75%

[1]Pager page: https://earthquake.usgs.gov/earthquakes/eventpage/us2000h8ty/impact (accessed Feb. 7 2019)
[2]Pager page: https://earthquake.usgs.gov/earthquakes/eventpage/us1000g3ub/impact (accessed Feb. 7 2019)

Figure 4. (a) Location of M_w 4.5+ earthquake epicenters in 2018 (source: https://earthquake.usgs.gov/); (b) Location of ~78,000 geotagged tweets in 2018; (c) Location of ~2,700 geotagged tweets after elimination of known automated tweets in 2018

Classification of Social Media Activity using Machine Learning

After heuristically analyzing subsets of tweets for various events, due to the large number of tweets associated with each event, machine learning approaches were implemented for the analyses. Tweets from two main earthquake events were analyzed to assess their content: The August 5th 2018 M_w 6.9 Lombok earthquake in Indonesia, and the September 5th 2018 M_w 6.6 Hokkaido earthquake in Japan. After separating original tweets from retweets and replies, the original tweets were further analyzed, since these were deemed more likely to have "original" content that would be relevant to the earthquake. Analyses of the original tweets involved the classification of tweets in the classes described earlier. For the classification, three different machine learning classifiers were used: Random Forest (RF), Dense Neural Network, and Embedding Neural Network.

The results are summarized in Table 3 for the two events using the Random Forest classifier. As shown in Table 3, the class associated with "impact" is the largest, representing 28% and 46% of the total volume of tweets for the Hokkaido and Lombok event respectively. Tweets associated with the intensity of the earthquake were classified in the "Felt Intensity" class and represented 10% of the tweets. Other classes, such as the "Supporting Messages" were about 10-15%. About 25-29% of the tweets remain unclassified using our methodology. Although, we expect that improvements could be made to classify the unclassified tweets, at this stage, our main focus was on the content of the tweets related to the impact of the earthquake.

Table 3. Classification of original tweets for two earthquake events

Class of Tweet	M_w 6.6 Hokkaido, Japan September 5 2018	M_w 6.9 Lombok, Indonesia August 5 2018
Impact	28%	46%
Supporting Messages	10%	14%
Felt Intensity	10%	10%
Funny Messages	0.3%	0.25%
Undetermined	29%	25%

