

Bayesian functional models for population forecasting

Han Lin Shang^{a,*}, Arkadiusz Wiśniowski^{a,b}, Jakub Bijak^{a,b}, Peter W.F. Smith^{a,b}, and
James Raymer^{a,c}

^a*ESRC Centre for Population Change, University of Southampton*

^b*Southampton Statistical Sciences Research Institute, University of Southampton*

^c*Australian Demographic and Social Research Institute, The Australian National University*

Abstract

We explore the functional modelling approach to population forecasting within the wider context of Bayesian predictions and model uncertainty. The functional modelling approach can be used to analyse and forecast many different age- and time-specific components for fertility, mortality and migration. For each of these demographic processes, we perform Bayesian model averaging across the outcomes of two functional models to take into account model uncertainty. We illustrate the method with a population forecast for the United Kingdom for 2010–2030. We conclude that regularities in age profiles of demographic processes, where available, provide important information for the forecasts and as such should be included in the forecasting process.

Keywords: Age schedules, Model averaging, Model selection, Functional models, Lee-Carter model, Model uncertainty, Population forecasting

1. Introduction

In this paper, we extend the population forecasting framework developed in [Wiśniowski et al. \(2013\)](#), which utilises the well-known model of [Lee and Carter \(1992\)](#), to explore the functional modelling approach to population forecasting. We additionally embed this framework in a wider context of Bayesian predictions and model uncertainty. The underlying functional modelling, originally suggested by [Hyndman and Ullah \(2007\)](#), allows for analysing and forecasting many different age- and time-specific components for each of the three main processes of population dynamics, namely fertility, mortality and migration. For each process we perform Bayesian model averaging across the outcomes of two functional

*Corresponding author

Email address: H.Shang@soton.ac.uk (Han Lin Shang)

models, in order to take into account the model uncertainty and a varying level of data support for different models. Finally, we combine the results in a joint cohort-component framework to obtain a hybrid Bayesian population forecast. To be more specific, we use Bayesian model averaging to assign weights to the two functional models. Having determined weights, we then use time-series models to produce point and interval forecasts.

The method is illustrated by a population forecast for the United Kingdom (UK) for 2010-2030, which we compare against the official population projection of the Office for National Statistics (ONS). We discuss the importance of various modifications of the basic functional approach, as well as to smoothing irregular age patterns of international migration. The proposed method is argued to offer more flexibility than those based on single forecasting models, whilst allowing for a coherent treatment of various types of uncertainty thanks to the overarching Bayesian statistical framework.

The rest of this paper is structured as follows. In Section 2 we present the proposed framework for population forecasting by using the functional models. The description of the applied cohort-component model of population renewal is followed by a brief introduction to the functional analysis, and a recapitulation of Bayesian model selection and averaging. In Section 3, we present the age-specific mortality, fertility, emigration rates and immigration counts for the UK. Selected results are shown in Section 4, for the example of the population of the UK, forecasted until 2030. Finally, Section 5 offers some concluding remarks and suggestions for further work.

2. Functional Models for Population Forecasting

2.1. Population Renewal Framework

A cohort component model is used for describing the evolution of an age-specific population (see also [Preston et al., 2001](#); [Wiśniowski et al., 2013](#)). For each age or age group, we need to estimate age-specific fertility, mortality, emigration rates and immigration counts. We work with single year of age and, instead of modelling net migration, we choose to model emigration rates and immigration counts separately for the reasons given in [Rees \(1986\)](#) and [Raymer et al. \(2012\)](#). The accurate estimation and forecast of mortality is important for the calculation of the survival rate s_i for years $i = 0, \dots, 89, 90+$. The survival rate measures the proportion of age i , which will survive to the next period of time. Second, the fertility rate, f_j , for $j = 13, \dots, 51$ measures the yearly average number of

surviving offspring per woman aged j . Third, the emigration rate η_i for $i = 0, \dots, 89, 90+$ measures the average yearly number of emigrants of age i relative to the population exposure. Finally, the immigration count, I_i , measures the total number of immigrants of age i .

Let $P_{i,t}^F$ and $P_{i,t}^M$ denote the number of females and males of age i at the beginning of year t . The relationship between consecutive periods of times can be expressed by means of a projection matrix, given by

$$\begin{bmatrix} \mathbf{P}_{t+1}^F \\ \mathbf{P}_{t+1}^M \end{bmatrix} = \begin{bmatrix} a\mathbf{b}_t^F & \mathbf{0} \\ \mathbf{s}_t^F & \mathbb{O} \\ (1-a)\mathbf{b}_t^M & \mathbf{0} \\ \mathbb{O} & \mathbf{s}_t^M \end{bmatrix} \times \begin{bmatrix} \mathbf{P}_t^F \\ \mathbf{P}_t^M \end{bmatrix} + \begin{bmatrix} \mathbf{I}_t^F \\ \mathbf{I}_t^M \end{bmatrix}, \quad (1)$$

where $\mathbf{P}_t^k$, $k = M$ or F , denotes the male and female population, respectively, for all ages at the beginning of year t , and $a = 1/(1 + 1.05)$ is the assumed proportion to female births in the population (Preston et al., 2001). The $\mathbf{b}_t^k = (0, \dots, b_{13,t}^k, \dots, b_{51,t}^k, \dots, 0)$ is a vector of life-table birth rates, which can be derived from the age-specific fertility rates as follows

$$b_{i,t}^k = \frac{1}{1 + 0.5\mu_{0,t}^k} \frac{1}{2} (f_{i,t} + s_{i,t}^F f_{i+1,t}),$$

where $f_{i,t} > 0$ represents the fertility rate at age i in year t ; $\mu_{0,t}^k$ represents the age-specific female or male mortality rate at age 0; $s_{i,t}^F$ represents female survival rate at age i in year t . For males and females, the age-specific survival rate can be estimated from age-specific mortality and emigration rates. It is defined by

$$s_{i,t}^k = \begin{cases} \frac{1-0.5(\mu_{i,t}^k + \eta_{i,t}^k)}{1+0.5(\mu_{i+1,t}^k + \eta_{i+1,t}^k)} & \text{if } i = 0, \dots, 89 \\ \frac{1-0.5(\mu_{i,t}^k + \eta_{i,t}^k)}{1+0.5(\mu_{i,t}^k + \eta_{i,t}^k)} & \text{if } i = 90+ \end{cases},$$

where $\mu_{i,t}$ represents the mortality rate at age i in year t ; $\eta_{i,t}$ represents the emigration rate at age i in year t . As shown in Preston et al. (2001), the survival rate matrix for all

ages can then be expressed as

$$\mathbf{s}_t^k = \begin{bmatrix} s_{0,t}^k & 0 & 0 & \dots & 0 \\ 0 & s_{1,t}^k & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & s_{88,t}^k & 0 & 0 \\ 0 & 0 & \dots & 0 & s_{89,t}^k & s_{90+,t}^k \end{bmatrix}.$$

Note that in Equation (1), $\mathbf{0} = (0, \dots, 0)$ is a vector of length 91, $\mathbb{O}$ is a matrix of zeros of size (90×91) , and $\mathbf{I}_t^F = (I_{0,t}^F, \dots, I_{90+,t}^F)'$ represents the vector of immigration counts at year t , and $'$ represents vector transpose.

