THE LANCET
Planetary Health

## Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Masselot P, Mistry M, Vanoli J, et al. Excess mortality attributed to heat and cold: a health impact assessment study in 854 cities in Europe. *Lancet Planet Health* 2023; published online March 16. https://doi.org/10.1016/ S2542-5196(23)00023-2.

# Excess mortality attributed to heat and cold: a health impact assessment study in 854 cities in Europe

## Appendix

**Table of Contents**

## A. Data

### A.1. City selection

The initial list of cities was taken from the Urban Audit dataset from Eurostat,[1] that contains 870 cities. From this list, we discarded 16 overseas cities that were far and with a very different climate from the main continent (Supplementary Table **S1**). The description of the final sample of 854 cities is provided in Table 1 of the main manuscript.

**Table S1:** List of overseas cities in the Urban Audit dataset discarded from the analysis.

| Country | City |
|---|---|
| France | Saint Denis, Fort-de-France, Mamoudzou, Cayenne, Saint-Louis |
| Iceland | Reykjavik |
| Portugal | Funchal, Ponta Delgada |
| Spain | Las Palmas, Sante Cruz de Tenerife, Telde, Ceuta, Melilla, Arrecife, Santa Lucía de Tirajana, Puerto de la Cruz |

**A.2. Description of the MCC dataset**

We linked the 854 selected cities to locations of the Multi-Country Multi-City (MCC) dataset. To perform the linkage, we looked at the intersection between the city polygons from Eurostat and a buffer of 10 km around the MCC point locations. When there were several matches, we selected the larger intersection. Eventually, a manual check was performed to correct mismatches. A display of the 232 cities with MCC counterparts is shown in Supplementary Figure **S1**, a summary of the MCC dataset is given in Supplementary Table **S2** and a description of country-specific mortality datasets is given below. Analysis on the MCC data was restricted to years 1990 to 2019 to avoid the COVID-19 years.

**Cyprus**

Daily mortality is represented by counts of all-cause deaths collected by the Health Monitoring Unit of the Ministry of Health of Cyprus. The ideas and opinions expressed herein are those of the author. Endorsement of these ideas and opinions by the Ministry of Health of Cyprus is not intended nor should it be inferred.

**Czechia**

Daily mortality is represented by counts of all-cause deaths obtained from the Czech Statistical Office and the Institute of Health Information and Statistics.

**Estonia**

Daily mortality is represented by counts of deaths for non-external causes (ICD-9: 0-799; ICD10: A00-R99) obtained from Estonian Causes of Death Registry.

**Finland**

Data were collected from the Helsinki Metropolitan Area between 1st of January 1994 and 31st of December 2014. Daily number of deaths were obtained from Statistics Finland and are represented by counts of deaths for nonexternal causes (ICD-9: 0-799; ICD-10: A00-R99).

**France**

Daily mortality is represented by counts of all-cause deaths provided by the French National Institute of Health and Medical Research (CepiDC).

**Germany**

Daily mortality, obtained from the Research Data Centres of the Federation and the Federal States of Germany (Forschungsdatenzentrum der Statistischen Ämter des Bundes und der Länder), is represented by counts of deaths for all causes.

**Greece**

Daily mortality is represented by counts of all-cause deaths provided by the Hellenic Statistical Authority.

**Ireland**

Daily mortality, provided by the Irish Central Statistics Office Northern and Ireland Social Research Agency, is represented by counts of deaths for non-external causes only (ICD-9: 0-799; ICD-10: A00-R99).

**Italy**

Daily mortality is represented by counts of all-cause deaths obtained from local mortality registries and from the rapid mortality surveillance system.

**Netherlands**

Daily mortality is represented by counts of all-cause deaths provided by Statistics Netherlands.

**Norway**

Daily mortality is represented by all-cause counts of deaths and is provided by the Norwegian Cause of Death registry.

**Portugal**

Daily mortality, obtained from Statistics Portugal, is represented by counts of deaths for non-external causes only (ICD-9: 0-799; ICD-10: A00-R99).

**Romania**

Daily mortality, obtained from the National Institute for Statistics (NIS) in Romania, is represented by counts of deaths for all causes. The mortality dataset includes the decedents with the stable residence (permanent) or normal residence (defined as the place/city where a person lived mostly in the last 12 months of his/her life) in the seven Romanian cities.

**Spain**

Daily mortality, obtained from Spain National Institute of Statistics, is represented by counts of deaths for nonexternal causes (ICD-9: 0-799; ICD-10: A00-R99).

**Sweden**

Daily mortality is represented by counts of all-cause deaths obtained from the Swedish Cause of Death Register at the Swedish National Board of Health and Welfare.

**Switzerland**

Daily mortality, provided by the Federal Office of Statistics (Switzerland), is represented by counts of non-external deaths other than accidents (ICD-10codes A00-R99, V01-V99, W00-X59).

**United Kingdom**

Daily mortality, gathered from Office for National Statistics, is represented by counts of deaths for all causes.

**Table S2:** Description of the linkage between the Urban Audit and the MCC datasets.

| Region | Country | Cities in Urban Audit | Cities in MCC | Total deaths in MCC cities | Available age groups | MCC data period |
|---|---|---|---|---|---|---|
| Northern | Denmark | 4 | 0 (0%) | - | - | - |
| | Estonia | 3 | 3 (100%) | 136,371 | - | 1997 - 2018 |
| | Finland | 9 | 1 (11%) | 130,395 | 00-64; 65-99 | 1994 - 2011 |
| | Ireland | 5 | 1 (20%) | 185,624 | - | 1990 - 2007 |
| | Latvia | 10 | 0 (0%) | - | - | - |
| | Lithuania | 6 | 0 (0%) | - | - | - |
| | Norway | 4 | 1 (25%) | 135,159 | 00-74; 75-99 | 1990 - 2016 |
| | Sweden | 14 | 3 (21%) | 760,527 | 00-14; 15-64; 65-74; 75-84; 85-99 | 1990 - 2016 |
| | United Kingdom | 135 | 103 (76%) | 6,494,919 | 00-44; 45-64; 65-74; 75-84; 85-99 | 1990 - 2019 |
| Western | Austria | 6 | 0 (0%) | - | - | - |
| | Belgium | 15 | 0 (0%) | - | - | - |
| | France | 72 | 18 (25%) | 1,753,573 | 00-64; 65-99 | 2000 - 2015 |
| | Germany | 127 | 12 (9%) | 3,105,865 | - | 1993 - 2015 |
| | Luxembourg | 1 | 0 (0%) | - | - | - |
| | Netherlands | 47 | 5 (11%) | 453,395 | - | 1995 - 2016 |
| | Switzerland | 12 | 8 (67%) | 243,638 | 00-64; 65-74; 75-84; 85-99 | 1995 - 2013 |
| Eastern | Bulgaria | 18 | 0 (0%) | - | - | - |
| | Czechia | 18 | 3 (17%) | 534,382 | 00-29; 30-59; 60-74; 75-99 | 1994 - 2019 |
| | Hungary | 19 | 0 (0%) | - | - | - |
| | Poland | 68 | 0 (0%) | - | - | - |
| | Romania | 35 | 8 (23%) | 951,146 | - | 1994 - 2016 |
| | Slovakia | 8 | 0 (0%) | - | - | - |
| Southern | Croatia | 7 | 0 (0%) | - | - | - |
| | Cyprus | 3 | 3 (100%) | 67,217 | 00-44; 45-64; 65-74; 75-84; 85-99 | 2004 - 2019 |
| | Greece | 14 | 1 (7%) | 287,969 | 00; 01-14; 15-64; 65-74; 75-84; 85-99 | 2001 - 2010 |
| | Italy | 87 | 16 (18%) | 645,420 | - | 2001 - 2010 |
| | Malta | 1 | 0 (0%) | - | - | - |
| | Portugal | 14 | 2 (14%) | 807,618 | 00-64; 65-99 | 1990 - 2012 |
| | Slovenia | 2 | 0 (0%) | - | - | - |
| | Spain | 90 | 44 (49%) | 588,840 | 00-04; 05-14; 15-44; 45-64; 65-74; 75-99 | 2009 - 2013 |

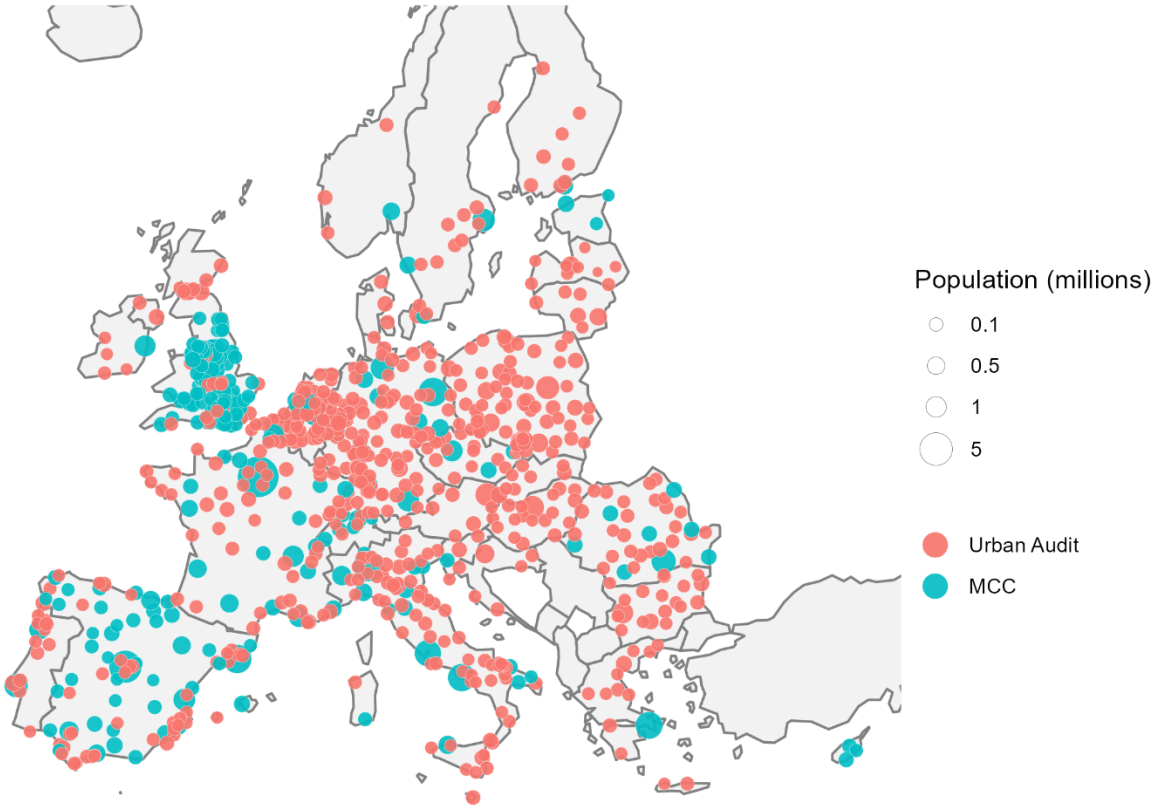


**Figure S1:** Map of the 854 Urban audit cities, including the subset of 232 MCC cities.

### A.3. City-level characteristics

This section provides details on the city-level characteristics. Supplementary Table S3 provides details on the variables used for impact measurement, while Supplementary Table S4 describes the variables used to derive the composite meta-predictors in the second-stage model through PLS (see Section B).

Data sources were searched in the following nested order of geographical aggregation: Urban Audit that exactly matches the cities, NUTS3 level that is usually representative of cities, NUTS2 level that corresponds to larger regions and can be less representative of the smallest cities. Environmental variables that were unavailable in Eurostat or with too many missing values were extracted from various external datasets. In this case, we extracted all pixels within cities boundaries and averaged their series.

