Bayesian functional models for population forecasting

Han Lin Shang
University of Southampton

Joint work with Peter W. F. Smith, Jakub Bijak, Arkadiusz Wiśniowski and James Raymer

CPC centre for population change

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Motivation

1. Accurately forecast population via cohort component projection model

Motivation

1. Accurately forecast population via cohort component projection model
2. Produce probabilistic forecasts

Motivation

1. Accurately forecast population via cohort component projection model
2. Produce probabilistic forecasts
3. Forecast age-specific male and female population together when possible

Motivation

1. Accurately forecast population via cohort component projection model
2. Produce probabilistic forecasts
3. Forecast age-specific male and female population together when possible
4. As a generalisation of Lee-Carter model, functional data analysis provides a flexible approach

Motivation

1. Accurately forecast population via cohort component projection model
2. Produce probabilistic forecasts
3. Forecast age-specific male and female population together when possible
4. As a generalisation of Lee-Carter model, functional data analysis provides a flexible approach
5. Bayesian model averaging between the independent and coherent functional models

Data: United Kingdom

1. Population and mortality data (1975–2009) are from Human Mortality Database (`http://www.mortality.org/`)

Data: United Kingdom

1. Population and mortality data (1975–2009) are from Human Mortality Database (`http://www.mortality.org/`)
2. Fertility data are from Human Fertility Database (`http://www.humanfertility.org/cgi-bin/main.php`)

Data: United Kingdom

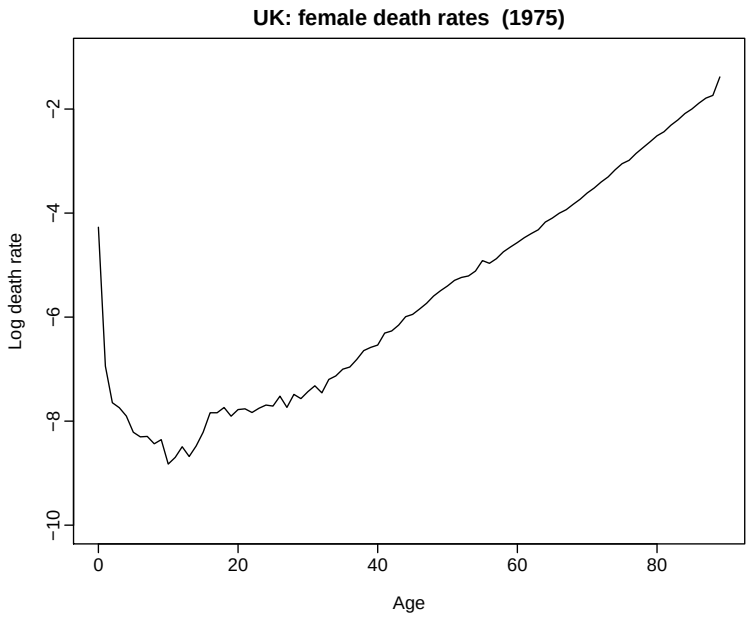
1. Population and mortality data (1975–2009) are from Human Mortality Database (`http://www.mortality.org/`)
2. Fertility data are from Human Fertility Database (`http://www.humanfertility.org/cgi-bin/main.php`)
3. Emigration and immigration data are from Office for National Statistics (`http://www.statistics.gov.uk/hub/index.html`)

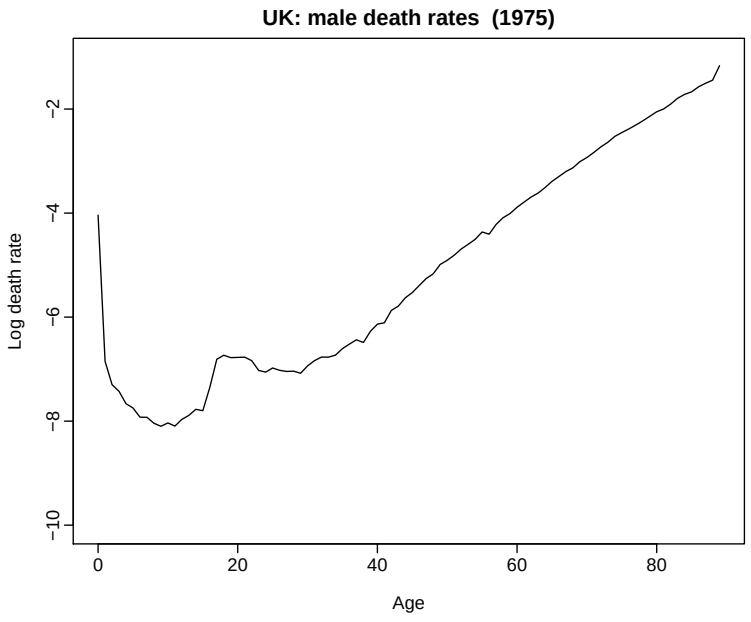
Data: United Kingdom

1. Population and mortality data (1975–2009) are from Human Mortality Database (`http://www.mortality.org/`)
2. Fertility data are from Human Fertility Database (`http://www.humanfertility.org/cgi-bin/main.php`)
3. Emigration and immigration data are from Office for National Statistics (`http://www.statistics.gov.uk/hub/index.html`)
4. Aim to forecast population by age and gender from 2010 to 2030

Graphical display of mortality rates

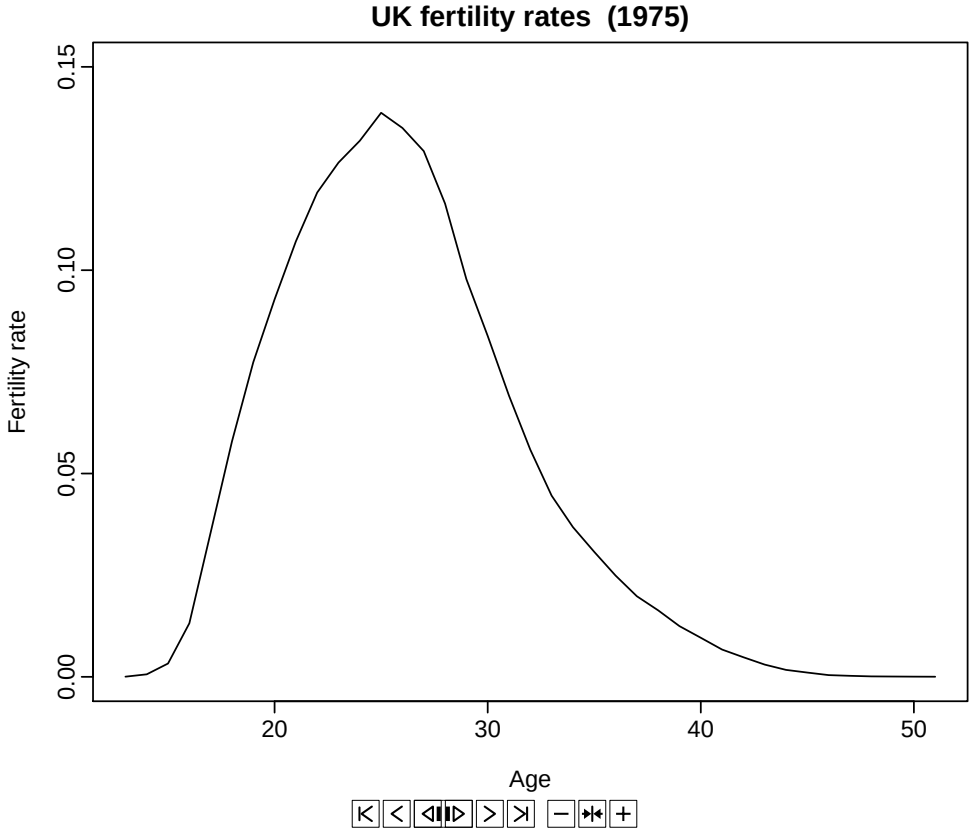
1. Mortality rates for single year of age (0–90+) between 1975 and 2009
2. Mortality rates dip in early childhood, climb in the teen years, stablise in the early 20s, and then steadily increase with age





