Secondary Publication

Zander, Kerstin K.; Rieskamp, Jonas; Mirbabaie, Milad; u. a.

Responses to heat waves : what can Twitter data tell us?

Date of secondary publication: 30.09.2025

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-110550x

Primary publication

Zander, Kerstin K.; Rieskamp, Jonas; Mirbabaie, Milad; u. a. (2023): Responses to heat waves : what can Twitter data tell us?, in: Natural Hazards, Dordrecht [u.a.]: Springer, Vol. 116, Nr. 3, pp. 3547–3564, doi: 10.1007/s11069-023-05824-2.

Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available under a Creative Commons license.

The license information is available online:
https://creativecommons.org/licenses/by/4.0/legalcode

Natural Hazards (2023) 116:3547–3564
https://doi.org/10.1007/s11069-023-05824-2

ORIGINAL PAPER

Responses to heat waves: what can Twitter data tell us?

Kerstin K. Zander[1] · Jonas Rieskamp[2] · Milad Mirbabaie[2] · Mamoun Alazab[3] · Duy Nguyen[1]

Received: 22 July 2022 / Accepted: 5 January 2023 / Published online: 13 January 2023
© The Author(s) 2023

Abstract
Increasing average temperatures and heat waves are having devasting impacts on human health and well-being but studies of heat impacts and how people adapt are rare and often confined to specific locations. In this study, we explore how analysis of conversations on social media can be used to understand how people feel about heat waves and how they respond. We collected global Twitter data over four months (from January to April 2022) using predefined hashtags about heat waves. Topic modelling identified five topics. The largest (one-third of all tweets) was related to sports events. The remaining two-thirds could be allocated to four topics connected to communication about climate-related heat or heat waves. Two of these were on the impacts of heat and heat waves (health impacts 20%; social impacts 16%), one was on extreme weather and climate change attribution (17%) and the last one was on perceptions and warning (13%). The number of tweets in each week corresponded well with major heat wave occurrences in Argentina, Australia, the USA and South Asia (India and Pakistan), indicating that people posting tweets were aware of the threat from heat and its impacts on the society. Among the words frequently used within the topic 'Social impacts' were 'air-conditioning' and 'electricity', suggesting links between coping strategies and financial pressure. Apart from analysing the content of tweets, new insights were also obtained from analysing how people engaged with Twitter tweets about heat or heat waves. We found that tweets posted early, and which were then shared by other influential Twitter users, were among the most popular. Finally, we found that the most popular tweets belonged to individual scientists or respected news outlets, with no evidence that misinformation about climate change-related heat is widespread.

Keywords Disaster communication · Extreme heat · Machine learning · Resilience · Risk perception · Social media · Topic modelling

✉ Kerstin K. Zander
kerstin.zander@cdu.edu.au

[1] Northern Institute, Charles Darwin University, Brinkin, Australia

[2] Paderborn University, Paderborn, Germany

[3] Faculty of Science and Technology, Charles Darwin University, Brinkin, Australia

1 Introduction

Globally, heat is a public health concern (Capon et al. 2019; Ebi et al. 2021; Jay et al. 2021) and "one of the most underappreciated hazards of climate change" (Nature Editorial 2021). Extreme heat has led to the deaths of 30 million people in the last three decades (Vicedo-Cabrera et al. 2021). The 2003 European heat wave, for example, killed 70,000 people (Robine et al. 2008). Heat affects both rural and urban populations where there is increasing exposure to morbidity and mortality from extreme heat and related heat illnesses (Hatvani-Kovacs et al. 2016; Wouters et al. 2017; Zander et al. 2018). Temperatures in some regions in the world are already exceeding life-threatening thresholds during heat waves and eventually such places will become inhabitable (Horton et al. 2021). Apart from killing people, often the elderly and those with chronic illnesses, heat decreases well-being through mild symptoms such as headaches and fatigue (Zander et al. 2017) and productivity (Zander et al. 2015; Andrews et al. 2018) and disrupts comfort and everyday life. Heat stress can also increase violence and crime (Stevens et al. 2019).

Given that average temperatures are predicted to be 2.4 degrees higher than pre-industrial levels by 2100, almost a full degree above the Paris limit (Climate Action Tracker 2021), and heat waves are set to become more frequent and longer (IPCC 2021), there is increasing recognition of the need for research on how to increase adaptation to the direct effects of heat on human society. Adaptation to heat includes changes in individual behaviour (e.g. drinking, resting, shifting activities to cooler times of the day), improving infrastructure at both household level and in public places, and adjusting institutional policies (Turek-Hankins et al. 2021). While there is research on how people are learning to cope with increased heat (Bakhsh et al. 2018; Jay et al. 2021; Zander et al. 2021), adaptation to heat is ultimately limited by physiological responses of human organism (Sherwood and Huber 2010; Asseng et al. 2021). One adaptation, therefore, is migration to cooler places to avoid negative health implications and a decline in well-being (Cattaneo and Peri 2016; Zander and Garnett 2020).

Notwithstanding the impacts of heat at a local level, it is unclear how many people suffer from heat globally. Studies on heat adaptation are limited and mostly confined to developed countries where heat is mostly treated as a health issue, particularly in urban areas (Turek-Hankins et al. 2021). Moreover, data are scarce since people do not necessarily report to medical doctors or hospitals when they only feel mild symptoms. To overcome the data shortage, researchers have used data from ambulance call-outs (Guo 2017) or occupational accidents (Varghese et al. 2019) in order to assess the relation between heat-related illnesses and ambient temperatures, but this has also required that heat-related illness be sufficiently severe to warrant medical intervention. A limited number of studies have collected primary data on heat adaptation through surveys (e.g. Hatvani-Kovacs et al. 2016; Rauf et al. 2017; Zander et al. 2018; Beckmann et al. 2021).

In this study, we aimed to explore the potential utility of a large and hitherto neglected dataset-freely available social media data—for communications about heat waves. One of the most frequently used social media platforms by researchers is Twitter (Lupton 2014). Twitter provides a certain percentage of all tweets, the messages sent on Twitter, including associated meta-data, free of charge to be used for non-commercialised research via a free accessible API. We collected Twitter data to explore (1) what kind of useful information social media data can provide from communications about heat waves, (2) what we can learn from Twitter conversations about how people are affected by and cope with heat waves, and (3) who engages and to what extent in conservations about heat waves.

Social media is now the most common means for disseminating information rapidly about disasters (Abedin and Babar 2018). Increasingly, the public turns to social media during crises to share the latest information, while agencies (governmental, non-governmental or non-institutional) use social media for formal communication of updated information, warnings and specific advice or give directions to the public and stakeholders such as volunteers (Bird et al. 2012; Abedin and Babar 2018). Many studies have shown that the number of social media messages from people close to events increases substantially immediately before and during their occurrence. This suggests that social media use, and in particular Twitter messages, are reliable indicators of public awareness (Bird et al. 2012; Kryvasheyeu et al. 2016; Wang et al. 2016; Silver and Andrey 2019). Globally, researchers have taken advantage of the increasing volume of social data provided by the public to improve understanding of how people respond to (Wang et al. 2016; Reynard and Shirgaokar 2019; Martín et al. 2020) and recover from (Jamali et al. 2020) sudden-onset disasters such as floods, bushfires, earthquakes and cyclones/hurricanes. Social media has also been used for early detection and tracking of such events (Wang et al. 2016) and to assess the level of damage afterwards (Kryvasheyeu et al. 2016).

