

Societal Attention to Heat Waves Can Indicate Public Health Impacts

EKATERINA BOGDANOVICH¹, LARS GUENTHER², MARKUS REICHSTEIN³, DOROTHEA FRANK⁴,
GEORG RUHRMANN⁵, ALEXANDER BRENNING⁶, JASPER M. C. DENISSEN⁷, AND RENÉ ORTH⁸

¹ Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Jena, Germany

² Institute for Journalism and Communication Studies, University of Hamburg, Hamburg, Germany

³ Institute for Communication Science, Friedrich Schiller University Jena, Jena, Germany

⁴ Department of Geography, Friedrich Schiller University Jena, Jena, Germany

(Manuscript received 31 December 2022, in final form 4 April 2023, accepted 19 May 2023)

ABSTRACT: Both the frequency and intensity of hot temperature extremes are expected to increase in the coming decades, challenging various socioeconomic sectors including public health. Therefore, societal attention data available in real time, such as Google search attention, could help monitor heat-wave impacts in domains with lagged data availability. Here, we jointly analyze societal attention and health impacts of heat waves in Germany at weekly time scales. We find that Google search attention responds similarly to hot temperatures as indicators of public health impacts, represented by excess mortality and hospitalizations. This emerges from piecewise linear relationships of Google search attention to and health impacts of temperature. We can then determine temperature thresholds above which both attention and public health are affected by heat. More generally, given the clear and similar response of societal indicators to heat, we conclude that heat waves can and should be defined from a joint societal and meteorological perspective, whereby temperatures are compared with thresholds established using societal data. A better joint understanding of societal attention and health impacts offers the potential to better manage future heat waves.

KEYWORDS: Social science; Heat wave; Health

1. Introduction

Heat waves have severe consequences for society and public health and are even referred to as “silent killers” (Bhattacharya 2003). While they probably cause less property damage than do hurricanes or floods, they tend to cause more fatalities than other types of extreme weather events (Hughes et al. 2016). For instance, the European heat wave in 2003 took more than 70 000 lives (Robine et al. 2008). Deaths are typically caused by hyperthermia, dehydration, respiratory disease, cerebrovascular disease, or heat stroke (Hajat et al. 2010). Some people are more vulnerable in this respect than others, for example, populations in urban areas or with preexisting health problems, but also children and elderly people (Gabriel and Endlicher 2011; Lindley et al. 2019). Climate change will very likely increase the frequency, intensity, and duration of heat waves (Meehl and Tebaldi 2004; IPCC 2021), highlighting the need to better understand the respective societal responses to hot temperatures.


Recently, digital data are facilitating the assessment of the societal response to various types of events; for instance, online search frequencies or social media data can serve as proxies for socioeconomic impacts (Maurer and Holbach 2016; Ripberger 2011; Scharnow and Vogelgesang 2011). This becomes more and more feasible since online news and information are becoming increasingly important; in 2017, 74% of

internet users in Germany (16–74 yr) read news on the internet (Eurostat 2019). The news about disasters attracts high public interest; therefore, there is a tendency among media to cover the most severe events (Yan and Bissell 2018). Previous studies have already analyzed multiple aspects of societal attention related to hydrometeorological extreme events. They found that attention as assessed by Google search frequencies or communication on Twitter scales with the magnitude of an event and/or distance from it (Li et al. 2018; Shelton et al. 2014; Kryvasheyev et al. 2016; De Albuquerque et al. 2015). In particular, attention to heat waves increases with rising temperatures (Singh et al. 2018; Jung et al. 2019; Jung and Uejio 2017), especially in urban areas (Grasso et al. 2017), where heat stress is typically higher because of the urban heat island effect (Tuholske et al. 2021). Online search frequencies were also shown to be an instrument to detect heat stress (Jung et al. 2019; Li et al. 2016).

In this study, we aim to explore the potential of societal attention to heat-wave conditions for (i) detecting and defining heat waves, which is traditionally done based on meteorological variables only (Seneviratne et al. 2012) while the joint consideration of societal data offers the opportunity for refined, domain-based definitions, and (ii) analyzing the potential of societal attention to predict actual heat-wave impacts, for example, on people’s health. For this purpose, we investigate the societal response to temperature using multiple temperature and societal metrics.

2. Data and methods

Our analysis focuses on Germany, where comprehensive societal and meteorological data are available and where heat waves have repeatedly had impacts on society. Mortality during heat waves in 2003 and 2015 in Germany was elevated by

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Ekaterina Bogdanovich, ebogdan@bgc-jena.mpg.de

over 70% (Heudorf and Schade 2014; Muthers et al. 2017). Our study period is 2010–19. While this is constrained by the concurrent availability of the analyzed data streams, this period includes several documented heat waves: 2010, 2013, 2015, 2018, and 2019 (An der Heiden et al. 2020; Matzarakis et al. 2020). We focus on the five warmest months in each year, from May to September. All variables are aggregated to the weekly time scale to match the limited temporal resolution of some of the employed datasets; week start is on Monday. All data sources and data periods can be found in Table A1 in appendix A.

a. Hydrometeorological variables

To determine the most relevant temperature variable for societal attention, we consider daily maximum, minimum, and average temperature data from the ERA5 reanalysis dataset (Hersbach et al. 2020). The data are gridded with a $0.5^\circ \times 0.5^\circ$ spatial resolution. We chose to focus on these variables because maximum temperature is considered in the heat-wave definition of the WMO (World Meteorological Organization 1992); minimum temperature represents nighttime temperatures, which can affect people's quality of sleep; and average temperatures may jointly capture the heat impacts of daytime and nighttime values.

As an additional variable, we consider apparent temperature (AT) as a widely used heat index, which relates more closely to how humans perceive temperature extremes (National Weather Service 2021) and is commonly used in heat warning systems. The calculation of AT is based on maximum temperature, relative humidity, and several coefficients, which describe the resistance of the human body to heat [the equation and the full list of coefficients can be found in Rothfus (1990, 1–2)]. The relative humidity in this context is obtained from the ERA5 dataset as well.

To match the limited temporal resolution of some of the employed datasets, all daily temperature variables are aggregated to the weekly time scale by computing the average, maximum, and minimum values of each variable. The name of a temperature variable consists of two parts: daily temperature (Tmean, Tmax, or Tmin) and weekly aggregation (mean, max, or min). For example, Tmean_max is the weekly maximum of daily average temperature (the abbreviations can be found in Table A2 in appendix A).

Last, we obtain country-level averages of these variables by averaging them across all respective grid cells, while using a weighting reflecting their population in 2010, which is derived from the United Nations–estimated grid-level population density (Center for International Earth Science Information Network 2018) (Fig. 1a).

b. Societal variables

We use Google search data from Google Trends to assess societal attention. Google is the most popular search engine in Germany and accounts for more than 90% of the German search engine market (Fig. A1c in appendix A), and about 75% of the German population are internet users (Fig. A1a in appendix A). People who have no access to the internet or do not use Google for their online searches are not included in

this study. We obtain daily time series of search interest for the topics “heat wave” and “heat stroke,” which also encompass similar phrases in any language (Google News Initiative 2022). Here, we use English language search topics to obtain the data for our analyses. Google Trends data represent the search interest for any topic as a fraction of the total number of searches at the specific time and location. The data are thereby adjusted for the increase in search volume over the study period. Further, the data are scaled between 0 and 100, where 100 is the maximum search interest in the particular topic (Google News Initiative 2022; Rogers 2016). Hereinafter, we refer to this Google search interest as societal attention. The daily Google search attention data are aggregated to a weekly time scale by calculating average weekly value (Figs. 1b,c). The summers of 2018 and 2019 were exceptionally warm in Germany (Xu et al. 2020), which caused a peak in attention.

It has been shown that mass media have the potential to affect societal attention to a particular topic; this impact is often described as agenda setting (McCombs and Shaw 1972; Shaw and Martin 1992). Therefore, mass media could potentially trigger the attention of online searches, in addition to the meteorological conditions. To study this relationship, we consider media attention via press data, which is defined as the number of articles mentioning heat waves during the study period. In particular, we use the German search term “(Deutschland) AND ((Hitzewelle) OR (extreme Hitze))” in the databases Factiva (<http://www.factiva.com/>) and WiSo (<https://www.wiso-net.de/>), to collect the number of heat wave mentions from three leading German newspapers (*Die Welt*, *Die Zeit*, and *Süddeutsche Zeitung*). The largest German newspaper, *Bild*, was not included since data were not available for the entire study period. The online versions of these newspapers are also considered as an additional variable. In total, there are 170 heat-wave mentions in print newspapers and 275 mentions in the respective online newspapers for the 2010–19 study period (Figs. 1d,e).