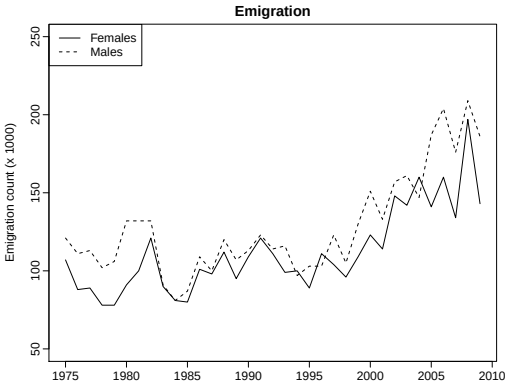
Graphical display of fertility rates

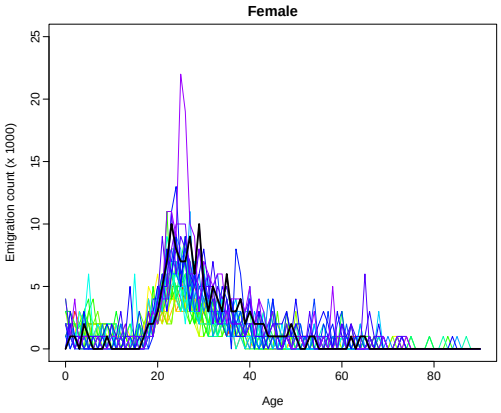
1. Age-specific fertility rates for ages (13–51) between 1975 and 2009
2. Increase in fertility rates at higher ages in more recent years caused by a tendency to postpone child-bearing

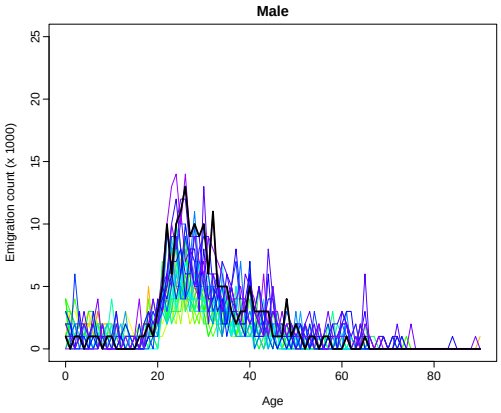


Graphical display of emigration rates

1. A slight increase over time in the emigration, but the patterns seem to be more volatile
2. Emigration data come from the International Passenger Survey

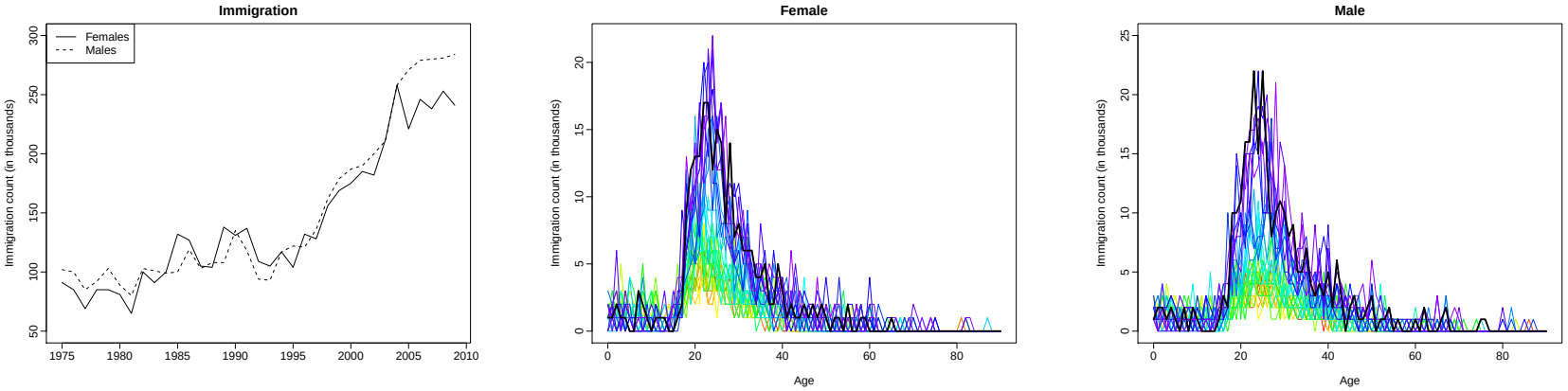






Graphical display of immigration counts

Immigration counts have been rapidly increasing from 1975 to 2009



Cohort component projection model

$$\begin{bmatrix} P^{\mathsf{F}}_{1,t+1} \\ P^{\mathsf{F}}_{2,t+1} \\ \vdots \\ P^{\mathsf{F}}_{z-1,t+1} \\ P^{\mathsf{F}}_{z,t+1} \end{bmatrix} = \begin{bmatrix} 0 & \dots & b^{\mathsf{F}}_{13,t} & \dots & b^{\mathsf{F}}_{51,t} & \dots & 0 \\ s^{\mathsf{F}}_{0,t} & 0 & 0 & & \dots & & 0 \\ 0 & s^{\mathsf{F}}_{1,t} & 0 & & \dots & & 0 \\ \vdots & & \ddots & & & & \vdots \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & \dots & & s^{\mathsf{F}}_{z-2,t} & 0 & 0 \\ 0 & 0 & \dots & & 0 & s^{\mathsf{F}}_{z-1,t} & s^{\mathsf{F}}_{z,t} \end{bmatrix} \times \begin{bmatrix} P^{\mathsf{F}}_{1,t} \\ P^{\mathsf{F}}_{2,t} \\ \vdots \\ P^{\mathsf{F}}_{z-1,t} \\ P^{\mathsf{F}}_{z,t} \end{bmatrix} + \boldsymbol{I}^{\mathsf{F}}_t,$$

- Growth matrix includes survival rates $s_{i,t},\ i = 0, \dots, 90+$ based on mortality and emigration, and fertility rates $b_{j,t},\ j = 13, \dots, 51$

