Environmental Research 261 (2024) 119667

Contents lists available at ScienceDirect

Environmental Research

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

The spatial distribution of heat related hospitalizations and classification of the most dangerous heat events in California at a small-scale level

Kristen Hansen [a,b], Armin Schwartzman [a], Lara Schwarz [a], Anais Teyton [a], Rupa Basu [c], Tarik Benmarhnia [d,*]

[a] Herbert Wertheim School of Public Health, University of California San Diego, La Jolla, CA, USA
[b] Axle Research and Technology, Rockville, MD, USA
[c] Office of Environmental Health Hazard Assessment, California Environmental Protection Agency, Oakland, CA, USA
[d] Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA

ARTICLE INFO

Keywords:
Climate change and health
Extreme heat and health
Spatial modeling
Bayesian hierarchical modelling

ABSTRACT

Many studies have explored the impact of extreme heat on health, but few have investigated localized heat-health outcomes across a wide area. We examined fine-scale variability in vulnerable areas, considering population distribution, local weather, and landscape characteristics. Using 36 different heat event definitions, we identified the most dangerous types of heat events based on minimum, maximum, and diurnal temperatures with varying thresholds and durations. Focusing on California's diverse climate, elevation, and population distribution, we analyzed hospital admissions for various causes of admission (2004–2013). Our matching approach identified vulnerable zip codes, even with small populations, on absolute and relative scales. Bayesian Hierarchical models leveraged spatial correlation. We ranked the 36 heat event types by attributable hospital admissions per zip code and provided code, simulated data, and an interactive web app for reproducibility. Our findings showed high variation in heat-related hospitalizations in coastal cities and substantial heat burdens in the Central Valley. Diurnal heat events had the greatest impact in the Central Valley, while nighttime extreme heat events drove burdens in the southeastern desert. This spatially informed approach guides local policies, prioritizing dangerous heat events to reduce the heat-health burden. The methodology is applicable to other regions, informing early warning systems and characterizing extreme heat impacts.

Competing financial interests

The authors declare they have no actual or potential competing financial interests.

1. Introduction

Extreme heat events result in the highest number of weather-related deaths in the US and globally, more than any other weather-related disaster (Nitschke et al., 2011; Robine et al., 2012; Ebi et al., 2021; Cole et al., 2023). Many epidemiological studies have found that extreme heat events (EHE) increased the risk of hospitalization for many diagnoses including hospitalizations from cardiovascular, respiratory, diabetes, fluid and electrolyte disorders, and renal failure(Vaidyanathan et al., 2019; Achebak et al., 2024). However, EHEs do not affect populations equally, and some communities are more vulnerable to heat (Smargiassi et al., 2009; Benmarhnia et al., 2015; Schinasi et al., 2018; Guo et al., 2023). Health risks associated with heat can vary across space (Vaneckova et al., 2010; Hondula et al., 2012; Hondula and Barnett, 2014; Benmarhnia et al., 2017), including within cities or counties (McElroy et al., 2020). Identifying such spatial heterogeneity in the impacts of EHEs can be particularly useful for targeting vulnerable areas and communities as well as to guide early warning systems that can greatly reduce the health impacts of extreme heat. However, studying the spatial heterogeneity of the effects of extreme heat requires important methodological considerations, including the handling of spatial information, the classification of extreme heat events, and the scale of estimated effects.

In addition, communities may be differentially impacted by various types, or "flavors", of EHEs (Anderson and Bell, 2011). There is no standardized way to classify heat events. For example, EHEs can be defined using different temperature metrics, such as minimum,

* Corresponding author. Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, 92093, CA, USA.
E-mail address: tbenmarhnia@ucsd.edu (T. Benmarhnia).

https://doi.org/10.1016/j.envres.2024.119667
Received 7 March 2024; Received in revised form 29 June 2024; Accepted 21 July 2024
Available online 26 July 2024

maximum, or diurnal temperature (representing the nighttime-daytime difference). Additionally, various lengths of heat exposure (single day versus multi-day heat events) may be considered. Finally, different percentile thresholds (e.g., 95th, 99th etc.) can be used to define extreme heat events. While it is expected that higher thresholds are associated with higher risks of hospital admission (on a relative scale), such events are less frequent (McElroy et al., 2020). Therefore, taking the occurrence of such extreme heat events into consideration is critical when estimating the total burden on the absolute scale (i.e., the total number of attributable hospitalizations) associated with different heat events (McElroy et al., 2020).

Moreover, while it is important to quantify spatial variation in the risks associated with heat on a relative scale (expressed through risk ratios or standardized mortality ratios for instance), it is also critical to quantify such risk on an absolute scale in terms of number of cases attributable to extreme heat (Gasparrini et al., 2015; Vaidyanathan et al., 2019). The relative scale allows us to see where people are more (or less) vulnerable to heat, whereas the absolute scale determines where the greatest number of people affected reside. These two scales can complement one another in the process of designing adaptation strategies. The absolute scale allows governments to design policies that proportionately target those areas with the highest burden to reduce the total number of hospitalizations due to heat. The relative scale allows policymakers to identify areas composed of vulnerable populations that are not typically emphasized in studies focusing on the state or county levels or when using traditional spatiotemporal methods. An additional challenge when assessing the fine-scale spatial variability of heat impacts is related to the precision of estimates. We propose a novel within-community matched design coupled with a Bayesian Hierarchical Model (Aguilera et al., 2020; Schwarz et al., 2021; Chen et al., 2024) to analyze spatio-temporal impacts of various EHEs at a fine scale.

Spatial heterogeneity in the type of heat events that drive the health burden can be notably explained by differences in population composition (Hondula and Barnett, 2014; Benmarhnia et al., 2017), local meteorological conditions (Guirguis et al., 2018), or landscape characteristics (Schinasi et al., 2018; Chakraborty et al., 2023). California constitutes an ideal region to study the spatial variation of the impact of heat on hospital admissions due to the high burden, high variation in population distribution, and high variance in climate throughout the state (Gershunov and Guirguis, 2012). Heat-related health impacts are well studied in California (Ostro et al., 2009; Sherbakov et al., 2018; Fard et al., 2023); however, to the best of our knowledge, there is no small region spatial estimation or statewide estimate of spatial variability of heat-related health impact across California.

We applied a comprehensive spatio-temporal approach, an extension of a within-community matched design, to study the spatial variability in the health impacts of heat events in California for unplanned hospitalizations for cardiovascular disease (CVD), respiratory disease, acute renal failure, dehydration, and heat illness. We extended this methodology by using Bayesian models to account for spatial autocorrelation, improve precision, and explore what drives small-scale impacts to extreme heat including the heat metric, heat event length, and heat event intensity. Identifying what heat event characteristics drive the greatest health burden on the relative and absolute scale can be used to prioritize specific areas and neighborhoods in policy planning to best protect populations from the effects of extreme heat.

