International Journal of Applied Earth Observation and Geoinformation 134 (2024) 104205

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Public responses to heatwaves in Chinese cities: A social media-based geospatial modelling approach

Mingxuan Dou [a], Yandong Wang [a], Mengling Qiao [b], Dongyang Wang [a], Jianya Gong [a,c], Yanyan Gu [a,*]

[a] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
[b] Department of Earth and Environmental Engineering, Columbia University, New York, NY 10027, USA
[c] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

ARTICLE INFO

Keywords:
Spatial disparities
Public attention
Heatwaves
Social sensing

ABSTRACT

Increasing exposure to heatwaves threatens public health, challenging various socioeconomic sectors in the coming decades. Prior studies mostly concentrated on the heatwaves occurring in specific regions by examining temperature durations, ignoring the fact that heatwaves typically swept across a large area. To comprehensively assess the effects of heatwaves, we jointly analyzed public attention to heatwaves using a dataset of over 10 million geo-located Weibo tweets across 321 cities in China. By considering spatial disparities, two kinds of public attention at city level, namely the number of heat-related tweets (NHTs) and the ratio of heat-related tweets (RHTs), were designed to indicate the severity and location of heatwave impacts, respectively. The heat cumulative intensity was used as a proxy for heatwaves, which exhibited more significant correlations with RHTs than NHTs. The multiscale geographically weighted regression (MGWR) model was employed to investigate the spatiotemporal variations of environment, demographic, and economic-social factors. Six city groups were clustered with MGWR coefficients that were consistent with the seven geographic subregions of China. This research provides a new perspective and methodology for public attention to heatwaves using geo-located social sensing data and highlights the need for actions to mitigate future heatwave stress in sensitive cities.

1. Introduction

Persistent temperature extremes can harm ecosystems and pose a direct threat to human health (Wei et al., 2023; Bogdanovich et al., 2023). Over the past decades, there have been increased long-lasting and geographically widespread heatwaves in various regions of the world, as evidenced by Southern Africa (Meque et al., 2022), Europe, the Pacific Northwest (White et al., 2023), and Eastern China (Jiang et al., 2023). Under the combined effects of high-intensity greenhouse gases and urban development, it is anticipated that heatwaves will become more frequent, intense, and extreme across the globe (Christidis et al., 2020). Heatwaves are now the deadliest natural disasters, known as the silent killer (Luber and McGeehin, 2008). Therefore, accurate and timely sensing of the region where the heatwave occurs, the population it affects, and the extent of its social attention are essential for heatwave preparedness and response.

Exposure to extreme heat can lead to a significant increase in mortality (Giorgini et al., 2017), mental health issues (Thompson et al., 2018), and substantial economic losses (Hsiang et al., 2017). In general, previous studies on heatwave exposure primarily utilized earth observations and statistical data to monitor ambient temperature and humidity, which makes it hard to quantify the thermal discomfort of the population (Cecinati et al., 2019). On the other hand, some researchers capture the heat effects on public health using sanitary-type indicators (e.g., heatwave-attributable mortality or morbidity and hospital admissions). However, due to sample sparsity and annual statistical scale, current datasets can hardly meet existing requirements for large-scale and timely estimates of public attention to heatwaves, especially in developing countries.

Social media has become an essential tool for assessing damage from major natural disasters such as earthquakes, floods, hurricanes, and wildfires, offering comprehensive geographical coverage and high data density (Huang et al., 2022; Feng et al., 2022; Xiao et al., 2015). Utilizing social sensing data provides three key advantages in evaluating heatwave effects. Firstly, the extensive reach of social media allows for analysis over large geographical areas. Secondly, user-generated content

* Corresponding author at: Wuhan University, 129 Luoyu Road, Wuhan 430079, China.
E-mail address: yyg@whu.edu.cn (Y. Gu).

https://doi.org/10.1016/j.jag.2024.104205
Received 12 October 2023; Received in revised form 10 May 2024; Accepted 1 October 2024
Available online 6 October 2024

related to heat, such as tweets, captures a holistic view of thermal comfort by reflecting not only meteorological conditions like temperature and humidity but also socio-economic and demographic factors affecting adaptability (Song et al., 2020; He et al., 2019). Lastly, social media metrics inherently account for the differential impact of heatwaves in various settings, factoring in population and infrastructure distribution, and amplifying public perceptions of heatwave severity in densely populated urban areas compared to remote regions.

Previous studies have proven the value and effectiveness of social media in understanding heat wave response and exposure, but the heterogeneity of social media activity and its influencing factors have received scant attention. Due to the complexity of human activities during heatwaves, it is necessary to have an in-depth comprehension of how a combination of meteorological variables, together with demographic and socio-economic variables, influence public attention towards heatwaves. Additionally, research on a large geographical scale can better reveal these differences, assist decision-makers in making informed decisions, and help urban residents reduce heat stress more effectively. However, few studies have examined geospatial disparities in how people sensed heatwaves on a large geographical scale.

In this context, this study aimed to explore public attention to heatwaves in mainland China using geo-located social media data, a region whose ecosystem and social development are vulnerable to extreme heat events. A BERT-based language model was employed to extract heat-related Weibo tweets from June to September 2022. The associations between heatwaves and public attention were quantified using regression models. Local relationships had been modeled by the multiscale geographical weighted regression (MGWR) model, and K-means++ was used to identify cluster city groups based on the coefficients of the MGWR. Three contributions were achieved: (1) establish a social media data analytics framework to track and mine public information related to heatwaves; (2) characterize the local drivers of heat perception and explore the spatial stationarity of public attention to heatwaves; and (3) classify nationwide city groups by heatwave perception, which were consistent with the seven geographic subregions of China.

2. Literature review

2.1. Social media as an indicator of natural hazards

Crowdsourced social media data is widely recognized as a valuable resource for information mining, situation awareness, and coordinating relief in the natural disaster management field (Resch et al., 2018; Ji et al., 2022; Rui, 2023). Based on its inherent advantages of vast volume and low cost, researchers have developed plenty of human-centered approaches to enhance emergency response and inform decision-making. The severity and extent of a natural hazard can be detected by tracing digital footprints on social media (Dou & Gu, 2022).

Kryvasheyeu et al. (2016) analyzed the response of Twitter users to Hurricane Sandy in the United States. Their results showed that the volume and composition of Twitter streams correspond directly to hurricane threats and damage distribution, but they noted that caution should be taken when developing practical tools. A similar study was conducted in South Carolina by Li et al. (2018), which leveraged flood-related tweets for flood mapping in near real time. The generated flood mapping was found to be consistent and comparable with official inundation maps. Meanwhile, Fang et al. (2019) examined the social media activities on the Weibo platform during the 2016 Wuhan rainstorm in China. They pointed out that flood-related Weibo activities were consistent with the disaster process in various aspects, including temporal variation, topic evolution, and spatial hotspots. In addition, other metrics derived from the content of tweets, such as the number of topics or sentiment level, can contribute a richer dimension to disaster assessment. For example, Li et al. (2021) demonstrated the value of parsing textual information from social media and spatial density to quantitatively estimate the geographically distributed damage in the 2019 Ridgecrest, California, earthquake.

Geo-located social media data is particularly pertinent because it contains a great deal of metadata, including text and images, along with precise geographic locations and temporal information. As a crowdsourcing approach, these data provide observations about actual events that occurred in the real world. Although social media may suffer from the data quality (e.g., unstructured, noisy, and high uncertainty), sample bias (e.g., uneven target population and local penetration), and fake information (Goodchild and Glennon, 2010). Hence, this study asserts that geolocated social media data with high spatial and temporal resolution remains a significant asset in understanding the public's perception of natural hazards, such as heatwaves, and in assessing the potential effects of heatwaves on local communities. The present study proposes the utilization of heat-related social media activities as a surrogate measure for assessing the occurrence and intensity of heatwaves.

2.2. Social sensing in heatwaves

As the frequency of extreme weather events rises, researchers are increasingly drawn to the use of social activities to examine the public's perspective on weather. Some scholars have studied the effect of weather conditions on sentiment expressed on social media (Li et al., 2014). Baylis et al. (2018) conducted a representative study using data from 75 US metropolitan areas from both Facebook and Twitter. They found that weather conditions such as cold/hot temperatures, precipitation, cloud cover, etc., can worsen expressions of sentiment regardless of whether weather-related terms are included. Furthermore, Zander et al. (2023) demonstrated that combining heat-related sentiment and monitored data can help to estimate a spatial temperature map at the district level. At both city and national level, Lyu et al. (2024) used Twitter data to map heat exposure in near real time.

Another stream of research tried to explore the relationship between heatwaves and human responses. As pioneers, Jung & Uejio (2017) explored heat-related tweets in five U.S. cities. Using the autoregressive integrated moving average (ARIMA) time series model, they found a significant positive correlation between heat exposure and the number of tweets in three of the five cities. People's sensitivity to heatwaves derived from social media and heat exposure derived from extreme temperatures were analyzed in Wang et al. (2021). With global Twitter conversations, Zander et al. (2023) investigated the intensity of Twitter activity and changes in individual behavior during hot days. A strong correspondence between the two can be identified in regions like Argentina, Australia, the USA, and South Asia.

Despite the success of these studies, there are still significant theoretical and technical problems with the use of social media data for heatwave-related research. Important technical problems include how to mine the information related to thermal perception in the massive data. With the increasing amount of data generated by social media, it is necessary to develop a social media data mining framework to extract useful information from social media data and conduct comparative analysis across events and spaces. In terms of theoretical challenges, a major issue is the social and geographical differences that influence public attention on social media. This disparity can have serious consequences for emergency management and disaster recovery capabilities (Zou et al., 2023; Yuan et al., 2021). To solve the above challenges, this study takes the summer heat wave event in China in 2022 as an example to explore the extraction, spatio-temporal patterns, and driving factors affecting public attention to heatwaves on social media platforms.

