MONASH University

Rob J Hyndman

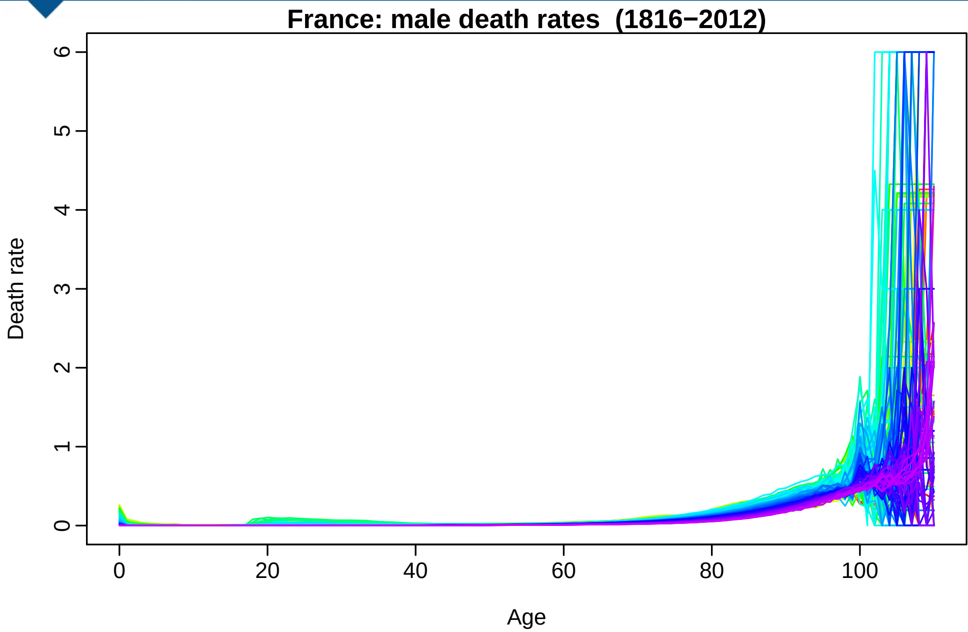
Functional time series

with applications in demography

3. Forecasting functional time series

Outline

Functional time series

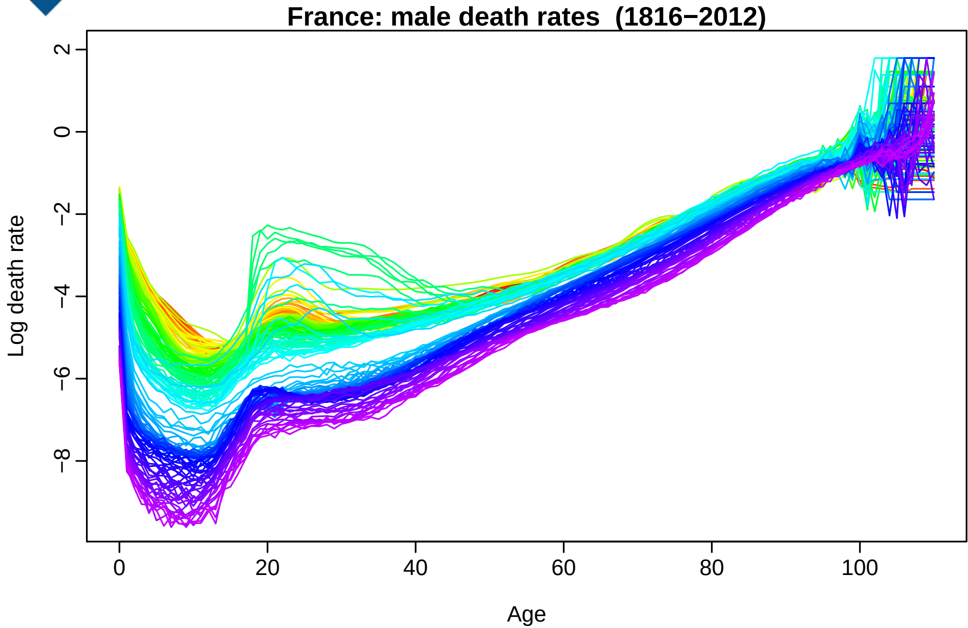


Functional time series

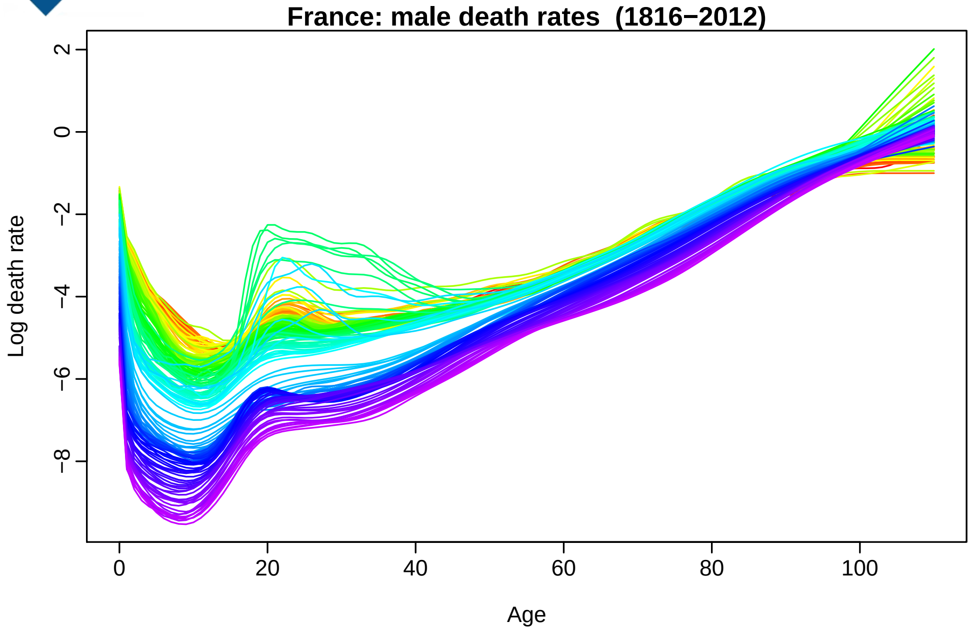
$$y_t(x_i) = g_\lambda(z_t(x_i)) = \begin{cases} \log[z_t(x_i)] & \text{if } \lambda = 0; \\ \lambda^{-1}\left[z_t^\lambda(x_i) - 1\right] & \text{otherwise.} \end{cases}$$
$$= s_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i}$$

- $z_t(x_i)$ is observed data for age x_i in year t, $i = 1, \dots, N, \quad t = 1, \dots, T$.
- λ chosen so that $\varepsilon_{t,i} \sim \text{NID}(0, 1)$.
- We estimate $s_t(x)$, a smooth function of x.
- We want to forecast **whole curve** $z_t(x)$ for $t = T + 1, \dots, T + h$.

Functional time series



Functional time series



Functional principal components

$$y_t(x_i) = s_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i},$$
$$s_t(x) = \mu(x) + \sum_{k=1}^{T-1} \beta_{t,k}\,\phi_k(x)$$

1. Estimate smooth functions $s_t(x)$ using weighted penalized regression splines.
2. Compute $\mu(x)$ as $\bar{s}(x)$ across years.
3. Compute $\beta_{t,k}$ and $\phi_k(x)$ using functional principal components.
4. To forecast $y_t(x_i)$, we need forecasts of $\{\beta_{t,k}\}$.

Functional time series model

$$y_{t,x} = s_t(x) + \sigma_t(x)\varepsilon_{t,x},$$
$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

- Only use the first K eigenfunctions.
- Cohort effects ignored.
- Check $r_t(x)$ to see if K large enough, and if cohort effects present.
- Robust PCA used because of wars.

Functional time series model

$$y_{t,x} = s_t(x) + \sigma_t(x)\varepsilon_{t,x},$$

$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

- Only use the first K eigenfunctions.
- Cohort effects ignored.
- Check $r_t(x)$ to see if K large enough, and if cohort effects present.
- Robust PCA used because of wars.

Functional time series model

$$y_{t,x} = s_t(x) + \sigma_t(x)\varepsilon_{t,x},$$
$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

- Only use the first K eigenfunctions.
- Cohort effects ignored.
- Check $r_t(x)$ to see if K large enough, and if cohort effects present.
- Robust PCA used because of wars.

Functional time series model

$$y_{t,x} = s_t(x) + \sigma_t(x)\varepsilon_{t,x},$$
$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

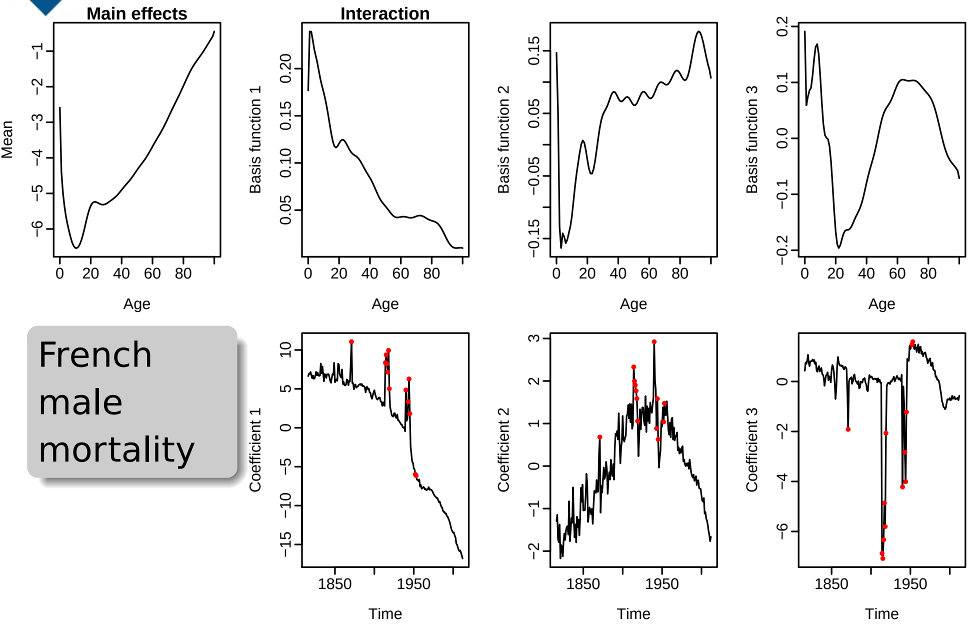
- Only use the first K eigenfunctions.
- Cohort effects ignored.
- Check $r_t(x)$ to see if K large enough, and if cohort effects present.
- Robust PCA used because of wars.

Functional time series model

$$y_{t,x} = s_t(x) + \sigma_t(x)\varepsilon_{t,x},$$
$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

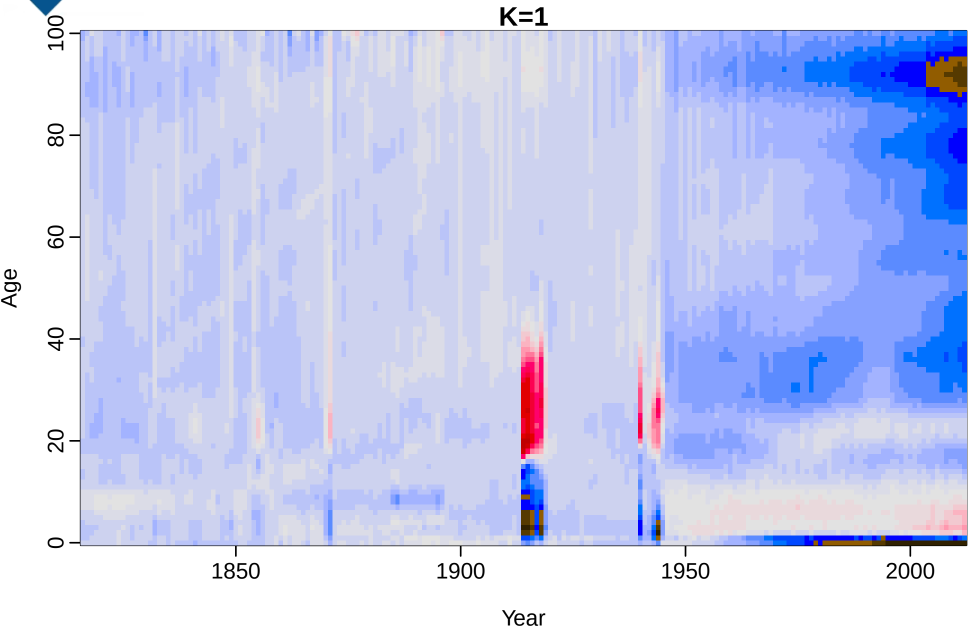
- Only use the first K eigenfunctions.
- Cohort effects ignored.
- Check $r_t(x)$ to see if K large enough, and if cohort effects present.
- Robust PCA used because of wars.

