Science Advances

AAAS

advances.sciencemag.org/cgi/content/full/2/3/e1500779/DC1

Supplementary Materials for

Rapid assessment of disaster damage using social media activity

Yury Kryvasheyeu, Haohui Chen, Nick Obradovich, Esteban Moro, Pascal Van Hentenryck, James Fowler, Manuel Cebrian

Published 11 March 2016, *Sci. Adv.* **2**, e1500779 (2016)
DOI: 10.1126/sciadv.1500779

The PDF file includes:

Fig. S2. Originality of content, expressed through the fraction of retweets in the stream of messages.
Fig. S3. Global popularity of local content.
Fig. S4. Comparison of predictive capacity of activity and sentiment.
Fig. S5. Comparison of activity-damage correlation strength for different precision levels of geo-location.
Fig. S6. Average sentiment trends over time: Comparison between sentiment metrics.

Supplementary Materials.

Table S1 List of keywords included in the analysis, with their corresponding message counts.

Keyword	Count
power	4 825 717
sandy	4 745 099
hurricane	4 680 290
weather	3 333 025
storm	2 555 196
gas	1 991 524
Governor	498 135
stay safe	484 732
recovery	431 591
climate	420 217
FEMA	329 789
flooding	264 132
no power	261 998
climate change	236 009
wall st	233 411
blackout	213 520
mta	206 504
frankenstorm	205 467
Cuomo	92 014
prayforusa	91 293

