Heliyon 9 (2023) e21987

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Patterns in negative emotions, sleep disorders, and temperature: Evidence from microblog big data

Xiaowen Li [a,b,*], Jun Zhang [b], Bing Li [c]

[a] College of Geography and Tourism, Anhui Normal University, Wuhu, 241000, China
[b] Department of Psychology, Chosun University, Gwangju, 61452, South Korea
[c] College of Art Design & Physical Education, Chosun University, Gwangju, 61452, South Korea

ARTICLE INFO

Keywords:
Microblog big data
Apparent air temperature
Natural language processing implementation
Emotion identification techniques
Negative binomial regression modeling

ABSTRACT

Existing studies have shown that temperature is related to mental illness and sleep disorders. However, few studies have explored the relationship between temperature and microblog negative emotions (MNE) and the predictive effect of MNE on sleep disorders. The present study elucidating the temperature patterns of MNE and sleep disorders, examines the predictive capability of these adverse emotions in precipitating sleep disorders, and operating within the schema of "climate-psychology-behavior". A negative binomial regression model (NBR) was formulated, amalgamating Temperature data, negative affective information procured from microblog, and sleep disorder records. Temperature and Apparent Air Temperature (AAT) were found to have a non-linear association with microblog negative emotions and sleep disorders, exhibiting a modest effect within a specified range, while extreme temperatures (both high and low) demonstrated substantial effects. In the constructed model, gender serves as a moderating factor, with females being more susceptible to temperature and AAT effects on MNE and sleep disorders than their male counterparts. Interestingly, AAT surfaced as a superior predictor compared to actual temperature. MNE were effective predictors of sleep disorders. Employing social media-centric models, as showcased in this study, augments the identification and prevention strategies targeting disease symptoms or pathologies within mental and public health domains.

1. Introduction

The classical heat hypothesis postulates that elevated temperatures could escalate the propensity of individuals towards agitation and heightened aggression [1]. This relationship is also reflected in the higher consumption of antidepressant medication during frigid winters as compared to temperate summers [2]. The study by Patz further reveals the ability of temperature fluctuations to alter psychological states and cognitive behaviors, predominantly evidenced by an upsurge in adverse emotions such as anxiety, depression, and sadness [3]. Such escalation in negative emotions, including anxiety and depression, potentially catalyzes psychological complications like sleep disorders [4]. Beyond mere temperature, thermal indicators, including the diurnal temperature range (DTR) and apparent air temperature (AAT), have captivated substantial scientific interest. Wang Y's investigation inferred a correlation between DTR and the onset of psychological disorders such as anxiety, depression, and schizophrenia [5]. AAT, a composite metric that

* Corresponding author. College of Geography and Tourism, Anhui Normal University, Wuhu, 241000, China.
E-mail addresses: psylxw@gmail.com (X. Li), 921324214@qq.com (J. Zhang), libing05112023@163.com (B. Li).

https://doi.org/10.1016/j.heliyon.2023.e21987
Received 6 June 2023; Received in revised form 28 October 2023; Accepted 1 November 2023
Available online 3 November 2023
2405-8440/

incorporates environmental mean temperature, relative humidity, and wind velocity, offers a more objective representation of human thermal sensation relative to average temperature. Prior investigations suggest that AAT exhibits a more pronounced correlation with mental disorders than other temperature-related variables [6]. Conversely, some academicians report a lack of significant association between temperature and emotional states [7]. Aside from potential inaccuracies in measurement, this absence of correlation could potentially be attributed to a curvilinear relationship between the two factors. Based on a thorough literature review, Wang Y posited the existence of a nonlinear association between climatic conditions and emotional states [8]. This view is supported by a research collective from Tianjin University, who proposed a nonlinear link between environmental variables and the dependent variable [9]. Larsen's research also concurs on a nonlinear correlation between emotional indices and temperature, indicating a lack of correlation at lower temperatures but a marked mood deterioration at temperatures exceeding 21° Celsius. Interestingly, the mood discrepancy within temperature brackets of roughly 16 °C–21 °C and 27 °C–32 °C is analogous to the average mood deviation between Sundays and Mondays [10].

Numerous studies have examined the relationship between relative temperature and mental health, but few have specifically focused on the correlation between temperature and negative emotions expressed on social media. For example, Baylis utilized Twitter and Facebook data from the United States to explore how temperature affects both positive and negative emotions. Their findings revealed that negative emotions increase at both lower temperatures and extremely high temperatures (above 30° Celsius), while negative emotions are less prevalent between 20 and 30° Celsius [11]. Building on this research, we propose hypothesis H_1: There is a nonlinear relationship between temperature variables such as DTR and AAT, and mental health indicators including MNE and sleep disorders.

Numerous academics propose that the impact of environmental factors such as temperature on individual mental health is influenced by a variety of demographic characteristics, including gender, age, and socioeconomic status [12]. Building on this understanding, we formulate hypothesis H_2: Gender specifically acts as a moderating factor in the relationship model that connects temperature with mental health indicators such as MNE and sleep disorders.