3. Material and methods

This study employs a four-step research strategy. Firstly, we collected Weibo data, weather data, and other geographical data from open-source platforms. The next step was to fine-tune the pre-trained Bert-based model and employ the most effective models to identify

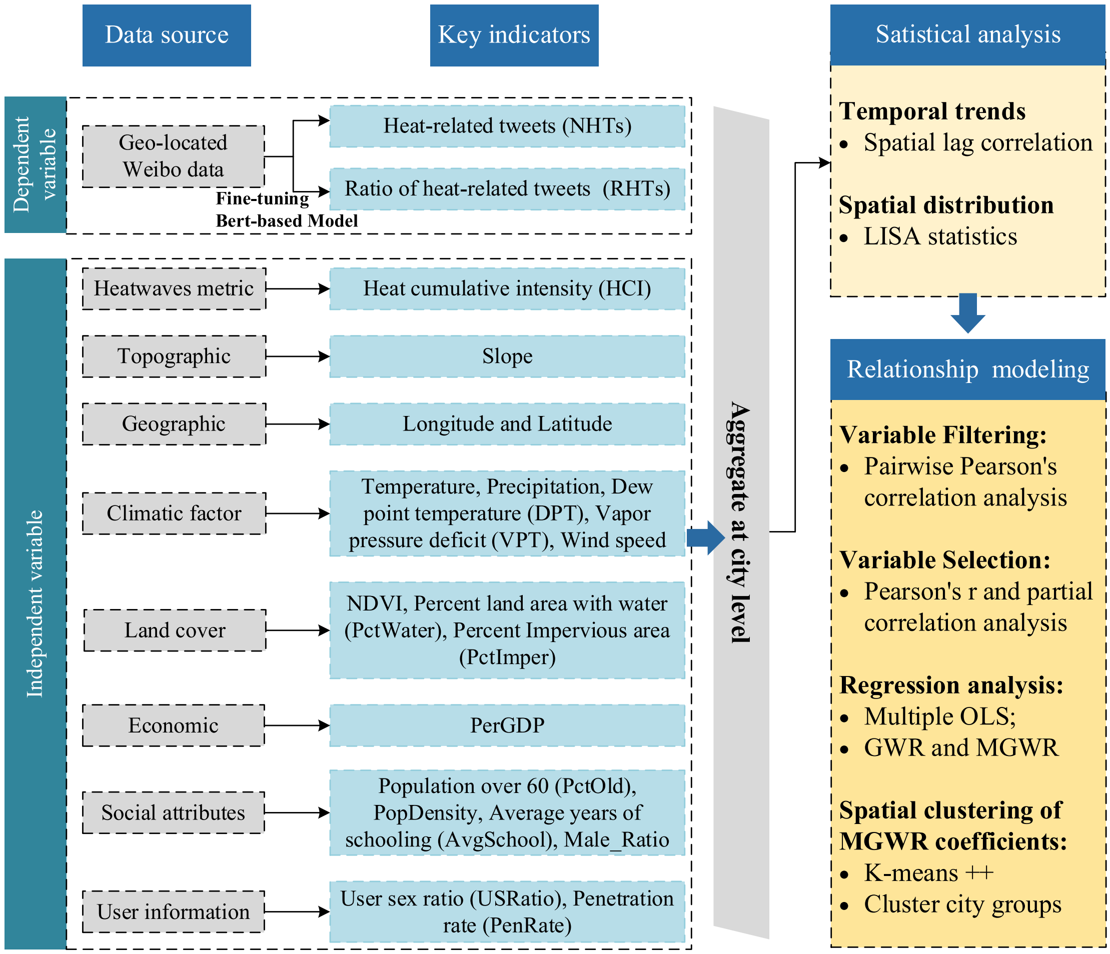


Fig. 1. Workflow of this study.

tweets related to heatwaves. Thirdly, we proposed and calculated indicators of public attention to heatwaves, such as the cumulative intensity of heat. The last step included analyzing the temporal-spatial distribution patterns of social activities and modeling the correlation between heatwaves and public attention (Fig. 1).

3.1. Study area and data

3.1.1. Study area

In the past seventy years, the average annual surface temperature in China has risen significantly, at a rate of 0.26 °C per decade. The warming rate of China is greater than the global average for the same period, which is an area of concern for global climate change. During the summer of 2022, the eastern region of China experienced the most severe heatwaves since 1961, lasting 79 days and encompassing an area of over 500 km^2. Using the entire nation as a research object, it is possible to effectively examine the influence mechanisms of various physical environments and social-economic factors on heat perception during heatwaves and provide regional heat-risk references.

3.1.2. Weibo data

Geolocated Weibo tweets within China from June to September 2022 were collected through the Weibo API, with 10 million tweets created by over 3 million unique users. To obtain tweets related to the heatwaves, keywords filtering method were adopted, including: hot (热), high temperature (高温), heatstroke (中暑), sunburn (晒), sunbaked (烤), sultry (闷).

Based on verification type, we categorized the Weibo users into two groups: ordinary users and media users. Specifically, the types of media users who verify include government, enterprise, media verification, and celebrities. During the registration process, a media account user can also choose a gender. Due to the lack of a reliable way to distinguish between ordinary users and media users on Weibo, we chose not to exclude a small proportion of media users accounts.

By conducting data collection, processing, and analysis, it is possible to obtain a comprehensive user profile that includes information like gender, age, geographic location, number of fans and followers, total number of Weibo posts, statistics on retweets, comments, as well as the heatwave perception expressed in text. Within the 3,647,842 Weibo users, the percentage of male and female users is 29.79 % male to 70.21 % female. In terms of geography and gender distribution, Weibo users (particularly females) in their 20 s are obviously more enthusiastic about discussion. As the heatwaves took place in China, the local users were enthusiastic about the discussion.

The demographic data came from Weibo user profiles at the individual level, including unique IDs and genders. Based on this information, two user information indicators were constructed, including penetration rate (number of Weibo users divided by the population of each city) and user sex ratio (number of male Weibo users divided by the total number of Weibo users).

3.1.3. Meteorological and environmental data

A series of explanatory variables as inspired from previous studies were collected to understand public response in extreme heat (Wu et al., 2024; Hass et al., 2021). The variables could be generally categorized into environmental, meteorological, and social (Table S1).

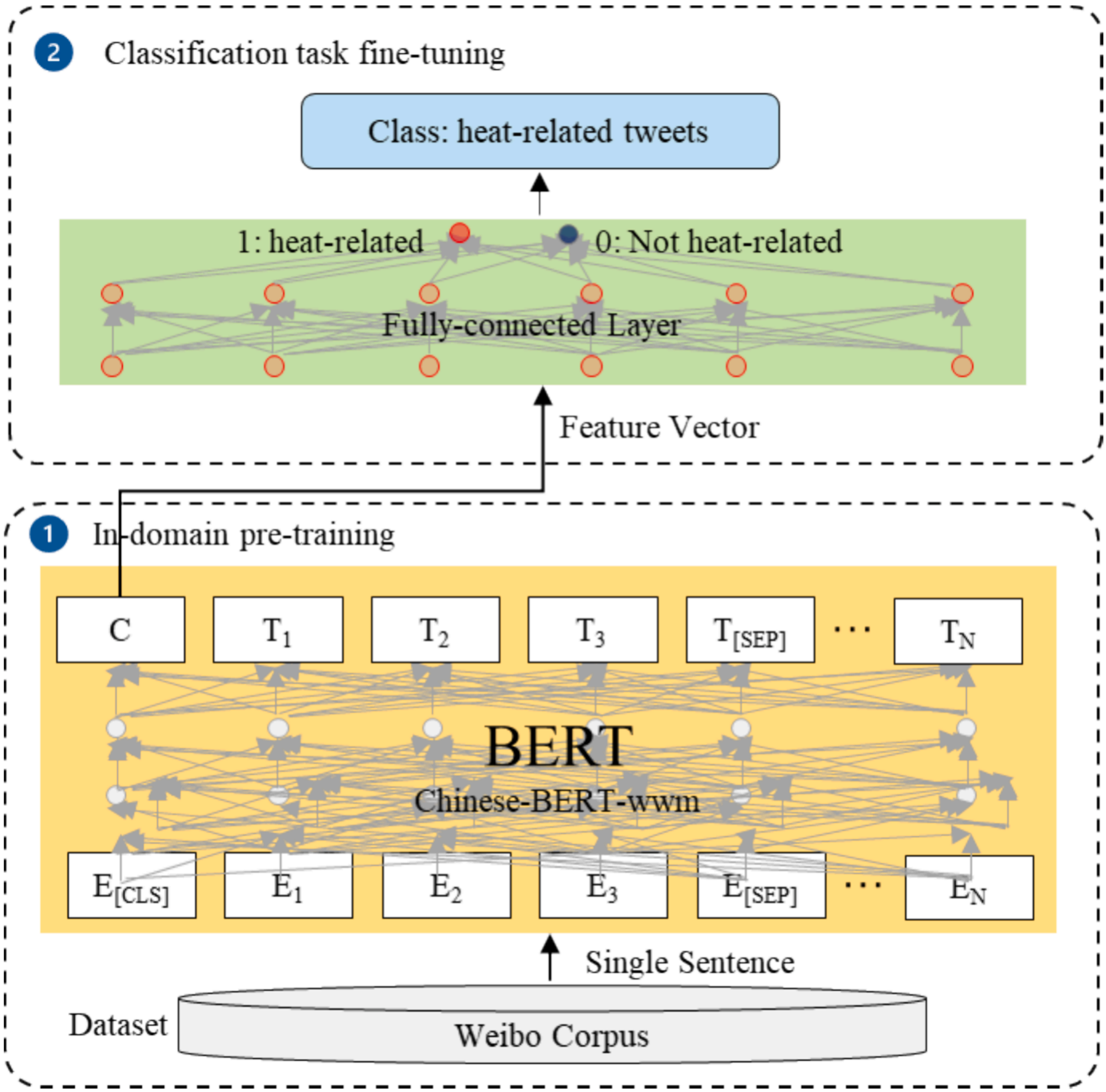


Fig. 2. Architecture of the BERT model for heat-related text classification.

