

Social Media and Severe Weather: Do Tweets Provide a Valid Indicator of Public Attention to Severe Weather Risk Communication?

JOSEPH T. RIPBERGER

*Center for Risk and Crisis Management, and Cooperative Institute for Mesoscale Meteorological Studies,
University of Oklahoma, Norman, Oklahoma*

HANK C. JENKINS-SMITH AND CAROL L. SILVA

Center for Risk and Crisis Management, University of Oklahoma, Norman, Oklahoma

DEVEN E. CARLSON

Department of Political Science, University of Oklahoma, Norman, Oklahoma

MATTHEW HENDERSON

Center for Risk and Crisis Management, University of Oklahoma, Norman, Oklahoma

(Manuscript received 2 May 2013, in final form 26 June 2014)

ABSTRACT

Effective communication about severe weather requires that providers of weather information disseminate accurate and timely messages and that the intended recipients (i.e., the population at risk) receive and react to these messages. This article contributes to extant research on the second half of this equation by introducing a “real time” measure of public attention to severe weather risk communication based on the growing stream of data that individuals publish on social media platforms, in this case, Twitter. The authors develop a metric that tracks temporal fluctuations in tornado-related Twitter activity between 25 April 2012 and 11 November 2012 and assess the validity of the metric by systematically comparing fluctuations in Twitter activity to the issuance of tornado watches and warnings, which represent basic but important forms of communication designed to elicit, and therefore correlate with, public attention. The assessment finds that the measure demonstrates a high degree of convergent validity, suggesting that social media data can be used to advance our understanding of the relationship between risk communication, attention, and public reactions to severe weather.

1. Introduction

Despite significant advances in radar technology and subsequent improvements in the accuracy of warnings, tornadoes remain a significant threat to the property, safety, and lives of U.S. residents. In 2011 alone, 1691 confirmed tornadoes in the continental United States were responsible for 550 fatalities, approximately 5400 injuries, and more than \$10 billion of property and crop damage (NOAA/NWS Storm Prediction Center 2012). Though anomalously high, the record-setting damage caused by tornadoes in 2011 serves as an important and

sobering reminder that scientific and technological advancements in our understanding of severe weather can only go so far—tornadoes occur in a social environment wherein the exposed population, not technology or technical experts, are responsible for risk mitigation and protective action. In such an environment, effective communication between providers of hazardous weather information (i.e., forecasters) and the public at large represents a critical variable that can limit the societal impact of tornadoes (Brooks and Doswell 2002; Doswell et al. 1999).

In short form, communication denotes an exchange of information among individuals, parties, and/or groups. In the context of severe weather, tornado watches and warnings represent a basic yet essential form of communication wherein National Weather Service (NWS)

Corresponding author address: Joseph T. Ripberger, 120 David L. Boren Blvd., Rm. 3106, Norman, OK 73072.
E-mail: jtr@ou.edu; joseph.ripberger@noaa.gov

forecasters inform a given population of the potential for (in the case of watches) or imminence of (in the case of warnings) tornado development. This communication is effective if the target population 1) is exposed to the information, 2) pays attention to it, and 3) understands it. It is ineffective if the target population does not receive, attend to, or comprehend the information being conveyed (Lindell and Perry 2012). In other words, the difference between effective and ineffective communication hinges upon exposure, attention, and comprehension, all three of which are preconditions for protective action behavior (Lindell and Perry 2012). People are unlikely to engage in the actions necessary to protect themselves from tornadoes (i.e., seek shelter) if they do not receive, attend to, or understand information, like tornado watches and tornado warnings, provided by weather professionals.

To date, researchers interested in public responses to severe weather have focused most of their efforts on the factors that influence message reception and comprehension. These efforts have yielded a number of important insights. For example, extant research on reception demonstrates that members of the public are significantly less likely to receive messages that are transmitted at nighttime (Schmidlin et al. 1998). Other studies have found that minority and native Spanish-speaking populations are less likely to receive critical information about severe weather than nonminority and native English-speaking populations (e.g., Ahlborn et al. 2012). Scholars have also shown that linguistic, ethnic, and racial differences influence the extent to which people understand the information conveyed in warnings (e.g., Aguirre 1988; Donner et al. 2012). In a separate but substantively similar line of research, scientists interested in comprehension have found that the content of the message itself can influence public comprehension. For example, people are more likely to understand messages about extreme or severe weather if they include information about the nature, location, guidance, time, and source of the hazard or risk (e.g., Sorensen 2000). Adding more complex information such as the probability of a particular event can also influence the likelihood that people will understand it (e.g., Morss et al. 2008; Joslyn and Savelli 2010).

Though useful, this research largely neglects the attention part of the effective communication equation. As a result, we know relatively little about the factors that influence public attention to messages issued by weather professionals. Who pays attention to tornado watches and warnings when they are issued? How does attention fluctuate across space and time? Do message content and/or mode of delivery influence the extent to which people pay attention to information about severe weather?

Scholars have yet to answer these and other important questions about public attention to severe weather communication for a variety of reasons, some of which involve data limitations. Simply put, researchers have neglected public attention because they have yet to develop an adequate measure of the concept. In an attempt to overcome this limitation, this paper proposes, develops, and validates a new indicator of public attention to severe weather communication that is based on the growing stream of “real time” data that members of the public publish on social media platforms, which in this case is Twitter.