At present, social media platforms engage approximately half of the global population, accounting for about a third of daily internet usage [13]. The diverse social media landscape includes text-oriented platforms such as blogs, forums, and microblogs, image-centered applications like Moments and Instagram, and multimedia channels including TikTok and YouTube. This diversity not only encompasses various formats—text, images, videos—but also extends to differences in anonymity, privacy, and potential sensitivity. The extensive public use of social media has revolutionized information dissemination and enabled avenues for data aggregation. Consequently, these platforms have emerged as valuable tools for researchers, allowing them to engage with study participants at minimal cost [14]. For instance, a survey conducted by Al Zou'bi across varied groups including undergraduates, graduate students, and researchers revealed that an impressive 76 % of respondents trust research findings based on social media data collection [15]. In the context of the present research, Microblog, a long-standing social media platform with an expansive user base, will be utilized. Its ability to provide pertinent demographic data significantly facilitates subsequent research undertakings.

Advancements in big data technologies and machine learning have enabled the utilization of the rapidly increasing data generated by social media platforms for predicting and automatically identifying signs of mental health and psychiatric disorders [16]. Research conducted by Liu demonstrates how microblog big data can overcome the data reporting delays typically encountered in traditional research methods, such as surveys and observational studies. This capability is further enhanced by the dynamic nature of social media platforms as data sources. Geotagged content, in particular, offers a novel supplement to traditional methods, extending visitor monitoring capabilities. It provides insights into temporal trends of emerging activities and shared content, as well as spatial patterns of visitor movements, all of which are beyond the reach of conventional approaches [17]. To summarize, we posit H_3: MNE are indicative of sleep disorders.

2. Data and methods

2.1. Data source

Data pertaining to temperature was acquired from the China Meteorological Data Service Center, encompassing daily meteorological data spanning the study period from 2018 to 2020. This amassed daily meteorological data underwent processing to identify and rectify missing values, inaccuracies, or outliers. Following this, parameters such as daily average temperature, DTR, and AAT were computed. AAT was derived using prevalent meteorological parameters, including average temperature, relative humidity, and wind speed, based on the subsequent formula: $AAT = 1.17T + 0.2e - 0.65V - 2.7$, $e = \frac{RH}{100} \times 6.105 \times \exp\frac{17.27T}{237.7+T}$.Within the formula, '*AAT*' symbolizes the perceived temperature (°C), '*T*' stands for the temperature (°C), '*e*' signifies water vapor pressure (hPa), '*V*' represents wind speed (m/sec), and '*RH*' corresponds to relative humidity (%). Evidently, increased values of relative humidity and wind speed contribute to a lower perceived temperature [18].

DTR is quantified as $DTR = T_{\max}$ - $T_{\min}$. In this equation, $T_{\max}$ denotes the highest temperature (°C) observed within a day, while $T_{\min}$ signifies the lowest temperature (°C) within the same period.

Data pertaining to sleep disorders was obtained from the electronic medical records across numerous hospitals in Hefei, Anhui Province, spanning the duration of 2018–2020. This collection encompasses both outpatient and inpatient records of individuals diagnosed with sleep disorders within the specified study period.

To collect MNE data, the first stage involved employing Python to scrape relevant text from Sina Weibo, guided by keywords such as "weather", "temperature", "high temperature", "cold", from the Hefei region for the period between 2018 and 2020. Post data cleaning efforts, including deduplication, a sum of 37.1 million valid entries was amassed. Subsequent to this, the collected text underwent natural language processing to convert it into machine-digestible data. This process encompassed the removal of stop words and word segmentation, employing a stop word list supplied by the Harbin Institute of Technology and the Jieba word segmentation library. This is a pivotal step in natural language processing, enabling the extraction of significant verbiage and the elimination of immaterial words. The final step of the processing entailed text vectorization [19,20]. Text vectorization forms the nucleus of natural language processing, facilitating the transformation of human language into machine-readable data. Common techniques for text vectorization include TF-IDF, word2vec, and deep learning models, with this study employing the word2vec method. Following this, a Bayesian classifier was utilized to categorize the machine-readable data and isolate negative emotions. The Bayesian classifier is currently perceived as an efficacious theoretical model in knowledge representation and inferencing [21]. The classifier utilized in this study for emotion recognition comprised two key phases: the initial phase involved classifier learning, wherein the classifier was structured using sample data. For this study, 108,000 data entries were manually labeled to construct the Bayesian classifier.According to the division standard of the Positive and Negative Affect Scale (PANAS), the emotions were divided into positive emotions, neutral emotions and negative emotions. Negative emotions including Distressed/Upset/Guilty/Scared/Hostile/Irritable/Ashamed/Nervous/Jittery/Afraid.In the process of manual labeling, in order to ensure the accuracy, two groups of mutual proofreading are adopted, that is, group A checks the labeling of group B, and group B checks the labeling of group A; The final phase involved classifier inference, wherein the conditional probability of class nodes was computed for classification and emotion categorization, as delineated in the ensuing formula:Presuming a text comprises n features, specifically $e_1, e_2, \ldots, e_n$, and has i categories, specifically $m_1, m_2, \ldots, m_i$, the attribute independence presumption instituted by the Bayesian classifier can be depicted via the subsequent formula: $p(m_i|e_1\ , e_2\ , \ \ldots\ , e_n)p(m_i)\ = p(e_1|m_i)p(e_2|m_i)\ldots p(e_n|m_i)p(m_i)$.The most plausible category for an entity is ascertained by the maximum posterior estimate, demonstrated in the succeeding equation: $C_{MAP}\ = p(e_1|m_i)p(e_2|m_i)\ldots p(e_n|m_i)p(m_i)$.Each probability value on the right side of the equation can be facilely calculated given the known conditions, thereby facilitating the calculation of the probability for the text belonging to each respective category. The category possessing the highest probability is selected as the ultimate classification outcome [22].