Our paper contributes to extending the applications of disaster related Twitter studies to heat waves which, although often deadly, are rarely perceived to be as severe and destructive as are some other natural hazards. To the best of our knowledge, social media analysis has been employed to explore how people are impacted by heat waves in only a handful of country-specific case studies (Japan: Murakami et al. 2020; India: Cecinati et al. 2019; Italy: Grasso et al. 2017; and the US: Jung and Uejio 2017; Moore et al. 2019). Our paper differs to this existing body of the literature in that we collected and analysed Twitter data globally over a period of two months. We also used an innovative method of topic modelling designed for short text such as Twitter tweets, a Dirichlet multinomial mixture model (GSDMM), which extends the commonly applied latent Dirichlet allocation (LDA) model.

2 Data and methods

2.1 Twitter data collection

The data in this research were collected using snscrape, a pre-built scraper for social networking services in Python. The TwitterSearchScraper method of snscrape's Twitter module (JustAnotherArchivist 2020) allows users to retrieve historical tweets based on a set of keywords. Here, we searched for historical heat-related tweets applying two inclusion criteria. Firstly, the tweets must contain any of the following ten keywords: 'heat wave', 'heatwave', 'sweltering heat', 'unbearable heat', 'torrid heat', 'heat warning', 'heat watch', 'heatstroke', 'heat stroke' and 'extreme temperature'. Secondly, these tweets must have been published over a period of four months from January through to April of 2022. The reason we chose this four-month time frame was that the global temperature during January–April 2022 was the fifth highest in the 143-year record (NOAA 2022; Figure S1 in Supplementary Materials) with above-average temperatures observed in Asia, South America, the Atlantic and Pacific Oceans and Australia.

Among all information available in the Tweet object of TwitterSeachScraper, we selected 13 variables to download (Table S1 in Supplementary Materials). The data collection process resulted in a raw data set containing 62,920 rows of data which corresponded

to 62,920 tweets that matched the two criteria described above. Figure 1 outlines the data extraction, processing, analysis and visualisation processes.

2.2 Data processing

Before any data analysis and topic modelling were performed, data processing was required to transform the raw data into an appropriate format. We applied common standards for processing Twitter text data with some adjustments to suit our data set (Xue et al. 2020). The detailed steps are provided in the supplementary materials.

In order to remove noise from the data set, tweets with irrelevant content were discarded after an initial analysis. We found that entertainment was a frequently occurring topic which we removed by filtering the words 'film', 'music' and 'movie'. This step removed another 1302 tweets, which brought the total number of tweets removed during data processing to 12,656 tweets. The final data set contained 50,264 tweets with a total of 351,677 words which were then used as inputs for the topic modelling.

2.3 Topic modelling and coherence score

Topic modelling is a popular and important unsupervised machine learning and natural language processing (NLP) algorithm. It is a tool designed to analyse and extract hidden topics from large datasets, in this case, Twitter tweets, which cannot easily be read and analysed objectively without a machine. Topic models classify articles according to their key topics and word frequencies and thereby reveal hidden themes (Chang et al. 2009). Articles within each topic have similar words that occur frequently together (Blei et al. 2003). However, the researchers' knowledge and experience are also needed to interpret the model outcome into meaningful themes (Karami et al. 2018).

There are different approaches to topic modelling. A common topic modelling approach is latent Dirichlet allocation (LDA), an unsupervised machine learning technique that generates topics based on the pattern of occurrence or co-occurrence of words in a document (Jacobi et al. 2016). LDA, however, does not perform well when the texts are very short,

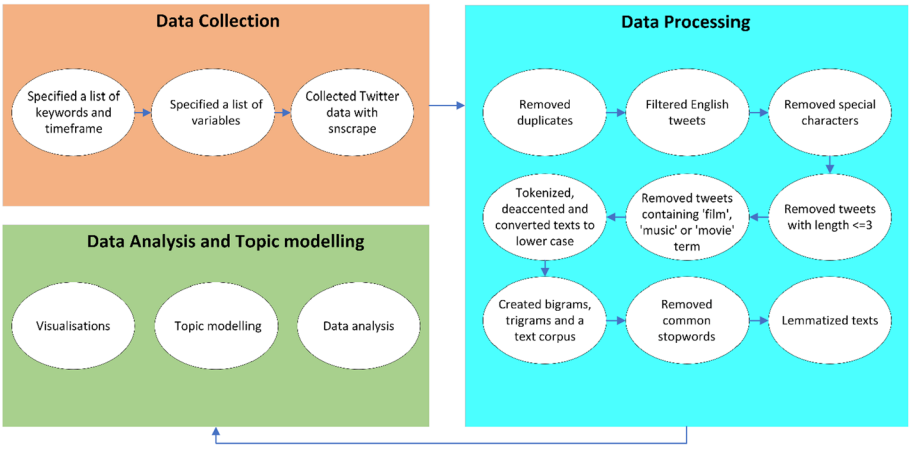

Fig. 1 Data collection, processing and analysis

such as is the case with Twitter tweets, due to the lack of information on co-occurrence of words (Qiang et al. 2020). We therefore used Gibbs sampling algorithm for the Dirichlet multinomial mixture (GSDMM) proposed in Yin and Wang (2014). Not only does GSDMM achieve a good balance in both completeness and homogeneity, but the model also obtains better clustering results in short text topic modelling tasks compared to LDA and other traditional topic modelling techniques (Agarwal et al. 2020; Qiang et al. 2020).

The GSDMM topic modelling algorithm has four parameters, namely k, α, β, and the number of iterations. Unlike other topic models, the k in GSDMM represents the maximum number of topics available during the clustering process, and it is set larger than the true number of topics. α and β are two real numbers with values set between 0 and 1. α relates to the probability that a document will be assigned to a new topic, while β relates to the probability that a document will be assigned to a topic with similar words (Yin and Wang 2014). The performance of the GSDMM model is stable after only five iterations (Yin and Wang 2014), but the model gives the best coherence score with fifty iterations (Omurca et al. 2021), which we used here. The implementation of GSDMM model in Python can be found at a public GitHut repository named rwalk/gsdmm (JustAnotherArchivist 2020).

Unlike α, β and the number of iterations, setting the number of topic (k) for GSDMM is less straightforward, since the true number of topic number is unknown. To determine the number of topics for the GSDMM model, the coherence score for seventeen different values of k, starting with $k=4$ and ending with $k=18$, was calculated by analysing the frequency of the hundred and fifty most used words and phrases (see Röder et al. 2015). With each value of k, the GSDMM model was run ten times, and the coherence score for each iteration was used to compute an average.

We visualised the outcomes of the GSDMM topic model with word clouds and used different metrics, including the number of tweets, retweets, likes, replies, as well as the number of uniform resource locators (URLs) included and the number of tweets which used hashtags, to compare the different topics. To assign the most likely topic to each tweet, we used the pre-built *MovieGroupProcess's choose_best_label* function.

