São Paulo 2025

SPATIO-TEMPORAL EXTENSIONS OF THE LEE–CARTER AND LI–LEE MODELS

Péter Vékás, Ph.D.
Associate Professor
Actuarial Degree Program Director
Corvinus University of Budapest

Introduction

- Mortality forecasting is central to actuarial work, supporting more accurate prediction of pension and life insurance liabilities and better management of longevity risk.
- A substantial body of research has followed the seminal paper by Lee and Carter (1992), including many multi-population extensions since Li and Lee (2005).
- Most models forecast time-dependent parameters using standard time series methods such as ARIMA.
- However, when applied to multiple countries, this approach becomes limiting as it ignores cross-sectional and spatial dependencies often present in mortality patterns across populations.

Introduction

- We draw on methods from panel and spatial econometrics to capture cross-sectional and spatial dependence and improve forecast accuracy.
- We apply these methods to country-specific mortality indices from the Lee–Carter and Li–Lee models, using data from 22 European countries.
- For each country and model, we identify the method that performs best.
- The pensions and life insurance industry may benefit from more accurate mortality forecasts.

The Lee–Carter and Li–Lee models

- The Lee–Carter (1992) model decomposes the logarithm of central death rate m_{xt} at age x and calendar year t as follows:

$$\ln m_{xt} = a_x + b_x k_t + \varepsilon_{xt}.$$

- The Li–Lee (2005) multi-population model assumes a common trend represented by the common mortality index K_t and country-specific fluctuations in country i represented by mean-reverting mortality indices k_{it}:

$$\ln m_{ixt} = a_{ix} + B_x K_t + b_{ix} k_{it} + \varepsilon_{ixt}.$$

- This setup ensures the long-term coherence (non-divergent behavior) of forecasts.
- Parameter constraints are used in both models to ensure identifiability.

Spatial autocorrelation

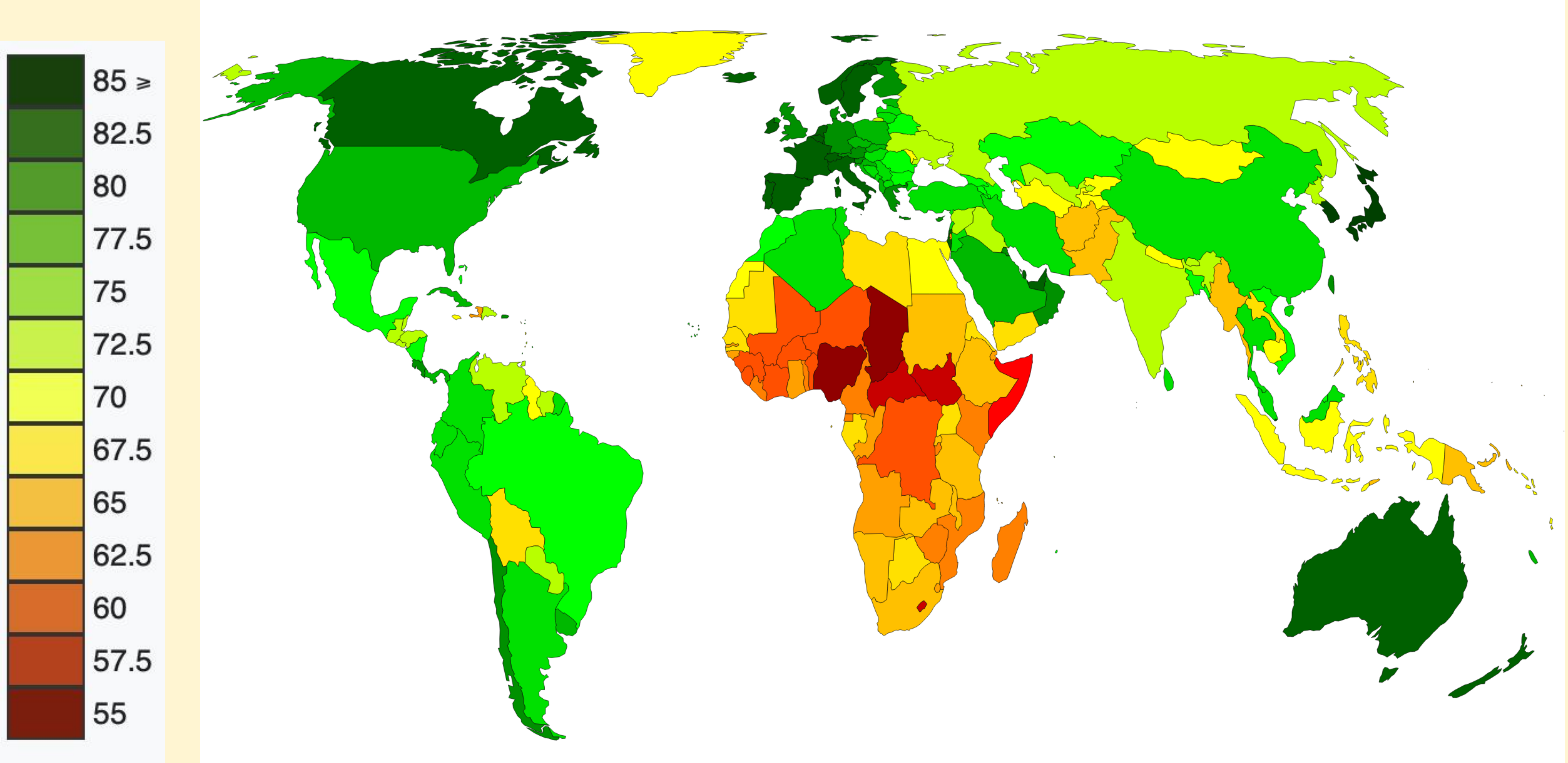
- Traits or events in one location can influence nearby areas more strongly than distant ones.
- Several actuarial applications presented in Brechmann and Czado (2014).
- Heuristically, spatial autocorrelation is the correlation between a variable and a weighted average of its values at "nearby" locations.
- The weighted average of nearby values is called a spatial lag:

$$\tilde{y}_i = \sum_{j=1}^{n} w_{ij} y_j$$

- Row-normalized spatial weights matrix:

$$\mathbf{W} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix} \quad w_{ii} = 0, \quad \sum_{j=1}^{n} w_{ij} = 1 \quad (i = 1, 2, \ldots, n)$$

Life expectancy at birth (UN, 2023)



Spatial and spatio-temporal autocorrelation

- We use two weighting schemes:
  (1) unweighted averages of first-order neighbors, and
  (2) inverse distances of capital cities.
- Moran's I measure of spatial autocorrelation is the slope of the regression line between the original values and their spatially lagged counterparts.
- In spatio-temporal autocorrelation, the value at a given time and location may be influenced by values at nearby times and locations.
- Two common specifications:
  - Contemporaneous: only current values at nearby locations exert influence.
  - Lagged: past values at nearby locations exert influence.

Spatio-temporal autocorrelation

$$\mathbf{W}_{space} = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{pmatrix} \qquad \mathbf{W}_{time} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

- Contemporaneous and lagged spatio-temporal weights matrices ($\otimes$ is the Kronecker-product) :

$$\mathbf{W}_{cont} = \mathbf{I}_{time} \otimes \mathbf{W}_{space} + \mathbf{W}_{time} \otimes \mathbf{I}_{space}$$

$$\mathbf{W}_{lagged} = \mathbf{W}_{time} \otimes (\mathbf{W}_{space} + \mathbf{I}_{space})$$

JOCO São Paulo 2025

Forecasting methods

1. Most commonly, the mortality index k_{it} is assumed to follow a Random Walk with Drift (RWD), assuming spatio-temporal independence of yearly increments of the mortality index:

$$\Delta k_{it} = \alpha_i + \xi_{it},$$
$$\xi_{it} \sim_{i.i.d.} \mathcal{N}(0, \sigma).$$

2. ARIMA models are also commonly used, allowing time dependence but no cross-sectional dependence:

$$\Delta^{(d_i)} k_{it} = \alpha_i + \sum_{\ell=1}^{p_i} \varphi_{i\ell} \Delta^{(d_i)} k_{i,t-\ell} + \sum_{\ell=1}^{q_i} \theta_{i\ell} \xi_{i,t-\ell} + \xi_{it},$$
$$\xi_{it} \sim_{i.i.d.} \mathcal{N}(0, \sigma_i).$$