2.2. Research methodology

The dependent variables in this study are the count of MNE and sleep disorders, both of which are count variables. Poisson models and negative binomial models are most commonly used in count models. The Poisson regression model has higher data requirements, requiring the conditional mean function and the variance function to be equivalent. The usual method for choosing between these two models is the *O-test*. The *O-test* also known as the over-dispersion test, can be used to determine whether there is over-dispersion in the data [23]. The over-dispersion test showed that $O = 7.02$ (>1.96), indicating over-dispersion in the data. The Poisson regression model is not suitable for this study; hence, the negative binomial regression (NBR) model is selected.

The negative binomial distribution for the dependent variable is defined by Equation: $f((y_i|x_i))\ = \frac{F(\theta+y_i)}{r(Y_i+1)(\theta)}\left(\frac{\lambda_i}{\lambda_i+\theta}\right)^{yi}\left(1-\frac{\lambda_i}{\lambda_i+\theta}\right)^{\theta yi}$, λ_i is conditional mean, $\lambda_i[1+(1/\theta)\lambda_i]$ is conditional variance. In this study, a negative binomial regression model was chosen to model the survey data by incorporating an unobservable individual effect into the conditional mean of the Poisson model: $1nY_i\ = X_i\beta+\ \varepsilon_i\ , i = 1, 2, \ldots, N$.It can reflect both the random error term in the classical regression equation and the heteroscedasticity caused by cross-sectional data. In the equation, Y represents the dependent variable, X represents the vector of explanatory variables and control variables, and β represents the vector of coefficients for the explanatory variables and control variables.

2.3. Research tools

The analysis is carried out using *R* software. Statistical analysis of the data is performed using packages such as “splines” and “dlnm” in *R* software (R 3.4.2). The test standard is set at $\alpha = 0.05$.

3. Research results

3.1. Descriptive statistical results

During the research period (2018–2020), a total of 37.1 million microblog posts were counted, of which 10.64 million (28.7 %) were negative in sentiment. On average, there were 33,828 posts per day ($SD = 6765$), of which 9716 ($SD = 1721$) were negative. Over the same period, there were 88,695 cases of sleep disorders, averaging 81 per day ($SD = 18$). The number of microblog negative emotions was highest on Mondays (daily average of 1278) and lowest over the weekends (Saturday, 987; Sunday, 1147). The frequency of sleep disorders was highest over the weekend (Saturday, 98; Sunday, 104) and lowest on Wednesdays (75 per day).

3.2. Temperature patterns of MNE and sleep disorders based on NBR

By constructing a NBR model, we examined the relationship between daily average temperature, DTR, AAT, and the number of microblog negative emotions and sleep disorders. To account for the nonlinear relationship, polynomial terms of average temperature, DTR, and AAT were examined. We controlled for confounding factors such as time trends, day of the week effects, and holidays. Additionally, to investigate the moderating effect of gender, interaction terms between gender and temperature, DTR, and AAT were included. The AIC was used to compare the model fit for the count of MNE posts. A smaller AIC indicates a better model fit [24].

Table 1
NBR model of daily temperature, DTR, AAT and the number of MNE.

