

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Analysis of the Twitter Response to Superstorm Sandy: Public Perceptions, Misconceptions, and Reconceptions of an Extreme Atmospheric Hazard

John A. Knox, Brendan Mazanec, Emily Sullivan,
Spencer Hall and Jared A. Rackley

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/64019>

Abstract

Superstorm Sandy was the second-costliest hurricane in U.S. history, causing catastrophic flooding and prolonged power outages in New Jersey and New York. The public's response to this extreme event on the social network Twitter is examined using statistical analysis and manual inspection of 185,000 "tweets" relating to Sandy. Sentiment analysis of tweets from Manhattan Island reveals a statistically significant trend toward negative perceptions, especially on the southern half of the island, as Sandy made landfall. Inspection of all tweets uncovered scientific misconceptions regarding hurricanes, and a surprising and disquieting anthropomorphic *reconception* of Sandy. This reconception, divorced from factual information about the storm, dominated the "Twittersphere" compared to official scientific information. The implications of such reconceptions for social media communication during future extreme events, and the utility of the methodology employed for analysis of other events, are discussed.

Keywords: hurricane, social media, scientific misconceptions, Superstorm Sandy, Twitter

1. Introduction

At the end of October 2012, Superstorm Sandy—a hurricane with the size of an extratropical cyclone—battered the mid-Atlantic coast of the United States. This chapter examines the

response of the public on the social media venue Twitter [1] to the approach, landfall, and immediate aftermath of the storm.

Twitter is one of the most popular social networks in the world, with approximately 180 million users at the time of Sandy and over 300 million users at the end of 2015 [2]. A novel feature of Twitter is the restriction of posts to 140 characters or fewer. This restriction encourages brevity and contractions among users, and limits in-depth discussion, unlike other social networks such as Facebook.

The social media reaction to Sandy has been examined from a variety of angles. Edwards et al. analyzed millions of geolocated “tweets” (i.e., posts on Twitter) and focused on the utility of Twitter to meet needs often provided heretofore by first responders and relief agencies [3]. Similarly, Chatfield et al., in a conference presentation, studied the ability of Twitter users to convey time-critical information during this disaster [4]. Lachlan et al. noted, however, that Twitter communications during the storm were used more for emotional release than for dissemination of information, and that messages from official organizations were largely absent [5]. In addition, a large automated effort to examine Twitter messages during Sandy is underway at the National Center for Atmospheric Research [6].

Our research presented in this chapter attempts to combine the best aspects of statistical analysis of a large dataset with qualitative insights gained by manual, not automated, examination of individual tweets. We have accomplished the latter via the creation of original software which makes visual inspection of Twitter messages easy and efficient. Our initial objective was to characterize misconceptions and their propagation in Twitter posts; however, our research revealed not only misconceptions but an intriguing, and disquieting, *reconception* of Sandy that threatened to drown out factual messages regarding the storm.

Below, we provide an overview of Superstorm Sandy, discuss in detail the two methodologies used to study the dataset of tweets, examine the results from both methodologies, and provide conclusions based on our quantitative and qualitative results.

2. Physical science overview

This overview of Sandy relies on the authoritative post-storm analysis of [7]. Sandy began as a tropical wave off the west coast of Africa on 11 October 2012. After traversing the tropical Atlantic and moving westward into the Caribbean without much growth, it reached tropical storm intensity (34 kt or 39 mph or 17 ms^{-1}) on 22 October south of Jamaica (**Figure 1**). At this point the storm received the name “Sandy,” from the list created by the World Meteorological Organization for Atlantic hurricanes which alternates between male and female names familiar to the cultures that border the tropical Atlantic. “Sandy” was selected a female name, given to the 18th named storm of the 2012 Atlantic hurricane season, following “Rafael.”

Two days later, after executing a loop south of Jamaica, Sandy reached Category 1 hurricane strength (64 kt or 74 mph or 33 ms^{-1}) on the Saffir-Simpson Hurricane Intensity Scale at 1200 UTC 24 October just off the southeast coast of Jamaica. It then crossed Jamaica, moving

northward, and became a major hurricane (100 kt or 115 mph or 51 ms^{-1}), i.e. a Category 3 on the Saffir-Simpson scale, just prior to landfall in Cuba. Jamaica, Cuba, Haiti, and the Dominican Republic were all impacted by Sandy, with 69 deaths and hundreds of thousands of homes destroyed, particularly in Cuba. Sandy expanded greatly in size after crossing Cuba, with the radius of tropical-storm-force winds doubling by the time it passed the Bahamas. However, its winds dropped in intensity, below the threshold for a hurricane.

