SUPPLEMENTAL MATERIALS

ASCE *Natural Hazards Review*

Information Dissemination, Diffusion, and Response during Hurricane Harvey: Analysis of Evolving Forecast and Warning Imagery Posted Online

Rebecca E. Morss, Robert Prestley, Melissa Bica,
and Julie L. Demuth

DOI: 10.1061/NHREFO.NHENG-1802

www.ascelibrary.org

Section S1. Additional Information about Data Collection and Original Image Tweet and Time Filtering

As summarized in the main text, this study utilizes a data set described in Bica et al. (2019). To obtain their data, Bica and colleagues first developed a list of 796 Twitter accounts for authoritative sources of hurricane risk information. Next, they collected all tweets from these 796 accounts from 17 August 2017 (when Harvey became a named tropical storm) to 10 October 2017 (when Nate dissipated). They also collected all tweets in reply to these accounts and all retweets and quote tweets of these accounts' tweets during this period. They then filtered these data to all *original tweets* posted by these 796 accounts (in other words, tweets that were not replies to, retweets of, or quote tweets of others' tweets) containing at least one still image, video, or animated GIF.

We began our study with the same data analyzed in Bica et al. (2019). However, while analyzing the data we found that, during their filtering and coding process, Bica et al. inadvertently omitted some of the authoritative sources' original tweets with imagery. Thus, we collaborated with Bica and colleagues to refilter their initial 2017 Atlantic hurricane season data to obtain the full set of original tweets with imagery posted between 00:00 UTC 17 August 2017 (19:00 CDT 16 August) and 15:00 UTC 2 September 2017 (10:00 CDT 2 September). Start and end times for the period studied were selected based on when the NWS started and ended issuing Harvey-related advisories, forecasts, and warnings. We then restarted analysis using this time-filtered data set, shown in Figure 2.

Section S2. Additional Information about Authoritative Source Coding, Categorization, and Filtering

Each authoritative source account with a tweet in the data set was coded based on the source's role at the time of Hurricane Harvey, using the coding scheme in Prestley and Morss (2024). The initial coding scheme used for geographic area of responsibility had five groups. We then combined several of the groups to develop the more compact categorization described in the article: National, Local Harvey (also referred to as Local), or Other. The initial coding scheme for professional role had eight groups, and again several were combined for the more compact categorization described in the article: Weather Media, News Media, NWS, Non-NWS Government, or Other.

Accounts were coded as organizational if they had an official organizational affiliation, indicated by use of an organization name as their username, an organizational logo or symbol as their profile picture, and/or first-person plural or third-person constructions in their Twitter bio. All other accounts were coded as individual, indicated by their use of a person's name as their username, a personal image as their profile picture, and/or first-person singular constructions in their Twitter bio. For individual accounts, we coded area of responsibility and role based on the organizational affiliation indicated in their Twitter bio. If their affiliation had changed by the time of source coding or was unclear, we coded these using other information available on the Internet.

This study aims to study forecast and warning communication with people who are in areas at risk from a specific hurricane threat, Harvey. Given the different goals of the Bica et al. (2019) study, many of the 796 authoritative source accounts in their data set communicate primarily with populations that were not at risk from Harvey. For example, many of the sources were local to Florida or Puerto Rico, where Hurricanes Irma and Maria made U.S. landfall in 2017. In addition, many of the accounts were affiliated with news media sources who communicate about a variety of topics. As a result, even after filtering the data to tweets posted during the period when Harvey was a threat, the data set includes many tweets unrelated to Harvey or other hurricanes. Thus, as summarized in the main text, we used the geographic scope and affiliation coding to filter the source accounts to those most focused on communicating about Harvey's threat with people in areas at risk from the storm.

We decided which sources to retain using knowledge about hurricane risk and social media communication, combined with results from analysis conducted using the initial Bica et al. (2019) data set (before we identified the missing data discussed in Section S1). For example, we anticipated that National and Local Harvey sources would tweet more information about Harvey's threat relevant to people at risk than sources local to other regions. This was supported by our initial analysis, which found that National and Local Harvey accounts tweeted Harvey forecast and warning imagery more actively than accounts with other geographic scopes. Therefore, we decided to include only National and Local Harvey sources in this study.

We also anticipated that NWS and Weather Media sources would tweet more about Harvey forecast and warning information than News Media or Other sources, who would tweet more about other topics. The initial analysis supported this for National sources: National NWS and Weather Media accounts posted most of the Harvey forecast and warning tweets from National sources. In addition, National News Media, Non-NWS Government, and Other accounts posted few image tweets containing Harvey forecast and warning information. Therefore, we decided to remove these last three types of National sources from the data set. Among Local Harvey sources, however, we found that all account types actively tweeted Harvey forecast and warning imagery; thus, we retained all types of Local sources.