Meteorological and environmental data were taken from different open data platforms. Climate-related data, including hourly air temperature, precipitation, wind speed, dew point temperature, and sea level pressure, were collected from the National Meteorological Science Data Center (https://data.cma.cn). Dew point temperature significantly influences the relative perception of heatwaves by impacting our body's ability to cool down through sweat evaporation (Cecinati et al., 2019). Higher dew points, indicating higher humidity, lead to a feeling of greater heat and discomfort, even at the same air temperature (Cvijanovic et al., 2023). Topographic data, including elevation and slope, were provided by Resources and Environmental Sciences (https://www.resdc.cn). Land cover data include the monthly global normalized difference vegetation index (NDVI) datasets with a spatial resolution of 1 km * 1 km, collected from MOD13A3 and downloaded from the EOSDIS website. Water area and impervious surface data with 30 m fine-scale spatial resolution were collected based on the China land cover dataset (CLCD), which was the first Landsat-derived annual product for China from 1990 to 2019 with high-resolution land cover maps (Yang and Huang, 2021).

In terms of demographic and socio-economic factors, nighttime light data was obtained by integrating DMSP-OLS and SNPP-VIIRS (Wu et al., 2021). Other social-economic variables, such as gross domestic product (GDP) per capita, percentage of population over 60 years old, and people with a college degree or higher, were collected from the 2019 China City Statistical Yearbook at the city level.

Table 1
Performance of prediction on the verification dataset.

Category	Number	Precision	Recall	F1-score
Heat-relevant	712	0.99	0.96	0.97
Not relevant	688	0.96	0.99	0.97

3.2. Methods

3.2.1. Classify Weibo tweets using BERT

The extraction of heat-related tweets is a typical text classification task in natural language processing, in which a given document is assigned to predefined categories according to its content. BERT is a pre-trained model that acquires extensive linguistic knowledge and contextual relationships through pre-training on large-scale text corpora. Previous studies have used the BERT model to screen social media data with high precision (Wang et al., 2021; Rui, 2023). However, training a model from scratch is time-consuming and computationally expensive. By employing fine-tuning techniques, BERT can be better adapted to specific downstream tasks such as sentiment analysis, question answering systems, and text classification, thereby enhancing the performance of these tasks.

In this study, two-stage fine-tuning techniques were designed to implement heat-related text classification: (1) in-domain pre-training: fine-tuning Chinese-BERT-wwm with Weibo corpus to extract Weibo text contextual embedding; (2) classification task fine-tuning: further fine-tuning pre-trained Chinese-BERT-wwm on heat-related tweet classification tasks. Fig. 2 is a high-level description of the BERT-model. Chinese BERT with whole word masking (Chinese-BERT-wwm) model, which is more effective for Chinese text embedding, was employed to fine-tuning solution. In this study, the input comprises a single tweet extracted from corpus. Initially, the tweet is tokenized into a sequence that includes a special CLS token, which represents the overall meaning of the sentence. Subsequently, these tokens are processed through the BERT model. The output from the model consists of vectors corresponding to each token, where the vector associated with the CLS token encapsulates the semantic representation of the entire sentence. The first stage of fine-tuning is to obtain an accurate representation of the CLS vector. In the subsequent phase, the BERT model training parameters of the first stage are fixed. The CLS vectors are then fed into a fully connected layer. This network model is trained through supervised learning, utilizing labels to classify texts related to heat. Four indicators were employed to evaluate the model's performance: accuracy score, recall, precision, and F1 score.

3.2.2. Public attention metrics

The study constructed two metrics: the number of heat-related tweets (NHTs) and the ratio of heat-related tweets (RHTs) to analyze the public attention to heatwaves. The NHTs reflected the social consequences of heatwaves and their extent. Heatwaves that occur in major cities were more noticeable than those that occur in desert regions. Similarly, more people would be impacted by heatwaves in densely populated metropolitan areas than in sparsely populated areas. The RHTs were defined as the number of tweets divided by the total number of tweets posted within the same period. The ratio of heat-related tweets had been proven to be an effective indicator that reflects the public attention to heat and was highly related to the intensity of heatwaves.

$$RHTs = \frac{\#\ heat - related\ tweets}{\#\ all\ tweets} \quad (1)$$

3.2.3. Heat-related metrics

There is no universal definition of a heatwave due to the variety of climatic conditions and socio-demographic characteristics across the world. Typically, researchers classify heatwaves based on their daily maximal temperature, duration, and spatial extent (Wu et al., 2023). For this study, we adopted the definition of a heatwave provided by the China Meteorological Administration, which is defined as at least three consecutive days with a maximum daily temperature exceeding 35 °C (Zhao et al., 2023).

To make easier comparisons among different cities, this study used the heat cumulative intensity as a useful metric that integrates the heatwave frequency and intensity into a single indicator. According to Perkins-Kirkpatrick and Lewis (2020), the formula for the heat

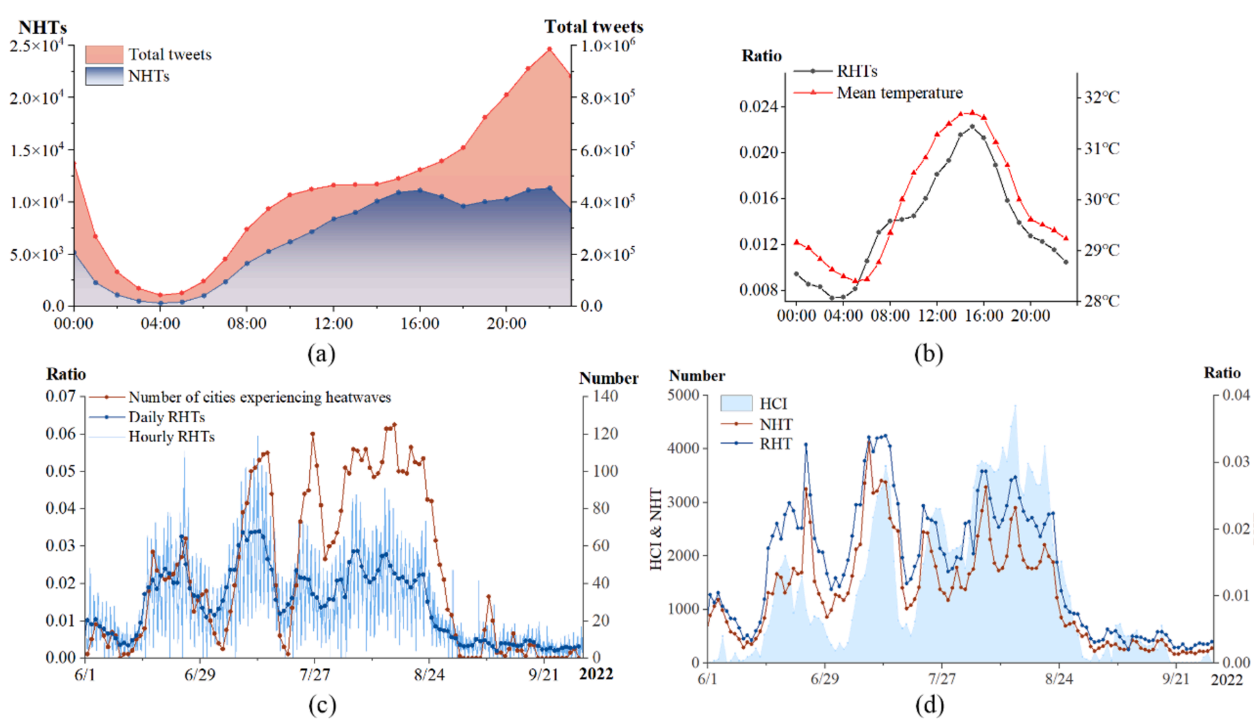


Fig. 3. Temporal trend of urban heat perception. (a) Temporal trend of total tweets and NHTs; (b) RHTs and mean temperatures observed at monitoring stations by hour; (c) Global trend of RHTs and the number of cities experiencing heatwaves; (d) Temporal correlation of RHT, NHT and HCI.

