Granular mortality modeling with temperature and epidemic shocks: a three-state regime-switching approach

Jens Robben, Karim Barigou, Torsten Kleinow

Insurance Data Science Conference
London - June 19-20, 2025

University of Amsterdam

RCLR Research Centre for Longevity Risk

Introduction

Introduction

- Mortality rates are seasonal — with the highest rates occurring during winter (Marti-Soler et al. 2014).

Introduction

- Mortality rates are seasonal — with the highest rates occurring during winter (Marti-Soler et al. 2014).
- External shocks, such as severe heat waves and epidemics, can lead to deviations from these expected seasonal patterns — known as excess mortality (Iuliano et al. 2018, Nielsen et al. 2011).

Introduction

- Mortality rates are seasonal — with the highest rates occurring during winter (Marti-Soler et al. 2014).
- External shocks, such as severe heat waves and epidemics, can lead to deviations from these expected seasonal patterns — known as excess mortality (Iuliano et al. 2018, Nielsen et al. 2011).
- Various methodologies have been proposed:
  - Distributed Lag (Non-Linear) Models, e.g., Schwartz (2000), Gasparrini et al. (2010) and Guibert et al. (2024),
  - Extreme value analysis, e.g., bivariate POT approach in Li & Tang (2022).
  - Machine learning methods, e.g., gradient boosting for the association temperature-mortality, e.g. Robben et al. (2024).
  - Jump processes for pandemics, e.g., Cox et al. (2006), Chen & Cox (2009).

Research goals

1. Introduce a novel weekly mortality modeling framework that integrates a granular age- and region-specific baseline with a regime-switching model that captures mortality shocks driven both by environmental and epidemic factors.

Research goals

1. Introduce a novel weekly mortality modeling framework that integrates a granular age- and region-specific baseline with a regime-switching model that captures mortality shocks driven both by environmental and epidemic factors.

2. Use of several data sources to identify environmental and epidemic shocks:
   - Daily gridded temperature data from the Copernicus Climate Data Store.
   - Epidemic data from the French Sentinelles network (influenza anomalies).
   - Hospital admission records from Santé Publique France to capture the COVID-19 excess mortality.

Research goals

1. Introduce a novel weekly mortality modeling framework that integrates a granular age- and region-specific baseline with a regime-switching model that captures mortality shocks driven both by environmental and epidemic factors.

2. Use of several data sources to identify environmental and epidemic shocks:
   - Daily gridded temperature data from the Copernicus Climate Data Store.
   - Epidemic data from the French Sentinelles network (influenza anomalies).
   - Hospital admission records from Santé Publique France to capture the COVID-19 excess mortality.

3. Quantify the different sources of uncertainty around the model's estimates and forecasts, and analyze in-sample and out-of-sample performance.

Research goals

1. Introduce a novel weekly mortality modeling framework that integrates a granular age- and region-specific baseline with a regime-switching model that captures mortality shocks driven both by environmental and epidemic factors.

2. Use of several data sources to identify environmental and epidemic shocks:
   - Daily gridded temperature data from the Copernicus Climate Data Store.
   - Epidemic data from the French Sentinelles network (influenza anomalies).
   - Hospital admission records from Santé Publique France to capture the COVID-19 excess mortality.

3. Quantify the different sources of uncertainty around the model's estimates and forecasts, and analyze in-sample and out-of-sample performance.

4. Short-term mortality forecasting based on temperature (RCP 2.6, RCP 4.5, and RCP 8.5 pathways) and influenza scenarios based on a SIRS model.

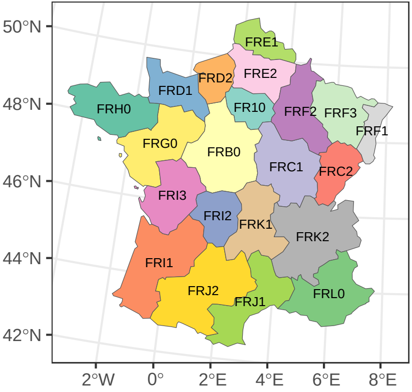
Data sources

Death counts

Eurostat: deaths by week, sex, 5-year age group and NUTS 2 region from France throughout the years 2013-2024 (21 regions).

