Temperature impacts on hate speech online: evidence from 4 billion geolocated tweets from the USA

Annika Stechemesser, Anders Levermann, Leonie Wenz

Lancet Planet Health 2022; 6: e714–25

Potsdam Institute for Climate Impact Research, Potsdam, Germany (A Stechemesser MSc, Prof A Levermann PhD, L Wenz PhD)**; Institute of Physics, Potsdam University, Potsdam, Germany** (A Stechemesser, Prof A Levermann)**; Lamont-Doherty Earth Observatory, Columbia University, New York, NY, USA** (Prof A Levermann)**; Mercator Research Institute on Global Commons and Climate Change, Berlin, Germany** (L Wenz)

Correspondence to: Dr Leonie Wenz, Potsdam Institute for Climate Impact Research, Potsdam 14473, Germany **leonie.wenz@pik-potsdam.de**

Summary

Background **A link between weather and aggression in the offline world has been established across a variety of societal settings. Simultaneously, the rapid digitalisation of nearly every aspect of everyday life has led to a high frequency of interpersonal conflicts online. Hate speech online has become a prevalent problem that has been shown to aggravate mental health conditions, especially among young people and marginalised groups. We examine the effect of temperature on the occurrence of hate speech on the social media platform Twitter and interpret the results in the context of the interlinkage between climate change, human behaviour, and mental health.**

Methods **In this quantitative empirical study, we used a supervised machine learning approach to identify hate speech in a dataset containing around 4 billion geolocated tweets from 773 cities across the USA between May 1, 2014 and May 1, 2020. We statistically evaluated the changes in daily hate tweets against changes in local temperature, isolating the temperature influence from confounding factors using binned panel-regression models.**

Findings **The prevalence of hate tweets was lowest at moderate temperatures (12 to 21°C) and marked increases in the number of hate tweets were observed at hotter and colder temperatures, reaching up to 12·5% (95% CI 8·0–16·5) for cold temperature extremes (–6 to –3°C) and up to 22·0% (95% CI 20·5–23·5) for hot temperature extremes (42 to 45°C). Outside of the moderate temperature range, the hate tweets also increased as a proportion of total tweeting activity. The quasi-quadratic shape of the temperature–hate tweet curve was robust across varying climate zones, income quartiles, religious and political beliefs, and both city-level and state-level aggregations. However, temperature ranges with the lowest prevalence of hate tweets were centred around the local temperature mean and the magnitude of the increases in hate tweets for hot and cold temperatures varied across the climate zones.**

Interpretation **Our results highlight hate speech online as a potential channel through which temperature alters interpersonal conflict and societal aggression. We provide empirical evidence that hot and cold temperatures can aggravate aggressive tendencies online. The prevalence of the results across climatic and socioeconomic subgroups points to limitations in the ability of humans to adapt to temperature extremes.**

Funding **Volkswagen Foundation.**

Introduction

In the context of rapid anthropogenic climate change,[1] the question of how the climate influences human aggression, which dates back to the ancient world,[2] is more prominent than ever. Previous research on the link between climate change and aggression has identified three main pathways of influence.[3,4] First, the direct physical discomfort from hot temperatures causes violence and aggression, which has been shown by theoretical and experimental studies and summarised in several key reviews.[5–8] Second, climate change worsens socioeconomic conditions that have been found to be indicators for aggression-prone behaviour among adults, such as economic deprivation, food insecurity during childhood, and low educational attainment.[9–12] Third, changing climatic conditions and increasing extreme weather events enforce group-level aggressions—eg, through more frequent ecomigration events.[3,4] Empirical studies have found that temperature anomalies are associated with a higher risk for armed conflict in ethically fractionalised countries,[13] higher incidence of civil war in Africa,[14] and intergroup and interpersonal conflicts in Africa and the Middle East.[15–19]

However, with more than 60% of the world population using the internet,[20] aggression and violence can also spread in the digital environment. Online hate is a prevalent problem, with four of ten Americans having personally experienced online harassment,[21] and disproportionally affects groups that already have an increased risk of marginalisation. As of 2015, almost three of four women globally had been exposed to or had experienced some form of online violence.[22] 25% of African Americans and 10% of Hispanic Americans report being affected by online harassment due to race or ethnicity compared with 3% of White Americans.[23] Compared with their peers, LGBTQ teenagers are four times more likely to report online sexual harassment.[24] Within the last 5 years, online harassment has become more severe while simultaneously

Research in context

Evidence before this study

Multidisciplinary studies have found an effect of temperature on aggression with empirical analyses suggesting that deviations from mild temperatures increase aggressive tendencies and conflict risk worldwide. However, conflict is nowadays not limited to the physical space alone, but is also prevalent online in the form of hate speech. People affected by hate speech have been shown to be more likely to have mental health problems or to experience an aggravation of pre-existing conditions. We analysed the effect of temperature on the occurrence of hate speech on the social media platform Twitter in the USA between 2014 and 2020.

Added value of this study

To our knowledge, this is the first empirical study assessing the impact of temperature on online hate speech in the USA. The use of datasets from Twitter enabled the analysis of unprompted aggressions since Twitter users express their opinions online without external encouragement. Furthermore, users can tweet from any location, reducing the barrier to expressing aggression in response to temporal discomfort. In a sample of around 4 billion geolocated tweets, more than 75 million hate tweets were identified using a supervised machine learning classifier. The statistical analysis revealed a quasi-quadratic dependence of hate speech on temperature with low prevalence of hate speech observed in moderate temperatures and sharp increases in hate speech in warmer and colder temperatures. This quasi-quadratic relationship was preserved in separate analyses of temperature and hate speech in different climate zones and in the context of socioeconomic differences (income, religious adherence, and electoral outcomes). The lowest prevalence of hate speech was observed at temperatures centred around the local mean temperature and the magnitude of the increases in hate speech in hot and cold temperatures varied across climate zones.

Implications of all the available evidence

The quasi-quadratic relationship identified in this study shows that extreme temperatures lead to more aggression online. In contrast to the majority of quantitative studies assessing physical violence, this conclusion was the same for hot temperatures and cold temperatures. Daily maximum temperatures of more than 30°C were consistently associated with substantial increases in hate speech across all climate zones and across all socioeconomic subgroups. This persistent association suggests limits to the capacity for temperature adaptation since the increases in aggression persisted even in regions where hot temperatures are common and across socioeconomic groups that have economic means to mitigate uncomfortable temperatures. Overall, the results presented in this study highlight the importance of climate change mitigation and adaptation against temperature extremes and the need to effectively combat hate speech online.

becoming more normalised within society,[25] urging the UN to publish an official call for action to stop online hate speech, especially against minority groups.[26] At the beginning of the COVID-19 pandemic, an increase in online Sinophobia was observed.[27]

Targeted aggressions on social media have been shown to have mental health impacts,[28,29] such as heightened anxiety, depression, and self-harm.[30] Psychological research regarding the effects of online harassment suggests that individuals becoming victims of hate speech feel unsafe using online services in general, excluding them from online services and opportunities.[21] Teenagers who experience aggression online are more likely to have problems at school and to exhibit delinquent behaviour offline.[31] In addition to the direct adverse effects for the individual, hate speech online is also predictive of hate crimes in the offline environment.[32] Substantial evidence indicates that online hate speech, aggression, and harassment undermine the public good of equal dignity for all,[33] highlighting the need to identify potential drivers of hate speech to develop optimal strategies to combat it. The links between temperature and aggression and temperature and physical conflicts indicate that temperature might also be a potential driver for hate speech.

