Insurance: Mathematics and Economics 78 (2018) 339–350

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

The double-gap life expectancy forecasting model☆

Marius D. Pascariu [a,*], Vladimir Canudas-Romo [b], James W. Vaupel [c]

[a] Institute of Public Health, University of Southern Denmark, Denmark
[b] School of Demography, Australian National University, Canberra, Australia
[c] Max–Planck Institute for Demographic Research, Germany

ARTICLE INFO

Article history:
Available online 2 October 2017

Keywords:
Life expectancy forecasting
Mortality modelling
Best-practice trends
Sex-gap
Time series models

ABSTRACT

Life expectancy is highly correlated over time among countries and between males and females. These associations can be used to improve forecasts. Here we propose a method for forecasting female life expectancy based on analysis of the gap between female life expectancy in a country compared with the record level of female life expectancy in the world. Second, to forecast male life expectancy, the gap between male life expectancy and female life expectancy in a country is analysed. We present these results for various developed countries. We compare our results with forecasts based on the Lee–Carter approach and the Cairns–Blake–Dowd strategy. We focus on forecasting life expectancy at age 0 and remaining life expectancy at age 65.

1. Introduction

The history of the evolution of life expectancy is of crucial importance for demographers and actuaries who want to develop more accurate forecasting models. Between 1840 and 2014 no more than seven countries have been the record holders of female life expectancy at birth; starting with Sweden and Norway in the 19th century and finishing with present day Japan. The competition among countries to reduce mortality levels resulted in a remarkable linear rise as presented by Oeppen and Vaupel (2002), or a segmented linear trend as suggested by Vallin and Meslé (2009). In developed countries, the linear trend in period life expectancy has proven itself to better fit trends in human mortality than more complex mathematical models based on age-specific death rates (White, 2002). The rate of change in age-specific death rates have less regular patterns over time than life expectancy, which is an age-aggregated measure. Thus, although life expectancy loses specificity it compensates in terms of accuracy. Furthermore, data highly aggregated by age give valuable information that can be used to tackle the issue of mortality forecasting from a clearer perspective.

Torri and Vaupel (2012) built on the idea that future human longevity is given by a general life expectancy trend. Their model at first forecasts the world's record life expectancy and then the gap between the record and the current life expectancy of a particular population of interest assuming a tendency towards convergence with the predicted record level. The Torri–Vaupel approach is promising but has the drawback that populations that lag behind record life expectancy cannot become the record holder; in addition the interdependence between the sexes is not recognized. Furthermore, no population's life expectancy can exceed the forecast record.

Between 1950 and 2014 the record holder for life expectancy at birth changed more than 15 times among 5 countries; and in the same manner the record holder for life expectancy at age 65 changed more than 10 times among 6 countries. This indicates that the record is not given by a single reference population. The case of Japan shows that a country with a very low level of life expectancy, which was the case immediately after World War II in this country, can improve at a fast pace, catch up with the low mortality populations and eventually become the record holder. How long a population can maintain the status of record holder is an open question. A method that can capture change in the record-holder is highly relevant. We propose such a method by using the trend-line of record life expectancy, instead of the actual record values. The use of trend-line implies that the best-practice country in a given year can be above the best-practice line. This fact was shown by Oeppen and Vaupel (2002).

The majority of the forecasting models used by demographers and actuaries tend to predict future longevity for specific countries separately for males and females. One reason could be that females, as a group, have a different mortality age-pattern from males. They live longer and the death rates for females are lower than those for males at all ages, even before birth and in almost every country in the world (Austad, 2006). The most pronounced discrepancy

☆ Article presented at Longevity 11: The Eleventh International Longevity Risk and Capital Markets Solutions Conference, 7-9 September 2015, Université Lyon 1, Lyon, France.

* Corresponding author.
E-mail address: mpascariu@health.sdu.dk (M.D. Pascariu).

https://doi.org/10.1016/j.insmatheco.2017.09.011

can be observed in the very old, among centenarians and supercentenarians (persons with an age of 110 and more) when women outnumber men by more than nine to one (Perls and Fretts, 1998). The sex gap in life expectancy widened and then shrank in the last half of the last century as the rate of improvement in female life expectancy exceeded that for males. Thus, the available evidence indicates the presence of behavioural as well as biological differences between the sexes, and social and psychological factors all play important roles in differentiating the mortality patterns for females and males. To simplify analysis an assumption generally made is that females and males are two different populations independent of each other.

Li and Lee (2005) introduced a method for forecasting death rates of different populations and for both sexes that are not expected to diverge, using an augmented common factor model. Hyndman et al. (2013) propose a method for coherent forecasting of mortality rates in different subpopulations based on functional principal components models of simple functions of rates. The product-ratio functional forecasting method models the geometric mean of subpopulation rates and the ratio of subpopulation rates to product rates. Raftery et al. (2013) also discuss the possibility of forecasting life expectancy using a two-sex model, and develop this idea with the introduction of an elegant model to obtain joint probabilistic projections of life expectancy for both sexes (Raftery et al., 2014). First, female life expectancy is forecast using a Bayesian hierarchical model and then the gap between female and male is modelled, recognizing in a formal way the correlation in mortality.

Further knowledge can be gained by integrating the idea of the life expectancy correlation between sexes and also between countries, into a single model. The main objective of this article is to present such a model.

The remainder of the article is organized as follows. First, in Section 2 the data used in fitting the model are presented. In Section 3 a new life expectancy projection model is proposed. In Section 4 a method to assess the performance of the model is given. Section 5 shows simulation results and illustrations of life expectancy in several countries by sex. The discussion and conclusion are in Section 6.

2. Data description

The data source used in this article is the Human Mortality Database (2017), which contains historical mortality data for 47 homogeneous populations in different countries and regions. HMD constitutes a reliable data source because it includes high quality historical mortality data that was subject to a uniform set of procedures, guaranteeing the cross-national comparability of the information.

For the purpose of our analysis we have focused on a subset of these data covering calendar years 1950–2014 and the 0–95 age range in 38 countries and regions, giving 76 sex-specific populations. The selected populations must have sufficient size to allow the fitting of a forecasting model and should be unique, meaning that a person included in one population should not be included in others. The selected countries are shown in Table 1 along with the dates used to define the fitting periods.

3. The method

The objective is to construct a model for forecasting life expectancy of female and male life expectancy at any age. The model is based on correlations existing among countries and between sexes. The method combines separate forecasts to obtain joint female and male life expectancies that are coherent with the best-practice trend and correlated.

The model construction follows four steps:

1. Best-practice life expectancy is identified in order to get a general sense of the direction and the rate of change in human mortality.
2. The gap between female life expectancy and the best-practice trend in the world is forecast using a classic time series model, thus determining future female life expectancy.
3. The gap between male and female life expectancy is forecast with the help of a linear model to obtain the country specific male life expectancy.
4. Prediction intervals are constructed from a multivariate normal distribution with mean zero and covariance matrix given by the residuals generated in the fitting of the three time series in the previous steps.

The core of the proposed double-gap model can be summarized by two equations: first future female life expectancy at age x, time t and country k, $e^f_{k,x,t}$, can be obtained as the difference between future best-practice life expectancy at that age and time, $e^{bp}_{x,t}$, and a predicted gap or distance, $D_{k,x,t}$, of the performance of the specific lagging country or region,

$$e^f_{k,x,t} = e^{bp}_{x,t} - D_{k,x,t}. \tag{1}$$

Similarly future life expectancy for the male population is modelled as the difference between future female life expectancy and the sex gap, $G_{k,x,t}$, in life expectancy,

$$e^m_{k,x,t} = e^f_{k,x,t} - G_{k,x,t}. \tag{2}$$

3.1. Step 1—the best-practice trend

The best-practice trend in life expectancy is defined as the predicted value of a linear model based on the female record life expectancy time series of the form

$$e^{record}_{x,t} = \alpha_{x0} + \alpha_{x1}t, +\epsilon^{(0)}_{x,t}, \text{ with } t = 1, 2, 3... \tag{3}$$

therefore,

$$e^{bp}_{x,t} = \alpha_{x0} + \alpha_{x1}t, \tag{4}$$

where $e^{record}_{x,t}$ denotes the record life expectancy at age x and time t, $e^{bp}_{x,t}$ is the best-practice trend, α_{xi} represent the parameters of the model fitted at age x, and the errors $\epsilon^{(0)}_{x,t}$ are independent and identically distributed random variables normally distributed with mean zero and variance $\sigma^{(0)}$. To predict future best-practice levels we will follow the past regularity observed in improvement in life expectancy and extrapolate directly the future trend.