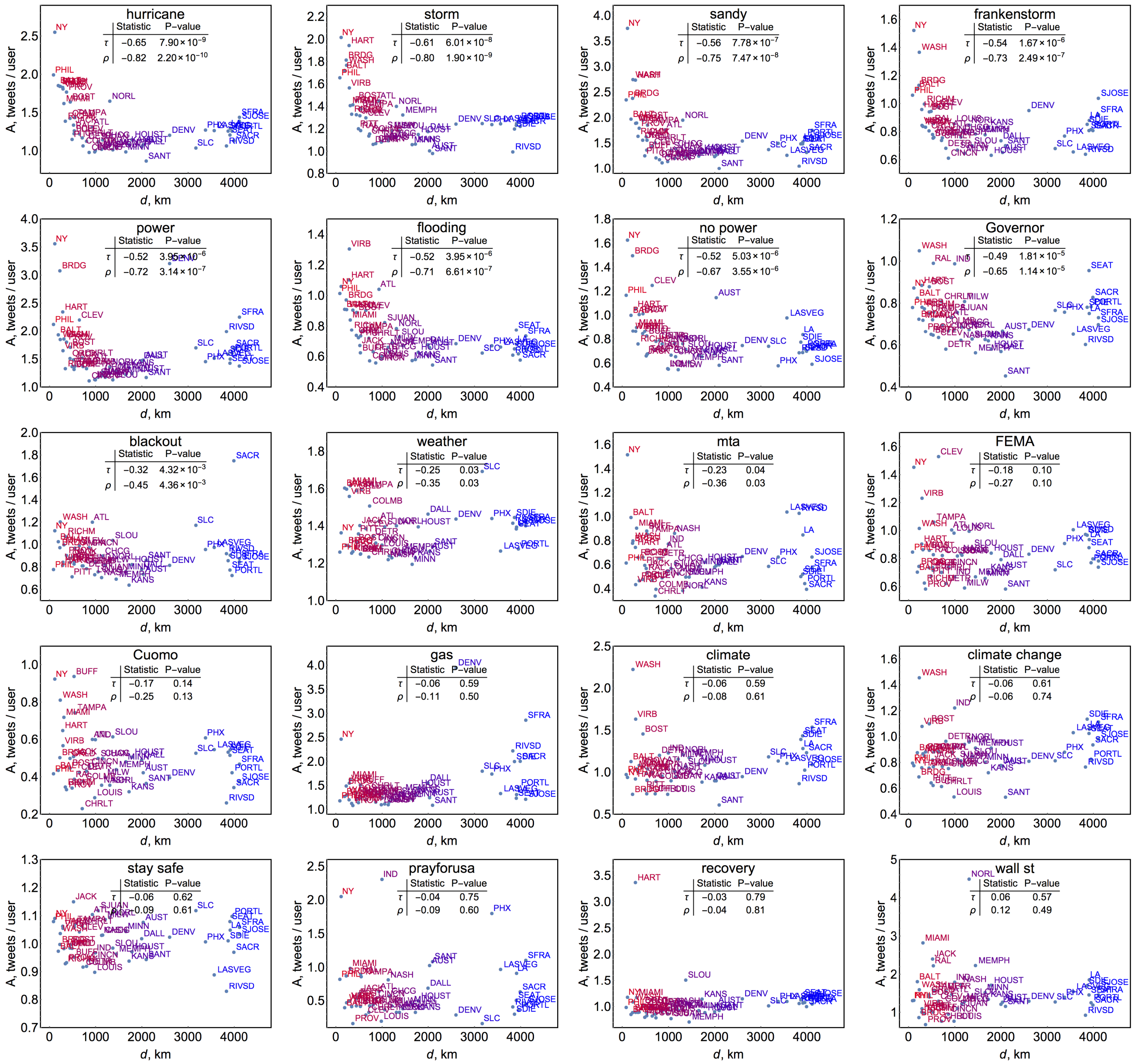

Figure S1 Normalized local activity on the topic as a function of distance to the hurricane path. For the words strongly related to Hurricane Sandy (top row) activity decreases with distance, and after the distance of 1200 – 1500 km proximity does not affect the level of activity. Because of this, and also because for some words (“gas”, “power”, “Governor”, etc.) trends for East and West coasts differ, we use rank correlation for East coast cities as a measure of relevance. The values of these correlation coefficients between activity and distance are shown in insets.

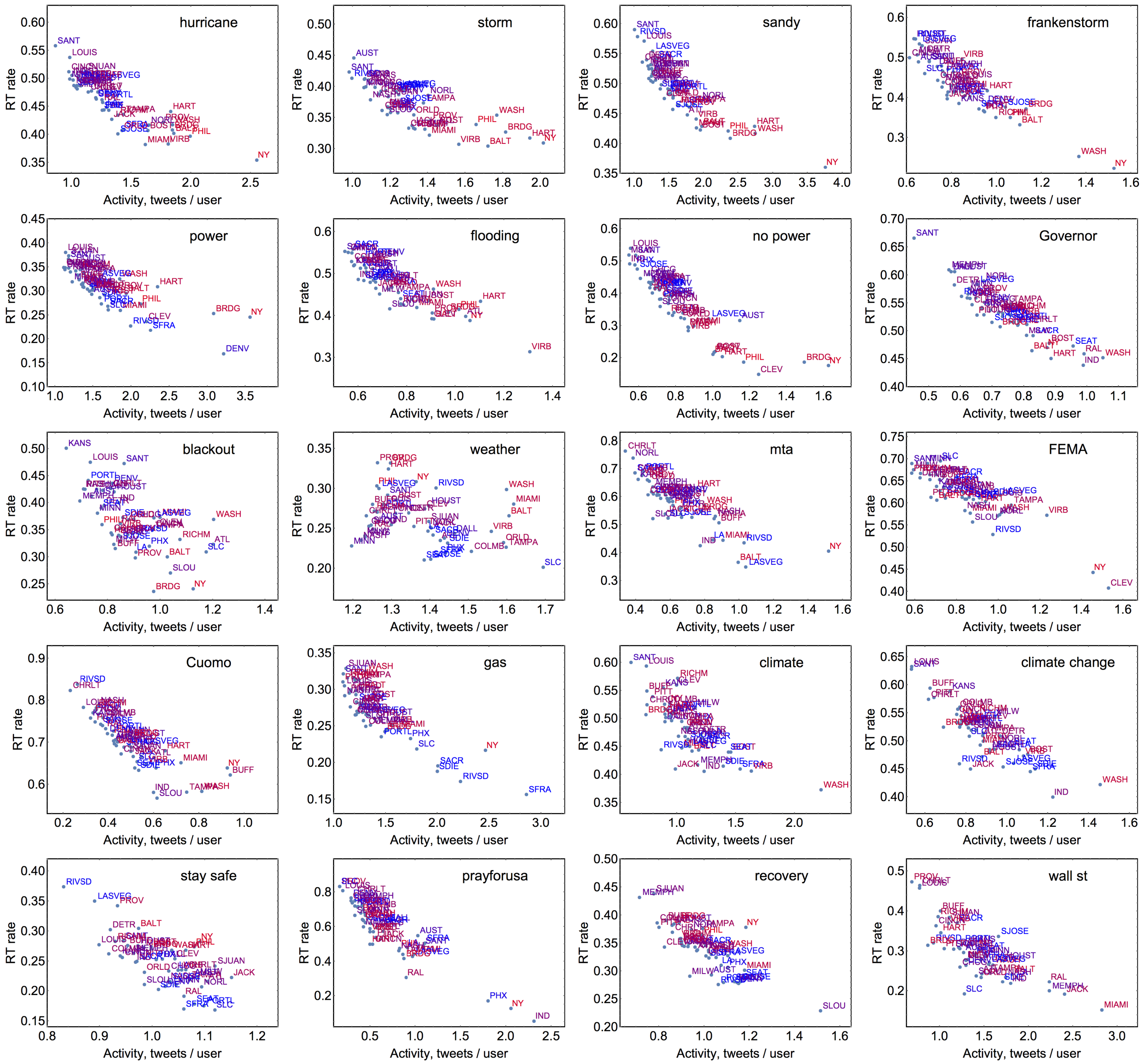

Figure S2 Originality of the content, expressed through the fraction of retweets in the stream of messages. The words are sorted by their relevance based on the strength of activity-proximity correlation, as in Fig S1. Most of the words show the inverse relationship between normalized activity and retweet rate. For the event-related keywords at the top this is more pronounced: cities follow the downward linear trend closely and affected places are concentrated at the "high activity/low retweet rate" end of the spectrum. Less relevant words demonstrate wider spread ("weather", "stay safe") and/or uniform distribution of affected places along the downward trend ("climate change").