2.2. Functional Models: Preliminaries

Recent advances in computer recording and storing facilities allow statisticians to analyse high-dimensional data that include many variables, such as mortality for different ages. The objective of the functional data analysis is to analyse a set of underlying functions, usually smooth and bounded within an interval; such functional data may be age-specific mortality or fertility rates (see for example, [Hyndman and Ullah, 2007](#); [Hyndman and Shang, 2009](#); [Hyndman et al., 2013](#)). [Ramsay and Silverman \(2005\)](#) and [Ferraty and Vieu \(2006\)](#) provided detailed surveys of many parametric and nonparametric techniques for analysing functional data, and some recent developments are collected in the edited books by [Ferraty and Romain \(2011\)](#) and [Ferraty \(2011\)](#).

The functional data modelling process ([Hyndman and Ullah, 2007](#)) can be summarised into the following steps:

1. **Box-Cox transformation.** Let $m_{i,t}$ represent the original data for age $i = 0, \dots, 89, 90+$ in year t . The Box-Cox transformation is applied to each component of population in order to alleviate heteroscedasticity and can be specified as

$$g_{i,t} = \begin{cases} \frac{1}{\xi}(m_{i,t}^\xi - 1) & \text{if } 0 < \xi \leq 1; \\ \ln(m_{i,t}) & \text{if } \xi = 0. \end{cases} \quad t = 1, 2, \dots, n,$$

where $g_{i,t}$ represents the transformed data at age i in year t .

2. **Pre-smoothing step.** Since the object in functional data analysis is a smooth function, we apply a smoothing technique to transform a set of discrete data points

to such a function. As a result, there is a change of notation from $g_{i,t}$ to $g_t(x_i)$, where x_i represents the discrete observations. It is assumed that there is an underlying continuous and smooth function $\tau_t(x)$ that is observed with error at discrete ages. Then, we can write

$$g_t(x_i) = \tau_t(x_i) + \sigma_t(x_i)\varepsilon_{i,t},$$

where $\sigma_t(x_i)$ models the variability for each age x_i in year t , and $\varepsilon \sim N(0, 1)$ is an independent and identically distributed random variable. The values of $\tau_t(x_i)$ and $\sigma_t(x_i)$ are estimated from data (see [Hyndman and Ullah, 2007](#), for detail).

For modelling age-specific mortality, we utilise penalised regression splines with a partial monotonic constraint for age above 65 (see [Ramsay, 1988](#); [Hyndman and Ullah, 2007](#), for more details). For modelling age-specific fertility, we use a weighted median smoothing B -spline, constrained to be concave (see [He and Ng, 1999](#)). For modelling age-specific emigration rates and immigration counts, a smoothing spline is used where the smoothing parameter is automatically determined by generalised cross validation (see [Wahba, 1990](#)).

The functional models can be either independent or coherent, where the latter one is done jointly for both sexes.

- 3a. **Decomposition step for the independent functional model.** By using functional principal component analysis (FPCA), a set of functions is decomposed into orthogonal functional principal components and their associated scores. The functional principal component decomposition is given by

$$\tau_t(x) = \mu(x) + \sum_{k=1}^K \beta_{t,k} \phi_k(x) + e_t(x), \quad x \in [0, 90+], \quad (2)$$

where $\mu(x)$ is the mean function; $\{\phi_1(x), \dots, \phi_K(x)\}$ is a set of the first K functional principal components; $\{\beta_{t,1}, \dots, \beta_{t,K}\}$ is a set of the corresponding principal component scores; $e_t(x) \sim N(0, \sigma^2)$ is the residual function with mean zero and finite variance; and $K < n$ is the number of retained components. In practice, mean function can be estimated by $\hat{\mu}(x) = \frac{1}{n} \sum_{t=1}^n \tau_t(x)$; $\{\hat{\phi}_1(x), \dots, \hat{\phi}_K(x)\}$ can be obtained from singular value decomposition; $\hat{\beta}_{t,k}$ and $\hat{\sigma}^2$ are drawn from their respective posterior distributions, where the prior densities of the variance parameters associated with the principal component scores and model error term are inverse

gamma distributions with hyperparameters 10^{-3} and 10^{-3} . Throughout the paper, we select the number of components that explain at least 99% of the total variation in data.

- 3b. **Decomposition step for the coherent functional model.** Equation (2) is designed to model a single population, such as for fertility for females. For analysing female and male mortality jointly, we adapt the multilevel functional data model to analyse two subpopulations that may be correlated (see for example, [Li and Lee, 2005](#); [Hyndman et al., 2013](#)). The basic idea is to decompose functions from different subpopulations into an aggregated average, a common trend, a sex-specific trend and measurement error. The common and sex-specific trends are modelled by projecting them onto the eigenvectors of covariance operators of the aggregated and sex-specific centred stochastic processes, respectively. For example, the smoothed female mortality rate at year t can be expressed as

$$\tau_t^F(x) = \mu(x) + w^F(x) + R_t(x) + U_t^F(x) + \varepsilon_t^F(x), \quad (3)$$

where each term in (3) can be estimated by

$$\begin{aligned} w^F(x) &= \mu^F(x) - \mu(x), \\ R_t(x) &\approx \sum_{k=1}^K \beta_{t,k} \phi_k(x), \\ U_t^F(x) &\approx \sum_{l=1}^L \gamma_{t,l}^F \psi_l^F(x). \end{aligned}$$

Following the work of [Crainiceanu and Goldsmith \(2010\)](#), we draw the principal component scores and the variance of model error from their posterior using the Bayesian paradigm. The values of K and L are determined by the principal components set to explain at least 95% and 90% of the total variations, respectively.

