# Hurricane Irma's Effect on Human Mobility

**Saurabh Thakrani**
University of Maryland, College Park
sthakran@umd.edu

**Yashika Bajaj**
University of Maryland, College Park
ybajaj@umd.edu

**Abstract**

Natural disasters are inevitable and may have a great impact on the local population. People tend to protect themselves during such situations and are likely to move from the disaster struck place to a safer place. With the use of technology and smartphones, it has become possible to get the digital footprint of a person via geo-tagged tweets. Our study focuses on analyzing the mobility pattern during hurricane Irma using the geo-tagged tweets and to find how people have moved in different counties in the state of Florida during the hurricane. This paper includes the related work which has been done for the analysis of human mobility as well as the study conducted by us to analyze the mobility pattern.

## Introduction

Tropical cyclones represent one of the costliest and devastating threats for populations in both the developed and developing world, with long-lasting consequences that extend over several years. The overall profile for natural hazard losses in the U.S. since 1960 finds that tropical cyclones represent roughly 26% of the total losses. While the projections on frequency and intensity of tropical cyclones remain inconclusive, there is increasing risk exposure as population and assets continue to shift to coastal areas. During tropical cyclones, evacuations are primarily ordered for those who live on the coast or in adjacent low-lying areas due to the anticipated storm surge. However, the compliance rates with evacuation orders are often significantly less than 100%, with many coastal residents preferring to ride out the storm rather than leave. The decisions to stay or go depend on several factors such as the perception of risk, prior experience, and the severity of the storm itself. Traditionally, evacuation rates were determined by post-evacuation questionnaire surveys months following the hurricane. Traffic counts have also been used, but these routinely underestimate the number of people evacuating. However, evaluating the tweets of the people in the affected areas can be a good option.

Twitter is the most popular micro-blogging service in the world. Millions of people use this online social network to connect socially with friends, family members and co-workers. They use it to inform others what they are doing, thinking or what is happening.

For our analysis we have filtered out the tweets for state of Florida and in particular the residents of Florida. Our emphasis is on the Hurricane Irma that was a category 5 Hurricane initially and it struck Florida as category 4 Hurricane on September 10. Our study focuses on unfolding the mobility patterns of the residents of Florida during this time. We wish to analyze the effectiveness of evacuation process by evaluating mobility patterns of the people by looking at the tweet locations in the pre, during and post hurricane period. We are using the geotagged tweets for this study.

### Related Work

The earlier work done in this field included hypothesis testing based out of assumptions seen on the patterns of mobility among the people during the natural disasters like hurricane, typhoons etc. There are also some studies about the mobility of people in the affected area during these natural disasters. Knowledge about the mobility patterns would help the government in sending relief to the locations in need as well as manage the traffic to avoid traffic jams. They would also help the government set up temporary communication centers so that the people can talk to their near and dear ones and tell about their well beings.

To know more about these patterns the mobility pattern during and after Hurricane Sandy—one of the largest tropical storms recorded in the Atlantic Ocean that affected millions of people — which struck the northeastern seaboard of the United States leading to significant injury and loss of human life in October 2012, was studied [1]. To check the travel trajectories a hypothesis testing was done. While the rest of the days had similar trajectories, there was a variation for the 1st day, along with that the short-distance trips increased and the long-distance trips decreased. Further the mobility patterns were being tested to see if they were chaotic during the hurricane or not? The center of mass and radius of gyration were taken into consideration and again 2 hypotheses were tested. The tests declared that people need to move to safer places during hurricane in order to sustain

Copyright © 2018, University of Maryland (www.umd.edu). All rights reserved.

