UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

QUANTITATIVE ANALYSIS OF SOCIAL MEDIA SENSITIVITY TO NATURAL DISASTERS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

CYRIL BEYNEY
Norman, Oklahoma
2015

To my parents, who have always believed in me and made this life experience possible.

Acknowledgments

Firstly, I want to acknowledge my advisor, Dr. Charles D. Nicholson, for always encouraging me and for making me feel confident in the process of doing this research, even if sometimes I thought I could not make it. It has been a privilege to work with him, in such a great environment at the University of Oklahoma. His recommendations and daily good mood were a great source of inspiration during my studies.

I gratefully thank my family for all their support, financial and moral, that made this experience possible. They believed in me and were always there to listen to me, even with the thousands miles between us. A shout out to my friends Mikael Perrin, Jeremy Pfeifer, Vera Bosco, Olivia Perret and Alexandre Xausa among the others for not only their support but also for making this experience amazing and unforgettable.

I would also acknowledge Dr. Kash Barker and Dr. F. Hank Grant for agreeing to be in my committee. Finally, I would like to thank Jennifer Covington and all the faculties and staff of the School of Industrial and Systems Engineering for all their help and kindness throughout my stay at the University of Oklahoma.

Contents

List of Tables

List of Figures

Abstract

The use of social media platforms such as Facebook or Twitter increased exponentially over the last years. They have become a major tool for people to communicate, and for news media to relay their content. In parallel with this growth, researchers have started to use social media in various domains of study, including the analysis of people's behavior on these online platforms during natural disasters. Diverse aspects of this behavior have been studied in the literature, including the detection of a disaster, the variations of the sentiment expressed, the differences regarding the distance to a disaster, or even the improvements of relevant content labeling. To the best of our knowledge, no study has been conducted that considers Twitter as a sensor with different sensitivity levels to different types of natural disasters, which brought us to analyze Twitter as a social sensor during different types of these events.

The data used in this study is streamed from the Twitter Public Stream and tweets have been pre-processed to keep those in English, with geo-coordinates and from the contiguous United States only; the text content has also been cleaned and tweets have been filtered using a list of general keywords that we built. We select 5 natural disasters of different types that occurred over the recent years, and for each of them we detect the shifts in behavior of the Twitter data in order to compare and contrast before, during and after the disaster the variations of tweet

frequency, the proximity to the center of the disaster, the variations of sentiment expressed and the variations of tweet frequency by level of social vulnerability.

The results obtained in the empirical analysis demonstrate that Twitter is indeed a social sensor with different sensitivity levels to natural disasters. As a matter of fact we observe, depending on the type of disaster, different patterns of tweet frequency and proximity-to-disaster; negative sentiment tweets also tend to cluster closer to the disaster during the disrupted period, and finally areas with high level of social vulnerability are generally more sensitive compared to the others. The lack of available data can sometimes be an issue, but this work is an important finding to define Twitter as a social sensor to natural disasters.

Chapter 1

Introduction

1.1. Purpose of the study

Nowadays, social media have become a major way to communicate, either about conversational topics or to relay more informational content such as news [27]. They are in this way an important tool for news media to spread information, and for research purposes to collect what people think, to detect some events even before they are publicly relayed by official agencies, and so on. Several recent studies are based on the use of social media during natural disasters, studying mainly either the mood of the people or the sensitivity of the public during a specific case study. This is a recent domain of research linked with the important growth of social media uses by the population. One of the first works in this area was in 2008 after the wildfires in South Carolina [25]. Since then, the main case studies in recent years have been related to Hurricane Sandy and the Haiti earthquakes. They were both major disasters that allowed scientists to conduct different types of research based on how people responded in social media during these events.

1.2. Twitter

Twitter is a social network platform launched in 2006. It provides a microblogging service that allows users to send, read, share ("retweet" or "RT") short messages up to 140 characters called "tweets". One of the major characteristics is that the feed is provided in real-time, according to the people that the users "follow", trends (keywords or labels called "hashtags") they are looking for, etc. These features allow researchers to measure what the crowd thinks, how people react to an event, how to detect event, etc. These social media based research became more and more common in the same time when social media like Twitter or Facebook became more and more popular and used as a way to communicate, express all kind of feelings or relay news. The other advantage of Twitter is that the research is facilitated by the tools provided by Twitter to stream or get data from the stream of tweets with many possibilities (keyword targeting, geo-location restrictions, and so on), allowing researchers to easily get and use these data in experiments for their studies.

1.3. Problem Definition

In the existing works on the uses of social media during natural disasters, the main questions involved are: *How a natural disaster could be detected on Twitter?* [8, 20] *How does the crowd react in terms of mood variations?* [4, 17, 22] *or regarding their distance to the disaster?* [15, 21] *How can we better define tweets that are relevant to a particular disaster?* [14, 27]. Furthermore, a wide range of the research already conducted in this area covers a specific case study (i.e. one natural disaster) to illustrate the study, and if several natural disasters are studied, it is not to compare the results between them. To the best of our knowledge, no study has been conducted that considers Twitter as a sensor with different sensitivity levels to different types of natural disasters.

Therefore, we would like in this work to study several natural disasters, compare the results on the same scales and draw the conclusions about Twitter sensitivity levels according to different metrics and different natural disasters. Consequently, the main research question of this work is *"How and under which circumstances Twitter can be a social sensor during a natural disaster?"*

- Which metrics could measure the variations of the sensitivity in Twitter data ? We define these metrics as following:
  - the **Frequency** over time of tweets posted
  - the **Emotion** of the crowd defined by a sentiment analysis
  - the **Proximity** to the disaster of tweets posted
- What is the sensitivity level of Twitter data during different types of natural disasters (i.e. Severe Winter Storms, Severe Thunderstorms, Wildfires, etc.) for each of the metrics defined?
- What is the sensitivity level of Twitter data in terms of duration of the disrupted state and the time to recover to a stable state?
- What is the sensitivity level of Twitter data in terms of the social vulnerability of the populations that use Twitter during natural disasters?

1.4. Structure of the Thesis

Firstly, this thesis details in Chapter 2 the background and literature review about social media analysis during natural disasters. Secondly, Chapter 3 presents the methodology of this work and the proposed solution for the quantitative analysis of Twitter sensitivity with the different metrics. Chapter 4 describes the analysis conducted on real case studies, interprets and discuses the results obtained to define its viability and sensibility according to the metrics. Finally, Chapter 5 concludes this thesis.

Chapter 2

Literature Review

This Chapter reviews the existing work in the area of social media analysis during natural disasters.

2.1. Sentiment Analysis

Sentiment Analysis, also called Opinion Mining, is a common tool used with social media to define the opinion of people about a product, a movie, etc. In the case of a natural disaster, several studies such as [4] or [17] classify tweets either as *Positive*, *Neutral* or *Negative*, according to the polarity of the emotion expressed. This is the most common classification in Sentiment Analysis. However, other studies in Sentiment Analysis such as [22] claim that it is not the most accurate way to classify emotions, and then try to classify them following the well-known classification from Ekman [9] who stated than the human can have seven different kind of emotions: Anger, Disgust, Fear, Happiness, Sadness, Surprise and Neutral. As classifiers, we found out in these papers but also in studies from [22] or [4] that Bayesian Models are widely tested and compared with Support Vector

Machines (SVM). Several similarities exist in the features used in these papers and the tools behind are often cited in the domain of Sentiment Analysis.

Positive and Negative weights: these weights are given according to several lexicons or tools:

- **AFINN**: Presented by Nielsen in 2011 [18], this sentiment lexicon is composed by 2,477 English words manually labeled between minus five (negative) and plus five (positive) by its author.

- **Hu and Liu lexicon**: Presented in 2004 [10], this sentiment lexicon is composed by 6,800 English words manually labeled either negative or positive by its authors.

- **SentiWordNet**: SentiWorldNet 3.0, the last version of this sentiment lexicon, is based on the lexical database WorldNet, which groups synonym words (nouns, verbs, adjectives and adverbs) into sets called "synsets" with a brief definition, these sets being linked to each others with semantic relations (i.e. Synonymy, Antonymy) [16]. In [1], the authors developed a semi-supervised technique that gives for each synset (~117,000) a positive, negative and objective (neutral) weight, which, compared to the AFINN lexicon, gives for a same word different scores depending on its meaning in a sentence.

- **SentiStrength**: Tool that defines the average sentiment strength of a sentence according to its words and the context around [26]. Informal language is recognized, and results can be given in different ways: two sentiment strengths (-1 not negative to -5 extremely negative, 1 not positive to 5 extremely positive), binary (positive/negative), trinary (positive/negative/neutral) and single scale (-4 to +4) results.

Syntactic features:

- Punctuation: Exclamation and interrogation marks can be counted and used as a feature representing the intensity of the emotion expressed.
- Emoticons/Smileys: They are widely used in online interactions to express emotions; can be detected, or even more detailed counted and classified as positives or negatives.

2.2. Event detection

If several studies focused partly on sentiment analysis during natural disasters, others target on event detection and more particularly on earthquake detection. The work by [20], which was one of the first ones in the area of event detection on social media, tends to focus on detecting typhoons and earthquakes in Japan by crawling tweets with specific keywords, and then using a Support Vector Machine algorithm to determine those truly related to a natural disaster. Later, another study focused on earthquake detection worldwide [8], arguing that the real-time aspect of social media allows to detect events before they are publicly announced by the U.S. Geological Survey (USGS). They implemented a short-term-average over long-term-average (STA/LTA) algorithm with three different levels of threshold, and were able to detect with a moderate threshold 48 earthquakes over 5,175 cataloged by the USGS, arguing that "the majority of [...] earthquakes are either too small to produce perceivable shaking or occur outside populated areas". Among the larger earthquakes when Twitter is a viable sensor, it has the ability to detect earthquakes faster than traditional systems and consequently is a powerful additional tool for weather related agencies.

2.3. Geo-location of the tweets

Since 2009, tweets can contain geographic meta-data indicating the position from which the tweet has been posted on the social media. These information can represent the exact location provided by the GPS component of the device, or a specific place the user can choose when posting a tweet. By default, these information are not available if the user does not grant access to Twitter. That being said, not all the tweets contain geo-location information. In fact, according to a study from 2013 [13], only 1.6% of the total stream of tweets contain the exact location where the tweet was posted. To mitigate this issue, in some works, such as [15], the authors geoparsed the text of the tweets to extract geographic content and use a named entity recognition technique in order to match locations up to the place/street level of accuracy.

In [4], the authors mapped the moods reflected in tweets during Hurricane Sandy. According to them and showed by their results, even if Hurricane Sandy has a regionally limited impact in terms of damages, people have been emotionally affected by the storm far away from the coast. Their maps display the population's response in space and time to the disaster measured through a sentiment analysis, showing that the closer people were to the point the storm made landfall the more they tweeted, negative sentiment tweets being always clustered in closer proximity to the storm.

Another interesting case study also showed that people tend to tweet differently depending on their proximity to the disaster but with some different conclusions [21]. The authors performed a trend analysis of the tweets when the Great East Japan earthquake hit Japan and the period around. They showed that the tweet frequency peaked dramatically when the earthquake hit. However, by comparing different regions, they found out that people posted fewer tweets in the closed areas heavily damaged compared to further areas with less damages. This is explained by the fact that people were not in a safe situation to tweet, or were not technically able to access the Internet.

2.4. Relevance of the data

In several studies, data available and used are often already labeled as relevant or not regarding a specific topic, e.g. natural disasters. However, in case of a new disaster, the data collected from social media such as Twitter are unlabeled. Therefore, some research focused on how they could label data as relevant. In [14], the authors used a domain adaptation algorithm based on Naïve Bayes and Expectation-Maximization to learn classifiers from source labeled data collected during a disaster to classify unlabeled data collected during another disaster (Hurricane Sandy and the Boston bombing are used in their experimentations). They compared the performance of this classifier with a classic Naïve Bayes classifier on the source labeled data only for three classifications, and the first one gives better results in tasks more specifics to the target disaster (i.e. is this tweet related to the target disaster?). However, they discuss the need of testing such algorithms on bigger datasets with other types of classifiers.

In [27], the authors stated that with the important increase of social media uses and consequently the increase of social media data, it would be profitable, especially during disasters, to only keep valuable information from them. Therefore, they developed a Bayesian model to classify tweets during Hurricane Sandy as either "Informational" (tweets that provide concrete and valuable information) or "Conversational" (that do not provide valuable information). In their approach, they compared a Bag-of-words technique ($\sim 3{,}000$ features) with a set of 9 features they defined and trained on a manually labeled set of 1,086 tweets.

- **Conversational content**: The presence of emoticons, Internet slang and abbreviations, curse words and abrupt sentences can be detected and are characteristics of a conversational text.

- **Informational content**: The presence of instructional keywords (such as "text" or "call"), phone numbers, retweets, and multiple sentences can be detected and are characteristics of informational content. Original URLs can also be extracted from shortened URLs to check the presence of news sources and informative text.

The authors used a Naives Bayes classifier with a 10-fold cross validation to train and test their model, and then compared it to a Bag-of-Words model and a combination of the two models. According to the results, their features obtain precision rates slightly under the ones with the Bag-of-words but they can still be used solely if computing and time resources are limited. Nevertheless, the precision rate for informational tweets is low for all the scenarios, due to unbalanced data at the beginning (139 informational against 943 conversational tweets).

