THESIS COVER – BACK/OUTSIDE

SDU MARIUS D. PASCARIU PH.D. Thesis – MODELLING AND FORECASTING MORTALITY

THESIS COVER – FRONT/OUTSIDE

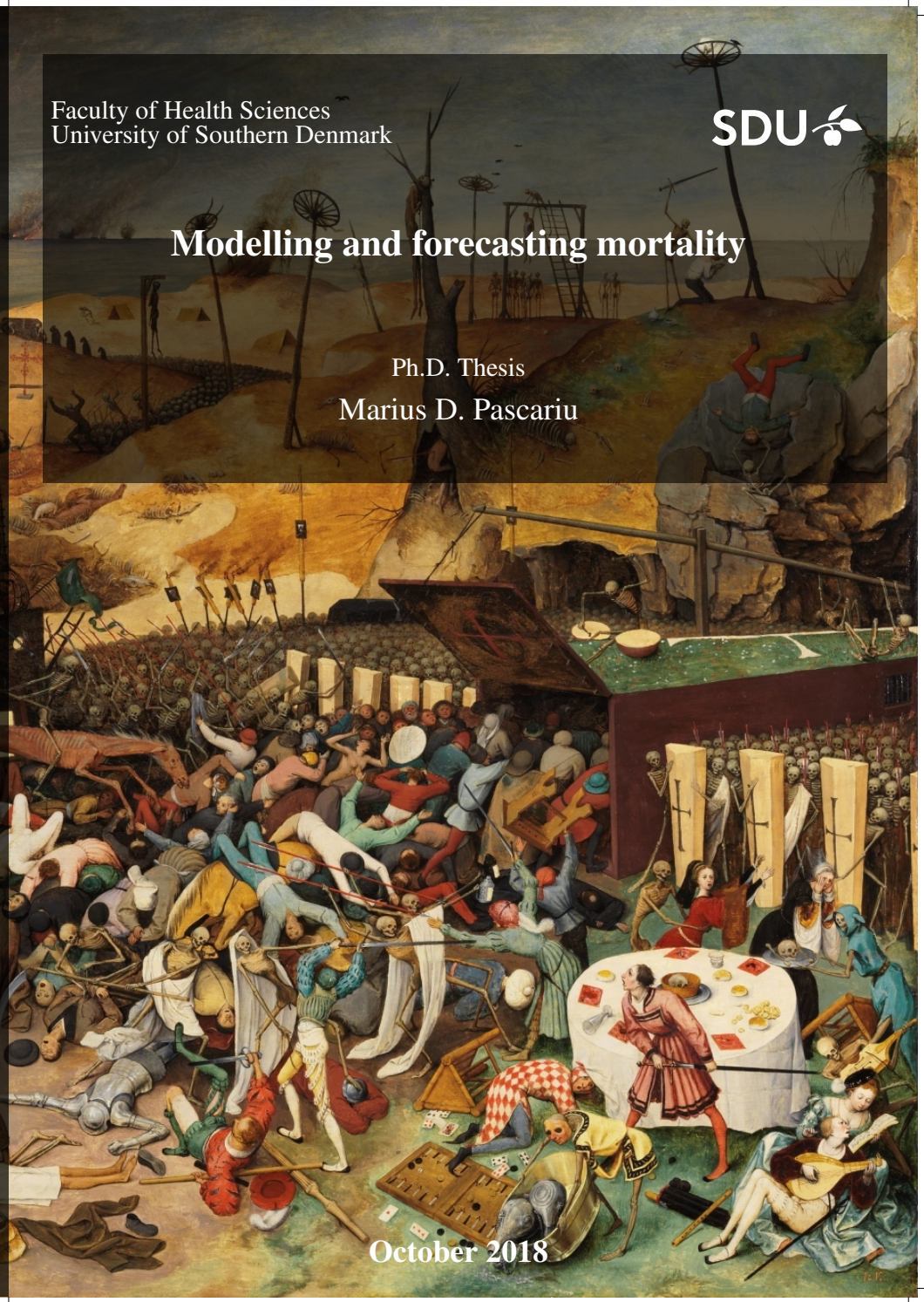
Faculty of Health Sciences
University of Southern Denmark

SDU

Modelling and forecasting mortality

Ph.D. Thesis

Marius D. Pascariu



October 2018

Printed by: Print & Sign, SDU
Odense, Denmark

The Triumph of Death

Bruegel "the Elder" Pieter. 1562 – 1563. Oil on panel.

Museo del Prado. Madrid. Spain.

The painting on the cover depicts a customary theme in medieval literature: the dance of Death. All of the social institutions are included in this composition and neither power nor devotion can save them. It is just a matter of time. Some attempt to struggle against their inevitable destiny while others are resigned to their fate. Only a pair of lovers, at the lower right, seem to remain unmoved by the current course of events. One might wonder: "for how long?"

Modelling and forecasting mortality

Ph.D. Thesis

by Marius D. Pascariu

University of Southern Denmark

Faculty of Health Sciences

Institute of Public Health

Unit of Epidemiology, Biostatistics and Biodemography

Odense, Denmark

October 2018

Academic Advisors

Professor **James W. Vaupel**, Ph.D.
Center on Population Dynamics
Institute of Public Health
University of Southern Denmark
Max-Planck Institute for Demographic Research, Rostock, Germany

Associate Professor **Vladimir Canudas–Romo**, Ph.D.
School of Demography
Australian National University, Canberra, Australia

Assessment Committee

Professor **Heather Booth**, Ph.D.
ANU College of Arts and Social Sciences
School of Demography
Australian National University, Australia

Professor **David Blake**, Ph.D.
Pensions Institute
Cass Business School
City University London, United Kingdom

Professor **Dorte Gyrd–Hansen**, Ph.D. (chair)
Danish Centre for Health Economics
Institute of Public Health
University of Southern Denmark, Denmark

Contents

Preface and acknowledgements

This thesis has been written during my appointment as Ph.D. Research Fellow at the Institute of Public Health within the University of Southern Denmark and during my research stay at the Australian National University, from March 2015 to October 2018. The work was conducted within the *Modelling and forecasting age-specific death at older ages* project [No.95–103–31186], with the financial support of the SCOR Corporate Foundation for Science.

Early in my professional career, while working as a life actuary for insurance companies, I learned that the decisions and risks we will adopt and experience in the future can be measured in present terms. Mortality is one of the risks we face every day, with enormous implications at different levels: for us as individuals, for our families and friends, for our community, and ultimately for society. The ability to decribe the length of life in mathematical terms started to fascinate me to the point that I decided to dedicate my time to explore this topic and to make this the focus of my research. I have had the fortune to cross paths with people who share my enthusiasm and interests. Many of them have had a significant impact in the development of my career. Notably, Adriana Lecu who is probably the first person who showed me what a life table is and she also convinced me to start an actuarial career. Professor James W. Vaupel strongly supported my academic work and ideas. He also convinced me to turn down a flashy job offer from Vienna and move to Odense to work under his supervision. These are decisions I do not regret.

In Denmark I met Professor Vladimir Canudas–Romo. He became my mentor and close advisor and under his constant guidance I was able to raise the quality of my publications. For this I offer my sincere gratitude to him.

I also thank my co-authors: Ugofilippo Basellini, José Manuel Aburto and Adam Lenart for constructive collaborations, critical revision of the manuscripts and for sharing their wisdom with me. Ugo demonstrated an impressive attention to details in all our common projects in such a way that would make you think that the ideas that passed his inspection will stand up to any criticism without losing their value. José was able to provide alterna-

tive solutions to the problems under investigation, thus validating our work. Adam is one of the first researchers I started sharing ideas with since my arrival in Odense. Even now I remember how intimidatingly smart his remarks were during our conversations. If one would like to become an accomplished data scientist he/she should aim at being Adam Lenart.

These years of research would have not been the same without all SDU collegues, who make our work place interesting, motivating and friendly. I must mention Anthony Medford, with whom I did not publish any work in the last four years but our endless discussions on actuarial topics motivated me to find practical applications for my work. I am still hoping for a joint collaboration Anthony! Special thanks to Marie–Pier Bergeron–Boucher, Søren Kjærgaard, Jonas Schöley, Silvia Rizzi, Rune Lindahl–Jacobsen, Jim Oeppen, and of course, Catalina Torres, who has been by my side during this process. Thank you for your kindness, patience, support and enthusiasm Catalina! Thanks to the entire SDU group.

