Modelling and short-term forecasting of seasonal mortality (arXiv pre-print)

Ainhoa-Elena Leger[a], aleg@sdu.dk, **Silvia Rizzi**[a], srizzi@sdu.dk, **Ugofilippo Basellini**[b], basellini@demogr.mpg.de

[a]University of Southern Denmark (Cpop), [b]Max Planck Institute for Demographic Research (MPIDR)

Context

- **Excess mortality** is used to quantify the toll of mortality shocks, such as infectious disease-related epidemics and pandemics.
- Excess mortality is the difference between observed and **expected mortality** (predicted).
- The coronavirus disease (COVID-19) pandemic boosted interest in **short-term mortality forecasts** in an epidemiological year.

Data

- Series of monthly death counts $\boldsymbol{y}$ and exposures $\boldsymbol{e}$ for 25 European countries from 1995 through June 2022 from Eurostat → the crude death rates (CDRs) are obtained as $\boldsymbol{y}/\boldsymbol{e}$.

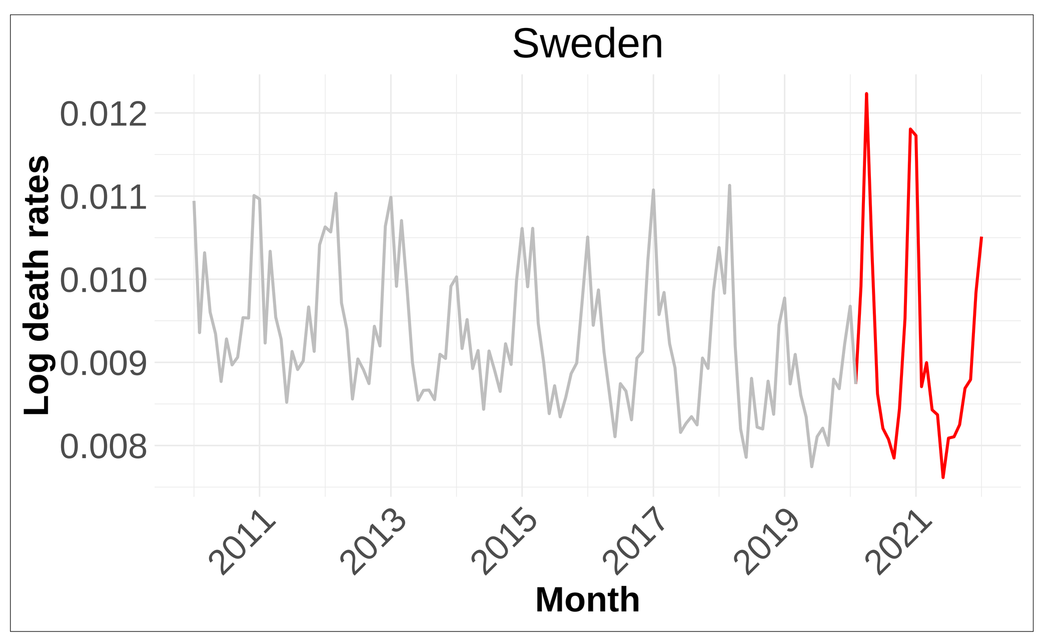

Figure: CDRs 2010–2022 (COVID-19 pandemic in red).

Literature

- The **Serfling model** (1963) first estimated a standard curve of expected seasonal mortality.
- Recent versions of the Serfling model are Poisson regressions (Thompson et al., 2009) and Quasi-Poisson regressions (EuroMOMO).
- Flexible versions of the model assume **smooth effects** to forecast the expected mortality (Aburto et al., 2021; Scortichini et al., 2020).
- The **modulation models** (Eilers et al., 2008) combine GLMs with P-spline smoothing.
- Few studies studied the sensitivity of excess mortality (Nepomuceno et al., 2022), but comparisons of forecasting approaches are lacking.

Objectives and reseach questions

- Extend the modulation models for forecasting mortality more flexibly.
- Which specification of the Serfling-Poisson model yields the most accurate predictions?

Methods - Models

The death counts y_t at months $t = 1, \dots, T$ are realisations of $Y_t \sim Poi(\mu_t)$. We combine the Serfling-Poisson model with P-splines (Currie, 2004) and compare it with the original model.

