

UC Irvine

UC Irvine Previously Published Works

Title

How Californians Tweet about Extreme Heat Events on Social Media: A Health Equity Perspective

Permalink

<https://escholarship.org/uc/item/02v2z66v>

Journal

Weather Climate and Society, 17(4)

ISSN

1948-8327

Authors

Sriperumbudur, Gomathi B

Fan, Yihang

Liu, Xiaozhen

et al.

Publication Date

2025-10-01

DOI

10.1175/wcas-d-24-0147.1

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

How Californians Tweet about Extreme Heat Events on Social Media: A Health Equity Perspective

GOMATHI B. SRIPERUMBUDUR,^a YIHANG FAN,^a XIAOZHEN LIU,^b MINGYU DEREK MA,^c WEI WANG,^c CHEN LI,^b
AND SUELLEN HOPFER^{ORCID}^a

^a *Health, Society and Behavior, Joe C. Wen School of Population and Public Health, University of California, Irvine, Irvine, California*

^b *Computer Science, University of California, Irvine, Irvine, California*

^c *Computer Science, University of California, Los Angeles, Los Angeles, California*

(Manuscript received 5 January 2025, in final form 26 June 2025, accepted 8 July 2025)

ABSTRACT: Extreme heat events are increasingly common and deadlier than any other weather phenomenon, significantly affecting mortality rates. In California, these events have intensified, impacting populations differently based on structural factors of health and location. Tools like the healthy places index (HPI), which combines 23 social and economic indicators, can guide targeted heat mitigation efforts together with social media monitoring. We examined Twitter (X) heat discourse between 2016 and 2022 across California and by HPI. From keyword-filtered tweets ($n = 4572$ in top 75th quartile HPI areas and $n = 2701$ in bottom 25th quartile HPI areas), we inductively identified eight heat discourse categories, listed in order of prevalence: perceived heat risk, life impact, coping, venting, heat warnings, community action, expressing relief, and climate change concern. Overall, extreme heat tweets correlated with heat events across these 7 years. Among high HPI areas, a greater proportion of heat tweets expressed positive sentiment, relief (when it cooled off), and coping. In contrast, among low HPI areas, a greater proportion of heat tweets expressed negative sentiment, venting, and heat warnings. Life impacts, perceived extreme heat, community action, and linking heat events with climate change were expressed with equal frequency across low and high HPI areas. Our results indicate that HPI serves as a robust geospatial heat equity tool for assessing extreme heat impacts across California and that extreme heat events are not experienced equally across communities.

SIGNIFICANCE STATEMENT: Increasingly, local governments are developing heat health action plans. A health equity lens recognizes that communities experience heat events differently depending on social determinants of health (e.g., housing, green space, air conditioning, income). Using the healthy places index (HPI), a composite measure of 23 community characteristics relevant to adaptive capacity to extreme heat events allows researchers to prioritize equity and measure how the most vulnerable communities respond to extreme heat. Communities in low HPI areas expressed greater negative sentiment and less coping. Social media monitoring of heat discourse by HPI provides real-time insight into heat risk experiences among the most vulnerable communities. Low HPI areas need to be prioritized for future public health research on adapting to extreme heat.

KEYWORDS: Community; Communications/decision making; Regional effects; Machine learning

1. Introduction

Rising temperatures have become a significant concern for public health because their frequency and severity have escalated in recent decades due to climate change (Cissé et al. 2022; Kiarisi et al. 2023; Palinkas et al. 2022; Schwarz et al. 2020). In the United States, incidents of extreme heat and related health emergencies have increased since the mid-1960s (Cissé et al. 2022; Watts et al. 2019). Predictions indicate a continued rise in extreme temperatures, strongly supported by numerous evidence-based simulations. The World Meteorological Organization highlighted that July 2024 set a new record as the hottest month ever documented globally, with notable impacts in the United States (Perez 2023). The definition of extreme heat varies by location, accounting for local acclimatization to regional climates rather than a fixed temperature measure (Guirguis et al. 2014). For instance, California

identifies days of extreme heat as when the daily peak temperature exceeds the 98th percentile of historic maximum temperatures for a location, based on the warmest month data from April to October between 1961 and 1990 (Cal-Adapt 2024). With this framework, the threshold for an extreme heat day in San Francisco is approximately 85.2°F, whereas it is 96.1°F in San Diego. Extreme heat is the foremost cause of weather-related deaths in the United States, outpacing total deaths from hurricanes, floods, and tornadoes (Knowlton et al. 2008). This extreme heat phenomenon disproportionately affects marginalized groups, including older adults and economically disadvantaged populations (Mendez 2020). The public often underestimates the health threats of extreme heat compared to other dramatic natural events like storm surges and wildfires (Ostro et al. 2009). This underestimation contributes to the low priority given to heat adaptation strategies in policymaking processes.

California faces unique challenges in dealing with extreme heat, given its diverse population and climate zones (Abedin and Babar 2018). The impact of extreme heat disproportionately

Corresponding author: Suellen Hopfer, shopfer@hs.uci.edu

DOI: 10.1175/WCAS-D-24-0147.1

© 2025 American Meteorological Society. This published article is licensed under the terms of the default AMS reuse license. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

Authenticated cgarrison@ametsoc.org | Downloaded 10/13/25 01:33 PM UTC

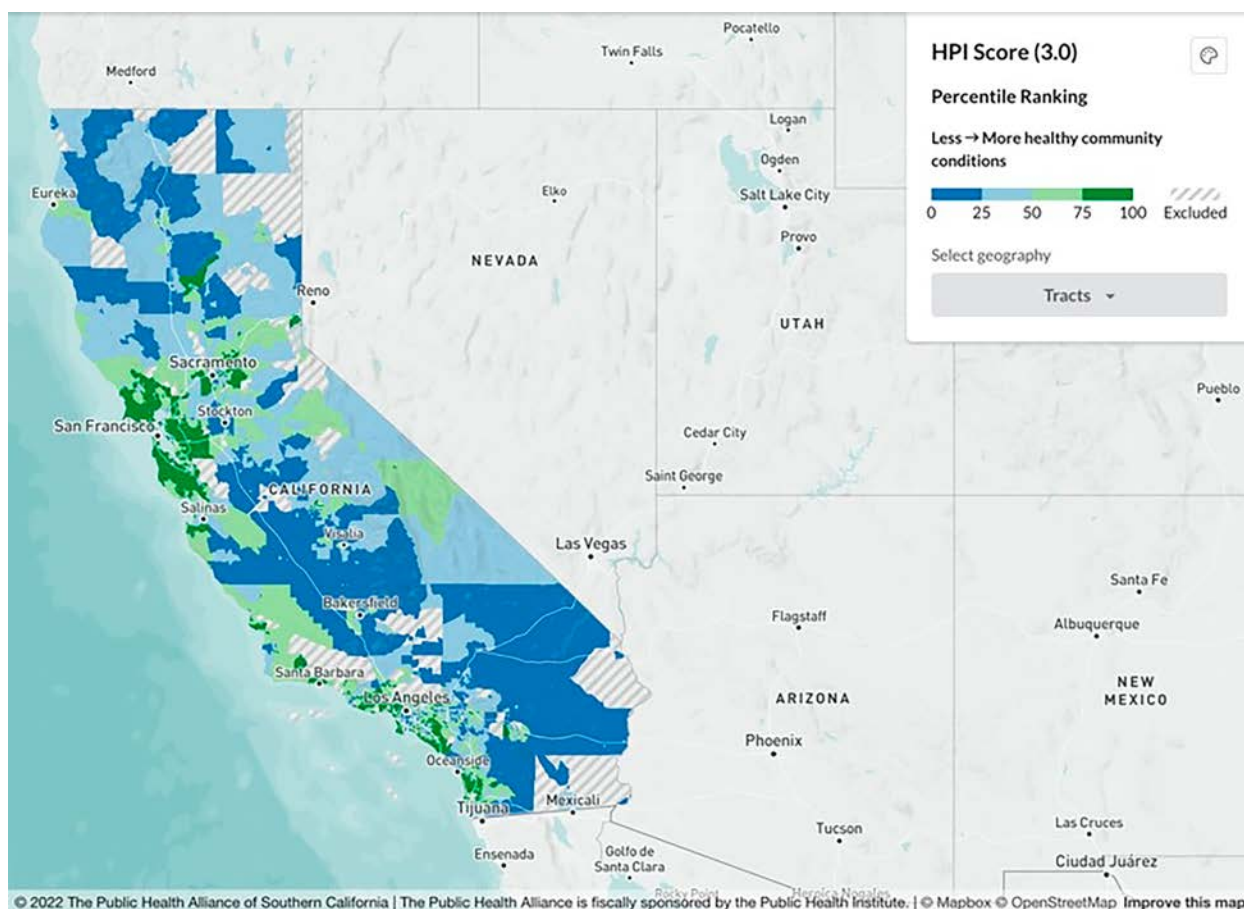


FIG. 1. HPI reference map. Courtesy of Public Health Alliance of Southern California.

affects low-income and marginalized populations due to not only their risk perceptions but also their minimal adaptive capacity and community resilience (Palinkas et al. 2022). Low-income communities, older adults, children, and those with underlying medical issues are among the populations most susceptible to extreme heat (Masri et al. 2022). The availability and use of air conditioning or cooling devices in residences play a crucial role, significantly diminishing the likelihood of heat-related health consequences (Gronlund and Berrocal 2020). The prevalence of home air conditioning varies widely in California, primarily influenced by regional climate differences, renter status, and housing type (Maizlish et al. 2019). For example, in Los Angeles County, around 20% of households lack air conditioning, whereas in the warmer inland Riverside County, less than 7% of households go without air conditioning (Gronlund and Berrocal 2020; Mann and Schuetz 2022).