To comprehensively study the societal response to heat waves, we consider variables related to human health (Fig. 1f,g): the weekly number of heat-affected hospitalizations, provided by the Federal Office of Statistics (Statistisches Bundesamt), and weekly mortality data, derived from Eurostat. To identify which age group is the largest contributor to total mortality, we consider mortality by various age groups: less than 20, 20–39, 40–59, 60–79, and over 80 yr old. The raw mortality data are detrended by subtracting the long-term linear trend. Excess mortality is calculated from the detrended data by subtracting the mean seasonal cycle.

c. Statistical analysis

To understand which meteorological variables are most strongly related to the Google search attention, we employ a random forest analysis. Therein, we predict the Google Trends time series using 11 temperature variables (a list of variables and their abbreviations can be found in Table A2 in appendix A) and determine their respective importance. This is done by comparing the Shapley additive explanations (SHAP) feature importances, which are calculated as the sum of the absolute SHAP values over the entire sample (Molnar 2020). SHAP

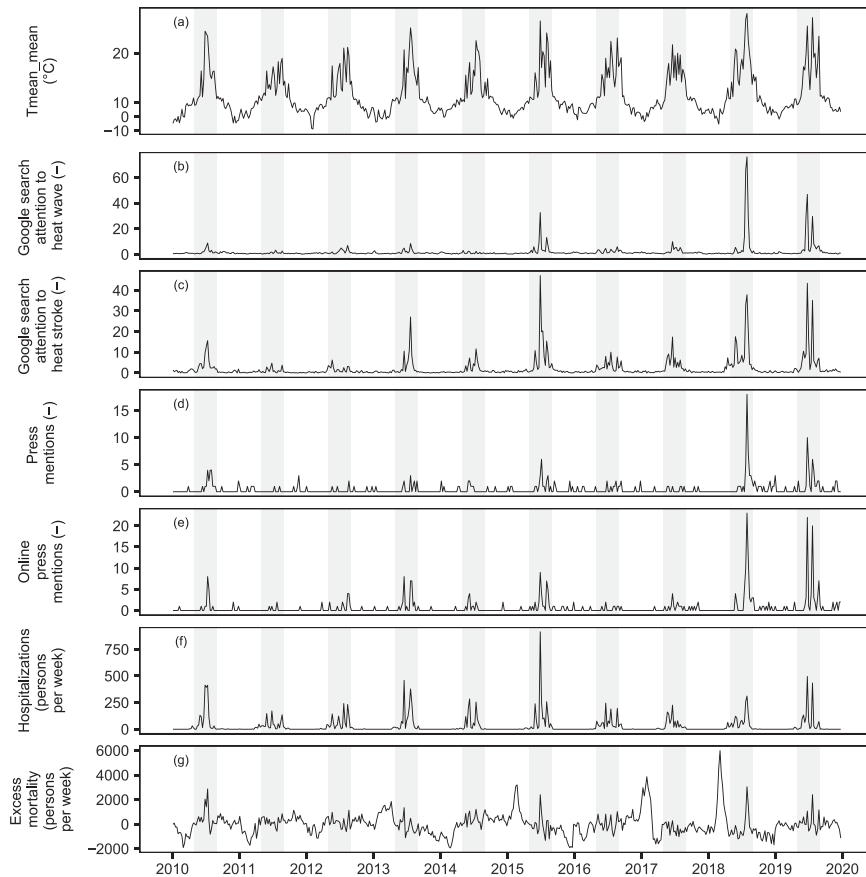


FIG. 1. Time series of the considered temperature and societal variables: (a) weekly averages of daily mean temperature, Google search attention to (b) *heat wave* and (c) *heat stroke*, press mentions of *heat wave* in (d) print and (e) online media, (f) heat-affected hospitalizations, and (g) human excess mortality. In contrast to hospitalizations and excess mortality, Google search attention is only an index that is proportional to the number of respective searches but does not provide the actual number of searches.

values assess for an individual predicted value of how much of it can be attributed to each predictor variable. The presence of high correlation between predictors may challenge the identification of the importance of variables (Dormann et al. 2013; Gregorutti et al. 2017); however, SHAP values have been shown to be robust and highly interpretable (Lundberg and Lee 2017). Additionally, to understand the relationship between Google search attention and other societal impact metrics, we perform a random forest analysis with SHAP feature importance calculation, where we predict the Google search attention using the data from other societal metrics. Furthermore, we test the role of legacy effects on the evolution of Google search attention in another random forest analysis, in which we use the three most important temperature variables as determined in the first analysis as explanatory variables, together with the attention of the previous week, average Google search attention of the two to three weeks before, and average Google search attention of the four to six weeks before. The random forest model and SHAP values calculation are repeated 100 times to infer the uncertainties in feature importances.

To assess potential temperature-related thresholds in the response of Google search attention to temperature, we employ piecewise linear regression. We only infer a threshold if the adjusted R^2 of the piecewise regression exceeds 0.4, which indicates that the two-piece regression model can substantially explain the attention–temperature relationship. The threshold is then defined as a breaking point between two linear models fitted to the data (Fig. 4a). The calculations are done with the temperature variable with the highest explanatory power derived from the first random forest analysis. To infer the threshold uncertainties, we apply bootstrapping. We resample the years and repeat the calculation of the threshold 500 times. This threshold analysis is done for each societal variable separately.

Finally, we analyze seasonal variations in the sensitivity of Google search attention to temperature for the three summer months June, July, and August. In contrast to the previous analyses, we disregard May and September here, because a sufficient number of heat waves in each month is needed for a meaningful assessment of the sensitivity. To determine the sensitivity, we (i) apply linear regression to the attention–temperature

relationships based on the data from all June months, all July months, and all August months, respectively. Thereby, (ii) we consider a fixed temperature range between the maximum temperature and 8° below that. This 8° interval provides a sufficient number of data points for the linear regression and ensures that only the warmest data points are considered. Then, (iii) we assess the slopes of these models, which indicate the sensitivity of attention to temperature. Last, (iv) we consider longer-term legacy effects (in addition to the analysis in Fig. B5 in appendix B) by distinguishing between warm and cold springs, that is, using only months from 5 yr with cold and warm springs, respectively. For this purpose, we calculate the average temperature of March, April, and May, using Tmax_mean as this is the temperature variable with the highest impact derived from the first random forest analysis. We do not use a random forest model here, because the number of data is not sufficient since we focus on individual months of the year.

3. Results and discussion

a. Relation of Google search attention to temperature variables and other societal metrics

Our results reveal that for heat-related Google search attention in Germany, the weekly average of daily maximum temperature (Tmax_mean) has the highest importance among the eleven considered temperature variables (Fig. 2a and Fig. B1a in appendix B). The relevance of the daily maximum temperatures might be related to the impact of heat stress during the daytime (Koppe et al. 2004). The similarly high importance of the weekly average of daily average temperature (Tmean_mean) indicates that besides daytime temperatures, nighttime temperatures also play a role because they can influence the recovery and sleep quality and, therefore, the vulnerability to daytime heat stress (Minor et al. 2022). Nevertheless, in general, daytime temperatures (Tmax) are slightly more important than nighttime temperatures (Tmin) in the importance ranking. The effect of minimum temperature is only indirect and detectable in Tmean, which includes nighttime temperatures. Google search attention to *heat stroke* is also related to Tmean rather than to minimum temperature alone (Fig. B2a). The effect of minimum temperature itself (Tmin_max) is stronger in the case of excess mortality (Fig. B3d). This is in line with previous research that points out the societal relevance of so-called tropical nights, in which the temperature does not fall below 20°C (Beckmann et al. 2021; Royé et al. 2021; Murage et al. 2017). Among the societal variables, media attention via (online) press mentions is most closely related to Google search attention (Fig. 2b and Fig. B1b in appendix B). This confirms previous findings regarding the presence of a strong relationship between media and public attention (McCombs and Shaw 1972). While slightly less prominent than media attention, health-related variables are also relevant predictors of Google search attention. Interestingly, considering different search terms in this context, we find varying relationships of Google search attention to other societal response indicators. In comparison with Google search attention to *heat wave*, we find that Google search attention to *heat stroke* is more related