Further work in this field focus on the "retweetability" of a tweet, arguing that informational and important tweets (posted by news channels, National Weather Services, or even first people on the ground) will be more retweeted and therefore are more relevant during emergencies [24]. Social media data from locals, also called people on the ground, can be valuable to emergency responders: [23] stated that locals are "uniquely positioned to share information that may not yet be available elsewhere in the information space [...] [and] may have knowledge about geographic or cultural features of the affected area that could be useful to those responding from outside the area.".

2.5. Social Vulnerability Index (SoVI)

The Social Vulnerability Index (SoVI) has been created in 2003 at the University of South Carolina in order to define, for the United States at county-level, the social vulnerability of communities to environmental hazards [5]. The first version of the index has been built on demographic and socioeconomic data from 1990 primarily from the U.S. Census Bureau, it contained 42 variables reduced to 11 factors, including personal wealth, age, race, etc. Across the time and the different versions of SoVI, several spatial changes can be observed regarding the most vulnerable communities [6]. The last version of the index, built from data collected between 2006 and 2010, follows new direction that "emphasize the constraints of family structure, language barriers, vehicle availability, medical disabilities, and health-care access in the preparation for and response to disasters". After running a Principal Component Analysis, scientists from the Hazards and Vulnerability Research Institute obtained seven significant components that explain 72% of the variance of the data. Once the index is calculated for each county, they are decomposed into percentiles. Scores in the top 20% correspond to the most vulnerable counties, where scores in the bottom 20% indicate the least vulnerable ones.

This index can be found in several applications, such as state-level hazard mitigation plan (e.g. Colorado and California), or the Sea Level Rise Coastal Impacts Viewer by the National Oceanic and Atmospheric Administration (NOAA) (see `http://coast.noaa.gov/slr/`). However, no study has been conducted on social media sensitivity to natural hazards with an emphasis on the social vulnerability level of the concerned areas and its potential effect on the results.

2.6. Resilience and Social Media

The term resilience originally refers to the ability to "bounce back" from a specific disruptive event [7]. Therefore, the concept of resilience can be decomposed in two steps: (i) preparedness measures to minimize the impacts of potential disruptive events and (ii) the ability to adapt and recover to failures in a timely manner. In [2], the authors defined the different states and transitions among them when a disruptive event e^j happens while a system is operating. In their Figure 2.1, S_0 denotes the original as-planned state, S_d denotes the disrupted state and S_f the recovered state.

The resilience of the system over time can then be defined as:

$$Resilience(t) = \frac{Recovery(t)}{Loss(t)}$$

If we focus the concept of resilience in terms of social media behavior during natural disasters we have the concept of community disaster resilience in the emergency management literature [7, 12]. The different terms employed in Figure 2.1, that is original stable state, disrupted state, recovered state and the related events will be used in the following Chapters.

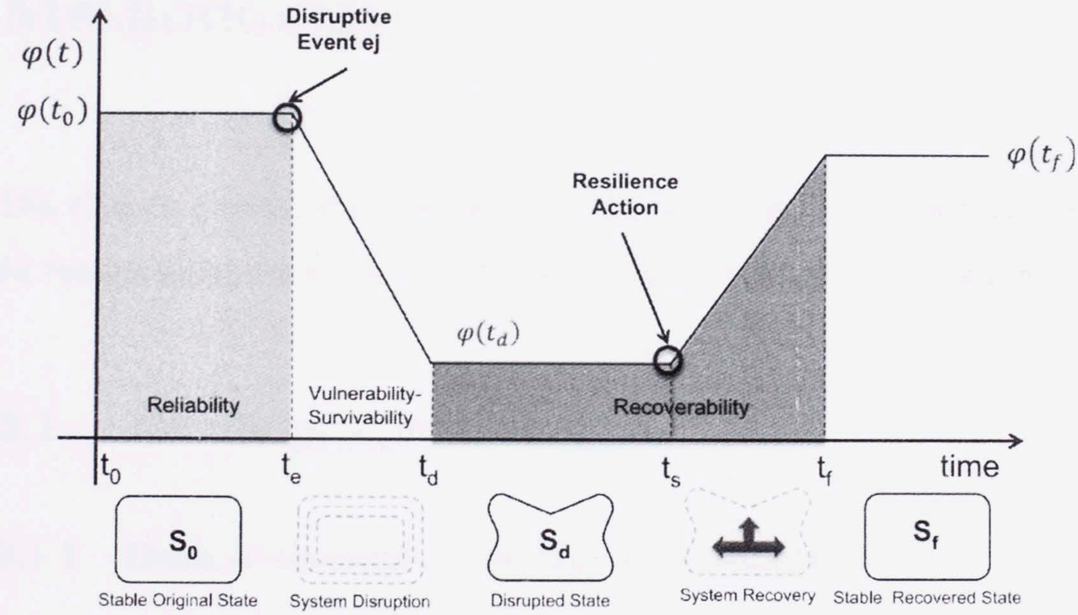


Figure 2.1: State transitions over time when a disruptive event happens

Chapter 3

Methodology

This Chapter presents the methodology of this work and the proposed solution for the quantitative analysis of Twitter sensitivity with the different metrics.

3.1. Tweets pre-processing

3.1.1 Data extraction, cleansing and selection

Twitter provides several APIs for developers and researchers, including a streaming API. In its last version, three types of streams are freely available with a developer account: Public, User and Site. User streams provide a stream of data and events specific to the authenticated user. Site streams allow services, such as websites or mobile push services, to receive real-time updates for a large number of users. Public streams are, according to Twitter, "suitable for following specific users or topics, and data mining". Two Public streams exist:

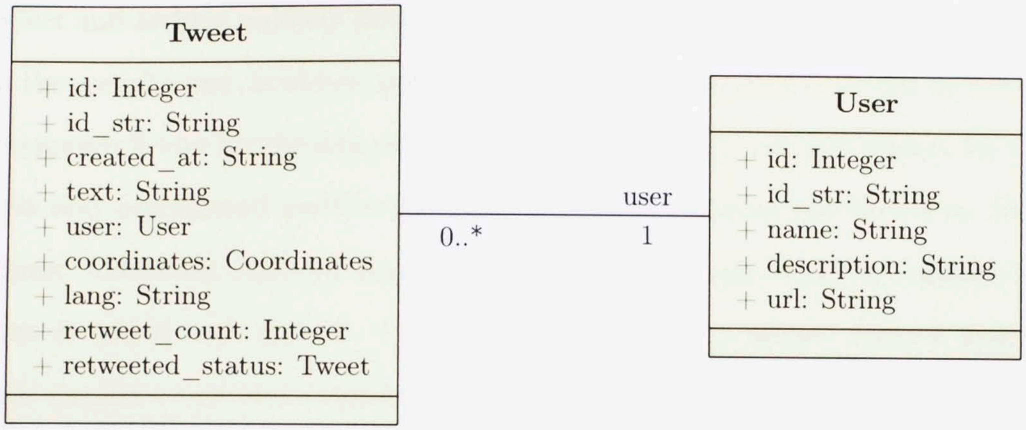


Figure 3.1: Simplified UML diagram of the Tweet JSON object

- the *Sample Stream*, which provides a random sample corresponding up to 1% of the total stream

- the *Filter Stream*, which provides tweets related to research criteria. Results cannot exceed 1% of the total stream.

Additionally, Twitter provides the *Firehose Stream*, which provides 100% of the total stream, however accessing this stream is expensive and not common for research.

The tweets are stored in JSON (JavaScript Object Notation) objects, a popular data format to transfer data on the Internet. Each object is represented by **attribute - value** pairs, where values can consist in basic data types such as string or integer, or a JSON sub-object. For each tweet, all the details regarding its author are included, see Figure 3.1. A broad range of information is also provided by Twitter for each tweet, the complete list being documented on their website: `https://dev.twitter.com/overview/api/tweets`.

The data used in this work come from the Sample Stream and have been collected during several months by the Archive Team (a group of volunteers who collect and archive publicly available Internet data). Datasets are freely available on the website `www.archive.org`. For each month, the data is stored as follows: one master folder for the month, sub-folders for each day, sub-sub-folders for each hour and compressed archives for each minute, containing the tweets in JSON format. Currently, data for the entire years 2012 through 2014 are available and more is added each month. This broad range of dates allows Twitter analysis with respect to several important natural disasters that occurred during these years. In this study, we filter the tweets as follows:

- Language: English
- Geo-tagged: Yes
- Location: Originating from the United States

These constraints have been established in order to compare the sensitivity of Twitter regarding natural disasters that occurred inside the United States, and using sentiment analysis tools on tweets written in English. To implement these filters, several steps must be taken:

- Data extraction: Extract all the *.json.bz2* files in the same folder for an easier further access
- First pre-process: Extract the content of each *.json.bz2* file, test each tweet if it is in English with geo-location information, clean the text (remove line breaks, punctuation, html links, etc.) and write it inside a *.csv* file.
- Geo-location pre-process: Test if each tweet kept in the previous step has been posted from the United States, and only keep those which validate this constraint.

Both first two steps have been implemented in Python with two functions: *extractMonthTweets()* (see Algorithm 1) and *processMonthTweets()* (see Algorithm 2), the variables *filenames* and *foldernames* containing respectively the list of files in memory to process (all named *archiveteam-twitter-stream-201X-XX.tar* depending on the year and month data have been collected) and the list of folders to extract the data in (all named *archiveteam-twitter-stream-201X-XX*, generated from the *filenames* list).

The textual content of a tweet can be noisy, in that it may include a variety of non-English content such as HTML links, tags, special characters, etc. Consequently, during the pre-processing of the data, we "scrub" the text, in order to have better results during the experimentation, especially for the sentiment analysis. Here are the different data-cleansing processes implemented for the tweet message:

- remove tags (tagging someone refers to including a @ character followed by the username, e.g. *@UofOklahoma*) and retweet entities (e.g. *RT @UofOklahoma*)
- remove HTML links
- remove all punctuation, special characters, numbers, line breaks and any additional whitespace

A few examples of scrubbed tweet messages are provided in Figure 3.2.

Algorithm 1 Step 1: Data extraction

```
procedure EXTRACTMONTHTWEETS(filename, filenames[], foldernames[])
    mainTar ← Open(filename, read)
    members ← getMembers(mainTar)
    for all member in members do
        if ! member.isDirectory() then
            mainTar.extract(member, foldernames[filenames.index(filename)])
        end if
    end for
    mainTar.close()
end procedure
```

Algorithm 2 Step 2: Data cleansing and pre-processing

```
procedure PROCESSMONTHTWEETS(foldername)
    tweetsBuffer ← [ ]
    exportFile ← open(foldername + '.csv', write)
    for all bz2file in foldername do
        jsonBZ2 ← open(bz2file, read)
        content ← read(jsonBZ2)
        for all line in content do
            tweet ← loadJSON(line)
            // Tweet processing and text cleansing
            tweetsBuffer.append(tweet)
            if len(tweetsBuffer) == 250 then
                write(exportFile, tweetsBuffer)
                tweetsBuffer ← [ ]
            end if
        end for
    end for
end procedure
```

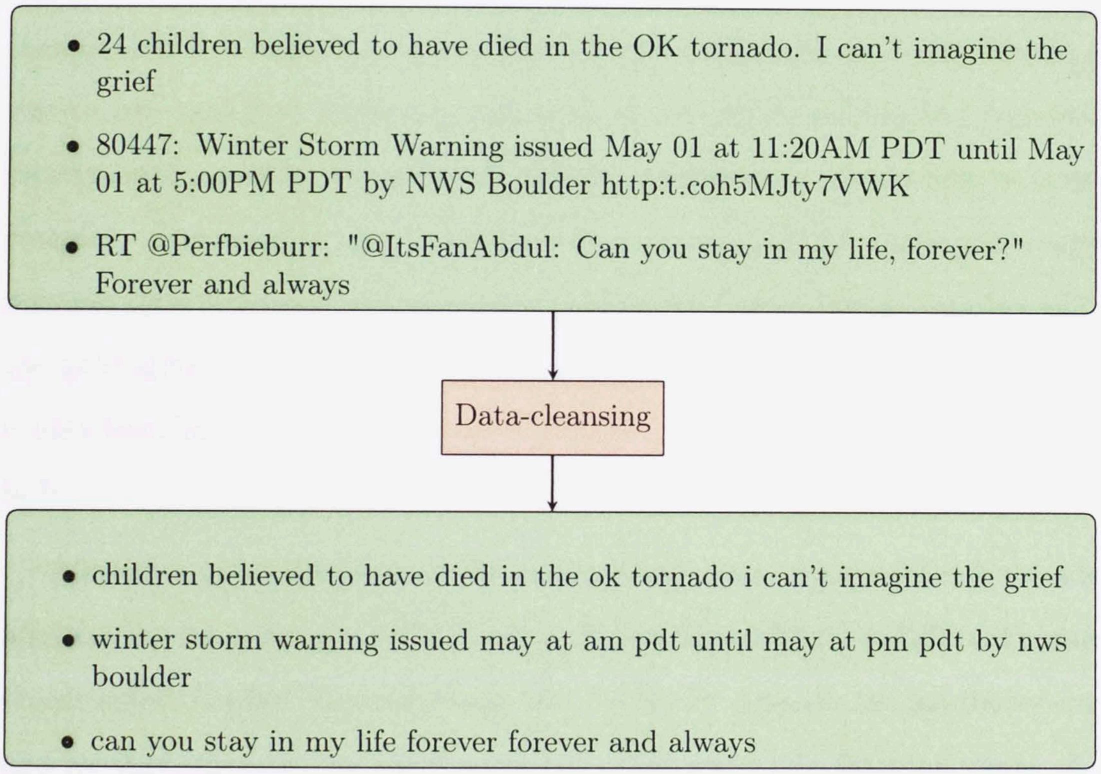


Figure 3.2: Result example of the function clearTweet()

3.1.2 Data geo-location selection

If we only keep tweets in English, those come from all over the world. As discussed in Chapter 2, a small percentage of the total stream of tweets contains the exact location where the tweet was posted. However, we define our study domain as tweets posted in the United States, and therefore we use these geo-tags to filter the data. Filtering coordinates can be done in several ways, some more complex than others. For instance, one can define a rectangular area and only keep the tweets that come from inside this basic area, or even inside another area with one location and a radius. Nevertheless, filtering according to the country is more complex. The website *Global Administrative Areas* (GADM - `www.gadm.org`) provides for a wide range of countries, including the United States, datasets with the geographic limits of their administrative areas. Data are available at the county level for the U.S. and in several formats, including data easily processable in R.