After source coding and filtering, we combined the geographic area and professional role codes into a single source categorization. We refer to the two National source types remaining in the data set as National NWS and National Weather Media. We refer to the five Local Harvey source categories as Local NWS, Local Weather Media, Local News Media, Local Non-NWS Government, and, since the three accounts coded as Local Harvey + Other were affiliated with Texas weather blogs, Local Weather Bloggers.

Section S3. Additional Information about Hurricane Risk Image and Tweet Content Coding and Filtering

Bica et al. (2019) coded the original tweets in their data set based on whether the tweets included hurricane risk images, defined as "including hurricane forecasts (e.g., the cone of uncertainty and ensemble or "spaghetti" models), observations of the hurricane (e.g., radar or satellite imagery) which inform risk assessment and forecasts, and evacuation information" (p. 1-2). They then filtered the data to only tweets containing such imagery. For tweets already coded by Bica et al., we retained this filtering. We then applied Bica et al.'s definition of hurricane risk images to code and filter the additional original tweets in our source-filtered data set (those inadvertently omitted in their study, described in Section S1). Combined, this yielded the hurricane risk image data set shown in Figure 2.

The tweets in the hurricane risk image data set were then coded for Harvey relevance and forecast/warning information. We tested the code definitions (described in the main text and provided in Prestley and Morss 2024) using the initial Bica et al. (2019) hurricane risk image data set filtered to the Harvey time period, before we identified the inadvertently omitted data. To test these coding definitions and calculate intercoder reliability, 2 researchers (RP and RM) independently coded 473 tweets (10% of the data set at that time), randomly selected.

While coding for Harvey relevance and forecast or warning information, we removed from the data set tweets that required interpretations of non-English language to make a coding judgment. We also coded each remaining tweet for language. Tweets were coded as Spanish if they included any content in an image or tweet text in Spanish (exclusively or along with English), and tweets that included only English language were coded as English. Inter-coder reliability for both language codes was excellent, with $\alpha > 0.95$. However, most of the Spanish tweets in the data set were posted by a few sources. Most of these tweets also contained Spanish language versions of specific NWS products, sometimes along with an English version of the same product. Since we could not disentangle the influence of Spanish language in this data set from the influence of source, image type, and Spanish alongside English, the language coding is not used in the article.

Section S4. Additional Information about Image Coding and Categorization

The image type and branding coding schemes used in the study are provided in Prestley and Morss (2024). For image type, the initial set of codes used is shown in Table S1. This included 14 codes representing visualizations that NWS and media sources commonly use to disseminate NWS graphical forecast and warning products and other hurricane risk information (see, e.g., Morss et al. 2022b). Other types of forecast, warning, and protective action imagery, including videos, were coded as Other Forecast. Because the data set included tweets with multiple images, as well as image tweets that had forecast or warning information only in the tweet text, some images did not contain any forecast or warning content; these were coded as Other Non-Forecast. This latter group contains primarily satellite, radar, and other storm-related imagery.

For branding, we coded each image based on whether it was an official NWS representation or not. Images were coded as NWS-branded if they either: 1) included a logo, symbol, or name of the NWS or one of its entities (e.g., the National Hurricane Center or a Weather Forecast Office), or 2) were a full or cropped version of an official NWS graphical or text product (same color scheme, symbolization, font, etc.). All other images were coded as Non-NWS-branded; these either: 1) included a logo, symbol, or name of an organization other than NWS, or 2) had no organizational branding (but were not images of official NWS products with the NWS logo cropped out). We did not consider the source account when coding for image branding. After coding, we found that few tweets contained multiple images with different branding. Thus, we conducted the analysis using a binary image branding scheme with mutually exclusive categories: a tweet was categorized as NWS-branded if any image was NWS-branded, and Non-NWS-branded otherwise.

As with the Harvey relevance and forecast/warning coding definitions, we tested and refined the image type and branding coding schemes using the initial Bica et al. (2019) hurricane risk image data set filtered to the Harvey time period, before identifying and filling in the missing data described in Section S1. First, 2 researchers (RP and RM) conducted 2 initial rounds of coding in which they independently coded 45 randomly selected tweets. After each round, they compared coding, discussed differences, and revised the coding schemes. The 2 researchers then independently coded an additional 135 randomly selected tweets. Intercoder reliability was assessed across the 225 tweets coded in the 3 test rounds, which represents 10% of the data set at that time. For branding, intercoder reliability was excellent, with $\alpha > 0.95$. For image type, $\alpha = 0.75-1.0$ for all categories that had data to assess intercoder reliability, as shown in Table S1. The two researchers adjudicated coding differences through discussion, and one researcher (RP) then coded the remainder of the data set.

