

Climate–Driven Mortality Forecasting

Weekly Seasonal Lee–Carter and Deep Learning Framework

Kenrick So¹ Karim Barigou¹ Jens Robben²

¹ISBA – LIDAM, UCLouvain

²Research Centre for Longevity Risk, University of Amsterdam

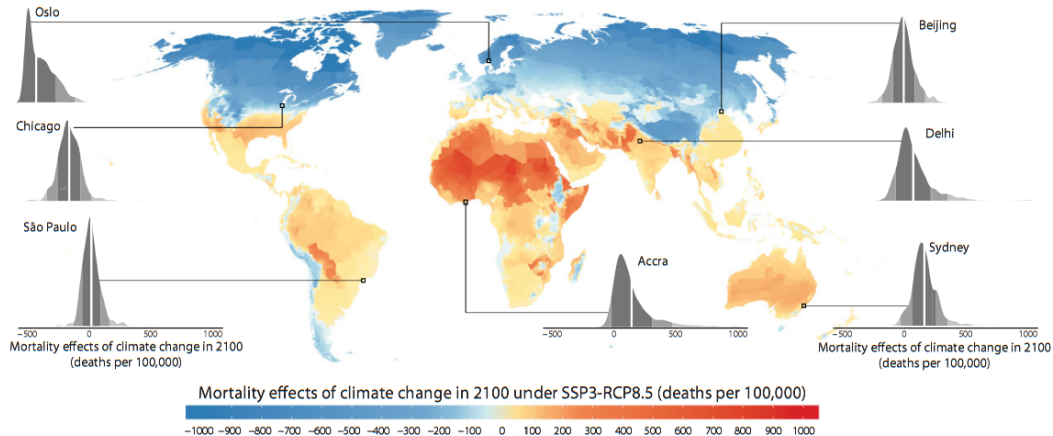


Figure: Forecasted **impact of climate change on the net death rate in 2100** under RCP8.5 emissions and SSP3 economic growth. Figure from [Carleton et al. \(2022\)](#).

Why it matters

- Mortality assumptions drive capital requirements and pension liabilities.
- Extreme weather causes immediate and **delayed mortality shocks**.
- Traditional mortality models struggle to capture these climate-driven shocks.

Why it matters

- Mortality assumptions drive capital requirements and pension liabilities.
- Extreme weather causes immediate and **delayed mortality shocks**.
- Traditional mortality models struggle to capture these climate-driven shocks.

Research gap

- DLNMs scale poorly to many interacting covariates ([Chen et al., 2019](#)).
- Deep learning mortality models omit climate information.
- **MortFCNet incorporates GRU without spatial structure** ([Zheng et al., 2025](#)).

Why it matters

- Mortality assumptions drive capital requirements and pension liabilities.
- Extreme weather causes immediate and **delayed mortality shocks**.
- Traditional mortality models struggle to capture these climate-driven shocks.

Research gap

- DLNMs scale poorly to many interacting covariates ([Chen et al., 2019](#)).
- Deep learning mortality models omit climate information.
- **MortFCNet incorporates GRU without spatial structure** ([Zheng et al., 2025](#)).

Objective

- Two-step framework: weekly seasonal Lee–Carter + learned climate residual.
- Compare **CNN-LSTM** and **GNN-LSTM** architectures.
- Extend to quantify uncertainty and produce interval forecasts.

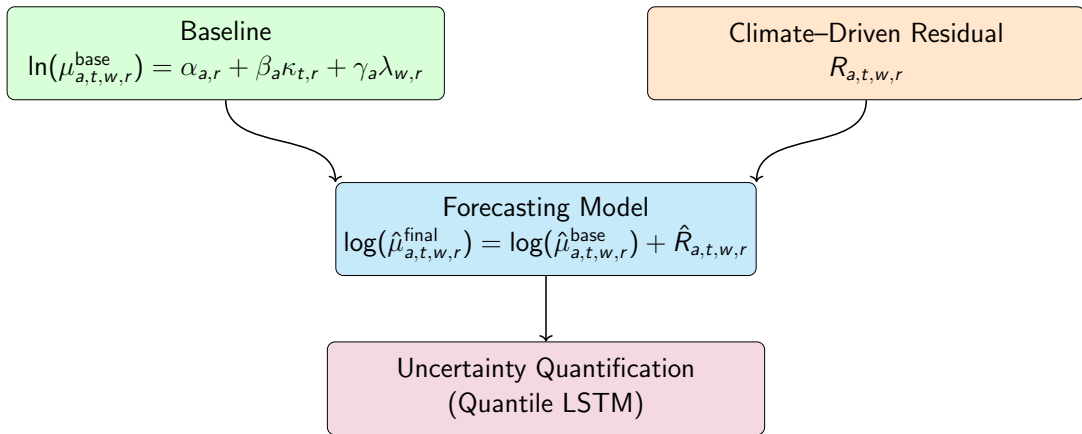


Figure: Our climate-forecasting framework captures **age- and region-specific climate effects** on excess mortality and is the first to incorporate **probabilistic forecasts**.