Algorithm 3 describes the coordinate filtering process. Firstly, it loads the administrative areas dataset. The result in R is a **SpatialPolygonsDataFrame** object called *GADM*. Secondly, from the *TWEETS* dataset, the attributes *lon* and *lat* that represent the coordinates are transformed into **SpatialPoints** objects with the function *coordinates()* in order to project them into the *GADM* object with the function *proj4string*; since the *GADM* object represents the limits of the United States, we can access the coordinates projected inside and finally get back all the data related to these remaining coordinates from *TWEETS*.

Algorithm 3 Step 3: Data geo-location filtering

procedure GEOLOCATIONFILTERING
  GADM ← load administrative areas
  TWEETS ← read("tweets.csv")
  Create Spatial Points from the coordinates in TWEETS
  Project the Spatial Points into the GADM
  Get the coordinates projected inside the GADM only
  USTWEETS ← coordinates + remaining data
end procedure

As shown in Figure 3.3, the tweets are correctly filtered to keep only those posted from the United States. For further filtering in the next chapter, the GADM object lists all the States and their respective counties with their geographic limits, which will be useful when dealing with the Social Vulnerability Index (SoVI). In a similar way, we can filter the tweets by State and county and then add attributes to our *USTWEETS* dataset to label our data based on the results.

3.1.3 Keyword filtering

As stated in Subsection 3.1.1, Twitter only provides up to 1% of the total stream, either if we select filters or if we do not. Moreover, as one can expect, messages contained in Twitter data are diverse. Consequently, there is a need to filter by keywords in order to limit the text messages to those which are germane to the topic of study, namely natural disasters. The goal of this research is to test the sensitivity of Twitter to natural disasters. Consequently, we build a common catalog (list of keywords) for all types of disasters, and no hashtag or keyword specific to an event is included (e.g. *"Moore Tornado"* or *"Hurricane Sandy"*), as it is usually used to analyze the trend of a particular event on social media. To build this list of keywords, we process in several steps.

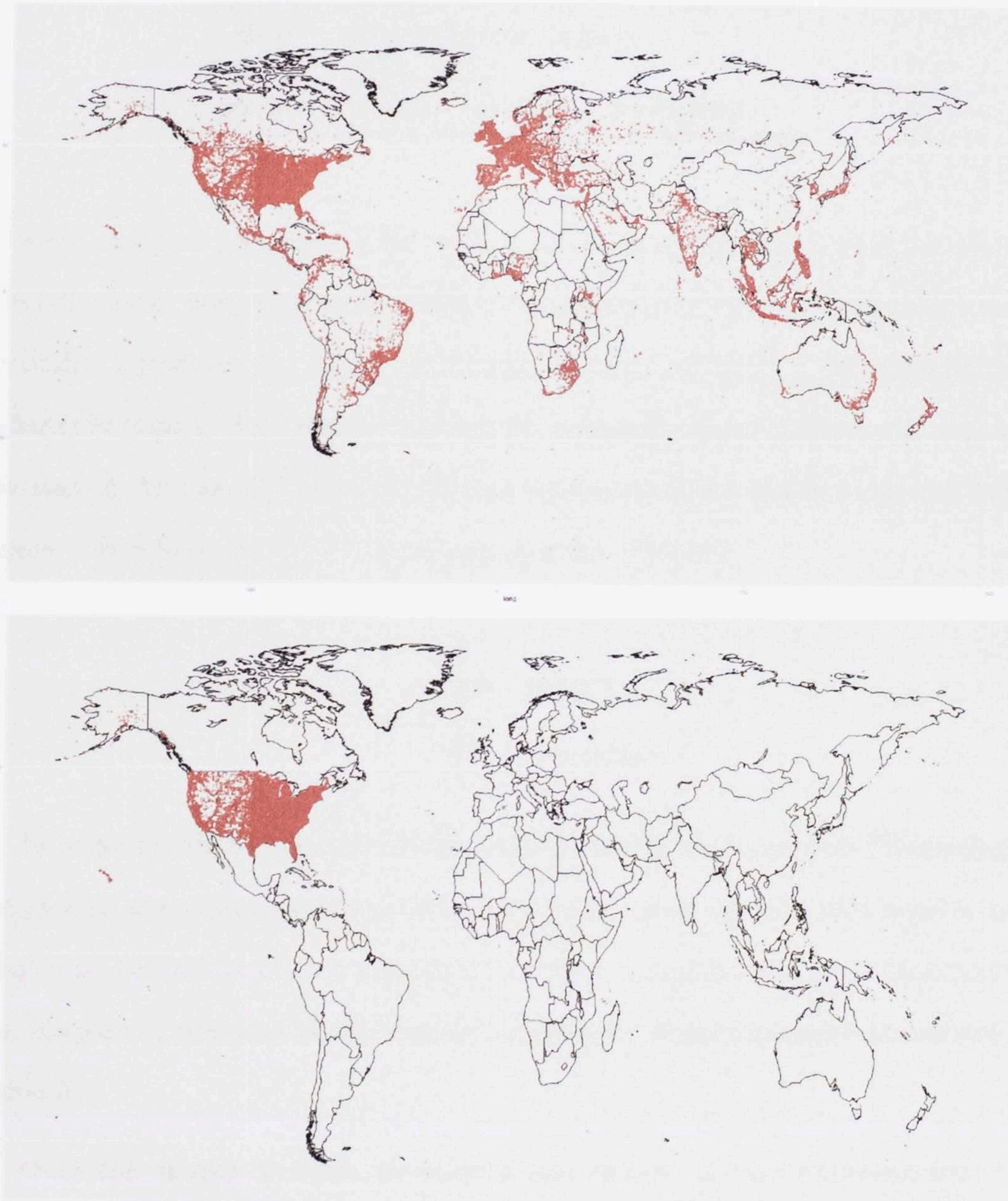


Figure 3.3: Tweet geo-locations before/after filtering

Table 3.1: Examples of word stemming

Stem	Related words
connect	connection, connected, connecting
argu	argue, arguing, argued
tall	taller, tallest
produc	product, produced, production

First, build a preliminary list with a dozen basic keywords, such as: *storm, tornado, flood, hail, wildfire* in order to "pre-filter" the tweets for the next step. Secondly, based on the tweets previously filtered, we build a **DocumentTermMatrix** object. The rows of the matrix represent the documents (in our case the text of the tweets), and the columns represent all the words in all the documents. Therefore, for the i^{th} document and the j^{th} word:

$$x_{i,j} = \begin{cases} 1, & \text{if } j \in i \\ 0, & \text{otherwise} \end{cases}$$

In order to reduce the size of this matrix, words are stemmed. Stemming is the process of reducing words to their word stem, which can be a root word or not. That way, several words are considered as their common stem and consequently the number of columns in the matrix is reduced. Some examples are shown in Table 3.1.

Once this matrix is built, we look at two things: (i) the *frequency list*: the sum of occurrences of each word, from which one can get the most used words in the corpus and derive those that are the most related to disaster events; (ii) the *Latent Dirichlet Allocation (LDA)*: LDA is a model that, for a given number of topics k, is looking all the documents for k topics across them and group words that have high probabilities to belong to the same topic [3].

Finally, *CrisisLex* is a lexicon presented in [19] aimed to improve "the recall in the sampling of Twitter communications that can lead to greater situational awareness during crisis situations". If several keywords and combinations of keywords were not useful in our study, it helped us building our catalog of keywords.

3.2. Metrics Definition

3.2.1 Frequency Analysis

The first step is to evaluate statistically any variation in the data and see if there is any correlation with the dates of the events. In order to achieve this goal, several *breakout detection* tools exist. Breakout detection is a statistical technique for detecting either sudden jumps in time series data (mean shift) or a gradual increase in the value of a metric between two stable states (ramp up). A recent technique developed by Twitter called Energy Divisive with Medians detects breakouts by detecting divergence in mean. According to the authors, their technique is "robust against the presence of anomalies [...] and is 3.5x faster than a state-of-the-art technique for breakout detection" [11]. This recent work is used on a daily basis at Twitter according to the authors and it has been tested at the University of Louisville School of Medicine to identify past influenza outbreaks from CDC data [28].

Using this tool, and for a disruptive event such as a natural disaster, three different breakouts are defined for the case studies: (i) the *disruptive time* t_d, when a change is observed in the normal flow, i.e. a significant increase in the tweet frequency when an event occurs; (ii) the *recovering time* t_r when a second change is observed, i.e. a significant decrease in the tweet frequency after the event; (iii) the *stabilization time* t_s when it stops decreasing to reach the recovered state. These dates will be used throughout the different analysis, and will help to define any variations in some metrics over time such as the tweet frequency. The idea here will be to compare between the different natural disasters how long the normal flow of tweets is disrupted and what is the recovery time for each of them.

3.2.2 Proximity Analysis

For each disaster, the location of its center is defined. Depending on the type of event, the center can either be the city or location a tornado first hit the ground (e.g. Moore, OK for the Moore tornado in 2013), the city that faced, or the areas around, most of the event damages (e.g. Buffalo, NY for the New York winter storm in 2014), or the closest nearby location of a wildfire (e.g. Black Forest wildfire in Colorado in 2014). Based on this location, we compute for each geo-tagged tweet its distance from the center of the event. This allows to define for any case study the median distance and the third quartile value of the distribution of filtered tweets during the event, and see if variations occur for different types of disasters.

3.2.3 Sentiment Analysis

A sentiment analysis is run on the filtered tweets to define whether the emotion expressed is positive, negative or neutral. This will be done with the **qdap** R package, which provides a *polarity*() function that, by default, uses the Hu and Liu lexicon [10] seen in Section 2.1 to compute a polarity score δ for a given text. Firstly, the algorithm checks the presence of any polarized word, positive or negative, based on this lexicon. Once polarized words are defined, a content cluster x_i^T is created for each of them, pulling by default the 4 words before and the 2 words after. These words, called valence shifters, are tagged either as negator x_i^N (e.g. "not"), amplifier x_i^a (e.g. "heavily") or de-amplifier x_i^d (e.g. "barely") based on related lexicons, or as neutral (x_i^0) otherwise. Then, each polarized word is weighted w, with a weight of -1 if it is a negative word, $+1$ otherwise. An additional weight is added, based on the number of valence shifters in x_i^T. Finally, the context clusters are summed and divided by $\sqrt{n}$, with n the total number of words. The result is the polarity score, δ, unbounded. Steps are detailed as follows:

$$w_{neg} = \left(\sum x_i^N\right) \mod 2$$

$$x_i^D = \max\left(\sum(-w_{neg} \cdot x_i^a + x_i^d), -1\right)$$

$$x_i^A = \sum(w_{neg} \cdot x_i^a)$$

$$x_i^T = (1 + x_i^A - x_i^D) \cdot w(-1)^{\sum x_i^N}$$

Finally:

$$\delta = \frac{\sum x_i^T}{\sqrt{n}}$$

After having computed the polarity score δ, the related emotion expressed by the text is defined as follows:

$$polarity = \begin{cases} \text{positive}, & \text{if } \delta > 0.0 \\ \text{neutral}, & \text{if } \delta = 0.0 \\ \text{negative}, & \text{otherwise} \end{cases}$$

Once this polarity will be defined for each tweet, we will look at the negative tweets and the skewness over time of the distance to the center of the disaster. The idea is to observe if negative sentiment tweets cluster in closer proximity to the disaster location during the event.

3.2.4 Social Vulnerability Approach

Using the GADM object as in Subsection 3.1.2, information regarding the county and the state tweets have been posted from can be determined and added to the data. Once each tweet is labeled with the state and county it comes from, we can relate to the Social Vulnerability Index detailed in Section 2.5, and see if the variations in terms of tweet frequency can be related to the variations of SoVI in the concerned areas. We will look at the national percentile rank NPR of each related county, and label the level of vulnerability as follows:

$$vulnerability = \begin{cases} \text{high}, & \text{if } NPR >= 80.0 \\ \text{low}, & \text{if } NPR <= 20.0 \\ \text{medium}, & \text{otherwise} \end{cases}$$

Chapter 4

Analysis and Results

This Chapter details the experiments evaluating the sensitivity of Twitter to different types of natural disasters. It describes the case studies on which the experimentation are conducted and their results.

4.1. Case studies

The goal of this study is to be as diverse as possible, therefore the case studies have to include several natural disasters of different types. They are grouped into categories used by the Federal Emergency Management Agency (FEMA) to classify disaster declarations: *Tornadoes*, *Winter Storms* and *Wildfires*. The events used as case studies are summarized in Table 4.1. In order of rows: two Tornado related events, one Wildfire and two Winter Storms. PA and IA stand for Public and Individual Assistance, listed by the FEMA (rounded, in million dollars).

