Institute and Faculty of Actuaries

PhD studentship output

Funded by the Institute and Faculty of Actuaries

Institute
and Faculty
of Actuaries

Coherent mortality forecasting: the weighted multilevel functional principal component approach

Ruhao Wu
Department of Mathematics
University of Leicester

September 2016

Coherent mortality forecasting: the weighted multilevel functional principal component approach

- Traditional independent mortality forecasting methods (Lee-Miller model, FDA model) tend to result in divergent forecasts for subpopulations

- Under closely related social, economic and biological backgrounds, mortality patterns of subpopulations within one large population are expected to be non-divergent in long run

- Desirable to model their mortality rates simultaneously while taking into account the heterogeneity among them

Multilevel FPCA

- In practice, sometimes a set of functional data comprise a number of subsets with strong correlations
- A two-way functional ANOVA model:

$$Y_{i,j}(x) = \mu(x) + \eta_j(x) + Z_i(x) + W_{i,j}(x)$$

- Using the Karhunen-Loève (KL) expansion:

$$Z_i(x) = \sum_k \beta_{i,k}\, \phi_k^{(1)}(x),\, W_{i,j}(x) = \sum_l \gamma_{i,j,l}\, \phi_l^{(2)}(x)$$

- Model expressed as:

$$Y_{i,j}(x) = \mu(x) + \eta_j(x) + \sum_k \beta_{i,k}\, \phi_k^{(1)}(x) + \sum_l \gamma_{i,j,l}\, \phi_l^{(2)}(x)$$

- $\{\phi_k^{(1)}(x): k = 1,2, \dots\}$, $\{\phi_l^{(2)}(x): l = 1,2, \dots\}$ orthonormal bases, $\{\beta_{i,k}: k = 1,2, \dots\}$ uncorrelated with $\{\gamma_{i,j,l}: l = 1,2, \dots\}$

Estimate the principal components

- $\widehat{K}_T(x_s, x_r) = \frac{1}{IJ}\sum_{i,j}\{Y_{i,j}(x_s) - \hat{\mu}(x_s) - \hat{\eta}_j(x_s)\}\{Y_{i,j}(x_r) - \hat{\mu}(x_r) - \hat{\eta}_j(x_r)\}$
- $\widehat{K}_B(x_s, x_r) = \frac{2}{IJ(J-1)}\sum_i \sum_{j_1<j_2}\{Y_{i,j_1}(x_s) - \hat{\mu}(x_s) - \hat{\eta}_{j_1}(x_s)\}\{Y_{i,j_2}(x_r) - \hat{\mu}(x_r) - \hat{\eta}_{j_2}(x_r)\}$
- $\widehat{K}_W(x_s, x_r) = \widehat{K}_T(x_s, x_r) - \widehat{K}_B(x_s, x_r)$

- Decompose $\widehat{K}_B(x_s, x_r)$ to obtain $\hat{\lambda}_k^{(1)}$, $\hat{\phi}_k^{(1)}(x)$
- Decompose $\widehat{K}_W(x_s, x_r)$ to obtain $\hat{\lambda}_l^{(2)}$, $\hat{\phi}_l^{(2)}(x)$

Weighted MFPCA for coherent mortality forecasting

- Observed $\{x_i, y_{t,j}(x_i)\}$, assume a underlying function $f_{t,j}(x)$ with error:
$$y_{t,j}(x_i) = f_{t,j}(x_i) + \sigma_{t,j}(x_i) e_{t,j,i}$$
- Incorporate weight into MFPCA, $w_t = \kappa(1-\kappa)^{n-t}$, a geometrically decaying weight with $0 < \kappa < 1$
- The entire weighted MFPCA model:
$$y_{t,j}(x_i) = \mu_j(x_i) + \sum_{k=1}^{N_1} \beta_{t,k} \phi_k^{(1)}(x_i) + \sum_{l=1}^{N_2} \gamma_{t,j,l} \phi_l^{(2)}(x_i) + \sigma_{t,j}(x_i) e_{t,j,i}$$
- Independent possibly non-stationary ARIMA models to extrapolate each of $\{\beta_{t,1}, \dots, \beta_{t,N_1}\}$; a univariate ARMA model with stationary restriction to extrapolate each of $\{\gamma_{t,j,1}, \dots, \gamma_{t,j,N_2}\}, j = 1, \dots m$