Finally, since periods of availability differed widely according to variables, cities, and regions, for each city and each variable we averaged all available years between 2000 and 2020. The availability of each meta-variable is shown in Supplementary Figure **S2**.

**Table S3:** List of other other city-level variables.

| Variable | Source | Description | Age groups |
|---|---|---|---|
| Death rate | NUTS3 | Death rate within age group | 00-04; 05-09; 10-14; 15-19; 20-24; 25-29; 30-34; 35-39; 40-44; 45-49; 50-54; 55-59; 60-64; 65-69; 70-74; 75-79; 80-84; 85-99; Total |
| Life expectancy | NUTS2 | Life expectancy at age | Birth; 5; 10; 15; 20; 25; 30; 35; 40; 45; 50; 55; 60; 65; 70; 75; 80; 85 |
| Population structure | NUTS3 | Proportion of the population in age group | 00-04; 05-09; 10-14; 15-19; 20-24; 25-29; 30-34; 35-39; 40-44; 45-49; 50-54; 55-59; 60-64; 65-69; 70-74; 75-79; 80-84; 85-99 |

**Table S4:** List of city-level meta-predictors.

| Variable | Source | Description |
|---|---|---|
| Total population | Urban Audit / Wikipedia | Note: population data were missing for several cities and were added from the Wikipedia pages for each city (Latvia: Valmiera; Belgium: Mechelen, Mouscron, La Louvière, Verviers) |
| Population above 65 | Urban Audit / NUTS3 | Percentage |
| Population density | Urban Audit / Wikipedia | Note: population density data were missing for several cities and were added from the Wikipedia pages for each city (Latvia: Valmiera; Belgium: Mechelen, Mouscron, La Louvière, Verviers; Ireland: Dublin, Limerick, Waterford; Norway: Bergen, Trondheim; UK: Glasgow, North Lanarkshire, Dundee) |
| Life expectancy | NUTS2 | At birth |
| Isolation | Urban Audit / NUTS2 | Proportion of single-person households |
| GDP | NUTS3 / Office of National Statistics | GDP per capita.<br>Note: GDP data for UK cities were unavailable in Eurostat and extracted from ONS |
| Unemployment rate | NUTS2 | Among active population (20-64 years old) for all education levels |
| Education level | NUTS2 | Proportion of active population (25-64 years old) with ISCED level >= 5 (higher education) |
| Deprivation rate | NUTS2 | Proportion of population under severe material deprivation condition |
| Hospital bed rates | NUTS2 | Number of hospital beds / inhabitants |
| Imperviousness | Copernicus High Resolution Layer | Percentage of soil sealing |
| Tree Cover Density | Copernicus High Resolution Layer | Level of Tree Cover Density (%) |
| Grassland | Copernicus High Resolution Layer | Proportion of grassland pixels |
| Water & Wetness | Copernicus High Resolution Layer | Average class between (1) permanent water, (2) temporary water, (3) permanent wetness and (4) temporary wetness |
| Small Woody Features | Copernicus High Resolution Layer | Small woody features density (%) |
| Elevation | AWS Terrain Tiles | Elevation at city centre |
| Coastal region type | Natural Earth | Lowest distance between coastal line and city centre |
| NDVI | Google Earth Engine - MODIS | |
| PM25 | Atmospheric Composition Analysis Group | https://sites.wustl.edu/acag/datasets/surface-pm2-5/ |
| NO2 | Atmospheric Composition Analysis Group | https://sites.wustl.edu/acag/datasets/surface-no2/ |
| Temperature range | Copernicus | Anuual temperature range |
| Mean temperature | Copernicus | Mean annual temperature |

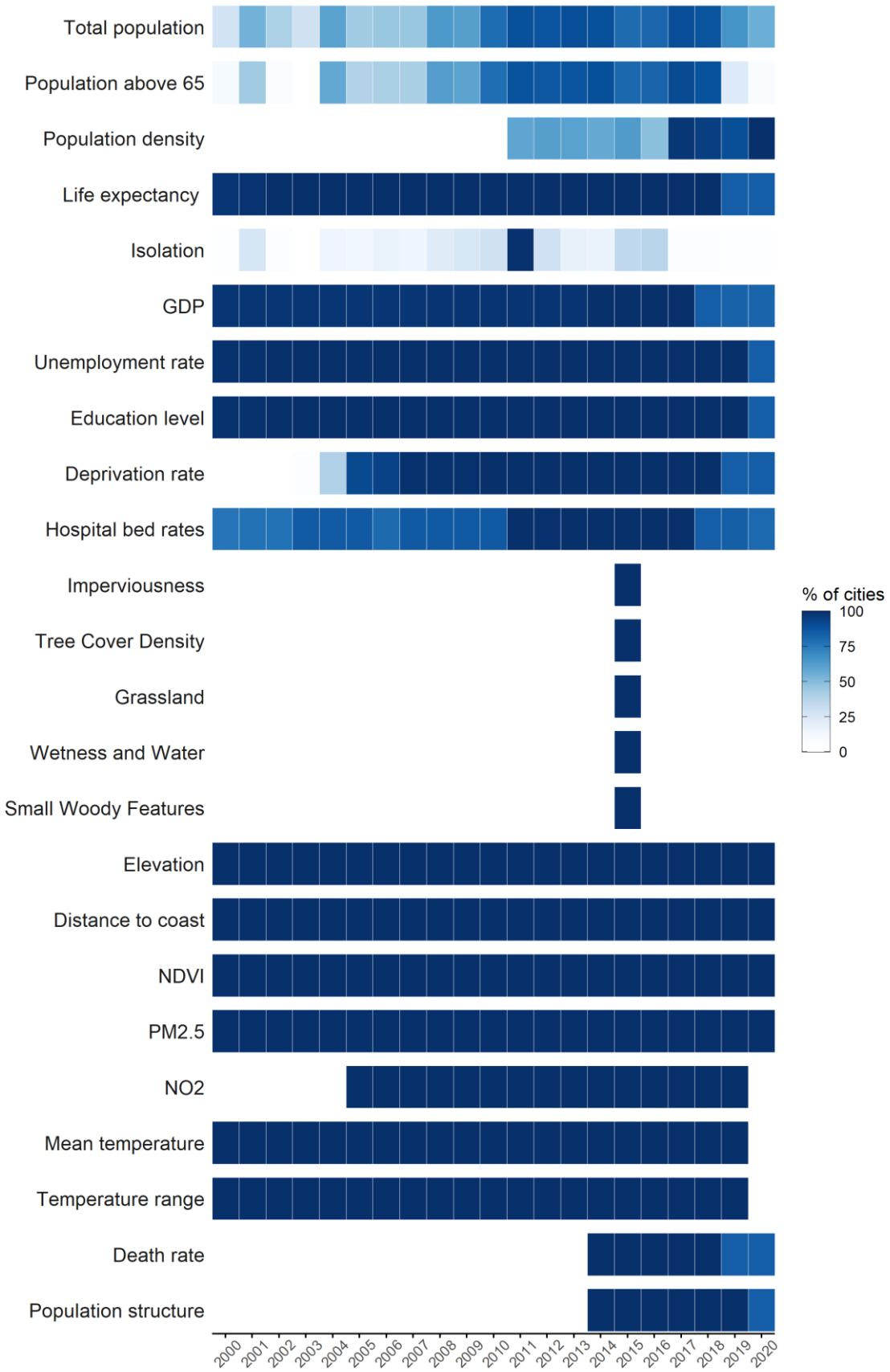


**Figure S2:** Proportion of cities with available values for each year and each meta-variable.

## B. Modelling details

We adopted a three-stage analysis design: i) city and age-group specific estimation of the overall cumulative exposure-response functions (ERF) in MCC cities, ii) creation of a predictive model by meta-regressing the first-stage ERF coefficients on age and the PLS components computed from 22 meta-predictors, and iii) prediction of risks to the whole list of cities using the second-stage model and health impact assessment. The steps of the analysis are illustrated in Figure S3.

In the first stage, we estimated age group-specific ERF through a quasi-Poisson time series model for each of the 232 MCC cities. We applied a distributed lag nonlinear model (DLNM)[2] with standard parameterization.[3] In the temperature dimension, we specified a quadratic B-spline with knots positioned at the 10$^{th}$, 75$^{th}$ and 90$^{th}$ percentile of the city-specific temperature distribution, and in the lag dimension a cubic natural spline with three knots positioned at equally-spaced log-values using a lag period of 21 days. Knots are positioned at percentiles to allow differences in risk that account for adaptation to the local climate. We also included indicators for the day of the week and a natural spline of time with seven degrees of freedom per year to control for time-varying confounding. Once the model was fit, we extracted the vector of five coefficients $\widehat{\boldsymbol{\theta}}_{ij}$ representing the overall cumulative exposure response function for age group $j$ of city $i$.[4]

In the second stage, we pooled the age-specific reduced first stage coefficients in a mulitvariate multilevel meta-regression[5] model as follows:

$$\widehat{\boldsymbol{\theta}}_{ij} = \boldsymbol{\gamma}_{r(i)} + ns(a_{ij}) + \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{b}_i + \boldsymbol{\varepsilon}_{ij} \quad (1)$$

where $\boldsymbol{\gamma}_{r(i)}$ is a term representing the region of city $i$, $ns(a_{ij})$ is a natural spline of the age $a_{ij}$ associated to $\widehat{\boldsymbol{\theta}}_{ij}$, $\boldsymbol{X}_i$ is a set of ten city-specific composite indices of vulnerability with $\boldsymbol{\beta}$ the associated coefficients, $\boldsymbol{b}_i$ is a city-specific random effect and $\boldsymbol{\varepsilon}_{ij}$ is the residual of the model. This model allows each coefficient in $\widehat{\boldsymbol{\theta}}_{ij}$ to be affected differently by the components of model (1).

As there was still residual heterogeneity once model (1) was fitted, in a fourth step we extracted the best linear unbiased prediction (BLUP) of the multivariate residuals $\widehat{\boldsymbol{b}}_i$ and interpolated them through kriging. The details of each of these four steps of model building are given in subsections B.1. to B.4. below.

Finally, the model built in the second stage was used to predict exposure-response functions (ERF) at all locations and for age groups: 20-44, 45-64, 65-74, 75-84 and 85+. These city-age ERFs were then used to assess the health impact of non-optimal temperature through the computation of the excess number of deaths, as well as raw and standardized excess mortality rates. Subsection B.5. below gives details on the quantification of the health impact.