Finally, many thanks to my family and friends in Romania. To my parents Maria and Victor Pascariu for the unconditional love, support and understanding. To my siblings Florin Pascariu and Silvia–Simona Coman for great moments together and the constant encouragement you gave me.

English summary

For the world as a whole, life expectancy has more than doubled over the past two centuries. This transformation of the duration of life has greatly enhanced the quantity and quality of people's lives. It has fuelled enormous increase in economic output and in population size, including an upsurge in the number of elderly. Understanding human mortality dynamics is of utmost importance in the context of rapid ageing and increasing length of life experienced by most populations nowadays. The present thesis highlights new and innovative methods for estimating and projecting future mortality levels among humans.

Three studies have been devised, which develop and analyse relevant statistical models for addressing uncertainty in future mortality. The studies are in the form of research manuscripts that are/will be published in scientific journals together with software packages ensuring the reproducibility of the results. In the first study, a method for forecasting life expectancy for females and males is developed. To forecast female life expectancy, the method is based on the analysis of the gap in life expectancy between females in a given country and females in record-holding countries. To forecast male life expectancy, the gap between male life expectancy and female life expectancy in a given country is analysed. In the second study, we explore a new approach inspired by indirect estimation techniques applied in demography, which can be used to estimate full life tables at any point in time, based on a given value of life expectancy at birth or at any other age. The third study makes use of the statistical properties of a probability density function in order to estimate the distribution of deaths of a population in the future. We employ time series methods for forecasting a limited number of central statistical moments and then reconstruct the future distribution of deaths using the predicted moments. The estimation of the density function is done using the maximum entropy approach.

The results show that mortality modelling can be tackled from different perspectives and higher accuracy of the future trajectories can be obtained when compared with the more traditional extrapolative methods based on age specific death rates or probabilities.

Danish summary

I løbet af de sidste to hundrede år har vi oplevet mere end en fordobling i den forventede levealder. Denne transformation af levealderen har haft stor betydning for menneskets livskvantitet og -kvalitet. Det har bidraget til en enorm økonomisk vækst, men også i forhold til befolkningsstørrelsen, som nu oplever en støt vækst i befolkningen af ældre. Det er yderst vigtigt, at vi forstår dynamikken bag menneskets dødelighed for at forstå den hurtige vækst i forventet levealder, som de fleste befolkningsgrupper oplever i dag. Denne afhandling sætter fokus på nye og innovative metoder til at forudsige den fremtidige levealder hos mennesker.

Der er gennemført tre studier, som udvikler og analyserer relevante statistiske modeller, der imødekommer usikkerhed vedr. det fremtidige niveau af dødelighed. Studierne er gennemført som forskningsmanuskripter, som enten er, eller vil blive, offentliggjort i videnskabelige tidsskrifter sammen med software, som sikrer gengivelse af de opnåede resultater. I det første studie udvikles en metode til at fremskrive den forventede levealder for mænd og kvinder. For at fremskrive kvinders forventede levealder baseres metoden på en analyse af afstanden imellem mænd og kvinders forventede levealder i et givent land. I det andet studie udforskes en ny tilgang, som er inspireret af indirekte estimationsteknikker anvendt indenfor demografi. Denne tilgang kan bruges til at estimere den komplette overlevelsestavle på et hvilket som helst tidspunkt, baseret på den forventede levealder ved fødsel eller ved en hvilken som helst alder. Det tredje studie gør brug af de statistiske egenskaber ved en tæthedsfunktion for at kunne estimere en befolkningsgruppes fremtidige dødsfaldsfordeling. Vi anvender tidsseriemetoder til at fremskrive et begrænset antal centrale statistiske momenter for derefter at rekonstruere den fremtidige fordeling af dødsfald ud fra de fremskrevne momenter. Estimering af tæthedsfunktionen baseres på maximal entropi metode.

Resultaterne viser, at modellering af dødelighed kan håndteres fra forskellige udgangspunkter og at der opnås større nøjagtighed for fremtidige dødelighedsforløb sammenholdt med mere traditionelle extrapolerende metoder, baseret på aldersspecifikke dødelighedsrater eller sandsynlighed for dødsfald.

Papers in the Thesis

Manuscripts included in this dissertation

Paper I

Pascariu M.D., Canudas–Romo V. and Vaupel J.W. (2018). The double-gap life expectancy forecasting model. *Insurance: Mathematics and Economics*; 78, 339–350. DOI: 10.1016/j.insmatheco.2017.09.011

Paper II

Pascariu M.D., Basellini U., Aburto J.M. and Canudas-Romo V. (2018). The Linear Link: Deriving Age-Specific Death Rates from Life Expectancy.

Paper III

Pascariu M.D., Lenart A. and Canudas–Romo V. (2018). Forecasting mortality using statistical moments.

Published Software 1: R Package

Pascariu M.D. (2017). MortalityLaws: Parametric Mortality Models, Life Tables and HMD. *The Comprehensive R Archive Network (CRAN)*. URL: `https://cran.r-project.org/web/packages/MortalityLaws`

Other co-authored works during the PhD, not included in the dissertation:

Paper IV

Pascariu M.D., Dańko M.J., Schöley J. and Rizzi S. (2018). ungroup: An R package for efficient estimation of smooth distributions from coarsely binned data. Journal of Open Source Software, 3(29), 937. DOI: https://doi.org/10.21105/joss.00937

Paper V

Bergeron Boucher M–P., Canudas–Romo V., Pascariu M.D. and Lindahl–Jacobsen R. (Forthcoming 2018). Modelling and forecasting sex differences in mortality: A sex–ratio approach. Genus: Journal of Population Sciences.

Published Software 2: R Package

Pascariu M.D. (2018). MortalityGaps: The Double–Gap Life Expectancy Forecasting Model. *The Comprehensive R Archive Network (CRAN)*. URL: https://cran.r-project.org/web/packages/MortalityGaps

Published Software 3: R Package

Schöley J., Pascariu M.D., Villavicencio F. and Dańko M.J. (2017). pash: Pace–Shape Analysis of Life–tables. *GitHub*. URL: https://github.com/jschoeley/pash

Abbreviations

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
ARIMA	Autoregressive Moving Average Model
CBD	Cairns–Blake–Dowd Model
DG	Double–Gap Life Expectancy Forecasting Model
LC	Lee–Carter Model
LL	Linear–Link Model
H	Shannon Entropy
HMD	Human Mortality Database
KPSS	Kwiatkowski–Phillips–Schmidt–Shin unit-root test
MaxEnt	Maximum–Entropy Method for Density Estimation
MEM	Maximum–Entropy Mortality Model
ME	Mean Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
sMAPE	Symmetric Mean Absolute Percentage Error
sMRAE	Symmetric Mean Relative Absolute Error
MASE	Mean Absolute Scaled Error
OLS	The ordinary least squares method for estimating the unknown parameters
SVD	The singular-value decomposition method

Mathematical notation

Mortality Laws

μ_x	Force of mortality at age x
m_x	Central death rate at age x
q_x	Probability of dying at age x
p_x	Surviving probability at age x
e_x	Remaining life expectancy at age x

Double–Gap Model

$D_{k,x,t}$	The gap between female life expectancy and the best–practice trend in the world at age x for country k and time t
$G_{k,x,t}$	The gap between male and female life expectancy
$e^f_{k,x,t}$	Female remaining life expectancy in country k, age x and time t
$e^m_{k,x,t}$	Male remaining life expectancy in country k, age x and time t
$e^{bp}_{x,t}$	Best-practice female life expectancy at age x and time t
$\mu_{k,x}, \phi, \theta, \beta$	parameters of the DG model
ϵ	residuals i.e. independent and identically distributed random variables normally with mean zero and variance σ

Linear–Link Model

$m_{x,t}$	Central death-rate at age x and time t
$e_{\theta,t}$	Life expectancy at age θ and time t
β, ν, k	Estimate parameters of the LL model
ε	Deviance residuals, independent and identically distributed random variables normally with mean zero and variance σ

Maximum Entropy Mortality Model

μ_n	The n–th statistical moment of the distribution of deaths
$f(x)$	Density function
$f_N(x)$	Approximations for $f(x)$ based on the first $N+1$ statistical moments
H	Shannon Entropy
$\mathcal{L}$	Lagrangian function
λ	Lagrange multipliers

Introduction

> *We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes.*
> – Pierre Simon Laplace (1825) – A Philosophical Essay on Probabilities

Longevity risk, defined as the risk that people live longer than expected, represents an important issue for current societies. Although longevity advancements increase the productive life span and welfare of millions of individuals, there are also increasing costs for pay-as–you-go (PAYG) and defined benefits pension systems, threatening the long-term solvency of financial institutions due to increases in unanticipated future liabilities. Additionally, if unhealthy life expectancy is extended as a result of improvements in mortality rates at advanced ages, public health expenditures are affected. To deal with these rapid changes and avoid negative consequence, academics and professionals alike have acknowledged the significance and necessity of accurately assessing longevity risk. Considerable effort has been dedicated to finding and developing better methods and mathematical models to predict the future. These are mainly statistical and epidemiological methods based on the observation of past trends and on the identification of determinants of the decline in physiological capacities with age. For instance, Pollard (1987) identified a variety of predictive models of age-specific death rates based on: projection by extrapolation of transformation of death-rates and mortality probabilities, projection by cause of death, projection by reference to model life tables, projection by reference to a *law of mortality*, projection by reference to another population, and combinations of these methods.

