Assessing Mortality Risk of Climate Change

Han Li[1] and Anastasios Panagiotelis[2]

[1] Centre for Actuarial Studies, Department of Economics
University of Melbourne

[2] Discipline of Business Analytics, Business School
University of Sydney

21 March 2024

Acknowledgment

This research is sponsored by the **SCOR Foundation for Science** through the project titled "*Who, When, and Where?Assessing Mortality Risk of Climate Change*".

Outline of the presentation

Outline

The link between climate change and human mortality

According to the WHO:

Between **2030** and **2050**, climate change is expected to cause approximately **250,000** additional deaths **per year**.

Weather-related catastrophes

Extreme temperatures

Climate-sensitive infectious diseases

How extreme temperatures affect mortality

- Globally, approximately **5 million** deaths were associated with non-optimal temperatures per year, that is **1** out of every **10** deaths in the world during 2000–2019.
- Extreme **heat** associated excess mortality is more **immediate**, while the increase in mortality following extreme **cold** is long **lasting**.
- The **elderly** are more **fragile** to extreme temperatures, and different **regions** react to extreme temperatures **differently**.

SCOR The Art & Science of Risk | Foundation for science | THE UNIVERSITY OF MELBOURNE | THE UNIVERSITY OF SYDNEY

The research questions

Our project is expected to provide insights into the following questions:

1. Who are the excess deaths? – *Find the age groups particularly sensitive to climate change.*
2. When do excess deaths occur? – *Determine if more excess deaths occur in winter or summer.*
3. Where are the excess deaths? – *Identify regions that are most vulnerable to climate change.*

General trends and patterns in mortality rates

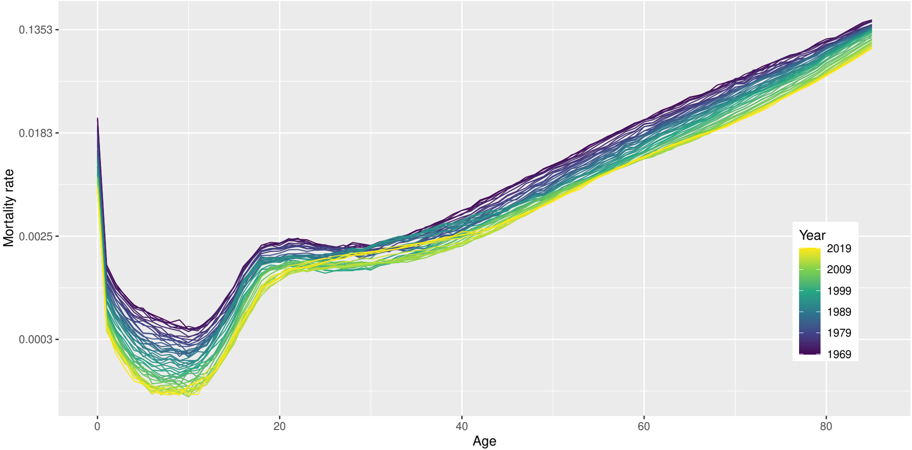
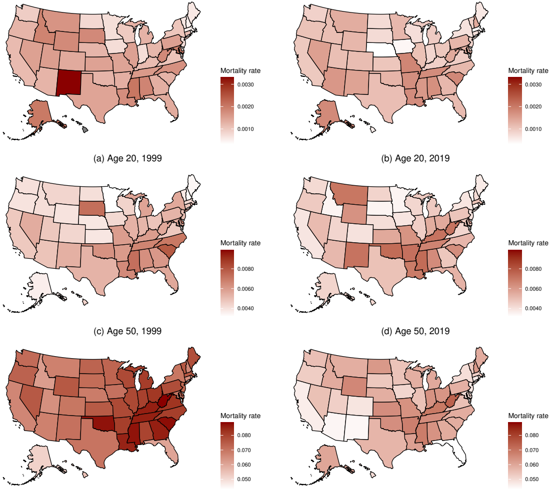


Figure 1: National-level male mortality for US: 1969–2019.

