

Dynamic Bayesian Mortality Surface Modelling with Local Age–Time Dependence

From Lee–Carter benchmarks to locally dependent state-space formulations

Mariane Branco Alves

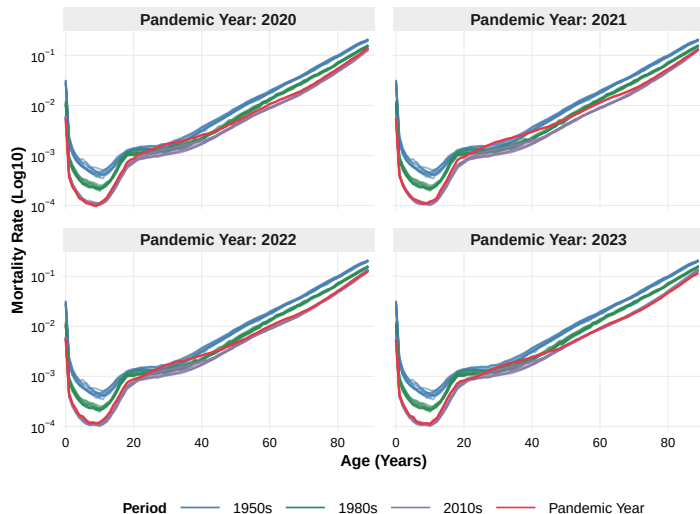
Longevity 2026

Joint work with Luís Philipe Craveiro Mendes and Rafael Santos Erbisti

LabMA – Universidade Federal do Rio de Janeiro — Universidade Federal Fluminense



Mortality changes gradually — until it doesn't



The forecasting problem

Mortality usually evolves smoothly across age and time, but exceptional periods can shift the entire surface.

Our question

Can we model mortality **locally across age and time**, update it efficiently, and still adapt when the regime changes?

We move from **low-dimensional age–period factor structures** to dynamic Bayesian formulations for the **whole age–time mortality surface**.

Local dependence

Neighbouring ages are allowed to share information directly.

Sequential learning

The mortality surface evolves as new information arrives.

Operational adaptation

Interventions can represent shocks and subsequent recovery.

Key message: The goal is not complexity for its own sake, but a mortality model that can be updated, interpreted and stress-tested as regimes evolve.

Bayesian Lee–Carter

$$y_{x,t} = \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t}, \quad \kappa_t = \kappa_{t-1} + \theta + \omega_t.$$

- ▶ Widely recognised mortality forecasting benchmark.
- ▶ Parsimonious age–period structure.
- ▶ Clear interpretation of age sensitivities β_x and period index κ_t .
- ▶ Coherent Bayesian uncertainty quantification.

Lee and Carter (1992); Pedroza (2006).

What remains restrictive?

$$\mathbb{E}(y_{x,t}) = \alpha_x + \beta_x \kappa_t.$$

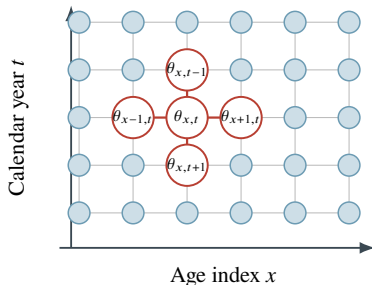
- ▶ Temporal change is driven by a common period index.
- ▶ Age-specific responses are governed by fixed sensitivities.
- ▶ Local dependence between neighbouring ages is not explicit.
- ▶ Local age–time deviations may be smoothed over.

Key message: Lee–Carter is our benchmark: the goal is to compare Bayesian surface formulations that balance local dependence, computational feasibility and observation-scale realism.

Formulation	What it adds	Main trade-off
GMRF-DLM / Lavine (1998)	Direct local dependence across age and time.	Rich local structure, but computationally demanding.
Matrix-variate DLM	Analytical sequential updating with age correlation.	Fast, but the correlation structure is more restrictive.
Poisson kDGLM + CAR	Deaths and exposures are modelled directly, with local age smoothing.	Natural count-data scale, but non-Gaussian updating is more involved.

Key message: Each formulation resolves a different part of the same problem: local dependence, computational feasibility and realistic observation modelling.

Bidimensional Markov structure for an age-time mortality



Each node represents a latent log-mortality rate
 $\theta_{x,t} = \log m_{x,t}$.

The Markov property makes $\theta_{x,t}$ depend only on adjacent ages and neighbouring years.

The precision matrix is sparse, supporting local smoothing and dynamic learning.

Mortality surface

We represent log-mortality as a latent age-time surface,

$$\theta_{x,t} = \log m_{x,t},$$

where t indexes calendar year and x indexes age.