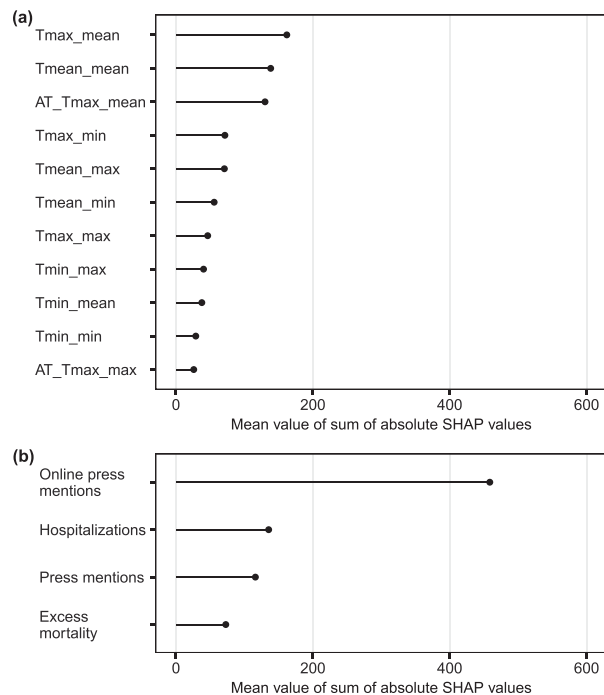


FIG. 2. Importance of different variables for predicting Google search attention to *heat wave* in Germany in a random forest model. The variable importance was derived using Shapley Additive explanations (SHAP) values. Predictor variables are (a) temperature variables and (b) societal variables. The first part of the temperature variable names denotes the type of considered daily temperature; the last part describes the type of aggregation to a weekly time scale (the abbreviations can be found in Table A2 of appendix A).

to health-related variables (Fig. B2 in appendix B). This suggests that the Google searches for these different terms are motivated differently; *heat-wave* searches may more often be driven by curiosity, whereas *heat stroke* searches may more often be associated with physical problems experienced or observed by an individual. The relationship between internet search queries and physical vulnerability was also demonstrated by Li et al. (2016). They showed that the search for heat stroke is an even better predictor of hospitalization in China than the temperature.

Overall, the relatively strong link between Google search attention and the other considered societal variables indicates that Google search attention is an informative proxy for aspects of the overall societal response, particularly in places where other data streams may be difficult to obtain. Next, we focus in more detail on the relationship between Google search attention and media attention in the case of heat waves. The random forest analysis can pinpoint relevant variables, but it cannot disentangle the effect of meteorological conditions from respective media coverage in driving Google search attention, such that the causal triggers of the search interest remain uncertain (Maurer and Holbach 2016; Ripberger 2011). To address this problem, we compare the typical temporal evolution of the considered variables during heat-wave

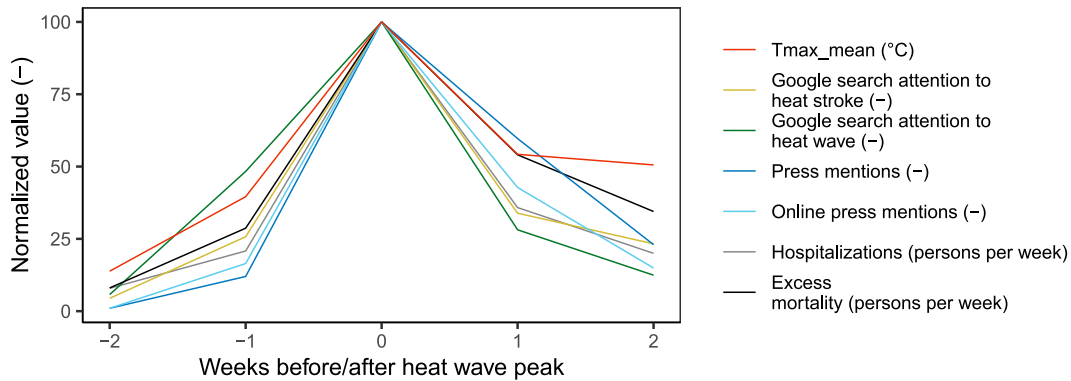


FIG. 3. Evolution of temperature and considered societal metrics during heat waves. Values are normalized for comparability and averaged across 10 events that are the hottest events in each year.

periods. Here, and for the remaining analyses, we use T_{\max_mean} as the temperature variable to describe heat waves as it is most closely related to Google search attention to *heat wave*. T_{\max_mean} plays a role in predicting other societal metrics as well (Fig. B3 in appendix B). Societal attention and impact variables peak synchronously with temperature during heat waves (Fig. 3). Prior to this, Google search attention responds more quickly to rising temperatures than press attention, indicating that Google search attention is probably more directly triggered by the heat-wave conditions than by respective media coverage. In addition, the Google search attention response is more in line with actual health impacts such as hospitalizations or elevated mortality. After the peak of the heat wave, health-related impacts decrease relatively quickly, while the heat waves still receive some attention in Google searches and even more in the press. While it has been shown that internet search queries can be triggered by news coverage (Granka 2010), our results indicate that press attention is slightly lagging rather than leading the Google searches in the case of heat waves. In fact, media attention is a result of the journalist's individual interpretation and assessment of the importance of events such as heat waves (Shoemaker et al. 2001), and Google search attention can actually reflect such perceived relevance. Since official heat-wave warnings and their reporting by the media can potentially drive Google search attention, we employ an additional analysis to test the role of heat-wave warnings in this context. For this purpose, we extend the analysis of the temporal evolution of temperature and societal variables to also include the number of heat warnings issued by the German National Weather Service (Matzarakis et al. 2020) (Fig. B4 in appendix B). The results show that societal interest typically rises before the issuing of heat warnings, and that this is largely independent of the spatial area of Germany for which heat warnings are issued.

Furthermore, we also test the role of legacy effects in the evolution of Google search attention (Fig. B5 in appendix B). To do so, we compare the relevance of the three most important temperature variables determined in Fig. 2a with Google search attention in previous weeks. We find that Google search attention of previous weeks is less important than concurrent temperatures, indicating that legacy effects do not play a major role.

b. Identifying temperature thresholds

When relating Google search attention to T_{\max_mean} , we find an exponential curve that can be approximated by a piecewise linear relationship (Fig. 4a). We verify the piecewise linear character of this relationship by comparing it with an ordinary single-slope linear regression and find a much higher explained fraction of variance (adjusted R^2 is 0.77 vs 0.23). A piecewise linear model also offers the opportunity to infer a threshold temperature at which the attention–temperature relationship changes significantly; we find 28.6°C to be the breaking point between the two linear models (Fig. 4a). Google search attention also shows a threshold behavior in relation to the other considered temperature variables. The threshold values are shown in Fig. B6 in appendix B.

This threshold analysis is done separately for each societal variable (Fig. 4b). The thresholds vary from 24.6° and 2.0°C for the median and interquartile range, respectively, for the hospitalizations to 28.6° and 1.7°C for the press mentions of heat wave. The thresholds for Google search attention are more in agreement with the thresholds for press mentions, corresponding to their relationship demonstrated by the random forest analysis (Fig. 2b).

