RESEARCH ARTICLE

Using social media to measure impacts of named storm events in the United Kingdom and Ireland

M. Spruce | R. Arthur | H. T. P. Williams

College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

Correspondence
Michelle Spruce, College of Engineering Mathematics and Physical Sciences, University of Exeter, Exeter, UK.
Email: ms886@exeter.ac.uk

Funding information
Economic and Social Research Council, Grant/Award Number: ES/P011489/1; Engineering and Physical Sciences Research Council, Grant/Award Number: PhD Studentship; Natural Environment Research Council, Grant/Award Number: NE/P017436/1

Abstract

Despite increasing use of impact-based weather warnings, the social impacts of extreme weather events lie beyond the reach of conventional meteorological observations and remain difficult to quantify. This presents a challenge for validation of warnings and weather impact models. This study considers the application of social sensing, the systematic analysis of unsolicited social media data to observe real-world events, to determine the impacts of named storms in the United Kingdom and Ireland during the winter storm season 2017–2018. User posts on Twitter are analysed to show that social sensing can robustly detect and locate storm events. Comprehensive filtering of tweets containing weather keywords reveals that ~3% of tweets are relevant to severe weather events and, for those, locations could be derived for about 75%. Impacts of storms on Twitter users are explored using the text content of storm-related tweets to assess changes in sentiment and topics of discussion over the period before, during and after each storm event. Sentiment shows a consistent response to storms, with an increase in expressed negative emotion. Topics of discussion move from warnings as the storm approaches, to local observations and reportage during the storm, to accounts of damage/disruption and sharing of news reports following the event. There is a high level of humour expressed throughout. This study demonstrates a novel methodology for identifying tweets which can be used to assess the impacts of storms and other extreme weather events. Further development could lead to improved understanding of social impacts of storms and impact model validation.

KEYWORDS
extreme weather, impacts, social media, social sensing, storms

1 | INTRODUCTION

It is well known that extreme weather events such as strong winds, heavy rain and snow cause impact and disruption to our daily lives (IPCC, 2014). However, there is little observational record of the specific impacts (e.g. damage to property, disruption to travel, danger to life, stress and anxiety) that occur as a result of these weather events. This information lies beyond the scope of traditional meteorological observations. The frequency

and intensity of extreme weather events has increased over recent years and is predicted to continue to increase (IPCC, 2014). Meanwhile, there has been a shift from forecasts that focus on meteorological conditions alone to forecasts that incorporate information about their associated impacts (Taylor, 2018). This impact-based forecasting strategy is endorsed by the World Meteorological Organization (WMO), who have produced guidance to support its development (WMO, 2015). Together, these trends create an urgent need to understand the ways in which extreme weather events affect people and property, to validate forecast models and warning systems.

Social media is increasingly used across the world (Statista, 2017) and this presents an opportunity to use the rich social information it creates to inform preparedness and response to natural hazard events. Many people routinely use social media to discuss weather conditions, particularly when weather patterns are unusual. During crisis events, such as periods of extreme weather, technological challenges in affected areas may slow official news correspondent reports, while social media reports may be more swiftly distributed (Spence et al., 2015). The public availability of data from some social media platforms, notably Twitter, opens the possibility to use social media data to understand how human activity is affected during an extreme weather event.

"Social sensing" using social media has been widely used for knowledge discovery in fields relating to public health, human behaviour, social influence and market analysis (Wang et al., 2015b). Social sensing broadly refers to a set of sensing and data collection models whereby data are collected from humans or personal devices (Wang et al., 2015a). In this paper, social sensing using unsolicited social media data is distinguished from solicited crowd-sourcing, where users voluntarily participate and report observations in a structured or semi-structured manner. Examples of solicited crowd-sourcing include the UK Met Office Weather Observations Website (WOW. Met Office Weather Observations Website, 2019), where the public can provide amateur weather observations, and the UK Snow Map (UK Snow Map, 2010), where Twitter users are asked to report snowfall observations using a particular hashtag (#uksnow). While solicited crowd-sourcing offers benefits in that data are more reliable and can be provided in a structured form by a set of dedicated volunteers, the volumes of data generated are typically low relative to the high volumes seen in unsolicited social media use; this can limit the usefulness of solicited data for understanding of wider impacts.

For social sensing using unsolicited social media, each individual user plays the role of a sensor. When a user publicly posts an item to a social media platform, they are providing a piece of sensor data. When grouped together by topic or location, large numbers of social media posts can therefore be used to develop an understanding of a range of issues. Social sensing of this nature has already been successfully used to detect natural hazards such as earthquakes (Sakaki et al., 2010), wildfires (Boulton et al., 2016) and floods (Brouwer et al., 2017; Tkachenko et al., 2017; Arthur et al., 2018; Rossi et al., 2018). A number of studies have used social media to understand impacts of hurricanes in the United States (Guan and Chen, 2014; Cervone et al., 2016; Kryvasheyeu et al., 2016; Morss et al., 2017; Kim and Hastak, 2018; Wu and Cui, 2018).

This study explores whether social sensing can help meteorologists to understand how human activity is affected during extreme weather events, in terms of both emotional impacts and other social impacts (e.g. disruption, damage) revealed by the topics of conversation during storm events. Some weather-related studies have begun to explore this opportunity. The effects of weather on mood have been shown using sentiment expressed in tweet text linked to weather conditions (Hannak et al., 2012; Caragea et al., 2014; Li et al., 2014; 2019; Baylis et al., 2018). The categorization of tweet content related to weather and natural hazards has also been explored using both manual methods (Spence et al., 2015; Halse et al., 2018) and automated methods (Alam et al., 2018). However, to date there has been little exploration of social sensing focused on social impacts of weather for the purposes of impact-based forecast validation.

In the present study, data from the social media platform Twitter were collected during the 2017/2018 UK and Ireland storm season (approximately October–March) to explore social sensing as a methodology for assessing the social impacts of storms. The research uses and builds on the social sensing methods described by Arthur et al. (2018) to extract, filter, locate and get useful meaning from social media data collected during this storm period. Sentiment analysis is used to look at the aggregated emotional response to storms and how this changes during the period of a storm event. Categorization of storm-related tweet content provides an indication of what kind of information can be determined from tweets, looking in particular for content related to social impacts. The aims of the study are (a) to establish a methodology for social sensing that can provide useful information about social impacts of storms and (b) to apply the methodology to explore the impact of storms in the United Kingdom and Ireland during winter 2017/2018. These objectives are intended to help develop social sensing as a source of impact observations suitable for validation of impact-based weather forecasting systems.

The paper is split into the following sections: Section 2 outlines the methods used for data collection,

filtering and content analysis; Section 3 reports the main findings of the analysis, focusing on sentiment and categorized impacts observed during storm events; finally Section 4 summarizes the main benefits and limitations of the social sensing approach as demonstrated in this study, and makes some suggestions for future research.

2 | DATA COLLECTION AND METHODS

This study uses a hybrid approach of methods from previous studies which successfully collected and found useful meaning from Twitter data relating to weather events or natural hazards (Lachlan et al., 2014; Arthur et al., 2018; Cowie et al., 2018; Halse et al., 2018). Social media data were collected, filtered for relevance and geo-located. The content of the resulting dataset was then analysed using sentiment analysis and automated categorization.

2.1 | UK/Ireland storm season 2017/2018

Since 2015 the Met Office in the United Kingdom and Met Éireann in Ireland have used a storm naming system to raise public awareness of the effects of stormy weather with the public and to increase preparedness in response to weather extremes. A storm is named if it is expected to cause "medium" or "high" impacts from wind and/or precipitation, i.e. storms will be named for weather systems which are expected to have an amber or red weather warning issued by Met Éireann and/or the Met Office's National Severe Weather Warning Service (https://www.metoffice.gov.uk/news/releases/2017/storm-names-for-2017-18-announced). Weather warnings are colour coded in response to their potential impact and likelihood; amber and red warnings are therefore issued for weather events which are both probable and likely to cause significant disruption.

