Forecasting functional time series

Rob J Hyndman

MONASH University

Mortality rates

Fertility rates

Some notation

Let $y_t(x)$ be the observed (smoothed) data in period t at age x, $t = 1, \dots, n$.

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

- Estimate $\mu(x)$ as mean $y_t(x)$ across years.

Some notation

Let $y_t(x)$ be the observed (smoothed) data in period t at age x, $t = 1, \dots, n$.

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

- Estimate $\mu(x)$ as mean $y_t(x)$ across years.
- Estimate $\beta_{t,k}$ and $\phi_k(x)$ using (robust) functional principal components.

Some notation

Let $y_t(x)$ be the observed (smoothed) data in period t at age x, $t = 1, \dots, n$.

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

- Estimate $\mu(x)$ as mean $y_t(x)$ across years.
- Estimate $\beta_{t,k}$ and $\phi_k(x)$ using (robust) functional principal components.
- $\beta_{t,k}$ are principal component scores

Some notation

Let $y_t(x)$ be the observed (smoothed) data in period t at age x, $t = 1, \dots, n$.

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

- Estimate $\mu(x)$ as mean $y_t(x)$ across years.
- Estimate $\beta_{t,k}$ and $\phi_k(x)$ using (robust) functional principal components.
- $\beta_{t,k}$ are principal component scores
- $\phi_k(x)$ are eigenfunctions

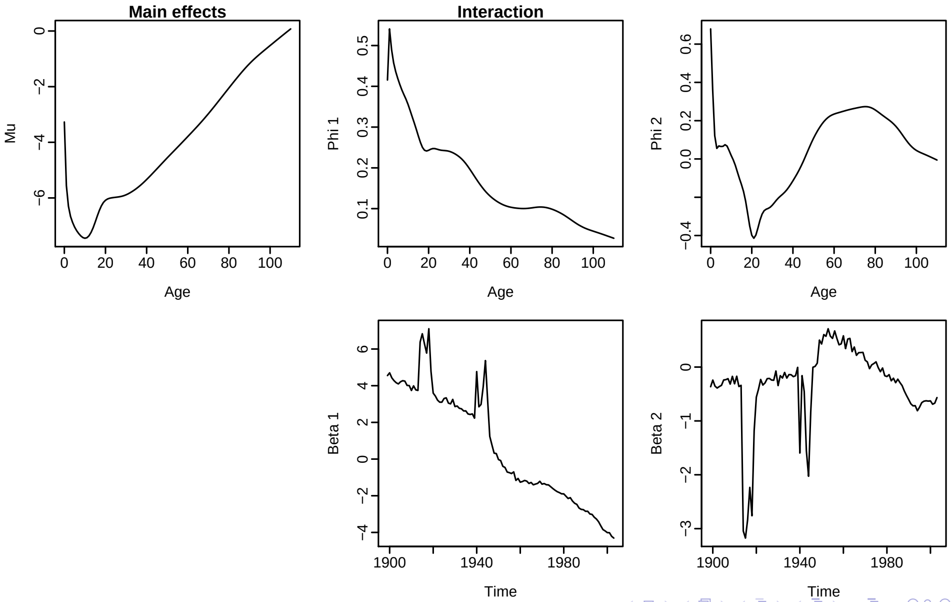
Some notation

Let $y_t(x)$ be the observed (smoothed) data in period t at age x, $t = 1, \dots, n$.

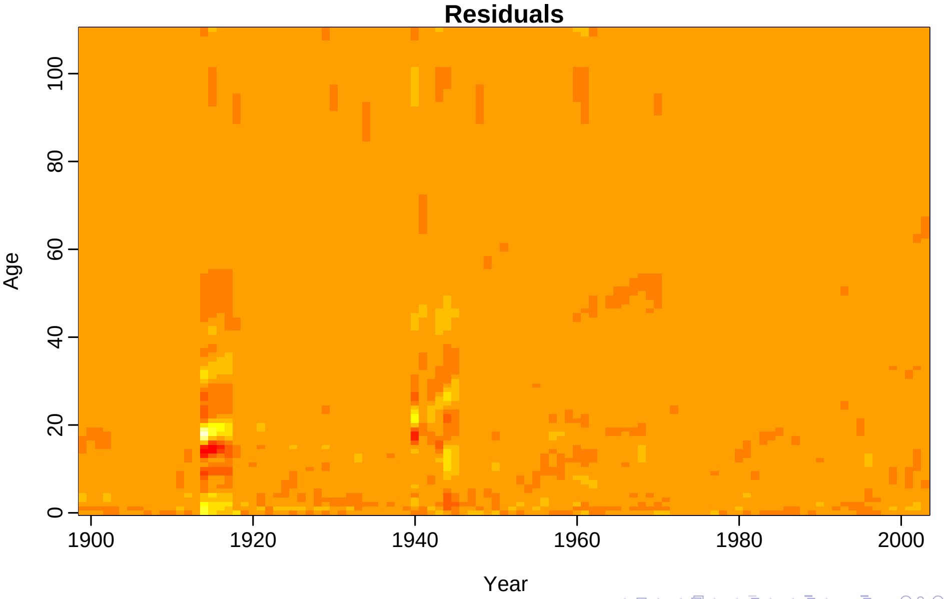
$$y_t(x) \;=\; \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + e_t(x)$$