Cohort component projection model

$$\begin{bmatrix} P^{\mathsf{F}}_{1,t+1} \\ P^{\mathsf{F}}_{2,t+1} \\ \vdots \\ P^{\mathsf{F}}_{z-1,t+1} \\ P^{\mathsf{F}}_{z,t+1} \end{bmatrix} = \begin{bmatrix} 0 & \dots & b^{\mathsf{F}}_{13,t} & \dots & b^{\mathsf{F}}_{51,t} & \dots & 0 \\ s^{\mathsf{F}}_{0,t} & 0 & 0 & & \dots & & 0 \\ 0 & s^{\mathsf{F}}_{1,t} & 0 & & \dots & & 0 \\ \vdots & & \ddots & & & & \vdots \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & \dots & & s^{\mathsf{F}}_{z-2,t} & 0 & 0 \\ 0 & 0 & \dots & & 0 & s^{\mathsf{F}}_{z-1,t} & s^{\mathsf{F}}_{z,t} \end{bmatrix} \times \begin{bmatrix} P^{\mathsf{F}}_{1,t} \\ P^{\mathsf{F}}_{2,t} \\ \vdots \\ P^{\mathsf{F}}_{z-1,t} \\ P^{\mathsf{F}}_{z,t} \end{bmatrix} + \boldsymbol{I}^{\mathsf{F}}_t,$$

- Growth matrix includes survival rates $s_{i,t},\ i = 0, \dots, 90+$ based on mortality and emigration, and fertility rates $b_{j,t},\ j = 13, \dots, 51$
- $\boldsymbol{I}^{\mathsf{F}}_t = \left[I^{\mathsf{F}}_{0,t}, \dots, I^{\mathsf{F}}_{+90,t}\right]^\top$ denotes female immigration counts

Functional data analysis

1. Received increasing amount of attention (Hyndman and Ullah, 2007; Lazar and Denuit, 2009; Yang et al., 2010; Dowd et al., 2010; Cairns et al., 2011; D'Amato et al., 2011)

Functional data analysis

1. Received increasing amount of attention (Hyndman and Ullah, 2007; Lazar and Denuit, 2009; Yang et al., 2010; Dowd et al., 2010; Cairns et al., 2011; D'Amato et al., 2011)
2. Let $m_{t,i}$ represent the original data for ages $i = 0, \ldots 90+$ in year t. The Box-Cox transformation is applied to each component of population to alleviate heteroscedasticity and non-normality

$$g_{t,i} = \begin{cases} \frac{1}{\lambda}(m_{t,i}^{\lambda} - 1) & \text{if } 0 < \lambda \leq 1 \\ \ln(m_{t,i}) & \text{if } \lambda = 0 \end{cases}$$

where λ is the Box-Cox transformation parameter

Smoothing step

3. We apply a smoothing technique to transform the set of discrete data points, $g_{t,i}$, to continuous function

$$g_t(x_i) = \tau_t(x_i) + \sigma(x_i)\epsilon_t, \qquad t = 1, \ldots, n,$$

where $\sigma(x_i)$ models the variability for each age x_i, and $\epsilon_t \overset{iid}{\sim} N(0,1)$

Independent decomposition step

To model female and male data independently,

$$\tau_t^{\mathsf{F}}(x) = \mu^{\mathsf{F}}(x) + \sum_{k=1}^{K} \beta_{t,k}^{\mathsf{F}} \phi_k^{\mathsf{F}}(x) + e_t^{\mathsf{F}}(x), \qquad x \in [0, 90+] \tag{1}$$

where

- $\mu^{\mathsf{F}}(x)$ represents the mean function for females

Independent decomposition step

To model female and male data independently,

$$\tau_t^{\mathsf{F}}(x) = \mu^{\mathsf{F}}(x) + \sum_{k=1}^{K} \beta_{t,k}^{\mathsf{F}} \phi_k^{\mathsf{F}}(x) + e_t^{\mathsf{F}}(x), \qquad x \in [0, 90+] \tag{1}$$

where

- $\mu^{\mathsf{F}}(x)$ represents the mean function for females
- $\{\phi_1^{\mathsf{F}}(x), \ldots, \phi_K^{\mathsf{F}}(x)\}$ is a set of first K functional principal components

Independent decomposition step

To model female and male data independently,

$$\tau_t^{\mathsf{F}}(x) = \mu^{\mathsf{F}}(x) + \sum_{k=1}^{K} \beta_{t,k}^{\mathsf{F}} \phi_k^{\mathsf{F}}(x) + e_t^{\mathsf{F}}(x), \qquad x \in [0, 90+] \tag{1}$$

where

- $\mu^{\mathsf{F}}(x)$ represents the mean function for females
- $\{\phi_1^{\mathsf{F}}(x), \ldots, \phi_K^{\mathsf{F}}(x)\}$ is a set of first K functional principal components
- $\{\beta_{t,1}^{\mathsf{F}}, \ldots, \beta_{t,K}^{\mathsf{F}}\}$ is a set of corresponding principal component scores

Independent decomposition step

To model female and male data independently,

$$\tau_t^{\mathsf{F}}(x) = \mu^{\mathsf{F}}(x) + \sum_{k=1}^{K} \beta_{t,k}^{\mathsf{F}} \phi_k^{\mathsf{F}}(x) + e_t^{\mathsf{F}}(x), \qquad x \in [0, 90+] \tag{1}$$

where

- $\mu^{\mathsf{F}}(x)$ represents the mean function for females
- $\{\phi_1^{\mathsf{F}}(x), \ldots, \phi_K^{\mathsf{F}}(x)\}$ is a set of first K functional principal components
- $\{\beta_{t,1}^{\mathsf{F}}, \ldots, \beta_{t,K}^{\mathsf{F}}\}$ is a set of corresponding principal component scores
- $e_t^{\mathsf{F}}(x) \sim N(0, v(x))$ is the residual function

Independent decomposition step

To model female and male data independently,

$$\tau_t^{\mathsf{F}}(x) = \mu^{\mathsf{F}}(x) + \sum_{k=1}^{K} \beta_{t,k}^{\mathsf{F}} \phi_k^{\mathsf{F}}(x) + e_t^{\mathsf{F}}(x), \qquad x \in [0, 90+] \tag{1}$$

where

- $\mu^{\mathsf{F}}(x)$ represents the mean function for females
- $\{\phi_1^{\mathsf{F}}(x), \ldots, \phi_K^{\mathsf{F}}(x)\}$ is a set of first K functional principal components
- $\{\beta_{t,1}^{\mathsf{F}}, \ldots, \beta_{t,K}^{\mathsf{F}}\}$ is a set of corresponding principal component scores
- $e_t^{\mathsf{F}}(x) \sim N(0, v(x))$ is the residual function
- $K < n$ is the number of retained components, which explain at least 99% of the total variation in data

Independent decomposition step

The terms in equation (1) can be estimated by

- $\widehat{\mu}^{\mathsf{F}}(x) = \frac{1}{n}\sum_{t=1}^{n} \tau_t^{\mathsf{F}}(x)$