In the 2017/2018 UK storm season, which generally runs from autumn to early spring, there were a number of named storms which affected the United Kingdom with expected medium or high impacts from wind and/or rain/snow (Table 1). The reason for naming storms is to improve public communication about weather events likely to cause significant impacts. Named storms are likely to attract attention from social media users because of their severity and the use of the names in official communication and forecasts. Named storms are also useful from a technical point of view, as one can search directly for the storm's name. Therefore this study mainly focuses on named storms and the impacts associated with them. Twitter data were collected for named storms for the

TABLE 1 Met Office record of storm names during the 2017/2018 storm season

Storm name	Date named	Date of impact on United Kingdom
Aileen[a]	September 12, 2017	September 12–13, 2017
Ex-Hurricane Ophelia	October 11, 2017 (named by NHC)	October 16–17, 2017
Brian	October 19, 2017	October 21, 2017
Caroline	December 5, 2017	December 7, 2017
Dylan	December 29, 2017	December 30–31, 2017
Eleanor	January 1, 2018	January 2–3, 2018
Fionn	January 16, 2018	January 16, 2018
David[a]	January 17, 2018 (named by Météo France)	January 18, 2018
Georgina	January 23, 2018	January 24, 2018
Emma	March 1, 2018 (named by the Portuguese Met Service)	February 28–March 3, 2018

Information taken from https://www.metoffice.gov.uk/barometer/uk-storm-centre.
[a]Twitter data for Storm Aileen and Storm David were not collected for this study.

duration of the 2017/2018 UK storm season from October 16, 2017 (when news of ex-Hurricane Ophelia hitting the United Kingdom was reported in the media) until March 10, 2018, post Storm Emma. Tweets containing keywords for weather related to a storm (e.g. wind, rain etc.) were also collected during this period. This was so that tweet activity which included weather terms only could be compared with tweets relating specifically to named storms.

Other countries' meteorological services may also name storms, using similar naming systems, so that some storms are already named before hitting the United Kingdom/Ireland. If a weather system has previously been named by another meteorological service, then it retains this name when it reaches the United Kingdom/Ireland. For example ex-Hurricane Ophelia was named by the US National Hurricane Centre (NHC), Storm David by Météo-France and Storm Emma by the Portuguese Met Service.

2.2 | Social media and Twitter data collection

At the end of 2017 it was estimated that there were 2.46 billion social media users around the world, reflecting the global usage of smartphones and mobile devices. The social media platform Twitter, having 330 million

monthly active users (Statista, 2017), is a social networking and microblogging service that allows registered users to interact via short published messages (tweets) up to 280 characters in length. Twitter makes user posts freely available via the Twitter API, making Twitter a popular source of observational data for both social and natural scientists (Williams *et al.*, 2013). Data collection using Twitter can be achieved using keywords or "hashtag" references to specific topics or events. However, suitable algorithms must be applied to filter the data to ensure that only relevant information is then taken forwards for analysis (Spence *et al.*, 2015). Locating the user who has posted an item to a social media platform is another challenge. At present only 1% to 2% of Twitter posts, for example, carry a Global Positioning System (GPS) location or specific location coordinates (Dredze *et al.*, 2013); therefore, other methods must be employed to infer the place of origin.

Using the methods outlined by Arthur *et al.* (2018), tweets relating to named storms and storm-associated weather conditions were collected using the Twitter Streaming API (via a Python script using the Twython package [McGrath, 2013]). This API returns all tweets up to a limit of 1% of the total volume of tweets at any point in time. Search keywords were used as an initial filter applied by the API to identify and download relevant tweets (Table 2). As tweets using these keywords are unlikely to reach the API limit, it is believed that most if not all relevant tweets are downloaded using this method (Morstatter *et al.*, 2013). Some storm names were prone to typing errors in tweets; therefore, some common variants were accounted for in the search terms used. Only tweets in the English language were collected, since the majority of the populations in this study (United Kingdom and Ireland) are English speaking. Tweets were collected over the time period October 16, 2017 to March 10, 2018. Each tweet was saved as a JSON object which is a lightweight data-interchange format often used for transmitting data from a server to a web application (https://www.json.org). Each JSON object contains the tweet text as well as a number of

meta-data fields relating to each tweet (i.e. timestamp, username, user location, geotag etc.).

The storm name collection keywords are shown in Table 2. Storm names were added to the "Storm Names" data collection in the days leading up to each storm event and therefore collections for each storm name do not cover the whole of the study period. As wind is the main weather type to cause impacts during a storm event, tweets relating to wind were collected as well as storm names. Precipitation also causes impacts during a storm event; however, weather warnings relating to each of the named storms predominantly related to the impact of winds, rather than precipitation. It is also likely that there were precipitation events (snow or heavy rain) not related to storm activity which makes the precipitation dataset less comparable with the storm dataset. Therefore, while tweets relating to precipitation were also collected and filtered for relevance, the crucial comparison is between the storm tweet collection and the wind tweet collection. More than 100 million tweets were collected from the API during the 2017/2018 storm season (see Table 3).

Figure 1 shows time series of the numbers of tweets containing the specified keywords collected *per* day during the period October 16, 2017 to March 10, 2018. This includes all tweets (including retweets) in the raw dataset prior to any filtering for relevance to named storms. The time period of each named storm in the collection period is shown by the grey bars. There appear to be associated peaks in Twitter activity relating to the Wind collection. Peaks in the Storm Names collection are less obviously associated with storm events, but inspection suggested that this collection contained some highly relevant content amongst a lot of irrelevant content, which is likely to confound the association. The Precipitation collection has some storm-associated peaks but also many peaks not associated with storm events.

This study is concerned with the social impact of storms as experienced by social media users. For this purpose, retweets are retained in most parts of the analysis, including counts and time series measuring total activity around storms, and sentiment analysis (where it is asserted that retweeting implies endorsement, approval or agreement with the sentiment expressed in the original tweet). For purposes of observing social impacts, retweets and "quote" tweets are removed as they do not represent original observations. This removal was performed using tweet metadata.

TABLE 2   Twitter collections referred to in this study

Collection	Keywords
Wind	wind, gale, windstorm, hurricane
Precipitation	rain, raining, rainy, rainstorms, rainstorm, hail, hailstones, hailstorm, hailing, hale, snow, blizzard, snowstorm
Storm names	storm, ophelia, ofelia, opelia, opehlia, opheliaireland, brian, caroline, dylan, eleanor, fionn, fion, georgina, emma

Only tweets containing one or more of the keywords shown were added to the initial unfiltered dataset for each collection.

2.3 | Filtering and location inference

After data collection, the first stage in processing the Twitter data was to apply a suitable relevance filter to remove any obviously irrelevant data. The various filters

TABLE 3 Total number of tweets collected within each collection and remaining after applying both relevance filter and location inference

Tweet collection	All tweets in raw data collection (unfiltered)	Tweets remaining after filtering for relevance		Tweets remaining after filtering for relevance and location inference		
	Number of tweets	Number of tweets	% (of all tweets)	Number of tweets	% (of all tweets)	% (of tweets after filtering for relevance)
1. Precipitation	67,448,047	3,264,573	4.8%	1,982,378	2.9%	60.7%
2. Wind	26,298,449	831,076	3.2%	472,586	1.8%	56.9%
3. All Storm Names	8,101,901	278,412	3.4%	214,220	2.6%	76.9%
Ophelia	897,054	214,730	23.9%	167,369	18.7%	77.9%
Brian	2,037,045	12,970	0.6%	9,439	0.5%	72.8%
Caroline	1,199,149	8,552	0.7%	4,993	0.4%	58.4%
Dylan	2,504,264	3,907	0.2%	2,410	0.1%	61.7%
Eleanor	555,433	11,872	2.1%	9,761	1.8%	82.2%
Fionn	43,936	1,260	2.9%	878	2.0%	69.7%
Georgina	104,327	894	0.9%	650	0.6%	72.7%
Emma	760,693	24,227	3.2%	18,720	2.5%	77.3%

FIGURE 1 Number of tweets collected *per* day during the collection period October 16, 2017 to March 10, 2018. The period of each named storm is shown by a grey bar
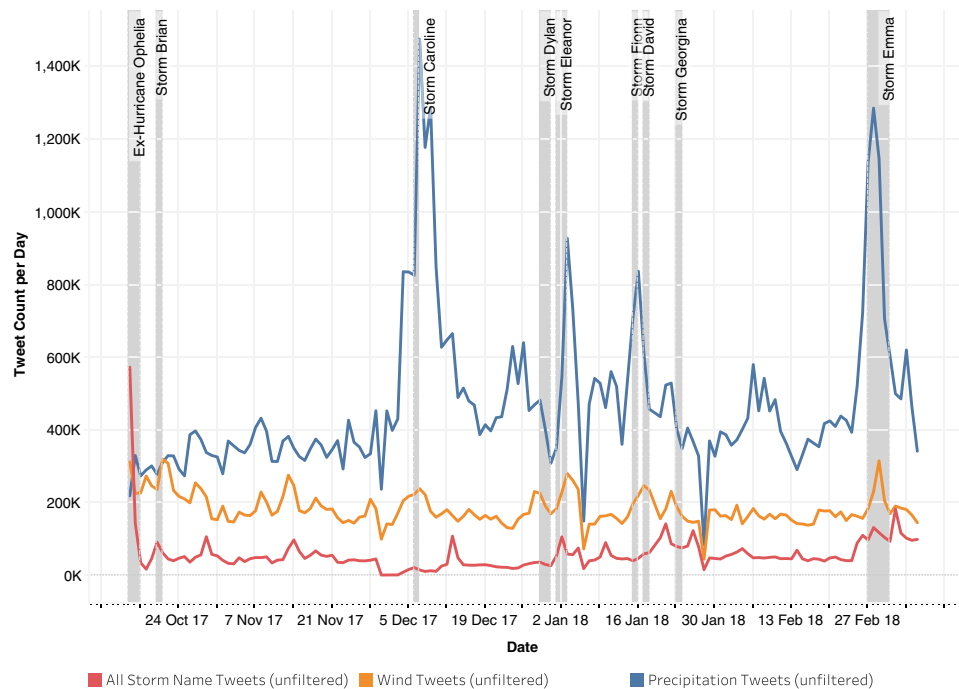
applied can be split into the following stages which are described in the order in which they were applied.

