SCOR Chair on Mortality Research

Mortality Projections

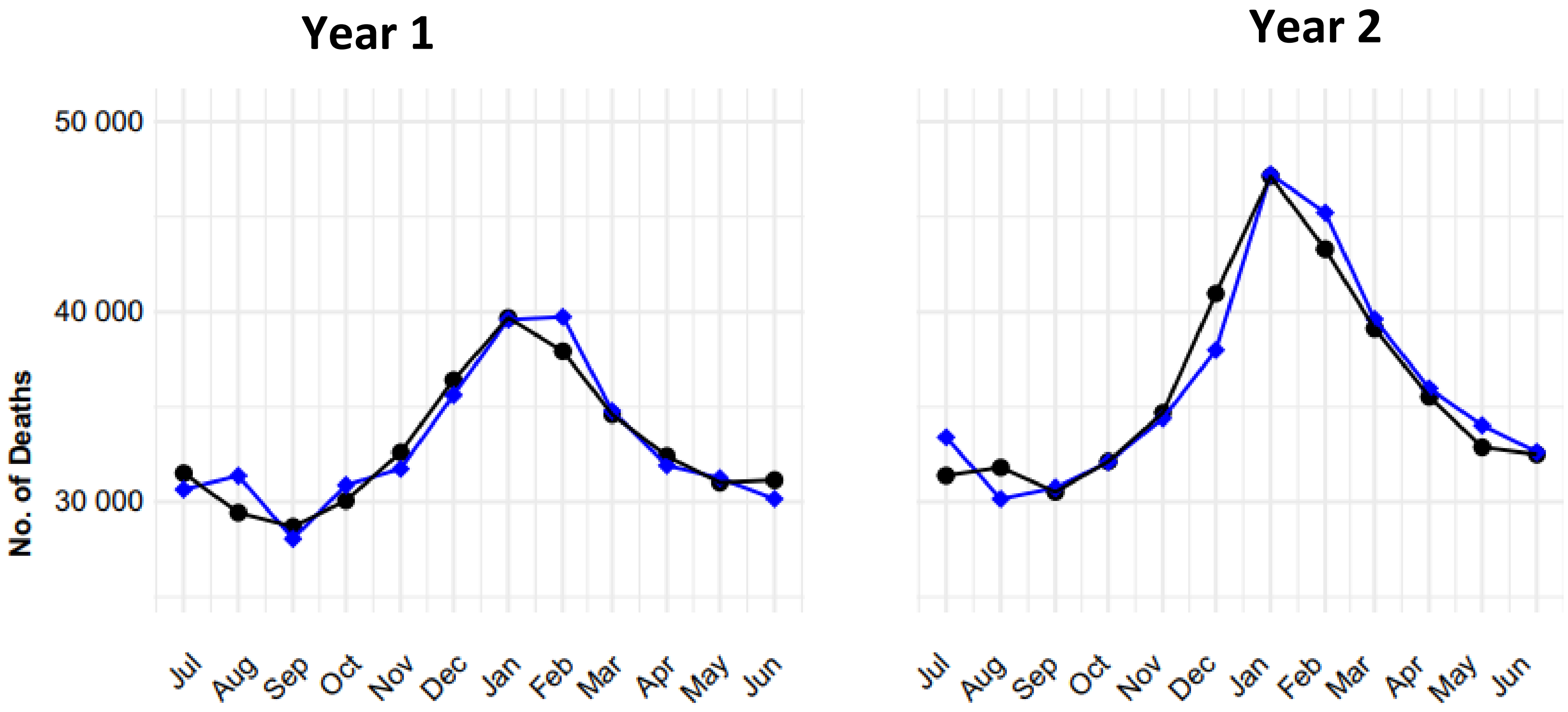
Friday 8th November

Smoothing trends and seasonality in short-term mortality forecasting

Ainhoa-Elena Leger[1], Silvia Rizzi[1], Ugofilippo Basellini[2]

[1] The Interdisciplinary Centre on Population Dynamics, University of Southern Denmark
[2] Max Planck Institute for Demographic Research (MPIDR)

Deaths in an epidemiological year show seasonality. The amplitude of the seasonality might differ across years.