Table S1. Set of initial image types used to test and refine the image type coding scheme and to assess intercoder reliability. N/A indicates that intercoder reliability could not be assessed for that image type because it did not appear in the data set used for testing and intercoder comparison.

Image type code in initial coding scheme	# of tweets in testing / intercoder comparison data set	Krippendorf's alpha
Watches and warnings	170	0.93
Cone of uncertainty	23	0.95
Tropical weather outlook	0	N/A
Spaghetti plots	3	0.93
Key messages	3	1.0
Rainfall forecasts	8	1.0
Excessive rainfall outlooks	0	N/A
River flood forecasts	2	1.0
Convective outlooks	10	0.97
Mesoscale discussions	2	1.0
Storm surge inundation	2	1.0
Arrival of tropical-storm-force winds	0	N/A
Probability of tropical-storm-force winds	1	0.75
Threat and impact	4	0.94
Other – forecast	18	0.82
Other – non-forecast	15	0.82

During the image coding process, we revised the initial image type coding scheme in two ways, described in the main text and shown in the left-hand column of Table S2. In addition, after completing the coding, we found that several of the image type codes were applied to fewer than 50 tweets in the data set. We retained one of these, Key Messages, in a separate category for analysis due to its embedded imagery and distinct diffusion characteristics. We placed tweets in the Evacuation/preparedness information code in either the Text or Other Forecast categories, depending on which definition fit their imagery. The remaining image type codes with few tweets were either combined with a related image type code or merged into the Other Forecast group, as shown in Table S2.

We also found that the data set included several commonly co-occurring image types, when a frequently used format of one image type included a smaller visual representation of another image type. One example is Cone images; many, especially those in the NWS format, contain embedded small depictions of NWS watches and warnings, i.e., Watch/Warning imagery (e.g., Figure 4a). A second example is Key Messages images, which typically contain two smaller embedded NWS forecast images (often a Cone image and another image type, e.g., Figure 5d). These common co-occurrences complicated interpreting the results, due to embedded image types inheriting the retweet counts of the parent image type.

For the analysis shown in the article, we therefore revised the image type categorization to be mutually exclusive. Images coded as two or more commonly co-occurring image types, due to embedded smaller image types as described above, were categorized according to the larger parent image type (e.g., Cone and Key Messages in the examples above). We placed tweets with two distinct types of imagery (other than Key Messages images) into a new Multiple category. Tweets with multiple images of the same type, e.g., two different visuals depicting Model Output, as shown in Figure 4c, were kept in that image type category.

Table S2. Set of revised image types used to code the full data set and their correspondence to the image categories used in the article after several image types were combined with other categories. Italicized codes were revised after the initial rounds of coding that used the scheme in Table S1: the Model output, Text product, and Evacuation/preparedness graphics codes were added, and the Watches and warnings code was subdivided into NWS Impact Watch/Warning and Watch/Warning. Code definitions are provided in Prestley and Morss (2024)

Image type code in revised coding scheme	**# of tweets in Harvey forecast and warning data set**	**Image type category in article**
NWS Watch and warning impact graphics	1215	NWS Impact Watch/Warning
Watches and warnings	927	Watch/Warning
Cone of uncertainty	401	Cone
Tropical weather outlook	57	Tropical Outlook
Spaghetti plots	39	Model Output
Computer model output	88	Model Output
Key messages	33	Key Messages
Rainfall forecasts	214	Rainfall
Excessive rainfall outlooks	28	Rainfall
River flood forecasts	126	River Flood
Convective outlooks	63	Convective
Mesoscale discussions	45	Convective
Text products	109	Text
Storm surge inundation	16	Other Forecast
Arrival of tropical-storm-force winds	13	Other Forecast
Probability of tropical-storm-force winds	11	Other Forecast
Threat and impact	31	Other Forecast
Evacuation/preparedness graphics	49	Other Forecast or Text
Other forecast	321	Other Forecast
Other non-forecast	303	Other Non-Forecast

The left column in Table S2 shows the revised image type coding scheme that we used to code the full data set, and the middle column shows how prevalent each of those codes was in the data set (prior to reorganization). The right column shows how the codes in the left column were reorganized into the image type categories used in the main text. The mutually exclusive categorization used in the article (shown in Table 2) contains 13 categories: the 12 categories in the right column of Table S2 and Multiple.

Section S5. Additional Information about Data Analysis

Figure S1 shows that as is typical for Twitter data, the distribution of retweets and replies is highly skewed, with many tweets having no or few retweets or replies. Thus, we used the median (rather than the mean) number of tweets as a summary statistic. Replies are also a metric of diffusion, attention, and response, but they are much less common than retweets in this data set. As a result, we observed limited variability in replies across tweet categories. Moreover, although the data set provided by Bica et al. (2019) contained the number of quote tweets associated with each original tweet, additional examination identified issues with these values. We therefore decided to focus the quantitative analyses shown in the article on retweets.