	The number of MNE			The number of MNE			The number of MNE		
	Coefficient (SE)	IRR (95 % CI)	*p*	Coefficient (SE)	IRR (95 % CI)	*p*	Coefficient (SE)	IRR (95 % CI)	*p*
Intercept	2.668(0.021)		<.001	2.787(0.024)		<.001	2.702(0.023)		<.001
Gender	0.138(0.005)	1.148(1.119,1.167)	<.001	0.156(0.006)	1.169(1.137,1.201)	<.001	0.141(0.004)	1.151(1.134,1.177)	<.001
Temperature	−0.028(0.006)		<.001						
Temperature ^{2}	0.002(0.0002)		<.001						
Temperature × Gender	0.003(0.00017)		<.001						
Temperature^2 × Gender	0.00001 (0.00001)		0.159						
DTR				−0.018(0.017)		0.145			
DTR ^{2}				0.003(0.005)		0.274			
DTR × Gender				0.004(0.011)		0.358			
DTR^2 × Gender				0.00003 (0.00002)		0.067			
AAT							−0.033(0.006)		<.001
AAT ^{2}							0.002(0.0002)		<.001
AAT × Gender							0.004(0.00016)		<.001
AAT^2 × Gender							0.00002 (0.00001)		0.023
Time	−0.00039 (0.00001)	0.99961 (0.99958, 0.99963)	<.001	−0.00044 (0.00002)	0.99961 (0.99958, 0.99963)	<.001	−0.00041 (0.00001)	0.99961 (0.99958, 0.99963)	<.001
Day (ref = Monday)									
Tuesday	−0.019 (0.015)	0.981 (0.953, 1.010)	0.102637252	−0.024(0.016)	0.976(0.948,1.005)	0.066807201	−0.021(0.017)	0.979(0.951,1.008)	0.108360531
Wednesday	−0.020 (0.015)	0.980 (0.952, 1.009)	0.09121122	−0.024(0.021)	0.976(0.948,1.005)	0.126548954	−0.025(0.023)	0.975(0.947,1.004)	0.138528013
Thursday	−0.033 (0.015)	0.968 (0.940, 0.996)	0.013903448	−0.036(0.018)	0.965(0.937,0.993)	0.022750132	−0.035(0.017)	0.966(0.938,0.994)	0.020
Friday	−0.055 (0.015)	0.946 (0.919, 0.974)	<.001	−0.057(0.020)	0.945(0.918,0.973)	0.002185961	−0.062(0.018)	0.940(0.913,0.968)	<.001
Saturday	−0.154 (0.015)	0.858 (0.833, 0.883)	<.001	−0.157(0.018)	0.855(0.831,0.881)	<.001	−0.157(0.016)	0.855(0.831,0.881)	<.001
Sunday	−0.101 (0.015)	0.904 (0.879, 0.931)	<.001	−0.103(0.021)	0.902(0.877,0.929)	<.001	−0.108(0.019)	0.898(0.873,0.925)	<.001
Holiday	−0.060 (0.024)	0.942 (0.898, 0.988)	<.001	−0.061(0.030)	0.941(0.897,0.987)	0.021009429	−0.069(0.034)	0.933(0.889,0.979)	0.021208184
AIC	18534			18943			17956		

Table 2
NBR model of daily temperature, DTR, AAT and the number of sleep disorders.

	sleep disorders			sleep disorders			sleep disorders		
	Coefficient (SE)	IRR (95 % CI)	*p*	Coefficient (SE)	IRR (95 % CI)	*p*	Coefficient (SE)	IRR (95 % CI)	*p*
Intercept	13.513(0.467)		<.001	14.106(0.477)		<.001	13.891(0.458)		<.001
Gender	1.090(0.039)	2.973(2.433,3.187)	<.001	1.140(0.051)	3.126(2.811,3.317)	<.001	1.102(0.044)	3.011(2.724,3.256)	<.001
Temperature	−0.098(0.012)		<.001						
Temperature ^{2}	0.006(0.0003)		<.001						
Temperature × Gender	0.005(0.00021)		<.001						
Temperature^{2} × Gender	0.00003 (0.00005)		0.274						
DTR				−0.043(0.037)		0.126			
DTR ^{2}				0.004(0.008)		0.309			
DTR × Gender				0.018(0.027)		0.667			
DTR^{2} × Gender				0.00012 (0.00018)		0.252			
AAT							−0.157(0.012)		<.001
AAT ^{2}							0.011(0.0003)		<.001
AAT × Gender							0.015(0.00027)		<.001
AAT^{2} × Gender							0.00019 (0.00013)		0.072
Time	−0.00001 (0.00001)	0.99999(0.99997, 1.00003)	0.157	−0.00001 (0.00002)	0.99999(0.99995, 1.00006)	0.309	−0.00002 (0.00003)	0.99999(0.99996, 1.00005)	0.252
Day (ref = Monday)									
Tuesday	−0.027 (0.013)	0.973 (0.948, 0.998)	0.01890433	−0.035(0.017)	0.966(0.941,0.991)	0.020	−0.031(0.017)	0.969(0.944,0.994)	0.034111623
Wednesday	−0.055 (0.013)	0.946 (0.922, 0.971)	<.001	−0.065(0.019)	0.937(0.913,0.962)	<.001	−0.063(0.021)	0.939(0.915,0.964)	0.001349898
Thursday	0.006 (0.013)	1.006 (0.981, 1.032)	0.322206167	−0.004(0.018)	0.996(0.971,1.022)	0.412070448	0.001(0.015)	1.001(0.976,1.027)	0.473423536
Friday	0.075 (0.013)	1.077 (1.051, 1.105)	<.001	0.073(0.016)	1.076(1.049,1.103)	<.001	0.074(0.020)	1.077(1.050,1.104)	<.001
Saturday	0.242 (0.013)	1.274 (1.243, 1.307)	<.001	0.239(0.017)	1.270(1.239,1.303)	<.001	0.236(0.022)	1.266(1.235,1.299)	<.001
Sunday	0.250 (0.013)	1.283 (1.252, 1.316)	<.001	0.241(0.018)	1.273(1.241,1.305)	<.001	0.241(0.015)	1.273(1.241,1.305)	<.001
Holiday	0.192 (0.020)	1.212 (1.165, 1.261)	<.001	0.182(0.023)	1.200(1.153,1.249)	<.001	0.186(0.022)	1.204(1.157,1.253)	<.001
AIC	10532			11268			10137		