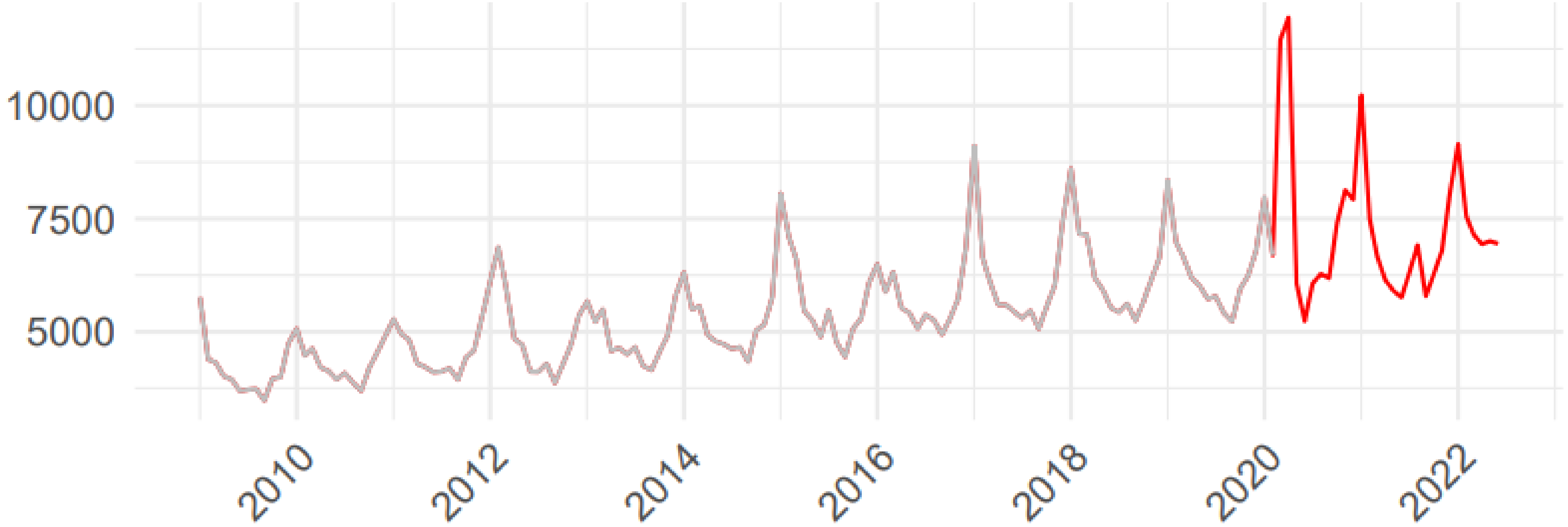
Leger and Rizzi 2024

Short-term mortality forecasts:
→ Realistic forecasts
→ Counterfactual forecasts

Context – Excess mortality in the literature

- **Excess mortality** is the difference between observed and expected deaths.
- It quantifies the mortality burden related to **mortality shocks**:
  - Seasonal influenza (Mølbak et al., 2015)
  - Heatwaves (Toulemon and Barbieri, 2008)
  - Pandemics (Ansart et al., 2009).
- In the framework of the Covid-19 pandemic, excess mortality preferred to the reported Covid-19 deaths.
- Estimating excess mortality implies estimating **the expected number of deaths** as a baseline level of mortality.

Monthly raw death counts, Spain, Men, Age 85+, Years 2009-2002



A “standard curve of expected seasonal mortality”

- **Serfling (1963)** described a linear regression model with a cyclical component.
- More recent approaches consider a Poisson Serfling regression.
- Suppose that the number of deaths $\mathrm{Y_i}$ is a realization of a Poisson distribution:

$$\mathrm{Y_i} \sim \mathrm{Poisson}(\mu_\mathrm{i}),\ \ \mu_\mathrm{i} = \mathrm{E}(\mathrm{Y_i})$$

- The log link function relates the mean μ_i to the linear predictor:

$$\log(\mu_\mathrm{i}) = \eta_\mathrm{i} = \tilde{\mathrm{x}}_\mathrm{i}^\mathrm{T}\beta$$

$$\mu_\mathrm{i} = \exp(\eta_\mathrm{i}) = \exp\left(\tilde{\mathrm{x}}_\mathrm{i}^\mathrm{T}\beta\right).$$

- The **Poisson Serfling regression** can be expressed as:

$$\log(\mu_\mathrm{t}) = \beta_0 + \beta_1\mathrm{t} + \beta_3\cos(\mathrm{wt}) + \beta_4\sin(\mathrm{wt})$$

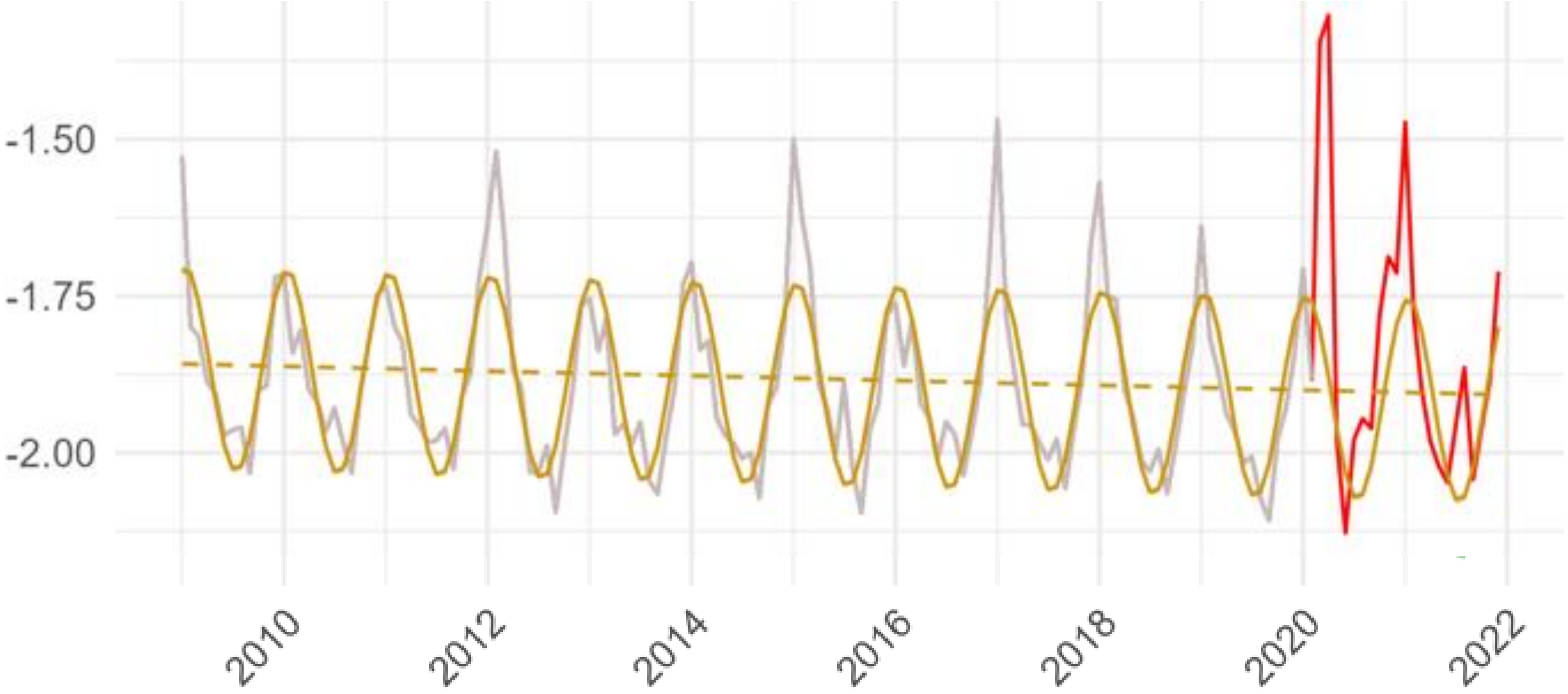
where $t = 1, \dots, T; w = 2\pi/p$ with $p = 12$ months.