2. Methods

2.1. Hospitalization data

We obtained all unscheduled hospitalization data in California for the years 2004–2013 from the Office of Statewide Health Planning and Development (OSHPD). The following primary diagnoses were evaluated, as listed in the International Classification of Disease codes, 9th Revision, Clinical Modification (ICD-9): acute myocardial infarction (MI) (410), acute renal failure (584), cardiac dysrhythmias (427), CVD (390–459), dehydration/volume depletion (276.5), essential hypertension (401), heat illness (992), ischemic heart disease (410–414), ischemic stroke (433–436), and all respiratory diseases (460–519). These diseases were chosen because they have previously been linked to extreme temperatures (Sherbakov et al., 2018; Vaidyanathan et al., 2019). For this analysis, all cardiovascular hospitalizations were grouped, leaving five hospitalization outcomes of interest. Data were aggregated into daily counts for each zip code. Zip code population data was obtained from the 2010 US Census and was mapped using Zip Code Tabulation Areas (ZCTAs). Climate regions in California are presented in Supplemental Fig. A1 and Population by zip code is presented in Fig. A2. We focused on unscheduled hospital admissions and emergency department visits data as there is robust evidence linking extreme heat with hospital admissions and emergency department visits but also because, for planning purposes, preventing heat-related morbidity will guarantee wider benefits than focusing on mortality outcomes.

2.2. Meteorological data and extreme heat events definition

Temperature data were downloaded from a publicly available dataset that collects weather data from National Oceanic and Atmospheric Administration Cooperative Observer Program (NOAA COOP) stations across the US and provides data at a gridded level (https://cal-adapt.org/). Daily maximum and minimum temperatures (°C) were derived from 1/16° (~6 km) gridded observed data from this dataset for all of California (Livneh et al., 2013). Population-weighted centroids for each ZCTA were linked to the nearest temperature measurements using the *geonear* function in Stata15 SE. Population data, including population-weighted centroids, were obtained from the 2010 US census.

There are many possible definitions for a heat event as evidenced by warning systems across the world. In this study, we considered 36 distinct EHE definitions based on zip code-specific temperature distributions. These include different metrics: heat index, maximum and minimum temperatures, and the difference between the maximum and minimum temperature for each day (i.e., diurnal temperature). We then considered duration as either 1, 2, or 3 days of heat exposure. Finally, we considered temperature extremity by including the 95th, 97.5th, and 99th percentiles for minimum and maximum temperatures and 1st, 2.5th, and 5th percentiles for diurnal temperatures. The 36 EHE definitions are all the possible combinations of the above factors. As an example, consider a 95th percentile maximum temperature, 1-day event. An EHE day is defined as one in which the daily maximum temperature is greater than or equal to the 95th percentile of the distribution of maximum temperatures during the warm season (May–September) for each zip code. The observed temperature threshold for this EHE is displayed in Fig. 1 in degrees Celsius. Fig. 1 displays temperatures for the four main geographic regions in California, showing that the Central Valley and southeastern desert are warmer, and the coast and eastern mountain ranges are generally cooler. The same procedure is applied to all 36 different EHE definitions.

2.3. Data analysis

2.3.1. The spatial within-community matched design

We used a spatio-temporal approach, which can be described as a spatial within-community temporally matched design (Goin et al., 2019). We have applied this approach to other environmental exposures in previous work (Aguilera et al., 2020; Schwarz et al., 2021; Chen et al., 2024; Do et al., 2024). This approach allows us to control for any time-fixed, measured, or unmeasured confounders at the zip code level. Our approach includes four sequential steps to quantify the heat impacts on both relative and absolute scales. A schematic of this approach can be found in Fig. 2. First, we adopted a procedure to match EHE days in each zip code to similar non-EHE days (i.e., days without a heat event and within the same calendar year and day of the week), and we produced a

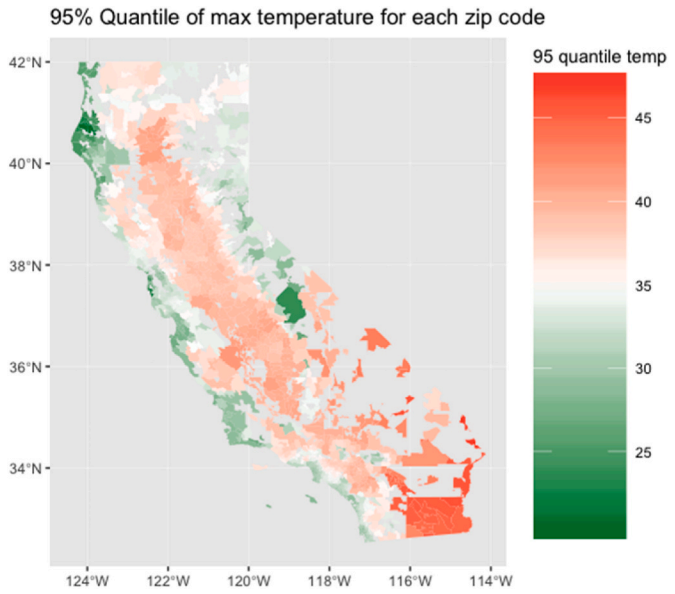

Fig. 1. Display of the 95th percentile threshold in degrees Celsius used in our analysis for each zip code in California. The mean threshold of 35° Celsius is shown in white.

contrast from a weighted average based on distance in time from the EHE day, given that the control non-EHE days fall within the same calendar year. All non-EHE days for each ZCTA between May and September are used as matched controls as long as they have temperatures above the 75th percentile for that ZCTA. They are then weighted based on how far in time each is from the event days using a kernel distance weight. This is a temporal version of an Inverse Distance Weighting (IDW) procedure, where our bandwidth is a single calendar year. Second, we calculated the absolute difference between EHE days and the weighted average of non-EHE days for each zip code. Third, we modeled the relationship between excess counts of hospitalizations and the population size of each zip code using a linear model, where the residuals (see details in Supplemental Fig. A3) represent the heat impacts relative to the population size. Said differently, this model is standardizing on population size. The model was defined as follows:

$$Y = X\beta + \epsilon$$

In this equation, Y is the excess count of hospitalizations (henceforth referred to as the absolute scale estimate), X represents the population value for each zip code in the 2010 census, ϵ is the error term, and β is the estimate of the population effect. We do not consider the effect estimate, but rather the difference between the predicted Y and the observed Y.