Mortality Data

- **Setting:** Metropolitan France (21 NUTS2 regions), 1990–2019.
- **Population:** Ages 65+, grouped into 5-year age bands.
- **Indexing:** a (age), t (year), w (ISO week), r (region).
- Observed mortality is defined as:

$$m_{a,t,w,r} = \frac{d_{a,t,w,r}}{E_{a,t,w,r}}$$

where $d_{a,t,w,r}$ denotes weekly deaths and $E_{a,t,w,r}$ exposure.

Climate Data

Source: E-OBS gridded observational dataset (Copernicus Climate Data Store), 0.10° resolution (≈ 10 km), aggregated to NUTS2 level and converted to ISO weeks.

Variable	Definition	Weekly Aggregation
$T_{\max}$	Daily maximum temperature (°C)	Max
T_{avg}	Daily average temperature (°C)	Mean
$T_{\min}$	Daily minimum temperature (°C)	Min
Hum	Relative humidity (%)	Mean
Rain	Total precipitation (mm)	Max
Wind	Wind speed (m/s)	Mean

Outline

1. Two-Step Forecasting Framework

2. Uncertainty Quantification

Baseline Model

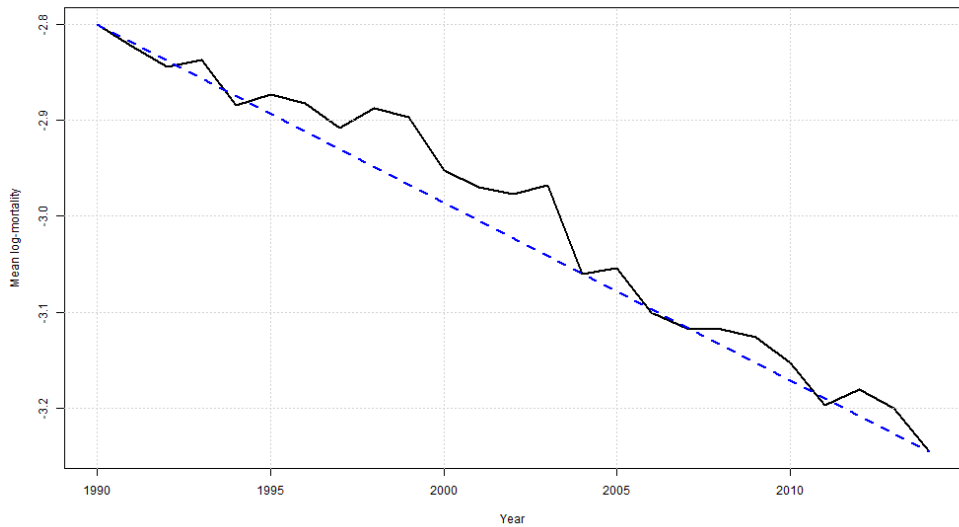
We follow [Robben and Barigou \(2025\)](#) using a **weekly seasonal Lee–Carter**,

$$\ln(\mu_{a,t,w,r}^{\text{base}}) = \alpha_{a,r} + \beta_a \kappa_{t,r} + \gamma_a \lambda_{w,r},$$

where

- $\alpha_{a,r}$: baseline log-mortality for age a , region r .
- β_a : age-specific sensitivity.
- $\kappa_{t,r}$: region-specific annual mortality trend.
- γ_a : age-specific seasonal effects.
- $\lambda_{w,r}$: intra-year seasonal mortality pattern.

Yearly National LC Mean Log-Mortality without Breakpoint



We follow [Van Berkum et al. \(2016\)](#) and decompose the annual mortality trend into national and regional deviations

$$\kappa_{t,r} = \underbrace{K_t}_{\text{national}} + \underbrace{u_{t,r}}_{\text{regional}} .$$

Minimize BIC to determine number of breakpoints and location over fitted log-mortality. The drift of the national factor is estimated by OLS over most recent post break point b_{m^*}

$$\pi = \frac{K_{t_{\max}} - K_{b_{m^*}+1}}{t_{\max} - b_{m^*}+1} \quad b_{m^*} < t \leq t_{\max}.$$

The national factor is forecasted using a random walk with drift,

$$K_t = K_{t-1} + \pi + \epsilon_t \quad \epsilon_t \sim N(0, \sigma^2).$$

Regional deviations are forecasted using an AR(1) with intercept ([Li and Lee, 2005](#)).