2.3.1 | Time zone filter

The raw data collection contains tweets from all global locations including the United States and other countries. Only tweets which relate to weather activity in the United Kingdom and Ireland are of interest for this study; therefore, the dataset was first filtered based on the time zone entity of each tweet to remove international tweets. The use of time zone as a proxy for the country level location of a tweet is discussed by Schulz *et al.* (2013) who found that over 80% of tweets can be accurately localized to a country using the time zone entity. Tweets with the

following time zones were therefore kept in the dataset: GMT, London, Europe/London, UTC, BST, GMT + 1, Dublin, Europe/Dublin, Edinburgh.

As of May 2018, in order to comply with General Data Protection Regulation requirements, Twitter has removed the time zone field from tweet metadata (Cowie *et al.*, 2018). Other methods for location inference (as described in Section 3.3.6 below) remain effective in the absence of time zone information. This filter removes approximately 90% of tweets in the raw data collection and therefore makes later processing steps more computationally efficient.

2.3.2 | Bot filter

"Bots" are automated user accounts that are set up to perform a particular function, such as collate/spread content from a set of sources, promote a particular view or deliver advertising. Automated tweets from bot accounts are highly unlikely to contain information relating to social impacts of weather activity, but the presence of this kind of content can distort the dataset. To remove bot content, the number of tweets by each user account was calculated for the entire dataset. User accounts with a disproportionately high number of tweets (in this case >1% of the total volume of tweets in the dataset) were identified as bot accounts; automated accounts tend to create significantly more tweets than human users. All tweets posted by bot accounts were then removed from the dataset. A further manual review of the remaining users generating a high proportion of tweets found some additional bot accounts which were also removed. This filter removes approximately 1% of tweets in the raw data collection.

2.3.3 | Weather station filter

Data collections containing weather-related terms include a high number of tweets automatically posted by amateur weather stations. As this study is focused on social impacts, these tweets are deemed irrelevant since they are not directly related to social impacts. A process was developed to remove them. Tweets from weather stations typically follow a fixed structure, e.g. "Wind 2.0 mph E Barometer 30.10 in Falling slowly Temperature 68.5 F Rain today 0.00 in Humidity 55." Here these were identified using a script that searches the text of a tweet and counts weather-related terms; if there were more than two weather-related terms the tweet was identified as a weather station tweet. This method was shown to work well by manual inspection. Tweets identified as being from weather stations using this method were removed from the dataset. This filter removes a very small number

of tweets in the raw data collection for named storms; however, it removes approximately 1% of tweets in the raw data collections for wind and precipitation.

2.3.4 | Irrelevant term filter

As for the weather station filter, this filter is more relevant to the data collections containing weather-related terms rather than storm names. There are many phrases in the English language which use weather-related terms but do not relate to weather, as well as some homographs for weather-related words; these are irrelevant to this study so tweets that contain them were removed using a look-up table method. A list of common terms or phrases which use weather-related terminology but are clearly not referring to a weather event (such as "wind up," "throw caution to the wind," "cook up a storm" etc.) were identified in tweet text and those tweets were removed from the dataset. This filter removes a very small proportion of tweets from the remaining raw data collection.

2.3.5 | Machine learning relevance filter

Although the previous stage removed much irrelevant content, an additional stage of filtering was still necessary to remove tweets which included the search keywords but were not relevant to wind, precipitation and storms. These included, for example, business advertising, links to articles on other topics, references to people and places who shared a name with the storm, and various other items of irrelevant content. Tweets in the Storm Names collection were particularly in need of additional filtering, since there are many celebrities or other individuals who share the same names as the storms studied here. To achieve this, the methods used successfully in previous studies (Arthur *et al.*, 2018; Cowie *et al.*, 2018) were employed.

A set of 6,000 tweets were randomly selected from the tweet collections. Each tweet in this set was then manually labelled as relevant or irrelevant. Manual coding was conservative, labelling as irrelevant tweets that were obviously unrelated to the study topic and also tweets which were ambiguous (i.e. providing insufficient information to decide on relevance). In total there were 1,495 tweets in the dataset labelled as relevant and 4,505 tweets labelled as irrelevant. The labelled dataset was then used as training data for a multinomial naïve Bayes classifier. As a first validation test for this approach, 25% of the data were held back as a validation set and a classifier was trained on the remaining 75% of cases; this classifier had accuracy (i.e. correctly identified the

relevance/irrelevance) of 92% on the held-back validation tweets, with an F1 score of 0.84. As a second test, to confirm the robustness of the approach, the same training/validation test was repeated with 6-fold cross-validation. The results of each test were combined to give an overall mean F1 score of 0.80 and the summed confusion matrix (also known as contingency table) below (where True is relevant and False is irrelevant):

$$\begin{pmatrix} & \text{Predicted} & & \\ & \text{No} & \text{Yes} & \\ \text{Actual No} & 4,301 & 204 & \\ \text{Yes} & 274 & 1,221 & \end{pmatrix}$$

This confusion matrix shows overall accuracy of 92%, with most tweets in the filtered dataset classified as not relevant. Accuracy was higher on the False class (4,301/4,505 = 95%) than on the True class (1,221/1,495 = 82%), with a slight tendency to misclassify relevant tweets as irrelevant. This could be attributed to the training dataset being unbalanced and biased towards irrelevant tweets. However, this is a conservative error that ensures tweets that are retained are highly likely to be relevant. This is probably due to the wide variety of tweets in the Storm Names collection which were not related to named storm discussion. The multinomial naïve Bayes classification approach was deemed to be accurate enough and sufficient for the purposes of this study based on the results discussed above. A new classifier was then trained on the entire set of manually coded tweets to take forward as the relevance filter for this study. As an additional check of the performance of this classifier, random manual checks of the data after this filter was applied to the whole tweet dataset confirmed that it was performing well.

The Bayesian filter described above removes a further 4–5% of tweets in the data collection for named storms and approximately 2% of tweets in the Wind and Precipitation data collection.

Table 3 shows the number and percentage of tweets remaining for each tweet collection after the stages of relevance filtering described in Sections 2.3.1–2.3.5 were applied. Overall there are 3–4% of tweets remaining after relevance filtering. Table S1 provides a more detailed breakdown of the number and percentage of tweets removed at each stage of relevance filtering for each tweet collection.

2.3.6 | Location inference

After relevance filtering was completed, each tweet in the dataset was also processed to identify if it can be located using information contained within the tweet. The spatial distribution of tweets relating to the weather would also give an indication of social impacts in particular locations.

As found in other studies, this study also finds that only ~1% of tweets contain geo-coordinates of the tweet origination. Therefore, a location inference method is required. Using the same location inference approach as the one outlined by Arthur *et al.* (2018), the filtered tweet dataset was examined for different kinds of geographical information: geo-coordinates (geotag), the place a user designated in the Twitter application when posting (place), the location given in the user profile (user location) and place names mentioned in the tweet text. This method is based on the location inference method validated by Schulz *et al.* (2013) who found 92% accuracy when inferred location was compared against tweets for which a geotag was known. Thus, there were four tweet elements examined for location information in the following order:

- Geotag: locate tweets using geotag (GPS coordinates)
- Place: location of tweet (polygon coordinates)
- User location: GPS coordinates; if not lookup text with Geonames db
- Place names mentioned in tweet text: DBpedia Spotlight (Mendes *et al.*, 2011) lookup identifies place names and coordinates

It was found that the most useful elements of a tweet which can be used to determine a location are the user location and place name mentioned in the tweet text. Table 3 shows the number and percentage of tweets in the filtered dataset for which a location could be found for each tweet collection. On average 77% of filtered tweets could be located using this inference method. Here "located" means that a tweet was allocated to a defined spatial area with high confidence. Table S2 provides more detail on the specific numbers and proportion of tweets located by each tweet element for each tweet collection.