2. Literature review

Public attention is an important concept that has attracted research from a variety of disciplines, ranging from public health and epidemiology to economics to political science. As such, scholars have developed a number of tools intended to measure the concept. Unfortunately, the subset of tools capable of measuring attention at the societal level, rather than the individual or small group level, is rather small and suffers from a number of significant limitations for analyzing real-time changes in attention. National and international surveys, which provide the most widely used indicators of attention, are useful but temporally static and conducted only periodically. They may tell us how many people are paying attention to something on a given day at a given time (or even a series of days/times if the survey is longitudinal in nature) but can tell us little about the day or the 5–10 min after a tornado watch or warning is issued, which is when fluctuations in attention are most likely to occur and possibly impact behavior.

Measuring attention at this level of temporal detail requires continuous and real-time data, which is difficult to collect via mainstream surveys, at least in the current environment. As such, researchers have directed their efforts toward the development of a new suite of tools that leverage the massive amount of information that members of the public transmit and/or broadcast via social media to measure fluctuations of public attention. The logic underlying such measures is rather simple—the more people talk about a particular issue, topic, or hazard (via Twitter, Facebook, Google+, and other social media sites), the more likely it is that they are paying attention to it. Thus, increased discussion of an issue, topic, or hazard is thought to indicate increased attention.¹

¹For similar research using Internet search query data, see Ginsberg et al. (2009), Wilson and Brownstein (2009), Chan et al. (2011), and Scharkow and Vogelgesang (2011), and Ripberger (2011).

Chew and Eysenbach (2010), for instance, show that fluctuations in public posts on Twitter (tweets) provided a valid and useful indicator of public attention to the H1N1 pandemic that emerged in April 2009. Other researchers have arrived at similar conclusions—that public activity on social media (Twitter in the case of these studies) provides a real-time indicator of attention to issues such as norovirus (Velasco et al. 2011), influenza (Lamos and Cristianini 2010), swine flu (Szomszor et al. 2012), political parties (Tumasjan et al. 2010), and even earthquakes (Sakaki et al. 2010). In light of these conclusions, we are optimistic that social media can be used to analyze public attention to severe weather communication. Before such analyses can proceed, however, we must step back and evaluate the reliability and validity of social media–based indicators of public attention to severe weather.

3. Data

To initiate this evaluation, we drew on methods used in the studies mentioned above to create an indicator of public attention to tornadoes—one type of severe weather—based on fluctuations in the number of tornado-related tweets posted on Twitter. To accomplish this, we developed a program that interfaces with Twitter’s streaming application programming interface (API) to continuously collect and archive tweets that contain the word “tornado.” In addition to the text of each tweet, the program archives the set of metadata provided by Twitter about the tweet itself and the user that created it. With respect to the former, the program archives information such as the date and time that the tweet was published and the number of times it was “retweeted” (reposted by another Twitter user). With respect to the latter, the program archives additional provided details such as the username of the author of each tweet, the user description and URL (link) the user has posted (when available), the number of people who follow that user, and the latitude and longitude of the user when they published the tweet (when available).²

Between 25 April and 11 November 2012, we used this program to collect and archive 3 030 919 tweets (and associated metadata) that were published by 1 747 541 different Twitter users, where 75.6% of these users contributed a single tweet, 89.7% published two or fewer tweets, and 94.4% of them produced three or

fewer tweets containing the word tornado during this time period. In other words, the vast majority of users responsible for the tweets we collected published less than four tweets during the time period of our analysis, suggesting that most of them are infrequent commenters on this sort of severe weather. However, a small portion of the users responsible for these data posted a relatively large number of tweets about tornadoes during this time period. For example, “Tornado Alerts” (an account managed by Simple Weather Alert, which broadcasts NWS alerts to various communities around the county) contributed 3390 tweets, making them the most active user in this database. Despite this seemingly large number, frequent commenters are responsible for a relatively small portion of the total tweets about tornadoes during this time period. The top 25 commenters posted a total of 35 095 tweets, which is less than 1.2% of the 3 030 919 tweets contained in this database. This indicates that the overwhelming majority of tweets we collected during this time period came from users that rarely comment about tornadoes, which lends preliminary credence to the notion that these data capture “public” rather than “expert” reflections on severe weather. On top of this, our decision to collect and examine tweets that contain the word “tornado” undoubtedly biases the data toward people who use the root word “tornado” when discussing tornadoes (i.e., English-, Spanish-, and Portuguese-speaking populations).

Having briefly discussed the Twitter users responsible for publishing the more than 3 million tweets contained in our dataset, Fig. 1 illustrates the temporal variation in the tweets they posted.³ As shown in the figure, Twitter users published an average of 15 952 tweets containing the word “tornado” per day (24-h UTC period) during this time period. In terms of tweets posted, the most active days were 8 September, 1 June, and 18 September 2012, when users posted 141 975, 114 489, and 101 574 tweets containing the word tornado, respectively. The least active days were 18 July and 27 April 2012, when users posted 4115 and 5414 tweets, respectively.

This brief description of our database demonstrates that the tweets we have collected: 1) generally consist of public rather than expert comments on severe weather and 2) vary rather substantially—in terms of frequency—over time. These findings are consistent with our contention that tornado-related Twitter activity

²The collection of these data and the protection of the user identity are undertaken using a protocol approved by the University of Oklahoma Internal Review Board for the protection of human subjects.

³Note that the “blank” spots in Fig. 1 represent days when our program was unable to collect and archive Twitter data because of errors in the data collection process. In most instances, these errors were caused by interruptions in our access to Twitter’s API. In the analysis that follows, data associated with these days are coded as missing (NA).