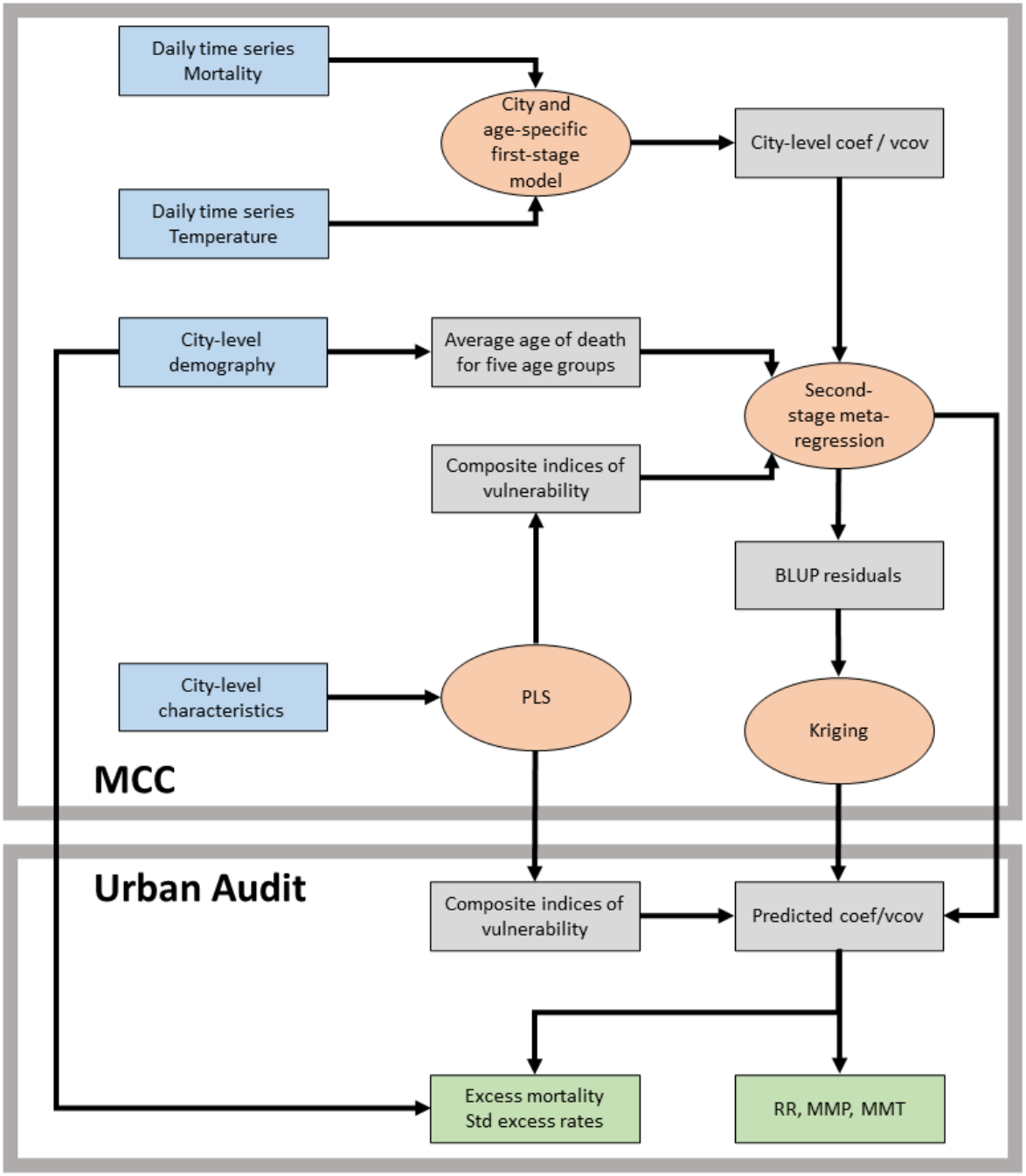


**Figure S3:** Flowchart describing the modelling framework. Squares represent input/output data while ellipses represent models/processes. Blue represents input data, grey derived data and green output data.

**B.1. Spatial background**

To represent a spatial background, we included indicators $\boldsymbol{\gamma}_{r(i)}$ in equation (1) representing the four European regions as defined by the United Nations M49 standard[6] (see Supplementary Figure **S4**). Note that to simplify the prediction, we included the regional factor with Helmert contrasts[7], where the intercept of the model represents the average effect across regions. This allowed us to simply discard the $\boldsymbol{\gamma}_{r(i)}$ coefficients to predict average risks encompassing all regions (such as those in Figure 1 of the main manuscript).

The regional background as described above was chosen over other alternative definitions, obtained through: i) separate natural splines (with 2df) of latitude and longitude, ii) a bi-dimensional spline (with 2x2df) of latitude and longitude, and iii) indicators of the Köppen-Geiger climate classification. The background based on UN regions was the most balanced and led to the model with the lowest AIC, as shown in Supplementary Table **S5**.

Region-specific exposure-response functions are shown in Supplementary Figure **S5**. It shows that relative risks associated with extreme cold are higher in the southern and western region. Risks associated with heat are lower in the southern region. The MMP is also lower in the eastern region, resulting in higher overall effect of heat.

**Table S5:** Comparison of the Akaike Information Criterion (AIC) for second-stage meta-regression models using different definitions of regional backgrounds.

| Background type | Model term | AIC |
|---|---|---|
| Null model | None | 3283 |
| Regions | $\boldsymbol{\gamma}_{r(i)}$ with $r(i)$ corresponding the UN M49 region | **3169** |
| Latitude and Longitude | $s(lat) + s(lon)$ | 3182 |
| Latitude and Longitude with interaction | $s(lat, long)$ | 3191 |
| Köppen-Geiger Climate classification | $\boldsymbol{\gamma}_{kg(i)}$ with $kg(i)$ corresponding to the climate class | 3249 |

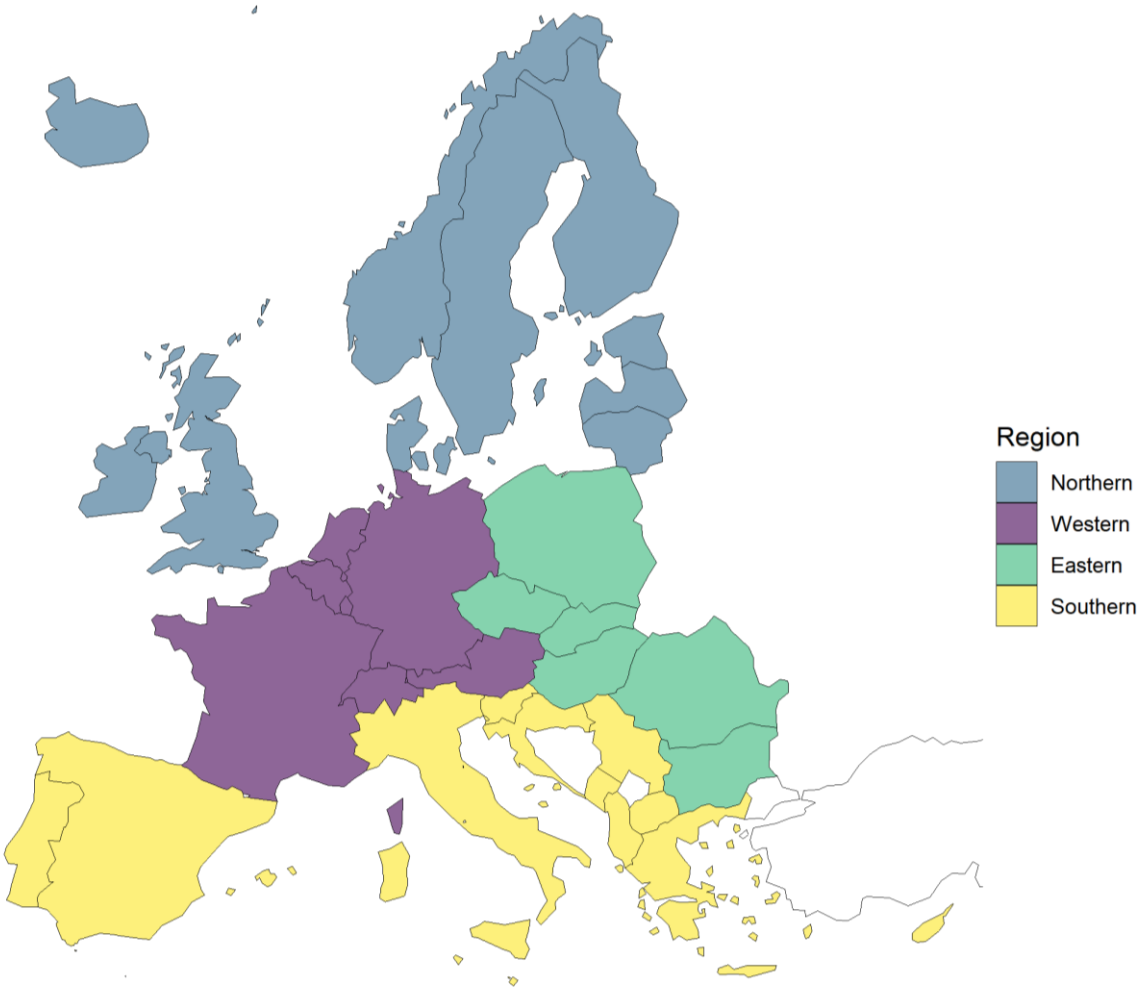


**Figure S4:** Map of the four European United Nations M49 regions.

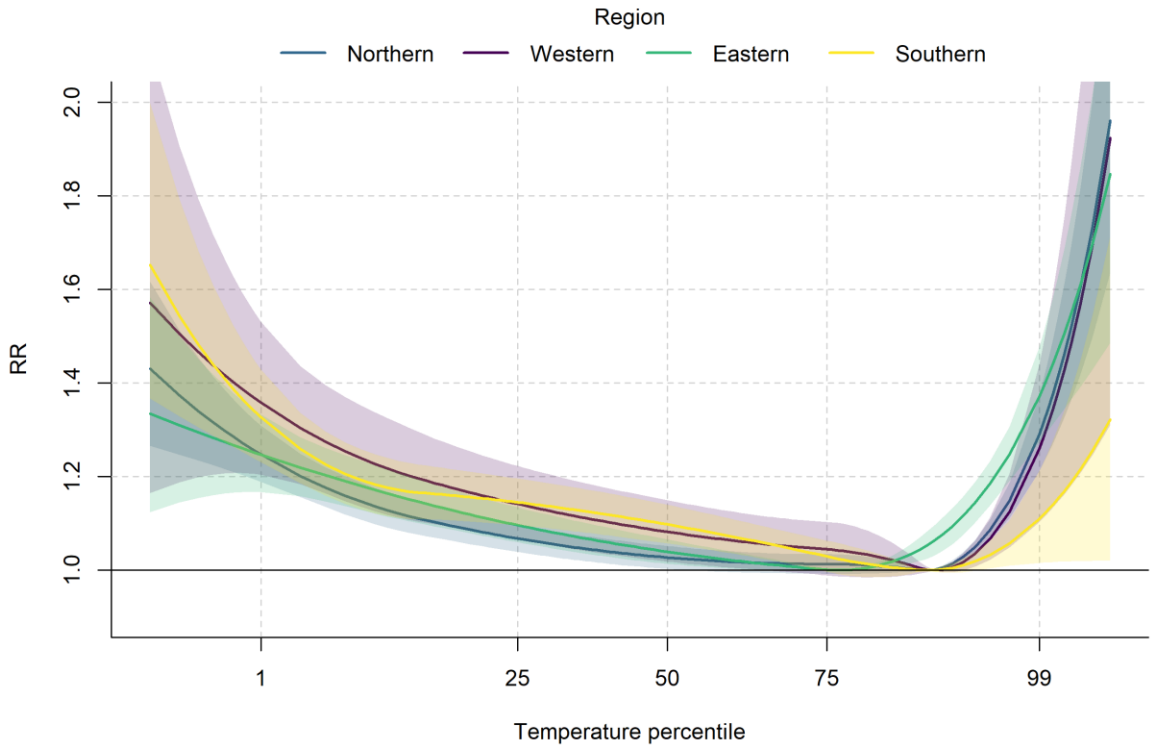


**Figure S5:** Region-specific exposure-response functions between temperature and mortality.

**B.2. Modelling differential risks by age**

The first-stage time series regression model was performed on each age group in the MCC dataset, with age groups differing by country (see Supplementary Table **S2**). However, the number of deaths recorded can be very low for the youngest age groups in smaller cities, resulting in unstable first-stage estimates and low statistical power.[8] Therefore, when the total death count in the series is below 5000 for an age group, we aggregated it with the older one when available.

Because the age groups differ between countries in the MCC dataset, we included age as a continuous age variable $a_{ij} \in [0; 100]$ in the second-stage model. For each city $i$ and each age group $j$ available in the first stage, we computed an associated *average age of death* $a_{ij}$ as a simple weighted mean of the ages within the group, i.e:

$$a_{ij} = \frac{\sum_{k=l}^{u} w_{ik} k}{\sum_{k=l}^{u} w_{ik}} \quad (2)$$

where $l$ and $u$ are the age boundaries of the group and $w_{ik}$ are the death rates associated with age $k$ in city $i$. Note that for the oldest age group, we instead used the life expectancy at the lower bound (*e.g.* at 85 for the group 85 and older). Death rates and life expectancy are extracted for each city from Eurostat (see Supplementary Table S3).