Focus on the age groups 65-69, . . . , 90+.

French NUTS 2 regions

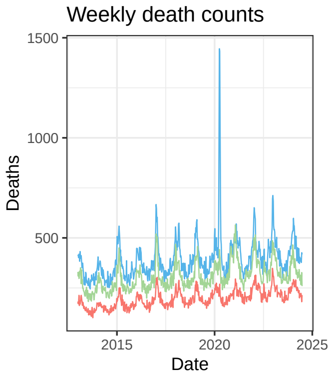


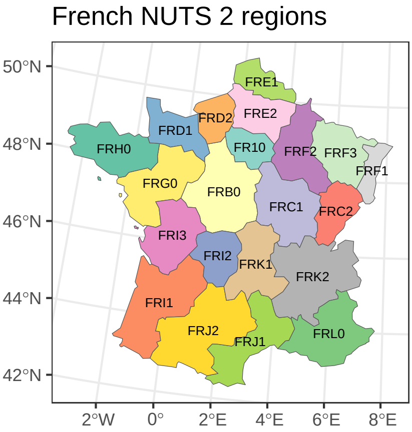
Death counts

Eurostat: deaths by week, sex, 5-year age group and NUTS 2 region from France throughout the years 2013-2024 (21 regions).

Focus on the age groups 65-69, . . . , 90+.

Seasonal trend:

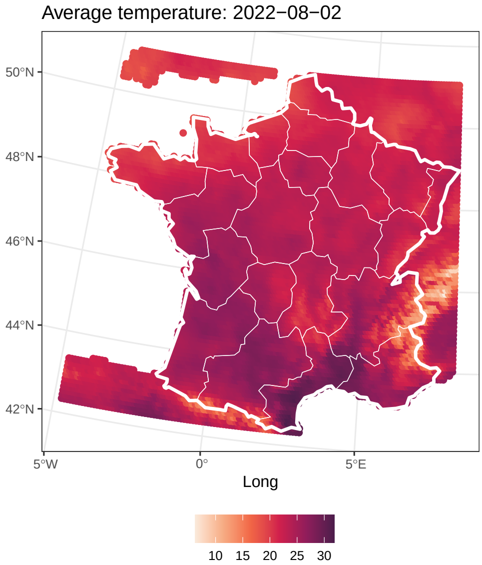




Weather data

E-OBS land-only, gridded meteorological data for Europe from the Copernicus Climate Data Store.

Daily, high-resolution gridded dataset, defined on a grid with spatial resolution of $0.10°$ (≈ 11 km).

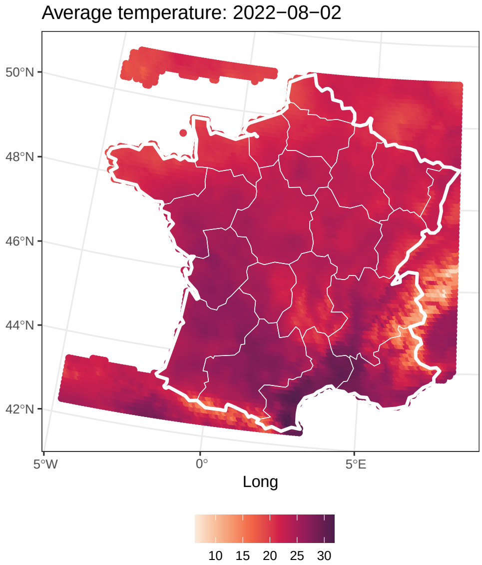


Weather data

E-OBS land-only, gridded meteorological data for Europe from the Copernicus Climate Data Store.

Daily, high-resolution gridded dataset, defined on a grid with spatial resolution of $0.10°$ (≈ 11 km).