Climate-Driven Excess Mortality

We model short-term deviations from baseline mortality using log-mortality residuals:

$$R_{a,t,w,r} = \log \left(\frac{d_{a,t,w,r}}{E_{a,t,w,r}} \right) - \log \left(\hat{\mu}_{a,t,w,r}^{\text{base}} \right).$$

We learn these residuals as a nonlinear function of climate covariates:

$$\hat{R}_{a,t,w,r} = f_{\theta}(\mathbf{x}_{\tau,r})$$

where $\mathbf{x}_{\tau,r} \in \mathbb{R}^d$ denotes the vector of weekly climate variables for region r at week τ .
The final mortality model:

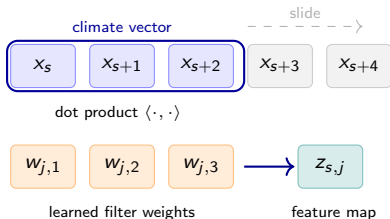
$$\log(\hat{\mu}_{a,t,w,r}^{\text{final}}) = \log(\hat{\mu}_{a,t,w,r}^{\text{base}}) + \hat{R}_{a,t,w,r}.$$

Convolutional Neural Network

A convolutional filter scans a window of consecutive weeks and learns recurring local temporal patterns

$$z_{s,j}(\mathbf{X}_{\tau,r}) = g \left(b_j + \sum_{\ell=1}^m \langle \mathbf{w}_{j,\ell}, \mathbf{x}_{\tau-T+s+\ell-1,r} \rangle \right).$$

The output $z_{s,j}$ is the activation produced when filter j is applied to the window beginning at position s .



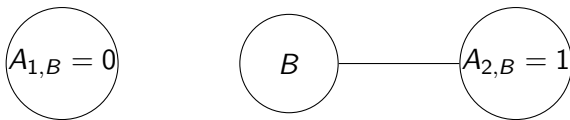
Graph Convolutional Network

A graph convolution propagates climate information across adjacent regions,

$$\mathbf{P}^{(\delta+1)} = g\left(\tilde{\mathbf{A}} \mathbf{P}^{(\delta)} \mathbf{W}^{(\delta)} + \mathbf{b}^{(\delta)}\right),$$

where

- $\mathbf{A} \in \{0, 1\}^{R \times R}$: adjacency matrix, $A_{ij} = 1$ if regions i and j share a border.
- $\tilde{\mathbf{A}} = \mathbf{D}^{-1/2} \mathbf{A} \mathbf{D}^{-1/2}$: **normalized adjacency** used for information passing.
- δ : current GCN layer
- $\mathbf{P}^{(\delta)} \rightarrow \mathbf{P}^{(\delta+1)}$: node representations before and after propagation.



Region embedding (Wuthrich and Merz, 2023)

$$\xi_{\mathcal{R}} : \mathcal{R} \rightarrow \mathbb{R}^{q_{\mathcal{R}}}$$

The model assigns each region a **learned low-dimensional vector** that captures regional effects not encoded by the CNN/GNN structure.

Region embedding (Wuthrich and Merz, 2023)

$$\xi_{\mathcal{R}} : \mathcal{R} \rightarrow \mathbb{R}^{q_{\mathcal{R}}}$$

The model assigns each region a **learned low-dimensional vector** that captures regional effects not encoded by the CNN/GNN structure.

Age representation

$$a_{\text{mid}} \in \{67.5, 72.5, 77.5, 82.5, 87.5, 92.5\}$$

We encode each 5-year age group by its midpoint to capture the **monotonic sensitivity between climate extremes and age**.

Long-Short Term Memory

The three gates control how past climate and mortality signals are filtered over time:

$$\mathbf{f}_{s,r,a} = \text{sig}(\mathbf{b}^{(f)} + \mathbf{W}^{(f)}\mathbf{u}_{s,r,a} + \mathbf{V}^{(f)}\mathbf{h}_{s,r,a}) \quad (\text{forget})$$

$$\mathbf{i}_{s,r,a} = \text{sig}(\mathbf{b}^{(i)} + \mathbf{W}^{(i)}\mathbf{u}_{s,r,a} + \mathbf{V}^{(i)}\mathbf{h}_{s,r,a}) \quad (\text{input})$$

$$\mathbf{o}_{s,r,a} = \text{sig}(\mathbf{b}^{(o)} + \mathbf{W}^{(o)}\mathbf{u}_{s,r,a} + \mathbf{V}^{(o)}\mathbf{h}_{s,r,a}) \quad (\text{output})$$

The candidate update represents new climate-driven information:

$$\tilde{\mathbf{h}}_{s,r,a} = \tanh(\mathbf{b}^{(c)} + \mathbf{W}^{(c)}\mathbf{u}_{s,r,a} + \mathbf{V}^{(c)}\mathbf{h}_{s,r,a})$$

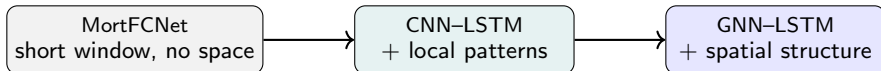
Memory evolves by combining past and new climate effects:

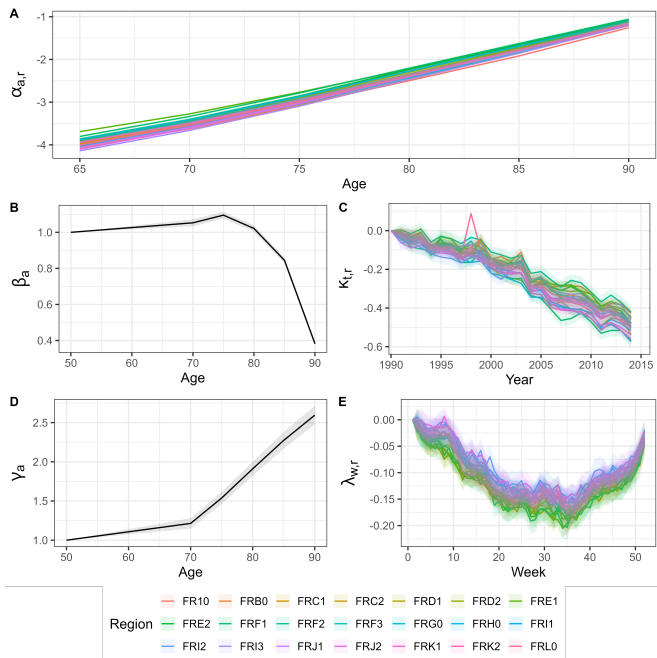
$$\mathbf{c}_{s+1,r,a} = \mathbf{f}_{s,r,a} \odot \mathbf{c}_{s,r,a} + \mathbf{i}_{s,r,a} \odot \tilde{\mathbf{h}}_{s,r,a}$$

Final output reflects lagged effects carried forward:

$$\mathbf{h}_{s+1,r,a} = \mathbf{o}_{s,r,a} \odot \tanh(\mathbf{c}_{s+1,r,a})$$

	MortFCNet (benchmark)	CNN-LSTM (proposed)	GNN-LSTM (proposed)
Temporal encoder	GRU	CNN + LSTM	GCN + LSTM
Look-back window	2 weeks	4 weeks	4 weeks
Local climate patterns	—	(conv. filters)	(adjacency $\tilde{\mathbf{A}}$)
Region representation	integer ID	learned embedding $\xi_{\mathcal{R}}$	learned embedding $\xi_{\mathcal{R}}$
Age input	—	standardized scalar a_{mid}	standardized scalar a_{mid}





Per-region Test MSE

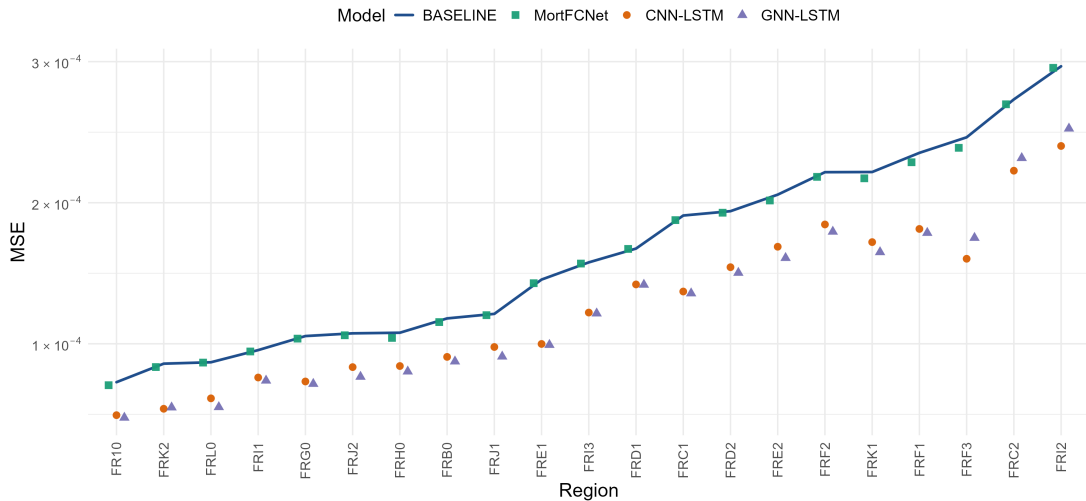


Figure: Low MSE regions have similar climate–mortality interactions (GNN–LSTM). High MSE regions have different spatial interaction, **adjacency constraints become restrictive**.