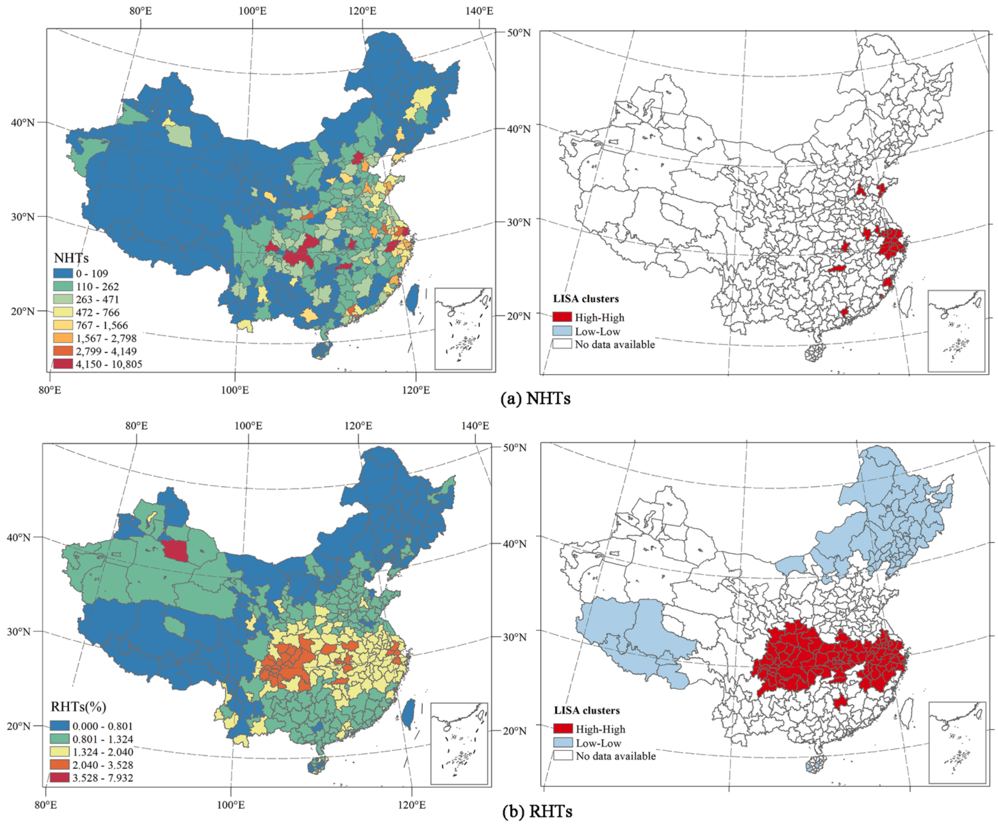


Fig. 4. Spatial distribution of heatwave perception and LISA clusters at city-level.

cumulative intensity can be expressed as follows:

$$Heat\ cumulative\ intensity = \sum_{1}^{d} (T_{max} - T_{35}) \bullet H_d \tag{2}$$

T_{max} the maximum daily temperature and T_{35} is the temperature threshold for heatwaves in this study. H_d is an indicator variable ($H_d = 1$ if heatwave happened, else $H_d = 0$.)

3.2.4. *Modeling geographic schema of heatwave perception*

3.2.4.1. Statistical analysis of RHTs and NHTs. The statistical modeling process mainly includes two dimensions: the temporal dimension and the spatial dimension. Regarding the temporal dimension, we first investigated the correlation between the ratio of heat-related tweets (RHTs) and the number of cities experiencing heatwaves throughout the study period. Then, the lag correlation between the RTHs and NTHs, as well as the number of cities enduring heatwaves, were analyzed. Simultaneously, daily RHTs and NHTs were conducted to obtain an overall picture of the temporal trend for heat-related Weibo tweets. In the spatial dimension, we utilized the Moran's I statistic to explore the global spatial autocorrelation of heat-related social media activities across China at city-level with a simple first-order queen-weighted matrix. The LISA statistics method was conducted to identify the hot spots and cold spots of heat-related activities and comprehend the spatial association of two heat-related indicators comprehensively.

3.2.4.2. MGWR model for RHTs. **Variable Selection:** There could be significant correlations between variables. Based on the correlation coefficient, variables had been filtered. To reduce the multicollinearity problem, a Pearson's correlation analysis was conducted to evaluate the correlations between pairwise indicators, including metrological and environmental variables as well as demographic and socio-economic dimensions with eight factors and 24 variables, as a guide for selecting variables for further analysis. Pearson's r and partial correlation coefficient were used to quantify the relationships between the underlying factors and the human response variables. The variables retained need to satisfy that both absolute values of the above two correlation coefficients are greater than 0.10 at the 0.05 level of significance.

Regression analysis: Based on the reserved variable, a set of hierarchical linear regression models with ordinary least squares (OLS) were constructed to understand the associations between the underlying factors and heat perception. The Global Moran's I residual test was taken to examine the assumptions about the independence test of explanatory variables. Since the global spatial autocorrelation exists as the result of the RHTs, we further constructed the geographically weighted regression to evaluate the spatial non-stationarity of estimates and explored the local variations in the aspects of strength and direction among variables. Furthermore, the MGWR model was used to investigate probable variables impacting heat perception since impact factors may play diverse roles throughout a geographical range (Mohammadi et al., 2023). The MGWR model was proposed by Fotheringham et al. (2017), whose goal is to address the difficulty of dealing with multi-scale spatial phenomena with conventional GWR techniques. MGWR offers a multi-scale framework whereby each explanatory variable is assigned its own geographic scale.

Spatial clustering of coefficients: Clustering techniques were commonly used to identify patterns in datasets. In this study, the K-Means++ algorithm was employed to divide the cities into k distinct clusters based on the similarity of the MGWR regression coefficients. Clustering aims to minimize the distance between sample points within the same cluster while maximizing the distance between clusters. The silhouette score was used to evaluate the clustering outcomes for the selected k values. When the shear score is close to 1, it indicates that the data instance is close to the center of its cluster, while when it is close to 0, it indicates that the sample is an outlier, and when it is -1, it indicates

Table 2
Correlations analysis between predictors and NHTs and RHTs.

Category	Variable	RHTs		NHTs	
		Correlation	Partial correlation	Correlation	Partial correlation
Heatwave	HCI	***0.761****	***0.635****	***0.228****	***0.258****
Topographic	Slope	0.051	−0.017	−0.058	0.093
Geographic	Lon	−0.092	−0.332	0.096	−0.032
	Lat	−0.291	0.113*	−0.106	0.043
Climatic factors	MeanMaxTemp	0.506	−0.074	0.183*	−0.103
	Precipitation	***−0.117****	***−0.139****	0.021	0.026
	DPT	***0.303****	***0.155****	0.197*	0.105
	VPD	−0.027	0.012	−0.030	−0.164*
	Wspeed	−0.066	0.029	0.046	−0.026
Land cover	NDVI	0.060	0.059	0.004	0.024
	PctWater	0.165*	0.083	0.264	0.040
	PctImper	−0.012	−0.019	0.315	0.094
Economic	PerGDP	0.179*	0.072	***0.519****	***0.135****
Social attributes	PctOld	0.093	0.121*	−0.059	−0.169*
	AvgSchool	0.014	−0.027	***0.455****	***0.278****
	PopDensity	***0.131****	***−0.161****	***0.451****	***−0.134****
	Male_Ratio	0.083	0.015	0.092	−0.231*
User information	USRatio	−0.089	0.097	−0.147	0.078
	PenRate	0.002	0.079	***0.486****	***0.180****

HCI: Heat cumulative intensity; MeanMaxTemp: Mean max temperature.
DPT: Dew point temperature; VPD: Vapor pressure deficit; Wspeed: Wind speed.
PctWater: Percent land area with water; PctImper: Percent Impervious area.
PctOld: Population over 60; AvgSchool: Average years of schooling.
USRatio: User sex ratio; PenRate: Penetration rate.
Note: *p-value < 0.05.

Table 3
OLS regression for explaining the NHTs with HCI.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	260.366	−567.234	−4752.571	133.853	−129.951	−2555.02
HCI	4.177					0.226
PerGDP		0.519				0.069
AvgSchool			9.119			0.182
PopDensity				0.451		0.241
PenRate					0.486	0.315
Adjust R^2	4.9 %	26.8 %	20.4 %	20.1 %	23.4 %	**42.3 %**

Table 4
OLS regression for explaining the RHTs with HCI.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.830	1.426	0.618	1.201	0.735
HCI	0.761				0.711
Precipitation		−0.117			−0.093
DPT			0.303		0.106
PopDensity				0.131	0.071
Adjust R^2	57.80 %	11 %	8.90 %	1.40 %	**58.80 %**

that the observation has been incorrectly assigned to a cluster. Spatial data analysis and modeling were performed by GWR4.0, ArcGIS 10.2, SPSS, and Python 3.

4. Results

4.1. Prediction accuracy and verification of heat-related tweets

We fine-tuned the Chinese-BERT-wwm model and ran a total of 10 epochs. Of the 7,000 labeled samples, 80 % were used as training samples and test samples, and 20 % were used as verification sets (1,400 samples). Table 1 displays the accuracy and efficacy of the model on verification sets. The prediction accuracy on the verification set has been raised by 2 % via the process of fine-tuning the Weibo corpus. Moreover, compared to linear regression, using a convolutional layer as a classifier can simultaneously achieve more balanced classification performance.

4.2. Spatiotemporal variation of public attention to heatwaves

4.2.1. Temporal patterns of urban heat perception

This study used previously classified datasets with about 157,892 heat-related Weibo tweets to explore temporal patterns of urban heat perception. Fig. 3 provides a comprehensive depiction of the statistical

Table 5
Statistics of GWR and MGWR models for RHTs.