2.3.2. Bayesian Hierarchical Models

For our fourth step, we used a spatial Bayesian Hierarchical model (BHM) to consider spatial autocorrelation and improve precision in our estimates. Spatial modeling leverages information from surrounding areas to improve the precision of the estimate at any point in space. We used the within-community matched design absolute scale estimates for each zip code as the response value in linear BHMs. For relative scale estimates, we used the residuals from the linear models described above, using the *spBayes* package in R. We used population weighted centroids provided by the US Census Bureau as the spatial unit for this analysis as this type of modeling requires a Spatial Points Data Frame. We fit an empirical semi-variogram to estimate the starting prior for the parameters: sill (σ^2), nugget (τ^2), and range (ϕ). Due to the shape of the empirical semi-variogram, we chose a Spherical correlation structure. The model was formulated as a two-stage model:

$$1^{st}\text{ stage}: Y_i|\theta, W \sim N(X\beta + W, \tau^2 I)$$

$$2^{nd}\text{ stage}: W|\sigma^2, \phi \sim N(0, \sigma^2 H)$$

Where W is the vector of spatial weights, and θ is the vector of estimated spatial parameters. The Y_i are our outcome values, which are independent but conditional on W. H represents the structure of the spatial covariance and X represents an intercept. The second stage model captures the spatial process of the data. We completed model specification by adding prior values and distributions to β and τ^2, and the hyper parameters ϕ and σ^2.

We used prior distributions of parameters to reduce sensitivity to the priors during the sampling process. We applied 1000 Markov chain Monte Carlo (MCMC) samples, with the final 250 kept after the burn-in period. The final recovered spatial weights were utilized as the estimates for excess hospitalizations in each zip code. We interpolated across space using multi-level B-splines to create a surface of estimates. Though the above methodology assumes isotropy, we acknowledge that the spatial correlation may not be stationary. Isotropy is the assumption that the spatial correlation has the same range in all directions at all points in a data set. Lastly, to represent the precision of the BHM estimates, we estimated the signal-to-noise ratio (SNR) using the resulting model output (weights for each ZCTA and standard deviations). An SNR is the

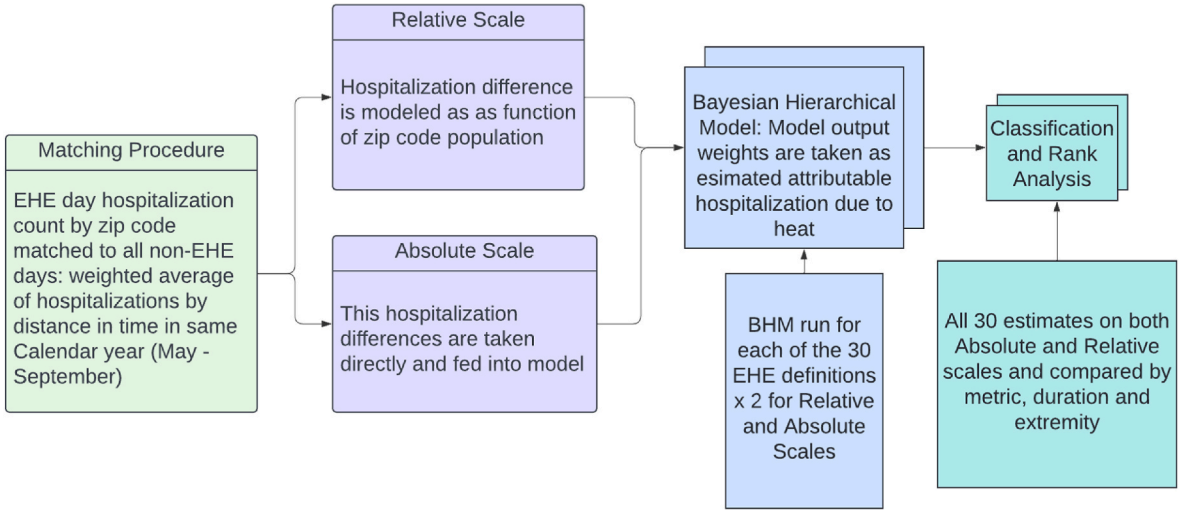

Fig. 2. Schematic of the methods procedure for within-community matched design and Bayesian extension.

estimated spatial prediction from the BHM divided by the standard error estimated from that BHM. The SNR was mapped for each ZCTA as an indication of areas where estimates are more (or less) precise. We accept that these SNR values are artificially large because of the violation of the isotropy assumption and thus the traditional cutoff for significance (which is 2.0) would be too small. We also present Bootstrap -T lower bound confidence interval for absolute excess counts in Supplemental Fig. A4.

We included randomly generated data and R syntax for reproducibility purposes [https://github.com/KristenHansen/SpatialHeatWaves].

2.3.3. Extreme heat events definition classification and rank

For the above analysis, each EHE definition is considered in isolation; however, it is important to consider all types of events holistically. For both the relative and absolute scales, we formed a dataset with our BHM estimates of excess hospitalization due to heat and standard errors consisting of 72 columns representing each EHE definition, and rows corresponding to our zip codes. We were then able to rank the extremity of the EHE for each zip code relative to those other events in the same zip code. We took the highest effect estimate found in all EHE definitions as the "most effective" or dangerous EHE. This allowed for the exploration of patterns for all our heat event classifications simultaneously: metric, duration, and extremity. We explored the spatial heterogeneity of the importance of metric, duration, and extremity through the highest ranked EHE by ZCTA and Climate Zone (as defined by the California Energy Commission (California Energy Commission, n.d.)). This serves as an exploratory analytical extension to the BHMs.

3. Results

We observed 131,461 total hospitalizations during EHE days for the five disease subgroups, and 98,562 hospitalizations for matched non-EHE days. CVD hospitalizations accounted for most of the total hospitalizations on both EHE and non-EHE days, with 58% and 66% on EHE and non-EHE days, respectively (see Table 1).

3.1. Spatial Bayesian Model results

Fig. 3 displays the results of the Spatial Bayesian Model for the example EHE definition (95th percentile in maximum temperature for one day). Fig. 3(a) shows the absolute-scale results from the Bayesian Hierarchical model on all-cause hospitalization count differences computed via a within community matched design (results for each ICD code separately are presented in Supplemental Figs. A5 and A6). We observe areas of high effect specifically in the Central Valley and the southeastern desert highlighted as having the highest associations with extreme heat. For Fig. 3(b), which represents the relative scale results, we use population as a covariate in the spatial Bayesian model. Similar patterns are found on both risk scales; however, the relative model (Fig. 3(b)) does display more detail and positive estimates for less populated regions, although the estimates are close to zero, as would be expected for a model where population is included.

However, Fig. 3 does not adequately display the results on the spatial scale that the analysis is computed. Fig. 4 displays the Signal-to-Noise Ratio (SNR) plots for the same EHE definition as Fig. 3 on both relative and absolute scales. The SNR results show higher precision in the absolute case as opposed to the relative. This is to be expected because relative case estimates are lower in magnitude. Results for all other EHE definitions can be viewed in the R Shiny application created to accompany this publication at https://kristenhansen.shinyapps.io/HeatCA.

In Table 2, there is a display of temperature and population data by climate zone as well as the SNR for the highest ranking EHE (the highest observed effect of all 36 EHE definitions). Of the ZCTAs, 11.8% do not have a single EHE definition that has precision in estimates (SNR <2), but this is only true for the very lowest population zip codes with incredibly low hospitalization rates. Only 5% of ZCTAs never saw an SNR above 1.