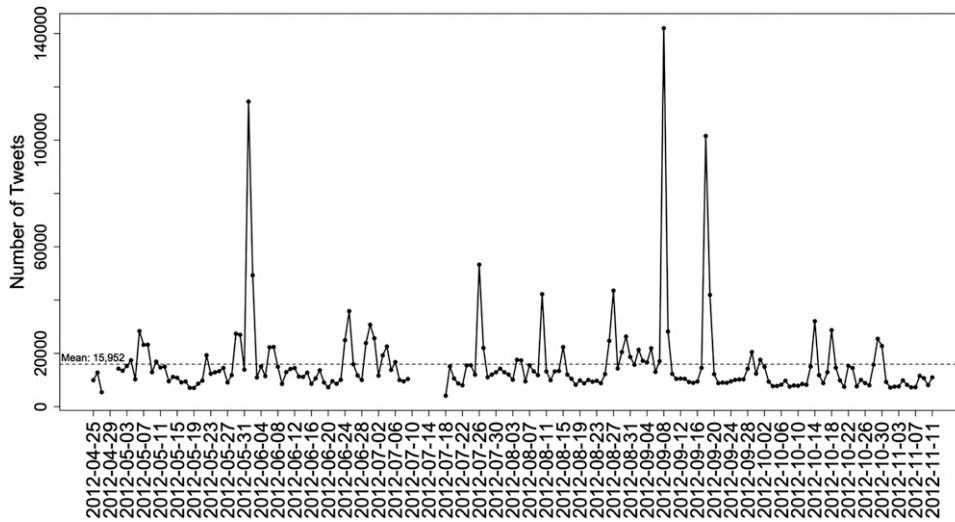


FIG. 1. The number of daily tweets containing the word “tornado” published between 25 April and 11 November 2012.

may provide a useful indicator of public attention to severe weather. But does it provide a valid indicator of public attention to severe weather communication?

4. Research design and analysis

To answer questions of this kind, researchers generally rely on a battery of validity tests that gauge the extent to which a newly proposed indicator accurately measures the concept that the analyst is attempting to capture. The easiest and best type of validity test simply compares the indicator against observations of the underlying concept. A new indicator is valid if it accurately measures the concept. Unfortunately, social scientists work in a world where it is difficult if not impossible to collect objective observations about an underlying concept. Doing so would require that we “get into the heads” of individual people and document whether or not they are paying attention to a given message at a given point in time. Given the impractical nature of this task, social scientists almost always conduct indirect assessments of validity by comparing a new measure against observable phenomena that are theoretically related to the latent construct that one is trying to measure. If the

new measure is systematically related to the phenomena in a way that is consistent with the theory, one can reasonably conclude that the new measure is valid.

In this project, we assess the validity of the indicator we have proposed by systematically comparing “tweets” about tornadoes to two types of risk communication—tornado watches and tornado warnings—that are designed to provoke public attention to this sort of severe weather. We accomplish this by comparing the number of watches and warnings issued—and the number of people impacted by those watches and warnings on a given day—to the number of tweets containing the word “tornado” that were published on that day. If tweets provide a valid indicator of public attention to severe weather communication, then daily tweet counts will be positively associated with the number of warnings/watches issued, as well as with the number of people affected by those watches and warnings. Moreover, if our measure is valid, then tornado-related Twitter activity should be sensitive to the differential severity and urgency associated with watches and warnings. Relative to watches, tornado warnings communicate greater risk, which means that they should theoretically stimulate higher levels of attention. Table 1 provides descriptive

TABLE 1. Descriptive statistics.

	Min	Mean	Median	Max	Std dev
Number of tweets	4115	15 952.21	12 061.50	141 975	15 383.66
Number of watches	0	0.44	0	7	0.98
Number of warnings	0	5.50	2	56	8.96
Population affected by watches (in millions)	0	1.76	0	55.79	5.69
Population affected by warnings (in millions)	0	0.26	0.03	8.71	0.85

statistics for number of tweets, watches, and warnings as well as affected population.

a. Analytical procedure

To compare the number of tweets posted on a given day against the number of tornado watches and warnings that were issued and the number of people affected, we compiled a database containing daily measures of the following variables: number of tweets containing the word “tornado,” number of tornado watches issued by the Storm Prediction Center (total $n = 89$ watches), and number of tornado warnings issued by storm-based polygon (total $n = 1105$ warnings). (Data on tornado warnings were collected using the NWS Watch/Warning Archive compiled and maintained by the Iowa Environmental Mesonet, which is accessible via <http://mesonet.agron.iastate.edu/request/gis/watchwarn.phtml>.) Then we estimated the total number of people affected per watch and warning issued each day. These estimates are relatively easy to calculate for tornado watches—we simply matched each county that was included in a watch to the most recent (2011) county-level population estimates provided by the U.S. Census Bureau and then aggregated by day (total $n = 354\,697\,498$ people; census data available at http://quickfacts.census.gov/qfd/download_data.html). Storm-based warnings, by comparison, do not align with political boundaries (i.e., counties), which made it more difficult to estimate the number of people affected by a given warning. Nevertheless, we proceeded by overlaying storm-based warning polygons on to a 2.5-arc-min grid that contains population estimates (via the 2010 U.S. Census) in each cell. (The grid we used was produced and distributed by the Columbia University Center for International Earth Science Information Network and is available at <http://sedac.ciesin.columbia.edu/data/set/gpw-v3-population-count-future-estimates>.) If a cell in the grid was inside or on the line of a warning polygon, then the population estimate associated with that cell was extracted and added to the population estimates associated with other cells that were inside or on the line of the same polygon. After repeating this process for each warning, we aggregated by day to estimate the total number of people affected per warning on each day that we included in our analysis (total $n = 51\,816\,954$ people). To facilitate interpretation, we scaled these estimates by dividing both “population affected” variables by 1 000 000.