2.4 Measuring the level of activity, popularity and influence of Twitter accounts

We applied a range of different metrics to measure the popularity, influence, and activity of different Twitter accounts. The most straightforward metric to measure activity is the TweetRank metric (Nagmoti et al. 2010), which is defined by the number of tweets posted by a Twitter account over a period of time. Measuring the popularity of Twitter accounts was done using the in-degree metric (Cha et al. 2010), as well as the FollowerRank (FR) (Nagmoti et al. 2010), which is the normalised version of the traditional in-degree metric. The FR of an account is defined by the number of its followers divided by the sum of the number of its followers and the number of its followees. The FR ranges from 0 and 1. Accounts with a FR close to 0 are not considered popular, while users with a FR close to 1 are extremely popular (Primo et al. 2021).

Compared to measuring activity and popularity, measuring the degree of influence of a Twitter account is more challenging since there is no unified definition for an influential user. There is, however, a large number of possible measures to estimate influence (Riquelme and González-Cantergiani 2016). We opted for three metrics proposed by Cha et al. (2010) to measure the level of influence of a Twitter account. These three metrics are indegree influence, which considers the number of followers of a Twitter account; retweet

influence, which considers the number of retweets a Twitter account received; and mention influence, which considers the number of times a Twitter account was mentioned by others.

Apart from the aforementioned metrics, we also measured the number of likes, previously termed 'favourites', and the number of replies a tweet received. Liking a tweet is a means to show acknowledgement and agreement, while replying to a tweet is a way for Twitter users to join a conversation and voice their opinions on the topic of discussion (Zhang et al. 2018). Hence, the number of likes and the number of replies can also be used to measure how popular and influential a Twitter account is.

3 Results

3.1 Description of Twitter activity and engagement

3.1.1 Number of heat or heat wave-related tweets

Overall, the number of heat or heat wave-related tweets increased by almost 30% from January to April 2022 (Fig. 2). Compared to January 2022, the number of tweets dropped slightly in February 2022, which can be attributed to the lower-than-average temperatures observed in many regions in the world (NOAA 2022). The number of tweet was the highest in April 2022, reaching 15,467.

The number of heat or heat wave-related tweets corresponded well to the occurrence of several heat waves during this time period. In the second week of January (between 10/01/2022 and 16/01/2022), a heat wave hit several countries in South America, including Argentina (Earth Observatory 2022). This led to a steep increase in the number of tweets containing the words 'argentina', 'buenos', 'buenos aires' or 'south america' compared to the week before the heat wave hit the country. In the third week of January (between 17/01/2022 and 23/01/2022), Western Australia experienced a week-long heat wave with "severe-intensity to extreme-intensity" (Bureau of Meteorology 2022). As a result, the number of tweets containing words such as 'australia', 'perth' or 'western australia' also increased significantly. In the second week of February (between 07/02/2022 and 13/02/2022), a heat wave hit several cities in California, which led to an increase in the number of tweets containing the words 'california', 'san francisco', 'sacramento', 'los angeles' or 'san diego' from five in week 5 to 167 in week 6.

In early March 2022, a heat wave hit several South Asian countries, including Pakistan and India. There was an increase in the number of tweets with the words 'pakistan', 'india'

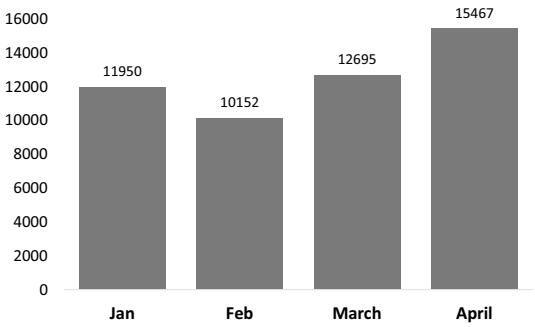

Fig. 2 Number of tweets per month, from January 2022 to April 2022

or 'asia' from the first week of March (week 9), and this number peaked in the last week of April (between 25/04/2022 and 01/05/2022) when India and Pakistan experienced record high temperatures. Compared to the heat waves in other regions earlier in 2022, the heat wave in South Asia gained higher media attention (e.g. The Guardian 2022) and the number of tweets with the words 'pakistan', 'india' or 'asia' peaked at a much higher number of 292.

3.1.2 Activity, popularity and influence of Twitter accounts

The ten most active Twitter accounts were personal accounts belonging to individual users (Table 1). The account with the highest number of heat or heat wave-related tweets was 'heatstroke_x', who posted 595 tweets over the four-month period. Despite the high number of tweets posted, 'heatstroke_x' received minimal engagement from their followers as illustrated through the number of replies, retweets and likes. Other accounts such as 'climateguyw' (120 tweets posted), and 'wsl' (117 tweets posted) had higher numbers of followers engaging with their tweets (Table 1).

Unlike the ten most active accounts, the ten Twitter accounts with the most followers belonged to large organisations. Eight out of the ten accounts with the most followers belonged to news outlets and broadcasting corporations, one account belonged to a sports association (NBA: National Basketball Association), and one to a government official (rashtrapatibhvn). Most of the highly active accounts were from the UK, USA and India (Table 2), three countries with a large English-speaking population. However, despite having many followers, the heat or heat wave-related tweets posted by these accounts did not receive a lot of engagement, as shown by the low number of retweets and likes (Table 2). All these ten accounts were very popular (FR values close to 1.0), meaning that they had many more followers than followees. Although there was no significant difference in the FR values across the ten accounts, the number of followers varied greatly. The two Twitter accounts 'TheEconomist' and 'BBCNews', for example, had the same FR value, but 'TheEconomist' had 83% more followers than 'BBCNews' (Table 2).

The ten accounts with the most retweets also belonged to personal accounts administered by individual Twitter users (Table 3). 'PlatformAdam', a Twitter account managed by a company that specialises in developing environmental data products, had the second-highest number of retweets, despite having posted only one relevant tweet.

Table 1 The ten accounts with the highest number of tweets

Username	Location	Replies	Tweets	Followers	Followees	Retweets	Likes	Mentions
heatstroke_x	N/A	166	595	1955	956	43	1719	228
climateguyw	USA	190	120	11,842	3166	1826	2845	41
wsl	N/A	353	117	1,979,152	1061	1550	10,065	15
sherisperry	USA	0	115	28	30	0	1	0
PaSWHQ	USA	0	111	3	1	0	1	0
QuipmoSurf	N/A	0	109	6846	8,241	2	2	0
SurfNews365	USA	0	106	3660	1,031	1	13	0
isobar_bot	N/A	0	97	9	10	0	0	0
ndmaindia	India	48	65	367,096	3144	281	552	3
LoganStewart51	N/A	0	61	2	18	0	1	0

Table 2 The ten accounts with the most followers

Username	Location	Replies	Tweets	FR	Followers	Followees	Retweets	Likes	Mentions
NBA	USA	122	8	0.99995	40,909,939	1902	334	2410	74
TheEconomist	UK	7	1	0.99999	27,118,183	151	11	18	2
FoxNews	USA	18	1	0.99999	23,161,086	262	14	58	13
rashtrapatibhvn	India	17	1	1.00000	20,558,986	2	241	1059	2
ABC	Australia	60	1	0.99997	17,746,603	450	53	121	3
ndtv	India	8	5	1.00000	17,624,661	15	33	236	12
BBCNews	UK	73	1	0.99999	14,818,570	97	18	78	16
timesofindia	India	8	5	0.99998	14,630,389	320	23	109	14
guardian	UK	4	1	0.99990	10,821,491	1055	13	32	16
BBCSport	UK	16	7	0.99994	10,631,926	644	28	272	5