In this study, we statistically analysed the effect of temperature on hate speech, using a dataset including more than 4 billion tweets posted on the social media platform Twitter between 2014 and 2020 with distinct geolocation at the US city level. In 2019, a fifth of the US population used Twitter,[34] making it suitable for an analysis of US online discourse. This is also reflected by a growing body of literature that uses social media data for a variety of quantitative research in the context of climate change to which our study contributes.[35–38] Using a supervised machine learning approach, we identified more than 75 million hate tweets within the dataset, corresponding to around 2% of the total tweet volume. They were aggregated to the daily level and city level as this roughly corresponds to the life span of tweets and resolves the local climate experienced by users. We then evaluated within-city changes in hate tweets against local changes in daily maximum temperature using fixed-effects panel regression models to control for unobserved confounders.

Methods

Data sources

All climate data used in this analysis were obtained from the ERA5, the fifth generation European Centre for Medium-Range Weather Forecasts global climate and weather reanalysis dataset.[39] Our main climate variable of interest was daily maximum temperature since it reflects the hottest temperature reached on a specific day. The usage of daily mean temperature could mask temperature extremes experienced by Twitter users. Furthermore, daily maximum temperatures typically occur between noon and the late afternoon, which is a popular time for Twitter use (appendix p 9). We further included total precipitation (m), cloud cover (%), and wind speed on a daily level (m/s) between May 1, 2014 and May 1, 2020 as control variables. The data were used on a $0{\cdot}25° \times 0{\cdot}25°$ grid. Cities were interpreted as latitude–longitude points since in most cases, they were contained entirely in one grid cell. The daily city time series for each climate variable corresponded to the time series associated with the cell that the city was located in.

The raw data consisted of more than 4 billion tweets, spanning a timeframe from May 1, 2014 to May 1, 2020. Tweets were sampled from the 1% Twitter stream using a bounding box around the USA to extract geolocated tweets in the country. Around 1–2% of all tweets are geolocated; thus, although the stream contained all of the geolocated tweets in this period, the data only represent a small proportion of total tweet volume. For simplicity, we refer to this dataset as US1420 hereafter.

To examine the persistence of the relationship between temperature and online hate speech, we did separate analyses for different climate zones, income quartiles, forms of religious adherence, and 2016 election results. The US Department of Energy[40] distinguishes eight climate zones, five of which were included in this analysis; data coverage for the subarctic, mixed-dry, and very cold zones was too sparse to enable meaningful analysis. Per-capita income data at the county level were provided by the US Bureau of Economic Analysis.[41] Each city per-capita income was approximated by the average county per-capita income. We classified income into quartiles, ranging from low income (US$26 400 to ≤45 300), medium-low ($45 301 to ≤52 100), medium-high ($52 101 to ≤62 200), to high income ($62 201 to ≤194 000). The data on religious beliefs were obtained from the 2010 US census.[42] In this study, we only differentiated between Catholic and Evangelical beliefs since the data density was too sparse to conduct a meaningful assessment for other religious beliefs at the city level. Data on the 2016 election outcome on county level were obtained from Harvard University.[43]

Machine learning approach for the detection of hate tweets

The definition of what is considered as hateful is often unclear.[44] For this analysis, we adopted the UN Strategy and Plan of Action definition of hate speech: "any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor".[44] To identify tweets containing hate speech in the raw dataset, we used a machine learning approach (appendix pp 3–8). We followed a standard Natural Language Processing pipeline (appendix p 4).[45,46] Based on three previously published datasets containing labelled hate tweets (HAR and HATE datasets),[47–49] we assembled a dataset for training and testing that was separate from the US1420 dataset. The training and testing data contained examples of hate tweets and non-hate tweets. The training dataset was used to teach a classifier to assign tweets to so-called hate and no hate classes on the basis of attributed values of the tweets. This classifier was then applied to the testing dataset. Since the tweets in the testing dataset were also labelled, we could assess the performance of the classifier by checking how many labels were correct. The classifier was then applied to the distinct US1420 dataset and assigned each tweet in the US1420 data to hate and no hate classes, detecting around 75 million hate tweets, corresponding to around 2% of the sample. A full list of the different classifiers and performances is included in the appendix (p 6).

See Online for appendix

Hate speech data

All hate tweets in the sample were temporally aggregated at the daily level because this roughly corresponds to both the time span at which users consume social media and experience weather considering the circadian rhythm of body temperature.[50] For example, hot temperatures experienced in the day might have effects on hate speech occurrence in the evening since users might have more time for social media usage later in the day. Similarly, the discomfort of a hot night with little sleep has been identified as a factor that promotes irritability and aggression.[51] Thus, hot temperatures could plausibly lead to an increase in hateful content during the morning or over the course of the following day.

To enable a precise match to local temperature data, tweets were aggregated at the city level using TIGER/Line shapefiles (2020) provided by the US Census Bureau.[52] To ensure sufficient data coverage, only cities with more than 50 000 inhabitants were included. On the basis of this criteria, 773 cities were included, and their distribution across the USA covered different climate zones and socioeconomic compositions (appendix p 9). Additionally, tweets were also aggregated at the state level, which was subsequently used as a robustness control. Details on the state-level analysis are included in the appendix (pp 23–26).

The US1420 dataset encompasses 6 years (May, 2014 to the end of April, 2020), two of which were leap years (2016, 2020). Considering that daily data were available for 773 cities, this yielded a total of 1 694 416 possible observations. We counted 1 694 416 observations in the US hate speech Twitter data; therefore our data were complete and we have no concerns about data sparsity in our analysis.

Statistical analysis

We applied a binned fixed-effects panel-regression model to estimate the relationship between temperature and online hate speech. This approach used exogenous variation in local weather to identify the effect of daily maximum temperature on the amount of daily hate on Twitter, assuming that variation in temperature is random conditional on a set of fixed effects.[53–55] Specifically, we included city:year fixed effects ($\mu_{c,y}$), to exploit the temperature variation within one city within 1 year (ie, Los Angeles in 2015) to estimate the effect of temperature on hate speech. These city-year indicator variables flexibly control for local differences (eg, in administration) and larger time trends across the sample period, thus limiting omitted variable bias that is inherent to inter-city comparisons and enhancing the confidence of the causal inference of the results.[54,56] We purposefully did not statistically control for month of year because this has been shown to lead to mis-specifications of the temperature-aggression relationship.[7,57,58] Since the temperature variation within one city in 1 month is small, the temperature signal would be stunted. The main result is robust against the inclusion of city:month:year fixed effects as an alternative to city:year fixed effects with smaller effect sizes (appendix p 16). Instead, we added a dummy variable controlling for holidays (F_d) to limit the bias from temperature-independent seasonal influences. Holidays were approximated using the closing days of the New York Stock Exchange, which captures major US holidays.[59]

The independent variable, daily maximum temperature, was discretised into 3°C bins covering –30°C to 55°C. This semi-parametric approach allows for non-linearities without specifying a fixed functional form or making other assumptions about the data.[60–62] A dummy variable B_i was introduced for each bin where:

$$B_i(T_{c,d}) = \begin{cases} 1 \text{ if } T_{c,d} \in B_i \\ 0 \text{ else.} \end{cases}$$

$T_{c,d}$ describes the daily maximum temperature for city *c*. To assess robustness, we considered bin widths of 1°C and 5°C (appendix pp 18–19).

We applied the natural logarithm to the dependent variable, $H_{c,d}$, corresponding to the daily number of hate tweets per city. The regression coefficients were subsequently converted into percentages. In the main model, we dealt with zeros by regressing $log(H_{c,d}+1)$ because this enables better interpretability. As a robustness check, we used an inverse hyperbolic sine transformation on the data, which is approximately equal to the natural logarithm but well defined at zero (appendix p 17); the results of this robustness check were consistent with the main analysis.

To further isolate the temperature response of the occurrence of hate tweets, other climate variables that might impact human behaviour were included. Specifically, we controlled for daily total precipitation ($P_{c,d}$), cloud cover ($C_{c,d}$), and wind speed ($S_{c,d}$).