- Estimate $\mu(x)$ as mean $y_t(x)$ across years.
- Estimate $\beta_{t,k}$ and $\phi_k(x)$ using (robust) functional principal components.
- $\beta_{t,k}$ are principal component scores
- $\phi_k(x)$ are eigenfunctions
- $e_t(x) \overset{\text{iid}}{\sim} \text{N}(0, v(x))$ is random error

Functional PC



Functional PC

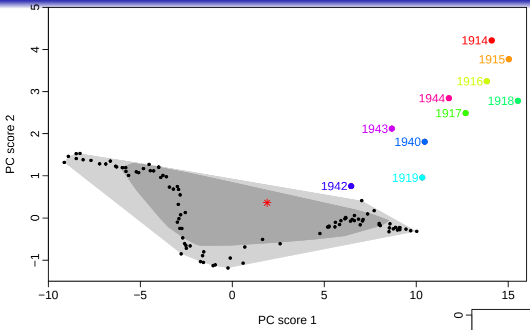


Outliers

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

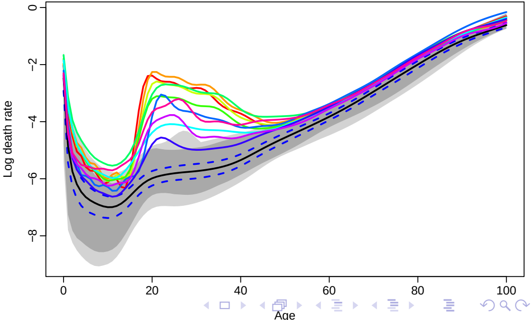
- Outliers in $\{y_t(x_i)\}$ will show up as outliers in $\{\beta_{t,1}, \beta_{t,2}, \dots\}$.

Bagplots

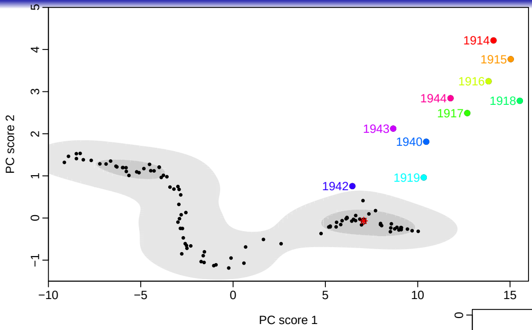


Red asterisk: Median
Dark grey: 50% region
Light grey: outer region

Bagplot for the French male mortality data.

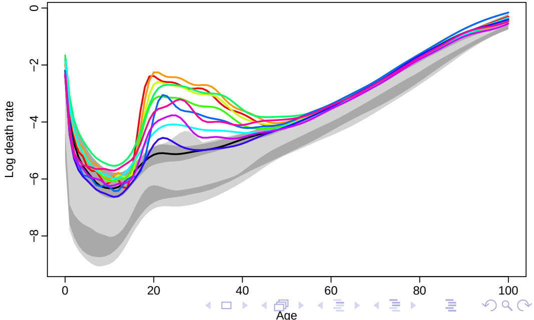


Functional HDR boxplot



Red asterisk: Mode
Dark grey: 50% region
Light grey: outer region

HDR boxplot for the French male mortality data.



Functional time series model

$$y_t(x) \;=\; \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\, \phi_k(x) + e_t(x)$$

Functional time series model

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

- The eigenfunctions $\phi_k(x)$ show the main regions of variation.

Functional time series model

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

- The eigenfunctions $\phi_k(x)$ show the main regions of variation.
- The scores $\{\beta_{t,k}\}$ are uncorrelated by construction. So we can forecast each $\beta_{t,k}$ using a univariate time series model.