To incorporate exposures $\mathrm{e_t}$ to model **death rates**, the regression becomes:

$$\log(\mu_\mathrm{t}) = \log(e_t) + \beta_0 + \beta_1\mathrm{t} + \beta_3\cos(\mathrm{wt}) + \beta_4\sin(\mathrm{wt})$$

Monthly log death rates, Spain, Men, Age 85+, Years 2009-2002

Poisson Serfling



Aim

- Methodological choices influence the estimates of excess death, e.g., the **Poisson Serfling** model imposes a rigid structure on the trend and seasonality.
- **Aim**: To propose and compare several approaches with varying degree of flexibility to forecasting mortality in the short term:
  - The **Poisson Serfling** with fixed trend and seasonality.
  - The **modulation model** with smooth trend and varying seasonality.

    We combine the modulation models for seasonal data (Eilers et al., 2008) and forecasting with P-splines smoothing developed in the long term (Currie et al., 2004).
  - A **smooth trend model** with fixed seasonality.

Modulation model

- We want to account for non-linear trends. We seek a smooth estimate of **μ** = (μi).
- **Eilers et al. (2008)** developed the **modulation models** that introduce a smooth trend function and time-varying coefficients:

$$\log(\mu_t) = \upsilon_t + f_t \cos(wt) + g_t \sin(wt)$$

where υ accounts for the smooth trend function; f and g are smooth functions that describe the amplitudes of the cosine and sine and are constructed by approximating B-spline basis.

- In the matrix formulation

$$\log(\mu) = B\alpha + CB\beta + SB\gamma = \eta$$

with B a B-spline regression matrix, $\upsilon = B\alpha$, $f = B\beta$, and $g = B\gamma$ and C and S diagonal matrices. A penalty on the B-splines coefficients for the trend and the modulation functions forces them to vary smoothly.

GLM **estimation** is performed with the Iterated Re-Weighted Least Squares (IRLS).

Smooth trend model

sdu.dk/cpop

- We propose a model with smooth trend and fixed seasonality:

$$\log(\mu_t) = \upsilon_{\mathrm{t}} + \beta_1 \cos(\mathrm{wt}) + \beta_2 \sin(\mathrm{wt}).$$

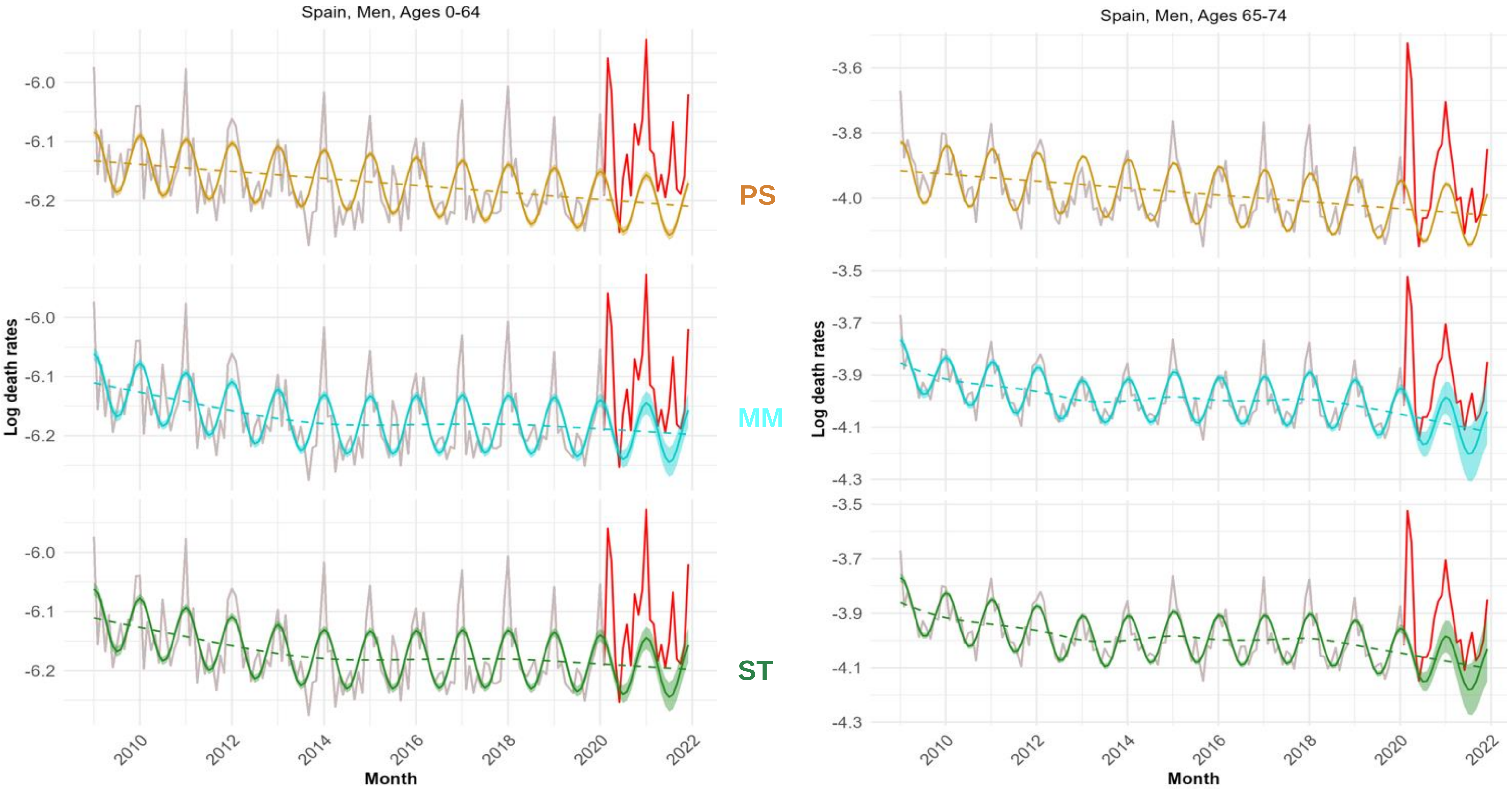
- The linear predictor models the trend component with the B-splines matrix B and time varying coefficients α, and the seasonal component with the vectors $c_{\mathrm{t}} = \cos(\mathrm{wt})$ and $s_{\mathrm{t}} = sin(\mathrm{wt})$ and $\eta = \left[\mathrm{B}\middle|c_t^{\prime}\middle|s_t^{\prime}\right][\alpha'|\beta^1|\beta^2] = \mathbf{B\theta}$.
- Same iterated GLM scoring algorithm can be used.