2.3.7 | Results of filtering and location inference

After applying the above methods of relevance filtering the number of tweets retained for analysis was substantially reduced. Figure 2 shows an example of this reduction for Storm Brian. Compared with the unfiltered data, the filtered dataset contains far fewer tweets. However, there is now a clear peak of Twitter activity of relevance to Storm Brian which coincides with the period of the
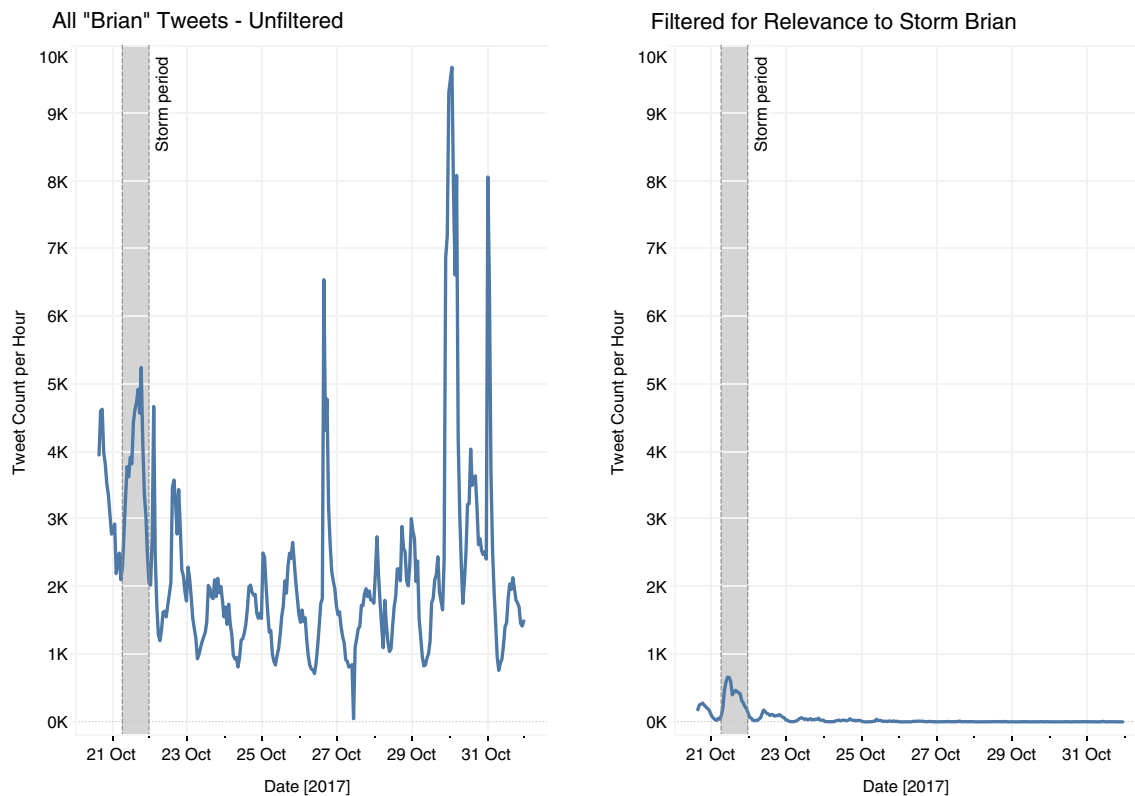
All "Brian" Tweets - Unfiltered

Filtered for Relevance to Storm Brian

FIGURE 2 "Brian" tweets: unfiltered (i.e. all tweets containing the word "brian") versus post filtering for tweets relevant to Storm Brian

storm (shown by the grey bar in the figure). The same is found for each of the named storms in the dataset (data not shown). Figure 3 shows tweets that were both located (using location inference) and relevant (passed the relevance filters). All other analysis uses all relevant tweets that are located to the United Kingdom and Ireland by time zone, but not necessarily precisely located using the inference process.

Results for the Precipitation, Wind and Storm Names collections, pre- and post-filtering and after location inference, can be found in Table 3. Typically, <5% of tweets are retained after filtering for relevance. Interestingly this was much higher (~24%) for the dataset relating to ex-Hurricane Ophelia. This is most likely because Ophelia is an uncommon name. Where a storm is named with a more common name (i.e. Brian, Caroline etc.) the percentage of tweets retained after filtering for relevance is much smaller because there is a higher background level of Twitter activity. Of the relevant tweets, typically 55–80% could be successfully geo-located using the inference method outlined above.

Figure 3 presents a case study of located tweets in England and Wales by county, as an example of the social sensing technique. This case study shows the spatial extent of tweet activity in England and Wales for

Storm Brian following application of location inference. Tweets located in Scotland, Northern Ireland and Ireland are not shown in this figure but were included in other analyses. Darker shading indicates where there was more Twitter activity for a particular area than average for that location, plotted as an exceedance probability. The probability of exceedance is a statistical metric describing the probability that a particular value will be met or exceeded (McMahan *et al.*, 2013). In this example, this provides the likelihood of recording a given number of tweets about storms in this particular location, based on the frequency distribution of observed counts across the whole storm collection dataset. This provides geographical information on where the storm is being most discussed on Twitter and therefore an indication of which areas of the country are likely to be most affected by the storm. In this example for Storm Brian, more significant tweet activity can be seen in the west, south and southwest of England and Wales. It also shows how the spatial pattern of tweets changes over time during the period leading up to, during and after the storm. As anticipated, there is a peak of activity on the day of the storm, which quickly reduces in the days afterwards.

Once both relevance filtering and location inference were completed, the dataset was then prepared for
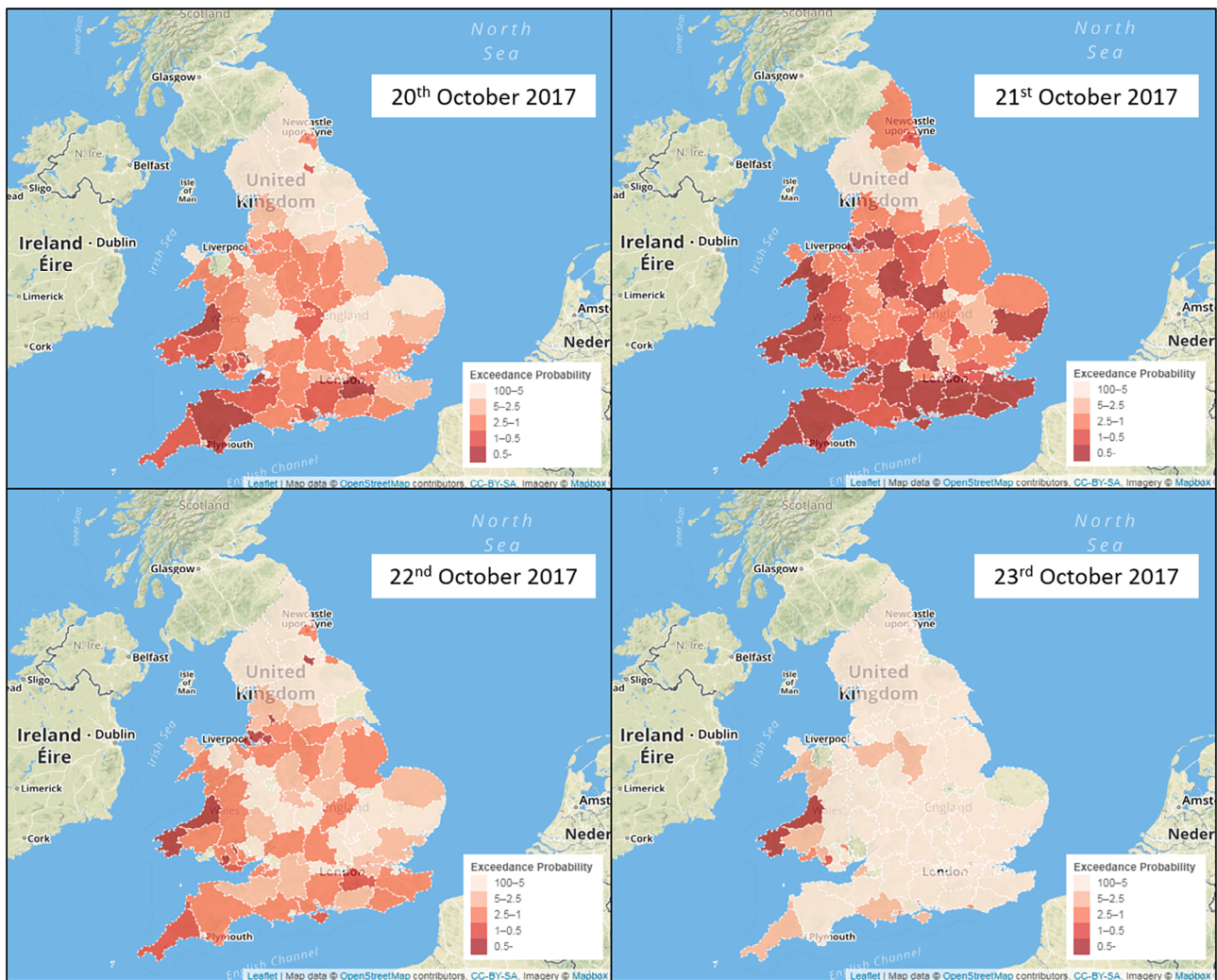
FIGURE 3 Storm Brian tweets (after filtering for relevance) located in England/Wales and grouped by county for each day of the storm period. Storm Brian hit the United Kingdom on October 21, 2017. Shading indicates the exceedance probability for the number of tweets observed by county (i.e. the likelihood of that activity level accounting for prevalence of tweet activity in that particular location). Data shown in this visualization are restricted to England and Wales only, but data analysed in this study extend to Scotland and Ireland

further analysis to determine information on social impact from the tweet data. All filtered tweets' text was used for sentiment and content analysis.