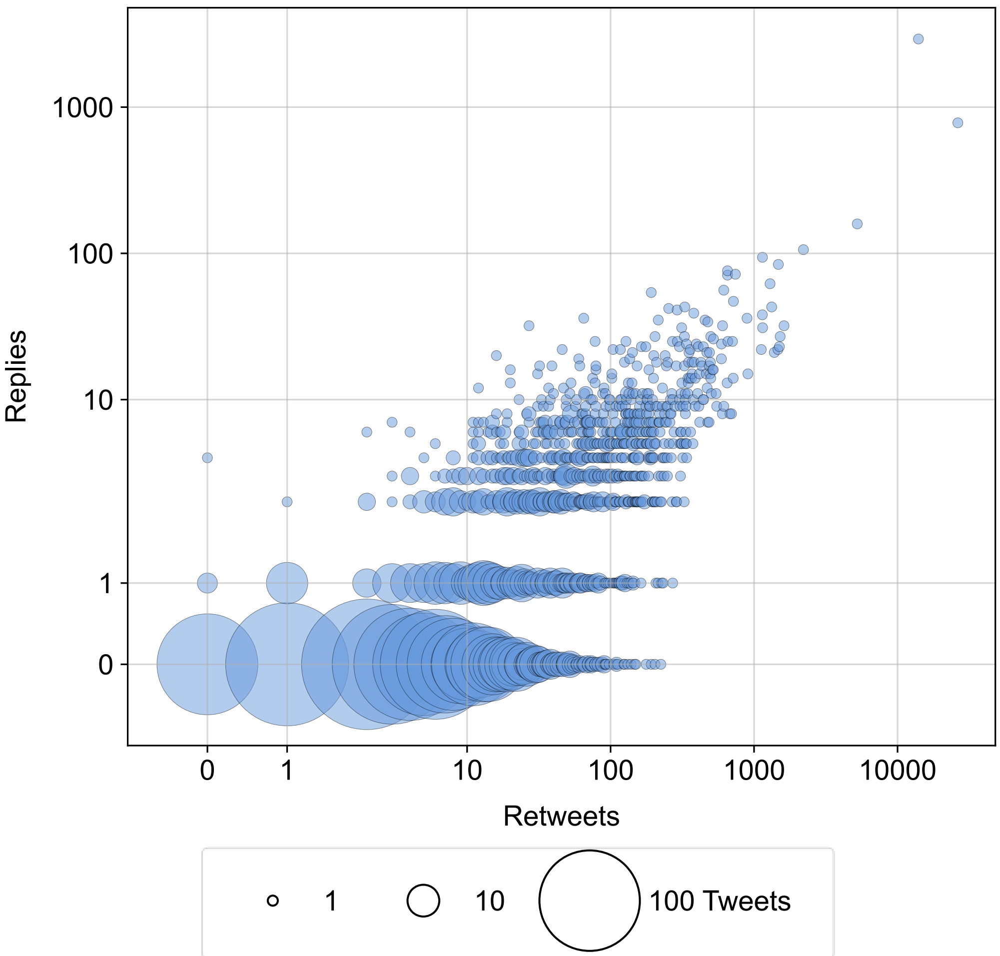

Figure S1. Scatter plot depicting the distribution of retweets and replies for the 3441 tweets in the forecast and warning data set, prior to removal of the 2 tweets with outlier diffusion (in the upper right corner). The size of each circle indicates the number of tweets in the data set that has the combination of retweets and replies indicated on the x and y axes, respectively. Note that both the x and y axis are on a logarithmic scale, with 0 added to show tweets with no retweets and/or replies.

As summarized in the main text, the data set includes two tweets with anomalous retweet and reply behavior. One of these outliers (uppermost circle in Figure S1) is a text image of an evacuation notice tweeted by @brazoriacounty (the official account of a coastal county near Houston, Texas) on 29 August. This tweet had 13,988 retweets and 2923 replies, more than 2.5

times as many retweets and 24 times as many replies as the most-diffused non-outlier tweet. As discussed in Bica et al. (2019), many of the replies to this tweet are not related to Harvey's threat; instead, they are about climate change or U.S. politics, including comments about the U.S. president at the time, Donald Trump (who retweeted the tweet). The other outlier (rightmost circle in Figure S1) is a Rainfall image tweeted by @nws. It had 26,323 retweets and 781 replies, more than 5 times as many retweets and 4 times as many replies as the most-diffused non-outlier tweet. Many of the replies to this tweet are also unrelated to Harvey's threat. This reply content indicates that much of these tweets' diffusion and response was unrelated to our research questions, which focus on forecast and warning communication with populations at risk. Thus, we removed them from the data set for subsequent quantitative analyses.