Key message

The Markovian structure makes mortality at each age and year borrow information locally from neighbouring ages and adjacent years.

Starting point

Lavine's formulation (1998) provides a formal bridge between conditionally Gaussian Markov random fields and a dynamic linear model representation.

Why this matters

A Markov random field specifies local dependence through neighbourhood relations, while a dynamic linear model provides a natural framework for Bayesian learning and forecasting over time.

Key message

This connection allows us to translate local Markovian dependence into a state-space formulation.

Mortality-specific adaptation

We adapt Lavine's generic Markovian–dynamic construction to mortality surfaces indexed jointly by calendar year and age:

$$\theta_{x,t} = \log m_{x,t}.$$

Local age–time dependence

The latent log-mortality at year t and age x borrows information from neighbouring ages and adjacent years:

$$(x-1, t), \quad (x+1, t), \quad (x, t-1), \quad (x, t+1).$$

Key message

The mortality curve is not updated age by age in isolation; the whole age–time surface is regularised through local dependence.

Adaptive learning

Evolution uncertainty may vary over time. At selected periods, uncertainty can be inflated to help the learning system absorb structural mortality shocks.

Intervention modelling

This provides a natural mechanism for accommodating abrupt changes in mortality regimes, such as pandemic and post-pandemic shifts.

Level and local slope

The state may include both the mortality level and its local temporal slope, allowing the model to track not only the level of mortality but also its rate of change.

Mortality improvement

The slope component provides a natural dynamic summary of the potential for mortality increase or decline at each time point.

Idea For each year t , we model the whole mortality curve jointly:

$$\mathbf{Y}_t = \{Y_{x,t} : x \in \mathcal{X}\}, \quad \Theta_t = \{\boldsymbol{\theta}_{x,t} : x \in \mathcal{X}\},$$

where Θ_t is a $p \times J$ state matrix, with one column per age.

Matrix-variate DLM

$$\mathbf{Y}_t \mid \Theta_t, \tau^2 \sim N_J\left(\Theta_t^\top F_t, \tau^2 R_\rho\right),$$

$$\Theta_t \mid \Theta_{t-1}, \tau^2 \sim MN_{p \times J}\left(G_t \Theta_{t-1}, W_t, \tau^2 R_\rho\right).$$

Age dependence

$$R_\rho(x, x') = \exp\left\{-\frac{|x - x'|}{\rho}\right\}.$$

$$\text{Cov}(\boldsymbol{\omega}_{x,t}, \boldsymbol{\omega}_{x',t}) = \tau^2 R_\rho(x, x') W_t.$$

Key message: The model preserves analytical sequential updating while allowing neighbouring ages to evolve dependently.

$$\begin{aligned}Y_{x,t} \mid \eta_{x,t} &\sim \text{Poisson}(E_{x,t}\eta_{x,t}), \\ \log \eta_{x,t} &= \theta_{1,t} + S_{x,t} + \varepsilon_{x,t}, \\ \theta_{1,t} &= \theta_{1,t-1} + \theta_{2,t-1} + \omega_{1,t}, \\ \theta_{2,t} &= \theta_{2,t-1} + \omega_{2,t}, \\ S_{x,t} &= S_{x,t-1} + \omega_{3,t}.\end{aligned}$$

Local CAR structure at the initial surface

$$S_{x,1} \mid S_{-x,1} \sim N\left(\rho \sum_{x' \in \partial_x} \frac{W_{x,x'}}{W_{x+}} S_{x',1}, \frac{\tau^2}{W_{x+}}\right), \quad \partial_x = \{x-1, x+1\}.$$

Key message: The formulation preserves the natural count-data scale while introducing local age dependence and dynamic evolution.

Why Poisson?

Deaths and exposures enter directly, avoiding a Gaussian approximation to empirical log-mortality rates.

Why CAR?

Neighbouring ages share information, stabilising irregular age-specific rates without forcing global factors.

The “but”

With a non-Gaussian observation model, standard Gaussian state-space updating is no longer directly available in closed form.

The kDGLM route

The kDGLM framework, developed in joint work with collaborators, propagates posterior information sequentially using analytic approximations and conjugate projections.