3.2. Extreme heat event definition rank results

Table 3 shows for the four highest ranked heat event definitions, the number of ZCTAs for which it is the most impactful regarding risk of hospitalizations for both the absolute and relative scales, as well as the proportion of zip codes with a specific EHE definition. We observed that the highest number of heat-related hospitalizations occurred during diurnal EHEs where the temperature difference between day and night was very small. This is particularly noticeable for the top two highest ranked EHEs on both the absolute and relative scales. However, the third and fourth most dangerous EHEs on an absolute scale were defined by diurnal and maximum temperatures, respectively. On a relative scale, the third and fourth most dangerous EHE definitions used minimum temperatures. This is because the low population desert regions of the state are highly affected by minimum temperature EHEs. There are fewer but a consistent number of highly effective EHE definitions that are based on the heat index metric. This lower number is potentially attributable to lower relative humidity in much of the state of California. We can see the spatial pattern of the top ranked EHEs in Fig. 5(a) for the entire state. Fig. 5(b) shows a more detailed map with all 36 definitions for the greater Los Angeles area. We can observe that minimum temperature EHEs are the most dangerous (in terms of their ranking) in the desert regions, diurnal EHEs are the most dangerous in the agricultural regions and coastal areas, and maximum temperature EHEs are often the most dangerous in urban regions and the mountains/forested areas. Heat index EHEs are seen as variable but most impactful in the Central Valley and central coast. There is considerable variability in urban regions, as can be seen in the Los Angeles plot, where diurnal temperature EHEs (and to a lesser extent minimum temperature EHEs) also represent a significant number of ZCTAs. Although all these regions do contain a mixture of the four metrics, the general pattern holds for both relative

Table 1
Descriptive statistics describing the hospitalizations on EHE days and matched non-EHE days in California, 2004–2013. The non-EHE days are the weighted average of all days in the warm season. Thus, this is not a true count but an average.

	Type of Day	Total observed	Mean observed per day per zip code	Percent Increase from matched total	Standard deviation
All-Cause Hospitalizations	EHE	131,461	0.95	33.3%	1.446
	Non-EHE	98,562	0.72	NA	1.148
Respiratory Disease	EHE	40,900	0.297	38.4%	0.647
	Non-EHE	29,551	0.216	NA	0.498
Cardiovascular Disease	EHE	76,644	0.559	17.2%	0.967
	Non-EHE	65,394	0.477	NA	0.803
Acute Renal Failure	EHE	7918	0.057	327%	0.251
	Non-EHE	2421	0.0176	NA	0.135
Dehydration	EHE	5215	0.038	445%	0.200
	Non-EHE	1171	0.009	NA	0.093
Heat Illness	EHE	784	0.006	3136%	0.077
	Non-EHE	25	0.0002	NA	0.014

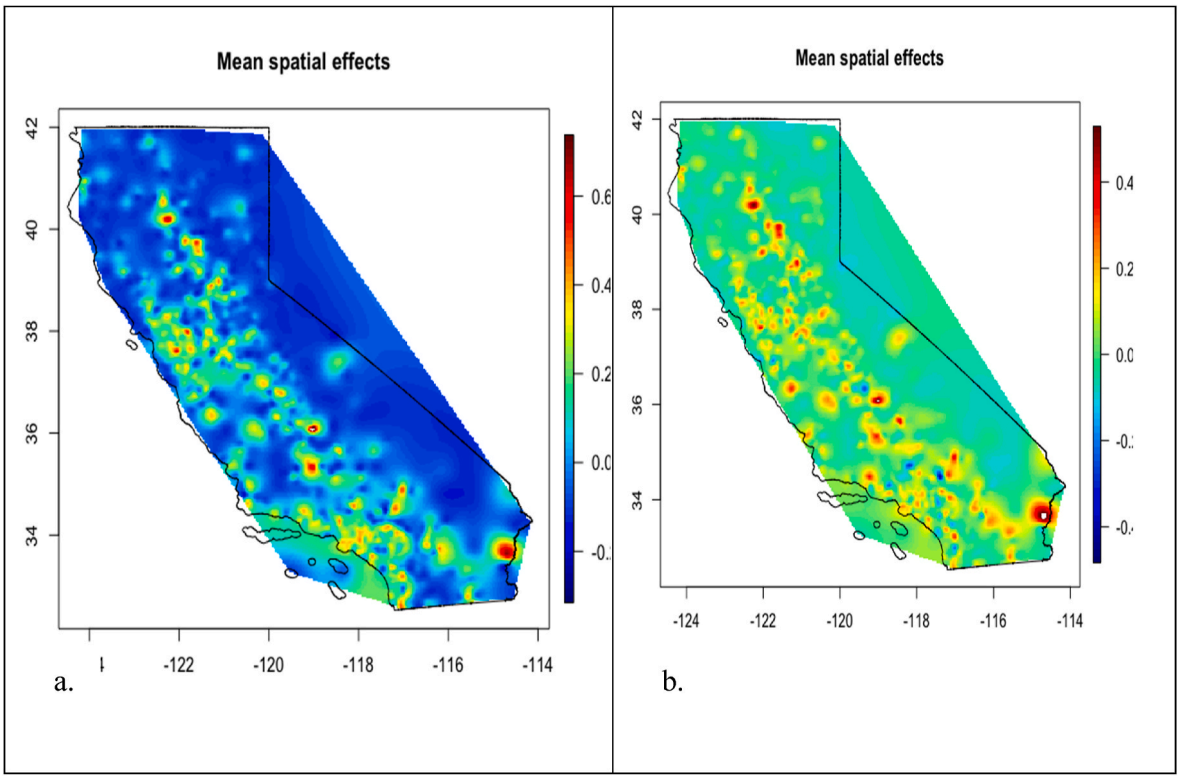

Fig. 3. Spatial distribution of effects of one heat event definition (>95th Tmax percentile for one day) on total hospitalization on the absolute (a, left) and the relative (b, right) scales.