Tables 1 and 2 present the results of the NBR, which examines the relationships between daily average temperature, DTR, AAT, the count of MNE and sleep disorders, while controlling for time trends, day of the week effects, and holidays. The models indicate that temperature, temperature squared, perceived temperature, and perceived temperature squared are significant in both models ($p < .001$), indicating a nonlinear relationship between daily average temperature, AAT, and the count of MNE and sleep disorders. Specifically, temperature and AAT exhibit a negative correlation with the count of MNE and sleep disorders, while temperature squared and perceived temperature squared show a positive correlation. This suggests a curvilinear relationship, where the association between temperature, AAT, and the count of MNE and sleep disorders is moderate within a certain range. Both high and low temperatures lead to an increase in the count of MNE and sleep disorders (Fig. 1). In terms of predictive performance, perceived temperature outperforms average temperature for both the count of MNE and sleep disorders (as indicated by lower AIC values). DTR does not show a significant relationship with the count of MNE or sleep disorders. Tables 1 and 2 demonstrate that the interaction terms between gender and temperature, as well as gender and perceived temperature, are significantly positive ($p < .001$). Heterogeneous effects do exist by gender, we can conclude that the impact of temperature on MNE is 0.003 greater for women than men, and the effect of AAT is 0.004 greater. The influence of temperature on sleep disorders is 0.005 greater for women, and the effect of AAT is 0.015 greater.

The count of MNE significantly decreases on weekends, with the lowest count observed on Saturdays compared to Mondays (*IRR* = *0.858, 0.855, 0.855,* respectively). In contrast, the count of sleep disorders is significantly higher on weekends compared to Mondays, with the highest count observed on Sundays (*IRR* = *1.283, 1.273, 1.273*, respectively). If it is a holiday, the count of MNE significantly decreases, while the count of sleep disorders significantly increases. One possible explanation for this phenomenon is the source of the data. The sleep disorder data in this study is derived from electronic records of public hospitals. For the majority of working individuals, it is more convenient to seek medical treatment at public hospitals on weekends or holidays.

3.3. Relationship between the count of MNE and sleep disorders

Table 3 illustrates the NBR analysis outcomes, which evaluated the correlation between the average daily temperature, DTR, AAT, and the count of MNE with the incidence of sleep disorders, whilst accounting for temporal trends, weekday effects, and holidays. The findings underscored a rise in the frequency of sleep disturbances with an increasing count of MNE (IRR = 1.099, 1.108, 1.103/per 1000 posts). A comparison between the AIC values of models in Table 3, which incorporated the count of MNE, and models in Table 2, which did not, revealed a lower AIC in the former. This suggests an improved model fit with the inclusion of MNE variable, thus highlighting the MNE as a robust predictor of sleep disorders.

4. Conclusion and outlook

This research integrates temperature, MNE, and sleep disorders within an NBR framework, contributing to the existing body of literature on the impact of temperatures and AAT on sleep disorders. The findings highlight a plateau effect within defined bounds, where extreme temperatures lead to more pronounced effects. Relative to temperature, AAT emerges as a stronger predictor, supporting the complex curvilinear relationship between climate and psychology, and refuting the idea of a simplistic linear correlation, as endorsed by previous studies. The human perception of climate involves a comfort zone, comprising factors such as temperature, humidity, sunshine, wind speed, and precipitation. Within this zone, individuals are more likely to experience positive emotions compared to outside it. This insight is aligned with other research findings, revealing a nonlinear relationship between emotional indicators and temperature, with temperatures above 21° Celsius causing a noticeable drop in mood [11]. Kovats systematically determined that the optimal working temperatures for individuals lie between 10 and 20 °C [25], while Maloney's study, utilizing a questionnaire method, explored the ideal temperature to amplify residents' happiness [26]. The environmental stress model, put forth