After collecting, calculating, and compiling these data, we used them to estimate two sets of negative binomial regression models. The first set of models regress daily tweet counts on the number of watches and warnings issued per day between 25 April 2012 and 11 November 2012. If tweets provide a valid indicator of

public attention to severe weather risk communication, we should find that these signals exert a strong and positive effect on Twitter activity. On days when there were a large number of tornado watches and/or warnings issued, we should see a correspondingly high number of tweets. On days when there were fewer tornado watches and/or warnings, by comparison, Twitter traffic should be less pronounced. The second series of models predict daily tweet counts as a function of the number of people affected by watches and warnings each day (in millions). Again, if Twitter activity provides a valid indicator of public attention to messages about severe weather, then the number of people affected by watches and warnings will exhibit a strong and positive relationship with tornado-related Twitter activity; on days when a large number of people were impacted by watches and/or warnings, the number of tornado-related tweets should be high, relative to the days when fewer people were affected.⁴

We used negative binomial regression to estimate these models because the outcome variable (tweets per day) is a zero-censored count variable. When modeling variables of this sort, analysts generally use Poisson regression if the conditional mean and variance are roughly equivalent (equidispersion) and negative binomial regression if the model exhibits signs of overdispersion (Greene 2011). In this case, our models exhibited signs of overdispersion⁵ so we used the following formulation of negative binomial regression to estimate our models:

$$P(y \mid \mu[x_i], \theta) = \frac{\Gamma(y_i + \theta)}{y! \Gamma(\theta)} \left(\frac{\theta}{\theta + \mu[x_i]} \right)^\theta \left(\frac{\mu[x_i]}{\theta + \mu[x_i]} \right)^y,$$

where y is the outcome variable (tweets per day), μ is the conditional mean, θ is the dispersion parameter, and Γ represents the gamma function from which θ is drawn. Consistent with standard practice, we fit these models using maximum likelihood estimation (MLE) that, in addition to the conditional mean, estimated the dispersion parameter (θ) used to fit each model. Each model converged upon a solution after a single iteration.

⁴ Like most regression techniques, negative binomial regression models are sensitive to autocorrelation and nonlinearity. To ensure that our models were robust to these sensitivities, we specified a number of alternative models that included 1- and 2-day distributed lag terms and a variety of transformed and polynomial terms. These additions did not improve the fit of the models, which suggests that they are robust to these sensitivities.

⁵ We used the regression-based tests outlined by Cameron and Trivedi (1990) to test for overdispersion.

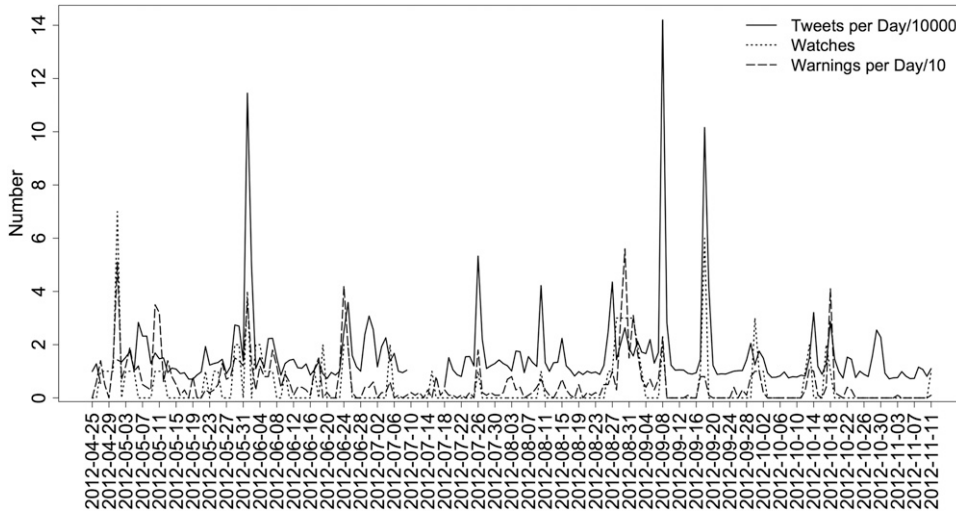


FIG. 2. A temporal comparison of daily tweet counts and the number of tornado watches issued and tornado warnings issued.

b. Findings

Figure 2 plots daily watch, warning, and tweet counts against one another over time. This figure reveals preliminary support for the supposition that daily tweet counts correspond with the number of tornado watches and warnings issued on a given day.

Unfortunately, visual inspection can be deceiving—the human eye has a tendency to seek out and therefore overestimate the occurrence of patterns within a given set of data. As such, we subjected these relationships to the more rigorous set of statistical tests described above. The results are summarized in the first three columns of Table 2. Model 1 provides results of a regression of daily tweet counts on daily watch counts, whereas model 2

presents results of a regression of tweet counts on warning counts and model 3 contains results of a regression of tweet counts on both watch and warning counts, so as to estimate the independent effect of warnings on Twitter activity when controlling for the number of watches issued on a given day (and vice versa).

As expected, the positive and statistically significant coefficients listed in models 1, 2, and 3 indicate that increases in the number of watches and warnings issued on a given day are independently and jointly associated with increases in the number of tornado-related tweets posted on that day. When both parameters are included in the model (model 3), each watch issued on a given day increased tweet rates by a factor of 1.19 (19.46%), whereas each warning increased tweet rates by a factor

TABLE 2. Negative binomial regression models predicting daily tweet counts. Values represent the coefficients and standard errors (in parentheses) derived from six different negative binomial regression models that were fitted using MLE; the outcome variable in all six models is tweets per day and the predictor variables are listed on the left side of the table. AIC = Akaike information criterion; BIC = Bayesian information criterion.