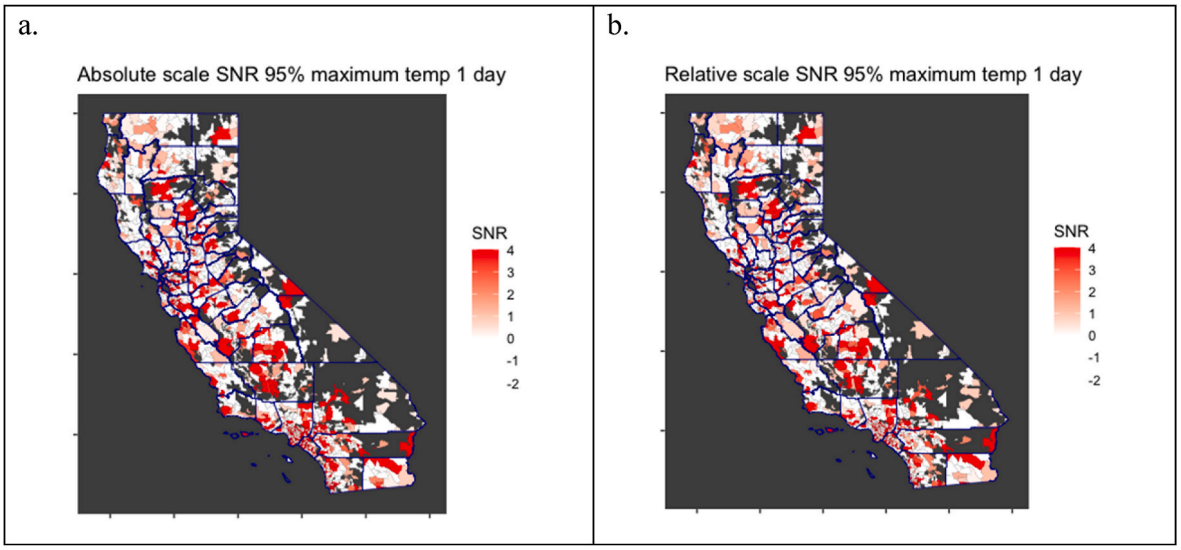

Fig. 4. Signal-to-Noise Ratio (SNR) values showing the effect on hospitalizations of one extreme heat event definition (>95th Tmax percentile for one day) on both the absolute (a, left) and relative (b, right) scale. Dark areas correspond to areas with no observed hospitalization for this extreme heat event definition.

and absolute scales. We find that the less extreme (95th percentile) EHE definitions are more frequent in the absolute scale, suggesting that longer duration, less extreme heat events may cause a higher number of hospitalizations overall. The SNR information for all 36 EHE definitions is accessible at the following link: https://github.com/KristenHansen/SpatialHeatWaves.

All results can be visualized for each county in California via an interactive web application: https://kristenhansen.shinyapps.io/HeatCA. We additionally present the results for the County of San Diego to better visualize the spatial heterogeneity in the supplemental materials (see Supplemental Figs. A7 and A8).

4. Discussion

In this study, we aimed to determine where the areas of health burden associated with extreme heat are and which extreme heat definitions have the greatest impact for different regions in California at the fine-scale level. Areas including the Central Valley region experience higher burden throughout the studied outcomes. Highest spatial variability can be seen within large urban areas: Los Angeles, San Diego, and

Table 2
Temperature, population and SNR results displayed by climate zone as defined by the California Energy Commision [31].

Climate Zone	Mean 95% Max Temp.	Mean 95% Min Temp	Number of ZCTAs	Median Population of ZCTA (2010)	Median SNR of highest rank (by ZCTA)
1	25.2	13.4	49	1222	3.2
2	31.7	14.7	90	3104	7.6
3	27.1	15.3	162	21726	14.6
4	32.4	17.5	78	22385	10.5
5	28.3	15.1	25	7646	8.7
6	29.8	19.9	93	27732	11.9
7	30.0	21.1	59	31779	16.7
8	31.9	21.0	115	36171	15.1
9	34.7	21.7	176	31103	14.3
10	36.9	21.9	103	39337	11.4
11	38.9	21.4	92	3787	6.4
12	38.1	20.4	251	12819	8.8
13	39.9	23.3	121	11661	8.7
14	38.5	23.7	70	2632	4.9
15	43.2	29.5	43	8099	6.6
16	32.8	15.2	207	633	2.5
All	34.1	19.5	1744	15719	9.2

Table 3
The number of ZCTAs for which the four most dangerous EHE definitions fall within the maximum temperatures, minimum temperatures, heat index, and low difference in temperatures definitions of EHE out of 1744 ZCTAs observed.

Scale	Rank	Maximum EHE (%)	Minimum EHE (%)	Diurnal EHE (%)	Heat Index EHE (%)
Relative	**First**	495 (28.3%)	338 (19.4%)	644 (36.9%)	267 (15.3%)
Relative	**Second**	469 (26.9%)	446 (25.6%)	549 (31.4%)	280 (16.0%)
Relative	**Third**	512 (29.3%)	479 (27.4%)	495 (28.3%)	258 (14.8%)
Relative	**Fourth**	471 (27.0%)	461 (26.4%)	496 (28.4%)	316 (18.1%)
Absolute	**First**	506 (29.0%)	289 (16.5%)	600 (34.4%)	349 (20.0%)
Absolute	**Second**	469 (26.9%)	359 (20.6%)	566 (32.4%)	350 (20.1%)
Absolute	**Third**	478 (27.4%)	422 (24.2%)	512 (29.3%)	332 (19.0%)
Absolute	**Fourth**	467 (26.8%)	421 (24.1%)	489 (28.0%)	367 (21.0%)

the San Francsico Bay Area.

Similarly, the heat event definitions that led to the most hospitalizations varied greatly depending on zip code. In general, of the different EHE types we considered, diurnal heat events had the greatest effect in the coastal and agricultural regions, minimum temperature heat events were most detrimental in the desert, maximum temperature heat events were the most impactful in forest and mountainous regions as well as urban areas, and heat index definitions appear most impactful with considerable variability in the Central Valley and central coast, where more extreme levels of humidity are observed. This pattern holds true on both the relative and absolute scales. We find that more extreme and longer heat events ranked highly on the relative scale, with duration being the driver on the absolute scale.

Previous studies have considered or accounted for spatial variation in the impacts of extreme heat by adjusting for spatial location, conducting geographic weighted regressions, or applying cluster analysis methods (Vaneckova et al., 2010; Hondula et al., 2012; Benmarhnia et al., 2017; Song et al., 2021). Spatial autocorrelation in heat effects can and has been studied by using a Gaussian smoother (Chen et al., 2015; Chien et al., 2016). Other studies have stratified by spatial units without considering spatial autocorrelation or spatial structure (Murage et al., 2020; Ingole et al., 2020; Gasparrini et al., 2022). One of the most common methods for estimating spatial variation in heat-related impacts is based on cluster analysis methods, the most readily used being Kulldorf (Vaneckova et al., 2010; Hondula et al., 2012; Benmarhnia et al., 2017). Kulldorf analysis identifies a significant excess of cases within a moving circular window, providing a measure of how unlikely it would be to encounter the observed excess of cases in a comparison region across space. However, such an approach, which is based on significance testing, dichotomizes the spatial units and identifies "significant" clusters where most of the cases occur. Thus, a significant cluster can be driven by the cluster-specific susceptibility to heat or the population size or density. For this reason, it becomes difficult to get a contrasted and comprehensive assessment of the spatial variability regarding heat-related health impacts. Other spatial approaches have been used, such as spatial point pattern analysis to estimate relative risks across the study region at a fine scale (Chen et al., 2015; Chien et al., 2016). Bayesian hierarchical models can also be particularly useful to account for spatial variation in heat vulnerability in data rich regions like large cities (Hondula and Barnett, 2014). For large geographic areas, such as countries or states, existing studies generally consider only large spatial units like US metropolitan areas (Anderson et al., 2013) or UK districts (Bennett et al., 2014; Gasparrini et al., 2022). A recent study also proposed a spatial extension of distributed non-linear models for continuous temperatures (Quijal-Zamorano et al., 2024) based on a Bayesian framework similar to our approach. Our approach also allows researchers to rank heat events according to their impacts on a fine spatial scale. Obtaining a classification of the types of heat events that are associated with the highest health burden by zip code is particularly useful in informing tailored interventions to optimize benefits associated with heat action plans and early warning systems.