Forecasting with P-splines (Currie et al., 2004)

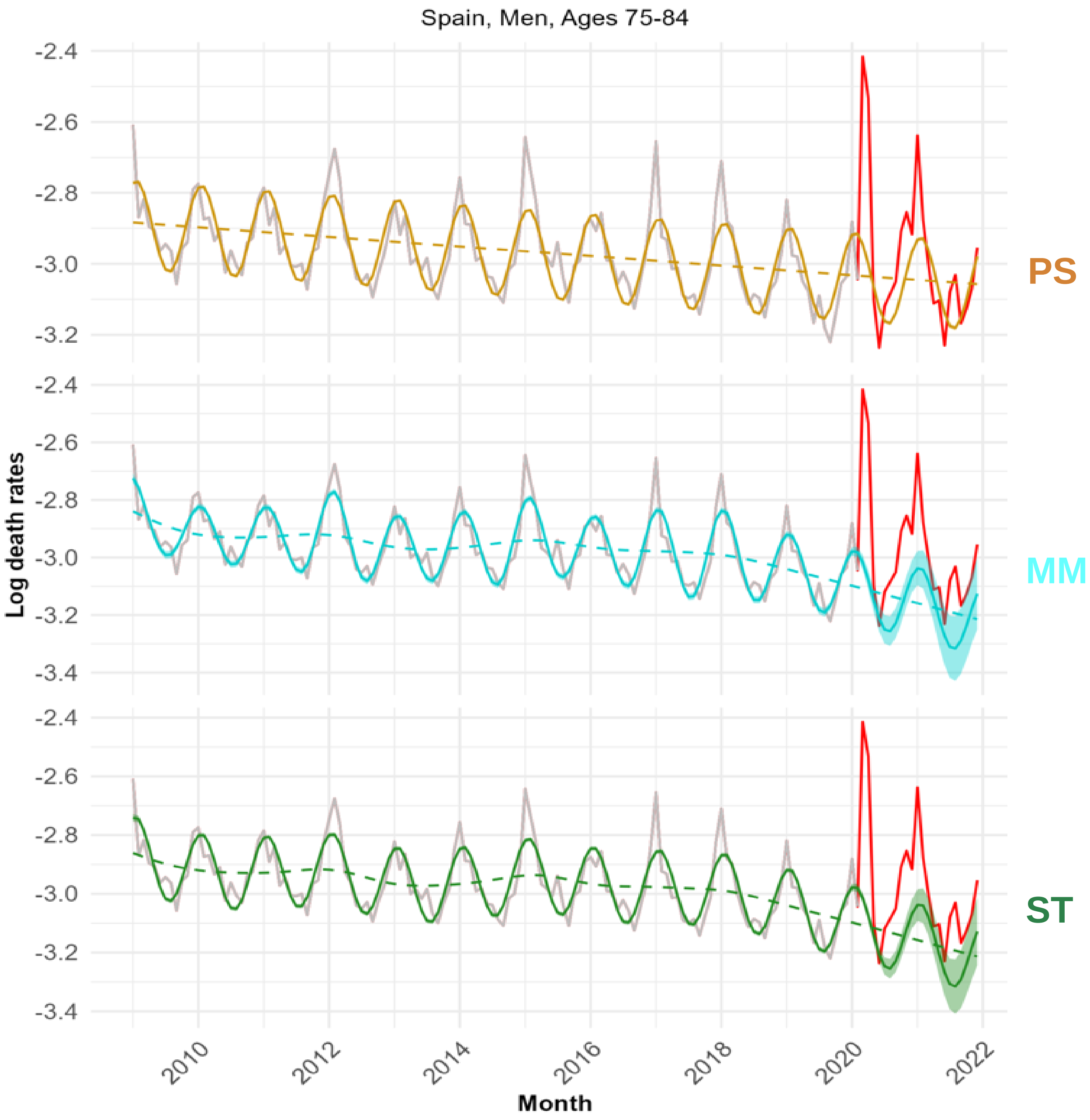
- Forecasting can be performed by adapting the missing data method (non seasonal) developed by **Currie et al. (2004)**.
- The forecasting of future values is treated as a missing value problem and the fitted and forecast values are estimated simultaneously.
- We have data y_1 for n_1 months and we wish to forecast n_2 months into the future.
  - ➢ We extend the set of knots and compute the regression matrix B for n_1 + n_2 months.
  - ➢ We define a weight matrix $\mathbf{V} = \mathrm{blockdiag}(\mathbf{I}; \mathbf{0})$ where $\mathbf{I}$ has size n_1 and $\mathbf{0}$ has size n_2.