To align with the NUTS 2-level mortality data:
$\Rightarrow$ Construction of population-weighted daily temperature averages by using gridded population.



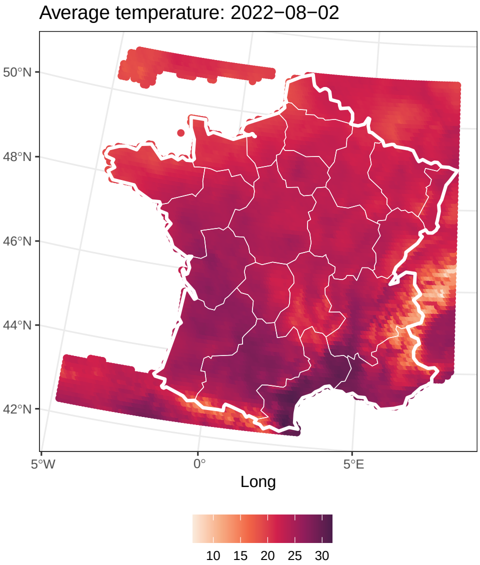
Weather data

E-OBS land-only, gridded meteorological data for Europe from the Copernicus Climate Data Store.

Daily, high-resolution gridded dataset, defined on a grid with spatial resolution of $0.10°$ (≈ 11 km).

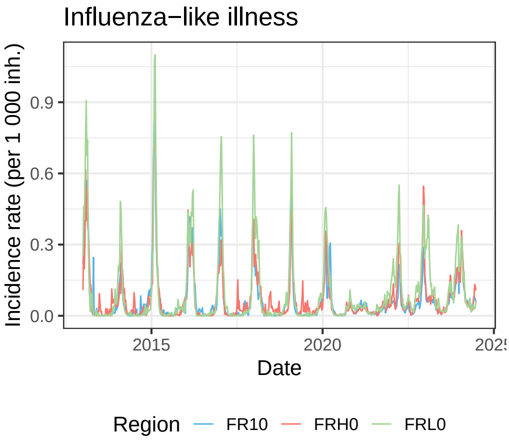
To align with the NUTS 2-level mortality data:
$\Rightarrow$ Construction of population-weighted daily temperature averages by using gridded population.

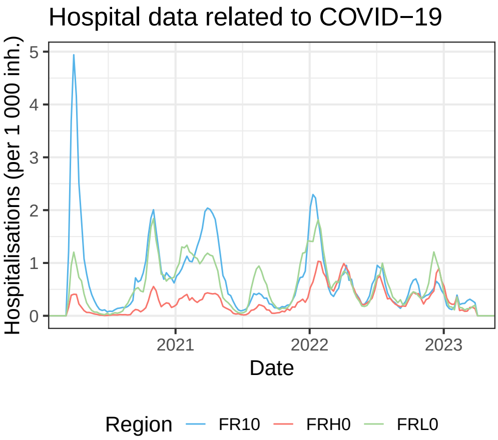
To align with the weekly time scale: hot- and cold-week index (frequency of hot/cold days) and weekly average of daily temperature anomalies (severity of hot/cold days).



Epidemic data: Influenza and COVID-19 hospitalizations

French Sentinelles Network: weekly influenza data from 1300 general practioners.
Santé Publique France: weekly COVID-19 hospitalizations.