their mobility. In yet another paper [2] the use of Twitter geotags to study human mobility has been presented and its efficiency over the call record data has been argued. There are also certain biases in this methodology like the potential sampling bias, communication modality and location biases for sending tweets. The tweets are also used as a means to study the mobility distances among the people. The parameter radius of gyration was used to study the geographical constraints in the movements. The probability density function was used to see if the frequent travelers among the cities are less directed and more diffused in their movements as well as to study the long-distance travelers. Another work [3] considered fifteen destructive cases across five types of natural disasters and analyzed the human movement data before, during, and after each event, comparing the perturbed and steady state movement data. The paper also talks about Lévy flight model and showed that individual movement trajectories exhibited similar shapes after being rescaled by the radius of gyration. There were multiple hypotheses which were tested in this paper - check if the power law governs human urban travels in multiple types of natural disaster, shifts in the distances of centers of movement during natural disasters are positively correlated with the values of the radius of gyration in steady states and values of radius of gyration during natural disasters are positively correlated with the ones in steady states. In another study [4], which dealt with analyzing human mobility across political borders, twitter data was used to track the movement of people from Kenya as the normal call and text services are constrained within the political boundaries. A series of temporal movement patterns were constructed using these geo-referenced tweets which included daily, monthly and total movements patterns by linking 24 hour, monthly and total tweets respectively. Estimates of spatial spread of movement was based on radius of gyration. Flow networks were constructed in which the nodes were represented by the centroids of the district level boundaries. The total no. of connections was used to investigate the centrality by using four measures: degree, betweenness, closeness and eigenvector. Another study [5] which aimed at analyzing social media data for disaster footprint and damage assessment used twitter data. In this study, machine learning technique i.e. Latent Dirichlet Allocation was used for semantic information extraction. This is an unsupervised, self-learning topic modelling approach. The semantic evaluation was performed by tokenization, removing stop words and handling synonyms. The data usage was restricted to a particular area as the semantic analysis would get affected by language variation. After applying the LDA modelling the spatial hot spot analysis was carried out by using the well-known Getis-Ord G method that detects local attribute clusters in the data. Another study [6] talks about using Twitter data to evaluate the evacuation process during the Hurricane Matthew. We have adopted somewhat similar approach but dividing the data into 3 parts according to dates to evaluate pre, during and post condition in the affected region. Also, [6] talks about the decrement in the number of tweets from the pre-evacuation process date to the post evacuation date which confirms our hypothesis for our paper. The article talks about using social media data to replace the traditional method of surveys that were being used to evaluate the evacuation process by the government. Further a county wise study was performed on the dataset to identify the state and county from which maximum tweets were done. Another study [7] focuses mostly on determining just the mobility pattern by comparing the geotagged tweets' location from the land use data of the cities considered. The study explores spatial and temporal patterns of these geotagged tweets. One more study [8] has revealed the importance of using GPS log data over traditional survey data to analyze the trajectories of human movement. The paper gives an idea to zoom into individuals to get the most out of data and see the trends. Another study [9] elaborates the use of social media data during floods to evaluate the impact on real-time basis. It talks about using the Facebook, Twitter, Flickr etc. data for disaster monitoring and by monitoring the paper points towards the physical characteristics for example the flood depth and location during floods. The results could be used in the first hours of an event to trigger action and allocate people and money for disaster response. Next [10] is a case study focusing on the remote sensing data that can be used to evaluate the disaster impact. The data can be compared to other satellite data etc. to compare the results. Another study [11] discusses the use of Twitter during disaster situations in particular, the modifications and the use in general that can prove to be effective during these cases. The paper has even justified the use of Twitter for such studies over Facebook. The paper has basically recommended a new version of Twitter exclusively for disasters. The other paper [12] compares the use of Twitter by the users during disaster and normal days. This paper focuses on the features of Twitter use in emergency and mass convergence situations and offers an examination of some Twitter-based behaviors. According to this paper the highest tweets are generated during the disaster events and the quantity of Twitter activity measured correlates to both size and significance of happenings. Also, the evaluation of tweets sensed the type of information that is being discussed in these tweets. Another study [13] is the paper with detail analysis of the social media platform used to enhance emergency situation awareness. According to [14], being able to infer the number of people in a specific area is of extreme importance for the avoidance of crowd disasters and to facilitate emergency evacuations. Another study [15] talks about the unevenness in the geodata. It suggests that well-educated people in the occupations of management, business, science, and arts are more likely to be involved in the generation of georeferenced tweets and photos.

## Research Questions

The previous study has shown that the Twitter geotagged data is very accurate in terms of location and hence predicting the mobility patterns. In our study we are considering the Twitter data which is a combination of two files, first the geo located tweets and second Twitter ids.

The idea is to study the mobility patterns before, during and after the hurricane Irma, and by that we mean to look into where people go, for how long and to evaluate: (1) how those behaviors relate to what is going on the ground (floods, destroyed homes, etc.) and (2) how those behaviors relate to the residents of Florida. We want to look into how the density of people changes in different counties of Florida with respect to the hurricane.