2.4 | Sentiment analysis

The "sentiment" of a tweet measures the net level of positive or negative emotion it expresses. In this case, following various studies that use sentiment analysis with tweets to examine collective mood related to weather conditions (Hannak *et al.*, 2012; Caragea *et al.*, 2014; Li *et al.*, 2014; 2019; Baylis *et al.*, 2018), sentiment analysis is used to infer the mood of Twitter users. By analysing the collective sentiment of tweets during the period of a storm

event, the aim is to get an indication of the emotional impact of the storm.

Tweet text was analysed using the sentiment analysis package TextBlob (Loria, 2010). This Python package is a popular lexicon-based sentiment analysis tool well suited to the relatively short text strings found in tweets. In preliminary work, TextBlob was tested against another leading sentiment package, VADER (Hutto and Gilbert, 2014), which gave comparable results. Since there was no substantive difference, TextBlob was preferred for ease of use with this dataset.

The TextBlob package returns a sentiment polarity value between -1 and 1, where <0 implies negative sentiment and >0 implies positive sentiment. The value returned is based on a sentiment classifier trained on a large dataset of text relating to movie reviews tagged as positive or negative. The

Category	Tweet Text Examples	Sentiment Score
Humour	"Brian? What kind of name is that for a storm? Everyone knows Brian is a snail."	0.60
	"Am I the only one to find it really hard to take a storm called #Brian seriously?	−0.21
	"And Brian? Really? Storm Rambo or Terminator would be far better than #StormBrian"	0.27
Damage	"This is the scene this morning as the waves have damaged the Harbour Office during Storm Brian."	−0.26
	"Storm Brian damage causes floodlight damage. Revised home game vs @ChesterCityFC"	−0.40
	"Scaffolding in Helsby High Street BLOWN OVER by #StormBrian high winds"	0.00
Disruption	"Train delay: National Rail have warned of delays due to high winds from Storm Brian"	−0.25
	"Storm Brian latest - tree blocks railway lines and hovercraft suspended"	−0.41
	"Major motorway was CLOSED after Storm Brian floods carriageway"	−0.02
Warnings	'#StormBrian could lead to travel disruption this weekend.'	−0.06
	'Storm Brian set to batter UK with heavy rain and 70mph winds.'	−0.20
	'Take care on the coast folks. Waves are quite high with #StormBrian'	0.16
Observations	"It's really windy out there!"	0.20
	"Storm Brian seems to have arrived now…"	0.00
	"Storm Brian just brought in the heaviest rain shower I've ever seen.....it really scared our 2 cats."	−0.29
News	"Storm Brian makes landfall on west coast"	0.00
	"Storm Brian: 'Weather bomb' storm set to batter Britain and Ireland - Earth News"	−0.11
	"Live updates on Storm Brian across Somerset at https://t.co/r49v2aUKFc"	0.00
Other	"Will I be ok putting my bin out tomorrow or will Brian still be raging ?"	−0.05
	"Had afternoon tea with this cutie in a cosy cottage as Storm Brian rattled the windows. #friends #halftermbreak"	0.00
	"No sign of Storm Brian up here, thankfully! We have a lovely day, a great one for a long walk on the beach"	0.42

FIGURE 4 Example of the types of tweets included in each category with sentiment score calculated using the TextBlob package. These are synthetic tweets rather than actual tweets, in order to protect user privacy

sentiment polarity score for each tweet is based on all words in the tweet text. Figure 4 provides examples of tweets with sentiment score calculated using TextBlob.

2.5 | Content analysis

Filtered tweets in the Storm Brian dataset at times of peak activity (October 20, 2017 to October 22, 2017) were manually analysed and placed into one of seven categories based on their content. Only tweets containing original content (i.e. excluding retweets and quotes) were analysed for their content. Categories were determined after an initial inspection of a subsample of filtered tweets, using a similar approach to a study on the volume and content of Tweets associated with Hurricane Sandy (Lachlan *et al.*, 2014). The categories used were:

• Humour—Tweet contains a joke, sarcastic remark or light-hearted commentary on experience of the storm event; does not provide any information about any impact as a result of the storm.
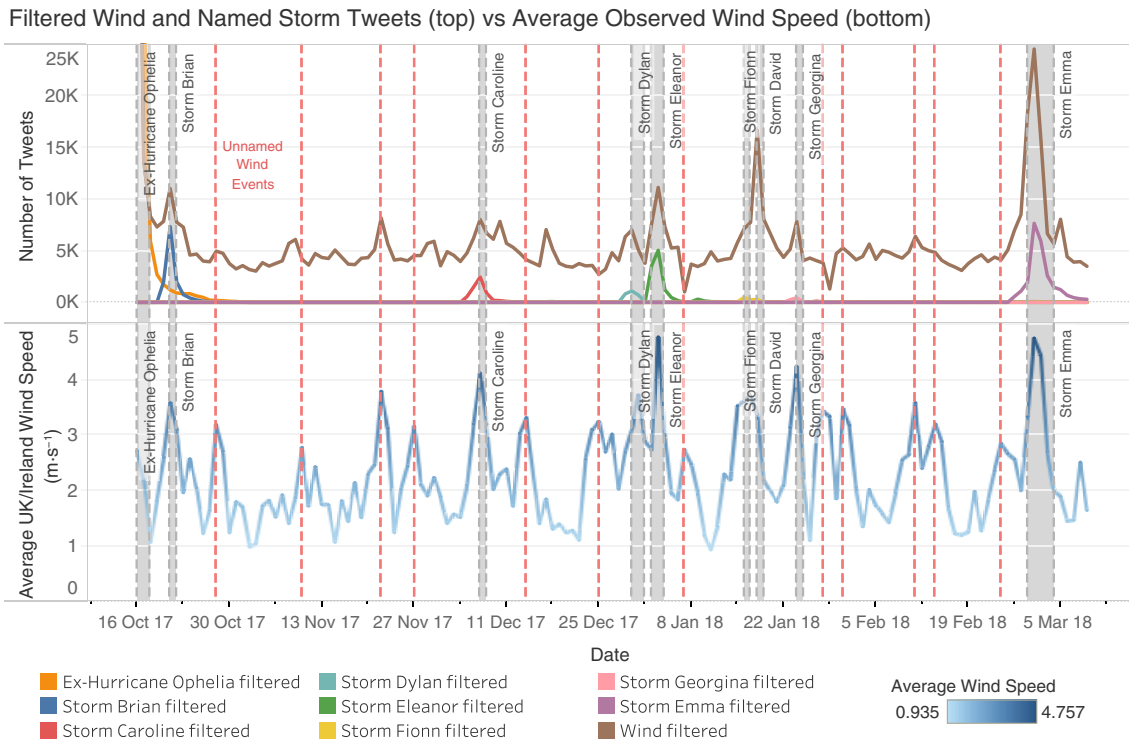
FIGURE 5 (Top) Time series of the number of tweets *per* day for named storm events (after filtering for relevance) versus the number of wind tweets *per* day for the 2017/2018 storm period. Ex-Hurricane Ophelia produced very high numbers of tweets in the Named Storm and Wind collections for October 16, 2017; that is why plotted counts are truncated for display. Tweet counts for each collection on this date are ~170k ("Ophelia") and ~60k ("wind") respectively. (Bottom) Time series of the average UK/Ireland wind speed for the same period. Peaks in wind speed are identified by dashed lines between the two plots to allow visual comparison of wind speed and peaks in wind tweet activity

- Damage—Tweet contains information about damage to persons or property.
- Disruption—Tweet contains information about disruption to daily life, e.g. train delays, road closures, not able to go to work.
- Observations—Tweet contains commentary on the weather occurring, e.g. "wind is very strong," "Storm Brian has arrived here in Balamory."
- Warnings—Tweet contains information and advice about the forthcoming storm, or a warning about danger to persons or property due to the storm.
- News—Tweet contains reference to a media report on the storm event.
- Other—Tweet content relating to the storm that does not fit into the above categories.

Figure 4 provides examples of the types of tweets used in each category.