- **Online learning:** update as new deaths and exposures arrive.
- **No full refit:** avoid simulation-based inference over the entire historical series at every time point.



kDGLM

R package on CRAN

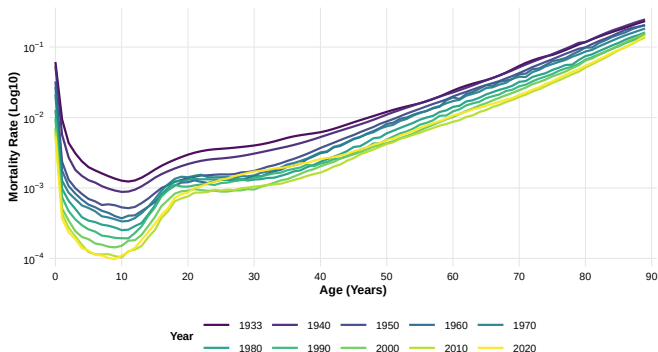
Developed jointly with collaborators at

LabMA/UFRJ

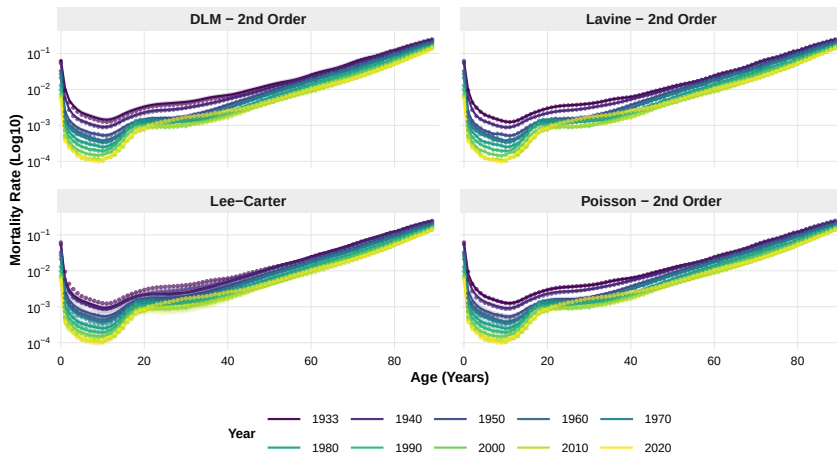
- ▶ Human Mortality Database.
- ▶ Annual deaths and exposures.
- ▶ Single ages 0–89.
- ▶ 1933–2024.

Empirical feature

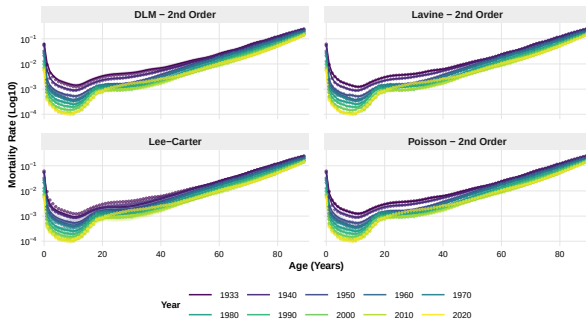
The curve evolves smoothly over age, but exceptional periods produce visible departures from historical patterns.



Fitted mortality surfaces across selected years



Key insight: All models recover the broad mortality curve; differences become important in local smoothing, uncertainty and dynamic behaviour under shocks.

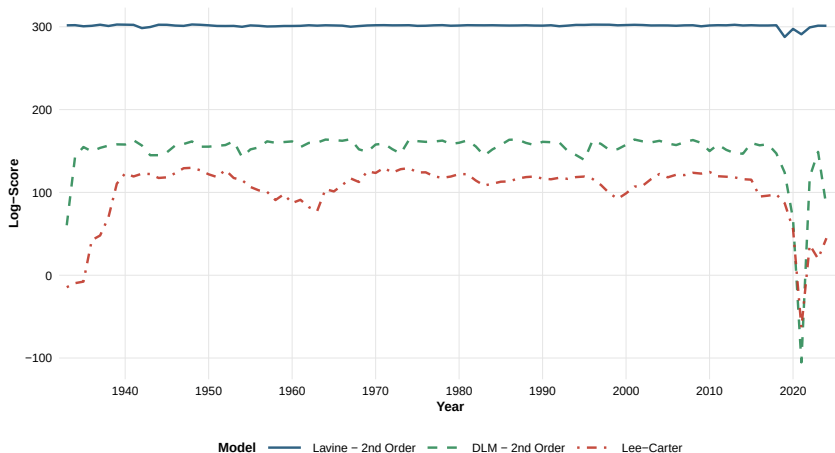


Model	IS	MAPE
Lavine – 2nd order	10.63	0.22
DLM – 2nd order	49.39	2.75
Poisson – 2nd order	46.27	2.47
Lee-Carter	39.89	7.52

Reading the evidence

The dynamic alternatives improve point accuracy relative to Lee-Carter, but no single metric captures the full trade-off.