The Double-Gap model in Eqs. (1) and (2) are applied here to life expectancies at birth and at age 65: Analysing the period between 1950 and 2014 we can observe that the record life expectancy at birth increased at a rate of 2.1 years per decade from 73.5 to 86.8, while at age 65 the improvement was on average 1.27 years per decade, captured in the parameter α_{x1} in Eq. (3). These rates of increase at age 65 imply a change from 16.3 years in 1950 in Iceland to 24.2 in 2014 in Japan. The linear fit is presented in Fig. 1.

3.2. Step 2—the gap to best-practice trend, $D_{k,x,t}$

One way to forecast the gap between the best-practice trend and country specific female life expectancy is to use the classic *ARIMA* model (Box and Jenkins, 1976). This is appropriate when the data set is sufficiently long and exhibits a stable and consistent pattern over time with few outliers.

Table 1
Selected HMD countries and years with available data used for the illustration.
Source: Human Mortality Database (2017).

Available data	Countries and regions
1950–2010	Bulgaria
1950–2011	Canada
1950–2012	Italy
1950–2013	Scotland, England & Wales, Iceland, New Zealand
1950–2014	Australia, Austria, Belgium, Czech Republic, Denmark, Finland, France, Hungary, Ireland, Japan, Netherlands, Norway, Portugal, Spain, Slovakia, Switzerland, Sweden, U.S.A.
1956–2014	East Germany, West Germany
1958–2014	Poland, Russia
1959–2013	Estonia, Latvia, Lithuania, Ukraine
1959–2014	Belarus
1970–2014	Taiwan
1981–2013	Greece
1983–2014	Israel, Slovenia

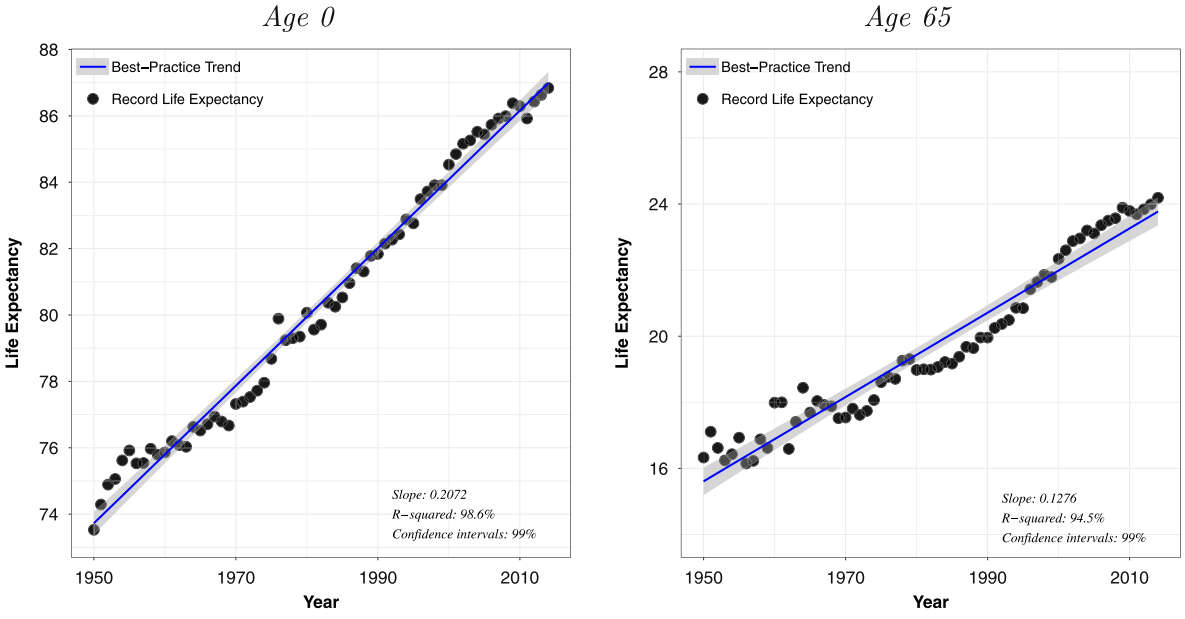

Fig. 1. The trend of record female life expectancy at birth and at age 65 between 1950 and 2014.

In general notation, we have an *ARIMA*(p, d, q) model, where p is the order of the autoregressive process, d indicates the order of integration, namely the number of times that the series must be differenced in order to make it stationary, and q is the order of the moving average process. The general form of an *ARIMA*(p, d, q) model for a stochastic process $D_{x,t}$ is given by

$$\nabla^d D_{k,x,t} = \underbrace{\mu_{k,x}}_{\text{Drift}} + \underbrace{\sum_{i=1}^{p} \phi_i \nabla^d D_{k,x,t-i}}_{\text{Regression}} + \underbrace{\epsilon^{(1)}_{k,x,t} + \sum_{j=1}^{q} \theta_j \epsilon^{(1)}_{k,x,t-j}}_{\text{Smoothed noise}} \tag{5}$$

where the response can be obtained from the linear regression of previous gaps plus additional *smoothed noise*. We denote with $\nabla^d D_{k,x,t}$ the stationary (transformed) time series used to fit the *ARIMA* model. The constant parameter $\mu_{k,x}$ is the drift, indicating the average change in the series over time; ϕ_i are the parameters of the auto-regressive part, and θ_j are the parameters of the moving average part. Finally $\epsilon^{(1)}_{k,x,t}$ is a sequence of independent and identically distributed random variables with mean zero and variance $\sigma^{(1)}$ (see Table 2).

For each country and period of time an appropriate model is fitted so that it captures the information given by the past pattern of the gap. We consider *ARIMA*(p, d, q) models where d is selected based on successive KPSS unit-root tests (Kwiatkowski et al., 1992). That is, we test the data for a unit root; if the test result is significant, we test the differenced data for a unit root; and so on until non-significant. Once the order of difference d is selected, we proceed to select the values of p and q by minimizing the AIC. Finally based on the historical trend we decide whether a drift should be allowed in the model.

Table 2
Estimated parameters of the ARIMA model for the gap between best-practice and country specific data at birth and at Age 65, 1950–2014.
Source: Authors' calculations based on data described in Table 1.

	Age	Rank	μ	ϕ_1	ϕ_2	θ_1
USA	0	(0, 1, 0)	–	–	–	–
	65	(0, 1, 0)	–	–	–	–
France	0	(1, 1, 0)	–	−0.3519	–	–
	65	(1, 1, 1)	–	−0.3048	–	−0.4533
Sweden	0	(2, 1, 1)	0.0283	−1.1521	−0.5065	0.9173
	65	(0, 1, 1)	0.0175	−0.6694	–	–

An analysis for the case of France over the 1950–2014 period indicated that the *ARIMA*(1, 1, 0) for age 0 and *ARIMA*(1, 1, 1) for age 65 are the most suitable models for describing the data. For the USA the random walk with no drift is found to be the most parsimonious model for both ages, but for Sweden, *ARIMA* models with a higher rank degree are needed. Estimated future values of the gap in 2050, together with 80% and 95% prediction intervals, are plotted in Fig. 2.