## Datasets

The dataset has been bought from Twitter consisting the geotagged tweets and the twitter ids of the individuals. The data consists of 485,675 tweets in total from the residents of Florida. We then extracted the user_ids, dates and coordinates from the given data to form a separate data frame. The coordinates were converted to separate latitude and longitude columns. In order to effectively analyze the pattern, we filtered out the tweets between the dates August 15, 2017 and September 30, 2017. Now, we wanted to see the number of tweets from each user_id during the time period we defined. We plotted a graph between the user_id and the number of tweets per user_id [Figure 1].

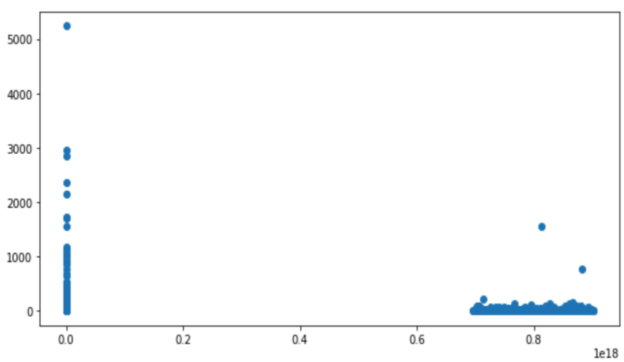


Fig. 1

It was visible that the maximum number of tweets from a user was more than 5000 which was abnormally huge. Also, there were other users who had tweeted in thousands in such a small period of time. Similarly, there were some users who had tweeted very less number of times. In order to conduct a good analysis, we filtered out the users with maximum of 600 tweets and minimum of 5 tweets per user_id [Figure 2]. This was done as a measure to eliminate fake twitter accounts as well as bots.

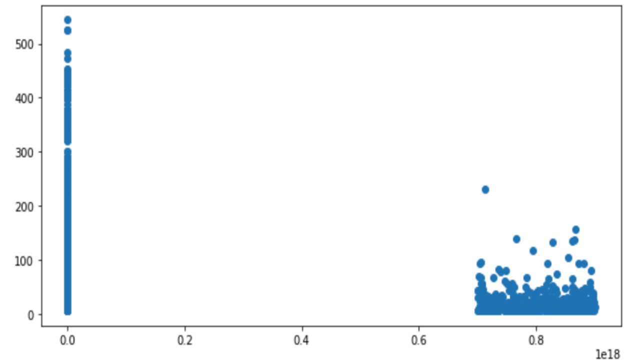


Fig. 2

The number of users decreased to 9290 unique users while the total number of tweets in the sample now were reduced to 219006.

## Methods

We are using the coordinates of these unique user_ids to evaluate their mobility. Basically, we are finding out the counties that these coordinates are a part of, and by that we intend to use them as a factor to analyze the mobility. If the county is same for all the three periods, chances are the users have not travelled to any new location. We want to use these locations of the users to evaluate their mobility. Now, to check this we will be comparing the counties obtained to two datasets found on FEMA website. These are the datasets of the information about the funds released for different counties and cities for the people to help them rebuild their houses etc. damaged in the hurricane.

Our idea is that the counties where these funds have been released were the ones affected by the hurricane and ideally the people should move away from these counties.

We will be making an origin-destination table that will show the density of users in each county before and after the hurricane. Converting this into a heat map will clearly show the information through the color density.

## Experimental Findings

We plotted the number of tweets on each day to have a close look on the pattern. It was seen that the number of tweets were more before the hurricane and it gradually decreased during and after the hurricane [Figure 3].