The average ages of death $a_{ij}$ were then included as a continuous linear term in the second stage model. Nonlinear terms of age through a natural spline with various knot placements were also investigated but resulted in highest AIC compared to the linear term, as shown by Supplementary Figure S6.

For predictions at age groups in the third stage, we repeated the process. For each city and each age group considered, we computed the average age of death within this age group using equation (2), or alternatively we used life expectancy at 85 for age group 85 and older.

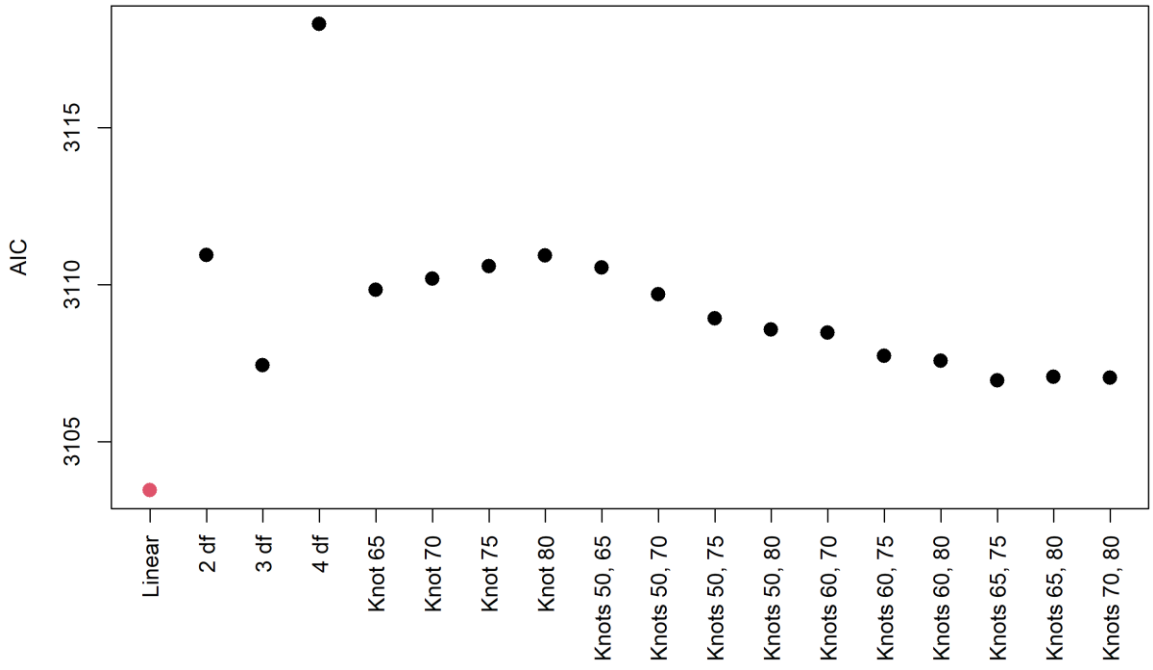


**Figure S6:** Akaike Information Criterion of the second-stage meta-regression model with different age term specifications. Red indicates the specification with the lowest AIC. 'df' refers to degrees of freedom, namely a specification in which the knots are regularly placed along with the age domain.

**B.3. Partial least-squares components**

We accounted for a full list of 22 variables $\boldsymbol{Z}_i$ representing diverse socio-economic and environmental characteristics. However, many of these variables are highly correlated as shown in Supplementary Figure S7. We therefore used Partial Least Squares (PLS)[9] to construct uncorrelated composite indices of vulnerability of the form $\boldsymbol{X}_i = \boldsymbol{Z}_i\boldsymbol{R}$ with the weights in $\boldsymbol{R}$ chosen such that that their variance and correlation with $\widehat{\boldsymbol{\theta}}_{\boldsymbol{ij}}$ are both maximal, and all components in $\boldsymbol{X}_{\boldsymbol{i}}$ are independent from each other. In other words, the weight matrix $\boldsymbol{R}$ solve the following optimisation problem:[10]

$$\widehat{\boldsymbol{R}} = \text{argmax corr}\left(\widehat{\boldsymbol{\theta}}_{ij}, \boldsymbol{X}_i\right)\text{var}(\boldsymbol{X}_i) \quad (3)$$

subject to the constraint that the components in $\boldsymbol{X}_i$ are uncorrelated from each other. This optimisation problem is solved iteratively, by extracting the coefficients of the first component from a singular value decomposition (SVD) of the covariance matrix between $Z_i$ and $\widehat{\boldsymbol{\theta}}_{ij}$, then the coefficients of the second component by the same process on the residuals left by the first SVD, and so on.[11]

Note that the PLS is equivalent to applying a shrinkage penalisation to the coefficients applied to each of the individual variables in $\boldsymbol{Z}_i$, with the shrinkage factor depending on the number of components of $\boldsymbol{X}_i$ included in the model.[10] Therefore, the effect of individual factors is biased in order to reduce the variance induced by the high-dimensionality and correlation between factors. We chose this number of components at 4 by minimizing the AIC, as shown in Supplementary Figure S8. These four components are used as composite indicators of vulnerability in the second-stage meta-regression, and are meant to capture the geographical variability in risks across the European cities.

Supplementary Figure **S9** shows the correlation between each of the select composite indices of vulnerability and each meta-predictor, Supplementary Figure **S10** displays the spatial distribution of the PLS components, and Supplementary Figure S11 illustrates the difference in risks at the higher and lower ends of each component distribution. The first component indicates decreased risk of heat and cold for locations with higher hospital bed rates and temperature range and lower amounts of small woody features. The second components is mostly associated with smaller, poorer and humid cities, with highly increased risks of cold but slightly decreased risks of heat. The third component represents highly populated cities with low amounts of small woody features, mostly common in the Mediterranean area. It is associated with highly increased risks of cold and marginally increased risks of heat. Finally, component 4 represents smaller and greener cities most common in northern Europe and Germany. It is associated with increased MMP, although with steeper effect of heat and lower risks of cold.


**Figure S7:** Correlation matrix between meta-predictors used in PLS.

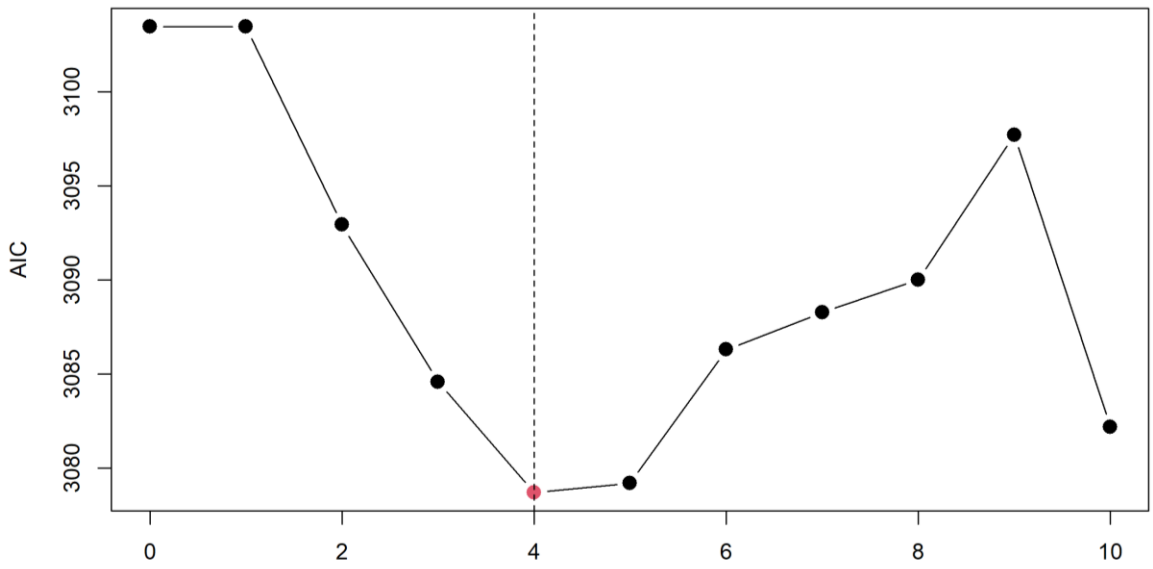


**Figure S8:** Akaike Information Criterion of the second-stage meta-regression model with a different number of PLS components in the model.