Independent decomposition step

The terms in equation (1) can be estimated by

- $\widehat{\mu}^{\mathsf{F}}(x) = \frac{1}{n}\sum_{t=1}^{n} \tau_t^{\mathsf{F}}(x)$
- $\{\widehat{\phi}_1^{\mathsf{F}}(x), \ldots, \widehat{\phi}_K^{\mathsf{F}}(x)\}$ are the sample principal components, obtained by singular value decomposition

Independent decomposition step

The terms in equation (1) can be estimated by

- $\widehat{\mu}^{\mathsf{F}}(x) = \frac{1}{n}\sum_{t=1}^{n} \tau_t^{\mathsf{F}}(x)$
- $\{\widehat{\phi}_1^{\mathsf{F}}(x), \ldots, \widehat{\phi}_K^{\mathsf{F}}(x)\}$ are the sample principal components, obtained by singular value decomposition
- $\widehat{\beta}_{t,k}^{\mathsf{F}}$ and $\widehat{v}(x)$ are drawn from their respective posterior distribution, where the likelihood is the Box-Cox normal and the prior densities of the variance parameters are inverse gamma distributions

Coherent decomposition step

4. We model female and male data jointly via multilevel functional principal component regression

$$\tau_t^{\mathsf{F}}(x) = \underbrace{\mu(x) + w^{\mathsf{F}}(x)}_{\mu^{\mathsf{F}}(x)} + R_t(x) + U_t^{\mathsf{F}}(x) + \varepsilon_t^{\mathsf{F}}(x), \qquad (2)$$

where

- $\mu(x)$ represents the overall mean function

Coherent decomposition step

4. We model female and male data jointly via multilevel functional principal component regression

$$\tau_t^{\mathsf{F}}(x) = \underbrace{\mu(x) + w^{\mathsf{F}}(x)}_{\mu^{\mathsf{F}}(x)} + R_t(x) + U_t^{\mathsf{F}}(x) + \varepsilon_t^{\mathsf{F}}(x), \qquad (2)$$

where

- $\mu(x)$ represents the overall mean function
- $w^{\mathsf{F}}(x)$ is the female deviation from the overall mean function

Coherent decomposition step

4. We model female and male data jointly via multilevel functional principal component regression

$$\tau_t^{\mathsf{F}}(x) = \underbrace{\mu(x) + w^{\mathsf{F}}(x)}_{\mu^{\mathsf{F}}(x)} + R_t(x) + U_t^{\mathsf{F}}(x) + \varepsilon_t^{\mathsf{F}}(x), \qquad (2)$$

where

- $\mu(x)$ represents the overall mean function
- $w^{\mathsf{F}}(x)$ is the female deviation from the overall mean function
- $R_t(x)$ models the common trend for female and male data

Coherent decomposition step

4. We model female and male data jointly via multilevel functional principal component regression

$$\tau_t^{\mathsf{F}}(x) = \underbrace{\mu(x) + w^{\mathsf{F}}(x)}_{\mu^{\mathsf{F}}(x)} + R_t(x) + U_t^{\mathsf{F}}(x) + \varepsilon_t^{\mathsf{F}}(x), \qquad (2)$$

where

- $\mu(x)$ represents the overall mean function
- $w^{\mathsf{F}}(x)$ is the female deviation from the overall mean function
- $R_t(x)$ models the common trend for female and male data
- $U_t^{\mathsf{F}}(x)$ models the gender-specific trend in the female data

Coherent decomposition step

4. We model female and male data jointly via multilevel functional principal component regression

$$\tau_t^{\mathsf{F}}(x) = \underbrace{\mu(x) + w^{\mathsf{F}}(x)}_{\mu^{\mathsf{F}}(x)} + R_t(x) + U_t^{\mathsf{F}}(x) + \varepsilon_t^{\mathsf{F}}(x), \qquad (2)$$

where

- $\mu(x)$ represents the overall mean function
- $w^{\mathsf{F}}(x)$ is the female deviation from the overall mean function
- $R_t(x)$ models the common trend for female and male data
- $U_t^{\mathsf{F}}(x)$ models the gender-specific trend in the female data
- $\varepsilon_t^{\mathsf{F}}(x)$ represents the error term with finite variance $\delta(x)$

Coherent decomposition step

4. We model female and male data jointly via multilevel functional principal component regression

$$\tau_t^{\mathsf{F}}(x) = \underbrace{\mu(x) + w^{\mathsf{F}}(x)}_{\mu^{\mathsf{F}}(x)} + R_t(x) + U_t^{\mathsf{F}}(x) + \varepsilon_t^{\mathsf{F}}(x), \qquad (2)$$

where

- $\mu(x)$ represents the overall mean function
- $w^{\mathsf{F}}(x)$ is the female deviation from the overall mean function
- $R_t(x)$ models the common trend for female and male data
- $U_t^{\mathsf{F}}(x)$ models the gender-specific trend in the female data
- $\varepsilon_t^{\mathsf{F}}(x)$ represents the error term with finite variance $\delta(x)$
- Similarly, equation (2) also holds for males

Coherent decomposition step

5. The terms in equation (2) can be estimated by

$$\widehat{\mu}(x) = \frac{1}{2n}\sum_{j=1}^{2}\sum_{t=1}^{n}\tau_t^{(j)}(x), \quad \widehat{w}^{\mathsf{F}}(x) = \widehat{\mu}^{\mathsf{F}}(x) - \widehat{\mu}(x),$$

$$\widehat{R}_t(x) \approx \sum_{k=1}^{K}\widehat{\beta}_{t,k}\widehat{\phi}_k(x), \quad \widehat{U}_t^{\mathsf{F}}(x) \approx \sum_{l=1}^{L}\widehat{\gamma}_{t,l}^{\mathsf{F}}\widehat{\psi}_l^{\mathsf{F}}(x),$$

where

1. $\left\{\widehat{\beta}_{t,1},\ldots,\widehat{\beta}_{t,K}\right\}$ and $\left\{\widehat{\gamma}_{t,1}^{\mathsf{F}},\ldots,\widehat{\gamma}_{t,L}^{\mathsf{F}}\right\}$ represent sample principal component scores of $\widehat{R}_t(x)$ and $\widehat{U}_t^{\mathsf{F}}(x)$

Coherent decomposition step

5. The terms in equation (2) can be estimated by

$$\widehat{\mu}(x) = \frac{1}{2n}\sum_{j=1}^{2}\sum_{t=1}^{n}\tau_t^{(j)}(x), \quad \widehat{w}^{\mathsf{F}}(x) = \widehat{\mu}^{\mathsf{F}}(x) - \widehat{\mu}(x),$$

$$\widehat{R}_t(x) \approx \sum_{k=1}^{K}\widehat{\beta}_{t,k}\widehat{\phi}_k(x), \quad \widehat{U}_t^{\mathsf{F}}(x) \approx \sum_{l=1}^{L}\widehat{\gamma}_{t,l}^{\mathsf{F}}\widehat{\psi}_l^{\mathsf{F}}(x),$$

where

1. $\left\{\widehat{\beta}_{t,1},\ldots,\widehat{\beta}_{t,K}\right\}$ and $\left\{\widehat{\gamma}_{t,1}^{\mathsf{F}},\ldots,\widehat{\gamma}_{t,L}^{\mathsf{F}}\right\}$ represent sample principal component scores of $\widehat{R}_t(x)$ and $\widehat{U}_t^{\mathsf{F}}(x)$
2. $\mathbf{\Phi} = \left\{\widehat{\phi}_1(x),\ldots,\widehat{\phi}_K(x)\right\}$ and $\mathbf{\Psi} = \left\{\widehat{\psi}_1^{\mathsf{F}}(x),\ldots,\widehat{\psi}_L^{\mathsf{F}}(x)\right\}$ represent the sample principal components of $\widehat{R}_t(x)$ and $\widehat{U}_t^{\mathsf{F}}(x)$