California has experienced persistent heat waves in addition to drought, with variations in heat-related mortality across the state. Many residents in the Central Coast region who were used to moderate summer temperatures do not perceive themselves as vulnerable to heat waves and often do not take preventive measures against heat stress (Guirguis et al. 2014; Knowlton et al. 2008). California has advocated equity

in disadvantaged communities through environmental justice principles (Zuñiga and Méndez 2024). In 2022, California issued an extreme heat action plan, aiming to enhance community resilience and address the varied impacts of extreme heat, particularly considering the diverse microclimates across the state (Newsom 2022). Communities in California are increasingly under pressure to adapt to climate change and increasing heat events, yet it is unclear how community residents are responding to these increasing threats.

a. Healthy places index as a health equity tool

The healthy places index (HPI), introduced in 2018 by the Public Health Alliance of Southern California, is a crucial tool for identifying vulnerable populations in extreme heat contexts (Maizlish et al. 2019). Aligned with the IPCC's (2021) conceptual model, which assesses heat vulnerability through exposure, sensitivity, and adaptability pathways, the HPI comprehensively addresses the impact of social determinants of health on extreme heat and associated health disparities. In contrast to tools like the social vulnerability index 2018 (Centers for Disease Control and Prevention and Agency for Toxic Substances and Disease Registry 2024; Flanagan et al. 2011) and the California EnviroScreen 4.0

(CalEPA 2023; Office of Environmental Health Hazard Assessment 2024), the HPI assigns the highest weight to economic factors (35%) relative to other factors, including education, transportation, social, housing, healthcare access, clean environment, and neighborhood greenspace (Maizlish et al. 2019). Importantly, the HPI emphasizes its utility as a guide for policy action. The HPI domains are aligned with the social determinants of health and policy action areas. Notably, the HPI's most disadvantaged quartile of census tracts exhibits a substantial overlap with indexes of economic deprivation. This unique focus makes the HPI particularly relevant for identifying vulnerable populations in California that face distinctive heat exposure risk (Maizlish et al. 2019). Residing in older- or poorer-quality homes, living in rental homes with limited control over temperature regulation, and making difficult choices between spending energy for cooling and other essential needs are some issues attributable to low HPI scores (Fig. 1).

b. Social media monitoring for extreme heat

Social media has become a useful data source for mining public opinion on various topics, including extreme weather and climate change (Grasso et al. 2017; Kryvasheyev et al. 2016; Young 2018). Twitter (now X), in particular, with its real-time reactive, user-generated, and user-responsive data, provides reliable information on how people react to extreme heat (Roxburgh et al. 2019). Most studies have focused on the volume of tweets as an indicator of the impact of extreme heat and the dynamics of information dissemination (Dzyuban et al. 2022; Zander et al. 2023). However, the nature and content of heat-related tweets remain poorly understood. A 2023 study used unsupervised machine learning to conduct topic modeling and identify the most frequent topics of global Twitter conversations related to heat waves during 4 months (Zander et al. 2023). Taking a different approach, another study categorized theme-specific tweets by combining grounded theory and a corpus-based method (Jung and Uejio 2017).

c. Sentiment and heat

Other social media heat studies have focused on sentiment detection in heat-related tweets. For example, Dzyuban et al. (2022) examined the relationships between social media sentiment and weather data in Singapore and Phoenix, Arizona, and found that negative sentiment was significantly related to high heat. The study did not, however, explore the themes of heat-related tweets nor did it examine content in depth. Baylis et al. (2018) similarly used social media posts to investigate the association between meteorological conditions and sentiment expressions, employing longitudinal data from more than 3.5 billion posts across 2009–16 in the United States. They found that hot temperatures and other extreme weather were related to negative sentiment, even when excluding weather-related posts. A previous study reported a significant relationship between temperature and tweet sentiment, but findings used an unfiltered tweet dataset, weakening the link between sentiment and study theme (Jung and Uejio 2017). Moreover, it is

important to note that conclusions regarding climate-related Twitter communication are specific to the region being studied, owing to unique climate patterns and various heat experiences across societies (Dzyuban et al. 2022; Grasso et al. 2017; Roxburgh et al. 2019). Given that relatively few studies have explored social media discourse on California's extreme heat, our study expanded on previous work by using sentiment and content analyses to detect scales of responses and perceptions of extreme heat in regions of California from Twitter data.

Extreme heat provokes alterations in emotions and behavior, often leading to sensations of anger, irritability, aggression, discomfort, stress, and fatigue. Concurrently, social media provides a platform for individuals to express their perceptions and experiences during weather occurrences, offering a snapshot of societal sentiment toward weather events in real time. Researchers found that extreme heat and precipitation events exacerbate negative emotional states online while reducing positive expressions, as indicated by an analysis of nearly 8 billion social media posts and corresponding weather data from 2015 to 2021 (Dzyuban et al. 2022; Obradovich and Minor 2022). Using social media to monitor public sentiment can help policymakers develop evidence-based strategies to mitigate the negative effects of extreme temperatures.

By examining tweets about heat throughout California in certain geographic areas designated by the HPI, a measure of community well-being and health equity, this study focused on the following research questions:

RQ1. How is the experience of extreme heat expressed in tweets across California from 2016 to 2022?

RQ2. Does the distribution of Twitter heat conversations differ in high and low HPI areas?

RQ3. Does the sentiment distribution of heat-related tweets differ between high and low HPI areas?

Hypothesis 1: We expect the proportion of positive sentiment tweets to be greater in high compared to low HPI areas in California.

$$H_0: \text{positive sentiment}^{\text{highHPI}} = \text{positive sentiment}^{\text{lowHPI}}$$

$$H_a: \text{positive sentiment}^{\text{highHPI}} > 0.5 \text{ positive sentiment}^{\text{lowHPI}}$$

Hypothesis 2: We expect the proportion of tweets expressing negative sentiment to be greater in low HPI areas compared to high HPI areas.

$$H_0: \text{negative sentiment}^{\text{highHPI}} = \text{negative sentiment}^{\text{lowHPI}}$$

$$H_a: \text{negative sentiment}^{\text{highHPI}} < 0.5 \text{ negative sentiment}^{\text{lowHPI}}$$

2. Methods

a. Study areas

This study used the HPI, a health-equity-focused data tool developed by the Public Health Alliance of Southern California, a coalition of 10 health departments in California. HPI is a composite measure of 23 community characteristics across eight policy action areas including economic, education, transportation, social, housing, healthcare, environment, and neighborhood that influence health and quality of life (Maizlish et al. 2019; University of California, Los Angeles 2022). It provides a health equity lens

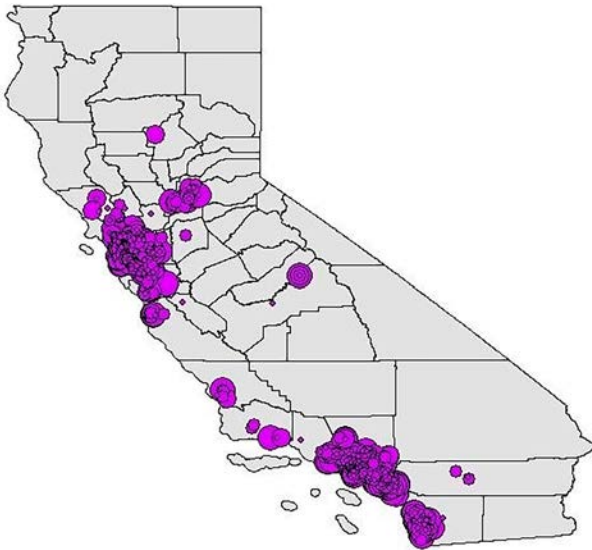


FIG. 2. Map representing the spatial distribution of tweets regarding heat waves in high HPI areas in California.