Functional time series model

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

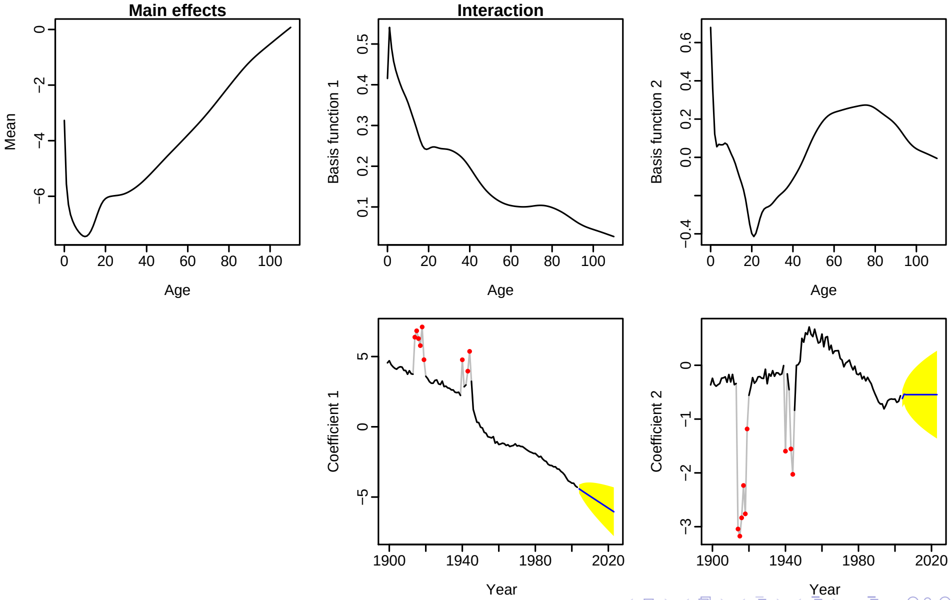
- The eigenfunctions $\phi_k(x)$ show the main regions of variation.
- The scores $\{\beta_{t,k}\}$ are uncorrelated by construction. So we can forecast each $\beta_{t,k}$ using a univariate time series model.
- Outliers are treated as missing values.

Functional time series model

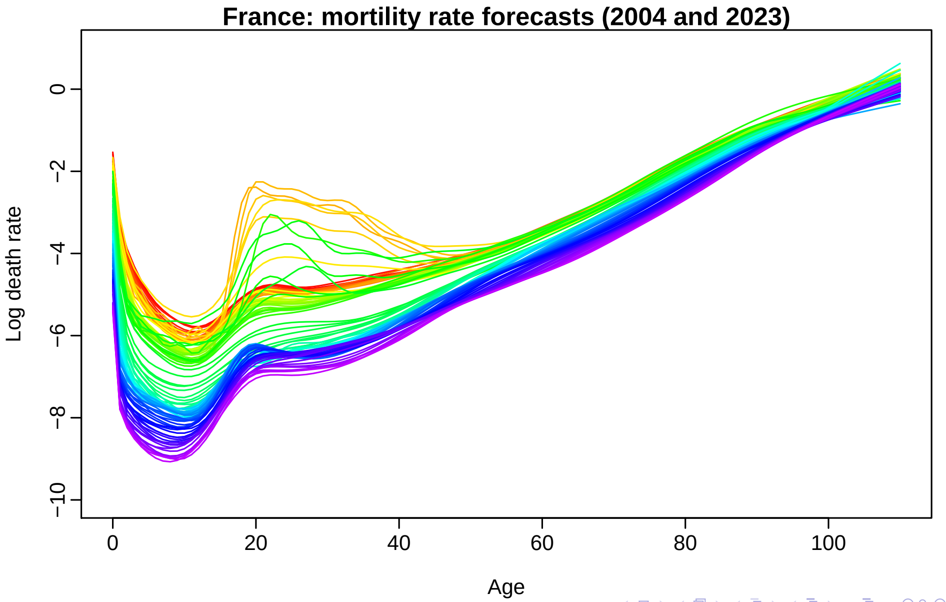
$$y_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\,\phi_k(x) + e_t(x)$$

- The eigenfunctions $\phi_k(x)$ show the main regions of variation.
- The scores $\{\beta_{t,k}\}$ are uncorrelated by construction. So we can forecast each $\beta_{t,k}$ using a univariate time series model.
- Outliers are treated as missing values.
- Univariate ARIMA models are used for forecasting.

