ANALYZING HUMAN INTEREST TO HURRICANE HARVEY USING LOCATION BASED SOCIAL MEDIA

AND TOBLER'S FIRST LAW OF GEOGRAPHY

by

Adailin Lebron Bengochea

A directed research report submitted to the Geography Department of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Applied Geography
with a specialization in Geographic Information Science

August 2018

Committee Members:

Dr. Yihong Yuan, Chair

Dr. Yongmei Lu, Committee Member

Table of Contents

1. Introduction

In 2017 the Unites States experienced sixteen separate billion-dollar natural disaster catastrophes, making it a historic year for weather and climatic disaster events. The events included: one drought, two floods, one freeze, eight severe storms, three tropical cyclones, and one wildfire event, all resulting in the deaths of 362 people in addition to significant economic impacts (NOAA 2018). Hurricane Harvey was one of the three tropical cyclones to make landfall in the United States and devastated the state of Texas.

Hurricane Harvey, considered as one of the costliest hurricanes in recent history, started as a low atmospheric pressure off the coast of Africa on August 13th, 2017 and later evolved to a Category 4 storm near Rockport, Texas due to the warm temperature conditions in the Gulf of Mexico. Hurricane Harvey was unlike any other storm because instead of continuing its projected course and loose intensity, the storm stalled over the Greater Houston area for five devastating days bringing record rainfall and destruction to the area. Houston is considered the fourth largest metropolitan city in the United States and is home to 2.3 million people. In the wake of the storm, the Texas Department of Public Safety (TxDPS) reported over 277,315 single family homes as damaged, of which 10,816 were completely destroyed; in addition, over 117,000 people were rescued and evacuated, and a total of 88 deaths were reported, which is a low number for such a devastating storm (Perryman 2017).

Public officials, disaster managers, flood victims and their families used social media, such as Twitter, Facebook, and Instagram during the turmoil days of pre and post Hurricane Harvey. Social media was used to share vital live time information including, but not limited to: warning communities of the storm dangers, issuing evacuation orders, requesting the need for assistance, and informing family and friends of the safety status of its users. For instance, a woman tweeted for help because 911 dispatchers were unresponsive, “I have 2 children with me and tge [sic] water is swallowing us up. Please send help,” and within an hour of her Tweet help arrived (Rhodan 2017). Additionally, the Texas county of Brazos used Twitter to inform the community of storm dangers, “GET OUT NOW” was Tweeted when one of the local county levees was breached due to large masses of rain caused by Hurricane Harvey (Kluger 2017). Furthermore, social media was a main contributor in helping raise funds post the storm event (Plummer 2017). For example, JJ Watt, a prominent Houston Texans defensive end football player, raised over 35 million dollars for the victims of Hurricane Harvey through the means of social media (YouCaring 2017).

The development of the World Wide Web provides easy access to social media and networking sites through GPS-enabled devices such as cell phones and tablets. This technology has facilitated the advent of big data analytics, allowing the collection of rich datasets linked to a geographic location and time stamp and providing researchers the possibility to explore a variety of human related topics, such as analyzing human behavior through semantics and analyzing human mobility patterns (Yuan, Liu, and Wei 2017). One of the many positive uses of social media has been for disaster management through establishing situational awareness before, during, and after a disaster by enabling people to communicate and share resources and information (Lindsay 2011). This is especially true for natural disasters because even though current meteorological technologies can predict advanced warnings for tropical cyclones and other natural disasters, such as winter storms and rainstorms, the technology is still incapable of accurately predicting paths and severity of these storms (Wang and Taylor 2016). In addition, many studies relating information diffusion theory have been conducted with the use of social media datasets to explore the exchange of information during emergency events. For example, Wang et al. (2016)'s explored how social media was used to distribute emergency information during the 2012 Beijing Rainstorm event (Section 2.3). This research can provide meaningful insights to disaster management and facilitate contingency planning by policymakers (Wang and Taylor 2016).

The purpose of this study is to explore human interest of Hurricane Harvey through twitter activity based on the geographic proximity to the event, i.e. the Greater Houston area where the storm stalled for five days. Understanding the impact distance has on human interest to a disaster will provide viable knowledge for emergency management officials, which may be used during the planning and response phase of a disaster. This study will also provide new methods for exploring how distance to a disaster impacts human behavior relating to the disaster. Since distance is quantifiable, producing a numerical value between any two points of interest can support the use of statistical analysis that can reveal patterns and trends applicable to the present and future research studies. The gravity model will be used to analyze the spatial interaction of Hurricane Harvey related Tweet since the model is effective in predicting the degree of interaction between geographical entities and its simplistic equation (Hardy, Frew, and Goodchild 2012). The results will also be related to Tobler's First Law of Geography which states, "everything is related to everything else, but near things are more related than distant things" (Tobler 1970), meaning locations closest to the Greater Houston area will generate more tweets relating to Hurricane Harvey versus locations further away from the Greater Houston area. Specifically, this study will answer the following question: Using twitter data generated during Hurricane Harvey, can Tobler's First

Law of Geography be applied in interpreting human interest to natural disasters based on distance to the event?

Based on the literature review analyzed for this study, there hasn't been sufficient study on the spatial interaction between distance and human interest to natural disasters. Therefore, this study proposes an empirical study for examining the effects of distance decay in relation to natural hazards and emergencies. This study can improve emergency management agendas by analyzing how distance influences human interest of natural disasters based on Tobler's First Law of Geography, assuming places closer in geography have a stronger connection (resulting in a higher tweet count) versus places further away. Some emergency management examples where this study may provide viable insights are: recognizing areas of high concern based on Tweet density, preparing social media response protocols to be used during a disaster, delivering information to areas where there is lack of journalistic coverage relating the disaster, and maximizing fundraising profits through the use of social media. In addition, the results of this study can provide for further research directions aiming at understanding the observed spatial interaction in this study.

2. Literature Review

2.1 Tobler's First Law of Geography

Waldo Tobler invoked the First Law of Geography: "everything is related to everything else, but near things are more related than distant things" in his study simulating population growth in Detroit in 1970 (Tobler 1970). Even though his study was in response to slow computers at the time, this law has been applied in an array of research relating to the spatial analysis of distance (L. Shi et al. 2017; Yuan, Liu, and Wei 2017). As stated in the introduction, the invention of the World Wide Web has allowed the access to rich datasets linked to a geographic location and time stamp allowing for researchers to explore an array of human related topics.

Furthermore, since the rise of geographic information science (GIS), TFL has become more useful now than ever as the quantity and quality of data are improving simultaneously with technology and TFL is essentially at the core of spatial autocorrelation statistics used in spatial analysis (Miller 2004). For example, Shi et al. (2017) uses Tobler's First Law (TFL) to analyze empirical geo-tagged mobile phone data to explore human mobility and the formation of social communities amongst individuals; this study indicates that different communities form based on the contributions of distance decay effect along with social phenomena such as the home-work separation and confirms individuals are more likely to form a social community with people physically nearby. Although this study does not use a social media network such as Twitter, it demonstrates the relevance of applying TFL to geo-tagged mobile phone data for measuring distance relationships amongst two entities.

Another usage where TFL was applied for analyzing distance relationship using a geo-referenced dataset is in a study conducted by Yuan, Liu and Wei (2017) where they model China's image in mass media using an open source dataset, "The Global Data on Events, Location and Tone" (GDELT); in addition, Flickr data and Airline Carrier data were used as complimentary datasets. The aim of the study was to explore the possibility of employing mass media data to quantify spatial connections between countries and how these connections evolve through time. This study uses geo-tagged datasets (Flickr and GDELT) to explore spatial relatedness to China. The results indicated that based on TFL, spatial adjacency facilitates international interactions; however, similarity must also be considered. For example, Nepal, Laos, and Bhutan were not strongly connected to China even though they were close in proximity. This study demonstrates how TFL can be applied to explore relatedness between a fixed location, China, and other countries, which is similar to the overall aim of this papers research study. This study will apply geo-enabled Twitter data for analyzing human interest to Hurricane Harvey based on the geo-referenced Twitter data. This will be achieved by conducting a spatial analysis using the gravity model (discussed in further detail in section 2.2) to demonstrate the spatial relationship between both Texas Counties and mainland United States to the Greater Houston area.