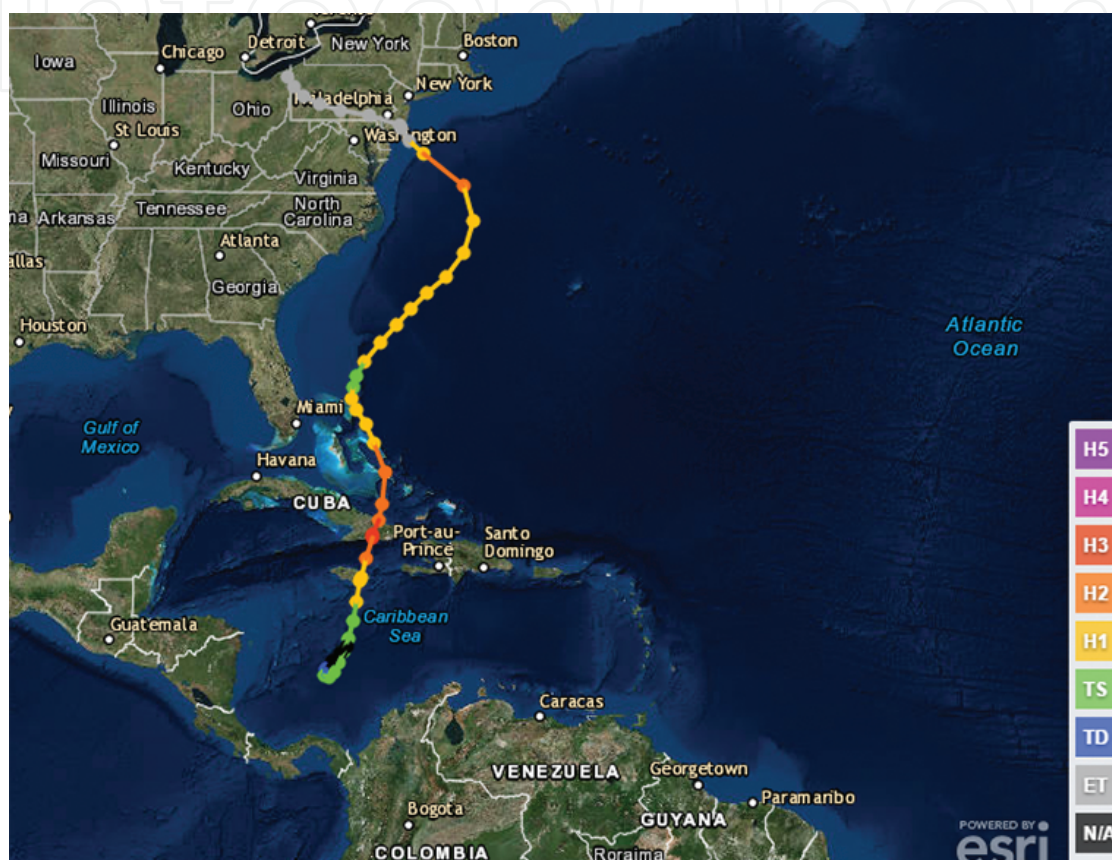
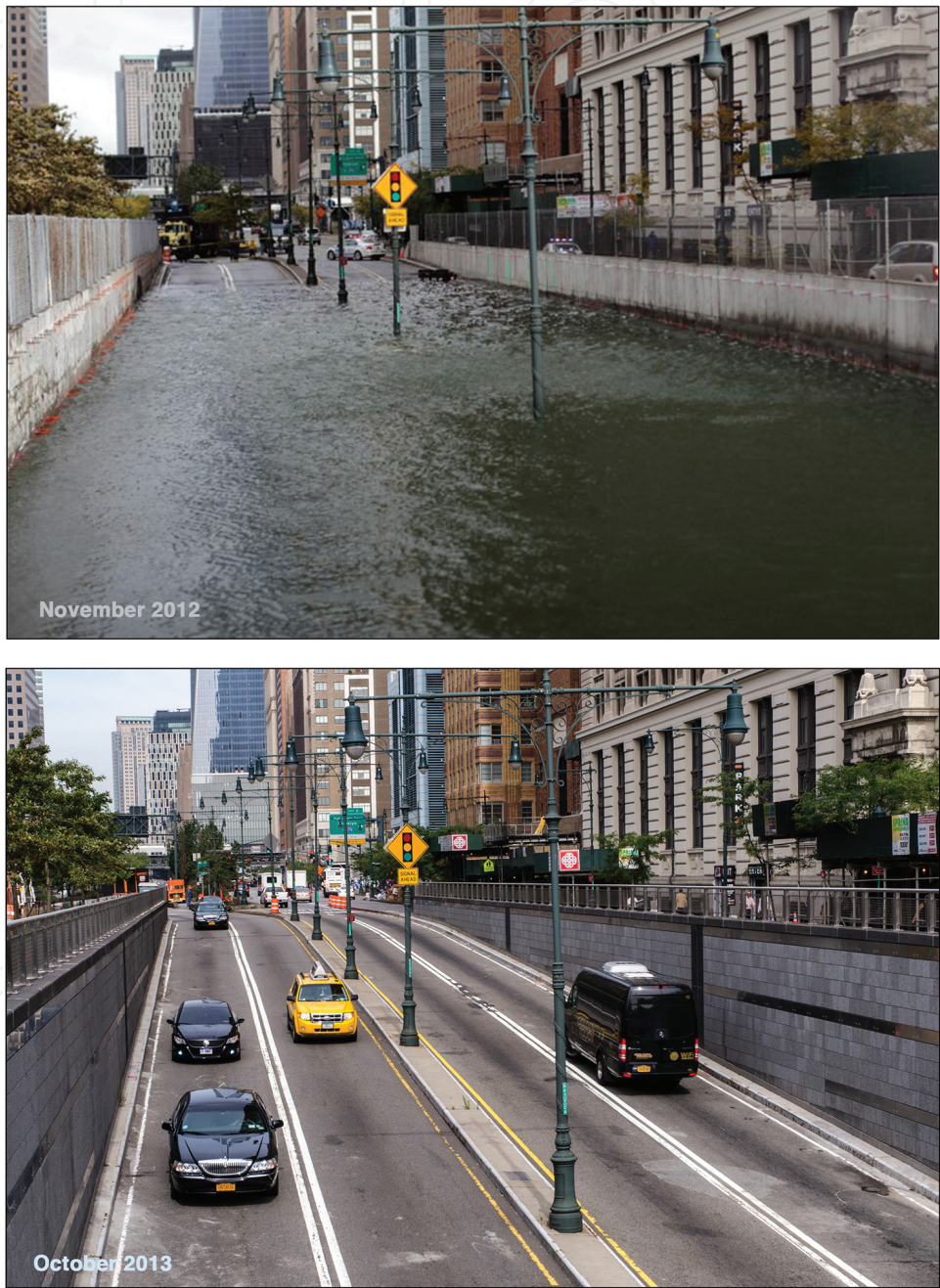


Figure 1. The track of Sandy from inception in the Caribbean Sea until its demise after extratropical transition (ET) over the northeastern United States. Saffir-Simpson categories are indicated by number and shading. Image adapted from <https://coast.noaa.gov/hurricanes/?redirect=301ocm>.