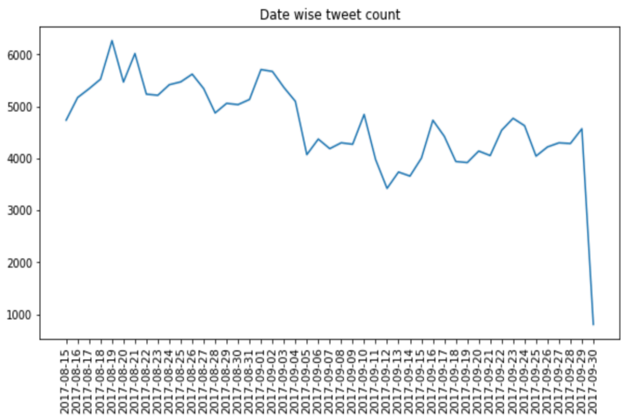


Fig. 3

Also, the maximum tweeting activity was observed between 7PM to 9PM as well as the maximum number of tweets were recorded on Friday. We divided the dataset into three parts depending on the dates. The dates for one of the datasets was August 15 – August 29, which was considered the pre-hurricane period; the next part was August 30 – September 13 which was considered as the period when the hurricane was developing and affecting different regions and finally September 14 – September 30, period called the post duration of the hurricane. This division helped us in considering only the users with tweets on all the three-time period, as only then it will be useful for us to focus on the mobility pattern. We found 6888 users who were present in all of the three time periods. We further wanted to make sure that a particular user has at least 2 tweets in each of the periods. After this filtration there were 4858 unique users left.
For all these users, the centroid of the coordinates was found for all the three time periods. Next, a Geo-Dataframe was formed for all the three time periods to find the counties for each of the centroid point. To find the counties, county shape file was used. The shape file helped in identifying the county name as well as the state for each of the points. We saw that out of 67 counties of Florida there were 63 counties in our dataset from where the tweets were done in the pre-hurricane period. The intensity of the number of tweets from each county of Florida was plotted for all the 3 time periods [Fig.4].

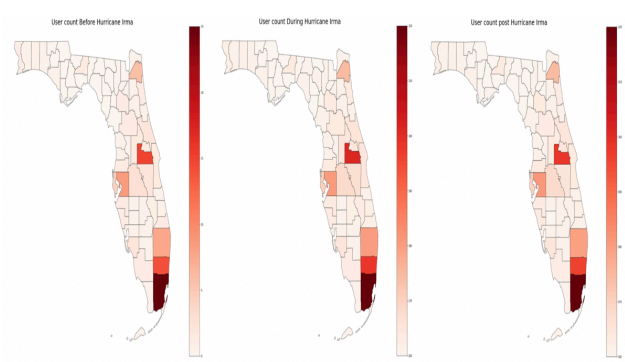


Fig. 4

It was seen that there was not much difference that we could see using the figures and therefore, we decided to look at the numbers.

After analyzing the found-out information we discovered that there were tweets among our dataset which were done by the users who were not in Florida. However, since the motive of this study was to study the patterns of the residents of Florida, we decided to filter out the people whose location during the pre-hurricane period was not in Florida and continued the analysis with the remaining 4717 unique users. Now, after comparing the names of the counties for each user-id we found that there were 1623 people who displaced from their origin counties in the during-hurricane period. Looking at the post-hurricane period we found 1512 people who had displaced from their origin position. This number was less as compared to the during-hurricane period and it can be implied that some of the people moved back to their origin counties in the post-hurricane period.

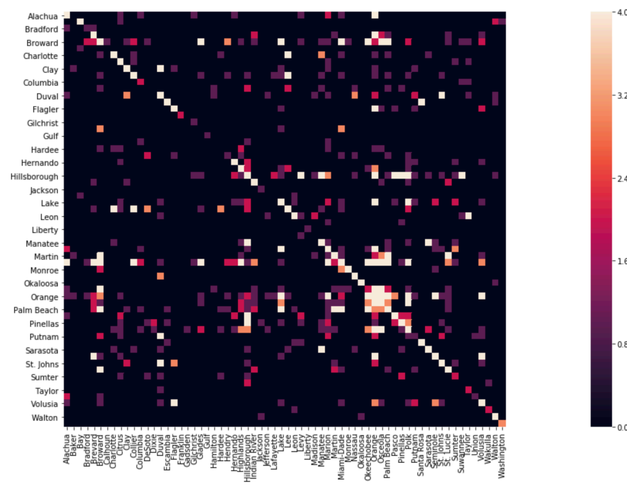

Fig. 5

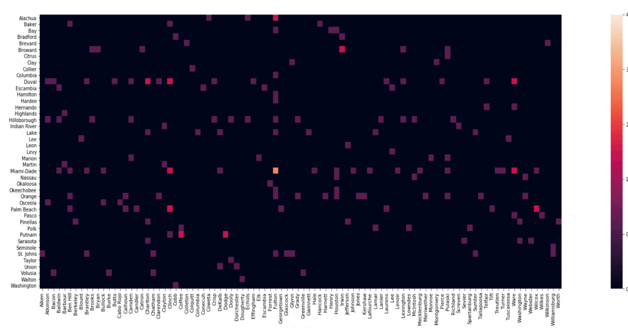

Fig. 6

Diving deep into the analysis, we found out that in the during-hurricane period, people had moved from 63 counties of Florida to 152 different counties. This implied that some of the people had moved to different states. We now created heat maps between the county found for pre-hurricane period and during the hurricane for all the user ids. [Fig.5] shows the heat map for the inter county displacement within