- parametric effect for the trend and seasonality (**SP**)

$$log(\mu_t) = log(e_t) + \beta_0 + \beta_1 t + \beta_3 cos(wt) + \beta_4 sin(wt)$$

- smooth trend and smooth seasonality (**SP-STSS**) (also known as modulation model),

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

- smooth trend and fixed seasonality (**SP-STFS**)

$$log(\mu_t) = log(e_t) + \upsilon_t + \beta_1 cos(wt) + \beta_2 sin(wt)$$

with $w = 2\pi/p$, period $p = 12$, smooth trend $\upsilon_t = \Sigma_j \alpha_j B_j(t)$, smooth modulation functions $f_t = \Sigma_j \beta_j B_j(t)$ and $g_t = \Sigma_j \gamma_j B_j(t)$, and $B_j(t)$ basis of B-splines.

Methods - Forecasts

- Forecasting is treated as a missing value problem. The IWLS algorithm is adapted to estimate the fitted values y_1 (n_1 months) and forecast values y_2 (n_2 months) simultaneously.

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

with $\boldsymbol{\breve{B}}_+$ and $\boldsymbol{P}_+$ extended regression matrix and penalty matrices, $\boldsymbol{\theta}$ regression parameters, $\boldsymbol{V} = blockdiag(\boldsymbol{I}_{n_1}; \boldsymbol{0}_{n_2})$ matrix weighting 1 the observations and 0 the forecasts.

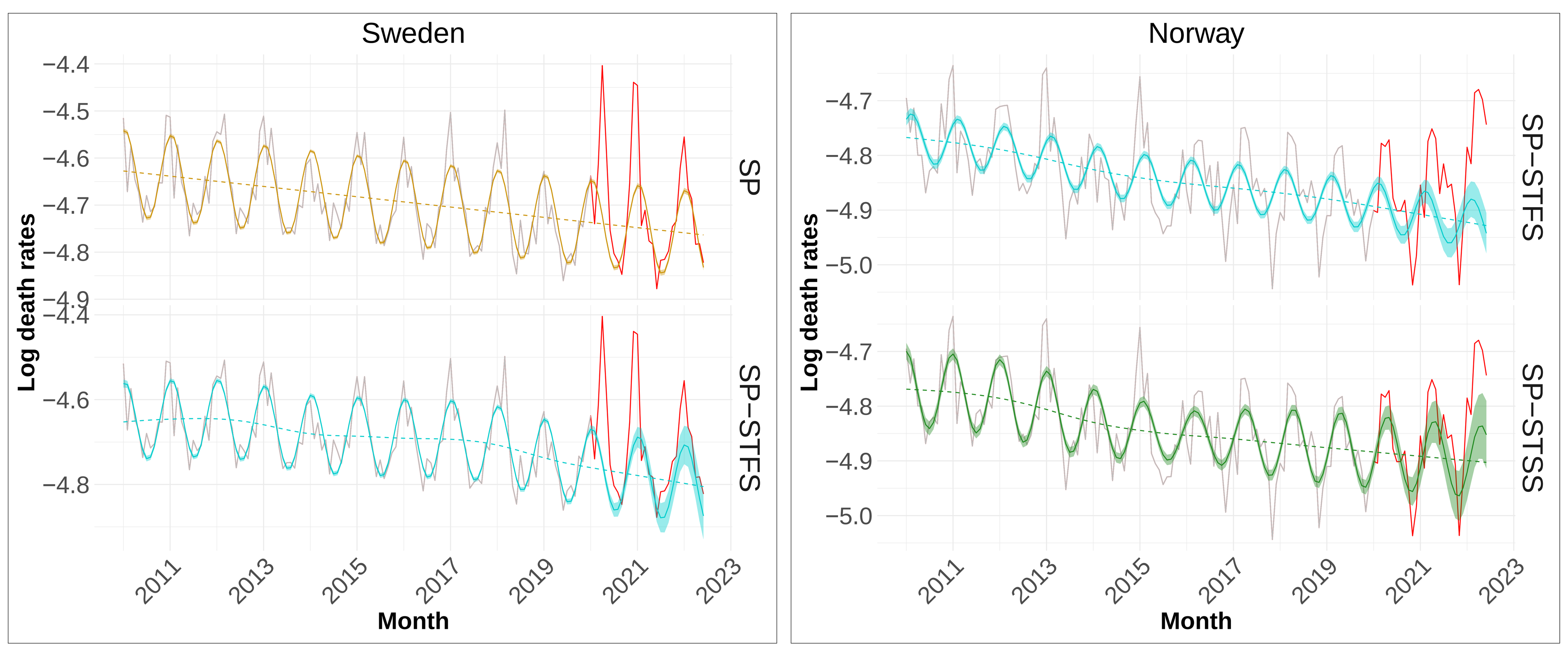

Figure: Modelling and forecasting of CDRs. Trend function (dashed) and modulation functions (solid).

