Centre for Population Change Working Paper

Number 41

A functional data analysis approach for forecasting population: A case study for the United Kingdom

Han Lin Shang

Peter W. F. Smith

Jakub Bijak

Arkadiusz Wisniowski

December 2013

ISSN2042-4116

ABSTRACT

Cohort component models are often used to model the evolution of an age-specific population, and are particularly useful to highlight which demographic component contributes the most to population change. Many methods have been proposed to forecast four demographic components, namely mortality, fertility, emigration and immigration. These existing methods are sometimes considered from a deterministic viewpoint, which in practice can be quite restrictive. The statistical method we propose is a multilevel functional data analytic approach, where the mortality and migration for females and males are modelled and forecasted jointly. The forecast uncertainty associated with each component is incorporated through bootstrapping. Using the historical data for the United Kingdom from 1975 to 2009, we found that the proposed method shows good in-sample forecast accuracy for the holdout data between years 2000 and 2009. Moreover, we produce out-of-sample population forecasts from 2010 to 2030, and compare our forecasts with those produced by the Office for National Statistics.

KEYWORDS

age-specific population forecasting; population projection matrix; Leslie matrix; functional data analysis; functional principal component analysis

EDITORIAL NOTE

Dr Han Lin Shang is a Research Fellow at the ESRC Centre for Population Change (CPC), working with his colleagues and co-authors of this paper on developing a dynamic population model for the UK.

Professor Peter Smith is Professor of Social Statistics and leads the CPC work package ‘modelling population growth and enhancing the evidence base for policy’. Peter’s research interests include graphical modelling, exact inference and models for longitudinal data.

Dr Jakub Bijak is a lecturer in Demography and CPC member. Jakub’s research interests are in migration and population forecasting, agent based population modelling and the Bayesian approach.

Dr Arkadiusz Wisniowski is a Research Fellow at the Centre with special interests in Bayesian statistics, population modelling and forecasting and time series econometrics.

Corresponding author: Han Lin Shang, H.Shang@southampton.ac.uk

The ESRC Centre for Population Change Working Paper Series is edited by Teresa McGowan.

ESRC Centre for Population Change

The ESRC Centre for Population Change (CPC) is a joint initiative between the Universities of Southampton, St Andrews, Edinburgh, Stirling, Strathclyde, in partnership with the Office for National Statistics (ONS) and the National Records of Scotland (NRS). The Centre is funded by the Economic and Social Research Council (ESRC) grant number RES-625-28-0001.

Website | Email | Twitter | Facebook | Mendeley

A FUNCTIONAL DATA ANALYSIS APPROACH FOR FORECASTING POPULATION: A CASE STUDY FOR THE UNITED KINGDOM

TABLE OF CONTENTS

1 INTRODUCTION

In recent decades, we have seen a considerable amount of development in the stochastic modelling and forecasting of population. This includes the pioneering work of Alho & Spencer (1985, 2005), Keilman (1990), Ahlburg & Land (1992), Lee & Tuljapurkar (1994), Lutz (1996), Bongaarts & Bulatao (2000), De Beer (2000), Alho et al. (2006), Raftery et al. (2012), Bryant & Graham (2013), among many others. Most of this body of work emphasise the advantages of stochastic modelling and forecasting over deterministic counterparts (for reviews, see Wilson & Rees 2005, Booth 2006). While the stochastic approach incorporates uncertainty into population estimates and forecasts, the deterministic approach provides only plausible scenarios representing high, medium and low. Despite the advantages of stochastic forecasts, they have been of limited use by official statistical agencies for several reasons (Lutz & Goldstein 2004). First, it is not straightforward to consider all different types of forecast uncertainties. Second, official statistical agencies are often constrained to a limited set of statistical methods with which they are familiar (see also Wiśniowski et al. 2013). While much has been done in the development of stochastic techniques, there remains a lot of work needed to produce probabilistic models that are usable at a detailed demographic level.

In this paper, we consider a functional data analytic approach for forecasting population (see also Hyndman & Booth 2008). As a generalisation of Lee & Carter's (1992) method, the functional data analysis views each component of population from a functional perspective, by combining ideas from nonparametric smoothing (Eubank 1999), functional principal component regression (Hyndman & Ullah 2007), and bootstrapping (Efron 2011). It has the following features: data are pre-processed before performing functional principal component analysis in order to reduce the effect of the noisy and missing observations. Each component of population is modelled as continuous functions of age, so that patterns of variation among years are captured by the functional principal components and their scores. By drawing from normal distributions, the simulated scores can be forecasted using a univariate time-series technique for each replication. Conditioning on the estimated mean and functional principal component functions, the probabilistic forecasts of future realisations can be obtained through bootstrapping.

Demographic events, such as fertility, mortality, and migration tend to exhibit strong regularities in their age patterns. Modelling the age profiles over time permits a relatively concise representation of the history of demographic patterns. The time-series methods can be utilised to extrapolate age profiles in future years. Throughout this paper, we focus on exploring two functional data models for forecasting age-specific fertility, mortality, emigration rates and immigration counts in a cohort component projection model. We also focus on exploring the consequences of choosing different Box-Cox transformation parameters for age-specific fertility, mortality, emigration and immigration in a cohort component projection model in terms of its in-sample measure of uncertainty. As an illustration, we use a time series from the United Kingdom (UK), consisting of age-specific data for single year of age from 1975 to 2009. We use the term "forecast" to refer to an outcome of a probabilistic exercise in predictions, as opposed to purely deterministic "projection" (Keilman 1990).

This functional data analytic approach was first implemented by Hyndman & Booth (2008) to forecast population in Australia. Our work differs from Hyndman & Booth (2008) in four aspects: (1) we use the multilevel functional data method to jointly model mortality and migration for females and males. In doing so, an improved point and interval forecast accuracy can be obtained, as evident from the in-sample population forecasts in Section 5; (2) we use a grid search method to find the optimal Box-Cox transformation parameter for mortality, fertility and migration; (3) we model the emigration and immigration separately, instead of the net migration in Hyndman & Booth (2008); (4) bootstrapping is the chosen method for constructing prediction interval, and it does not depend on normality assumption.

This article is organised as follows. In Section 2, we present a general background to population forecasting, define the sex- and age-specific population projection matrix, and highlight the issues pertaining to the stochastic modelling and forecasting of each demographic component by age and sex. In Section 3, we introduce the functional principal component regression to forecast the age profiles of each demographic component. Illustrated by the UK's data given in Section 4, we evaluate the in-sample accuracy using a holdout sample, and present out-of-sample forecasts from 2010 to 2030 in Section 5. Conclusions are given in Section 6, along with some ideas on how the methods presented here can be further extended.

2 COHORT COMPONENT POPULATION PROJECTION MODEL

We consider the cohort component model 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 survived offspring of women aged j. Third, the emigration rate η_i for $i = 0, \dots, 89, 90+$, measures the average number of migrants at age i. Finally, the immigration count, I_i, measures the total number of immigrants at age i.

As in Wiśniowski et al. (2013), let $P^{\mathrm{F}}_{i,t}$ and $P^{\mathrm{M}}_{i,t}$ 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

$$\left[\begin{array}{c}\mathbf{P}^{\mathrm{F}}_{t+1}\\ \hline \mathbf{P}^{\mathrm{M}}_{t+1}\end{array}\right] = \left[\begin{array}{c|c} a\mathbf{b}^{\mathrm{F}}_t & \mathbf{0}\\ \mathbf{s}^{\mathrm{F}}_t & \mathbb{O} \\ \hline (1-a)\mathbf{b}^{\mathrm{M}}_t & \mathbf{0}\\ \mathbb{O} & \mathbf{s}^{\mathrm{M}}_t\end{array}\right] \times \left[\begin{array}{c}\mathbf{P}^{\mathrm{F}}_t\\ \hline \mathbf{P}^{\mathrm{M}}_t\end{array}\right] + \left[\begin{array}{c}\mathbf{I}^{\mathrm{F}}_t\\ \hline \mathbf{I}^{\mathrm{M}}_t\end{array}\right],$$

where $\mathbf{P}^k_t$, $k = \mathrm{M}$ 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 $\mathbf{b}^k_t = (0, \dots, b^k_{13,t}, \dots, b^k_{51,t}, \dots, 0)$ is a vector of life-table birth rates, which can be derived from the age-specific fertility rates as follows

$$b^k_{i,t} = \frac{1}{1 + 0.5\mu^k_{0,t}} \frac{1}{2} \left(f_{i,t} + s^{\mathrm{F}}_{i,t} f_{i+1,t}\right),$$

where $f_{i,t} > 0$ represents the fertility rate at age i in year t; $\mu^{\mathrm{k}}_{0,t}$ represents the age-specific female or male mortality rate at age 0; $s^{\mathrm{F}}_{i,t}$ 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

$$s^k_{i,t} = \begin{cases} \frac{1-0.5\left(\mu^k_{i,t}+\eta^k_{i,t}\right)}{1+0.5\left(\mu^k_{i+1,t}+\eta^k_{i+1,t}\right)} & \text{if } i = 0, \dots, 89 \\ \frac{1-0.5\left(\mu^k_{i,t}+\eta^k_{i,t}\right)}{1+0.5\left(\mu^k_{i,t}+\eta^k_{i,t}\right)} & \text{if } i = 90+ \end{cases},$$

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

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

In (2), $\mathbf{0} = (0, \dots, 0)$ is a vector of length 91 and $\mathbb{O}$ is a matrix of zeros of size (90×91), and $\mathbf{I}^{\mathrm{F}}_t = (I^{\mathrm{F}}_{0,t}, \dots, I^{\mathrm{F}}_{90,t})'$ represents the vector of immigration counts at year t.

The cohort-component population projection model is very useful and easy to implement for delineating the evolution of a population (see for example, Rogers 1975, Kajin et al. 2012). The projection model is also helpful for describing the long-term evolution, where the composition of the population achieves an equilibrium (Dublin & Lotka 1925, Rogers 1995, Preston et al. 2001, de la Horra et al. 2013). However, in practice, the difficulty of the projection model is that it is not possible to know the true values of each component of a population. Therefore, we are required to forecast each component. Wiśniowski et al. (2013) take a Bayesian viewpoint to model various sources of uncertainty in each demographic component of population. Instead of modelling rates, their Bayesian method models the Poisson counts (see also Czado et al. 2005). Instead of using more than one principal component, they include a cohort effect (see also Renshaw & Haberman 2006). In order to capture the correlation between genders, they use vector autoregression model of order 1. In this paper, we tackle this problem by applying two functional data models and highlight their differences in terms of forecast accuracy.

3 FUNCTIONAL DATA ANALYTIC APPROACH

The functional data analytic approach has received an increasing amount of attention in the field of demographic forecasting (see also Lazar & Denuit 2009, Yang et al. 2010, Dowd et al. 2010, Cairns et al. 2011, D'Amato et al. 2011), since its first application to demographic forecasting by Hyndman & Ullah (2007). The objective of functional approach is to analyse a set of functions, usually smooth and bounded within an interval; such functional data may be age-specific mortality or fertility rates (see for example, Hyndman & Ullah 2007, Hyndman & Shang 2009, Hyndman et al. 2013). While Ramsay & Silverman (2005) and Ferraty & Vieu (2006) provided detailed surveys of many parametric and nonparametric techniques for analysing functional data, some recent developments are collected in the edited books by Ferraty & Romain (2011) and Ferraty (2011).

In the functional data analysis literature, popular approaches for nonparametric functional estimation include kernel regression methods (see for example, Ferraty & Vieu 2006), models that assume the functions of interest can be represented as linear combinations of, for example, fourier, wavelet, spline or eigenbasis functions (see for example, Ramsay & Silverman 2005), and models that assume the functions are realisations of stochastic processes (see for example, Bosq 2000). In this paper, we consider that each function can be approximated by a set of functional principal components and their corresponding scores.