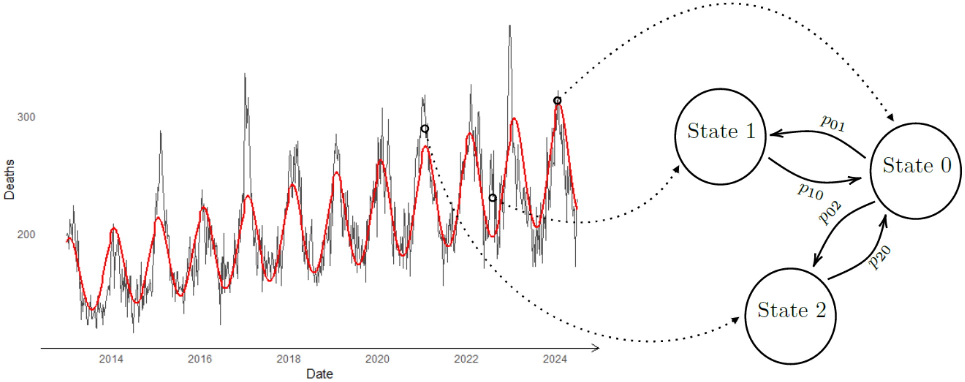


Model specification and calibration

Model specification

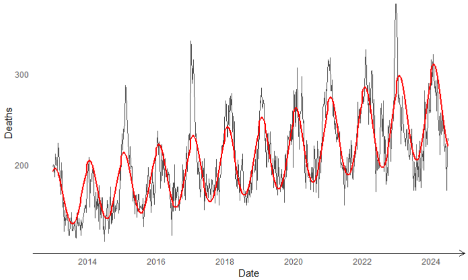
We propose a three-state regime-switching model:

- State 0 (Baseline state) : Weekly, region-specific and age-specific baseline mortality.
- State 1 (Environmental shock state) : Deviations due to extreme temperatures.
- State 2 (Respiratory shock state) : Deviations due to influenza and COVID-19.



Weekly, region- and age-specific baseline mortality model

A weekly, region and age-group-specific baseline mortality model to capture overall seasonal trends across all regions.



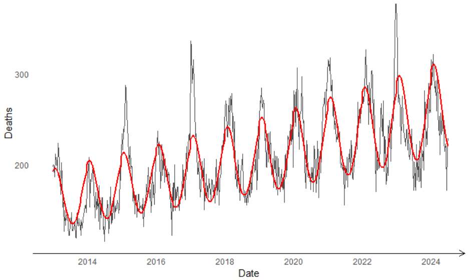
Weekly, region- and age-specific baseline mortality model

A weekly, region and age-group-specific baseline mortality model to capture overall seasonal trends across all regions.

Incorporate seasonality through Fourier terms:

$$D_{x,t}^{(r)} \sim \text{Poisson}\left(E_{x,t}^{(r)}\,\mu_{x,t}^{(r)}\right),$$

$$\log \mu_{x,t}^{(r)} = \gamma_{x,0}^{(r)} + \gamma_{x,1}^{(r)} t + \gamma_{x,2}^{(r)} \sin\left(\frac{2\pi w(t)}{52.18}\right) + \gamma_{x,3}^{(r)} \cos\left(\frac{2\pi w(t)}{52.18}\right) + \gamma_{x,4}^{(r)} \sin\left(\frac{2\pi w(t)}{26.09}\right) + \gamma_{x,5}^{(r)} \cos\left(\frac{2\pi w(t)}{26.09}\right),$$



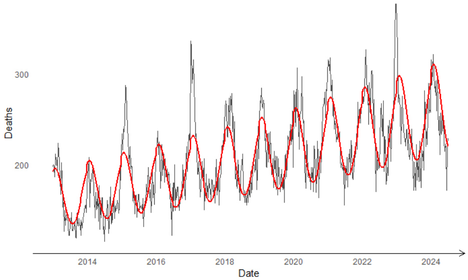
Weekly, region- and age-specific baseline mortality model

A weekly, region and age-group-specific baseline mortality model to capture overall seasonal trends across all regions.

Incorporate seasonality through Fourier terms:

$$D_{x,t}^{(r)} \sim \text{Poisson}\left(E_{x,t}^{(r)}\,\mu_{x,t}^{(r)}\right),$$

$$\log \mu_{x,t}^{(r)} = \gamma_{x,0}^{(r)} + \gamma_{x,1}^{(r)} t + \gamma_{x,2}^{(r)} \sin\left(\frac{2\pi w(t)}{52.18}\right) + \gamma_{x,3}^{(r)} \cos\left(\frac{2\pi w(t)}{52.18}\right) + \gamma_{x,4}^{(r)} \sin\left(\frac{2\pi w(t)}{26.09}\right) + \gamma_{x,5}^{(r)} \cos\left(\frac{2\pi w(t)}{26.09}\right),$$



Fit one Poisson GLM jointly on all regions, and add a penalty term to obtain smooth variations in the estimated $\boldsymbol{\gamma}_{x,p} = (\gamma_{x,p}^{(r)})_{r\in\mathcal{R}}$ across neighbouring regions.