Coherent decomposition step

5. The terms in equation (2) can be estimated by

$$\widehat{\mu}(x) = \frac{1}{2n}\sum_{j=1}^{2}\sum_{t=1}^{n}\tau_t^{(j)}(x), \quad \widehat{w}^{\mathsf{F}}(x) = \widehat{\mu}^{\mathsf{F}}(x) - \widehat{\mu}(x),$$

$$\widehat{R}_t(x) \approx \sum_{k=1}^{K}\widehat{\beta}_{t,k}\widehat{\phi}_k(x), \quad \widehat{U}_t^{\mathsf{F}}(x) \approx \sum_{l=1}^{L}\widehat{\gamma}_{t,l}^{\mathsf{F}}\widehat{\psi}_l^{\mathsf{F}}(x),$$

where

1. $\left\{\widehat{\beta}_{t,1},\ldots,\widehat{\beta}_{t,K}\right\}$ and $\left\{\widehat{\gamma}_{t,1}^{\mathsf{F}},\ldots,\widehat{\gamma}_{t,L}^{\mathsf{F}}\right\}$ represent sample principal component scores of $\widehat{R}_t(x)$ and $\widehat{U}_t^{\mathsf{F}}(x)$
2. $\mathbf{\Phi} = \left\{\widehat{\phi}_1(x),\ldots,\widehat{\phi}_K(x)\right\}$ and $\mathbf{\Psi} = \left\{\widehat{\psi}_1^{\mathsf{F}}(x),\ldots,\widehat{\psi}_L^{\mathsf{F}}(x)\right\}$ represent the sample principal components of $\widehat{R}_t(x)$ and $\widehat{U}_t^{\mathsf{F}}(x)$
3. K and L are the number of retained components, which explain at least 95% and 90% of the total variations in data

Forecasting step of coherent decomposition

6. Conditional on the estimated principal components $\mathbf{\Phi}$ and $\mathbf{\Psi}$ and smooth functions $\mathbf{\Gamma}$, point forecasts of $g^{\mathsf{F}}_{n+h}(x)$ are

$$\begin{aligned}\widehat{g}^{\mathsf{F}}_{n+h|n}(x) &= \mathsf{E}\left[g^{\mathsf{F}}_{n+h}(x)\middle|\mathbf{\Phi},\mathbf{\Psi},\mathbf{\Gamma}\right] \\ &= \widehat{\mu}(x) + \widehat{w}^{\mathsf{F}}(x) + \sum_{k=1}^{K}\widehat{\beta}_{n+h|n,k}\widehat{\phi}_k(x) + \sum_{l=1}^{L}\widehat{\gamma}^{\mathsf{F}}_{n+h|n,l}\widehat{\psi}^{\mathsf{F}}_l(x)\end{aligned}$$

where $\widehat{\beta}_{n+h|n,k}$ and $\widehat{\gamma}^{\mathsf{F}}_{n+h|n,l}$ are the forecasted principal component scores, obtained from a univariate time-series method, such as the ARIMA

Prediction interval

Prediction interval of $g^{\mathsf{F}}_{n+h|n}(x)$ can be obtained using parametric bootstrap samples, given by

$$\widehat{g}^{b,\mathsf{F}}_{n+h|n}(x) = \widehat{\mu}(x) + \widehat{w}^{\mathsf{F}}(x) + \sum_{k=1}^{K} \widehat{\beta}^{b}_{n+h|n,k}\widehat{\phi}_k(x) + \sum_{l=1}^{L} \widehat{\gamma}^{b,\mathsf{F}}_{n+h|n,l}\widehat{\psi}^{\mathsf{F}}_l(x) + \widehat{\varepsilon}^{b}_{n+h}(x) + \widehat{\sigma}(x)\widehat{\epsilon}^{b}_{n+h}, \qquad b = 1, \ldots, B,$$

where

1. $\widehat{\beta}^{b}_{n+h|n,k}$ is the forecast of bootstrapped principal component scores $\{\widehat{\beta}^{b}_{1,k}, \ldots, \widehat{\beta}^{b}_{n,k}\}$

Prediction interval

Prediction interval of $g^{\mathsf{F}}_{n+h|n}(x)$ can be obtained using parametric bootstrap samples, given by

$$\widehat{g}^{b,\mathsf{F}}_{n+h|n}(x) = \widehat{\mu}(x) + \widehat{w}^{\mathsf{F}}(x) + \sum_{k=1}^{K} \widehat{\beta}^{b}_{n+h|n,k}\widehat{\phi}_k(x) + \sum_{l=1}^{L} \widehat{\gamma}^{b,\mathsf{F}}_{n+h|n,l}\widehat{\psi}^{\mathsf{F}}_l(x) + \widehat{\varepsilon}^{b}_{n+h}(x) + \widehat{\sigma}(x)\widehat{\epsilon}^{b}_{n+h}, \qquad b = 1, \ldots, B,$$

where

1. $\widehat{\beta}^{b}_{n+h|n,k}$ is the forecast of bootstrapped principal component scores $\{\widehat{\beta}^{b}_{1,k}, \ldots, \widehat{\beta}^{b}_{n,k}\}$
2. $\widehat{\gamma}^{b,\mathsf{F}}_{n+h|n,l}$ is the forecast of bootstrapped principal component scores $\{\widehat{\gamma}^{b,\mathsf{F}}_{1,l}, \ldots, \widehat{\gamma}^{b,\mathsf{F}}_{n,l}\}$

Prediction interval

Prediction interval of $g^{\mathsf{F}}_{n+h|n}(x)$ can be obtained using parametric bootstrap samples, given by

$$\widehat{g}^{b,\mathsf{F}}_{n+h|n}(x) = \widehat{\mu}(x) + \widehat{w}^{\mathsf{F}}(x) + \sum_{k=1}^{K} \widehat{\beta}^{b}_{n+h|n,k}\widehat{\phi}_k(x) + \sum_{l=1}^{L} \widehat{\gamma}^{b,\mathsf{F}}_{n+h|n,l}\widehat{\psi}^{\mathsf{F}}_l(x) + \widehat{\varepsilon}^{b}_{n+h}(x) + \widehat{\sigma}(x)\widehat{\epsilon}^{b}_{n+h}, \qquad b = 1, \ldots, B,$$

where

1. $\widehat{\beta}^{b}_{n+h|n,k}$ is the forecast of bootstrapped principal component scores $\{\widehat{\beta}^{b}_{1,k}, \ldots, \widehat{\beta}^{b}_{n,k}\}$
2. $\widehat{\gamma}^{b,\mathsf{F}}_{n+h|n,l}$ is the forecast of bootstrapped principal component scores $\{\widehat{\gamma}^{b,\mathsf{F}}_{1,l}, \ldots, \widehat{\gamma}^{b,\mathsf{F}}_{n,l}\}$
3. $\widehat{\varepsilon}^{b}_{n+h}(x)$ is drawn from $N(0, \widehat{\delta}(x))$