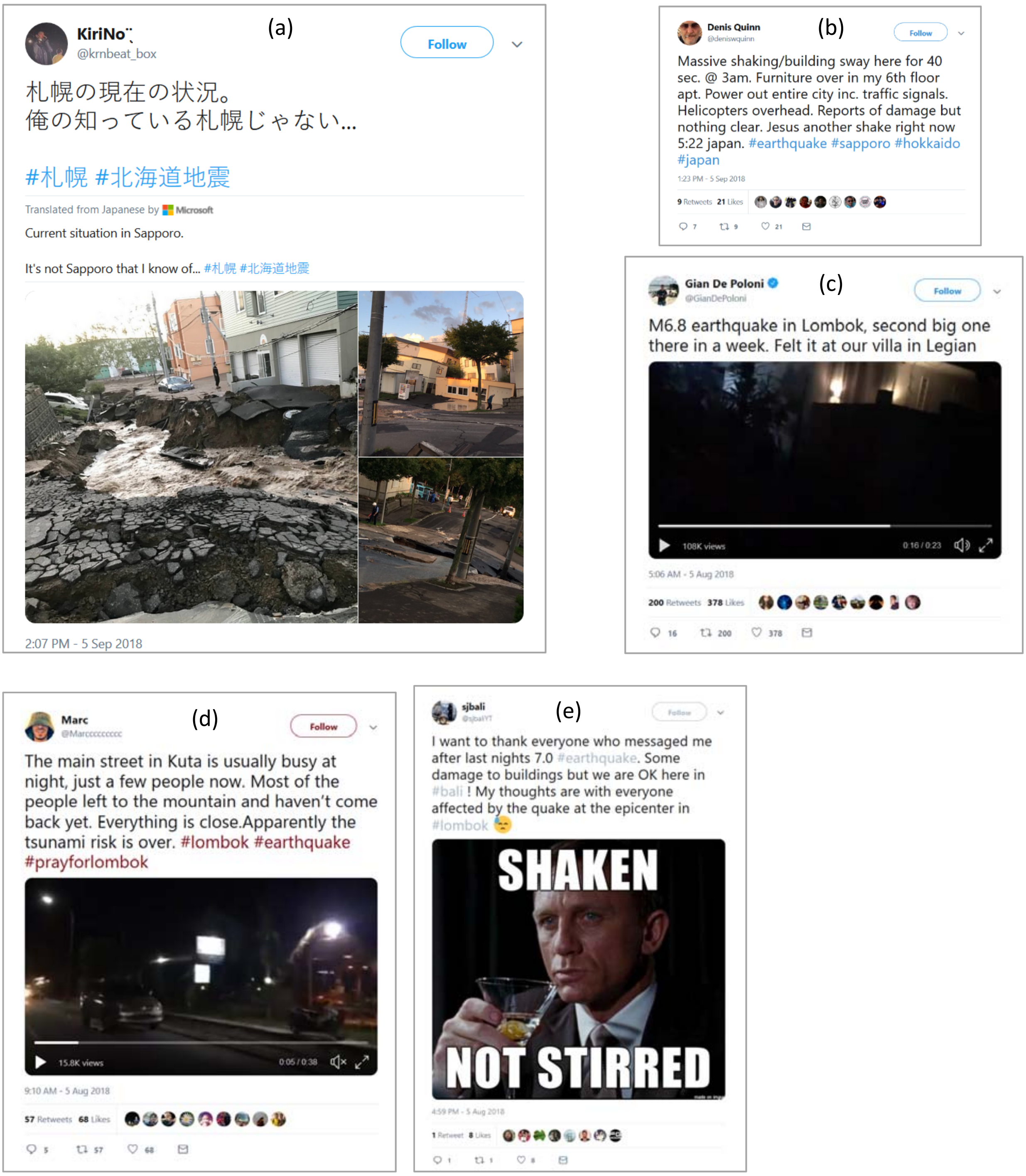

Figure 5. Example of tweets from (a, b) the Hokkaido earthquake in Japan; (c-e) the Lombok earthquake in Indonesia that have useful content related to the impact of the earthquake on infrastructure.

Tweet Content Related to Earthquake Impact

As mentioned earlier, the objective of this study is to identify content that is truly valuable in assessing the impact of an earthquake on infrastructure. Fig. 5 illustrates examples of tweets that were classified as "Impact" of an earthquake and were posted by individuals, and not local or international news media (e.g., Reuters, CNN, etc.). It is important to note that this valuable content is not accessible elsewhere, as it did not attract the attention of media. Fig. 5a and 5b are tweets from the Hokkaido earthquake, whereas Fig. 5c-e are from the Lombok earthquake in Indonesia. The content was generated nearly immediately after the earthquake (within hrs) by individuals who happened to be in the area affected by the earthquake. They provide an assessment of the conditions (Fig. 5b and 5e), photos (Fig. 5a) and video documentation (Fig. 5c-d). Note that the content generated is not particularly popular, i.e., in some cases a handful of reactions occur on this content. However, for scientific purposes, being able to have immediate access to that information is truly valuable.

Classification of Tweets Related to Earthquake Impact on Infrastructure

A word cloud analysis of the tweets in the class "Impact" was conducted and an example is shown in Fig. 6 for the Hokkaido earthquake and the Lombok earthquake. Tweets that were classified in the "Impact" class, were further classified in sub-classes using vocabularies that we developed. Each sub-class has its unique vocabulary that contains words that are relevant to it. A grading scheme was developed by which a tweet is classified in the sub-classes below based on the number and type of words that fit its vocabulary, but a tweet can also belong in more classes if it has strong enough vocabulary resemblance. Tweets that do not have any resemblance with the vocabulary of a sub-class remain "Unclassified". The following sub-classes were used:

- Natural hazards: Tweets on the type of natural hazard such as tsunamis, landslides and aftershocks;
- Impact on infrastructure: Tweets on the consequences of an earthquake on infrastructure;
- Casualties: Tweets on the consequences of the earthquake on people;
- Policy: Tweets on the economic and policy consequences from an earthquake;
- Unclassified: Tweets not classified in the above sub-classes.

The results of the classification of the "Impact" tweets are shown in Table 4. Not surprisingly, tweets related to casualties and natural hazards are the largest. A smaller subset belongs to the category "impact on infrastructure", and "policy" and a comparable size subset remained "unclassified".

Figure 6. Stemmed word cloud analyses of "Impact on infrastructure" class of tweets for the Hokkaido earthquake in Japan (left) and the Lombok earthquake in Indonesia (right).