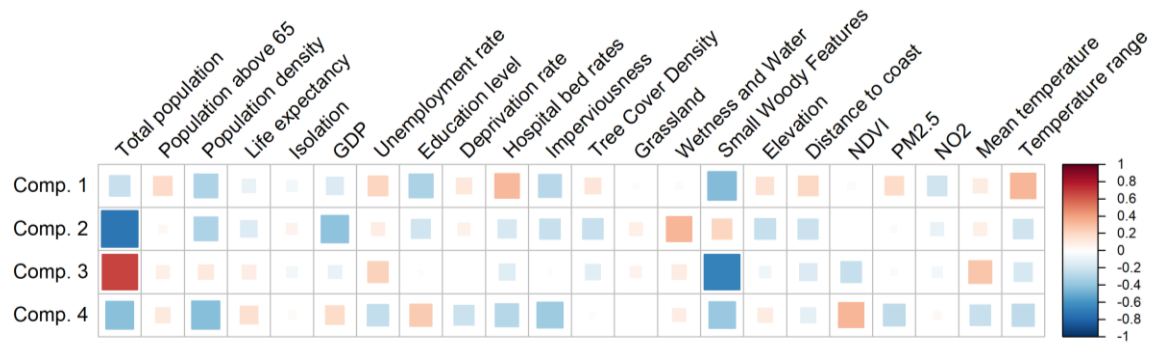


**Figure S9:** Correlation between PLS components and individual meta-predictors.

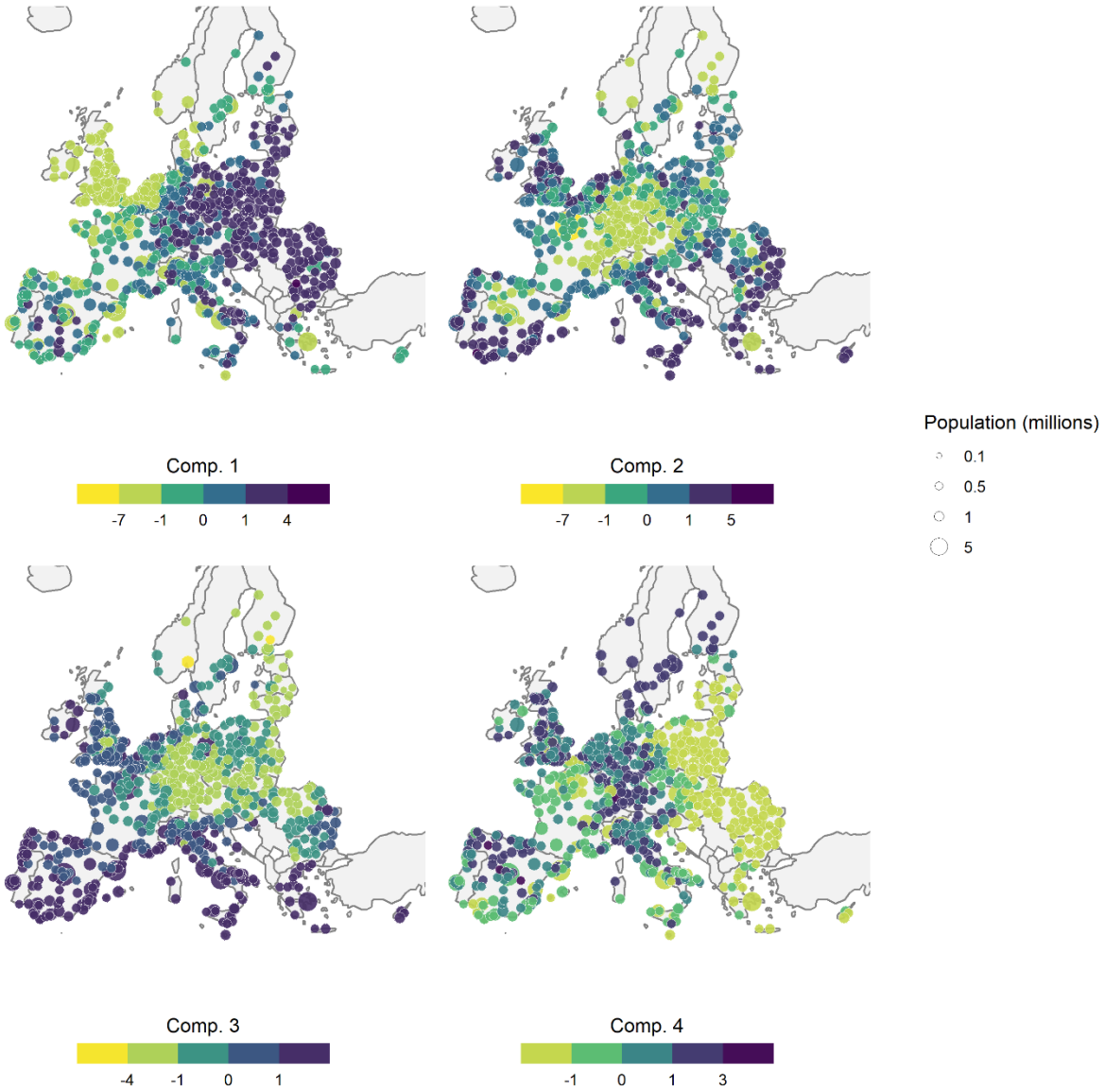


**Figure S10:** Geographical distribution of the values of each PLS component used in the second stage meta-regression.

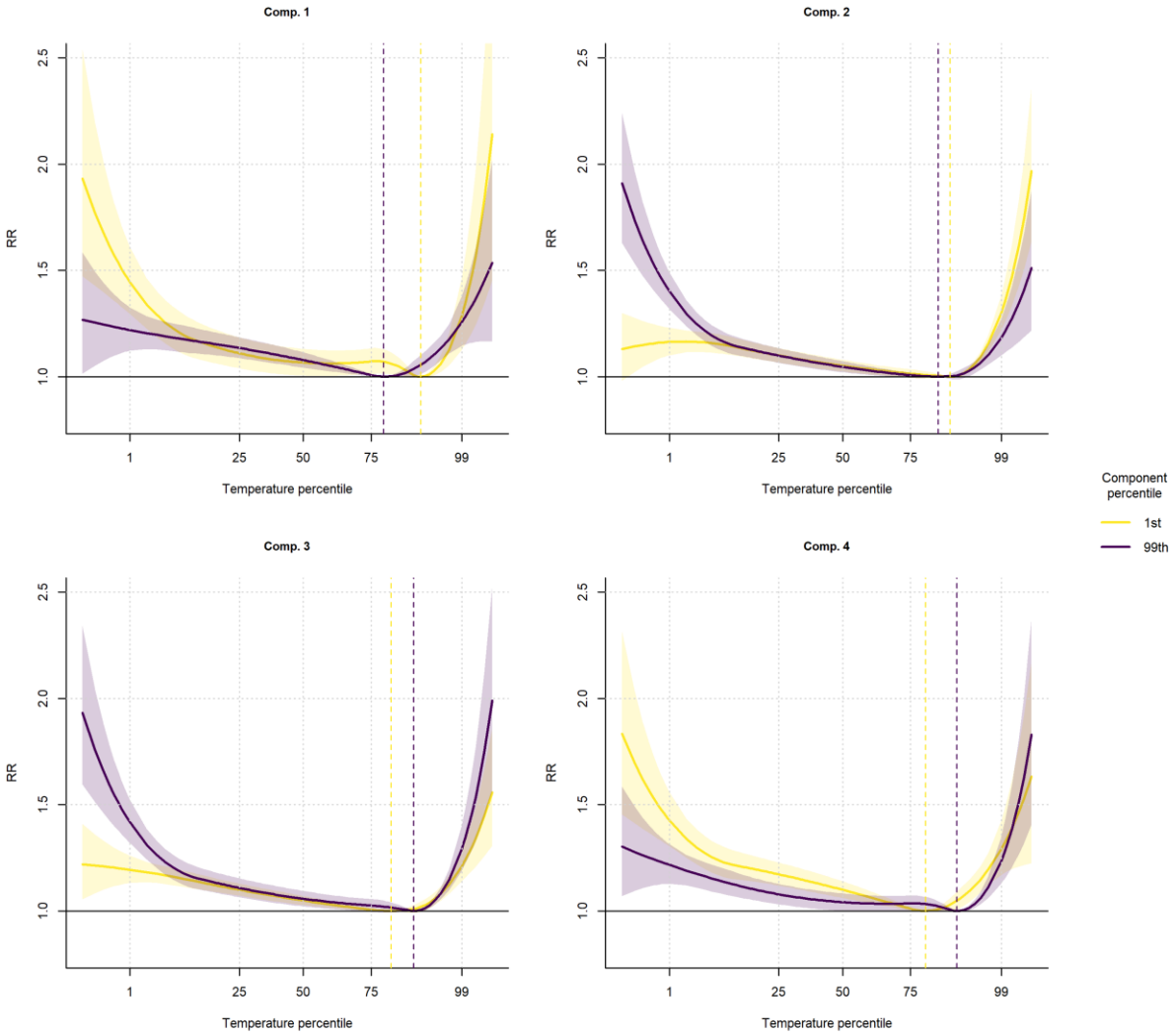


**Figure S11:** Predicted exposure-response relationship at extreme percentiles of the distribution of each PLS component.

**B.4. Kriging interpolation of the random effects**

Supplementary Figure **S12** shows the fixed-effects predictions versus the best linear unbiased predictions (BLUP) from the second-stage model (1), indicating some divergences. This can lead to biases in the predicted risks across geographical areas. An example is shown for the city of Munich in Supplementary Figure **S13**, for which the prediction substantially over-estimates risk associated with heat. This divergence captured by the random effects can represent geographical clusters of vulnerability that are not explained by the PLS components and the spatial background. We therefore added an additional spatial interpolation step to model such local deviations in risks, using BLUP residuals $\widehat{\boldsymbol{b}}_i$ from MCC locations.

Spatial interpolation was performed by kriging. Briefly, assuming for now that $b_i$ is univariate, Kriging predicts $b_i$ at any point in a spatial domain as a linear combination of all measured values. i.e. the kriged surface at a new location $I$ is:

$$\hat{b}_I = \sum_i w_{Ii} b_i \tag{4}$$

It was shown that the best linear unbiased prediction $\hat{b}_I$ uses weights defined as

$$\boldsymbol{w}_I = [w_{I1}, \dots, w_{In}] = v^T V^{-1} \tag{5}$$

where $v$ is the vector containing the covariance of between $b_I$ and all observed locations $b_i$, and $V$ is the (co)variance matrix of the observed $b_i$.[12] Note that formulae (4) and (5) assume that there is no trend in $b_i$ and that residuals are Gaussian. Both assumptions are respected by the definition of random effects $b_i$ that are assumed Gaussian with null mean by design.[5] The formulas (4) and (5) can straightforwardly be extended to the case of multivariate $\boldsymbol{b}_i$.[13]

As there is only one realization of $b_i$ at each observed location $i$, the covariances in $v$ and $V$ cannot be directly estimated. Therefore, ordinary Kriging assumes that the covariance between two locations $i$ and $j$ follows a covariance function $\gamma(h)$ that depends only on the distance $h$ between the two locations. This covariance function $\gamma(h)$ can then be estimated by a variogram, which measures the variance of the difference of measurement according to their distance.[14] A parametric function can then be fitted to the variogram, and then be used to define the values in $v$ and $V$ in (5). Note that in the case of multivariate $\boldsymbol{b}_i$ several variogram have to be fitted, one for each dimension in $\boldsymbol{b}_i$ and one (co)variogram for each pair of dimensions in $\boldsymbol{b}_i$.

Supplementary Figure S14 shows the empirical (co)variogram for $\boldsymbol{b}_i$. These variogram displays important nugget effects (the variogram at distance 0) and ranges of around 500 km, indicating smooth surfaces.[15] This is confirmed by the kriged surfaces shown in Supplementary Figure **S15**. It also shows that the fifth dimension, corresponding to the last coefficient of the ERF, is the one with the highest magnitude, i.e. the residual spatial heterogeneity is more important for heat.

Once kriging had been performed, we extracted the predictions of $\widehat{\boldsymbol{b}}_i$ for each city in the Urban Audit dataset. This local deviation in risk was then added to the prediction made by the second-stage meta-regression model for each city. As fixed and random effects are independent by design in the mixed-effects framework[5], the final covariance matrix of predicted spline coefficients is simply the sum of covariances matrices from the meta-regression model and the kriging. Note that Kriging is an exact interpolator, i.e. the prediction at observed locations (here MCC locations) is equal to the observed value.[15] Therefore, at these locations the result is exactly equal to the BLUP of the second-stage meta-regression model.

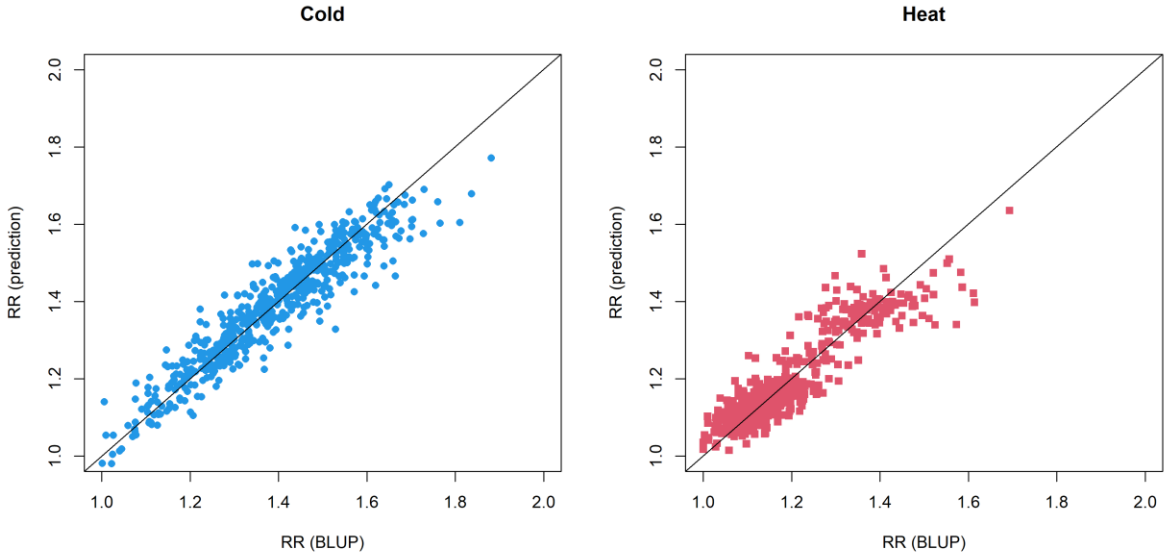


**Figure S12:** Relative risks (RRs) for cold (1st percentile) and heat (99th percentile) obtained as fixed-effects predictions versus BLUP.