Predictor	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Number of watches	0.310* (0.035)		0.178* (0.044)			
Number of warnings		0.038* (0.004)	0.024* (0.005)			
Population affected by watches (in millions)				0.060* (0.005)		0.031* (0.006)
Population affected by warnings (in millions)					0.429* (0.037)	0.271* (0.040)
Intercept	9.481* (0.039)	9.405* (0.041)	9.392* (0.039)	9.485* (0.033)	9.486* (0.034)	9.450* (0.031)
AIC	3894.091	3894.558	3874.598	3851.371	3859.893	3818.649
BIC	3903.832	3904.299	3887.586	3861.112	3869.634	3831.637
Dispersion parameter (θ)	4.248	4.238	4.720	5.241	5.025	6.230
Number of cases	190	190	190	190	190	190

* p value < 0.001.

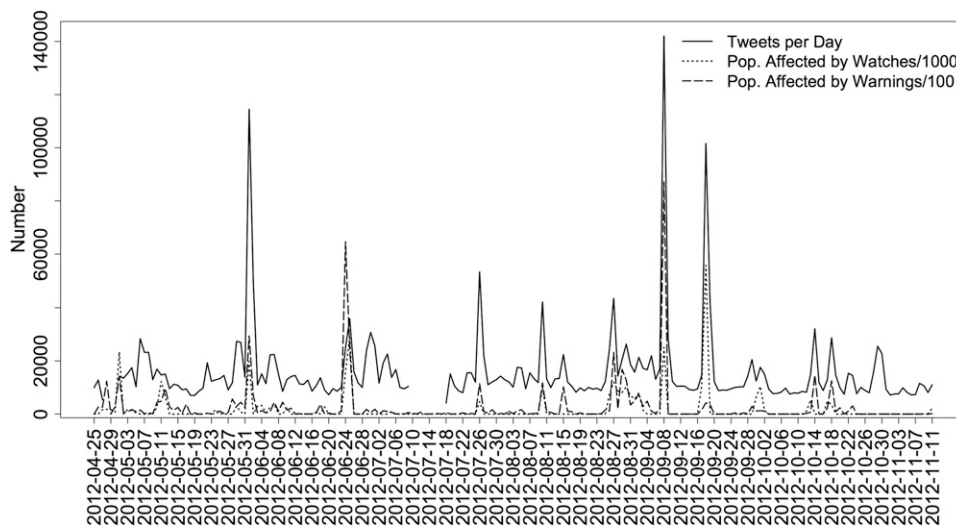


FIG. 3. A temporal comparison of daily tweet counts and the number of people affected by tornado watches and tornado warnings.

of 1.02 (2.43%).⁶ To add some perspective, model 3 predicts that Twitter users will publish approximately 11 992 tweets containing the word “tornado” on “non-active” days where no tornado watches and no tornado warnings are issued; 16 151 tweets on a “moderately active” day where a single watch and five warnings are issued; and 27 262 tweets on “highly active” days, where three watches are issued by the Storm Prediction Center and 12 warnings are issued by NWS forecast offices.

Having analyzed the relationship between daily watch, warning, and tweet counts, we turn now to the number of people affected by watches and warnings each day. Again, Fig. 3 lends preliminary but insufficient support for the notion that the number of people affected by watches and warnings on a given day will—if tweets provide a valid indicator of public attention—correspond with the number of tornado-related tweets that day.

To further analyze these patterns, we subjected the relationships depicted in Fig. 3 to the statistical comparisons summarized in models 4, 5, and 6 of Table 2. Model 4 presents results of a regression of daily tweet counts on the estimated number of people affected by watches each day (in millions), model 5 provides results of a regression of tweet counts on the number of people affected by warnings (in millions), and model 6 contains results of a regression of tweet counts on the estimated number of people affected by watches and warnings (in millions).

⁶ Incident rate ratios were calculated by exponentiating the coefficients in the model.

Again, the positive and statistically significant coefficient estimates listed in models 4, 5, and 6 indicate that increases in the number of people affected by watches and warnings are independently and jointly associated with increases in tornado-related tweets. Each unit increase—corresponding to one million people—in the number of people affected by tornado watches in a given day increases tweet rates by a factor of 1.03 (~3.2%). The same unit increase in the number of people affected by warnings increases tornado-related Twitter activity by a factor of 1.31 (~31.9%). To provide some perspective on the substantive magnitude of this difference, model 6 predicts that approximately 13 527 tweets will be published on an average “watch only day” where 2 000 000 people are affected by tornado watches and no one is affected by tornado warnings. This number increases to 21 872 tweets on an average “warning only day” where the same number of people are affected by warnings and no one is affected by watches. This difference is unsurprising given the differential urgency associated with tornado watches and warnings. Relative to watches, tornado warnings communicate greater risk, which should—and do—correspond with higher levels of public attention.

c. Follow-up analysis and findings

The preceding analysis demonstrates that the number of tweets containing the word “tornado” in a given day increases with the number of tornado watches and warnings issued on that day and with the number of people affected by those forms of risk communications; it also makes clear that the relationship between Twitter activity and tornado warnings is more pronounced than the

relationship between tweets and tornado watches, a less urgent form of risk communication. These findings are consistent with our contention that these messages can be used to monitor public attention to severe weather communication. However, the preceding analysis does not rule out the possibility that the spikes in Twitter activity were caused by tornadoes themselves rather than the watches and warnings that were generally issued before the tornadoes occurred. Thus, it could be that tornado-related Twitter activity provides a valid measure of public attention to the hazard itself, not the risk communication that precedes it. To assess this possibility, we perform two additional sets of analyses.