the state of Florida while [Fig. 6] shows the heat map for interstate displacement. We again created heat maps between the counties in pre-hurricane period and the post-hurricane period which looked almost similar to the ones showed in [Fig.5] and [Fig. 6].
We further found that out of 1623 people, 177 had moved to 10 different states in the during-hurricane period with the majority in the state of Georgia. Similarly, in the post-hurricane period, 119 out of 1512 people were found to be present in 8 different states with the majority in Georgia. Thus, it can be said that majority of the people who moved out of Florida preferred Georgia which is a neighboring state to Florida as shown in [Fig. 7]. The figure shows only those states where people had displaced.

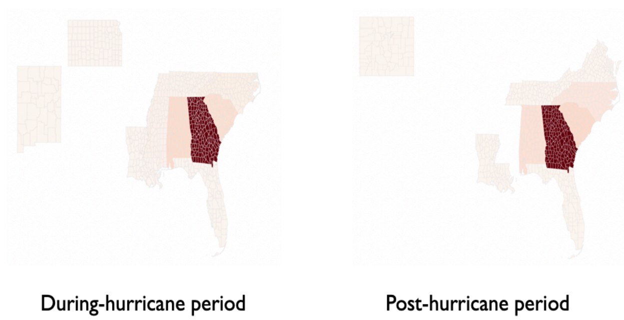


Fig. 7

Next, we wanted to visualize the difference in the number of people coming to a county and going from a county during the two time periods mentioned above for the state of Florida. We created 3 csv files consisting of user-id and the counties for all the 3 time periods and compared them in Tableau [Fig. 8].

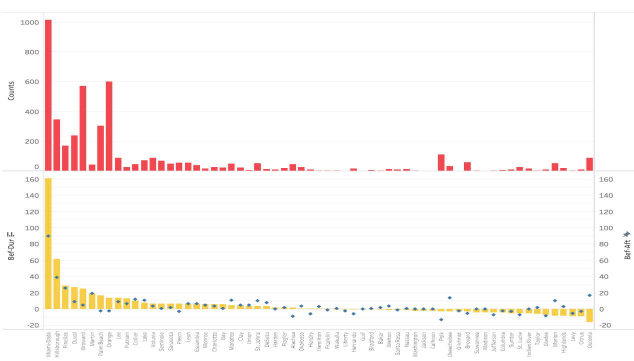

Fig. 8

Here, the red bars show the count of users for each county of Florida in the pre-hurricane period. The yellow bars show the difference between the number of users in pre-hurricane period and during-hurricane period for each county such that a positive value for a yellow bar means those many users have moved from that county. A negative value for yellow bar means that those many users have come into that county. Similarly, the blue dots show the difference between the number of users in pre-hurricane period and post-hurricane period for each county. On comparing the yellow bars and the blue dots for each county, we could see that the number of displaced people from many of the counties have dropped in the post-hurricane period which suggests that many people came back to their origin counties after the hurricane was over.

## Results

It was observed that a high number of tweets came from select counties such as Broward, Palm Beach, Miami-Dade, Orange and Hillsborough. It was seen that majority of the people stayed at their origin counties even when the hurricane hit the state of Florida. Only 1623 out of 4717 people migrated in the during-hurricane period out of which 177 moved to different states and the rest within Florida. In the post-hurricane period, 1512 people were found to be migrated out of which 119 had moved to different states. This showed that most of the people moved within Florida. Out of the people who had migrated to other states mostly went to Georgia. Maximum mobility was recorded from Miami, Palm Beach and Broward counties. We also observed that some of the people returned back to their origin counties after the hurricane. Looking at the results we could see a potential future work that can be stated to make the project an ideal study.