Regional differences in mortality experience



(a) Age 20, 1999

(b) Age 20, 2019

(c) Age 50, 1999

(d) Age 50, 2019

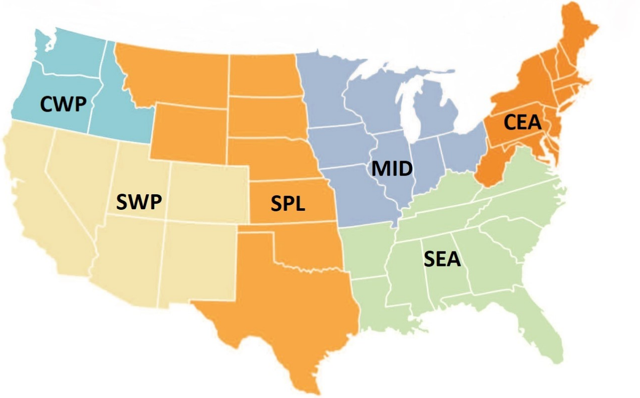
(e) Age 80, 1999

(f) Age 80, 2019

Figure 2: Male mortality rates across the US in 1999 (left panel) and 2019 (right panel).

Geographical resolutions: continental level

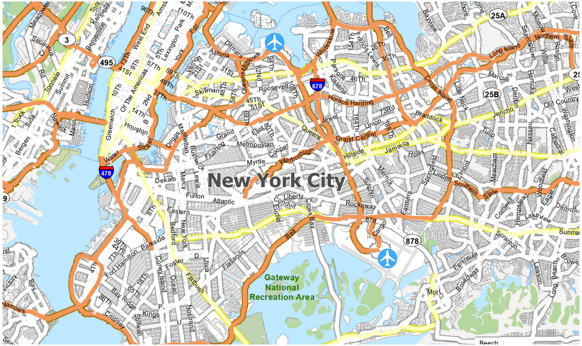
Li, H., Tang, Q., 2022. Joint extremes in temperature and mortality: A bivariate POT approach. *North American Actuarial Journal*, 26(1), 43–63.



Source: Actuaries Climate Index Executive Summary (2018), Page 4, Figure 2

Geographical resolutions: city level

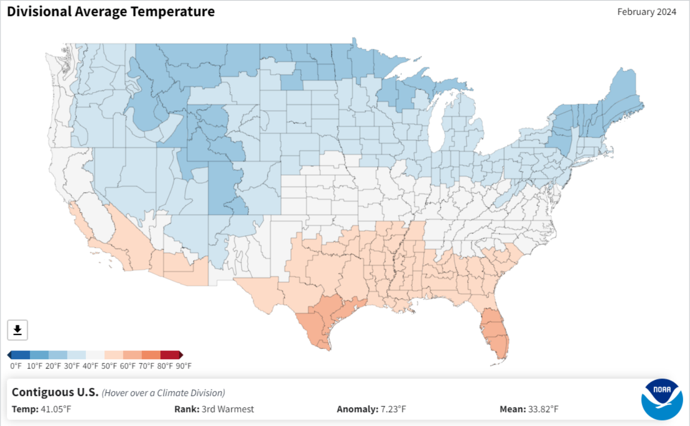
Gasparrini, A., Armstrong, B., & Kenward, M. G., 2010. Distributed lag non-linear models. *Statistics in Medicine*, 29(21), 2224–2234.



Source: GISGeography, `https://gisgeography.com/new-york-city-map`

Geographical resolutions: climate division level

Vose, Russell S., et al., 2014. Improved historical temperature and precipitation time series for US climate divisions. *Journal of Applied Meteorology and Climatology*, 53(5), 1232–1251.



Source: NOAA, `https://www.ncei.noaa.gov`

Outline

Death data

The US county-level monthly death data for the time period 1968–2018 are obtained from two main sources listed as follows.

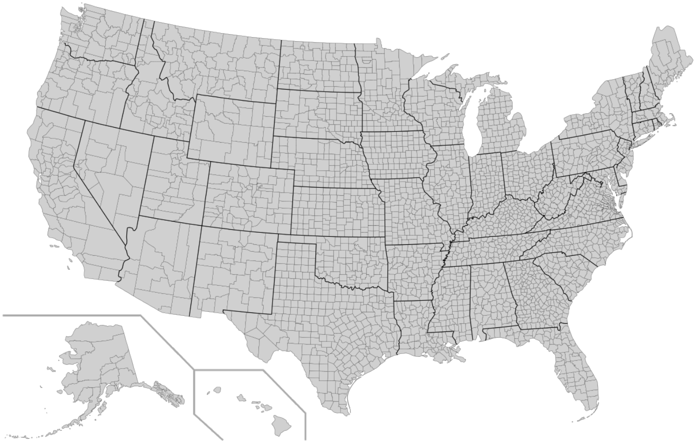
- *The National Center for Health Statistics (NCHS)*
- *Centers for Disease Control and Prevention (CDC) WONDER*

Note that:

1) As there is no publicly available data on monthly age-specific population exposure, particularly at the county level, we assume constant exposure throughout the year.