Key message: Accuracy matters, but model choice must also account for uncertainty, computational cost and robustness to shocks.

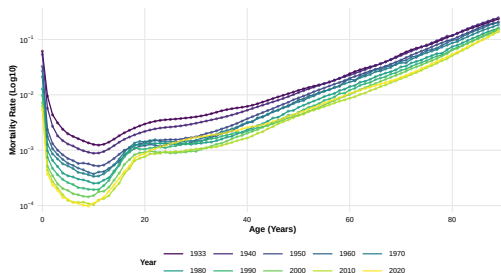


Key insight: All comparable proposed models were better than Lee-Carter according to log-score.

Better fit — but at what computational cost?



Model	Execution time
DLM – 2nd order	0.08 s
Poisson – 2nd order	7.48 s
Lee–Carter	18.81 min
Lavine – 2nd order	60.17 h



Operational reading

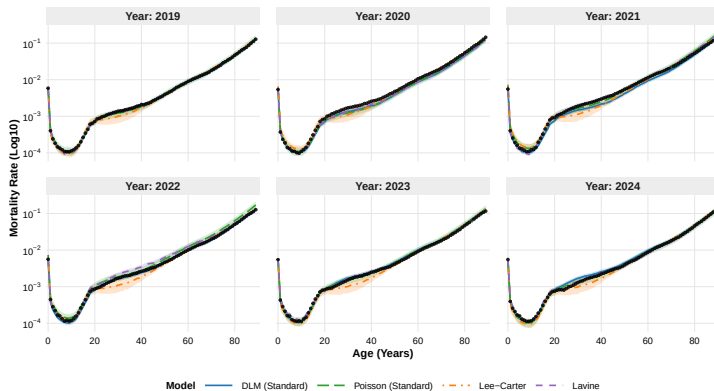
The matrix-variate DLM gives the strongest speed–dependence compromise when repeated updating and forecasting are required.

Caution

Excellent fit does not guarantee better forecasts. Very tight fitting may signal overfitting and must be judged together with predictive behaviour.

Key message: The richest local structure is not automatically the most useful operational model.

When smooth dynamics are not enough



Key message: Baseline dynamics capture gradual evolution, but abrupt mortality shocks require an explicit mechanism for regime change.

At an intervention time, inject a mean shift and additional uncertainty:

$$\eta_t \sim N(h_t, H_t), \quad \eta_t \perp \omega_t.$$

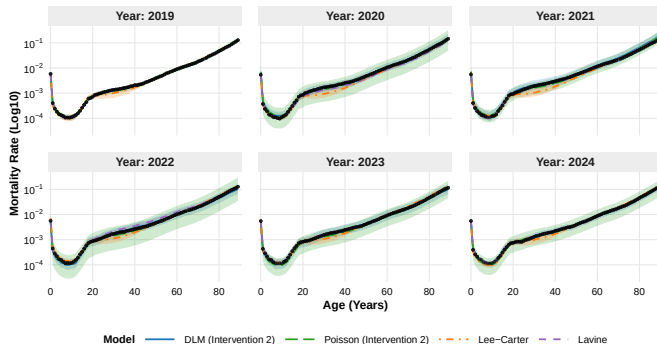
The modified prior is

$$\theta_t^* \mid D_{t-1}, I_t^* \sim N(a_t + h_t, R_t + H_t).$$

Interpretation

- ▶ h_t : anticipated direction and size of the shock.
- ▶ H_t : extra uncertainty about the new regime.
- ▶ After intervention, sequential learning continues within the same DLM structure.

Key message: Intervention is a formal forecasting device for shock and recovery, not an ad hoc correction of the fitted curve.



MAPE, 2019–2024

DLM (interv.)	7.06%
Poisson (interv.)	7.56%
Lavine	9.01%
Lee-Carter	13.52%

Risk-management implication

Shock and recovery mechanisms support stress testing and short-term updates of longevity assumptions.

Key message: Intervened dynamic models track the pandemic path and substantially reduce pandemic-period error relative to Lee-Carter.

For each age x , we display the total percentage reduction in fitted mortality between the first and last year of the period: $100 \times I_{x,t_0:t_1} = 100 \times \left(1 - \frac{m_{x,t_1}}{m_{x,t_0}}\right)$, $\ell_{x,t} = \log m_{x,t}$.

$$I_{x,t_0:t_1} = 1 - \exp\{\ell_{x,t_1} - \ell_{x,t_0}\}.$$

Lee–Carter

$$\ell_{x,t} = a_x + b_x k_t$$
$$I_{x,t_0:t_1}^{LC} = 1 - \exp\{b_x(k_{t_1} - k_{t_0})\}.$$

A parsimonious factor structure: age-specific changes are driven by a common period factor k_t , modulated by b_x .