Interestingly, we find generally lower temperature thresholds, indicating higher sensitivity to heat, for the health-related indicators than for the interest-related indicators. This may be due to the fact that different groups of people are sampled with the different indicators; people sampled with health-related societal metrics may be older or/and more vulnerable to heat relative to the group of Google search users. In particular, age groups 60–79 and over 80 yr old are the main contributors to the total mortality during our study period (Fig. A2 in appendix A), which supports that elderly people are more sensitive to heat stress (Koppe et al. 2004; Romanello et al. 2021). Previous studies have shown that thresholds can also vary across different regions. For instance, Grasso et al. (2017) found a higher temperature threshold in social media activity for southern Italian cities in comparison with northern ones. Karwat and Franzke (2021) demonstrated higher temperature thresholds in mortality for Spanish cities in comparison with German and French ones. In this study, we

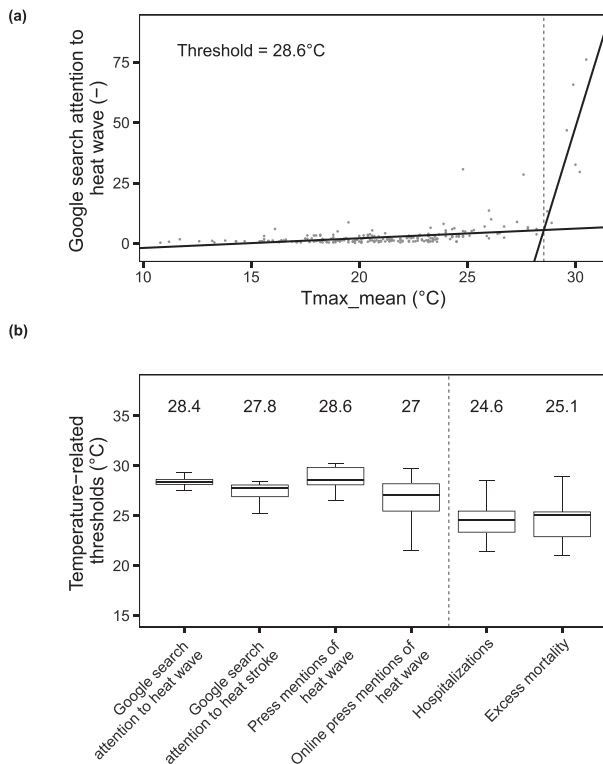


FIG. 4. Temperature thresholds above which an amplified heat-wave response is found. (a) Example analysis of the relationship of Google search attention to *heat wave* to weekly mean of daily maximum temperatures. The vertical dashed line indicates the determined threshold temperature. (b) Temperature thresholds determined with all considered societal metrics. Boxes show medians and uncertainty after bootstrapping the data (25th and 75th percentile). The dashed line separates media and health-related variables.

analyze data for Germany at the country level and do not distinguish between regions. Further analysis is needed to understand the spatial variability of temperature thresholds.

Moving beyond long-term mean relationships as in the previous analyses, we assess seasonal variations in the attention–temperature relationship. This is done with a different methodology based on linear regression slopes of attention–temperature relationships using exclusively data from specific months of the year, because there are too few data for inferring piecewise regressions as used above. We focus here on June, July, and August while including the spring months by differentiating between cold and warm spring conditions. The analysis is done separately for years with warm and cold springs. The results show variations in the attention–temperature relationship across months, with a higher Google search attention sensitivity to temperature in June than in July and August (Fig. B7 in appendix B), especially after cold springs. While the Google search interest is not strongly related to previous weeks (Fig. B5 in appendix B), it seems that temperature during the previous months preconditions the heat-wave response. People are probably not adapted to high temperatures after the cold months, especially if these were colder than usual. The further adaptation to hot temperatures is expressed by a decreasing slope in

relationships between temperature and Google search attention in July and August. This is confirmed by an increase in mortality during the first heat wave in summer (Anderson and Bell 2011). Further, Singh et al. (2018) also showed greater sensitivity of Google search frequencies to temperature in India if the temperature difference between winter and summer was higher.

Our findings have to be seen in light of some potential limitations. First, Google algorithms are not transparent and change through time (Nuti et al. 2014; Lazer et al. 2014). However, Google Trends have been employed previously and have been shown to be a useful tool for understanding societal attention to health-related hazards (Singh et al. 2018; Kam et al. 2019). Moreover, we find similar thresholds from Google search attention as for other societal metrics, which also underlines its usefulness in studying societal responses to hot extremes.

Second, elderly people and children, who are most vulnerable to heat waves (Romanello et al. 2021), may be underrepresented in the Google Trends data. However, the number of elderly people who are using the internet is increasing. In 2019, 80% of people in the 60–69 age group and 52% in the 70+ age group were internet users (Fig. A1). Moreover, their (younger) relatives and friends could search on Google for them to get information about heat waves and help them to reduce the health effects of heat (e.g., heat stroke). About 30% of Europeans over 65 yr old do not use the internet, but this number is decreasing (Fig. A1b in appendix A). However, about 85% of people aged between 16 and 74 yr are active internet users in Germany (Eurostat 2022), and Google is the most popular search engine in Germany, which accounts for over 90% of the search engine market (Fig. A1c in appendix A), so our data should still be sufficiently representative. Future research could follow up on our study by considering the health status and socioeconomic background of the analyzed population.

4. Conclusions

Overall, we find that (i) Google search attention is an informative proxy for aspects of the societal response to heat waves as assessed from various societal indicators, (ii) search attention is affected by both day- and nighttime temperatures, (iii) the relationship with temperature can be modeled piecewise linear with emerging temperature thresholds above which Google search attention increases strongly, (iv) lower thresholds are found for health-related indicators when compared with interest-related indicators, meaning that the vulnerability increases earlier than the attention, and (v) the sensitivity of Google search attention to temperature is higher in early summer than in late summer. Therefore, we acknowledge that Google search attention might also be confounded to some extent by human behavior and not exclusively driven by meteorological conditions.

The clear threshold behavior of the societal response to (hot) temperatures calls for further investigations to better understand the space–time variability of this relationship and the involved thresholds. The approach developed and presented in this study could be applied across space–time scales to identify their impact on the resulting heat-wave characterization. Further, our results can help to translate and interpret ordinary

weather forecast information to take regional vulnerabilities into account. In this context, taking into consideration temperature thresholds derived with past societal response data can improve heat warnings. Moreover, Google search attention is available in real time in contrast to health-related metrics, and for these reasons, Google search attention can be a useful early warning indicator.

Overall, we find that in Germany, societally relevant heat-wave conditions are characterized by mean daily temperature above 18°–22° depending on the considered variable. This suggests that heat waves can and should be determined also from a societal perspective. Further, we find that Google search attention contains valuable information for predicting actual heat-wave impacts on public health as evidenced by the similar time evolution during heat waves (Fig. 3) and similar underlying temperature thresholds (Fig. 4). This way, societal attention data in our study allow us to conclude that early warnings of heat waves do not manage to timely raise people's attention, which otherwise would be reflected in lower temperature thresholds as compared with health-related variables. In general, such enhanced understanding of the bulk societal response to heat waves will be increasingly valuable in the future as hot extremes are amplifying in the context of climate change.

Acknowledgments. We thank the German Federal Office of Statistics (Statistisches Bundesamt–Destatis) for providing hospitalization data, Ulrich Weber for obtaining

and processing ERA5 data, and Kelley De Polt for providing Google Trends data. We acknowledge data on heat warnings for Germany that were obtained online (https://opendata.dwd.de/climate_environment/health/historical_alerts/heat_warnings/), with kind advice from Stefan Muthers from the German Weather Service. Author Bogdanovich acknowledges support from the International Max Planck Research School for Global Biogeochemical Cycles. Author Orth is funded by the German Research Foundation (Emmy Noether Grant 391059971). Further, this work was supported by funding from the MYRIAD-EU project from the European Union's Horizon 2020 research and innovation programme under Grant 101003276.

Data availability statement. The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

APPENDIX A

Data Information

Table A1 provides the data sources and periods, and temperature variable names and abbreviations are listed in Table A2. Figure A1 shows internet usage in Germany, and Fig. A2 shows the percentage of total mortality by age group.

TABLE A1. Data sources.

Data	Source	Reference	Time period
Google search attention to <i>heat wave</i>	Google Trends	https://trends.google.com/trends/?geo=DE	2010–19
Google search attention to <i>heat stroke</i>	Google Trends	https://trends.google.com/trends/?geo=DE	2010–19
Press mentions	Factiva, WiSo	http://www.factiva.com/ ; https://www.wiso-net.de/dosearch	2010–19
Online press mentions	Factiva, WiSo	http://www.factiva.com/ ; https://www.wiso-net.de/dosearch	2010–19
Hospitalizations	Federal Office of Statistics	https://www.destatis.de/DE/Home/_inhalt.html	2010–19
Mortality	Deaths by week and sex, Eurostat	https://ec.europa.eu/eurostat/databrowser/view/DEMO_R_MWK_TS/default/table?lang=en	2010–19
Temperature	ERA5	Hersbach et al. (2020)	2010–19
Relative humidity	ERA5	Hersbach et al. (2020)	2010–19

TABLE A2. Temperature abbreviations.