Forecasting methods

3. Dynamic Panel Linear Model (DPLM, Blundell and Bond, 1998, estimated by Generalized Method of Moments), equivalent to AR models with equal parameters across countries, with no normality assumption:

$$\Delta^{(d)}\,\mathbf{k}_t = \alpha + \sum_{\ell=1}^{p} \rho_\ell \Delta^{(d)}\mathbf{k}_{t-\ell} + \xi_t.$$

4. Vector Autoregression (VAR, estimated using Elastic Net Regression following Guibert, Lopez and Piette, 2019), allowing cross-sectional dependence:

$$\Delta^{(d)}\mathbf{k}_t = \boldsymbol{\alpha} + \sum_{\ell=1}^{p} \mathbf{A}_\ell \Delta^{(d)}\mathbf{k}_{t-\ell} + \xi_t.$$

Forecasting methods

5. Spatio-Temporal ARIMA (STARIMA, Pfeifer and Deutsch, 1980, where $W^{(m)}$ is the m-th order spatial lag operator), estimated by Kalman filter, allowing spatio-temporal dependence:

$$z_{it} = \alpha + \sum_{\ell=1}^{p} \sum_{\mathrm{m}=0}^{\lambda_\ell} \phi_{\ell m} W^{(m)} z_{i,t-\ell} + $$

$$+ \sum_{\ell=1}^{p} \sum_{m=0}^{\lambda_\ell} \theta_{\ell m} W^{(m)} \xi_{i,t-\ell} + \xi_{it},$$

$$z_{it} = \frac{\Delta^{(d_i)} k_{it} - \mathrm{mean}\left(\Delta^{(d_i)} k_{it}\right)}{\mathrm{sd}\left(\Delta^{(d_i)} k_{it}\right)}.$$

Forecasting methods

6. Spatial Dynamic Panel (SDP, Lee and Yu, 2010, estimated using Quasi Maximum Likelihood, where $\mathbf{W}$ is a spatial weights matrix), allowing spatio-temporal dependence:

$$\Delta^{(d)}\mathbf{k}_t = \boldsymbol{\alpha} + \rho_{space}\, \mathbf{W}\Delta^{(d)}\mathbf{k}_t + \rho_{\text{time}}\Delta^{(d)}\mathbf{k}_{t-1} + \rho_{space-time}\, \mathbf{W}\Delta^{(d)}\mathbf{k}_{t-1} + \xi_t.$$

Forecasting methods

7. Eigenvector Spatio-Temporal Filter (ESTF, Griffith, 2010, where $\boldsymbol{S}_{space-time}$ and $\boldsymbol{S}_{space}$ are matrices of eigenvectors of $\boldsymbol{W}_{space-time}$ and $\boldsymbol{W}_{space}$, respectively, and $\mathbf{D}$ is the matrix of country dummy variables), allowing spatio-temporal dependence:

$$\Delta^{(d)}\,\mathbf{k}_t = \boldsymbol{\alpha} + \mathbf{S}_{space-time}\,\boldsymbol{\beta} + \mathbf{S}_{space}\,\boldsymbol{\gamma} + \mathbf{D}\boldsymbol{\phi} + \boldsymbol{\xi}_t\,.$$

- Estimated using Ordinary Least Squares with Stepwise or LASSO selection to reduce the number of basis vectors.

Summary of forecasting methods

Method	Dependence	Estimation	Hyperparameters	Criterion
RWD	temporal	MLE	–	BIC
ARIMA	temporal	MLE	$p_i\ ,q_i)\ i\ =\ 1,2,\ldots,N($	BIC
DPLM	temporal	GMM	p	significance
VAR	temporal and cross-sectional	ENR	p	CV-MSE
STARIMA	spatio-temporal	KF	$p\ ,\lambda_\ell\ (\ell\ =\ 1,2,\ldots,p)\ ,$ $q\ ,\mu_\ell\ (\ell = 1,2,\ldots,q)$	BIC
SDPLM	spatio-temporal	QML	$\boldsymbol{W}$, time lag, space-time lag, Lee-Yu transf. ∈ {0, 1}	BIC
ESTF	spatio-temporal	stepwise or LASSO	$\boldsymbol{W}$, spec. ∈ {cont., lagged}, selection ∈ {stepwise, LASSO}	BIC and Deviance Ratio