Categorization of tweets was performed manually by two human coders after initial discussion and agreement of the coding scheme. In total 5,961 tweets relating to Storm Brian were manually categorized. A subsample of 100 randomly chosen tweets from the filtered tweet data was used for an inter-coder reliability check. Cohen's kappa (κ) was used to determine the agreement between the two coders' judgement on the category of each tweet in the subsample. There was near perfect agreement between the two coders with $\kappa = 0.889$, $p < 0.0005$. This provided confidence in the categorization coding scheme used.

Note that both text and pictures in tweets were used to assign a category, but not emojis as these were removed from the dataset to simplify text analysis processes.

3 | RESULTS

3.1 | Combined time series plot

Tweet counts in the filtered datasets for wind and storm names were plotted over time (Figure 5). The time period for each storm is also shown. Peaks in the volume of tweets coincide with the (UK Met Office recorded) date of impact of storms shown in Table 1. Peaks in the volume of wind tweets also coincide with peaks in the volume of storm name tweets.
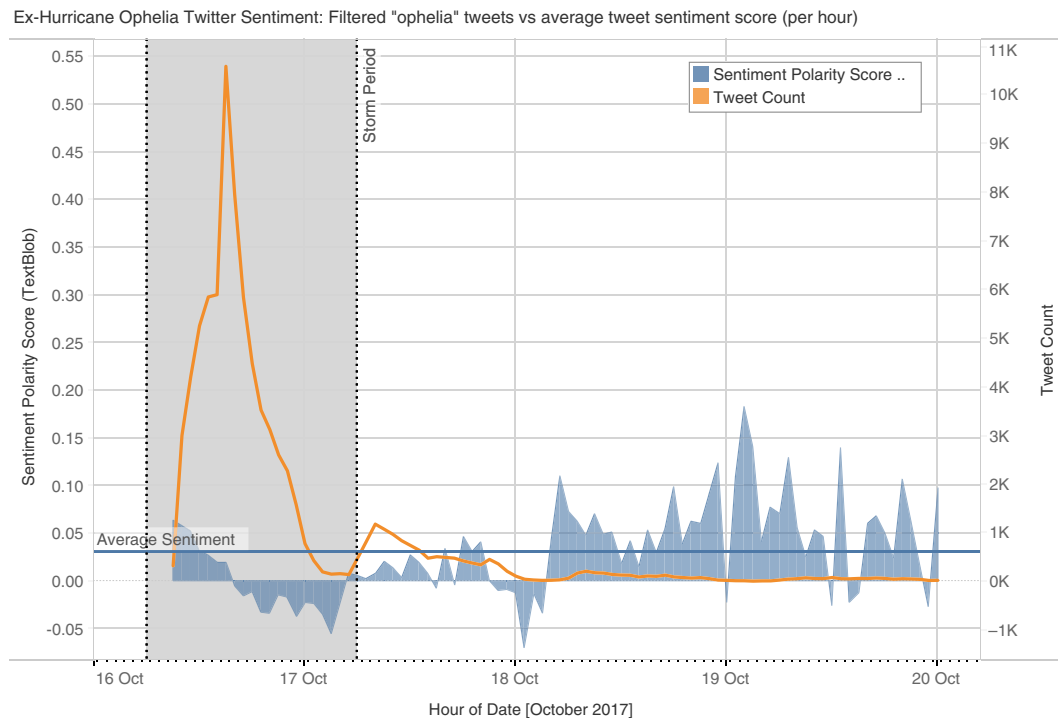
FIGURE 6 Sentiment polarity score for "Ophelia" tweets versus tweet count—line graph shows tweet count; area graph shows sentiment polarity score, aggregated over 2 hr windows. The period of the storm is shown by the grey shaded bar. There is a clear trend in sentiment, which drops during the storm period and then rises following the storm; see also Figure 7. Due to the smaller numbers of tweets, the variance of the sentiment scores increases significantly with time

Figure 5 also shows that there were peaks in tweets relating to wind events which occurred at a time when there was not a named storm event (indicated by 'Unnamed Wind Event(s)' in the figure). Of the 12 peaks in wind speed not attributed to a named storm event, manual inspection of the time series identifies that four of these peaks correspond to peaks in wind tweet volume while eight appear not to. This shows that there were wind-related events being talked about on Twitter at these times and could suggest that the weather was sufficiently windy to generate discussion on Twitter, but not enough for a named storm event. This shows that social media may have some success in detecting smaller wind events that are not named storms.

The storms which saw the greatest wind speed and impacts (Brian, Caroline, Eleanor, Emma) also appear to have larger volumes of tweets than the lesser known/less impactful storms (Dylan, Fionn, Georgina).

3.2 | Sentiment

To understand the emotional response to storm events during the period of the storm, the average sentiment by hour was plotted against the tweet volume over time (Figure 6). For ex-Hurricane Ophelia there is a very clear

drop in sentiment (i.e. tweets become less positive and even negative) during and following the peak of tweet activity, before rising again after the storm has passed.

The distribution of sentiment in filtered tweets is shown as a histogram of average hourly sentiment in each of the Twitter collections (Figure 7). Average sentiment of tweets in the United Kingdom during 2017 was shown in another study (using the same sentiment analysis methods) to be 0.13 (Arthur and Williams, 2018); this reference value is shown in Figure 7 for comparison. For each tweet collection the distribution of tweet sentiment peaks around an average sentiment score lower than the UK average sentiment. The tweet collection with the lowest average sentiment is the Storm Names collection, with the Wind and Precipitation collections showing relatively higher values, albeit still below the UK baseline. This suggests that wind and rain have an adverse effect on sentiment, with more extreme weather (storms) associated with more extreme low sentiment.
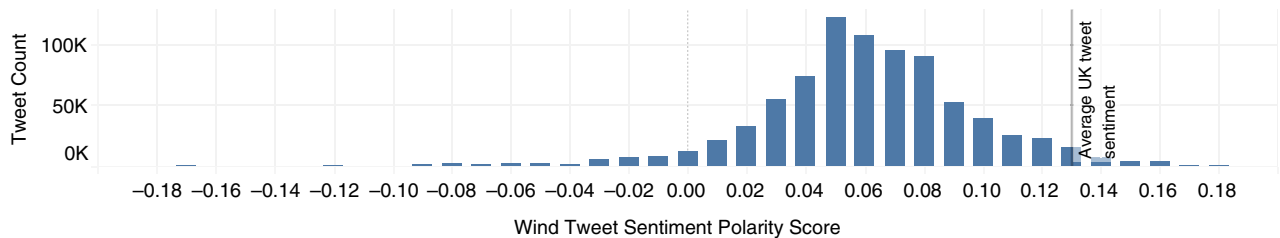
3.3 | Content analysis

For each storm, filtered named storm tweets in the day before, during and after each named storm event were manually reviewed and categorized. The results for Storm

Precipitation Tweet Sentiment
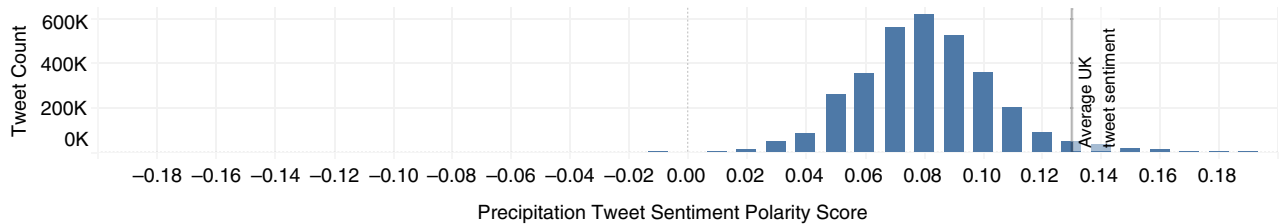
Wind Tweet Sentiment
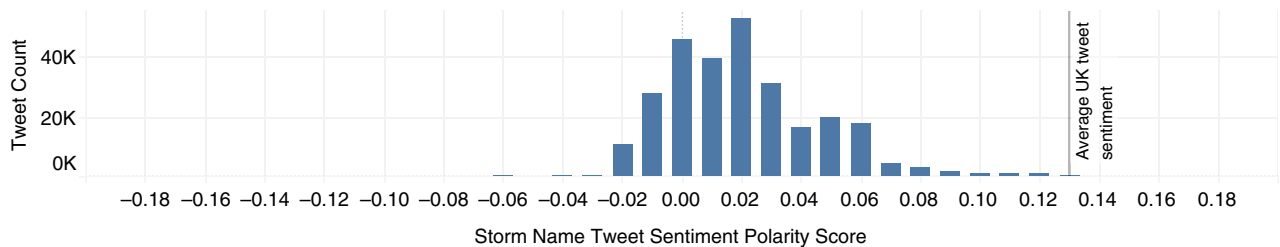

Storm Names Tweet Sentiment

FIGURE 7 Average tweet sentiment score *per* hour histograms for Precipitation, Wind and Storm Name collections. Average sentiment score is normally distributed around a mean average sentiment score. The average UK tweet sentiment (Arthur and Williams, 2018) is shown for comparison

Brian from October 20, 2017 to October 22, 2017 are shown in Figure 8. Similar patterns were observed for other named storms (data not shown). There is a clear temporal trend to the types of content posted by Twitter users as the storm passes through. In early stages, warnings are prevalent, but these show a distinct drop in volume as the main effects of the storm begin to be felt (in the early hours of October 21, 2017). In contrast, tweets relating to observations of the weather occurring and reports of damage/disruption begin to increase as the storm passes through. News reports also increase in frequency in the day after the storm. The level of humour expressed throughout the storm period is somewhat more consistent, remaining around 25% of tweets. Tweets categorized as 'other' include tweets which cannot be categorized under any of the other headings, e.g. commentary on sports results, business advertising, very short tweets with no information. There appears to be no obvious trend in volumes of these tweets.