We also included controls for weekdays (W_d) since previous literature suggests that Twitter usage differs especially between weekdays and weekends.[63] The overall model thus reads as:

$$\log(H_{c,d}+1) = \sum_{i=0}^{b} \alpha_i B_i(T_{c,d}) + \beta P_{c,d} + \gamma C_{c,d} + \delta S_{c,d} + \zeta F_d + \eta W_d + \mu_{c,y} + \varepsilon_{c,d}.$$

The regression coefficients α_i, corresponding to the respective bins *i*, can be interpreted as the percentage change in hate speech in that bin. Percentage changes are in relation to the omitted bin, which corresponds to the minimum bin in all analyses (15–18°C for the main panel regression). In the principle figures, only coefficients whose corresponding 3°C temperature bins contained at least 1·5% of the data are shown. We used errors clustered at the state level in our analyses.

To assess the change in hate as a proportion of the total tweet volume, we computed the hate share as follows:

$$HS_{c,d} = \frac{H_{c,d}}{AT_{c,d}} \times 100;$$

where $AT_{c,d}$ denotes all tweets in the sample on day *d* in city *c*. The hate share was used to analyse the relationship between temperature and hate tweets while considering the general potential influences of weather on the tweet volume.

Role of the funding source
The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results
A strong non-linear relationship was identified between daily maximum temperature and the percentage change in hate tweets (figure 1). Fewest hate tweets occurred between temperatures of 15°C and 18°C. The number of hate tweets remained comparably low for the directly adjacent temperature bins, but sharply increased for temperatures warmer than 27°C and colder than 6°C. On cold days with maximum temperatures between –6°C and –3°C, the number of hate tweets was approximately 12·5% (95% CI 8·0–16·5) higher than on days in the 15 to 18°C temperature range, and on hot days (42 to 45°C), the number of hate tweets was more than 22·0% (20·5–23·5) higher than days in the 15 to 18°C temperature range. On average, a city's temperature varied across 12·6 bins per year. The changes for all bins and the effects of the control variables are shown in the appendix (p 15). For temperatures higher than 27°C and lower than 9°C, all respective regression coefficients were statistically significant at the highest level ($p \leq 0{\cdot}001$). The only bin that was not significant at any significance level ($p>0{\cdot}1$) was the 18 to 21°C bin, suggesting that the variation was below our detection level of uncertainty.

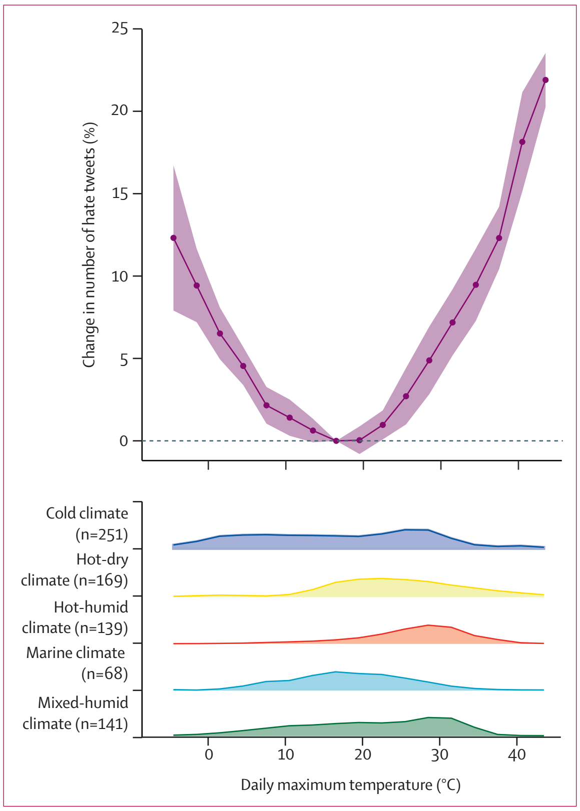

Figure 1: **Binned panel-regression model of the relationship between daily maximum temperature and the percentage change in number of hate tweets**
The percentage change in hate tweets shown is relative to the 15–18°C bin. Purple dots show the regression coefficient for each 3°C bin. Purple shaded areas denote 95% CIs. Errors were clustered at the state level. The marginal distributions show the mean percentage of days per bin for each city and year for five major climate zones across the USA. In the bottom panel, the numbers in parentheses on the x-axis indicate the number of cities included in each climate zone.

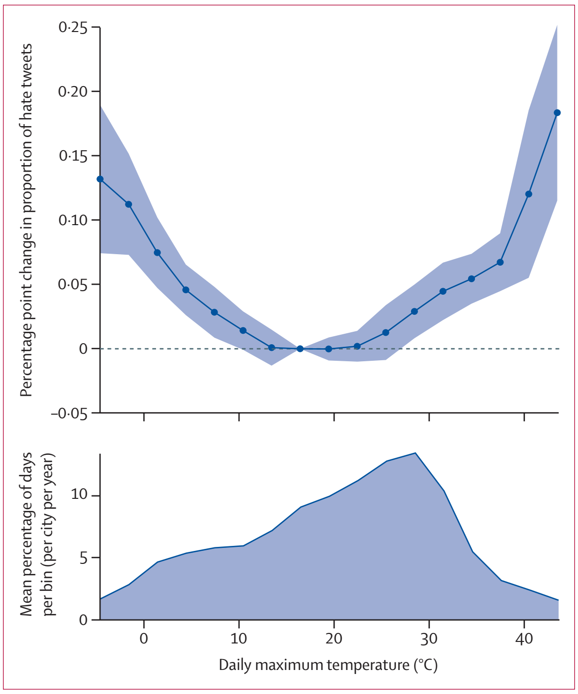

Figure 2: **Relationship between daily maximum temperature and the percentage point change in geolocated hate tweets as a proportion of all geolocated tweets at the city level**
In the omitted bin (15–18°C), the mean proportion of hate tweets at the city level amounted to around 1·5% of all tweets. A percentage point change of up to 0·13 percentage points for cold temperatures and up to 0·18 percentage points for hot temperatures therefore corresponds to increases of around 8·6% and 12·1% in number of hate tweets, respectively. This indicates that the hate response to extreme temperatures is not just a reflection of a general effect of temperature on tweet volume.

In addition to the number of hate tweets, we also used the follower-weighted number of hate tweets as the dependent variable to approximate the daily reach of hate speech. We weighed each hate tweet by the number of followers of its author. The shape of the curve was preserved, but the heat responses increased by up to 26·5% (95% CI 23·0–30·0) at high temperatures (42–45°C; appendix p 22).

Results of additional analyses using varying bin widths (1°C or 5°C) were consistent with the main findings (appendix pp 18–19).

In addition to the analysis for 773 US cities, we aggregated tweets to the state level and applied the same binned panel-regression approach. At this level of granularity, we identified a non-linear temperature–hate tweet relationship with the lowest prevalence of hate tweets observed at moderate temperatures and increases observed at temperatures lower than 9°C or higher than 18°C (appendix pp 23–26). Furthermore, since the COVID-19 pandemic triggered an increase in hate speech,[27] we also did a robustness check excluding tweets after the outbreak; we found the same non-linear shape and comparable increases in hate tweet occurrence for extreme temperatures as in the main analysis (appendix p 14).