2) The CDC WONDER online database censors death counts between 0 and 9 due to privacy constraints for years 1989 onward.

A map of US counties



Source: Wikipedia, `https://en.wikipedia.org/wiki/County_(United_States)`

Temperature data

We collect the corresponding monthly temperature data at the climate division level, from the National Oceanic and Atmospheric Administration (NOAA) for the contiguous United States.

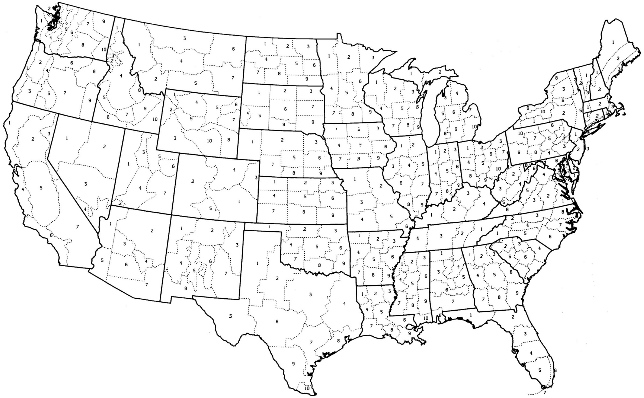
- *NOAA Monthly U.S. Climate Divisional Database*

Note that:

1) There are 344 climate divisions, and they are larger than counties but smaller than states.

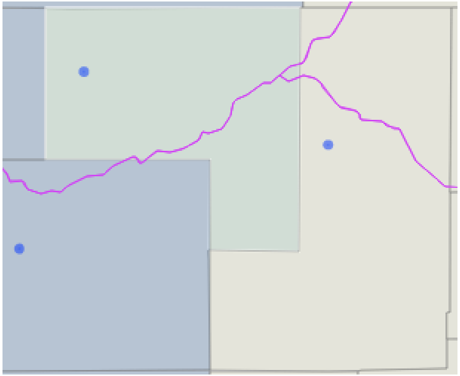
2) The borders of climate divisions overlap those of counties, some care will need to be taken to combine the deaths database with the weather database.

A map of US climate divisions



Source: NOAA, `https://www.ncei.noaa.gov/access/monitoring/dyk/us-climate-divisions`

Matching counties into climate divisions



Pink line: Climate division border

Grey line: County border

Blue dot: County population center

A quick glance at the data

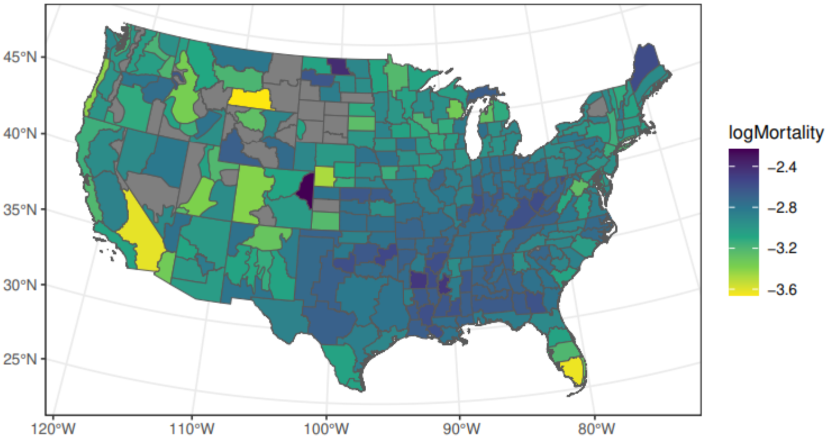


Figure plots log mortality, ages 75–84, January 2018.

Outline

Challenges

Quantifying mortality risk due to temperatures is challenging:

1. The relationship is non-linear;
2. This non-linear relationship may also vary with geography;
3. There is a need to control for time trend and age patterns that can confound the relationship.