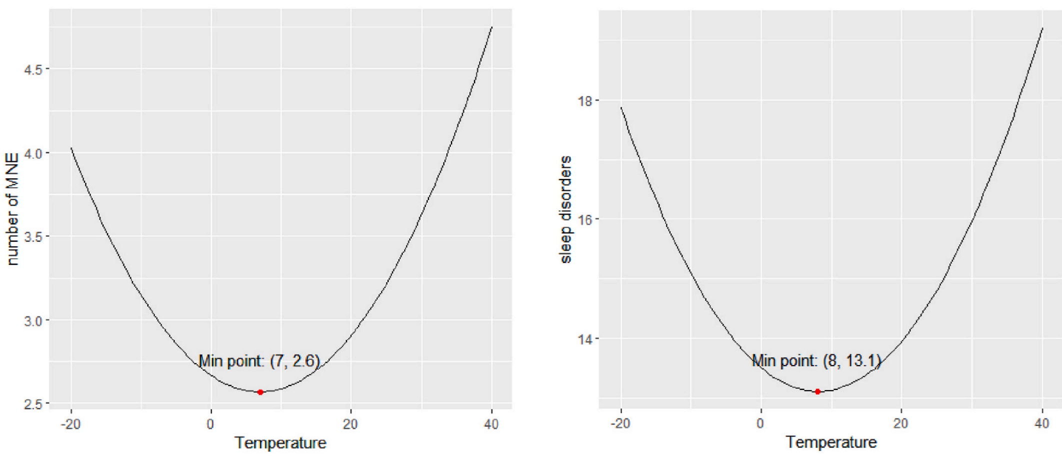

Fig. 1. Regression model fit plot.

Table 3
NBR model of daily temperature, DTR, AAT, MNE and the number of sleep disorders.

	Sleep disorders			Sleep disorders			Sleep disorders		
	Coefficient (SE)	IRR (95 % CI)	*p*	Coefficient (SE)	IRR (95 % CI)	*p*	Coefficient (SE)	IRR (95 % CI)	*p*
Intercept	8.979(0.367)		<.001	8.843(0.337)		<.001	8.531(0.317)		<.001
Anger count/1000	0.094(0.020)	1.099(1.092,1.105)	<.001	0.103(0.018)	1.108(1.102,1.115)	<.001	0.098(0.019)	1.103(1.094,1.109)	<.001
Temperature	−0.107(0.031)		<.001						
Temperature ^{2}	0.00536(0.00173)		0.001						
DTR				−0.084(0.059)		0.077			
DTR ^{2}				0.00414(0.00367)		0.130			
AAT							−0.137(0.031)		<.001
AAT ^{2}							0.00483(0.00143)		<.001
Time	−0.00004 (0.00001)	0.99996(0.99993, 0.99999)	<.001	−0.00007 (0.00002)	0.99993(0.99989, 0.99997)	<.001	−0.00011 (0.0003)	0.99989(0.99993, 0.99999)	<.001
Day (ref = Monday)									
Tuesday	−0.030 (0.013)	0.970 (0.946, 0.996)	0.010508128	−0.032(0.021)	0.969(0.945,0.995)	0.013	−0.035(0.016)	0.966(0.942,0.992)	0.009
Wednesday	−0.057 (0.013)	0.944 (0.920, 0.969)	<.001	−0.059(0.019)	0.943(0.918,0.967)	<.001	−0.059(0.016)	0.943(0.918,0.967)	<.001
Thursday	0.002 (0.013)	1.002 (0.977, 1.028)	0.438865521	−0.004(0.019)	0.996(0.971,1.022)	0.763	−0.004(0.017)	0.996(0.971,1.022)	0.707
Friday	0.067 (0.013)	1.070 (1.043, 1.097)	<.001	0.064(0.017)	1.066(1.04,1.094)	<.001	0.061(0.016)	1.063(1.037,1.091)	<.001
Saturday	0.224 (0.013)	1.251 (1.220, 1.284)	<.001	0.217(0.020)	1.242(1.211,1.275)	<.001	0.216(0.015)	1.241(1.210,1.274)	<.001
Sunday	0.237 (0.013)	1.268 (1.237, 1.300)	<.001	0.234(0.017)	1.264(1.234,1.297)	<.001	0.231(0.018)	1.260(1.230,1.293)	<.001
Holiday	0.186 (0.020)	1.204 (1.158, 1.253)	<.001	0.178(0.022)	1.195(1.149,1.244)	<.001	0.181(0.026)	1.198(1.152,1.247)	<.001
AIC	8932			9083			8731		

by Evans and colleagues, further explains this phenomenon. It characterizes temperature as a human stressor and threat, instigating negative emotions as a stress response to perceived threats under extreme temperatures [27]. This investigation asserts a negligible effect of the DTR on sleep disorders, a finding that contrasts with the study by Yi, W et al. [28]. A potential explanation for this divergence might be that the study was conducted in a typical subtropical monsoon climate, where extreme variations in DTR are infrequent, except for occasional occurrences during the spring and autumn seasons. This is consistent with findings by Küller, who highlight region-specific temperature perception, demonstrating a distinct latitudinal gradient. Specifically, lower latitudes correlate with increased thermal tolerance and decreased cold tolerance, along with enhanced cold sensitivity among the population [29]. Further, the analysis model reveals a moderating influence of gender on these effects, with females showing higher sensitivity to both temperature and AAT than males. This is observed both in terms of MNE and the frequency of sleep disorders, aligning with the findings of Tawatsupa et al. [30]. Considering the problem from the perspective of psychological distress, men have a stronger ability to control emotions, which can reduce the psychological distress brought about by external factors, compared to women.