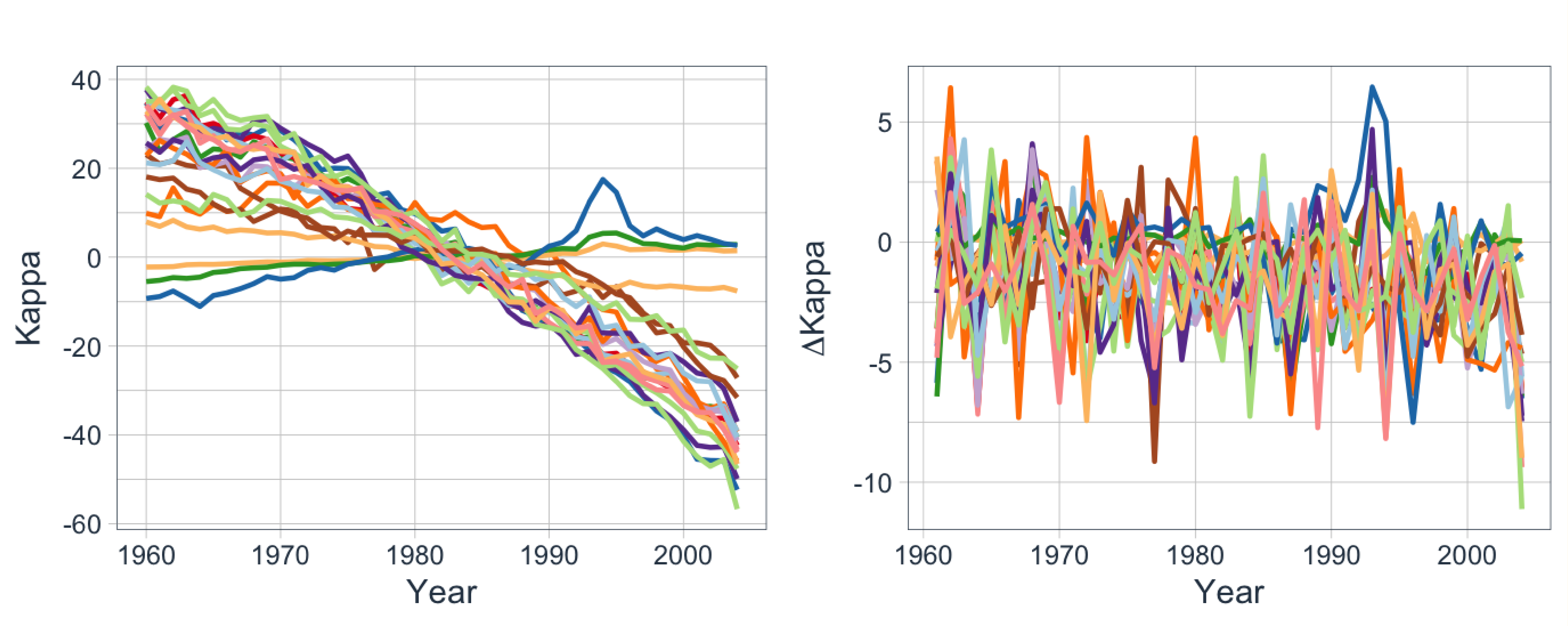
Data

- Unisex death counts and exposures by country, age, and calendar year from the Human Mortality Database (HMD) for all N = 22 European countries having data for all years between 1960 and 2019.
- We removed data for ages above 99 years due to low exposures.
- We divided the data into a training (1960-2004) and a test period (2005-2019).
- We used data only up to 2019 to avoid testing on the years of COVID-19, which would have led to a bias towards models that tend to overestimate mortality.