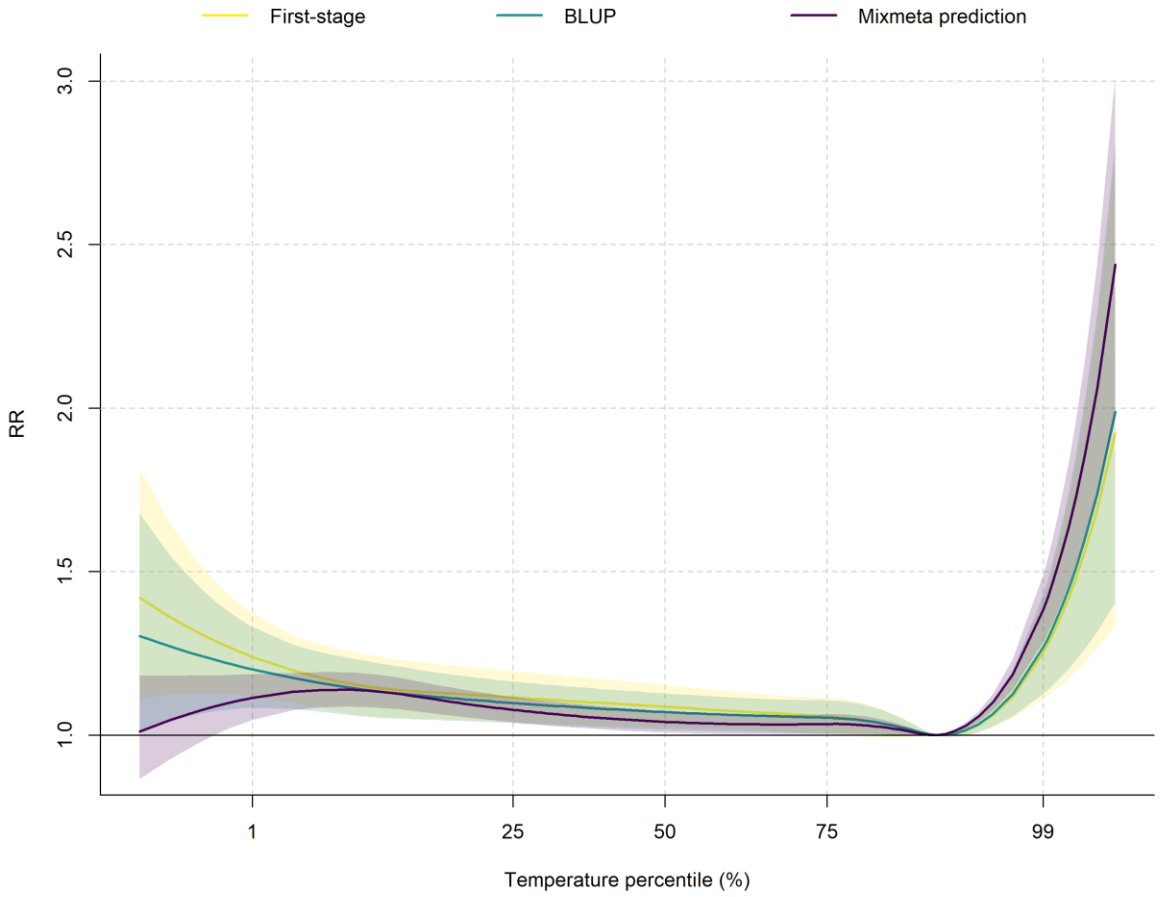


**Figure S13:** Example of the discrepancy between BLUP estimate and prediction by model (1), using as an example the city of Munich (Germany) for the age group 65 and older.

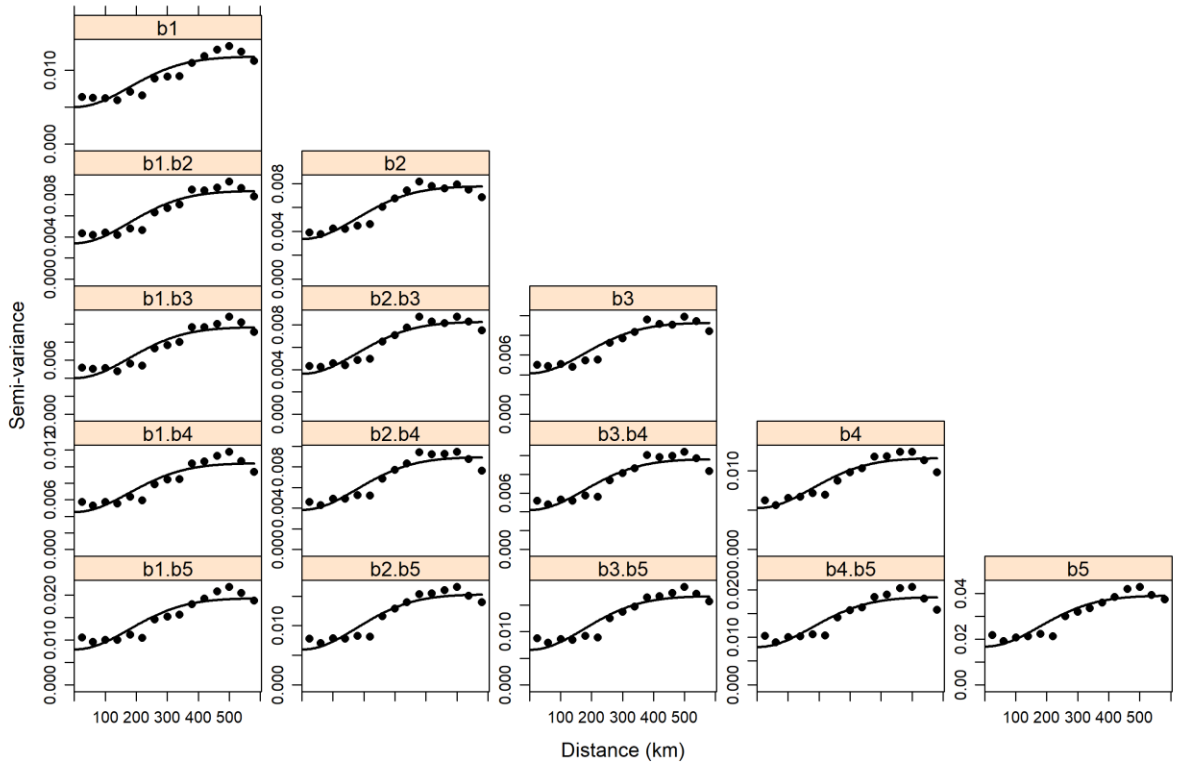


**Figure S14:** Empirical variograms showing the variance of the difference between values versus the distance, and fitted models.

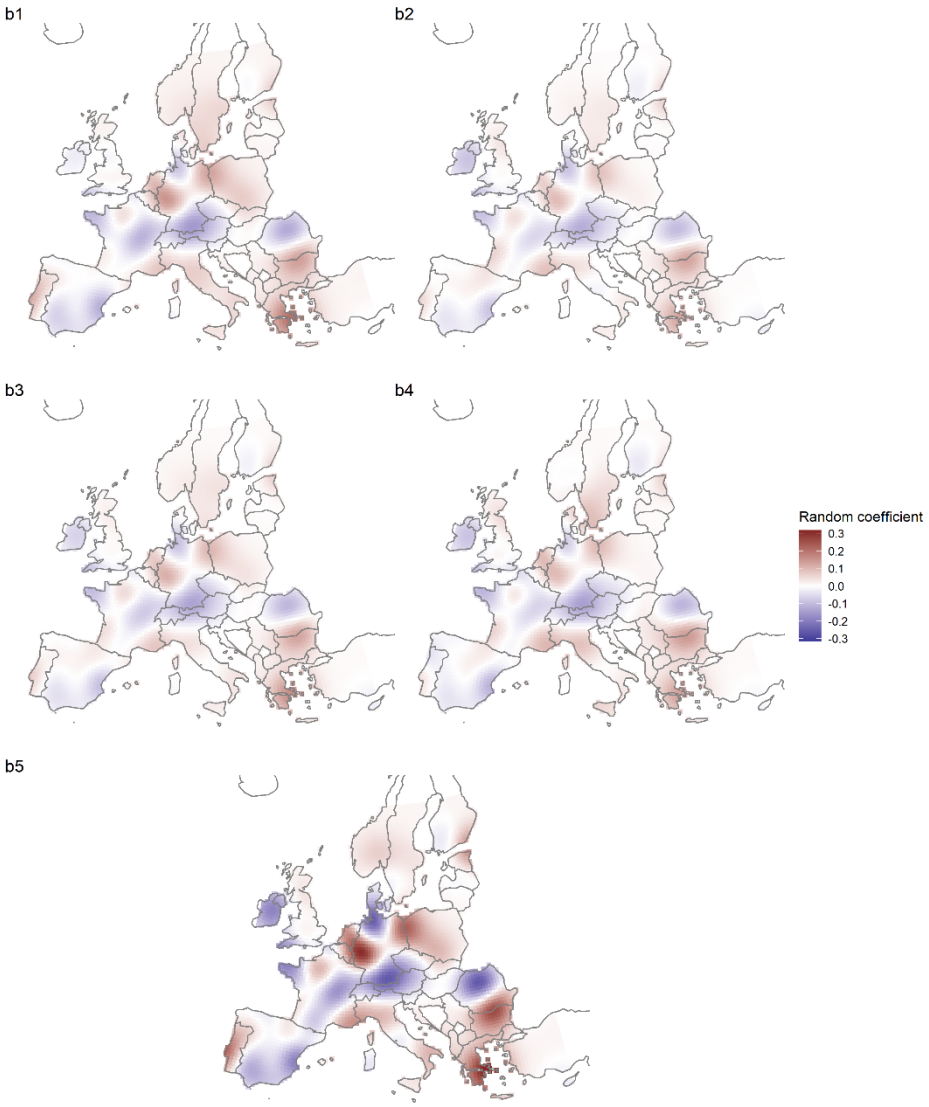


**Figure S15:** Interpolated surfaces for each BLUP residual.

**B.5. Computation of excess mortality and uncertainty**

The final model built in the second stage was used to predict ERFs for each of the 801 cities and five groups: 20-44, 45-64, 65-74, 75-84, 85+. We excluded the population less than 20 years old as the low counts for the youngest age groups make the estimates less reliable. We then computed the excess number of deaths and mortality rates from these ERFs, using classical formulae for attributable risk.[16] Since mortality series were not available for all cities and age groups, the baseline mortality were computed from the crude annual death rates obtained from city and age-specific death rates (see Supplementary Table **S3**). Consistently with previous studies, heat and cold were defined relative to the minimum mortality temperature (MMT). These numbers were transformed into rates by dividing them by the population. In Figure 3 of the main manuscript, numbers are divided by the total population (all aged) to obtain a breakdown of ages at which temperature-related excess mortality occurs. When reported at the country or regional level, excess deaths are first summed across cities and then divided by the total country or regional-level population. For standardized excess mortality rates, age-group-specific numbers were first divided by the age -specific population, and these rates were then weighted-averaged using the Standard European Population for 2013. To obtain country level rates, the same methodology is applied, using country sums of excess mortality numbers and population.

Uncertainty on these numbers were obtained by 1000 Monte-Carlo simulations performed at the meta-regression level. More specifically, we sampled from a multivariate normal distribution centred on the estimated meta-coefficient ($\widehat{\boldsymbol{\beta}}$, $\widehat{\boldsymbol{\gamma}}_{r(i)}$, and internal coefficients related to the spline expansion $ns(a_{ij})$) and (co)variance matrix derived from the second-stage model. These 1000 sampled meta-regression coefficients were then used to predict 1000 sets of city and age-group-specific spline coefficients $\widehat{\boldsymbol{\theta}}_{ij}^{s}$, to which we added 1000 sampled kriged $\widehat{\boldsymbol{b}}_i^s$ also from a multivariate normal distribution ($s = 1, \ldots, 1000$). These 1000 sets of final spline coefficients $\widehat{\boldsymbol{\theta}}_{ij}^{s} + \widehat{\boldsymbol{b}}_i^s$ were then used to generate 1000 sets of exposure-response functions and standardized rates. Note that, as the random effect $\widehat{\boldsymbol{b}}_i$ is assumed independent from the other components of the model, the (co)variance matrices from the fixed effect and random effect simulations can simply be added. We then computed all health impact measures mentioned above for each simulated sample $s$ and obtained 95% empirical confidence intervals by extracting the 2.5$^{th}$ and 97.5$^{th}$ quantiles of the simulated values of excess numbers, rates, and standardized rates.