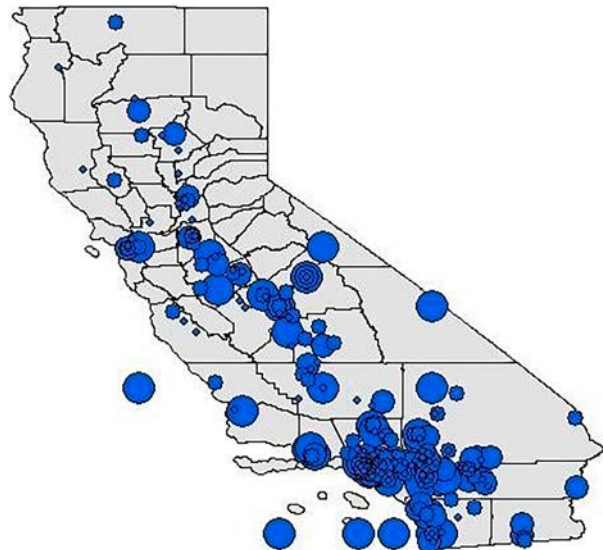


FIG. 3. Map representing the spatial distribution of tweets regarding heat waves in low HPI areas in California.

with a single composite score ascribed to geographies that enables comparisons of locations considering multiple social factors. HPI is displayed as a score divided into four quartiles, with the highest quartile (75th percentile and above) reflecting the areas of California with the best performance in social drivers of health. High HPI areas reflect relatively healthy community conditions, including clean air and water, more air conditioning availability, better education, employment conditions, and other metrics, which may help residents better deal with extreme heat. The lowest quartile (25th percentile and below) reflects geographies with low scores on key measures, implying the least healthy communities. These indicators include poor air quality, fewer homes with air conditioning, less tree canopy, and so on, which may intensify residents' experience of a heat wave. The two quartiles offer a good contrast of important aspects that define the health and quality of life of residents and their ability to adapt to extreme heat. Therefore, comparing the two quartiles offers an opportunity to examine whether and how the experience of heat differs depending on location.

b. Data sources

Three sources of data were used in this paper.

1) TWEET DATA

The source of tweets was a private archive of historical tweets. For November 2015 through December 2022, we retrieved daily tweets using Twitter's public application programming interface with a random sample rate of 5% (Twitter.com 2010). We amassed approximately 2.2 billion tweets in this archive. Each tweet contained the following attributes: (i) text content, (ii) time it was posted, (iii) information about the user who posted the tweet, and (iv) information about the user's approximate geolocation with a granularity coarser than census tract (i.e., state, county, and city).

2) HPI DATA

We retrieved census tract-level HPI information from the HPI website, hosted by the UCLA Luskin Center (see <https://www.healthyplacesindex.org>).

3) CALIFORNIA CENSUS TRACT DATA

To determine the census tract to which each tweet belonged, we retrieved the 2019 TIGER/Line shapefile for census tracts in California from the federal government's open data site (see <https://data.gov>).

c. Tweet extraction and filtering

The tweets analyzed in this paper were extracted and filtered using a five-step process.

Step 1: From the historical tweet archive dataset, we used three filters to retrieve tweets relevant to the topic of this paper: spatial, date, and extreme heat keyword filters. A spatial filter was first applied to retrieve tweets from California only. Then, a date filter was used to preserve only tweets between 2016 and 2022 and for each year, between April and October, to capture tweets expressed in real time during hotter months. Finally, we extracted tweets relevant to heat waves by using a filter based on the following keywords: "heat wave," "heatwave," "extreme heat," and "high temperature." We initially used 25 keywords related to heat (see the appendix for the iterative process of establishing keywords) refining the keyword filtering process by searching tweets containing heat and place (car, home, office, school), period of time (e.g., summer, June), adjectives (unbearable, dry), and adaptive capacity (e.g., AC use) (Jung and Uejio 2017). The final four keywords, however, yielded the most relevant tweets with the least noise. Applying the four keyword filters resulted in 19351 tweets.

TABLE 1. Performance of RoBERTa-large vs GPT-4 on 295 tweet test cases.

Code type proportion	Code	RoBERTa-large	GPT-4
		F1 average	F1 average
0.25	Perceived heat risk	0.61	0.39
0.22	Life impact	0.77	0.70
0.20	Coping strategies	0.80	0.86
0.17	Venting	0.80	0.78
0.14	Heat warning or PSA	0.77	0.83
0.14	Community action	0.85	0.83
0.08	Expressing relief	0.74	0.69
0.05	Climate change concern	0.73	0.78

Step 2: We applied a machine learning–based heat wave relevance filter. Because not all tweets from step 1 were considered to be relevant, we filtered the tweets to remove noisy data. The noise consisted of tweets referring to music bands, movies, shows, and songs with the name heatwave that were removed (e.g., “you would not believe how many Heat Waves covers and remixes there are on Spotify” refers to a song title; “What a show!! @kendricklamar kendricklamar #hiphop #sandiego #kdot #heatwave” refers to a musical performance; “#NowPlaying Heatwave–Always and Forever” refers to the movie Heatwave). We applied a supervised transformer-based machine learning model trained on 100 expert-annotated data pairs consisting of tweet content and a binary label indicating whether the tweet is relevant to extreme heat. The model was initialized from a pretrained RoBERTa-large model (Liu et al. 2019) with a linear layer to produce the classification result. We evaluated the trained model on a standout test set with expert annotation of 87 tweets, and the model yielded an accuracy of 93% and a recall of 100% for relevant tweets.

Step 3: We determined the location of each tweet. Tweet data before this step did not contain granular information on location based on census tracts. Each tweet had a place identifier referring to a rectangle bounding box retrievable from Twitter. Each bounding box encloses a space between two latitudes and two longitudes, indicating the approximate location of the tweet poster. We wrote a program to incorporate the bounding box information into the tweets and used the coordinates of the central point in each bounding box as the location of each tweet.

Step 4: We tagged tweets with census tract–based HPI information. We wrote a program to tag the census tract to which each tweet belonged using the geolocation coordinates. We considered a tweet to belong to a census tract if its coordinate fell in the boundary of the census tract and tagged the identifier of the census tract to the tweet. Then, we added HPI-related information to each tweet based on its census tract identifier.

Step 5: We split tweets based on HPI percentile, which was census tract–based HPI information. Each census tract was associated with an HPI percentile (0%–100%, with a higher percentile indicating a healthier place). We used the highest HPI quartile ($\geq 75\%$) and the lowest HPI quartile ($\leq 25\%$) as filters to split relevant tweets into two

subsets, representing tweets posted from census tracts with high and low HPI scores, respectively. This step resulted in two final tweet datasets—high HPI tweets (4572 tweets) and low HPI tweets (2701 tweets)—used for further analysis (see Figs. 2 and 3 for the spatial distribution of social media heat discourse across California low and high HPI areas).

d. Machine learning predictions

We developed a mixture of expert machine learning prediction models to produce and identify the codes of a given tweet. Both large language modeling (LLM) with in-context demonstration and supervised trained transformer-based classification-based models were used to leverage their benefits for different categories. We identified eight content (heat) categories, and each tweet could be associated with more than one category. We treated the task as eight binary classification tasks and developed a model for each category separately and then merged the prediction results. For the LLM method, we prompted Generative Pre-trained Transformer 4 (GPT-4) with the definition and inclusion and exclusion criteria of a particular heat category, along with several in-context demonstrations of known tweets and whether they should be linked to the category. We then parsed the natural language output of LLM to produce the predicted label of whether the tweet met the query category. For the supervised learned method, we trained a sequence classification model based on a pretrained RoBERTa-large model on expert annotations (tweet, classification flag) pairs, in which the classification flag indicated whether the tweet belonged to a particular code category. We evaluated both methods on a separate testing set with human annotations. The LLM method was used to assign categories of “heat warning or public service announcement (PSA),” “coping strategies,” and “climate change concern,” and the supervised learned method was used for all other categories (see Table 1).

e. Data analysis

This mixed-methods study employed the content analysis of heat tweets to identify categories of social media users who tweeted about heat. The content analysis was conducted both manually and by machine learning. Subsequently, to determine whether high and low HPI mean tweet frequencies significantly differed (for the eight identified heat conversation categories), z tests were performed.