As it moved northward in the western Atlantic, Sandy regained strength and became a menace to the Atlantic coast of the United States. Due to interactions with an approaching trough over North America and the influence of warm Gulf Stream waters underneath it, Sandy intensified as it moved north and reached a secondary maximum of 85 kt (98 mph or 44 ms^{-1}) at 1200 UTC on 29 October about 220 nautical miles (405 km) southeast of Atlantic City, New Jersey. Under the influence of a strong high-pressure system to its north, Sandy executed a highly unusual left turn that brought the storm inland over New Jersey instead of heading eastward out to sea. At 2330 UTC on 29 October, Sandy made landfall just northeast of Atlantic City with an estimated sustained wind of 70 kt (81 mph or 36 ms^{-1}) and a minimum central pressure of 945 mb. It was one of the largest and most intense hurricanes ever observed in the mid-Atlantic,

particularly for so late in hurricane season. Due to the increasingly extratropical nature of the storm as it approached the coast, hurricane experts reclassified Sandy as a “post-tropical” storm shortly before landfall. However, in terms of hazards such as wind, rain, and storm surge, Sandy was virtually indistinguishable from a hurricane. In this paper we adopt the descriptor “Superstorm” for Sandy in order to reflect its dual nature.



Battery Park Underpass - New York City, New York

Figure 2. The Battery Park underpass at the southern tip of Manhattan Island immediately after Sandy (top), and what it normally looks like (bottom). Photo courtesy K.C.Wilsey/FEMA.

Sandy's impacts in the United States were unusually severe and widespread, owing to its nearly 1000-mile swath of gale-force winds near the time of landfall. It was the deadliest tropical cyclone to hit outside of the southern U.S. in 40 years, with 72 U.S. deaths directly attributed to Sandy. A majority of these deaths (41 out of 72, or 57%) were due to storm surge along the Atlantic coast, but 20 of the deaths (28%) were due to falling trees. In addition, 87 indirect deaths were caused by Sandy, mostly related to loss of power during cold weather. At least 650,000 homes were damaged or destroyed in the U.S., approximately half in the state of New York near or along the coastline. So many homes perished because of the exceptional storm surge, aided by a full-moon tide, which reached a record 9.40 feet (2.87 m) above normal at the Battery on the southern tip of Manhattan Island (**Figure 2**). Similar high tides occurred just to the right of Sandy's eye at landfall; before it failed, the Sandy Hook tide gauge recorded a surge of 8.57 feet (2.61 m) above normal. Throughout the region, from New Jersey to Connecticut, barrier islands were inundated and in some cases breached; coastal areas were flooded to depths of several feet.

Superstorm Sandy caused approximately \$50 billion in damage and was the second-costliest hurricane in U.S. history, topped only by Hurricane Katrina in 2005. About 8.5 million people lost electrical power during the storm, most of them in the hard-hit regions of New Jersey and New York; some customers were without power for months. It was the worst disaster in the history of the New York subway system. The flooding and power outages due to Sandy also closed the New York Stock Exchange for 2 days, the NYSE's longest closure in 124 years. It is in the context of this extreme event that we now turn to the public's response on Twitter.

3. Data and methodology

As a reminder, data for this study are Twitter posts, which are uncensored public utterances on a social media platform. Readers are advised of more-than-occasional strong language that is inevitably included in this narrative.

According to Pew Research, there were over 20 million tweets about Superstorm Sandy from October 27, 2012 through November 1, 2012 [8]. An analysis of that volume of data was beyond the scope of this project. Instead, we chose to isolate subsets of this immense trove of tweets and eventually created our own software to examine, both qualitatively and quantitatively, a sizable subset of the full trove. The intent was to create an intermediate level of breadth of tweets and depth of analysis of the tweets, rather than either to crunch statistics on a huge dataset or else to scrutinize in fine detail a small number of tweets. Our approach merges both statistical analysis with informed qualitative impressions based on the personal reading of thousands of tweets, made efficient via our software.

To this end, a third-party Twitter export service, GNIP Company, was used to acquire the necessary Twitter data needed for this study. The service used the two keywords "sandy" and "superstorm" to sift through all of the tweets posted between October 25, 2012, and November 2, 2012, tag the tweets that contained one or both of those keywords, and export the tagged tweets to two JavaScript Object Notation (JSON) files. A JSON file is a representation of JavaScript Objects in text form. These files are referred to below as "the full dataset" or "the

complete dataset,” and contain approximately 185,000 individual tweets. To our knowledge this wealth of data makes our research one of the more comprehensive analyses to date of social media during the Sandy event.

Two methodologies were pursued in the course of our research, and are discussed in detail below.

3.1. Sentiment analysis methodology

3.1.1. Data

Initial analysis of the dataset was begun using the open source software OpinionFinder [9], which can identify subjectivity and positive or negative sentiment in phrases. This software is used widely in multiple disciplines [10]. Tweets are classified on a numeric scale with positive tweets set to greater than zero, negative tweets set to less than zero and neutral tweets set to zero. Problems inherent in the OpinionFinder classifying system are described in Ref. [11]. This study assumes that OpinionFinder correctly identifies tweets as positive, negative, or neutral, but also notes the tendency of OpinionFinder to over-classify tweets as neutral. Fortunately, a neutral zero does not skew the data. Numerous other studies have used OpinionFinder analysis of tweets to conduct research, including some work on Superstorm Sandy [12].

A small subset of the full dataset, tweets on Manhattan Island, was examined. The point of this component of the research was to track the evolution of sentiment over time; therefore, classified tweets were divided into equal time intervals. The first analysis uses 17 time intervals of 12 hours each, while the second compares a before-event interval of to a during/after event period. Twelve-hour intervals were chosen for simplicity while maintaining high enough time resolution to denote change. The two 12-hour periods from 1200 UTC 29 October to 1200 UTC 30 October also encompass an appropriate time period of the direct impacts from Sandy on Manhattan. The time of 1200 UTC 29 October was chosen as the mark between pre-event and event based on Manhattan weather observations.