Abbreviation	Full name
Tmax_max	Weekly max of daily max temperature, °C
Tmax_mean	Weekly avg of daily max temperature, °C
Tmax_min	Weekly min of daily max temperature, °C
Tmean_max	Weekly max of daily avg temperature, °C
Tmean_mean	Weekly avg of daily avg temperature, °C
Tmean_min	Weekly min of daily avg temperature, °C
Tmin_max	Weekly max of daily min temperature, °C
Tmin_mean	Weekly avg of daily min temperature, °C
Tmin_min	Weekly min of daily min temperature, °C
AT_Tmax_max	Apparent temperature weekly values, calculated based on Tmax_max (°C)
AT_Tmax_mean	Apparent temperature weekly values, calculated based on Tmax_mean (°C)

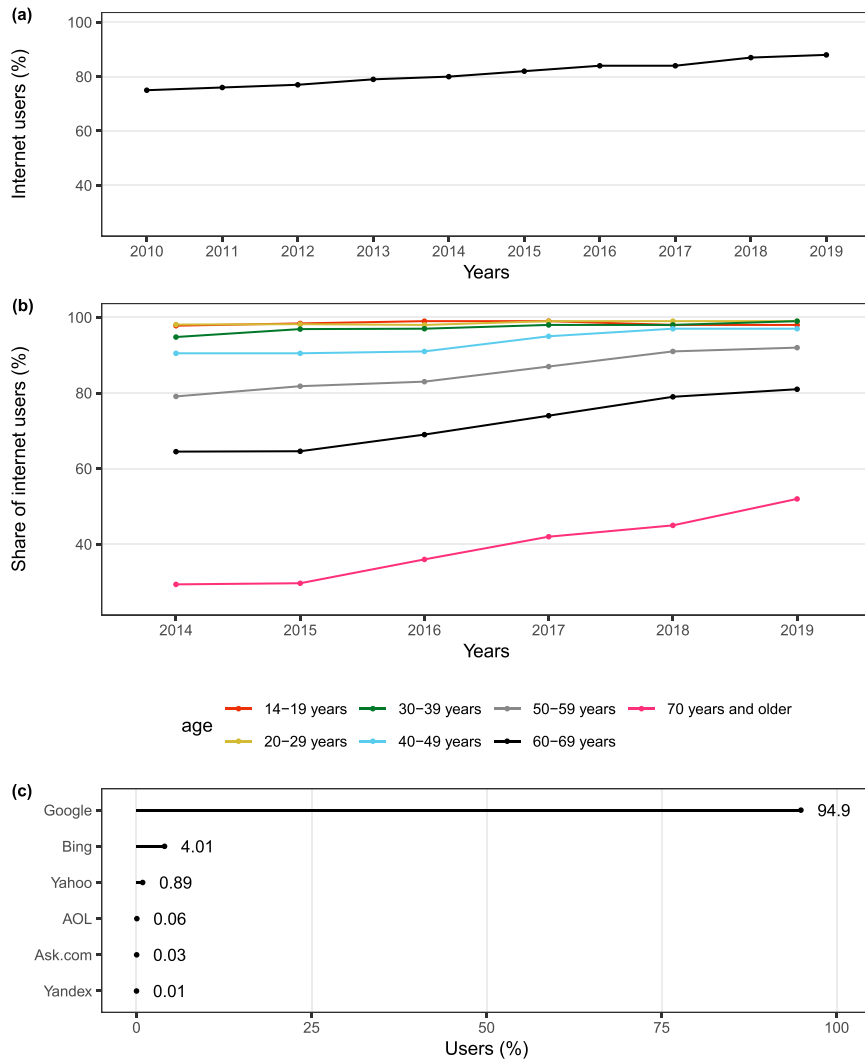


FIG. A1. Internet usage in Germany: (a) percent of internet users in Germany during the period 2010–19 (14 yr and older; [Statista 2021b](#)), (b) share of internet users by age group ([Statista 2021a](#)), and (c) share of search engine market in Germany in 2018 ([Statista 2018](#)).

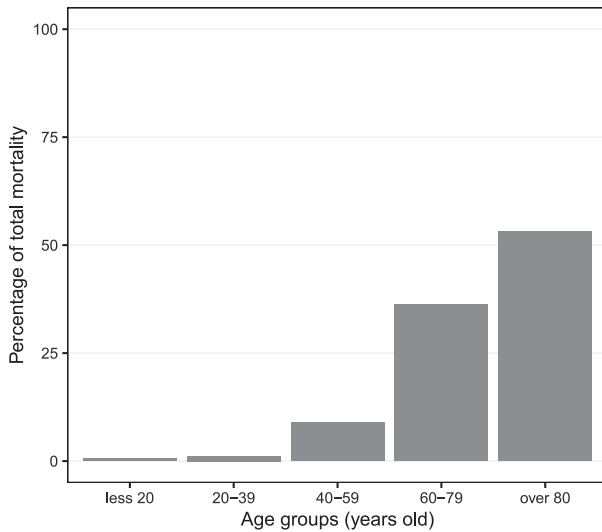


FIG. A2. Percentage of total mortality by age group. The cumulative absolute mortality during the period 2010–19 (April–September) is considered.

APPENDIX B

Additional Results

Figures B1 and B2 show the importance of various variables for predicting Google search attention to the terms *heat wave* and *heat stroke*, respectively, in Germany with a random forest model. Figure B3 presents importance of variables for identifying the most relevant temperature variable for various societal metrics. Figure B4 shows evolution of temperature, considered societal metrics, and number of warnings during heat waves for the five years for which the percentage of heat warnings is highest and the five years for which the percentage of heat warnings is lowest. Figure B5 shows variable importance for predicting Google search attention to *heat wave* in Germany with a random forest model using main temperature variables and time-lagged Google search attention as predictors. Figure B6 shows temperature thresholds in Google search attention to *heat wave* for six important temperature variables. Figure B7 presents seasonally varying sensitivity of Google search attention to temperature.

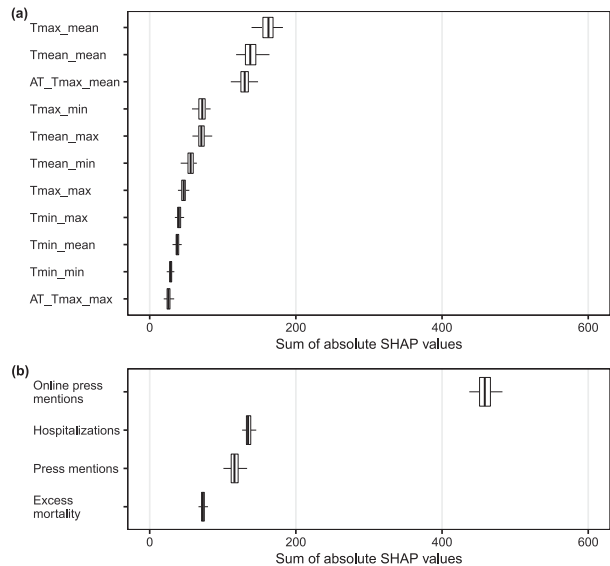


FIG. B1. Importance of different variables for predicting Google search attention to *heat wave* in Germany with a random forest model. The variable importance was derived using SHAP values. Predictor variables are (a) temperature variables and (b) societal variables. The first part of the temperature variable names denotes the type of considered daily temperature; the last part describes the type of aggregation to a weekly time scale (the abbreviations can be found in Table A2). Boxes show medians and uncertainty after bootstrapping the data (25th and 75th percentile).