3.1 AN INDEPENDENT FUNCTIONAL DATA MODEL

Differing from the Lee & Carter's (1992) model, the proposed functional data analytic approach can be described as follows:

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

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

   where $g_{i,t}$ represents the transformed data at age i in year t. To select the optimal transformation parameter ς, we examine a range of plausible values from $\{0, 0.1, \dots, 1\}$. Based on the in-sample coverage probability reported in Section 5, we found that the log transformation ($\varsigma = 0$) is appropriate for mortality, the Box-Cox transformation with parameter $\varsigma = 0.4$ is appropriate for fertility (see also Hyndman & Booth 2008), and $\varsigma = 0.5$ is appropriate for migration.

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 a smooth 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

$$g_t(x_i) = \tau_t(x_i) + \sigma_t(x_i)\varepsilon_{i,t}, \qquad t = 1, \dots, n,$$

   where $\sigma_t(x_i)$ models the variability for each age x_i in year t; and $\varepsilon_{i,t} \sim N(0, 1)$ is an independent and identically distributed random variable.

   For modelling age-specific mortality, we utilised penalised regression splines with a partial monotonic constraint for age above 65 (see Ramsay 1988, Hyndman & Ullah 2007, for more details). For modelling age-specific fertility, we used a weighted median smoothing B-spline, constrained to be concave (see He & 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).

3. Decomposition step. Higher order terms of the functional principal component decomposition improves the Lee-Carter model because these additional components capture non-random patterns, which are not explained by the first functional principal component (Booth et al. 2002, Renshaw & Haberman 2003, Koissi et al. 2006). By using functional principal component analysis (FPCA), a set of functions is decomposed into orthogonal functional principal components and their associated scores. For a survey on FPCA, consult Hall (2011) and Shang (2013). The functional principal component decomposition is given by

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

   where $\mu(x)$ is the mean function estimated by $\widehat{\mu}(x) = \frac{1}{n}\sum_{t=1}^{n} \tau_t(x)$; $\{\phi_1(x), \dots, \phi_K(x)\}$ is a set of the first K functional principal components; $\{\beta_{t,1}, \dots, \beta_{t,K}\}$ is a set of principal component scores and $\beta_{t,k} \sim N(0, \lambda_k)$ where λ_k is the kth eigenvalue of the covariance operator (see Appendix A for details); $e_t(x) \sim N(0, \sigma^2)$ is the error function with mean zero and finite variance; and $K < n$ is the number of retained components. Hyndman & Booth (2008) found that $K = 6$ is sufficient to capture a substantial amount of variance in the data.

   In order to capture the main sources of model uncertainty, we consider the parametric bootstrap method of Crainiceanu & Goldsmith (2010) by sampling $\beta_{t,k}$ and $e_t(x)$ from Gaussian distributions with zero mean and estimated variances. In Appendix B, we present the modified WinBUGS code for estimating the variance parameters.

4. Forecasting step. A wide range of time-series models may be used to forecast principal

component scores. Conditioning on the smoothed functions $\mathcal{I} = \{\tau_1(x), \ldots, \tau_n(x)\}$ and the estimated set of functional principal components $\mathcal{B} = \{\phi_1(x), \ldots, \phi_K(x)\}$, the h-step-ahead probabilistic forecast of $m_{n+h}(x)$ can be obtained as

$$\widehat{g}^{(b)}_{n+h|n}(x) = \text{E}[g_{n+h}(x)|\mathcal{I}, \mathcal{B}] = \widehat{\mu}(x) + \sum_{k=1}^{K} \widehat{\beta}^{(b)}_{n+h|n,k}\phi_k(x) + \widehat{e}^{(b)}_{n+h|n}(x) + \widehat{\sigma}_{n+h}(x)\widehat{\epsilon}^{(b)}_{n+h},$$

for $b = 1, \ldots, B$, where $B = 1000$ represents the number of iterations. $\widehat{\beta}^{(b)}_{n+h|n,k}$ 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 & Khandakar (2008) based on corrected Akaike Information Criterion, $e^{(b)}_{n+h|n}(x)$ is simulated from a normal distribution with zero mean, $\widehat{\sigma}_{n+h}(x)$ represents the estimated variance from the historical observations, and $\widehat{\epsilon}^{(b)}_{n+h}$ is simulated from a standard normal distribution.

3.2 A COHERENT FUNCTIONAL DATA MODEL

Individual forecasts of females and males, even when based on similar extrapolative procedures are likely to imply increasing divergence in each component of population in the long run, counter to the expected and observed trend toward convergence (Wilson 2001). This motivates the development of coherent forecasting methods, such as the augmented common factor Lee-Carter model (Li & Lee 2005) or the product-ratio model (Hyndman et al. 2013). Here, we introduce a new estimation method for the coherent functional data model, and apply it to population forecasting.

We present the problem in the context of forecasting male and female age-specific mortality rates, despite the fact that the methodology can easily be generalised to migration. The multilevel functional data method can model multiple sets of functions that may be correlated among groups (see also Crainiceanu et al. 2009, Crainiceanu & Goldsmith 2010, Staicu et al. 2010, Zipunnikov et al. 2011). Staicu et al. (2010) draw the close connection between the multilevel functional data models and hierarchical modelling, and suggest Bayesian method using Markov chain Monte Carlo sampling algorithm to estimate the parameters (see also Crainiceanu & Goldsmith 2010).

In this paper, we adapt the multilevel functional data model to analyse two samples of functions, observed for two subpopulations that may be correlated. The basic idea is to decompose functions from different subpopulations into an aggregated average, a sex-specific deviation from the 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. To express our idea mathematically, the smoothed female mortality rate at year t can be written as

$$\tau_t^{\text{F}}(x) = \mu(x) + \omega^{\text{F}}(x) + R_t(x) + U_t^{\text{F}}(x) + \varepsilon_t^{\text{F}}(x), \tag{1}$$

where $\mu(x)$ represents the overall mean function, $\omega^{\text{F}}(x)$ is the female deviation from the overall mean function, $R_t(x)$ models the common trend for female and male mortality rates, $U_t^{\text{F}}(x)$ models the sex-specific trend in the female mortality, and $\varepsilon_t^{\text{F}}(x)$ represents the error term. To ensure identifiability, we assume $R_t(x)$, $U_t^{\text{F}}(x)$ and $\varepsilon_t^{\text{F}}(x)$ are uncorrelated, and $\varepsilon_t^{\text{F}}(x)$ is white noise with finite variance σ^2.

Because the centred stochastic processes R and U are unknown in practice, the population eigenvalues and eigenfunctions can only be approximated through realisations of $\{R_1(x), \dots, R_n(x)\}$ and $\{U_1(x), \dots, U_n(x)\}$. The sample mean and sample covariance are thus given by

$$\widehat{\mu}(x) = \frac{1}{2n}\sum_{j=1}^{2}\sum_{t=1}^{n}\tau_t^{(j)}(x), \tag{2}$$

$$\widehat{\omega}^{\text{F}}(x) = \widehat{\mu}^{\text{F}}(x) - \widehat{\mu}(x), \tag{3}$$

$$\widehat{R}_t(x) \approx \sum_{k=1}^{K}\widehat{\beta}_{t,k}\widehat{\phi}_k(x), \tag{4}$$

$$\widehat{U}_t^{\text{F}}(x) \approx \sum_{l=1}^{L}\widehat{\gamma}_{t,l}^{\text{F}}\widehat{\psi}_l^{\text{F}}(x), \tag{5}$$

where $\{\widehat{\beta}_{t,1}, \dots, \widehat{\beta}_{t,K}\}$ and $\{\widehat{\gamma}_{t,1}^{\text{F}}, \dots, \widehat{\gamma}_{t,L}^{\text{F}}\}$ represents the sample principal component scores of $\widehat{R}_t(x)$

and $\widehat{U}_t^{\text{F}}(x)$, respectively; $\{\widehat{\phi}_1(x), \dots, \widehat{\phi}_K(x)\}$ and $\{\widehat{\psi}_1^{\text{F}}(x), \dots, \widehat{\psi}_L^{\text{F}}(x)\}$ are the corresponding orthogonal sample eigenfunctions; K and L represent the retained number of components. Following the work of Crainiceanu & Goldsmith (2010), we use a cumulative percentage of total variation to determine K and L. Concretely, the optimal K and L are determined to be the first few components that explain at least 99% and 90% of total variations in the common trend and sex-specific trend.

Inserting (2) to (5) into (1), we obtain

$$g_t^{\text{F}}(x) = \widehat{\mu}(x) + \widehat{\omega}^{\text{F}}(x) + \sum_{k=1}^{K} \widehat{\beta}_{t,k}\widehat{\phi}_k(x) + \sum_{l=1}^{L} \widehat{\gamma}_{t,l}^{\text{F}}\widehat{\psi}_l^{\text{F}}(x) + \varepsilon_t^{\text{F}}(x) + \sigma_t(x)\varepsilon_t, \tag{6}$$

where $\widehat{\beta}_{t,k} \sim N(0, \widehat{\lambda}_k^{(1)})$, $\widehat{\gamma}_{t,l}^{\text{F}} \sim N(0, \widehat{\lambda}_l^{\text{F}})$ and $\varepsilon_t^{\text{F}}(x) \sim N(0, \widehat{\sigma}^2)$. Note that (6) has been given in Hyndman & Ullah (2007, eq.15) as future research, but to the best of our knowledge, this has not been studied in population forecasting.

We formulate (6) from a linear mixed model (Pinheiro & Bates 2000), where the mechanism of Bayesian inference can be used to estimate $\lambda_k^{(1)}, \lambda_l^{\text{F}}, \sigma^2$ and then draw random samples $\beta_{t,k}, \gamma_{t,l}^{\text{F}}$ from the posterior distribution (Crainiceanu & Goldsmith 2010).

Conditioning on the estimated basis functions $\mathbf{\Phi} = \left[\widehat{\phi}_1(x), \dots, \widehat{\phi}_K(x)\right]$, $\mathbf{\Psi} = \left[\widehat{\psi}_1^{\text{F}}(x), \dots, \widehat{\psi}_L^{\text{F}}(x)\right]$ and smoothed functions $\boldsymbol{\mathcal{I}} = \left[\tau_1^{\text{F}}(x), \dots, \tau_n^{\text{F}}(x)\right]$, the point forecasts of $g_{n+h}^{\text{F}}(x)$ are given by

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

where $\widehat{\beta}_{n+h|n,k}$ and $\widehat{\gamma}_{n+h|n,l}^{\text{F}}$ are the forecasted principal component scores, obtained from a univariate time-series method, such as the autoregressive integrated fraction moving average (ARFIMA) method of Hyndman et al. (2013).

The prediction interval of $g_{n+h|n}^{\text{F}}(x)$ can be obtained using parametric bootstrap samples, given by

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

for $b = 1,\dots,B$, where $\widehat{\beta}^{b}_{n+h|n,k}$ is the forecast of principal component scores $\left\{\widehat{\beta}^{b}_{1,k},\dots,\widehat{\beta}^{b}_{n,k}\right\}$ drawn from $N\left(0,\widehat{\lambda}^{(1)}_k\right)$; similarly, $\widehat{\gamma}^{\mathrm{F}}_{n+h|n,l}$ is the forecast of bootstrapped principal component scores $\left\{\widehat{\gamma}^{b,\mathrm{F}}_{1,l},\dots,\widehat{\gamma}^{b,\mathrm{F}}_{n,l}\right\}$ drawn from $N\left(0,\widehat{\lambda}^{\mathrm{F}}_l\right)$; $\widehat{\varepsilon}^{b}_{n+h}(x)$ is drawn from $N\left(0,\widehat{\sigma}^2\right)$; $\widehat{\sigma}_{n+h}(x)$ represents the estimated variance for each age, and $\widehat{\epsilon}^{b}_{n+h}$ is simulated from a standard normal distribution; and $B = 1000$ represents the number of bootstrap replications. The prediction intervals can be constructed from percentiles of these mortality forecasts.

4 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 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 for age and time. By analysing the changes in mortality as a function of both age x and time t, we have seen that mortality has shown a gradual decline over time. To have an idea of this evolution, we present the logarithm of 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.