Classified tweets for each time period were then aggregated using ArcMap to census tracts in Manhattan, with each tract taking the sum of the classified tweets (i.e., tracts that have a higher proportion of positive tweets to negative tweets obtain a more positive rank). Aggregating the data points to census tracts made it more manageable and help to smooth out pockets of many tweets versus areas with fewer tweets. Census tracts also help to distribute the data by population, because tracts are roughly similar in population count. The census tracts in Manhattan are small enough in size that it can be assumed residents within the area experienced similar effects from Sandy.

Bias is inherent in using data such as tweets. Here, the database is obviously biased towards those with access to the internet via desktop (and especially access to smartphones, because power outages occurred), and is presumably also biased toward a younger demographic that utilizes Twitter. For this study, we assume that the opinions reflected in the tweets adequately reflect the opinion of others in the area that did not have access to Twitter.

3.1.2. Statistical analysis

The first step in analyzing the evolution of the opinions over time was to determine if there is a statistically significant change. To accomplish this, an analysis of variance (ANOVA) model was run using the RStudio statistical IDE [13]. ANOVA can effectively determine if one or more of the time intervals has a statistically different positive/negative opinion. In our work, the null hypothesis was that no change in opinion occurred over the 9 days.

To further assess an evolution in opinion, a second test was used to determine if there was a difference in opinion between two time periods: before Sandy's impacts were felt in Manhattan and after the effects began to be felt. In order to determine which test is appropriate for measuring the difference between these time pairs, that data were tested for normality using the Kolmogorov-Smirnov (K-S) Goodness-of-Fit test. If the data were determined to be non-normal by the K-S test, then the Wilcoxon Matched Pairs Signed-Ranks test was then used to determine a difference. (The Wilcoxon Matched Pairs Signed-Ranks test is appropriate for determining whether there are significant differences in a pair of non-parametric data such as this.) Again, the null hypothesis was that there was no change in opinion before and after Sandy's impacts. Finally, sums of the classified opinions for each of the 12 time intervals were plotted using R's barplot command. This allowed for visualization of the data and further determines directionality of the alternate hypothesis.

3.2. TweetReviewer methodology

3.2.1. Data

After our initial OpinionFinder effort, we decided to explore the full dataset in more breadth and depth. The JSON files were then imported to a MySQL database using an import program written in PHP to accomplish this more in-depth analysis.

3.2.2. Quantitative and qualitative analysis

Once the data were in the MySQL database, a tool was needed to aid in reviewing the tweets, sorting them as relevant (pertaining to Superstorm Sandy in some manner) or irrelevant (not referring to Superstorm Sandy) to the project, and bookmarking the tweets of interest. The program TweetReviewer was built by the second author on the .Net 4.0 Framework and written in C# for this purpose (**Figure 3**).

A set of filter words were created and plugged into the program to help determine the relevance of the numerous tweets. The filter words are listed in the Appendix. These filters were used to aid the researchers as they went through the tweet dataset by hand to determine the relevance of the tweets the filters did not tag. The tweets deemed relevant either through the filter tagging process or by hand, were sorted together. The non-relevant tweets were also marked accordingly and separated from the relevant tweets. However, none of the non-relevant tweets were deleted. The entire database of tweets was archived and retained.



Figure 3. TweetReviewer software graphical user interface.

Once the relevant and non-relevant tweets were separated, new keywords were searched using MySQL's search capabilities to identify the number of times that particular keyword was used. This process helped provide a clearer picture of the perception of Superstorm Sandy through the lens of Twitter.

The posts have been analyzed spatially as well as temporally, with word counts and word clouds as a function of day and time scrutinized as quantitative measures of public responses. We have also examined tweets individually by the thousands using the database, which has permitted the authors to develop qualitative insights into the public response that would be difficult or impossible without simple visualization and bookmarking tools.

4. Results

4.1. Sentiment analysis

Analysis using ANOVA allowed for a rejection of the null hypothesis that no change in opinion occurred over the 9 days examined. ANOVA showed that there was a significant difference between the opinions of the different 12-hour time intervals. As seen in **Table 1**, based on an F -value of 3.971 calculated using [14] and a p -value of 0.0464, a difference was found between the time intervals that are significant at the 95% confidence level.

The second set of tests included the Kolmogorov-Smirnov (K-S) Goodness-of-Fit test, and Wilcoxon Matched Pairs Signed-Ranks test. Results from the K-S tests, shown in **Table 2**, demonstrate that the data are not normally distributed. p -Values for both "before" and "after" data are vanishingly small (< 0.001), indicating that the null hypothesis can be rejected and a non-parametric test statistic would be more appropriate, such as the Wilcoxon Matched Pairs Signed-Ranks test.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	1	2.1	2.0868	3.971	0.0464*
Residuals	4792	2518.4	0.5255		

Table 1. ANOVA results for Manhattan Island sentiment analysis. An asterisk indicates that the result is significant at the 95% confidence level.

<i>Before</i>	<i>After</i>
$D = 0.2183, p\text{-value} < 2.2\text{e-}16$	$D = 0.1573, p\text{-value} < 2.2\text{e-}16$

Table 2. Lilliefors (Kolmogorov-Smirnov) Normality test results for the Manhattan Island sentiment analysis.

The results of the Wilcoxon Matched Pairs Signed-Ranks test give $p < 0.001$ (**Table 3**). A p -value this small indicated that, once again, the null hypothesis can be rejected. We can confidently say that opinions for the time period before Sandy’s effects were felt in Manhattan differ from the period during and after effects were felt.

<i>Data: Before vs. After</i>	
$V = 18,766$	$p\text{-value} = 9.881\text{e-}07$

Table 3. Wilcoxon Matched Pairs Signed-Ranks test results for the Manhattan Island sentiment analysis.