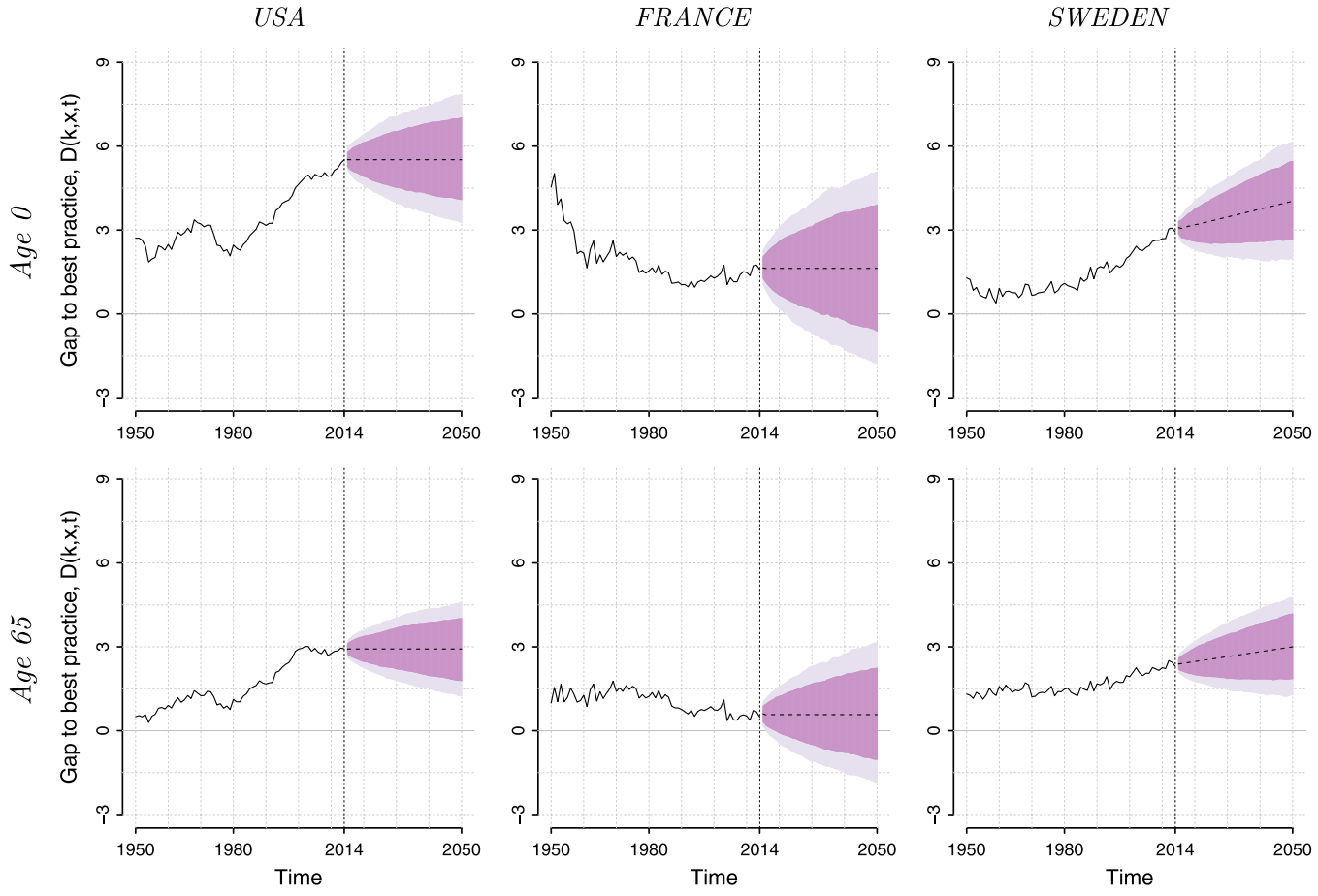

Fig. 2. The forecast gap between the best-practice trend and country-specific female life expectancy at birth and at age 65, with associated 80% and 95% prediction intervals, 1950–2050.

$$G^*_{k,x,t} = \begin{cases} \beta_0 + \underbrace{\beta_1 G_{k,x,t-1} + \beta_2 G_{k,x,t-2}}_{\text{Autoregressive model}} + \underbrace{\beta_3 (e^f_{k,x,t} - \tau)_+}_{\substack{\text{Level associated with} \\ \text{life expectancy when the gap} \\ \text{starts narrowing}}} + \epsilon^{(2)}_{k,x,t} & \text{if,} \quad e^f_{k,x,t} \le A, \\ \underbrace{G_{k,x,t-1} + \epsilon^{(3)}_{k,x,t}}_{\text{Random walk}} & , \quad \text{otherwise.} \end{cases} \tag{6}$$

Box I.

The forecast gaps for France show that the French female population could surpass the best practice trend in the future. This information is given by the lower side of the 80% and 95% prediction limits which are below zero. The forecasts for Sweden suggest a continuation of the historical trend where improvement in life expectancy at birth and age 65 is lower than the pace given by our selected benchmark, namely the best-practice trend. However the speed of divergence is slow, approximately one year of life expectancy in a 40 year forecasting horizon. For the USA, the forecasts suggest little change.

3.3. Step 3—the sex gap model, $G_{k,x,t}$

To predict the gap in life expectancy between females and males, $G_{k,x,t}$, at a given age x for specified country k at time t we apply a method that consists of a linear model and a random walk process with no drift.

The linear model takes into account the gap in the previous two years and an additional term that relates to female life expectancy. This term is given by $(e^f_{k,x,t} - \tau)_+$ where τ is the level of life expectancy at the time when the sex gap is expected to stop widening and start narrowing. The notation $(z)_+$ represents the maximum value between zero and z. The linear model is fitted over all ages lower than the level of female life expectancy, A. The levels of τ and A are determined from historical data by maximizing the resulting maximum likelihoods of our linear model over integer values of τ and A. In the statistical software ***R*** the linear model can be fitted using the *crch* package (Messner and Zeileis, 2015).

Because there is little evidence to make any assumptions about future pattern of the female–male gap at advanced ages (Raftery et al., 2014) the random walk model will be used to further fit and predict the evolving gap if life expectancy surpasses the obtained limit A, Eq. (6) is given in Box I.

As a further check we ensure that the modelled gap will always be between the observed historical minimum and maximum values of the female–male gap,

$$G_{k,x,t} = min\{max\{G^*_{k,x,t}, L\}, U\}, \tag{7}$$

Table 3
Estimated parameters for sex-gap forecast models for life expectancy at birth and Age 65.
Source: Authors' calculations based on data described in Table 1

Parameters	Estimate *Age 0*	Estimate Age 65	$Pr(>\|t\|)$ *for both Ages*
β_0	0.21257	0.14052	<2e−16
β_1	0.82184	0.64807	<2e−16
β_2	0.15971	0.32943	<2e−16
β_3	−0.02690	−0.01442	<2e−16
τ	75	15	
A	86	24	
L	0.99	0.33	
U	13.68	5.24	

where L and U are the minimum and maximum observed gaps, respectively. The errors $\epsilon^{(2)}_{k,x,t}$ and $\epsilon^{(3)}_{k,x,t}$ are independent and identically distributed random variables normally distributed with mean zero and variance $\sigma^{(2)}$ and $\sigma^{(3)}$, respectively.

The presented method is similar with the linear model used by Raftery et al. (2014). In order to obtain joint probabilistic forecasts of life expectancies for female and male populations, Raftery et al. modelled the relation between the two by projecting the sex-gap using a linear regression with different levels of female life expectancy as covariates. The model is applied to World Population Prospects 2008 set of quinquennial data starting in 1950 (United Nations, 2009).

We chose to adopt a modified version of the Raftery model because several covariates in the original model, which was constructed for projecting 5 years intervals, were not statistically significant for a 1-year step projection model. Also, we decided not to impose any dependency of an initial life expectancy in our model as in the original Raftery model. This decision was taken because an important number of time series in the Human Mortality Database start after 1950 as shown in Table 1.

The model is fitted using the data from all the countries in order to obtain the coefficient values and then it is used to forecast the gap for each country separately, using country specific female life expectancy.

Estimates of the model parameters are provided in Table 3 for the models fitted at age 0 and age 65, respectively. The parameter β_0 denotes the intercept level, which could be interpreted as a biological gap between the sexes; β_1 and β_2 represent the effect of the previous two gaps at time $t-1$ and $t-2$, influencing the range of possible values for the new gap. Together the first three parameters, β_0, β_1 and β_2 explain the majority of the gap trend. The negative β_3 parameter gives the speed of the convergence between the female and male life expectancies. As shown in Table 3, the life expectancies at birth are converging faster than those at age 65.

The forecast values of the sex gap in the USA, together with 80% and 95% prediction intervals based on the 1950–2014 data, can be observed in Fig. 3. In all three countries, and indeed in many other developed countries, the sex gap increased between 1950 and about 1980, and then decreased to 2014. The models for age 0 suggest a continuation of the descending trend until the beginning of 2030 where the gap will remain approximately constant.

The transition from a decreasing gap to stagnation coincides with the shift from the linear model to the random walk model described in Eq. (6). For instance in France, where currently life expectancy is higher than in the USA, the period of time needed to reach a value of life expectancy of 86 years for female population is shorter i.e. resulting in a projection with a shorter period of time with a decreasing sex-gap. In USA and Sweden the forecast gap in 2050 is approximately 3 years but in France it is 6 years for life expectancy at birth. For life expectancy at age 65 the models forecast very little change. Also, even if it is not impossible, the models suggest that it is highly unlikely that the sex-gap would become negative and a higher life expectancy for males would be experienced in any of the three countries either at age 0 or 65.

3.4. Step 4—dealing with correlated prediction intervals

Our approach to forecasting combines different models that generate separate predictions. Because our aim is to obtain coherent results we construct prediction intervals from a multivariate normal distribution with mean zero and covariance matrix given by the residuals generated in the fitting of the three time series in the previous steps.

The multivariate normal distribution of the three-dimensional random vector of residuals $\xi = [\epsilon^{(0)}_{x,t}, \epsilon^{(1)}_{k,x,t}, \epsilon^{(2,3)}_{k,x,t}]$ can be written,

$$\xi \sim \mathcal{N}_3(\mu, \Sigma), \tag{8}$$

with the mean vector,

$$\mu = \left[E(\epsilon^{(0)}_{x,t}) = 0, E(\epsilon^{(1)}_{k,x,t}) = 0, E(\epsilon^{(2,3)}_{k,x,t}) = 0\right],$$

and 3×3 covariance matrix,

$$\Sigma = \left[Cov(\epsilon^{(0)}_{x,t}, \epsilon^{(1)}_{k,x,t}, \epsilon^{(2,3)}_{k,x,t})\right].$$

The time series of errors obtained by fitting the random walk model in the Raftery model, $\epsilon^{(3)}_{k,x,t}$ (see Eq. (6)), is usually of a very short length over the 1950 and 2014 period. This is because in most countries the level of life expectancy at age x is below the determined level A, during the entire period of time. Therefore, in practice the random walk model is used only for forecasting in most of the countries. The assumption adopted in order to keep the model simple is that variance σ of $\epsilon^{(3)}_{k,x,t}$ equals the variance observed in $\epsilon^{(2)}_{k,x,t}$.