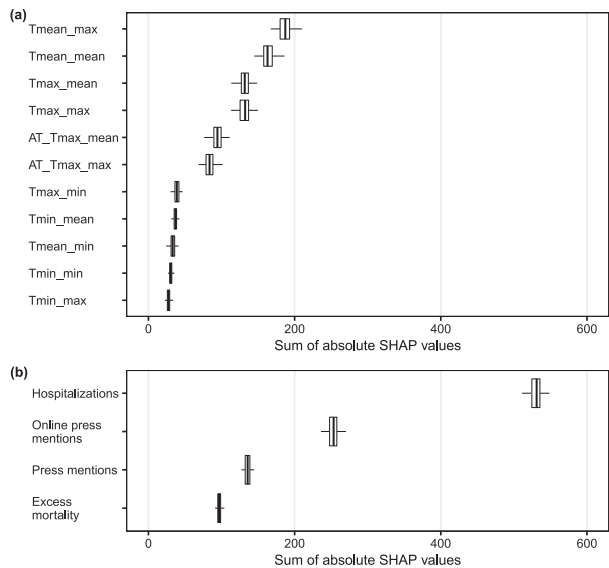


FIG. B2. As in Fig. B1, but for *heat stroke*.

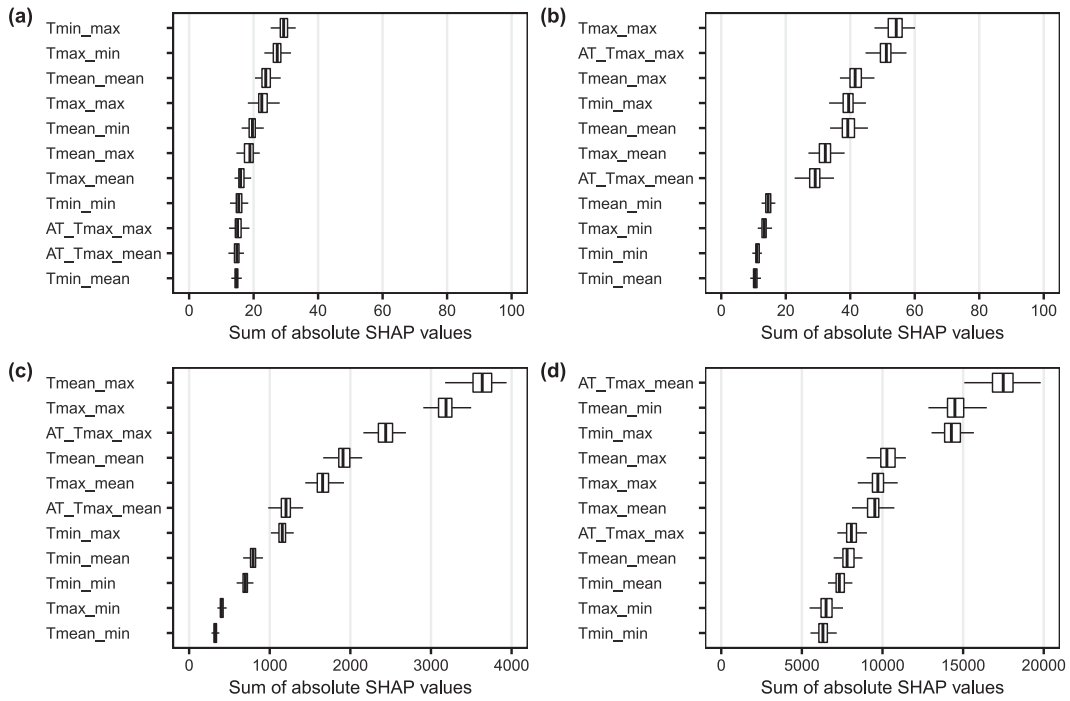


FIG. B3. Similar to Fig. B1, but for identifying the most relevant temperature variable for the following societal metrics: (a) press mentions, (b) online press mentions, (c) hospitalizations, and (d) excess mortality.

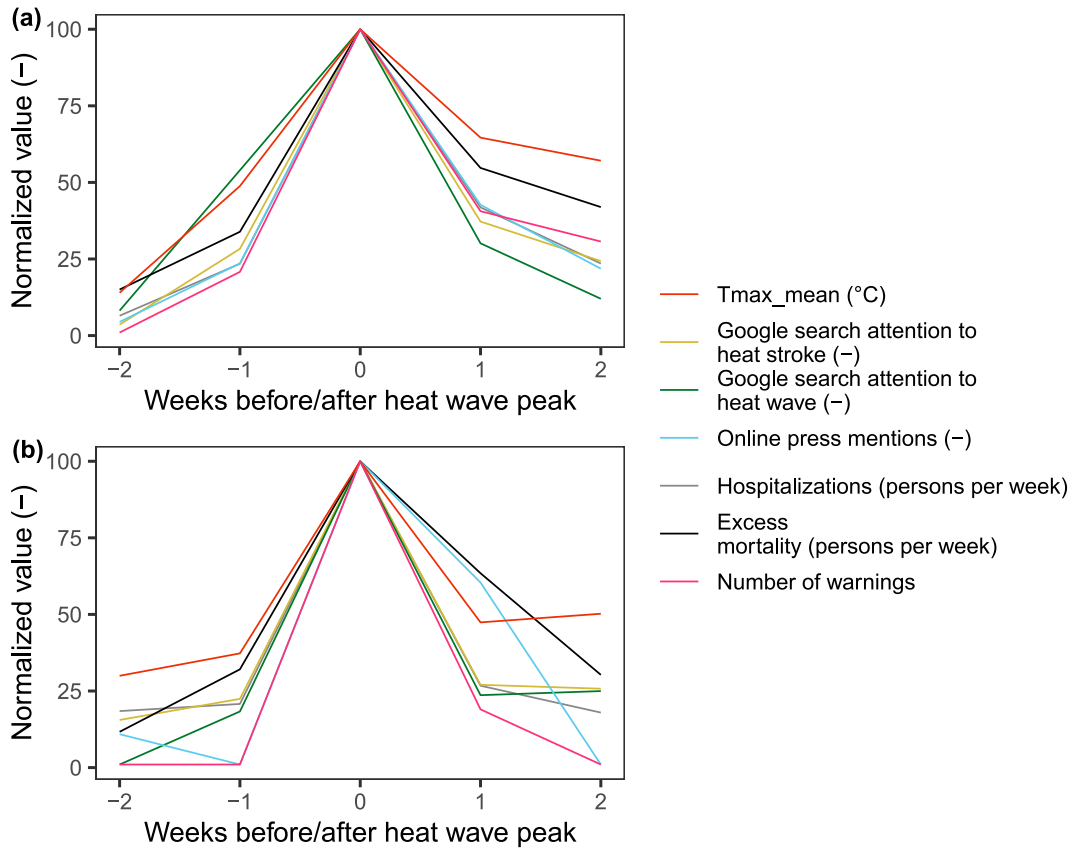


FIG. B4. Evolution of temperature, considered societal metrics, and number of warnings during heat waves for (a) the five years for which the percentage of heat warnings is highest and (b) the five years for which the percentage of heat warnings is lowest. The number of warnings is produced for at least 50% of German districts. Press mentions are excluded because of insufficient data. Values are normalized for comparability and averaged across 10 events that are the hottest events in each year.

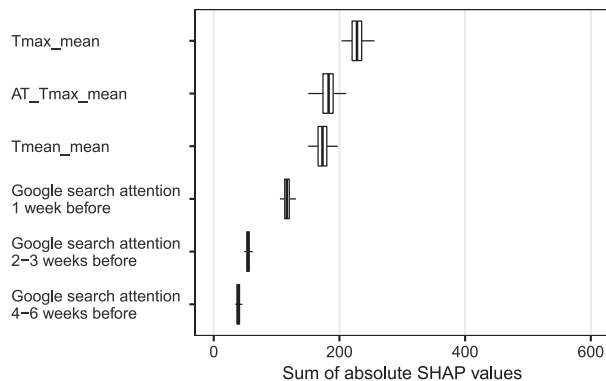


FIG. B5. Similar to Fig. B1, but for predicting Google search attention to *heat wave* in Germany with a random forest model using main temperature variables and time-lagged Google search attention as predictors.