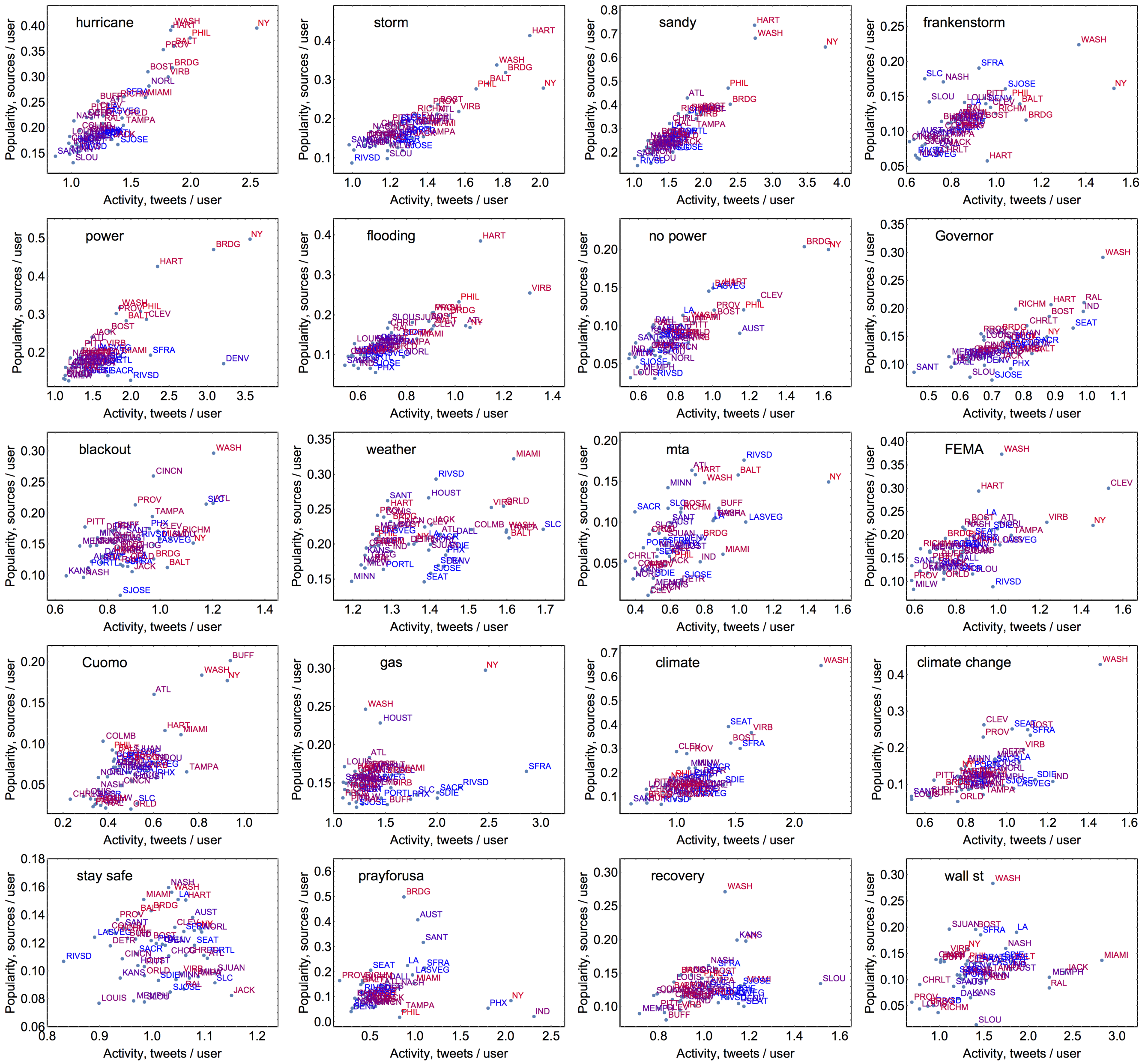

Figure S3 Global popularity of local content. Words are sorted, as in Figure S1, according to relevance based on the strength of activity-distance correlation for East Coast cities. Event-related keywords at the top show a strong linear relationship between activity and global popularity of messages generated locally. As relevance of the word to the disaster weakens, so does the correlation between activity and popularity.

Table S2 Ranking of the keywords included into analysis according to strength of the correlation between the distance and activity for East Coast cities. Event-related words on the top are used for further analysis of activity, including its relationship to damage.