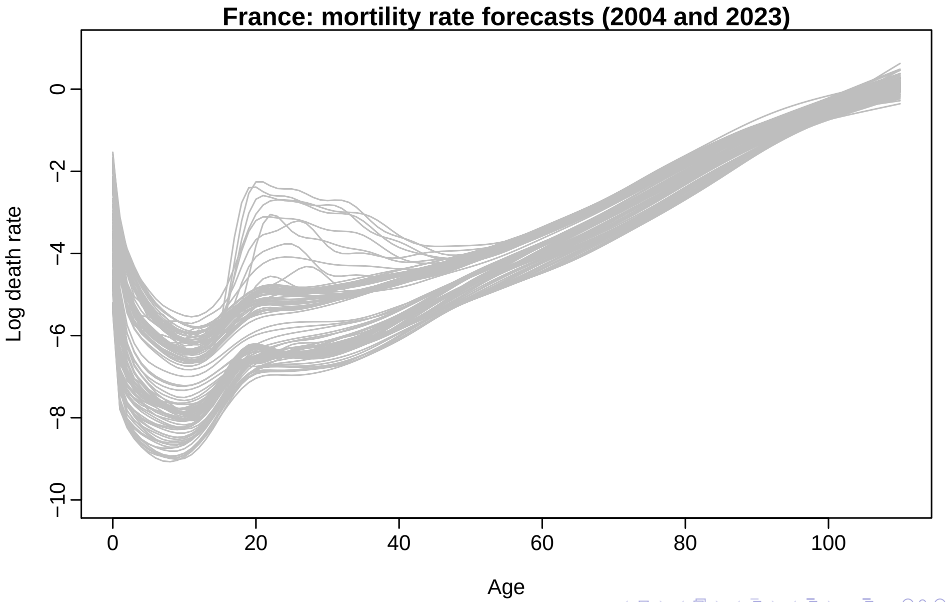
Forecasting the PC scores



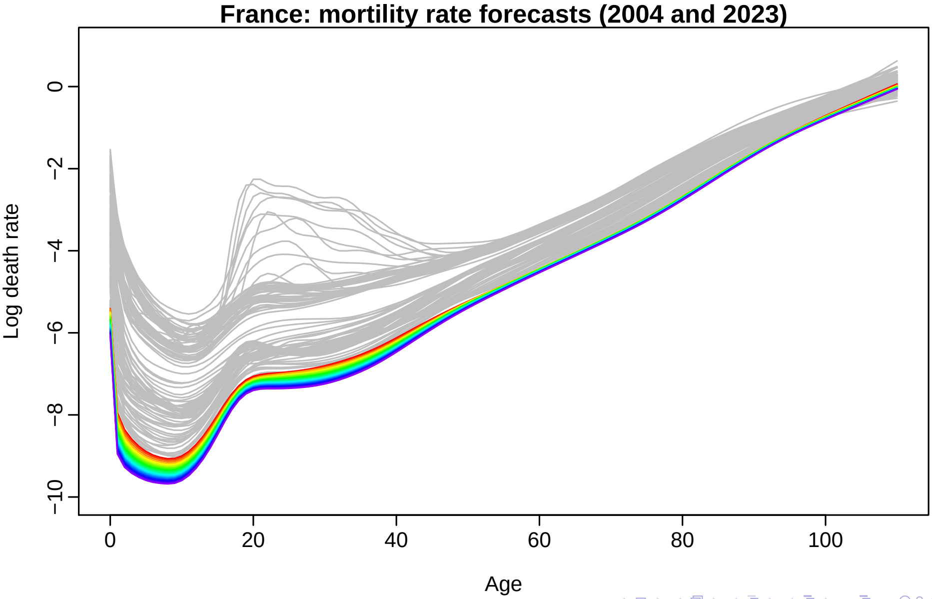
Forecasts of $y_t(x)$



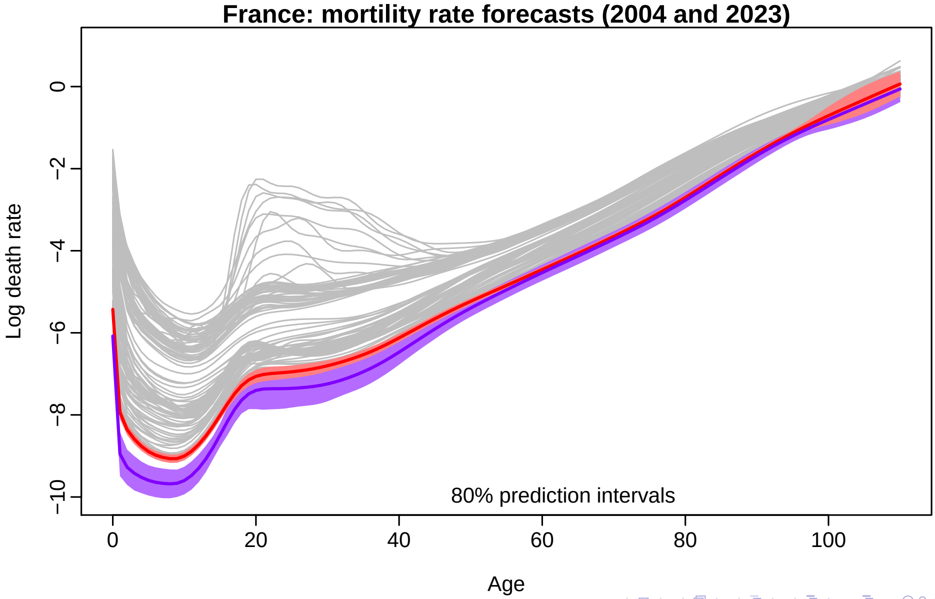
Forecasts of $y_t(x)$



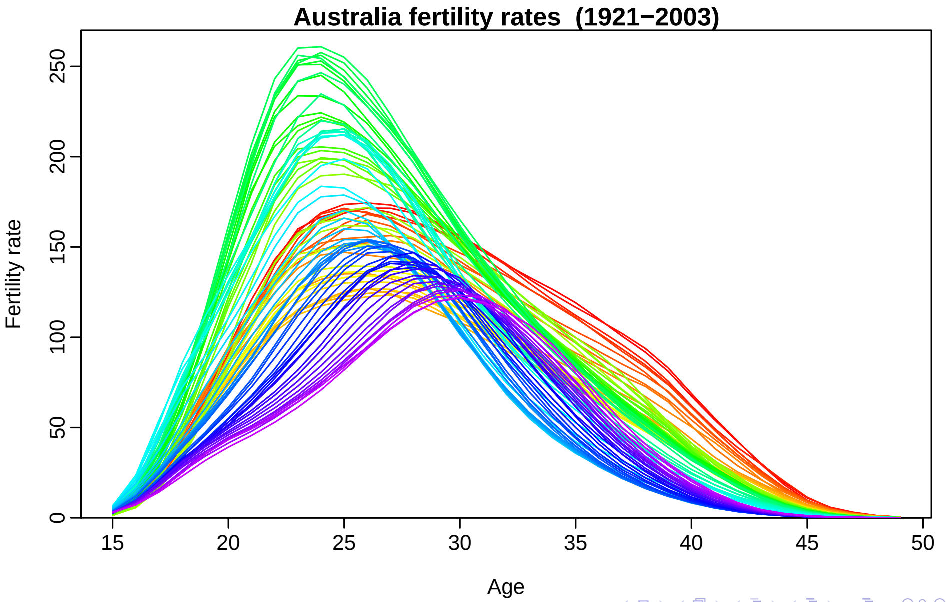
Forecasts of $y_t(x)$



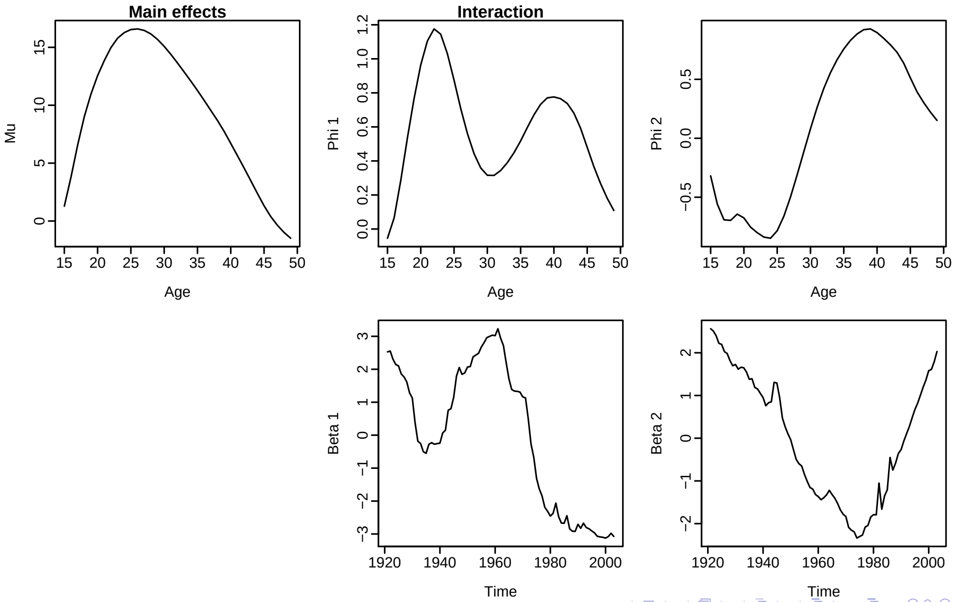
Forecasts of $y_t(x)$



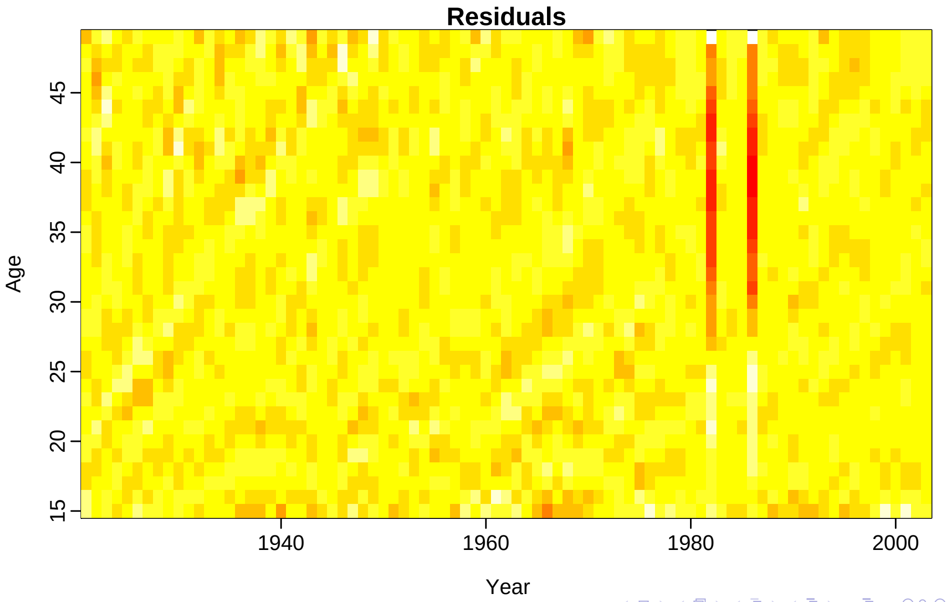
Fertility application



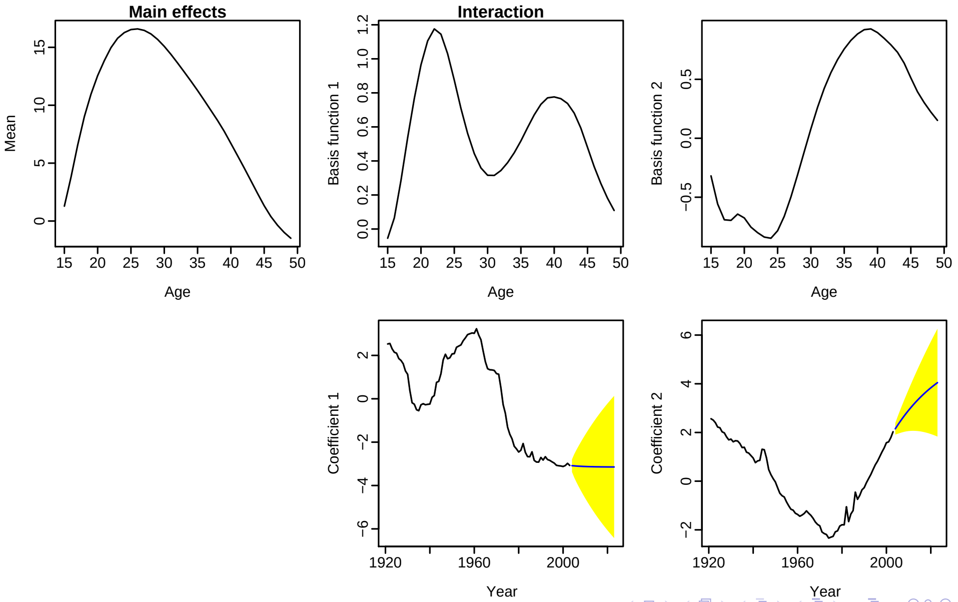
Fertility model



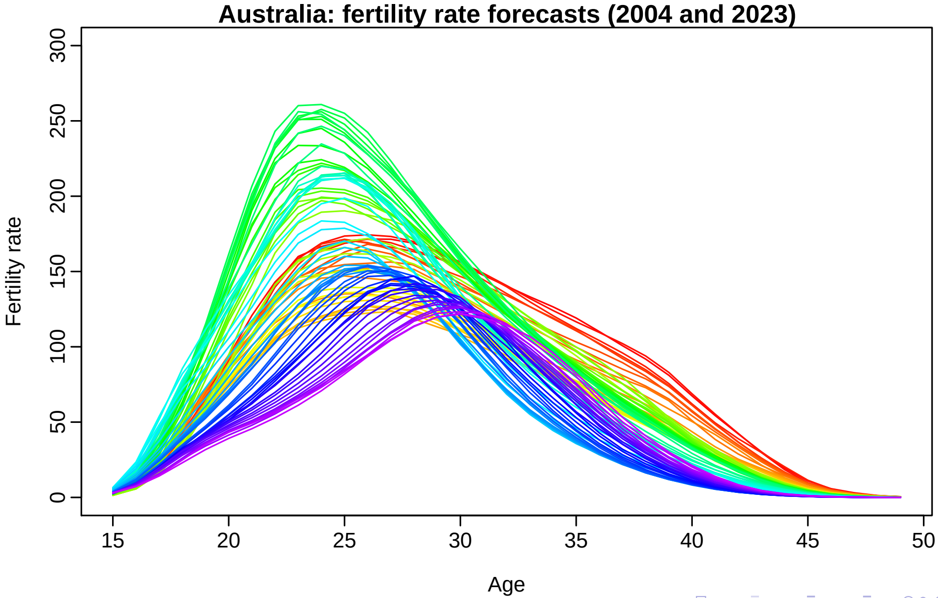
Fertility model



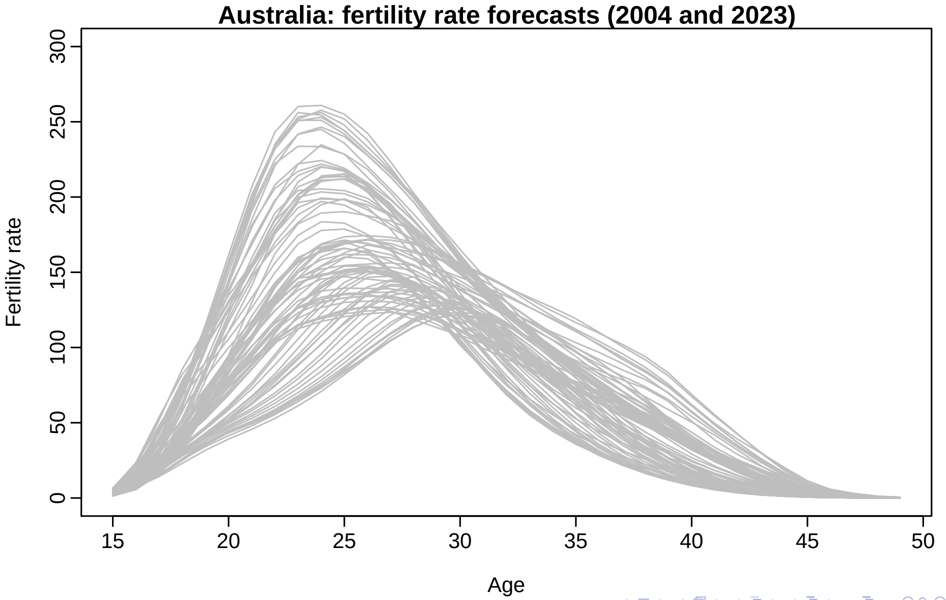
Fertility model



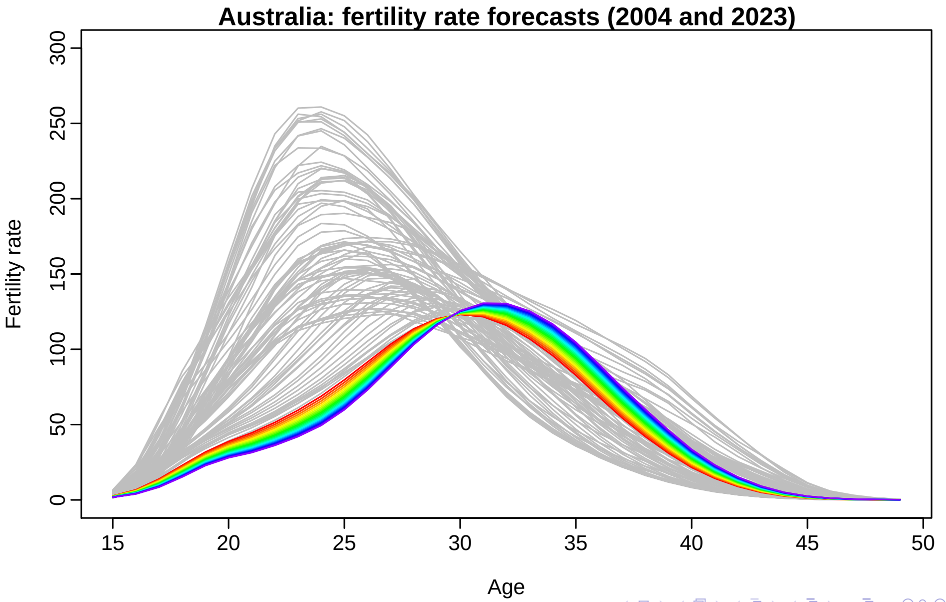