Table 4.1: List of the different case studies analyzed

Event	Location	Start Date	Peak Date	End Date	PA	IA	Casualties/Damages
Moore Tornado	Moore, OK	5/18/2013	5/20/2013	6/2/2013	47.6	15.2	24 deaths, \$2 billion
AR/MS Tornadoes	Vilonia, AR Louisville, MS	4/27/2014	4/27/2014 4/28/2014	4/30/2014	9.9 87.9	3.0 5.9	16 deaths, \$223 million 10 deaths
Black Forest Fire	Black Forest, CO	6/11/2013	-	6/20/2013	7.4	-	2 deaths, 22.31 mile^2 burned
NY Winter Storm	Buffalo, NY	11/13/2014	-	11/26/2014	30.6	-	
SC Winter Storm	Half of the state	2/10/2014	-	2/15/2014	219.2	-	

- **Moore Tornado**: This tornado has been rated as an EF5, the highest level on the Enhanced Fujita scale. It hit the ground of Oklahoma in May 20, during the May 18–21, 2013 tornado outbreak that occurred mainly in Midwestern United States. Another important tornado touched the ground the day before in Norman, close to the city of Moore, and has been rated EF4.

- **Louisville/Vilonia Tornadoes**: These two tornadoes are part of the tornado outbreak that occurred mainly in Southern United States between April 27-30 2014, including Arkansas, Mississippi and Tennessee. They have both been rated as EF4. We will refer to the name Louisville Tornado in the analysis.

- **Black Forest Fire**: Forest fire that started on June 11, 2013 near Colorado Springs. Around 500 houses have burned in what is called the most destructive wildfire in Colorado history. It took a week to fully contain the fire.

- **NY Winter Storm**: Severe snow storm that occurred in the state of New York between 13-21 November 2014. Several counties have been concerned, mainly in the area around the city of Buffalo, NY with 5 to 7 feet of snowfall.

- **SC Winter Storm**: Severe snow and ice storm that affected the East Coast of the United States between 11-17 February 2014. Several states, including Georgia, North and South Carolina were impacted. We focused on South Carolina, where the damages were the highest, with several counties affected and up to 27.5 inches of snowfall in some areas. The city of Columbia, in the center of the state, will be chosen for the proximity analysis.

Table 4.2: Results of the keyword filtering on the tweets with our lexicon

Time Period	Raw tweets	Filtered tweets	Percentage
May 1 - 31 2013	516,750	3,997	0.77
June 1 - 30 2013	524,749	3,537	0.67
February 1-28 2014	627,909	7,578	1.21
April 15 - May 15 2014	712,879	4,664	0.65
November 1-30 2014	479,187	4,233	0.88

The list of keywords built regarding the methodology seen in 3.1.3 consists of 64 keywords related to the vocabulary of natural disasters, consequences and emergency response. There is no keyword related to a specific event, as we want to measure the sensitivity of the crowd regarding the use of general terms, and thereby not having possibly biased results by using keyword related to one particular event. The results of the filtering are displayed in Table 4.2, and we will use these tweets filtered by keywords in the following experiments. The number of tweets for each time period is given, before and after the filtering. The tweets before the filtering are in English, from the United States, while the tweets after filtering are from the contiguous United States only (tweets labeled as from Hawaii and Alaska are not considered for the analysis) and must contain one of the identified keywords. See Appendix A for the exhaustive list of keywords. The percentage of filtered tweets varies from 0.67% to 1.21% of the raw tweets, with an average of 0.84%.

4.2. Frequency Analysis

Using the breakout detection tool, we detect t_d, t_r and t_s for each case study and the results are listed in Table 4.3. The last column provides the *total observed time*

Table 4.3: Breakout dates and total observed time for each natural disaster

Event	t_d	t_r	t_s	t_{obs}
Moore Tornado	05/19/13 6PM	05/22/13 12AM	05/24/13 06AM	108
Louisville Tornado	04/27/14 6PM	04/29/14 12AM	04/30/14 12PM	66
Black Forest Fire	06/12/13 6AM	06/13/13 06PM	06/14/13 12PM	54
NY Winter Storm	11/13/14 5AM	11/17/14 11AM	11/20/14 11PM	186
SC Winter Storm	02/12/14 6PM	02/14/14 12PM	02/17/14 12AM	104

within these breakouts, that is $t_{obs} = t_s - t_d$, in hours. Tweets have been grouped by 6 hour time windows to determine the frequency over time and the breakouts. These dates will be used in further analysis to retrieve the tweets posted during the peak activity related to each event. Observed times vary from 54 hours for the Black Forest Fire to 186 hours for the Winter Storm in New York, with a total average of 103.6 hours. By grouping the disasters by their type, Winter Storms are on average the longest observed natural disasters, followed by the Tornadoes and finally the Wildfires.

Table 4.4 summarizes the tweet frequency variations for each case study. This is done by computing the average frequency during the 7 days prior to t_d, defined as the original stable state frequency ν_0, and comparing this frequency with the disrupted frequency ν_d between t_d and t_r in order to establish the rate increase. The significance of the difference between the two frequency means is statistically verified with a independent two-sample t-test (with unequal sample sizes and either equal or unequal variances based on the results of Levene's test for homogeneity of variances) and the p-value p_1 obtained. The value ν_s, defined as the recovered stable state frequency, computed during a week after t_s, is also provided to see if the system goes back to the same state as before the disruptive event or not. The p-value p_2 is obtained similarly as p_1 except it corresponds to

Table 4.4: Results of the frequency analysis for each natural disaster

Event	ν_0	ν_d	ν_s	% inc.	p_1	p_2	t_{rec}
Moore Tornado	31.67 (24.96)	75.9 (50.60)	27.71 (18.09)	139.68	0.0232	0.5059	54
Louisville Tornado	30.26 (15.91)	75.5 (44.32)	29.03 (13.64)	149.51	0.0541	0.7604	36
Black Forest Fire	26.85 (13.47)	83 (29.02)	24.21 (12.37)	209.10	0.0112	0.4527	18
NY Winter Storm	25.93 (10.43)	52.28 (33.77)	40.57 (25.56)	101.64	0.0046	0.0082	84
SC Winter Storm	91.31 (50.06)	206.5 (122.86)	64.68 (35.63)	126.16	0.0332	0.0303	60

the test for difference between ν_0 and ν_s (in other words, is the recovered state statistically equivalent to the original state?). The last column provides the recovery time, that is the time to recover to a stable state, defined as $t_{rec} = t_s - t_r$, in hours.

As one might expect, the recovery time t_{rec} is proportional to the total observed time t_{obs}, and except for the Black Forest Fire, the t_{rec} values are respectively equal or greater than $t_{obs}/2$, that is it takes more time to recover to a stable state than shifting to a disrupted state. Finally, Figures 4.1 - 4.3 depict for each disaster the breakouts detected with the tweet frequency per 6 hours. In some cases, the three breakouts were distinctively detected (and are clearly visible in the plots, such as for the Moore Tornado), while in other cases, such as the NY Winter Storm, more than three grouped breakouts were detected. In that case we use the extreme breakouts for t_d and t_s and the middle breakout for t_r that correspond to the group of breakout points best reflective of the disaster dates listed in Table 4.1.

All cases except the Louisville Tornado reflect statistically significant frequency changes from the stable period to the disrupted period. While the p_1 value for the Louisville Tornado exceeds 0.05, we feel this is likely due to the lack of available data. We note that there is a statistical difference in the variance of this same data before and after the disruptive event. Again, by grouping the disasters by their type, the Winter Storm events are those with the smallest increase (101.64 and 126.16%), then the Tornado events with 139.698 and 149.51% and finally the Black Forest Fire has the highest rate increase with 209.10%. Interestingly, the Black Forest Fire has both the highest increase and the smallest recovery time. Regarding the recovered state of each disaster, the null hypothesis of difference in means (i.e. the system has recovered to a stable state statistically equivalent to the original state) is not rejected for the Tornado events and the Black Forest Fire. However, it is rejected for both Winter Storms, therefore, the frequency increase for the NY Winter Storm and the frequency decrease for the SC Winter Storm that are observed during the 7 days following t_s are statistically significant.