- 4a. **Forecasting step for the independent functional model.** Conditioning on the smoothed functions $\mathcal{I} = \{\tau_1(x), \dots, \tau_n(x)\}$ and the estimated set of functional principal components $\mathbf{B} = \{\hat{\phi}_1(x), \dots, \hat{\phi}_K(x)\}$, the h -step-ahead probabilistic forecast

of $m_{n+h}(x)$ can be obtained as

$$\begin{aligned}\widehat{m}_{n+h|n}^{(b)}(x) &= \mathbb{E}[m_{n+h}(x)|\mathcal{I}, \mathbf{B}] \\ &= \widehat{\mu}(x) + \sum_{k=1}^K \widehat{\beta}_{n+h|n,k}^{(b)} \widehat{\phi}_k(x) + \widehat{e}_{n+h|n}^{(b)}(x) + \widehat{\sigma}_{n+h}(x) \widehat{\epsilon}_{n+h}^{(b)},\end{aligned}$$

where $b = 1, \dots, B$, and $B = 1000$ represents the number of iterations, $\widehat{\beta}_{n+h|n,k}^{(b)}$ denotes the h -step-ahead forecast of $\beta_{n+h,k}$ using a univariate time series model, such as the optimal autoregressive integrated moving average (ARIMA) model selected by the automatic algorithm of Hyndman and Khandakar (2008) based on an information criterion, such as the corrected Akaike information criterion. Further, $\widehat{e}_{n+h|n}^{(b)}$ is simulated from a normal distribution with zero mean, $\widehat{\sigma}_{n+h}(x)$ represents the estimated variance from the historical observations, and $\widehat{\epsilon}_{n+h}^{(b)}$ is simulated from a standard normal distribution. By using the parametric bootstrap method, the prediction interval for the multilevel functional data model can be constructed similarly.

- 4b. **Forecasting step for the coherent functional model.** Conditioning on the smoothed functions $\mathcal{I} = \{\tau_1^F(x), \dots, \tau_n^F(x)\}$ and the estimated set of functional principal components $\mathbf{B} = \{\widehat{\phi}_1(x), \dots, \widehat{\phi}_K(x)\}$ and $\mathbf{L} = \{\widehat{\psi}_1^F(x), \dots, \widehat{\psi}_L^F(x)\}$, the h -step-ahead probabilistic forecast of $m_{n+h}(x)$ can be obtained as

$$\begin{aligned}\widehat{m}_{n+h|n}^{(b)}(x) &= \mathbb{E}[m_{n+h}(x)|\mathcal{I}, \mathbf{B}, \mathbf{L}] \\ &= \widehat{\mu}(x) + \widehat{w}^F(x) + \sum_{k=1}^K \widehat{\beta}_{n+h|n,k}^{(b)} \widehat{\phi}_k(x) + \sum_{l=1}^L \widehat{\gamma}_{n+h|n,l}^F \widehat{\psi}_l^F(x) + \widehat{e}_{n+h|n}^{(b)}(x) + \widehat{\sigma}_{n+h}(x) \widehat{\epsilon}_{n+h}^{(b)},\end{aligned}$$

2.3. Model Selection and Averaging

The model-averaging approach combines two or more forecasts from a set of candidate (nested or non-nested) models. Because these models may reflect different assumptions, model structures and degrees of model complexity, it is expected that more robust result can be obtained than the most accurate method alone. The idea of model averaging has been studied intensively in statistics, dating back to the seminal work by Bates and Granger (1969). Since then, a flurry of articles have appeared on this topic; see Clemen (1989) for a review from a frequentist viewpoint and Hoeting et al. (1999) for a review from a Bayesian

viewpoint. Recent developments in model averaging are collected in the monograph by [Claeskens and Hjort \(2008\)](#). In demographic forecasting, there has been limited usage, but notable exceptions include [Smith and Shahidullah \(1995\)](#), [Ahlburg \(1998, 2001\)](#) and [Sanderson \(1998\)](#) in the context of census tract forecasting, [Bijak \(2010\)](#) for migration forecasting, and [Shang \(2012\)](#) in the context of mortality and life expectancy forecasting.

Let $M_1, M_2, \dots, M_R$ be a set of R possible models, and let $\theta_1, \theta_2, \dots, \theta_R$ be the vector of parameters associated with each model. Denote Δ as the quantity of interest, such as a combined forecast of age-specific mortality, then its posterior distribution given data D is

$$\begin{aligned} \Pr(\Delta|D) &= \sum_{r=1}^R \Pr(\Delta|M_r, D) \Pr(M_r|D) \\ &= \sum_{r=1}^R \Pr(\Delta|M_r, D) \underbrace{\frac{\Pr(D|M_r)}{\sum_{l=1}^R \Pr(D|M_l) \Pr(M_l)}}_{\text{weight}}, \end{aligned} \quad (4)$$

where $\Pr(D|M_r) = \int \Pr(D|\theta_r, M_r) \Pr(\theta_r|M_r) d\theta_r$, $\Pr(\theta_r|M_r)$ is the prior density of θ_r under model M_r , $\Pr(D|\theta_r, M_r)$ is the likelihood, and $\Pr(M_r)$ is the prior probability that M_r is the true model. From a Bayesian viewpoint, Equation (4) is an average of the posterior distributions under different models considered, weighted by their posterior model probabilities ([Hoeting et al., 1999](#)).

In the statistical literature, there exist a number of methods for assigning weights, some of which include the Bates-Granger averaging ([Bates and Granger, 1969](#)), Granger-Ramanathan averaging ([Granger and Ramanathan, 1984](#)), information criterion averaging ([Burnham and Anderson, 2002](#)), Bayesian model averaging ([Raftery et al., 2005](#)). In this paper, we consider the harmonic mean estimator proposed by [Newton and Raftery \(1994\)](#) to approximate the marginal likelihood of each model.

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 (MCMC) iterations. The marginal likelihood can be written by

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

where T represents the total number of MCMC iterations; 1000 iterations are retained

after 5000 iterations used for burn in. For discussions on the efficiency of this estimator, see [Raftery \(1996\)](#) and [Raftery et al. \(2007\)](#)

3. Data

The historical UK population data include observations from 1975 to 2009, from which we aim to forecast population by age and sex from 2010 to 2030. To obtain such forecasts, it is essential to accurately estimate and forecast age-specific fertility, mortality and emigration rates, as well as immigration counts. The fertility data were obtained from the [Human Fertility Database \(2013\)](#), while the mortality data were obtained from the [Human Mortality Database \(2013\)](#). The emigration rates and immigration counts were obtained directly from the Office for National Statistics (ONS). The UK population has been obtained from [Human Mortality Database \(2013\)](#). The UK mid-year population estimate for 2009, used as a baseline for prediction, has been obtained from the ONS.

We consider mortality rates for single year of age for ages from 0 to 90+. For each gender in a given calendar year, the mortality rates, given by the ratio between the “number of deaths” and the “exposure to risk”, are arranged in a matrix by age and year. By analysing the changes in mortality as a function of both age x and year t , we have seen that mortality has shown a gradual decline over year. To have an idea of this evolution, we present the log mortality rates for ages 0-90+ from 1975 to 2009 in [Figure 1](#). Mortality rates dip in early childhood, climb in the teen years, stabilise in the early 20s, and then steadily increase with age. Some years exhibit sharp increases in mortality between the late teens and early 20s. In general, we notice that for both females and males, mortality rates are decreasing over time, especially for ages between 0 and 10. Males exhibit considerably higher mortality in young adulthood than females.