Variable	GWR					MGWR				
	BW	Mean	Min	Max	STD	BW	Mean	Min	Max	STD
Intercept	65	−0.043	−0.877	1.04	0.491	46	0.015	−0.66	0.868	0.406
HCI	65	0.383	−0.118	1.351	0.360	86	0.332	0.038	1.299	0.301
Precipitation	65	−0.152	−0.798	0.544	0.275	256	−0.156	−0.218	−0.104	0.033
DPT	65	0.293	−1.141	1.569	0.344	319	0.32	0.313	0.332	0.004
PopDensity	65	0.129	−0.411	0.654	0.191	319	0.052	0.045	0.066	0.005
Performance	$R^2 = 0.874$, Adj. $R^2 = 0.852$ AIC = 344.503, AICc = 362.523 BIC = 529.013					$R^2 = 0.855$, Adj. $R^2 = 0.841$ AIC = 348.598, AICc = 354.568 BIC = 457.887				

trends in social media usage during heatwave periods. Examining the hourly social media activity throughout the day, the overall tweet activity and NHTs reach their nadir at 4:00. Notably, NHTs reach their peak activity around 16:00, and their other peak is approximately 22:00 (Fig. 3a). This phenomenon was potentially attributed to the heightened influence of high-temperature heatwaves on nocturnal human activities, especially disturbing individual sleep, manifested in sleep deprivation and sleep exploitation. Regarding RHTs, the highest frequency occurs at 16:00, aligning with the trend of average temperature (Fig. 3b). The mean temperature is the average temperature of 2419 stations from June to September 2022 at 2419 stations from the National Climate Center, China Meteorological Administration for 24 h. Compared to the indicator of heat-related sentiment proposed in Murakami et al. (2016), hourly mean RHTs in this study showed a more consistent association with temperature trends. For daily activity from June 1 to September 31, the number of cities experiencing heatwaves shows a correlated fluctuation with heat-related tweet activity, with a correlation coefficient exceeding 0.7 (Fig. 3c). Additionally, both NHTs and RHTs demonstrate a high correlation with the HCI (Fig. 3d). The correlation coefficients were 0.78 for HCI and RHT, and 0.75 for HCI and NHT. RHTs correlated more strongly with heat related indexes (e.g., HCI and the number of cities experiencing heatwaves) than NHTs.

4.2.2. Spatial distribution and clustering of urban heat perception

Two metrics (NHTs and RHTs) of human heatwave perception across China were presented in Fig. 4. Both city-level NHTs and RHTs exhibited significant positive global spatial autocorrelation (Moran's I = 0.03, z-score = 4.73, p-value < 0.001 for NHTs, and I = 0.22, z-score = 31.60, p-value < 0.001 for RHTs). Through LISA clusters, the High-High LISA clusters were observed for both NHTs and RHTs, as shown in Fig. 4. High NHTs were mainly distributed in densely populated areas in the eastern and southern regions of China, and less in the western regions. The average value of NHTs was 422. High-High clusters for NHTs were mainly distributed in municipalities and provincial capitals in the east and south regions (e.g., Beijing, Shanghai). High-High clusters for NHTs appeared in the provincial capitals and municipalities. RHTs displayed a significant spatial autocorrelation with an average value of 1.22 %. Low-Low clusters for HWTs were identified in the north and southwest of China. High-High for HTWs were primarily found in the Chengdu-Chongqing regions and the middle and lower reaches of the Yangtze River, including the Wuhan and Suzhou regions.

4.3. Associations between public attention and heatwaves

4.3.1. Heat-related variables from OLS model

Due to the availability of GDP data for only 321 cities from the year 2019, subsequent analyses in this study are confined to these cities (Fig. S1). We conducted a correlation analysis and ordinary least squares multiple linear regression models with corresponding potential explanatory variables to explain NHT and RHT (Table 2). For NHTs, five variables—HCI, PerGDP, average years of schooling, population density, and penetration rate—are significant at the 0.05 level, with both the correlation coefficients and partial correlation coefficients exceeding 0.1. Among these, HCI accounts for the smallest proportion of variation in NHTs, explaining only 4.9 % of the variance (R^2 = 4.9 %, Table 3). HCI, precipitation, dew point temperature, and population density, were significantly correlated to RHTs (Table 4). The most important factor was HCI, which accounts for 57.8 % of the variation of RHTs. The supplementary materials provided the correlation between cumulative heat wave intensity and NHT at different threshold temperatures to increase the robustness of the results (Table S2).

4.3.2. Spatial heterogeneity of heat-related tweets

The spatial autocorrelations of OLS residuals of RHTs and NHTs were examined, and the results indicated that RHT residuals are not independent and can be fully explained by variables (Moran's I < 0.001, z-score = 0.224, p-value = 0.822 for NHTs, and Moran's I = 0.22, z-score = 14.370, p-value < 0.001 for RHTs). Therefore, the GWR and MGWR models were conducted for RHTs to eliminate modeling errors induced by the spatial autocorrelation of variables.

In Table 5, we compared the GWR and MGWR models' bandwidth (BW) statistics and regression coefficients. Cross-validation and the Gaussian kernel were used to optimize bandwidth selection for the GWR and MGWR models. The GWR model allocated a constant bandwidth (BW = 65, or 20 % of the total sample) to each independent variable. As shown in Fig. 5, the Moran's I of the standard residuals of the GWR and MGWR models were −0.011 (p-value = 0.989) and −0.015 (p-value = 0.437), indicating the successful eradication of spatial randomness and spatial autocorrelation of the residuals. The GWR and MGWR models explained more variance than the OLS model, and their respective adjusted R^2 values were 0.874 % and 0.855 %. Fig. 5 shows that the spatial variation of local prefecture-level cities R^2 generally decreased from west to east.

The MGWR model, in contrast, assigns a distinct bandwidth to each variable. Clearly, the standard deviations of GWR coefficients were greater than those of MGWR coefficients, and the minimum and maximum values of GWR coefficients frequently have opposite signatures. In the GWR model, for instance, the statistical intensity range for the cumulative heat wave was between −0.118 and 1.351, with a standard deviation of 0.360. The statistical value of the coefficient in MGWR varied between 0.038 and 1.299. Therefore, the proportion of individuals posting microblogs about heat was proportional to the intensity of heatwaves. Consequently, the negative coefficient of GWR may run counter to prior expectations. Furthermore, due to overfitting, GWR's narrow bandwidth leads to high coefficient variability and extreme values in its coefficients. This made the interpretation of GWR's results uncertain. In summary, the MGWR model used in this study can effectively prevent symbol inversion and produce more accurate estimation results.

Fig. 6 illustrated the spatial visualization of GWR and MGWR coefficients and significance (i.e., |t-values| greater than 1.645). The coefficient of MGWR was statistically more significant than that of GWR.