To further examine the temperature–hate tweet relationship against the potential influence of weather on the overall tweet volume, we considered hate tweets as a proportion of all geolocated tweets in the sample. If the overall tweet volume is temperature-dependent such that warmer and colder temperatures result in a general increase in the number of tweets, it is possible that the proportion of hate tweets is almost constant. We tested for this by computing the proportion of geolocated hate tweets at the city level as a proportion of all US geolocated tweets in the respective city for each day. The binned panel-regression model used throughout the analysis was then applied with the daily hate tweet proportion as the dependent variable and daily maximum temperature and other weather controls as the independent variables; the resulting temperature–hate tweet response function (figure 2) had the same non-linear shape as observed in figure 1. An increase in tweets of around 0·13 percentage points (95% CI 0·07–0·19) was observed for cold temperatures and an increase of around 0·18 percentage points (0·12–0·25) for warm temperatures. The mean proportion of hate tweets on all days that fell within the omitted bin (15–18°C) amounted to around 1·5%. Thus, the percentage point increases observed translated to approximately 8·6% more hate tweets in cold temperatures and around 12·1% more hate tweets in warm temperatures. This result is evidence that not only the volume of hate speech on Twitter increases in more extreme temperatures, but also that the proportion of hate in all tweets rises. Analysis of state-level aggregation confirmed these results (appendix p 26). The remaining analyses were conducted for both the number of hate tweets (figures 3, 4) and the proportion of hate tweets (appendix pp 20–21) as the dependent variable.

Our dataset comprised cities with diverse climatic conditions. These differences in mean temperature and temperature variability mean that the panel analysis was to a larger extent informed by cities with higher temperature variability. The local temperature extremes might differ from the temperature extremes in the panel analysis. To investigate the potential impact of these local differences in climate on the overall temperature–hate speech relationship, we did separate analyses for five distinct climate zones (figure 3). The general quasi-quadratic shape of the US-wide response curve (figure 1) was preserved across all climate zones for which there was sufficient data coverage (figure 3A). The temperature-dependent minimum number of hate tweets and the strength of the increase in hate tweets for warm and cold temperatures differed in accordance with the individual temperature distribution of the climate zone. In the cold climate zone, which spans most of the north of the contiguous USA, a broad range of temperatures were observed annually. Accordingly, the percentage change in hate tweets was low between 6°C and 24°C. The maximum increase in hate tweets of 17·5% (95% CI 5·5 to 29·0) compared with the omitted bin (15 to 18°C) was observed between temperatures of 39°C and 42°C, which only occur rarely in this climate zone. This is likely to explain the large confidence interval. Temperatures between 24°C and 33°C were more common in this climate zone. For the 30°C to 33°C temperature bin, hate tweets increased by around 7·0% (4·5 to 9·5). At cold temperatures (–6 to –3°C), hate tweets increased by more than 12·0% (7·5 to 17·0). A similar pattern was observed in the hot-dry and mixed-humid zones across a smaller temperature range. For the mixed-humid climate zone, the maximum increase amounted to more than 11·0% (8·0 to 14·0) on cold days (0 to 3°C) and around 9·0% (6·5 to 11·0) on warm days (33 to 36°C) relative to the omitted bin (21 to 24°C). 3°C temperature bins outside this range contained less than 1·5% of days. For the hot-dry climate zone, the increase in hate tweets in hot temperatures (42°C to 45°C) was most pronounced (almost a 24% increase [22·5 to 25·5]). In cold temperatures (0 to 3°C), the number of hate tweets increased by 10·0% (–0·5 to 20·0) relative to the omitted bin (18 to 21°C); however, temperatures lower than 9°C were fairly uncommon in this zone and tweets only increased by 8·0% (7·0 to 8·5) in the 9 to 12°C bin. For the marine climate zone, the increase in tweets in response to hot temperatures was stronger than the increase in response to cold temperatures, reaching up to 11·0% (7·5 to 15·0) at temperatures higher than 33°C. For the hot-humid climate zone, the curve had a V-shape with sharp increases on both sides of the omitted bin (24 to 27°C), with a 10·5% increase (4·5 to 16·5) in hate tweets in cold temperatures (6 to 9°C) and a 15·0% (13·0 to 17·0) increase in hot temperatures (36 to 39°C).

For all climate zones, we observed that the lowest incidence of hate tweets (omitted bin) coincided with the mean temperature across the time period (figure 3A). This could suggest that the hate tweet increases are dependent on temperatures we are used to. However, independent of the local mean temperature and distributions, hot temperatures of more than 30°C led to significant increases in hate tweets of at least 7%, pointing to potential limits in adaptation to hot temperatures.

When using the proportion of hate tweets as the dependent variable, the shape of the curves were largely preserved (appendix p 20).

Previous research suggests that economic factors influence the occurrence of hate speech.[64] Other studies point to partisan differences in acceptance of hate speech and its usage by politicians[65,66] or discuss religion not only as a target but also as a source for different types of hate speech.[67] Therefore, it is possible that these characteristics affect the here identified temperature–hate speech relationship. To address this, we did separate analyses for different income quartiles, religious beliefs, and across 2016 voting patterns.

Across all socioeconomic subgroups considered, the relationship between daily maximum temperature and hate tweets had a distinct and robust non-linear shape, which is in its general form independent of economic, religious, and political differences. Independent of income quartiles, increases in the number of hate tweets at cold and warm temperature extremes fell generally between 10% and 15%. The only exception was that the medium–low income group experienced higher maximum temperatures than the other groups and and increases in hate tweets of up to 22·0% (95% CI 20·5–23·5) were observed in this temperature range (42 to 45°C; figure 4). For all income quartiles, the mean temperature coincided approximately with the minimum bin. The non-linear shape of the temperature–hate speech relationship was also preserved when data were categorised into cities with predominantly Catholic or Evangelical religious beliefs (figure 4) with similar responses observed for extreme temperatures. The observed relationship also holds independently of the 2016 election results (figure 4). The increase in hate tweets in colder temperatures was slightly more pronounced in cities belonging to counties that had a Democratic majority in the 2016 election, reaching up to 13·5% (7·0–20·5) at temperatures between –6°C and