Furthermore, the use of TFL in addition to the aforementioned gravity model was also used by Liu et al. (2014) to analyze the relatedness of toponym co-occurrences in textual web documents for same-level geographical areas in China. This study examined the textual connotation in web documents by counting the geographic component, i.e. Chinese province names, and relating their co-occurrence via distance and similarities. Results exhibited the basic assumption of TFL, toponyms that are spatially close have similar properties allowing for a regionalized method of examining Chinas geographic organization. The geographic component of each web produced text was analyzed and a distance relationship was created via the gravity model, proving TFL in conjunction with the gravity model is a good approach for analyzing a spatial relationship between two entities.

2.2 The Gravity Model

The gravity model is derived from the universal law of gravity that states objects apply force to each other based on the direct proportionality to their mass and inverse proportion to the distance between them; basically, the force between two objects becomes weaker as distance increases (Koçaslan 2017). The gravity model is widely used in economics due to its success in explaining different types of flows, such as migration, commuting, tourism, and commodity shipping (Bergstrand 1985). Since the model ultimately measures distance, the gravity model is attracting more attention in geographical spatial analysis. This study applies the gravity model in conjunction with TFL: “everything is related to everything else, but near things are more related than distant things” (Tobler 1970). By using the gravity model and TFL this study explores the spatial interaction of Hurricane Harvey Tweet messages. The aim of this study is to demonstrate that places closer in distance are more related than places further away by verifying locations in close proximity to the Greater Houston area ultimately produce more Tweets relating to Hurricane Harvey (strong attraction) versus locations further away (weak attraction).

The gravity model is shown below (Eq. 1):

$$I_{ij} = K \frac{P_i P_j}{D_{ij}^{\beta}} \tag{1}$$

where P_i and P_j are the conceptual sizes (relative importance) of locations i and j, D_{ij} represents the distance separating the geographic mean centers of i and j, and I_{ij} denotes the interaction/connection between i and j. β is the distance friction coefficient and investigates the role of distance; larger values represent higher degree of distance decay (Yuan 2017). The constant K does not affect the model fitting, it serves to adjust the magnitude of interaction.

Since the gravity model ultimately measures distance, it is commonly used in research to demonstrate spatial interaction between two geographic locations. For example, a recent study explored the distance decay effect of two data sources: mass media (GDELT data) and a location based social media in China (Weibo data); this study used a composed gravity model to analyze connections between Chinese provinces through mass media and social media (Yuan 2017). Based on the results for the best fit coefficient, Yuan demonstrates that mass media displays a weak distance decay effect, while the social media site demonstrates a strong distance decay. In addition, Yuan ends the model construction section of the paper by listing different aspects of uncertainty in the study. Uncertainties included the natural variability in human activities, the potential impact of spatial autocorrelation, and data imprecision solely due to signal strength in the GPS device. Lastly, Yuan discusses the limitations of models and algorithms, stressing not all models align with datasets and vice versa. Although not directly related to analyzing human interests of an event based on distance, this study utilizes the gravity model to measure the connection or degree of interaction between places based on distance as well as applying a Chinese social networking site closely related to Twitter, with geo-tagged information for recording location. This research successfully applies the gravity model to investigate the spatial interaction between geographic locations.

In addition, the gravity model in conjunction with location-based data is commonly used to explore human mobility patterns since it also takes into consideration the direction of flow. Yuan and Medel (2016) use the gravity model to analyze how distance impacts travel behavior for international and intercountry travel via user uploaded Flickr data, a photo and video sharing networking site that maintains the geographic integrity of the uploaded media based on the location where it was taken. Results indicate that distance of travel is reflected by the size and geographical location of the home county and that travel flows to different counties are based on the proximity to home country (Yuan and Medel 2016). Liu et al. (2014) also used the gravity model along with an unspecified major Chinese location based social network service to explore check in data to analyze spatial interactions between cities via human inter-urban trip patterns; this study confirmed distance plays a significant role in human mobility patterns and spatial interactions. Prediction of travel behavior is also a common human mobility topic explored via the gravity model. Pappalardo, Rinzivillo, and Simini (2016) examined how the gravity model and cell phone data can be used to explore an individual's movements and provide reasoning for exploring new places. Results demonstrate that individual mobility patterns are driven by the individuals force to return to the same place and the individuals collective force during an explorative phase. The aforementioned papers validate the feasibility and applicability of applying the gravity model in conjunction with location based data to explore the distance decay and provides inductive reasoning for the use of the gravity model in this study.

The majority of the literature reviewed relating to the gravity model focuses on either demonstrating distance relatedness between two places or modelling human mobility. However, this study will apply the gravity model to static data represented by tweet mean centers in each state, counties in Texas, and the Greater Houston area (see Section 3.1 for additional details) to explain the relationship between distance and human interest (analyzed by tweets relating to the event). The results of this study will provide viable knowledge for disaster management personnel by investigating how people's concern/interest to an event is affected based on distance to the event. Local and government officials can use the results of this study for future contingency plans for disaster management. Some examples of how the results of this research can be used are: improving disaster relief by delivering information to locations where there is lack of journalistic coverage relating to the event and/or tracking financial budget through fundraising and the sharing of information via social media. In addition, the results of this study can provide for further research directions in investigating the reason for the observed distance relationship in this study.

2.3 Social Networks and Natural Disasters

Social media is a valuable tool in disaster management as well as for understanding human behaviors. Lindsay (2011) summarized a report on the uses, future options, and policy considerations of social media and disasters. According to Lindsay (2011), social media ranks as the fourth most popular source to access and share emergency information. It can be used for multiple applications, such as warning others of unsafe areas or situations, informing friends and family of safety status, and raising funds for disaster relief. Lindsay (2011) categorizes the use of social media in two ways: to communicate and share information, and to be as an emergency management tool. In addition, this report claims that citizen posts are more relevant and searched than emergency management agency or organizations; Lindsay (2011) supports this statement with an example from the 2007 Virginia Tech shooting event where most warning messages came from students and unofficial sources. This report also stresses the importance of social media used as a means for requesting aid and employs an example from the 2011 Japanese earthquake where individuals tweeted for help when phone call access was unavailable. Considerations to be taken during emergency management are also addressed in this report, such as the possibility of malicious usage during a disaster. Lindsay concludes the report by stating that as social media becomes more popular, more people will turn to it as the main source of information. This report illustrates the appropriate uses of social media during and after a disastrous event and provides support for stressing the relevance of this study.

A relevant study which represents the value of analyzing social media data relating to a natural disaster is the study conducted by Xiao et al. (2015) which examines the spatial heterogeneity of Twitter posts in New York City after the major disastrous event of Hurricane Sandy. In this study, the usage of social media was categorized as: 1) real time dissemination of information, meaning users shared event related information in real time instead of waiting for media coverage, 2) establishing situational awareness, where citizens are viewed as sensors which can be monitored during the evolution of a situation and can provide direction for policy makers during emergency situations, and 3) as a mean of backchannel communications, meaning that citizens can provide information when all other means of information are unavailable. This study also discusses the uncertainties associated with social media usage during an extreme event, such as biases in the data, data accuracy, and as mentioned by Lindsay (2011) the malicious use of social media to redirect emergency respondents. Xiao's (2015) study uses a regression analysis to analyze the generation of tweets; results drawn from this study indicate that population, damage levels, and wealth status play a significant role in the generation of tweets. This study also only examines tweets relating to Hurricane Sandy from New York City; this study could be improved by incorporating nationwide tweets relating to the event.