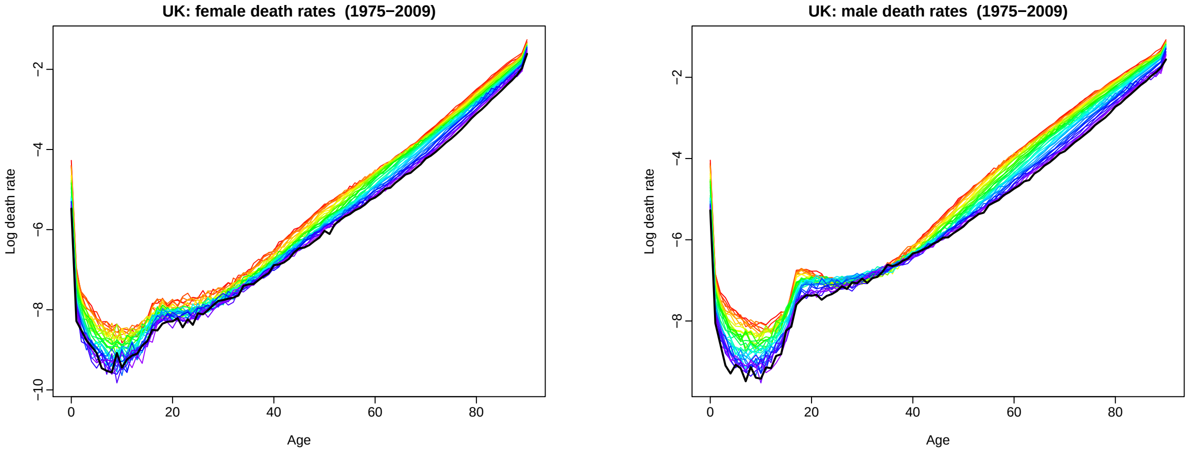


Figure 1. UK female and male age-specific log mortality rates (1975-2009). The distant past years are shown in red, with the most recent years in violet. Curves are ordered chronologically according to the colours of the rainbow. The black line represents the mortality in 2009

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 or higher education (Ní Bhrolcháin & Beaujouan 2012).

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

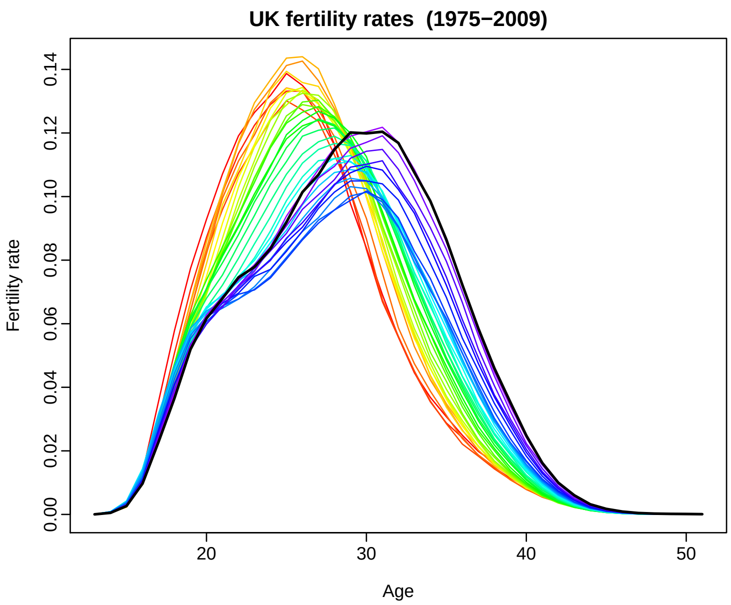


Figure 2. UK age-specific fertility rates between ages 13 and 51. The black line represents the fertility pattern in 2009

5 RESULTS

We present the results of forecasting each component of the UK population, using the independent functional data model described in Section 3.1. For each component, we examine the goodness-of-fit of the model to the data, and forecasts of future age-specific patterns.

5.1 FORECASTS OF MORTALITY

We present the first two fitted functional principal components which capture more than 96% of the total variation in the female mortality, and their associated scores in Figure 4. Throughout this section, the functional principal component decomposition is carried out using the independent functional data model, for example. The functional principal components are modelling different movements in log mortality rates. By inspecting the peaks and troughs, we found that $\phi_1(x)$ primarily models mortality changes in children and those over 60, while $\phi_2(x)$ models the difference in adults between ages 20 and 40. The log mortality rates for children and those over 60s have dropped over the whole data period, and the difference in adults between ages 20 and 40 is also continuing to drop. These phenomena are captured by the time-series model in forecasting

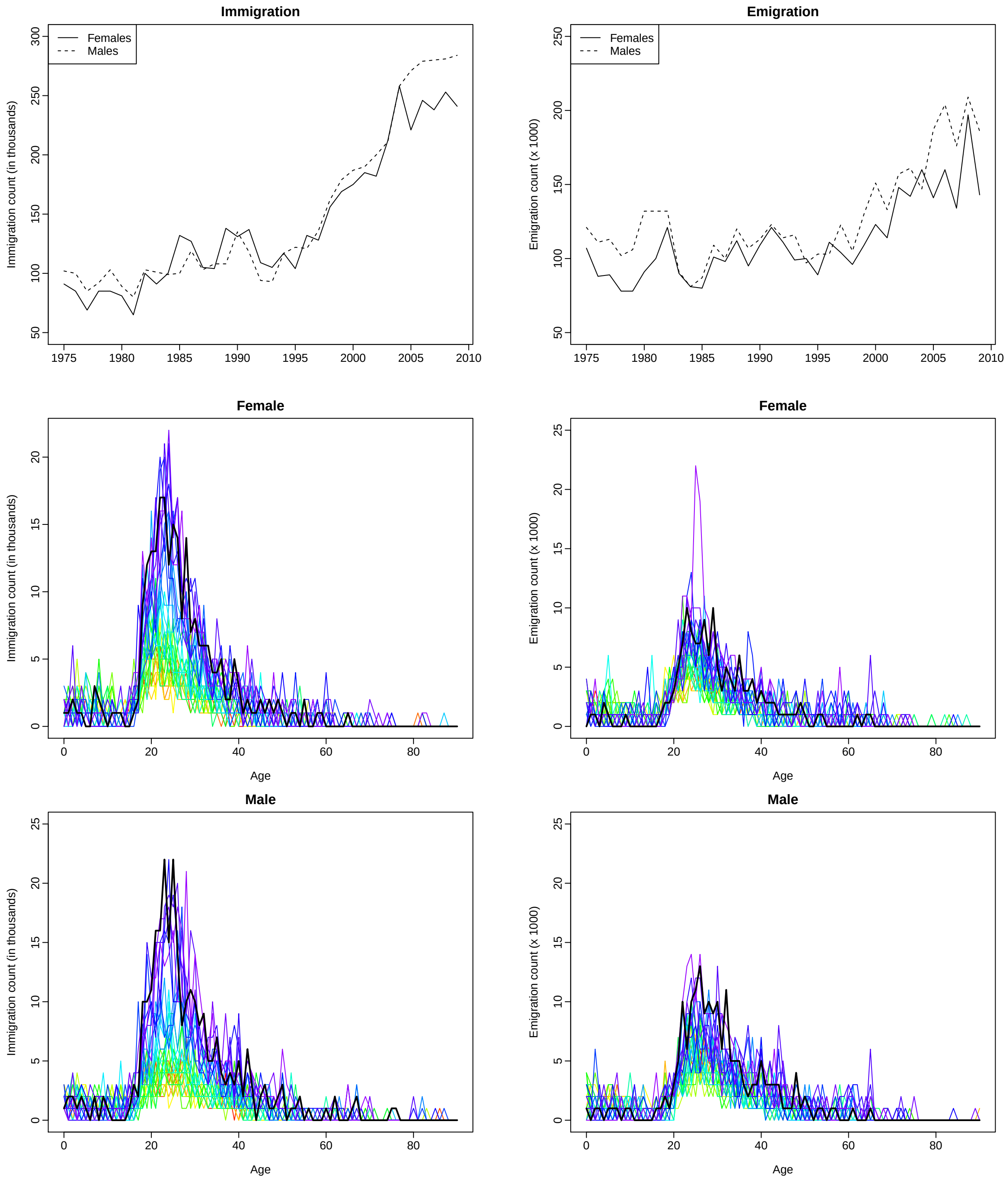


Figure 3. Total and age-specific emigration and immigration counts for the UK from 1975 to 2009. The black line represents the migration data in 2009

continuing decreasing trend for these coefficients over the next 26 years.

In Figures 5a and 5b, we present the model fit to the data for 2009. Based on the historical

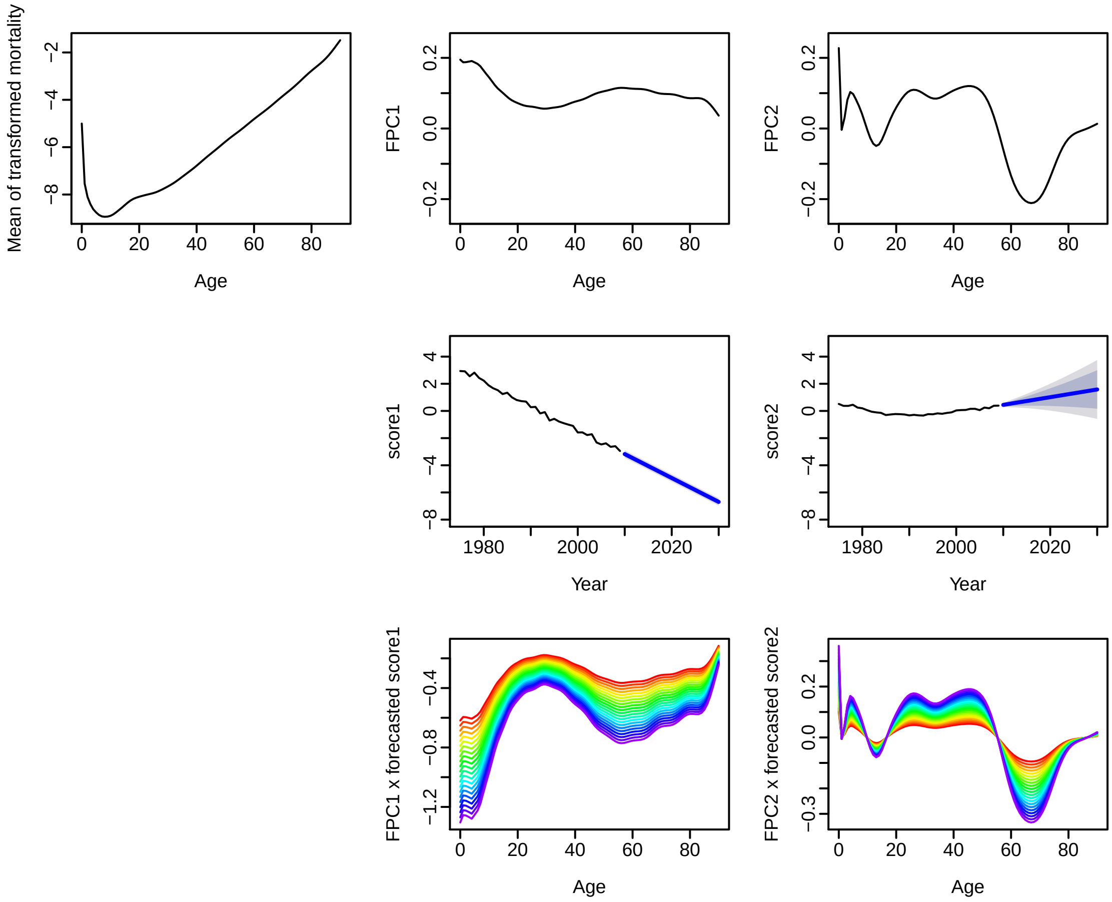


Figure 4. Functional principal components and associated scores for the UK male mortality data. A decomposition of $K = 6$ has been used for analysis, although we display only the first two components. The solid line represents the point forecasts of scores, where the dark and light grey regions represent the 80% and 95% point-wise prediction intervals, respectively. In the bottom panel, the forecasted principal component scores are multiplied by the fixed functional principal components, then adding the mean of transformed mortality to produce forecasts

mortality from 1975 to 2009, we produce the probabilistic forecasts of age-specific log mortality rates. As shown in Figures 5c and 5d, the mortality rates are continuing to decline, especially for elderly people. The decline seems to be more rapid for males than females. As a result of decreases in mortality, we observe an increase in life expectancy for both females and males shown in Figures 5e and 5f.