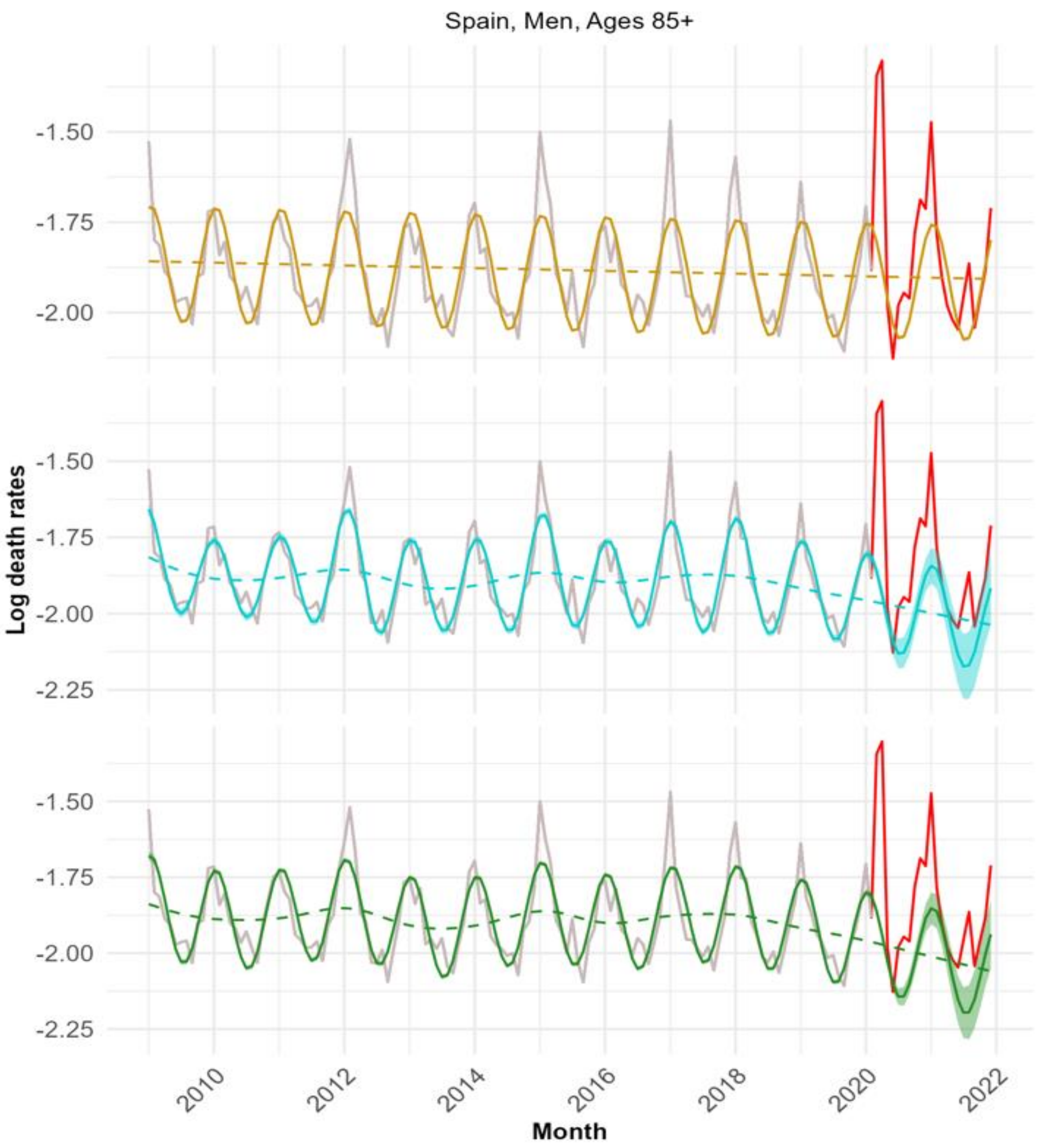
Data and application

- Monthly death counts by sex and age-group for:
  - **Denmark** from 2007 to 2022 (Statistics Denmark)
  - **Spain** from 2009 to 2022 (Instituto Nacional de Estadística)
  - **Sweden** from 2000 to 2022 (Statistics Sweden).
- To obtain monthly mortality forecasts for **1-3 epidemic years**, comparing:
  - **the Poisson Serfling (PS)**
  - **the modulation models (MM)**
  - **the smooth trend model (ST)**.


Spain, Men, Ages 0-64
Spain, Men, Ages 65-74
Log death rates
Log death rates
Month
Month
PS
MM
ST
-6.0
-6.1
-6.2
-3.6
-3.8
-4.0
-3.5
-3.7
-3.9
-4.1
-4.3
2010
2012
2014
2016
2018
2020
2022

sdu.dk/cpop





sdu.dk/cpop

CPop SDU SCOR FOUNDATION FOR SCIENCE

Goodness of fit on death rates. BIC, Men. MM performs best

Age	5 years series PS	5 years series MM	5 years series ST	10 years series PS	10 years series MM	10 years series ST
Denmark						
0-64	101.29	**100.9**	101.00	199.13	195.31	**195.27**
65-74	**98.7**	98.71	98.74	209.96	**195.5**	195.51
75-84	144.47	**141.8**	141.92	294.18	**284.07**	284.3
85	152.25	**149.78**	151.58	298.42	**291.78**	291.95
Sweden						
0-64	114.55	**114.49**	114.60	231.62	**225**	225.35
65-74	122.52	**121.5**	122.53	254.96	**238.59**	239
75-84	197.56	**194.06**	196.30	381.87	**379.99**	380.42
85	254.4	**241.28**	248.51	503.66	**492.54**	493.67
Spain						
0-64	433.64	**426.46**	430.84	961.97	**911.79**	914.05
65-74	446.49	**417.17**	428.51	1101.62	**956.8**	963.81
75-84	1223.97	**1036.26**	1109.52	2645.99	**2261.72**	2450.14
85	1809.11	**1430.97**	1574.19	3752.08	**3094.7**	3408.09