Weighted MFPCA for coherent mortality forecasting

- $\hat{\beta}_{(t+h),k}$ denote the *h*-step ahead forecast of $\beta_{(t+h),k}$ and $\hat{\gamma}_{(t+h),j,l}$ denote the *h*-step ahead forecast of $\gamma_{(t+h),j,l}$
- The *h*-step ahead forecast of $y_{t,j}(x)$ is obtained as:

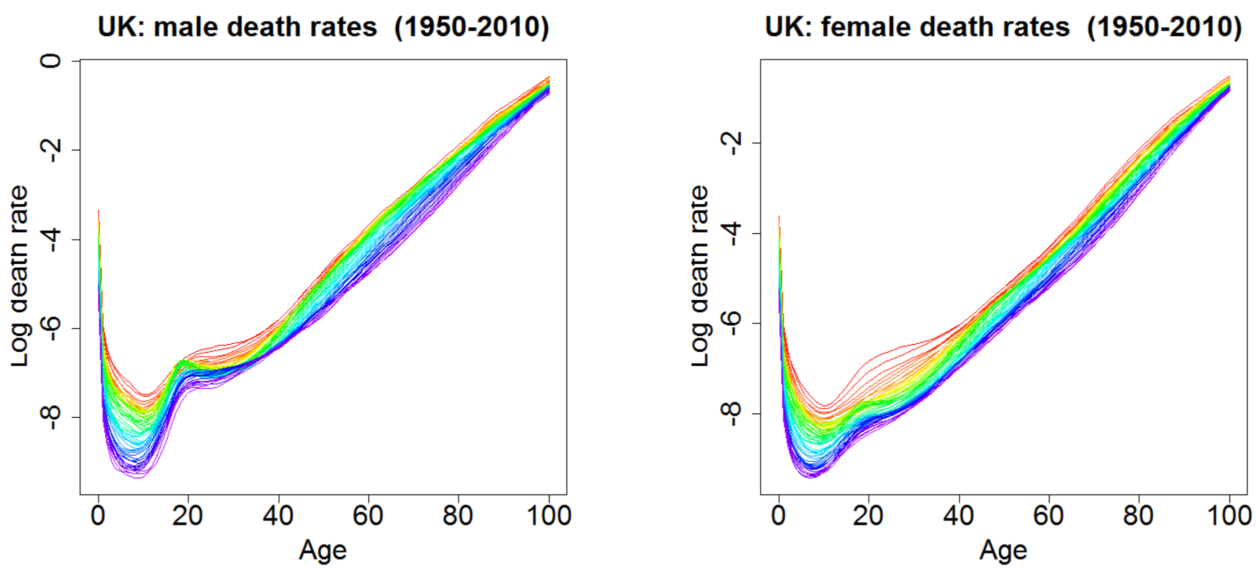
$$\hat{y}_{(t+h),j}(x) = \hat{\mu}_j(x) + \sum_{k=1}^{N_1} \hat{\beta}_{(t+h),k}\hat{\phi}_k^{(1)}(x) + \sum_{l=1}^{N_2} \hat{\gamma}_{(t+h),j,l}\hat{\phi}_l^{(2)}(x)$$

- The forecasting variance can be obtained by adding up the variance of each component:

$$var\{y_{(t+h),j}(x)\} = \hat{\sigma}_{\mu_j}^2(x) + \sum_{k=1}^{N_1} u_{(t+h),k}\{\hat{\phi}_k^{(1)}(x)\}^2 + \sum_{l=1}^{N_2} v_{(t+h),j,l}\{\hat{\phi}_l^{(2)}(x)\}^2 + \{\sigma_{(t+h),j}(x)\}^2$$