If the performance of the proposed methods on the short-term can be subject to debate, in the long run Oeppen and Vaupel (2002) showed that demographers and actuaries consistently underestimated the increase in life expectancy by embracing unrealistic assumptions or often ignoring the long term trends.

In order to identify, elucidate and quantify future longevity, one must employ forecasting models that adequately capture the effects of all the moving parts of the mechanism driving the changes in life expectancy. Even if this is, like Laplace's vision, an unrealistic and impossible endeavour, tackling the stochasticity of the evolution of mortality from multiple perspectives is what validates our predictions.

1.1 History of Mortality Modelling

Modelling human mortality has been an important and active area of research for demographers, insurance mathematicians and medical scholars since Graunt (1662) first examined mortality in London to produce the first publication that was concerned mostly with public health statistics. Graunt's work showed that, while individual life-length was uncertain, there was a more predictable pattern of mortality in groups and causes of death. Halley (1693) showed how to actually construct a non-deficient mortality table from empirical birth-death data and even succeeded in presenting a method to perform a life annuity calculation based on this table. Such early tables were empirical and calculation was time consuming. Theoretical mortality modelling first began with DeMoivre (1725), who postulated a uniform distribution of deaths model, and showed simplified annuity calculation methods. Taking a biological approach to mathematical modelling, Gompertz (1825) assumed that the force of mortality μ_x in adulthood shows a nearly exponential increase, where the two parameters of his model are positive and vary with the level of mortality and the rate of increase in mortality with age. The Gompertz model and its modified version by Makeham (1867), where an additional constant is added to take into account the background mortality due to causes unrelated to age, were widely used as the standard models for adult mortality in humans (Kirkwood, 2015; Olshansky and Carnes, 1997); and then extended further to animal species in general (Sacher, 1977).

Figure 1.1: *Mortality modelling timeline*

Early discovery of important notions in the field of mortality modelling were made by Danish scientists like Oppermann (1870) and Thiele (1871), with exposure to actuarial science and active in insurance companies at that time. However, the recognition of their work came only many years later because of a combination of factors, ranging from the publication of the new ideas in inaccessible places and an uncommon language like Danish, to the lack of interdisciplinary collaboration (Hoem, 1983).

After one century of developments, the structure of the mathematical models became increasingly complex and capable of accurately capturing all the spectrum of the mortality intensity experienced by humans. For example Heligman and Pollard (1980) proposed an eight-component mortality model that fits the entire age range,

$$\mu_x = A^{(x+B)^C} + De^{-E(\log x - \log F)^2} + GH^x. \tag{1.1}$$

Figure 1.2 shows how this model can fit the entire age-range by decomposing the age-pattern of mortality into three pieces, each part with a relatively small number of parameters to control it. There are three parameters (A, B and C) to describe child mortality, three to describe a very flexible accident hump (D, E and F) typically occurring in young adulthood, and finally two parameters (G and H) to describe mortality at older ages. The

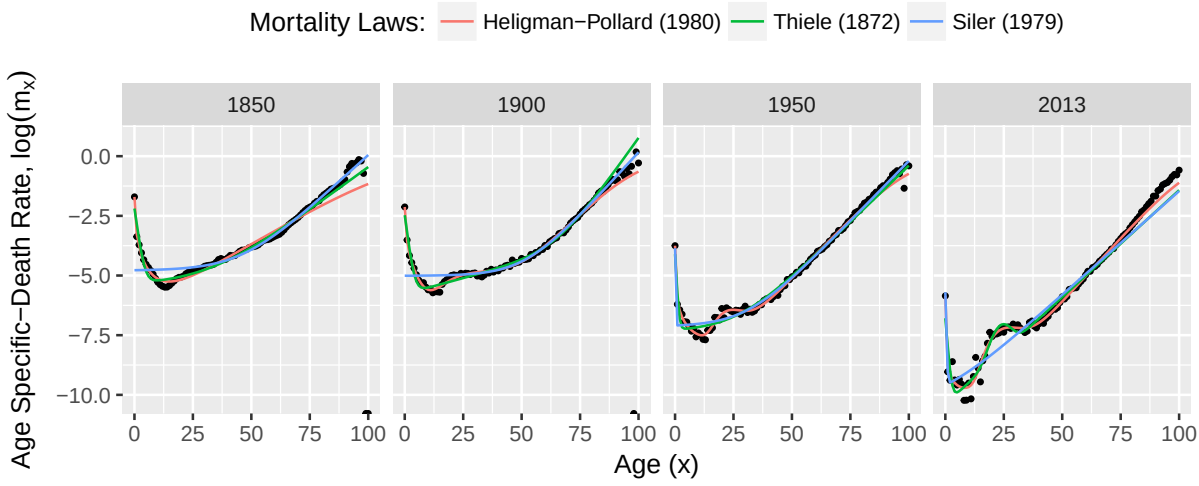


Figure 1.2: *Observed and fitted death rates between age 0 and 80 for male population in Sweden. The mortality is extrapolated up to age 100.*

main disadvantage of this model is that in its traditional form is difficult to fit and it does not account for uncertainty.

Siler (1983) developed a five-parameter competing-hazard model in order to capture mortality during "immaturity", adulthood and senescence and to facilitate inter-specific comparison, by assigning location and dispersion parameters respectively for the three important sections of the mortality curve. Thatcher et al. (1998) performed studies to fit different mathematical models to different reliable data sets on adult and oldest-old mortality (aged 80 and above) covering the few recent decades. They evaluated the comparative compatibility of those models to the data, established the logistic model as the best mathematical model of human adult mortality, replacing the widely used Gompertz model and Makeham model. The logistic model assumes that the force of mortality μ_x is a logistic function of age x.

Author	**Publication**	**Model**
De Moivre	1725	$\mu(x) = 1/(\omega - x)$
Gompertz	1825	$\mu(x) = Ae^{Bx}$
Gompertz	–	$\mu(x) = \frac{1}{\sigma} exp\left\{\frac{x-M}{\sigma}\right\}$
Inverse–Gompertz	–	$\mu(x) = \frac{1}{\sigma} exp\left\{\frac{x-M}{\sigma}\right\} / \left(exp\left\{e^{\frac{-(x-M)}{\sigma}}\right\} - 1\right)$
Makeham	1867	$\mu(x) = Ae^{Bx} + C$