The age-specific fertility rates are defined as the number of live births during a given calendar year, according to the age of the mother among the female resident population of the same age at 30th June. Age-specific fertility rates between ages 13 and 51 from 1975 to 2009 are presented in [Figure 2](#). We notice that there is an increase in fertility rates at higher ages in more recent years caused by a tendency to postpone child bearing while women are pursuing careers ([Ní Bhrolcháin and Beaujouan, 2012](#)).

The total flows of emigration rates and immigration counts are presented in the top row of [Figure 3](#). We notice that migration for both genders follow a similar trend over

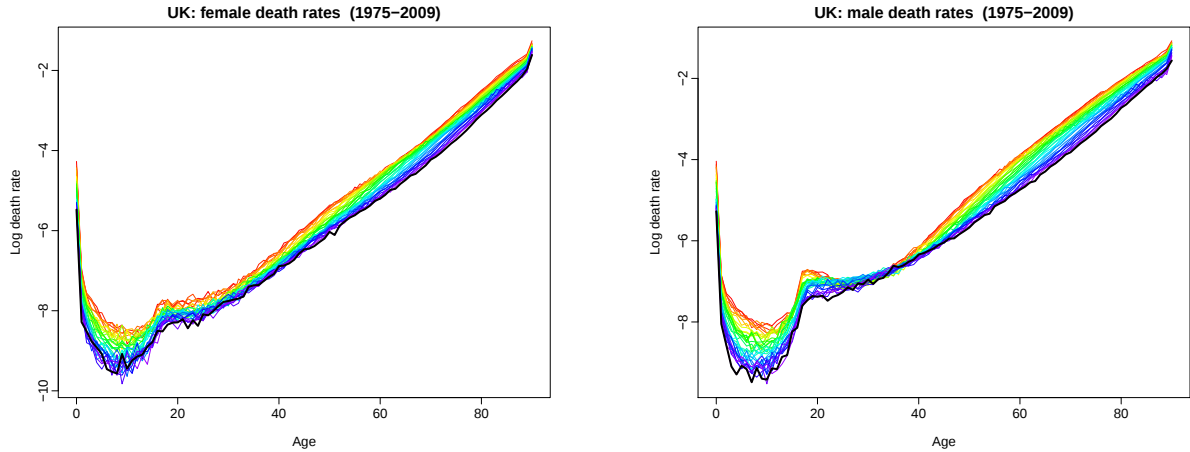


Figure 1: UK female and male age-specific log mortality rates (1975-2009). The thick black line represents the mortality in 2009

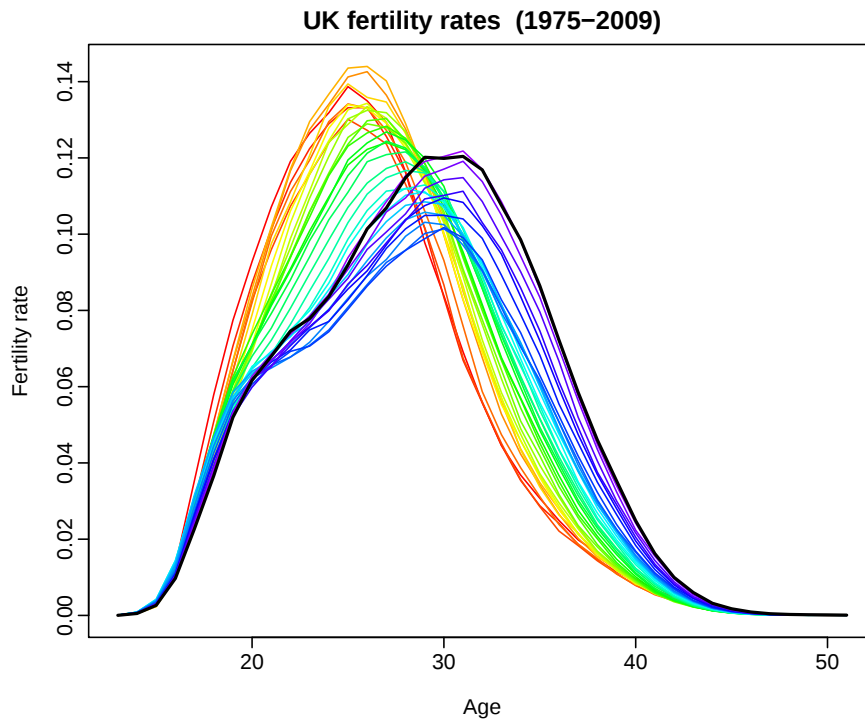


Figure 2: UK age-specific fertility rates between ages 13 and 51. The thick black line represents the fertility pattern in 2009

years. The immigration counts have been rapidly increasing since 1990 up until 2005. By contrast, there is a slight increase over time in the emigration data, but the patterns seem to be more volatile. One explanation for such a volatility stems from the fact that the data on emigration in the UK come from the International Passenger Survey (IPS), which has several pitfalls as explained in [Raymer et al. \(2012\)](#), for example. The other explanation

for such a volatility may due to political and economic developments and changes in legislation (Alders et al., 2007). In particular, larger irregularities appear when the data are disaggregated by single year of age, as illustrated for immigration and emigration in the middle and bottom rows of Figure 3.