## Limitations and Future Work

The above results made us realize that however, there was maximum mobility recorded from Miami, Palm Beach and Broward counties, they were not the most affected areas from the hurricane. According to information collected by us, Collier and Lee counties were the ones where the hurricane hit directly resulting in floods in Lee and landslide in Collier. However, the weather forecast predicted the hurricane to hit Miami and the nearby areas. Now, we can conclude multiple things from our findings which are all assumptions and would require a future analysis. First, maximum mobility was recorded from all the unaffected counties or not so much affected counties may be due to the wrong prediction. Second, we analyzed the mobility from the two affected counties but did not find any significant movement, this can be due to many reasons which can all be termed as assumptions once again; we did not have enough data from these counties or the population of these counties is much lesser than the Miami, Palm Beach and Broward counties which can be a factor. Third, there were certain counties except the three mentioned that can be seen in Fig. 4 which had high number of people tweeting from that counties. This can be due to the reason that these counties had shelter locations set up by the government or special assistance provided for the evacuees.

We can make the project much more realistic by implementing text analysis on the tweets and considering only those users who were talking about the hurricane, shelter, etc. This might help us finding some legitimate trends. We can also check the user-ids with more than 600 tweets and see if they were bots, humans, social workers, etc. We can even apply a model on the above results to predict the mobility of the evacuees thus making the future evacuation process much more effective.

## References


1. Wang Q, Taylor JE (2014) Quantifying Human Mobility Perturbation and Resilience in Hurricane Sandy. PLoS ONE 9(11): e112608

2. Jurdak R, Zhao K, Liu J, AbouJaoude M, Cameron M, Newth D (2015) Understanding Human Mobility from Twitter. PLoS ONE 10(7): e0131469.

3. Wang Q, Taylor JE (2016) Patterns and Limitations of Urban Human Mobility Resilience under the Influence of Multiple Types of Natural Disaster. PLoS ONE 11(1): e0147299.

4. Blanford JI, Huang Z, Savelyev A, MacEachren AM (2015) Geo-Located Tweets. Enhancing Mobility Maps and Capturing Cross-Border Movement. PLoS ONE 10(6): e0129202.

5. Bernd Resch, Florian Usländer & Clemens Havas (2017): Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment, Cartography and Geographic Information Science, DOI: 10.1080/15230406.2017.1356242

6. Hughes, A.L. and Palen, L. (2009) 'Twitter adoption and use in mass convergence and emergency events', *Int. J. Emergency Management*, Vol. 6, Nos. 3/4, pp.248–260.

7. Li, Y.; Li, Q.; Shan, J. (2017) Discover Patterns and Mobility of Twitter Users—A Study of Four US College Cities. ISPRS Int. J. Geo-Inf. 2017, 6, 42.

8. Yue Y., Lan T., Yeh A., Li Q. (2014) Zooming into individuals to understand the collective: A review of trajectory-based travel behaviour studies. https://doi.org/10.1016/j.tbs.2013.12.002

9. Eilander D., Trambauer P., Wagemaker J., Loenen A. (2016) Harvesting Social Media for Generation of Near Real-time Flood Maps. https://doi.org/10.1016/j.proeng.2016.07.441

10. Guido Cervone, Elena Sava, Qunying Huang, Emily Schnebele, Jeff Harrison & Nigel Waters (2016) Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study, International Journal of Remote Sensing, 37:1, 100-124, DOI: 10.1080/01431161.2015.1117684

11. Hossman T., Legendre F., Carta P., Gunningberg P., Rohner C., (2011) Twitter in disaster mode: opportunistic communication and distribution of sensor data in emergencies.

12. Botta F., Moat H., Preis T. (2015) Quantifying crowd size with mobile phone and Twitter data https://doi.org/10.1098/rsos.150162

13. Yin J., Lampert A., Cameron M., Robinson B., Power R. (2015) Using social media to enhance emergency situation awareness. Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)

14. Martín Y, Li Z, Cutter SL (2017) Leveraging Twitter to gauge evacuation compliance: Spatiotemporal analysis of Hurricane Matthew. PLoS ONE 12(7): e0181701. https://doi.org/10.1371/journal.pone.0181701

15. Linna Li , Michael F. Goodchild & Bo Xu (2013) Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr, Cartography and Geographic Information Science, 40:2, 61-77, DOI: 10.1080/15230406.2013.777139