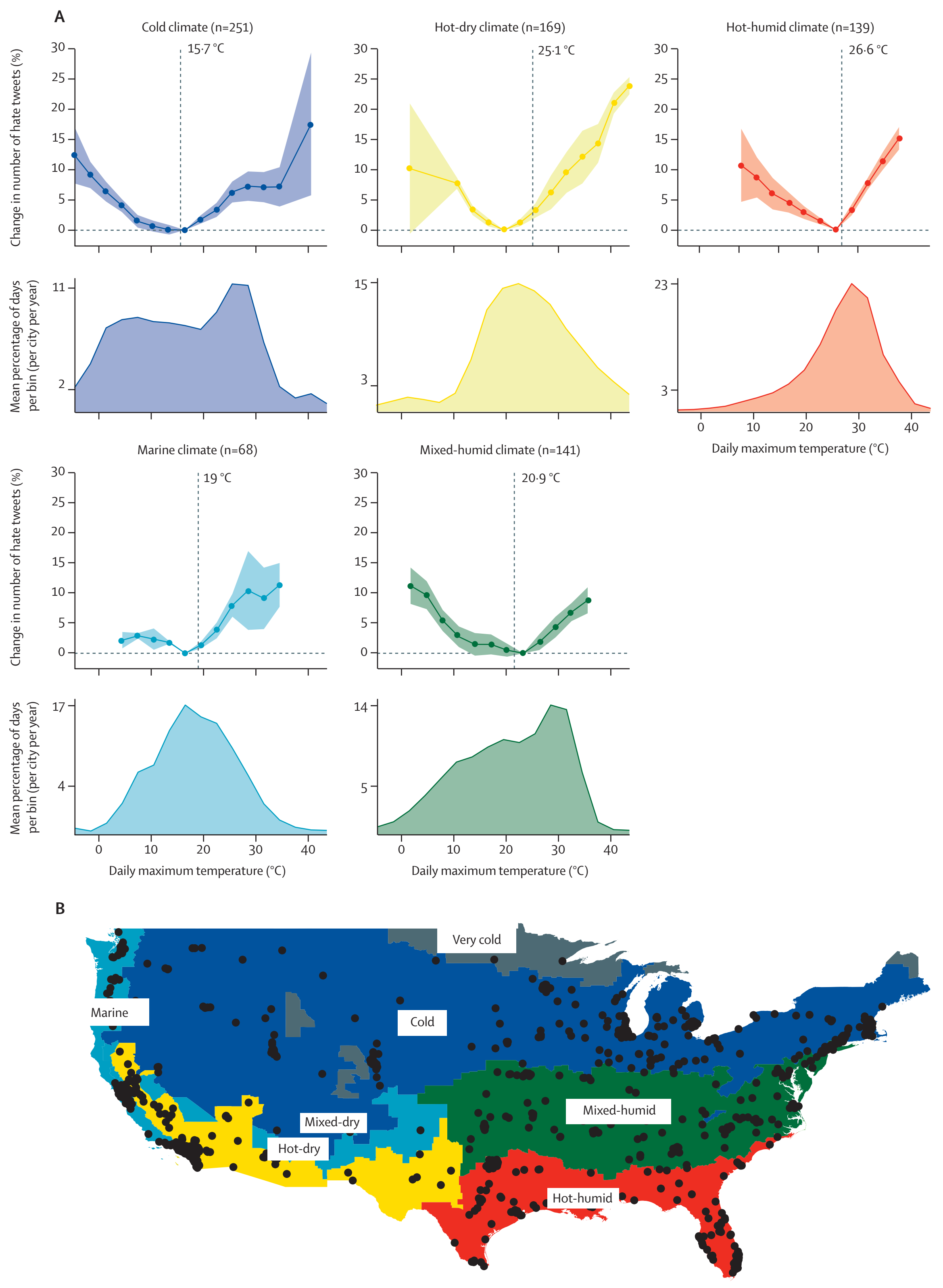

Figure 3: **Relationship between temperature and hate tweets across five climatic zones in the USA**
(A) Relationship between daily maximum temperature (x-axis) and the percentage change in hate tweets relative to the minimum bin (y-axis) for 773 US cities distributed across five climate zones as assessed with a binned panel-regression approach. Dots represent the respective regression coefficients for each bin and the shaded areas denote 95% CIs. Errors were clustered at the state level. For all climate zones, a similar U-shape can be observed where there is enough data with low values for moderate temperatures and high values for cold and warm temperatures. Dotted vertical lines indicate the mean daily maximum temperature across the sample period. The marginal distributions (x-axes) show the mean percentage of days per bin for each city and year for the respective climate zone. The numbers in parentheses show the number of cities included in each climate zone. (B) Geospatial distribution of climate zones across the USA (with the exception of Alaska and Hawaii); black dots show the 773 cities included in the analysis.

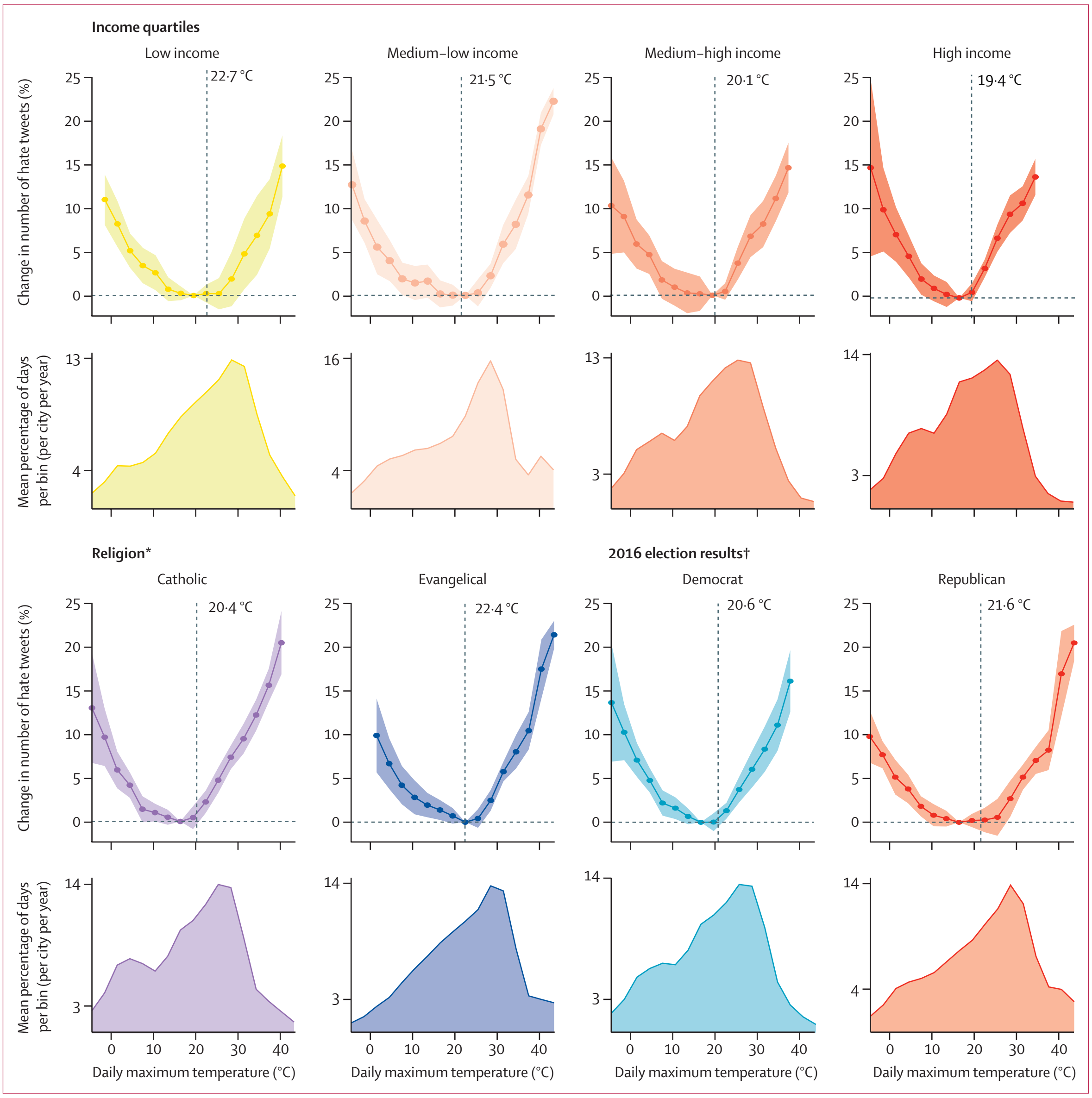

Figure 4: **The relationship between daily maximum temperature and hate tweets by socioeconomic factors**
Dots show regression coefficients for each 3°C bin. Shaded areas show 95% CIs. Errors were clustered at state level. The marginal distributions (x-axes) show the mean percentage of days per bin for each city and year for the respective climate zone. Dotted vertical lines indicate the mean daily maximum temperature across the sample period. A similar U-shape can be observed for all socioeconomic subgroups with low values for daily maximum temperatures in the range of 15–27°C and high values for temperatures outside of this range. *The data on religious beliefs were obtained from the 2010 US census.[42] †Data on the 2016 election outcome at the county level were obtained from Harvard University.[43]

–3°C and only around 10·0% (7·0–13·0) in cities belonging to counties that had a Republican majority in the 2016 election. By contrast, the maximum increase in hate tweets in hot temperatures amounted to around 16·0% (12·5–19·5) in cities with a Democratic majority (at temperatures of 36 to 39°C) and 20·5% (18·5–22·5) in cities with a Republican majority (at temperatures of 42 to 45°C). Although the mean temperatures were within 1°C of each other, Republican majority cities had more extremely hot days, which is likely to explain the differing heat response. The analysis using the proportion of hate tweets as the dependent variable is included in the appendix (p 21). Overall, the results were preserved for the hate-share analysis.