The predictive efficacy of MNE in forecasting sleep disorders has been observed, and a parallel temperature pattern is evident in both MNE and sleep disturbances. Social psychology and media communication theories can provide an explanation for this phenomenon. For instance, the Media and Health Theory posits a positive correlation between exposure to content on social media platforms and an increased inclination to engage in the depicted behaviors [31]. These platforms, offering an extensive and accessible social network to those struggling with sleep disturbances, serve as a continual source of personalized and socially integrated content, facilitating communication and imitation. The Cultivation Theory, on the other hand, suggests that consistent exposure to behaviors through media might distort an individual's perception of reality, potentially inflating the frequency and pervasiveness of these behaviors in the real world [32]. This distorted perception could increase susceptibility to behaviors represented in media [33]. Supporting this notion, the Social Learning Theory emphasizes that observing specific behaviors through social media can influence imitative actions or thoughts. Bandura's insights propose that media communication can "educate, empower, motivate, and guide" audiences, catalyzing significant shifts in human psychology and behavior. In addition, the study by Kim and Bentley reinforces these theories, demonstrating that factors like user engagement (e.g., comment count) and perceived social support (e.g., received 'likes') can alter user behavior. Collectively, these studies offer a novel empirical framework to explore how contemporary collectivism has evolved in the digital age, with individuals exhibiting more "herd-like" tendencies in decision-making as their interdependence increases [34].

This investigation employs an innovative model driven by social media data, pertinent for the detection and deterrence of mental health disorders and symptoms in the fields of public health and mental health. Considering the burgeoning volume of social media data owing to its pervasive utilization by the populace and advancements in big data mining and machine learning technologies, this evidence-based framework rooted in social media big data plays a vital role in enhancing preventive measures and recovery initiatives for mental health. Despite its strengths, this study presents certain limitations and areas for potential enhancement:1. According to scholars both domestically and internationally, the impact of climatic factors on mental health is modulated by demographic variables. However, in this study, due to data feasibility, only gender was incorporated as a moderating variable. Age and socioeconomic status may also function as possible moderating variables.

2 Emotional states are inherently nuanced, multifaceted, and operate on multiple levels. Given the constraints of the Bayesian classifier used in this study, text recognition was categorized into three emotional states: positive, neutral, and negative. However, with the advent of deep learning models, further delineation and recognition of emotions have become feasible. Subsequent research endeavors aim to achieve more granular emotion recognition through the application of advanced deep learning models.
3 The nature of acquisition and processing of large social media datasets could potentially predispose these data to Type *I* errors [35], leading to false identifications and an overproduction of inferences derived from the data. Future research could enhance the efficacy of social media data-driven methods by employing a hybrid approach that amalgamates survey research with longitudinal follow-up survey recruitment strategies.

Ethics statement

The study protocol was approved by the Ethics Committee of Anhui Normal University (number 2020XJ44). All participants provided voluntary written informed consent before study procedures.

Funding

This research received no external funding.

Data availability statement

No, Research-related data are not deposited in public repositories. Data will be made available on request.

Additional information

No additional information is available for this paper.

CRediT authorship contribution statement

Xiaowen Li: Writing – review & editing, Writing – original draft, Project administration. **Jun Zhang:** Data curation, Conceptualization. **Bing Li:** Software.

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.