	Kendall rank		Spearman rank	
keyword	τ	P-value	ρ	P-value
hurricane	-0.65	$7.90\ 10^{-9}$	-0.82	$2.20\ 10^{-10}$
storm	-0.61	$6.01\ 10^{-8}$	-0.80	$1.90\ 10^{-9}$
sandy	-0.56	$7.78\ 10^{-7}$	-0.75	$7.47\ 10^{-8}$
frankenstorm	-0.54	$1.67\ 10^{-6}$	-0.73	$2.49\ 10^{-7}$
power	-0.52	$3.95\ 10^{-6}$	-0.72	$3.14\ 10^{-7}$
flooding	-0.52	$3.95\ 10^{-6}$	-0.71	$6.61\ 10^{-7}$
no power	-0.52	$5.03\ 10^{-6}$	-0.67	$3.55\ 10^{-6}$
Governor	-0.49	$1.81\ 10^{-5}$	-0.65	$1.14\ 10^{-5}$
blackout	-0.32	$4.32\ 10^{-3}$	-0.45	$4.36\ 10^{-3}$
weather	-0.25	0.03	-0.35	0.03
mta	-0.23	0.04	-0.36	0.03
FEMA	-0.18	0.10	-0.27	0.10
Cuomo	-0.17	0.14	-0.25	0.13
gas	-0.06	0.59	-0.11	0.50
climate	-0.06	0.59	-0.08	0.61
climate change	-0.06	0.61	-0.06	0.74
stay safe	-0.06	0.62	-0.09	0.61
prayforusa	-0.04	0.75	-0.09	0.60
recovery	-0.03	0.79	-0.04	0.81
wall st	0.06	0.57	0.12	0.49

Table S3 Activity-damage correlations across keywords, in order of decreasing strength. Note that event-related keywords (on the basis of activity-distance relationship) are also most predictive of damage. For the final analysis we use the following selection: “sandy”, ”hurricane”, ”storm”, ”power”, “FEMA” and “flooding”.

		Kendall rank		Spearman rank		Pearson	
keyword	ZCTAs	τ	P-value	ρ_S	P-value	ρ_P	P-value
sandy	469	0.38	$3.07\ 10^{-35}$	0.54	$8.17\ 10^{-37}$	0.61	$2.52\ 10^{-49}$
hurricane	420	0.33	$3.29\ 10^{-24}$	0.47	$1.27\ 10^{-24}$	0.53	$1.04\ 10^{-31}$
power	496	0.33	$3.23\ 10^{-28}$	0.46	$1.35\ 10^{-27}$	0.50	$1.19\ 10^{-32}$
no power	395	0.28	$1.06\ 10^{-16}$	0.40	$8.27\ 10^{-17}$	0.43	$4.42\ 10^{-19}$
flooding	74	0.28	$4.90\ 10^{-4}$	0.40	$4.06\ 10^{-4}$	0.50	$5.49\ 10^{-6}$
FEMA	122	0.24	$6.49\ 10^{-5}$	0.34	$1.33\ 10^{-4}$	0.33	$2.13\ 10^{-4}$
storm	386	0.23	$1.10\ 10^{-11}$	0.34	$7.99\ 10^{-12}$	0.36	$3.17\ 10^{-13}$
gas	455	0.16	$1.62\ 10^{-7}$	0.25	$8.11\ 10^{-8}$	0.30	$1.06\ 10^{-10}$
blackout	113	0.14	0.02	0.21	0.03	0.17	0.07
recovery	98	0.14	0.04	0.20	0.05	0.30	$2.41\ 10^{-3}$
frankenstorm	37	0.17	0.14	0.29	0.08	0.18	0.29
climate	33	0.18	0.14	0.27	0.13	0.30	0.09
Governor	135	0.08	0.17	0.12	0.16	0.18	0.04
Cuomo	27	0.18	0.19	0.28	0.15	0.44	0.02
stay safe	86	0.08	0.28	0.10	0.37	0.10	0.38
weather	369	0.03	0.39	0.04	0.40	0.14	$7.52\ 10^{-3}$
climate change	22	0.10	0.52	0.14	0.53	0.02	0.95
mta	19	0.09	0.60	0.11	0.65	0.18	0.46
prayforusa	3	0.33	0.60	0.50	0.67	0.59	0.60
wall st	26	0.00	0.98	0.01	0.96	-0.01	0.95