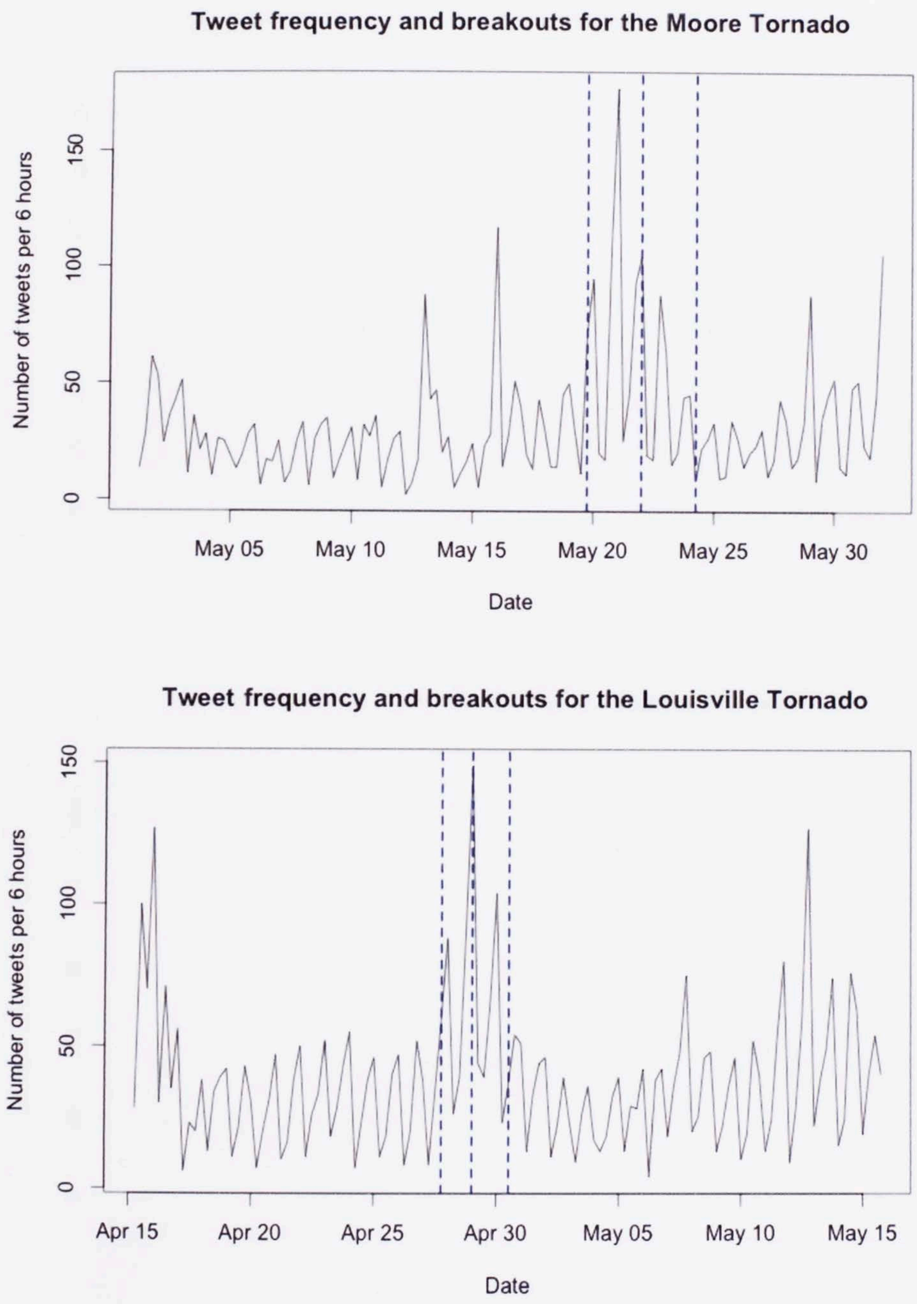


Figure 4.1: Tweet frequency for the Tornado events

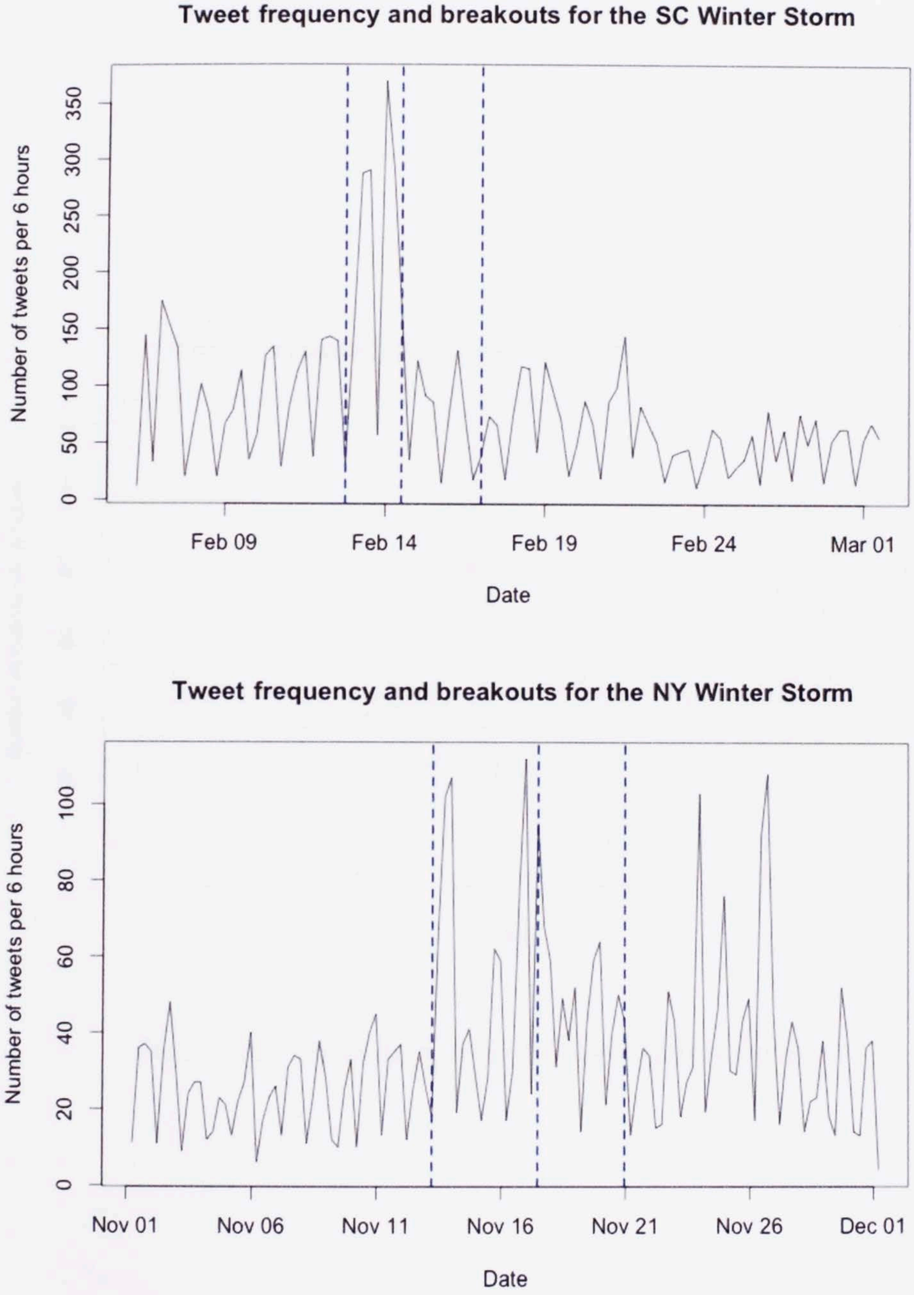


Figure 4.2: Tweet frequency for the Winter Storm events

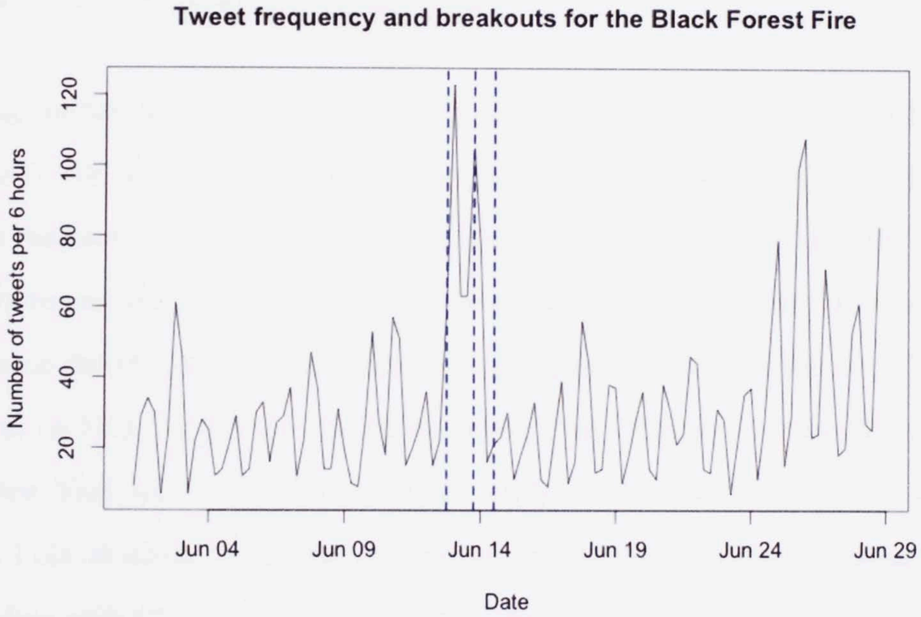


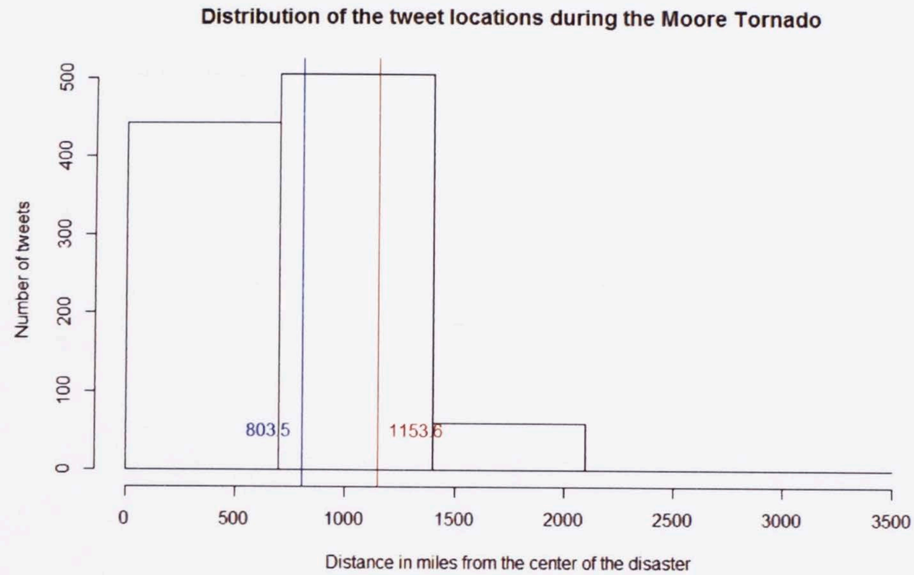
Figure 4.3: Tweet frequency for the Wildfire events

Table 4.5: Results of the proximity analysis for each natural disaster

Event	Median	Q3	Stdev
Moore Tornado	803.51	1153.59	432.12
Louisville Tornado	517.75	776.42	466.11
Black Forest Fire	1231.31	1479.69	350.97
NY Winter Storm	459.98	1121.30	645.79
SC Winter Storm	537.63	674.51	456.74

4.3. Proximity Analysis

Using the breakouts detected in Section 4.2, we look at the tweets posted between t_d and t_s for each disaster (see Table 4.3) and more precisely at the distribution of their distance from the center of the related disaster. Table 4.5 lists for each case study the median, the third quartile value (Q3) and the standard deviation of the distance distribution. Median values range from 459.98 miles for the NY Winter Storm to 1231.1 miles for the Black Forest Fire. However, if the Winter Storm in New York has the smallest median value, it has the third highest Q3 value with 1121.30 miles, where the smallest Q3 value is for the Winter Storm in South Carolina with 674.51 miles and a median of only 537.63 miles. The Winter Storms have less than 80 miles difference in their median values (despite an important gap in their Q3 values). However, it is not the case for the Tornadoes. We observe a high median value for the Moore Tornado compared to a much lower median value for the Louisville Tornado. Finally, Figures 4.4 and 4.5 depict for each disaster the histogram of the distance distribution of tweets posted between t_d and t_s. The median value is represented in blue and the third quartile value in red.