Prediction interval

Prediction interval of $g^{\mathsf{F}}_{n+h|n}(x)$ can be obtained using parametric bootstrap samples, given by

$$\widehat{g}^{b,\mathsf{F}}_{n+h|n}(x) = \widehat{\mu}(x) + \widehat{w}^{\mathsf{F}}(x) + \sum_{k=1}^{K} \widehat{\beta}^{b}_{n+h|n,k}\widehat{\phi}_k(x) + \sum_{l=1}^{L} \widehat{\gamma}^{b,\mathsf{F}}_{n+h|n,l}\widehat{\psi}^{\mathsf{F}}_l(x) + \widehat{\varepsilon}^{b}_{n+h}(x) + \widehat{\sigma}(x)\widehat{\epsilon}^{b}_{n+h}, \qquad b = 1, \ldots, B,$$

where

1. $\widehat{\beta}^{b}_{n+h|n,k}$ is the forecast of bootstrapped principal component scores $\{\widehat{\beta}^{b}_{1,k}, \ldots, \widehat{\beta}^{b}_{n,k}\}$
2. $\widehat{\gamma}^{b,\mathsf{F}}_{n+h|n,l}$ is the forecast of bootstrapped principal component scores $\{\widehat{\gamma}^{b,\mathsf{F}}_{1,l}, \ldots, \widehat{\gamma}^{b,\mathsf{F}}_{n,l}\}$
3. $\widehat{\varepsilon}^{b}_{n+h}(x)$ is drawn from $N(0, \widehat{\delta}(x))$
4. $\widehat{\epsilon}^{b}_{n+h}$ is simulated from $N(0, 1)$

Bayesian model averaging

1. Let M_1 and M_2 be two functional models considered

Bayesian model averaging

1. Let M_1 and M_2 be two functional models considered
2. θ_1 and θ_2 be vector of parameters associated with each model

Bayesian model averaging

1. Let M_1 and M_2 be two functional models considered
2. θ_1 and θ_2 be vector of parameters associated with each model
3. Denote Δ as the quantity of interest, such as a combined forecast of age-specific mortality, then its posterior distribution given data D is

$$\begin{aligned}\Pr(\Delta|D) &= \sum_{r=1}^{2} \Pr(\Delta|M_r, D)\Pr(M_r|D) \\ &= \sum_{r=1}^{2} \Pr(\Delta|M_r, D) \underbrace{\frac{\Pr(D|M_r)\Pr(M_r)}{\sum_{l=1}^{2} \Pr(D|M_l)\Pr(M_l)}}_{\text{weight}}\end{aligned}$$

Harmonic estimator of marginal likelihood

1. Given equal prior probability for the two models, the posterior odds can be obtained by the ratios of two marginal likelihoods averaged over all Markov chain Monte Carlo iterations

Harmonic estimator of marginal likelihood

1. Given equal prior probability for the two models, the posterior odds can be obtained by the ratios of two marginal likelihoods averaged over all Markov chain Monte Carlo iterations
2. The marginal likelihood introduced by Newton & Raftery (1994) can be written as

$$\Pr(D|M_r) \approx \left(\frac{1}{T} \sum_{t=1}^{T} \Pr\left(D|\theta_{M_r}^{(t)}, M_r\right)^{-1} \right)^{-1}$$

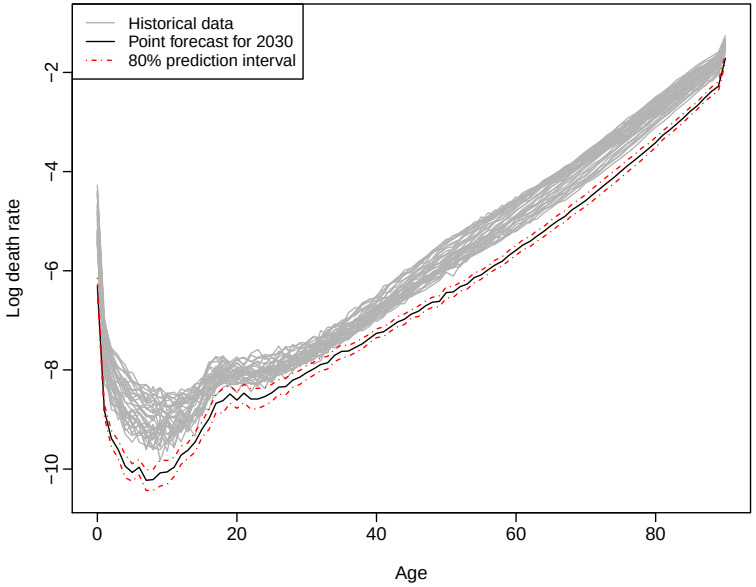
Harmonic estimator of marginal likelihood

1. Given equal prior probability for the two models, the posterior odds can be obtained by the ratios of two marginal likelihoods averaged over all Markov chain Monte Carlo iterations
2. The marginal likelihood introduced by Newton & Raftery (1994) can be written as

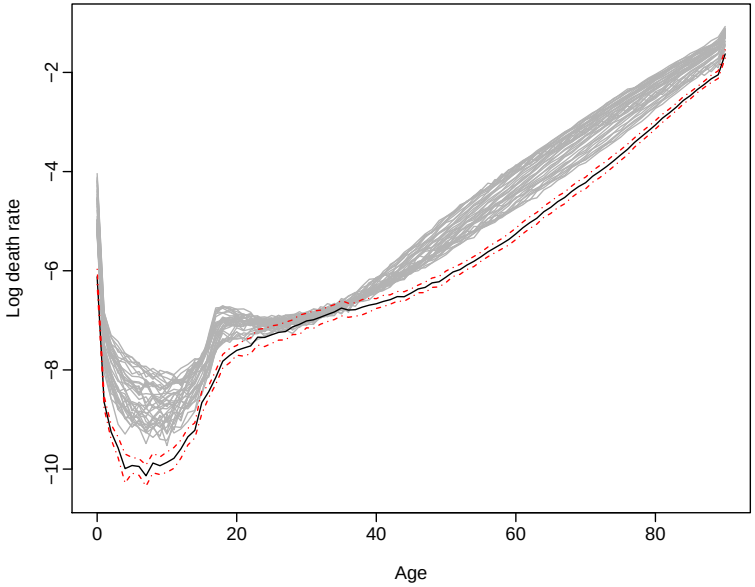
$$\Pr(D|M_r) \approx \left(\frac{1}{T} \sum_{t=1}^{T} \Pr\left(D|\theta_{M_r}^{(t)}, M_r\right)^{-1} \right)^{-1}$$