References

[1] C.A. Anderson, Heat and violence, Curr. Dir. Psychol. Sci. 10 (2001) 33–38.
[2] K.B. Murray, F. Di Muro, A. Finn, et al., The effect of weather on consumer spending, J. Retail. Consum. Serv. 17 (2010) 512–520.
[3] J.A. Patz, D. Campbell-Lendrum, T. Holloway, et al., Impact of regional climate change on human health, Nature 438 (2005) 310–317.
[4] C. Noelke, M. McGovern, D.J. Corsi, et al., Increasing ambient temperature reduces emotional well-being, Environ. Res. 151 (2016) 124–129.
[5] Y. Wang, M.N. Eliot, P. Koutrakis, et al., Ambient air pollution and depressive symptoms in older adults: results from the MOBILIZE Boston study, Environ. Health Perspect. 122 (6) (2014) 553–558.
[6] S. Vida, M. Durocher, T.B. Ouarda, et al., Relationship between ambient temperature and humidity and visits to mental health emergency departments in Québec, Psychiatr. Serv. 63 (11) (2012) 1150–1153.
[7] J.A. Denissen, L. Butalid, L. Penke, M.A.G. van Aken, The effects of weather on daily mood: a multilevel approach, Emotion 8 (2008) 662–667.
[8] Y. Wang, H. Chen, Human beings are born with the energy of heaven and earth: the influence of meteorology on human psychology and behavior, Adv. Psychol. Sci. 25 (6) (2017) 1077–1092.
[9] Z. Mi, G. Tu, X. Wang, Application of artificial neural networks in thermal comfort experimental research, J. Tianjin Univ. 4 (37) (2004) 331–335.
[10] M.E. Larsen, T.W. Boonstra, P.J. Batterham, et al., We feel: mapping emotion on twitter, IEEE J. Biomed. Health Inform. 19 (4) (2015) 1246–1252.
[11] P. Baylis, N. Obradovich, Y. Kryvasheyeu, H. Chen, L. Coviello, E. Moro, M. Cebrian, J. Fowler, Weather impacts expressed sentiment, PLoS One 13 (4) (2018), e0195750.
[12] A.M. Abbasi, M. Motamedzade, M. Aliabadi, et al., The impact of indoor air temperature on the executive functions of the human brain and the physiological responses of the body, Health Promot. Perspect. 9 (1) (2019) 55–64.
[13] S. Lee, M. Xenos, Social distraction? Social media use and political knowledge in two U.S. Presidential Elections, Comput. Hum. Behav. 90 (2019) 18–25.
[14] D. Arigo, S. Pagoto, L. Carter-Harris, et al., Using Social Media for Health Research: Methodological and Ethical Considerations for Recruitment and Intervention Delivery, SAGE Publications, 2018, p. 205.
[15] H.W. Al Zou'bi, M. Khatatbeh, K.H. Alzoubi, et al., Attitudes and knowledge of adolescents in Jordan regarding the ethics of social media data use for research purposes, J. Emp. Res. Hum. Res. Ethics 15 (1) (2020) 87–96.
[16] J. Nicholas, S. Onie, M.E. Larsen, Ethics and privacy in social media research for mental health, Curr. Psychiatr. Rep. 22 (12) (2020).
[17] S. Liu, S.D. Young, A survey of social media data analysis for physical activity surveillance, J. Forens. Legal Med. 57 (2018) 33–36.
[18] M. Min, The Impact of Temperature Changes and Environmental Pollutants on the Admission of Patients with Mental and Behavioral Disorders in Yancheng City, Anhui Medical University, 2020.
[19] X. Liu, D. Tian, Online comment sentiment analysis based on the fusion of deep learning and machine learning, Softw. Guide 18 (2) (2019) 1–4.
[20] H. Weng, T. Ma, Analysis of the spatiotemporal differentiation characteristics and influencing factors of temperature perception among the Chinese population based on Weibo data, J. Geo-Inform. Sci. 24 (10) (2022) 2058–2070.
[21] F. Yin, X. Zhang, Q. Feng, et al., Research on the classification method of university students' attention to hot events on Weibo based on Bayesian classification, Softw. Guide 15 (6) (2016) 112–114.
[22] D. Zhang, Y. Wang, C. Zhao, et al., Bayesian decision-based model for extremely short text classification, J. Chongqing Univ. Sci. Technol. (Nat. Sci. Ed.) 20 (4) (2018) 82–86.
[23] G.P. Zhu, L.L. Zhu, M.C. Meng, et al., Application of zero-inflated negative binomial regression model in the study of influencing factors of comorbid diseases, Chin. J. Dis. Control Prev. 22 (10) (2018) 1063–1066.
[24] H.R. Stevens, P.L. Graham, P.J. Beggs, et al., In cold weather we bark, but in hot weather we bite: patterns in social media anger, aggressive behavior, and temperature, Environ. Behav. (2020) 46–49.
[25] R.S. Kovats, S. Hajat, Heat stress and public health: a critical review, Annu. Rev. Publ. Health 29 (2008) 41–55.
[26] M.P. Maloney, M.P. Ward, G.N. Braucht, A revised Scale for the measurement of ecological attitude and knowledge, Am. Psychol. 30 (7) (1975) 787–790.
[27] S. Cohen, G.W. Evans, D. Stokols, D.S. Krantz, Behavior, Health, and Environmental Stress, Springer Science and Business Media, 2013.
[28] W. Yi, X. Zhang, R. Pan, et al., Quantifying the impacts of temperature variability on hospitalizations for schizophrenia: a time series analysis in Hefei, China, Sci. Total Environ. 696 (2019), 133927.
[29] R. Küller, The impact of light and colour on psychological mood: a cross-cultural study of indoor work environments, Ergonomics 49 (14) (2006) 1496–1507.
[30] B. Tawatsupa, L.L. Lim, T. Kjellstrom, S.A. Seubsman, A. Sleigh, The Thai Cohort Study team, The association between overall health ， psychological distress ， and occupational heat stress among a large national cohort of 40 ，913 Thai workers, Glob. Health Action 3 (1) (2010) 5034.
[31] S.J. Kim, L.A. Marsch, J.T. Hancock, et al., Scaling up research on drug abuse and addiction through social media big data, J. Med. Internet Res. 19 (10) (2017) e353.
[32] H. Bilandzic, The complicated relationship between media and morality, Media Psychol. 23 (1) (2011) 46–51.
[33] I. Ajzen, M. Fishbein, Attitude-behavior relations: a theoretical analysis and review of empirical research, Psychol. Bull. 84 (5) (1977) 888–918.
[34] R.A. Bentley, M.J. O'Brien, W.A. Brock, Mapping collective behavior in the big-data era, Behav. Brain Sci. 37 (1) (2014) 63–76.
[35] S.D. Young, Behavioral insights on big data: using social media for predicting biomedical Outcomes, Trends Microbiol. 22 (11) (2014) 601–602.