Makeham	–	$\mu(x) = \frac{1}{\sigma} exp\left\{\frac{x-M}{\sigma}\right\} + C$
Oppermann	$\leq$ 1870	$\mu(x) = Ae^{Bx} + (C + Dx)e^{-Ex}$
Oppermann	1870	$\mu(x) = Ax^{-\frac{1}{2}} + B + Cx^{\frac{1}{2}}$
Thiele	1871	$\mu(x) = A_1 e^{-B_1 x} + A_2 e^{-\frac{1}{2}B_2(x-C)^2} + A_3 e^{B_3 x}$
Wittstein & Bumstead	1883	$\mathrm{q(x)} = \frac{1}{B} A^{-(Bx)^N} + A^{-(M-x)^N}$
Steffenson	1930	$\log_{10} s(x) = 10^{-A\sqrt{x}-B} + C$
Perks	1932	$\mu(x) = (A + BC^x)/(BC^{-x} + 1 + DC^x)$
Harper	1936	$log_{10} s(x) = A + 10^{B\sqrt{x}+Cx+D}$
Weibull	1951	$\mu(x) = \frac{1}{\sigma}\left(\frac{x}{M}\right)^{\frac{M}{\sigma}-1}$
Inverse–Weibull	–	$\mu(x) = \frac{1}{\sigma}\left(\frac{x}{M}\right)^{-\frac{M}{\sigma}-1} / \left(exp\left\{\left(\frac{x}{M}\right)^{-\frac{M}{\sigma}}\right\} - 1\right)$
Van der Maen	1943	$\mu(x) = A + Bx + Cx^2 + I/(N - x)$
Van der Maen	1943	$\mu(x) = A + Bx + I/(N - x)$
Quadratic	–	$\mu(x) = A + Bx + Cx^2$
Beard	1971	$\mu(x) = KAe^{Bx}/(1 + Ae^{Bx})$
Beard–Makeham	1971	$\mu(x) = KAe^{Bx}/(1 + Ae^{Bx}) + C$
Gamma–Gompertz	1979	$\mu(x) = Ae^{Bx}/(1 + \frac{AG}{B}(e^{Bx} - 1))$
Siler	1983	$\mu(x) = A_1 e^{-B_1 x} + A_2 + A_3 e^{B_3 x}$
Heligman–Pollard	1980	$q(x)/p(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + GH^x$
Heligman–Pollard	1980	$q(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1+GH^x}$
Heligman–Pollard	1980	$q(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1+KGH^x}$
Heligman–Pollard	1980	$q(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^{x^K}}{1+GH^{x^K}}$
Rogers–Planck	1984	$q(x) = A_0 + A_1 e^{-Ax} + A_2 e^{\left\{B(x-U) - e^{-C(x-U)}\right\}} + A_3 e^{Dx}$
Martinelle	1987	$\mu(x) = (Ae^{Bx} + C)/(1 + De^{Bx}) + Ke^{Bx}$
Carriere	1992	$S(x) = \psi_1 S_1(x) + \psi_2 S_2(x) + \psi_3 S_3(x)$
Carriere	1992	$S(x) = \psi_1 S_1(x) + \psi_4 S_4(x) + \psi_3 S_3(x)$
Kostaki	1992	$q(x)/p(x) = A^{(x+B)^C} + De^{-E_i(\ln x - \ln F)^2} + GH^x$
Kannisto	1998	$\mu(x) = Ae^{Bx}/(1 + Ae^{Bx})$
Kannisto–Makeham	–	$\mu(x) = Ae^{Bx}/(1 + Ae^{Bx}) + C$

Table 1.1: *Main parametrization functions for human mortality*

The above models describe mortality at a fixed point in time; however, actual mortality is stochastic and evolving continuously. Thus, while the mortality models described above are static, the parameters must be fitted periodically to accommodate changes in mortality patterns.

1.2 Mortality Forecasting

Since Malthus (1798), demographic forecasting became more prominent, and the desire to anticipate future changes in the composition and structure of populations lead to the development of cohort-component forecasting methods. The first to use these techniques was Cannan (1895), who prepared a cohort component forecast for England and Wales. By the end of the 1920's, such forecasts had also been made for the Soviet Union by Tarasov in 1922 (De Gans, 1999), for the Netherlands by Wiebols (1925), for Sweden by Wicksell (1926), and for the United States by Whelpton (1928).

At the beginning of the twentieth century, when demographers were developing the cohort-component forecasting system, statisticians established the foundations of the so called stationary processes. This theory was based on a linear transformation of white noise. Although the main features of the theory were essentially perfected by the beginning of the 1950's (Doob and Doob, 1953), their practical application in statistics did not become standard until the publication of Box and Jenkins (1970). Early examples of their use in demography include Saboia (1974, 1977). The popularity of these methods is related to their great flexibility, and to the fact that they allow for the incorporation of effects of changes in behavioural and socio-economic variables in forecasts and permit the construction of confidence intervals (Tabeau et al., 2001).

Until the 1980s, the mathematical models used in directly forecasting mortality rates or life expectancy were relatively simple and involved a fair degree of subjective judgement. Over the past thirty years, a number of new approaches have been developed for forecasting

mortality using stochastic models such as Alho (1990); Alho and Spencer (1991); McNown and Rogers (1989, 1992); Bell and Monsell (1991), and Lee and Carter (1992, henceforth LC). The main advantage of stochastic models is that the output is not a single figure but a distribution. LC proposes a log-bilinear model for mortality rates incorporating both age and year effects:

$$\ln m(x,t) = a(x) + b(x)k(t) + \varepsilon(x,t) \tag{1.2}$$

where, $m(x,t)$ is the observed central death rate at age x in year t, $a(x)$ represents the average age-specific pattern of mortality, $b(x)$ is a pattern of deviations from the age of profile as the mortality index $k(t)$ varies, and finally $\varepsilon(x,t)$ denotes the residual term at age x and time t.

Lee and Carter's work has been widely cited and it is also used for long-run forecasts of age-specific mortality rates by the U.S. Bureau of the Census as a benchmark model, see Hollmann et al. (1999). Moreover, since that paper, most other models attempting to assess both time and age evolution of mortality have started with the LC framework. Numerous papers since Lee and Carter (1992) have tried to improve upon their model by adding more principal components, or a cohort effect, or any range of similar statistical quantities. Booth et al. (2006) modify the LC model by optimally choosing the time period over which to fit the model and adjust the index of mortality, $k(t)$, to fit the total number of deaths in each year. De Jong and Tickle (2006) reduced the number of parameters in LC to model mortality rates as a smoothed state space model. Yang et al. (2010) take a further step and use multiple principal components to expand the LC model. Chen and Cox (2009) introduce jumps into modelling the state variable, found in Lee and Carter (1992), to increase goodness of fit measures and price insurance linked securities. Deng et al. (2012) use a more advanced jump diffusion model to fit the temporal state variable and Li et al. (2011) identify non-linearities in the temporal state variable. A cohort effect, which incorporates the year of birth into the model, is added to the LC model in Renshaw and Haberman (2006). And more recently, Mitchell et al. (2013) used a methodology similar to LC to model the changes in log mortality rates rather than levels of log mortality rates.

On the other hand, a totally different approach to mortality modelling emerged at the beginning of 2002 when two important articles were published: Oeppen and Vaupel (2002) and White (2002). Oeppen and Vaupel showed that life expectancy at birth in the record-holding countries has increased linearly since 1840. White (2002) also found a linear trend in sexes combined life expectancy in 21 industrial nations from 1955 and 1995. Figure 1.3 shows some details of the probable trajectories of limits and convergence for average life expectancy over the past four centuries. The vertical bars show the inter-quartile range of life expectancy for countries containing half the World's population. These emphasize three massive changes: rapid improvement in life expectancy, greater symmetry in the distribution, and the globalization of mortality experience (Oeppen, 2006).