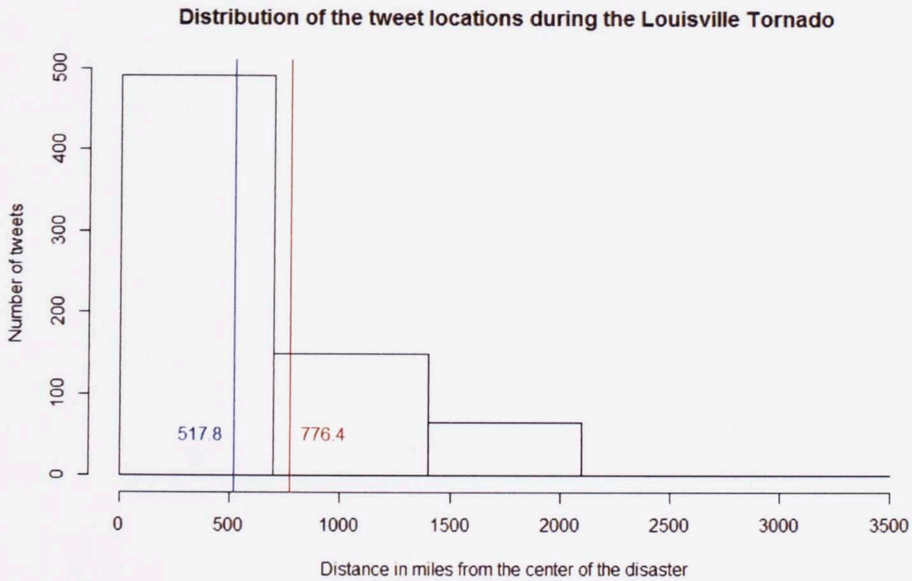


Figure 4.4: Distance distribution for the Tornado events

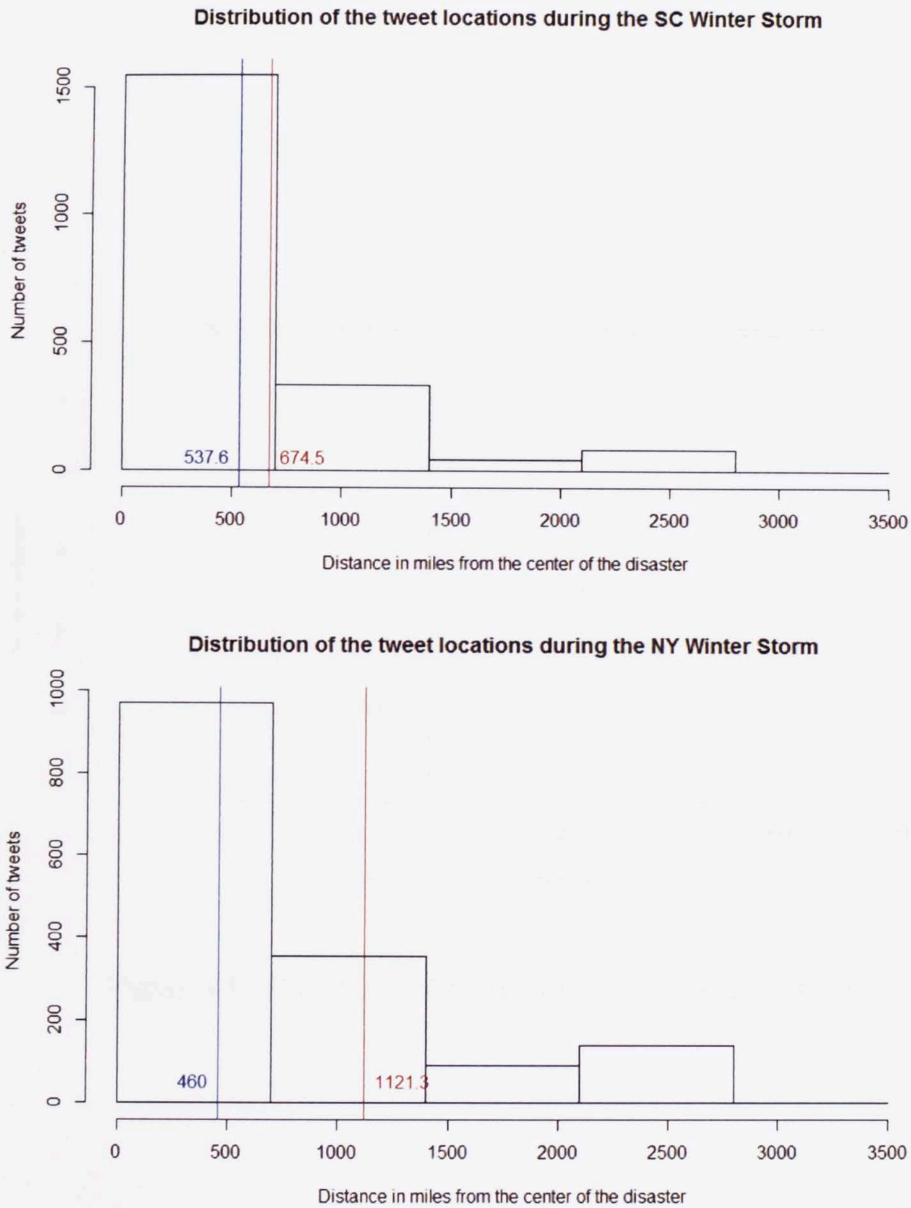


Figure 4.5: Distance distribution for the Winter Storm events

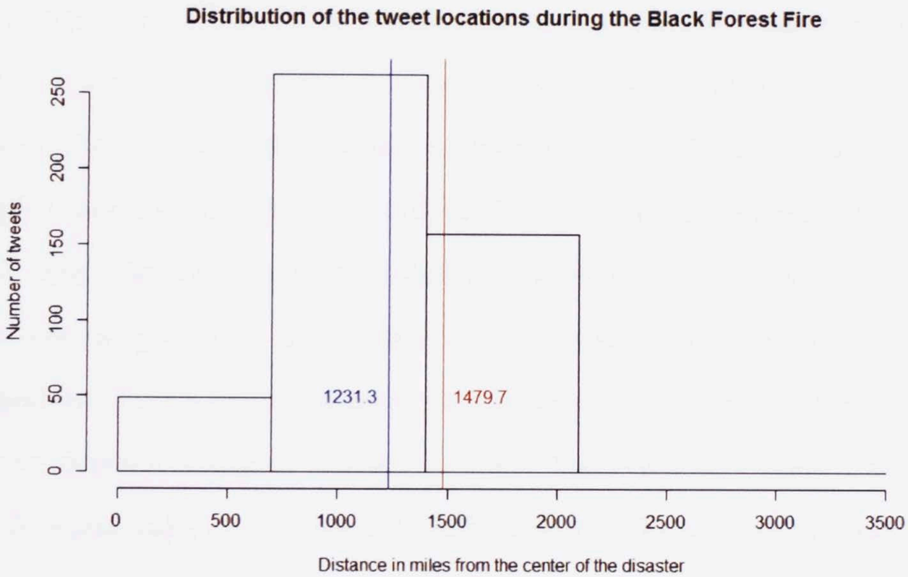


Figure 4.6: Distance distribution for the Wildfire events

4.4. Sentiment Analysis

Using the breakouts detected in Section 4.2, we look at the tweets posted between t_d and t_s for each disaster (see Table 4.3) and more precisely at the tweets labeled as expressing a negative sentiment after the sentiment analysis seen in Subsection 3.2.3. Table 4.6 lists for each case study the variations of the skewness of the distance distribution of these negative tweets. The value γ_0 is defined as the skewness of the distance distribution of negative sentiment tweets during the 7 days prior to t_d. This value is compared to the disrupted skewness γ_d between t_d and t_r in order to establish the rate increase. The significance of the difference between the two skewness means is statistically verified with a independent two-sample t-test and the p-value p'_1 obtained. The value γ_s, defined as the recovered stable state skewness, computed during a week after t_s, is also provided to see if the system goes back to the same state as before the disruptive event or not. We perform a data transformation for the skewness values in order to only have positive values to compute the rate increase. Each skewness value γ_i from the list γ of skewness values over time is transformed as follows: $\gamma_i \leftarrow \gamma_i + \min(\gamma) + 1.0$. The p-value p'_2 is also obtained similarly as p'_1 except it corresponds to the test for difference between γ_0 and γ_s. Since we are filtering again the data by the polarity of tweets, thereby diminishing the amount of available data over time, tweets have been grouped by 12 hour time windows.

Similarly to [4], the idea is to verify if during natural disasters, negative sentiment tweets tend to cluster closer to the disaster, that is having a higher positive skewness in their distance-to-disaster distribution during the disaster. Regarding the results obtained in Table 4.6, we observe a statistically significant increase of the skewness for the Louisville Tornado and both Winter Storms.

Table 4.6: Results of the sentiment analysis for each natural disaster

Event	γ_0	γ_d	γ_s	**% inc.**	p'_1	p'_2
Moore Tornado	2.409 (0.525)	1.839 (0.469)	2.193 (0.569)	-23.67	0.0153	0.3069
Louisville Tornado	2.230 (0.318)	2.750 (0.206)	2.215 (0.504)	23.32	0.0036	0.9264
Black Forest Fire	2.011 (0.430)	1.486 (0.543)	2.737 (0.378)	-26.12	0.0847	6.65e-05
NY Winter Storm	2.340 (0.412)	2.832 (0.288)	2.936 (0.711)	21.02	8.54e-04	0.0116
SC Winter Storm	2.4817 (0.504)	3.0425 (0.465)	2.5337 (0.605)	22.60	0.0179	0.8067

On the other hand, the skewness for the Moore Tornado actually decreases by 23.67%. Despite the fact that we do observe clustering of negative tweets for different types of disasters, however, these increases are within the same range of ~21-24%. We finally observe another decrease of 26.12% for the Black Forest Fire, nevertheless its related p'_1 value shows that it is not statistically significant, likely due to the brief disrupted period and thereby a lack of data. The p'_2 values obtained show that the skewness goes back to a statistically equivalent recovered state to the original stable state for both Tornado events and the SC Winter Storm. The negative sentiment tweet skewness shifts to an even higher positive skewness for the Black Forest Fire and the NY Winter Storm. Figures 4.7 - 4.9 depict for each disaster the skewness variations per 12 hours.

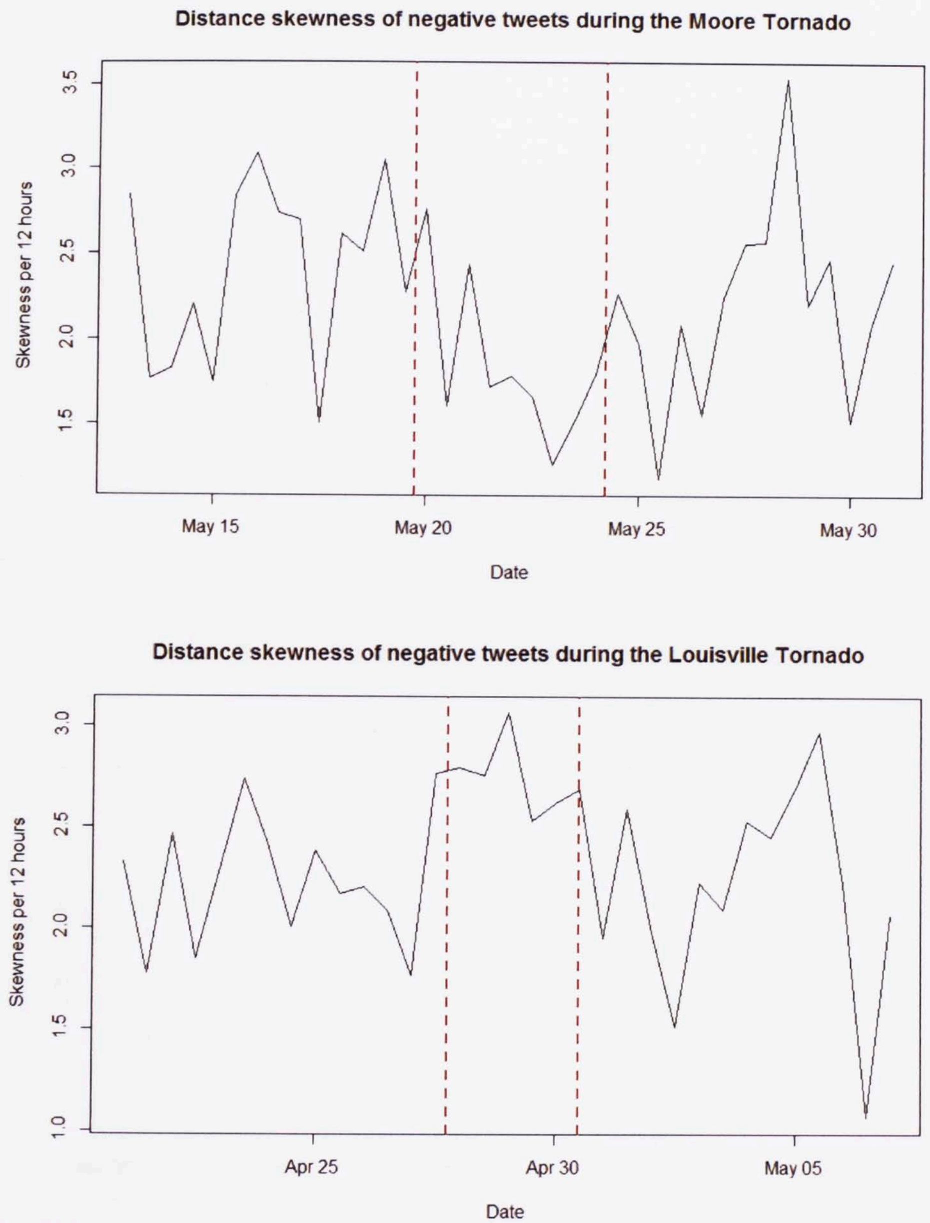


Figure 4.7: Skewness of negative tweet distance distribution for the Tornado events

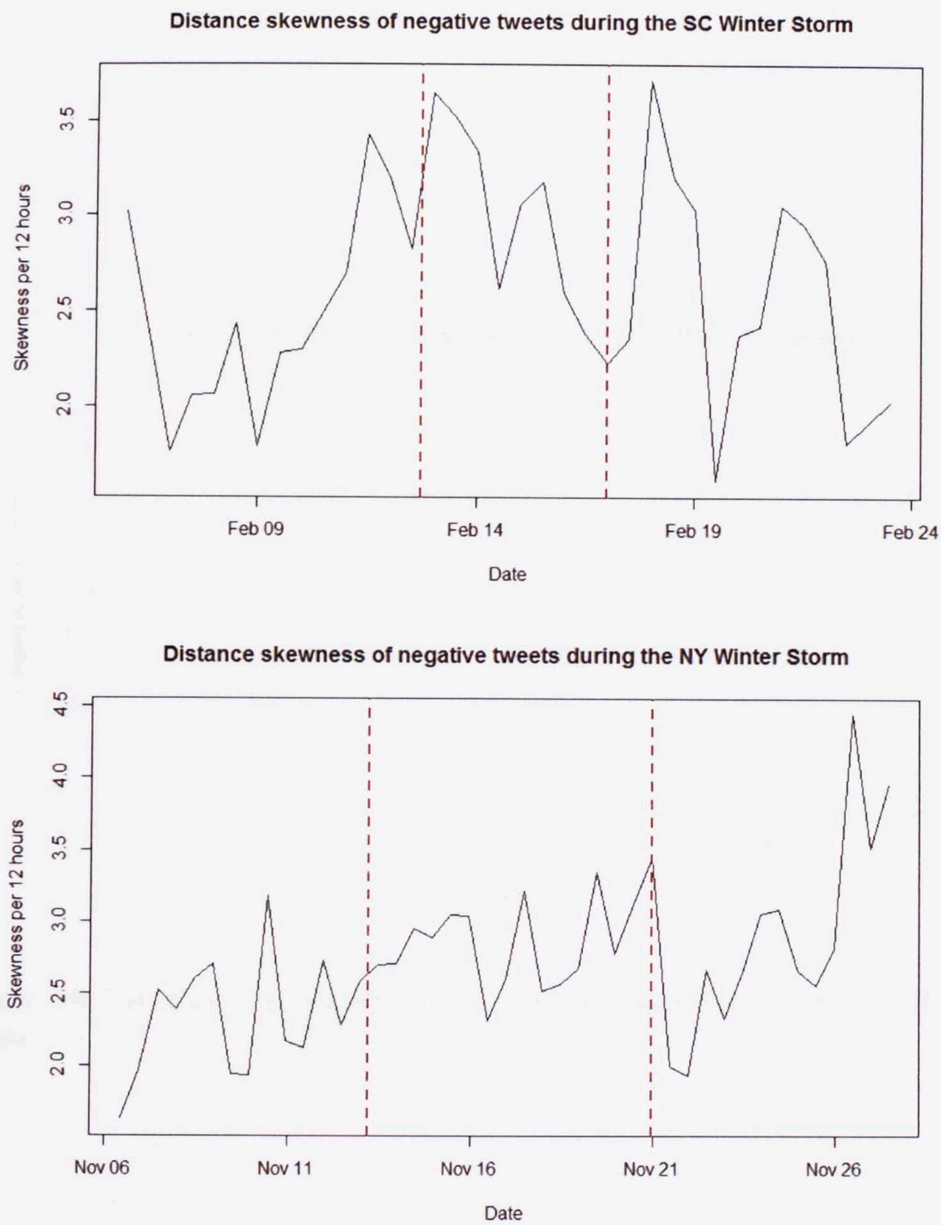


Figure 4.8: Skewness of negative tweet distance distribution for the Winter Storm events

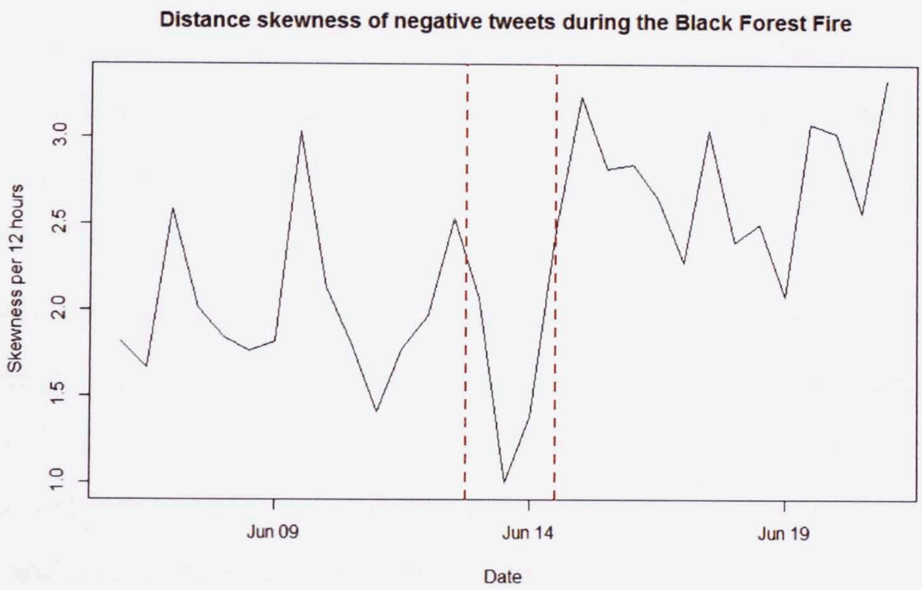


Figure 4.9: Skewness of negative tweet distance distribution for the Wildfire events

4.5. Social Vulnerability Approach

Each tweet is labeled with a social vulnerability level defined as low, medium or high based on the SoVI [5] and the county where it comes from. Once again, we use the breakouts listed in Table 4.3 and define for each level of social vulnerability the original frequency (ν_0^{low}, ν_0^{med} and ν_0^{high}) during 7 days prior to the disaster, the disrupted frequency (ν_d^{low}, ν_d^{med} and ν_d^{high}) between t_d and t_r, the rate increase and the p-value from the t-test to determine whether or not the difference in means is statistically significant (p_{low}, p_{med} and p_{high}). Results are listed in Table 4.7 with the statistically significant differences highlighted, and Table 4.8 lists for each case study the percentages of tweets by social vulnerability level of the counties they are from.