Modelling mortality deviations from the baseline

Explain observed deviations from the baseline deaths using region-specific environmental and epidemiological features:

$$\hat{b}_{x,t}^{(r)} := \hat{E}\left[D_{x,t}^{(r)}\right] = E_{x,t}^{(r)}\,\hat{\mu}_{x,t}^{(r)},$$

Death counts are modeled for $i = 0, 1, 2$ by

$$D_{x,t}^{(r)} \mid S_t^{(r)} = i \sim \text{POI}\left(\hat{b}_{x,t}^{(r)} \cdot \exp\left[\left(\boldsymbol{z}_t^{(r)}\right)^\top \boldsymbol{\alpha}_{i,x}\right]\right),$$

where

- $S_t^{(r)}$ is a region- and time-dependent Markov chain.
- Region- and time-dependent covariate vector $\boldsymbol{z}_t^{(r)}$, with state- and age-specific $\boldsymbol{\alpha}_{i,x}$.

Motivation: Extreme temperatures can have a larger impact on people aged 90+ compared to those aged 65-69.

Modelling mortality deviations from the baseline

Transition probabilities are given by

$$p_t^{ij}\left(\boldsymbol{z}_t^{(r)};\boldsymbol{\beta},u_r\right)=\begin{cases}\dfrac{\exp\left(\left(\boldsymbol{z}_t^{(r)}\right)^\top\boldsymbol{\beta}_{ij}+U_r\right)}{1+\sum_{j'\in\mathcal{J}_i}\exp\left(\left(\boldsymbol{z}_t^{(r)}\right)^\top\boldsymbol{\beta}_{ij'}+U_r\right)} & j\neq 0\\[2ex] \dfrac{1}{1+\sum_{j'\in\mathcal{J}_i}\exp\left(\left(\boldsymbol{z}_t^{(r)}\right)^\top\boldsymbol{\beta}_{ij'}+U_r\right)} & j=0,\end{cases}$$

Motivation: If very high temperatures are observed at time t, the probability of moving to state 1 should increase. We include a spatial effect to account for regional disparities by including an ICAR model:

$$\boldsymbol{U}=(U_1,U_2,\dots,U_R)\sim\mathcal{N}\left(\boldsymbol{0},\left[\tau\cdot(\boldsymbol{D}-\boldsymbol{W})\right]^{-1}\right),$$

Calibration: Expectation-Maximization algorithm.

Case study on 21 French NUTS 2 regions

State-Specific Poisson Model Specifications

State 1: Models impact of heatwave-related shocks:

$$\log \mathbb{E}\left[D_{x,t}^{(r)} \mid S_t^{(r)} = 1\right] = \log \hat{b}_{x,t}^{(r)} + \sum_{a \in \mathcal{X}_{\text{red}}} \Big(\alpha_{1,1}^{(a)} \mathsf{TA}_t^{(r)} + \alpha_{1,2}^{(a)} \mathsf{TA}_{t-1}^{(r)} + \alpha_{1,3}^{(a)} \mathsf{TA}_{t-2}^{(r)} + \alpha_{1,4}^{(a)} \mathsf{HI}_t^{(r)} + \alpha_{1,5}^{(a)} \mathsf{HI}_{t-1}^{(r)} + \alpha_{1,6}^{(a)} \mathsf{HI}_{t-2}^{(r)} \Big) \mathbb{1}\left\{x \subset a\right\}.$$