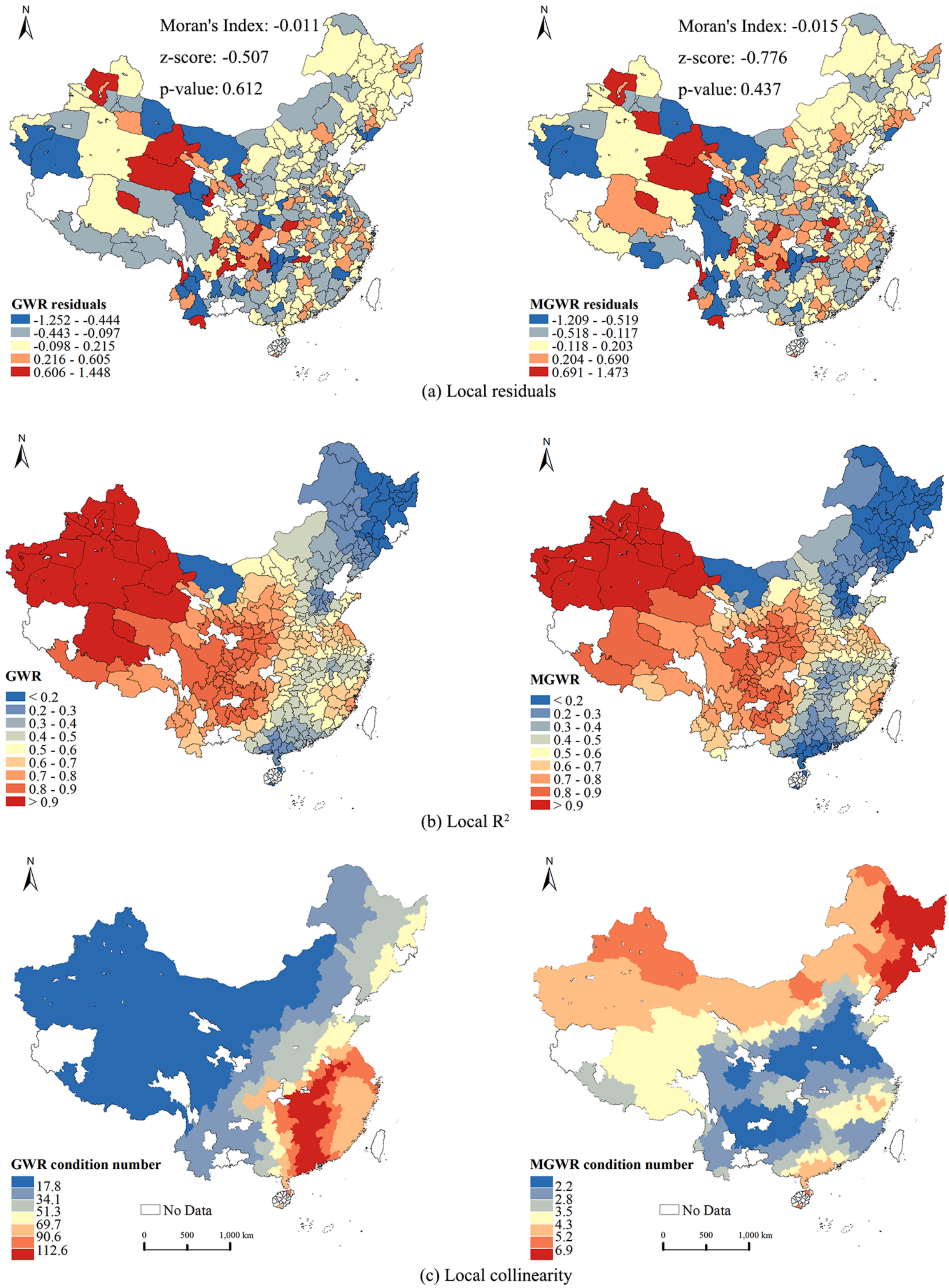
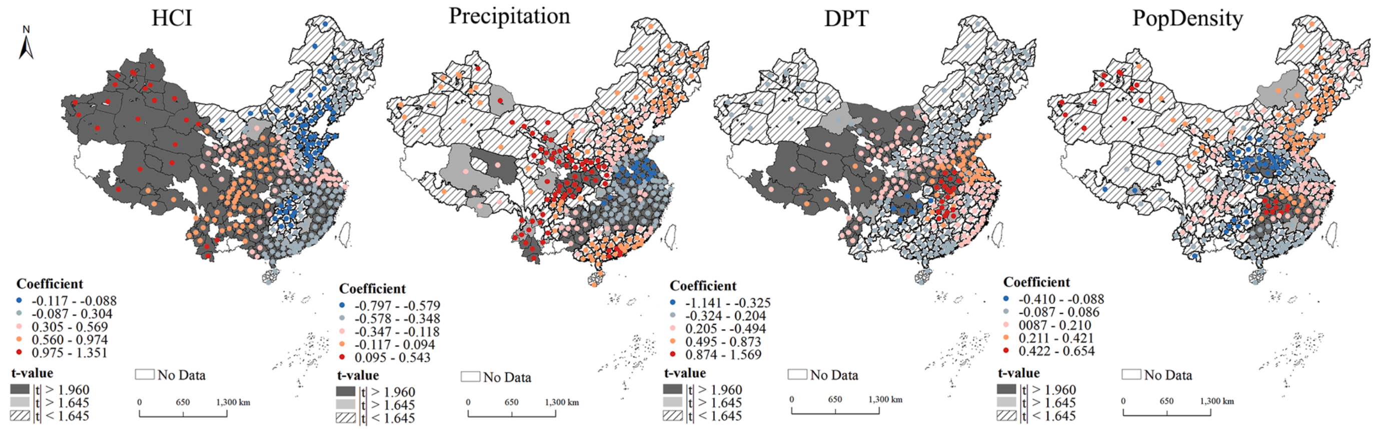
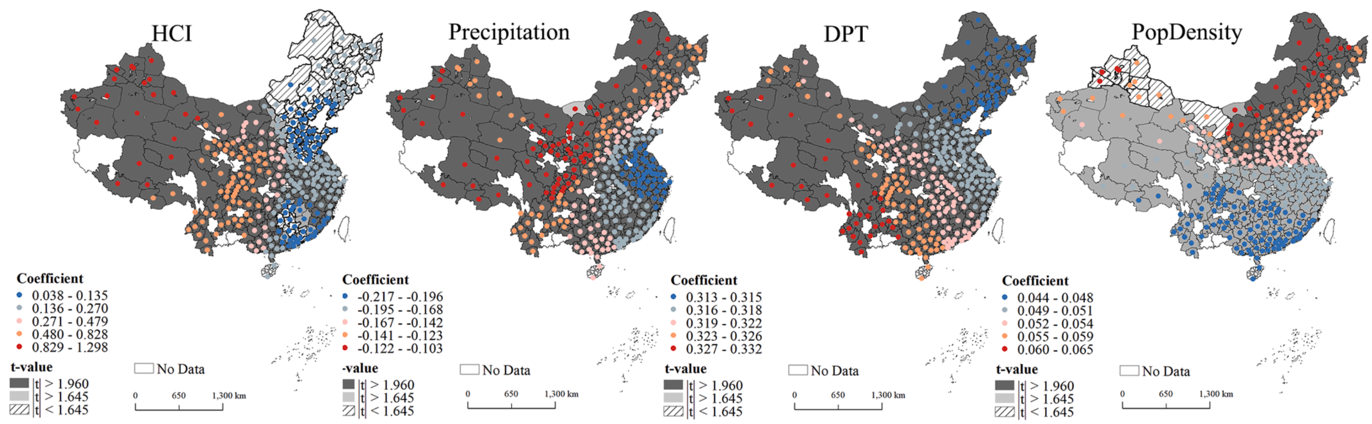


Fig. 5. The comparison of effectiveness tests for GWR (left) and the MGWR (right) models.



(a) GWR Coefficients



(b) MGWR Coefficients

Fig. 6. Spatial distributions of (a) GWR and (b) MGWR coefficients.

The MGWR coefficients displayed across the maps for HCI, Precipitation, DPT, and Population Density reveal distinct spatial patterns that suggest region-specific interactions between these variables and public heat response. Higher HCI coefficients are concentrated in central and western areas, especially Xinjiang and Tibet. Precipitation coefficients are generally negative, with less negative values in the southeast. DPT shows a uniformly positive association, particularly pronounced in the southwest. The population density coefficient is the highest in the north, which indicates that population factors play an important role in heat perception. These patterns underscore the regional variability and the importance of geographic-specific analyses in understanding the impacts of environmental and demographic factors.

4.3.3. Cluster city groups based on public attention to heatwaves

The cities in the study area were clustered into six groups using the K-means++ method based on four factors of the MGWR model. The cluster city groups were consistent with the seven geographic subregions of mainland China, which are North China, Northwest China, Southwest China, Central China, South China, Northeast China, and Eastern China, and subdivided according to climate and socioeconomic development (Fig. 7a and b). Cluster city groups 1, 2, 3, and 5 were located in North China, Northwest China, Northeast China, and Eastern China. It is worth noting that cluster city group 4 combines Central China and South China into one cluster.

Based on the average coefficient of grouping, the primary determinant that affects RHTs was the HCI metric. The HCI affected China the least in the east and the most in the west, which includes South China and Northwest China. The second reason was rainfall. Rainfall affects Eastern China the most, while it affected Northwest China and Southwest China the least. Overall, the cluster city groups based on the level of public attention towards heatwaves revealed notable regional disparities in inhabitants' sensitivity to heat. Moreover, there were varying degrees of variance in sensitivity observed throughout the seven geographic subregions of mainland China. These associations were probably due to regional disparities in the rate of climate variable fluctuations, such as precipitation and temperature. It was worth noting that ongoing warming trends have further impacted the contemporary climate system in mainland China.

5. Discussions

5.1. The innovation of heatwave perception from Weibo

From a new perspective, this study introduced the theoretical framework of heatwave perception using geolocated Weibo data. Each user of Weibo was considered a data source, and a large language model was utilized to analyze heat-related Weibo tweets. This study fine-tuned the Chinese-BERT-wwm model on a large microblog corpus and trained annotated text using a convolutional layer classifier. This data processing process enhances classification accuracy and efficacy significantly. Subsequently, the impact of heatwaves on residents can be fully revealed at a more comprehensive level and at a finer temporal resolution (daily scale). Based on high-relevance tweets, we revealed the associations between public attention and heatwaves. Individual perception lag may explain the one-day spatial latency between social activity and cities experiencing heatwaves, which may be due to the perception lag of individuals. The objective was to investigate the spatiotemporal variations of environment, demographic, and economic-social factors