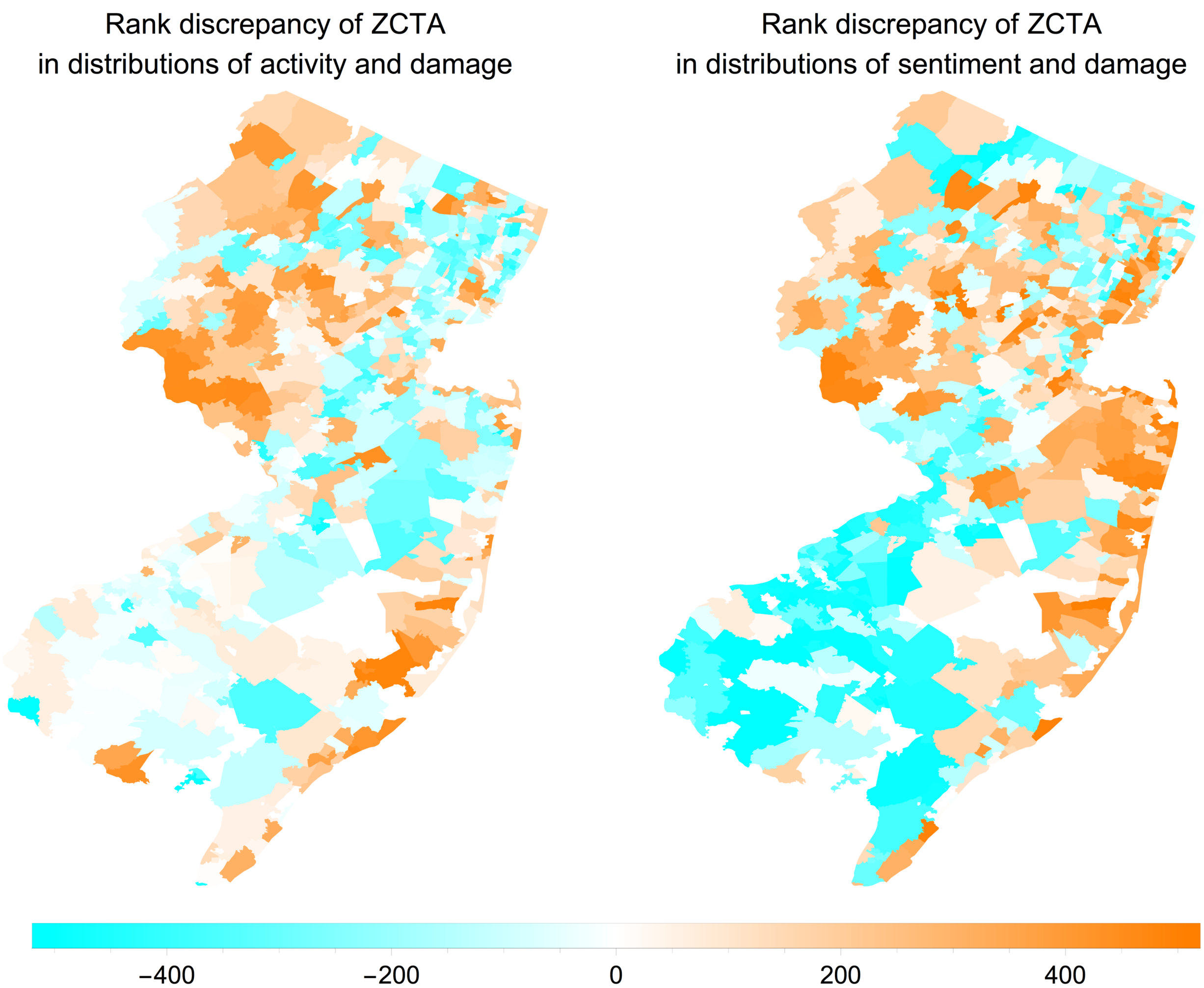

Figure S4 Comparison of predictive capacity of activity and sentiment. The figure maps all zipcode tabulation areas shaded by color, with saturation reflecting discrepancy of the area's rank in two corresponding distributions. For instance, if a particular zipcode is 5th in the ranking of activity, but 100th in the ranking of damage the discrepancy is equal to 95. Discrepancies are normalized by the maximum observed deviation. The stronger the correlation is between the distributions the more uniform and light the map would be, as is the case for activity-vs.-damage map on the left.