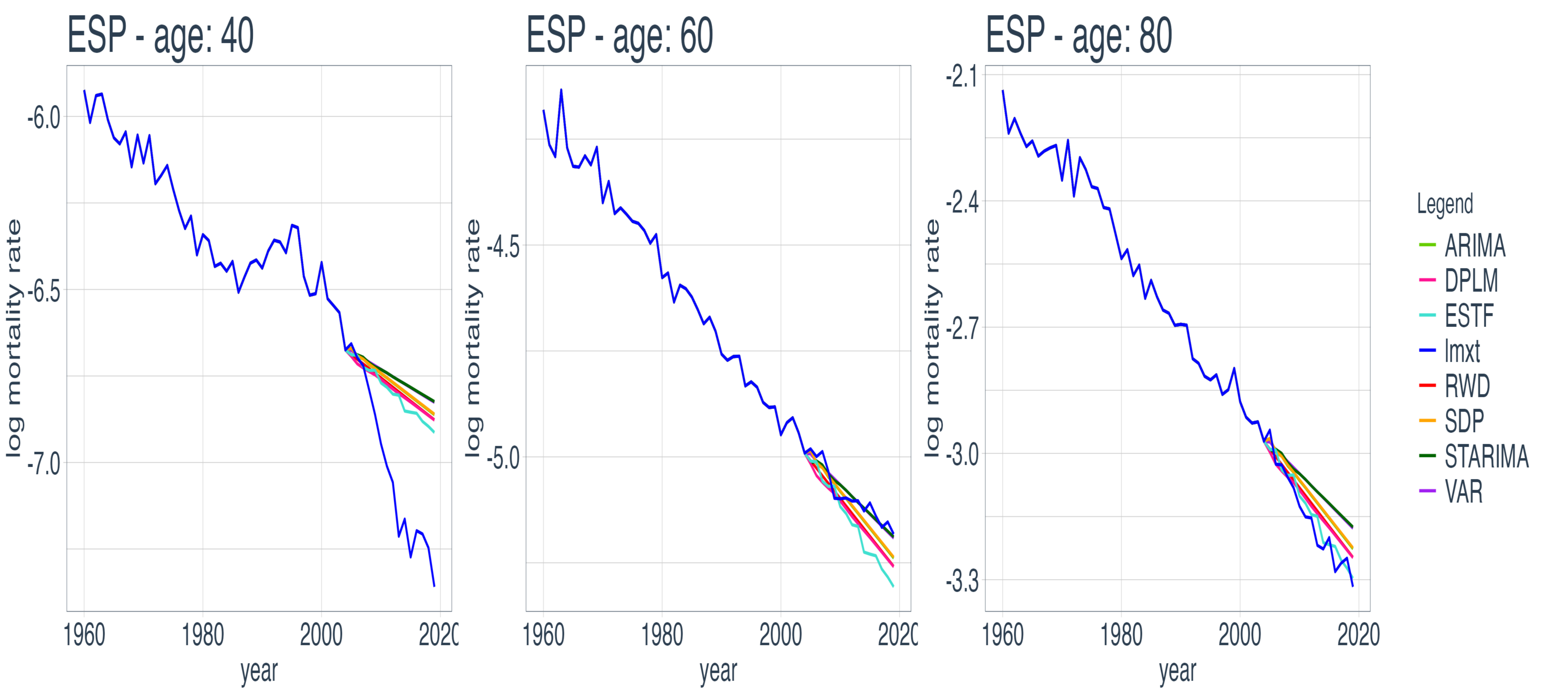
Modeling steps

- We used R (R Core Development Team) for all calculations.
- We estimated the LC and LL models for all countries on the training data (1960 through 2004) to extract the mortality index series k_{it} using the Poisson assumption of Brouhns et al. (2006).
- We differenced k_{it} once for stationarity, as indicated by the second-generation panel unit root test of Costantini and Lupi (2013).
- We estimated the parameters of all seven techniques on the differenced series on the training set and selected their best hyperparameters by optimizing the associated criteria.
- We forecasted the series into the test set, computed the forecasted mortality rates, and computed the Mean Squared Error (MSE) of the logarithmic rates.
- We also computed the unweighted average of the forecasts (ensemble).
- We considered only mean-reverting specifications for LL for coherence.

Results – Lee–Carter mortality index series and their first-order differences



Example: Lee–Carter forecasts for Spain under the 7 methods



JOCO São Paulo 2025

Robust model selection

- Model performance can depend heavily on the train-test boundary, so we designed a robust model selection procedure across multiple splits.
- We generated forecasts using 3 different train-test splits per country (with the last year in the training period being 2003, 2004, and 2005).
- For each split, we computed:

$$\text{Under}performance(Model) = MSE(Model) - MSE(Best\ Model)$$

to measure how much a model underperforms the best one.