1) CONTENT ANALYSIS OF EXTREME HEAT TWITTER DATA

The content analysis identified how users tweeted about extreme heat. Descriptive codes were inductively identified from a random sample of tweets across the 7-yr study period (Tracy 2019). The coding process consisted of five steps: 1) Two coders independently reviewed about 800 randomly sampled tweets across years to gain familiarity with tweets. 2) Initial codes were generated inductively by describing how tweets referenced extreme heat. Two coders independently coded tweets until theoretical saturation (Bowen 2008). 3) Coders compared and contrasted inductive codes, discussing any potential additional or distinct codes from extreme heat communication literature and California's extreme heat action plan (Newsom 2022). 4) Coders reviewed, revised, and refined all candidate conceptual codes and applied them to a randomly sampled batch of 421 heat-related tweets from April to September in 2015–22 to confirm the saturation of code categories. 5) Coders finalized the codebook, including defining the inclusion and exclusion criteria and listing examples for each code.

Globally, exclusion criteria included tweets using heatwave in reference to music, movies, stores, or merchandise; referring to heat and heat waves in geographical areas outside of California; and using languages other than English. For example, for tweets categorized as venting, to distinguish them from other codes such as perceptions of extreme heat and coping, tweets with neutral or positive expressions were excluded. Tweets using exclamations, swear words, or negative connotations were included in the venting category. Each tweet could be labeled with a maximum of three codes. This ensured that the entirety of the tweet expression was considered while eliminating overrepresentation of any code category. All discrepancies in coding between the two coders were discussed, with disagreements resolved by consensus. A third coder was involved for consensus when needed during the coding process.

A sample of 421 random tweets from both high and low HPI areas was analyzed for content individually by two coders, who then inductively derived descriptive codes. Each tweet could be coded as reflecting more than one heat category. To ensure inclusion of all possible code descriptions, we created multiple subcategories. The subcategories with overlapping content descriptions were then put under a larger umbrella code after extensive iterations. These larger categories were supplied to machine learning software, which helped us automate the coding process for the entire tweet dataset.

Overlapping codes were categorized into eight main codes. Inter-coder reliability (Cohen's kappa, or κ) was computed for each identified extreme heat codes: broadcasting heat warnings ($\kappa = 0.97$), coping ($\kappa = 0.94$), life impact ($\kappa = 0.92$), venting ($\kappa = 0.94$), perceived heat ($\kappa = 0.93$), discussing community action ($\kappa = 0.93$), linking climate change to heat events ($\kappa = 0.89$), and expressing relief ($\kappa = 0.87$). Interrater reliability demonstrated near-perfect agreement in the coding process (Hallgren 2012; Zapf et al. 2016).

2) DISTRIBUTION OF EXTREME HEAT TWEET PROPORTIONS FOR LOW AND HIGH HPI AREAS

To compare the average extreme heat tweets by category, we computed proportions to account for the imbalance in sample size between high ($n = 4572$) and low ($n = 2701$) HPI geographic regions. To examine whether the proportions of heat categories differed between low and high HPI areas, z scores were computed and two-sample z -score analyses were conducted for each extreme heat category. Tweets classified as neutral were excluded from the sentiment analysis. Null and alternative hypotheses were developed for each category. For three categories—coping, venting, and perceived extreme heat—we developed directional hypotheses. The hypotheses are laid out here rather than the introduction given that our first research question involved discovering the actual heat discourse categories first. We expected coping strategies to be discussed at higher frequencies in high HPI areas, venting to be discussed to a greater degree in low HPI areas, and perceived heat to be discussed to a greater extent in low HPI areas.

(i) Broadcasting heat warnings

H_0 : broadcasting heat warnings^{highHPI} = broadcasting heat warnings^{lowHPI}

H_a : broadcasting heat warnings^{highHPI} $\neq$ broadcasting heat warnings^{lowHPI}

(ii) Coping strategies

H_0 : coping strategies^{highHPI} = coping strategies^{lowHPI}

H_a : coping strategies^{highHPI} > 0.5 coping strategies^{lowHPI}

(iii) Life impact

H_0 : life impact^{highHPI} = life impact^{lowHPI}

H_a : life impact^{highHPI} $\neq$ life impact^{lowHPI}

(iv) Venting

H_0 : venting^{highHPI} = venting^{lowHPI}

H_a : venting^{highHPI} < 0.5 venting^{lowHPI}

(v) Perceived extreme heat

H_0 : perceived extreme heat^{highHPI} = perceived extreme heat^{lowHPI}

H_a : perceived extreme heat^{highHPI} < 0.5 perceived extreme heat^{lowHPI}

(vi) Discussing community action

H_0 : discussing community action^{highHPI} = discussing community action^{lowHPI}

H_a : discussing community action^{highHPI} $\neq$ discussing community action^{lowHPI}

(vii) Linking climate change to extreme heat

H_0 : linking climate change to extreme heat^{highHPI} = linking climate change to extreme heat^{lowHPI}

H_a : linking climate change to extreme heat^{highHPI} $\neq$ linking climate change to extreme heat^{lowHPI}

(viii) Expressing relief

H_0 : expressing relief^{highHPI} = expressing relief^{lowHPI}

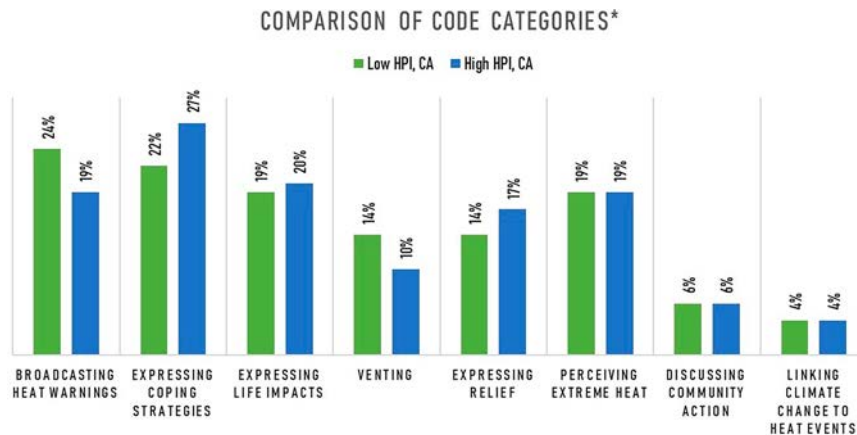


FIG. 4. Comparison of tweets in each code in high and low HPI areas. The percentages represent the proportion of tweets per respective HPI area. Because tweets could fall in multiple categories, the totals exceed 100%.

H_a : expressing relief^{highHPI} $\neq$ expressing relief^{lowHPI}

(ix) *Positive sentiment*

H_0 : positive sentiment^{highHPI} = positive sentiment^{lowHPI}

H_a : positive^{highHPI} > 0.5 positive sentiment^{lowHPI}

(x) *Negative sentiment*

H_0 : negative sentiment^{highHPI} = negative sentiment^{lowHPI}

H_a : negative sentiment^{highHPI} < 0.5 negative sentiment^{lowHPI}

3. Results

Results are reported in three sections: the content analysis of the eight categories that emerged regarding public tweets about extreme heat, comparisons across low and high HPI for the eight extreme heat categories, and the sentiment analysis of high and low HPI areas in California.

a. Content analysis

1) BROADCASTING EXTREME HEAT WARNINGS

The tweets containing weather forecast information, including warnings, specific weather-related information, and numeric temperatures, were coded as broadcasting extreme heat. They included tweets with signal words such as “warn,” “warnings,” “alert,” and “stay safe” and terminology referencing hazard impact, guidance, and forecast. Tweets containing prediction information, references to news about heat, temperatures, and tips related to extreme heat events were all coded as broadcasting extreme heat. The category included tweets from news agencies, news anchors, reporters, journalists, and social media personalities, along with the general population. It implies that the impact of extreme heat was significant enough to warrant expression on social media.

2) COPING DURING EXTREME HEAT

A subgroup of tweets expressed individual-level coping with high temperatures. Tweets included actions to seek relief from

extreme heat such as spending time indoors, protecting eyes and skin, staying hydrated, and using air conditioning, alongside strategies seeking relief from heat while outdoors like visiting pools or beaches. References to costs, maintenance, and presence or absence of air conditioning systems at home, in vehicles, at work, or in other public places were also included in this category.