In terms of tweets providing information on social impacts of the storm, those tweets categorized as damage

or disruption are likely to provide information on the specific impacts experienced by Twitter users. For the example of Storm Brian in Figure 8, 1,020 tweets were categorized as damage or disruption. This means that approximately 17% of filtered tweets for Storm Brian provide information on impacts ranging from damage to property, road closures and power outages.

4 | DISCUSSION

The widespread use of Twitter during extreme weather events, such as named storms in the United Kingdom and Ireland, has created an opportunity to use this rich data source to find useful information. In particular, it offers a potential "social sensing" mechanism by which observations of social impacts of extreme weather can be gathered and measurements which are not available from traditional meteorological observations. The demand for such information is evidenced by the recent rise in impact-led forecasting across the meteorological sciences.

FIGURE 8 Categorized tweets by date/hour for Storm Brian in the period leading up to, during and the day after the storm passed over the United Kingdom and Ireland. Tweets are categorized and plotted as a percentage of all tweets in that hour to account for the expected variation in tweet volumes over each 24 hr period. The number of tweets in that hour is also shown by the line graph

This study presents an analysis of data collected from Twitter during the 2017/2018 storm season in the United Kingdom and Ireland. Various computational techniques were used to filter and extract only those tweets of relevance to wind, precipitation and named storm events. The volume of storm-related weather (wind/rain) tweets increases substantially during storm events. Tweets referring to named storms, after careful filtering to exclude irrelevant content, show clear spikes of activity corresponding to the storm event. Analysis of content shows systematic trends in both sentiment and topics expressed in tweets relating to storms.

Sentiment analysis of tweet content showed clear and consistent emotional impacts of named storms. Average sentiment in weather-related tweets during a named storm event was much less positive than the expected baseline for "normal" Twitter activity. Consistent across multiple storms, collective sentiment was shown to fall significantly as the extreme weather associated with the storm begins to be experienced, before recovering after the storm passes. Furthermore, sentiment is consistently lower in tweets relating to storms than in tweets about wind or rain; however, sentiment for all these weather conditions is lower than the baseline expectation. While sentiment analysis is a crude measure of the psychological aspects of extreme weather, the strength and consistency of the results shown here suggest that these weather events have a substantive adverse impact on social wellbeing.

Categorization of filtered tweets based on their topic and/or content showed another consistent pattern in the type of information being posted on Twitter during the period of a named storm weather event. In the period leading up to a storm it was found that tweets were mainly giving warnings and information about potential impacts. During the storm, tweets contain information about how people are being affected by the storm, such as tweets on disruption and damage. After the storm, tweets continue to report observations and damage/disruption, but also begin to share links to news reports covering the storm. Surprisingly, the proportion of tweets categorized as "humour" remains quite consistently large throughout the period of a storm, with many tweets making light of the given name of each storm and sharing humorous comments about its impacts, rather than commenting directly on the weather. The patterns shown here suggest that further investigation of content might

allow robust measurements of damage and disruption associated with storm events, with some refinements to the method to control for noise and bias. Common sources of noise and bias in social media data include linguistic variation (e.g. regional dialect, slang), tangential content (e.g. tweets related to the storm but not its direct impact, i.e. humour, other) and tweets providing misleading or false information. This kind of impact measurement is hard to obtain by other methods and has clear value for validation of weather hazard impact models. Combined with the location inference method this could be developed to provide information on both how and where the biggest impacts as a result of the storm are experienced.

An interesting finding of this study is the existence of peaks of Twitter activity relating to wind and precipitation that are not related to named storm events. Inspection shows that these peaks reflect genuine discussion of weather conditions, showing high levels of public engagement and concern with weather, similar in some cases to those observed for named storms. This finding may have implications for the design of storm-naming systems and wider understanding of when public information should be issued by meteorological agencies.

There are a number of methodological caveats and limitations to this study. After filtering tweets for relevance to storm events, there were relatively small numbers of tweets retained in the data collections for some of the named storms. The relatively small size of the dataset in these cases makes it difficult to identify patterns in tweet discussion confidently.

With regard to sentiment analysis, the tool used in this study (TextBlob) has a predefined training corpus based on a dataset of movie reviews. Therefore, it is likely that there may be some uncertainty over the accuracy of some of the sentiment scores assigned to tweets in the storm dataset. To enhance the sentiment analysis of tweets relating to an extreme weather event, it is suggested that a bespoke training corpus based on example tweets from the filtered dataset in this study be created to identify positivity and negativity in tweets relating to the weather. This would provide more confidence in the relevance of the data being used for sentiment scoring.

Aside from improvements to the methods used here, future work might increase understanding of the power and scope of social sensing for weather hazard/impact monitoring by looking at content in different ways. An obvious extension to the work performed in this study is to go into further depth regarding the identification of particular kinds of hazard and/or impact, for example by separating travel disruption from damage to property from risks to health. Whether this approach can provide

accurate quantification in terms of counting instances of particular impacts is an open research question. The results reported here suggest that clear patterns can be obtained at a reasonable level of granularity. An extension might consider validation of each tweet against the observed weather conditions for that date/time and grid square; this might allow epidemiological study of how different weather conditions (both chronic and episodic) affect behaviour and wellbeing, alongside the more straightforward opportunity to validate the accuracy of individual users as social sensors. Related to impact-based weather forecasting, the volume of activity generated by events categorized as red/amber/yellow might be analysed to study the match between severity judged by meteorological organizations and severity as reported by the general population.

What this study has shown is how social media can be used to provide another layer of information about the social impacts of extreme weather, both emotionally and physically, spatially and temporally, in a way that has not been available before. Being able to determine more specific information about social impacts not available in weather observation data means that impact-based warnings for the public can be tailored towards high impact events. It also provides a method of validation of information provided by meteorological agencies in weather warnings for the public.

ORCID

M. Spruce https://orcid.org/0000-0002-1744-8165

REFERENCES

Alam, F., Ofli, F., Imran, M. and Aupetit, M. (2018) *A Twitter tale of three hurricanes: Harvey, Irma, and Maria*. Available at: http://arxiv.org/abs/1805.05144 [Accessed 28th February 2019].

Arthur, R., Boulton, C.A., Shotton, H. and Williams, H.T.P. (2018) Social sensing of floods in the UK. *PLoS One*, 13(1), e0189327. https://doi.org/10.1371/journal.pone.0189327.

Arthur, R. and Williams, H. T. P. (2018) *The human geography of Twitter*. Available at: http://arxiv.org/abs/1807.04107 [Accessed 7th March 2019].

Baylis, P., Obradovich, N., Kryvasheyeu, Y., Chen, H., Coviello, L., Moro, E., Cebrian, M. and Fowler, J.H. (2018) Weather impacts expressed sentiment. *PLoS One*, 13(4), e0195750. https://doi.org/10.1371/journal.pone.0195750.

Boulton, C. A., Shotton, H. and Williams, H. T. P. (2016) *Using social media to detect and locate wildfires*. Available at: https://

ore.exeter.ac.uk/repository/handle/10871/28583 [Accessed 18th September 2018].

Brouwer, T., Eilander, D., van Loenen, A., Booij, M.J., Wijnberg, K. M., Verkade, J.S. and Wagemaker, J. (2017) Probabilistic flood extent estimates from social media flood observations. *Natural Hazards and Earth System Sciences*, 17(5), 735–747. https://doi. org/10.5194/nhess-17-735-2017.

Caragea, C., Squicciarini, A., Stehle, S., Neppalli, K. and Tapia, A. (2014) *Mapping moods: geo-mapped sentiment analysis during Hurricane Sandy*. Available at: http://asquicciarini.ist.psu.edu/ pdf/iscram14.pdf [Accessed 18th September 2018].