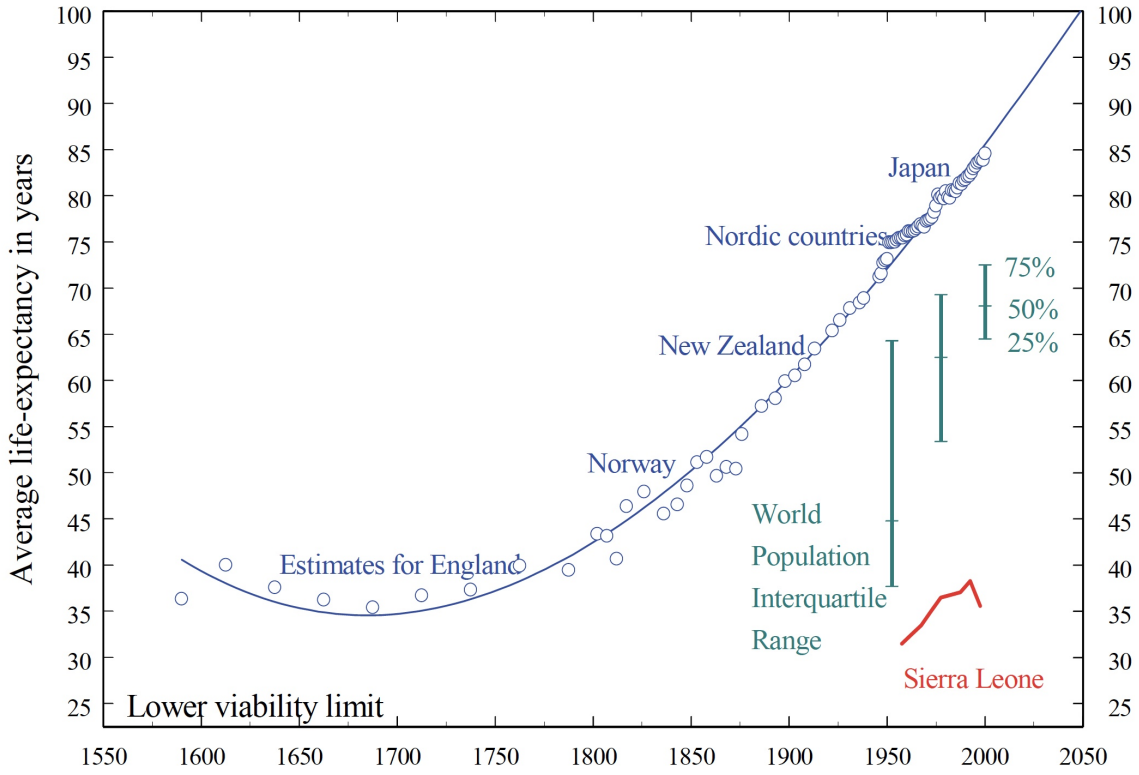


Figure 1.3: *Limits and convergence for national average female life expectancy at birth*
Source: Oeppen (2006)

According to Oeppen and Vaupel (2002), this linear rise may be the most remarkable accomplishment of mass human endeavour ever achieved. Yet little research has been done on why the revolution in human longevity started about 200 hundred years ago, why the revolution started in Scandinavia, and why progress in increasing record life expectancy

has been so steady. Both findings, Oeppen–Vaupel and White, challenge the view that risks of death are fundamental to understanding the future trends of mortality, in a world dominated by LC type of forecasting models where the main objective is to project the age-specific death rates and afterwards to derive the other life table measures.

Torri and Vaupel (2012) started to build on this methodology and published a model that at first forecasts the world's record life expectancy and then the gap between the record and the current life expectancy of a particular country/population of interest assuming a convergence to the forecast record level. Also, the United Nations and US Census Bureau produces many of its current deterministic projections by extrapolating broad summaries of population processes, and then breaking them down into age-specific rates using model schedules and relational models, to yield the age- and sex-specific fertility, mortality, and migration rates that are required by the standard cohort component population projection method. And more recently Raftery et al. (2013) propose an approach to forecasting life expectancy directly using a random walk model with a non-constant drift and a Bayesian hierarchical model, which allows the estimatimation of different rates of improvement in life expectancy for each country.

1.3 Aims

The overall rationale behind this thesis is to provide alternative solutions for demographers and actuaries to the problem of mortality modelling and forecasting, to estimate mortality indices and actuarial life tables using the forecasts and to evaluate the applicability of the solution in a range of settings.

Aim I: To develop a method for forecasting female and male life expectancy based on analysis of the gap between female life expectancy in a country compared with a given benchmark and also by assessing the sex-gap evolution. (Paper I).

Aim II: To explore the relationship between life expectancy and age-specific deaths rates and develop a statistical model inspired by indirect estimation techniques applied in demography , which can be used to estimate full life tables at any point in time, based on a given value of life expectancy at birth or any other age. (Paper II).

Aim III: To propose a new method of predicting the death distributions of a certain population by making use of the properties of statistical moments given by a density function. (Paper III).

Aim IV: To demonstrate the applicability and accuracy of the developed methods by creating user-friendly open-source software packages. (R packages).

1.4 Methods and Data

The present thesis groups a series of new methods and applications in order to reach the defined aims. A brief description of these methods is provided in this section. The detailed models are presented in the thesis manuscripts.

1.4.1 Forecasting Life Expectancy

Life expectancy is highly correlated over time among countries and between males and females. The objective is to construct a model for forecasting life expectancy making use of the correlations existing among countries and between sexes. 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 used as benchmark in the manuscript I is that proposed by Oeppen and Vaupel (2002) for females in the record-holding countries. This trend was used due to its remarkable linear regularity at age 0.

To predict future life expectancy levels, the benchmark given by the best-practice life expectancy is identified in order to get a general sense of the direction and the rate of change in human mortality. The gap between female life expectancy in a given population and the best-practice trend in the world, $D_{k,x,t}$, is forecast using a classic time series model, thus determining future female life expectancy. The gap between male and female life expectancy, $G_{k,x,t}$ is forecast with the help of a mixed model to obtain the country specific male life expectancy.

$$\nabla^d D_{k,x,t} = \underbrace{\mu_{k,x}}_{\text{Drift}} + \underbrace{\sum_{i=1}^{p} \phi_i \nabla^d D_{k,x,t-i}}_{\text{Regression}} + \underbrace{\epsilon^{(1)}_{k,x,t} + \sum_{j=1}^{q} \theta_j \epsilon^{(1)}_{k,x,t-j}}_{\text{Smoothed noise}} \tag{1.3}$$

$$G^*_{k,x,t} = \begin{cases} \beta_0 + \underbrace{\beta_1 G_{k,x,t-1} + \beta_2 G_{k,x,t-2}}_{\text{Autoregressive model}} + \underbrace{\beta_3 (e^f_{k,x,t} - \tau)_+}_{\substack{\text{Level associated with} \\ \text{life expectancy when the gap} \\ \text{starts narrowing}}} + \epsilon^{(2)}_{k,x,t} & \text{if,} \quad e^f_{k,x,t} \leq A, \\ \underbrace{G_{k,x,t-1} + \epsilon^{(3)}_{k,x,t}}_{\text{Random walk}} & , \quad \text{otherwise.} \end{cases} \tag{1.4}$$

Once the two gaps are identified and their future trend determined, 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,

$$e^f_{k,x,t} = e^{bp}_{x,t} - D_{k,x,t}. \tag{1.5}$$

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,

$$e^m_{k,x,t} = e^f_{k,x,t} - G_{k,x,t}. \tag{1.6}$$

The current model is not restricted to the usage of a particular benchmark, countries or regions. One may decide to use a different trend depending on the best performing model for each case based on their past evaluation. Also, 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. 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.

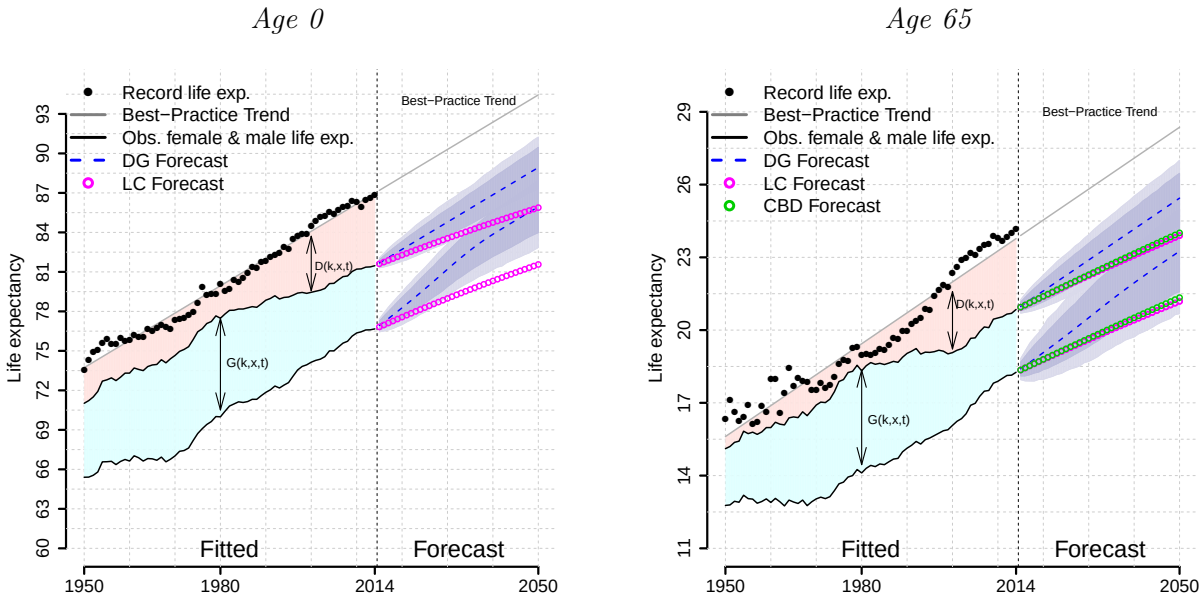


Figure 1.5: *Actual and forecast life expectancy at birth and at age 65 generated by the DG, LC and CBD models for females and males in the USA, 1950-2050. Prediction intervals at the 80% and 95% level are shown only for the DG model.*

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, 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.