3) LIFE IMPACT OF EXTREME HEAT

A subgroup of tweets captured a range of life impacts from extreme heat. Tweets talked about disruptions to daily life including inability to go outside, inability to perform outdoor leisure or work activities, cancellation of events and sports games, and other postponements due to heat. Other tweets mentioned impacts on health, concerns about food safety, impacts to work or income, changes in the work environment, and even the closure of businesses. The life impact category extended to conversations about heat effects on animals and pets, houseplants, green spaces, and other living things. Many tweets in this category occurred in combination with other tweet categories, such as broadcasting, venting, and coping.

4) VENTING

Venting tweets contained negative expressions regarding the extreme heat experience. These tweets contained words, tone, or punctuation reflecting anger, disgust, and other negative expressions. For example, they used expletives, capital letters, and exclamation marks to imply strong negative feelings.

5) PERSONALLY PERCEIVED EXTREME HEAT

Tweets that made statements about extreme heat without any explicit positive or negative connotation were categorized as perceiving extreme heat. This included descriptions of high temperatures without alluding to obvious positive or negative perceptions or coping methods, life consequences, or weather forecast information. These tweets also contained

TABLE 2. The content analysis of Twitter (X) extreme heat categories. Each tweet could be coded for more than one category.

Tweet category	Tweet example	Low HPI (<i>n</i> = 2701)	High HPI (<i>n</i> = 4572)
		<i>n</i> (%)	<i>n</i> (%)
Broadcasting heat warnings	FYI, @NWSLosAngeles issued an extended heat wave warning for this weekend, so make sure to stay safe and hydrated at graduation @CalPoly!	645 (0.24)	891 (0.19)
Expressing coping strategies	Glad I could get the AC in my car fixed as soon as the heatwave started to come to an end	604 (0.22)	1252 (0.27)
Expressing life impacts	Volleyball and Tennis vs BHS tomorrow (Tuesday) has been canceled due to concerns regarding the extreme heat. Stay hydrated TITANS	510 (0.19)	933 (0.20)
Venting	This heat wave is ridiculous, #sanfrancisco. May it never end!	389 (0.14)	452 (0.10)
Expressing relief	Awww, it's cool = now, 70° #heatwave	375 (0.14)	758 (0.17)
Perceiving extreme heat	Extreme heat!	504 (0.19)	857 (0.19)
Discussing community action	Heat waves are becoming far too frequent. And it's an equity issue. If you're poor and marginalized, you're likely to be much more vulnerable to extreme heat.	171 (0.06)	256 (0.06)
Linking climate change to heat events	@GreedRwc @VWins16 @PhysicalMeta @sfchronicle Its been 90 years since the last hurricane and not much documented prior to that. On top of record heatwave, record drought, and typical wildfires that have accelerated in size and frequency over the past 20 years. You're blind and/or ignorant if you think it's all coincidence	97 (0.04)	184 (0.04)

temperature comparisons between two geographical locations or merely a hashtag such as #heatwave. Even though they did not contain a specific stance, their presence on social media indicates acknowledgment of extreme heat as a significant life experience.

6) DISCUSSING COMMUNITY ACTION

These tweets not only contained past and present concerns but also discussed expected or aspirational actions as a community. These tweets addressed current community action or commented on implicitly expected community action, covering the following aspects: concern about high-risk populations or communities, such as older, vulnerable, and low-income people; community and state adaptation strategies like social support such as cooling centers, provision of free water and food, donation of supplies, creation of water kiosks, and fundraisers; critical infrastructure damage such as power outages and lack of air conditioning; and current community activities and adaptation strategies such as responses to brownouts, sociopolitical issues, and potential policy changes. Although these tweets were sparse in number, they provide crucial insights from users and contained information about local community initiatives and setting expectations to combat extreme heat. For example, they addressed power bills; limits and outages; impacts on the power grid; heat-related damage to roadways, runways, and railways; heat-related power outages; and initiatives such as solar panels.

7) LINKING CLIMATE CHANGE AND INCREASINGLY FREQUENT EXTREME HEAT EVENTS

These tweets linked weather disasters and extreme heat events to global climate change. Some tweets talked about long-term impacts, strategies, and reckoning of international experiences of extreme weather events. They predicted long-term or future trends by comparing past experiences and used a global perspective to comment on California's extreme weather.

8) EXPRESSING RELIEF

Contrary to our expectation that heat-related tweets would primarily have negative themes, we found tweets expressing positive experiences during heat waves. We found that expressions of relief from heat related to descriptions of escaping from heat waves. This code included tweets expressing positive sentiment during extreme heat events. This included expressions of gratitude for residing in a colder place relative to a hotter area or expressions of relief for lower temperatures. To rule out various factors that led to positive expressions, we excluded positive tweets discussing personal events unrelated to extreme heat.

b. Frequency of tweet heat topics by HPI area

All eight heat conversation categories were tweeted in both high and low HPI areas (see Table 2 and Fig. 4). Two of the eight categories showed significant differences between low

TABLE 3. Differences in tweet frequency of heat categories between low and high HPI areas in California.

Code	z	One- or two-sided t test	p	Interpretation
Broadcasting heat warnings	5.158	Two	<0.001	Significant; the null hypothesis is rejected
Coping strategies	4.605	One	<0.001	Significant; the null hypothesis is rejected
Life impact disruptions	1.5511	Two	0.123	Cannot reject the null hypothesis
Venting	-0.0452	One	0.516	Cannot reject the null hypothesis
Perceived extreme heat	0.1057	One	0.460	Cannot reject the null hypothesis
Discussing community action	-1.2456	Two	0.215	Cannot reject the null hypothesis
Linking climate change to heat events	0.9287	Two	0.362	Cannot reject the null hypothesis
Expressing relief	0.0269	Two	0.992	Cannot reject the null hypothesis

and high HPI areas: broadcasting heat warnings and coping. Broadcasting extreme heat warnings were tweeted significantly more frequently in low HPI areas compared with high HPI areas ($z = 5.16$, $p < 0.001$). By contrast, coping strategies were tweeted significantly more in high HPI areas ($z = 4.61$, $p < 0.001$). Coping constituted 27% of tweets in high HPI areas and 22% in low HPI areas (see Table 3). For all other heat conversation categories, tweets were expressed at relatively the same proportions in both HPI areas.

c. Sentiment analysis

We also classified a sample of tweets based on sentiment expression into positive, negative, and neutral themes. Tweets classified as positive had overarching positive language suggestive of positive experiences of heat. The negative classification included tweets that contained negative expressions of extreme heat experience. Overall, we found significantly higher proportions of positive sentiment tweets among high HPI areas ($z = 4.54$, $p < 0.001$; Table 4) and significantly higher proportions of negative tweets among low HPI areas ($z = -5.52$, $p < 0.001$; Table 4).

4. Discussion

Extreme heat events pose an often underestimated threat to human health (Mayrhuber et al. 2018). Effective heat communication warning systems reduce mortality related to heat, yet it is unclear if they reach the public and how especially vulnerable communities respond (Lowe et al. 2011). In light of increasing heat events, social media may provide real-time

insights into public experiences. Our study examined how California residents experienced and discussed extreme heat on social media in California during a 7-yr period. Top heat topics in tweets included perceived heat risk, life impacts, and coping, followed by venting, heat warnings, need for community action, relief from heat, and climate change concerns. This answered our first research question regarding how the general public tweets about heat.

Social media platforms have challenged the position of traditional media sources such as radio, television, newspapers, and magazines as mediators of the discourse on topics of concern (Roxburgh et al. 2019). The most prevalent heat discourse involved perceiving heat risk, which included declarations of heat, similar to findings by Lambrecht et al. (2021) in their longitudinal study of the heat discourse on Facebook in Phoenix, Arizona. Our 7-yr heat tweet dataset expressed a range of life impacts, including disruptions and cancellations in leisure and work, food supply challenges (spoiling from heat), and energy impacts (school cancelled due to power outages and extreme heat advisories). Fewer tweets addressed public health concerns, such as heat at night that affected sleep and poor housing conditions. Zander et al. (2023) found that 26% of relevant tweets discussed the impacts of heat and heat waves in a 4-month period in 2022 via topic modeling. Our study's assessment of tweets about the life impact of heat indicates that they represented 22% of all tweets during the summer months across 7 years. Although our study found that the greatest number of tweets were in the categories of perceptions of heat risk, life impact, and coping, Zander et al. (2023) found only 13% of their tweets involved perceptions and warnings. Their study also discovered 17% discourse attributing extreme weather events to climate change, compared to only 4% in our tweet dataset. Zander et al. (2023) also reported 20% of tweets related to health impacts, whereas our analysis found negligible health impact discourse (the few relevant tweets were categorized as life impacts). This difference may be attributed to our use of LLM to interpret tweet content, allowing for more contextualized interpretation rather than topic modeling.