Understanding human emotional response to major events is also a valuable tool for emergency management. Li et al. (2017) explores the semantics of Twitter messages in relation to three extreme events to demonstrate how understanding human emotional responses based on social media data can help governments and health agencies gain insight to managing relief missions and meet readiness requirements. Results showed that people express different emotions to distinct types of events, such as anxiety to the Swine Influenza

Virus and anger to the Lybia crisis. In addition, this study also indicates the need to produce quick and specific information regarding an event to reduce anxiety levels in the public. Although this study only addresses the semantics of Twitter messages, it still stresses the importance of how Twitter data can be utilized for emergency management. In addition, the incorporation of a geographic component in this study could provide further aid in emergency management by demonstrating different emotional responses in relation to the event location.

Furthermore, as stated in the introduction, another research agenda where social media has widely been used is in the study of information diffusion theory. Dong et al. (2018) analyzed information diffusion on China's largest social networking site, Weibo, for two earthquake events in China. The study researched information diffusion from three main perspectives: individual characteristics, types of social relationships between social media users, and the topology measurement of real interaction networks. An interesting observation made in this study is the basic assumption that demographics, language and geographic distribution of online users are important "sensors" in disaster relief, surveillance and control; in addition, it was noted that there were few foreign users that participated in posts relating to the event, nearly all of which lived in China. This observation adds context to the presented research question since it proves that people that do not live in China or are not "from" China do not display the same level of interest as a person that is of Chinese descent living in China. In addition, this study present ideas on how to enhance rescue efforts during earthquake events.

Another example of a study that used information diffusion theory is Wang et al. (2016)'s study where social media was used to explore how emergency information was distributed during the 2012 Beijing Rainstorm event. This study also used Weibo as the social media networking site and found that social network activity fluctuated during stages of the rainstorm event while also varying among different topics relating to before, during, and after the rainstorm. An interesting finding in this study was areas with a large concentration of Weibo activity where linked to more intense rainfall events, confirming the distance relationship of TFL where areas “closer” to the rainfall were more connected, produce more activity, than areas further away. This study also demonstrated how exploring the spread of emergency information from social media can be used to ease responses during emergency events.

Although most of the studies related to social media and natural disasters only analyze the semantics and frequency of the messages shared within the social networking sites, some studies have incorporated a geographical component in their reasoning for Twitter activity. For example, Gao and Liu (2015) sought to better understand the dynamic changes in human behaviors corresponding to different event types by analyzing Twitter and Google Trends. Their study revealed that human concerns about events reflect the attraction of the event as well as subsequent discussion relating to the event. In other words, people will lose interest to certain event related information unless they are closely connected to the event geographically. This study focus is more on tweet counts and less on location; however, it does consider the role that a region plays when noting individual interest in event-related information and therefore can be a good representation of how relatedness via distance to an event influences Twitter activity.

In addition, Wang and Tayler (2016) also use information diffusion theory in their study relating the influence of fifteen separate natural disasters on human mobility patterns before, during, and after each event using individual movement data from Twitter. Results showed that natural disaster influence human movements in urban areas and that human movements are governed by power laws; however, the impact varies based on the severity and length of each event causing more erratic behavior for extreme events and losing the correlation observed in the study. This study analyzed information diffusion in Twitter during the three main phases of an event and took the geographic context of the information to explore human mobility pattern to better understand how this information can be used during a natural disaster to aid contingency planning for policymakers.

Lastly, another example where a geographic component is inconspicuously incorporated is Lu and Brelsford’s (2014) study which explored how patterns of social media interactions are influenced by extreme events. This study analyzed Twitter data generated before and after the 2011 earthquake and tsunami in Japan. Results showed that all Twitter users increased their online activity after the event, however only Japanese speakers expanded their social network to a much higher degree as opposed to English speakers. It was also noted that Japanese speakers were only sharing earthquake related content. This study also provides insight into how a social media network responds to stress and can provide aid in disaster preparation, warning, and recovery.

3. Methodology

3.1 Dataset and pre-processing

This research utilized Twitter data, an online social networking media which allows users to share messages limited to 140 characters called “Tweets”, to extract data generated during the Hurricane Harvey time frame, specifically, 18 August 2017 through 1 September 2017 (one week before to one week after Hurricane Harvey made landfall in Texas). Although the data only comprises of 1% of the total Tweets [1]generated during the specified time frame, it should still provide a valuable sample pool for this study with more than 140 million active users, publishing over 400 million Tweets a day (Kumar, Morstatter, and Liu 2014). The relevant information gathered and extracted via the Twitter API include: time stamp which includes date and time (created_at); message with hashtags (text); individual user ID (screen_name); option for GPS enable tracking (geo_enable); country origin of user account (country_code & country); longitude and latitude (coordinates); and city and state information (location). For this study, all the tweets relating to Hurricane Harvey in the United States mainland[2] were analyzed; pgAdmin 4, a database management software, was used to query and extract Hurricane Harvey related Tweets.

This study examined Hurricane Harvey Tweets at two different spatial scales: US States and Texas Counties to demonstrate the interest levels at both a local and state governmental level. Initially this study was to include the counties from neighboring Texas States as well (New Mexico, Oklahoma, Arkansas, and Louisiana); however, Hurricane Harvey related Tweets at the county level for these States were too dispersed and would have caused inaccurate results, therefore only the Texas Counties were modeled.

First step in data pre-processing was to eliminate all Tweets without location information and all Tweets that fell outside of the United States mainland. There were a total of 11,558,601 tweets in the dataset[1], of which 1,879 were omitted due to the lack of spatial information from the user disabling the geo_enable function. Next, all records outside of the

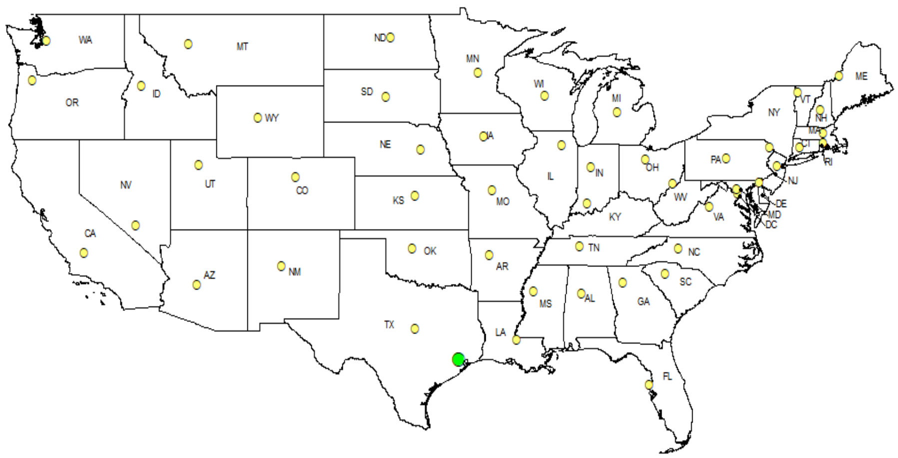
[1] The Twitter API returns all matching documents up to a volume equal to the streaming cap, which is currently set to 1% of the total current volume of Tweets published on Twitter (Kumar, Morstatter, and Liu 2014).
[2] This study will analyze all collected Tweets in the 48 states in the United States mainland in addition to Washington, DC., Alaska, Hawaii, the US territories and minor islands have been omitted due to lack of Twitter activity.

United States (our study area) were removed; there were a total of 10,834,742 tweets having ‘US’ as their county code. A total of 621,218 were removed from the dataset because of different country codes. In addition, there were 783 tweets unaccounted for due to the lack of country code, however those tweets had coordinates as part of their attribute information and were left in the dataset to be analyzed during the mapping stage. In sum, 10,835,462 of the Tweets generated during the specified timeframe were analyzed in this study (Table 1). Next was the mapping phase which was composed of importing the selected Tweets into ArcMap, a geospatial processing program used primarily to view, edit, create and analyze geospatial data; all Tweets were mapped via their coordinate attribute information. Next, the attributes were populated for the unaccounted Tweets that did not provide a country code or state attribute information. Of the 783 Null value records, 63 were removed from the pre-processed dataset because the Tweet coordinates fell outside of the Unites States mainland. After all attributes were populated a Mean Center analysis was applied. For states, the mean center was calculated based on the state value in the attribute table, whereas for the Texas counties the mean center was calculated based on a select by location query using the US Census county data and the state value in the attribute table. This analysis in turn produced a point layer for the Tweet geographic centers of each location in the study (Fig. 1). The point layer was then used to calculate the distance variable for the gravity model between each location and the Greater Houston area (Table 2 & Table 3).