Functional time series model

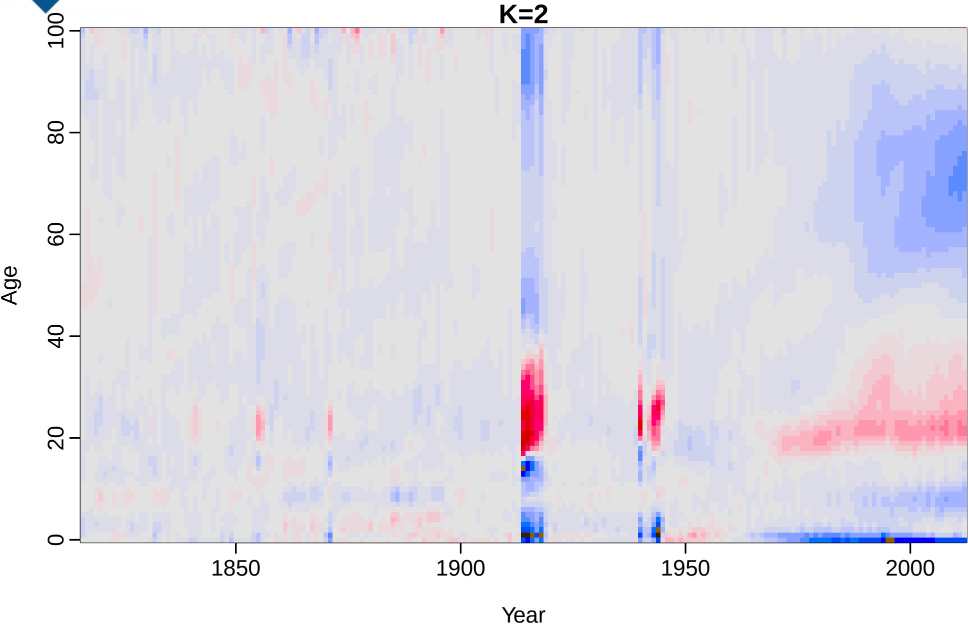


French male mortality

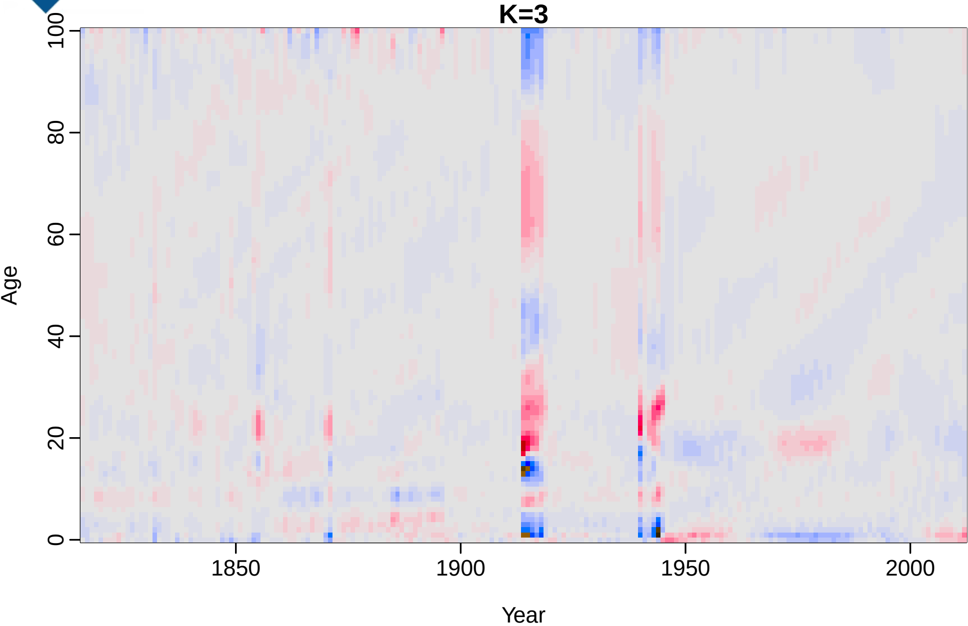
Functional time series model



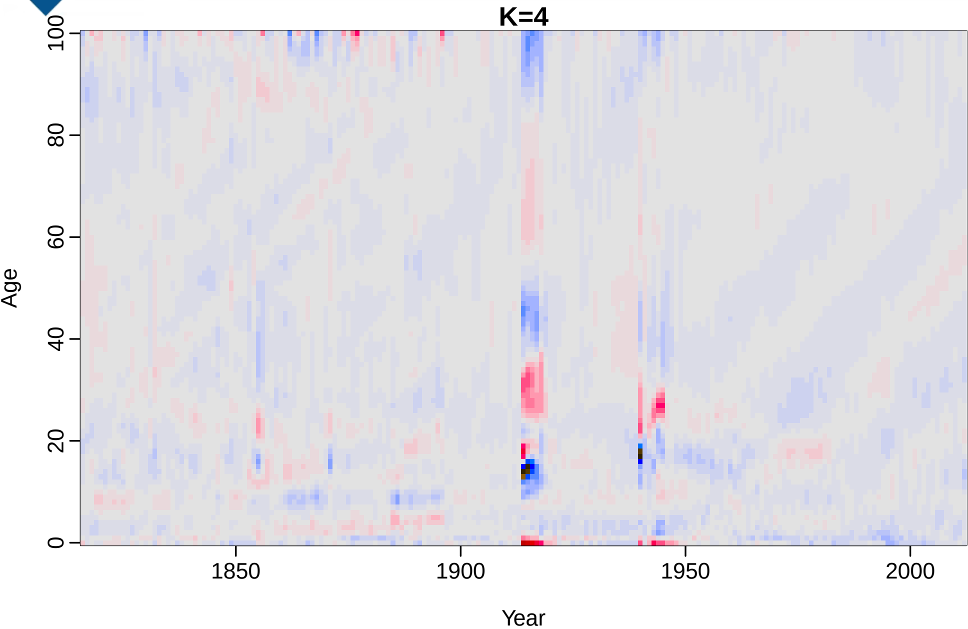
Functional time series model



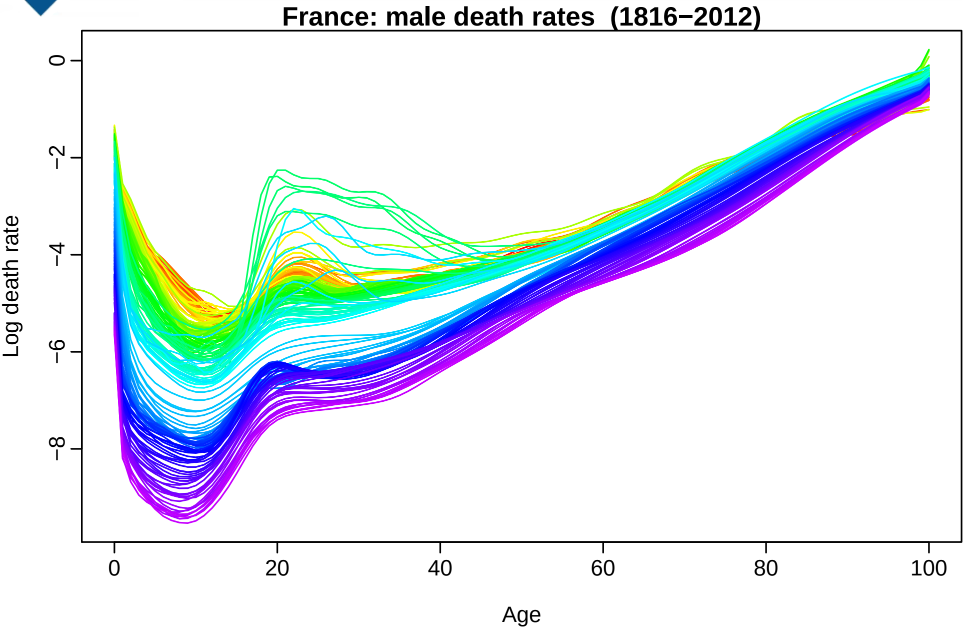
Functional time series model



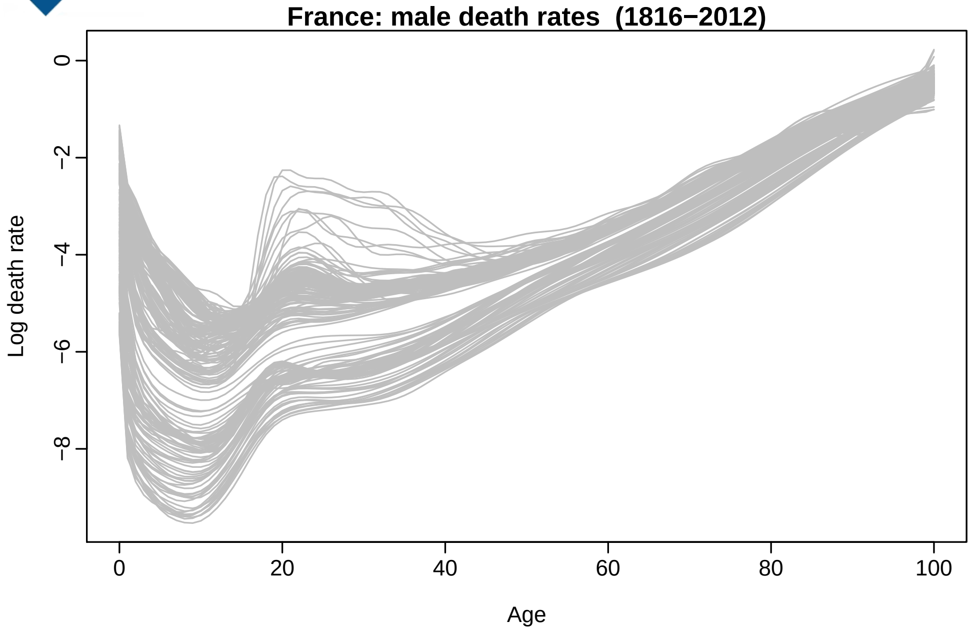
Functional time series model



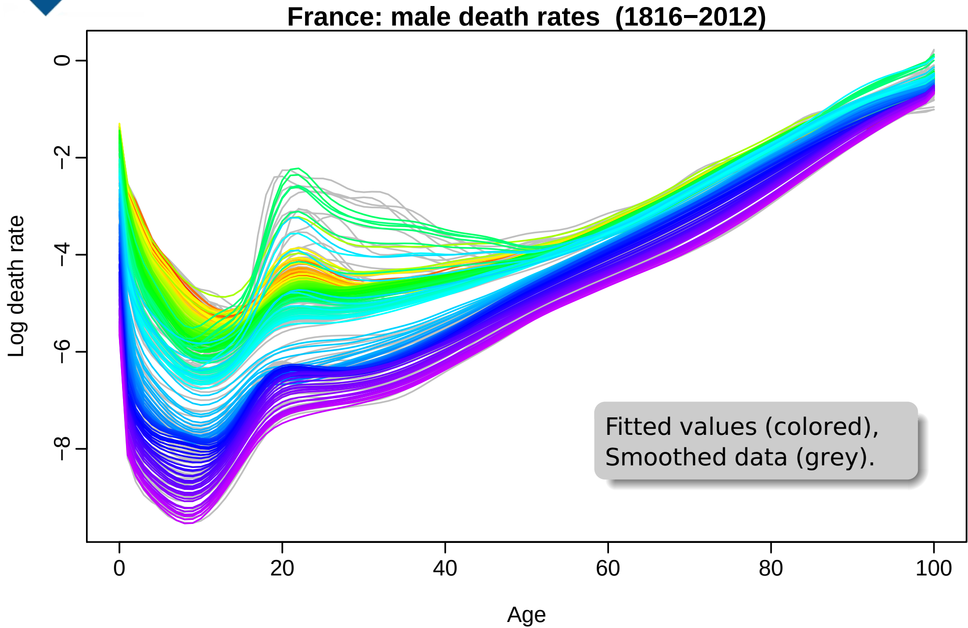
Functional time series model



Functional time series model

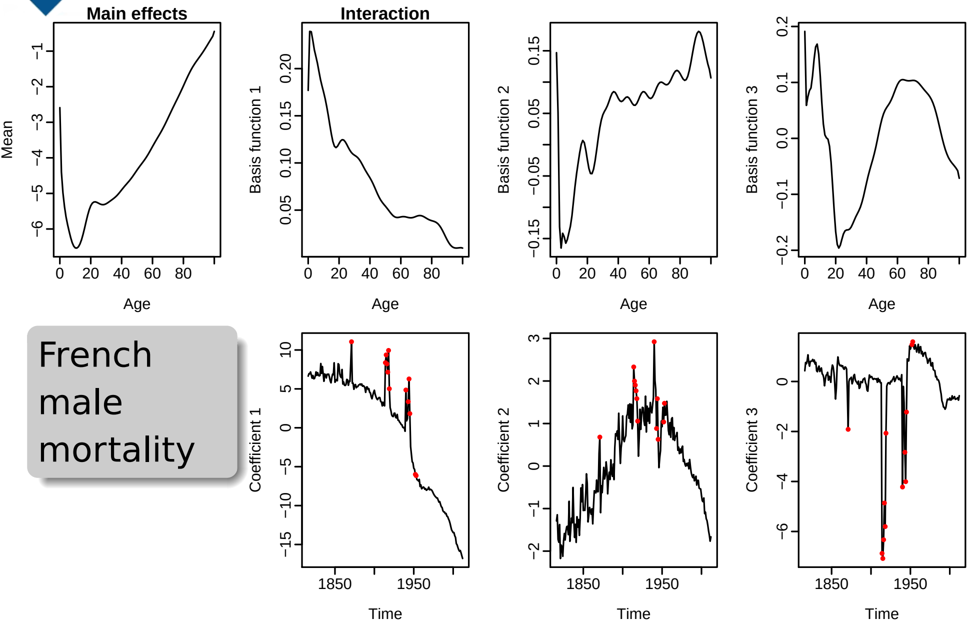


Functional time series model



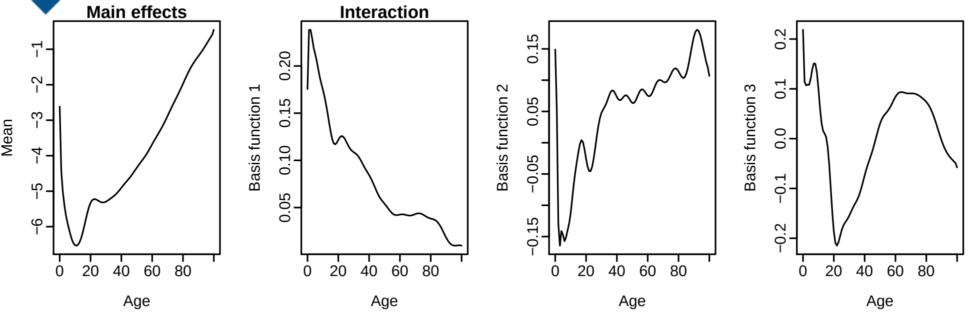
Outline

Functional principal components



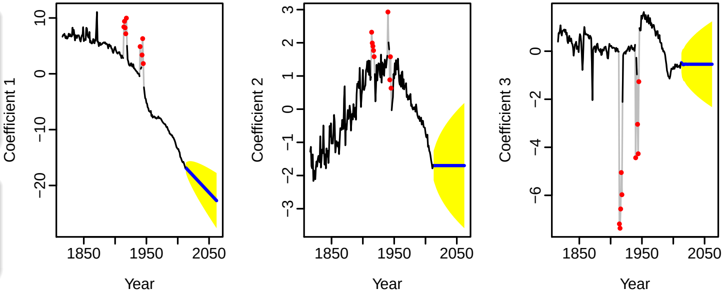
French male mortality

Functional principal components



French male mortality

ARIMA(2,1,0)+c
ARIMA(3,1,0)
ARIMA(4,1,0)



Forecasts

$$y_{t,x} = s_t(x) + \sigma_t(x)\varepsilon_{t,x},$$

$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

Forecasts

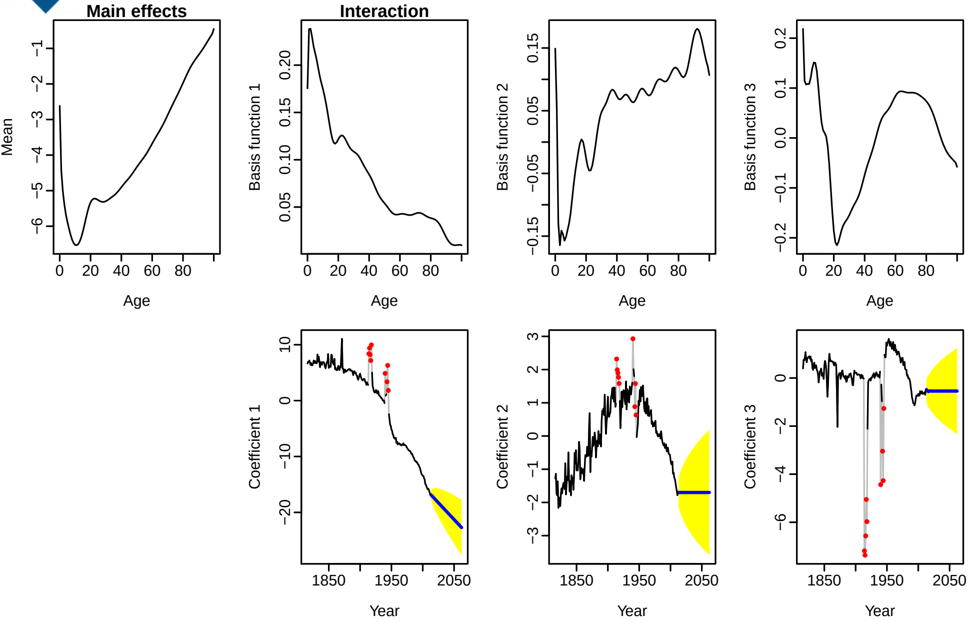
$$y_{t,x} = s_t(x) + \sigma_t(x)\varepsilon_{t,x},$$
$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

$$\mathrm{E}[y_{n+h,x} \mid \boldsymbol{y}] = \hat{\mu}(x) + \sum_{k=1}^{K} \hat{\beta}_{n+h,k}\,\hat{\phi}_k(x)$$