Our eight identified extreme heat social media categories coincide with a climate change vulnerability assessment to determine real-time exposure, sensitivity, and adaptive capacity to increasing frequency and intensity of extreme heat events resulting from climate change (see Fig. 5; Manangan et al. 2014). The vulnerability assessment framework informs various stages of building resilience against climate change effects

TABLE 4. The sentiment analysis of extreme heat tweets in high and low HPI areas. * $p < 0.001$.

		High	Low
Sentiment		n (%)	n (%)
Positive*	Grateful to be missing the Artic Blast on East coast! Although we have a heat wave in SF.	1008 (22)	330 (12)
Negative*	I absolutely hate the SUMMER. . .rather be cold than hot anyway!#summer #summer17#Heatwave2016	1735 (38)	1296 (48)
Neutral	Record heat wave in the forecast! 80F!! Lol!!	1829 (40)	1075(40)

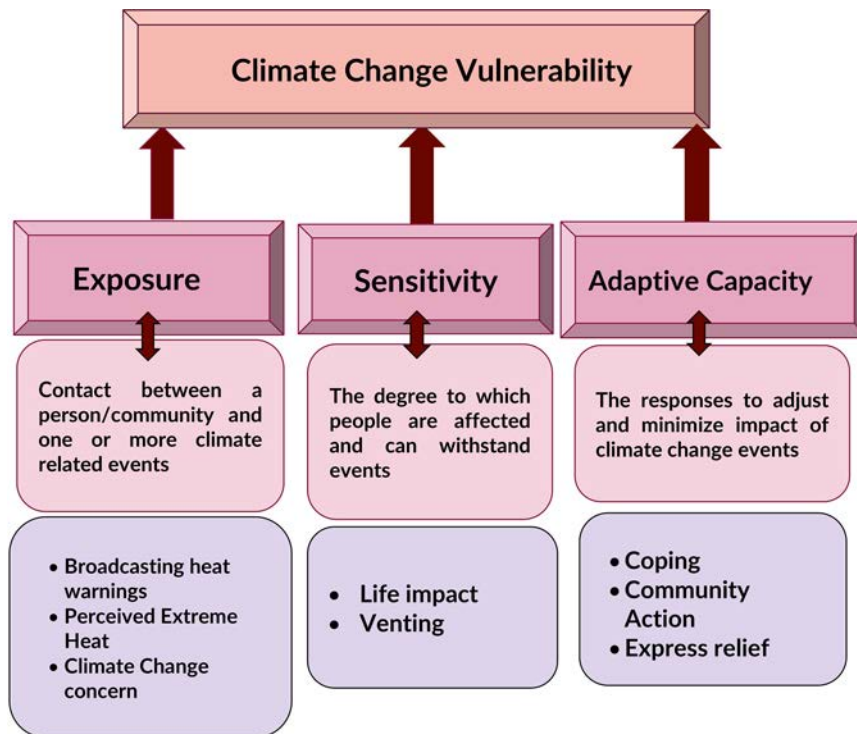


FIG. 5. Extreme heat codes contextualized in the climate change vulnerability framework. Adapted from Manangan et al. (2014).

and aids health departments and agencies in preparing and responding to climate change and related rise in extreme heat events. This correlation not only affords relevance to the identified codes but also alludes to their practical utility.

Given that extreme heat events are experienced differently across geographic areas and vulnerable communities (Manangan et al. 2014; Mayrhuber et al. 2018), we examined whether heat discourse differed between low and high HPI areas. All eight heat discourse themes occurred in both high and low HPIs in California, with some notable differences. High HPI areas had more tweets with positive sentiment and about coping, whereas low HPI areas had more tweets with negative sentiment and about venting and heat warnings. These findings align with California's extreme heat action plan (Newsom 2022), which calls for deeply engaging heat-affected communities and taking different actions based on the level of heat threat and community adaptive capacity (Dunlap 2024).

Results regarding the differential heat discourse in low and high HPI areas regarding coping, warnings, and sentiment have implications for future heat communication. Low HPI areas, representing vulnerable communities, may need more communication about coping and local government support. Zottarelli et al.'s (2022) national evaluation of 10 local government websites presenting heat information found omission of critical protective action especially for vulnerable communities. Low HPI areas may have enough warnings, given that they may be living with recurring and extended heat events. Communication may need to emphasize community-level

adaptive capacity and resources in addition to individual-level coping. The statistically higher prevalence of expressions about coping strategies in high HPI areas and contrastingly higher prevalence of broadcasting heat warnings in low HPI areas require further exploration and research, given that they suggest vulnerabilities of low HPI populations to heat events.

Our study aimed to highlight the utility of HPI as a tool to identify extreme heat vulnerabilities and contextualize population-level vulnerabilities in a composite index. By combining HPI with social media monitoring, we can better understand experiences of extreme heat and how structural determinants may influence these experiences. This approach can inform the design of targeted public health interventions to mitigate adverse health outcomes, particularly in marginalized and vulnerable communities (Bernhard et al. 2015; Harlan et al. 2012; Reid et al. 2009).

The Integrated Climate Adaptation and Resiliency Program (2024) at California's Office of Planning and Research identified the lack of an integrated tool to understand climate change in California. The office also highlighted the need to identify data from communities that are vulnerable to climate challenges to ensure equitable distribution of adaptation resources. Our study is a step toward applying the HPI as a social equity tool to identify vulnerable populations and automate social media expressions to determine what people have to say about extreme heat. Automated language processing tools applied in the context of social equity indexes may be a good way to not only

identify vulnerabilities but also see how effects of climate change and extreme heat are experienced and expressed.

In their multicity study, [Young et al. \(2021\)](#) showed that the volume of heat-related tweets increased as temperatures rose. They also found that sentiment changes were predictable in many cities. This is similar to the findings in our study, except that we extended beyond geography to examine sentiment by low and high HPI areas. We found significant correlations not only between the number of tweets and temperature hikes but also between sentiment and HPI score.

A study on mass media and climate change found that variations in temperatures and extreme weather events prompted people to connect these personal experiences to climate change ([Kirilenko and Stepchenkova 2014](#)), similar to findings in our study. [Kirilenko and Stepchenkova \(2014\)](#) corroborated that tweets can be used to gauge climate change perceptions. Our study went a step further by including social equity factors in perceptions of climate change. Using machine learning to automate the tweet analysis could have practical applications for policymakers to not only gauge sentiment but also identify zip code or census tract-level expressions of concern in a timely manner, possibly even in real time. Because of their real-time nature, social media broadcasts by the public may be employed to predict areas of need or at risk of extreme heat for local government to focus interventions. Alternatively, social media discourse may be used to gauge social and life impacts of heat.

Limitations

Even though personal experiences and related social media expressions are accessible measures, they can introduce bias in understanding how extreme heat events are experienced by the public ([Weber and Stern 2011](#)). Given that Twitter's user base is mostly young adults with access to the internet and smartphones and who have higher incomes on average ([Steinert-Threlkeld 2018](#)), the generalizability of heat event perceptions to California's other demographics including older, more rural, lower income, lower educational attainment, and non-English speakers limits the generalizability especially in relation to communities vulnerable to extreme heat. This particular social media platform may overlook those who are particularly vulnerable and have fewer resources to cope with extreme heat.

A second limitation related to our geotagging method of using the central point (centroid) of a bounding box to geolocate tweets introduces four potential biases: urban–rural representation bias, edge effect misassignment, population density oversimplification, and algorithmic structural bias ([Malik et al. 2021](#); [Yin et al. 2022](#)). These potential biases (of assuming uniform tweet distribution, uneven population density) could amplify underrepresentation of rural populations in analyses and systematically disadvantage regions with low-density tweets (e.g., rural regions). While population-weighted centroids can be used to mitigate urban–rural representation bias, this is recommended as a mitigation strategy when bounding boxes reflect > 5 mi. Most of CA's bounding boxes were less than 10 mi in size.