The spatial heterogeneity in heat risks in California has been documented previously. Such variability can be explained by climate zones, land use, and topography. Previous research has shown how coastal areas in California are experiencing more humid nighttime events in the context of climate change (Gershunov and Guirguis, 2012). This intensification in heat wave activity is due directly to mean warming, especially ocean warming and increasing atmospheric moisture in coastal areas. In parallel, desert heat waves are expected to become progressively and relatively less intense while coastal heat waves are projected to intensify even relative to background warming. There are 16 distinct climate zones in California that, by definition, have distinct temperature distributions, and the heat patterns vary substantially. Heavy irrigation has also been shown to influence the flavors of heat events and the variance structure of temperature observations.

There are some limitations in our study that should be highlighted. We are restricted to the zip code level, which presents multiple issues. Zip codes have high variability in population size and geographic area and are not regularly shaped. They suffer from what is often referred to as the modifiable areal unit problem. There are low rates of daily hospitalizations, as would be expected, but this leads to low precision along with areas of very low population. We used a Bayesian Hierarchical model (BHM), which improved precision in estimation because it utilizes spatial autocorrelation. Although the BHM improves precision over other types of modeling, it has a restrictive isotropic assumption, which may not hold under scrutiny. To account for this, we used priors that were not particularly restrictive and ended up with acceptance rates around 15%. Spatial correlation structure will be affected by topography and climate. For this reason, we also took a relatively small distance as our range parameter. In future work, it would be important to further explore the use of anisotropic models for spatial data. Here, we consider spatial dependency in the hierarchical models; however, with the isotropy assumption, we are potentially misspecifying the model, thereby increasing bias in the estimates. Further analysis should strive for higher precision in estimation without using incorrect assumptions, perhaps a Bayesian framework that allows for anisotropy. We considered lagged effects considering up to 3 days. However, we did not include longer lags

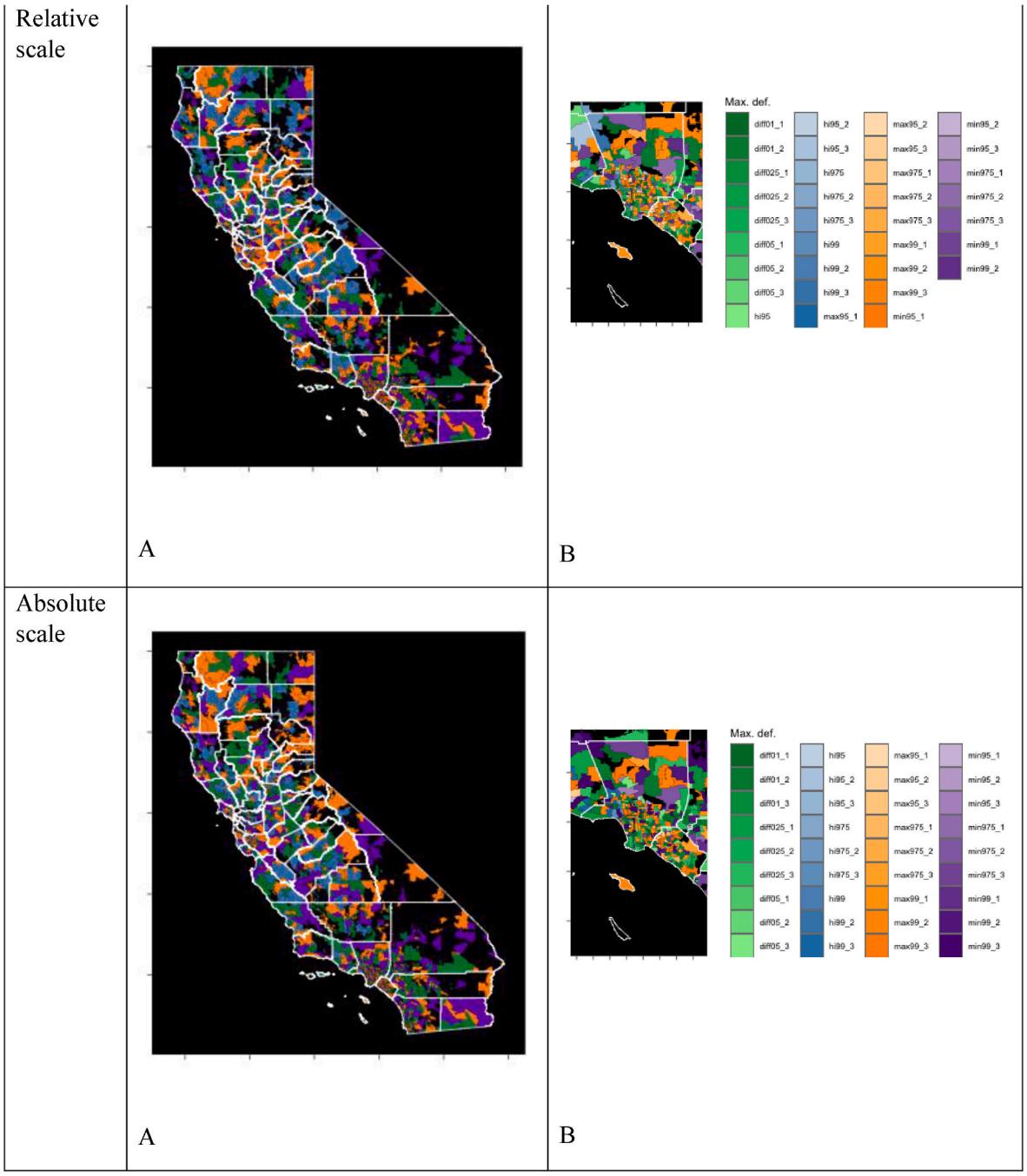

Fig. 5. The Extreme Heat Event (EHE) definitions that have the highest number of hospitalizations attributable to heat for each ZCTA on both the absolute and relative-to-population risk scales. Fill colors represent the different metrics: Green represents diurnal, orange represents maximum temperature, blue represents heat index, and purple represents minimum temperature heat events. Maps on the left are statewide considering only metric. On the right (Fig. 5B) is a portion of this map zoomed into the greater Los Angeles area showing more detail and each EHE definition, darker colors are more extreme. Dark areas correspond to areas with no observed hospitalization for this extreme heat event definition.

because of the lack of power at the zip code level. Future studies may also focus on the drivers of such spatial heterogeneity in the rankings of the most dangerous EHEs such as population characteristics, environmental factors, or the existence of implemented heat action plans.