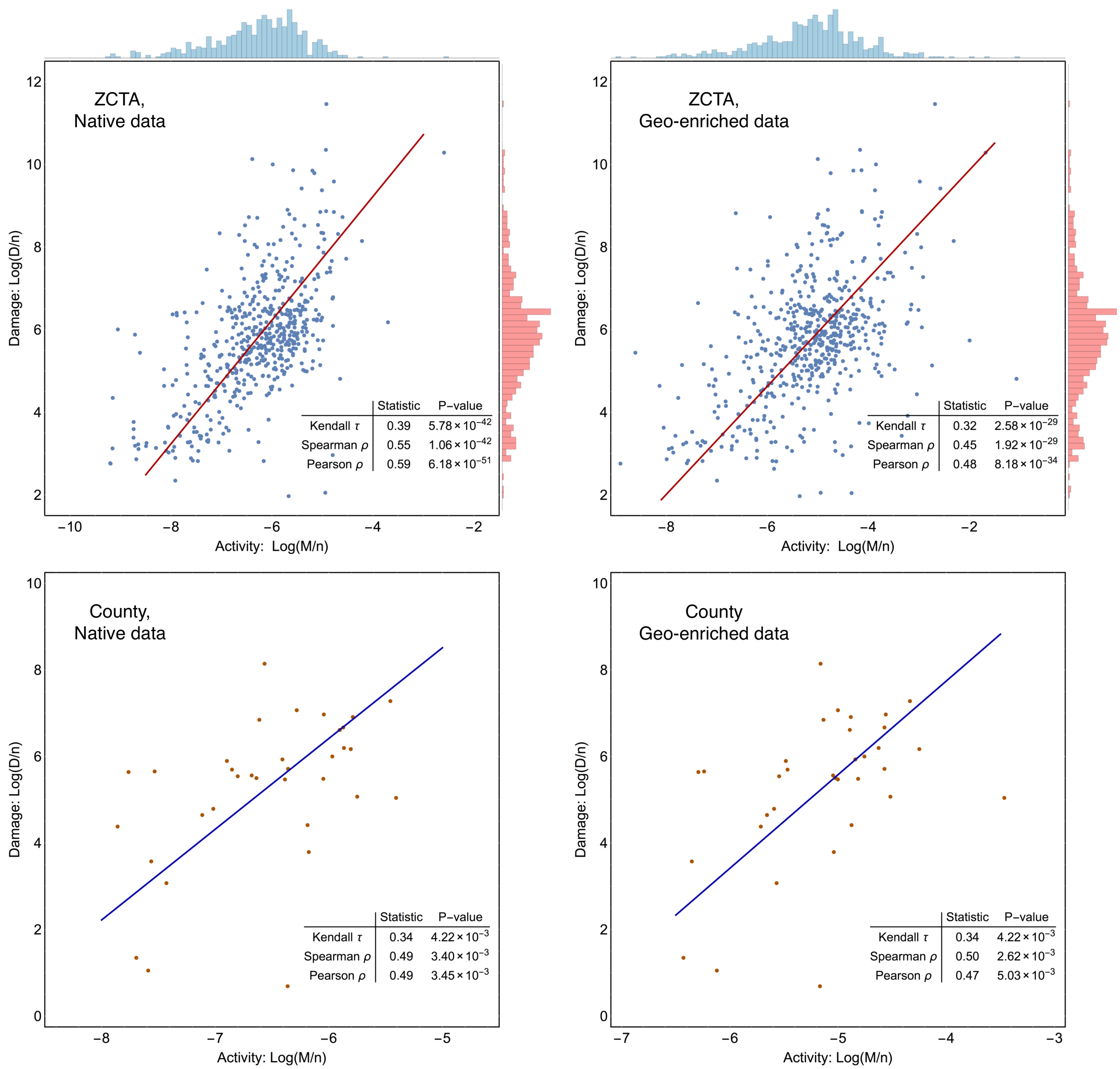

Figure S5 Comparison of the activity-damage correlation strength for different precision of geo-location. The figure shows that the correlation is strongest for fine-resolution (ZCTA) analysis of natively geo-coded data (compare the top left panel versus top right panel with corresponding analysis of geo-enriched data). At the coarser spatial resolution of counties the precision of geo-location and amount of data has little effect. We conclude that natively geo-coded data gives the best predictive power for damage assessment; however, in the absence of sufficient traffic (sparsely populated areas, or small-scale disasters) using geo-enrichment gives a viable analysis option.

Table S4 Effect of normalization variable choice on the strength of activity-damage relationship (ZCTA-resolution).

	Variables normalized by …			
	… Census population		… "Twitter population"	
Correlation measure	statistic	P-value	statistic	P-value
Kendall τ_K	0.39	$5.78\ 10^{-42}$	0.36	$7.88\ 10^{-36}$
Spearman ρ_S	0.55	$1.06\ 10^{-42}$	0.51	$8.07\ 10^{-36}$
Pearson ρ_P	0.59	$6.18\ 10^{-51}$	0.57	$2.48\ 10^{-47}$

Table S5 County level estimates of damage: modeling (Hazus-MH) and *ex-post* data on insurance and FEMA Individual Assistance grants.

				Damage estimates	
County	Population	Tweets	Users	*ex-post*, $M	Hazus-MH, $M
Atlantic	275422	388	216	954.	1630.
Bergen	918888	2589	1056	729.	1070.
Burlington	451336	405	246	54.6	164.
Camden	513539	275	190	147.	103.
Cape May	96304	167	98	29.3	740.
Cumberland	157785	61	47	12.7	128.
Essex	787744	1871	859	844.	375.
Gloucester	289586	172	120	6.29	151.
Hudson	652302	1963	838	314.	3600.
Middlesex	823041	2334	950	406.	776.
Monmouth	629384	2686	922	919.	1930.
Ocean	580470	1780	570	587.	3240.
Passaic	502885	1035	429	41.8	34.2
Salem	65774	28	19	18.6	167.
Union	543976	1734	642	87.2	395.
Bronx	1408473	732	421	50.6	635.
Kings	2565635	2855	1662	660.	5470.
Nassau	1349233	2525	1042	1590.	6860.
New York	1619090	7259	4022	252.	4820.
Orange	374512	305	176	39.2	22.7
Putnam	99607	172	72	0.2	0.405
Queens	2272771	2299	1334	832.	3650.
Richmond	470728	1286	481	353.	1880.
Rockland	317757	400	188	83.3	86.8
Suffolk	1499273	2472	1046	569.	2720.
Ulster	181791	92	52	0.524	8.03
Westchester	961670	1262	640	237.	1320.