In the first set, we estimate two negative binomial regression models that expand upon the models 3 and 6 in Table 2 by statistically controlling for the occurrence of tornadoes. We accomplish this by 1) adding an indicator of daily tornado count to the model presented in model 3, which (as specified) models daily tweet count as a function of the number of tornado watches that were issued on that day and the number of warnings issued, and 2) adding a measure that documents the number of people affected by tornadoes on a given day (in millions) to the model presented in model 6, which (as specified) predicts daily tweet count as a function of the number of affected by tornado watches and warnings each day. (Data on tornadoes were collected using the Severe Weather Database compiled and maintained by the Storm Prediction Center, which is accessible via www.spc.noaa.gov/wcm/#data.) To approximate the number of people affected by tornadoes on a given day, we began by overlaying tornado segments onto the same 2.5-arc-min population grid that we used to approximate the number of people affected by storm-based warnings. If a tornado segment intersected a cell in the grid, then the population estimate associated with that cell was extracted and added to the population estimates associated with other cells that the tornado segment intercepted. After repeating this process for each tornado, we aggregated by day and then divided that number by 1 000 000 to estimate the total number of people affected by tornadoes on each day (in millions) that we included in our analysis (total $n = 1\,553\,444$ people). If tornadoes, rather than watches or warnings, were responsible for the spike in tweets on days where watches and warnings were issued, then the coefficients associated with tornadoes will be positive and statistically significant while the watch and warning coefficients would be rendered insignificant. The reverse would be true if watches and/or warnings, as opposed to tornadoes, were responsible for the increases in tornado-related Twitter activity. Finally, it could be the case that the three variables—watches, warnings, and tornadoes—prompted independent

levels of Twitter activity. It could be, for example, that watches elicit a nominal level of activity, warnings prompt significantly more, and then tornadoes prompt some number of tweets beyond that attributable to the prestorm communication. If this is the case, then the coefficients associated with all three sets of variables will be positive and statistically significant.

“Controlling” for the occurrence of tornadoes represents a statistical approach to disentangling the causal relationship between tweets, watches, warnings, and tornadoes. Though useful, a compelling research design is often more persuasive than a statistical solution to problems of this sort. Thus, the second set of follow-up analyses represent research design solutions that leverage the fact that a substantial portion of tornado watches and warnings are false alarms—they are issued but not accompanied by the occurrence of a tornado. For our purposes, these days are significant because we can be reasonably confident that tornado-related tweets that were published on these days were prompted by something other than tornadoes, such as the communication that preceded them.

To explore this possibility, the second set of follow-up analyses re-estimate models 3 and 6 in Table 2; this time, however, we restrict our analytic sample to two types of days—those on which there were no tornado watches, tornado warnings, or tornadoes ($n = 50$) and those on which at least one warning was issued but no watches were issued and no tornadoes occurred ($n = 36$). Doing so allows us to isolate the effect of tornado warnings on Twitter activity by removing the potentially confounding influence of watches and tornadoes. Positive and statistically significant relationships between daily tornado warning counts, the number of people affected by those warnings, and Twitter activity would provide strong evidence that warnings—not watches or tornadoes—are responsible for increases in the number of tweets published that contain the word “tornado.” In theory, one could use a similar procedure to estimate the independent effects of watches on daily tweet counts. Unfortunately, the number of days on which at least one tornado watch was issued but no warnings were issued and no tornadoes occurred is too small to produce a reliable estimate ($n = 1$).

Models 7 and 8 in Table 3 present the results of our first set of follow-up analyses. Mirroring the results presented in Table 2, the coefficients associated with the watch and warning variables are positive and statistically significant in both models. Again, however, the warning coefficients are quite a bit larger than the watch coefficients, suggesting that the effect of watches on the number of tornado-related tweets that were published each day was less pronounced than the effect of warnings.

TABLE 3. Negative binomial regression models predicting daily tweet counts. Values represent the coefficients and standard errors (in parentheses) derived from four different negative binomial regression models that were fitted using MLE; the outcome variable in all six models is tweets per day and the predictor variables are listed on the left side of the table.

Predictor	Model 7	Model 8	Model 9	Model 10
Number of watches	0.184* (0.045)			
Number of warnings	0.027* (0.005)		0.061* (0.016)	
Number of tornadoes	-0.012 (0.012)			
Population affected by watches (in millions)		0.031* (0.006)		
Population affected by warnings (in millions)		0.243** (0.074)		1.136* (0.281)
Population affected by tornadoes (in millions)		0.746 (1.606)		
Intercept	9.394* (0.040)	9.451* (0.031)	9.310* (0.045)	9.327* (0.043)
AIC	3875.728	3820.411	1646.103	1644.408
BIC	3891.963	3836.646	1653.396	1651.701
Dispersion parameter (θ)	4.740	6.237	7.411	7.556
Number of cases	190	190	84	84

* p value < 0.001.

In contrast, neither the count nor the population coefficients associated with the newly added tornado variables are statistically different from zero, suggesting that tornadoes themselves do not produce increases in Twitter activity above and beyond those attributable to the watches and warnings that, in most cases, preceded them.⁷

Models 9 and 10 in Table 3 present the results of our second set of follow-up analyses. Consistent with previous results, model 9 reveals a strong, positive, and statistically significant relationship between daily warning and tweet counts. Model 10 demonstrates a similar relationship—tweets containing the word “tornado” were more common on days when a large number of people were affected by warnings. Again, the fact that these findings hold absent the occurrence of a tornado provides strong evidence that tornado-related Twitter activity is prompted by warnings, not tornadoes.

When considered in tandem, these results suggest that tornado watches and warnings—two important forms of risk communication—produce a measurable increase in the number of tweets regarding tornadoes that is independent of tornadoes, the hazard they were designed to proceed. In so doing, they reinforce our contention

that social media in general, and Twitter in particular, can be used to monitor public attention to severe weather communication.