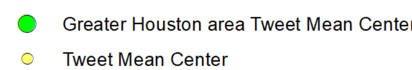
Table 1: Total Tweets in Study, Accountability of tweets during the Pre-processing stage, the 783 tweets that had Null values in the country code still had coordinate information and were processed during the mapping phase of the project.

Tweet Source	**Tweet Count**
Collected from Twitter API	11,558,601
Imported to pgAdmin	11,456,743
US Country Code	10,834,742
Other Country Code	621,218
Null value for Country Code	783
Pre-processing Dataset	10,835,525
Final Analysis Dataset	10,835,462

(a)



(b)



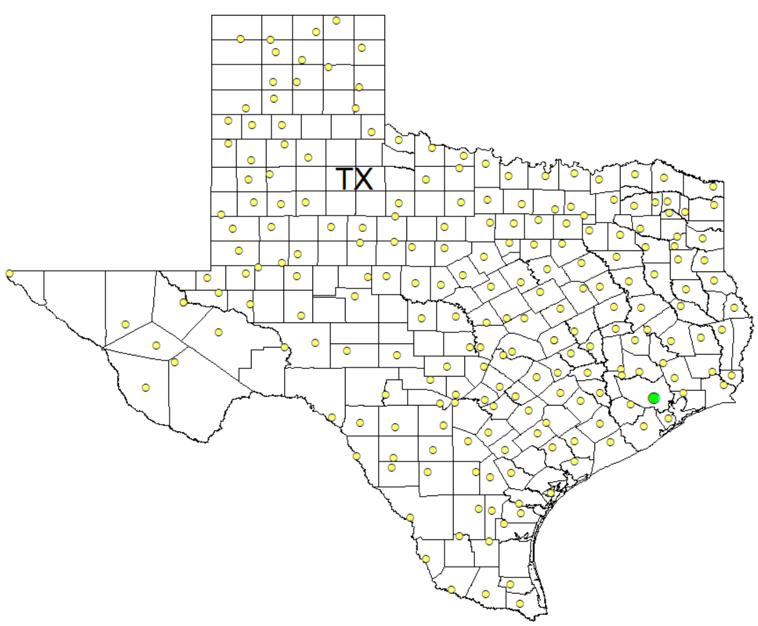


Fig. 1: Twitter messages mean center in (a) the United States mainland; (b) Texas counties (some counties in Texas did not have Twitter data).

Table 2: State level analysis. Distance variables calculated based on mean center in relation to the Greater Houston area.

State	Abbreviation	Distance (km)	State	Abbreviation	Distance (km)
Alabama	AL	856.76	North Carolina	NC	1,494.28
Arkansas	AR	673.75	North Dakota	ND	2,042.94
Arizona	AZ	1,674.56	Nebraska	NE	1,326.96
California	CA	2,362.17	New Hampshire	NH	2,588.01
Colorado	CO	1,489.04	New Jersey	NJ	2,221.40
Connecticut	CT	2,383.50	New Mexico	NM	1,230.15
Washington DC*	DC	1,946.34	Nevada	NV	2,112.00
Delaware	DE	2,086.86	New York	NY	2,241.38
Florida	FL	1,215.25	Ohio	OH	1,656.64
Georgia	GA	1,110.95	Oklahoma	OK	750.26
Iowa	IA	1,400.92	Oregon	OR	2,971.98
Idaho	ID	2,480.39	Pennsylvania	PA	2,000.95
Illinois	IL	1,466.80	Rhode Island	RI	2,507.15
Indiana	IN	1,426.70	South Carolina	SC	1,365.00
Kansas	KS	1,053.80	South Dakota	SD	1,688.70
Kentucky	KY	1,236.12	Tennessee	TN	1,014.38
Louisiana	LA	379.65	Texas	TX	335.46
Massachusetts	MA	2,533.78	Utah	UT	1,946.55
Maryland	MD	1,953.23	Virginia	VA	1,768.07
Maine	ME	2,778.68	Vermont	VT	2,542.51
Michigan	MI	1,787.07	Washington*	WA	3,033.84
Minnesota	MN	1,788.34	Wisconsin	WI	1,712.35
Missouri	MO	1,075.48	West Virginia	WV	1,673.93
Mississippi	MS	627.18	Wyoming	WY	1,898.65
Montana	MT	2,476.66			

Table 3: County level analysis. Note: The counties in this list all had Hurricane Harvey related Tweet messages, there were 40 counties in Texas that did not have Twitter data and 101 counties that did not have Hurricane Harvey related Tweets.

County	Distance (km)	County	Distance (km)	County	Distance (km)	County	Distance (km)	County	Distance (km)
Anderson	233.51	Cooke	472.49	Hays	264.27	Lubbock	775.23	Tarrant	399.85
Angelina	187.10	Coryell	294.06	Hemphill	838.60	Matagorda	112.62	Taylor	532.60
Aransas	255.42	Dallas	371.99	Henderson	284.12	Maverick	522.37	Tom Green	542.10
Atascosa	329.26	Dawson	721.27	Hidalgo	479.70	McLennan	283.00	Travis	269.69
Austin	93.26	Denton	406.54	Hood	390.41	Medina	364.11	Upshur	319.26
Bastrop	203.51	Dimmit	468.94	Hopkins	378.34	Midland	712.49	Uvalde	445.45
Bee	285.03	Duval	371.08	Houston	178.40	Montgomery	58.79	Victoria	197.80
Bell	269.53	Ector	727.45	Hunt	394.78	Nacogdoches	220.32	Walker	118.00
Bexar	338.60	El Paso	1,109.02	Hutchinson	880.32	Navarro	291.69	Waller	70.84
Bowie	427.49	Ellis	340.76	Jefferson	120.37	Nolan	579.30	Washington	124.07
Brazoria	56.65	Erath	397.84	Jim Hogg	430.70	Nueces	319.72	Webb	480.25
Brazos	148.80	Fayette	159.62	Jim Wells	354.19	Orange	138.89	Wharton	104.68
Brewster	815.08	Fort Bend	41.98	Johnson	361.65	Panola	283.68	Wichita	571.16
Brown	422.33	Freestone	236.88	Karnes	273.66	Parker	416.21	Williamson	257.09
Burleson	166.58	Frio	391.10	Kaufman	347.18	Polk	116.24	Wise	440.72
Burnet	311.07	Galveston	44.83	Kendall	338.14	Potter	882.00	Zapata	504.05
Caldwell	235.83	Gillespie	356.87	Kerr	382.42	Reeves	814.49	Zavala	460.53
Calhoun	182.65	Gonzales	215.28	Kleberg	355.66	Robertson	188.01		
Cameron	462.08	Grayson	461.40	Lampasas	320.09	Rockwall	376.68		
Chambers	50.30	Guadalupe	274.51	Lavaca	190.37	Rusk	275.24		
Cherokee	243.51	Hale	789.09	Liberty	53.32	San Jacinto	71.25		
Childress	701.45	Hardin	111.24	Limestone	245.91	San Patricio	309.82		
Collin	400.08	Harris	0.00	Live Oak	320.33	San Saba	370.44		
Comal	288.67	Harrison	324.16	Llano	327.41	Smith	296.24		

Once the Tweet records were organized and all attributes accounted for, the messages were queried to extract Hurricane Harvey related context. PgAdmin 4, a database management software, was utilized for this phase of the study. A Structured Query Language (SQL) code was created using the LIKE operator in a WHERE clause statement to search for Hurricane Harvey related keywords and Twitter hashtags, a system used to categorize Tweets. During this phase, some keywords were extracted in Spanish because it was noted during the visual observation of Twitter messages that a sizable portion of Tweets were in Spanish (Table 4). After the Hurricane Harvey messages were queried and extracted, 10% of the messages per each state were quality checked to verify Hurricane Harvey related content.