- We evaluated each model's performance by computing across splits:
  (A) the average of underperformance scores,
  (B) the maximum of underperformance scores (more conservative).

Number of wins by model and forecasting method (A: average, B: maximum underperformance)

Method	LC (A)	LC (B)	LL (A)	LL (B)
RWD	1	1	4	4
ARIMA	3	3	1	1
VAR	2	2	4	3
DPLM*	2	2	**6**	**5**
STARIMA*	2	2	**6**	**6**
SDPLM*	2	3	0	0
ESTF*	**7**	**7**	0	1
AVERAGE*	3	2	1	2

Spatio-temporal

*Not yet used in the actuarial literature.

Spatio-temporal clusters

- Local Indicators of Spatial Association (LISA) reveal significantly elevated spatial autocorrelation in the Baltic region (Estonia, Latvia, and Lithuania) and in Central Western Europe (France, Germany, and Switzerland).
- Additionally, the dominant eigenvectors from the ESTF method separate the British Isles (UK and Ireland) and Scandinavia (Finland, Norway, and Sweden) from the rest of continental Europe.
- These clusters have plausible geographical and historical explanations. For instance, the Baltic countries form a distinctive group due to their shared Soviet legacy and the severe demographic crisis they experienced during the 1980s.

Takeaways

- We bring spatial and panel econometric tools into mortality forecasting — a natural but underused extension.
- The methods we propose outperform standard time series models across most countries in both LC and LL frameworks.
- We reveal geographically interpretable longevity clusters, showing that mortality is not just temporal but spatially connected.
- Forecasts are made probabilistic and actuarially usable via Poisson-based parametric bootstrapping.
- These methods support better-informed decisions in pensions and life insurance, with models that reflect the real structure of mortality data.

São Paulo 2025

Thank you! Obrigado!

Questions?

IAA SECTIONS

IBA INSTITUTO BRASILEIRO DE ATUÁRIA

References

- Álvaro-Meca, A. et al. (2011). The Convergence of European Mortality in Both Sexes in the Near Future: A Spatio-Temporal Approach. Population Review, 50(2).
- Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. The Review of Economic Studies, 58(2):277–297.
- Astuti, A. M., Setiawan, Zain, I., and Purnomo, J. D. T. (2020). A Review of Panel Data on Spatial Econometrics Models. Journal of Physics: Conference Series, 1490(1):012032.
- Baiocchi, G. and Distaso, W. (2003). GRETL: Econometric Software for the GNU Generation. Journal of Applied Econometrics, 18(1):105–110.

References

- Basellini, U., Camarda, C. G., and Booth, H. (2023). Thirty years on: A review of the Lee–Carter method for forecasting mortality. *International Journal of Forecasting*, 39(3):1033–1049.
- Bennett, J. E., Li, G., Foreman, K., Best, N., Kontis, V., Pearson, C., Hambly, P., and Ezzati, M. (2015). The future of life expectancy and life expectancy inequalities in England and Wales: Bayesian spatiotemporal forecasting. *Lancet*, 386(9989):163–170. Epub 2015 Apr 29.
- Bivand, R., Millo, G., and Piras, G. (2021). A Review of Software for Spatial Econometrics in R. *Mathematics*, 9(11).
- Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

References

- Brechmann, E. and Czado, C. (2014). *Spatial Modeling*. In: Frees, E.W., Derrig, R.A. and Meyers G. (ed.). *Predictive Modeling Applications in Actuarial Science*. Cambridge University Press.
- Carracedo, P., Debón, A., Iftimi, A., et al. (2018). Detecting spatio-temporal mortality clusters of European countries by sex and age. *International Journal for Equity in Health*, 17(1):38.
- Cheysson, F. (2016). *starma: Modelling Space Time AutoRegressive Moving Average (STARMA) Processes*. *R package version 1.3*, URL: https://CRAN.R-project.org/package=starma.
- Costantini, M. and Lupi, C. (2013). A Simple Panel-CADF Test for Unit Roots. *Oxford Bulletin of Economics and Statistics*, 75(2):276–296.
- Elhorst, J. (2014). *Spatial econometrics: from cross-sectional data to spatial panels*. Springer.