5. Discussion and conclusions

Severe weather occurs in a social environment wherein effective communication between providers of hazardous weather information (i.e., forecasters) and the public can mean the difference between life and death. At its core, effective communication requires that two conditions be met: 1) that providers of weather information disseminate their message (i.e., issue a tornado watch or warning) and 2) that the intended recipients of their message (i.e., the population threatened by the convective conditions) receive, pay attention to, and understand it. Though all three parts of the second condition are important, extant research generally focuses on receipt and comprehension, while neglecting attention. This is unfortunate because attention mediates the relationship between information and action. If people are not paying attention when a tornado watch or warning is issued, for example, it is unlikely that they will deliberately engage in the sort of action necessary to protect themselves if a tornado were to occur (i.e., seek shelter).

This neglect of public attention to messages about severe weather likely stems from a variety of factors, some of which involve data limitations. Simply put, we know relatively little about public attention to severe weather communication because researchers in the field have yet to develop an adequate measure of the concept. Motivated by this limitation, we drew upon recent research in other fields to propose and develop a new

⁷ As suggested by Gelman and Hill (2006), we used analysis of variance (deviance) to confirm this “null” finding. As expected, adding the tornado parameters to models 7 and 8 did not produce a statistically significant improvement in model fit. It is possible, however, that our sample size ($n = 190$) is too small to detect an independent and significant effect for tornadoes, which are (by design) correlated with the other variables in the model (warnings and watches). As such, readers should interpret this null finding for tornadoes with some caution.

indicator of public attention to messages about one type of severe weather (tornadoes) based on public usage of one type of social media (Twitter). After introducing this measure, we subjected it to a series of tests and found that the new indicator is systematically related to a specific type of communication—tornado watches and warnings—in a manner that is consistent with our limited understanding of the concept. Tornado-related tweets are relatively frequent on days when a large number of tornado watches are issued (compared to days where few, if any, warnings or watches are issued) but even more frequent on days when a relatively large number of warnings are issued. A similar pattern characterizes the temporal relationship between Twitter activity and the number of people affected by warnings and/or watches—on days when a large number of people are included in watches, the number of tornado-related tweets is high (relative to the days where few, if any, people are affected), but it is even higher on days when a large number of people are impacted by tornado warnings. On top of this, the evidence we presented suggests daily Twitter activity is more responsive to tornado watches and warnings than tornadoes themselves. Again, this is precisely what we would expect of an indicator that purports to measure public attention to severe weather communication, not the weather itself.

Accordingly, we believe that these results provide preliminary but strong evidence that social media data can be used to develop valid indicators of real-time attention to severe weather communication that will afford future researchers an unprecedented opportunity to answer critical yet unstudied questions about the relationship between communication, attention, and public responsiveness to severe weather. For example, measures of this type might help us to answer important questions about the relative effectiveness of experimental communication strategies such as the “impact based” warning system that was recently implemented by NWS offices in parts of Kansas and Missouri. The logic motivating this product is rather straightforward—the public will pay more attention to a given warning if that warning indicates that the potential risk of the approaching storm exceeds some threshold. But is this system effective? Do people pay more attention to severe thunderstorm and tornado warnings when they contain high-threat language? One way to answer this question involves the construction and use of a measure similar to the one we employed in this paper. If the experimental warning system is effective, then public attention to tornado warnings—as indicated by social media trends in the regions that have employed the new system—will vary as a function of the threat tag assigned to warning issued.

This example represents one of the many questions that real-time indicators of public attention to severe weather communication will allow researchers to explore in the near future. This paper uses social media to develop and validate a crude version one such indicator. Our hope is that our discussion and this indicator—in tandem with other indicators—will provide analysts with the inspiration and tools necessary to go forth and develop better, more refined, indicators. Among other things, we suggest that researchers interested in this pursuit work to develop a system that is capable of separating the signal (tweets from attentive members of the public) from the noise (everything else) in near-real time so that providers of severe weather information can meaningfully track the effect of their messages on public attention as they disseminate them. Accomplishing this will require careful consideration of the dynamics within social media communities. On Twitter, for example, messages occasionally go “viral”—they are transmitted and retransmitted (retweeted) many times in a short period of time. This happens for a number of reasons, most of which are related to the content of the message and/or the centrality of the user who posted the original message within an expansive and active network of other users (Hong et al. 2011; Hansen et al. 2011; Hoang et al. 2011). The development of a real-time system capable of separating the signal from the noise will have to take these and other dynamics into account.

Future research, development, and refinement notwithstanding, we hope that our research showcases the benefits of empirically measuring and studying public attention to severe weather risk communication.

Acknowledgments. Funding for this project was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreements NA11OAR4320072 and NA12OAR4590120, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author (s) and do not necessarily reflect the views of NOAA or the U.S. Department of Commerce.

REFERENCES

- Aguirre, B. E., 1988: The lack of warnings before the Saragosa tornado. *Int. J. Mass Emerg. Disasters*, **6**, 65–74.
- Ahlborn, L., J. M. Franc, and D. S. Med, 2012: Tornado hazard communication disparities among Spanish-speaking individuals in an English-speaking community. *Prehosp. Disaster Med.*, **27**, 98–102, doi:10.1017/S1049023X12000015.
- Brooks, H. E., and C. A. Doswell, 2002: Deaths in the 3 May 1999 Oklahoma City tornado from a historical perspective. *Wea. Forecasting*, **17**, 354–361, doi:10.1175/1520-0434(2002)017<0354:DITMOC>2.0.CO;2.