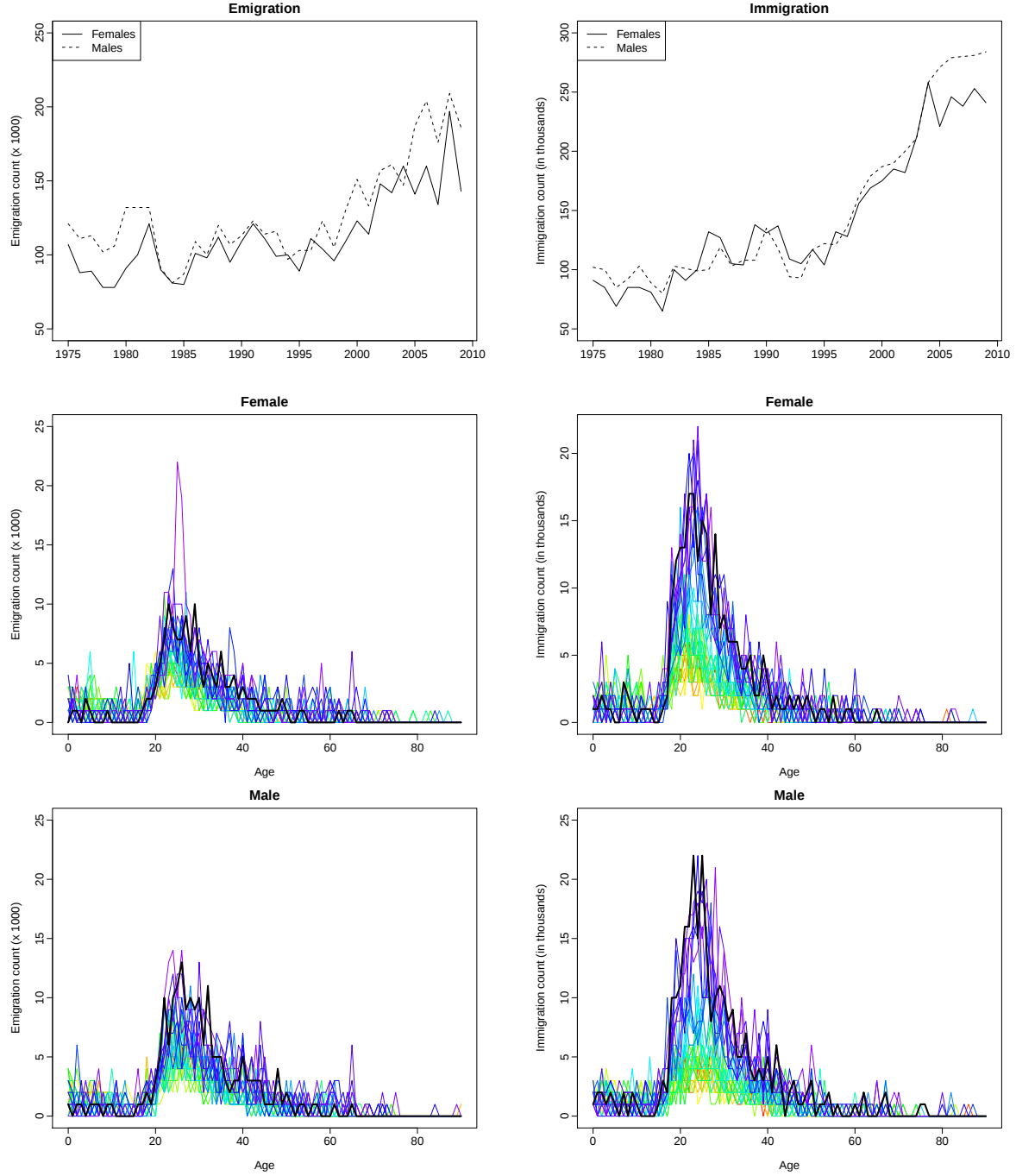


Figure 3: Total and age-specific emigration and immigration counts for the UK from 1975 to 2009. The black line represents the migration data in 2009

4. Selected Results

Since the coherent functional model can not be applied to model fertility, we use independent functional time-series model for fertility. For both mortality and emigration rates, we found that the coherent functional model outperforms the independent functional model. For immigration counts, we found that the independent functional model has a larger marginal likelihood than the coherent functional model. The log marginal likelihood (LML) and weights associated with each model is given in Table 1.

	Mortality		Emigration		Immigration	
	LML	weight	LML	weight	LML	weight
Coherent functional model	11022.38	1	19284.05	1	-9335.64	0
Independent functional model	10267.61	0	18951.01	0	-7602.9	1

Table 1: Log marginal likelihood and weight associated with each model for forecasting age-specific mortality rates, emigration rates and immigration counts

4.1. Forecasts of mortality rates

Using the historical mortality data from 1975 to 2009, Figure 4 presents the point forecast of log mortality rates in 2030, along with the 80% prediction interval. We found that the log mortality rates for children and those over 60s have dropped over the historical data presented in gray, and this pattern is likely to continue. Also, the mortality differences in adults between ages 20 and 40 is continuing to drop, especially for females.

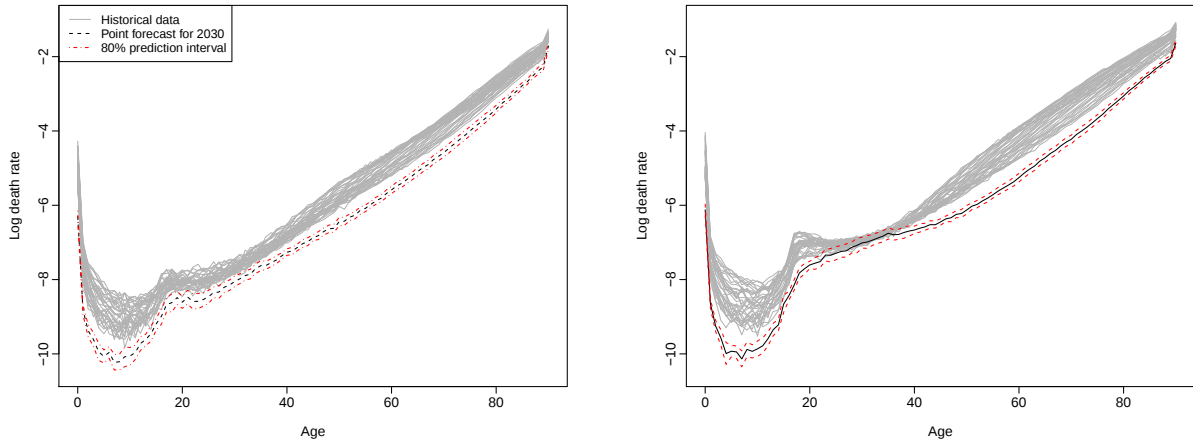


Figure 4: Forecasted age-specific mortality rates in 2030 for both females and males, based on the historical data from 1975 to 2009

4.2. Forecasts of fertility rates

Based on the historical fertility data from 1975 to 2009, we produce the probabilistic forecasts of age-specific fertility in Figure 5. The greatest forecast change is a continuing decrease in fertility rates for ages between 17 and 30, but a continuing increase in fertility for ages between 30 and 40. The resulting total fertility rates are presented in Figure 5b. It seems that the total fertility rates will decrease until 2015, and then increase thereafter. The slight declining, yet uncertain, fertility rates signal a possibility of another period of postponement, which may be linked to difficult economic conditions in terms of budgetary austerity in the UK in the 2020s (Kreyenfeld et al., 2012).