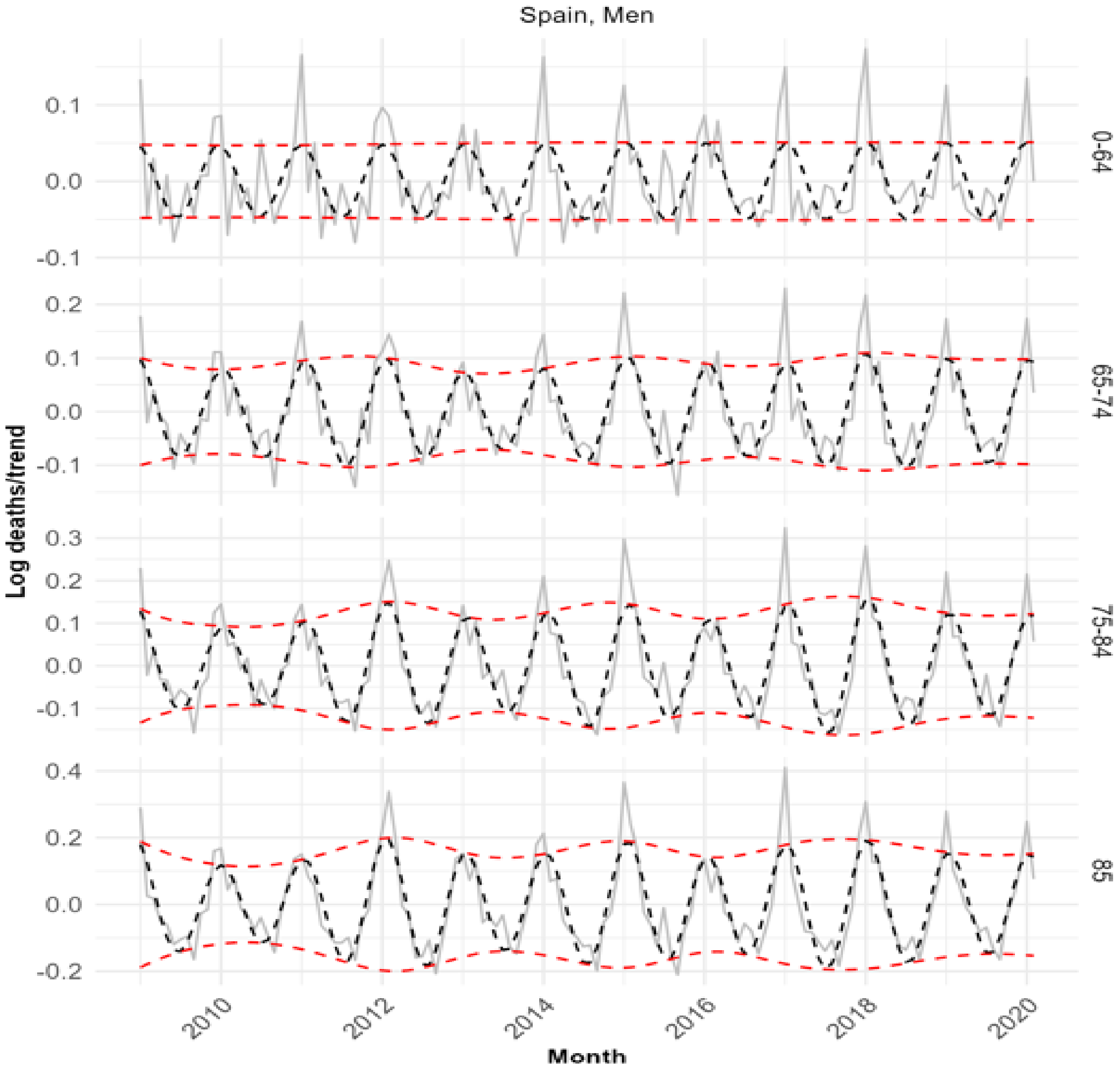
Accuracy of death rates. RMSE (x 1000) and MAPE, Men. ST performs best

	5 years series						10 years series					
	RMSE			MAPE			RMSE			MAPE		
Age	PS	MM	ST	PS	MM	ST	PS	MM	ST	PS	MM	ST
Denmark												
0-64	**0.14**	0.15	0.14	5.13	5.117	**5.03**	0.17	0.15	**0.15**	6.324	**5.699**	5.709
65-74	1.2	1.2	**1.19**	4.623	4.623	**4.621**	1.69	1.25	**1.25**	6.725	**4.846**	4.847
75-84	3.52	3.38	**3.36**	4.891	4.798	**4.754**	**3.29**	3.55	3.56	**4.526**	4.835	4.851
85	12.93	13.34	**12.85**	5.81	5.988	**5.756**	**13.07**	14.05	13.83	**5.801**	6.396	6.282
Sweden												
0-64	0.1	0.1	**0.1**	4.667	4.671	**4.665**	0.1	0.1	**0.09**	5.181	4.822	**4.818**
65-74	1	1	**0.99**	4.637	4.651	**4.589**	1.13	**0.92**	0.92	5.398	4.517	**4.499**
75-84	**2.83**	3.01	2.85	**4.131**	4.421	4.225	**2.66**	2.66	2.66	**4.159**	4.228	4.219
85	**10.82**	11.3	10.83	**4.856**	5.097	4.906	10.25	10.51	**10.14**	4.782	4.829	**4.728**
Spain												
0-64	0.1	0.1	**0.1**	3.648	3.724	**3.64**	**0.08**	0.09	0.08	**2.54**	3.765	3.682
65-74	1.08	**1.04**	1.06	4.523	**4.29**	4.463	0.76	**0.74**	0.75	**3.267**	3.4	3.4
75-84	3.72	3.72	**3.7**	5.488	5.117	**4.998**	3.78	**2.54**	2.72	7.726	**3.683**	3.865
85	**13.15**	13.82	14.01	**6.137**	6.531	6.745	10.85	**10.27**	12.42	6.383	**5.325**	6.652

A rolling window of 5 or 10 years (from first available year by country to 2019) is used to forecast one year ahead.