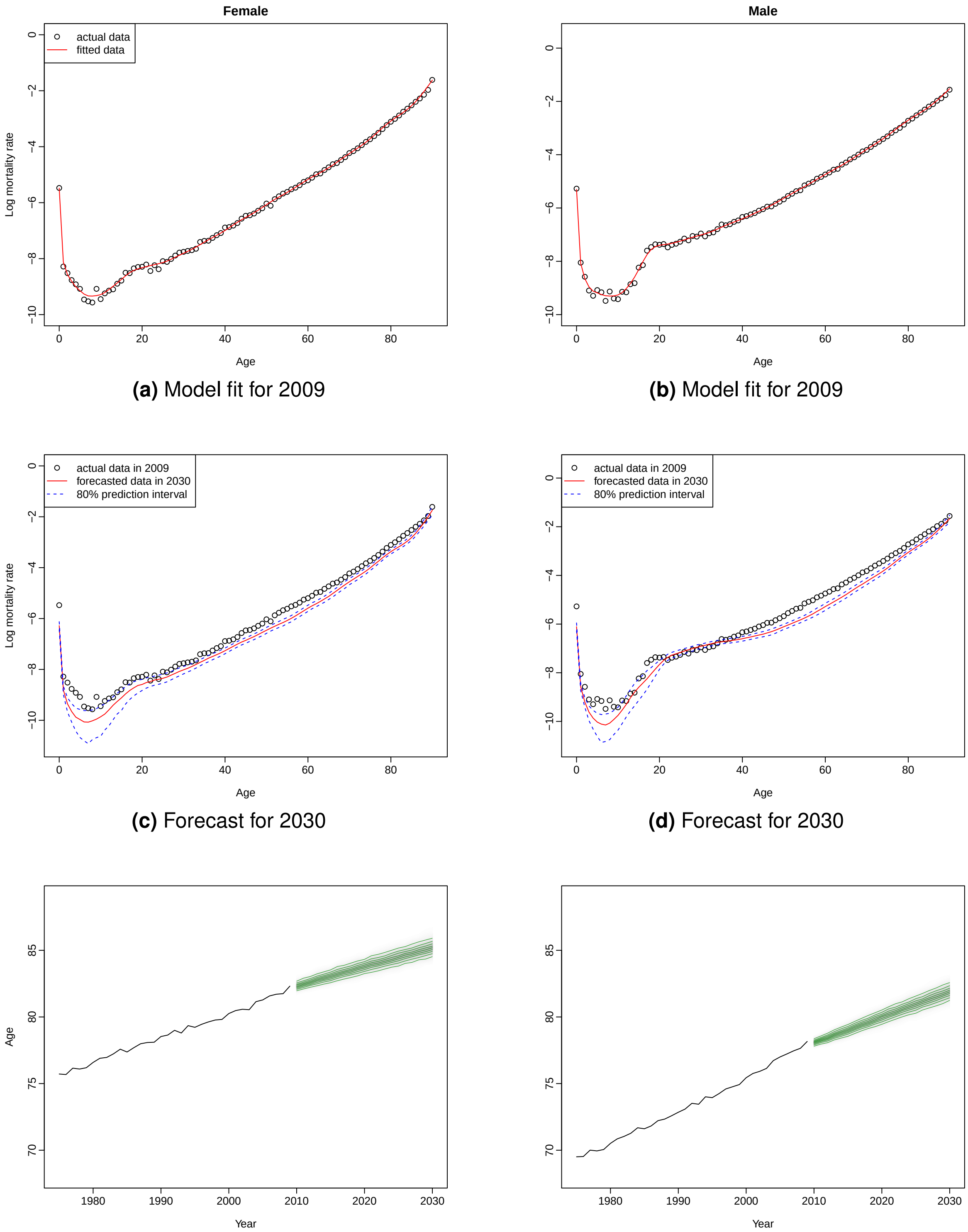
To support our rationale, the in-sample prediction of life expectancy from 2000 to 2009 is analysed. It turns out that 68% of the observed mortality rates for years 2000-2009 fall into the 80% prediction interval. For the 95% prediction interval, 86% observations fall into it. With the same data set, the in-sample coverage probabilities are similar to those obtained via Bayesian log-linear model with Poisson counts (see Wiśniowski et al. 2013, for more details).

Furthermore, for a set of possible Box-Cox transformation parameters, we calculate their in-sample empirical coverage probabilities, and determine the optimal transformation parameter ς with a minimal difference between the nominal and empirical coverage probabilities. In the case of mortality, $\varsigma = 0$. The table of the in-sample coverage probabilities for different Box-Cox transformation parameters can be provided upon request from the authors.

5.2 FORECASTS OF FERTILITY

Akin to mortality, the first two fitted functional principal components that capture around 92% of the total variation and their associated scores are presented in Figure 6. The functional principal components are modelling the fertility rates of mothers in different age groups: $\phi_1(x)$ represents mothers between their late 20s and 30s; $\phi_2(x)$ represents young mothers in their late teens and those between ages 20 and 40. The forecasted coefficients associated with each functional principal component demonstrate the recent social effects. The fertility trend between the late 20s and 30s are predicted to continue in the near future.

In Figure 7a, we present the model fit to the data for 2009. Based on the historical fertility from 1975 to 2009, we produce the probabilistic forecasts of age-specific fertility. As shown in Figure 7b, the greatest forecast change is a continuing decrease in fertility rates for ages between



(a) Model fit for 2009

(b) Model fit for 2009

(c) Forecast for 2030

(d) Forecast for 2030

(e) Life expectancy at birth

(f) Life expectancy at birth

Figure 5. Model fit for mortality to 2009 data, forecasted age-specific mortality for 2030, and forecasted life expectancy at birth until 2030

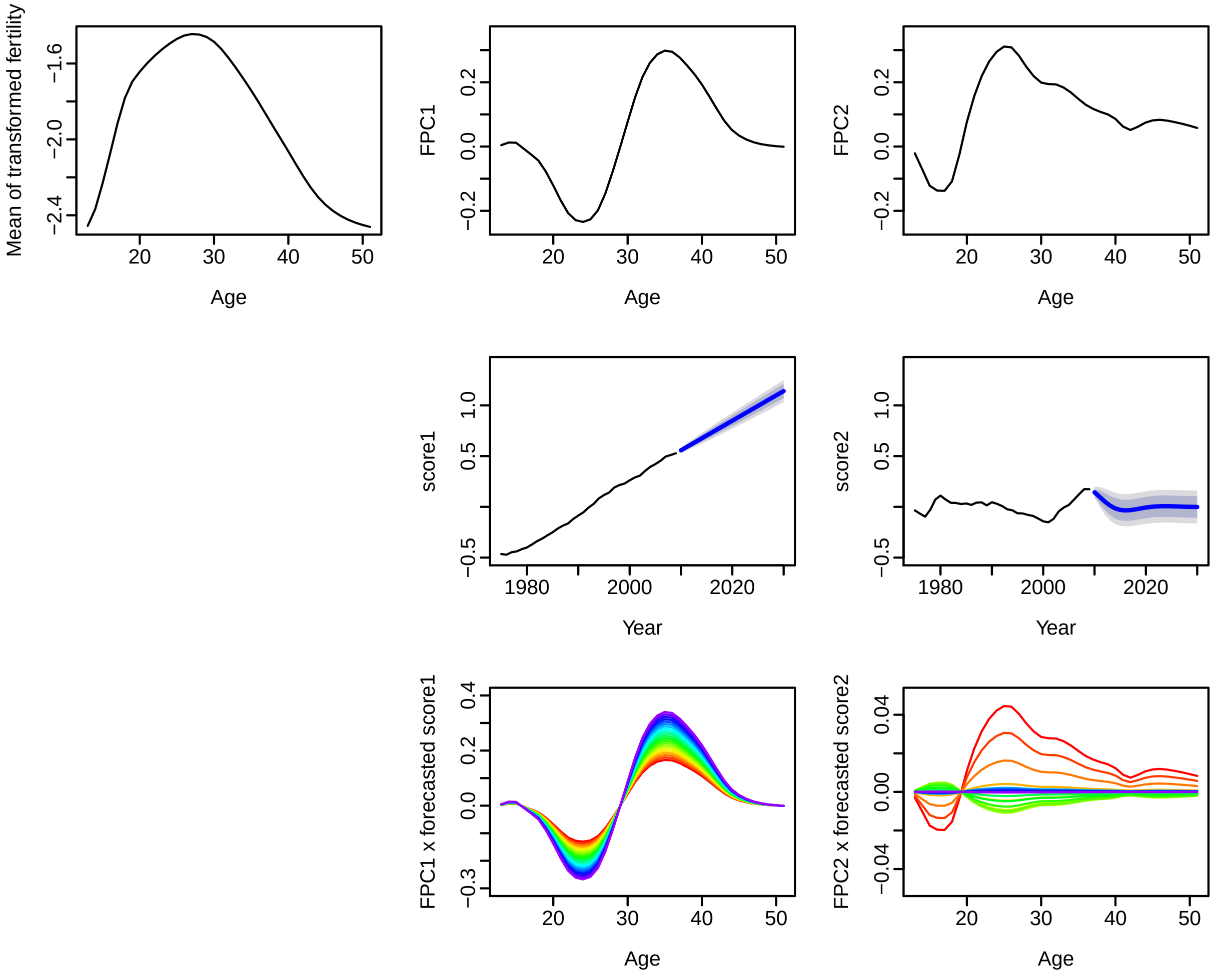
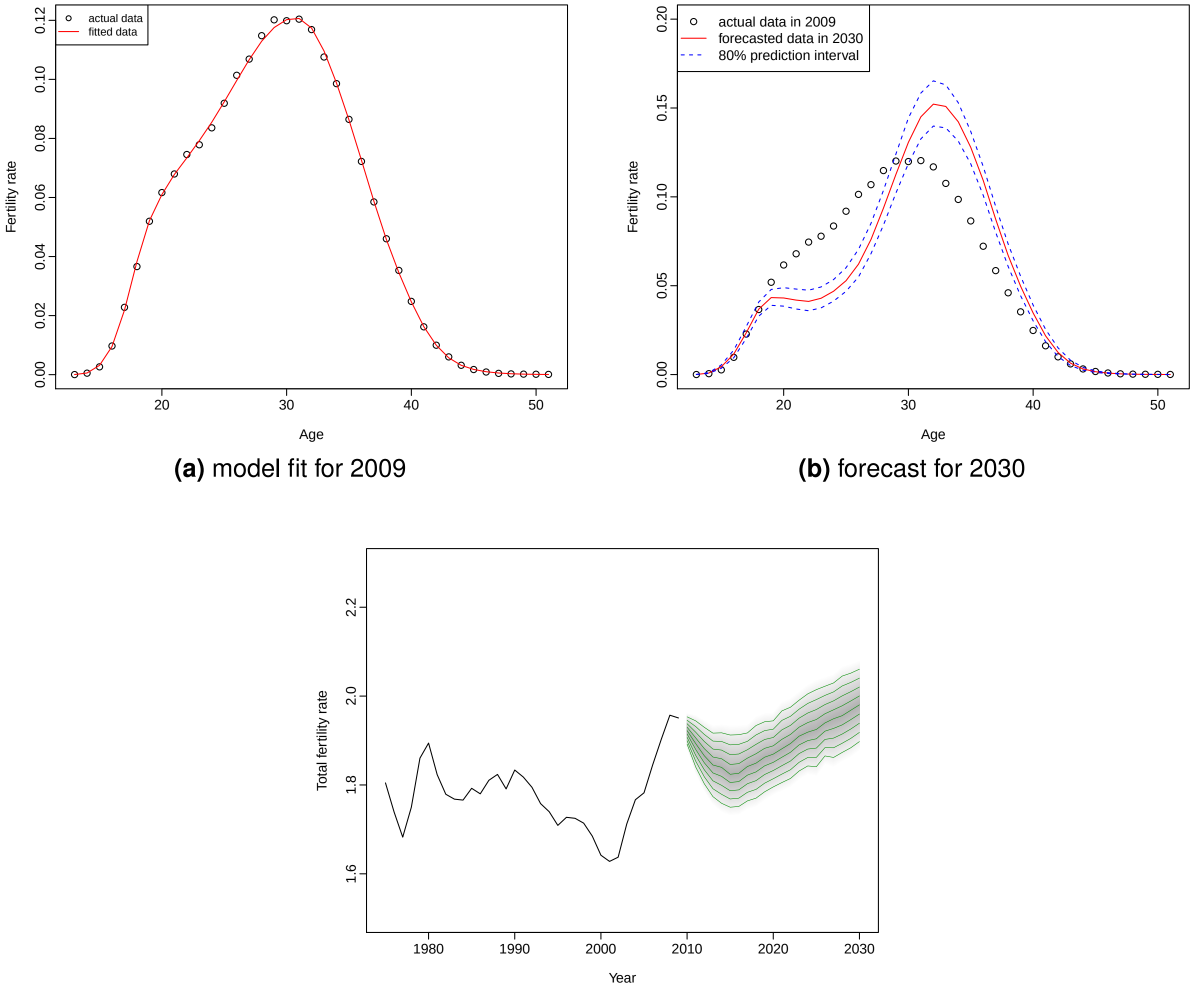