São Paulo 2025

References

- Gibbs, Z., Groendyke, C., Hartman, B., and Richardson, R. (2020). Modeling County-Level Spatio-Temporal Mortality Rates Using Dynamic Linear Models. *Risks*, 8(4).
- Greco, F. and Scalone, F. (2013). *A Space-Time Extension of the Lee–Carter Model in a Hierarchical Bayesian Framework: Modelling and Forecasting Provincial Mortality in Italy* (working paper). https://api.semanticscholar.org/CorpusID:55707767
- Griffith, D. (2012). Space, time, and space-time eigenvector filter specifications that account for autocorrelation. *Estadistica Española (Spanish Statistical Magazine)*, 54:4–27.
- Griffith, D. A. (2010). Modeling spatio-temporal relationships: retrospect and prospect. *Journal of Geographical Systems*, 12(2):111–123. 15

References

- Guibert, Q., Loisel, S., Lopez, O., and Piette, P. (2020). *Bridging the Li-Carter's gap: a locally coherent mortality forecast approach* (working paper).
- Guibert, Q., Lopez, O., and Piette, P. (2019). Forecasting mortality rate improvements with a high- dimensional VAR. *Insurance: Mathematics and Economics*, 88:255–272.
- Hassan, A. (2021). *Spatial data analysis: applications to population health*. PhD thesis, Université de Lille.
- HMD (2024). *Human Mortality Database*. Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). www.mortality.org. Accessed: 21-06-2024.
- Hurlin, C. and Mignon, V. (2007). *Second Generation Panel Unit Root Tests* (working paper). Working Papers halshs-00159842, HAL.

References

- Hyndman, R., Booth, H., Tickle, L., Maindonald, J., Wood, S., and Team, R. C. (2023). *demography: Forecasting Mortality, Fertility, Migration and Population Data*. Comprehensive R Archive Network. R package version 2.0.
- Hyndman, R. J. and Khandakar, Y. (2008). Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software*, 27(3):1–22.
- Islam, M. D., Li, B., Lee, C., and Wang, X. (2021). *Incorporating Spatial Information in Machine Learning: the Moran Eigenvector Spatial Filter Approach* (working paper).
- Kleiber, C. and Lupi, C. (2011). *Panel unit root testing with R*
- (working paper). https://api.semanticscholar.org/CorpusID:197629737
- Lee, R. D. and Carter, L. R. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*, 87(419):659–671.

References

- 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.
- Liu, Z. (2021). *Bayesian Poisson Log-normal Model with Regularized Time Structure and Spatial Framework for Mortality Projection of Multi-population*. PhD dissertation, Clemson University. All Dissertations. 2859.
- Ocaña-Riola, R. and Mayoral-Cortés, J. M. (2010). Spatio-temporal trends of mortality in small areas of Southern Spain. *BMC Public Health*, 10:26.
- Pfeifer, P. E. and Deutsch, S. J. (1980). A Three-Stage Iterative Procedure for Space-Time Modeling. *Technometrics*, 22(1):35–47.
- Pickle, L. W. (2000). Exploring spatio-temporal patterns of mortality using mixed effects models. *Statistics in Medicine*, 19(17-18):2251–2263.

References

- R Core Team (2021). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.
- Robben, J. (2021). *MultiMoMo: Multi-population mortality models. R package version 1.0.0.*
- Salima, B. A. (2018). *Spatial econometrics on panel data* (working paper). https://api.semanticscholar.org/CorpusID:214675854
- Tuljapurkar, S., Li, N., and Boe, C. (2000). A universal pattern of mortality decline in the G7 countries. *Nature*, 405(6788):789–792.
- Ugarte, M. D., Goicoa, T., and Militino, A. F. (2010). Spatio-temporal modeling of mortality risks using penalized splines. *Environmetrics*, 21(3-4):270–289.
- Vazzoler, S. (2021). *sparsevar: Sparse VAR/VECM Models Estimation*. R package version 0.1.0. https://CRAN.R-project.org/package=sparsevar

References

- Yu, J., de Jong, R., and fei Lee, L. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics*, 146(1):118–134.
- Yu, J. and Lee, L. (2010). Estimation of Unit Root Spatial Dynamic Panel Data Models. *Econometric Theory*, 26(5):1332–1362.