FR = Follower rank

Table 3 The ten accounts with the highest retweet counts

Username	Location	Replies	Tweets	Followers	Followees	Retweets	Likes	Mentions
irn_2h	Japan	21	1	42,268	49	4508	21,446	0
PlatformAdam	N/A	256	1	7079	657	3301	7665	7
billmckibben	USA	105	1	398,538	1430	2235	4378	4
khanthefatima	India	219	1	73,712	1519	1974	7517	2
climateguyw	USA	190	120	11,842	3166	1826	2845	41
ZeeStudios_	India	13	1	94,259	152	1791	4676	0
wsl	N/A	353	117	1,979,152	1,061	1550	10,065	15
jin_bugs	USA	29	1	2103	1794	1437	5755	0
rameshlaus	India	22	2	743,160	764	1283	3462	0
vamsikaka	India	5	1	572,976	841	1082	2902	0

'PlatformAdam' received 3301 retweets for their tweet posted on April 29 that read: "The current extreme #heatwave in #Pakistan and #India as seen today, on the fourth intense hot day, by #Copernicus #Sentinel3 LST (Land Surface Temperature, not Air!). LST collected on April 29 shows max value exceeding 62 °C/143°F. Gaps due to cloud/snow/nodata. #ClimateEmergency".

Another notable Twitter account in the top ten was 'khanthefatima' with a tweet posted on April 21 about the heat wave in India, reading: "Almost every journalist going to report on Jahangirpuri is returning unwell. The heat is, with no exaggeration, unbearable. Now think of all those who lost their homes today, many of whom would be fasting. Parched and empty stomach, the roof over their head has also been stolen". This tweet received many likes and replies (Table 2).

3.2 Topic modelling development

The α and β parameters for the GSDMM model were set at 0.1 and 0.01, respectively (see Sect. 2.3). To determine the number of topics (k), we estimated multiple GSDMM models. The plot of the average coherence scores for fifteen different values of k showed no major variance (Figure S2 in Supplementary Materials). The GSDMM model achieved the highest average coherence score of 0.469 with $k=6$, closely followed by a 0.462 coherence score with $k=12$. However, upon further inspection, we discovered many semantic similarities between the six clusters of words generated by GSDMM with $k=6$. As a result, five topics were selected for our final GSDMM model.

3.3 Topic analysis and labelling

To label each topic, we investigated the one hundred most frequently used words within. Topic 1 was the largest cluster with 16,930 tweets (33%), followed by topic 2 with 10,299 tweets (20%), topic 4, 3 and 5 (Fig. 3). Four of the five topics (66% of all tweets) could be assigned to communications about climate-related heat, while tweets belonging to topic 1 could not.

In topic 1, 'watch' and 'heat' were the two most frequently used words. Words such as 'game', 'play, and 'team' being the third, fourth and fifth most frequently used words suggested that this topic was about sports events and labelled accordingly. Words such as 'basketball', along with 'buck, 'bull', 'hawk' and 'raptor', names of major basketball teams in the NBA, indicated that many tweets in this topic were about basketball. We could not detect a reference to temperatures, and 'heat' was mostly used in relation to a basketball team called 'Miami Heat'.

Topic 2 was the second largest cluster with 10,299 tweets (20%). Similar to topic 1, 'heat' and 'watch' were among the most frequently used words in this cluster. However, unlike tweets related to sport events (cluster 1), tweets in topic 2 also included words such as 'unbearable', 'sweltering', 'feel' and 'die', indicating that people talk about how the heat strained on them. Other frequently used words included 'sleep', 'work', 'body', 'sweat', 'heatstroke', 'people' and 'life', which indicated that the tweets in this topic were about the impacts of heat on the health and well-being of people. Hence, the topic was named

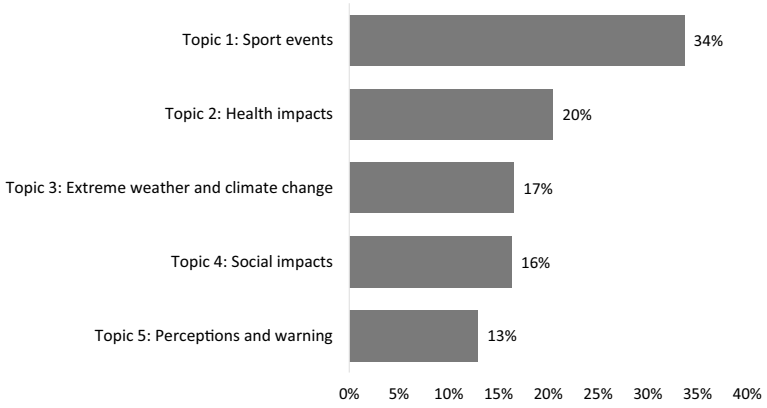

Fig. 3 Topic distribution and labelling

'Health impacts'. Moreover, words such as 'wear' and drink' suggested people talking about actions needed to keep healthy during extreme heat.

In topic 3 (17%), 'extreme, 'heat' and 'temperature' were the three words with the highest frequencies. Words such as 'climate', 'climate_change', 'wildfire', 'drought', 'flood', 'water' and 'record' were also frequently used, suggesting that this topic consisted of tweets about extreme temperature and weather conditions attributable to climate change. The topic was therefore named 'Extreme weather and climate change'. This topic had the highest number of retweets, and more URLs and hashtags were included in tweets in this topic than in other topics.

Topic 4 (16%) consisted of many frequently used heat wave-related words and adjectives indicating feelings and frustration, including 'sweltering_heat', 'unbearable', 'fuck', 'unbearable_heat' and 'hot'. Combined with other frequently used words such as 'work', 'school', 'kid', 'family', 'hospital', and 'heatstroke', we assumed that this cluster mainly consisted of tweets expressing opinions on how extreme temperatures and heat waves impact the daily lives of people, as compared to their health which was the main theme of topic 2. We also found the words 'air_conditione', 'cool', 'stay', and 'electricity' in this topic, indicating cooling strategies. The word 'car' is among the top 20 words, and 'child' among the top 50, indicating that users talk about extreme heat in cars and the consequences. We labelled this topic 'Social impacts'.

Topic 5, the smallest cluster (13%), contained many words that also appeared in topics 2 and 4. However, this topic also included the words 'warn' and 'warning' more frequently than in the other topics. Other words such as 'check', 'know', 'think', 'listen' and 'feel' made us label this topic 'Perceptions and warning'. The word clouds in Fig. 4 show the most frequently used words in tweets belonging to each of the five topics.

4 Discussion

4.1 What we learned from the Twitter communications

Our exploratory study is the first of its kind to analyse global Twitter communication about heat waves. We aimed to explore (1) whether social media communications data can provide useful information about heat waves, (2) what we can learn from Twitter conversations about how people are affected by and cope with heat waves, and (3) who engages with conservations about heat waves, and to what extent. By addressing these three aims, six lessons emerged that are worth highlighting.

First of all, the intensity of Twitter activity corresponds well with occurrences of major heat waves, such as those identified in California, Argentina, Western Australia and South Asia. Previous studies have shown that the number of social media messages from people close to events increases substantially immediately before and during their occurrence, suggesting that social media conversations are reliable indicators of natural hazard occurrences, in particular of sudden hazards such as floods (Bird et al. 2012), tornados (Silver and Andrey 2019), or fires (Wang et al. 2016). Our study is among the very few, besides Grasso et al. (2017), to establish this relationship between occurrence and increasing Twitter activity related to heat, which is often referred to as a slow onset hazard (Zander et al. 2019; Oppermann et al. 2021). The increasing number of Twitter messages during heat waves can also be interpreted as an indicator of public awareness, which relates to the third lesson learned.