Figure 6. Functional principal components and their scores for the fertility data. A decomposition of $K = 6$ has been used for analysis, although we display only the first two components. The solid line represents the point forecasts of scores, where the dark and light grey regions represent the 80% and 95% point-wise prediction intervals

17 and 30, but a continuing increase in fertility for ages between 30 and 40. The resulting total fertility rates are presented in Figure 7c. It seems that the total fertility rate will decrease until 2015, and then increase thereafter. The slight declining, yet uncertain, fertility rates signal a possibility of another period of postponement, which can be linked to difficult economic conditions in terms of budgetary austerity in the UK in the second decade of the 21st century (Kreyenfeld et al. 2012).



(a) model fit for 2009

(b) forecast for 2030

(c) total fertility rate

Figure 7. Model fit for fertility to 2009 data, forecasted age-specific fertility for 2030, and forecasted total fertility rate until 2030

The in-sample prediction of total fertility rate from 2000 to 2009 is analysed. It turns out 61% of the observed total fertility rate from 2000 to 2009 fall within the 80% prediction interval. For the 95% prediction interval, 73% observations fall into it. With the same data set, the in-sample coverage probabilities are similar to those obtained via Bayesian log-linear model with Poisson counts (see Wiśniowski et al. 2013, for detail). Furthermore, the optimal Box-Cox transformation parameter for the fertility data is 0.4.

5.3 FORECASTS OF MIGRATION

Migration consists of two parts, namely emigration (outflow) and immigration (inflow). We present the first two fitted functional principal components that capture around 79% of the total variation in both the female emigration rate and female immigration counts, along with their associated scores in Figures 8 and 9. The functional principal components are modelling different age patterns. While $\phi_1(x)$ primarily models the migration of those in their 20s and 40s, $\phi_2(x)$ models the contrasts in migration between 20s and 30s. The future migration for the 20s and 40s is predicted to continue to follow a similar pattern, but the difference between 20s and 30s has levelled off. This phenomenon is reflected by the time-series model in projecting no change to these coefficients over the next 26 years.

In Figures 10a, 10b, 11a and 11b, we present the model fit to the migration data in 2009. Based on historical migration from 1975 to 2009, we produce the probabilistic forecasts of age-specific emigration rates and immigration counts. As shown in Figures 10c, 10d, 11c and 11d, the greatest forecast change is a continuing increase in emigration rates and immigration counts for ages between 20 and 40. The resulting female mean emigration rates across ages and female total immigration counts across ages are presented in Figures 10e, 10f, 11e and 11f. They seem to be increasing between 2010 and 2030.

5.4 POPULATION FORECASTS

Using the population in 1999 as a baseline, the age composition of the in-sample forecasted population from 2000 to 2009 is presented in Figure 12. Compared with the holdout data (a percentage of historical data that is reserved for use in measuring the forecast accuracy of a method), we found that age profiles of the forecasted population obtained from the independent

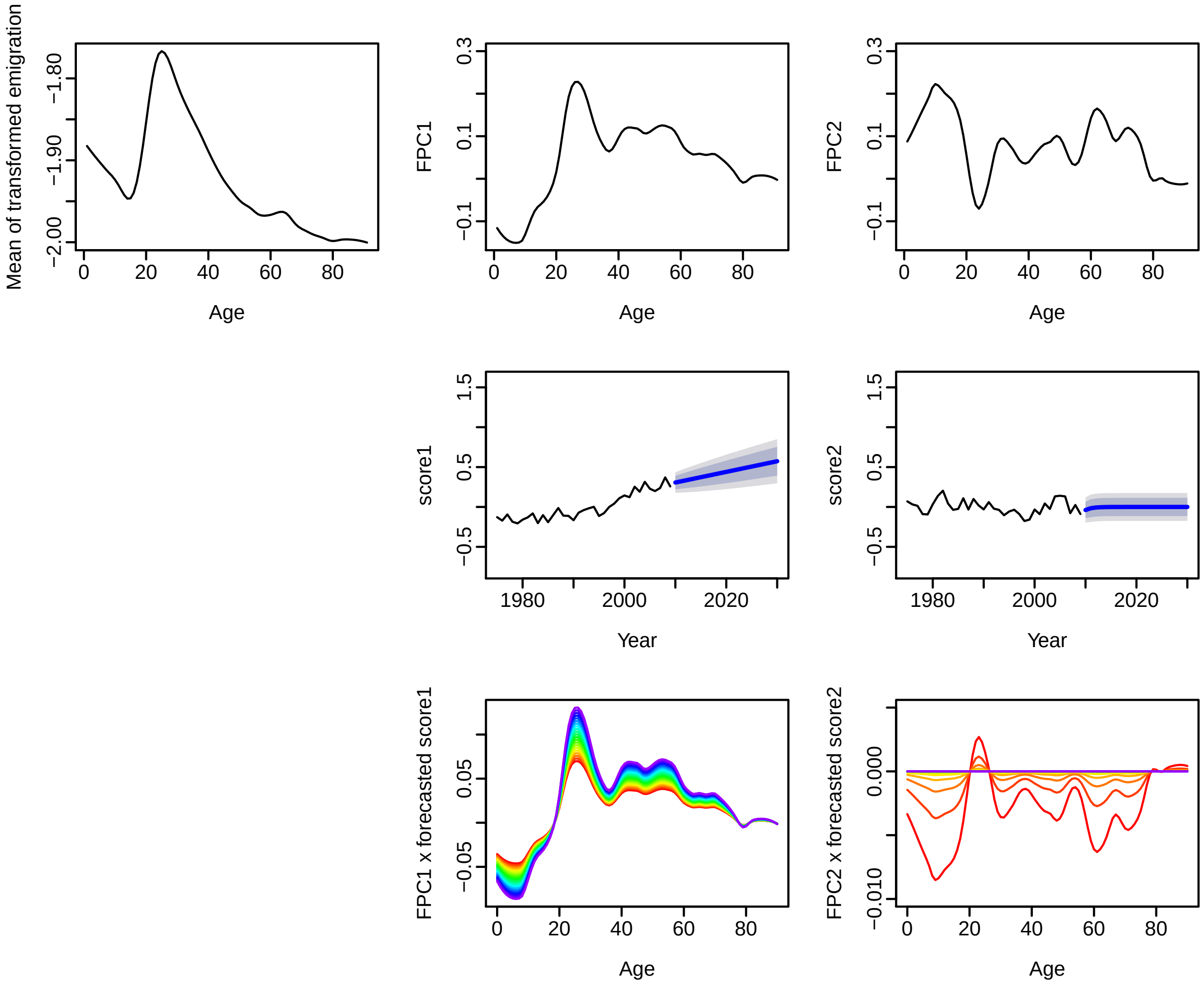


Figure 8. Functional principal components and their scores for the emigration rates. A decomposition of $K = 6$ has been used for analysis, although we display only the first two components. The solid line represents the point forecasts of scores, where the dark and light grey regions represent the 80% and 95% point-wise prediction intervals