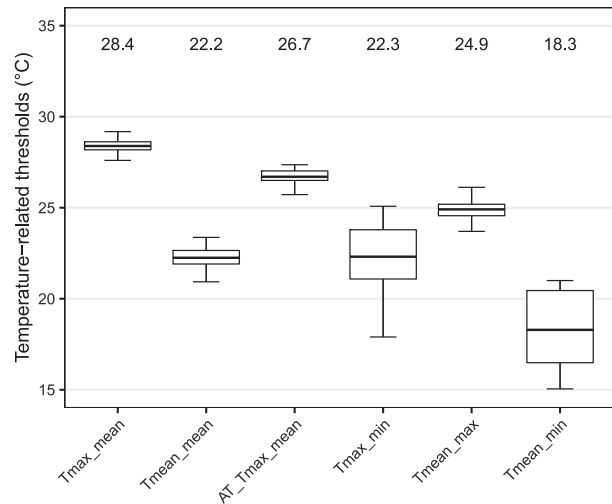


FIG. B6. Temperature thresholds in Google search attention to *heat wave* for six important temperature variables. The first part of the temperature variable names denotes the type of considered daily temperature; the last part describes the type of aggregation to a weekly time scale (the abbreviations can be found in Table A2). Boxes show medians and uncertainty after bootstrapping the data (25th and 75th percentile).

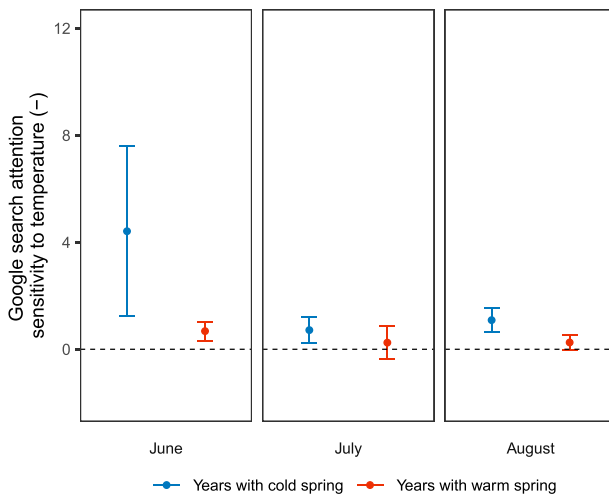


FIG. B7. Seasonally varying sensitivity of Google search attention to temperature. Points represent linear regression slopes of attention–temperature relationships, and error bars are confidence intervals as computed by the linear regression. Results in blue are obtained with data from years after comparatively cold springs; results in red are derived with data from years after warm springs.

REFERENCES

- An der Heiden, M., S. Muthers, H. Niemann, U. Buchholz, L. Grabenhenrich, and A. Matzarakis, 2020: Heat-related mortality: An analysis of the impact of heatwaves in Germany between 1992 and 2017. *Dtsch. Arzteblatt Int.*, **117**, 603–609, <https://doi.org/10.3238/arztebl.2020.0603>.
- Anderson, G. B., and M. L. Bell, 2011: Heat waves in the United States: Mortality risk during heat waves and effect modification by heat wave characteristics in 43 U.S. communities. *Environ. Health Perspect.*, **119**, 210–218, <https://doi.org/10.1289/ehp.1002313>.
- Beckmann, S. K., M. Hiete, and C. Beck, 2021: Threshold temperatures for subjective heat stress in urban apartments—Analysing nocturnal bedroom temperatures during a heat wave in Germany. *Climate Risk Manage.*, **32**, 100286, <https://doi.org/10.1016/j.crm.2021.100286>.
- Bhattacharya, S., 2003: European heatwave caused 35,000 deaths. *New Scientist*, <https://www.newscientist.com/article/dn4259-the-2003-european-heatwave-caused-35000-deaths/>.
- Center for International Earth Science Information Network, 2018: Gridded Population of the World, version 4 (GPWv4): Population density, revision 11. NASA Socioeconomic Data and Applications Center, accessed 10 October 2020, <https://doi.org/10.7927/H49C6VHW>.
- De Albuquerque, J. P., B. Herfort, A. Brenning, and A. Zipf, 2015: A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *Int. J. Geogr. Inf. Sci.*, **29**, 667–689, <https://doi.org/10.1080/13658816.2014.996567>.
- Dormann, C. F., and Coauthors, 2013: Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 27–46, <https://doi.org/10.1111/j.1600-0587.2012.07348.x>.
- Eurostat, 2019: Culture statistics—2019 edition. European Commission Publications Office, 204 pp., <https://data.europa.eu/doi/10.2785/118217>.
- , 2022: Digital economy and society statistics—Households and individuals. Accessed 7 July 2022, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals#Internet_usage.
- Gabriel, K. M., and W. R. Endlicher, 2011: Urban and rural mortality rates during heat waves in Berlin and Brandenburg, Germany. *Environ. Pollut.*, **159**, 2044–2050, <https://doi.org/10.1016/j.envpol.2011.01.016>.
- Google News Initiative, 2022: Basics of Google Trends. Accessed 10 July 2022, <https://newsinitiative.withgoogle.com/resources/lessons/basics-of-google-trends/>.
- Granka, L. A., 2010: Measuring agenda setting with online search traffic: Influences of online and traditional media. *2010 Annual Meeting of the American Political Science Association*, Washington, DC, APSA, <https://ssrn.com/abstract=1658172>.
- Grasso, V., A. Crisci, M. Morabito, P. Nesi, and G. Pantaleo, 2017: Public crowdsensing of heat waves by social media data. *Adv. Sci. Res.*, **14**, 217–226, <https://doi.org/10.5194/asr-14-217-2017>.
- Gregorutti, B., B. Michel, and P. Saint-Pierre, 2017: Correlation and variable importance in random forests. *Stat. Comput.*, **27**, 659–678, <https://doi.org/10.1007/s11222-016-9646-1>.
- Hajat, S., M. O'Connor, and T. Kosatsky, 2010: Health effects of hot weather: From awareness of risk factors to effective health protection. *Lancet*, **375**, 856–863, [https://doi.org/10.1016/S0140-6736\(09\)61711-6](https://doi.org/10.1016/S0140-6736(09)61711-6).
- Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Heudorf, U., and M. Schade, 2014: Heat waves and mortality in Frankfurt am Main, Germany, 2003–2013. *Z. Gerontol. Geriatr.*, **47**, 475–482, <https://doi.org/10.1007/s00391-014-0673-2>.
- Hughes, L., E. Hanna, and J. Fenwick, 2016: The silent killer: Climate change and the health impacts of extreme heat. Climate Council of Australia, 36 pp., https://openresearch-repository.anu.edu.au/bitstream/1885/187127/2/01_Hughes_The_silent_killer%253A_Climate_2016.pdf.
- IPCC, 2021: *Climate Change 2021: The Physical Science Basis*. Cambridge University Press, 2391 pp.
- Jung, J., and C. K. Uejio, 2017: Social media responses to heat waves. *Int. J. Biometeor.*, **61**, 1247–1260, <https://doi.org/10.1007/s00484-016-1302-0>.
- , —, C. Duclos, and M. Jordan, 2019: Using web data to improve surveillance for heat sensitive health outcomes. *Environ. Health*, **18**, 59, <https://doi.org/10.1186/s12940-019-0499-x>.
- Kam, J., K. Stowers, and S. Kim, 2019: Monitoring of drought awareness from Google Trends: A case study of the 2011–17 California drought. *Wea. Climate Soc.*, **11**, 419–429, <https://doi.org/10.1175/WCAS-D-18-0085.1>.
- Karwat, A., and C. L. E. Franke, 2021: Future projections of heat mortality risk for major European cities. *Wea. Climate Soc.*, **13**, 913–931, <https://doi.org/10.1175/WCAS-D-20-0142.1>.
- Koppe, C., S. Kovats, G. Jendritzky, and B. Menne, 2004: Heatwaves: Risks and responses. WHO Regional Office for Europe, 123 pp., <https://apps.who.int/iris/handle/10665/107552>.
- Kryvasheyeu, Y., H. Chen, N. Obradovich, E. Moro, P. Van Hentenryck, J. Fowler, and M. Cebrian, 2016: Rapid assessment of disaster damage using social media activity. *Sci. Adv.*, **2**, e1500779, <https://doi.org/10.1126/sciadv.1500779>.
- Lazer, D., R. Kennedy, G. King, and A. Vespignani, 2014: The parable of Google Flu: Traps in big data analysis. *Science*, **343**, 1203–1205, <https://doi.org/10.1126/science.1248506>.
- Li, T., F. Ding, Q. Sun, Y. Zhang, and P. L. Kinney, 2016: Heat stroke internet searches can be a new heatwave health