1.4.2 Estimating Age-Specific Death Rates

In predicting demographic processes, such as human mortality, methods involving extrapolation of mortality rates or probabilities are the most common approaches. However, methods based on extrapolating life expectancy directly are very appealing because they offer the same, or higher, level of forecast accuracy but with the advantage of being parsimonious, focusing on one variable rather than several.

We have introduced a simple method, the Linear-Link model, to derive the entire schedule of age-specific death rates, based on a single value of life expectancy and prior knowledge

of human mortality patterns. Our method can be regarded as a decomposition approach of the human mortality curve between the general age pattern, β_x, and an age-specific speed of improvement, ν_x. The method is inspired by: (1) the Log-quadratic model (Wilmoth et al., 2012) in the sense of using a leading indicator in determining the age pattern of mortality; (2) the model introduced by Ševčíková et al. (2016) by adopting an inverse approach to death rates estimation starting from life expectancy; (3) the Lee–Carter model (1992) using the same interpretation of mortality improvement over time and age; and finally (4) the Li et al. (2013) method to model the rotation of age patterns of mortality decline for long-term projections.

The model is based on the observed linearity between age-specific death rates, m_x, and life expectancy at a certain age, e_θ. It can be seen as a method that links the life expectancy at age θ at any point in time to a mortality curve estimated from the death rates m_x's that return a life expectancy level of e_θ. To gain precision in the fitting of the death-rates, the Linear-Link model can be extended by including additional parameters:

$$\begin{aligned} \log m_{x,t} &= \beta_x \log e_{\theta,t} + \nu_x k + \varepsilon_{x,t} \quad \text{for} \quad x \geq \theta, \\ \sum_{x=\theta}^{\omega} \nu_x &= 1, \quad \text{and} \quad \nu_x \geq 0, \end{aligned} \tag{1.7}$$

where ν_x is the speed of mortality improvement over time at age x, k is an estimated correction factor independent of time and $\varepsilon_{x,t}$ are independent and identically distributed random variables normally distributed with mean zero and variance σ^2.

The method can be useful in three different situations: future target life expectancy, life tables for countries with deficient data and historical life table construction.

First, the model can be used in forecasting practice when the level of life expectancy is forecast first. We showed that this model can accurately reconstruct a Lee–Carter forecast starting from a single value of life expectancy at birth. This is important, because the Linear-link model offers the possibility of taking advantage of the more regular pattern of the life expectancy evolution. It is much easier and more parsimonious, from a technical

perspective to forecast one time series of expectation of life than to extrapolate 100 or 110 series of death probabilities corresponding to each age group. In the same manner adult mortality can be estimated based on a value of life expectancy at an advanced age, say age 65.

Second, the method can be used to build model life tables and to estimate the current age patterns of mortality in poor-data countries or regions, like Sub-Saharan Africa. In this case, the parameters of the model are estimated based on a collection of historical life tables from several regions or populations. Once the parameters have been estimated, and implicitly the model life table, they remain fixed. The relevant mortality curve is simply calibrated in accordance with a single value of life expectancy at birth or any other age instead of child mortality as in the case of Wilmoth et al. (2012). In our analysis, we show examples using high quality data from developed countries in order to demonstrate the efficiency of the model, and to be able to assess the accuracy of the mortality curve reconstruction. However, the estimation procedure and the steps of the algorithm are the same for this case too.

Third, the Linear-Link model can be a useful tool in a variety of research contexts of historical demography like backward projections and estimation of mortality levels in historical populations. Due to the existence of scarce non-standardized population data in the past and population censuses only for the more recent times, the very possibility of projecting mortality backward is of theoretical interest (Ediev, 2011).

1.4.3 Forecasting the age-at-death distribution

In addition to investigating the evolution of age-specific death rates or the pattern of life expectancy in the future, the problem of identifying possible future longevity levels in a given population can be tackled by analysing the distribution of deaths and the change in its location-shape measures over time. For a distribution, the collection of all the statistical moments uniquely determines its density function.

We consider the classical moment problem where a positive density $f(x)$ is sought from knowledge of its power moments. The method proposed here assesses the evolution of

observed moments of the distribution of deaths in order to forecast them by employing multivariate time series models and reconstructing the forecast distribution using the maximum entropy approach (*MaxEnt*) developed by Mead and Papanicolaou (1984). *MaxEnt* offers a definite procedure for the construction of a sequence of approximations for the true density based on the information entropy given by the density. The procedure aims at constructing specific sequences of functions $f_N(x)$ which eventually converge to the true distribution $f(x)$ as the number of moments used, N, approaches infinity

$$\mu_n = \int_a^\omega x^n f_N(x) dx, \quad n = 0, 1, 2 \ldots, N, \tag{1.8}$$

where μ_n represents the n-th statistical moment of a continuous density function. We denote the estimated density $f_N(x)$ in order to indicate that this density was generate based on a finites number of moments, N.

Taking advantage of the regularity of human mortality, the reconstruction of a density function can be obtained by imposing a prior restriction on the class of function where the solution is sought. In this way, only a small number of moments, usually 3 to 6, are needed in order to determine a good fit. As strategy for finding the local maxima of the entropy functional $\mathcal{L} = \mathcal{L}(f)$, we employ the method of Lagrange multipliers, λ_n for the n-th moment:

$$\mathcal{L} = H + \sum_{n=0}^{N} \lambda_n \left[\hat{\mu}_n - \mu_n\right], \tag{1.9}$$

where H denotes the information entropy measure of the estimated density.

Reconstructing the density function from a set of predicted moments has the advantage of allowing accelerating/decelerating rates of mortality improvement over age and time, identifying in this way the source of longevity risk.

1.4.4 Data

The data source used in this thesis is the Human Mortality Database (2017a), which contains homogeneous historical mortality data for populations in 43 different countries and territories. The HMD constitutes a reliable data source because it includes high quality data that were subject to a uniform set of procedures, thus maintaining the cross-national comparability of the information.

For the purpose of our analyses we have focused on a subset of these data covering mainly calendar years 1950–2015 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 the models and should be unique, meaning that a person included in one population should not be included in others.

The predictive power of the methods is demonstrated by performing out–of–sample forecasts and estimations in the 0–100 age range. Data at higher ages might be unreliable or too sparse for different populations, which would make it difficult to differentiate between data related problems and modelling issues.

1.5 Overview of the thesis and discussion

The evolution of human mortality is a complex process that is driven by a large number of factors and can not be explained by a single statistical model. Different approaches can be taken to predicting future mortality. In the current thesis, we introduce three innovative methods that make use of the observed trends in several demographic indicators.

Manuscript I presents a model built on the idea of a persistent trend in life expectancy over long periods of time that is driving the future development in longevity across countries and populations in a coherent manner. The available data in the Human Mortality Database (HMD) suggests a linear increase in record life expectancy at birth in the world since the mid-19th century. Since then, a few countries have occupied the first position (Bengtsson, 2006), maintaining their leadership for several years only to be surpassed by other countries that used to lag behind. The evolution of mortality for an individ-

ual country can be characterised by periods of rapid developments, stagnation or even deterioration of longevity levels. However, despite all these possible realizations, history shows that it is reasonable to assume that mortality is driven by a benchmark–like record female life expectancy. We implemented this idea in the so called *Double–Gap* model and tested it using data from all the countries listed in HMD. Detailed results shown for the USA, France and Sweden suggests that DG is capable of generating comparable predictive power with the two most commonly used forecasting models, the Lee–Carter and the Cairns–Blake–Dowd models.

In Manuscript II, an indirect estimation technique is employed to estimate the level of mortality at a given point in time by deriving the age-specific mortality estimates from the observed link between life expectancy and death rates. Life expectancy is an age-aggregated measure and 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. Used in the context of forecasting, the *Linear–Link* model demonstrated capabilities in replicating the Lee–Carter forecasts with the advantage of generating graduated mortality estimates.