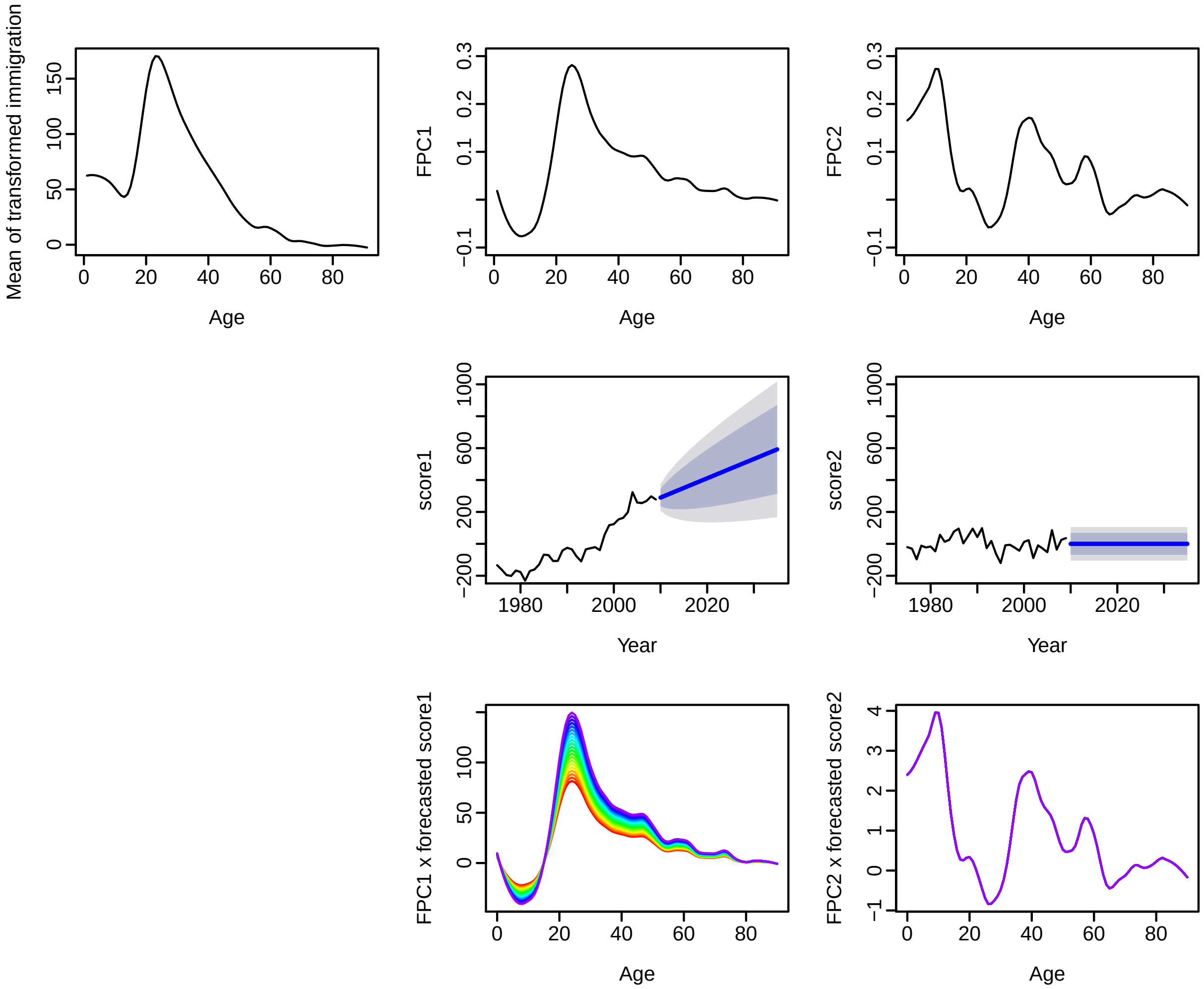
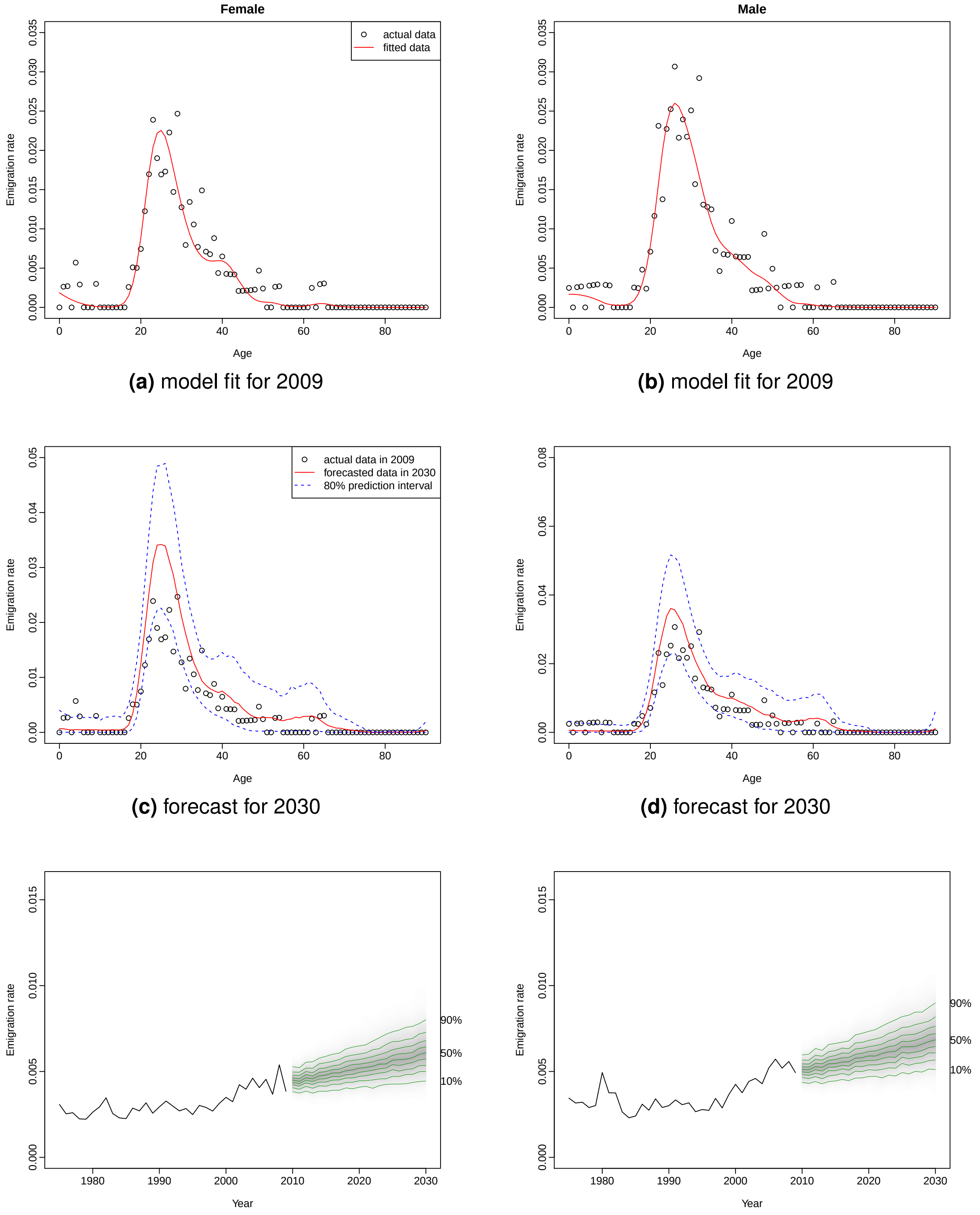


Figure 9. Functional principal components and their scores for the immigration counts. A decomposition of $K = 6$ has been used for analysis, although we display only the first two components. The solid line represents the point forecasts of scores, where the dark and light grey regions represent the 80% and 95% point-wise prediction intervals



(a) model fit for 2009

(b) model fit for 2009

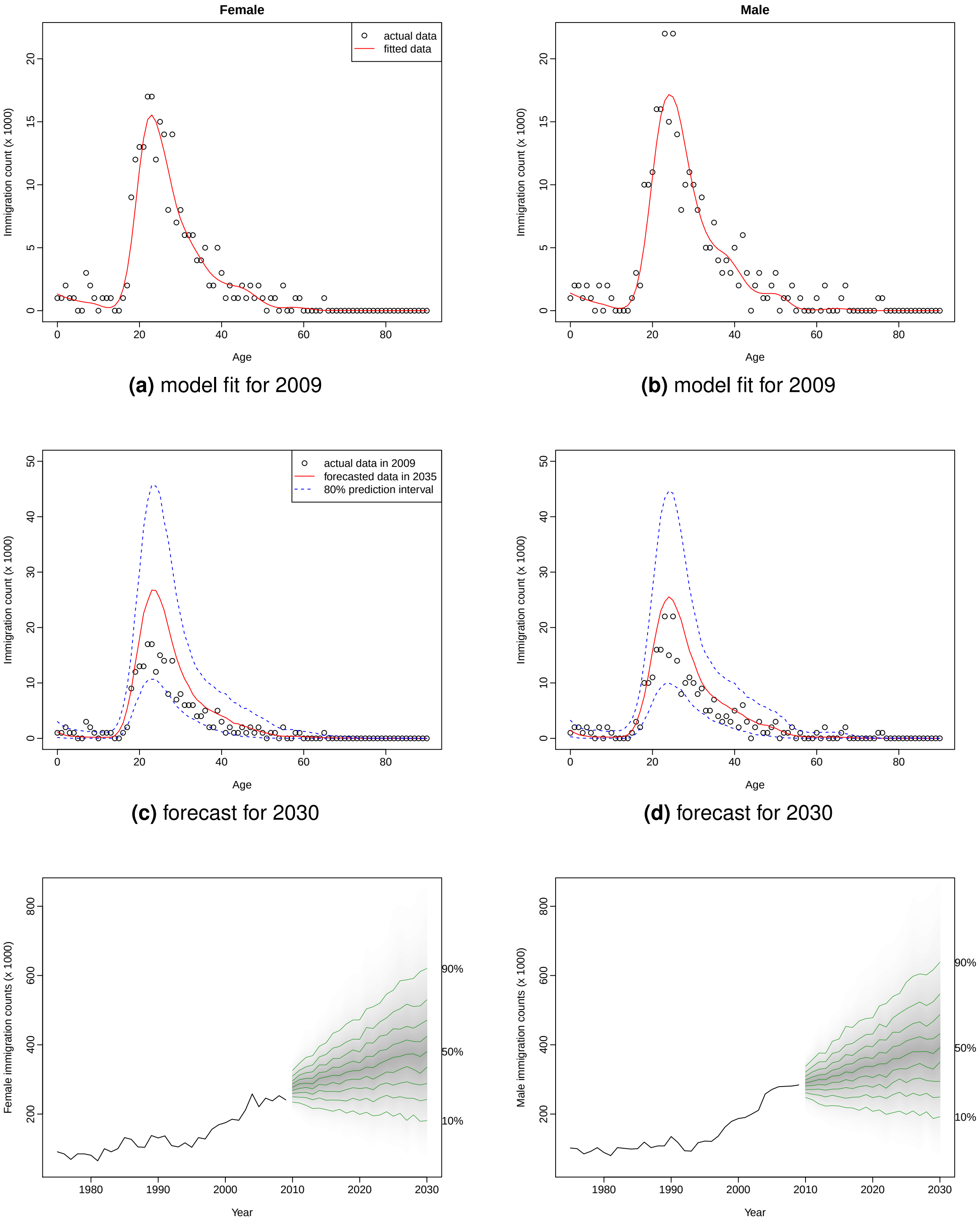
(c) forecast for 2030

(d) forecast for 2030

(e) mean emigration rate

(f) mean emigration rate

Figure 10. Model fit for female and male emigration to 2009 data, forecasted age-specific emigration rate for 2030, and forecasted female and male mean emigration rates across ages between 2010 and 2030



(a) model fit for 2009

(b) model fit for 2009

(c) forecast for 2030

(d) forecast for 2030

(e) total immigration count

(f) total immigration count

Figure 11. Model fit for female and male immigration to 2009 data, forecasted age-specific immigration counts for 2030, total female and male immigrant counts between 2010 and 2030

and coherent functional time-series methods are similar in shape to the actual population, except for the newborns between 2005 and 2009. This may be due to the unforeseen increase in fertility and increase in female immigrants due to European Union expansion (Waller et al. 2012).

In order to compare the point and interval forecast accuracy between the two functional methods, we use mean absolute forecast error (MAFE) and coverage probability difference (CPD) (see also Shang et al. 2011). The MAFE is the average of absolute error, |*actual* - *forecast*|, across years in the forecasting period and ages, it measures forecast precision regardless of sign. The CPD is the absolute difference between the nominal coverage probability and empirical coverage probability for each age, averaged over years in the forecasting period. They are given by

$$\text{MAFE}_h = \frac{1}{(11-h)\times 91}\sum_{k=1}^{11-h}\sum_{i=0}^{90+}\left|f_k(x_i) - \widehat{f}_k(x_i)\right|, \tag{7}$$

$$\text{CPD}_h = \frac{1}{11-h}\sum_{k=1}^{11-h}\left|\delta - \frac{1}{91}\sum_{j=0}^{90+} I\left[\widehat{f}_{k,lb}(x_i) < f_k(x_i) < \widehat{f}_{k,ub}(x_i)\right]\right|, \tag{8}$$

where $I[\cdot]$ represents the indicator function taking value 0 or 1, *lb* and *ub* represent the lower and upper bounds of the prediction interval, respectively; δ represents the nominal coverage probability, customarily 0.8. Values of MAFE_h and CPD_h closer to 0 indicate greater accuracy of the point and interval forecasts.

From Table 1, we see that the multilevel functional data method does not only have a smaller overall MAFE than the independent functional time-series method, but it also performs better in terms of CPD.

With the population in 2009 as a baseline, the age composition of forecasted population in 2030 is presented in Figure 13. Forecasts of the total male and female populations are presented on the right panel. We found that the age profile of the population in 2030 is mainly driven by future migration and, to some extent, fertility. The largest uncertainties are associated with the number of newborns, as well as the population at ages between 20 and 45, for both males and females. It is also expected that the number of older people will be increasing in 2030, as well as the working age group between 20 and 45.