Cervone, G., Sava, E., Huang, Q., Schnebele, E., Harrison, J. and Waters, N. (2016) Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study. *International Journal of Remote Sensing*, 37(1), 100–124. https://doi.org/10.1080/01431161.2015.1117684.

Cowie, S., Arthur, R. and Williams, H.T.P. (2018) @choo: tracking pollen and hayfever in the UK using social media. *Sensors*, 18 (12), 4434. https://doi.org/10.3390/s18124434.

Dredze, M., Paul, M.J., Bergsma, S. and Tran, H. (2013) *Carmen: a Twitter geolocation system with applications to public health*. Available at: https://pdfs.semanticscholar.org/9bc4/ 6fb12f2c7fae0e9e56e734e6efb9ca07fd98.pdf [Accessed 18th September 2018].

Guan, X. and Chen, C. (2014) Using social media data to understand and assess disasters. *Natural Hazards*, 74(2), 837–850. https://doi.org/10.1007/s11069-014-1217-1.

Halse, S., Montarnal, A., Tapia, A. and Bénaben, F. (2018) *Bad weather coming: linking social media and weather sensor data*. In: Tomaszewski B, Boersma K. (Eds.) ISCRAM 2018: *Proceedings of the 15th International Conference on Information Systems for Crisis Response and Management, ISCRAM 2018*. Rochester Institute of Technology, pp. 507–515.

Hannak, A., Anderson, E., Feldman Barrett, L., Lehmann, S., Mislove, A. and Riedewald, M. (2012) *Tweetin' in the rain: exploring societal-scale effects of weather on mood*. Available at: www.aaai.org [Accessed 24th January 2019].

Hutto, C.J. and Gilbert, E. (2014) *VADER: a parsimonious rule-based model for sentiment analysis of social media text*. ICWSM. Available at: https://www.semanticscholar.org/paper/VADER% 3A-A-Parsimonious-Rule-Based-Model-for-Analysis-Hutto-Gilbert/ a6e4a2532510369b8f55c68f049ff11a892fefeb [Accessed 16th November 2018].

IPCC. (2014) In: Field, T.E.B., Barros, V.R., Dokken, D.J., Mach, K. J., Mastrandrea, M.D., Chatterjee, A.N.L.M., Ebi, K.L., Estrada, Y.O., Genova, R.C., Girma, B., Kissel, E.S., MacCracken, L.L.W.S. and Mastrandrea, P.R. (Eds.) *Climate Change 2014: Impacts, Adaptation and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge and New York, NY: Cambridge University Press.

Kim, J. and Hastak, M. (2018) Social network analysis: characteristics of online social networks after a disaster. *International Journal of Information Management*, 38(1), 86–96. https://doi. org/10.1016/J.IJINFOMGT.2017.08.003.

Kryvasheyeu, Y., Chen, H., Obradovich, N., Moro, E., Van Hentenryck, P., Fowler, J. and Cebrian, M. (2016) Rapid assessment of disaster damage using social media activity. *Science Advances*, 2(3), e1500779. https://doi.org/10.1126/sciadv. 1500779.

Lachlan, K.A., Spence, P.R., Lin, X. and Del Greco, M. (2014) Screaming into the wind: examining the volume and content of tweets associated with Hurricane Sandy. *Communication Studies*, 65(5), 500–518. https://doi.org/10.1080/10510974.2014. 956941.

Li, H., Jadidi, Z., Chen, J. and Jo, J. (2019) The use of machine learning for correlation analysis of sentiment and weather data. In: Kim JH. et al (Eds.), *Robot Intelligence Technology and Applications 5*. Cham: Springer International Publishing, pp. 291–298. https://doi.org/10.1007/978-3-319-78452-6_25

Li, J., Wang, X. and Hovy, E. (2014) *What a nasty day: exploring mood–weather relationship from Twitter*. Available at: http:// arxiv.org/abs/1410.8749 [Accessed 18th September 2018].

Loria, S. (2010) *TextBlob: simplified text processing*. Available at: https://textblob.readthedocs.io/en/dev/ [Accessed 1st March 2018].

McGrath, R. (2013) *Twython — Twython 3.6.0 documentation*. Available at: https://twython.readthedocs.io/en/latest/ [Accessed 26th March 2019].

McMahan, A.C., Grover, C.N. and Vignola, F.E. (2013) *Exceedance – an overview | ScienceDirect Topics*. Available at: https://www. sciencedirect.com/topics/engineering/exceedance [Accessed 4th April 2019].

Mendes, P.N., Jakob, M., Garcia-Silva, A. and Bizer, C. (2011) DBpedia Spotlight: shedding light on the web of documents. In: *Proceedings of the 7th International Conference on Semantic Systems (I-Semantics)*. Available at: https://www.dbpedia-spotlight.org [Accessed 4th April 2019].

Morss, R.E., Demuth, J.L., Lazrus, H., Palen, L., Barton, C.M., Davis, C.A., Snyder, C., Wilhelmi, O.V., Anderson, K.M., Ahijevych, D.A., Anderson, J., Bica, M., Fossell, K.R., Henderson, J., Kogan, M., Stowe, K. and Watts, J. (2017) Hazardous weather prediction and communication in the modern information environment. *Bulletin of the American Meteorological Society*, 98(12), 2653–2674. https://doi.org/10.1175/BAMS-D-16-0058.1.

Morstatter, F., Pfeffer, J., Liu, H. and Carley, K.M. (2013) Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's firehose. In: *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*. AAAI press, 400–408.

Rossi, C., Acerbo, F.S., Ylinen, K., Juga, I., Nurmi, P., Bosca, A., Tarasconi, F., Cristoforetti, M. and Alikadic, A. (2018) Early detection and information extraction for weather-induced floods using social media streams. *International Journal of Disaster Risk Reduction*, 30, 145–157. https://doi.org/10.1016/j. ijdrr.2018.03.002.

Sakaki, T., Okazaki, M. and Matsuo, Y. (2010) Earthquake shakes Twitter users. In: *Proceedings of the 19th International Conference on World Wide Web – WWW '10*. New York, NY: ACM Press, p. 851. https://doi.org/10.1145/1772690.1772777

Schulz, A., Hadjakos, A., Paulheim, H., Nachtwey, J. and Uhlhäuser, M. (2013) A multi-indicator approach for geo-localization of Tweets. In: *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*. Palo Alto, CA: AAAI press, pp. 573–582.

Spence, P.R., Lachlan, K.A., Lin, X. and del Greco, M. (2015) Variability in Twitter content across the stages of a natural disaster: implications for crisis communication. *Communication Quarterly*, 63, 171–186. https://doi.org/10.1080/01463373.2015.1012219.

Statista. (2017) *Number of social media users worldwide 2010–2021 | Statista*. Available at: https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/ [Accessed 18th September 2018].

Taylor, A.L. (2018) Communicating high impact weather: improving warnings and decision making processes. *International Journal of Disaster Risk Reduction*, 30, 1–4. https://doi.org/10.1016/J.IJDRR.2018.04.002.

Tkachenko, N., Jarvis, S. and Procter, R. (2017) Predicting floods with Flickr tags. *PLoS One*, 12(2), e0172870. https://doi.org/10.1371/journal.pone.0172870.

UK Snow Map. (2010) #uksnow Map. Available at: https://uksnowmap.com/ [Accessed 19th March 2019].

Wang, D., Abdelzaher, T. and Kaplan, L. (2015a) *Social Sensing: Building Reliable Systems on Unreliable Data*, 1st edition. San Francisco, CA: Morgan Kaufmann Publishers. https://doi.org/10.1016/C2013-0-18808-3.

Wang, Y., Wang, T., Ye, X., Zhu, J. and Lee, J. (2015b) Using social media for emergency response and urban sustainability: a case study of the 2012 Beijing rainstorm. *Sustainability*, 8(12), 25. https://doi.org/10.3390/su8010025.

Williams, S.A., Terras, M.M. and Warwick, C. (2013) What do people study when they study Twitter? Classifying Twitter related academic papers. *Journal of Documentation*, 69(3), 384–410. https://doi.org/10.1108/JD-03-2012-0027.

WMO. (2015) *WMO guidelines on multi-hazard impact-based forecast and warning services*. Geneva, Switzerland. Available at: https://library.wmo.int/pmb_ged/wmo_1150_en.pdf [Accessed 26th March 2019].

WOW. Met Office Weather Observations Website. (2019) Available at: https://wow.metoffice.gov.uk/ [Accessed 19th March 2019].

Wu, D. and Cui, Y. (2018) Disaster early warning and damage assessment analysis using social media data and geo-location information. *Decision Support Systems*, 111, 48–59. https://doi.org/10.1016/j.dss.2018.04.005.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Spruce M, Arthur R, Williams HTP. Using social media to measure impacts of named storm events in the United Kingdom and Ireland. *Meteorol Appl.* 2020;27: e1887. https://doi.org/10.1002/met.1887