Table 4: Keywords used to extract Hurricane Harvey related messages.

Hurricane and Region related	Specific Harvey Hashtags
Houston	#HurricaneHarvey
Hurricane	#HoustonStrong
Huracan *(Hurricane in Spanish)*	#Harvey2017
Flood	#TexasStrong
Inundacion *(Flood in Spanish)*	#PrayingForTexas
Category 4	#HurricaneHarvery
Gulf Coast	

The final step in data pre-processing was to normalize the conceptual values that were inputted in the gravity model (Table 5 & Table 6). The first step of this process was to record the total number of Hurricane Harvey related Tweets for each state and Texas county. Next, the percentage of Hurricane Harvey tweets was divided by the total population of each state and Texas county, based on 2010 Census Bureau data; this computation generated the normalized tweet value that was inserted in the gravity model (Eq. 2).

$$Normalized\ Value = \left(\frac{Hurricane\ Harvey\ Tweets}{2010\ State\ Census\ Population}\right) \tag{2}$$

Table 5: Normalized value process for US States. 2010 Census population information derived from the United States Census Bureau. Total tweet count is based on PgAdmin query results.

State	Abbreviation	2010 Census Total Population	Total Tweets	Hurricane Harvey Tweets	Percentage	Normalized Value
Alabama	AL	4,779,736	141,059	470	0.33%	0.00010
Arizona	AZ	6,392,017	235,084	762	0.32%	0.00012
Arkansas	AR	2,915,918	59,305	264	0.45%	0.00009
California	CA	37,253,956	1,624,495	1,175	0.07%	0.00003
Colorado	CO	5,029,196	141,743	637	0.45%	0.00013
Connecticut	CT	3,574,097	101,555	318	0.31%	0.00009
Delaware	DE	897,934	31,012	72	0.23%	0.00008
Florida	FL	18,801,310	716,638	3,136	0.44%	0.00017
Georgia	GA	9,687,653	384,825	1,389	0.36%	0.00014
Idaho	ID	1,567,582	26,295	65	0.25%	0.00004
Illinois	IL	12,830,632	359,911	1,166	0.32%	0.00009
Indiana	IN	6,483,802	200,379	550	0.27%	0.00008
Iowa	IA	3,046,355	69,580	193	0.28%	0.00006
Kansas	KS	2,853,118	77,833	218	0.28%	0.00008
Kentucky	KY	4,339,367	111,452	322	0.29%	0.00007
Louisiana	LA	4,533,372	239,069	1,030	0.43%	0.00023
Maine	ME	1,328,361	22,416	125	0.56%	0.00009
Maryland	MD	5,773,552	231,085	728	0.32%	0.00013
Massachusetts	MA	6,547,629	199,301	721	0.36%	0.00011
Michigan	MI	9,883,640	251,746	613	0.24%	0.00006
Minnesota	MN	5,303,925	110,737	336	0.30%	0.00006
Mississippi	MS	2,967,297	80,418	316	0.39%	0.00011
Missouri	MO	5,988,927	141,382	578	0.41%	0.00010
Montana	MT	989,415	14,462	58	0.40%	0.00006

Table 5: Continued.

State	Abbreviation	2010 Census Total Population	Total Tweets	Hurricane Harvey Tweets	Percentage	Normalized Value
Nebraska	NE	1,826,341	49,571	145	0.29%	0.00008
Nevada	NV	2,700,551	170,215	411	0.24%	0.00015
New Hampshire	NH	1,316,470	26,957	86	0.32%	0.00007
New Jersey	NJ	8,791,894	271,734	758	0.28%	0.00009
New Mexico	NM	2,059,179	44,616	178	0.40%	0.00009
New York	NY	19,378,102	753,088	2,327	0.31%	0.00012
North Carolina	NC	9,535,483	296,168	1,188	0.40%	0.00012
North Dakota	ND	672,591	12,154	36	0.30%	0.00005
Ohio	OH	11,536,504	423,115	939	0.22%	0.00008
Oklahoma	OK	3,751,351	114,906	513	0.45%	0.00014
Oregon	OR	3,831,074	127,655	395	0.31%	0.00010
Pennsylvania	PA	12,702,379	335,739	1,081	0.32%	0.00009
Rhode Island	RI	1,052,567	29,940	87	0.29%	0.00008
South Carolina	SC	4,625,364	133,266	496	0.37%	0.00011
South Dakota	SD	814,180	12,972	23	0.18%	0.00003
Tennessee	TN	6,346,105	199,941	860	0.43%	0.00014
Texas	TX	25,145,561	1,446,150	12,689	0.88%	0.00050
Utah	UT	2,763,885	66,152	265	0.40%	0.00010
Vermont	VT	625,741	12,241	79	0.65%	0.00013
Virginia	VA	8,001,024	283,795	1,067	0.38%	0.00013
Washington*	WA	6,724,540	208,445	588	0.28%	0.00009
West Virginia	WV	1,852,994	49,271	156	0.32%	0.00008
Wisconsin	WI	5,686,986	93,699	301	0.32%	0.00005
Wyoming	WY	563,626	10,394	20	0.19%	0.00004
Washington DC*	DC	601,723	91,496	683	0.75%	0.00114

Table 6: Normalized values for Texas counties with Hurricane Harvey message content;

County	2010 Census Population	Total Tweets	Hurricane Harvey	Percentage	Normalized Value
Anderson	58,458	681	1	0.15%	0.00002
Angelina	86,771	1,752	16	0.91%	0.00018
Aransas	23,158	457	7	1.53%	0.00030
Atascosa	44,911	443	2	0.45%	0.00004
Austin	28,417	182	2	1.10%	0.00007
Bastrop	74,171	616	16	2.60%	0.00022
Bee	31,861	698	2	0.29%	0.00006
Bell	310,235	5,414	37	0.68%	0.00012
Bexar	1,714,773	121,424	886	0.73%	0.00052
Bowie	92,565	1,621	3	0.19%	0.00003
Brazoria	313,166	6,379	59	0.92%	0.00019
Brazos	194,851	19,956	199	1.00%	0.00102
Brewster	9,232	374	2	0.53%	0.00022
Brown	38,106	817	2	0.24%	0.00005
Burleson	17,187	120	2	1.67%	0.00012
Burnet	42,750	261	4	1.53%	0.00009
Caldwell	38,066	748	5	0.67%	0.00013
Calhoun	21,381	162	2	1.23%	0.00009
Cameron	406,220	12,041	79	0.66%	0.00019
Chambers	35,096	706	8	1.13%	0.00023
Cherokee	50,845	247	2	0.81%	0.00004
Childress	7,041	61	1	1.64%	0.00014
Collin	782,341	21,332	82	0.38%	0.00010
Comal	108,472	5,912	53	0.90%	0.00049
Cooke	38,437	289	2	0.69%	0.00005
Coryell	75,388	8,306	32	0.39%	0.00042
Dallas	2,368,139	42,237	108	0.26%	0.00005
Dawson	13,833	85	1	1.18%	0.00007
Denton	662,614	123,193	737	0.60%	0.00111
Dimmit	9,996	295	2	0.68%	0.00020
Duval	11,782	464	6	1.29%	0.00051
Ector	137,130	5,347	10	0.19%	0.00007
El Paso	800,647	24,044	37	0.15%	0.00005
Ellis	149,610	3,899	14	0.36%	0.00009
Erath	37,890	3,590	13	0.36%	0.00034
Fayette	24,554	268	4	1.49%	0.00016
Fort Bend	585,375	36,704	407	1.11%	0.00070
Freestone	19,816	951	8	0.84%	0.00040