Manuscript III goes deep into the statistical moments theory and proves that the problem of future development in longevity can be studied from a compositional perspective. A reduction in the probability of dying at young ages does not mean that fewer people will die. It only means that, for most people, death will occur at a later stage in life. Over a long enough time frame death is certain. The central moments of the distribution of deaths give an accurate description of the timing of death experience in a certain cohort or in a given year. If moments are extrapolated, one can use the maximum entropy method to reconstruct the future distribution of deaths. The main advantages of this approach is that it can anticipate where the most rapid change in mortality will occur on the age scale.

No statistical model in general and none of the three methods described in this thesis 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. The choice of the historical frame to be fitted is as important as the choice of the model.

Reproducible research

The presented methods and algorithms are implemented in the format of open source software libraries written in the `R` programming language (R Core Team, 2018). The `R` packages containing the source code, original data and usage examples can be downloaded and installed from the Comprehensive R Archive Network (`CRAN`) or from the author's GitHub repository (`https://github.com/mpascariu`). All the results presented in the thesis are reproducible.

Chapter 2

The Double-Gap Life Expectancy Forecasting Model

Marius Dan Pascariu
Vladimir Canudas-Romo
James W. Vaupel

Insurance: Mathematics and Economics 78: 339–350 (2017).

Corresponding Author: Marius D. Pascariu
Institute of Public Health
University of Southern Denmark
J.B. Winslows Vej 9B, 5000 Odense, Denmark
E-mail: mpascariu@health.sdu.dk

The Double-Gap Life Expectancy Forecasting Model

Marius D. Pascariu

Institute of Public Health, University of Southern Denmark, Odense, Denmark

Vladimir Canudas-Romo

School of Demography, The Australian National University, Canberra, Australia

James W. Vaupel

Max–Planck Institute for Demographic Research, Germany
Institute of Public Health, University of Southern Denmark, Odense, Denmark

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.

Keywords:

Life expectancy forecasting; Mortality modelling; Best practice trends; Sex-gap

2.1 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 recordholder 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 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. Coherent two-population modelling of age-specific death-rates have been done by Jarner and Kryger (2011); Cairns et al. (2011); Li and Hardy (2011) and Dowd et al. (2011).

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 2.2 the data used in fitting the model are presented. In Section 2.3 a new life expectancy projection model is proposed. In Section 2.4 a method to assess the performance of the model is given. Section 2.5 shows simulation results and illustrations of life expectancy in several countries by sex. The discussion and conclusion are in Section 2.6.

2.2 Data description

The data source used in this article is the Human Mortality Database (2017a), 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 2.1 along with the dates used to define the fitting periods.

2.3 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 bestpractice 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.

Table 2.1: *Selected HMD countries and years with available data used for the illustration*

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

Source: Human Mortality Database (2017a)

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,

$$e^f_{k,x,t} = e^{bp}_{x,t} - D_{k,x,t}. \tag{2.1}$$

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,

$$e^m_{k,x,t} = e^f_{k,x,t} - G_{k,x,t}. \tag{2.2}$$

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

$$e^{record}_{x,t} = \alpha_{x0} + \alpha_{x1}t, + \epsilon^{(0)}_{x,t}, \text{with } t = 1, 2, 3... \tag{2.3}$$

therefore,

$$e^{bp}_{x,t} = \alpha_{x0} + \alpha_{x1}t, \tag{2.4}$$

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.

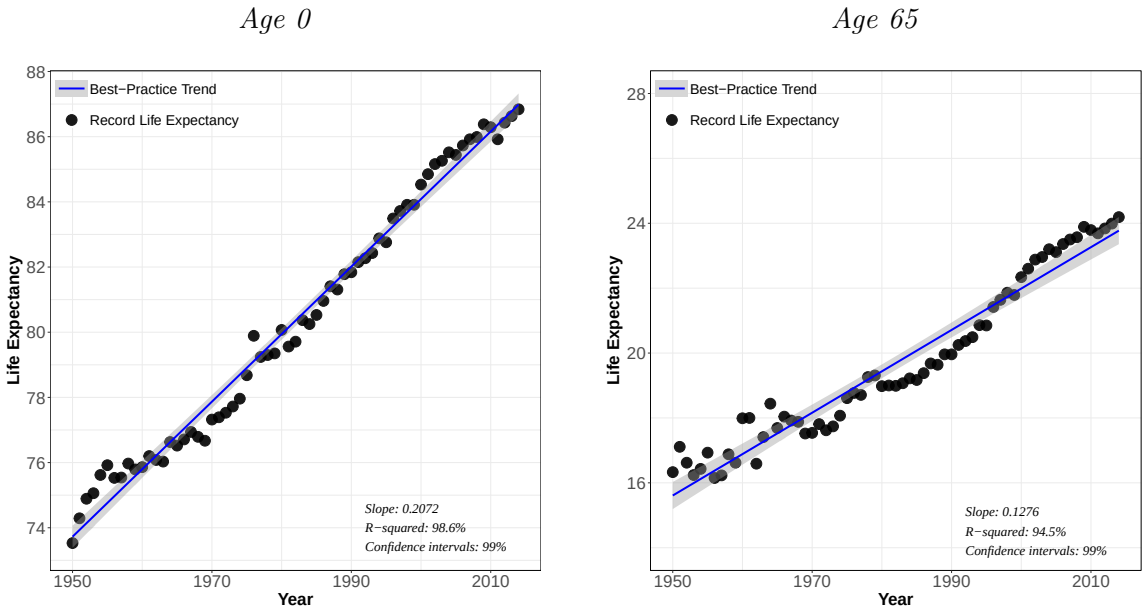


Figure 2.1: *The trend of record female life expectancy at birth and at age 65 between 1950 and 2014*

The Double-Gap model in equations (2.1) and (2.2) 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 equation (2.3). These rates of increase imply a change from 16.3 years in 1950 in Iceland to 24.2 in 2014 in Japan. The linear fit is presented in Figure 2.1.

2.3.2 Step 2 - The gap to best-practice trend

One way to forecast the gap between the best-practice trend and country specific female life expectancy, $D_{k,x,t}$, 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.

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 $D_{k,x,t}$ is given by:

$$\nabla^d D_{k,x,t} = \underbrace{\mu_{k,x}}_{\text{Drift}} + \underbrace{\sum_{i=1}^{p} \phi_i \nabla^d D_{k,x,t-i}}_{\text{Regression}} + \underbrace{\epsilon^{(1)}_{k,x,t} + \sum_{j=1}^{q} \theta_j \epsilon^{(1)}_{k,x,t-j}}_{\text{Smoothed noise}} \tag{2.5}$$

where the response can be obtained from 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 $\sigma^{(1)}$.

Table 2.2: *Estimated parameters of the ARIMA model for the gap between best-practice and country specific data at birth and at age 65, 1950-2014.*

	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

Source: Authors' calculations based on data described in Table 2.1

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.

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 parsimo-

nious 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 Figure 2.2.

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.