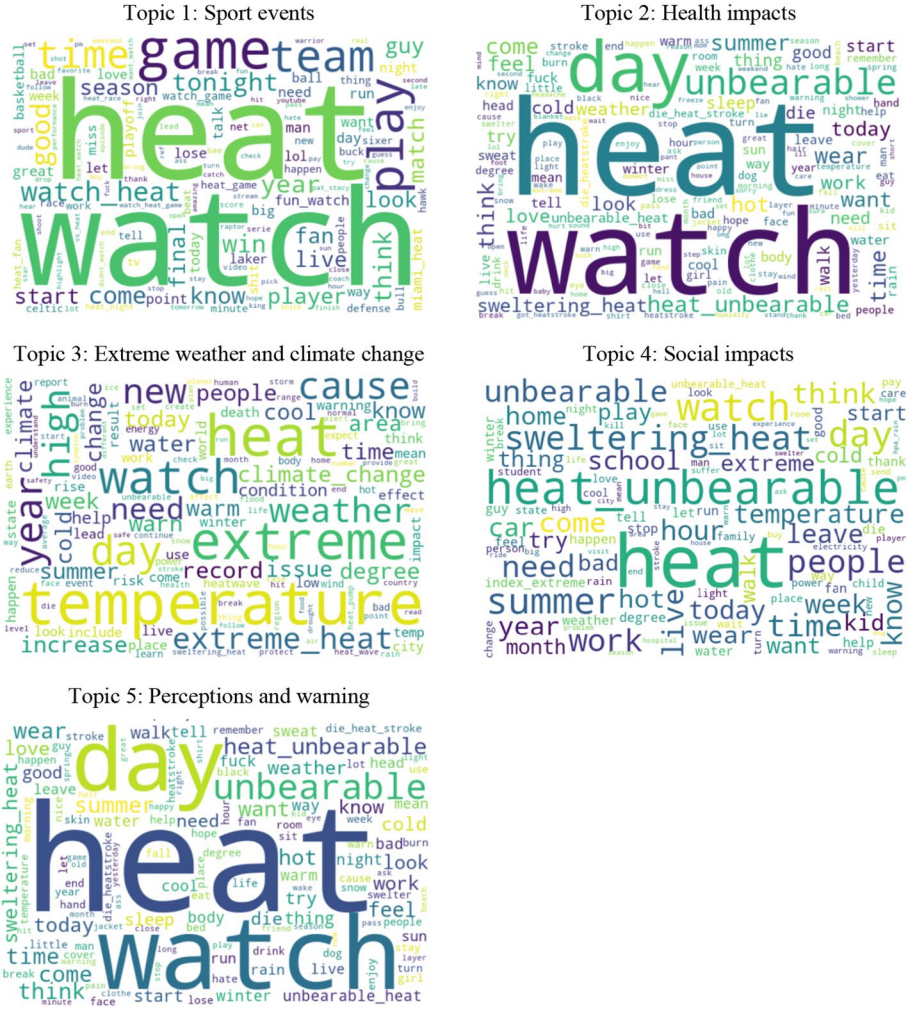

Fig. 4 Word clouds showing most frequent words in each topic

Second, we learnt little from the Twitter conversations about peoples' behaviour during hot days and heat waves. Among the most frequent 100 words in each topic, we identified only a few words that appeared to indicate adaptation, coping, and heat relief strategies. Tweets belonging to topic 4 ('Social impacts') included words such as 'air_conditione', 'cool' and 'stay', which may indicate people choose to stay in artificially cooled rooms on hot days, one of the main heat relief strategies in Australia (Zander et al. 2021), the USA (Lee and Shaman 2020) and many emerging countries, including in Asia (He et al. 2021). The word 'electricity' was also used frequently within that topic, suggesting that people may converse about high energy use in this context (see, e.g. Davis and Gertler 2015). The topic about the health impacts that frequently included words such as 'drink' and 'wear' also indicates some awareness about hydration during hot days, and the role of appropriate clothing. People seem to talk more about how they feel than what they might do to cope with heat. This expression of feelings can be interpreted as a consciousness of heat and heat waves as a potential threat,

an awareness that can heighten peoples' resilience since people who are aware of a threat are more likely to take adaptation actions, as shown for sudden onset hazards such as floods (Kellens et al. 2011; Liu et al. 2018) and also for heat waves (Zografos et al. 2016; Esplin et al. 2019; Howe et al. 2019). While we may have gathered more information about adaptation and coping strategies had we included additional search keywords such as 'prepare', 'plans' and other heat relief measures (see Huang and Xiao 2015), such words should have emerged in the topic modelling had they been present.

Third, we detected little fake news or account activities by users who disseminate false information about climate change and related heat, as can sometimes happen on social media (Allcott and Gentzkow 2017). The most prolific accounts belonged to individual scientists or scientific organisations, such as climateguyw or PlatformAdam, the official Twitter account of the residence of the President of India (rashtrapatibhvn), and large media corporations such as the ABC, Fox News, The Economist, and Reuters. This suggests that people rely strongly on scientific information and traditional news channels during extreme weather events (Freeman et al. 2018) and that these channels are not having to compete with misinformation.

Fourth, we found that no emergency agencies were among the ten most prolific Twitter accounts, whether in terms of the number of tweets, retweets or likes. For sudden hazards such as floods and fires, accounts managed by government emergency services used by many people, being considered a trusted source of warnings and updated information (Bird et al. 2012; Abedin and Babar 2018). Government emergency services did not feature highly among tweets posted before or during heat waves. This might be because government emergency and disaster management authorities do not consider heat, a slow-onset hazard, as their responsibility. For people affected by heat who prefer to receive alerts and information through official government Twitter accounts, rather than media channels or personal accounts, this may represent a gap in government services.

Fifth, we encountered some 'noise' in the collected data. After an initial analysis, we realised that a large number of tweets mentioned a movie named 'Heatwave' or a song named 'Heat Waves', neither of which was relevant to climate-related heat or heat waves. We therefore decided to clean the dataset by deleting all tweets referring to movies, films and music. Although this might distort the true picture of what people talk about on Twitter, we deemed this the correct approach in the light of our aims. Although we chose our search keywords carefully, this type of noise is unavoidable when collecting global Twitter data and we recommend multiple rounds of cleaning before starting topic modelling. Once the data have been cleaned, useful information and lessons emerged, as discussed here, even though we were left with one topic (sport events; 34% of all tweets) unrelated to climate-related heat and heat waves.

Sixth, when comparing different metrics of popularity and activity, we uncovered surprising details about the engagement of users, as discussed in the next section. We found significant effects of the content, the timeframe of the tweeting and the likelihood of a tweet being retweeted, as discussed in more detail in the next section. This knowledge can be useful for emergency authorities and organisations wanting to disseminate information about extreme heat situations rapidly and effectively to a large number of people.