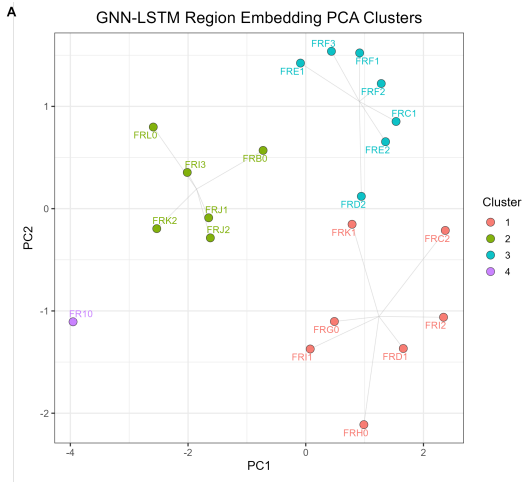
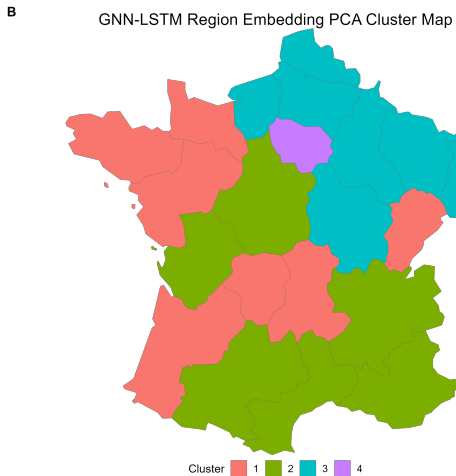


Figure: Project the region embeddings into two dimensions using PCA and applied k-means clustering to identify similar regions. Clusters of region embeddings show Paris, North, South, and West of France.

FRK2

Age group: 90+

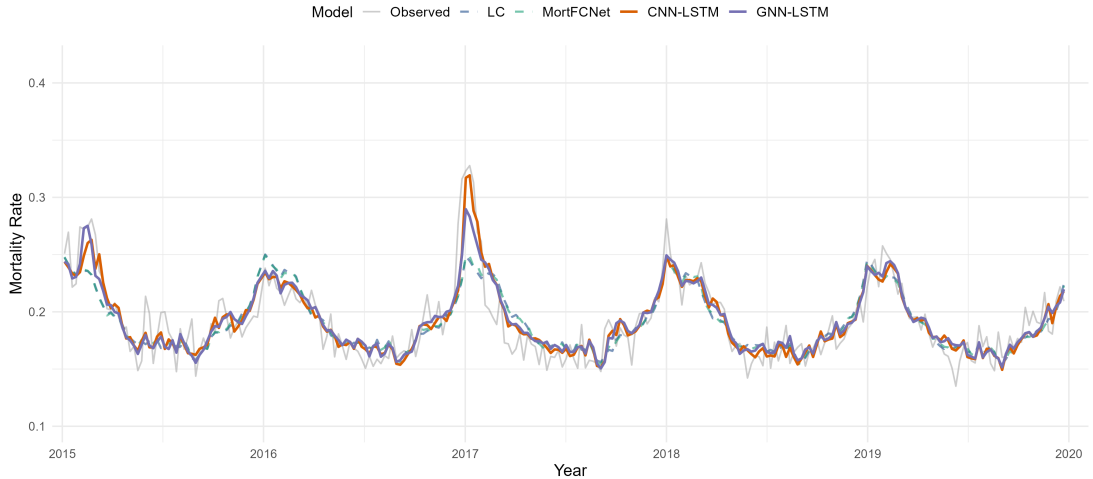


Figure: Our two architectures capture **peaks in mortality** particularly well during winter months.

Climate Feature Importance (CNN-LSTM)

Top 15 climate predictors ranked by permutation importance

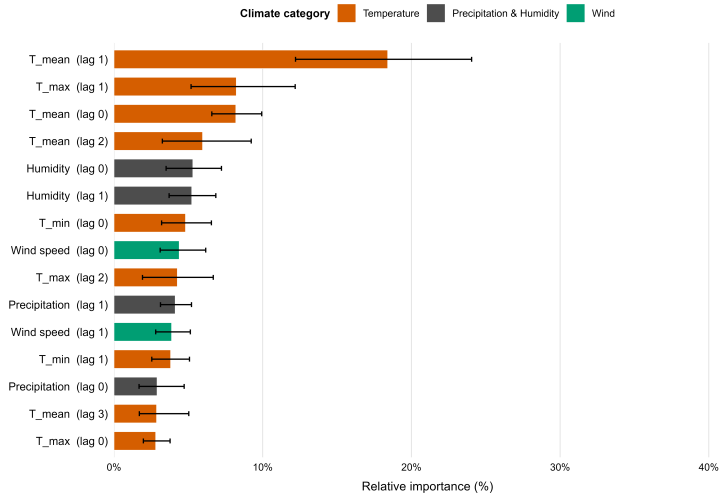


Figure: Mean temperature and its lags are the main driver of climate–mortality dynamics.

Outline

1. Two-Step Forecasting Framework

2. Uncertainty Quantification

1. Baseline Parameter Uncertainty

- Resample death counts from negative binomial distribution.
- Bootstrap the baseline model.

1. Baseline Parameter Uncertainty

- Resample death counts from negative binomial distribution.
- Bootstrap the baseline model.

2. Baseline Forecast Uncertainty

- Simulate random walk with drift and AR(1) processes.

1. Baseline Parameter Uncertainty

- Resample death counts from negative binomial distribution.
- Bootstrap the baseline model.