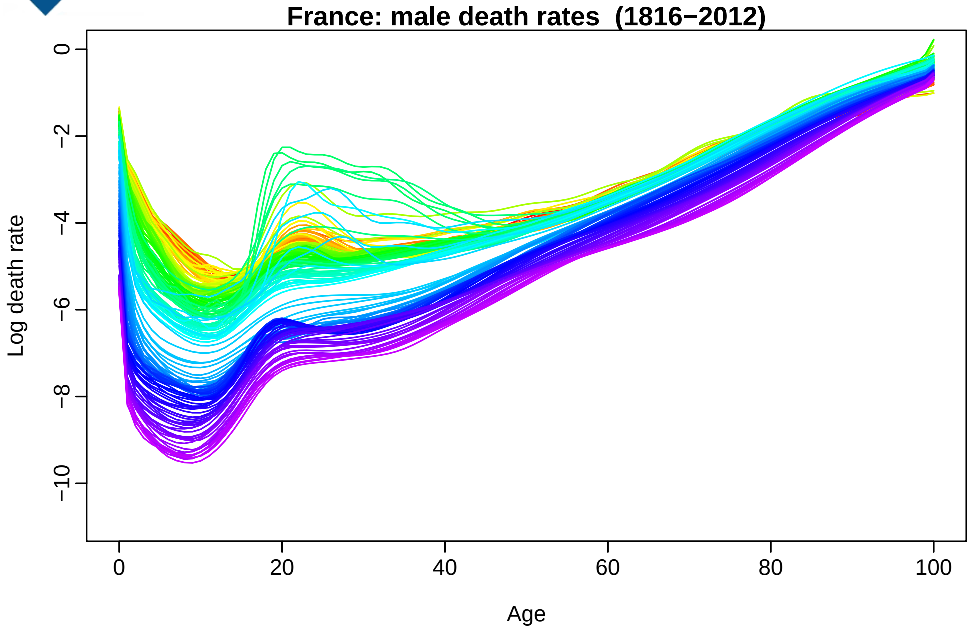
$$\mathrm{Var}[y_{n+h,x} \mid \boldsymbol{y}] = \hat{\sigma}_{\mu}^2(x) + \sum_{k=1}^{K} v_{n+h,k}\,\hat{\phi}_k^2(x) + \sigma_t^2(x) + v(x)$$

where $v_{n+h,k} = \mathrm{Var}(\beta_{n+h,k} \mid \beta_{1,k}, \dots, \beta_{n,k})$
$v(x) = \mathrm{Var}(r_t(x))$ and $\boldsymbol{y} = [y_{1,x}, \dots, y_{n,x}]$.