3. For mortality and emigration rates, coherent functional model has weight one. For immigration counts, independent functional model has weight one

Forecasts of mortality



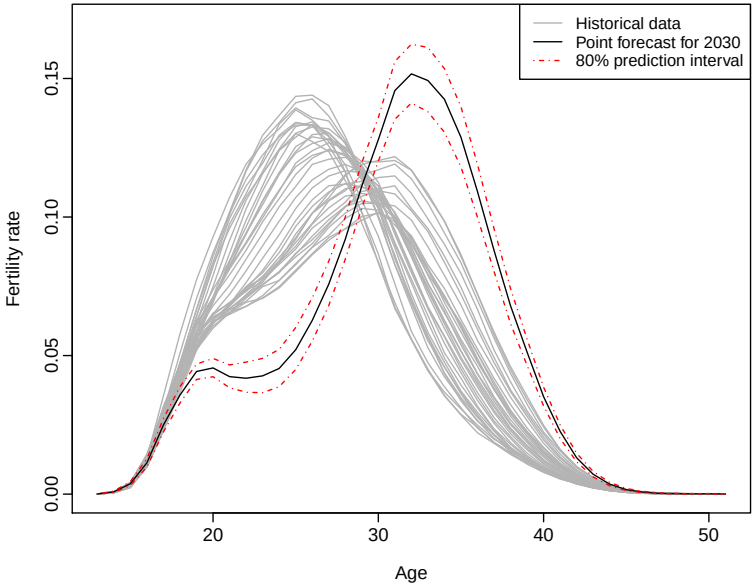
(a) Female

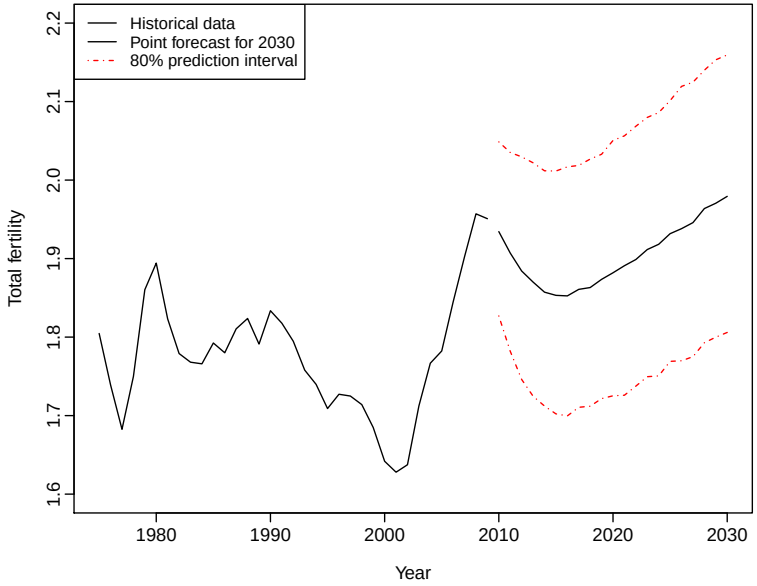


(b) Male

Log mortality rates are likely to drop in general, except for men aged around 30

Forecasts of fertility

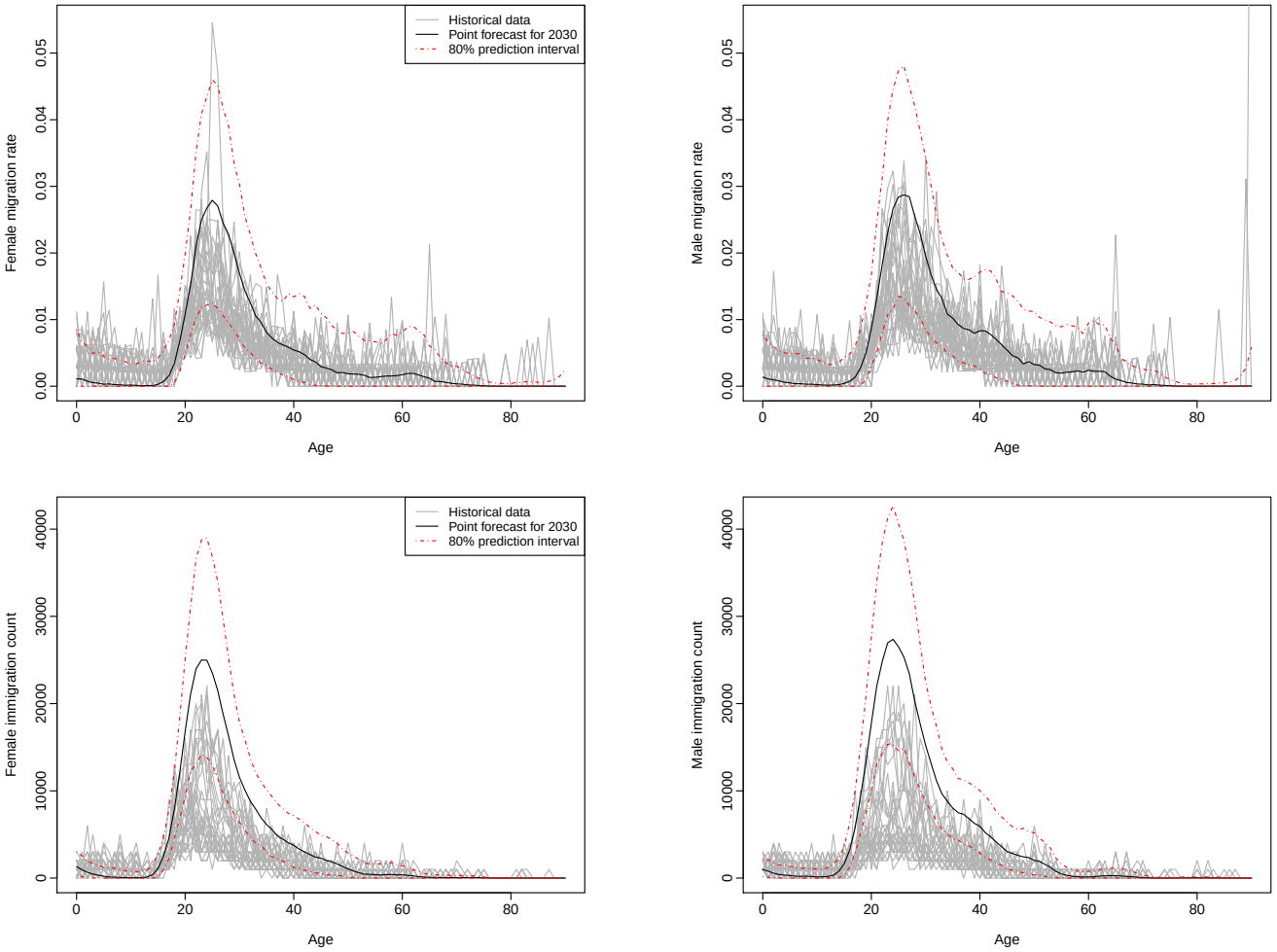




1) Total fertility rate will decrease until 2015, then increase thereafter

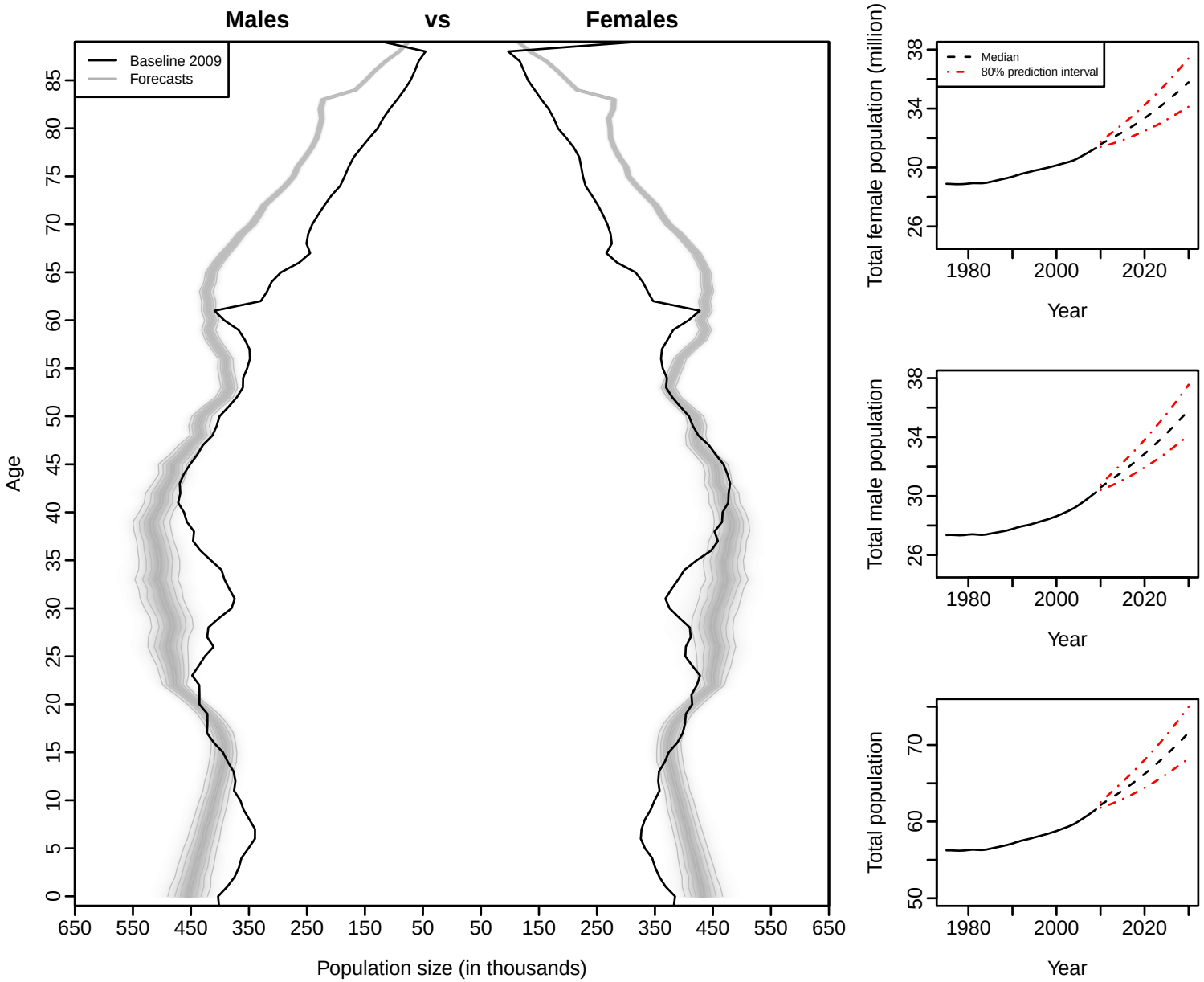
2) The slight declining, yet uncertain, fertility rates signal a possibility of another period of postponement

Forecasts of migration



A continuing increase in emigration rates and immigration counts for ages between 20 and 40

Forecasts of out-of-sample population in 2030



Discussion

1. Present an independent and a coherent functional model to forecast population via cohort component projection model

Discussion

1. Present an independent and a coherent functional model to forecast population via cohort component projection model
2. The coherent one models age-specific female and male data jointly, both functional models estimate uncertainty via parametric bootstrapping

Discussion

1. Present an independent and a coherent functional model to forecast population via cohort component projection model
2. The coherent one models age-specific female and male data jointly, both functional models estimate uncertainty via parametric bootstrapping
3. We aim to propose a fully Bayesian functional data approach

Discussion

1. Present an independent and a coherent functional model to forecast population via cohort component projection model
2. The coherent one models age-specific female and male data jointly, both functional models estimate uncertainty via parametric bootstrapping
3. We aim to propose a fully Bayesian functional data approach