The model

We propose the following model:

$$y_{i,x,t} = \mu_{i,x,t} + \varepsilon_{i,x,t} \quad \varepsilon \overset{i.i.d}{\sim} N(0,\sigma^2), \tag{1}$$

$$\mu_{i,x,t} = \alpha + \gamma_i^{(cd)} + \gamma_{\lfloor t/12 \rfloor}^{(yr)} + \gamma_{t \bmod 12}^{(mth)} + \gamma_x^{(age)} + \gamma_{\lfloor t/12 \rfloor, x}^{(yr-age)} + f_i(\text{Temp}_{i,t}), \tag{2}$$

where

- $y_{i,x,t}$ is (annualized) log mortality in climate division i at time t for age group x.
- $\gamma^{(cd)}$, γ^{yr}, γ^{mth}, γ^{age} are random effects for climate division, year, month, and age respectively.
- γ^{yr-age} is a two way random effect for year and age.
- $f_x(\text{Temp}_{i,t})$ is a climate division specific regression function measuring the effect of temperature on mortality

Data suppression

Log mortality is observed monthly but annualized

$$y_{i,x,t} = \log\left(\frac{\text{Deaths}_{i,x,t}}{\text{Population}_{i,x,t}} \times \frac{\text{Days in Year}_t}{\text{Days in Month}_t}\right)$$

When there are fewer than 10 deaths, we know that $y_{i,x,t}$ is upper bounded by

$$u_{i,x,t} = \log\left(\frac{10}{\text{Population}_{i,x,t}} \times \frac{\text{Days in Year}_t}{\text{Days in Month}_t}\right)$$

Data augmentation

In the Bayesian framework, the model becomes

$$y^*_{i,x,t} = \mu_{i,x,t} + \varepsilon_{i,x,t},$$

where the latent variable y^* is defined as

$$y^*_{i,x,t} \begin{cases} = y_{i,x,t} & \text{if not suppressed} \\ \sim N(\mu_{i,x,t}, \sigma^2) I(y^*_{i,x,t} < u_{i,x,t}) & \text{if suppressed} \end{cases}$$

The latent values corresponding to suppressed observation are treated like any other parameters that can be estimated.

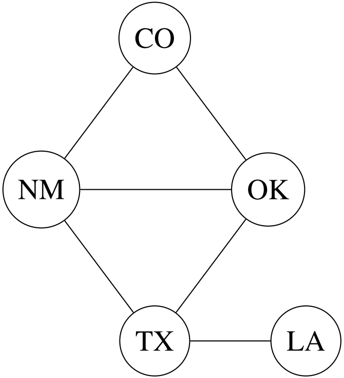
Priors

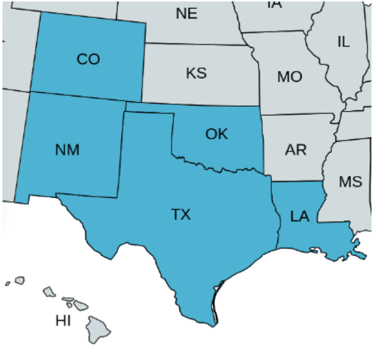
- The idea is to shrink "neighboring" random effects together
  - Neighboring regions;
  - Neighboring years (*e.g.* 2005 and 2006);
  - Neighboring months (March and April);
  - Neighboring age groups (*e.g.* 55–64 and 64–75).
- This is done with an intrinsic conditional autoregressive (ICAR) prior (Besag and Kooperberg, 1995).
- Best understood with graph theory.

ICAR prior and graphs

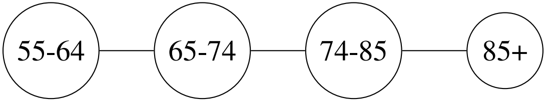
- A graph has nodes and edges.
  - In our case, nodes will correspond to random effects (denoted γ).
- Edges connect two nodes.
  - Edges will correspond to differences between random effects (denoted δ).
- ICAR prior assumes $\delta \sim N(0, \tau^2)$.
- Differences between neighbors shrunk to zero.
- Equivalent to assuming $\gamma \sim N(\bar{\gamma}, \tau^2)$ where $\bar{\gamma}$ is the neighborhood average.