Table 6: Continued;

County	2010 Census Population	Total Tweets	Hurricane Harvey	Percentage	Normalized Value
Frio	17,217	719	7	0.97%	0.00041
Galveston	291,309	7,831	146	1.86%	0.00050
Gillespie	24,837	197	1	0.51%	0.00004
Gonzales	19,807	471	5	1.06%	0.00025
Grayson	120,877	1,832	11	0.60%	0.00009
Guadalupe	131,533	2,873	7	0.24%	0.00005
Hale	36,273	1,235	18	1.46%	0.00050
Hardin	54,635	13,614	57	0.42%	0.00104
Harris	4,092,459	354,505	5160	1.46%	0.00126
Harrison	65,631	2,417	2	0.08%	0.00003
Hays	157,107	23,656	149	0.63%	0.00095
Hemphill	3,807	19	2	10.53%	0.00053
Henderson	78,532	864	6	0.69%	0.00008
Hidalgo	774,769	29,224	143	0.49%	0.00018
Hood	51,182	1,078	7	0.65%	0.00014
Hopkins	35,161	681	8	1.17%	0.00023
Houston	23,732	219	4	1.83%	0.00017
Hunt	86,129	4,862	1	0.02%	0.00001
Hutchinson	22,150	300	2	0.67%	0.00009
Jefferson	252,273	9,507	43	0.45%	0.00017
Jim Hogg	5,300	633	8	1.26%	0.00151
Jim Wells	40,838	2,200	7	0.32%	0.00017
Johnson	150,934	1,135	6	0.53%	0.00004
Karnes	14,824	759	2	0.26%	0.00013
Kaufman	103,350	2,308	7	0.30%	0.00007
Kendall	33,410	2,259	10	0.44%	0.00030
Kerr	49,625	866	4	0.46%	0.00008
Kleberg	32,061	3,442	33	0.96%	0.00103
Lampasas	19,677	148	2	1.35%	0.00010
Lavaca	19,263	202	2	0.99%	0.00010
Liberty	75,643	968	8	0.83%	0.00011
Limestone	23,384	955	4	0.42%	0.00017
Live Oak	11,531	45	2	4.44%	0.00017
Llano	19,301	159	2	1.26%	0.00010
Lubbock	278,831	15,937	37	0.23%	0.00013
Matagorda	36,702	503	5	0.99%	0.00014
Maverick	54,258	1,110	2	0.18%	0.00004
McLennan	234,906	8,038	42	0.52%	0.00018

Table 6: Continued.

County	2010 Census Population	Total Tweets	Hurricane Harvey	Percentage	Normalized Value
Medina	46,006	793	6	0.76%	0.00013
Midland	136,872	3,495	11	0.31%	0.00008
Montgomery	455,746	21,325	292	1.37%	0.00064
Nacogdoches	64,524	5,675	15	0.26%	0.00023
Navarro	47,735	1,495	6	0.40%	0.00013
Nolan	15,216	239	3	1.26%	0.00020
Nueces	340,223	993	25	2.52%	0.00007
Orange	81,837	4,067	14	0.34%	0.00017
Panola	23,796	308	3	0.97%	0.00013
Parker	116,927	2,579	6	0.23%	0.00005
Polk	45,413	544	3	0.55%	0.00007
Potter	121,073	3,678	14	0.38%	0.00012
Reeves	13,783	361	7	1.94%	0.00051
Robertson	16,622	293	10	3.41%	0.00060
Rockwall	78,337	2,195	3	0.14%	0.00004
Rusk	53,330	204	4	1.96%	0.00008
San Jacinto	26,384	359	2	0.56%	0.00008
San Patricio	64,804	17,930	362	2.02%	0.00559
San Saba	6,131	23	2	8.70%	0.00033
Smith	209,714	10,226	33	0.32%	0.00016
Tarrant	1,809,034	118,653	389	0.33%	0.00022
Taylor	131,506	4,197	29	0.69%	0.00022
Tom Green	110,224	4,134	10	0.24%	0.00009
Travis	1,024,266	69,261	879	1.27%	0.00086
Upshur	39,309	2,948	17	0.58%	0.00043
Uvalde	26,405	879	4	0.46%	0.00015
Victoria	86,793	2,467	72	2.92%	0.00083
Walker	67,861	7,997	60	0.75%	0.00088
Waller	43,205	9,996	86	0.86%	0.00199
Washington	33,718	1,181	10	0.85%	0.00030
Webb	250,304	4,917	29	0.59%	0.00012
Wharton	41,280	1,286	21	1.63%	0.00051
Wichita	131,500	4,920	19	0.39%	0.00014
Williamson	422,679	15,886	135	0.85%	0.00032
Wise	59,127	453	4	0.88%	0.00007
Zapata	14,018	158	4	2.53%	0.00029
Zavala	11,677	286	2	0.70%	0.00017

3.2 Analysis Framework

As stated in Section 2, the gravity model was used in this study to measure the relatedness/distance decay of human interest (normalized Tweet values) based on distance from each location (mean center of Tweets) to the affected Hurricane Harvey location (the mean center of Tweets in the Greater Houston area, Texas). This model was chosen due to "its effectiveness in measuring the degree of relatedness" (Yuan 2017). The objective of this study was to detect how distance to an event correlates to Tweet responses relating to said event. Since this study focuses on the spatial interaction of the Greater Houston area, P_i in Eq 1 is viewed as a constant so that the fitted model is modified as:

$$I_{ij} = K \frac{P_j}{D_{ij}^{\beta_1}} \quad (3)$$

where P_j is the conceptual sizes (the total number of Tweets for each location, state and county), D_{ij} represents the distance separating the geographic mean centers of i and j, and I_{ij} denotes the interaction/connection between i and j (the normalized value for Hurricane Harvey Tweets). K is still the same constant as Eq 1 and β is the distance friction coefficient that investigates the role of distance (Yuan 2017). The specific parameters are:

I_{ij} The normalized value of Hurricane Harvey Tweets in each location

P_j The total number of Tweets in each location (state and county)

D_{ij} The distance separating each location (mean center of Tweets) to the Greater Houston area (mean center of Tweets in the Greater Houston area)

The gravity model addresses the question of how distance decay plays a role in human interest to Hurricane Harvey, the natural disaster in question for this study (i.e. do locations "closest" to the event generate more Tweets relating to said event?). A comparative analysis was conducted at the Texas county extent and US mainland extent to demonstrate varying degrees of distance decay; the gravity model is used to demonstrate the best fit distance friction coefficient β for the comparative analysis.

3.3 Model Parameter Fitting

The best fit coefficients for β was calculated based on the evaluation of the goodness of fit (R^2) of the observed values I_{ij}. This study experimented with various β values and the β value that reached the highest R^2 was considered the best fit. Past studies relating human mobility patterns and transportation demonstrate that higher β values indicate a strong distance decay effect (Yuan, Lui, and Wei 2017). The fitted model results are indicated in Section 4.1.

4. Analysis Results and Discussion

4.1 Comparative Analysis via the Gravity Model Fitting

As discussed in Section 3.2, the gravity model was used to calculate the best-fit β value for a state level analysis and a county level analysis. Figure 2 displays the correlation between the goodness of fit (R^2) and the fit β value for the comparative analysis.