Forecasts of $y_t(x)$



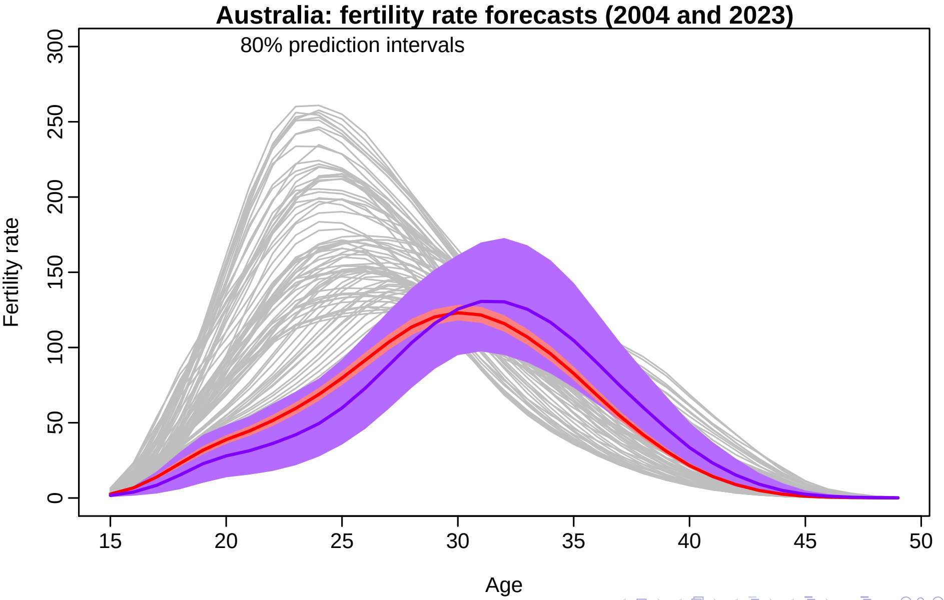
Forecasts of $y_t(x)$



Forecasts of $y_t(x)$



Forecasts of $y_t(x)$



Research team

MONASH University

- Bircan Erbas
- Han Lin Shang
- Farah Yasmeen
- Shahid Ullah
- Abdul-Aziz Hayat

Research team

MONASH University

- Bircan Erbas
- Han Lin Shang
- Farah Yasmeen
- Shahid Ullah
- Abdul-Aziz Hayat

Things we are working on

- cohort effects
  (our current models only allow period effects)

Research team

MONASH University

- Bircan Erbas
- Han Lin Shang
- Farah Yasmeen
- Shahid Ullah
- Abdul-Aziz Hayat

Things we are working on

- cohort effects