State 2: Models mortality shocks from influenza activity and COVID-19 hospitalizations:

$$\log \mathbb{E}\left[D_{x,t}^{(r)} \mid S_t^{(r)} = 2\right] = \log \hat{b}_{x,t}^{(r)} + \sum_{a \in \mathcal{X}_{\text{red}}} \Big(\alpha_{2,1}^{(a)} \mathsf{IA}_{t,t-1}^{(r)} + \alpha_{2,2}^{(a)} \mathsf{IA}_{t-2,t-3}^{(r)} + \alpha_{2,3}^{(a)} \mathsf{CI}_{t,t-1}^{(r)} + \alpha_{2,4}^{(a)} \mathsf{CI}_{t-2,t-3}^{(r)} + \alpha_{2,5}^{(a)} \mathsf{HA}_{t,t-1}^{(r)} + \alpha_{2,6}^{(a)} \mathsf{HA}_{t-2,t-3}^{(r)} \Big) \mathbb{1}\left\{x \subset a\right\}.$$

Modelling regime transition probabilities

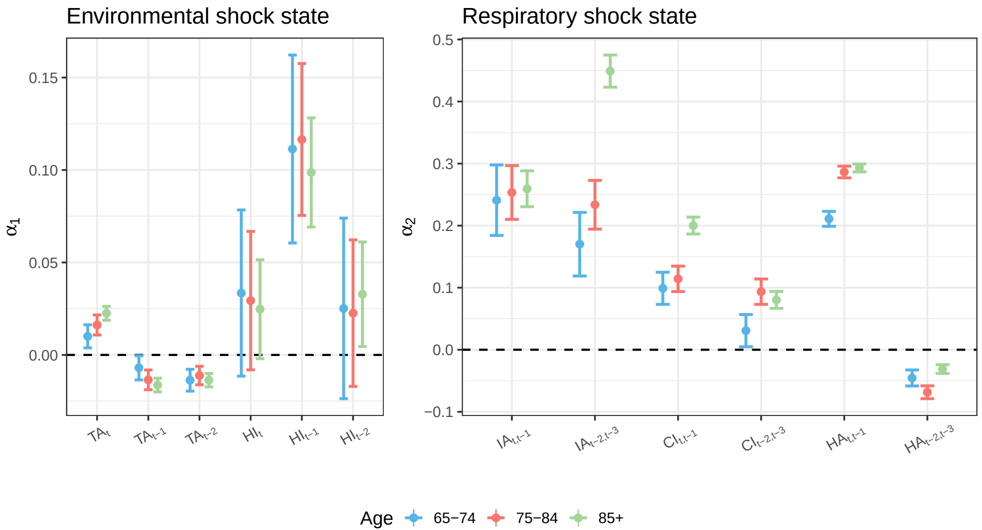
Assumptions: Transition probabilities are independent of age x and regional variations r are accounted for using a spatial effect U_r modelled by an ICAR process.

Transition Probabilities:

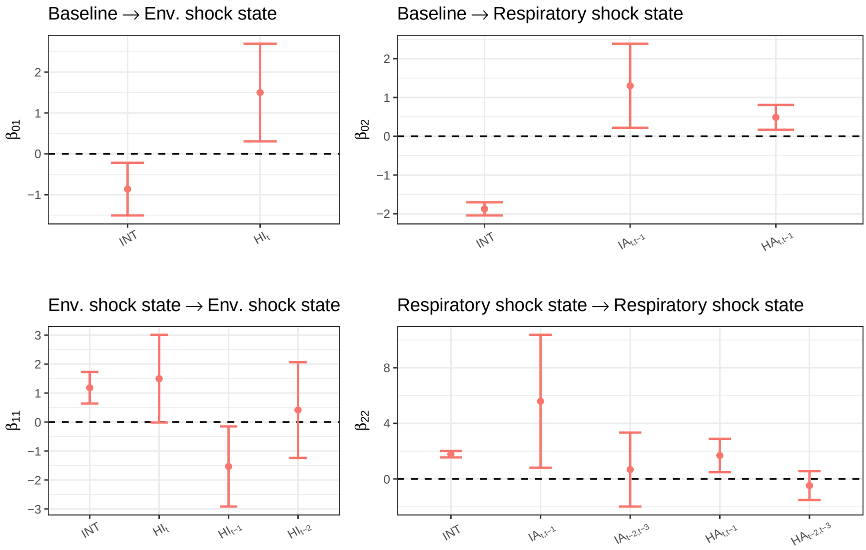
$$
\begin{aligned}
\operatorname{logit} p_t^{01}\left(\boldsymbol{z}_t^{(r)};\, \boldsymbol{\beta}_0, u_r\right) &= \beta_{01,0} + \beta_{01,1}\mathsf{HI}_t + U_r \\
\operatorname{logit} p_t^{02}\left(\boldsymbol{z}_t^{(r)};\, \boldsymbol{\beta}_0, u_r\right) &= \beta_{02,0} + \beta_{02,1}\mathsf{IA}_{t,t-1}^{(r)} + \beta_{02,2}\mathsf{HA}_{t,t-1}^{(r)} + U_r \\
\operatorname{logit} p_t^{11}\left(\boldsymbol{z}_t^{(r)};\, \boldsymbol{\beta}_1, u_r\right) &= \beta_{11,0} + \beta_{11,1}\mathsf{HI}_t + \beta_{11,2}\mathsf{HI}_{t-1} + \beta_{11,3}\mathsf{HI}_{t-2} + U_r \\
\operatorname{logit} p_t^{22}\left(\boldsymbol{z}_t^{(r)};\, \boldsymbol{\beta}_2, u_r\right) &= \beta_{22,0} + \beta_{22,1}\mathsf{IA}_{t,t-1}^{(r)} + \beta_{22,2}\mathsf{IA}_{t-2,t-3}^{(r)} \\
&\quad + \beta_{22,3}\mathsf{HA}_{t,t-1}^{(r)} + \beta_{22,4}\mathsf{HA}_{t-2,t-3}^{(r)} + U_r
\end{aligned}
$$

Features: Short-term features for transitions to shock states; mid-term lagged features for state persistence: HI (hot index), IA (influenza anomaly), HA (hospital admissions).

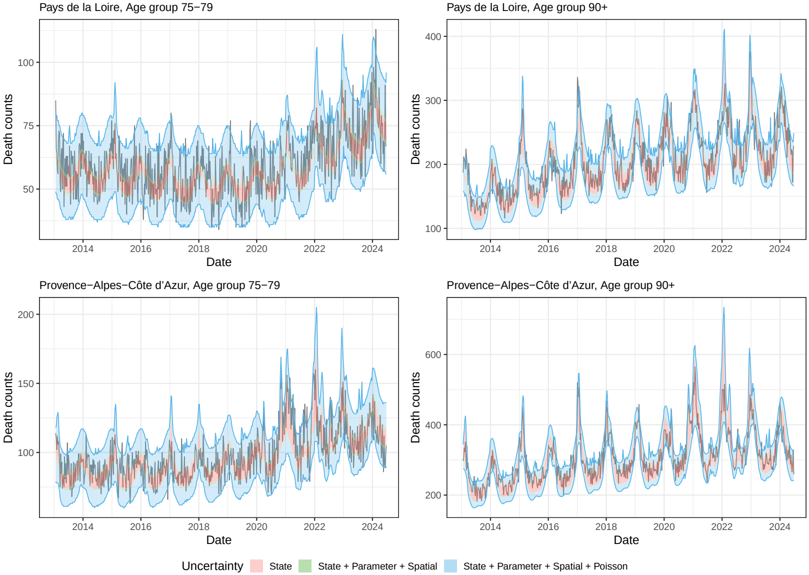
Results: Parameter estimates in both states