From the results, we observe that not all the case studies have a statistically significant increase from mediumly vulnerable areas, like for the Louisville Tornado or the SC Winter Storm. Both Winter Storms have a significant increase in areas with low and high vulnerability, and the NY Winter Storm observes significant shifts in all areas. Interestingly, in four of five disasters a significant increase in the highly vulnerable areas is observed. For the areas with medium social vulnerability, we observe less significant changes – three of five disasters demonstrate statistically significant changes, although the frequency shift for the SC Winter Storm is nearly significant. For the counties with low social vulnerability, we observe even less statistical evidence of a behavioral change: a significant increase in three of five disasters, however, one of which (the Louisville Tornado) is close to the significance threshold. The statistically significant rate increases observed for high vulnerability areas range between 62-166% while those observed in medium and low vulnerability range from 98% to 306%. That is, due to the inherent variability in the data by SoVI level, it seems that Twitter is more sensitive to changes in the more vulnerable areas.

The proportion of tweets for each of the three vulnerability levels are relatively consistent across each of the disaster scenarios under investigation. Approximately 11.5 – 15.5% of the tweets originate from low vulnerability areas, whereas 53.5 – 60% and 28.5 – 32.5% of the tweets originate from medium and highly vulnerable areas, respectively. Figures 4.10 - 4.12 depict for each disaster the tweet frequency per 6 hours by social vulnerability level.

Table 4.7: Results of the social vulnerability approach for each natural disaster

Event	ν_0^{low}	ν_d^{low}	% inc.	p_{low}	ν_0^{med}	ν_d^{med}	% inc.	p_{med}	ν_0^{high}	ν_d^{high}	% inc.	p_{high}
Moore Tornado	3.48	8.1	132.66	0.0726	19.19	48.9	**154.88**	**0.0141**	9.0	18.7	107.78	0.0689
Louisville Tornado	4.19	17.0	**306.19**	**0.0496**	15.89	42.0	164.34	0.0704	10.19	16.5	**62.0**	**0.0201**
Black Forest Fire	3.85	12.6	227.12	0.0710	13.59	50.2	**269.32**	**0.0115**	9.41	20.2	**114.72**	**6.047e-06**
NY Winter Storm	3.78	7.5	**98.53**	**0.0150**	13.52	29.22	**116.16**	**0.0039**	8.56	15.44	**80.52**	**0.0146**
SC Winter Storm	14.69	38.50	**162.04**	**0.0250**	53.12	105.38	98.39	0.0508	23.50	62.50	**165.96**	**0.0440**

Table 4.8: Percentage of tweets by level of social vulnerability

Event	Low (%)	Medium (%)	High (%)
Moore Tornado	11.59	60.11	28.30
Louisville Tornado	14.71	53.33	31.96
Black Forest Fire	13.34	54.26	32.39
NY Winter Storm	13.54	55.83	30.63
SC Winter Storm	15.55	55.17	29.28

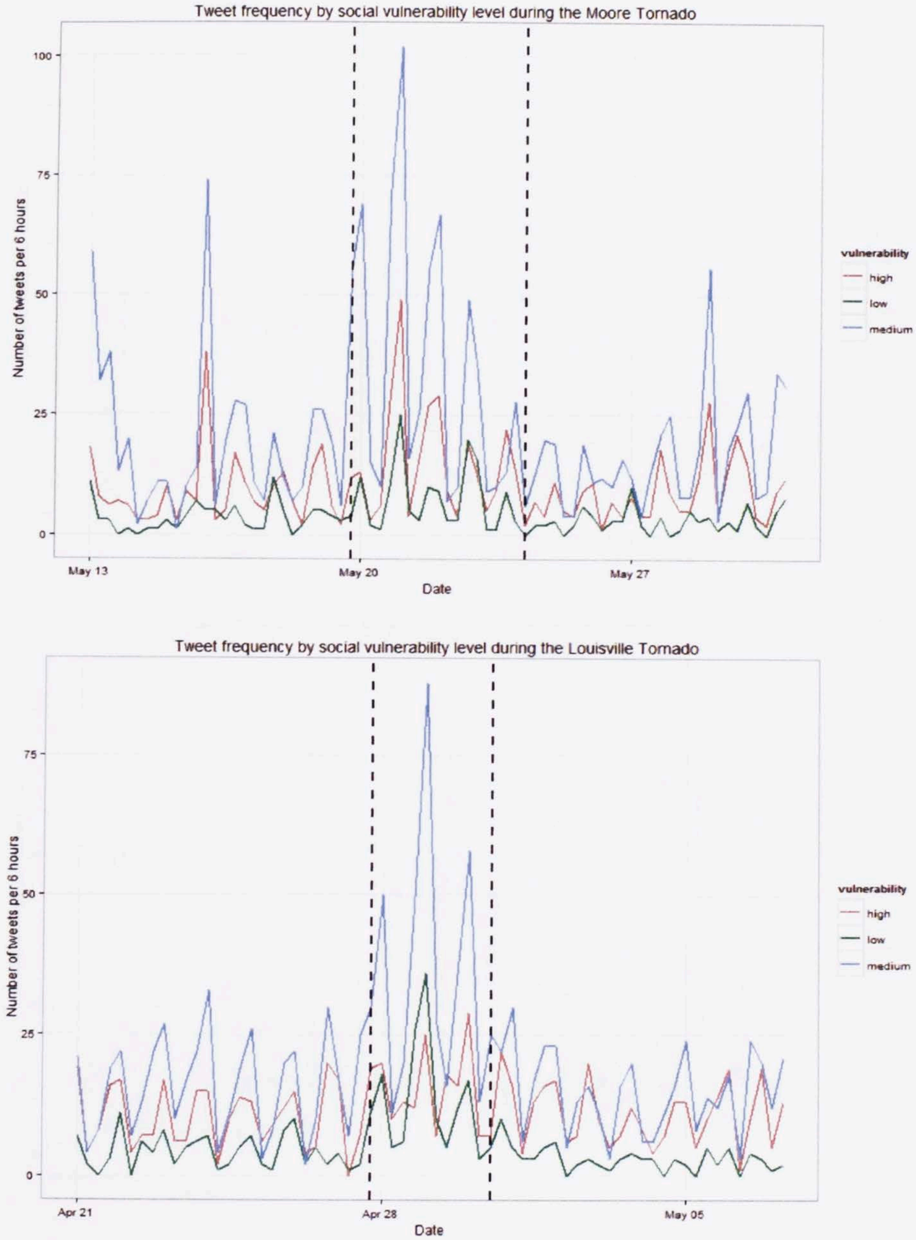


Figure 4.10: Tweet frequency by social vulnerability level for the Tornado events

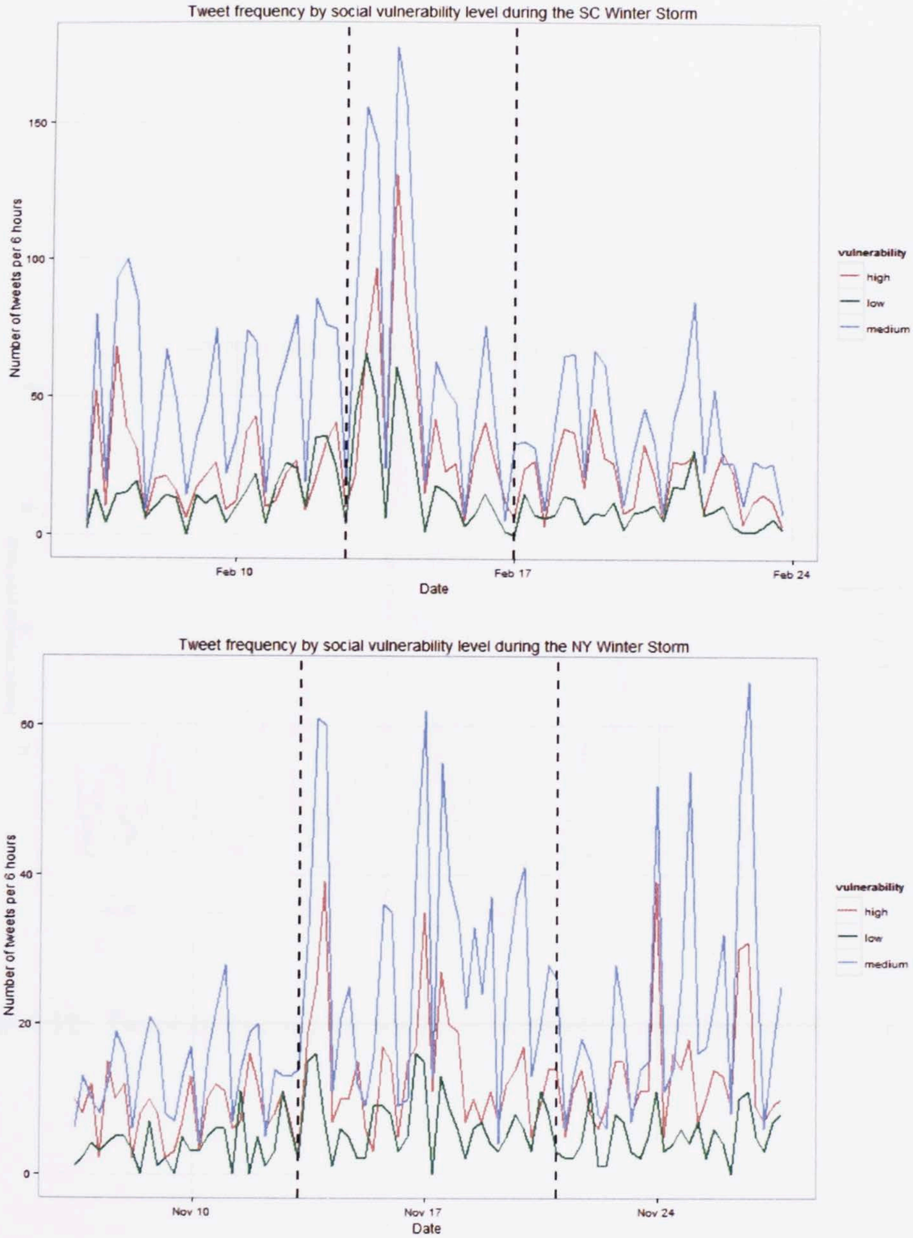


Figure 4.11: Tweet frequency by social vulnerability level for the Winter Storm events

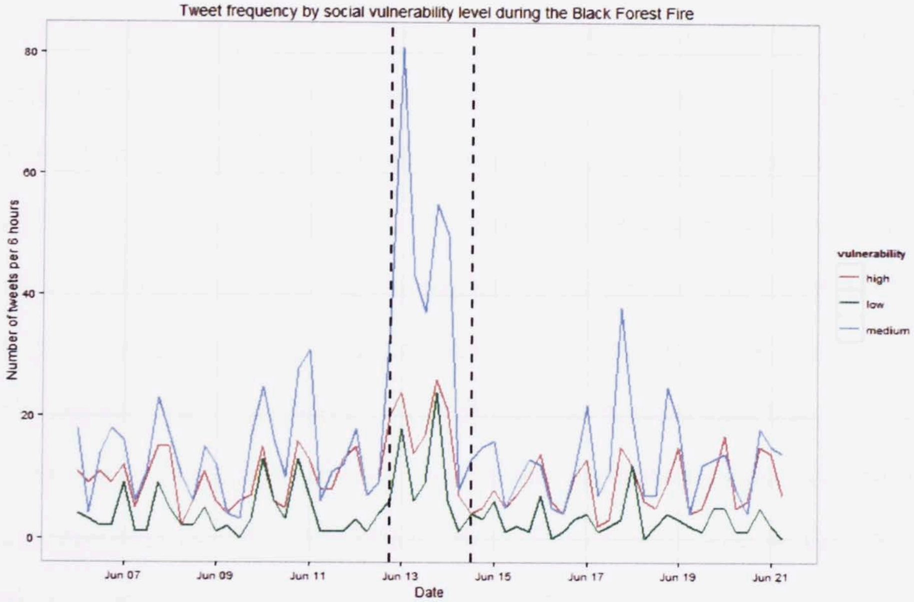


Figure 4.12: Tweet frequency by social vulnerability level for the Wildfire events

4.6. Discussion

First, regarding the results obtained in the frequency analysis, and more particularly regarding the breakout detection and recovery times, we observe different patterns. On average, the total observed time of disruption is the longest for Winter Storms, followed by the Tornado events and lastly the Wildfires. Tornado breakouts can be brief events compared to Winter Storms, so this order makes sense. However, for the Black Forest Fire, we do observe a significant variation in terms of tweet frequency but the breakout detected for the end of the event is not even close to the end date of the disaster, as if people only talked about it when the fire started and then stopped, maybe because they were not directly impacted by the disaster. A second increase can be observed after the fire has been contained and can be explained by people and news media talking about the damages and casualties caused by the fire once the final numbers were publicly known. We also observe for both Winter Storms that they do not recover to a statistically equivalent state as before the disruptive event. We interpret the increase observed for the NY Winter Storm in a similar way as the Black Forest Fire, that is the storm stopped several days after the breakout detected on social media, and therefore this increase is possibly due to post-event traditional media coverage. The decrease of frequency observed for the SC Winter Storm can be interpreted in the opposite way. Indeed, we observe that the beginning of the storm occurred before we detect it on Twitter data, hence we do have pre-events that influence the frequency before the first breakout. As previously stated, an interesting finding is that the recovery time detected is proportional to the total observed time, and except for the Black Forest Fire, we observe that it takes more time to recover to a stable state than shifting to a disrupted state. Consequently, we conclude that the tweet frequency along with the breakout detection is sensitive to different types of natural disasters.