Climate division graph



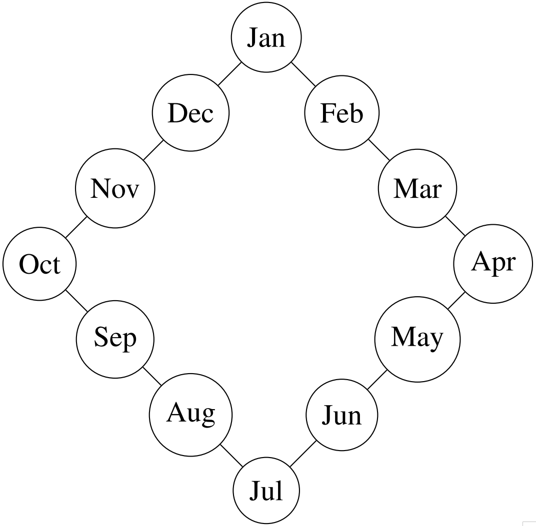


Age graph



Graph for year effects is similar

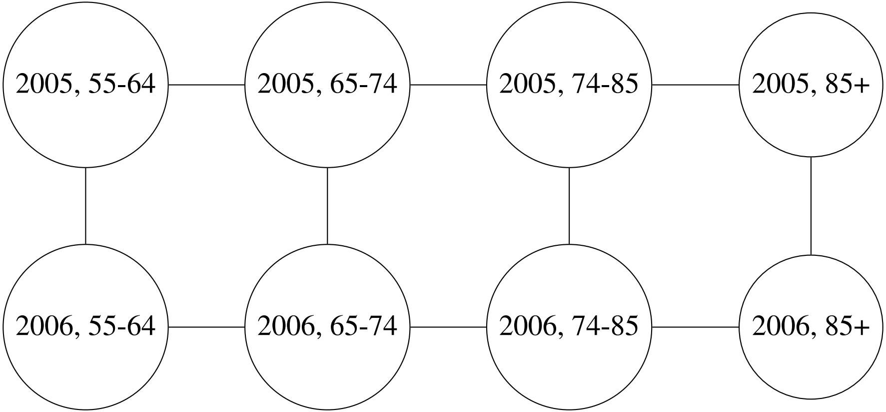
Month graph



Two-way effects

- For two way effects construct the Cartesian product of two graphs,
- Consider age year effect,
- Nodes become Age year pairs, *e.g.* (2005, 55–64),
- Edge between nodes if
  - Year is the same and age are neighbours, *e.g.*, (2005,55–64) and (2005, 65–74).
  - Year are neighbours and age is the same, *e.g.*, (2005, 55–64) and (2006, 55–64).

Two-way effects



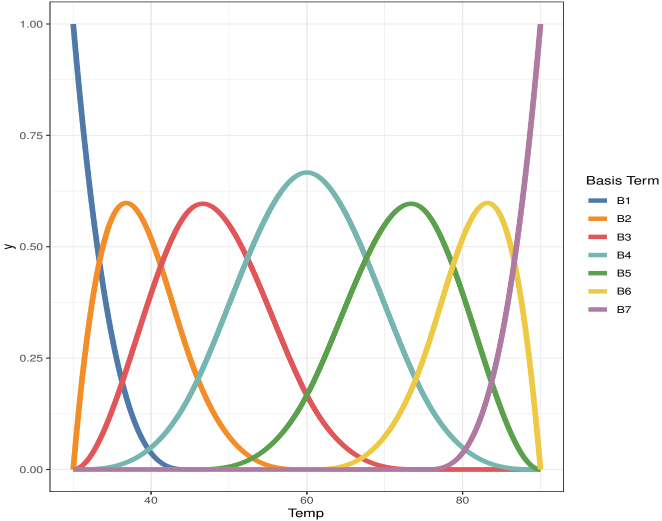
Regression effects

Regression effects for temperature modelled non-parametrically using B-Splines

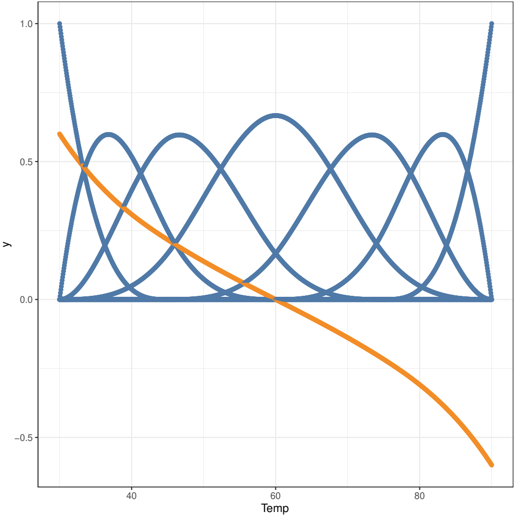
$$f(\text{Temp}) = \sum_{j=1}^{p} B_j(\text{Temp})\beta_j.$$

The functions $B_j(.)$ are known as a B-Spline basis, and a known pre-computed functions, while β_j are coefficients. This makes a non-linear regression into a linear regression.