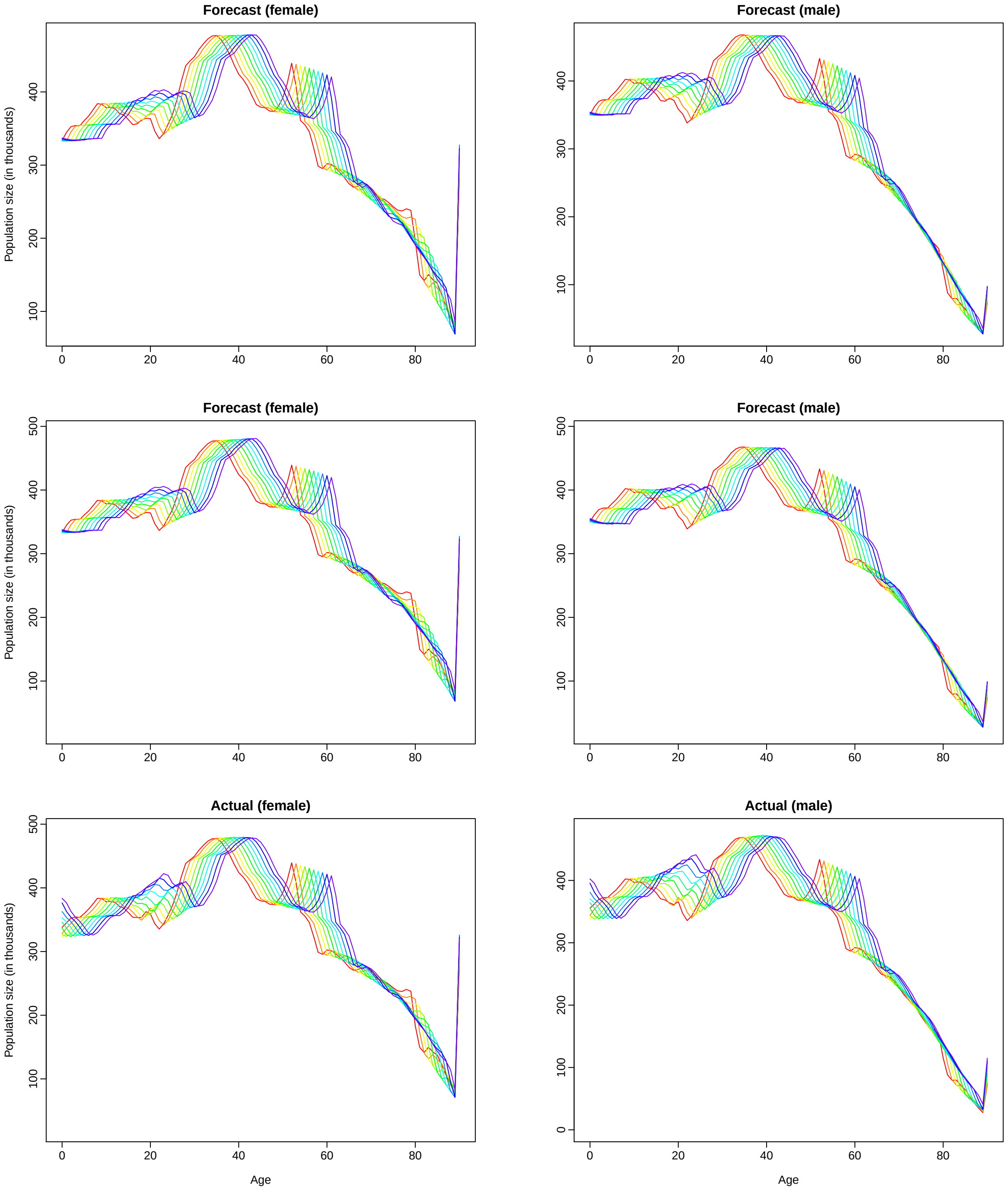


Figure 12. Forecasted age-specific female and male population size from 2000 to 2009. In the top row, we present the forecasts obtained from the independent functional time-series method. In the middle row, we present the forecasts obtained from the multilevel functional time-series method. In the bottom row, we present the actual holdout data

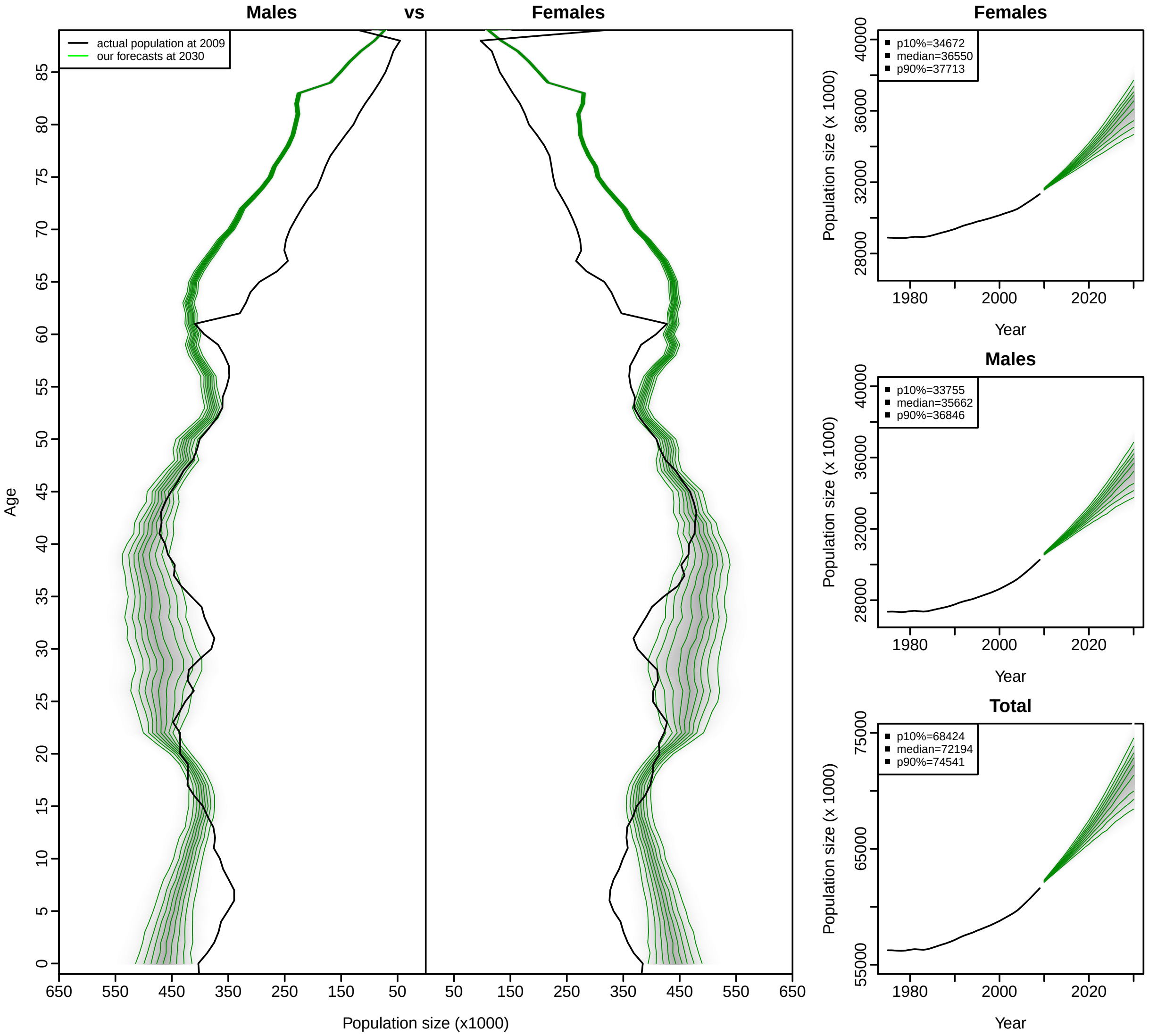


Figure 13. Forecasted age profiles of males and females (on the left) and forecasted population sizes of females, males and total (on the right)

horizon	MAFE				CPD			
	Female		Male		Female		Male	
h	Coherent	Ind	Coherent	Ind	Coherent	Ind	Coherent	Ind
1	1033.12	1062.78	941.16	1024.67	0.12	0.14	0.05	0.03
2	1694.88	1743.50	1557.71	1641.11	0.17	0.18	0.16	0.13
3	2132.37	2198.55	2158.08	2223.17	0.18	0.24	0.29	0.28
4	2296.70	2458.21	2618.12	2823.55	0.17	0.21	0.25	0.33
5	2424.59	2551.71	2972.96	3261.54	0.12	0.16	0.22	0.29
6	2708.18	2714.55	3543.36	3827.76	0.06	0.12	0.21	0.26
7	3249.01	3059.97	4403.38	4674.53	0.07	0.05	0.25	0.26
8	3908.40	3616.24	5805.94	5878.05	0.09	0.07	0.32	0.27
9	4662.73	4343.90	7291.31	7356.01	0.12	0.10	0.34	0.34
10	5347.75	5008.54	8386.71	8597.10	0.12	0.11	0.35	0.37
Mean	2945.78	2875.80	3967.87	4130.75	0.12	0.14	0.24	0.26

Table 1. Point and interval forecast accuracy comparison between the coherent (also known as multilevel) and independent functional data models. MAFE symbolizes mean absolute forecast error given in (7), while CPD represents coverage probability difference given in (8)

With 2009 as a baseline population, the median size of 2030 population is expected to reach 71.9 million, which is by 10.3 million larger than the population of 61.6 million in 2009. As shown in Table 2, we also compare our forecasts of total population with the 5-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 the recent policy of the UK government to reduce net migration to the levels below 100 thousand per year.

	Point forecast		80% prediction interval
	ONS	PB	PB
2010	62.262	62.193	(61.870, 62.557)
2015	64.776	64.210	(63.117, 65.386)
2020	67.173	66.449	(64.511, 68.430)
2025	69.404	69.040	(65.988, 72.019)
2030	71.392	71.859	(67.338, 75.937)

Table 2. Comparison of population forecasts (in million) between the ONS and the proposed parametric bootstrap (PB) method

6 CONCLUSION AND FUTURE RESEARCH

In this paper, 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. We combine the forecasts of mortality, fertility, emigration and immigration into the forecast of population, through a cohort component projection model. The advantage of our functional models 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. 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.

There are many ways in which the methodology presented here may be further extended, and here we mention four: (1) it may be possible to extend the multilevel functional data model to forecast age- and sex-specific population at subnational level; (2) it may be possible to develop a fully Bayesian functional principal component regression to forecast population, without conditioning on the fixed functional principal components; (3) the uncertainty of the baseline population size used for the projection could be incorporated into the forecast; and (4) we aim to explore the possibility of model averaging, which can adequately combine the forecasts obtained from several population projection methods. Model averaging may improve the point and interval forecast accuracy, as demonstrated in the context of mortality forecasting by Shang (2012). This work provides a natural foundation for such extensions.

APPENDIX A: FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS FROM A KERNEL OPERATOR VIEWPOINT

Throughout the paper, functional principal component decomposition provides a data-driven way to reduce the infinite-dimensional object to finite dimensions. Here, we state without proof the underlying concept of functional principal component analysis, and refer to Tran (2008) for detailed proofs.

Let f be a random variable $f : \Omega \to L^2(\mathcal{I})$, such as $f \in L^2(\Omega)$. f can also be seen as a stochastic process defined on a compact set $\mathcal{I}$, with finite variance $\int_{\mathcal{I}} \mathrm{E}(f^2) < \infty$. Let μ be the mean function of f, without loss of generality, let $f^c = f - \mu$ be a centered stochastic process.

Definition 1 *(Covariance operator) The covariance function of f is defined to be the function $K : \mathcal{I} \times \mathcal{I} \to R$, such that*

$$\begin{aligned} K(u, v) &= Cov[f(u), f(v)] \\ &= E\{[f(u) - \mu(u)][f(v) - \mu(v)]\}. \end{aligned}$$

By assuming f is a continuous and square-integrable covariate function, the function K induces the kernel operator $L^2(\mathcal{I}) \to L^2(\mathcal{I})$, $\phi \to K\phi$, given by

$$(K\phi)(u) = \int K(u, v)\phi(v)dv.$$

Lemma 1 *(Mercer's Lemma)*

Assume that K is continuous over $\mathcal{I}^2$, there exists an orthonormal sequence (ϕ_k) of continuous function in $L^2(\mathcal{I})$ and a non-increasing sequence (λ_k) of positive numbers, such that

$$K(u, v) = \sum_{k=1}^{\infty} \lambda_k \phi_k(u)\phi_k(v), \qquad u, v \in \mathcal{I}.$$

Theorem 1 *(Karhunen-Loève expansion)*

With Mercer's lemma, a stochastic process f can be expressed as

$$\begin{aligned} f(u) &= \mu(u) + \sum_{k=1}^{\infty} \sqrt{\lambda_k}\xi_k\phi_k(u), \\ &= \mu(u) + \sum_{k=1}^{\infty} \beta_k\phi_k(u), \end{aligned}$$

where $\xi_k = \frac{1}{\sqrt{\lambda_k}} \int_I f^c(v)\phi_k(v)dv$ is an uncorrelated random variable with zero mean and unit variance. In practice, we keep the first K terms, which reduces the infinite-dimensional object to finite dimensions.