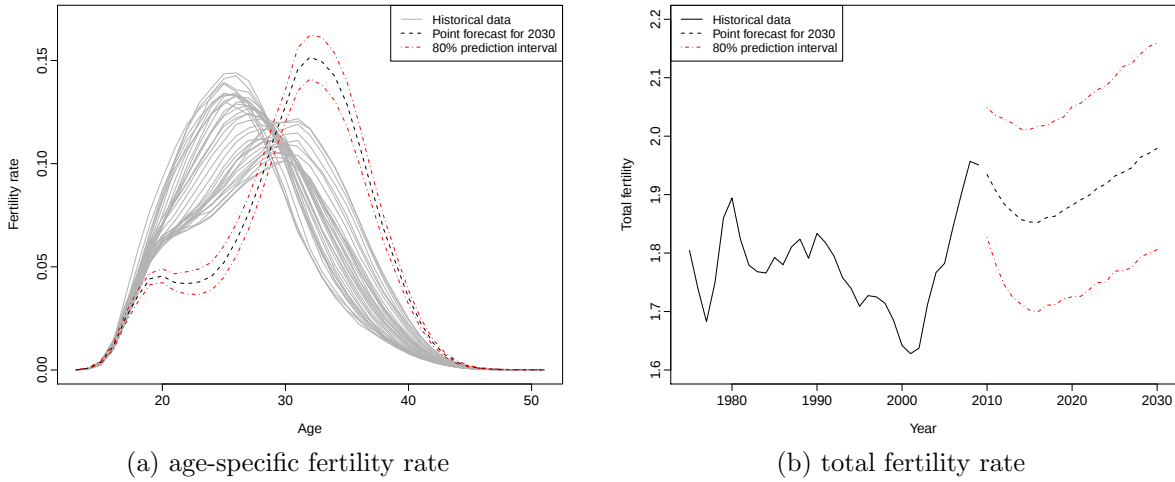


Figure 5: Forecasted age-specific fertility rates in 2030. The black dotted line represents the point forecast, where the dashed red lines represent the 80% prediction interval

4.3. Forecasts of emigration rates and immigration counts

Since the raw migration data are rather noisy, we apply the smoothing spline to obtain smooth curves, where the amount of smoothing is determined by generalised cross validation. Based on the historical emigration rates and immigration counts from 1975 to 2009, we produce the probabilistic forecasts of age-specific emigration rates and immigration counts. As shown in Figure 6, the greatest forecast change is a continuing increase in emigration rates and immigration counts for ages between 20 and 40.

4.4. Forecasts of population

With the population in 2009 as a baseline, the age composition of forecasted population in 2030 is presented in the first panel of Figure 7. Forecasts of the total female and male

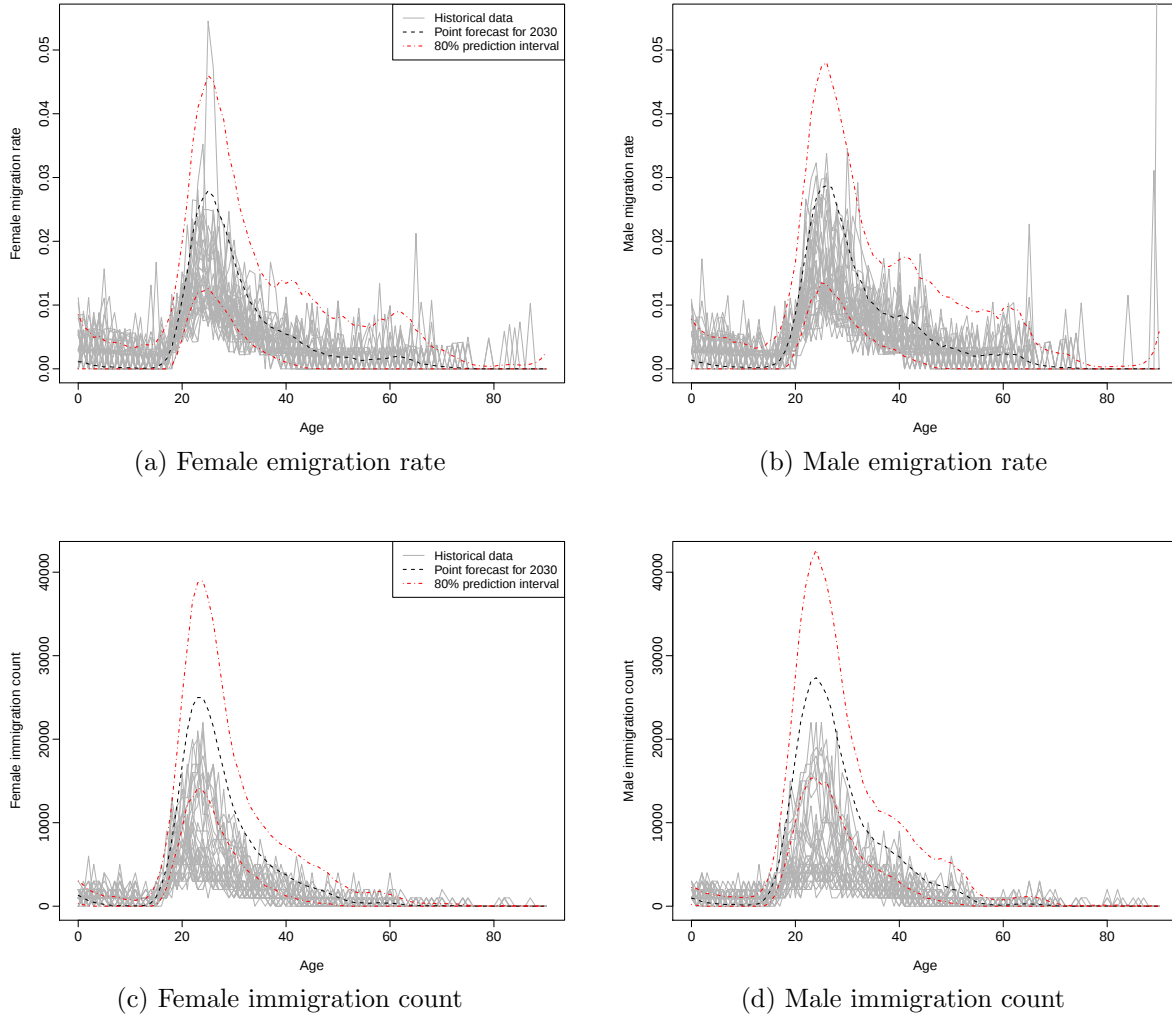


Figure 6: Forecasted age-specific emigration rates and immigration counts in 2030 for both females and males, based on the historical data from 1975 to 2009

populations are presented in the second panel of Figure 7. We found that the age profile of the population in 2030 is mainly driven by future migration and, to some extent, fertility. The median size of 2030 population is expected to reach 71.6 million, which is 10 million larger than the population of 61.6 million in 2009. We expect that the total population could exceed 70 million between 2028 and 2029. As shown in Table 2, we also compare our forecasts of total population with the five-year official forecasts prepared by the ONS, with 2010 as a baseline population. For each year considered, the ONS forecasts fall into our 80% prediction interval. The differences in forecasts may be due to the fact that the ONS assumes a constant net migration at the level of 200 thousand annually ([Office for](#)

[National Statistics, 2011](#)). It is impossible to verify whether such an assumption will hold, given high volatility in migration and recent policy of the UK government to reduce net migration to the levels below 100 thousand per year.