  (our current models only allow period effects)
- synergy and differences between groups
  (e.g., multi-racial mortality, breast-screen effects)

Research team

MONASH University

- Bircan Erbas
- Han Lin Shang
- Farah Yasmeen
- Shahid Ullah
- Abdul-Aziz Hayat

Things we are working on

- cohort effects
  (our current models only allow period effects)
- synergy and differences between groups
  (e.g., multi-racial mortality, breast-screen effects)
- choosing a basis better suited to prediction *(rather than explaining historical variation)*

Applications so far

- Fertility rates

Applications so far

- Fertility rates
- All-cause mortality rates

Applications so far

- Fertility rates
- All-cause mortality rates
- Population

Applications so far

- Fertility rates
- All-cause mortality rates
- Population
- Cancer mortality and incidence rates

Applications so far

- Fertility rates
- All-cause mortality rates
- Population
- Cancer mortality and incidence rates
- Yield curves in finance

Applications so far

- Fertility rates
- All-cause mortality rates
- Population
- Cancer mortality and incidence rates
- Yield curves in finance
- Seasonal electricity demand

Applications so far

- Fertility rates
- All-cause mortality rates
- Population
- Cancer mortality and incidence rates
- Yield curves in finance
- Seasonal electricity demand
- Seasonal El Niño sea-surface temperatures

Applications so far

- Fertility rates
- All-cause mortality rates
- Population
- Cancer mortality and incidence rates
- Yield curves in finance
- Seasonal electricity demand
- Seasonal El Niño sea-surface temperatures

Applications so far

- Fertility rates
- All-cause mortality rates
- Population
- Cancer mortality and incidence rates
- Yield curves in finance
- Seasonal electricity demand
- Seasonal El Niño sea-surface temperatures

More information

Papers and R code available at

www.robhyndman.info