The distributions of future values of country-specific life expectancies at age x are estimated by combining simulated future paths of the two gaps and the best-practice level through Monte-Carlo simulation.

4. Accuracy of forecasting prediction

To assess the performance of our model we look at differences between observed and forecast life expectancy and summarize the forecast accuracy. We carry out a back testing exercise in the spirit of Booth et al. (2006), Jarner and Kryger (2011) and Haberman et al. (2014). Four historical periods used for fitting are considered in our data set: 1950–1985,[1] 1950–1990, 1950–1995 and 1950–2000; and using the rest of the years until 2014 as the window of evaluation.

Let $e_{k,x,t}$ denote the observed remaining life expectancy at age x, time t and country k and $\hat{e}_{k,x,t}$ denote the forecast of $e_{k,x,t}$. Then we define the forecast error as follows:

$$\omega_{k,x,t} = e_{k,x,t} - \hat{e}_{k,x,t}. \tag{9}$$

Two measures are considered: mean error (ME) and mean absolute percentage error (MAPE). The mean error is a scale-dependent measure that is useful when comparing different methods applied to the same data set. Calculating the mean error of a forecast is straightforward as it indicates the degree of "optimism" or "pessimism" of the predicted values. However, any scale-dependent measure is sensitive to outliers. Most recommended in the scientific literature is MAPE (Hanke et al., 2001; Bowerman et al., 2004) which is scale-independent and can therefore be used to compare forecast performance across different sets of data.

$$\begin{aligned} ME &= mean\left(\omega_{k,x,t}\right), \\ MAPE &= mean\left(|100 \times \frac{\omega_{k,x,t}}{e_{k,x,t}}|\right), \end{aligned} \tag{10}$$

where the notation $mean(z)$ denotes the sample mean of $\{z\}$ over the period of interest.

[1] Greece, Israel and Slovenia were not evaluated on the 1950–1985 interval because of insufficient data, but were considered in the other three scenarios.

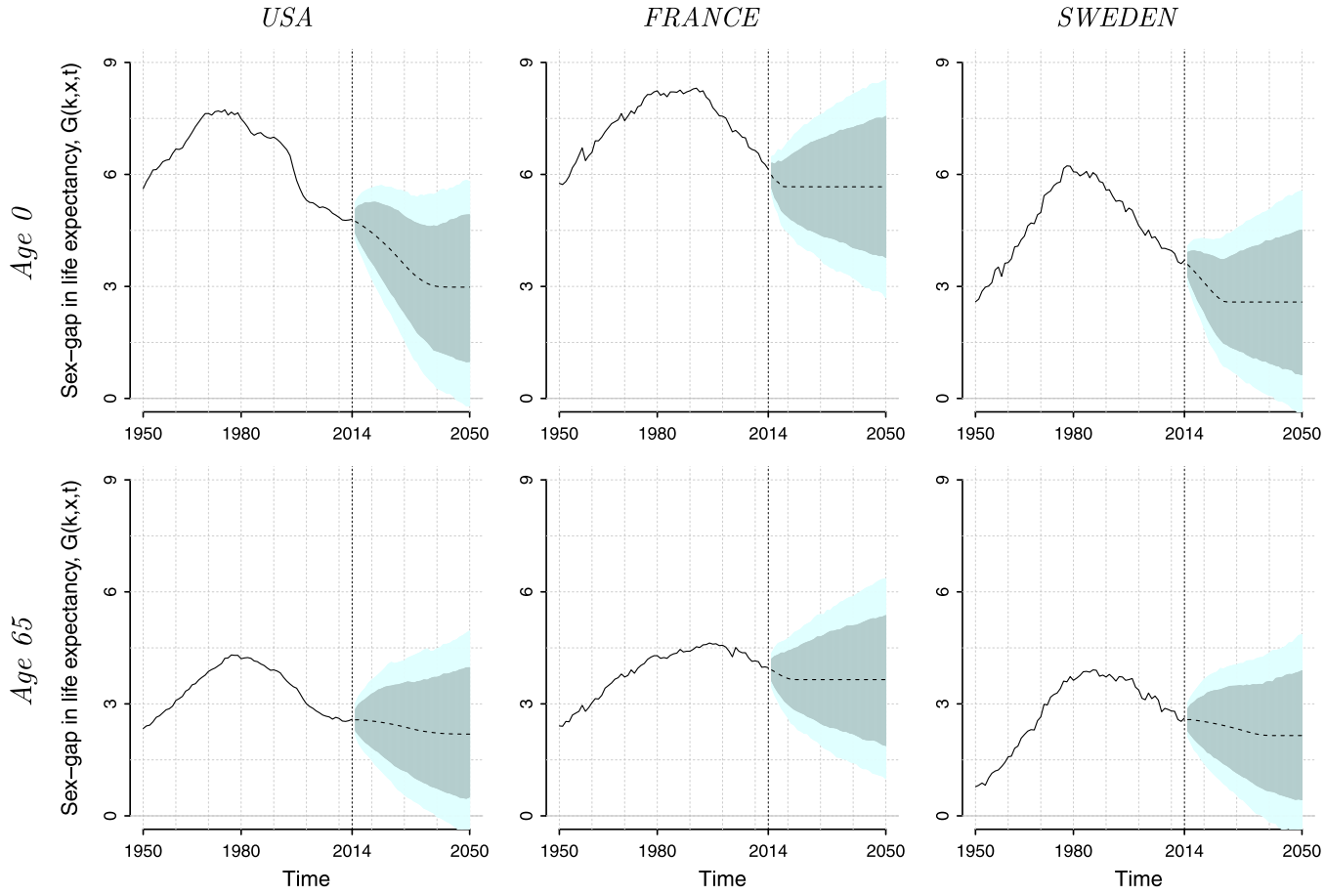

Fig. 3. The forecast gap between female and male life expectancy at birth and at age 65, with associated 80% and 95% prediction intervals, 1950–2050.

5. Results and illustrations

We estimate the distribution values of country specific life expectancies at birth and at age 65 by combining simulated future paths of the gaps and the best-practice trend. The forecast future life expectancies for the three selected countries with different patterns in the two gaps observed in the last 60 years, along with corresponding 80% and 95% prediction intervals, are shown in Table 4.

We compare our results with the values generated by the Lee–Carter model (Lee and Carter, 1992) and the Cairns–Blake–Dowd model (Cairns et al., 2006). The Lee–Carter model (LC) is the first stochastic extrapolative model to be developed and can be used to predict the central mortality rates $m_{x,t}$, for all ages. The Cairns–Blake–Dowd (CBD) is a stochastic model designed for modelling mortality at higher ages and builds on the observation that log death rates are approximately linear at ages above 40. Both approaches are well-established methods in mortality forecasting and can be easily implemented in ***R*** statistical software using the *StMoMo* package (Villegas et al., 2015). Comparison with the CBD model is performed only at age 65. The Lee–Carter model is fitted to ages 0–95 and 65–95, and the CBD is fitted over the 65–95 age range. Both models generate a matrix of forecast death rates. The forecast life expectancies are computed using standard life table calculations.

In order to obtain a complete series of death rates for all the ages up to 110 and to be able to accurately compute the life expectancies the Kannisto old age mortality model is used (Thatcher et al., 1998) which uses a logistic function fitted for death rates at ages above 80. However, if the predicted death rate at the highest age, in our case 95, is sufficiently large ($\geq$ 0.4) a constant force of mortality could be assumed. The difference in life expectancies between the two methods is insignificant.

In 2050, US forecast female life expectancy at birth is 88.93 years and 25.44 years at age 65 according to the Double-Gap model (henceforth DG). The Lee–Carter (LC) model predicts more pessimistic results, namely 85.88 years expectation of life at birth and 23.9 years at age 65. Using the DG we estimate an increase in life expectancy at birth of 7.46 years for females and 9.27 years for males, and an improvement in life expectancy at age 65 of 4.59 years for females and 5 years for males. Therefore, US male life expectancy forecast increases faster in the following 40 years than female life expectancy. In general DG model is more optimistic than the LC model, the forecast results for French, Swedish, and US populations over this horizon of time are higher than the LC forecasts.