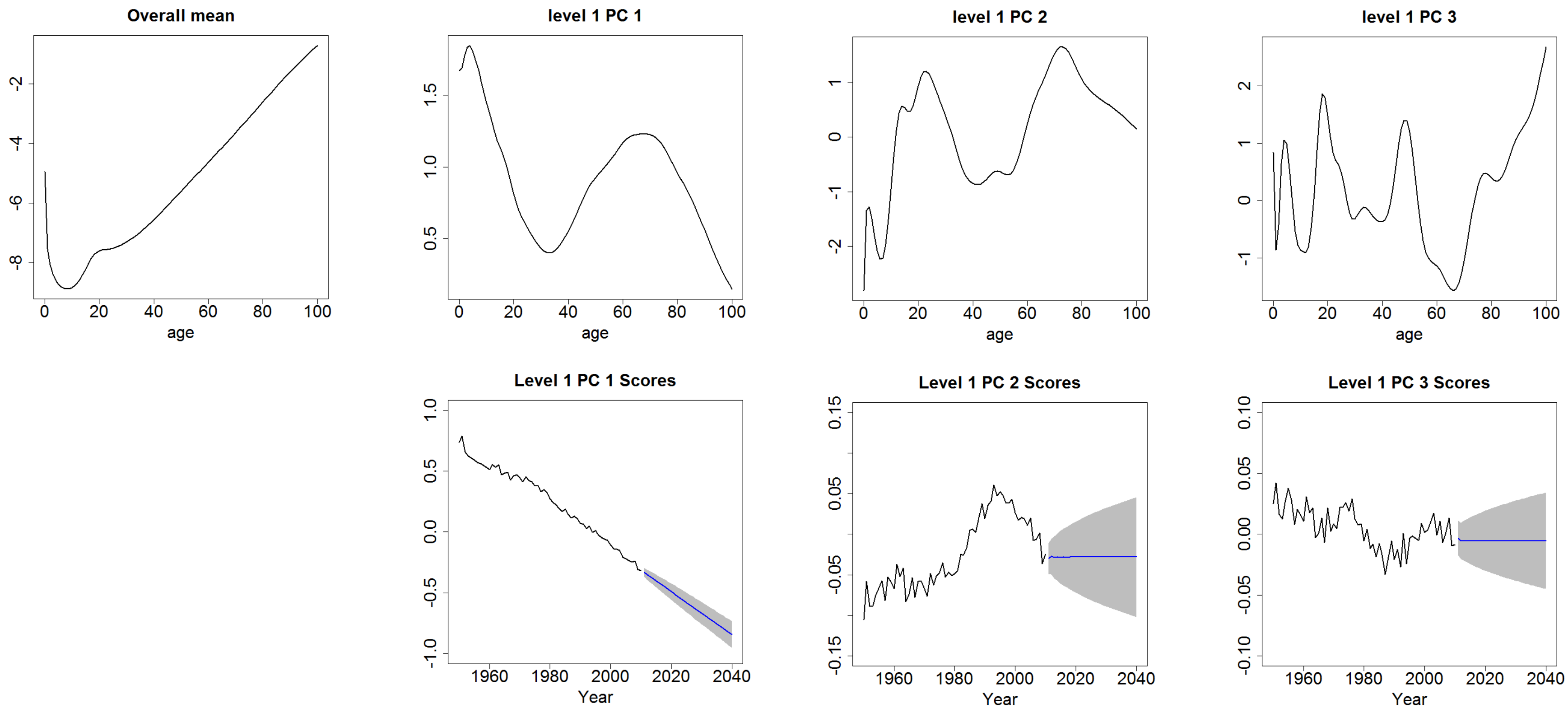
Coherent forecasting for the male and female mortality in the UK

- The smoothed log death rates for male and female in the UK from 1950 to 2010, viewed as functional data series