Regarding the results obtained during the proximity analysis, the first trend we observe is that, regarding Winter Storms, the median distance is about the same for both events. However, the difference between their Q3 values is important. We explain it by the fact that, despite the highest damages occurring in the state of New York, this storm originated from the Pacific Northwest before hitting the East coast of the U.S., while the SC Winter Storm mainly hit the Southern/East coast of the United States. We also observe that, for the Black Forest Fire, tweets are widely spread in the country, which can be related to the fact that since a forest fire does not directly hit a large population area (at least not at the beginning), tweets are less clustered close to the disaster. Finally, for the Tornado events, we observe for the Louisville Tornado values slightly higher than for the Winter Storms. However, for the Moore Tornado, we do have values way above the Louisville Tornado. We explain it by the fact that this specific event, due to the high damages it caused as an EF5 tornado and the casualties, has been uniquely relayed in media and has been an important topic all over the country. We conclude that the proximity is in indeed sensitive to different types of natural disasters.

Regarding the results obtained during the sentiment analysis, in the four cases where the negative sentiment tweet skewness variation was statistically significant during the natural disaster, we observe in three of them an increase of the skewness of the distance distribution, that is tweets that express a negative sentiment cluster closer to the disaster. The increase is relatively constant in these cases, around 22% regardless the type of disaster. Nevertheless, the fourth case, which is the Moore Tornado, observes a decrease of this skewness, that is negative tweets clustered further during the disaster. This can be explained by the same reasons we observed higher values for this case in the proximity analysis. Furthermore, the skewness does not go back to normal for two disasters, the Black Forest Fire and the NY Winter Storm. In both cases, the skewness shifts to an even higher positive skewness during the recovered stable period. Therefore, either the breakout detected should be modified to account for the full disaster or, more likely, longer term events require more complex analyses. Such analyses should account for end-of-event traditional media spikes, recapping final casualty and damage estimates. Finally, we conclude that the sentiment is indeed sensitive to different types of natural disasters.

Lastly, regarding the results obtained during the social vulnerability approach, we clearly observe the most statistically significant changes in highly vulnerable areas compared to counties with medium and low social vulnerability in which we observe, respectively, less and even less significant changes. We are drawing the conclusion that the sensitivity level to natural disasters in terms of tweet frequency increases as the social vulnerability of the concerned areas increases, that is areas highly vulnerable are likely to be more sensitive than those less vulnerable. Therefore the social vulnerability level is sensitive to different types of natural disasters.

Chapter 5

Conclusion

The principal motivation to conduct this research is, to the best of our knowledge, the gap in literature regarding Twitter sensitivity levels to different types of natural disasters. Consequently, we built a set a general keywords, defined several case studies and studied the sensitivity of Twitter to these events on different metrics: the tweet frequency with breakout detection, the proximity-to-disaster and the sentiment expressed. Additionally, a novel contribution is the analysis of the frequency sensitivity by level of social vulnerability of the population.

The results of our empirical analysis show that indeed, each of these metrics is sensitive to natural disasters and the type of disaster. In the Discussion Section, we also explain that the traditional media coverage of an event may play an important role. The breakout detection method successfully captures the beginning and ending of shorter duration events (e.g. Moore Tornado), however, the end breakout dates do not correlate well with the end of longer duration events (e.g. Black Forest Fire). This is evidence to support the hypothesis that the intensity of the social media response is not sustained for longer events which may loose the attention of the public over time. However, renewed interest is evidenced in post-disaster social media spikes, possibly due to traditional media coverage recapping damage and/or casualty reports.

An important challenge is the potential lack of available relevant data. We filter Twitter data with general disaster related keywords, as opposed to event related keywords to help generalize the efforts and not to focus on single events (possibly biasing the results). We also filter based on language, presence of geo-tags and geo-location, which may lead to an insufficient amount of available relevant data.

To summarize, we hypothesize the true effects of Twitter as a social sensor to natural disasters as a function f of the informative content related to natural disasters plus several potential sources of error as depicted in Equation 5.1. There is a need to account for three of these factors by minimizing their respective error in order to better process and analyze Twitter data.

$$\begin{aligned} T_{sensor} = f\Big(T_{\substack{informative,\\ content}}&,\ Filtering,\ Breakout\ Detection, \\ &Sentiment\ Detection,\ \epsilon \Big) \end{aligned} \tag{5.1}$$

Filtering is the most important factor. The validity of using Twitter as a sensor is dependent on the ability to extract informative content (signal) from the vast amount of unrelated content (noise) in social media. The goal with filtering is to isolate the relevant data from the noise without eliminating the signal. Having a sufficient amount of data is important, especially when the lexicon does not include keywords specifically related to a particular event. Filtering by geo-tagging is also a non-negligible source of data reduction.

Several models and tools to detect shifts in time series data exist, the goal of *Breakout Detection* is to detect the breakouts that best reflect the disaster related social media response and thereby mitigating the modeling error. The tool used in this study was developed by Twitter and is fast and robust against the presence of anomalies. However, further enhancements to such tools will benefit signal detection.

Regarding the *Sentiment Detection* factor, analyzing the sentiment expressed in a text is a difficult task that requires sophisticated tools and lexicon in order to minimize the misclassification. In this analysis we use the Hu and Liu lexicon and compute a polarity score for each tweet. Sentiment analysis in general is still an active area of research.

ϵ represents the irreducible error, that is the variations of human activity in social media.

We believe that the informative content is a function g of the following factors:

$$T_{\substack{informative \\ content}} = g\,(Proximity,\ Social\ Vulnerability,\ Technology) \tag{5.2}$$

The *Proximity* or distance-to-disaster influences individuals to post content about a natural disaster. People in closer proximity to the disaster are more likely to post.

The level of *Social Vulnerability* is also a factor which influences sensitivity to natural disasters. More vulnerable areas observe more significant variations in Twitter behavior. We use in this study the SoVI in order to define this vulnerability at the county level.

Both access and adoption to *Technology* are factors that can influence the sensitivity. Indeed, populations need to have a device with power and connected to the Internet in order to use Twitter. It is also dependent on their willingness to use such online platforms to communicate, and their willingness to provide informative content, including personal sentiment, needs, and possibly advice and recommendations to affected individuals. In this study, we assume both the access and adoption to be sufficient.

In this work, we contribute to the *Filtering* of Twitter data by using a set of general keywords created by using Latent Dirichlet Allocation among other techniques. We use a recent and powerful tool developed by Twitter for the *Breakout Detection*, as well as efficient and common polarity estimation techniques for the *Sentiment Detection*. Finally, we analyze the impact of the *Proximity* on the public and we bring a new contribution by contrasting the sensitivity to natural disasters based on the *Social Vulnerability* level of the population. The sources of error and areas for potential future research include the improved ability to filter by keywords, to better detect statistically significant shifts in tweeting and an enhanced ability to analyze and understand sentiment.

References

[1] Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May 2010.

[2] Kash Barker, Jose Emmanuel Ramirez-Marquez, and Claudio M. Rocco. Resilience-based network component importance measures. *Reliability Engineering & System Safety*, 117(0):89–97, 2013.

[3] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003.

[4] Cornelia Caragea, Anna Squicciarini, Sam Stehle, Kishore Neppalli, and Andrea Tapia. Mapping moods: Geo-mapped sentiment analysis during hurricane sandy. In *Proceedings of the 11th International ISCRAM Conference*, University Park, PA, United States, May 2014.

[5] Susan L Cutter, Bryan J Boruff, and W Lynn Shirley. Social vulnerability to environmental hazards. *Social science quarterly*, 84(2):242–261, 2003.

[6] Susan L Cutter and Christina Finch. Temporal and spatial changes in social

vulnerability to natural hazards. *Proceedings of the National Academy of Sciences*, 105(7):2301–2306, 2008.

[7] Neil Dufty. Using social media to build community disaster resilience. *The Australian Journal of Emergency Management*, 27(1):40–45, February 2012.

[8] Paul S Earle, Daniel C Bowden, and Michelle Guy. Twitter earthquake detection: earthquake monitoring in a social world. *Annals of Geophysics*, 54(6):708–715, 2012.

[9] Paul Ekman. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200, 1992.

[10] Minqing Hu and Bing Liu. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177. ACM, 2004.

[11] Nicholas A James, Arun Kejariwal, and David S Matteson. Leveraging cloud data to mitigate user experience from "breaking bad". 2014.

[12] Mark E Keim and Eric Noji. Emergent use of social media: a new age of opportunity for disaster resilience. *American journal of disaster medicine*, 6(1):47–54, January/February 2011.

[13] Kalev Leetaru, Shaowen Wang, Guofeng Cao, Anand Padmanabhan, and Eric Shook. Mapping the global twitter heartbeat: The geography of twitter. *First Monday*, 18(5), 2013.

[14] Hongmin Li, Nicolais Guevara, Nic Herndon, Doina Caragea, Kishore Neppalli, Cornelia Caragea, Anna Squicciarini, and Andrea H Tapia. Twitter mining for disaster response: A domain adaptation approach. In *Proceedings*

of the 12th International ISCRAM Conference, Kristiansand, Norway, May 2015.

[15] Stuart E Middleton, Lee Middleton, and Stefano Modafferi. Real-time crisis mapping of natural disasters using social media. *Intelligent Systems, IEEE*, 29(2):9–17, March/April 2014.

[16] George A Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.

[17] Ahmed Nagy and Jeannie Stamberger. Crowd sentiment detection during disasters and crises. In *Proceedings of the 9th International ISCRAM Conference*, pages 1–9, Vancouver, Canada, April 2012.

[18] Finn Årup Nielsen. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. In *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages*, pages 93–98, Heraklion, Crete, May 2011.

[19] Alexandra Olteanu, Carlos Castillo, Fernando Diaz, and Sarah Vieweg. Crisislex: A lexicon for collecting and filtering microblogged communications in crises. In *Proceedings of the AAAI Conference on Weblogs and Social Media (ICWSM'14)*, Ann Arbor, MI, United States, June 2014.

[20] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World Wide Web*, pages 851–860, Raleigh, NC, United States, April 2010. ACM.

[21] Takeshi Sakaki, Fujio Toriumi, and Yutaka Matsuo. Tweet trend analysis in

an emergency situation. In *Proceedings of the Special Workshop on Internet and Disasters*, page 3. ACM, 2011.

[22] Axel Schulz, Tung Dang Thanh, Heiko Paulheim, and Immanuel Schweizer. A fine-grained sentiment analysis approach for detecting crisis related microposts. In *Proceedings of the 10th International ISCRAM Conference*, Baden-Baden, Germany, May 2013.

[23] Kate Starbird, Grace Muzny, and Leysia Palen. Learning from the crowd: Collaborative filtering techniques for identifying on-the-ground twitterers during mass disruptions. In *Proceedings of the 9th International ISCRAM Conference*, Vancouver, Canada, April 2012.

[24] Kate Starbird and Leysia Palen. Pass it on?: Retweeting in mass emergency. In *Proceedings of the 7th International ISCRAM Conference*, Seattle, WA, United States, May 2010.

[25] Jeannette Sutton, Leysia Palen, and Irina Shklovski. Backchannels on the front lines: Emergent uses of social media in the 2007 southern california wildfires. In *Proceedings of the 5th International ISCRAM Conference*, pages 624–632, Washington, DC, United States, May 2008.

[26] Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558, 2010.

[27] Brandon Truong, Cornelia Caragea, Anna Squicciarini, and Andrea H. Tapia. Identifying valuable information from twitter during natural disasters. In *Proceedings of the 77th Annual Meeting of the Association for*

Information Science and Technology, Seattle, WA, United States, November 2014.

[28] Timothy Wiemken. Breakoutdetection: Using the new twitter tool for influenza surveillance. `http://blog.ctrsu.org/?p=96`, October 2014. [Online; accessed 2-December-2015].

Appendix A

Keywords

shelter	deadly	tragedy	dead missing
flooding	evacuated	fatalities	severe
snow	people killed	survivor	magnitude
hail	people dead	evacuees	emergency
victims	people trapped	public safety	flash flood
storm	people died	destroyed	surviving
explosion	donation	casualties	injuries
prayers	damage	fire fighters	warning
tornado	devastating	firefighters	threat
affected	inundated	volunteers	seismic
injured	crisis	ravaged	power supplies
wildfire	redcross	tragic	power outage
forest fire	red cross	impacted	destruction
donate	bushfire	torrential	rescuers
responders	breaking news	recover	dramatic
first responders	disaster	terrifying	hurricane