2. Baseline Forecast Uncertainty

- Simulate random walk with drift and AR(1) processes.

3. Residual Climate Uncertainty

- Quantile LSTM (pinball loss) estimates the τ -quantiles of the climate-driven residual.
- Parameterize a split-normal distribution and draws a residual $R_{a,t,w,r}^{(b)}$ from it.

Uncertainty Quantification

- Prediction Interval Coverage Probability: share of observations inside the predicted interval.
- Mean Prediction Interval Width: average interval width.

Uncertainty Quantification

- Prediction Interval Coverage Probability: share of observations inside the predicted interval.
- Mean Prediction Interval Width: average interval width.

Model	PICP	MPIW
Baseline	0.386	0.0084
MortFCNet	0.839	0.0239
CNN-LSTM	0.919	0.0333
GNN-LSTM	0.905	0.0285

Table: GNN-LSTM achieves target coverage with lower interval width than CNN-LSTM (target: 90%).

CNN-LSTM — Forecast Uncertainty

Region: FRK2 | Age group: 90

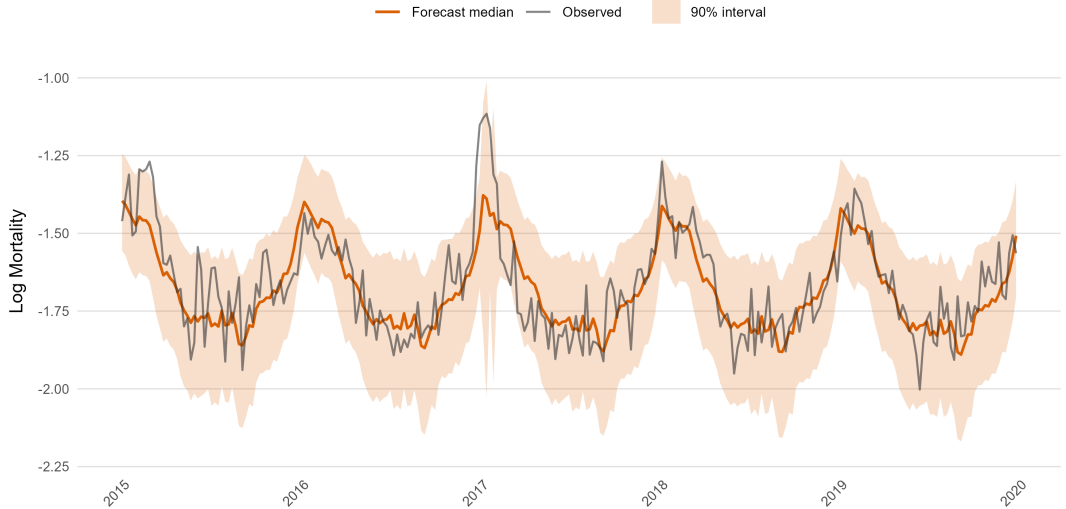


Figure: Quantile LSTM (CNN-LSTM) captures mortality peaks within 90% prediction interval.

Methods

- Developed a two-step framework that separates mortality and climate effects.
- CNN/GNN-LSTM captures climate-mortality dynamics across regions and over time.
- Quantile LSTM generates climate-sensitive prediction intervals.

Results

- Outperforms benchmark models in forecast accuracy across age groups and regions.
- Prediction interval coverage that captures mortality peaks.

Applications

- Supports long-term planning for pension funds and life insurers.
- Enables longevity stress-testing under climate scenarios.

Preprint:

bit.ly/DeepMortalityForecasting



References I

- Carleton, T., Jina, A., Delgado, M., Greenstone, M., Houser, T., Hsiang, S., Hultgren, A., Kopp, R. E., McCusker, K. E., Nath, I., Rising, J., Rode, A., Seo, H. K., Viaene, A., Yuan, J., and Zhang, A. T. (2022). Valuing the global mortality consequences of climate change accounting for adaptation costs and benefits. *The Quarterly Journal of Economics*, 137(4):2037–2105.
- Chen, Y.-H., Mukherjee, B., and Berrocal, V. J. (2019). Distributed lag interaction models with two pollutants. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 68(1):79–97. doi:[10.1111/rssc.12297](https://doi.org/10.1111/rssc.12297).
- Li, N. and Lee, R. (2005). Coherent mortality forecasts for a group of populations: An extension of the lee-carter method. *Demography*, 42(3):575–594. doi:[10.1353/dem.2005.0021](https://doi.org/10.1353/dem.2005.0021).
- Robben, J., Antonio, K., and Kleinow, T. (2026). The short-term association between environmental variables and mortality: evidence from europe. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 189(2):1131–1153. doi:[10.1093/jrsssa/qnaf052](https://doi.org/10.1093/jrsssa/qnaf052).
- Robben, J. and Barigou, K. (2025). A penalized distributed lag non-linear Lee-Carter framework for regional weekly mortality forecasting. arxiv.org/abs/2509.24087.
- Van Berkum, F., Antonio, K., and Vellekoop, M. (2016). The impact of multiple structural changes on mortality predictions. *Scandinavian Actuarial Journal*, 2016(7):581–603. doi:[10.1080/03461238.2014.987807](https://doi.org/10.1080/03461238.2014.987807).