4. It remains a future research to include cohort effect, at least in some of the four demographic components

Discussion

1. Present an independent and a coherent functional model to forecast population via cohort component projection model
2. The coherent one models age-specific female and male data jointly, both functional models estimate uncertainty via parametric bootstrapping
3. We aim to propose a fully Bayesian functional data approach
4. It remains a future research to include cohort effect, at least in some of the four demographic components
5. Thank you

References:

1. Hyndman, R., and Ullah, M. (2007), 'Robust forecasting of mortality and fertility rates: A functional data approach', *Computational Statistics & Data Analysis*, **51**(10), 4942-4956.
2. Lazar, D., and Denuit, M. M. (2009) 'A multivariate time series approach to projected life tables', *Applied Stochastic Models in Business and Industry*, **25**(6), 806-823.
3. Yang, S. S., Yue, J. C., and Huang, H.-C. (2010) 'Modeling longevity risks using a principal component approach: A comparison with existing stochastic mortality models', *Insurance: Mathematics and Economics*, **46**(1), 254-270.
4. Dowd, K. Cairns, A. J. G., Blake, D., Coughlan, G. D., Epstein, D. and Khalaf-Allah, M. (2010) 'Evaluating the goodness of fit of stochastic mortality models', *Insurance: Mathematics and Economics*, **47**(3), 255-265.
5. Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D. and Khalaf-Allah, M. (2011) 'Mortality density forecasts: An analysis of six stochastic mortality models', *Insurance: Mathematics and Economics*, **48**(3), 355-367.
6. D'Amato, V., Piscopo, G., and Russolillo, M. (2011) 'The mortality of the Italian population: Smoothing techniques on the Lee-Carter model', *The Annals of Applied Statistics*, **5**(2A), 705-724.
7. Newton, M. A. and Raftery, A. E. (1994) 'Approximate Bayesian inference with the weighted likelihood bootstrap', *Journal of the Royal Statistical Society. Series B*, **56**(1), 3-48.