A third limitation relates to our keyword-filtering data retrieval process. Using different keywords may result in capturing a different set of tweets. We outline and make transparent our use of heat-related keywords and strategy to make our data retrieval process transparent and to be able to replicate our study. Additionally, while energy impacts are considerable during extreme heat events, we decided to exclude tweets related to blackouts and brownouts because these keywords resulted in the dataset being too noisy with unrelated tweets.

5. Conclusions

Our research demonstrates the value of integrating place-based health equity tools like HPI with social media monitoring to enhance our understanding of extreme heat experiences and their relationship to structural- and place-based determinants. This approach can inform the development of targeted public health interventions, particularly for marginalized and vulnerable communities in low HPI areas, while leveraging community assets to address persistent and extreme heat events. Our findings revealed significant differences between low and high HPI areas, with low HPI areas expressing less coping strategies, more venting and negative sentiment, and receiving more heat warnings. These insights suggest that heat mitigation campaigns in low HPI areas may benefit from focusing on coping mechanisms.

This study addresses a gap in the literature by mapping heat discourse across neighborhoods with varying place-based drivers of heat adaptive capacity. The HPI framework identifies potential areas for policy action, while real-time heat discourse can validate and refine public health interventions. The combination of social media monitoring and automated language processing tools offers a valuable method for capturing real-time experiences and sentiment. Future research can build on this approach by using place-based social equity tools like HPI to target heat warnings, provide actionable recommendations, and plan community-level mitigation measures against extreme heat and weather events. This integrated approach has the potential to significantly improve our response to climate-related health challenges and reduce disparities in heat vulnerability.

Acknowledgments. This work was supported by the National Science Foundation Award IIS-2107150.

Data availability statement. The dataset—i.e., tweet identifiers generated and analyzed for the current study—is available upon reasonable request.

APPENDIX

Heat Discourse Twitter (X) Keywords Used in Filter Strategy from Twitter Archive

This appendix presents the keyword sets used in attempts for trials of deriving relevant extreme heat discourse tweets.

a. Attempt 1: Filtering the Twitter archive using the following keywords

(extreme heat, air conditioning, car heat, dry heat, heat intensifies, August heat, heat damn, heat extremely, heat home, heat index, heat intensifies, heat July, heat June, heat killing, heat may, heat office, heat related, heat school, heat September, heat sun, heat unbearable, heat warning, heat watch, heat winds, heat exhaustion, heatstroke, heat-stroke, humidity heat, melt heat, overheat, overheated, sleep heat, steam heat, summer heat, sunstroke, sunstroke, this heat, unbearable heat, cooling center, coolingctr) (place: CA, place: California).

b. Attempt 2: With the use of the following keywords

(heatwave, heat wave, sweltering heat, unbearable heat, heat warning, heat watch, heat stroke, heat stroke, heatstroke, extreme heat, extreme temperature, summer heat, sunstroke, heat index, heat intensifies, July heat, August heat, September heat, school heat, sleep heat, cooling center, energy assistance, air conditioning, power outage, brownout, black out) (place: CA, place: California).

c. Attempt 3: With the use of the following keywords

(heat wave, extreme heat, extremeheat, heat, hot, hotter, temperature, humid, humidity) (place: CA, place: California).

REFERENCES

- Abedin, B., and A. Babar, 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, <https://doi.org/10.1007/s10796-017-9789-4>.
- Baylis, P., N. Obradovich, Y. Kryvasheyev, H. Chen, L. Coviello, E. Moro, M. Cebrian, and J. H. Fowler, 2018: Weather impacts expressed sentiment. *PLOS ONE*, **13**, e0195750, <https://doi.org/10.1371/journal.pone.0195750>.
- Bernhard, M. C., S. T. Kent, M. E. Sloan, M. B. Evans, L. A. McClure, and J. M. Gohlke, 2015: Measuring personal heat exposure in an urban and rural environment. *Environ. Res.*, **137**, 410–418, <https://doi.org/10.1016/j.envres.2014.11.002>.
- Bowen, G. A., 2008: Naturalistic inquiry and the saturation concept: A research note. *Qual. Res.*, **8**, 137–152, <https://doi.org/10.1177/1468794107085301>.
- Cal-Adapt, 2024: Cal-Adapt climate tools, extreme heat days & warm nights. UC Berkeley and USGS, accessed 16 June 2025, <https://cmip5.cal-adapt.org/data/download/>.
- CalEPA, 2023: CalEnviroScreen 10th anniversary StoryMap commemorating 10 years of CalEnviroScreen. <https://storymaps.arcgis.com/stories/413d2b6be94c42ce85ada08499623a2a>.
- Centers for Disease Control and Prevention and Agency for Toxic Substances and Disease Registry, 2024: What is the CDC/ATSDR social vulnerability index?, 1 pp., <https://www.atsdr.cdc.gov/place-health/media/pdfs/2024/07/SVI-Fact-Sheet-H.pdf>.
- Cissé, G., and Coauthors, 2022: Health, wellbeing, and the changing structure of communities. *Climate Change 2022: Impacts, Adaptation and Vulnerability*, H.-O. Pörtner et al., Eds., Cambridge University Press, 1041–1170.
- Dunlap, L., 2024: Communicating heat risk: A guide to inclusive, effective, and coordinated public information campaigns. UCLA Rep., 27 pp., <https://innovation.luskin.ucla.edu/wp-content/uploads/2024/03/Communicating-Heat-Risk.pdf>.
- Dzyuban, Y., G. N. Y. Ching, S. K. Yik, A. J. Tan, P. J. Crank, S. Banerjee, R. X. Y. Pek, and W. T. L. Chow, 2022: Sentiment analysis of weather-related tweets from cities within hot climates. *Wea. Climate Soc.*, **14**, 1133–1145, <https://doi.org/10.1175/WCAS-D-21-0159.1>.
- Flanagan, B. E., E. W. Gregory, E. J. Hallisey, J. L. Heitgerd, and B. Lewis, 2011: A social vulnerability index for disaster management. *J. Homeland Secur. Emerg. Manage.*, **8**, 3, <https://doi.org/10.2202/1547-7355.1792>.
- Grasso, V., A. Crisci, M. Morabito, P. Nesi, and G. Pantaleo, 2017: Public crowdsensing of heat waves by social media data. *Adv. Sci. Res.*, **14**, 217–226, <https://doi.org/10.5194/asr-14-217-2017>.
- Gronlund, C. J., and V. J. Berrocal, 2020: Modeling and comparing Central and room air conditioning ownership and cold-season in-home thermal comfort using the American Housing Survey. *J. Exposure Sci. Environ. Epidemiol.*, **30**, 814–823, <https://doi.org/10.1038/s41370-020-0220-8>.
- Guirguis, K., A. Gershunov, A. O. Tardy, and R. Basu, 2014: The impact of recent heat waves on human health in California. *J. Appl. Meteor. Climatol.*, **53**, 3–19, <https://doi.org/10.1175/JAMC-D-13-0130.1>.
- Hallgren, K. A., 2012: Computing inter-rater reliability for observational data: An overview and tutorial. *Quant. Methods Psychol.*, **8**, 23–34, <https://doi.org/10.20982/tqmp.08.1.p023>.
- Harlan, S. L., J. H. Deplet-Barreto, W. L. Stefanov, and D. B. Petitti, 2012: Neighborhood effects on heat deaths: Social and environmental predictors of vulnerability in Maricopa County, Arizona. *Environ. Health Perspect.*, **121**, 197–204, <https://doi.org/10.1289/ehp.1104625>.
- Integrated Climate Adaptation and Resiliency Program, 2024: Vulnerable Communities Platform. <https://vcp.opr.ca.gov/>.
- IPCC, 2021: Mitigating climate change must be a priority for public health. *Lancet Public Health*, **6**, e620, [https://doi.org/10.1016/S2468-2667\(21\)00190-0](https://doi.org/10.1016/S2468-2667(21)00190-0).
- Jung, J., and C. K. Uejio, 2017: Social media responses to heat waves. *Int. J. Biometeor.*, **61**, 1247–1260, <https://doi.org/10.1007/s00484-016-1302-0>.
- Kiarsi, M., M. Amiresmaili, M. R. Mahmoodi, H. Farahmandnia, N. Nakhaee, A. Zareian, and H. Aghababaeian, 2023: Heat waves and adaptation: A global systematic review. *J. Thermal Biol.*, **116**, 103588, <https://doi.org/10.1016/j.jtherbio.2023.103588>.
- Kirilenko, A. P., and S. O. Stepchenkova, 2014: Public microblogging on climate change: One year of twitter worldwide. *Global Environ. Change*, **26**, 171–182, <https://doi.org/10.1016/j.gloenvcha.2014.02.008>.
- Knowlton, K., M. Rotkin-Ellman, G. King, H. G. Margolis, D. Smith, G. Solomon, R. Trent, and P. English, 2008: The 2006 California heat wave: Impacts on hospitalizations and emergency department visits. *Environ. Health Perspect.*, **117**, 61–67, <https://doi.org/10.1289/ehp.11594>.
- Kryvasheyev, Y., H. Chen, N. Obradovich, E. Moro, P. Van Hentenryck, J. H. Fowler, and M. Cebrian, 2016: Rapid assessment of disaster damage using social media activity. *Sci. Adv.*, **2**, e1500779, <https://doi.org/10.1126/sciadv.1500779>.
- Lambrecht, K., B. J. Hatchett, K. VanderMolen, and B. Feldkircher, 2021: Identifying community values related to heat: Recommendations for forecast and health risk communication. *Geosci. Commun.*, **4**, 517–525, <https://doi.org/10.5194/gc-4-517-2021>.