References II

- Wuthrich, M. V. and Merz, M. (2023). *Statistical Foundations of Actuarial Learning and its Applications*. Springer International Publishing. doi:10.1007/978-3-031-12409-9.
- Zheng, H., Wang, H., Zhu, R., and Xue, J.-H. (2025). Fine-grained mortality forecasting with deep learning. *Annals of Actuarial Science*, page 1–27. doi:10.1017/s1748499525100171.

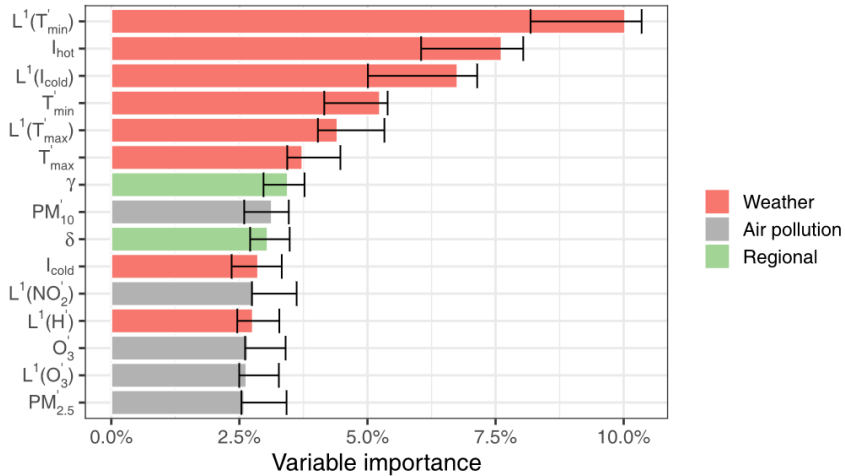


Figure: **Weather covariates** have a strong association with weekly mortality rates. Figure from [Robben et al. \(2026\)](#).

Model	Train MSE	Test MSE
Baseline	3.13×10^{-4}	1.65×10^{-4}
MortFCNet	3.03×10^{-4}	1.62×10^{-4}
CNN-LSTM	2.38×10^{-4}	1.26×10^{-4}
GNN-LSTM	2.20×10^{-4}	1.25×10^{-4}

Table: Our proposed architectures achieve **24% lower test MSE** than the baseline.

Age Group	Baseline	MortFCNet	CNN-LSTM	GNN-LSTM
65-69	6.40×10^{-6}	9.41×10^{-6}	6.12×10^{-6}	5.71×10^{-6}
70-74	1.05×10^{-5}	1.32×10^{-5}	1.01×10^{-5}	9.85×10^{-6}
75-79	2.34×10^{-5}	2.53×10^{-5}	2.22×10^{-5}	2.16×10^{-5}
80-84	5.23×10^{-5}	5.20×10^{-5}	4.72×10^{-5}	4.54×10^{-5}
85-89	1.74×10^{-4}	1.70×10^{-4}	1.44×10^{-4}	1.39×10^{-4}
90+	7.22×10^{-4}	7.03×10^{-4}	5.29×10^{-4}	5.30×10^{-4}

Table: Explicitly incorporating region information improves forecasted MSE for most age groups

Temperature Response Comparison - FRK2 (Age 90+)

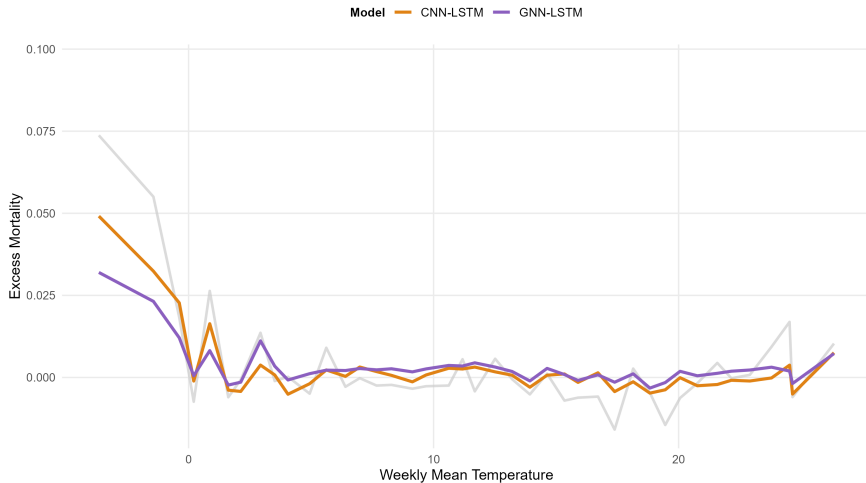


Figure: Both models capture the climate–mortality relationship during **extreme temperatures**.