- Cameron, A. C., and P. K. Trivedi, 1990: Regression-based tests for overdispersion in the Poisson model. *J. Econometrics*, **46**, 347–364.
- Chan, E. H., S. Vikram, C. Conrad, and J. S. Brownstein, 2011: Using web search query data to monitor dengue epidemics: A new model for neglected tropical disease surveillance. *PLoS Neglected Trop. Dis.*, **5**, e1206, doi:10.1371/journal.pntd.0001206.
- Chew, C., and G. Eysenbach, 2010: Pandemics in the age of Twitter: Content analysis of tweets during the 2009 H1N1 outbreak. *PLoS ONE*, **5**, e14118, doi:10.1371/journal.pone.0014118.
- Donner, W. R., H. Rodriguez, and W. Diaz, 2012: Tornado warnings in three southern states: A qualitative analysis of public response patterns. *J. Homeland Security Emerg. Manage.*, **9**, doi:10.1515/1547-7355.1955.
- Doswell, C. A., A. R. Moller, and H. E. Brooks, 1999: Storm spotting and public awareness since the first tornado forecasts of 1948. *Wea. Forecasting*, **14**, 544–557, doi:10.1175/1520-0434(1999)014<0544:SSAPAS>2.0.CO;2.
- Gelman, A., and J. Hill, 2006: *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, 625 pp.
- Ginsberg, J., M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant, 2009: Detecting influenza epidemics using search engine query data. *Nature*, **457**, 1012–1014, doi:10.1038/nature07634.
- Greene, W. H., 2011: *Econometric Analysis*. 7th ed. Prentice Hall, 1232 pp.
- Hansen, L. K., A. Arvidsson, F. A. Nielsen, E. Colleoni, and M. Etter, 2011: Good friends, bad news—Affect and virality in Twitter. *Future Info. Technol.*, **185**, 34–43, doi:10.1007/978-3-642-22309-9_5.
- Hoang, T.-A., E.-P. Lim, P. Achananuparp, J. Jiang, and F. Zhu, 2011: On modeling virality of Twitter content. *Digital Libraries: For Cultural Heritage, Knowledge Dissemination, and Future Creation*, Springer, 212–221, doi:10.1007/978-3-642-24826-9_27.
- Hong, L., O. Dan, and B. D. Davison, 2011: Predicting popular messages in twitter. *Proc. 20th Int. Conf. on the World Wide Web*, Hyderabad, India, Association for Computing Machinery, 57–58, doi:10.1145/1963192.1963222.
- Joslyn, S., and S. Savelli, 2010: Communicating forecast uncertainty: Public perception of weather forecast uncertainty. *Meteor. Appl.*, **17**, 180–195, doi:10.1002/met.190.
- Lampos, V., and N. Cristianini, 2010: Tracking the flu pandemic by monitoring the social web. *Second Int. Workshop on Cognitive Information Processing*, Elba, Italy, IEEE, 411–416, doi:10.1109/CIP.2010.5604088.
- Lindell, M. K., and R. W. Perry, 2012: The Protective Action Decision Model: Theoretical modifications and additional evidence. *Risk Anal.*, **32**, 616–632, doi:10.1111/j.1539-6924.2011.01647.x.
- Morss, R. E., J. L. Demuth, and J. K. Lazo, 2008: Communicating uncertainty in weather forecasts: A survey of the U.S. public. *Wea. Forecasting*, **23**, 974–991, doi:10.1175/2008WAF2007088.1.
- NOAA/NWS Storm Prediction Center, 2012: United States tornadoes of 2011. [Available online at www.spc.noaa.gov/wcm/2011-NOAA-NWS-tornado-facts.pdf.]
- Ripberger, J. T., 2011: Capturing curiosity: Using internet search trends to measure public attentiveness. *Policy Stud. J.*, **39**, 239–259, doi:10.1111/j.1541-0072.2011.00406.x.
- Sakaki, T., M. Okazaki, and Y. Matsuo, 2010: Earthquake shakes Twitter users: Real-time event detection by social sensors. *Proc. 19th Int. Conf. on World Wide Web*, New York, NY, ACM, 851–860, doi:10.1145/1772690.1772777.
- Scharkow, M., and J. Vogelgesang, 2011: Measuring the public agenda using search engine queries. *Int. J. Public Opin. Res.*, **23**, 104–113, doi:10.1093/ijpor/edq048.
- Schmidlin, T. W., P. S. King, B. O. Hammer, and Y. Ono, 1998: Risk factors for death in the 22–23 February 1998 Florida tornadoes. Quick Response Rep. 106. [Available online at www.colorado.edu/hazards/research/qr/qr106/qr106.html.]
- Sorensen, J., 2000: Hazard warning systems: Review of 20 years of progress. *Nat. Hazards Rev.*, **1**, 119–125, doi:10.1061/(ASCE)1527-6988(2000)1:2(119).
- Szomszor, M., P. Kostkova, and E. de Quincey, 2012: #Swineflu: Twitter predicts swine flu outbreak in 2009. *Electronic Healthcare*, M. Szomszor and P. Kostkova, Eds., *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Vol. 69, Springer, 18–26, doi:10.1007/978-3-642-23635-8_3.
- Tumasjan, A., T. O. Sprenger, P. G. Sandner, and I. M. Welpe, 2010: Predicting elections with Twitter: What 140 characters reveal about political sentiment. *Proc. Fourth Int. AAAI Conf. on Weblogs and Social Media*, Washington, DC, AAAI, 178–185. [Available online at www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewFile/1441/1852.]
- Wilson, K., and J. S. Brownstein, 2009: Early detection of disease outbreaks using the internet. *Can. Med. Assoc. J.*, **180**, 829, doi:10.1503/cmaj.1090215.