Additional analyses of the tweets classified in the "Impact" class demonstrated that the tweets can be used to provide an indication of the impact of the earthquake on specific infrastructure. For example, for the Hokkaido, Japan, earthquake, the word "landsides" (and its variations) was among the most popular words with 3151 occurrences (out of 5964 tweets in the "Impact" class), consistent with the large number of landslides observed in that event. The word "Typhoon" was also very popular, with 2115, as the typhoon had played a key role in the damage observed during the earthquake. The term "dead" and "kill" had only 1139 and 1031 occurrences respectively. For comparison, in the far more deadly Lombok event, the word "kill" and "dead" were among the most common terms with 4009 and 2905 occurrences out of 7991 tweets in the "Impact" class. The threat of a tsunami which was a serious consideration for the Lombok event was also at the top of the list with 3735 occurrences, while it was not a popular term for the Hokkaido event where there was no tsunami threat. Finally, landslides were not a primary concern in the Lombok earthquake and indeed the term was only used 35 times.

Although more work is needed, this analysis indicates that indeed social media may provide an understanding of the impact of an earthquake on infrastructure immediately after the earthquake.

Table 4. Classification of original tweets related to earthquake impact for two earthquake events

Sub-Class of Tweet	M_w 6.6 Hokkaido, Japan September 5 2018	M_w 6.9 Lombok, Indonesia August 5 2018
Natural Hazards	7509 (34.7%)	4404 (22.3%)
Casualties	7472 (34.5%)	11394 (57.6%)
Impact on Infrastructure	1621 (7.5%)	955 (4.8%)
Policy	2140 (9.9%	623 (3.1%)
Unclassified	2917 (13.5%)	2405 (12.2%)

CONCLUSIONS

The social media activity in Twitter that is associated with earthquakes was studied for the year 2018. It was observed that social media activity is correlated to earthquake activity spatially and temporally. The observed activity for a specific event is a function of a number of factors, including the magnitude of the earthquake, internet accessibility, and other social factors. However, the most important factor was the location of the earthquake and its intensity with respect to population density. Even small events in urban areas, had significant social media response. The social media response immediately after the earthquake was also investigated and it was found that events with significant impact on infrastructure have a continued social media reaction, compared to events that did not. Social media activity was also spatially correlated to earthquake activity. Finally, a classification of tweets using machine learning techniques and word analyses, demonstrated that there is significant content related to the impact of the earthquake on infrastructure that could be data-mined to collect information about the event in the immediate aftermath of an earthquake and can be used to pinpoint expected damage.

REFERENCES

Dashti, S., Palen, L., Heris, M. P., Anderson, K. M., Anderson, T. J., & Anderson, S. (2014, May). Supporting disaster reconnaissance with social media data: A design-oriented case study of the 2013 Colorado floods. In ISCRAM.

Gao, H., Barbier, G., Goolsby, R., & Zeng, D. (2011). Harnessing the crowdsourcing power of social media for disaster relief. Arizona State Univ Tempe.

Kaigo, M. (2012). Social media usage during disasters and social capital: Twitter and the Great East Japan earthquake. *Keio Communication Review*, 34(1), 19-35.

Kim, J., & Hastak, M. (2018). Social network analysis: Characteristics of online social networks after a disaster. *International Journal of Information Management*, 38(1), 86-96.

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. *Science advances*, 2(3), e1500779.

Lachlan, K. A., Spence, P. R., Lin, X., & Del Greco, M. (2014). Screaming into the wind: Examining the volume and content of tweets associated with Hurricane Sandy. *Communication Studies*, 65(5), 500-518.

Middleton, S. E., Middleton, L., & Modafferi, S. (2014). Real-time crisis mapping of natural disasters using social media. *IEEE Intelligent Systems*, 29(2), 9-17.

Murthy, D., & Longwell, S. A. (2013). Twitter and disasters: The uses of Twitter during the 2010 Pakistan floods. *Information, Communication & Society*, 16(6), 837-855.

Murthy, D., & Gross, A. J. (2017). Social media processes in disasters: Implications of emergent technology use. *Social science research*, 63, 356-370.

Sutton, J., Spiro, E. S., Johnson, B., Fitzhugh, S., Gibson, B., & Butts, C. T. (2014). Warning tweets: serial transmission of messages during the warning phase of a disaster event. *Information, Communication & Society*, 17(6), 765-787.

Yin, J., Karimi, S., Lampert, A., Cameron, M., Robinson, B., & Power, R. (2015, June). Using social media to enhance emergency situation awareness. *In Twenty-Fourth International Joint Conference on Artificial Intelligence*.