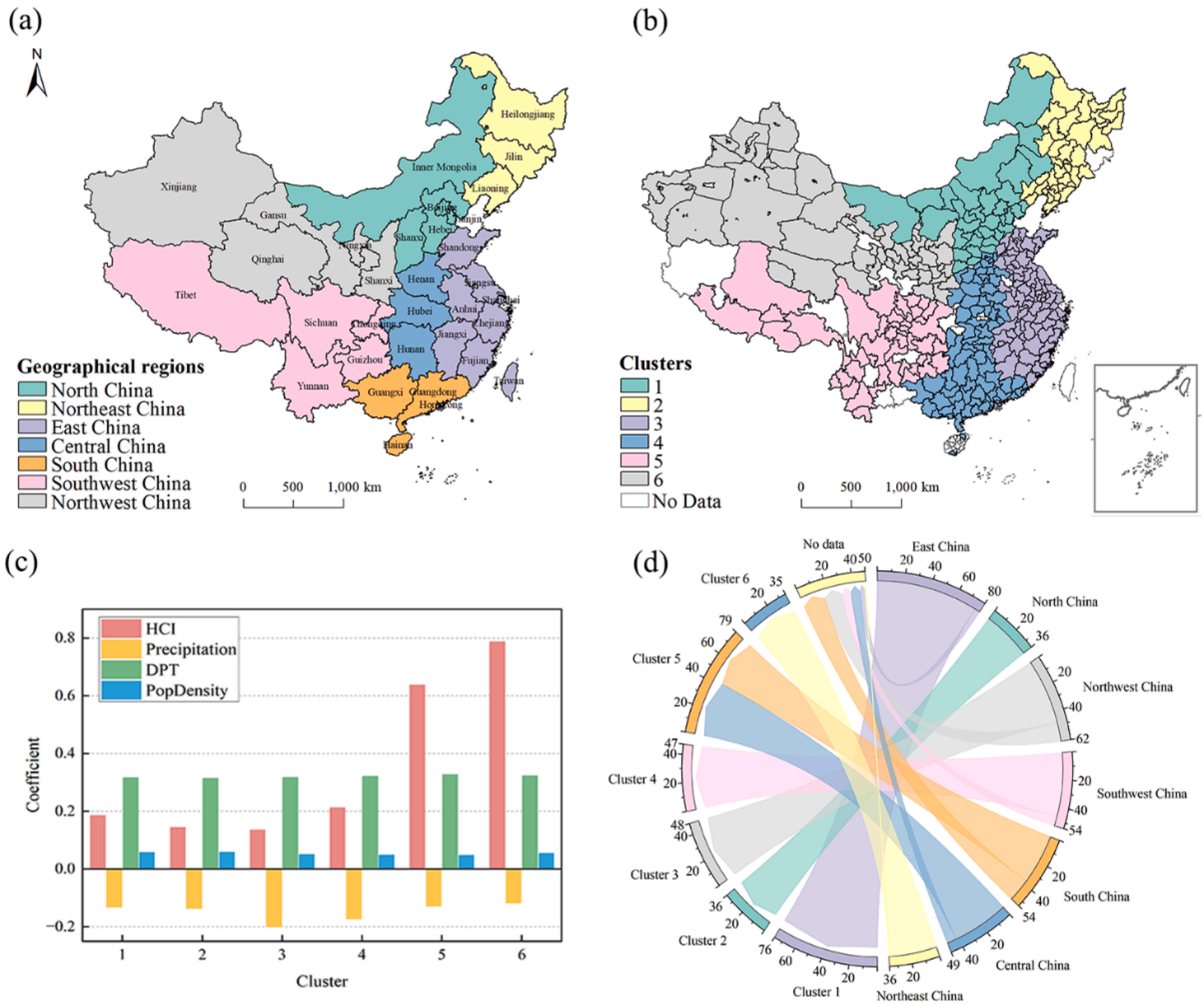


Fig. 7. (a) Seven subregions of China, (b) spatial distribution of MGWR coefficient clusters, (c) mean value of MGWR coefficient clusters, and (d) chord plots described the correspondence between geographic regions and MGWR coefficient clusters.

leveraging public attention derived from Weibo to heatwaves in geographic subregions across China.

As predicted, the results of cluster city groups showed that heat cumulative intensity (HCI) was associated with social activities. The HCI could explain the 57.8 % variation for RHTs but 4.9 % for NTHs. Specifically, NTHs were primarily influenced by socioeconomic and demographic attributes such as per capita GDP, population density, and penetration. As for RHTs, by contrast, it depended on environmental factors (i.e., HCI, precipitation, dew point temperature) and population density.

5.2. Geographic and demographic disparities of heatwave perception

Remote areas, such as midwest cities in China, require further investigation because their perception of heatwaves and behavioral responses may differ from those in big cities. Geographic disparities among remote areas and big cities can be better understood by incorporating place-based factors (e.g., check-in behaviors) and identifying underlying disparities (e.g., GDPs). By examining the local heat risk perception and behavior, this study can gain a deeper understanding of the situation and develop appropriate adaptive responses. For example, certain cities on China's mainland have higher land surface temperatures in regions with a higher concentration of low-income and minority individuals. Individuals from remote areas may have had a lower income, necessitating their engagement in income-generating activities during periods of extreme heatwaves. As a result, these vulnerable communities in remote areas require additional support and tailored adaptation plans. To accurately understand the impact of local climate and place-based factors (e.g., urban, rural) on geographic disparities in heatwave exposure and perception, it is necessary to have large sample sizes that encompass a wide variety of individuals and geographic regions.

As for demographic disparities, we stratified tweets into two different groups based on gender of poster (male and female) to evaluate the gender discrepancy in heatwave perception. First, we grouped tweets by gender, and then counted heat-related tweets from male posters/total tweets (MRHTs) and heat-related tweets from female posters/total tweets (FRHTs). Then the variables in Table 2 were employed to perform stepwise regression with MRHTs and FRHTs respectively. The regression results obtained are shown in Table S3. This study also emphasizes the necessity for further investigation of vulnerable populations (senior citizens), including those who are socially isolated, as well as specific demographics such as elders and decision-makers. Addressing these spatial and demographic gaps, as well as enacting tailored measures, may help China protect all citizens during heatwaves.

5.3. Comparative assessment

Previous studies have recognized heatwaves as a major health risk, but there has been less research on the public's perception of heatwaves using location-based social media data. Our findings indicated that heatwave perceptions can fundamentally compel economic and social actions in China to address the risks from global or regional warming. This finding was comparable to the results of the Gosling et al. (2009) study, which reviewed heatwaves during the period 2000–2007 and found that China was one of the most severely impacted regions. Our observations also indicate that female-young individuals, and individuals residing in remote regions, had a vulnerable perception of heatwaves. These observations were consistent with certain findings from previous investigations (Pappalardo et al., 2023; Liu et al., 2013; Semenza et al., 2008). Therefore, more research is required to enhance heatwave perception in various populations. This study has the potential to provide valuable insights for enhancing a population's adaptive capacity during heatwaves through location-based social media data.

5.4. Limitations

The geospatial approach used in this study may obscure the city-level associations between public attention and heatwaves when examined at a more detailed spatial–temporal resolution. Moreover, it is important to note that this study made efforts to mitigate sample selection bias. We also excluded 45 cities that lacked GDP data for the year 2019, leaving a final sample size of 321 cities for subsequent analysis. The present analysis focused on excluding cities that are predominantly characterized by economic underdevelopment, as they tend to have a smaller number of urban areas and a lower overall count of social media users, total tweets, NHTs, and RHTs. Furthermore, the PerGDP data utilized in this study was sourced from the 2019 China City Statistical Yearbook. Conversely, the demographic and socio-economic data, including PctOld and PopDensity, were obtained from the seventh National Population Census. Meanwhile, the data pertaining to social media users was gathered in 2022. A future study has the potential to improve these calculations by incorporating the latest GDP data or data that is consistent on a yearly basis.

6. Conclusions

This study provides a new perspective for understanding the associations between public attention and heatwaves in 321 Chinese cities. We use the MGWR model to quantify the impact of physical and socioeconomic factors on heatwave awareness, combining geolocated Weibo with environmental data. The results of the regression analysis have revealed that heat-sensitive cities were grouped based on the coefficients of MGWR. Our findings highlight differences in the direction and magnitude of the spatial effects of environmental, demographic, and socioeconomic factors on public thermal perception.

Given the disparity in public attention towards heatwaves using social media data, future research on heatwave perception should include an examination of remote regions considering geographic and demographic disparities. This is particularly important given previous studies on thermal comfort exposure, which have shown the existence of heat exposure disparities both within and between cities. The MGWR model's coefficients clearly align the relationship between city groups and human consciousness with geographical divisions. Future investigations should also examine whether changes in municipal city group patterns and the distribution of public attention through time are associated with heatwaves.

CRediT authorship contribution statement

Mingxuan Dou: Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Yandong Wang:** Resources, Supervision. **Mengling Qiao:** Data curation, Software. **Dongyang Wang:** Data curation, Formal analysis, Methodology. **Jianya Gong:** Funding acquisition, Resources, Software, Supervision. **Yanyan Gu:** Methodology, Project administration, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 42101464), the China Postdoctoral Science Foundation (Grant No. BX20220237), and Open Research Fund Program of LIESMARS (Grant No. 21I03).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2024.104205.

References

Baylis, P., Obradovich, N., Kryvasheyeu, Y., Chen, H., Coviello, L., Moro, E., Fowler, J.H., 2018. Weather impacts expressed sentiment. PLoS One 13 (4), e0195750. https://doi.org/10.1371/journal.pone.0195750.
Bogdanovich, E., Guenther, L., Reichstein, M., Frank, D., Ruhrmann, G., Brenning, A., Denissen, J.M.C., Orth, R., 2023. Societal attention to heat waves can indicate public health impacts. Weather Clim. Soc. 15 (3), 557–569. https://doi.org/10.1175/WCAS-D-22-0147.1.
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 (5), 762. https://doi.org/10.3390/ijerph16050762.
Christidis, N., McCarthy, M., Stott, P.A., 2020. The increasing likelihood of temperatures above 30 to 40 C in the United Kingdom. Nat. Commun. 11 (1), 3093. https://doi.org/10.1038/s41467-020-16834-0.
Cvijanovic, I., Mistry, M.N., Begg, J.D., Gasparrini, A., Rodó, X., 2023. Importance of humidity for characterization and communication of dangerous heatwave conditions. Npj Clim. Atmos. Sci. 6 (1), 33. https://doi.org/10.1038/s41612-023-00346-x.
Dou, M., Gu, Y., 2022. Community-level social topic tracking of urban emergency: a case study of COVID-19. Ann. Am. Assoc. Geogr. 112 (7), 1926–1941. https://doi.org/10.1080/24694452.2022.2045182.
Fang, J., Hu, J., Shi, X., Zhao, L., 2019. Assessing disaster impacts and response using social media data in China: a case study of 2016 Wuhan rainstorm. Int. J. Disaster Risk Reduct. 34, 275–282. https://doi.org/10.1016/j.ijdrr.2018.11.027.
Feng, Y., Huang, X., Sester, M., 2022. Extraction and analysis of natural disaster-related VGI from social media: review, opportunities and challenges. Int. J. Geograph. Inf. Sci. 36 (7), 1275–1316. https://doi.org/10.1080/13658816.2022.2048835.
Fotheringham, A.S., Yang, W., Kang, W., 2017. Multiscale geographically weighted regression (MGWR). Ann. Am. Assoc. Geogr. 107 (6), 1247–1265. https://doi.org/10.1080/24694452.2017.1352480.
Giorgini, P., Di Giosia, P., Petrarca, M., Lattanzio, F., Stamerra, C.A., Ferri, C., 2017. Climate changes and human health: a review of the effect of environmental stressors on cardiovascular diseases across epidemiology and biological mechanisms. Curr. Pharm. Des. 23 (22), 3247–3261. https://doi.org/10.2174/1381612823666170317143248.
Goodchild, M.F., Glennon, J.A., 2010. Crowdsourcing geographic information for disaster response: a research frontier. Int. J. Digital Earth 3 (3), 231–241. https://doi.org/10.1080/17538941003759255.
Gosling, S.N., Lowe, J.A., McGregor, G.R., Pelling, M., Malamud, B.D., 2009. Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. Clim. Change 92, 299–341. https://doi.org/10.1016/j.envres.2021.111173.
Hass, A.L., Runkle, J.D., Sugg, M.M., 2021. The driving influences of human perception to extreme heat: a scoping review. Environ. Res. 197, 111173.
He, C., Ma, L., Zhou, L., Kan, H., Zhang, Y., Ma, W., Chen, B., 2019. Exploring the mechanisms of heat wave vulnerability at the urban scale based on the application of big data and artificial societies. Environ. Int. 127, 573–583. https://doi.org/10.1016/j.envint.2019.01.057.
Hsiang, S., Kopp, R., Jina, A., Rising, J., Delgado, M., Mohan, S., Houser, T., 2017. Estimating economic damage from climate change in the United States. Science 356 (6345), 1362–1369. https://doi.org/10.1126/science.aal4369.