Table S6 Strength of activity-damage correlations for different damage estimates.

	Damage is estimated …			
	… by modeling (Hazus-MH)		… from insurance and FEMA claims	
Correlation measure	statistic	P-value	statistic	P-value
Kendall τ_K	0.28	$3.9\ 10^{-2}$	0.37	$7.16\ 10^{-3}$
Spearman ρ_S	0.44	$2.22\ 10^{-2}$	0.53	$4.92\ 10^{-3}$
Pearson ρ_P	0.33	$8.84\ 10^{-2}$	0.46	$1.55\ 10^{-2}$

Table S7 Predictive power of sentiment, analyzed at different spatial resolutions and normalized either by the area Census population or local Twitter user count (“Twitter population”).

	Kendall τ for sentiment-damage relationship			
	ZCTA		County	
Normalization	statistic	P-value	statistic	P-value
Census population	-0.031	0.294	-0.28	0.018
Twitter users count	-0.04	0.169	-0.34	0.005

Table S8 List of the disasters considered in the study with description of the damage data available for analysis. The source of insurance data in every case is a financial regulatory department of the corresponding state's government. We request the data using Open Public Records Act or Freedom of Information Act (OPRA or FOI). From all requests made, only New Jersey provided the data aggregated at ZCTA-level. New York and Arkansas provided county-level totals. Several states denied requests citing unavailability of data, or failed to respond. We did not request the data from the remaining states, as it became clear that majority of them are denied or aggregated coarsely. Therefore, in the main part of the manuscript damage analysis is performed at ZCTA-level for New Jersey, and county-level for New Jersey and New York. Additional analysis for other events relies only on FEMA data, in the absence of insurance data.

Year	State	FEMA declaration	OPRA/FOI	Damage data		
				FEMA	Insurance	
					ZCTA	county
2012	New York	DR-4085: hurricane	total by counties	v	-	v
	New Jersey	DR-4086: hurricane	total by ZCTAs	v	v	v
2013	Illinois	DR-4116: severe storms, straight-line winds and flooding	-	v	-	-
	Oklahoma	DR-4117: severe storms and tornadoes	-	v	-	-
	Alaska	DR-4122: flooding	-	v	-	-
	Colorado	DR-4145: severe storms, flooding, landslides and mudslides	-	v	-	-
	Illinois	DR-4157: severe storms, straight-line winds and tornadoes	-	v	-	-
2014	Washington	DR-4168: flooding and mudslides	no data	v	-	-
	Arkansas	DR-4174: severe storms, tornadoes and flooding	total by counties	v	-	v
	Mississippi	DR-4175: severe storms, tornadoes and flooding	no response	v	-	-
	Alabama	DR-4176: severe storms, tornadoes, straight-line winds and flooding	no response	v	-	-
	Florida	DR-4177: severe storms, tornadoes, straight-line winds and flooding	no data	v	-	-
	California	DR-4193: earthquake	no response	v	-	-
	Michigan	DR-4195: severe storms and flooding	no data	v	-	-

Table S9 The effect of the activity threshold filter on the strength of relationship between Twitter activity and damage. To minimize the potential confounding effect of media coverage and capture the activity of average users we implement the following steps. First, we use natively geocoded messages, leaving out media accounts themselves for the lack of coordinates, along with the primary form of response to their messages (retweets are not geocoded). Second, we remove large clusters of stationary messages. Finally, we filter messages by the threshold in the total activity (removing the messages from accounts that exceed the threshold). Results presented below demonstrate that our findings hold even for users who only post one message.

	Activity threshold					
	100	50	20	10	5	1
Kendall τ_K	0.39	0.40	0.40	0.41	0.39	0.34
Spearman ρ_S	0.55	0.55	0.55	0.56	0.54	0.48
Pearson ρ_P	0.59	0.59	0.59	0.60	0.57	0.50

Table S10 Mutual correlations between sentiment metrics on the level of individual messages.