## C. Additional results

### C.1. Additional plots

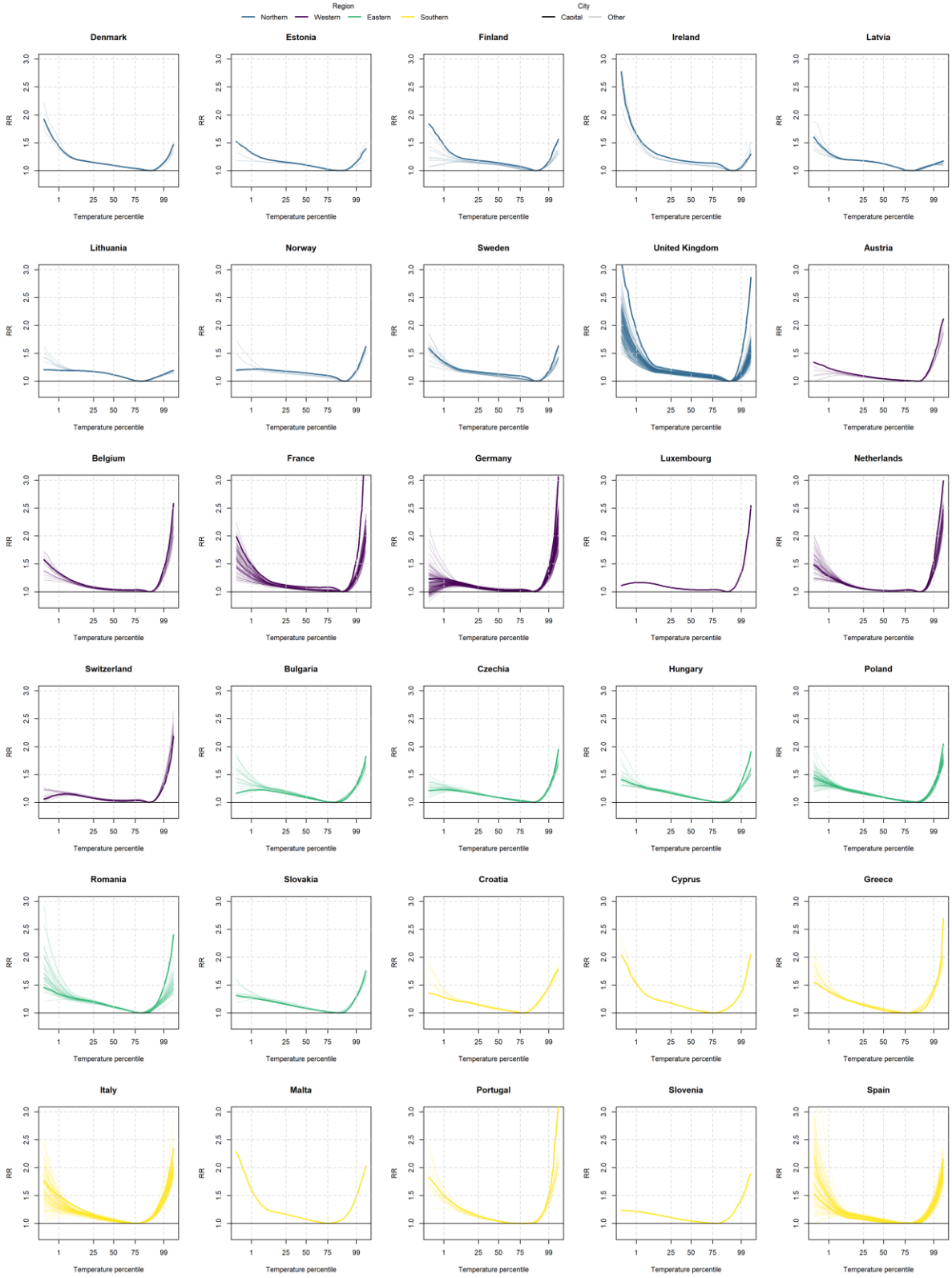


**Figure S16:** Overall cumulative exposure-response functions by country at 65 years old. The thicker line represents the captial city.

**C.2. Sensitivity to August 2003**

The month of August 2003 is excluded from the main analysis because of the high sensitivity of some results to this month, in particular Paris that saw unprecedented mortality during this period. We show here some results obtained with August 2003 remaining in the analysis. Figure S**17** shows the RR of capital cities (as in Figure 2 of the main manuscript) and shows that keeping August 2003 results in exceptionally high RR of heat for Paris, and less strinkingly for Lisbon. Figure S**18** shows excess deaths and presents key differences with the version in the main manuscript, including higher impacts for France, and mucher lower impacts of heat for Latvia and Lithuania. Finally, Table S**6** shows country level results, and particularly shows that including August 2003 results in almost 10% increase of cold-related annual death across Europe. Indeed, including this event shifts significantly the MMP in some places.

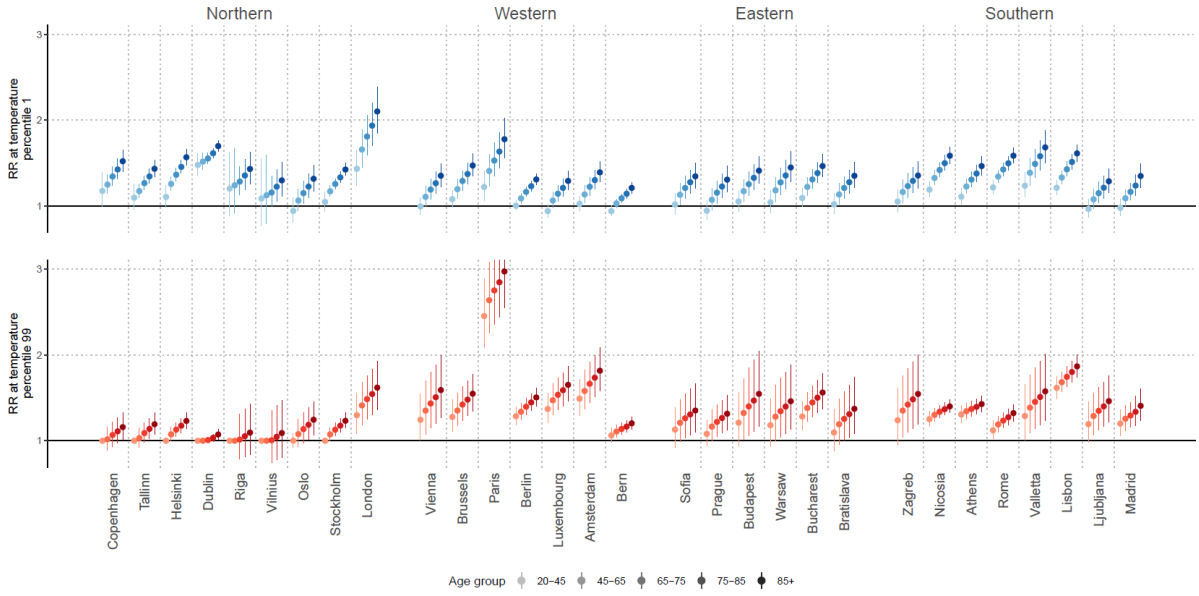


**Figure S17**: Cold (1st percentile of temperature, in blue) and heat (99th percentile of temperature, in red) relative risks (RRs) in capital cities for five age groups with August 2003 kept in the analysis.

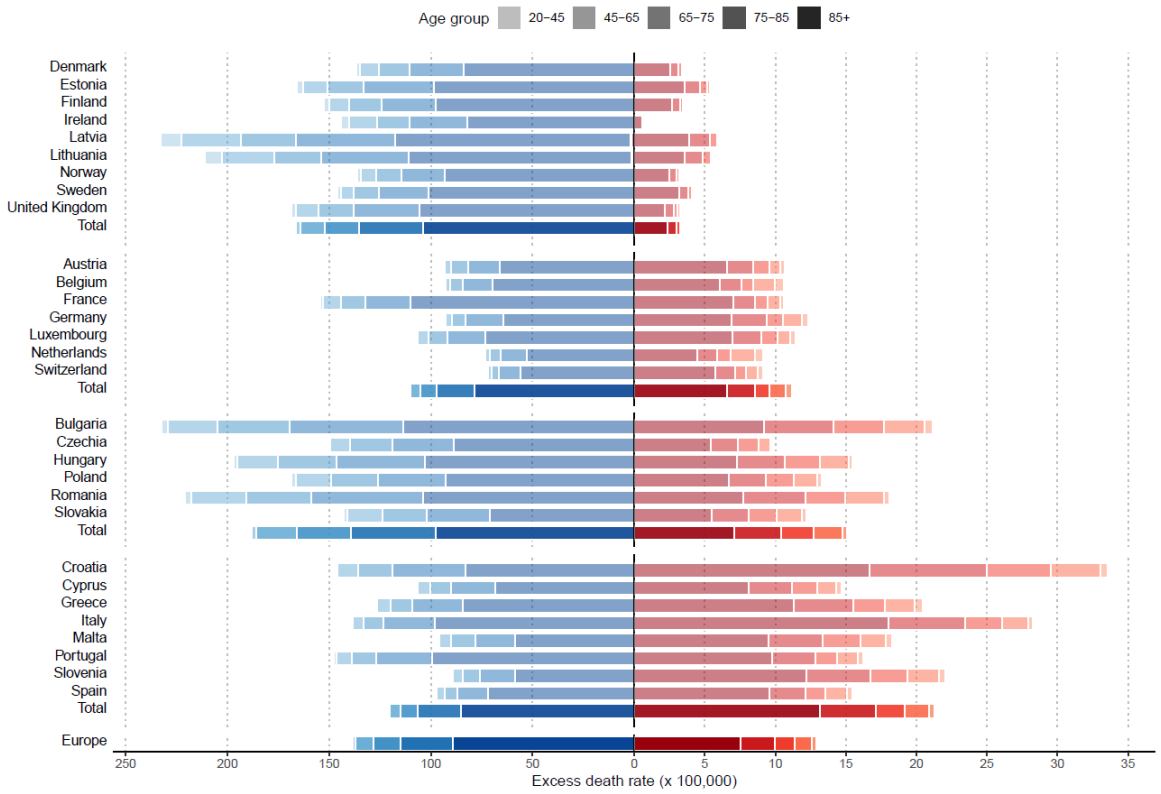


**Figure S18**: Country-level cold (in blue) and heat (in red) annual raw death rates broken down by age group, with August 2003 kept in the analysis.