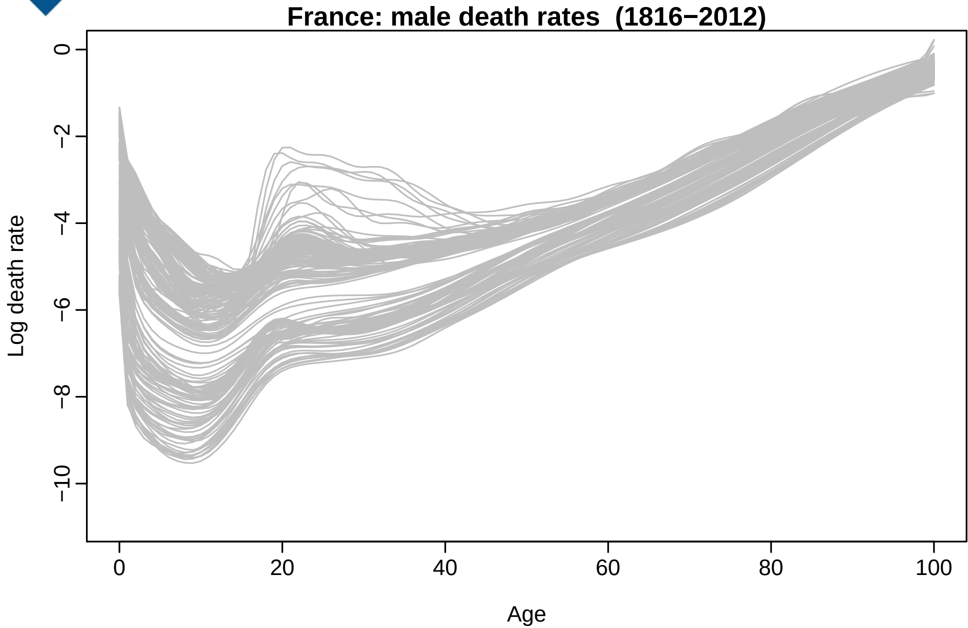
Forecasting the PC scores



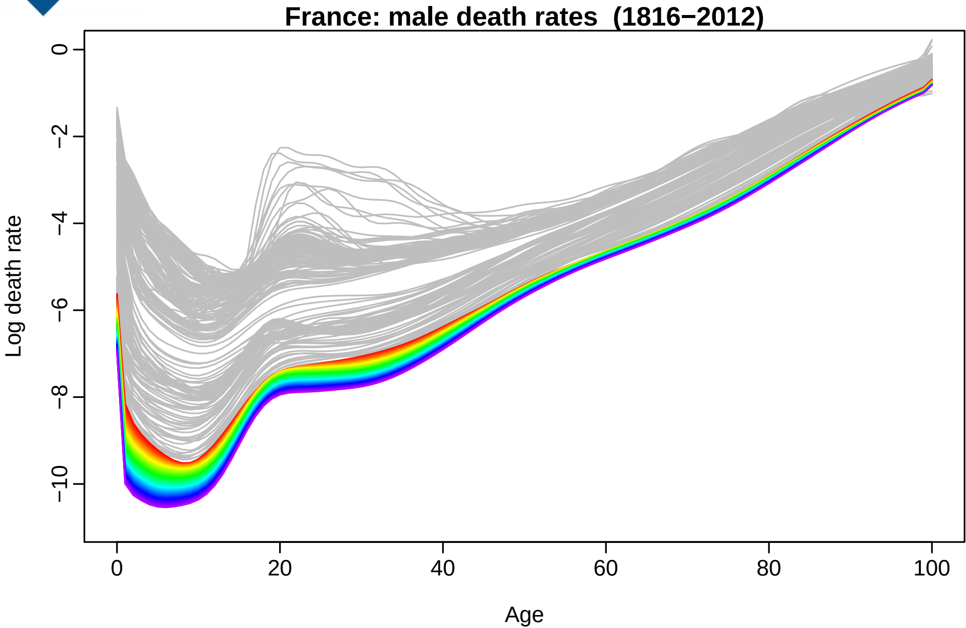
Forecasts of $s_t(x)$



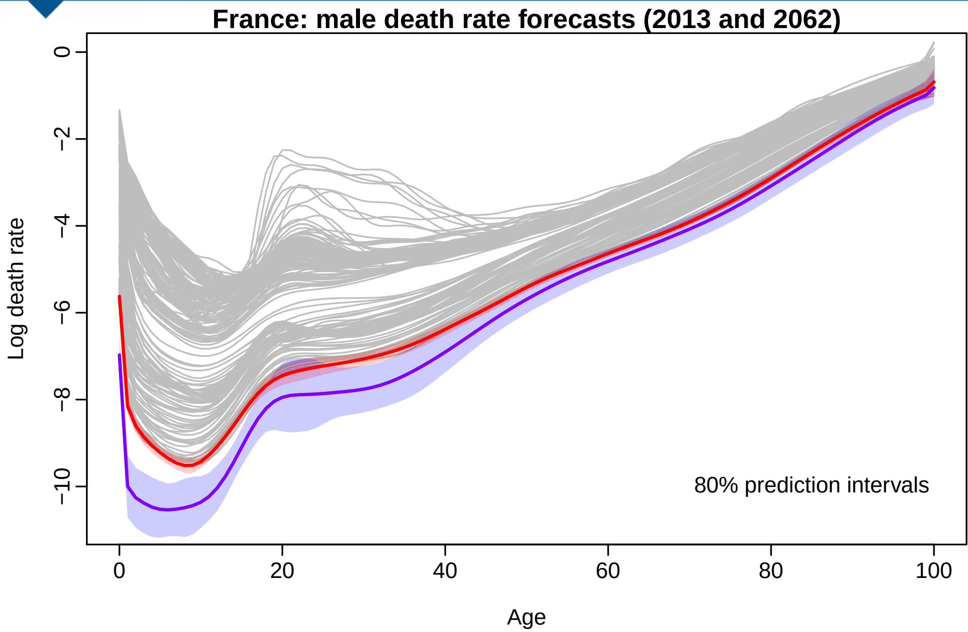
Forecasts of $s_t(x)$



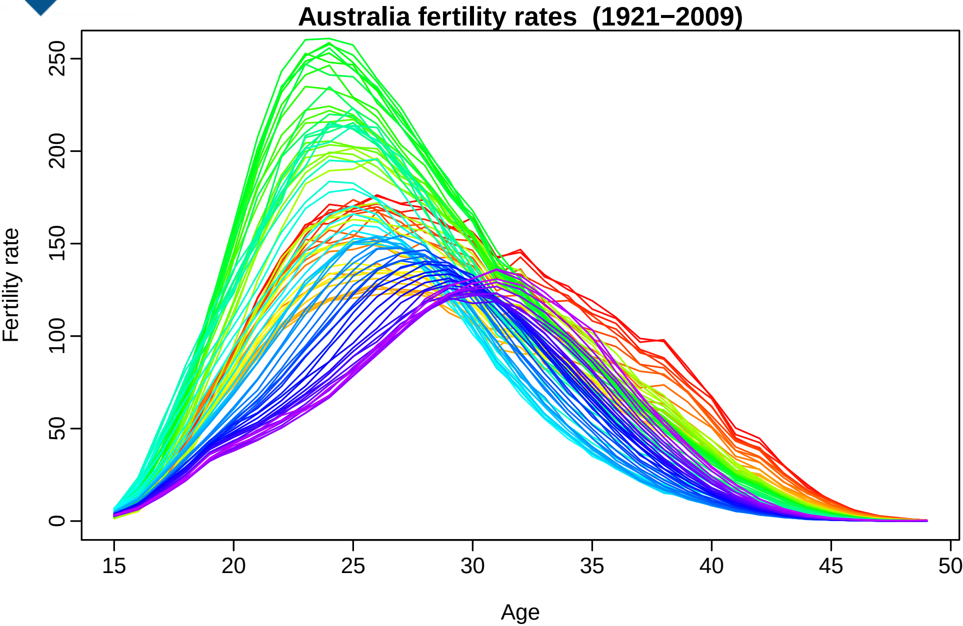
Forecasts of $s_t(x)$



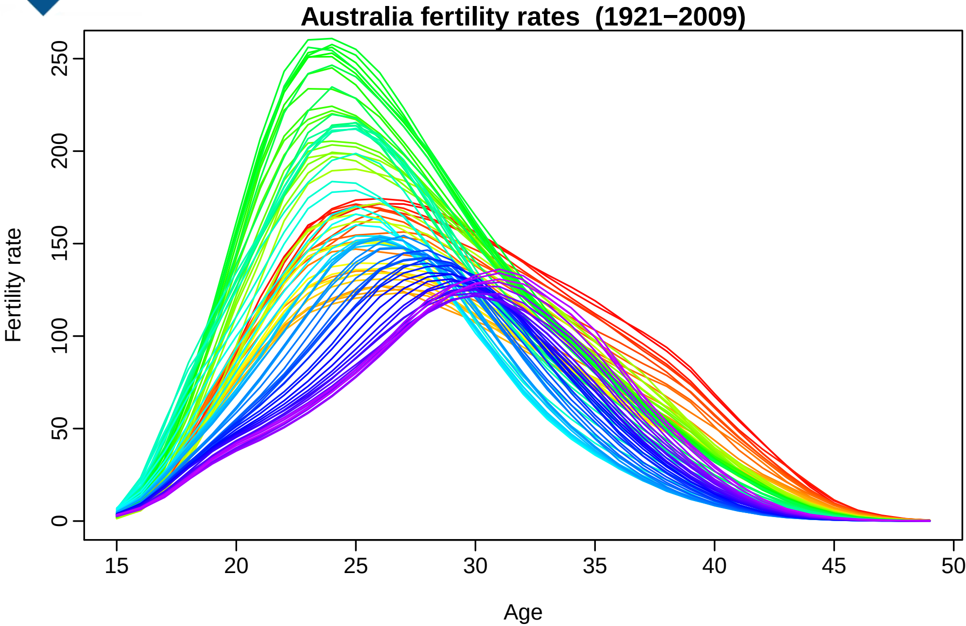
Forecasts of $s_t(x)$



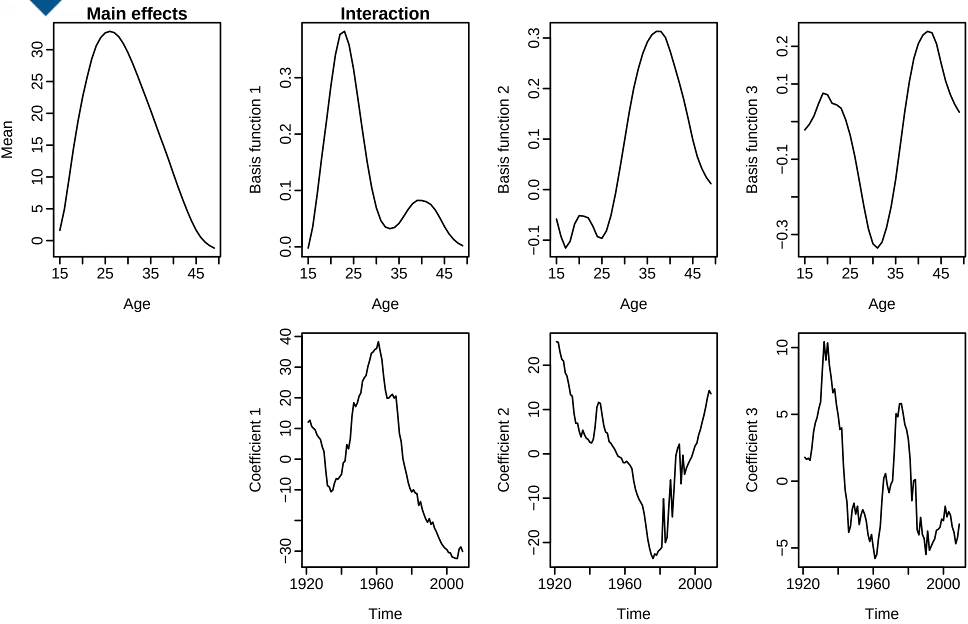
Australian fertility



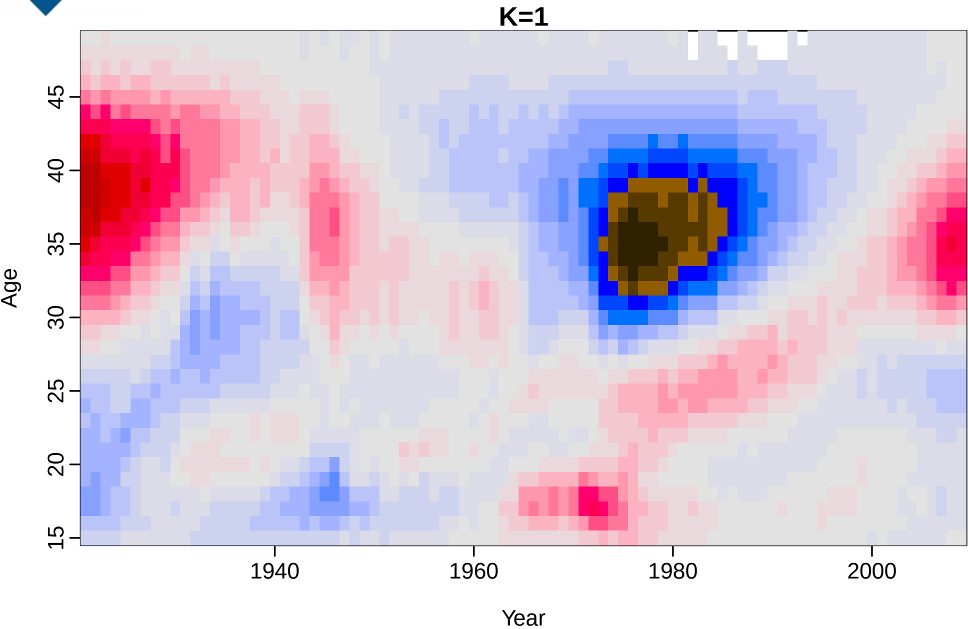
Australian fertility: smoothed



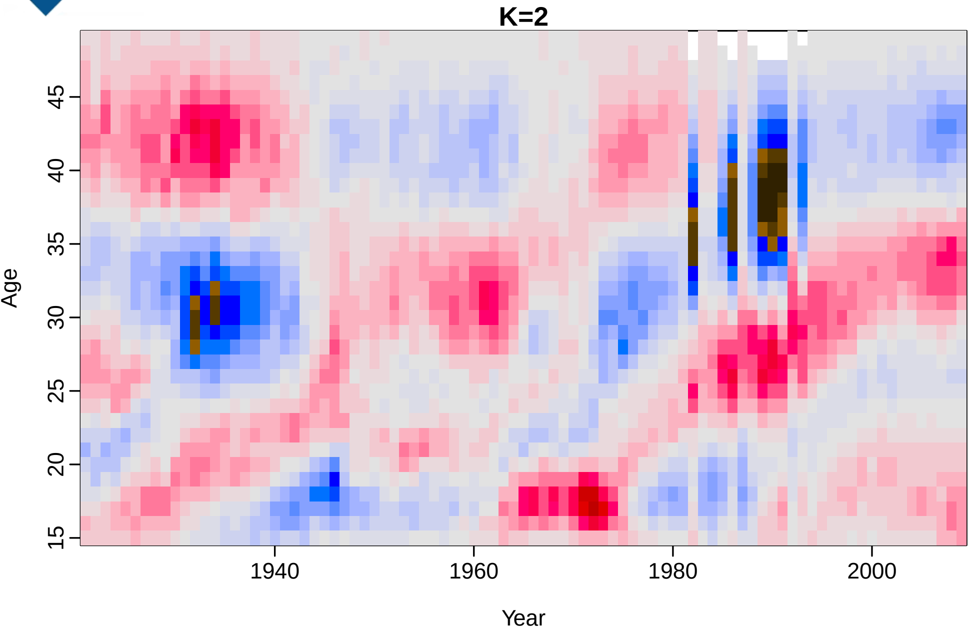
Australian fertility: PCA



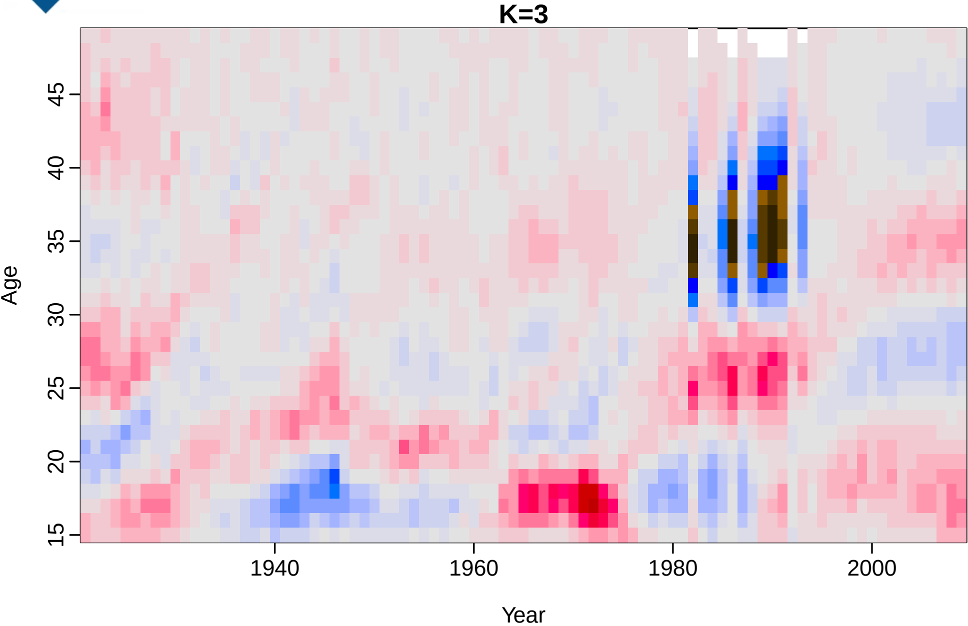
Australian fertility: residuals



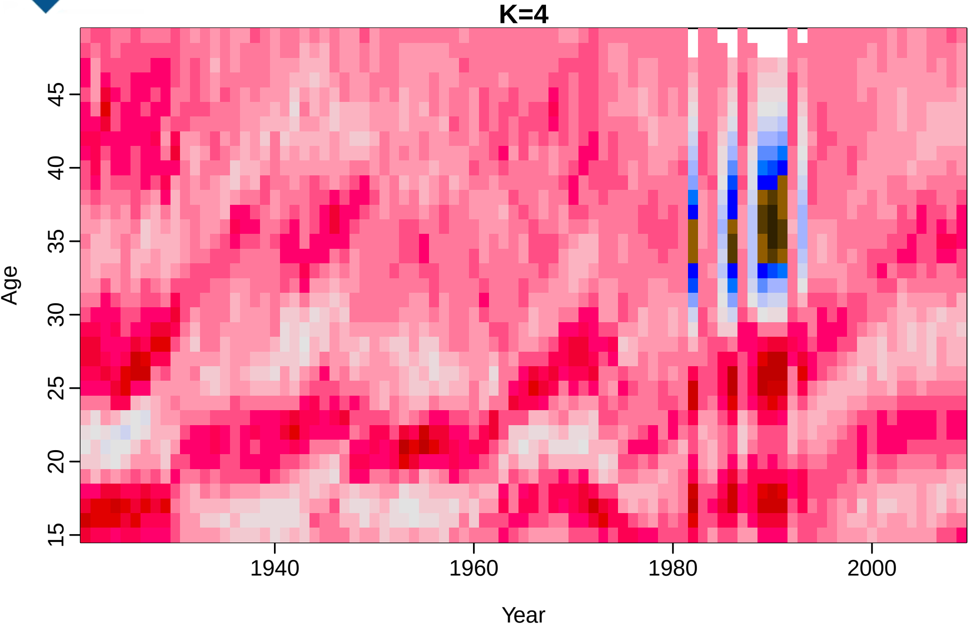
Australian fertility: residuals



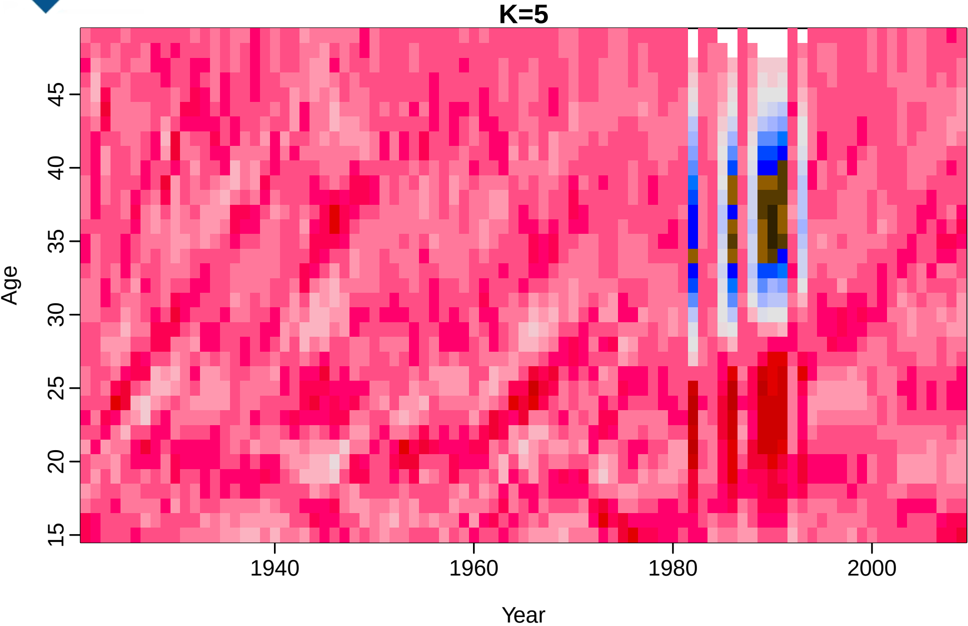
Australian fertility: residuals



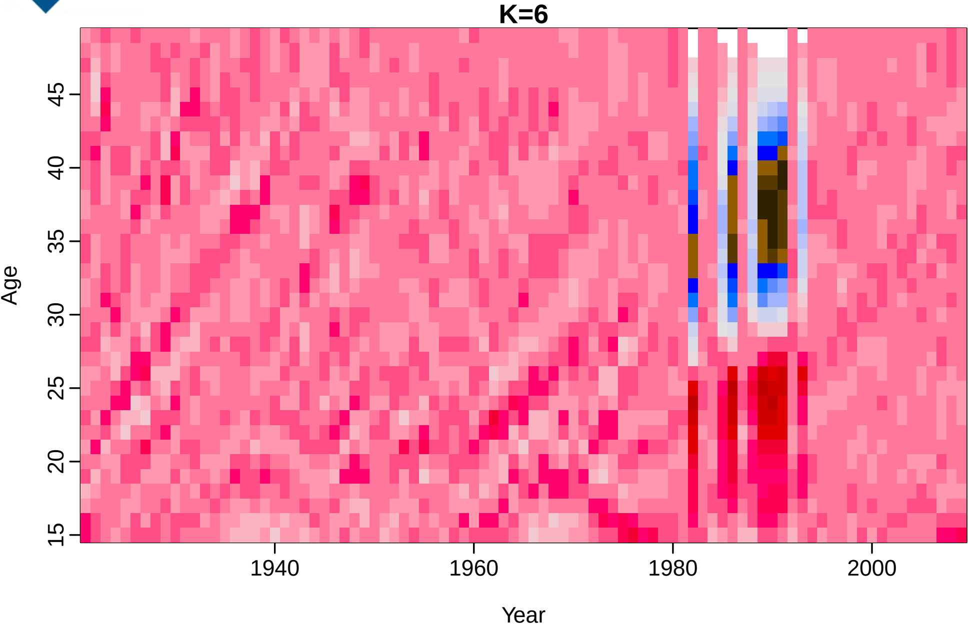
Australian fertility: residuals



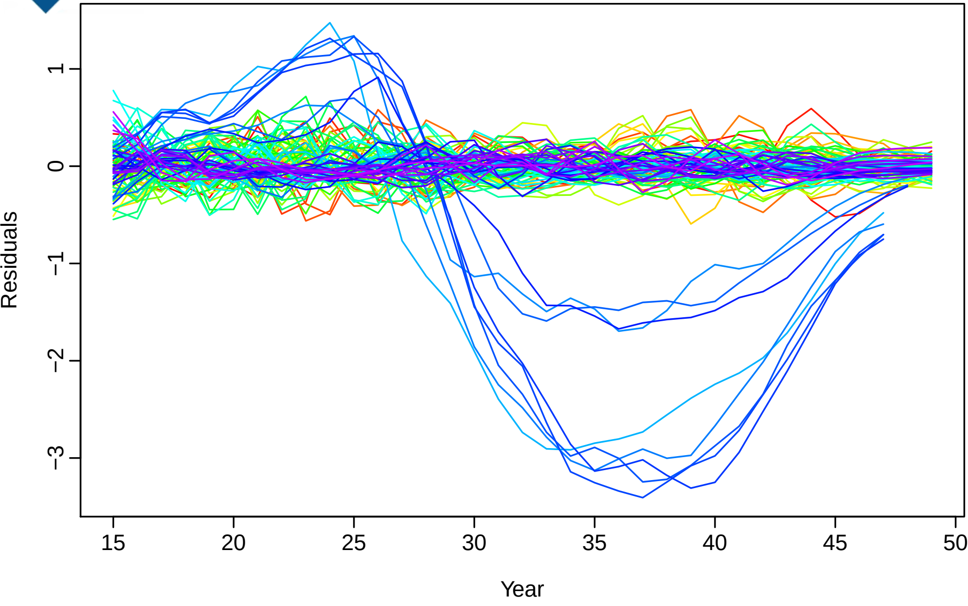
Australian fertility: residuals



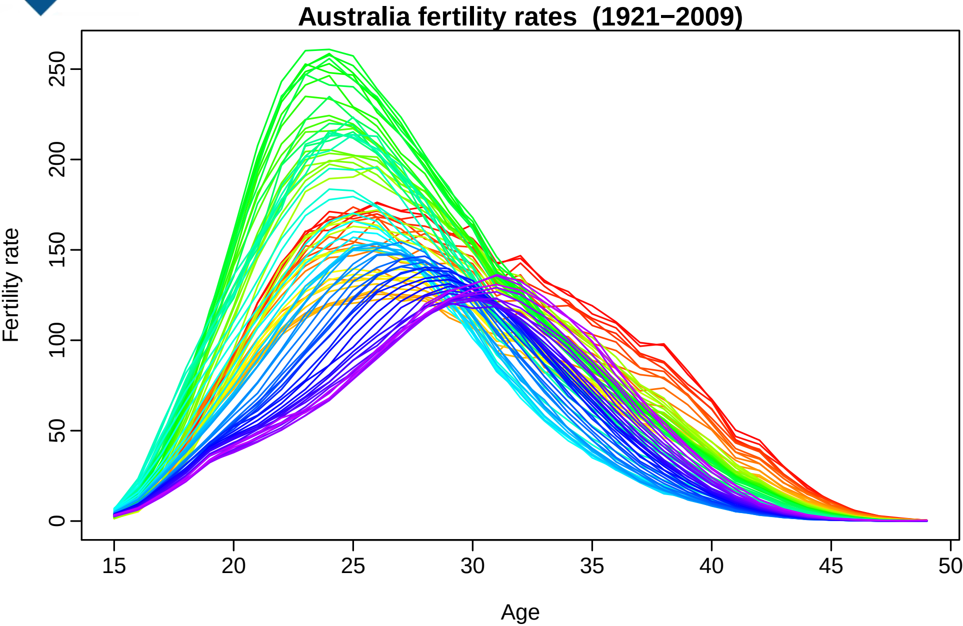
Australian fertility: residuals



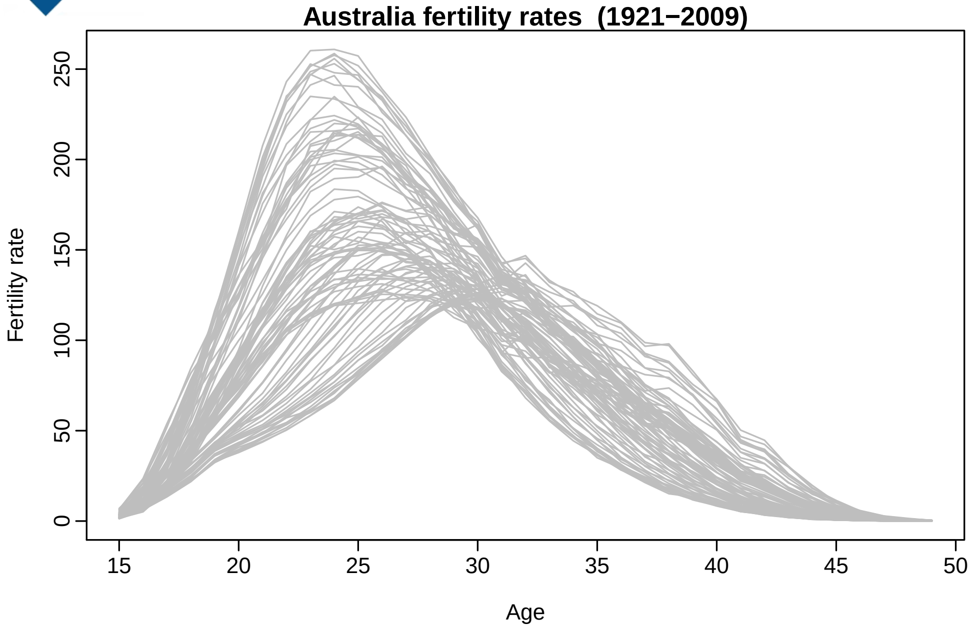
Australian fertility: residuals



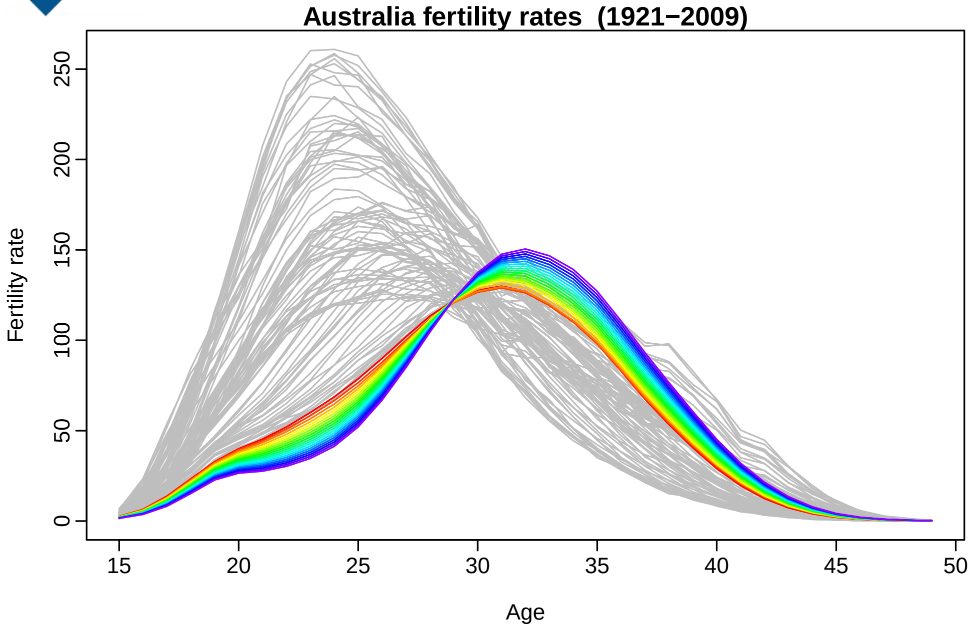
Australian fertility: forecasts



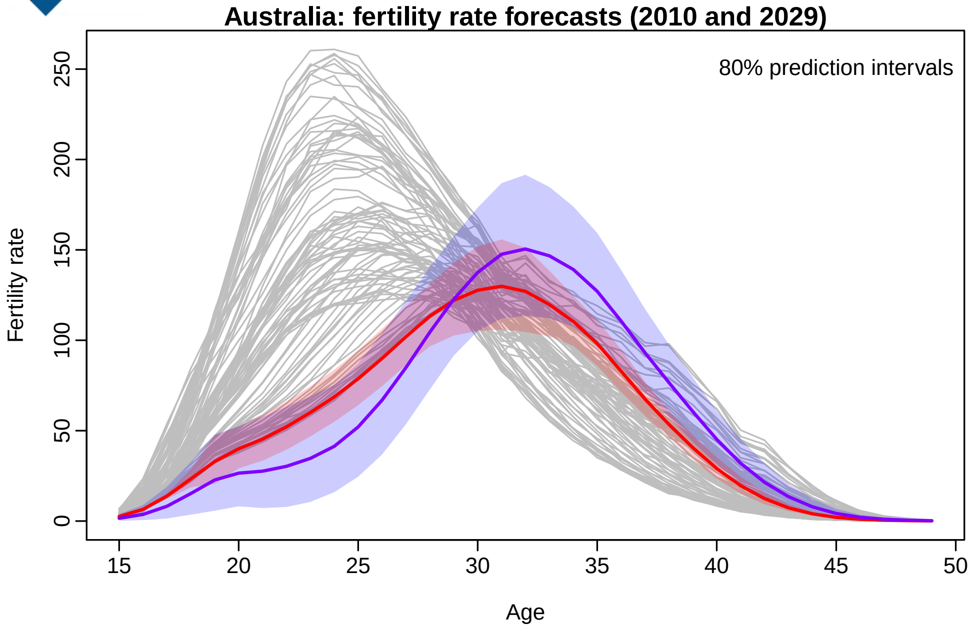
Australian fertility: forecasts



Australian fertility: forecasts



Australian fertility: forecasts



Outline

Life expectancy

$m(x)$ = mortality rate at age x.

Life expectancy at birth

$$e_0 = \int_0^\infty \exp\left[\int_0^x m(u)du\right] dx$$

- Approximated using life table methods.
- Iterate for $x = 0, 1, \dots,$, starting with $\ell_0 = 1$:

$q_x = m_x/(1+0.5m_x)$	Prob of death at age x
$d_x = \ell_x q_x$	Propn deaths at age x
$\ell_{x+1} = \ell_x - d_x$	Propn survive to age x
$L_x = \ell_x - 0.5d_x$	Propn survive to age $x+0.5$

Life expectancy

$m(x) =$ mortality rate at age x.

Life expectancy at birth

$$e_0 = \int_0^\infty \exp\left[\int_0^x m(u)du\right] dx$$

- Approximated using life table methods.
- Iterate for $x = 0, 1, \dots,$, starting with $\ell_0 = 1$:

$$\begin{aligned} q_x &= m_x/(1+0.5m_x) && \text{Prob of death at age } x \\ d_x &= \ell_x q_x && \text{Propn deaths at age } x \\ \ell_{x+1} &= \ell_x - d_x && \text{Propn survive to age } x \\ L_x &= \ell_x - 0.5d_x && \text{Propn survive to age } x+0.5 \end{aligned}$$

Life expectancy

$m(x)$ = mortality rate at age x.

Life expectancy at birth

$$e_0 = \int_0^\infty \exp\left[\int_0^x m(u)du\right] dx$$

- Approximated using life table methods.
- Iterate for $x = 0, 1, \dots,$, starting with $\ell_0 = 1$:

$$\begin{aligned} q_x &= m_x/(1+0.5m_x) && \text{Prob of death at age } x \\ d_x &= \ell_x q_x && \text{Propn deaths at age } x \\ \ell_{x+1} &= \ell_x - d_x && \text{Propn survive to age } x \\ L_x &= \ell_x - 0.5d_x && \text{Propn survive to age } x+0.5 \end{aligned}$$