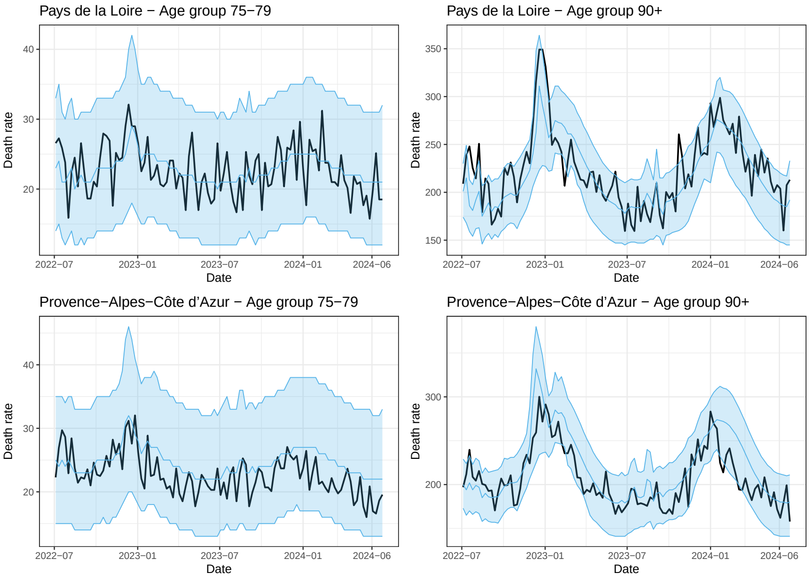
Results: Parameter estimates in transition probabilities



Uncertainty in in-sample predictions



Out-of-sample backtesting - Calibration: 2013 to mid 2022



Conclusion

Conclusion

Main results

- We proposed a three-state regime-switching weekly mortality model incorporating both the impact of temperature and epidemic shocks on mortality.
- We quantified the uncertainty in in- and out-of-sample predictions, and examine how different temperature and influenza scenarios influence mortality.
- Highest impact for the oldest age group and presence of harvesting effects.

Limitations and extensions

- Public Health Interventions: Adaptation measures like early-warning systems, cooling centers, and improved healthcare access can mitigate effects.
- Future Research: Extend analysis to morbidity data for better preparedness of hospitals and public healthcare systems.

References I

Chen, H. & Cox, S. H. (2009), 'Modeling Mortality With Jumps: Applications to Mortality Securitization', *Journal of Risk and Insurance* **76**(3), 727–751.

Cox, S. H., Lin, Y. & Wang, S. (2006), 'Multivariate Exponential Tilting and Pricing Implications for Mortality Securitization', *Journal of Risk and Insurance* **73**(4), 719–736.

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

Guibert, Q., Pincemin, G. & Planchet, F. (2024), 'Impacts of climate change on mortality: An extrapolation of temperature effects based on time series data in France', *arXiv preprint arXiv:2406.02054* .

Iuliano, A. D., Roguski, K. M., Chang, H. H., Muscatello, D. J., Palekar, R., Tempia, S., Cohen, C., Gran, J. M., Schanzer, D., Cowling, B. J. et al. (2018), 'Estimates of global seasonal influenza-associated respiratory mortality: a modelling study', *The Lancet* **391**(10127), 1285–1300.

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

Marti-Soler, H., Gonseth, S., Gubelmann, C., Stringhini, S., Bovet, P., Chen, P.-C., Wojtyniak, B., Paccaud, F., Tsai, D.-H., Zdrojewski, T. et al. (2014), 'Seasonal variation of overall and cardiovascular mortality: a study in 19 countries from different geographic locations', *PLOS One* **9**(11), e113500.

References II

Nielsen, J., Mazick, A., Glismann, S. & Mølbak, K. (2011), 'Excess mortality related to seasonal influenza and extreme temperatures in Denmark, 1994-2010', *BMC Infectious Diseases* **11**, 1–13.

Robben, J., Antonio, K. & Kleinow, T. (2024), 'The short-term association between environmental variables and mortality: evidence from Europe', *arXiv preprint arXiv:2405.18020* .

Schwartz, J. (2000), 'The distributed lag between air pollution and daily deaths', *Epidemiology* **11**(3), 320–326.