B-Spline basis

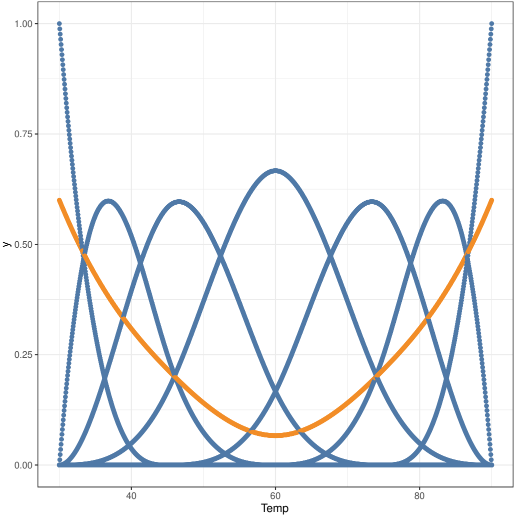


B-Spline basis



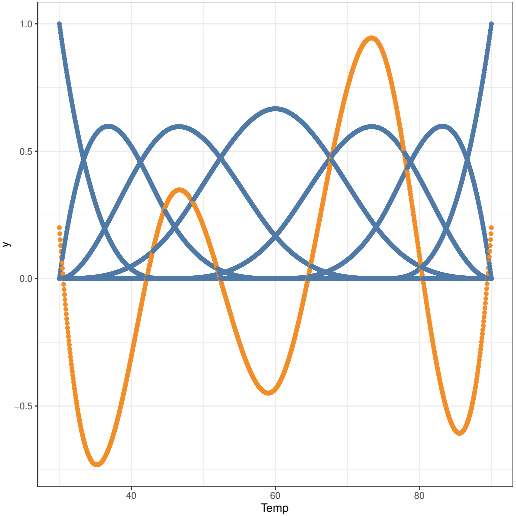
$$\boldsymbol{\beta} = \begin{pmatrix} 0.6 \\ 0.4 \\ 0.2 \\ 0.0 \\ -0.2 \\ -0.4 \\ -0.6 \end{pmatrix}$$

B-Spline basis



$$\boldsymbol{\beta} = \begin{pmatrix} 0.6 \\ 0.4 \\ 0.2 \\ 0.0 \\ 0.2 \\ 0.4 \\ 0.6 \end{pmatrix}$$

B-Spline basis



$$\boldsymbol{\beta} = \begin{pmatrix} 0.2 \\ -1.8 \\ 1.8 \\ -1.8 \\ 2.8 \\ -1.8 \\ 0.2 \end{pmatrix}$$

Prior on B-Spline coefficients

- Use a very large number of basis functions (*e.g.* 20 functions).
- Penalise differences between coefficients that correspond to adjacent splines.
- Approximates a penalty on integrated squared gradient of the function.
- Used in our framework for main regression effect.
- Bayesian version of penalized splines (Eilers and Marx, 1996).

Novelty

- We also have climate division specific temperature effects.
- The basis coefficients corresponding to the same B-spline across neighboring regions are also shrunk together (their difference is shrunk to zero).
- To the best of our knowledge this is a **novel innovation** in our work.
- Employ the same prior as for two way effects.

Recap-Priors

$$\delta_{(dom)} \sim N(0, \tau^2_{(dom)})$$
$$\tau_{(dom)} \sim \text{Cauchy}(0, 1)$$
$$p(\alpha, \sigma) \propto 1/\sigma^2$$

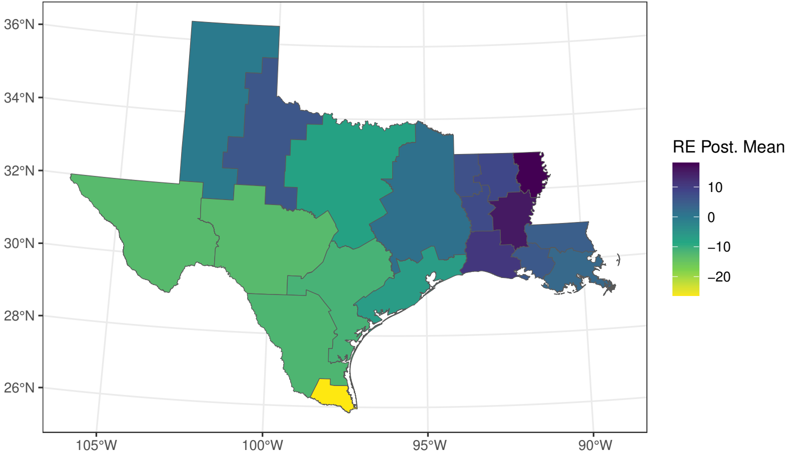
Here (dom) refers to either yearly, age, monthly, climate division, year-age effects, or B-spline basis coefficients (main and climate division specific). All γ and β are can be found deterministically from δ.