The sex-gap forecast given by the DG model is narrower in all the three countries than the predicted values of the LC and CBD model. The DG model has the advantage of modelling the female and male population together taking into account the coherence and correlation between the two, while for LC and CBD separate projections are needed resulting in trajectories with a divergent trend between female and male life expectancy. At age 65 the sex-gaps forecast by the LC and CBD model are similar.

A visual representation of the results already presented in Table 4 is given in Fig. 4 in connection with the historical female record life expectancy and the extension of the best-practice trend. In the long term the DG forecast trajectories of life expectancy follow the trend given by the best-practice line. On the other hand the LC and CBD projected trajectories tend to diverge for all three countries and sexes.

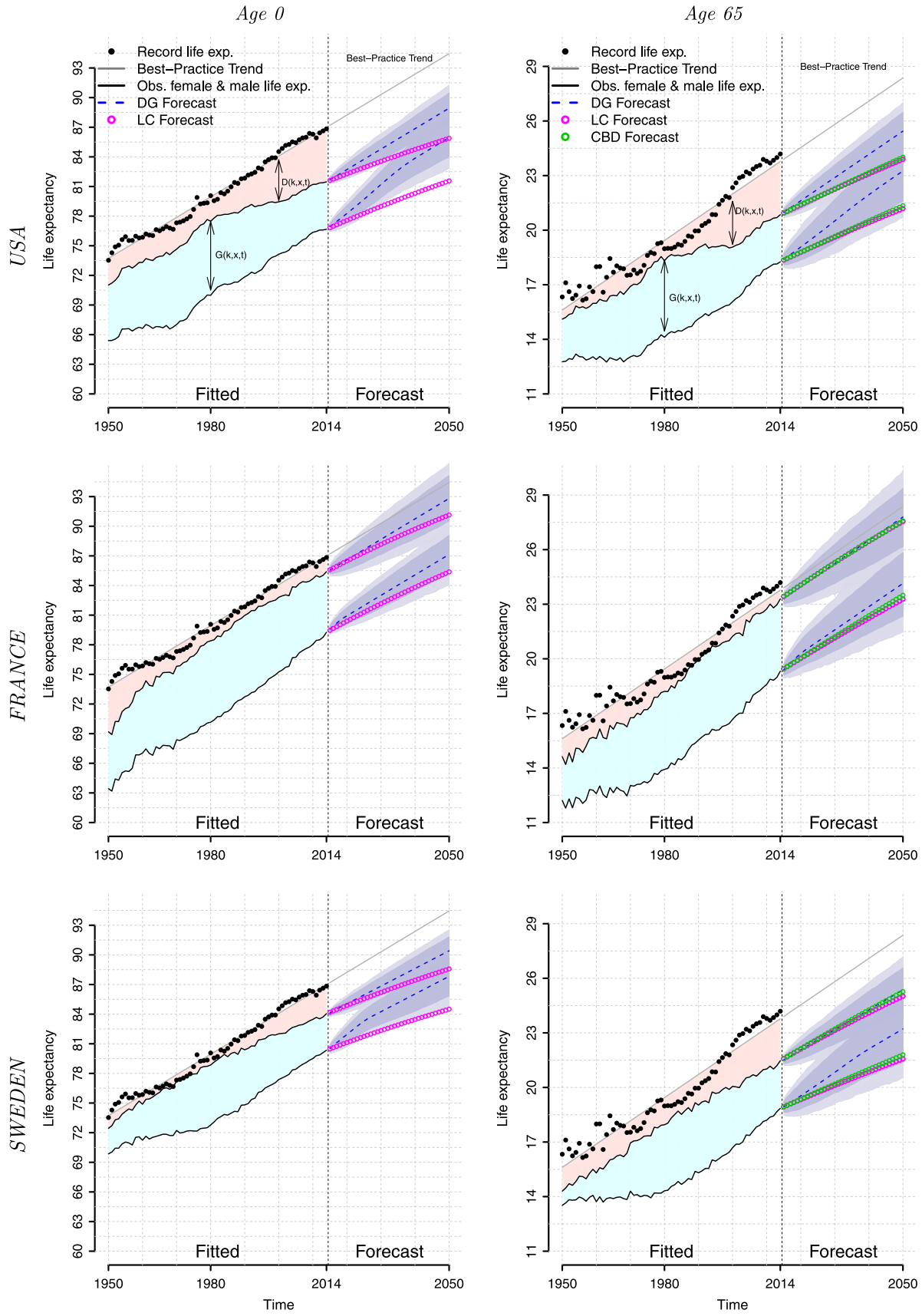

Fig. 4. Actual and forecast life expectancy at birth and at age 65 generated by the DG, LC and CBD models for females and males, 1950–2050. Prediction intervals at 80% and 95% levels are shown only for the DG model.

Table 4
Forecasts of life expectancy in 2050 produced by the Double-Gap (DG) Lee–Carter (LC) and Cairns–Blake–Dowd (CBD) models, with 80% and 95% prediction intervals. The models were evaluated on data from the period 1950–2014.
Source: Authors' calculations based on data described in Table 1

	Model		Age 0			Age 65		
			Females	Males	Sex gap	Females	Males	Sex gap
USA	DG	$\hat{e}_{x,2050}$	88.93	85.94	2.99	25.44	23.26	2.19
		80% PI	(87.41-90.46)	(83.93-87.83)	(1.10-5.00)	(24.36-26.53)	(21.46-24.97)	(0.47-3.98)
		95% PI	(86.64-91.18)	(82.94-88.94)	(0.01-5.99)	(23.81-27.14)	(20.63-25.94)	(−0.49-4.81)
	LC	$\hat{e}_{x,2050}$	85.88	81.57	4.31	23.90	21.19	2.71
		80% PI	(84.72-86.85)	(80.52-82.52)	–	(22.93-24.84)	(20.27-22.07)	–
		95% PI	(84.26-87.27)	(79.97-83.05)	–	(22.45-25.36)	(19.80-22.48)	–
	CBD	$\hat{e}_{x,2050}$	–	–	–	24.01	21.34	2.67
		80% PI	–	–	–	(22.73-25.40)	(20.13-22.58)	–
		95% PI	–	–	–	(22.16-26.12)	(19.55-23.32)	–
		$e_{x,2014}$	81.47	76.67	4.80	20.85	18.26	2.59
France	DG	$\hat{e}_{x,2050}$	92.82	87.15	5.67	27.79	24.14	3.65
		80% PI	(90.60-95.12)	(85.18-89.08)	(3.74-7.64)	(26.18-29.3)	(22.36-25.91)	(1.88-5.43)
		95% PI	(89.43-96.27)	(84.19-90.07)	(2.75-8.63)	(25.38-30.14)	(21.40-26.73)	(1.06-6.39)
	LC	$\hat{e}_{x,2050}$	91.14	85.38	5.76	27.55	23.28	4.27
		80% PI	(89.62-92.8)	(83.78-86.80)	–	(25.75-29.17)	(21.39-24.85)	–
		95% PI	(88.53-93.53)	(83.02-87.46)	–	(24.67-30.00)	(20.23-25.84)	–
	CBD	$\hat{e}_{x,2050}$	–	–	–	27.58	23.49	4.09
		80% PI	–	–	–	(24.77-30.67)	(20.93-26.46)	–
		95% PI	–	–	–	(23.54-32.68)	(19.83-28.26)	–
		$e_{x,2014}$	85.40	79.26	6.14	23.29	19.32	3.97
Sweden	DG	$\hat{e}_{x,2050}$	90.41	87.84	2.57	25.37	23.22	2.15
		80% PI	(89.03-91.79)	(85.84-89.92)	(0.49-4.58)	(24.24-26.50)	(21.48-24.94)	(0.42-3.88)
		95% PI	(88.25-92.51)	(84.81-90.95)	(−0.53-5.61)	(23.63-27.13)	(20.50-25.84)	(−0.48-4.86)
	LC	$\hat{e}_{x,2050}$	88.58	84.50	4.08	25.02	21.57	3.45
		80% PI	(87.37-89.69)	(83.32-85.45)	–	(23.90-26.03)	(20.45-22.57)	–
		95% PI	(86.69-90.15)	(82.69-85.90)	–	(23.13-26.44)	(19.90-23.19)	–
	CBD	$\hat{e}_{x,2050}$	–	–	–	25.27	21.79	3.48
		80% PI	–	–	–	(23.67-27.06)	(20.27-23.54)	–
		95% PI	–	–	–	(22.92-28.23)	(19.58-24.57)	–
		$e_{x,2014}$	84.05	80.35	3.70	21.47	18.85	2.62

Note: The uncertainty in the sex-gap in the case of forecasts generated by the LC and CBD is not available. Sex-specific LC and CBD models are fitted and used to forecast female and male life expectancy.

Prediction intervals given by the DG model indicate that the female French population has the highest probability, among the three countries, of surpassing the best-practice trend and becoming the new world record holder for life expectancy at birth or at age 65.