4.2 Factors impacting the retweetability of tweets

We learnt five lessons by applying metrics that can identify the characteristics of the Twitter accounts that disseminate heat or heat wave-related information most rapidly. First, echoing Cha et al. (2010), we found that the number of followers is a poor

indicator of an account's influence. The ten accounts with the most followers had less engagement from their followers than accounts belonging to individual users with fewer followers.

Second, we found that tweets may have similar content but can trigger different levels of engagement. For example, a tweet about extreme temperatures in India and Pakistan posted on 29 April by PlatformAdam was retweeted 9.22 times more and had 17 times more replies than a tweet with similar content posted by LicypriyaK a day later, an account with almost 100 times more followers than PlatformAdam (167,000 compared to 7100 followers). This difference appears to have been the timing of the tweet—the earlier a tweet is posted about an imminent heat wave, the more likely it is to be retweeted. Given the importance of information timeliness in disaster communication (Taylor et al. 2007; Son et al. 2019), this finding was not surprising. Users retweeting the earliest tweets providing timely updates and alerts might be perceived as credible sources of information and thus generate higher engagement among followers.

Third, tweets about extreme weather and climate change (topic 4), which also included weather updates, were more likely to be shared than tweets about social and health impacts despite having fewer tweets. Tweets discussing the health and social impacts of heat or heat waves attracted the attention and gained approval from Twitter users (indicated by having more 'likes' and 'replies') but failed to become viral. This may mean that Twitter users consider tweets about 'Extreme weather and climate change' to be interesting, relevant and thus worth sharing. This insight is consistent with Abdullah et al. (2015) and Lee and Yu (2020), who show that content relevance impacts retweetability. Furthermore, from manual inspection of the content of the tweets in each topic, we found that numbers and prepositions were used more frequently in tweets that shared information about extreme weather. Hence, the frequency of the use of numbers, quantifiers and prepositions (Lee and Yu 2020) could also have impacted the retweetability of heat or heat wave-related tweets. Since retweeting allows for quick dissemination of information about significant events (Zhao et al. 2011), it is important to understand what factors might influence the retweeting behaviour of Twitter users. We can therefore apply this understanding to improve the effectiveness of disaster communication, especially during heat waves.

Fourth, besides the differences in content relevance, we also found that URLs and hashtags were more commonly found in tweets on the 'Extreme weather and climate change' topic than other topics, which was reflected through the higher proportion of tweets that included URLs and hashtags (Table 4). This result is consistent with studies by Suh et al. (2010) and Bruns and Stieglitz (2012), who concluded that the inclusion of URLs or hashtags in tweets positively affects the retweetability of disaster tweets.

Fifth, our analysis of tweet engagement suggests that the social network of the author of a heat or heat wave-related tweet also affects retweetability. We found that PlatformAdam's tweet was retweeted by Karl_Lauterbach, Germany's Federal Minister of Health with 1 million Twitter followers, as well as cathmckenna, the former Canadian Climate Minister with nearly 182,000 followers. A tweet with the same content posted by LicypriyaK on the following day, however, was not retweeted by any notable Twitter account. Through retweeting, the user shows agreement with the content and its general credibility (Son et al. 2020). Therefore, in addition to being the first account to raise awareness about the extreme temperatures in India on 29 April 2022, receiving the approval from other influencers in their social network might be the reason PlatformAdam's tweet had such high engagement counts. This is consistent with the results in Xu and Yang (2012), who suggested that the social relationship of a Twitter user is crucial to the likelihood of their tweets being retweeted.

Table 4 Number of tweets, retweets, likes, replies and links by topic

Topic	Tweets	Retweets	Likes	Replies	URLs	Tweets with hashtags
Sport events	16,930	26,592	181,748	11,582	6382	2842
Health impacts	10,299	9590	95,680	7820	2330	820
Extreme weather and climate change	8328	26,820	89,121	6422	4500	2265
Social impacts	8202	14,495	85,784	7985	2416	1187
Perceptions and warning	6505	12,077	84,759	5484	2042	864

URLs = Uniform resource locators

4.3 Study limitations

Our study has three limitations. The first relates to the limited data collection which happened because we only searched for Twitter tweets. Social media users employing other channels such as Facebook or Instagram to talk about heat or heat waves in the relevant time horizon were not included. However, we are confident that we captured many relevant conversations since Twitter is the outlet used most frequently by people seeking and communicating information in disaster situations (Simon et al. 2015), although Facebook has greater potential for conveying detail (Anikeeva et al. 2015). Also related to the limited representativeness of the data collected is the fact that we only searched tweets written in English. This means users from English-speaking countries with a large population are overrepresented, including the USA, UK and India. This is a common problem of research on social media use with very few exploring tweets in languages other than English (e.g. in German: Netzel et al. 2021).

The second limitation is related to the timeframe over which the data for our study were collected. The data presented in this study covered a four-month period from January to April 2022. The months after this have also seen higher-than-average temperatures relative to the same period from previous years, in particularly in Europe. We suggest that tweets be collected for analysis over years, not months, to reduce the impact of single non-climate change-related events (such as the launch of a movie called 'Heatwave' in India, or specific sport events, as was the case in our study).

The third limitation is related to the GSDMM topic model. At the core of GSDMM is the maximum number of clusters (topics) K_{Max} assumed to be present in the entire text corpus. While the GSDMM model can perform well with a K_{Max} higher than the true number of clusters (Yin and Wang 2014), it is challenging to find an appropriate K_{Max} when the true number of clusters is unknown (Yin and Wang 2016). In our paper, we assumed that the value of K_{Max} was 18 and then evaluated the performance of the GSDMM topic model with different values of k ranging from 4 to 18. However, the value we selected for K_{Max} may actually be less than the actual K_{Max}, and so the optimal k was not in the evaluation range.

5 Conclusions

This study contributes to disaster communication research by analysing social media communications (Twitter tweets) to improve understanding of the global impacts of, and responses to, extreme heat and heat waves. Our study, one of the few to investigate social

media use during extreme heat, analysed 62,920 tweets from January to April 2022 and found a strong correspondence between the number of tweets in each week and major heat wave occurrences in Argentina, Australia, the USA and South Asia. This was consistent with previous findings that social media use is a good indicator of the intensity of natural hazards. An unsupervised text analysis (topic modelling) identified five topics. Four of these topics and most tweets (66%) related to climate-related extreme heat and heat waves. Approximately one-third of tweets (34%) were about sport events, mainly basketball games in the USA, and could not clearly be attributed to communications about climatic events. The four relevant emerging topics were 'Health impacts' (20%), 'Extreme weather and climate change' (17%), 'Social impacts (16%)', and 'Perceptions and warning' (13%). Our analysis suggests that many Twitter users are aware of heat or heat waves and that social media can help to disseminate information about extreme temperatures and weather updates. Many users employ Twitter to share how they feel about unbearable heat and Twitter communications were also used, albeit to a lesser extent, to talk about how to cope and relief from heat stress. We also identified the circumstances under which tweets are likely to be retweeted, including the role of timeliness and social networks, which can help when organisations need to disseminate information in an emergency.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11069-023-05824-2.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. The funding was provided by Charles Darwin University.