Life expectancy

$m(x)$ = mortality rate at age x.

Life expectancy at birth

$$e_0 = \int_0^\infty \exp\left[\int_0^x m(u)du\right] dx$$

- Approximated using life table methods.
- Iterate for $x = 0, 1, \dots,$, starting with $\ell_0 = 1$:

Variations for $x = 0$ and upper age group.

$q_x = m_x/(1 + 0.5m_x)$	Prob of death at age x
$d_x = \ell_x q_x$	Propn deaths at age x
$\ell_{x+1} = \ell_x - d_x$	Propn survive to age x
$L_x = \ell_x - 0.5d_x$	Propn survive to age $x + 0.5$

Life expectancy

$m(x)$ = mortality rate at age x.

Life expectancy at birth

$$e_0 = \int_0^\infty \exp\left[\int_0^x m(u)du\right] dx$$

- Approximated using life table methods.
- Iterate for $x = 0, 1, \dots,$, starting with $\ell_0 = 1$:

Variations for $x = 0$ and upper age group.

$q_x = m_x/(1 + 0.5m_x)$	Prob of death at age x
$d_x = \ell_x q_x$	Propn deaths at age x
$\ell_{x+1} = \ell_x - d_x$	Propn survive to age x
$L_x = \ell_x - 0.5d_x$	Propn survive to age $x + 0.5$

Approximate life expectancy at birth

$$e_0 = \sum_{x=0}^{\infty} L_x$$

Life expectancy

$m(x)$ = mortality rate at age x.

Life expectancy at birth

$$e_0 = \int_0^\infty \exp\left[\int_0^x m(u)du\right] dx$$

- Approximated using life table methods.
- Iterate for $x = 0, 1, \dots,$, starting with $\ell_0 = 1$:

Variations for $x = 0$ and upper age group.

$$q_x = m_x/(1 + 0.5m_x) \quad \text{Prob of death at age } x$$

$$d_x = \ell_x q_x \quad \text{Propn deaths at age } x$$

$$\ell_{x+1} = \ell_x - d_x \quad \text{Propn survive to age } x$$

$$L_x = \ell_x - 0.5d_x \quad \text{Propn survive to age } x + 0.5$$

Approximate remaining life expectancy at age u

$$e_u = \sum_{x=u}^{\infty} L_x$$

Life expectancy forecasts

$$\log[m_t(x_i)] = s_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i},$$

$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

Period life expectancy

- Computed from $m_t(x)$ for a given t.
- Forecast $m_{T+h}(x)$.
- Compute $e_{0,T+h}$.
- Prediction intervals by simulation
  - $r_t(x)$ resampled
  - $\varepsilon_{t,i} \sim \mathrm{N}(0,1)$
  - $\beta_{t,k}$ simulated from ARIMA model

Life expectancy forecasts

$$\log[m_t(x_i)] = s_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i},$$

$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

Period life expectancy

- Computed from $m_t(x)$ for a given t.
- Forecast $m_{T+h}(x)$.
- Compute $e_{0,T+h}$.
- Prediction intervals by simulation
  - $r_t(x)$ resampled
  - $\varepsilon_{t,i} \sim \mathrm{N}(0,1)$
  - $\beta_{t,k}$ simulated from ARIMA model

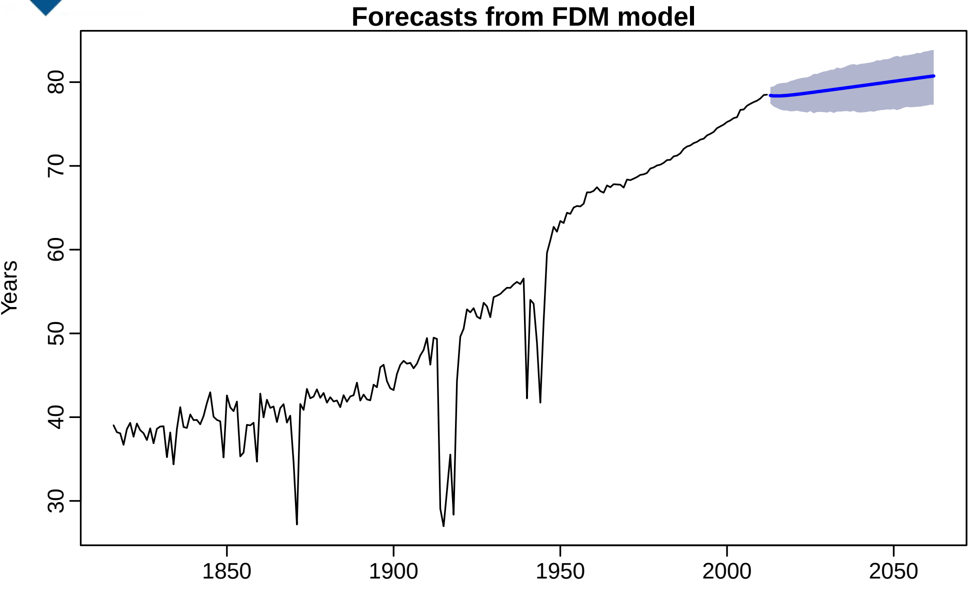
Life expectancy forecasts

$$\log[m_t(x_i)] = s_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i},$$
$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