In addition to the analysis of socioeconomic differences, we analysed the temperature–hate tweet response by administrative unit of US census divisions, which are frequently used for data collection and analysis (appendix pp 27–28). All temperature–hate tweet response curves at the census division level had the non-linear shape observed across all analyses with increases for cold days falling between 5% and 30% and increases for hot days reaching between 6% and 30%, depending on the census division.

Discussion

In this study, we found that hate speech increased in absolute volume, and also as a proportion of total tweeting activity, at temperature extremes. The quasi-quadratic shape of the temperature–hate tweet curve was robust across varying climate zones, income quartiles, and religious and political beliefs.

We restricted our analysis to geolocated tweets, since geographical information is necessary to match the tweets with climate and socioeconomic data for further analysis. Geolocation is an opt-in feature for the user, meaning that users have to specifically enable location services and choose to geolocate their individual tweet. We cannot assume that geolocated users necessarily represent all Twitter users,[68] which could be a limitation of this study. Studies have shown that female users are more likely to enable geolocation than males,[69,70] individuals of Asian or Latino ethnicity are more likely to enable geotagging than other ethnic groups,[70] the average user enabling location services is 0·55 years older than users opting out,[69] and iPhone users are slightly more likely to geotag than Android users.[71] However, there is no indication that these biases are significantly correlated with temperature and geolocated Twitter data have successfully been used in combination with climate variables in a number of studies.[35,37,72,73]

We only conducted the analysis for English tweets, which could present a further limitation considering the linguistic plurality of the USA. Assessing the language composition of the raw US1420 data, we found that 93% of tweets were in the English language (appendix pp 12–13) with the remainder of tweets distributed across a number of languages. The data density was therefore too low in other languages to conduct a separate analysis. Additionally, data from the American Community survey[74] showed that in each US state, at least 90% of the people questioned speak English only or English to a high standard.

Although in general it is impossible to accurately classify all hate speech due to its many incarnations, subtlety,[75] sarcasm, and context dependence, the robustness checks we conducted on the quality of the dataset suggest that the data present a representative sample of hateful discourse on Twitter. However, some of the expressions included in tweets have a different connotation based on the cultural context in which they are used. Specifically, some proportion of the tweets classified as hate contain the N-word with the spelling variant ending in “a” which has, in contrast to the spelling with “er”, according to some sources been reappropriated as a type of endearment in some communities.[76] However, the use of the word and its variants remains highly controversial. In the examples observed in our data (appendix p 7), the context is typically aggressive and derogatory. However, we cannot be sure that all instances in the dataset containing this particular slur are genuinely hateful. Furthermore, bot accounts, which have been estimated to contribute up to 10% of hateful content in some datasets comprising hate tweets,[77] might further bias our data. However, bot accounts are likely to only cause random errors rather than systematic errors in data since their activity is not temperature dependent, which suggests that while bot accounts introduce noise in the dataset, the bias is likely to be small.

Daily maximum temperature might not always match the temperature experienced by the Twitter user due to residential heating or cooling. This potential measurement error in the independent variables is, however, more likely to attenuate than to increase the magnitude of our estimates.

The results of this study provide evidence that extreme temperatures are associated with more hate speech on the social media platform Twitter. The quasi-quadratic relationship identified empirically confirms the hypothesis formulated by Anderson and colleagues,[8] which stated that uncomfortably hot and cold temperatures increase aggressive tendencies. Furthermore, the findings are supported by the general aggression model formulated by Anderson and Bushman in 2002.[7] Instead of using violent behaviour to assess temperature effects, the usage of hate speech allows the assessment of verbal aggressions. The results are also consistent with the relationship found between temperature and cyber racism in Europe.[78]

The analysis of different climate zones shows that the general relationship between temperature and hate speech was maintained across different climates although the temperature range with the lowest occurrence of hate speech in each climate zone shifted slightly in accordance with local temperature conditions. This could indicate that the response is dependent on the temperatures individuals are used to. However, since daily maximum temperatures of more than 30°C were consistently associated with substantial increases in hate speech, there are likely to be limits to the capacity for temperature adaptation. Additionally, the observation of similar effects and effect sizes across income quartiles could indicate further limits in adaptation: even among individuals in the highest income quartile, who are likely to be able to spend money on heat mitigation strategies such as air conditioning or more comfortable transportation, an increase in hate speech was observed on hot days. The similar responses observed across religious beliefs and political preference indicate that increased aggressive tendencies in hot and cold weather are not a question of mindset but subject to a more universal temperature influence. However, a limitation of our analysis was that the groupings based on income, religion, and partisan were not perfect since cities are never perfectly homogeneous. For example, in a city with predominantly Democratic voters as of 2016, all racist tweets could originate from a Republican minority or vice versa. To ameliorate potential inferential issues, individual-level measures of socioeconomic categories would be needed. In addition to the subgroups considered in the analysis, analyses of the differences between city and rural regions could provide further insights; however, this was not possible in this study due to sparsity of data in rural areas. With the progression of rapid, anthropogenic climate change, extreme weather such as heatwaves and cold spells will become more frequent.[79] In the USA, population-weighted heat exposure in metropolitan regions is projected to rise by 12–30 times by the end of the century under a high emissions scenario compared with population-weighted heat exposure at the start of the century.[80] Population-weighted cold exposure is projected to be 1·3–2·2 times larger than that at the start of the century. Assuming little adaptation and similar communication patterns, this would mean that hate expressed online could increase under future global warming. Calculations for temperature shifts and cyber racism in Europe show that the number of days hotter than a locally comfortable climate, weighted by the strength of the relationship between temperature and cyber racism, increases by around 50% in some parts of Europe.[78]

Both aggression and climate change have been found to have negative impacts on mental health. Climate change-induced risks for mental health[81] include increased climate anxiety,[82] a greater risk for depression especially in young people,[83,84] and increased suicide risks.[73,85] Hate speech has been shown to cause heightened anxiety, depression, and self-harm and a feeling of unsafety in online spaces.[28–30] Our results contribute to this literature by identifying the effect of temperature on hate speech as a new impact channel through which climate change could affect aggression and mental health.

Overall, the results presented in this study highlight not only the importance of climate change mitigation and adaptation against temperature extremes, but also the need to effectively combat hate speech online and to provide resources for people who are affected. Further work is needed to understand the nature of online abuse, to analyse the most prevalent types of hate, which topics it relates to, who is targeted, and who authors it.

Contributors

AS and LW designed the study. AS processed the climate and Twitter data. All authors contributed to the interpretation and presentation of the results. AS wrote the manuscript with contributions from LW. AS and LW revised the manuscript.

Declaration of interests

We declare no competing interests.

Data sharing

In compliance with Twitter terms of service restrictions and due to privacy concerns, the US1420 dataset will not be shared in a public repository. The HAR dataset used to train the classifier is available on request to the corresponding author. The HATE dataset is available online. The tweet IDs of the dataset are also available online. All climate data used in this analysis originated from the ERA-5 re-analysis dataset, which can be downloaded from the Climate Data Store. The mapping of US climate zones is available from the US Office of Energy Efficiency and Renewable Energy. The data underlying the socioeconomic subgroups are publicly available and sources are included in the Methods section. Scripts will be made publicly available immediately after publication.

For the **HATE dataset** see https://github.com/t-davidson/hate-speech-and-offensive-language

For **tweet IDs of the dataset** see https://github.com/ZeerakW/hatespeech

For the **Climate Data Store** see https://cds.climate.copernicus.eu/#!/home

For the **mapping of US climate zones** see https://www.energy.gov/eere/buildings/building-america-climate-specific-guidance

Acknowledgments

This study was funded by the Volkswagen Foundation. We thank Kelsey Barton-Henry for fruitful discussion.