Declarations

Conflict of interest The authors have not disclosed any competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abdullah NA, Nishioka D, Tanaka Y, Murayama Y (2015) User's action and decision making of retweet messages towards reducing misinformation spread during disaster. J Inf Process 23:31–40

Abedin B, Babar A (2018) Institutional vs. non-institutional use of social media during emergency response: a case of Twitter in 2014 Australian bush fire. Inf Syst Front 20:729–740

Agarwal N, Sikka G, Awasthi LK (2020) Evaluation of web service clustering using dirichlet Multinomial Mixture model based approach for dimensionality reduction in service representation. Inf Process Manag 57:102238

Allcott H, Gentzkow M (2017) Social media and fake news in the 2016 election. J Econ Perspect 31:211–235

Andrews O, Le Quere C, Kjellstrom T, Lemke B, Haines A (2018) Populations exposed to extreme heat under climate change: implications for workability and survivability. Lancet Planet Health 2:e540–e547

Anikeeva O, Steenkamp M, Arbon P (2015) The future of social media use during emergencies in Australia: insights from the 2014 Australian and New Zealand disaster and emergency management conference social media workshop. Aust J Emerg Manag 30:22–26
Asseng S, Spänkuch D, Hernandez-Ochoa IM, Laporta J (2021) The upper temperature thresholds of life. Lancet Planet Health 5:e378–e385
Bakhsh K, Rauf S, Zulfiqar F (2018) Adaptation strategies for minimizing heat wave induced morbidity and its determinants. Sustain Cities Soc 41:95–103
Beckmann SK, Hiete M, Schneider M, Beck C (2021) Heat adaptation measures in private households: an application and adaptation of the protective action decision model. Humanit Soc Sci Commun 8:227
Bureau of Meteorology (2022) Australian in January 2022. http://www.bom.gov.au/clim_data/IDCKGC1AR0/202201.summary.shtml. Accessed June 2022
Bird D, Ling M, Haynes K (2012) Flooding Facebook - the use of social media during the Queensland and Victorian floods. Aus J Emerg Manage 27:27–33
Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 3:993–1022
Capon A, Jay O, Ebi K, Lo S (2019) Heat and health: a forthcoming Lancet Series. Lancet 394:551–552
Cattaneo C, Peri G (2016) The migration response to increasing temperatures. J Dev Econ 122:127–146
Cecinati F, Matthews T, Natarajan S, McCullen N, Coley D (2019) Mining social media to identify heat waves. Int J Environ Res Public Health 16:762
Cha M, Haddadi H, Benevenuto F, Gummadi K (2010) Measuring user influence in Twitter: the million follower fallacy. In: Proceedings of the international AAAI conference on web and social media vol 4, pp 10–17
Chang J, Boyd-Graber J, Gerrish S, Wang C, Blei DM (2009) Reading tea leaves: how humans interpret topic models. Neural Inf Process Syst 22:288–296
Climate Action Tracker (2021) Warming projections global update November 2021
Davis LW, Gertler PJ (2015) Contribution of air conditioning adoption to future energy use under global warming. Proc Natl Acad Sci USA 112:5962–5967
Earth Observatory (2022) Southern hemisphere scorchers. https://earthobservatory.nasa.gov/images/149331/southern-hemisphere-scorchers [accessed June 2022]
Ebi KL et al (2021) Hot weather and heat extremes: health risks. Lancet 398:698–708
Esplin ED, Marlon JR, Leiserowitz A, Howe PD (2019) "Can you take the heat?" Heat-induced health symptoms are associated with protective behaviors. Weather Clim Soc 11:401–417
Freeman J, Hess K, Waller L (2018) Communication life line? ABC emergency broadcasting in rural/regional Australia. Commun Res Pract 4:342–360
Grasso V, Crisci A, Morabito M, Nesi P, Pantaleo G (2017) Public crowdsensing of heat waves by social media data. Adv Sci Res 14:217–226
Guo Y (2017) Hourly associations between heat and ambulance calls. Environ Pollut 220:14248–21428
Hatvani-Kovacs G, Belusko M, Skinner N, Pockett J, Boland J (2016) Heat stress risk and resilience in the urban environment. Sustain Cities Soc 26:278–288
He B-J, Zhao D, Xiong K, Qi J, Ulpiani G, Pignatta G, Prasad D, Jones P (2021) A framework for addressing urban heat challenges and associated adaptive behavior by the public and the issue of willingness to pay for heat resilient infrastructure in Chongqing. China Sustain Cities Soc 75:103361
Horton RM, de Sherbinin A, Wrathall D, Oppenheimer M (2021) Assessing human habitability and migration. Science 372:1279–1283. https://doi.org/10.1126/science.abi8603
Huang Q, Xiao Y (2015) Geographic situational awareness: mining tweets for disaster preparedness, emergency response, impact, and recovery. ISPRS Int J Geo-Inf 4:1549–1568
Howe PD, Marlon JR, Wang X, Leiserowitz A (2019) Public perceptions of the health risks of extreme heat across US states, counties, and neighborhoods. Proc Natl Acad Sci USA 116:6743–6748
IPCC (2021) Summary for policymakers. In: Masson-Delmotte V et al (eds) Climate change 2021: the physical science basis. Contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change, In Press. https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SPM_final.pdf [accessed Dec 2022]
Jacobi C, van Atteveldt W, Welbers K (2016) Quantitative analysis of large amounts of journalistic texts using topic modelling. Digit Journal 4:89–106
Jamali M et al (2020) Social media data and housing recovery following extreme natural hazards. Int J Disaster Risk Reduct 51:101788
Jay O (2021) Reducing the health effects of hot weather and heat extremes: from personal cooling strategies to green cities. Lancet 398:709–724
Jung J, Uejio CK (2017) Social media responses to heat waves. Int J Biometeorol 61:1247–1260
JustAnotherArchivist (2020) snscrape, GitHub. https://github.com/JustAnotherArchivist/snscrape [accessed June 2022]