	Topsy	LIWC	SentiStrength
Topsy	1.00	0.68	0.53
LIWC		1.00	0.60
SentiStrength			1.00

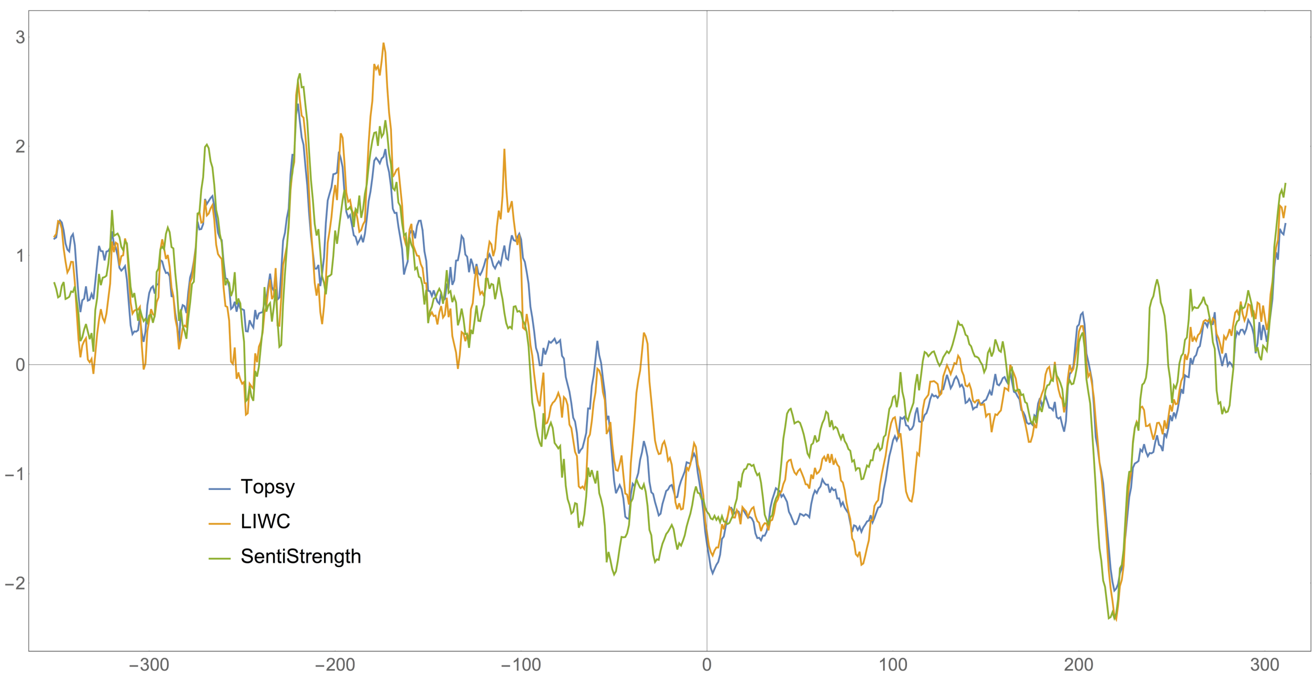

Figure S6 Average sentiment trends over time – comparison between sentiment metrics. All of the metrics reliably track the evolution of aggregate mood over time, as demonstrated by the trends for mean sentiment of all messages aggregated into hourly bins (raw scores for each method are rescaled and standardized).

Table S11 Top-ranking words by frequency of occurrence in positive and negative messages. Messages are grouped by sentiment scores; filtered to delete URLs, hashtags and usernames, converted to lowercase, and then word frequency analysis is performed. Top-30 most frequently occurring words in the positive and negative categories are presented below.

Positive messages		**Negative messages**	
Word	**Rate (%)**	**Word**	**Rate (%)**
power	5.15	power	6.82
hurricane	3.68	hurricane	3.33
storm	2.17	storm	1.86
good	1.39	lost	1.00
like	1.34	shit	0.99
lol	1.17	fuck	0.99
hope	0.97	bad	0.67
thanks	0.92	bitch	0.49
safe	0.92	damn	0.42
thank	0.81	lose	0.40
love	0.78	crazy	0.37
great	0.55	like	0.33
happy	0.47	dark	0.27
ok	0.46	sucks	0.23
best	0.38	come	0.23
please	0.37	worst	0.22
lmao	0.35	hell	0.20
fun	0.31	hate	0.18
come	0.29	charge	0.18
haha	0.27	worse	0.18
nice	0.26	mad	0.18
calm	0.25	wtf	0.17
wow	0.24	outside	0.17
glad	0.23	fucked	0.17
lucky	0.21	dead	0.15
awesome	0.21	scary	0.15
perfect	0.19	emergency	0.13
fine	0.18	wow	0.13
strong	0.17	ill	0.13
party	0.16	live	0.13

Table S12 Sentiment as a predictor of damage, comparison between metrics. In general, analysis by each method shows that a) sentiment at the fine resolution of ZCTAs is not predictive of damage (low ranking correlation and P-value >0.05); b) for the coarser spatial aggregation (counties) there is a weak relationship between sentiment and damage, which we report in the manuscript for the method with the highest statistical significance (Topsy).

	ZCTA		**County**	
	Kendall τ_K	P-value	Kendall τ_K	P-value
Topsy	-0.031	0.294	**-0.283**	**0.018**
LIWC	0.010	0.738	-0.144	0.230
SentiStrength	0.056	0.054	-0.023	0.847