References

1 Intergovernmental Panel on Climate Change. Climate change 2013: the physical science. https://www.ipcc.ch/report/ar5/wg1/ (accessed May 14, 2020).

2 Hopkins D, Martindale C, Vance N, et al. The Oxford history of classical reception in English literature. Oxford: Oxford University Press, 2020.

3 Miles-Novelo A, Anderson CA. Climate change and psychology: effects of rapid global warming on violence and aggression. *Curr Clim Change Rep* 2019; **5:** 36–46.

4 Plante C, Allen JJ, Anderson CA. Effects of rapid climate change on violence and conflict. *ORE Climate Science* 2017; published online April 26. https://doi.org/10.1093/acrefore/9780190228620.013.344.

5 Anderson CA. Temperature and aggression: effects on quarterly, yearly, and city rates of violent and nonviolent crime. *J Pers Soc Psychol* 1987; **52:** 1161–73.

6 Anderson CA. Temperature and aggression: ubiquitous effects of heat on occurrence of human violence. *Psychol Bull* 1989; **106:** 74–96.

7 Anderson CA, Bushman BJ. Human aggression. *Annu Rev Psychol* 2002; **53:** 27–51.

8 Anderson CA, Anderson KB, Dorr N, DeNeve KM, Flanagan M. Temperature and aggression. *Adv Exp Soc Psychol* 2000; **32:** 63–133.

9 Intergovernmental Panel on Climate Change. Climate change 2022: impacts, adaptation, and vulnerability. https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii (accessed Aug 2, 2022).

10 Carleton TA, Hsiang SM. Social and economic impacts of climate. *Science* 2016; **353:** aad9837.

11 Kotz M, Wenz L, Stechemesser A, Kalkuhl M, Levermann A. Day-to-day temperature variability reduces economic growth. *Nat Clim Chang* 2021; **11:** 319–25.

12 Kalkuhl M, Wenz L. The impact of climate conditions on economic production. Evidence from a global panel of regions. *J Environ Econ Manage* 2020; **103:** 102360.

13 Schleussner C-F, Donges JF, Donner RV, Schellnhuber HJ. Armed-conflict risks enhanced by climate-related disasters in ethnically fractionalized countries. *Proc Natl Acad Sci USA* 2016; **113:** 9216–21.

14 Burke MB, Miguel E, Satyanath S, Dykema JA, Lobell DB. Warming increases the risk of civil war in Africa. *Proc Natl Acad Sci USA* 2009; **106:** 20670–74.

15 Helman D, Zaitchik BF. Temperature anomalies affect violent conflicts in African and Middle Eastern warm regions. *Glob Environ Change* 2020; **63:** 102118.

16 O'Loughlin J, Witmer FDW, Linke AM, Laing A, Gettelman A, Dudhia J. Climate variability and conflict risk in east Africa, 1990–2009. *Proc Natl Acad Sci USA* 2012; **109:** 18344–49.

17 Burke M, Hsiang SM, Miguel E. Climate and conflict. *Annu Rev Econ* 2015; **7:** 577–617.

18 Ash K, Obradovich N. Climatic stress, internal migration, and Syrian civil war onset. *J Conflict Resolut* 2020; **64:** 3–31.

19 Gleick PH. Water, drought, climate change, and conflict in Syria. *Weather Clim Soc* 2014; **6:** 331–40.

20 Internet World Stats. Internet usage statistics. https://www.internetworldstats.com/stats.htm (accessed Oct 29, 2020).

21 Pew Research Center. Online harassment 2017. https://www.pewresearch.org/internet/2017/07/11/online-harassment-2017/ (accessed Oct 27, 2020).

22 UN Women. Urgent action needed to combat online violence against women and girls, says new UN report. Sept 24, 2015. https://www.unwomen.org/news/stories/2015/9/cyber-violence-report-press-release (accessed Feb 11, 2021).

23 Pew Research Center. 1 in 4 black Americans have faced online harassment because of their race, ethnicity. https://www.pewresearch.org/fact-tank/2017/07/25/1-in-4-black-americans-have-faced-online-harassment-because-of-their-race-or-ethnicity/ (accessed Oct 27, 2020).

24 GLSEN. Out online: the experiences of LGBT youth on the internet. https://www.glsen.org/news/out-online-experiences-lgbt-youth-internet (accessed Jan 15, 2021).

25 Pew Research Center. What we learned about online harassment in 2020 and how it has changed. https://www.pewresearch.org/fact-tank/2021/01/13/qa-what-weve-learned-about-online-harassment/ (accessed Jan 15, 2021).

26 Office of the United Nations High Commissioner for Human Rights. Hate speech: UN experts make joint call for action by states and social media firms. https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=25037&LangID=E (accessed Feb 11, 2021).

27 Stechemesser A, Wenz L, Levermann A. Corona crisis fuels racially profiled hate in social media networks. *EClinicalMedicine* 2020; **23:** 100372.

28 McCready AM, Rowan-Kenyon HT, Barone NI, Alemán AMM. Students of color, mental health, and racialized aggressions on social media. *J Stud Aff Res Pract* 2021; **58:** 179–95.

29 Stewart-Tufescu A, Salmon S, Taillieu T, Fortier J, Afifi TO. Victimization experiences and mental health outcomes among grades 7 to 12 students in Manitoba, Canada. *Int J Bullying Prev* 2021; **3:** 1–12.

30 Skilbred-Fjeld S, Reme SE, Mossige S. Cyberbullying involvement and mental health problems among late adolescents. *Cyberpsychology* 2020; published online Feb 21. https://doi.org/10.5817/CP2020-1-5.

31 Hinduja S, Patchin JW. Offline consequences of online victimization. *J Sch Violence* 2007; **6:** 89–112.

32 Williams ML, Burnap P, Javed A, Liu H, Ozalp S. Corrigendum to: Hate in the machine: anti-black and anti-Muslim social media posts as predictors of offline racially and religiously aggravated crime. *Br J Criminol* 2020; **60:** 242.

33 Waldron J. The harm in hate speech. Cambridge, MA: Harvard University Press, 2012.

34 Pew Research Center. 10 facts about Americans and Twitter. https://www.pewresearch.org/fact-tank/2019/08/02/10-facts-about-americans-and-twitter/ (accessed Oct 28, 2020).

35 Moore FC, Obradovich N, Lehner F, Baylis P. Rapidly declining remarkability of temperature anomalies may obscure public perception of climate change. *Proc Natl Acad Sci USA* 2019; **116:** 4905–10.

36 Wang J, Obradovich N, Zheng SA. 43-million-person investigation into weather and expressed sentiment in a changing climate. *One Earth* 2020; **2:** 568–77.

37 Baylis P. Temperature and temperament: evidence from Twitter. *J Public Econ* 2020; **184:** 104161.

38 Moore FC, Obradovich N. Using remarkability to define coastal flooding thresholds. *Nat Commun* 2020; **11:** 530.

39 Dee DP, Uppala SM, Simmons AJ, et al. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 2011; **137:** 553–97.

40 US Department of Energy. Building America best practices series: volume 7.3. Guide to determining climate regions by county. https://www.energy.gov/sites/prod/files/2015/10/f27/ba_climate_region_guide_7.3.pdf (accessed Aug 2, 2022).

41 US Bureau of Economic Analysis. Personal income by county, metro, and other areas. https://www.bea.gov/data/income-saving/personal-income-county-metro-and-other-areas (accessed Feb 11, 2021).

42 The Association of Religion Data Archives. U.S. religion census: religious congregations and membership study, 2010 (county file). https://www.thearda.com/Archive/Files/Downloads/RCMSCY10_DL.asp (accessed Nov 3, 2020).