Changes of the modulation component over time are negligible

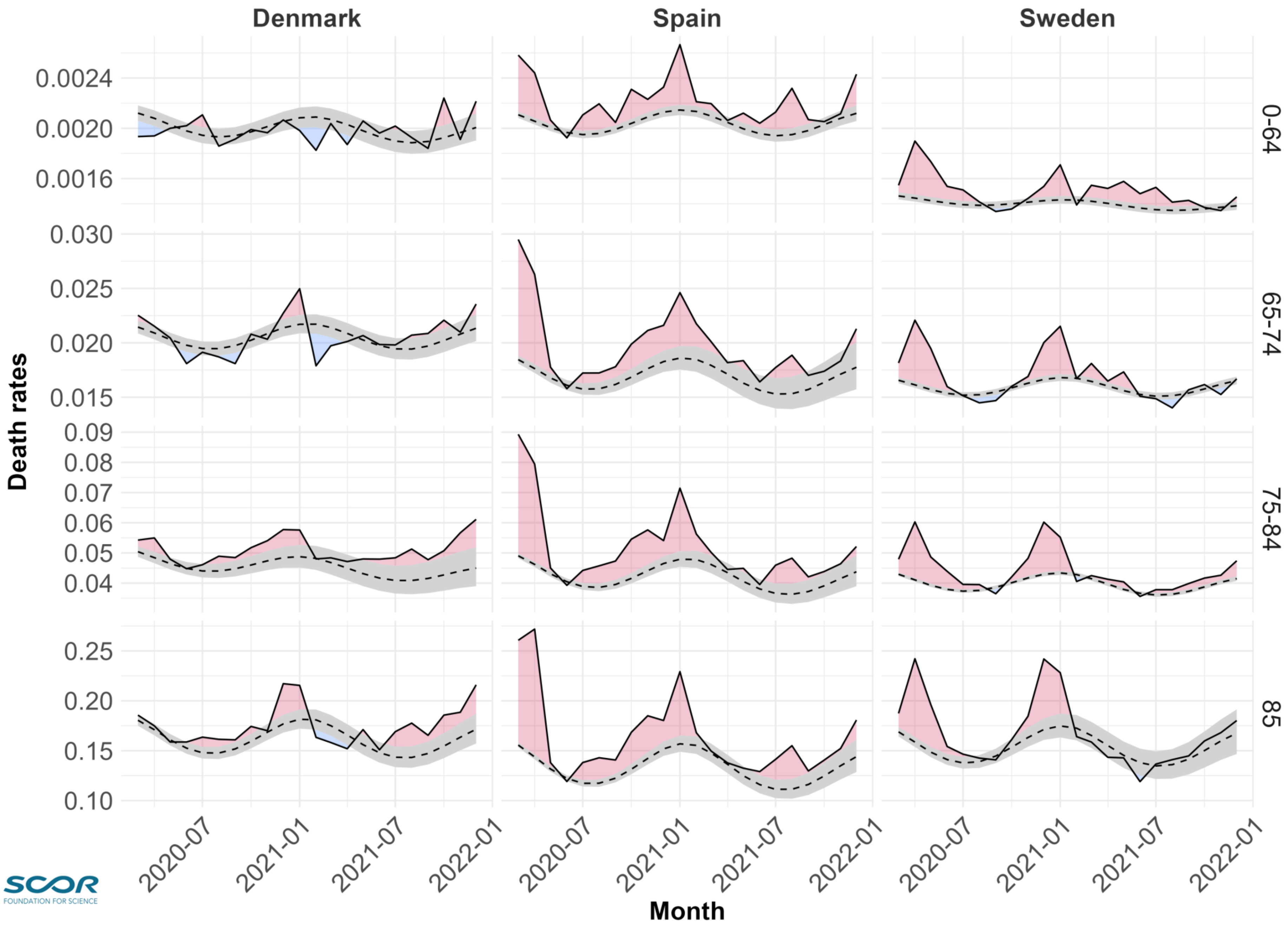
sdu.dk/cpop



- gray line = detrended series
- dashed black line = modulated component
- red dashed line = amplitude

CPop SDU SCOR FOUNDATION FOR SCIENCE

ST model Excess mortality, Men



sdu.dk/cpop

CPop SDU SCOR FOUNDATION FOR SCIENCE

november 2024

Contribution

- Forecasts of baseline seasonality in mortality are useful for quantifying the impact of seasonal influenza, heatwaves, past and future pandemics.
- Smoothness of the trend is a desirable feature when modeling the expected seasonality for all-cause death data.
- Further application to cause-of-death data.
- Consider mortality surface to model jointly age and years/months.

Ongoing work

- How does the Covid-19 health shock impact cohort life expectancy?
- Methodology based on the penalized composite link model for killing off cohorts (Rizzi et al., 2020).
- Counterfactual cohort life expectancy across HMD countries.

Thank you.