In-sample forecasts are performed using the DG, LC and CDB models in order to test the performance of the three models. Four forecasting horizons are selected starting with 1985 until 2014. The forecast values are compared with the historical values of life expectancy. Table 5 offers an overall performance of the forecast in the USA, France and Sweden but also over the 38 Human Mortality Database (HMD) countries and regions. DG performs better than both LC or CBD in terms of mean errors (ME) and mean absolute percentage errors (MAPE) when all the countries are considered. However at age 65 the difference between the models is minor especially in the male population.

Table 6 presents an in-depth overview of the accuracy measures for both sexes. DG is consistently less biased than LC for male life expectancy at birth in the three selected countries, but not for females. The CBD model is found to be more accurate than the LC model for age 65 in the male populations. However, there is no model that consistently performs better over all forecasting windows and populations in the study. Some models exhibit a particularly good or bad behaviour for certain historical trends due to the specific constraints of these models. These results show that the DG is capable of generating comparable predictive power with the two most commonly used forecasting models.

More visual results for 18 countries are presented in Figs. 5 and 6 in Appendix.

Table 5
Accuracy measures for the forecast life expectancy at birth and at age 65. Four evaluation periods are considered: 1985–2014, 1990–2014, 1995–2014 and 2000–2014. The results are averaged over the four periods.

Countries	Model	Age 0		Age 65	
		ME	MAPE	ME	MAPE
38 HMD Countries	DG	**−0.198**	**1.728**	**0.632**	**4.745**
	LC	1.099	1.907	0.748	5.294
	CBD	–	–	0.725	5.264
USA, France & Sweden	DG	**−0.285**	**0.619**	**0.414**	**3.433**
	LC	0.540	1.032	0.449	3.611
	CBD	–	–	0.421	3.518

6. Discussion

Our approach to forecast life expectancy combines separate forecasts to obtain joint male and female life expectancies that are coherent with the best-practice trend. The trend proposed in the current article is based on the record level of female life expectancy; this trend was used due to its remarkable linear regularity at age 0. The current model is not restricted to the usage of this particular benchmark, and countries or regions might decide to use a different trend depending on the best performing model for each case based on their past evaluation. In some cases, if the data allow other trends can be adopted, for example a super-population composed from Scandinavian countries if the goal is to forecast the life expectancy in one of these populations. Or the model can be applied to the USA in order to forecast life expectancy in each

Table 6
Accuracy measures for the forecast life expectancy at birth and at age 65, by sex. Four evaluation periods are considered: 1985–2014, 1990–2014, 1995–2014 and 2000–2014. The results are averaged over the four periods.

Country	Model	Female population				Male population			
		Age 0		Age 65		Age 0		Age 65	
		ME	MAPE	ME	MAPE	ME	MAPE	ME	MAPE
38 HMD	DG	**−0.309**	1.400	**0.450**	**3.616**	**−0.088**	**2.056**	**0.814**	**5.874**
	LC	0.510	**1.082**	0.482	3.629	1.689	2.732	1.014	6.960
	CBD	–	–	0.469	3.660	–	–	0.981	6.869
USA	DG	−0.912	1.135	−0.278	**1.955**	**−0.061**	**0.369**	**0.848**	**4.894**
	LC	**−0.414**	**0.666**	**−0.255**	2.394	0.926	1.240	0.904	5.195
	CBD	–	–	−0.272	2.388	–	–	0.871	5.007
France	DG	0.139	0.349	0.664	3.099	**0.112**	**0.509**	0.951	5.344
	LC	**0.031**	**0.304**	**0.305**	**1.640**	1.305	1.692	0.892	4.969
	CBD	–	–	0.314	1.672	–	–	**0.840**	**4.680**
Sweden	DG	−0.688	0.834	−0.340	1.664	**−0.298**	**0.517**	**0.641**	**3.639**
	LC	**−0.196**	**0.276**	**−0.245**	**1.272**	1.586	2.016	1.096	6.193
	CBD	–	–	−0.279	1.420	–	–	1.052	5.944

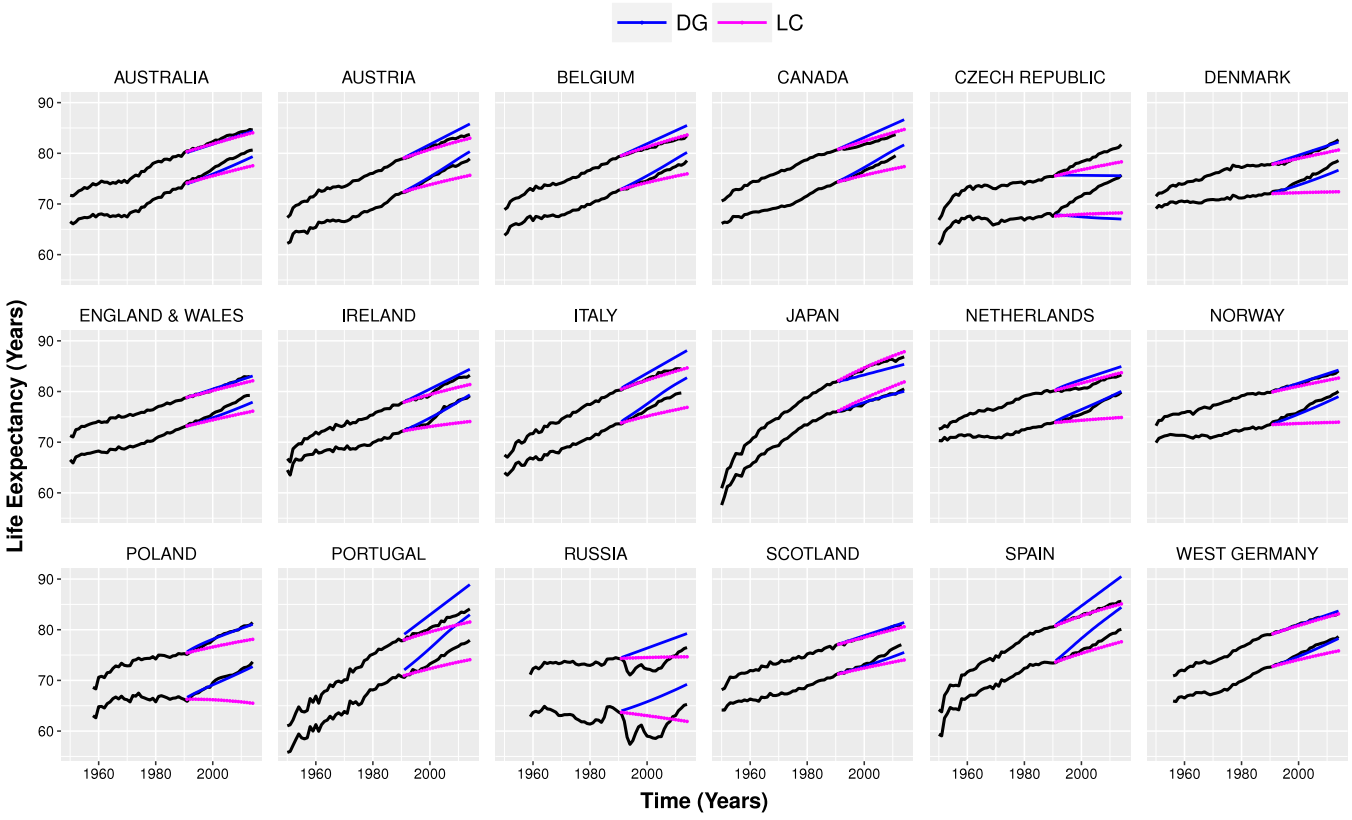

Fig. 5. Comparison of actual life expectancy at birth in 1990–2014 with forecasts generated by the Double-Gap and Lee–Carter models for 18 countries and regions.

American states and jurisdictions with the record US total female population as "best-practice" (Whelpton et al., 1948).

No forecasting model is meant to be used in prediction into an indefinite future. The rate of increase in life expectancy may vary depending on the selected historical period. Therefore, the choice of the historical frame to be fitted is as important as the choice of the model. For example, predicting life expectancy at age 65 based on a trend starting in the 19th century would underestimate the future improvements in human mortality. Also one might ponder the suitableness of the use of a linear trend at age 65. The fluctuations in the relative rate of improvement experienced after age 65 in the last decades (as seen in Fig. 1) suggest the current model can benefit from further research in this direction.

Starting with 1850, not only a rapid improvement in life expectancy has been taking place but also a compression of mortality experience or in other terms a "globalization" of improvements in mortality. After 1950 cross-sectional convergence in life expectancy between different countries is noticeable, with the main contribution being made by countries with a higher level of mortality (Oeppen, 2006). This is because of the increasing "communication" between the countries and continents and a much faster transfer of technology and innovations that help increase life expectancy in all countries. Our proposed method models the gap whether there is convergence or not and even allows countries with a higher level of mortality to become the record holder in terms of longevity at some point in the future.