Inference

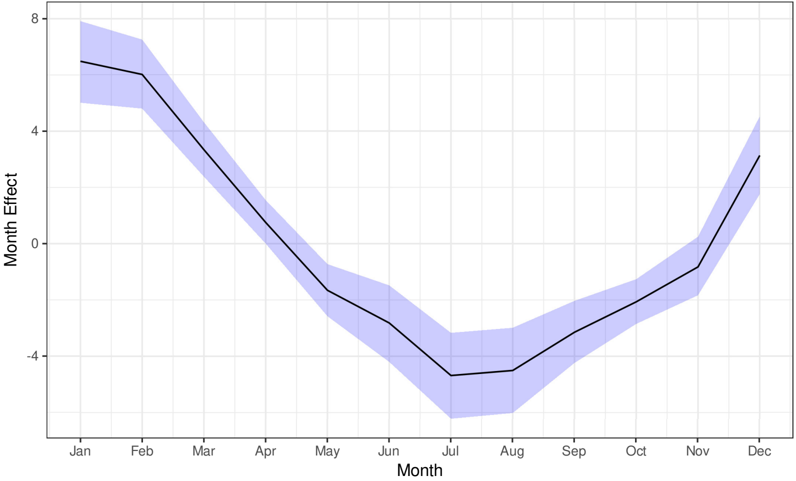
- Inference carried out using the No U Turn Sampler (NUTS).
- The implementation is in STAN (Carpenter et al., 2017).
- Samples from posterior by proposing new values of the parameter vector.
- Idea is to treat posterior as a measure of energy and evolve a proposal for the parameter vector using Hamiltonian dynamics.