The analysis demonstrates different results. The Texas County level analysis displays a weak distance decay (β= 0.675, R^2= 0.2418) compared to the States level analysis which demonstrated a much stronger distance decay. The States level analysis was modeled twice, first including the State of Texas (β= 2.275, R^2= 0.3449) and again without the State of Texas (β= 3.095, R^2= 0.108); the model was ran twice because in the initial run, which included the State of Texas, Texas was deemed as an outlier (Fig. 3a). Figure 3 plots the estimated and observed co-occurrences with the best fit β value. As Figure 3 demonstrates both models depict DC as an outlier, presenting the predicted values significantly lower than the observed values. This observation is potentially due to the fact Washington D.C. is an administrative unit and is therefore not well represented by the gravity model shown in this study. Washington D.C. was included in the study because of the quantity of Tweets present in the dataset. The results in this study indicate that the gravity model is a good fit for demonstrating human interest to a specific event while also relating to TFL since locations closest to the Greater Houston area generated more Twitter messages relating to Hurricane Harvey ("near things are more related than distant things"). One possible explanation for the weaker distance decay in the Texas county data is that Hurricane Harvey caused statewide impact, so the concerns for this catastrophe does not strictly follow classic distance decay within Texas. On the other hand, the FCC reports that 39% of residents in rural America lack internet broadband access (Strover 2018), and the Twitter data displayed more than half of the Texas counties either had no Twitter data or did not return any Hurricane Harvey related Tweet messages. This lack of data also influenced the result of our modeling calculation. Section 4.2 discusses further on data uncertainties present in this study.

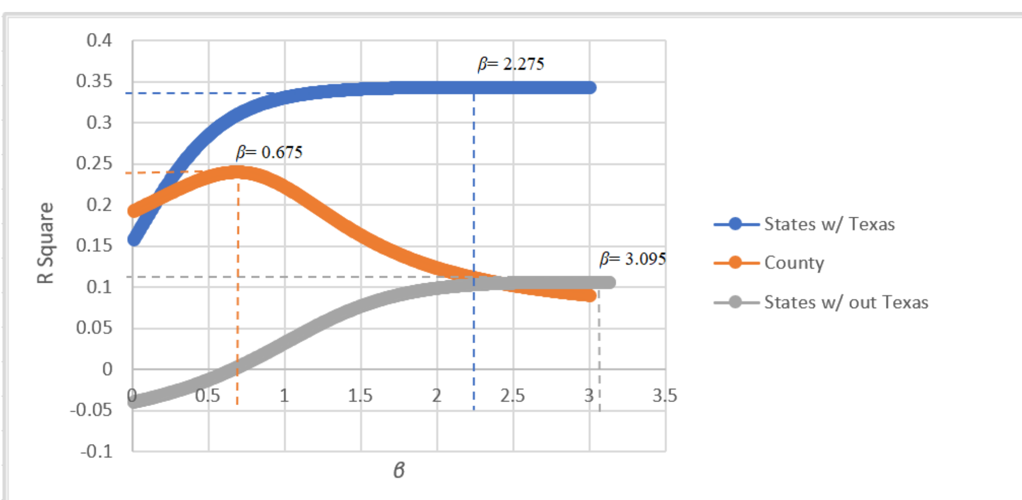
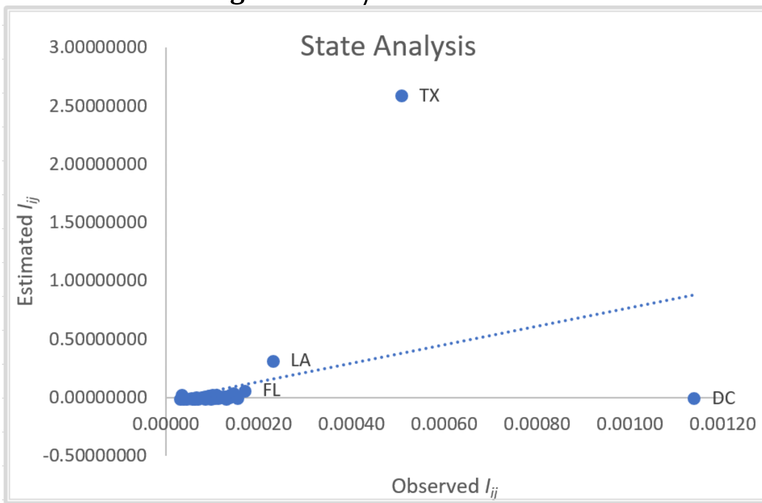
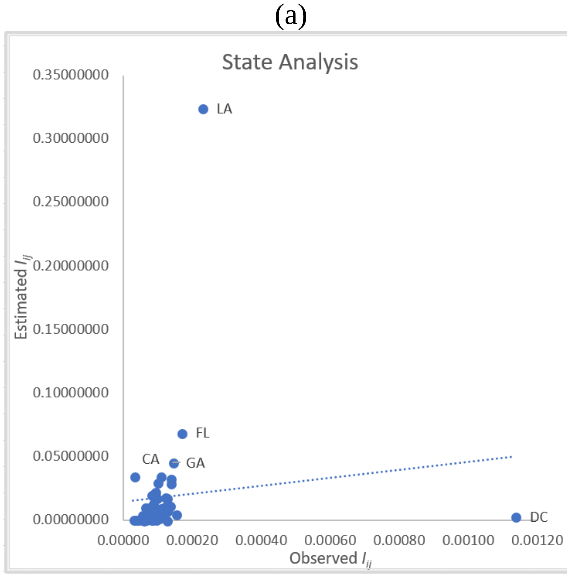


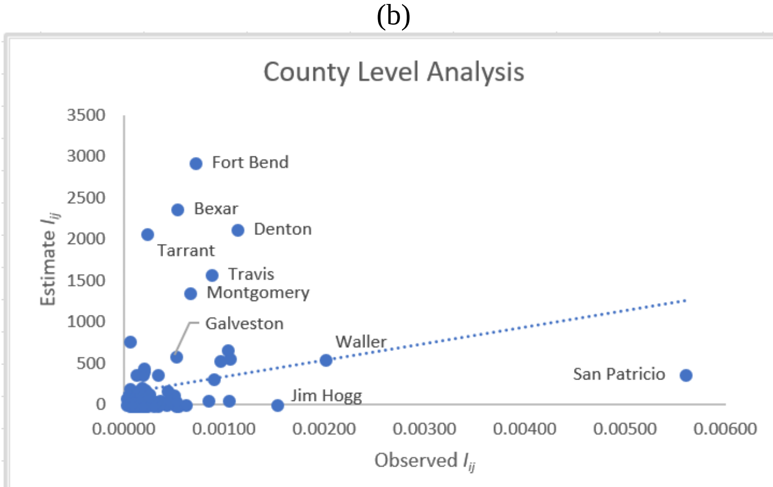
Fig. 2: Fitted β values and R^2.



(a)



(b)



(c)

Fig. 3: Correlation between the estimated and observed I_{ij} (a) States including Texas; (b) States excluding Texas; (c) Counties.

4.2 Discussion of Uncertainty

This section discusses data uncertainties that had potential to arise throughout the course of this study. Some uncertainties that are to be expected include but are not limited to:

- *Inaccuracy/imprecision with the limitation of data*: The strength and availability of global positioning system (GPS) varies resulting in inaccurate positional sampling which contribute to data uncertainty. For example, the 63 records omitted from this study were positioned in the middle of the Pacific Ocean.
- *Textual errors in Twitter messages*: Human typographical errors can inhibit the query results for message content (Harvey related tweets) or state information (not including or misspelling a state name) producing false model outputs. For example, Twitter noted a number of users misspelled the "HurricaneHarvey" hashtag as "HurricaneHarvery" and it produced a widespread sharing of the false hashtag, this hashtag was included in the SQL query code (Table 3).
- *Credibility of Twitter data*: Social media data contains large amount of spam information which can affect the credibility of the data.
- *Limitations of available data:* Twitter is widely used among the Unites States, however the application relies heavily on internet access. Some US locations have limited internet access, such as the rural counties in Texas; this limit on data can potentially have negative influences on the results of this study.
- *Imperfection of models and algorithms:* The results of this study are highly dependent of the gravity model, however the use of a different model or different

data may hinder dissimilar results. Future research should be adopted to validate the findings in this study.