Life expectancy is an age-aggregated measure but deeper knowledge can be obtained by converting the obtained life expectancy level into age-schedules of death rates and actuarial life tables by exploiting the regularities of age patterns of mortality. In

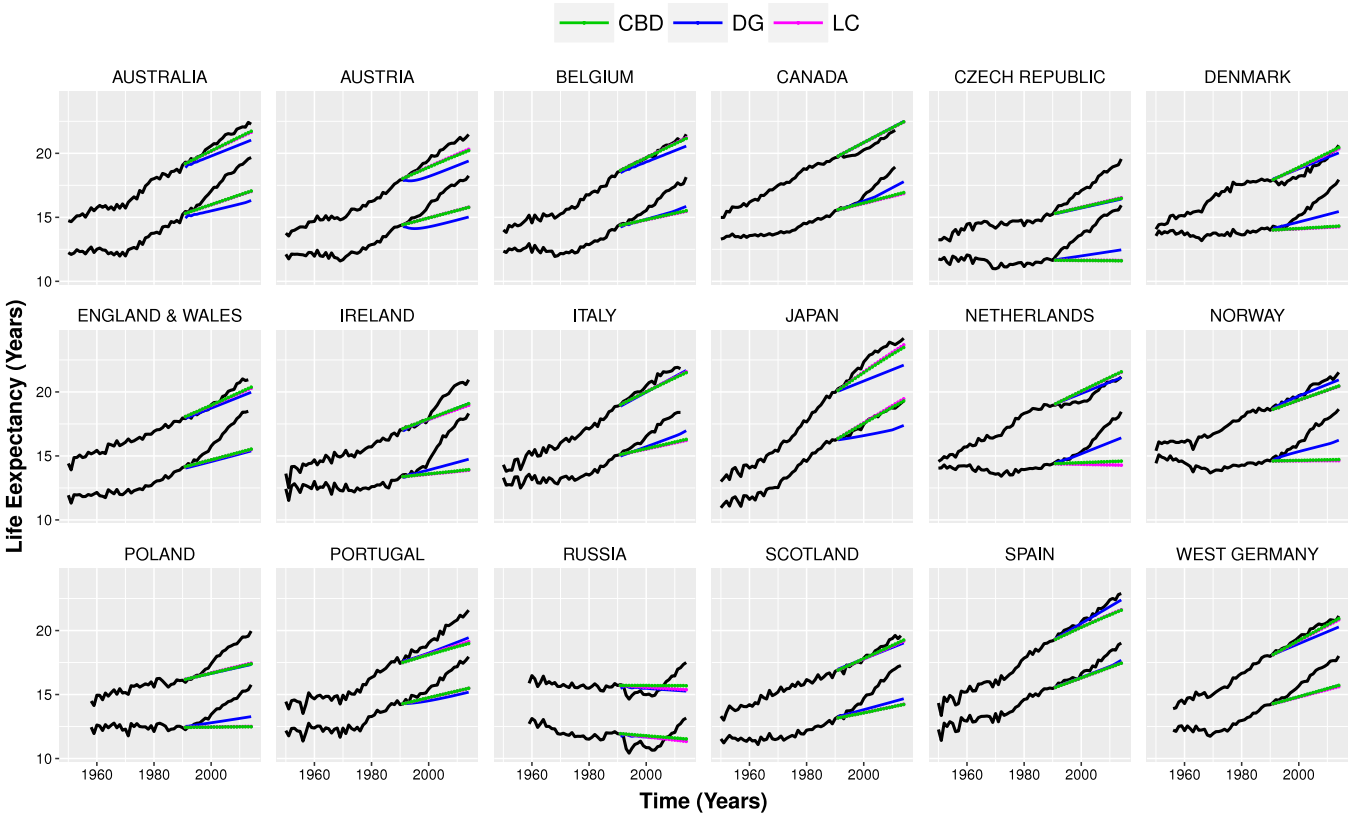

Fig. 6. Comparison of actual life expectancy at age 65 in 1990–2014 with forecasts generated by the Double-Gap, Lee–Carter and CBD models for 18 countries and regions.

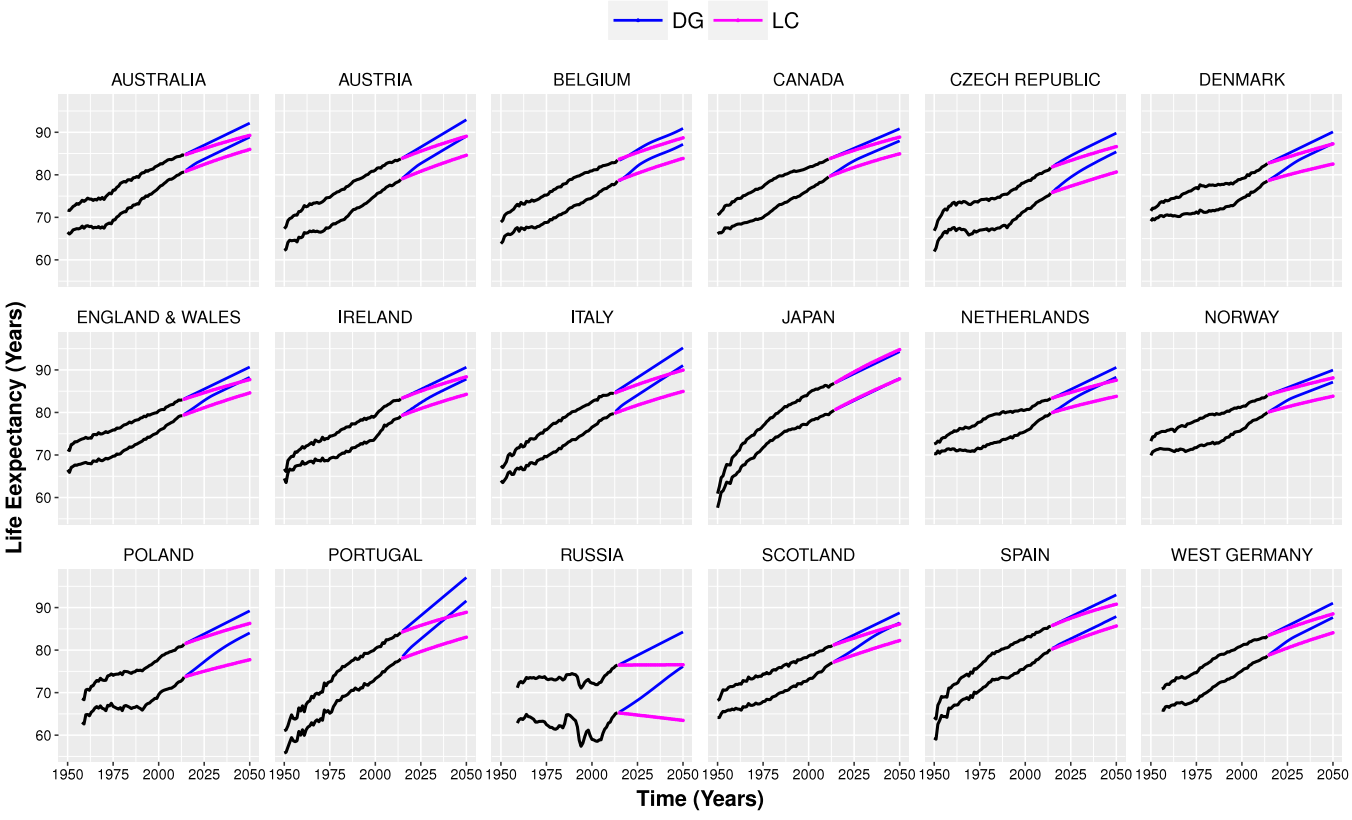

Fig. 7. Forecast life expectancy at birth in 2014–2050 for 18 countries and regions.