Conclusions

The comparison of Serfling-Poisson models revealed that accounting for smooth trends over longer reference periods is a desirable feature for short-term mortality predictions.

Results

- Using a **10-year fitting period** leads to a better forecasting accuracy than a 5-year fitting period.
- The preferred model according to mean MAPE is the **STFS**, followed by the SP, and the STSS model.

Table: Mean RMSE and MAPE on one-year ahead historical forecasts of CDRs (x 1000) based on a rolling-window scheme.

	5 years series						10 years series					
	RMSE			MAPE			RMSE			MAPE		
Country	SP	STSS	STFS	SP	STSS	STFS	SP	STSS	STFS	SP	STSS	STFS
Austria	0.52	0.53	0.53	4.08	4.12	4.06	0.51	0.51	0.5	4.1	3.93	3.9
Bulgaria	0.98	1.01	1.01	5	5.24	5.19	0.98	0.96	0.97	4.85	4.8	4.83
Croatia	0.8	0.83	0.82	4.93	5.15	5.06	0.77	0.78	0.78	4.24	4.41	4.43
Czechia	0.56	0.56	0.55	3.89	3.82	3.68	0.54	0.53	0.52	3.71	3.63	3.59
Denmark	0.5	0.51	0.5	3.85	3.88	3.85	0.48	0.48	0.47	3.76	3.64	3.56
Estonia	0.71	0.72	0.72	4.39	4.47	4.46	0.69	0.65	0.65	4.33	4.13	4.07
Finland	0.49	0.5	0.5	3.82	3.9	3.93	0.47	0.47	0.46	3.59	3.59	3.56
France	0.52	0.54	0.53	4.13	4.46	4.38	0.47	0.49	0.47	3.81	4.09	3.85
Germany	0.59	0.61	0.59	3.84	4.06	3.93	0.6	0.59	0.58	3.91	3.69	3.62
Greece	0.76	0.76	0.75	5.38	5.35	5.28	0.74	0.74	0.75	5.12	5.26	5.27
Hungary	0.8	0.83	0.82	4.3	4.58	4.46	0.74	0.74	0.74	3.8	3.95	3.94
Iceland	0.65	0.65	0.65	8.07	8.09	8.09	0.59	0.59	0.59	7.62	7.59	7.59
Ireland	0.49	0.48	0.49	5.36	5.38	5.46	0.46	0.4	0.4	5.48	4.86	4.86
Italy	0.71	0.78	0.77	5.33	5.88	5.84	0.68	0.73	0.72	4.91	5.35	5.3
Lithuania	0.84	0.82	0.81	5.2	4.99	4.96	1.02	0.87	0.87	6.2	5.21	5.22
Luxembourg	0.63	0.64	0.64	6.84	7.03	7.03	0.57	0.57	0.57	6.24	6.3	6.3
Netherlands	0.47	0.48	0.48	3.98	4.2	4.17	0.49	0.46	0.45	4.25	3.9	3.85
Norway	0.52	0.5	0.51	4.82	4.68	4.82	0.49	0.46	0.49	4.96	4.49	4.97
Poland	0.53	0.57	0.55	3.76	4.03	3.97	0.56	0.54	0.54	3.96	3.71	3.71
Portugal	0.84	0.89	0.88	5.79	6.31	6.32	0.82	0.84	0.83	5.54	5.82	5.77
Romania	0.83	0.85	0.85	5.24	5.36	5.35	0.73	0.76	0.78	4.24	4.68	4.72
Slovenia	0.6	0.62	0.61	4.89	4.93	4.93	0.58	0.58	0.58	4.35	4.36	4.32
Spain	0.66	0.69	0.68	5.53	5.67	5.56	0.64	0.62	0.62	5.38	5	4.99
Sweden	0.51	0.51	0.5	3.85	3.87	3.8	0.45	0.46	0.45	3.47	3.52	3.5
Switzerland	0.45	0.46	0.46	4.01	4.16	4.14	0.4	0.4	0.4	3.58	3.57	3.5
N countries*	17	2	6	18	1	6	7	4	14	9	3	13