5. Conclusion

This study proposes a method for discovering patterns that can provide viable insights into different research fields, specifically emergency management agendas. By utilizing Twitter data generated before, during, and after Hurricane Harvey, this study was able to explore the distance decay of human interest to a specific event, i.e. Hurricane Harvey which greatly affected the Greater Houston area. Results indicate that distance plays a weak role in the Texas County level analysis when compared to the State level analysis demonstrating that Texas State residents reacted strongly to the disaster even if they were not directly impacted by the storm. Meanwhile, the States level analysis revealed that distance displays a much stronger role. States furthest from Texas were less likely to generate Hurricane Harvey related Tweets. Results also indicate TFL can be used to guide the examination of human interests to specific events: places further away are less related (less interested) versus places closer which are more related and display more interest to said event.

Since this study proposes new theoretical foundation for exploring human interest to a specific event, in the case of this study the specific event is the natural disaster that was Hurricane Harvey. The results of this study will provide viable knowledge to emergency management officials that can and should be used for future contingency planning for disaster management agendas. Some emergency management examples where understanding the dynamics of human interests reflected in social media to natural disasters can provide insight to are: recognizing areas of high concern based on Tweet density, preparing social media response protocols to be used during a disaster, delivering information to areas where there is lack of journalistic coverage relating the disaster, and maximizing fundraising profits through the use of social media. In addition, this study opens a window of opportunity to new research directions which can further enhance emergency management agendas.

Future research directions include but are not limited to incorporating a temporal scale to this study to demonstrate how time affects Twitter responses to said event and investigating the cause of distance decay. A similar study was conducted by Teng (2017) where Tweets related to the Dallas mass shooting in 2016 were analyzed, results determined distance played a role on the impact of information diffusion, but the influence decreased as time passed. This study could be replicated to analyze other major events and validate results of the presented study. This study should also be replicated using other social media data to investigate if results are similar across social media platforms.

6. References

Afiune, G. 2017. State says Harvey's death toll has reached 88. The Texas Tribune. https://www.texastribune.org/2017/10/13/harveys-death-toll-reaches-93-people/.

Bergstrand, J. H. 1985. The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence. The Review of Economics and Statistics 67 (3):474–481.

Billion-Dollar Weather and Climate Disasters: Overview. National Climatic Data Center. https://www.ncdc.noaa.gov/billions/ (last accessed 2018).

Chen, Y. 2015. The distance-decay function of geographical gravity model: Power law or exponential law? Chaos, Solitons & Fractals 77:174–189.

Ciaccia, C., and Fox News. 2017. Tropical Storm Harvey: Is Twitter becoming the new 911? Fox News. http://www.foxnews.com/tech/2017/08/28/topical-storm-harvey-is-twitter-becoming-new-911.html.

Click here to support Houston Flood Relief Fund. YouCaring. https://www.youcaring.com/victimsofhurricaneharvey-915053 (last accessed 2018).

Dong, R., L. Li, Q. Zhang, and G. Cai. 2018. Information Diffusion on Social Media During Natural Disasters. IEEE Transactions on Computational Social Systems 5 (1):265–276.

Gao, C., and J. Liu. 2015. Uncovering Spatiotemporal Characteristics of Human Online Behaviors during Extreme Events. Plos One 10 (10):1–14.

Hardy, D., J. Frew, and M. F. Goodchild. 2012. Volunteered geographic information production as a spatial process. International Journal of Geographical Information Science 26 (7):1191–1212.

Hurricane Aftermath: Franchising Sees a Surge in Kindness and Business. Franchising USA. https://franchisingusamagazine.com/hurricane-aftermath-franchising-sees-surge-kindness-and-business (last accessed 2018).

Kang, C., Y. Liu, D. Guo, and K. Qin. 2015. A Generalized Radiation Model for Human Mobility: Spatial Scale, Searching Direction and Trip Constraint. Plos One 10 (11):1–11.

Kluger, J. 2017. Houston's Challenge to Rebuild Itself After Hurricane Harvey. Time. http://time.com/4931061/houston-after-harvey/.

Kocaslan, G. 2017. The Role of Distance in the Gravity Model: From the View of Newton, International Economics and Quantum Mechanics. NeuroQuantology 15 (2):208–214.

Kumar, S., F. Morstatter, and H. Liu. 2014. Twitter Data Analytics. New York, NY: Springer New York.

Li, X., Z. Wang, C. Gao, and L. Shi. 2017. Reasoning human emotional responses from large-scale social and public media. Applied Mathematics and Computation 310:182–193.

Lindsay, B. R. 2011. Social media and disasters: current uses, future options, and policy considerations. Washington, D.C.: Congressional Research Service.

Liu, Y., F. Wang, C. Kang, Y. Gao, and Y. Lu. 2013. Analyzing Relatedness by Toponym Co-Occurrences on Web Pages. Transactions in GIS 18 (1):89–107.

Liu, Y., Z. Sui, C. Kang, and Y. Gao. 2014. Uncovering Patterns of Inter-Urban Trip and Spatial Interaction from Social Media Check-In Data. PLoS ONE 9 (1):1–11.

Lu, X., and C. Brelsford. 2014. Network Structure and Community Evolution on Twitter: Human Behavior Change in Response to the 2011 Japanese Earthquake and Tsunami. Scientific Reports 4 (1):1–11.

Miller, H. J. 2004. Toblers First Law and Spatial Analysis. Annals of the Association of American Geographers 94 (2):284–289.

New Media. 2009. US Census Bureau 2010 Census. Visit Census.gov. https://www.census.gov/2010census/ (last accessed 2018).

Pappalardo, L., S. Rinzivillo, and F. Simini. 2016. Human Mobility Modelling: Exploration and Preferential Return Meet the Gravity Model. Procedia Computer Science 83:934–939.

Perryman, R. M. 2017. Weathering the Storm: The Impact of Hurricane Harvey on Texas. Perryman Report & Texas 34 (7):1–8.

Rhodan, Maya. 2017. 'Please Send Help.' Hurricane Harvey Victims Turn to Twitter and Facebook. Time. http:// http://time.com/4921961/hurricane-harvey-twitter-facebook-social-media/.

Shi, L., G. Chi, X. Liu, and Y. Liu. 2015. Human mobility patterns in different communities: a mobile phone data-based social network approach. Annals of GIS 21 (1):15–26.

Strover, S. 2018. Reaching rural America with broadband internet service. The Conversation. http://theconversation.com/reaching-rural-america-with-broadband-internet-service82488.

Teng, Y. 2017. Diffusion of Twitter Messages on Dallas Mass Shooting: Patterns and Factors.

Tobler, W. R. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography 46:234.

US Department of Commerce, and NOAA. 2018. Hurricane Harvey Info. National Weather Service. https://www.weather.gov/hgx/hurricaneharvey (last accessed 2018).

US Department of Commerce, and NOAA. 2018. Major Hurricane Harvey - August 25-29, 2017. National Weather Service. https://www.weather.gov/crp/hurricane_harvey (last accessed 2018).

Wang, Q., and J. E. Taylor. 2016. Patterns and Limitations of Urban Human Mobility Resilience under the Influence of Multiple Types of Natural Disaster. Plos One 11 (1):1–14.

Wang, Y., T. Wang, X. Ye, J. Zhu, and J. Lee. 2015. Using Social Media for Emergency Response and Urban Sustainability: A Case Study of the 2012 Beijing Rainstorm. Sustainability 8 (1):1–17.

Wei, J., and D. Marinova. 2015. The orientation of disaster donations: differences in the global response to five major earthquakes. Disasters 40 (3):452–475.

Xiao, Y., Q. Huang, and K. Wu. 2015. Understanding social media data for disaster management. Natural Hazards 79 (3):1663–1679.

Yuan, Y., and M. Medel. 2016. Characterizing International Travel Behavior from Geotagged Photos: A Case Study of Flickr. Plos One 11 (5):1–18.

Yuan, Y. 2017. Exploring the Spatial Decay Effect in Mass Media and Location-Based Social Media: A Case Study of China. Advances in Geocomputation Advances in Geographic Information Science :133–142.

Yuan, Y., Y. Liu, and G. Wei. 2017. Exploring inter-country connection in mass media: A case study of China. Computers, Environment and Urban Systems 62:86–96.