Karami A, Gangopadhyay A, Zhou B, Kharrazi H (2018) Fuzzy approach topic discovery in health and medical corpora. Int J Fuzzy Syst 20:1334–1345
Kellens W, Zaalberg R, Neutens T, Vanneuville W, De Maeyer P (2011) An analysis of the public perception of flood risk on the Belgian coast. Risk Anal 31:1055–1068
Kryvasheyeu Y et al (2016) Rapid assessment of disaster damage using social media activity. Sci Adv 2:e1500779
Lee VW, Shaman J (2017) Heat-coping strategies and bedroom thermal satisfaction in New York City. Sci Total Environ 574:1217–1231
Lee CH, Yu H (2020) The impact of language on retweeting during acute natural disasters: uncertainty reduction and language expectancy perspectives. Ind Manag Data Syst 120:1501–1519
Liu D, Li Y, Shen X, Xie Y, Zhang Y (2018) Flood risk perception of rural households in western mountainous regions of Henan Province, China. Int J Disaster Risk Reduct 27:155–160
Lupton D (2014) 'Feeling better connected': academics' use of social media. News and Media Research Centre, University of Canberra, Canberra
Martín Y et al (2020) Using geotagged tweets to track population movements to and from Puerto Rico after hurricane Maria. Popul Environ 42:4–27
Moore F et al (2019) Rapidly declining remarkability of temperature anomalies may obscure public perception of climate change. Proc Natl Acad Sci USA 116:4905–4910
Murakami D, Yamagata Y (2020) Spatial heat-wave assessments using geotagged Twitter data. In: Yamagata Y, Seya H (eds) Spatial analysis using big data. Elsevier Academic Press, London, pp 227–238
Nagmoti R, Teredesai A, De Cock M (2010) Ranking approaches for microblog search. IEEE/WIC/ACM International conference on web intelligence and intelligent agent technology, vol 1, pp 153–157
Editorial N (2021) Cities must protect people from extreme heat. Nature 595:331–332
Netzel L, Heldt S, Denecke M (2021) Analyzing Twitter communication about heavy precipitation events to improve future risk communication and disaster reduction in Germany. Urban Water J 18:310–331
NOAA (2022) State of the climate: global climate report for April 2022. https://www.ncdc.noaa.gov/sotc/global/202204 [accessed Dec 2022]
Omurca Sİ, Ekinci E, Yakupoğlu E, Arslan E, Çapar B (2021) Automatic detection of the topics in customer complaints with artificial intelligence. Balkan J Electr Comput Eng 9:268–277
Oppermann E, Kjellstrom T, Lemke B, Otto M, Lee JKW (2021) Establishing intensifying chronic exposure to extreme heat as a slow onset event with implications for health, wellbeing, productivity, society and economy. Curr Opin Environ Sustain 50:225–235
Primo F, Romanovsky A, de Mello R, Garcia A, Missier P (2021) A customisable pipeline for the semi-automated discovery of online activists and social campaigns on Twitter. World Wide Web 24:1235–1271
Qiang J, Qian Z, Li Y, Yuan Y, Wu X (2020) Short text topic modeling techniques, applications, and performance: a survey. IEEE Trans Knowl Data Eng Early Access 34(3):1427–1445
Rauf S, Bakhsh K, Abbas A, Hassan S, Ali A, Kächele H (2017) How hard they hit? Perception, adaptation and public health implications of heat waves in urban and peri-urban Pakistan. Environ Sci Pollut Res 24:10630–10639
Reynard D, Shirgaokar M (2019) Harnessing the power of machine learning: can Twitter data be useful in guiding resource allocation decisions during a natural disaster? Transp Res D Transp Environ 77:449–463
Riquelme F, González-Cantergiani P (2016) Measuring user influence on Twitter: a survey. Inf Process Manage 52:949–975
Robine J-M, Cheung SLK, Le Roy S, Van Oyen H, Griffiths C, Michel J-P et al (2008) Death toll exceeded 70,000 in Europe during the summer of 2003. C R Biol 331:171–178
Röder M, Both A, Hinneburg A (2015) Exploring the space of topic coherence measures. In: Cheng X, Li H, Gabrilovich E, Tang J (eds) In: Proceedings of the eighth ACM international conference on web search and data mining, ACM, pp 399–408
Sherwood SC, Huber M (2010) An adaptability limit to climate change due to heat stress. Proc Natl Acad Sci USA 107:9552–9555
Silver A, Andrey J (2019) Public attention to extreme weather as reflected by social media activity. J Conting Crisis Manag 27:346–358
Simon T, Goldberg A, Adini B (2015) Socializing in emergencies—a review of the use of social media in emergency situations. Int J Inf Manage 35:609–619
Son J, Lee HK, Jin S, Lee J (2019) Content features of tweets for effective communication during disasters: a media synchronicity theory perspective. Int J Inf Manag 45:56–68
Son J, Lee J, Oh O, Lee HK, Woo J (2020) Using a heuristic-systematic model to assess the Twitter user profile's impact on disaster tweet credibility. Int J Inf Manag 54:102176

Stevens HR, Beggs PJ, Graham PL, Chang H-C (2019) Hot and bothered? Associations between temperature and crime in Australia. Int J Biometeorol 63:747–762
Suh B, Hong L, Pirolli P, Chi EH (2010) Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. In: 2010 IEEE second international conference on social computing, pp 177–184
Taylor JG, Gillette SC, Hodgson RW, Downing JL, Burns MR, Chavez DJ, Hogan JT (2007) Informing the network: improving communication with interface communities during wildfire. Human Ecol Rev 14:198–211
The Guardian (2022) 'We are living in hell': Pakistan and India suffer extreme spring heatwaves. 2 May 2022. https://www.theguardian.com/world/2022/may/02/pakistan-india-heatwaves-water-electricity-shortages [accessed Dec 2022]
Turek-Hankins LL et al (2021) Climate change adaptation to extreme heat: a global systematic review of implemented action. Oxf Open Clim Chang. https://doi.org/10.1093/oxfclm/kgab005
Varghese BM et al (2019) Heatwave and work-related injuries and illnesses in Adelaide, Australia: a case-crossover analysis using the excess heat factor (EHF) as a universal heatwave index. Int Arch Occup Environ Health 92:263–272
Vicedo-Cabrera AM et al (2021) The burden of heat-related mortality attributable to recent human-induced climate change. Nat Clim Change 11:492–500
Wang Z, Ye X, Tsou MH (2016) Spatial, temporal, and content analysis of Twitter for wildfire hazards. Nat Hazards 83:523–540
Wouters H et al (2017) Heat stress increase under climate change twice as large in cities as in rural areas: a study for a densely populated midlatitude maritime region. Geophys Res Lett 44:8997–9007
Xu Z, Yang Q (2012) Analyzing user retweet behavior on Twitter. In: IEEE/ACM international conference on advances in social networks analysis and mining vol 1, pp 46–50
Xue J, Chen J, Chen C, Zheng C, Li S, Zhu T (2020) Public discourse and sentiment during the COVID 19 pandemic: using Latent Dirichlet Allocation for topic modeling on Twitter. PLoS ONE 15:e0239441
Yin J, Wang J (2014) A dirichlet multinomial mixture model-based approach for short text clustering. In: Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining. pp 233–242
Yin J, Wang J (2016) A text clustering algorithm using an online clustering scheme for initialization. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. pp 1995–2004
Zander KK, Botzen WJW, Oppermann E, Kjellstrom T, Garnett ST (2015) Heat stress causes substantial labour productivity loss in Australia. Nat Clim Change 5:647–651
Zander KK, Moss SA, Garnett ST (2017) Drivers of self-reported heat stress in the Australian labour force. Environ Res 152:272–279
Zander KK, Matthew S, Garnett ST (2018) Exploring heat stress relief measures among Australian labour force. Int J Environ Res Public Health 15:401
Zander KK, Richerzhagen C, Garnett ST (2019) Human mobility intentions in response to heat in urban South East Asia. Glob Environ Change 56:18–28
Zander KK, Garnett ST (2020) The importance of climate to emigration intentions from a tropical city of Australia. Sustain Cities Soc 63:102465
Zander KK, Shalley F, Tan G, Dyrting S, Taylor A (2021) "Run air-conditioning all day": adaptation pathways to increasing heat in the Northern territory of Australia. Sustain Cities Soc 74:103194
Zhang Y, Li J, Song Y, Zhang C (2018) Encoding conversation context for neural keyphrase extraction from microblog posts. In: Proceedings of the 2018 conference of the North American chapter of the association for computational linguistics: human language technologies, vol 1 (Long Papers) pp 1676–1686
Zhao WX (2011) Comparing Twitter and traditional media using topic models. In advances in information retrieval. ECIR Lecture Notes in Computer Science, vol 6611. Springer, Berlin, Heidelberg, pp 338–349
Zografos C, Anguelovski I, Grigorova M (2016) When exposure to climate change is not enough: exploring heatwave adaptive capacity of a multi-ethnic, low-income urban community in Australia. Urban Clim 17:248–265

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.