**Table S6:** Country-level annual excess number of deaths, attributable fractions, raw and age-standardised rates for cold and heat for the population aged 20 and above with August 2003 in the analysis. Brackets indicate 95% empirical confidence intervals.

| Region | Country | Excess deaths | | Attributable fraction (%) | | Excess death rates (x 100,000) | | Standardized excess death rates (x 100,000) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Cold | Heat | Cold | Heat | Cold | Heat | Cold | Heat |
| Northern | Denmark | 1,256 (995 - 1,491) | 31 (4 - 52) | 8.23 (6.52 - 9.77) | 0.20 (0.03 - 0.34) | 137 (108 - 163) | 3 (0 - 6) | 157 (125 - 188) | 4 (1 - 7) |
| | Estonia | 730 (529 - 933) | 23 (7 - 38) | 8.52 (6.18 - 10.90) | 0.27 (0.09 - 0.45) | 166 (120 - 212) | 5 (2 - 9) | 177 (128 - 227) | 6 (2 - 9) |
| | Finland | 2,227 (1,575 - 2,840) | 51 (1 - 89) | 9.27 (6.55 - 11.81) | 0.21 (0.00 - 0.37) | 153 (108 - 194) | 4 (0 - 6) | 165 (118 - 211) | 4 (0 - 7) |
| | Ireland | 1,571 (1,367 - 1,772) | 7 (-1 - 16) | 12.00 (10.44 - 13.52) | 0.06 (-0.01 - 0.12) | 144 (125 - 162) | 1 (-0 - 1) | 213 (185 - 241) | 1 (-0 - 2) |
| | Latvia | 2,035 (1,128 - 2,796) | 49 (-55 - 129) | 9.44 (5.23 - 12.98) | 0.23 (-0.26 - 0.60) | 233 (129 - 320) | 6 (-6 - 15) | 236 (132 - 324) | 6 (-6 - 15) |
| | Lithuania | 2,181 (1,225 - 3,077) | 54 (-56 - 145) | 8.89 (5.00 - 12.55) | 0.22 (-0.23 - 0.59) | 211 (118 - 298) | 5 (-5 - 14) | 221 (125 - 310) | 6 (-6 - 15) |
| | Norway | 1,128 (819 - 1,432) | 27 (6 - 46) | 10.28 (7.46 - 13.05) | 0.25 (0.06 - 0.41) | 136 (99 - 173) | 3 (1 - 5) | 174 (126 - 220) | 4 (1 - 7) |
| | Sweden | 4,098 (3,318 - 4,887) | 117 (59 - 167) | 8.77 (7.10 - 10.46) | 0.25 (0.13 - 0.36) | 146 (118 - 174) | 4 (2 - 6) | 151 (122 - 180) | 4 (2 - 6) |
| | United Kingdom | 44,685 (39,180 - 50,009) | 864 (614 - 1,095) | 10.28 (9.02 - 11.51) | 0.20 (0.14 - 0.25) | 168 (148 - 188) | 3 (2 - 4) | 190 (167 - 213) | 4 (3 - 5) |
| | **Total** | **59,911 (52,793 - 66,932)** | **1,224 (774 - 1,582)** | **10.00 (8.81 - 11.17)** | **0.20 (0.13 - 0.26)** | **166 (147 - 186)** | **3 (2 - 4)** | **187 (165 - 209)** | **4 (2 - 5)** |
| Western | Austria | 1,823 (1,063 - 2,570) | 208 (145 - 266) | 5.71 (3.33 - 8.05) | 0.65 (0.45 - 0.83) | 94 (55 - 132) | 11 (7 - 14) | 104 (61 - 146) | 12 (8 - 15) |
| | Belgium | 2,289 (1,525 - 3,023) | 261 (208 - 310) | 4.90 (3.27 - 6.47) | 0.56 (0.45 - 0.66) | 93 (62 - 123) | 11 (8 - 13) | 90 (59 - 119) | 10 (8 - 12) |
| | France | 27,708 (22,393 - 32,502) | 1,911 (1,679 - 2,135) | 9.31 (7.52 - 10.92) | 0.64 (0.56 - 0.72) | 154 (125 - 181) | 11 (9 - 12) | 140 (113 - 164) | 10 (9 - 11) |
| | Germany | 21,744 (15,294 - 27,717) | 2,880 (2,385 - 3,345) | 4.82 (3.39 - 6.14) | 0.64 (0.53 - 0.74) | 93 (65 - 118) | 12 (10 - 14) | 90 (63 - 114) | 12 (10 - 14) |
| | Luxembourg | 75 (55 - 95) | 8 (7 - 9) | 7.73 (5.74 - 9.87) | 0.83 (0.69 - 0.96) | 107 (79 - 136) | 11 (10 - 13) | 136 (101 - 174) | 14 (12 - 17) |
| | Netherlands | 4,141 (2,084 - 6,185) | 514 (389 - 639) | 4.46 (2.25 - 6.67) | 0.55 (0.42 - 0.69) | 73 (37 - 110) | 9 (7 - 11) | 85 (43 - 127) | 10 (8 - 12) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Switzerland | 1,427 (972 - 1,853) | 180 (134 - 223) | 4.56 (3.10 - 5.92) | 0.58 (0.43 - 0.71) | 72 (49 - 94) | 9 (7 - 11) | 72 (49 - 93) | 9 (7 - 11) |
| | **Total** | **59,206 (46,788 - 70,794)** | **5,962 (5,113 - 6,741)** | **6.21 (4.91 - 7.43)** | **0.63 (0.54 - 0.71)** | **111 (88 - 132)** | **11 (10 - 13)** | **107 (84 - 128)** | **11 (9 - 12)** |
| Eastern | Bulgaria | 6,038 (4,749 - 7,271) | 549 (302 - 799) | 9.72 (7.64 - 11.70) | 0.88 (0.49 - 1.29) | 233 (183 - 280) | 21 (12 - 31) | 279 (220 - 335) | 25 (14 - 36) |
| | Czechia | 3,797 (2,531 - 5,104) | 247 (126 - 363) | 8.17 (5.45 - 10.99) | 0.53 (0.27 - 0.78) | 150 (100 - 202) | 10 (5 - 14) | 184 (123 - 246) | 12 (6 - 17) |
| | Hungary | 5,525 (4,181 - 6,815) | 434 (274 - 594) | 9.16 (6.93 - 11.30) | 0.72 (0.46 - 0.98) | 197 (149 - 243) | 15 (10 - 21) | 221 (167 - 272) | 17 (11 - 24) |
| | Poland | 17,535 (13,312 - 21,622) | 1,384 (891 - 1,878) | 8.96 (6.80 - 11.05) | 0.71 (0.46 - 0.96) | 169 (128 - 208) | 13 (9 - 18) | 193 (147 - 238) | 15 (10 - 20) |
| | Romania | 13,270 (10,559 - 15,970) | 1,086 (698 - 1,491) | 10.27 (8.17 - 12.36) | 0.84 (0.54 - 1.15) | 221 (176 - 266) | 18 (12 - 25) | 269 (214 - 323) | 22 (14 - 30) |
| | Slovakia | 1,254 (952 - 1,542) | 107 (62 - 149) | 8.88 (6.74 - 10.91) | 0.76 (0.44 - 1.05) | 143 (109 - 176) | 12 (7 - 17) | 210 (160 - 257) | 18 (10 - 24) |
| | **Total** | **47,420 (36,720 - 57,337)** | **3,806 (2,453 - 5,138)** | **9.34 (7.23 - 11.29)** | **0.75 (0.48 - 1.01)** | **188 (146 - 227)** | **15 (10 - 20)** | **222 (173 - 268)** | **18 (12 - 24)** |
| Southern | Croatia | 1,620 (1,066 - 2,165) | 370 (251 - 479) | 7.11 (4.68 - 9.50) | 1.62 (1.10 - 2.10) | 147 (97 - 197) | 34 (23 - 43) | 168 (111 - 225) | 38 (26 - 49) |
| | Cyprus | 492 (384 - 600) | 67 (47 - 86) | 8.04 (6.28 - 9.82) | 1.10 (0.77 - 1.41) | 108 (84 - 131) | 15 (10 - 19) | 161 (126 - 196) | 21 (15 - 28) |
| | Greece | 4,335 (3,080 - 5,536) | 695 (445 - 960) | 6.45 (4.58 - 8.24) | 1.03 (0.66 - 1.43) | 127 (91 - 163) | 20 (13 - 28) | 118 (84 - 150) | 19 (12 - 26) |
| | Italy | 23,917 (18,259 - 29,558) | 4,858 (3,941 - 5,784) | 6.94 (5.30 - 8.57) | 1.41 (1.14 - 1.68) | 139 (106 - 172) | 28 (23 - 34) | 115 (88 - 142) | 24 (19 - 28) |
| | Malta | 165 (112 - 212) | 31 (19 - 42) | 7.15 (4.88 - 9.22) | 1.35 (0.81 - 1.83) | 97 (66 - 125) | 18 (11 - 25) | 128 (87 - 165) | 24 (14 - 32) |
| | Portugal | 5,101 (3,996 - 6,296) | 562 (484 - 642) | 7.76 (6.08 - 9.58) | 0.86 (0.74 - 0.98) | 148 (116 - 182) | 16 (14 - 19) | 144 (113 - 178) | 16 (14 - 18) |
| | Slovenia | 275 (159 - 395) | 68 (44 - 91) | 5.34 (3.10 - 7.69) | 1.32 (0.85 - 1.76) | 89 (52 - 129) | 22 (14 - 30) | 97 (56 - 139) | 24 (15 - 32) |
| | Spain | 18,266 (12,495 - 23,803) | 2,894 (2,178 - 3,619) | 5.53 (3.78 - 7.20) | 0.88 (0.66 - 1.09) | 97 (67 - 127) | 15 (12 - 19) | 87 (59 - 114) | 14 (11 - 18) |
| | **Total** | **54,170 (40,077 - 67,785)** | **9,545 (7,591 - 11,520)** | **6.41 (4.75 - 8.03)** | **1.13 (0.90 - 1.36)** | **121 (89 - 151)** | **21 (17 - 26)** | **107 (79 - 134)** | **19 (15 - 23)** |
| **Total** | | **220,706 (197,262 - 243,103)** | **20,538 (17,628 - 23,085)** | **7.60 (6.79 - 8.37)** | **0.71 (0.61 - 0.79)** | **138 (124 - 152)** | **13 (11 - 14)** | **139 (124 - 153)** | **13 (11 - 15)** |

## D. References


1. Eurostat. Cities (Urban Audit). https://ec.europa.eu/eurostat/web/cities/background.
2. Gasparrini, A., Armstrong, B. & Kenward, M. G. Distributed lag non-linear models. *Statistics in Medicine* **29**, 2224–2234 (2010).
3. Gasparrini, A. *et al.* Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *The Lancet* **386**, 369–375 (2015).
4. Gasparrini, A. & Armstrong, B. Reducing and meta-analysing estimates from distributed lag non-linear models. *BMC Medical Research Methodology* **13**, 1 (2013).
5. Sera, F., Armstrong, B., Blangiardo, M. & Gasparrini, A. An extended mixed-effects framework for meta-analysis. *Statistics in Medicine* **38**, 5429–5444 (2019).
6. United Nations Statistical Commission. Standard country or area codes for statistical use (M49). https://unstats.un.org/unsd/methodology/m49/#fn1.
7. Chambers, J. M. & Hastie, T. J. *Statistical Models in S*. (Chapman and Hall/CRC, 1991).
8. Armstrong, B. G., Gasparrini, A., Tobias, A. & Sera, F. Sample size issues in time series regressions of counts on environmental exposures. *BMC Medical Research Methodology* **20**, 15 (2020).
9. Wold, S., Sjöström, M. & Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* **58**, 109–130 (2001).
10. Frank, I. E. & Friedman, J. H. A Statistical View of Some Chemometrics Regression Tools. *Technometrics* **35**, 109–135 (1993).
11. de Jong, S. SIMPLS: An alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems* **18**, 251–263 (1993).
12. Matheron, G. Principles of geostatistics. *Economic Geology* **58**, 1246–1266 (1963).
13. Pebesma, E. J. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* **30**, 683–691 (2004).
14. Bachmaier, M. & Backes, M. Variogram or semivariogram? Understanding the variances in a variogram. *Precision Agric* **9**, 173–175 (2008).
15. Cressie, N. A. C. Spatial Prediction and Kriging. in *Statistics for Spatial Data* 105–209 (John Wiley & Sons, Ltd, 1993). doi:10.1002/9781119115151.ch3.
16. Gasparrini, A. & Leone, M. Attributable risk from distributed lag models. *BMC Medical Research Methodology* **14**, 55 (2014).