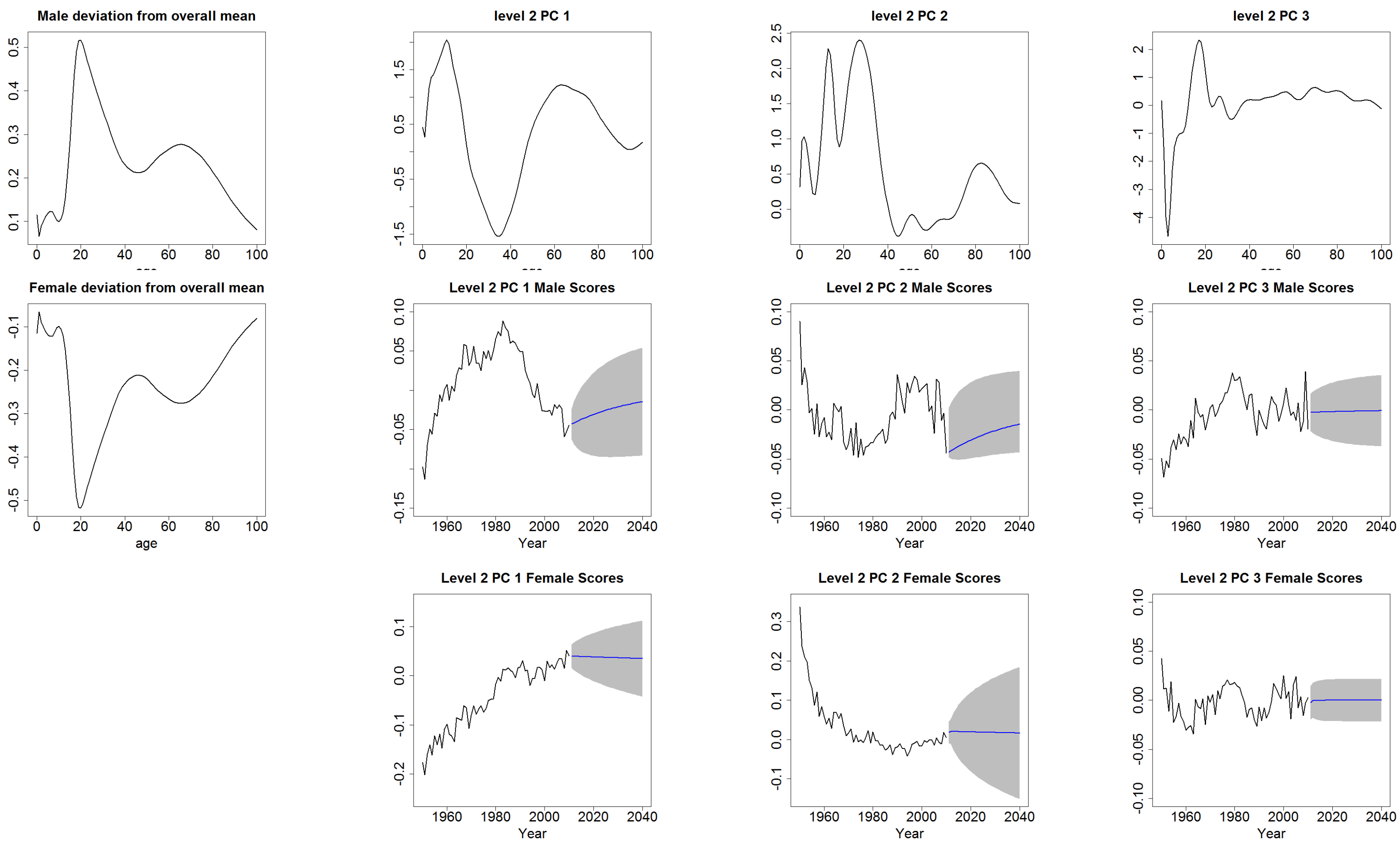
Coherent forecasting for the male and female mortality in the UK

- Level-one decomposition:



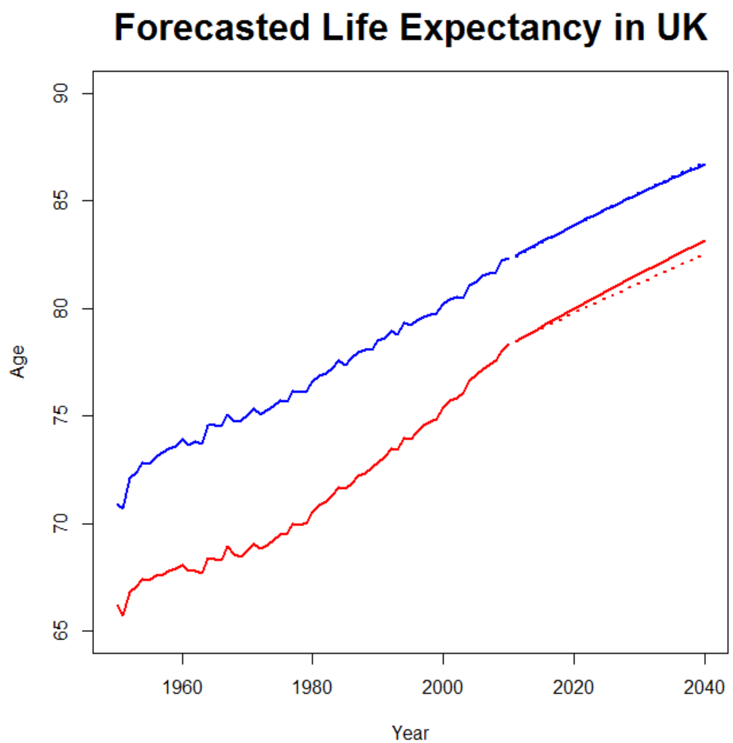
Coherent forecasting for the male and female mortality in the UK

- Level-two decomposition:



Coherent forecasting for the male and female mortality in the UK

- The 30-year forecasts of the male and female life expectancies at birth by weighted MFPCA model and the independent model



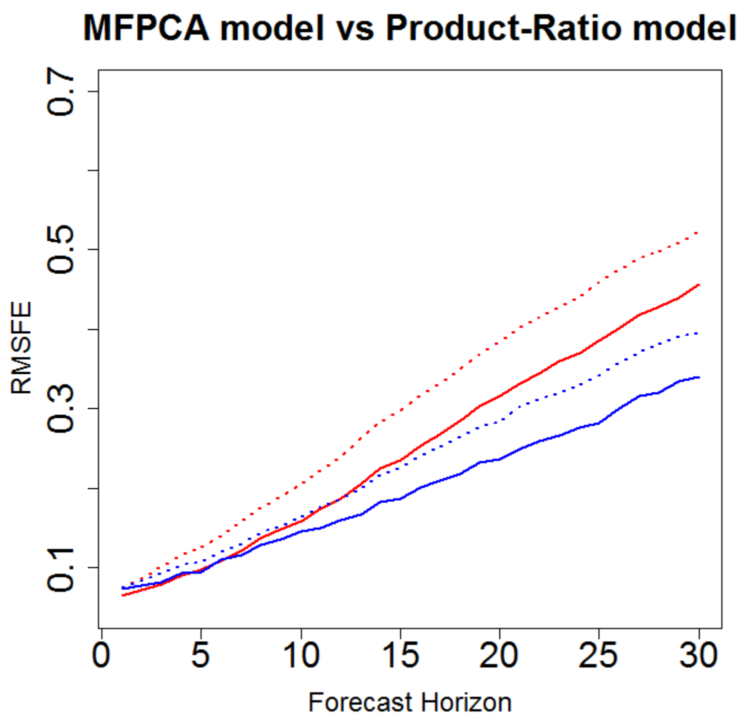
Comparing accuracy with the Product-Ratio model and the independent model