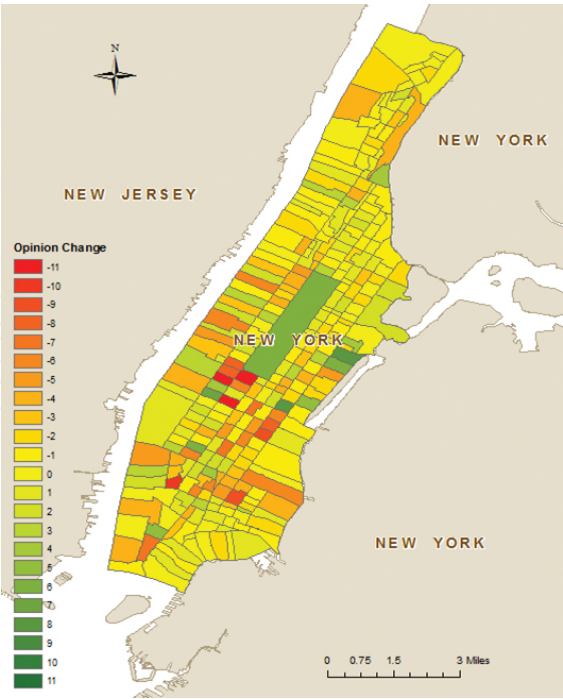


Figure 4. Sentiment analysis of opinion change in tweets on Manhattan Island from before Sandy to during/after Sandy’s impact on the area.

Figure 4 demonstrates this change geographically across Manhattan Island. The most negative changes occurred south of Central Park, broadly consistent with the region that experienced a power blackout due to Sandy [15]. Lesser negative trends, and some positive trends, were found most often north of Central Park, further from the blackout and from the Atlantic coast.

In addition, sums of opinions were plotted over time using R’s barplot command. The resulting diagram (**Figure 5**) allows a clear visualization of the evolution of opinions over time. The bar graph shows that overall opinions through 1200 UTC on 28 October tended to be positive, with the exception of slightly negative tweets from 1200 UTC to 1200 UTC on 26 October–27 October. After 1200 UTC 28 October, tweets became much more negative, reaching their negative peak at 1200 UTC on 29 October through 1200 UTC 30 October. Tweets remained negative in sentiment until the end of the period.

Overall, the results indicate that public opinion of Sandy did change in the Manhattan area. It appears that public opinion, expressed through tweets, was more lighthearted before effects from Sandy were felt in Manhattan. This correlates with observed skepticism from many in the path of the storm prior to actual impacts (see the following section). Interestingly, tweets become more negatively skewed 24 hours prior to actual impacts to the local Manhattan area. This could be related to non-meteorological impacts from Sandy that occurred prior to the storm’s direct impacts. As stores, subways, and other services closed, public opinion may have begun to shift. These early negative tweets could also be the result of evacuation orders, or news of destruction as Sandy swept up the East Coast. Further tests would be needed to determine the impacts of these and many other variables. It is clear, however, that the ratio of positive to negative opinions became most negative at the time when Sandy’s impacts were being felt in Manhattan (1200 UTC–1200 UTC 29–30 October).

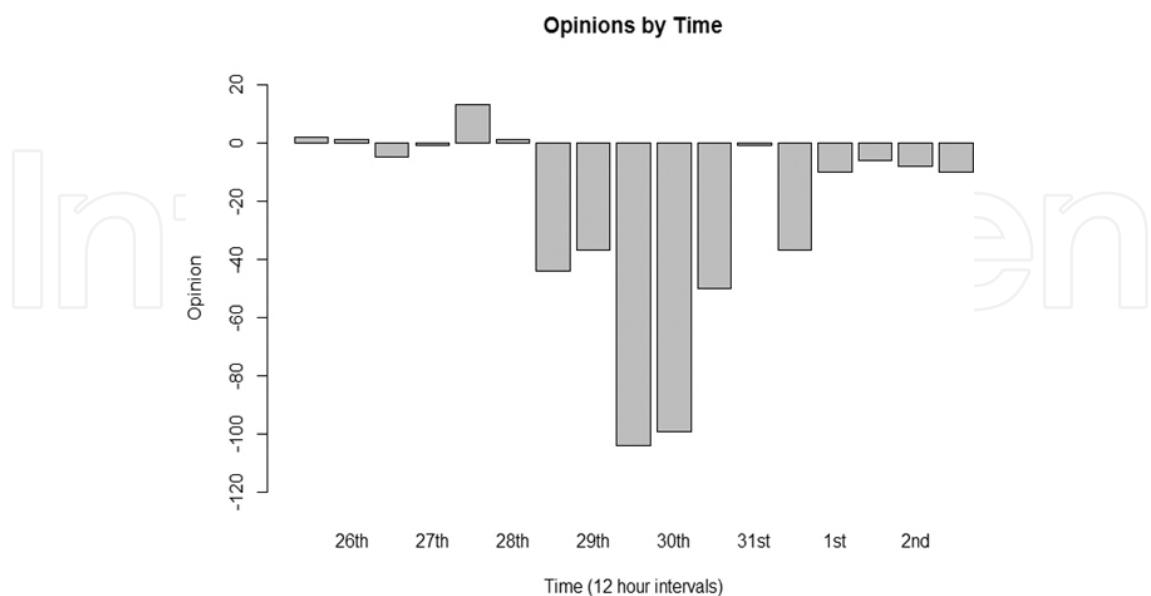


Figure 5. Temporal analysis of sentiment (positive or negative) on Manhattan Island in 12-hour increments from 25 October 2012 through 2 November 2012.

4.2. TweetReviewer analysis

The results uncovered by sentiment analysis failed to capture other insights that the researchers felt were important but were outside the scope of OpinionFinder. This led to the creation and use of TweetReviewer software to allow us to examine, both quantitatively and qualitatively, other aspects of the full database.

While analyzing the full database, it was determined there were two major common categories of interest with regard to the public’s perceptions of Superstorm Sandy: scientific misconceptions, and an anthropomorphic reconception of Sandy that, in the parlance, “went viral.”

4.2.1. Misconceptions

There were many misconceptions regarding Hurricane Sandy and hurricanes in general in the dataset. One quotidian misconception was the common misspelling of the word “hurricane.” In fact, the word “hurricane” was misspelled so often that the list of filtered words had to be adapted to accommodate the many misspellings (e.g., see #51 in Appendix). More substantively, there was a general lack of understanding of what defines a hurricane. These misconceptions can be separated into four categories: *strength*, *category*, *size*, and *duration*.