Poisson–CAR smoothed surface

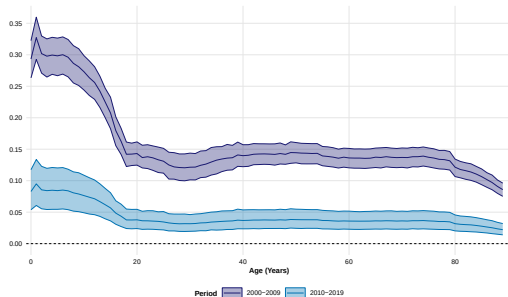
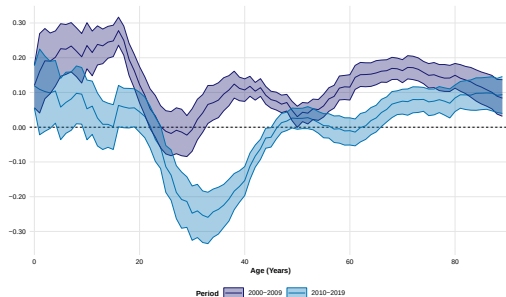
$$\ell_{x,t}^* = \log m_{x,t}^* = \theta_{1,t} + S_{x,t}$$

$$I_{x,t_0:t_1}^{PC} = 1 - \exp\{(\theta_{1,t_1} - \theta_{1,t_0}) + (S_{x,t_1} - S_{x,t_0})\}.$$

Improvements are computed from the smoothed age–time mortality surface, excluding unexplained heterogeneity.

Reading

This is a first-to-last relative change, not an annualised rate. Positive values indicate mortality reduction; negative values indicate mortality deterioration.



- ▶ Lee-Carter gives smooth, mostly positive total improvements in both periods.
- ▶ Poisson-CAR reveals stronger local age-specific variation, based on the smoothed surface.
- ▶ In 2010-2019, negative values at young-adult ages indicate fitted mortality deterioration.

Young-adult deterioration

For U.S. all-cause mortality, the negative improvements estimated by the Poisson–CAR model at young-adult ages in 2010–2019 are compatible with documented increases in working-age mortality related to opioids, other drugs, alcohol and substance-related causes.

Caution

This is a substantive hypothesis, not a causal conclusion: the model uses all-cause mortality and does not include cause-of-death information.

Take-away

Lee–Carter gives a stable global summary of mortality decline, while Poisson–CAR can reveal local age–time deviations that deserve substantive validation.

Sources: CDC/NCHS Data Brief No. 394, 2020; CDC Health, United States, Drug Overdose Deaths; National Academies, *High and Rising Mortality Rates Among Working-Age Adults*, 2021.

What we gain

- ▶ Mortality is modelled as an evolving age–time surface.
- ▶ Local dependence helps reveal age-specific dynamics.
- ▶ Sequential Bayesian updating keeps the model operational.

What the data suggest

- ▶ Dynamic formulations improve predictive accuracy relative to Lee–Carter.
- ▶ Intervention mechanisms improve pandemic-period forecasts.
- ▶ Poisson–CAR reveals local improvement

What remains a trade-off

- ▶ Richer local structure increases computational and modelling complexity.
- ▶ Parsimonious models are stable and interpretable, but may miss local deviations.
- ▶ Local deterioration patterns require substantive validation.

Take-home message

Dynamic Bayesian surface models do more than improve forecast accuracy: they provide a coherent framework for updating, interpreting and stress-testing mortality curves under gradual evolution and abrupt regime change.

Thank you

mariane@im.ufrj.br

-  Alves, M. B., Migon, H. S., Santos Jr., S. V. and Marotta, R. (2026). *An Efficient Sequential Approach for k -Parametric Dynamic Generalised Linear Models*. arXiv:2201.05387v5.
-  Banerjee, S., Carlin, B. P. and Gelfand, A. E. (2014). *Hierarchical Modeling and Analysis for Spatial Data*. CRC Press.
-  Lavine, M. (1998). Another look at conditionally Gaussian Markov random fields. In *Bayesian Statistics 6*, Oxford University Press, 371–387.
-  Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting U.S. mortality. *JASA*, 87, 659–671.
-  Pedroza, C. (2006). A Bayesian forecasting model: predicting U.S. male mortality. *Biostatistics*, 7, 530–550.
-  Santos Jr., S. V., Alves, M. B. and Migon, H. S. (2024). *kDGLM: a R package for Bayesian analysis of Generalized Dynamic Linear Models*. arXiv:2403.13069.
-  West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. Springer.