- Liu, Y., and Coauthors, 2019: RoBERTa: A robustly optimized BERT pretraining approach. arXiv, 1907.11692v1, <https://doi.org/10.48550/arXiv.1907.11692>.
- Lowe, D., K. L. Ebi, and B. Forsberg, 2011: Heatwave early warning systems and adaptation advice to reduce human health consequences of heatwaves. *Int. J. Environ. Res. Public Health*, **8**, 4623–4648, <https://doi.org/10.3390/ijerph8124623>.
- Maizlish, N., and Coauthors, 2019: California Healthy Places Index: Frames matter. *Public Health Rep.*, **134**, 354–362, <https://doi.org/10.1177/0033354919849882>.
- Malik, M., H. Lambda, C. Nakos, and J. Pfeffer, 2021: Population bias in geotagged tweets. *Proc. Int. AAAI Conf. Web Soc. Media*, **9**, 18–27, <https://doi.org/10.1609/icwsm.v9i4.14688>.
- Manangan, A. P., and Coauthors, 2014: Assessing health vulnerability to climate change: A guide for health departments. CDCP Climate and Health Tech. Rep., 24 pp., https://www.cdc.gov/climate-health/media/pdfs/AssessingHealthVulnerabilitytoClimateChange_1.pdf.
- Mann, R., and J. Schuetz, 2022: As extreme heat grips the globe, access to air conditioning is an urgent public health issue. *Brookings*, 25 July, <https://www.brookings.edu/articles/as-extreme-heat-grips-the-globe-access-to-air-conditioning-is-an-urgent-public-health-issue/>.
- Masri, S., Y. Jin, and J. Wu, 2022: Compound risk of air pollution and heat days and the influence of wildfire by SES across California, 2018–2020: Implications for environmental justice in the context of climate change. *Climate*, **10**, 145–165, <https://doi.org/10.3390/cli1010145>.
- Mayrhuber, E. A.-S., and Coauthors, 2018: Vulnerability to heatwaves and implications for public health interventions—A scoping review. *Environ. Res.*, **166**, 42–54, <https://doi.org/10.1016/j.envres.2018.05.021>.
- Mendez, M., 2020: *Climate Change from the Streets: How Conflict and Collaboration Strengthen the Environmental Justice Movement*. Yale University Press, 304 pp.
- Newsom, G., 2022: Protecting Californians from extreme heat: A state action plan to build community resilience. California Natural Resources Agency Rep., 72 pp., <https://resources.ca.gov/-media/CNRA-Website/Files/Initiatives/Climate-Resilience/2022-Final-Extreme-Heat-Action-Plan.pdf>.
- Obradovich, N., and K. Minor, 2022: Identifying and preparing for the mental health burden of climate change. *JAMA Psychiatry*, **79**, 285–286, <https://doi.org/10.1001/jamapsychiatry.2021.4280>.
- Office of Environmental Health Hazard Assessment, 2024: CalEnviroScreen 4.0. OEHHHA, accessed 15 June 2025, <https://oehha.ca.gov/calenviroscreen/report/calenviroscreen-40>.
- Ostro, B. D., L. A. Roth, R. S. Green, and R. Basu, 2009: Estimating the mortality effect of the July 2006 California heat wave. *Environ. Res.*, **109**, 614–619, <https://doi.org/10.1016/j.envres.2009.03.010>.
- Palinkas, L. A., and Coauthors, 2022: Vulnerable, resilient, or both? A qualitative study of adaptation resources and behaviors to heat waves and health outcomes of low-income residents of urban heat islands. *Int. J. Environ. Res. Public Health*, **19**, 11090, <https://doi.org/10.3390/ijerph191711090>.
- Perez, E., 2023: It's official: July 2023 was the warmest month ever recorded. *UN News*, 8 August, <https://news.un.org/en/story/2023/08/1139527>.
- Reid, C., M. S. O'Neill, C. J. Gronlund, S. J. Brines, D. G. Brown, A. V. Diez-Roux, and J. S. Schwartz, 2009: Mapping community determinants of heat vulnerability. *Environ. Health Perspect.*, **117**, 1730–1736, <https://doi.org/10.1289/ehp.0900683>.
- Roxburgh, N., D. Guan, K. J. Shin, W. Rand, S. Managi, R. Lovelace, and J. Meng, 2019: Characterising climate change discourse on social media during extreme weather events. *Global Environ. Change*, **54**, 50–60, <https://doi.org/10.1016/j.gloenvcha.2018.11.004>.
- Schwarz, L., and Coauthors, 2020: The health burden of fall, winter and spring extreme heat events in Southern California and contribution of Santa Ana winds. *Environ. Res. Lett.*, **15**, 054017, <https://doi.org/10.1088/1748-9326/ab7f0e>.
- Steinert-Threlkeld, Z. C., 2018: *Twitter as Data*. Cambridge University Press, 116 pp.
- Tracy, S. J., 2019: Advanced data analysis: The art and magic of interpretation. *Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact*, John Wiley and Sons, 203–226.
- Twitter.com, 2010: Streaming API documentation. <https://dev.twitter.com/docs>.
- University of California, Los Angeles, 2022: Healthy Places Index (HPI) 3.0 Data dictionary. <https://www.healthypacesindex.org>.
- Watts, N., and Coauthors, 2019: The 2019 report of The Lancet Countdown on health and climate change: Ensuring that the health of a child born today is not defined by a changing climate. *Lancet*, **394**, 1836–1878, [https://doi.org/10.1016/S0140-6736\(19\)32596-6](https://doi.org/10.1016/S0140-6736(19)32596-6).
- Weber, E. U., and P. C. Stern, 2011: Public understanding of climate change in the United States. *Amer. Psychol.*, **66**, 315–328, <https://doi.org/10.1037/a0023253>.
- Yin, J., Y. Gao, and G. Chi, 2022: An evaluation of geo-located Twitter data for measuring human migration. *Int. J. Geogr. Inf. Sci.*, **36**, 1830–1852, <https://doi.org/10.1080/13658816.2022.2075878>.
- Young, J. C., R. Arthur, M. Spruce, and H. T. P. Williams, 2021: Social sensing of heatwaves. *Sensors*, **21**, 3717, <https://doi.org/10.3390/s21113717>.
- Young, S. D., 2018: Social media as a new vital sign: Commentary. *J. Med. Internet Res.*, **20**, e161, <https://doi.org/10.2196/jmir.8563>.
- Zander, K. K., J. Rieskamp, M. Mirbabaie, M. Alazab, and D. Nguyen, 2023: Responses to heat waves: What can twitter data tell us? *Nat. Hazards*, **116**, 3547–3564, <https://doi.org/10.1007/s11069-023-05824-2>.
- Zapf, A., S. Castell, L. Morawietz, and A. Karch, 2016: Measuring inter-rater reliability for nominal data – Which coefficients and confidence intervals are appropriate? *BMC Med. Res. Methodol.*, **16**, 93, <https://doi.org/10.1186/s12874-016-0200-9>.
- Zottarelli, L. K., S. A. Blake, and M. T. Garza, 2022: Communicating heat-health information to the public: Assessing municipal government extreme heat event website content. *Wea. Climate Soc.*, **14**, 311–321, <https://doi.org/10.1175/WCAS-D-21-0019.1>.
- Zuñiga, M. E., and M. Méndez, 2024: The emergence of environmental justice in general plans: Lessons from California's senate bill 1000. *Urban Aff. Rev.*, **60**, 1123–1161, <https://doi.org/10.1177/10780874231175841>.