5. Conclusion

In this study, we used a flexible spatio-temporal method of analysis to detect areas with the highest heat-related burden and to determine the EHE definitions that drive heat related hospitalizations in different regions of California. We observed high variability in the type of EHE definitions that were deemed most impactful depending on zip code. Nevertheless, by understanding heat burden at such a fine scale, we can improve current warning systems and guide policy toward the locations and vulnerable subgroups that are most adversely affected by specific types of extreme heat events.

CRediT authorship contribution statement

Kristen Hansen: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Armin Schwartzman:** Writing –

review & editing, Supervision, Methodology, Investigation, Conceptualization. **Lara Schwarz:** Writing – review & editing, Validation, Software, Investigation, Data curation. **Anais Teyton:** Writing – review & editing, Writing – original draft, Validation, Software, Resources, Methodology. **Rupa Basu:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Conceptualization. **Tarik Benmarhnia:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Data availability

The authors do not have permission to share data.

Acknowledgments

We would like to thank Maren Hale for her help in revising this manuscript. This research was supported by the Office of Environmental Health Hazard Assessment No. 18-E0012, the National Institute on Aging (RF1AG080948) and the National Science Foundation's Coastlines and People project, "Heat waves in the Southern California coastal zone: Their oceanic and atmospheric drivers, human health impacts, and sustainable adaptation" (2209058).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2024.119667.

References

Achebak, H., Rey, G., Chen, Z., Lloyd, S.J., Quijal-Zamorano, M., Méndez-Turrubiates, R. F., Ballester, J., 2024. Heat exposure and cause-specific hospital admissions in Spain: a nationwide cross-sectional study. Environmental Health Perspectives 132 (5), 057009. https://doi.org/10.1289/EHP13254.

Aguilera, R., Hansen, K., Gershunov, A., Ilango, S.D., Sheridan, P., Benmarhnia, T., 2020. Respiratory hospitalizations and wildfire smoke: a spatiotemporal analysis of an extreme firestorm in San Diego County, California. Environmental Epidemiology 4 (5), e114. https://doi.org/10.1097/EE9.0000000000000114.

Anderson, G.B., Bell, M.L., 2011. Heat waves in the United States: mortality risk during heat waves and effect modification by heat wave characteristics in 43 U.S. Communities. Environmental Health Perspectives 119 (2), 210–218. https://doi.org/10.1289/ehp.1002313.

Anderson, G.B., Dominici, F., Wang, Y., McCormack, M.C., Bell, M.L., Peng, R.D., 2013. Heat-related emergency hospitalizations for respiratory diseases in the Medicare population. Am. J. Respir. Crit. Care Med. 187 (10), 1098–1103. https://doi.org/10.1164/rccm.201211-1969OC.

Benmarhnia, T., Deguen, S., Kaufman, J.S., Smargiassi, A., 2015. Review article: vulnerability to heat-related mortality: a systematic review, meta-analysis, and meta-regression analysis. Epidemiology 26 (6), 781–793. https://doi.org/10.1097/EDE.0000000000000375.

Benmarhnia, T., Kihal-Talantikite, W., Ragettli, M.S., Deguen, S., 2017. Small-area spatiotemporal analysis of heatwave impacts on elderly mortality in Paris: A cluster analysis approach. Sci. Total Environ. 592, 288–294. https://doi.org/10.1016/j.scitotenv.2017.03.102.

Bennett, J.E., Blangiardo, M., Fecht, D., Elliott, P., Ezzati, M., 2014. Vulnerability to the mortality effects of warm temperature in the districts of England and Wales. Nat. Clim. Change 4 (4), 269–273.

California Energy Commission. California building climate zones (n.d. https://gis.data.ca.gov/documents/CAEnergy::building-climate-zones/. (Accessed 3 March 2020).

Chakraborty, T.C., Newman, A.J., Qian, Y., Hsu, A., Sheriff, G., 2023. Residential segregation and outdoor urban moist heat stress disparities in the United States. One Earth 6 (6), 738–750. https://doi.org/10.1016/j.oneear.2023.05.016.

Chen, C., Schwarz, L., Rosenthal, N., Marlier, M.E., Benmarhnia, T., 2024. Exploring spatial heterogeneity in synergistic effects of compound climate hazards: extreme heat and wildfire smoke on cardiorespiratory hospitalizations in California. Sci. Adv. 10 (5), eadj7264 https://doi.org/10.1126/sciadv.adj7264.

Chen, K., Huang, L., Zhou, L., Ma, Z., Bi, J., Li, T., 2015. Spatial analysis of the effect of the 2010 heat wave on stroke mortality in Nanjing, China. Sci. Rep. 5 (1), 10816 https://doi.org/10.1038/srep10816.

Chien, L.-C., Guo, Y., Zhang, K., 2016. Spatiotemporal analysis of heat and heat wave effects on elderly mortality in Texas, 2006-2011. Sci. Total Environ. 562, 845–851. https://doi.org/10.1016/j.scitotenv.2016.04.042.

Cole, R., Hajat, S., Murage, P., Heaviside, C., Macintyre, H., Davies, M., Wilkinson, P., 2023. The contribution of demographic changes to future heat-related health burdens under climate change scenarios. Environ. Int. 173, 107836 https://doi.org/10.1016/j.envint.2023.107836.

Do, V., Chen, C., Benmarhnia, T., Casey, J.A., 2024. Spatial heterogeneity of the respiratory health impacts of wildfire smoke PM2.5 in California. GeoHealth 8 (4), e2023GH000997. https://doi.org/10.1029/2023GH000997.

Ebi, K.L., Capon, A., Berry, P., Broderick, C., Dear, R. de, Havenith, G., Honda, Y., Kovats, R.S., Ma, W., Malik, A., Morris, N.B., Nybo, L., Seneviratne, S.I., Vanos, J., Jay, O., 2021. Hot weather and heat extremes: health risks. Lancet 398 (10301), 698–708. https://doi.org/10.1016/S0140-6736(21)01208-3.

Fard, P., Chung, M. K.Jake, Estiri, H., Patel, C.J., 2023. Spatio-temporal interpolation and delineation of extreme heat events in California between 2017 and 2021. Environ. Res. 237, 116984 https://doi.org/10.1016/j.envres.2023.116984.

Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Tobias, A., Tong, S., Rocklöv, J., Forsberg, B., Leone, M., Sario, M.D., Bell, M.L., Guo, Y.-L.L., Wu, C., Kan, H., Yi, S.-M., Coelho, M. de S.Z. S., Saldiva, P.H.N., et al., 2015. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. Lancet 386 (9991), 369–375. https://doi.org/10.1016/S0140-6736(14)62114-0.

Gasparrini, A., Masselot, P., Scortichini, M., Schneider, R., Mistry, M.N., Sera, F., Macintyre, H.L., Phalkey, R., Vicedo-Cabrera, A.M., 2022. Small-area assessment of temperature-related mortality risks in England and Wales: a case time series analysis. Lancet Planet. Health 6 (7), e557–e564. https://doi.org/10.1016/S2542-5196(22)00138-3.