The strength of a hurricane is most commonly associated with the Saffir-Simpson Hurricane Wind Scale categories. Overall, the Twitter users had a fairly decent grasp of the numerical scale as far as the numbers were concerned, but the understanding of what those numbers stood for was almost entirely absent. Throughout the dataset (Table 4) were tweets, many from Florida, that downplayed Hurricane Sandy claiming there was little to worry about because it was “only cat 1.” Instead of acknowledging the force of the hurricane, users discriminated against the lower numbers even though the Saffir-Simpson Scale is based on wind only. It does not take into account the damage from other impacts like the storm surge and rainfall, and neither did the Twitter users. A few tweets mentioned Hurricane Sandy as having a storm surge more often associated with a Category 3 or Category 4 hurricane, such as this (tweet #76357 in the full database):

“Irene was a category 3. Sandy is a 1 but with the storm surge it’s supposed to act as a 4 at most.”

	Category 1	Aftermath	Apocalypse	Worse	Power
25 Oct. 2012	2	0	1	6	10
26 Oct. 2012	13	1	4	28	28
27 Oct. 2012	5	1	4	21	55
28 Oct. 2012	17	7	343	61	182
29 Oct. 2012	26	8	1136	213	1326
30 Oct. 2012	14	212	398	220	1748
Total	77	229	1886	549	3349

Table 4. Tweet word counts related to hurricane strength.

	Tornado	Touchdown	Landfall	How Long
25 Oct. 2012	13	0	10	0
26 Oct. 2012	21	0	11	1
27 Oct. 2012	33	1	10	1
28 Oct. 2012	126	7	22	2
29 Oct. 2012	213	4	142	13
30 Oct. 2012	170	2	43	10
Total	576	14	238	27

Table 5. Tweet word counts related to hurricane size.

	Over/over yet	Finish	Finally	Bring it	Passed/past
25 Oct. 2012	28	1	0	2	2
26 Oct. 2012	69	2	3	13	3
27 Oct. 2012	76	2	6	35	6
28 Oct. 2012	244	17	17	112	22
29 Oct. 2012	784	38	60	191	60
30 Oct. 2012	785	30	67	23	82
Total	1986	90	153	376	175

Table 6. Tweet word counts related to hurricane duration.

The size and duration of Hurricane Sandy was also frequently misunderstood (Tables 5 and 6). Very few tweets implied that the users fully grasped the sheer size of any hurricane, let alone the immense size of Sandy. They appeared to assume that a hurricane was a small storm that would be bad, much like a supercell or “derecho-sandy thing” (tweet #4885) that would “touchdown” or “touch land” (tweet #100800), wreak havoc, close schools, and leave, all within a matter of hours. There were several tweets that asked if the storm was “over yet?” Hardly any of the Twitter users seemed to grasp the fact that a hurricane is actually a huge storm spanning hundreds of miles which can last for days. Some tweets even called Hurricane Sandy a tornado (tweets #38455 and #49032) or, as one CEO put it, a “tornadocaine” (tweet #5267).

Among many visual misconceptions propagated during Superstorm Sandy, one of the most prominent and widely shared was a Photoshopped image of a supercell thunderstorm over the Statue of Liberty in New York City (e.g., tweet #60774). The supercell thunderstorm was indicated to be Sandy—another sign of confusion regarding the vast differences in size, strength, and duration between hurricanes and tornadic thunderstorms. This misconception

became so popular on social media that a story about it appeared in the Philadelphia media [16].

4.2.2. *The reconception of Sandy*

As Superstorm Sandy approached the mid-Atlantic coast, a number of Twitter users began to do something peculiar, at least from the perspective of scientists or emergency managers. Instead of relaying factual information regarding the storm, accounts pretending to be the personification of Hurricane Sandy started appearing, as venues for posting jokes about the hurricane from a first-person perspective. With the help of these accounts, the Twitter community took the idea of personifying Sandy and expanded upon it. Without any visible evidence of premeditation regarding the nature of this anthropomorphic Sandy, the Twitter community banded together and simply accepted a persona of their creation without question. They also hurled insults at the hurricane based on its fabricated persona and its perceived status as a female.

Sandy was suddenly no longer an impersonal, inanimate hurricane; she was an “independent sassy black hurricane who don’t need no man” (tweets #13365 and 29 subsequent tweets), who grew up in the ghetto, went to school with (Hurricane) Irene, cursed up a storm, voted Democratic, knew how to work the pole, had an avid and colorful sex life, and was the brunt of many cruel taunts related specifically to female anatomy. **Table 7** provides a sampling of word counts related to this persona in the days leading up to landfall.

	Bitch	Whore	Twerk	Pole	Shark	Black	Nigga
25 Oct. 2012	8	0	0	0	2	0	0
26 Oct. 2012	35	2	0	1	0	1	7
27 Oct. 2012	61	3	4	0	0	8	3
28 Oct. 2012	253	17	25	2	6	22	29
29 Oct. 2012	614	41	42	16	32	96	68
30 Oct. 2012	382	31	14	21	74	168	39
Total	1353	94	85	40	114	295	146

Table 7. Tweet word counts related to the invented persona of Sandy.

The various Twitter accounts created to portray this user-crafted persona encouraged its incredibly fast proliferation and popularity across the website. They posted tweets which consisted mainly of rather dirty jokes, inappropriate suggestions, racial slurs, female-specific insults, and Republican-specific insults. One of the aforementioned tweets talked about the Sandy persona tossing a trailer at a woman in a minivan simply because she had a bumper sticker of a Republican presidential candidate on her car (tweet #21587).

It is possible that the timing of the hurricane’s landfall, within a week of the 2012 presidential elections when tensions between the Democratic and Republican parties were already high,

amplified the politically slanted comments. But the outright cruelty and twisted content went beyond simple political reasons. To make matters worse, these offensive tweets were shared hundreds of times by users who were apparently from a variety of ages, races, and political beliefs.