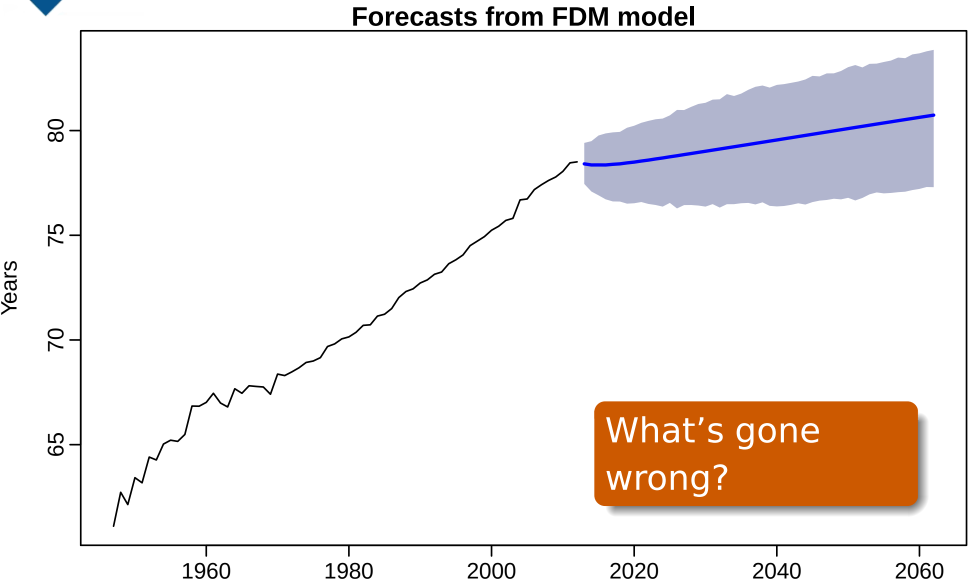
Period life expectancy

- Computed from $m_t(x)$ for a given t.
- Forecast $m_{T+h}(x)$.
- Compute $e_{0,T+h}$.
- Prediction intervals by simulation
  - $r_t(x)$ resampled
  - $\varepsilon_{t,i} \sim \mathrm{N}(0,1)$
  - $\beta_{t,k}$ simulated from ARIMA model

Life expectancy forecasts

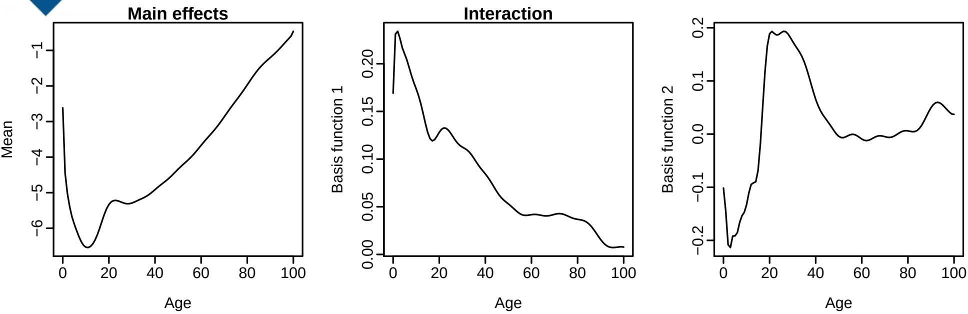


Life expectancy forecasts

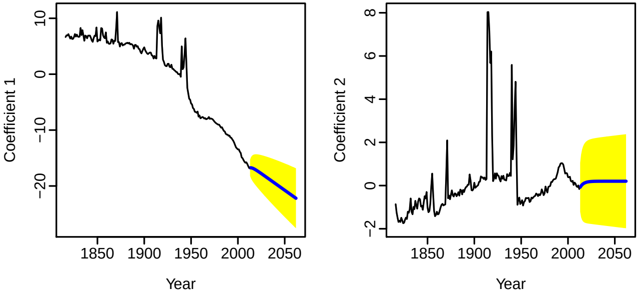


What's gone wrong?

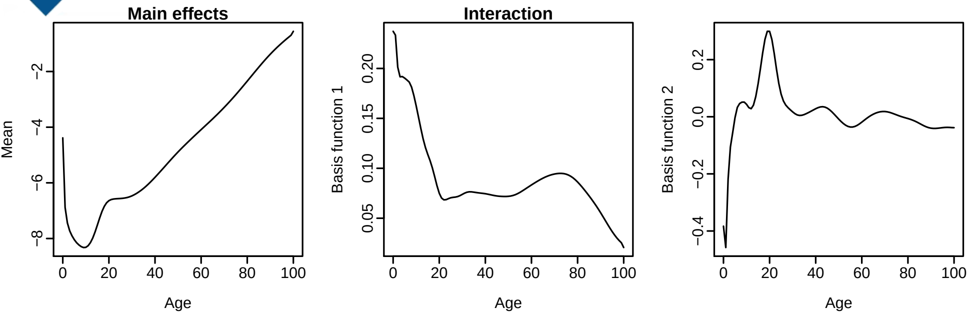
Mortality rate forecasts



French male mortality

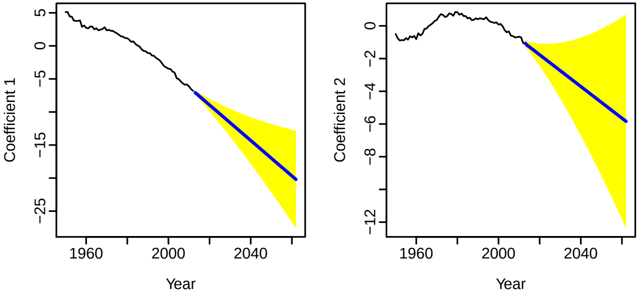


Mortality rate forecasts

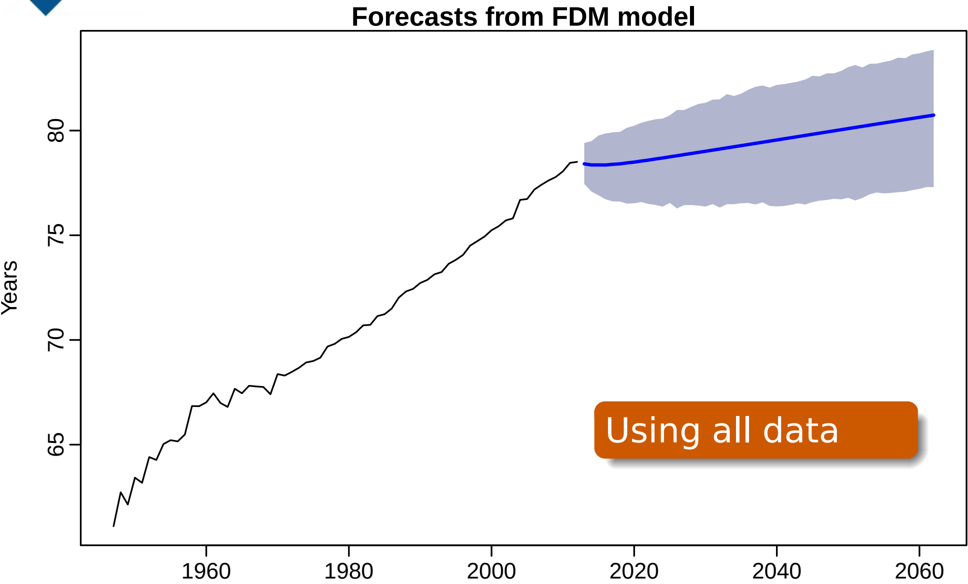


French male mortality

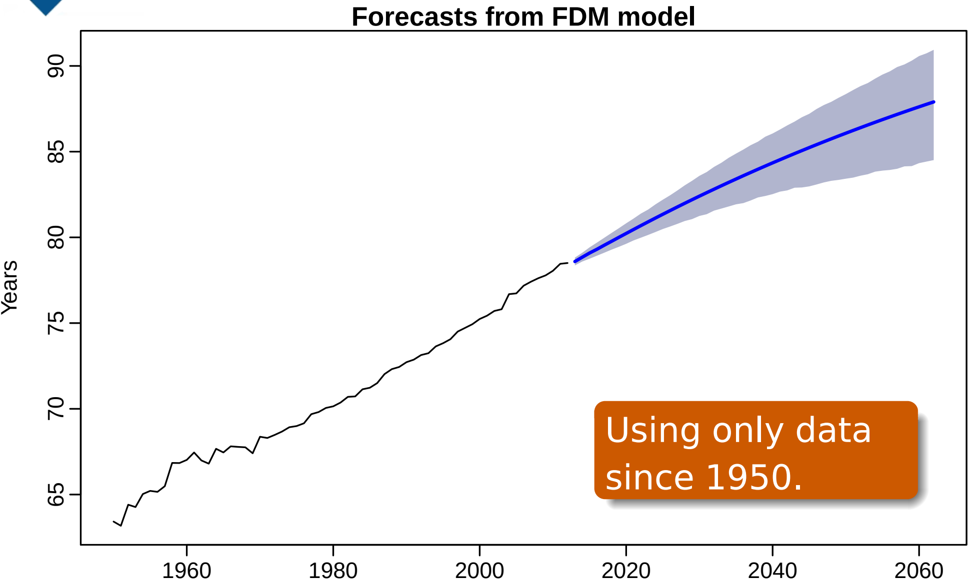
Using only data since 1950.



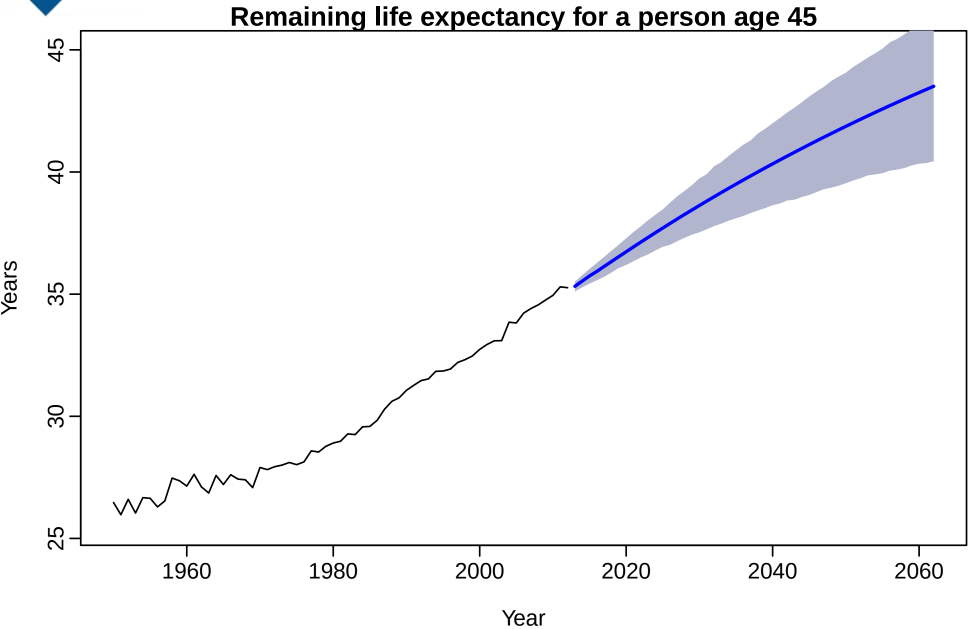
Life expectancy forecasts



Life expectancy forecasts



Life expectancy forecasts



Life expectancy forecasts

$$\log[m_t(x_i)] = s_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i},$$

$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

Cohort life expectancy

- Computed from $m_{s+x}(x)$ for a given s.
- Combine observed $m_{s+x}(x)$ where $s+x \leq T$ with forecast $m_{s+x}(x)$ for $s+x > T$.
- Compute $e^*_{0,s}$.
- Prediction intervals by simulation
  - $r_t(x)$ resampled
  - $\varepsilon_{t,i} \sim \mathrm{N}(0,1)$
  - $\beta_{t,k}$ simulated from ARIMA model

Life expectancy forecasts

$$\log[m_t(x_i)] = s_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i},$$

$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

Cohort life expectancy

- Computed from $m_{s+x}(x)$ for a given s.
- Combine observed $m_{s+x}(x)$ where $s + x \leq T$ with forecast $m_{s+x}(x)$ for $s + x > T$.
- Compute $e^*_{0,s}$.
- Prediction intervals by simulation
  - $r_t(x)$ resampled
  - $\varepsilon_{t,i} \sim \mathrm{N}(0, 1)$
  - $\beta_{t,k}$ simulated from ARIMA model

Life expectancy forecasts

$$\log[m_t(x_i)] = s_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i},$$

$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

Cohort life expectancy

- Computed from $m_{s+x}(x)$ for a given s.
- Combine observed $m_{s+x}(x)$ where $s + x \leq T$ with forecast $m_{s+x}(x)$ for $s + x > T$.
- Compute $e^*_{0,s}$.
- Prediction intervals by simulation
  - $r_t(x)$ resampled
  - $\varepsilon_{t,i} \sim \mathrm{N}(0, 1)$
  - $\beta_{t,k}$ simulated from ARIMA model

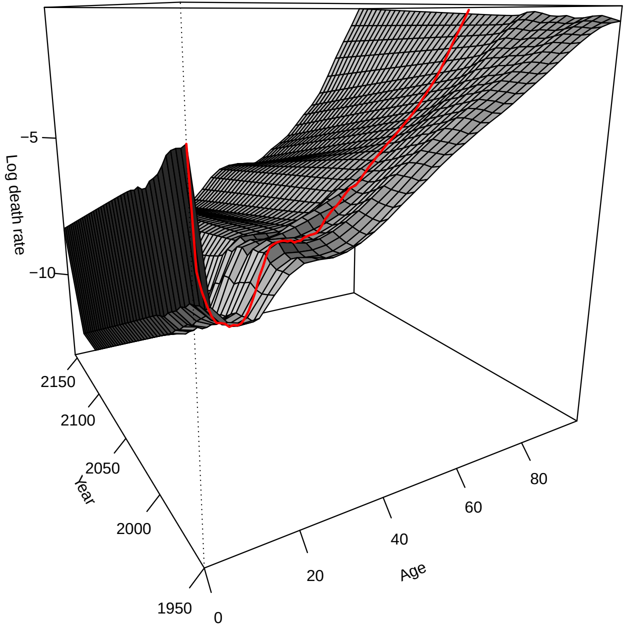
Life expectancy forecasts

$$\log[m_t(x_i)] = s_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i},$$
$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

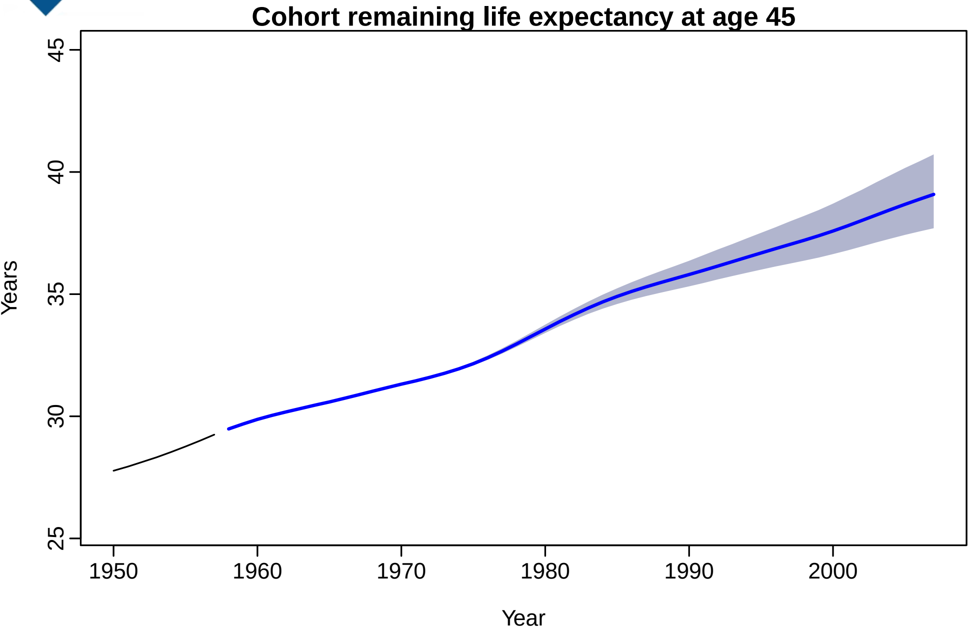
Cohort life expectancy

- Computed from $m_{s+x}(x)$ for a given s.
- Combine observed $m_{s+x}(x)$ where $s + x \leq T$ with forecast $m_{s+x}(x)$ for $s + x > T$.
- Compute $e^*_{0,s}$.
- Prediction intervals by simulation
  - $r_t(x)$ resampled
  - $\varepsilon_{t,i} \sim \mathrm{N}(0, 1)$
  - $\beta_{t,k}$ simulated from ARIMA model