Gershunov, A., Guirguis, K., 2012. California heat waves in the present and future. Geophys. Res. Lett. 39 (18) https://doi.org/10.1029/2012GL052979.

Goin, D.E., Gomez, A.M., Farkas, K., Zimmerman, S., Matthay, E.C., Ahern, J., 2019. Exposure to community homicide during pregnancy and adverse birth outcomes: a within-community matched design. Epidemiology 30 (5), 713–722. https://doi.org/10.1097/EDE.0000000000001044.

Guirguis, K., Basu, R., Al-Delaimy, W.K., Benmarhnia, T., Clemesha, R.E.S., Corcos, I., Guzman-Morales, J., Hailey, B., Small, I., Tardy, A., Vashishtha, D., Zivin, J.G., Gershunov, A., 2018. Heat, disparities, and health outcomes in San Diego County's diverse climate zones. GeoHealth 2 (7), 212–223. https://doi.org/10.1029/2017GH000127.

Guo, C., Ge, E., Lee, S., Lu, Y., Bassill, N.P., Zhang, N., Zhang, W., Lu, Y., Hu, Y., Chakraborty, J., Emeny, R.T., Zhang, K., 2023. Impact of heat on emergency hospital admission in Texas: geographic and racial/ethnic disparities. J. Expo. Sci. Environ. Epidemiol. https://doi.org/10.1038/s41370-023-00590-6.

Hondula, D.M., Barnett, A.G., 2014. Heat-related morbidity in brisbane, Australia: spatial variation and area-level predictors. Environmental Health Perspectives 122 (8), 831–836. https://doi.org/10.1289/ehp.1307496.

Hondula, D.M., Davis, R.E., Leisten, M.J., Saha, M.V., Veazey, L.M., Wegner, C.R., 2012. Fine-scale spatial variability of heat-related mortality in Philadelphia County, USA, from 1983-2008: a case-series analysis. Environ. Health 11 (1), 16. https://doi.org/10.1186/1476-069X-11-16.

Ingole, V., Marí-Dell'Olmo, M., Deluca, A., Quijal, M., Borrell, C., Rodríguez-Sanz, M., Achebak, H., Lauwaet, D., Gilabert, J., Murage, P., Hajat, S., Basagaña, X., Ballester, J., 2020. Spatial variability of heat-related mortality in barcelona from 1992–2015: a case crossover study design. Int. J. Environ. Res. Publ. Health 17 (7). https://doi.org/10.3390/ijerph17072553. Article 7.

Livneh, B., Rosenberg, E., Lin, C., Nijssen, B., Mishra, V., Andreadis, K., Maurer, E., Lettenmaier, D., 2013. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: update and extensions. J. Clim. 26, 9384–9392. https://doi.org/10.1175/JCLI-D-12-00508.1.

McElroy, S., Schwarz, L., Green, H., Corcos, I., Guirguis, K., Gershunov, A., Benmarhnia, T., 2020. Defining heat waves and extreme heat events using sub-regional meteorological data to maximize benefits of early warning systems to population health. Sci. Total Environ. 721, 137678 https://doi.org/10.1016/j.scitotenv.2020.137678.

Murage, P., Kovats, S., Sarran, C., Taylor, J., McInnes, R., Hajat, S., 2020. What individual and neighbourhood-level factors increase the risk of heat-related mortality? A case-crossover study of over 185,000 deaths in London using high-resolution climate datasets. Environ. Int. 134, 105292 https://doi.org/10.1016/j.envint.2019.105292.

Nitschke, M., Tucker, G.R., Hansen, A.L., Williams, S., Zhang, Y., Bi, P., 2011. Impact of two recent extreme heat episodes on morbidity and mortality in Adelaide, South Australia: a case-series analysis. Environ. Health 10 (1), 42. https://doi.org/10.1186/1476-069X-10-42.

Ostro, B.D., Roth, L.A., Green, R.S., Basu, R., 2009. Estimating the mortality effect of the July 2006 California heat wave. Environ. Res. 109 (5), 614–619. https://doi.org/10.1016/j.envres.2009.03.010.

Quijal-Zamorano, M., Martinez-Beneito, M.A., Ballester, J., Marí-Dell'Olmo, M., 2024. Spatial Bayesian distributed lag non-linear models (SB-DLNM) for small-area exposure-lag-response epidemiological modelling. Int. J. Epidemiol. 53 (3), dyae061 https://doi.org/10.1093/ije/dyae061.

Robine, J.-M., Michel, J.-P., Herrmann, F.R., 2012. Excess male mortality and age-specific mortality trajectories under different mortality conditions: a lesson from the

heat wave of summer 2003. Mech. Ageing Dev. 133 (6), 378–386. https://doi.org/10.1016/j.mad.2012.04.004.

Schinasi, L.H., Benmarhnia, T., De Roos, A.J., 2018. Modification of the association between high ambient temperature and health by urban microclimate indicators: a systematic review and meta-analysis. Environ. Res. 161, 168–180. https://doi.org/10.1016/j.envres.2017.11.004.

Schwarz, L., Hansen, K., Alari, A., Ilango, S.D., Bernal, N., Basu, R., Gershunov, A., Benmarhnia, T., 2021. Spatial variation in the joint effect of extreme heat events and ozone on respiratory hospitalizations in California. Proc. Natl. Acad. Sci. USA 118 (22), e2023078118. https://doi.org/10.1073/pnas.2023078118.

Sherbakov, T., Malig, B., Guirguis, K., Gershunov, A., Basu, R., 2018. Ambient temperature and added heat wave effects on hospitalizations in California from 1999 to 2009. Environ. Res. 160, 83–90. https://doi.org/10.1016/j.envres.2017.08.052.

Smargiassi, A., Goldberg, M.S., Plante, C., Fournier, M., Baudouin, Y., Kosatsky, T., 2009. Variation of daily warm season mortality as a function of micro-urban heat islands. J. Epidemiol. Community 63 (8), 659–664. https://doi.org/10.1136/jech.2008.078147.

Song, J., Yu, H., Lu, Y., 2021. Spatial-scale dependent risk factors of heat-related mortality: a multiscale geographically weighted regression analysis. Sustain. Cities Soc. 74, 103159 https://doi.org/10.1016/j.scs.2021.103159.

Vaidyanathan, A., Saha, S., Vicedo-Cabrera, A.M., Gasparrini, A., Abdurehman, N., Jordan, R., Hawkins, M., Hess, J., Elixhauser, A., 2019. Assessment of extreme heat and hospitalizations to inform early warning systems. Proc. Natl. Acad. Sci. U.S.A. 116 (12), 5420–5427. https://doi.org/10.1073/pnas.1806393116.

Vaneckova, P., Beggs, P.J., Jacobson, C.R., 2010. Spatial analysis of heat-related mortality among the elderly between 1993 and 2004 in Sydney, Australia. Soc. Sci. Med. 70 (2), 293–304. https://doi.org/10.1016/j.socscimed.2009.09.058.