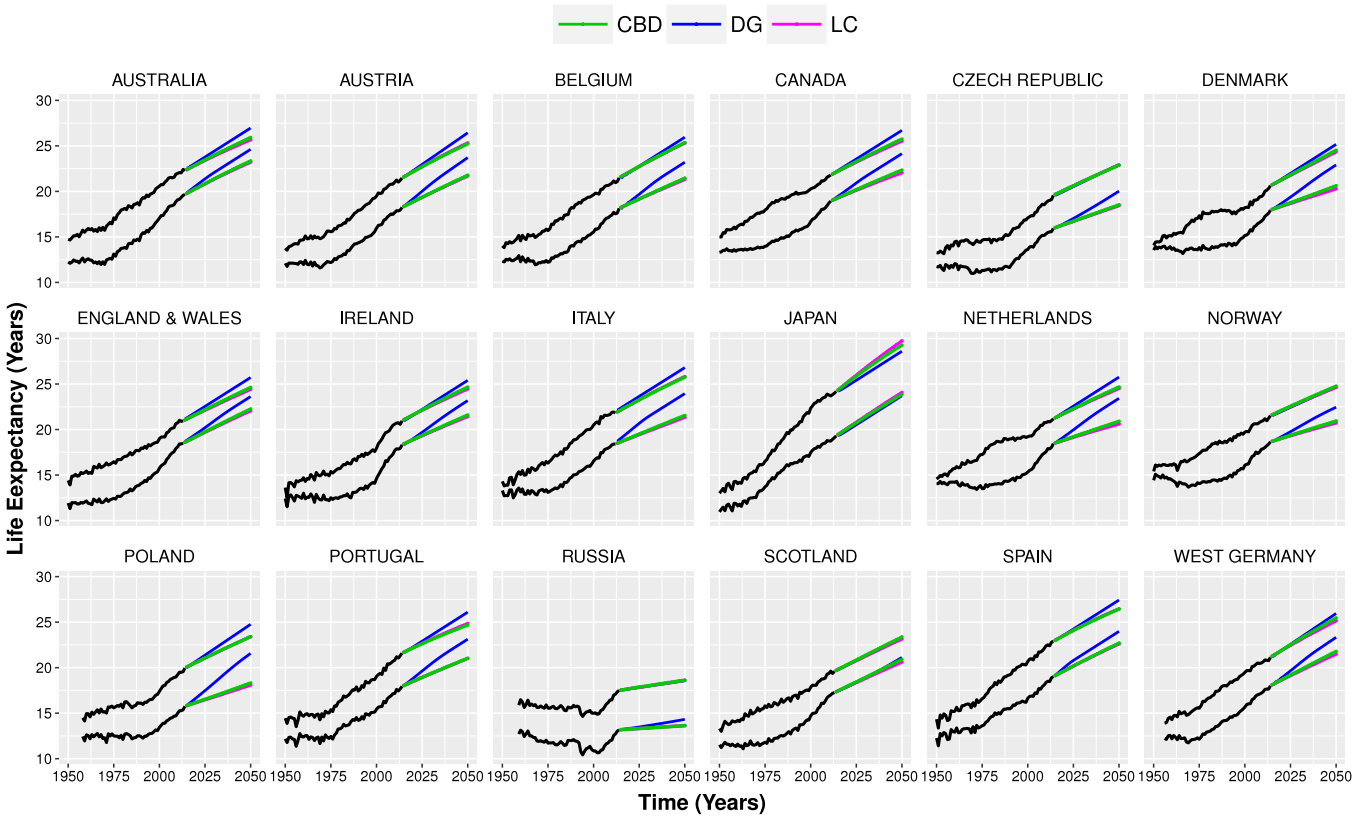

Fig. 8. Forecast life expectancy at age 65 in 2014–2050 for 18 countries and regions.

actuarial science the use of life tables, and other models reflecting life contingencies, is motivated by the need to determine insurance and pension risks, net premiums, and benefits. Although beyond the current project scope, a further step in our research is to transform forecast life expectancy into deaths rates and probabilities using indirect estimation techniques (Brass et al., 1971; Wilmoth et al., 2012) or by reconstruction of the empirical distribution of deaths from its statistical moments following the maximum entropy approach (Mead and Papanicolaou, 1984).

Having simple methods to predict future mortality levels is of high importance because of the growing significance this field is acquiring in society. Justified by the accuracy and simplicity demonstrated in the present article, the Double-Gap model represents an addition to the existing family of forecasting models. Today when so many models exist the researcher should probably not work simply with one model or approach to modelling the future, but with a combination of them. Thus, the Double-Gap model should be considered as a promising available forecasting tool.

Acknowledgements

We thank Heather Booth and researchers at the Max–Planck Center on the Biodemography of Aging for helpful comments. This work was conducted within the "Modelling and forecasting age-specific death at older age" Project – under the management of the University of Southern Denmark, Department of Public Health – with financial support of SCOR Corporate Foundation for Science and the AXA Research Fund. Researchers operated independently from the funders on this work, and the funding organizations had no role in the study design, data collection, analysis, interpretation, writing the report or the decision to submit it for publication.

Appendix

A.1. In-sample forecasts for 18 countries and regions

See Figs. 5 and 6.

A.2. Out-of-sample forecasts for 18 countries and regions

See Figs. 7 and 8.

References

Austad, S.N., 2006. Why women live longer than men: Sex differences in longevity. Gender Med. 3 (2), 79–92.

Booth, H., Hyndman, R., Tickle, L., De Jong, P., 2006. Lee-Carter mortality forecasting: A multi-country comparison of variants and extensions.

Bowerman, B., O'Connell, R., Koehler, A., 2004. Forecasting: Methods and Applications. Thomson Brooks/Cole, Belmont, CA.

Box, G.E., Jenkins, G.M., 1976. Time Series Analysis, Control, and Forecasting, Vol. 3226. Holden Day, San Francisco, CA, p. 10 3228.

Brass, W., et al., 1971. On the scale of mortality. Biol. Asp. Demogr. 69–110.

Cairns, A.J., Blake, D., Dowd, K., 2006. A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. J. Risk Insurance 73 (4), 687–718.

Haberman, S., Kaishev, V., Millosaovich, P., Villegas, A., Baxter, S., Gaches, A., Gunnlaugsson, S., Sison, M., 2014. Longevity basis risk: A methodology for assessing basis risk. Institute and Faculty of Actuaries Sessional Research Paper.

Hanke, J.E., Reitsch, A.G., Wichern, D.W., 2001. Business Forecasting. Prentice Hall, NJ.

Human Mortality Database, 2017. University of California, Berkeley (USA), and Max Planck Institutefor Demographic Research (Germany). Available at www.mortality.orgorwww.humanmortality.de (data downloaded on 06/06/17).

Hyndman, R.J., Booth, H., Yasmeen, F., 2013. Coherent mortality forecasting: the product-ratio method with functional time series models. Demography 50 (1), 261–283.

Jarner, S.F., Kryger, E.M., 2011. Modelling adult mortality in small populations: The SAINT model. Astin Bull. 41 (02), 377–418.

Kwiatkowski, D., Phillips, P.C., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? J. Econometrics 54 (1–3), 159–178.

Lee, R.D., Carter, L.R., 1992. Modeling and forecasting U.S. mortality. J. Amer. Statist. Assoc. 87 (419), 659–671.

Li, N., Lee, R., 2005. Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. Demography 42 (3), 575–594.

Mead, L.R., Papanicolaou, N., 1984. Maximum entropy in the problem of moments. J. Math. Phys. 25 (8), 2404–2417.

Messner, J., Zeileis, A., 2015. Censored Regression with Conditional Heteroscedasticity, Version: 0.9-1. CRAN - Electronic Resource.

Oeppen, J., 2006. Life expectancy convergence among nations since 1820: Separating the effects of technology and income. In: Perspectives on Mortality Forecasting. III. the Linear Rise in Life Expectancy: History and Prospects. In: Social Insurance Studies, pp. 55–82 (3).

Oeppen, J., Vaupel, J.W., 2002. Broken limits to life expectancy. Science 296 (5570), 1029–1031.

Perls, T.T., Fretts, R.C., 1998. Why women live longer than men - what gives women the extra years? Sci. Am. (2), 100–103.

Raftery, A.E., Chunn, J.L., Gerland, P.,Ševčíková, H., 2013. Bayesian probabilistic projections of life expectancy for all countries. Demography 50 (3), 777–801.

Raftery, A.E., Lalic, N., Gerland, P., 2014. Joint probabilistic projection of female and male life expectancy. Demogr. Res. 30, 795.

Thatcher, A.R., Kannisto, V., Vaupel, J.W., 1998. The Force of Mortality at Ages 80 to 120. Odense University Press, Odense, Denmark.

Torri, T., Vaupel, J.W., 2012. Forecasting life expectancy in an international context. Int. J. Forecast. 28 (2), 519–531.

United Nations, 2009. World Population Prospects: The 2008 Revision. New York, NY: United Nations.

Vallin, J., Meslé, F., 2009. The segmented trend line of highest life expectancies. Popul. Dev. Rev. 35 (1), 159–187.

Villegas, A., Millossovich, P., Kaishev, V., 2015. 'StMoMo': An R Package for stochastic mortality modelling.

Whelpton, P.K., Eldridge, H.T., Seigel, J.S., 1948. Forecasts of the Population of the United States. U.S. Government Publishing Office.

White, K.M., 2002. Longevity advances in high-income countries, 1955–96. Popul. Dev. Rev. 28 (1), 59–76.

Wilmoth, J., Zureick, S., Canudas-Romo, V., Inoue, M., Sawyer, C., 2012. A flexible two-dimensional mortality model for use in indirect estimation. Popul. Stud. 66 (1), 1–28.