* Number of countries in which the model performs the best

Illustration - COVID-19 pandemic

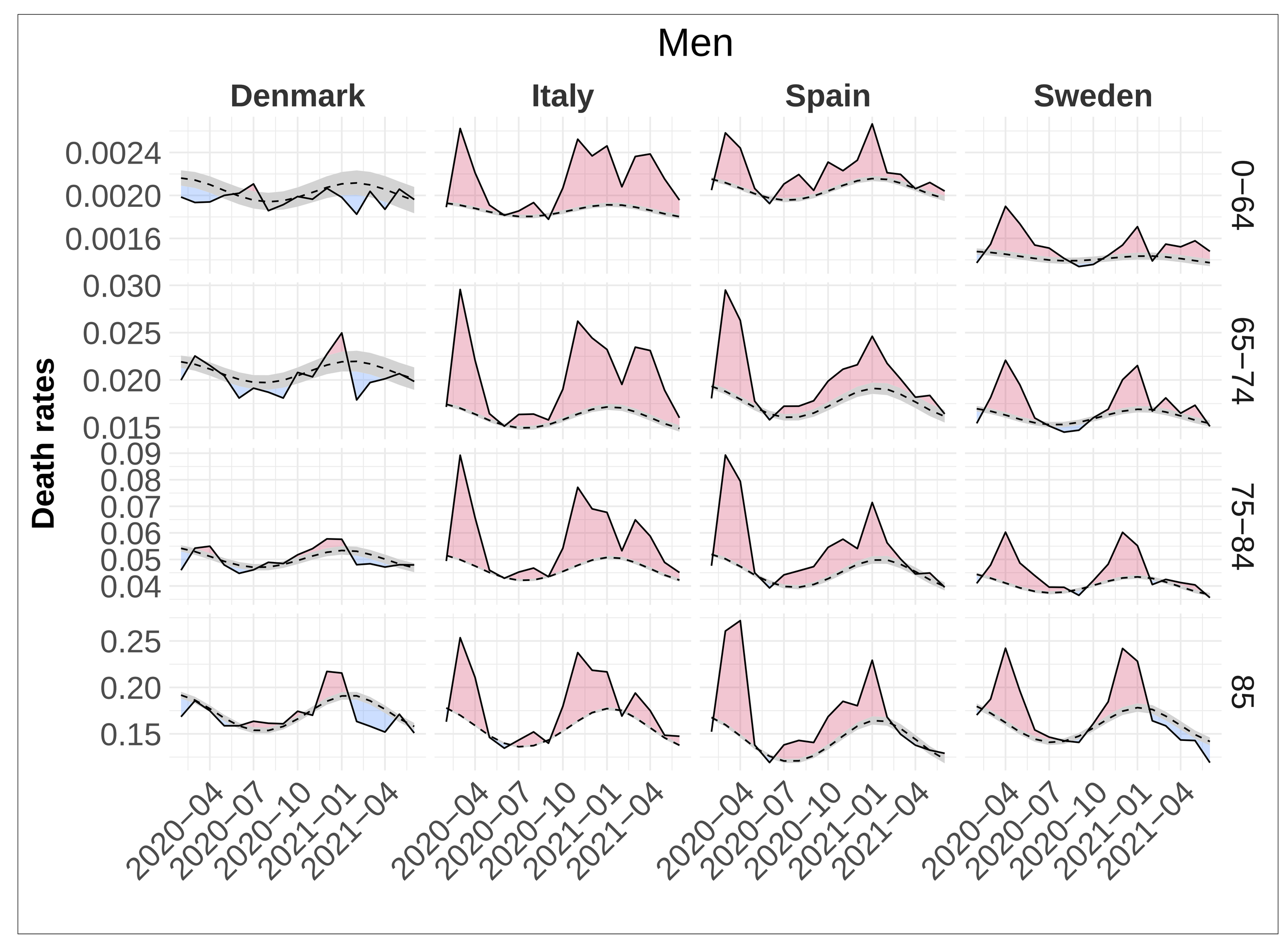

Figure: Excess CDRs (x 1000) in Denmark, Italy, Spain, Sweden.