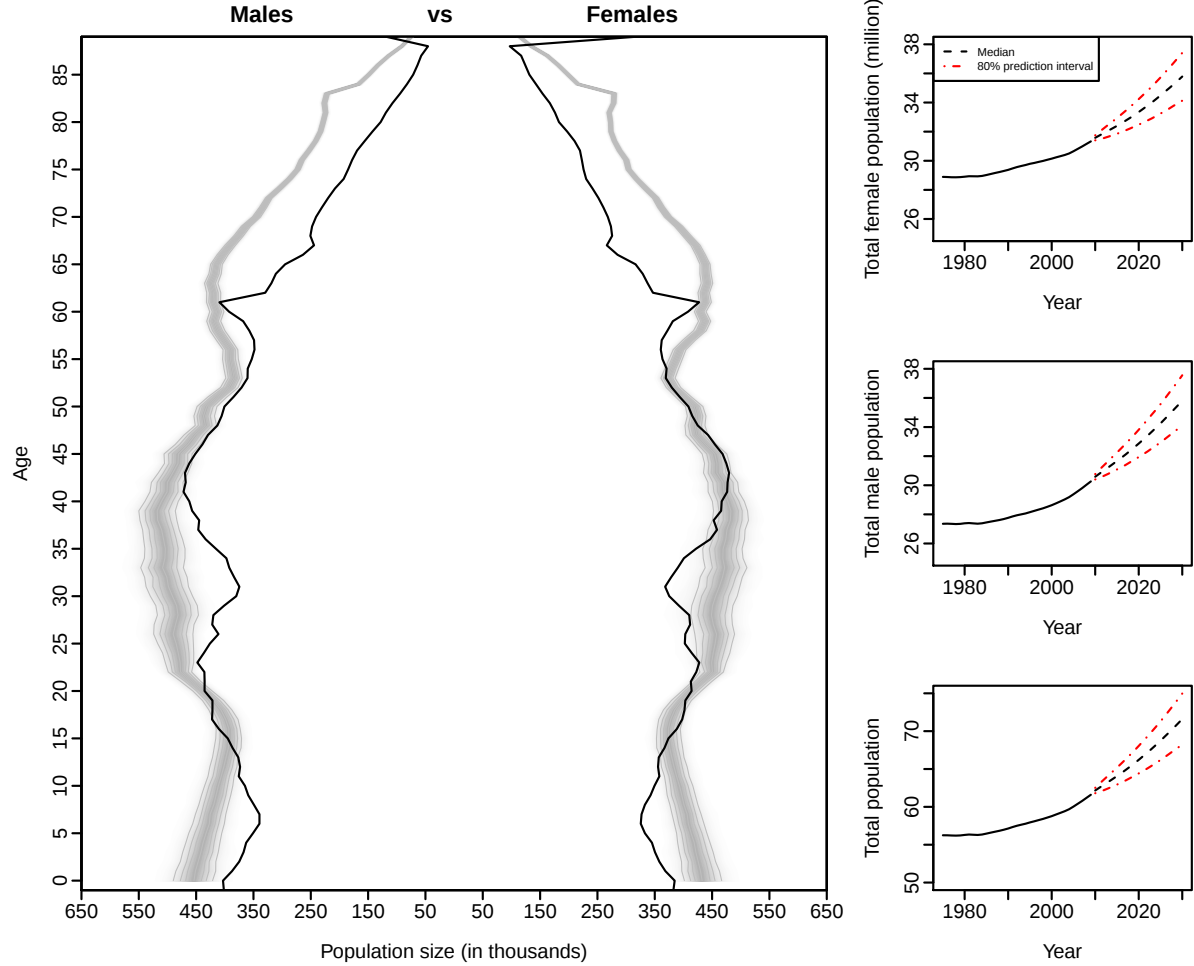


Figure 7: Forecasted age-specific population pyramid in 2030, along with the forecasted population sizes of females, males and total

Year	ONS	Median forecast		80% prediction interval	
		Proposed	Wiśniowski et al. (2013)	Proposed	Wiśniowski et al. (2013)
2010	62.262	62.165	62.238	(61.801, 62.512)	(62.171, 62.302)
2015	64.776	64.061	64.802	(62.941, 65.184)	(64.243, 65.386)
2020	67.173	66.214	67.905	(64.415, 68.049)	(66.481, 69.429)
2025	69.404	68.736	—	(66.193, 71.293)	—
2030	71.392	71.599	—	(68.238, 74.968)	—

Table 2: Comparison of population forecasts (in million) between the ONS, the proposed method and [Wiśniowski et al. \(2013\)](#) method based on the applications of the Lee-Carter model

5. Conclusion and future research

We present the independent and coherent functional time-series models for estimating and forecasting age schedules of the four demographic components of changes in the UK. By using a smoothing technique, the two functional models decompose the smoothed data into a number of functional principal components and their scores. The principal component scores and the variance parameter of the model residual are drawn from their posterior. For each sample, we produce time-series forecasts of principal component scores. The forecasted curves are obtained by multiplying the forecasted scores with fixed functional principal components and mean function.

Using the harmonic mean estimator, we compute the marginal likelihoods of the two functional models and assign the corresponding weights to the forecasts of mortality rates, emigration rates and immigration counts. Forecast population is then obtained through a cohort component projection model. The advantage of our approach can be attributed to: (1) the use of a smoothing technique to smooth out noisy or missing observations; (2) the use of higher order functional principal components to extract patterns in the data; (3) accounting for the uncertainties embedded in fertility, mortality and migration for each age and gender. Also, the advantage of the multilevel functional data model is that it incorporates correlation between two genders and thus allows each component of population to be modelled jointly, except fertility.

Since our models are conditional on the estimated functional principal components and mean function, it is likely that uncertainty is underestimated. In future, we aim to propose a fully Bayesian functional data analysis approach, in which the modelling and forecasting steps are considered together. Furthermore, our method models period age-specific rates, and it remains a future research to include cohort effect, at least in some of the four demographic components.

An additional contribution of this paper consists in evaluating an application of the proposed methodology to a situation of relatively good, yet still not perfect data availability. Given the regularities in age profiles of fertility, mortality and migration, disaggregation of the relevant data by age and sex provides important additional information for the forecasts. We argue that, data permitting, the population forecasting should follow bottom-up, from the age-specific rates, which describe the underlying processes more fully, rather than top-down, from summary aggregates as total fertility rates or life expectancies. For that

purpose, the functional approach coupled with a coherent analysis of model uncertainty offers a very natural way of maximising the use of the available information.