Outline

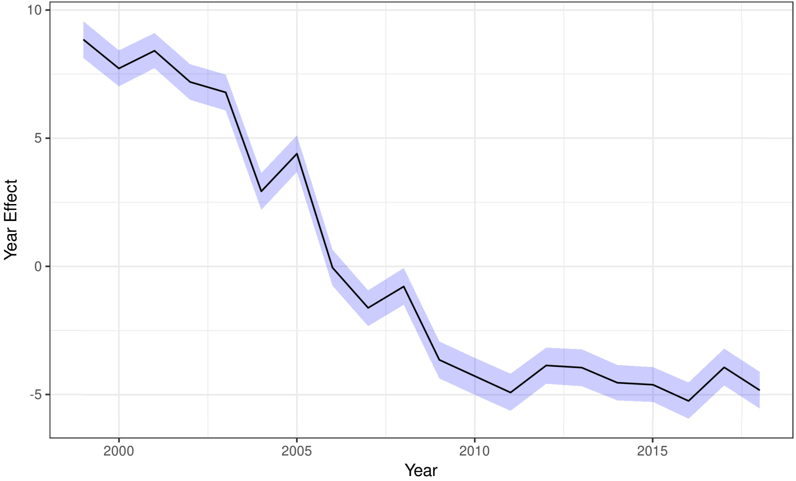
Results - Spatial effect



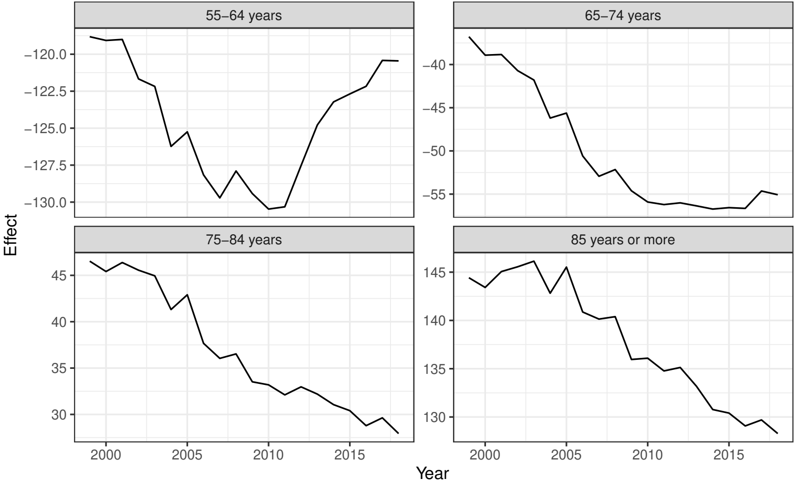
Results - Monthly effect



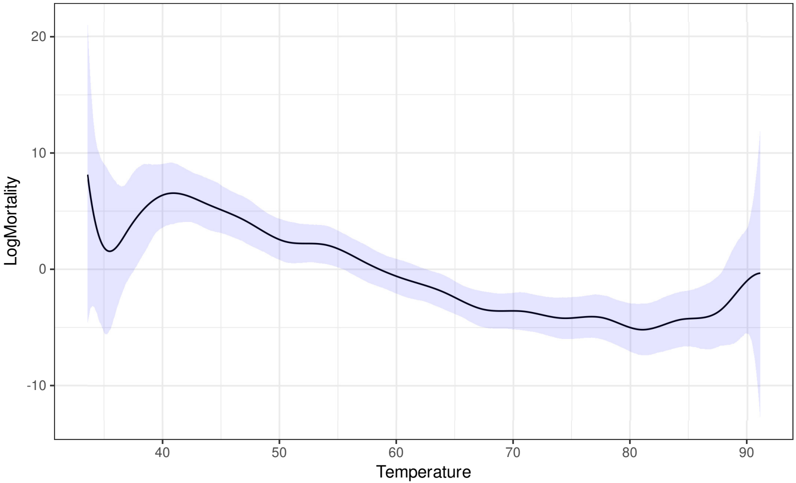
Results - Yearly effect



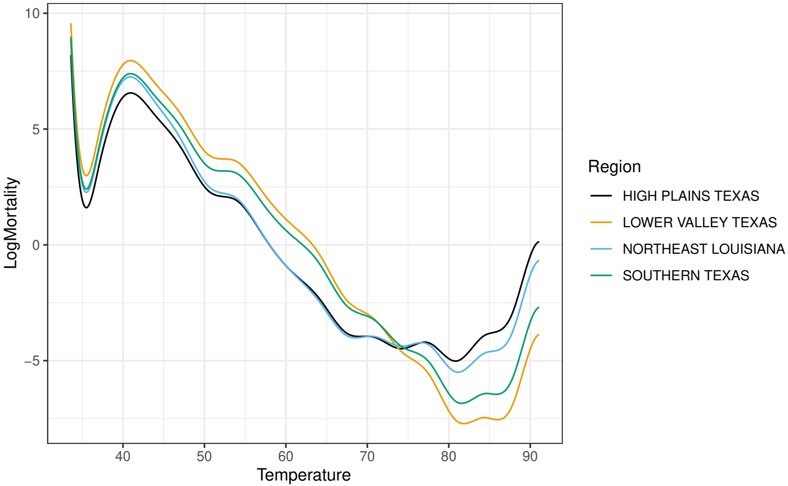
Results - Age-year effect



Results - Regression



Results - Regression



Conclusions

Our insights into the following questions:

1. Who are the excess deaths? – *Elderly people particularly the "oldest old"*.
2. When do excess deaths occur? – *Extreme cold temperatures in winter months and extreme heat during summer months*.
3. Where are the excess deaths? – *Geographic & Socioeconomic*.

References

Besag, J., & Kooperberg, C. (1995). On conditional and intrinsic autoregressions. *Biometrika*, 82(4), 733–746.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76.

Eilers, P. H., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, 11(2), 89–121.

Gasparrini, A., Armstrong, B., & Kenward, M. G., 2010. Distributed lag non-linear models. *Statistics in Medicine*, 29(21), 2224–2234.

Li, H., Tang, Q., 2022. Joint extremes in temperature and mortality: A bivariate POT approach. *North American Actuarial Journal*, 26(1), 43–63.

Vose, Russell S., et al., 2014. Improved historical temperature and precipitation time series for US climate divisions. *Journal of Applied Meteorology and Climatology*, 53(5), 1232–1251.

Questions and discussions

Contact emails: han.li@unimelb.edu.au
anastasios.panagiotelis@sydney.edu.au