The tweets referring to the Hurricane Sandy persona consistently referred to her as a “bitch” and joked about her coming to “blow” the entire east coast and make everyone “wet” or simply told her to “fuck off.” This was such an exceedingly common theme that the words “bitch,” “blow,” “wet,” and “fuck” became keywords for both the filters during the process of going through the tweets as well as for searching their word counts. Of these four words, “wet” was used the least common, only garnering a word count of 544 times used. The word “blow” was used 1129 times, “fuck” was used 3655 times, and “bitch” led all epithets in the full database with a word count of 4335 times used with regard to Hurricane Sandy.

These word counts are, of course, small compared to the total number of words used in the full database. To give a sense of how the persona of Sandy dominated the “Twittersphere,” **Table 8** presents the fractional representation of scientific/hurricane-related terms in the full database versus the top three persona terms. As shown in the table, the persona was many times more popular on Twitter than were factual reports about Sandy.

NWS	MPH	Category/cat	Landfall	Windspeed
0.022	0.133	0.387	0.079	0.0001

Table 8. Ratio of word count of scientific terms to the word counts of the three most common Sandy-persona terms.

5. Conclusions

Our results shed light on public perceptions, misconceptions, and reconceptions of an extreme atmospheric hazard. Superstorm Sandy was a virtually unprecedented event along the Atlantic coast of the United States, in terms of intensity, size, and path. What did the public make of this event on Twitter?

From our analysis of Manhattan Island data, it is plausible that residents underrated the storm’s ferocity until the last 24 hours before landfall. The largest sentiment swings occurred, quite naturally, in and near the regions most affected by the storm: the coastline and the south Manhattan neighborhoods blacked out due to a power plant failure.

Our more fine-grained analysis using the TweetReviewer software revealed additional aspects of the public’s reaction to Sandy. Profound confusion exists regarding the size, strength, and duration of hurricanes. The public seems to confuse hurricanes with tornadoes; this confusion during Sandy extended to the Photoshopped image of a supercell thunderstorm over the Statue of Liberty. Expectations that Sandy would be as brief and intense as a tornado were not met; in particular, the unusually large extent of Sandy (the largest storm in terms of diameter of gale-force winds since records began in 1988; see [7]) was not well understood by those on

Twitter. The more subtle point that Sandy could be “only a Category 1” and still do extensive damage due to storm surge was also not grasped.

What the “Twittersphere” did seem to eagerly grasp was a user-generated anthropomorphic “Sandy” who dealt out death and destruction like a villain in a superhero comic or movie. This personified “Sandy” was then made the object of race-based and gender-based slurs that were widely perceived to be amusing, rather than offensive if they had been made in public. This created an increased level of noise that tended to drown out the true signal: reports of important developments in Hurricane Sandy’s category changes, the watches, advisories, and warnings, and tweets with legitimate scientific information that could have better informed the public. Our results thus align more with those of [5] than with the more positive, life-saving impacts of Twitter found by other researchers. Our analysis of geo-located tweets may bias our results somewhat in this respect, however (R. Morss, pers. comm., January 2016).

When much of New York City lost electricity, the theme of attacking the hurricane based on its manufactured persona quickly fell by the wayside in favor of complaining about the power loss and realizing Sandy was in fact a hurricane and not a fictitious creation. The persona continued to play a substantial role in the dataset, but the newest point of interest was the loss of power. In fact, the word “power” was found 8649 times in our database, making it more popular than “bitch.”

We conclude, with some surprise, that until the hurricane interfered directly with people’s personal lives, Twitter users seemed content with obsessing over the invented persona of Hurricane Sandy. Rather than bemoaning this flight from reality, however, we encourage a more proactive response among emergency management personnel, meteorologists, and others who communicate directly with the public. Perhaps it is possible that this behavior can be utilized for the benefit of the public. If the public could quickly create and propagate a persona for a hurricane, it stands to reason that official outlets could do the same. They could create a Twitter account for a new hurricane, create their own persona for it, and use this to disseminate the important relevant information to the public in a format that is more easily digestible than more esoteric scientific criteria.

For example, a large swath of the U.S. population is familiar with superheroes on some level, whether DC, Marvel (or for even younger audiences, Pokémon). What if scientists began using popular superheroes to help describe the strength of a storm? As outlandish as it sounds, this is something the general public would grasp fairly easily, simply because they are already familiar with the characters. For instance, intense heat could be described as being on par with an attack by Marvel’s Human Torch, a member of the Fantastic 4 superhero group. A powerful electrical storm could be compared to Marvel superhero Thor wielding Mjolnir in battle. The incredible storm surge of a hurricane such as Sandy could be compared to the DC supervillain Ocean Master or New Wave in combat. Similar translations of specific hazards into personae would be possible in Pokémon, reaching even younger audiences.

These examples may come across as childish, but the public would easily grasp a general idea of the intensity of the heat, the electrical storm, and the storm surge from them. More importantly, children would easily comprehend these examples. A parent can disregard weather

alerts, but not if their children continually bother them. If this method of using pop-culture references to explain something as confusing as weather can bridge the communication gap to the next generation, then they could learn to listen to and obey weather alerts. This “if you can’t beat ‘em, join ‘em” approach to anthropomorphizing atmospheric hazards could capture some of the social media energy that might otherwise propel a completely non-factual (and offensive) personification to prominence, as occurred with Sandy.

Finally, we advocate the use and/or development of software such as our TweetReviewer as a means for visually inspecting thousands of tweets easily and efficiently. As our research indicates, actually reading the tweets provides insights that are unlikely to be gained by mere statistical crunching on datasets. The filtering capabilities of TweetReviewer enable the user to focus on relevant tweets and screen out non-relevant tweets, significantly accelerating the process and permitting human analysis of relatively large datasets.

Acknowledgements

The first author gratefully acknowledges the support of an M.C. Michael Award from the University of Georgia, which funded much of the research presented here.