Ainhoa-Elena Leger
aleg@sdu.dk @LegerAinhoa
Silvia Rizzi
srizzi@sdu.dk @silviriz
Ugofilippo Basellini
basellini@demog.mpg.de @ugobas

Appendix

B-splines

B-splines of degree q:

- Consist of q+1 polynomial pieces;
- Polynomial pieces join at q inner knots;

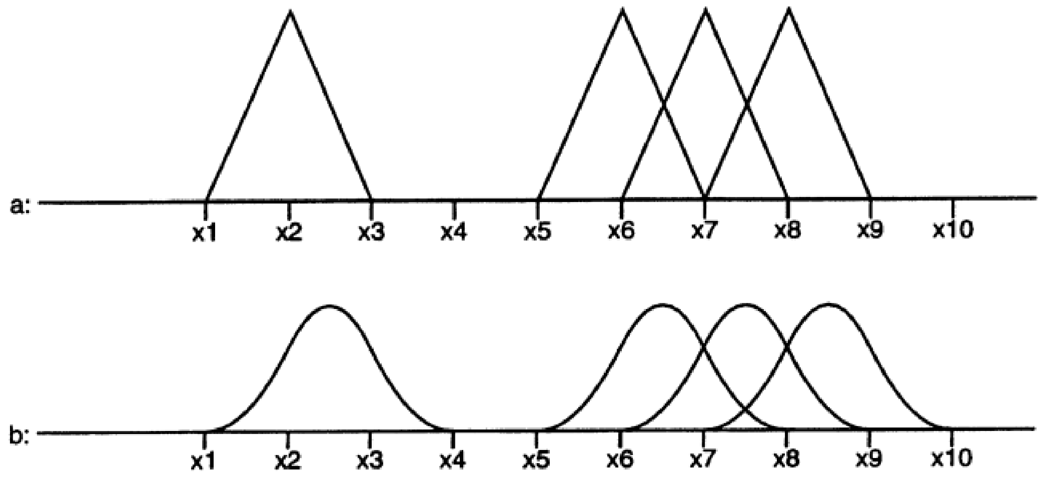


FIG. 1. *Illustrations of one isolated B-spline and several overlapping ones* (a) *degree* 1; (b) *degree* 2.

Eilers and Marx, 1996

Estimation of the modulation model

- The GLM estimation is performed with the Iterated Weighted Least Squares (IWLS).
- The estimate $\hat{\beta}$ is obtained iteratively

$$\beta^{(t+1)} = \beta^{(t)} + I(\beta^{(t)})^{-1} l'\left(\beta^{(t)}\right)$$

which can be re-written as

$$I(\beta^{(t)})\beta^{(t+1)} = I(\beta^{(t)})\beta^{(t)} + l'\left(\beta^{(t)}\right)$$

$$(X^T \widetilde{W}^{(t)} X)\beta^{(t+1)} = X^T \widetilde{W}^{(t)} X\ \beta^{(t)} + X^T(y - \tilde{\mu})$$

- The modulation model can be expressed as $\eta = [\mathrm{B}|\mathrm{CB}|\mathrm{SB}][\alpha'|\beta'|\gamma'] = \mathbf{B\theta}$ and the iterated GLM scoring algorithm can be used.

$$(\mathbf{B}^T \widetilde{\mathbf{M}}^{(\mathrm{t})}\mathbf{B}\ + \mathbf{P})\boldsymbol{\theta}^{(\mathrm{t+1})} = \mathbf{B}^T \widetilde{\mathbf{M}}^{(\mathrm{t})}\mathbf{B}\boldsymbol{\theta}^{(\mathrm{t})} + \mathbf{B}^T(y - \tilde{\mu}).$$

Estimation of the modulation model

We minimize the penalized Poisson deviance defined as

$$d^*(y;\mu) = 2\sum_{t=1}^{T} log(y_t/\mu_t) + \lambda_1 \|\boldsymbol{D\alpha}\|^2 + \lambda_2 \|\boldsymbol{D\beta}\|^2 + \lambda_2 \|\boldsymbol{D\gamma}\|^2$$

where the matrix $\boldsymbol{D} = \Delta^d$ constructs dth order differences of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. For instance, Δ^1 is a matrix $(J-1) \times J$ of first differences and $\boldsymbol{\Delta^1\alpha}$ is the vector with elements $\boldsymbol{\alpha_{j+1}} - \boldsymbol{\alpha_j}$, for $j = 1, \ldots, J-1$. By repeating this computation on $\boldsymbol{\Delta^1\alpha}$, we arrive at higher differences like $\boldsymbol{\Delta^2\alpha}$, where $\boldsymbol{\Delta^2}$ the $(J-2) \times J$ matrix of second-order differences of a J-vector.

The linear re-expressions 2 and 3 allow all of the parameters associated with each components to be estimated simultaneously as GLMs. Estimation of the coefficients is performed via the penalized version of the Iterated Weighted Least Squares (IWLS)

$$(\boldsymbol{\breve{B}'\widetilde{M}^{(t)}\breve{B}} + \boldsymbol{P})\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\breve{B}'\widetilde{M}^{(t)}\breve{B}\theta^{(t)}} + \boldsymbol{\breve{B}'}(\boldsymbol{y} - \boldsymbol{\tilde{\mu}}), \qquad (4)$$

where $\boldsymbol{\breve{B}}$ is the regression matrix, $\boldsymbol{\widetilde{M}} = diag(\boldsymbol{\mu})$ is the matrix of weights, and $\boldsymbol{P} = \Lambda\boldsymbol{D'D}$ is the penalty term. The positive penalty hyper-parameter $\Lambda = diag(\lambda_1, \lambda_2, \lambda_2)$ balance smoothness against fit to the data and allows for different penalty for the trend (λ_1) and modulation functions (λ_2). The penalty matrix can also be constructed to consider differents order of the differences for the trend and modulation functions as $\boldsymbol{P} = blockdiag(\lambda_1\boldsymbol{D_1'D_1}, \lambda_2\boldsymbol{D_2'D_2}, \lambda_2\boldsymbol{D_2'D_2})$. For instance, Carballo (2019) suggests λ_1 and λ_2 equal to 2 and 1, respectively.

The modulation model present some similarities with some other models in the literature. The smooth trend component $\boldsymbol{v}$ can be seen as a generalized additive model (GAM, Hastie & Tibshirani, 1990) and the seasonal components $\boldsymbol{f}$ and $\boldsymbol{g}$ as a varying-coefficient model (VCM, Hastie & Tibshirani, 1993). The advantage of P-splines over GAM and VCM is that they avoid both the backfitting algorithm and complex knot selection schemes.