Huang, X., Wang, S., Zhang, M., Hu, T., Hohl, A., She, B., Li, Z., 2022. Social media mining under the COVID-19 context: progress, challenges, and opportunities. Int. J. Appl. Earth Obs. Geoinf. 113, 102967. https://doi.org/10.1016/j.jag.2022.102967.

Ji, H., Wang, J., Meng, B., Cao, Z., Yang, T., Zhi, G., Zhang, J., 2022. Research on adaption to air pollution in Chinese cities: evidence from social media-based health sensing. Environ. Res. 210, 112762. https://doi.org/10.1016/j.envres.2022.112762.

Jiang, J., Liu, Y., Mao, J., Wu, G., 2023. Extreme heatwave over Eastern China in summer 2022: the role of three oceans and local soil moisture feedback. Environ. Res. Lett. 18 (4), 044025. https://doi.org/10.1088/1748-9326/acc5fb.

Jung, J., Uejio, C.K., 2017. Social media responses to heat waves. Int. J. Biometeorol. 61, 1247–1260. https://doi.org/10.1007/s00484-016-1302-0.

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

Li, L., Bensi, M., Cui, Q., Baecher, G.B., Huang, Y., 2021. Social media crowdsourcing for rapid damage assessment following a sudden-onset natural hazard event. Int. J. Inf. Manag. 60, 102378. https://doi.org/10.1016/j.ijinfomgt.2021.102378.

Li, J., Wang, X., Hovy, E., 2014. What a nasty day: exploring mood-weather relationship from twitter. In: Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management (CIKM '14). Association for Computing Machinery, New York, NY, USA, pp. 1309–1318, https://doi.org/10.1145/2661829.2662090.

Li, Z., Wang, C., Emrich, C.T., Guo, D., 2018. A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods. Cartogr. Geogr. Inf. Sci. 45 (2), 97–110. https://doi.org/10.1080/15230406.2016.1271356.

Liu, T., Xu, Y.J., Zhang, Y.H., Yan, Q.H., Song, X.L., Xie, H.Y., Ma, W.J., 2013. Associations between risk perception, spontaneous adaptation behavior to heat waves and heatstroke in Guangdong province, China. BMC Public Health 13, 1–14. https://doi.org/10.1186/1471-2458-13-913.

Luber, G., McGeehin, M., 2008. Climate change and extreme heat events. Am. J. Prev. Med. 35 (5), 429–435. https://doi.org/10.1016/j.amepre.2008.08.021.

Lyu, F., Zhou, L., Park, J., Baig, F., Wang, S., 2024. Mapping dynamic human sentiments of heat exposure with location-based social media data. Int. J. Geogr. Inf. Sci. 1–24. https://doi.org/10.1080/13658816.2024.2343063.

Meque, A., Pinto, I., Maúre, G., Beleza, A., 2022. Understanding the variability of heatwave characteristics in southern Africa. Weather Clim. Extremes 38, 100498. https://doi.org/10.1016/j.wace.2022.100498.

Mohammadi, A., Pishgar, E., Fatima, M., Lotfata, A., Fanni, Z., Bergquist, R., Kiani, B., 2023. The COVID-19 mortality rate is associated with illiteracy, age, and air pollution in urban neighborhoods: a spatiotemporal cross-sectional analysis. Trop. Med. Infect. Dis. 8 (2), 85. https://doi.org/10.3390/tropicalmed8020085.

Murakami, D., Peters, G.W., Yamagata, Y., Matsui, T., 2016. Participatory sensing data tweets for micro-urban real-time resiliency monitoring and risk management. IEEE Access 4, 347–372. https://doi.org/10.1109/ACCESS.2016.2516918.

Pappalardo, S.E., Zanetti, C., Todeschi, V., 2023. Mapping urban heat islands and heat-related risk during heat waves from a climate justice perspective: a case study in the municipality of Padua (Italy) for inclusive adaptation policies. Landsc. Urban Plan. 238, 104831. https://doi.org/10.1016/j.landurbplan.2023.104831.

Perkins-Kirkpatrick, S.E., Lewis, S.C., 2020. Increasing trends in regional heatwaves. Nat. Commun. 11 (1), 3357. https://doi.org/10.1038/s41467-020-16970-7.

Resch, B., Usländer, F., Havas, C., 2018. Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment. Cartogr. Geogr. Inf. Sci. 45 (4), 362–376. https://doi.org/10.1080/15230406.2017.1356242.

Rui, J., 2023. Exploring the association between the settlement environment and residents' positive sentiments in urban villages and formal settlements in Shenzhen. Sustain. Cities Soc. 98, 104851. https://doi.org/10.1016/j.scs.2023.104851.

Semenza, J.C., Hall, D.E., Wilson, D.J., Bontempo, B.D., Sailor, D.J., George, L.A., 2008. Public perception of climate change: voluntary mitigation and barriers to behavior change. Am. J. Prev. Med. 35 (5), 479–487. https://doi.org/10.1016/j.amepre.2008.08.020.

Song, J., Huang, B., Kim, J.S., Wen, J., Li, R., 2020. Fine-scale mapping of an evidence-based heat health risk index for high-density cities: Hong Kong as a case study. Sci. Total Environ. 718, 137226. https://doi.org/10.1016/j.scitotenv.2020.137226.

Thompson, R., Hornigold, R., Page, L., Waite, T., 2018. Associations between high ambient temperatures and heat waves with mental health outcomes: a systematic review. Public Health 161, 171–191. https://doi.org/10.1016/j.puhe.2018.06.008.

Wang, J., Meng, B., Pei, T., Du, Y., Zhang, J., Chen, S., Zhi, G., 2021. Mapping the exposure and sensitivity to heat wave events in China's megacities. Sci. Total Environ. 755, 142734. https://doi.org/10.1016/j.scitotenv.2020.142734.

Wei, H., Chen, B., Wu, S., Xu, B., 2023. Impact of early heat anomalies on urban tree cooling efficiency: evidence from spring heatwave events in India. Int. J. Appl. Earth Obs. Geoinf. 120, 103334. https://doi.org/10.1016/j.jag.2023.103334.

White, R.H., Anderson, S., Booth, J.F., Braich, G., Draeger, C., Fei, C., West, G., 2023. The unprecedented Pacific Northwest heatwave of June 2021. Nat. Commun. 14 (1), 727. https://doi.org/10.1038/s41467-023-36289-3.

Wu, H., Huang, B., Zheng, Z., Sun, R., Hu, D., Zeng, Y., 2023. Urban anthropogenic heat index derived from satellite data. Int. J. Appl. Earth Obs. Geoinf. 118, 103261. https://doi.org/10.1016/j.jag.2023.103261.

Wu, Y., Shi, K., Chen, Z., Liu, S., Chang, Z., 2021. Developing improved time-series DMSP-OLS-like data (1992–2019) in China by integrating DMSP-OLS and SNPP-VIIRS. IEEE Trans. Geosci. Remote Sens. 60, 1–14. https://doi.org/10.1109/TGRS.2021.3135333.

Wu, H., Xu, Y., Zhang, M., Su, L., Wang, Y., Zhu, S., 2024. Spatially explicit assessment of the heat-related health risk in the Yangtze River Delta, China, using multisource remote sensing and socioeconomic data. Sustain. Cities Soc., 105300

Xiao, Y., Huang, Q., Wu, K., 2015. Understanding social media data for disaster management. Nat. Hazards 79, 1663–1679. https://doi.org/10.1007/s11069-015-1918-0.

Yang, J., Huang, X., 2021. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. EarthSystem Sci. Data 13 (8), 3907–3925. https://doi.org/10.5194/essd-13-3907-2021.

Yuan, F., Li, M., Liu, R., Zhai, W., Qi, B., 2021. Social media for enhanced understanding of disaster resilience during Hurricane Florence. Int. J. Inf. Manag. 57, 102289. https://doi.org/10.1016/j.ijinfomgt.2020.102289.

Zander, K.K., Rieskamp, J., Mirbabaie, M., Alazab, M., Nguyen, D., 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.

Zhao, L., Li, X., Zhang, Z., Yuan, M., Sun, S., Qu, S., Lin, A., 2023. Developing a novel framework to re-examine half a century of compound drought and heatwave events in mainland China. Sci. Total Environ. 874, 162366. https://doi.org/10.1016/j.scitotenv.2023.162366.

Zou, L., Liao, D., Lam, N.S., Meyer, M.A., Gharaibeh, N.G., Cai, H., Li, D., 2023. Social media for emergency rescue: an analysis of rescue requests on Twitter during Hurricane Harvey. Int. J. Disaster Risk Reduct. 85, 103513. https://doi.org/10.1016/j.ijdrr.2022.103513.