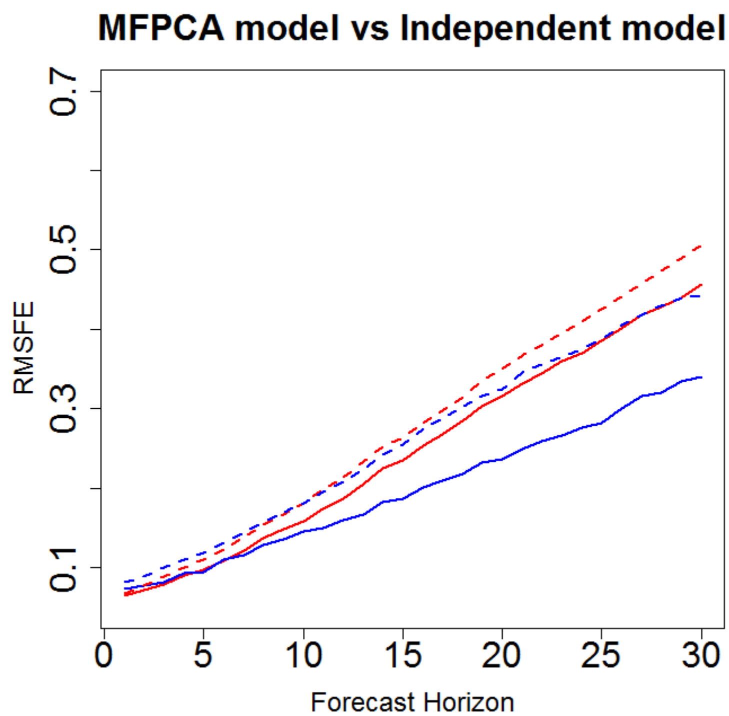
- Use the UK male and female mortality data from 1950 to 1973+t as observations and forecast the mortality rates for years 1973+$t+1$, …, 1973+$t+30$, for $t = 0, \dots, 9$

- For a specific forecast horizon h ($h = 1, \dots, 30$), the out-of-sample root mean square forecast error (RMSFE) for the j^{th} subpopulation is defined as:

$$RMSFE_j(h) = \sqrt{\frac{1}{10p}\sum_{t=0}^{9}\sum_{i=1}^{p}\{y_{(1973+t+h),j}(x_i) - \hat{y}_{(1973+t+h),j}(x_i)\}^2}$$

Comparing accuracy with the Product-Ratio model and the independent model





Comparing accuracy with the Product-Ratio model and the independent model

- Compute the average RMSFE (over forecast horizon and sex) for 9 developed countries, including Australia, USA, UK, France, Japan, Spain, Canada, Netherlands and Italy

	MFPCA	Product-Ratio	Independent
AUS	0.2774	0.2757	0.2806
USA	0.1568	0.1247	0.1614
UK	0.1916	0.2672	0.2685
FRA	0.2483	0.2188	0.2362
JPN	0.3616	0.3551	0.3614
ESP	0.2855	0.2766	0.3404
CAN	0.2353	0.2039	0.2451
NLD	0.2415	0.2383	0.2851
ITA	0.2512	0.2572	0.2694

Comparing accuracy with the Product-Ratio model and the independent model

- Compute the short-term RMSFE (average RMSFE for 1 to 10-year horizon)

	MFPCA	Product-Ratio	Independent
AUS	0.1548	0.1592	0.1617
USA	0.0927	0.0886	0.0927
UK	0.1065	0.1271	0.1243
FRA	0.1115	0.1116	0.1139
JPN	0.1262	0.1507	0.1522
ESP	0.1615	0.1618	0.1801
CAN	0.1406	0.1305	0.1425
NLD	0.1547	0.1586	0.1697
ITA	0.1182	0.1220	0.1243

Vs. the Product-Ratio model

- MFPCA: a group mean, a decomposition of level-one function and level-two function
- P-R model: a group mean, a decomposition of product function and ratio function

- Advantages:

1. No need to pre-processing the data as done in the Product-Ratio method
2. No need to assume the subpopulations have equal variances
3. The percentage of variance explained by each principal component at both levels can be calculated explicitly and easily