Appendix

Filter words used in our TweetReviewer software to narrow the database to the most relevant posts are:

1. “red cross” or “redcross”
2. “#sandy” if either “springs” or “beach” is not present.
3. “#hurricanesandy” or “#hurricainesandy”
4. “hurricane sandy”
5. “jersey shore”
6. “#njsandy”
7. “hurricane death megatron”
8. “sandy aftermathpocalypse”
9. “halloweenpocalypse”
10. “#postsandy”
11. “#aftersandy”
12. “#fucksandy”
13. “#fuckyousandy”
14. “#fusandy”

15. “#fyouandy”
16. “fuck sandy”
17. “fuck you sandy”
18. “#ctsandy”
19. “#bptsandy”
20. “#nhsandy”
21. “#dcsandy”
22. “#wvsandy”
23. “#nysandy”
24. “#ny1sandy”
25. “#nycsandy”
26. “#masandy”
27. “#vasandy”
28. “#mdsandy”
29. “#vtsandy”
30. “#tosandy”
31. “#6abcsandy”
32. “#lifeaftersandy”
33. “storm sandy”
34. If the tweet contains both “Sandy” and “Katrina”
35. If the tweet contains both “Sandy” and “Irene”
36. “#thankshurricanesandy”
37. “#thankssandy”
38. “#thankyousandy”
39. “#damnyousandy”
40. “#damnitsandy”
41. “post-sandy”
42. “post sandy”
43. “#frankenstorm”
44. “superstorm”

45. If the tweet contains both “Obama” and “Fema”
46. “fema”
47. “#damnsandy”
48. “#darnsandy”
49. “#ihatesandy”
50. If the tweet contains both “prayers” and “affected by sandy”
51. “@hurricanesandy”
52. “@ahurricanesandy”
53. “@sandydahurricane”
54. “@sandyshurricane”
55. “#survivingsandy” or “#survingsandy”
56. “#stormsandy”
57. “sandy storm”
58. “Frankenstorm”
59. “hurricane”
60. “bitch sandy”
61. “#pray”
62. “#hurricane”
63. If the entire tweet is “superstorm sandy xuo”

Author details

John A. Knox^{1*}, Brendan Mazanec¹, Emily Sullivan¹, Spencer Hall¹ and Jared A. Rackley²

*Address all correspondence to: johnknox@uga.edu

¹ University of Georgia, Athens, GA, USA

² Oak Ridge National Laboratory, Oak Ridge, TN, USA

References

- [1] Twitter. [Internet]. 2016. Available from: <https://www.twitter.com> [Accessed: 2016-04-03]
- [2] Statista, The Statistics Portal. [Internet]. 2016. Available from: <http://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/> [Accessed: 2016-04-03]
- [3] Edwards JF, Mohanty S, Fitzpatrick P, Martin D. Assessment of social media usage during Hurricane Sandy and the development of a Twitter-based, web application for real-time communication of storm-related information. 96th American Meteorological Society Annual Meeting. New Orleans, LA, 11 January 2016. https://www.hpc.msstate.edu/publications/docs/2016/01/14285social_media_tool_storm_info_ams.pdf [Accessed 2016-04-03]
- [4] Chatfield AT, Scholl HJ, Brajawidagda U. #Sandy Tweets: Citizens' co-production of time-critical information during an unfolding catastrophe. Proceedings of the 47th Hawaii International Conference on System Science, IEEE Computer Society, Waikoloa, Hawaii. 2014. pp. 1947-1957.
- [5] Lachlan KA, Spence PR, Lin X, Del Greco M. Screaming into the wind: Twitter use during Hurricane Sandy. *Communication Studies*. 2014; 65 (5): 500-518.
- [6] Demuth JL, Morss RE, Palen L, Anderson K, Anderson J, Kogan M, Stowe K, Bica M. Understanding dynamic communication, risk perception, and decisions during Hurricane Sandy through analysis of Twitter data, Society for Risk Analysis annual meeting, Arlington, VA, 9 December 2015.
- [7] Blake ES, Kimberlain TB, Berg RJ, Cangialosi JP, Beven III JL. Tropical cyclone report: Hurricane Sandy (AL182012) 22-29 October 2012. [Internet]. 2013. Available from: http://www.nhc.noaa.gov/data/tcr/AL182012_Sandy.pdf [Accessed: 2014-04-03]
- [8] Pew Research Center. [Internet]. 2013. Available from: <http://www.pewresearch.org/fact-tank/2013/10/28/twitter-served-as-a-lifeline-of-information-during-hurricane-sandy/> [Accessed: 2016-04-03]
- [9] OpinionFinder. 2.x Release Page [Internet]. 2011. Available from: http://mpqa.cs.pitt.edu/opinionfinder/opinionfinder_2/ [Accessed: 2016-04-03]
- [10] Bollen J, Mao H, Zeng X-J. Twitter mood predicts the stock market. *Journal of Computational Science*. 2010; 2 (1): 1-8. doi:10.1016/j.jocs.2010.12.007
- [11] Wilson T, Hoffman P, Somasundaran S, Kessler J, Wiebe J, Choi Y, Cardie C, Riloff E, Patwardhan S. OpinionFinder: A system for subjectivity analysis. Proceedings of Human Language Technologies Conference/Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005), Vancouver, Canada, 2005. pp. 35-36.

- [12] George Mason University. The GeoSocial Gauge [Internet]. 2016. Available from: <http://geosocial.gmu.edu/projects/superstorm-sandy/> [Accessed: 2016-04-03]
- [13] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL <https://www.R-project.org/>.
- [14] Free Statistics Calculators. Version 4.0 [Internet]. 2016. Available from: <http://www.danielsoper.com/statcalc3/calc.aspx?id=4> [Accessed: 2016-04-03]
- [15] New York Magazine. From The Editors [Internet]. 2012. Available from: <http://nymag.com/nymag/letters/hurricane-sandy-editors-letter-2012-11/> [Accessed: 2016-04-03]
- [16] Mathis J. That Awesome Hurricane Sandy Photo of the Statue of Liberty? Fake [Internet]. 2012. Available from: <http://www.phillymag.com/news/2012/10/29/awesome-hurricane-sandy-photo-statue-liberty-fake/> [Accessed: 2016-04-03]

IntechOpen