- warning surveillance indicator. *Sci. Rep.*, **6**, 37294, <https://doi.org/10.1038/srep37294>.
- Li, Z., C. Wang, C. T. Emrich, and D. Guo, 2018: A novel approach to leveraging social media for rapid flood mapping: A case study of the 2015 South Carolina floods. *Cartogr. Geogr. Inf. Sci.*, **45**, 97–110, <https://doi.org/10.1080/15230406.2016.1271356>.
- Lindley, S. J., P. A. Cook, M. Dennis, and A. Gilchrist, 2019: Biodiversity, physical health and climate change: A synthesis of recent evidence. *Biodiversity and Health in the Face of Climate Change*, M. Marselle et al., Eds., Springer, 17–46.
- Lundberg, S. M., and S.-I. Lee, 2017: A unified approach to interpreting model predictions. *Proc. 31st Int. Conf. on Neural Information Processing Systems*, Long Beach, CA, NIPS, 4768–4777, https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.
- Matzarakis, A., G. Laschewski, and S. Muthers, 2020: The Heat Health Warning System in Germany—Application and warnings for 2005 to 2019. *Atmosphere*, **11**, 170, <https://doi.org/10.3390/atmos11020170>.
- Maurer, M., and T. Holbach, 2016: Taking online search queries as an indicator of the public agenda: The role of public uncertainty. *Journalism Mass Commun. Quart.*, **93**, 572–586, <https://doi.org/10.1177/1077699015610072>.
- McCombs, M. E., and D. L. Shaw, 1972: The agenda-settings function of mass media. *Public Opin. Quart.*, **36**, 176–187, <https://doi.org/10.1086/267990>.
- Meehl, G. A., and C. Tebaldi, 2004: More intense, more frequent, and longer lasting heat waves in the 21st century. *Science*, **305**, 994–997, <https://doi.org/10.1126/science.1098704>.
- Minor, K., A. Bjerre-Nielsen, S. S. Jonasdottir, S. Lehmann, and N. Obradovich, 2022: Rising temperatures erode human sleep globally. *One Earth*, **5**, 534–549, <https://doi.org/10.1016/j.oneear.2022.04.008>.
- Molnar, C., 2020: SHAP (SHapley Additive exPlanations). Interpretable machine learning, 9.6, <https://christophm.github.io/interpretable-ml-book/shap.html>.
- Murage, P., S. Hajat, and R. S. Kovats, 2017: Effect of nighttime temperatures on cause and age-specific mortality in London. *Environ. Epidemiol.*, **1**, e005, <https://doi.org/10.1097/EE9.000000000000005>.
- Muthers, S., G. Laschewski, and A. Matzarakis, 2017: The summers 2003 and 2015 in south-west Germany: Heat waves and heat-related mortality in the context of climate change. *Atmosphere*, **8**, 224, <https://doi.org/10.3390/atmos8110224>.
- National Weather Service, 2021: What is the heat index? Accessed 24 September 2021, <https://www.weather.gov/ama/heatindex>.
- Nuti, S. V., B. Wayda, I. Ranasinghe, S. Wang, R. P. Dreyer, S. I. Chen, and K. Murugiah, 2014: The use of Google Trends in health care research: A systematic review. *PLOS ONE*, **9**, e109583, <https://doi.org/10.1371/journal.pone.0109583>.
- Ripberger, J. T., 2011: Capturing curiosity: Using internet search trends to measure public attentiveness. *Policy Stud. J.*, **39**, 239–259, <https://doi.org/10.1111/j.1541-0072.2011.00406.x>.
- Robine, J.-M., S. L. K. Cheung, S. Le Roy, H. Van Oyen, C. Grif-fiths, J.-P. Michel, and F. R. Herrmann, 2008: Death toll exceeded 70,000 in Europe during the summer of 2003. *C. R. Biol.*, **331**, 171–178, <https://doi.org/10.1016/j.crv.2007.12.001>.
- Rogers, S., 2016: What is Google Trends data—And what does it mean? Medium, <https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8>.
- Romanello, M., and Coauthors, 2021: The 2021 report of the *Lancet* Countdown on health and climate change: Code red for a healthy future. *Lancet*, **398**, 1619–1662, [https://doi.org/10.1016/S0140-6736\(21\)01787-6](https://doi.org/10.1016/S0140-6736(21)01787-6).
- Rothfusz, L. P., 1990: The heat index “equation” (or, more than you ever wanted to know about heat index). NWS Tech. Rep. SR 90-23, 2 pp., https://wonder.cdc.gov/wonder/help/Climate/ta_htindx.PDF.
- Royé, D., and Coauthors, 2021: Effects of hot nights on mortality in southern Europe. *Epidemiology*, **32**, 487–498, <https://doi.org/10.1097/EDE.0000000000001359>.
- Scharkow, M., and J. Vogelgesang, 2011: Measuring the public agenda using search engine queries. *Int. J. Public Opin. Res.*, **23**, 104–113, <https://doi.org/10.1093/ijpor/edq048>.
- Seneviratne, S. I., and Coauthors, 2012: Changes in climate extremes and their impacts on the natural physical environment. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, C. B. Field et al., Eds., Cambridge University Press, 109–230, <https://doi.org/10.7916/d8-6nbt-s431>.
- Shaw, D. L., and S. E. Martin, 1992: The function of mass media agenda setting. *Journalism Mass Commun. Quart.*, **69**, 902–920, <https://doi.org/10.1177/107769909206900410>.
- Shelton, T., A. Poorthuis, M. Graham, and M. Zook, 2014: Mapping the data shadows of Hurricane Sandy: Uncovering the sociospatial dimensions of ‘big data’. *Geoforum*, **52**, 167–179, <https://doi.org/10.1016/j.geoforum.2014.01.006>.
- Shoemaker, P. J., M. Eichholz, E. Kim, and B. Wrigley, 2001: Individual and routine forces in gatekeeping. *Journalism Mass Commun. Quart.*, **78**, 233–246, <https://doi.org/10.1177/107769900107800202>.
- Singh, T., C. Siderius, and Y. Van der Velde, 2018: When do Indians feel hot? Internet searches indicate seasonality suppresses adaptation to heat. *Environ. Res. Lett.*, **13**, 054009, <https://doi.org/10.1088/1748-9326/aaba82>.
- Statista, 2018: Market share of the leading search engines in Germany in 2018. Accessed 10 July 2020, <https://www.statista.com/statistics/445002/market-shares-leading-search-engines-germany/>.
- , 2021a: Share of internet users in Germany in 2014 to 2020, by age. Accessed 10 May 2021, <https://www.statista.com/statistics/790407/internet-usage-by-age-germany/>.
- , 2021b: Share of internet users in Germany from 2001 to 2021. Accessed 10 May 2021, <https://www.statista.com/statistics/380514/internet-usage-rate-germany/>.
- Tuholske, C., K. Caylor, C. Funk, A. Verdin, S. Sweeney, K. Grace, P. Peterson, and T. Evans, 2021: Global urban population exposure to extreme heat. *Proc. Natl. Acad. Sci. USA*, **118**, e2024792118, <https://doi.org/10.1073/pnas.2024792118>.
- World Meteorological Organization, 1992: International meteorological vocabulary. 2nd ed. WMO/OMM/BMO-182, 802 pp., https://library.wmo.int/doc_num.php?explnum_id=4712.
- Xu, P., L. Wang, Y. Liu, W. Chen, and P. Huang, 2020: The record-breaking heat wave of June 2019 in central Europe. *Atmos. Sci. Lett.*, **21**, e964, <https://doi.org/10.1002/asl.964>.
- Yan, Y., and K. Bissell, 2018: The sky is falling: Predictors of news coverage of natural disasters worldwide. *Commun. Res.*, **45**, 862–886, <https://doi.org/10.1177/0093650215573861>.