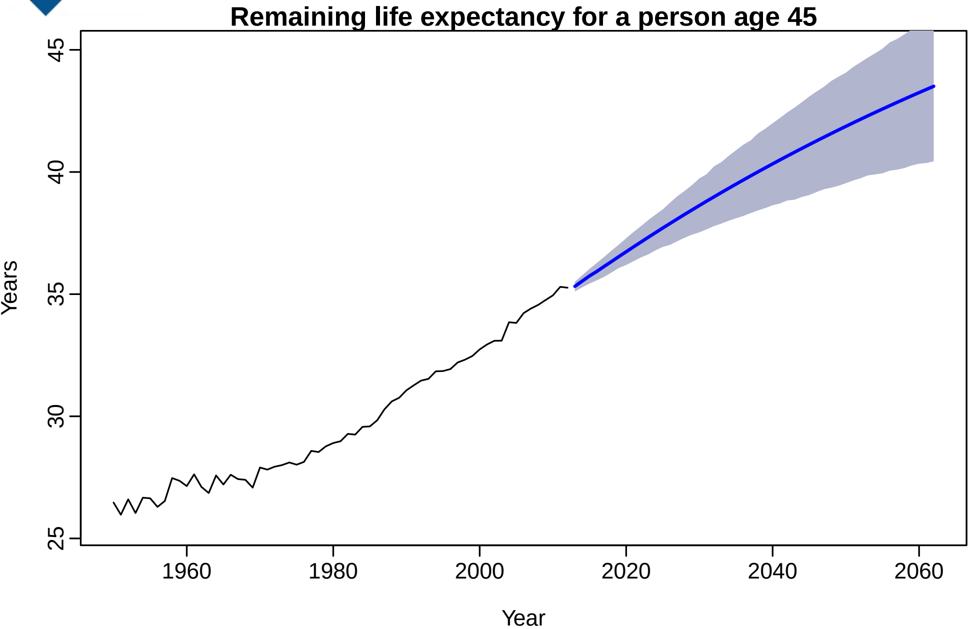
Cohort life expectancy



Cohort life expectancy forecast



Period life expectancy forecast



Outline

Exponentially weighted functional PCA

$$y_{t,x} = s_t(x) + \sigma_t(x)\varepsilon_{t,x},$$
$$s_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + r_t(x)$$

1. Exponentially decreasing weights:
$$w_t = \kappa(1-\kappa)^{T-t}, \qquad 0 < \kappa < 1.$$
2. $\hat{\mu}(x) = \sum_{t=1}^{T} w_t s_t(x)$
3. The function $\phi_k(x)$ which minimizes
$$\text{MISE} = \frac{1}{T}\sum_{t=1}^{T} w_t \int r_t(x)^2 dx$$
is the kth principal component
(computed recursively, $k = 1, 2, \dots$).

Exponentially weighted functional PCA

Computationally equivalent approach

- $\boldsymbol{W} = \text{diagonal}(w_1, \dots, w_T), \quad w_t = \kappa(1-\kappa)^{T-t}$
- Discretize $s_t^*(x) = s_t(x) - \hat{\mu}(x)$ on a dense grid of q equally spaced points.
- Denote discretized $s_t^*(x)$ as $T \times q$ matrix $\boldsymbol{G}^*$ and let $\boldsymbol{G} = \boldsymbol{W}\boldsymbol{G}^*$.
- SVD of $\boldsymbol{G} = \Phi\Lambda\boldsymbol{V}'$ where $\phi_k(x)$ is kth column of Φ.
- $\beta_{t,k}$ is (t,k)th element of $\boldsymbol{G}\Phi$.

Exponentially weighted functional PCA

Computationally equivalent approach

- $\boldsymbol{W} = \text{diagonal}(w_1, \dots, w_T), \quad w_t = \kappa(1-\kappa)^{T-t}$
- Discretize $s_t^*(x) = s_t(x) - \hat{\mu}(x)$ on a dense grid of q equally spaced points.
- Denote discretized $s_t^*(x)$ as $T \times q$ matrix $\boldsymbol{G}^*$ and let $\boldsymbol{G} = \boldsymbol{W}\boldsymbol{G}^*$.
- SVD of $\boldsymbol{G} = \Phi\Lambda\boldsymbol{V}'$ where $\phi_k(x)$ is kth column of Φ.
- $\beta_{t,k}$ is (t, k)th element of $\boldsymbol{G}\Phi$.

Exponentially weighted functional PCA

Computationally equivalent approach

- $\boldsymbol{W} = \text{diagonal}(w_1, \dots, w_T)$, $\quad w_t = \kappa(1-\kappa)^{T-t}$
- Discretize $s_t^*(x) = s_t(x) - \hat{\mu}(x)$ on a dense grid of q equally spaced points.
- Denote discretized $s_t^*(x)$ as $T \times q$ matrix $\boldsymbol{G}^*$ and let $\boldsymbol{G} = \boldsymbol{W}\boldsymbol{G}^*$.
- SVD of $\boldsymbol{G} = \Phi\Lambda\boldsymbol{V}'$ where $\phi_k(x)$ is kth column of Φ.
- $\beta_{t,k}$ is (t,k)th element of $\boldsymbol{G}\Phi$.

Exponentially weighted functional PCA

Computationally equivalent approach

- $\boldsymbol{W} = \text{diagonal}(w_1, \dots, w_T)$, $\quad w_t = \kappa(1-\kappa)^{T-t}$
- Discretize $s_t^*(x) = s_t(x) - \hat{\mu}(x)$ on a dense grid of q equally spaced points.
- Denote discretized $s_t^*(x)$ as $T \times q$ matrix $\boldsymbol{G}^*$ and let $\boldsymbol{G} = \boldsymbol{W}\boldsymbol{G}^*$.
- SVD of $\boldsymbol{G} = \Phi\Lambda\boldsymbol{V}'$ where $\phi_k(x)$ is kth column of Φ.
- $\beta_{t,k}$ is (t,k)th element of $\boldsymbol{G}\Phi$.

Exponentially weighted functional PCA

Computationally equivalent approach

- $\boldsymbol{W} = \text{diagonal}(w_1, \dots, w_T)$, $\quad w_t = \kappa(1-\kappa)^{T-t}$
- Discretize $s_t^*(x) = s_t(x) - \hat{\mu}(x)$ on a dense grid of q equally spaced points.
- Denote discretized $s_t^*(x)$ as $T \times q$ matrix $\boldsymbol{G}^*$ and let $\boldsymbol{G} = \boldsymbol{W}\boldsymbol{G}^*$.
- SVD of $\boldsymbol{G} = \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{V}'$ where $\phi_k(x)$ is kth column of $\boldsymbol{\Phi}$.
- $\beta_{t,k}$ is (t,k)th element of $\boldsymbol{G}\boldsymbol{\Phi}$.

Outline

Alternative approaches

1. **Hyndman-Ullah (2007) methods**
   - **HU**: as described
   - **HU50**: only data from 1950
   - **HUrob**: robust PCA
   - **HUrob50**: robust PCA and only data from 1950
   - **HUw**: exponentially weighted PCA
2. **Lee-Carter (1992) methods**
   - **LCnone**: $K = 1$, no smoothing, random walk + drift model.
   - **LC**: LCnone + adjusted $\beta_{t,1}$ to number of deaths
   - **TLB** (Tuljapurkar-Li-Boe, 2000): LC + data from 1950.
   - **LM** (Lee-Miller 2001): LCnone + data from 1950, adjusted $\beta_{t,1}$ to life expectancy, bias adjustment.
   - **BMS** (Booth-Maindonald-Smith, 2002): LCnone + fitting period determined from data, adjusted $\beta_{t,1}$ using Poisson GLM, bias adjustment.

Alternative approaches

1. **Hyndman-Ullah (2007) methods**
   - **HU**: as described
   - **HU50**: only data from 1950
   - **HUrob**: robust PCA
   - **HUrob50**: robust PCA and only data from 1950
   - **HUw**: exponentially weighted PCA
2. **Lee-Carter (1992) methods**
   - **LCnone**: $K = 1$, no smoothing, random walk + drift model.
   - **LC**: LCnone + adjusted $\beta_{t,1}$ to number of deaths
   - **TLB** (Tuljapurkar-Li-Boe, 2000): LC + data from 1950.
   - **LM** (Lee-Miller 2001): LCnone + data from 1950, adjusted $\beta_{t,1}$ to life expectancy, bias adjustment.
   - **BMS** (Booth-Maindonald-Smith, 2002): LCnone + fitting period determined from data, adjusted $\beta_{t,1}$ using Poisson GLM, bias adjustment.

Human Mortality Database

www.mortality.org

HMD Main Menu

Registration
- New User
- Change Password
- User's Agreement

Project
- FAQ
- Overview
- History

People
- Acknowledgements
- Research Teams
- HMD Publications

Methods
- Brief Summary
- Full Protocol
- Special Methods

Data
- What's New
- Explanatory Notes
- Data Availability
- Zipped Data Files
- Citation Guidelines

Links
- Max Planck Institute
- UC Berkeley
- UC Berkeley Demography
- Human Life Table Database
- Canadian HMD

General
- Contact us

The Human Mortality Database

John R. Wilmoth, *Director*	University of California, Berkeley
Vladimir Shkolnikov, *Co-Director*	Max Planck Institute for Demographic Research
Magali Barbieri, *Associate Director*	University of California, Berkeley and INED, Paris

The Human Mortality Database (HMD) was created to provide detailed mortality and population data to researchers, students, journalists, policy analysts, and others interested in the history of human longevity. The project began as an outgrowth of earlier projects in the Department of Demography at the University of California, Berkeley, USA, and at the Max Planck Institute for Demographic Research in Rostock, Germany (see history). It is the work of two teams of researchers in the USA and Germany (see research teams), with the help of financial backers and scientific collaborators from around the world (see acknowledgements).

We seek to provide open, international access to these data. At present the database contains detailed population and mortality data for the following 37 countries or areas:

Australia	Finland	Lithuania	Spain
Austria	France	Luxembourg	Sweden
Belarus	Germany	Netherlands	Switzerland
Belgium	Hungary	New Zealand	Taiwan
Bulgaria	Iceland	Norway	U.K.
Canada	Ireland	Poland	U.S.A.
Chile	Israel	Portugal	Ukraine
Czech Republic	Italy	Russia	
Denmark	Japan	Slovakia	
Estonia	Latvia	Slovenia	

Human Mortality Database

www.mortality.org

www.mortality.org

HMD

...an Mortality Database

Registra...
- New User
- Change Password
- User's Agreement

Project
- FAQ
- Overview
- History

People
- Acknowledgements
- Research Teams
- HMD Publications

Methods
- Brief Summary
- Full Protocol
- Special Methods

Data
- What’s New
- Explanatory Notes
- Data Availability
- Zipped Data Files
- Citation Guidelines

Links
- Max Planck Institute
- UC Berkeley
- UC Berkeley Demography
- Human Life Table Database
- Canadian HMD

General
- Contact us

John R. Wilmoth, ***Director***	University of California, Berkeley
Vladimir Shkolnikov, ***Co-Director***	Max Planck Institute for Demographic Research
Magali Barbieri, ***Associate Director***	University of California, Berkeley and INED, Paris

The Human Mortality Database (HMD) was created to provide detailed mortality and population data to researchers, students, journalists, policy analysts, and others interested in the history of human longevity. The project began as an outgrowth of earlier projects in the Department of Demography at the University of California, Berkeley, USA, and at the Max Planck Institute for Demographic Research in Rostock, Germany (see history). It is the work of two teams of researchers in the USA and Germany (see research teams), with the help of financial backers and scientific collaborators from around the world (see acknowledgements).

We seek to provide open, international access to these data. At present the database contains detailed population and mortality data for the following 37 countries or areas:

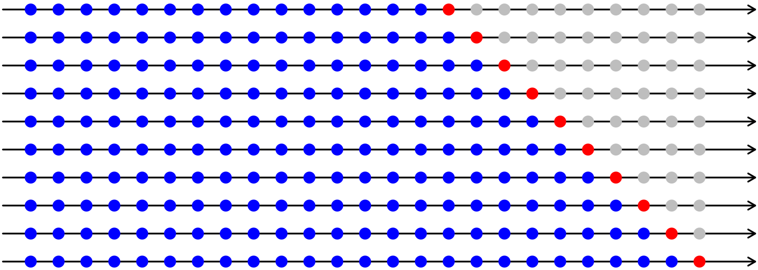
Australia	Finland	Lithuania	Spain
Austria	France	Luxembourg	Sweden
Belarus	Germany	Netherlands	Switzerland
Belgium	Hungary	New Zealand	Taiwan
Bulgaria	Iceland	Norway	U.K.
Canada	Ireland	Poland	U.S.A.
Chile	Israel	Portugal	Ukraine
Czech Republic	Italy	Russia	
Denmark	Japan	Slovakia	
Estonia	Latvia	Slovenia	

First year of training data

Country	LC	LCnone	TLB	LM	BMS[f]	BMS[m]	HU	HU50	HUrob	HUrob50	HUw
Australia	1921	1921	1950	1950	1953	1954	1921	1950	1921	1950	1921
Canada	1921	1921	1950	1950	1952	1948	1921	1950	1921	1950	1921
Denmark	1835	1835	1950	1950	1948	1948	1835	1950	1835	1950	1835
England	1841	1841	1950	1950	1952	1948	1841	1950	1841	1950	1841
Finland	1878	1878	1950	1950	1954	1954	1878	1950	1878	1950	1878
France	1816	1816	1950	1950	1947	1954	1816	1950	1816	1950	1816
Iceland	1838	1838	1950	1950	1838	1838	1838	1950	1838	1950	1838
Italy	1872	1872	1950	1950	1954	1954	1872	1950	1872	1950	1872
Netherlands	1850	1850	1950	1950	1946	1947	1850	1950	1850	1950	1850
Norway	1846	1846	1950	1950	1951	1948	1846	1950	1846	1950	1846
Scotland	1855	1855	1950	1950	1936	1948	1855	1950	1855	1950	1855
Spain	1908	1908	1950	1950	1952	1952	1908	1950	1908	1950	1908
Sweden	1751	1751	1950	1950	1948	1952	1751	1950	1751	1950	1751
Switzerland	1876	1876	1950	1950	1950	1950	1876	1950	1876	1950	1876

Evaluation design

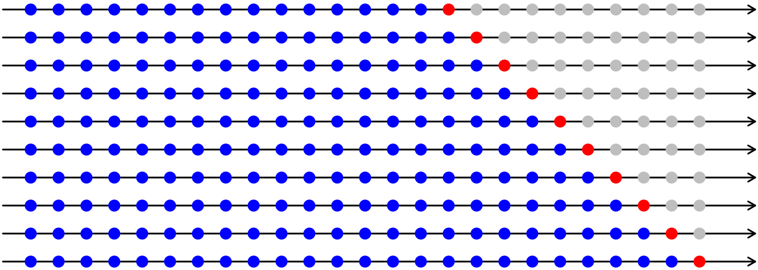
- Rolling forecast origin. Initial training set: 30 years



- Compute one-step forecasts for each training set, and compare to test data.
- Repeat until training period ends in 2003.
- Average MAE weighted by population size.