APPENDIX B: CODE USED FOR ESTIMATING VARIANCE PARAMETERS IN THE PARAMETRIC BOOTSTRAPPING

Statistical software WinBUGS (Lunn et al. 2009) originally designed for performing Bayesian analysis is used here to estimate variances in the principal component scores and error function, and sample principal component scores and error function, for the independent functional data model. Below is a modified version of WinBUGS code originally given in Crainiceanu & Goldsmith (2010).

```
model {
   for (i in 1:N_subj)
   {
         for(t in 1:N_obs)
         {
                W[i,t] ~ dnorm(m[i,t], taueps)
                m[i,t] <- xi[i,1] * E[t,1] + xi[i,2] * E[t,2] +
                          xi[i,3] * E[t,3] + xi[i,4] * E[t,4] +
                          xi[i,5] * E[t,5] + xi[i,6] * E[t,6]
         }
         for(k in 1:dim_space)
         {
                xi[i,k] ~ dnorm(0, ll[k])
         }
  }
  for(k in 1:dim_space)
  {
        ll[k] ~ dgamma(1.0E-3, 1.0E-3)
        lambda[k] <-1/ll[k]
  }
  taueps ~ dgamma(1.0E-3, 1.0E-3)
  lambda_taueps <- 1/taueps
}
```

1. N_subj is the number of subjects (sample size)
2. N_obs in the number of observations within subjects (dimension)
3. W[i, t] represents the data for age t and subject i
4. m[i, t] is the smoothed mean of W[i,t]

5. taueps is the precision of the error function
6. lambda_taueps is the variance of the error function
7. xi[i,k] is the score of the ith subject on the kth functional principal component
8. E[t,k] is the kth functional principal component evaluated at age *t* The matrix E is N_obs x K and is loaded as non-missing value data
9. dim_space is the number of functional principal components
10. ll[k] are the precisions of the scores xi[i,k]
11. lambda[k] are the variances of the scores xi[i,k]
12. all precision priors are Gamma priors with mean 1 and variance 1000

REFERENCES

Ahlburg, D.A. and Land, K.C. (1992) Population forecasting: guest editors' introduction. *International Journal of Forecasting,* 8 (2), 289–299.

Alho, J., Alders, M., Cruijsen, H., Keilman, N., Nikander, T. and Pham, D.Q. (2006) New forecast: Population decline postponed in Europe. *Statistical Journal of the United Nations Economic Commission for Europe,* 23, 1–10.

Alho, J.M. and Spencer, B.D. (1985) Uncertain population forecasting. *Journal of the American Statistical Association,* 80 (390), 306–314.

Alho, J.M. and Spencer, B.D. (2005) *Statistical Demography and Forecasting*, Springer, New York.

Bongaarts, J. and Bulatao, R.A. (eds) (2000) *Beyond Six Billion: Forecasting the World's Population*, National Academy Press, Washington DC.

Booth, H. (2006) Demographic forecasting: 1980-2005 in review. *International Journal of Forecasting,* 22 (3), 547–581.

Booth, H., Maindonald, J. and Smith, L. (2002) Applying Lee-Carter under conditions of variable mortality decline. *Population studies,* 56 (3), 325–336.

Bosq, D. (2000) *Linear Processes in Function Spaces: Theory and Applications*, Springer, New York.

Bryant, J.R. and Graham, P.J. (2013) Bayesian demographic accounts: Subnational population estimation using multiple data sources. *Bayesian Analysis,* 8 (2), 1–32.

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.

Crainiceanu, C.M. and Goldsmith, J.A. (2010) Bayesian functional data analysis using WinBUGS. *Journal of Statistical Software,* 32 (11).

Crainiceanu, C.M., Staicu, A.M. and Di, C.Z. (2009) Generalized multilevel functional regression. *Journal of the American Statistical Association,* 104 (488), 1550–1561.

Czado, C., Delwarde, A. and Denuit, M. (2005) Bayesian Poisson log-bilinear mortality projections. *Insurance: Mathematics and Economics,* 36 (3), 260–284.

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.

De Beer, J. (2000) Dealing with uncertainty in population forecasting, *Working paper,* Department of Population, Statistics Netherlands. http://www.cbs.nl/nr/rdonlyres/7dc466f9-fe4c-48dc-be9030216b697548/0/dealingwithuncertainty.pdf

de la Horra, J., Marín, J.M. and Rodríguez-Bernal, M.T. (2013) Bayesian inference and data cloning in population projection matrices, *Working paper 02,* Universidad Carlos III de Madrid.
http://e-archivo.uc3m.es/bitstream/10016/16113/1/ws130102.pdf

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.

Dublin, L.I. and Lotka, A.J. (1925) On the true rate of natural increase: As exemplified by the Population of the United States, 1920. *Journal of the American Statistical Association,* 20 (151), 305–339.

Efron, B. (2011) The bootstrap and Markov-chain Monte Carlo. *Journal of the Biopharmaceutical Statistics,* 21 (6), 1052–1062.

Eubank, R.L. (1999) *Nonparametric Regression and Spline Smoothing*, Marcel Dekker, New York.

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.

Hall, P. (2011) Principal component analysis for functional data: methodology, theory and discussion, *in* 'The Oxford Handbook of Functional Data Analysis', Oxford University Press, Oxford, pp. 210–234.

He, X. and Ng, P. (1999) COBS: Qualitatively Constrained Smoothing via Linear Programming. *Computational Statistics,* 14, 315Ð337.

Human Fertility Database (2013) *Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria)*. Accessed on 8/January/2013. 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 on 8/January/2013. http://www.mortality.org/

Hyndman, R. and Booth, H. (2008) Stochastic population forecasts using functional data models for mortality, fertility and migration. *International Journal of Forecasting*, 24 (3), 323–342.

Hyndman, R.J., Booth, H. and Yasmeen, 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 discussions). *Journal of the Korean Statistical Society,* 38 (3), 199–211.

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.

Kajin, M., Almeida, P.J.A.L., Vieira, M.V. and Cerqueira, R. (2012) The state of the art of population projection models from the Leslie matrix to evolutionary demography. *Oecologia Australis,* 16 (1), 13–22.

Keilman, N. (1990) *Uncertainty in National Population Forecasting: Issues, Backgrounds, Analyses, Recommendations*, Swets & Zeitlinger, Amsterdam.

Koissi, M., Shapiro, A.F. and Högnäs, G. (2006) Evaluating and extending the Lee-Carter model for mortality forecasting: Bootstrap confidence interval. *Insurance: Mathematics and Economics,* 38 (1), 1–20.

Kreyenfeld, M., Andersson, G. and Pailhé, A. (2012) Economic uncertainty and family dynamics in Europe. *Demographic Research,* 27, 835–852.

Lazar, D. Denuit, M. M. (2009) A multivariate time series approach to projected life table. *Applied Stochastic Models in Business and Industry,* 25 (6), 806–823.

Lee, R.D. and Carter, L.R. (1992) Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association,* 87 (419), 659–671.

Lee, R.D. and Tuljapurkar, S. (1994) Stochastic population projections for the United States: Beyond high, medium and low. *Journal of the American Statistical Association*, 89 (428), 1175–1189.

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.

Lunn, D., Spiegelhalter, D., Thomas, A. and Best, N. (2009) The BUGS project: Evolution, critique and future directions. *Statistics in Medicine,* 28 (25), 3049–3067.

Lutz, W. (ed.) (1996) *The Future Population of the World: What Can We Assume Today*, Earthscan, London.

Lutz, W. and Goldstein, J.R. (2004) Introduction: How to deal with uncertainty in population forecasting', *International Statistical Review* 72 (1), 1–4.

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. http://www.ons.gov.uk/ons/rel/npp/national-population- projections/ 2010-based-reference-volume--series-pp2/index.html

Pinheiro, J.C. and Bates, D.M. (2000) *Mixed Effects Models in S and S-PLUS*, Springer Verlag, New York.

Preston, S.H., Heuveline, P. and Guillot, M. (2001) *Demography: Measuring and Modeling Population Processes*, Blackwell, UK.

Raftery, A.E., Li, N., Ševčiková, H., Gerland, P. and Heilig, G.K. (2012) Bayesian probabilistic population projections for all countries. *Proceedings of the National Academy of Sciences of the United States of America,* 109 (35), 13915–13921.

Ramsay, J.O. (1988) Monotone regression splines in action. *Statistical Science,* 3 (4), 425–441.

Ramsay, J. and Silverman, B. (2005) *Functional Data Analysis*, 2nd edition, 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 P. Rees, (eds) *Population Structures and Models*, George Allen & Unwin, UK.

Renshaw, A. and Haberman, S. (2006) A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics,* 38 (3), 556–570.

Renshaw, A. and Haberman, S. (2003) Lee-Carter mortality forecasting: A parallel generalized linear modelling approach for England and Wales mortality projections. *Journal of the Royal Statistical Society Series C (Applied Statistics),* 52 (1), 119–137.

Rogers, A. (1975) *Introduction to Multiregional Mathematical Demography*, John Wiley & Sons, New York.

Rogers, A. (1995) *Multiregional demography: Principles, Methods and Extensions*, John Wiley & Sons, New York.

Shang, H.L. (2012) Point and interval forecast of age-specific life expectancies: A model averaging approach. *Demographic Research,* 27, 593–644.

Shang, H.L. (2013) A survey of functional principal component analysis. *AStA Advances in Statistical Analysis,* in press.

Shang, H.L., Booth, H. and Hyndman, R.J. (2011) Point and interval forecasts of mortality rates and life expectancy: A comparison of ten principal component methods. *Demographic Research,* 25 (5), 173–214.

Staicu, A.M., Crainiceanu, C.M. and Carroll, R.J. (2010) Fast methods for spatially correlated multilevel functional data. *Biostatistics,* 11 (2), 177–194.

Tran, N.M. (2008) An introduction to theoretical properties of functional principal component analysis. *Honours thesis*, The University of Melbourne. http://www.stat.berkeley .edu /~tran/pub/honours_thesis.pdf

Wahba, G. (1990) *'*Spline Models for Observational Data'*,* Society for Industrial and Applied Mathematics Philadelphia.

Waller, L., Berrington, A. and Raymer, J. (2012) Understanding recent migrant fertility in the United Kingdom. *Working paper 27,* ESRC Centre for Population Change. http://cpc.geodata.soton.ac.uk/publications/2012_Understanding_recent_migrant_fertility_WP27_Waller_et_al.pdf

Wilson, C. (2001) On the scale of global demographic convergence 1950-2000. *Population and Development Review,* 27 (1), 155–172.

Wilson, T. and Rees, P. (2005) Recent developments in population projection methodology: A review. *Population, Space and Place,* 11 (5), 337–360.

Wiśniowski, A., Smith, P.W.F., Bijak, J. and Raymer, J. (2013) Bayesian cohort component population forecasts, *in* 'Paper presented at Population Association of America', New Orleans, LA. http://paa2013.princeton.edu/papers/130993

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.

Zipunnikov, V., Caffo, B., Yousem, D.M., Davatzikos, C., Schwartz, B.S. and Crainiceanu, C.M. (2011) Multilevel functional principal component analysis for high-dimensional data. *Journal of Computational and Graphical Statistics*, 20 (4), 852–873.

ESRC Centre for Population Change
Building 58, Room 2001
Faculty of Social and Human Sciences
University of Southampton
SO17 1BJ

T: +44 (0)2380 592579
E: cpc@soton.ac.uk
www.cpc.ac.uk

To subscribe to the CPC newsletter and keep up-to-date with research activity, news and events, please register online: www.cpc.ac.uk/newsletter

You can also follow CPC on Twitter, Facebook and Mendeley for our latest research and updates:

www.facebook.com/CPCpopulation

www.twitter.com/CPCpopulation

www.mendeley.com/groups/3241781/centre-for-population-change

The ESRC Centre for Population Change (CPC) is a joint initiative between the University of Southampton and a consortium of Scottish universities including St Andrews, Edinburgh, Stirling and Strathclyde, in partnership with the Office for National Statistics and National Records of Scotland.