43 Pettigrew S. November 2016 general election results (county-level). https://doi.org/10.7910/DVN/MLLQDH (accessed Aug 2, 2022).

44 UN. United Nations strategy and plan of action on hate speech. 2019. https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_EN.pdf (accessed March 24, 2021).

45 Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction, 2nd edn. New York, NY: Springer-Verlag, 2009.

46 Shalev-Shwartz S, Ben-David S. Understanding machine learning: from theory to algorithms. Cambridge, MA: Cambridge University Press, 2014.

47 Golbeck J, Ashktorab Z, Banjo RO, et al. A large labeled corpus for online harassment research. Proceedings of the 2017 ACM on Web Science Conference. June 25, 2017. https://doi.org/10.1145/3091478.3091509 (accessed Aug 2, 2022).

48 Waseem Z, Hovy D. Hateful symbols or hateful people? Predictive features for hate speech detection on Twitter. https://aclanthology.org/N16-2013.pdf (accessed Aug 2, 2022).

49 Davidson T, Warmsley D, Macy M, Weber I. Automated hate speech detection and the problem of offensive language. *arXiv* 2017; published online March 11. http://arxiv.org/abs/1703.04009 (preprint).

50 Refinetti R, Menaker M. The circadian rhythm of body temperature. *Physiol Behav* 1992; **51:** 613–37.

51 Kamphuis J, Meerlo P, Koolhaas JM, Lancel M. Poor sleep as a potential causal factor in aggression and violence. *Sleep Med* 2012; **13:** 327–34.

52 United Stated Census Bureau. TIGER/line shapefiles. https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html (accessed Oct 13, 2020).

53 Auffhammer M. Quantifying economic damages from climate change. *J Econ Perspect* 2018; **32:** 33–52.

54 Hsiang SM. Climate econometrics. Cambridge, MA: National Bureau of Economic Research, 2016. https://doi.org/10.3386/w22181 (accessed Aug 2, 2022).

55 Dell M, Jones BF, Olken BA. What do we learn from the weather? The new climate-economy literature. Cambridge, MA: National Bureau of Economic Research, 2013. https://doi.org/10.3386/w19578 (accessed Aug 2, 2022).

56 Kolstad CD, Moore FC. Estimating the economic impacts of climate change using weather observations. *Rev Environ Econ Policy* 2020; **14:** 1–24.

57 Bushman BJ, Wang MC, Anderson CA. Is the curve relating temperature to aggression linear or curvilinear? Assaults and temperature in Minneapolis reexamined. *J Pers Soc Psychol* 2005; **89:** 62–66.

58 Bushman BJ, Wang MC, Anderson CA. Is the curve relating temperature to aggression linear or curvilinear? A response to Bell (2005) and to Cohn and Rotton (2005). *J Pers Soc Psychol* 2005; **89:** 74–77.

59 NYSE. Holidays and trading hours. https://www.nyse.com/markets/hours-calendars (accessed May 20, 2022).

60 Deschênes O, Greenstone M. Climate change, mortality, and adaptation: evidence from annual fluctuations in weather in the US. *Am Econ J Appl Econ* 2011; **3:** 152–85.

61 Deryugina T, Hsiang SM. Does the environment still matter? Daily Temperature and income in the United States. Cambridge, MA: National Bureau of Economic Research, 2014. https://doi.org/10.3386/w20750 (accessed Aug 2, 2022).

62 Zhang J, Jiang H, Liu G, Zeng W. A study on the contribution of industrial restructuring to reduction of carbon emissions in China during the five Five-Year Plan periods. *J Clean Prod* 2018; **176:** 629–35.

63 Longley PA, Adnan M, Lansley G. The geotemporal demographics of Twitter usage. *Environ Plan A* 2015; **47:** 465–84.

64 Denti D, Faggian A. Where do angry birds tweet? Income inequality and online hate in Italy. *Camb J Regions Econ Soc* 2021; **14:** 483–506.

65 Zannettou S, Elsherief M, Belding E, Nilizadeh S, Stringhini G. Measuring and characterizing hate speech on news websites. *arXiv* 2020; published online May 16. https://doi.org/10.48550/arXiv.2005.07926 (preprint).

66 Armstrong GM, Wronski J. Framing hate: moral foundations, party cues, and (in)tolerance of offensive speech. *J Soc Polit Psych* 2019; **7:** 695–725.

67 Moon R. Putting faith in hate: when religion is the source or target of hate speech. Cambridge, MA: Cambridge University Press, 2018. https://doi.org/10.1017/9781108348423 (accessed Aug 2,2022).

68 Karami A, Kadari RR, Panati L, Nooli SP, Bheemreddy H, Bozorgi P. Analysis of geotagging behavior: do geotagged users represent the Twitter population? *ISPRS Int J Geoinf* 2021; **10:** 373.

69 Sloan L, Morgan J. Who tweets with their location? Understanding the relationship between demographic characteristics and the use of geoservices and geotagging on Twitter. *PLoS One* 2015; **10:** e0142209.

70 Wood-Doughty Z, Smith M, Broniatowski DA, Dredze M. How does Twitter user behavior vary across demographic groups? https://aclanthology.org/W17-2912/ (accessed Aug 2, 2022).

71 Huang B, Carley KM. A large-scale empirical study of geotagging behavior on Twitter. https://doi.org/10.1145/3341161.3342870 (accessed Aug 2, 2022).

72 Baylis P, Obradovich N, Kryvasheyeu Y, et al. Weather impacts expressed sentiment. *PLoS One* 2018; **13:** e0195750.

73 Burke M, González F, Baylis P, et al. Higher temperatures increase suicide rates in the United States and Mexico. *Nat Clim Chang* 2018; **8:** 723–29.

74 United States Census Bureau. American Community Survey. https://data.census.gov/cedsci/table?t=Language%20Spoken%20at%20Home&g=0100000US%240400000&tid=ACSST5Y2020.S1601 (accessed May 5, 2022).

75 Nielsen LB. Subtle, pervasive, harmful: racist and sexist remarks in public as hate speech. *J Soc Issues* 2002; **58:** 265–80.

76 Smith HL. Has nigga been reappropriated as a term of endearment? *Am Speech* 2019; **94:** 420–77.

77 Ziems C, He B, Soni S, Kumar S. Racism is a virus: anti-Asian hate and counterhate in social media during the COVID-19 crisis. *arXiv* 2020; published online May 25. http://arxiv.org/abs/2005.12423 (preprint).

78 Stechemesser A, Wenz L, Kotz M, Levermann A. Strong increase of racist tweets outside of climate comfort zone in Europe. *Environ Res Lett* 2021; **16:** 114001.

79 Intergovernmental Panel on Climate Change. Climate change 2021: the physical science basis. https://www.ipcc.ch/report/ar6/wg1/ (accessed Aug 2, 2022).

80 Broadbent AM, Krayenhoff ES, Georgescu M. The motley drivers of heat and cold exposure in 21st century US cities. *Proc Natl Acad Sci USA* 2020; **117:** 21108–17.

81 Obradovich N, Migliorini R, Paulus MP, Rahwan I. Empirical evidence of mental health risks posed by climate change. *Proc Natl Acad Sci USA* 2018; **115:** 10953–58.

82 Ingle HE, Mikulewicz M. Mental health and climate change: tackling invisible injustice. *Lancet Planet Health* 2020; **4:** e128–30.

83 Wu J, Snell G, Samji H. Climate anxiety in young people: a call to action. *Lancet Planet Health* 2020; **4:** e435–36.

84 Majeed H, Lee J. The impact of climate change on youth depression and mental health. *Lancet Planet Health* 2017; **1:** e94–95.

85 Carleton TA. Crop-damaging temperatures increase suicide rates in India. *Proc Natl Acad Sci USA* 2017; **114:** 8746–51.