Evaluation design

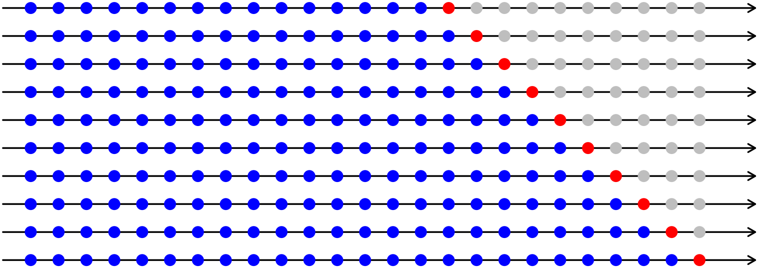
- Rolling forecast origin. Initial training set: 30 years



- Compute one-step forecasts for each training set, and compare to test data.
- Repeat until training period ends in 2003.
- Average MAE weighted by population size.

Evaluation design

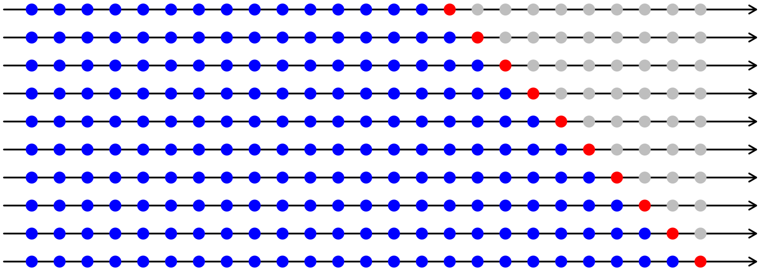
- Rolling forecast origin. Initial training set: 30 years



- Compute one-step forecasts for each training set, and compare to test data.
- Repeat until training period ends in 2003.
- Average MAE weighted by population size.

Evaluation design

- Rolling forecast origin. Initial training set: 30 years



- Compute one-step forecasts for each training set, and compare to test data.
- Repeat until training period ends in 2003.
- Average MAE weighted by population size.

Evaluation: male log mortality

One-step-ahead MAE

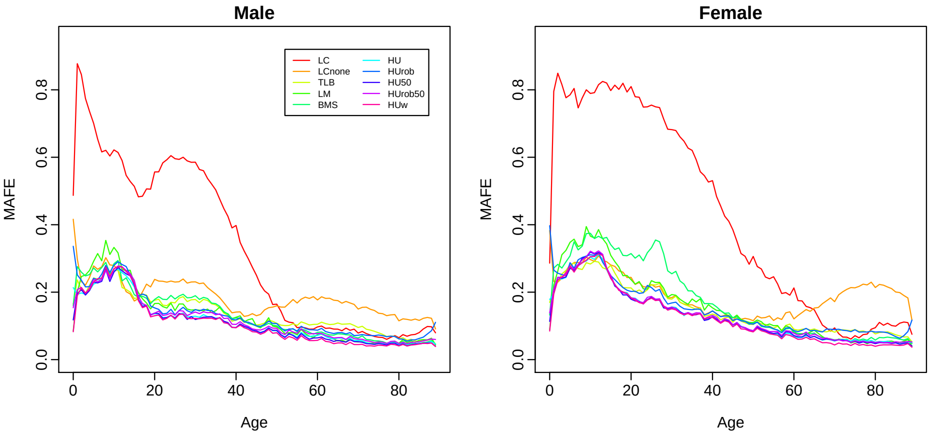
Country	LC	LCnone	TLB	LM	BMS	HU	HU50	HUrob	HUrob50	HUw
Australia	0.388	0.183	0.110	0.092	0.093	0.079	0.079	0.105	0.092	0.076
Canada	0.246	0.133	0.088	0.069	0.079	0.060	0.061	0.090	0.077	0.057
Denmark	0.185	0.180	0.152	0.157	0.145	0.127	0.127	0.143	0.136	0.123
England	0.545	0.175	0.085	0.058	0.088	0.062	0.054	0.083	0.063	0.052
Finland	0.445	0.211	0.150	0.156	0.140	0.141	0.133	0.147	0.137	0.131
France	0.450	0.180	0.083	0.054	0.102	0.064	0.050	0.115	0.061	0.050
Iceland	0.328	0.326	0.328	0.407	0.327	0.332	0.350	0.337	0.344	0.335
Italy	0.283	0.168	0.111	0.065	0.111	0.075	0.062	0.113	0.083	0.061
Netherlands	0.148	0.133	0.110	0.093	0.111	0.081	0.078	0.095	0.084	0.078
Norway	0.235	0.171	0.151	0.150	0.140	0.118	0.120	0.133	0.127	0.120
Scotland	0.649	0.215	0.141	0.152	0.131	0.124	0.121	0.130	0.121	0.118
Spain	0.243	0.158	0.112	0.067	0.113	0.065	0.064	0.072	0.075	0.058
Sweden	0.191	0.171	0.142	0.140	0.137	0.118	0.115	0.146	0.123	0.114
Switzerland	0.225	0.183	0.131	0.136	0.131	0.116	0.113	0.132	0.122	0.112
Weighted ave	0.351	0.168	0.103	0.075	0.105	0.074	0.067	0.102	0.080	**0.065**

Evaluation: female log mortality

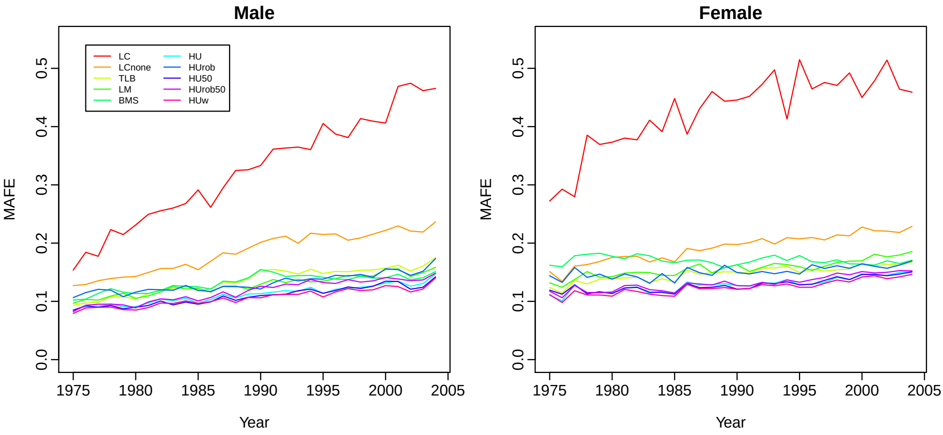
One-step-ahead MAE

Country	LC	LCnone	TLB	LM	BMS	HU	HU50	HUrob	HUrob50	HUw
Australia	0.288	0.179	0.104	0.114	0.103	0.091	0.091	0.104	0.095	0.089
Canada	0.201	0.128	0.074	0.084	0.075	0.069	0.072	0.073	0.073	0.068
Denmark	0.508	0.203	0.175	0.201	0.177	0.159	0.156	0.183	0.161	0.152
England	0.372	0.119	0.083	0.071	0.069	0.066	0.063	0.083	0.069	0.057
Finland	0.636	0.254	0.193	0.213	0.224	0.171	0.173	0.184	0.174	0.168
France	0.517	0.168	0.081	0.066	0.131	0.058	0.059	0.109	0.063	0.055
Iceland	0.381	0.367	0.375	0.414	0.434	0.353	0.358	0.354	0.355	0.343
Italy	0.363	0.138	0.095	0.078	0.098	0.072	0.068	0.103	0.075	0.066
Netherlands	0.377	0.159	0.101	0.114	0.110	0.091	0.089	0.099	0.091	0.088
Norway	0.564	0.204	0.169	0.189	0.199	0.152	0.153	0.180	0.156	0.151
Scotland	0.526	0.203	0.176	0.192	0.166	0.150	0.154	0.174	0.155	0.145
Spain	0.437	0.162	0.130	0.083	0.140	0.072	0.073	0.084	0.082	0.068
Sweden	0.389	0.176	0.145	0.181	0.285	0.147	0.138	0.223	0.141	0.139
Switzerland	0.400	0.232	0.166	0.189	0.172	0.146	0.148	0.157	0.150	0.145
Weighted ave	0.398	0.155	0.102	0.094	0.117	0.080	0.079	0.105	0.084	**0.075**

Mean Absolute Forecast Error



Mean Absolute Forecast Error



Evaluation: male life expectancy

One-step-ahead MAE

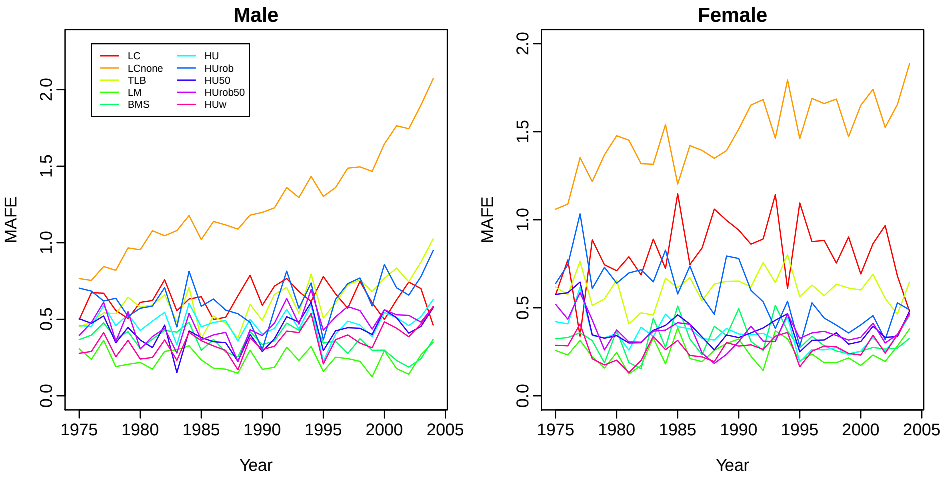
Country	LC	LCnone	TLB	LM	BMS	HU	HU50	HUrob	HUrob50	HUw
Australia	0.616	1.836	0.693	0.282	0.371	0.330	0.313	0.513	0.322	0.286
Canada	0.194	1.326	0.484	0.150	0.263	0.190	0.223	0.395	0.288	0.157
Denmark	0.381	0.380	0.574	0.223	0.420	0.297	0.339	0.560	0.554	0.230
England	0.936	1.837	0.486	0.176	0.235	0.491	0.198	0.633	0.274	0.255
Finland	0.587	1.861	0.580	0.188	0.282	0.545	0.327	0.475	0.412	0.390
France	0.921	2.298	0.285	0.128	0.172	0.522	0.188	1.032	0.210	0.291
Iceland	0.832	0.854	1.032	0.854	0.915	1.512	2.081	1.636	1.921	1.526
Italy	0.608	1.411	0.921	0.201	0.244	0.606	0.229	0.937	0.310	0.286
Netherlands	0.348	0.534	0.769	0.190	0.575	0.327	0.235	0.333	0.284	0.244
Norway	0.606	0.725	0.895	0.212	0.559	0.336	0.329	0.413	0.434	0.259
Scotland	1.303	1.728	0.446	0.204	0.285	0.448	0.357	0.473	0.329	0.246
Spain	0.526	1.264	0.661	0.186	0.254	0.302	0.432	0.441	0.419	0.177
Sweden	0.518	0.657	0.626	0.148	0.235	0.409	0.303	0.626	0.397	0.315
Switzerland	0.472	0.921	0.296	0.178	0.213	0.326	0.283	0.454	0.399	0.297
Weighted ave	0.657	1.554	0.586	**0.179**	0.266	0.432	0.260	0.677	0.311	0.254

Evaluation: female life expectancy

One-step-ahead MAE

Country	LC	LCnone	TLB	LM	BMS	HU	HU50	HUrob	HUrob50	HUw
Australia	0.412	1.493	0.371	0.253	0.275	0.261	0.245	0.308	0.264	0.240
Canada	0.294	1.056	0.351	0.111	0.105	0.147	0.194	0.198	0.192	0.101
Denmark	1.366	1.199	0.543	0.246	0.432	0.203	0.236	0.457	0.251	0.200
England	0.828	1.208	0.279	0.171	0.197	0.189	0.183	0.464	0.207	0.144
Finland	0.901	1.997	0.499	0.211	0.288	0.235	0.285	0.310	0.285	0.216
France	0.983	2.212	0.293	0.186	0.288	0.221	0.208	0.950	0.209	0.154
Iceland	0.655	1.777	2.289	0.869	1.030	1.724	2.074	1.566	1.755	1.393
Italy	0.571	1.655	0.639	0.184	0.247	0.270	0.222	0.526	0.290	0.171
Netherlands	0.811	1.216	0.396	0.177	0.238	0.220	0.206	0.241	0.219	0.172
Norway	1.065	1.254	0.494	0.181	0.242	0.287	0.347	0.586	0.298	0.206
Scotland	1.158	1.442	0.593	0.242	0.293	0.363	0.401	0.781	0.355	0.231
Spain	0.680	1.720	1.031	0.194	0.238	0.245	0.258	0.377	0.330	0.181
Sweden	0.942	0.958	0.325	0.185	0.304	0.287	0.214	1.074	0.213	0.206
Switzerland	0.793	1.575	0.404	0.180	0.244	0.202	0.230	0.308	0.174	0.188
Weighted ave	0.733	1.572	0.488	0.183	0.239	0.229	0.223	0.528	0.248	**0.167**

Mean Absolute Forecast Error



Outline

Selected references

Hyndman, Ullah (2007). “Robust forecasting of mortality and fertility rates: A functional data approach”. *Computational Statistics & Data Analysis* **51**(10), 4942–4956.

Hyndman, Shang (2009). “Forecasting functional time series (with discussion)”. *Journal of the Korean Statistical Society* **38**(3), 199–221.

Shang, Booth, Hyndman (2011). “Point and interval forecasts of mortality rates and life expectancy : a comparison of ten principal component methods”. *Demographic Research* **25**, 173–214.

Hyndman (2014). *demography: Forecasting mortality, fertility, migration and population data*. `cran.r-project.org/package=demography`