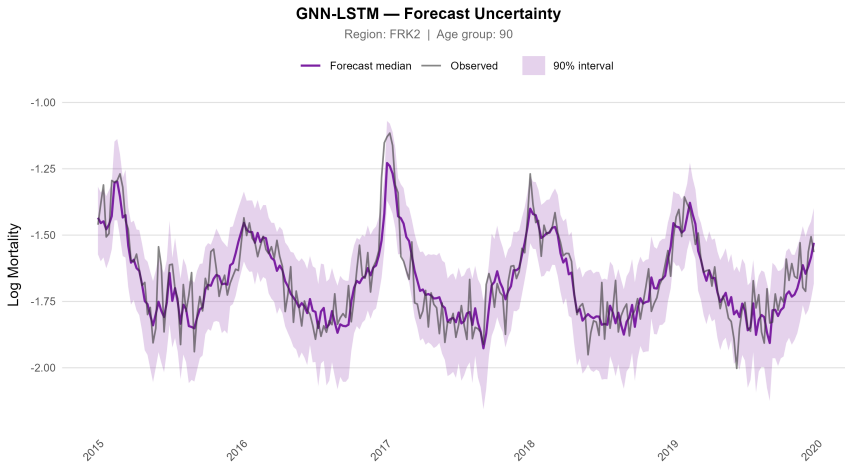
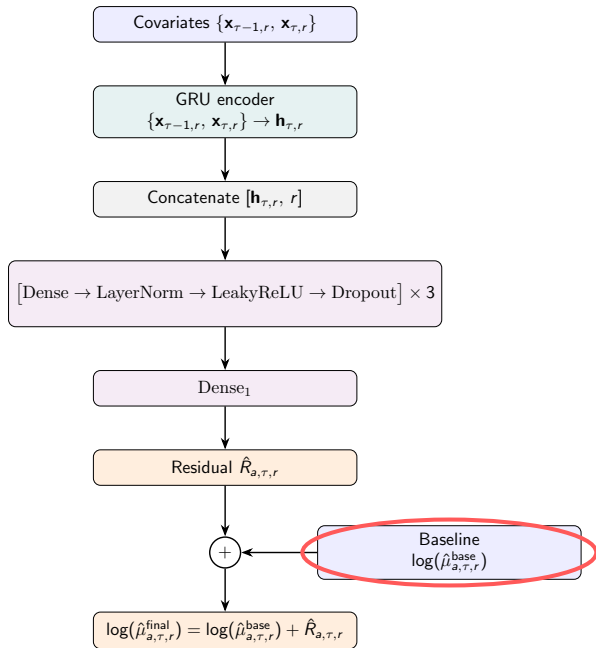
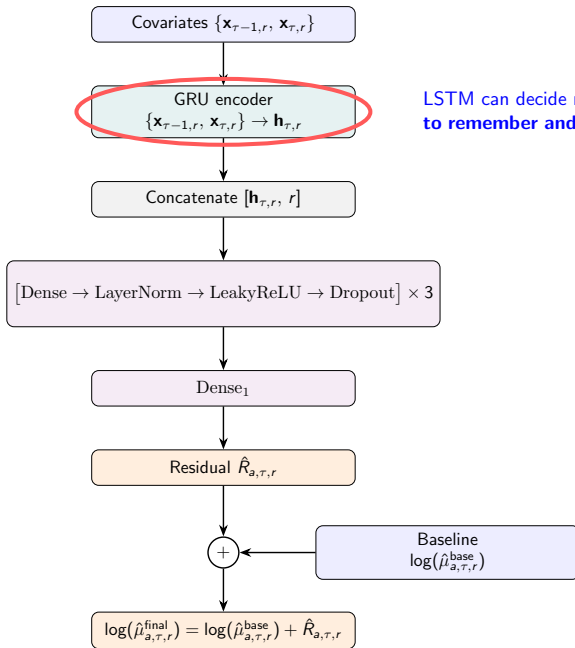


Figure: The GNN-LSTM has narrowed prediction intervals compared to the CNN-LSTM



Zheng et al. (2025) Baseline:

$$\begin{aligned}
 \ln(\mu_{t,w}^{(r)}) &= \theta_0^{(r)} + \theta_1^{(r)}t \\
 &+ \theta_2^{(r)} \sin\left(\frac{2\pi w}{52.18}\right) + \theta_3^{(r)} \cos\left(\frac{2\pi w}{52.18}\right) \\
 &+ \theta_4^{(r)} \sin\left(\frac{2\pi w}{26.09}\right) + \theta_5^{(r)} \cos\left(\frac{2\pi w}{26.09}\right)
 \end{aligned}$$



LSTM can decide more precisely what **patterns to remember and forget** at each time step