Acknowledgements

This work was carried out as a part of the Economic and Social Research Council (ESRC) Centre for Population Change (RES-625-28-0001) modelling strand. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors, and should not be attributed in any manner to the institutions with which they are affiliated.

References

- Ahlburg, D. A. (1998), Using economic information and combining to improve forecast accuracy in demography, Working paper, Industrial Relations Center, University of Minnesota, Minneapolis.
- Ahlburg, D. A. (2001), Population forecasting, *in* ‘Principles of Forecasting’, Kluwer Academic Publishers, New York, pp. 557–575.
- Alders, M., Keilman, N. and Cruijsen, H. (2007), ‘Assumptions for long-term stochastic population forecasts in 18 European countries’, *European Journal of Population* **23**(1), 33–69.
- Bates, J. M. and Granger, C. W. J. (1969), ‘The combination of forecasts’, *Operational Research Quarterly* **20**(4), 451–468.
- Bijak, J. (2010), *Forecasting International Migration in Europe: A Bayesian View*, Springer Series on Demographic Methods and Population Analysis, Springer, Dordrecht, NL.
- Burnham, K. P. and Anderson, D. R. (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, Springer, Springer.
- Claeskens, G. and Hjort, N. L. (2008), *Model Selection and Model Averaging*, Cambridge University Press, Cambridge.
- Clemen, R. T. (1989), ‘Combining forecasts: A review and annotated bibliography’, *International Journal of Forecasting* **5**(4), 559–583.

- Crainiceanu, C. M. and Goldsmith, J. A. (2010), ‘Bayesian functional data analysis using WinBUGS’, *Journal of Statistical Software* **32**(11).
- Ferraty, F., ed. (2011), *Recent Advances in Functional Data Analysis and Related Topics*, Springer, Heidelberg.
- Ferraty, F. and Romain, Y., eds (2011), *The Oxford Handbook of Functional Data Analysis*, Oxford University Press, Oxford.
- Ferraty, F. and Vieu, P. (2006), *Nonparametric Functional Data Analysis*, Springer, New York.
- Granger, C. W. J. and Ramanathan, R. (1984), ‘Improved methods of combining forecasts’, *Journal of Forecasting* **3**(2), 197–204.
- He, X. and Ng, P. (1999), ‘COBS: Qualitatively Constrained Smoothing via Linear Programming’, *Computational Statistics* **14**, 315–337.
- Hoeting, J. A., Madigan, D., Raftery, A. E. and Volinsky, C. T. (1999), ‘Bayesian model averaging’, *Statistical Science* **14**(4), 382–401.
- Human Fertility Database (2013), *Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria)*. Accessed at 8/January/2013.
URL: <http://www.humanfertility.org/cgi-bin/main.php>
- Human Mortality Database (2013), *University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany)*. Accessed at 8/January/2013.
URL: <http://www.mortality.org/>
- Hyndman, R. J., Booth, H. and Yasmeeen, F. (2013), ‘Coherent mortality forecasting: the product-ratio method with functional time series models’, *Demography* **50**(1), 261–283.
- Hyndman, R. J. and Khandakar, Y. (2008), ‘Automatic time series forecasting: the forecast package for R’, *Journal of Statistical Software* **27**(3).
- Hyndman, R. J. and Shang, H. L. (2009), ‘Forecasting functional time series (with discussion)’, *Journal of the Korean Statistical Society* **38**(3), 199–211.

- Hyndman, R. J. and Ullah, S. (2007), ‘Robust forecasting of mortality and fertility rates: A functional data approach’, *Computational Statistics and Data Analysis* **51**(10), 4942–4956.
- Kreyenfeld, M., Andersson, G. and Pailhé, A. (2012), ‘Economic uncertainty and family dynamics in Europe’, *Demographic Research* **27**, 835–852.
- Lee, R. D. and Carter, L. R. (1992), ‘Modeling and forecasting U.S. mortality’, *Journal of the American Statistical Association* **87**(419), 659–675.
- Li, N. and Lee, R. (2005), ‘Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method’, *Demography* **42**(3), 575–594.
- 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.
- Ní Bhrolcháin, M. and Beaujouan, E. (2012), ‘Fertility postponement is largely due to rising educational enrolment’, *Population Studies* **66**(3), 311–327.
- Office for National Statistics (2011), National population projections, 2010-based reference volume: Series pp2, Technical report, Office for National Statistics, Population projections unit.
- URL: <http://www.ons.gov.uk/ons/rel/npp/national-population-projections/2010-based-reference-volume--series-pp2/index.html>
- Preston, S. H., Heuveline, P. and Guillot, M. (2001), *Demography: Measuring and Modeling Population Processes*, Blackwell, UK.
- Raftery, A. E. (1996), Hypothesis testing and model selection via posterior simulation, in W. R. Gilks, S. Richardson and D. Spiegelhalter, eds, ‘Markov chain Monte Carlo in practice’, Chapman & Hall/CRC.
- Raftery, A. E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005), ‘Using Bayesian model averaging to calibrate forecast ensembles’, *Monthly Weather Review* **133**(5), 1155–1174.

- Raftery, A. E., Newton, M. A., Satagopan, J. M. and Krivitsky, P. N. (2007), Estimating the integrated likelihood via posterior simulation using the harmonic mean identity, *in* J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith and M. West, eds, ‘Bayesian Statistics’, Oxford University Press.
- Ramsay, J. O. (1988), ‘Monotone regression splines in action’, *Statistical Science* **3**(4), 425–441.
- Ramsay, J. O. and Silverman, B. W. (2005), *Functional Data Analysis*, 2nd edn, Springer, New York.
- Raymer, J., Abel, G. J. and Rogers, A. (2012), ‘Does specification matter? Experiments with simple multiregional probabilistic population projections’, *Environment and Planning A* **44**(11), 2664–2686.
- Rees, P. (1986), Developments in the modelling of spatial populations, *in* R. Woods and P. Rees, eds, ‘Population Structures and Models’, George Allen & Unwin, UK.
- Sanderson, W. C. (1998), ‘Knowledge can improve forecasts! A review of selected socio-economic population projection models’, *Population and Development Review* **24**(supplement), 88–117.
- Shang, H. L. (2012), ‘Point and interval forecasts of age-specific life expectancies: A model averaging approach’, *Demographic Research* **27**, 593–644.
- Smith, S. K. and Shahidullah, M. (1995), ‘An evaluation of population projection errors for census tracts’, *Journal of the American Statistical Association* **90**(429), 64–71.
- Wahba, G. (1990), *Spline Models for Observational Data*, Society for Industrial and Applied Mathematics, Philadelphia.
- Wiśniowski, A., Smith, P. W. F., Bijak, J. and Raymer, J. (2013), Bayesian cohort component population forecasts, *in* ‘Paper for the Population Association of America 2013 Annual Meeting’, New Orleans.