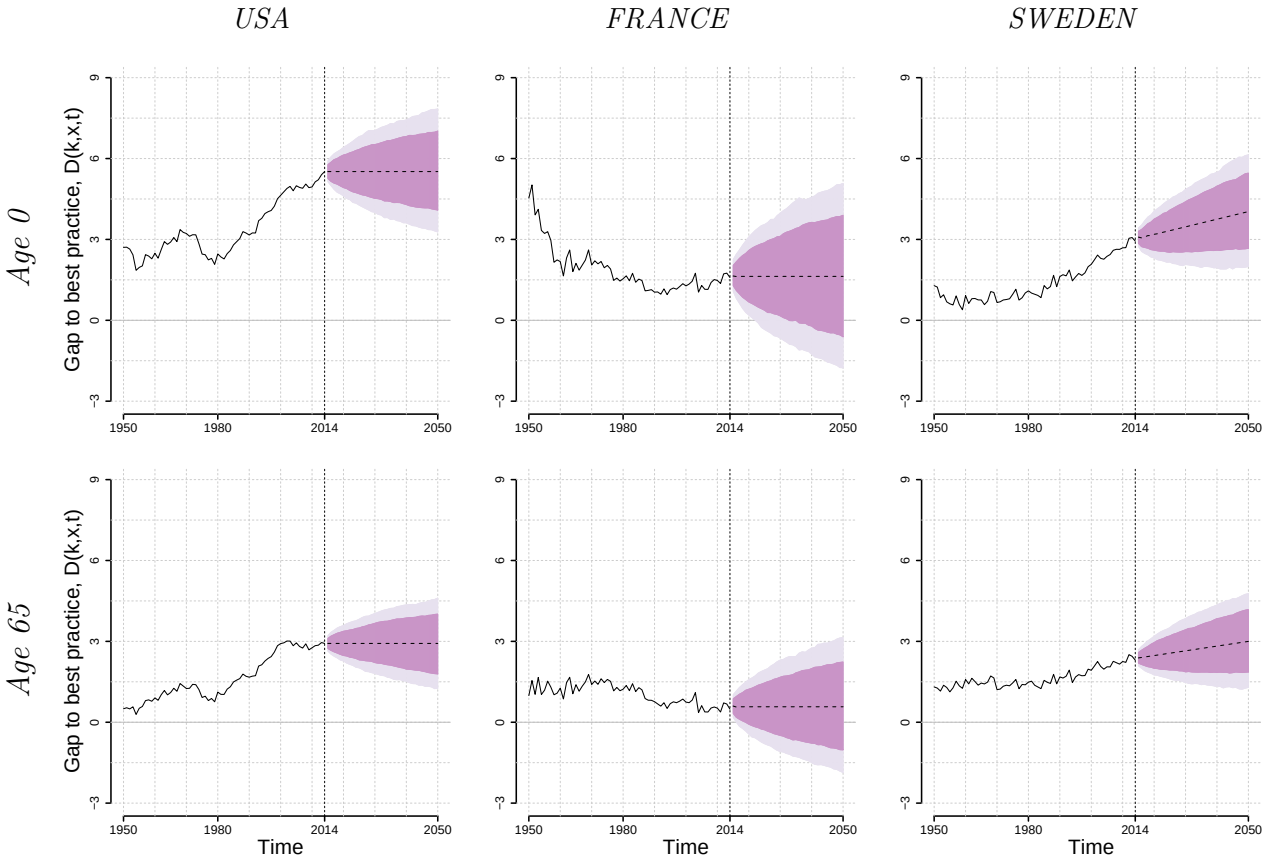


Figure 2.2: *The forecast gap between the best-practice trend and country-specific female life expectancy at birth and at age 65, with associated 80% and 95% prediction intervals, 1950-2050.*

2.3.3 Step 3 - The sex gap model

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).

$$G^*_{k,x,t} = \begin{cases} \beta_0 + \underbrace{\beta_1 G_{k,x,t-1} + \beta_2 G_{k,x,t-2}}_{\text{Autoregressive model}} + \underbrace{\beta_3 (e^f_{k,x,t} - \tau)_+}_{\substack{\text{Level associated with} \\ \text{life expectancy when the gap} \\ \text{starts narrowing}}} + \epsilon^{(2)}_{k,x,t} & \text{if,} \quad e^f_{k,x,t} \leq A, \\ \underbrace{G_{k,x,t-1} + \epsilon^{(3)}_{k,x,t}}_{\text{Random walk}} & \text{, otherwise.} \end{cases} \tag{2.6}$$

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.

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,

$$G_{k,x,t} = min\{max\{G^*_{k,x,t}, L\}, U\}, \tag{2.7}$$

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 2.1.

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.

Table 2.3: *Estimated parameters for sex-gap forecast models for life expectancy at birth and age 65*

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	

Source: Authors' calculations based on data described in Table 2.1

Estimates of the model parameters are provided in Table 2.3 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 2.3, 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 Figure 2.3. 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 equation (2.6). 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 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.

2.3.4 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.

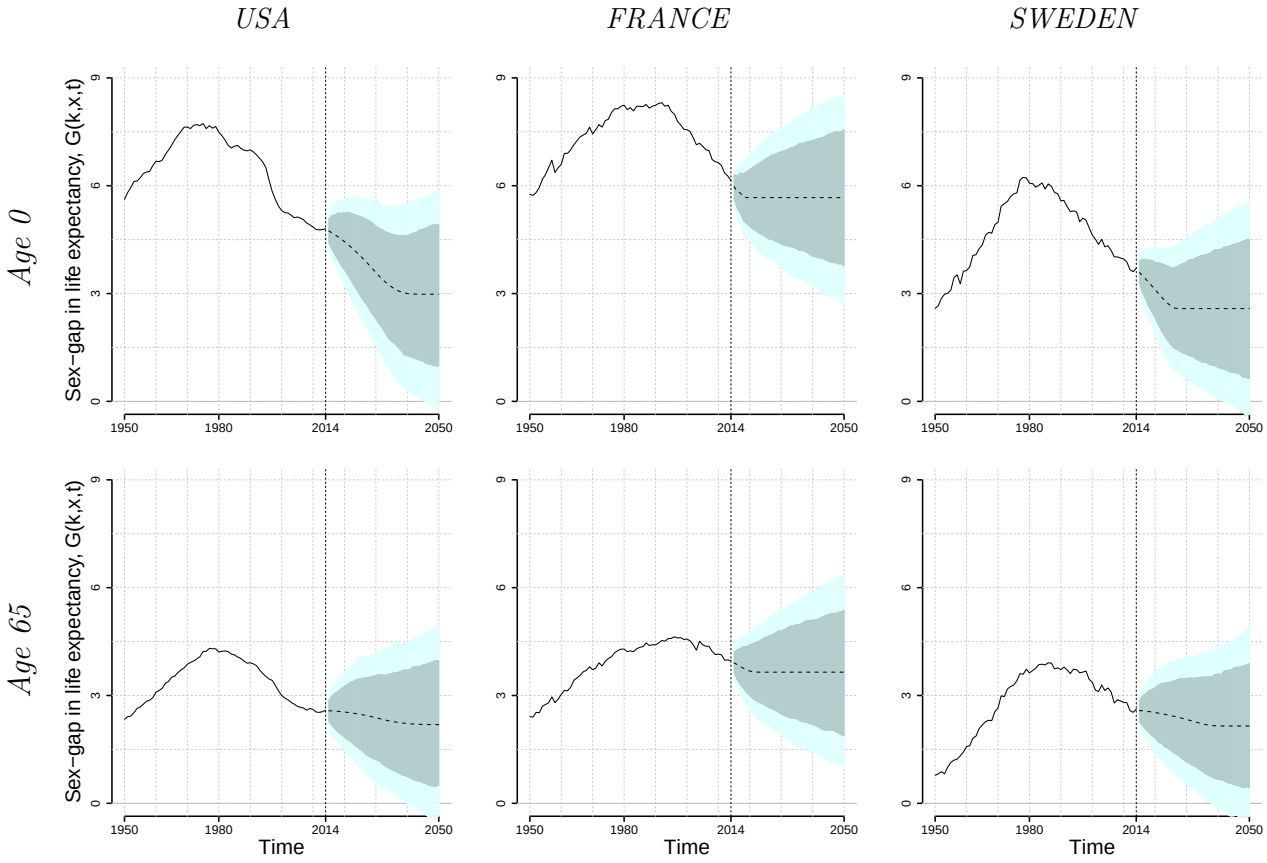


Figure 2.3: *The forecast gap between female and male life expectancy at birth and at age 65, with associated 80% and 95% prediction intervals, 1950-2050*

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,

$$\xi \sim \mathcal{N}_3(\mu, \Sigma), \tag{2.8}$$

with the mean vector,

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

and 3×3 covariance matrix,

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

The time series of errors obtained by fitting the random walk model in the Raftery model, $\epsilon^{(3)}_{k,x,t}$ (see Equation 2.6), 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 the $\epsilon^{(2)}_{k,x,t}$.

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.

2.4 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:

$$\omega_{k,x,t} = e_{k,x,t} - \hat{e}_{k,x,t}. \tag{2.9}$$

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

[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.

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.

$$ME = mean\Big(\omega_{k,x,t}\Big),$$
$$MAPE = mean\Big(|100 \times \frac{\omega_{k,x,t}}{e_{k,x,t}}|\Big), \tag{2.10}$$

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

2.5 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 2.4.

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 (≥ 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 sexgaps forecast by the LC and CBD model are similar.

A visual representation of the results already presented in Table 2.4 is given in Figure 2.4 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.

Table 2.4: *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.*

			AGE 0			AGE 65		
MODEL			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.

Source: Authors' calculations based on data described in Table 2.1

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.

Out-of-sample forecasts are performed using the DG, LC and CDB models in order to test the performance of the three models. Four forecasting horizon are selected starting with 1985 until 2014. The forecast values are compared with the historical values of life expectancy. Table 2.5 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 2.5: *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.*

		AGE 0		AGE 65	
COUNTRIES	MODEL	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

Table 2.6 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 that 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.

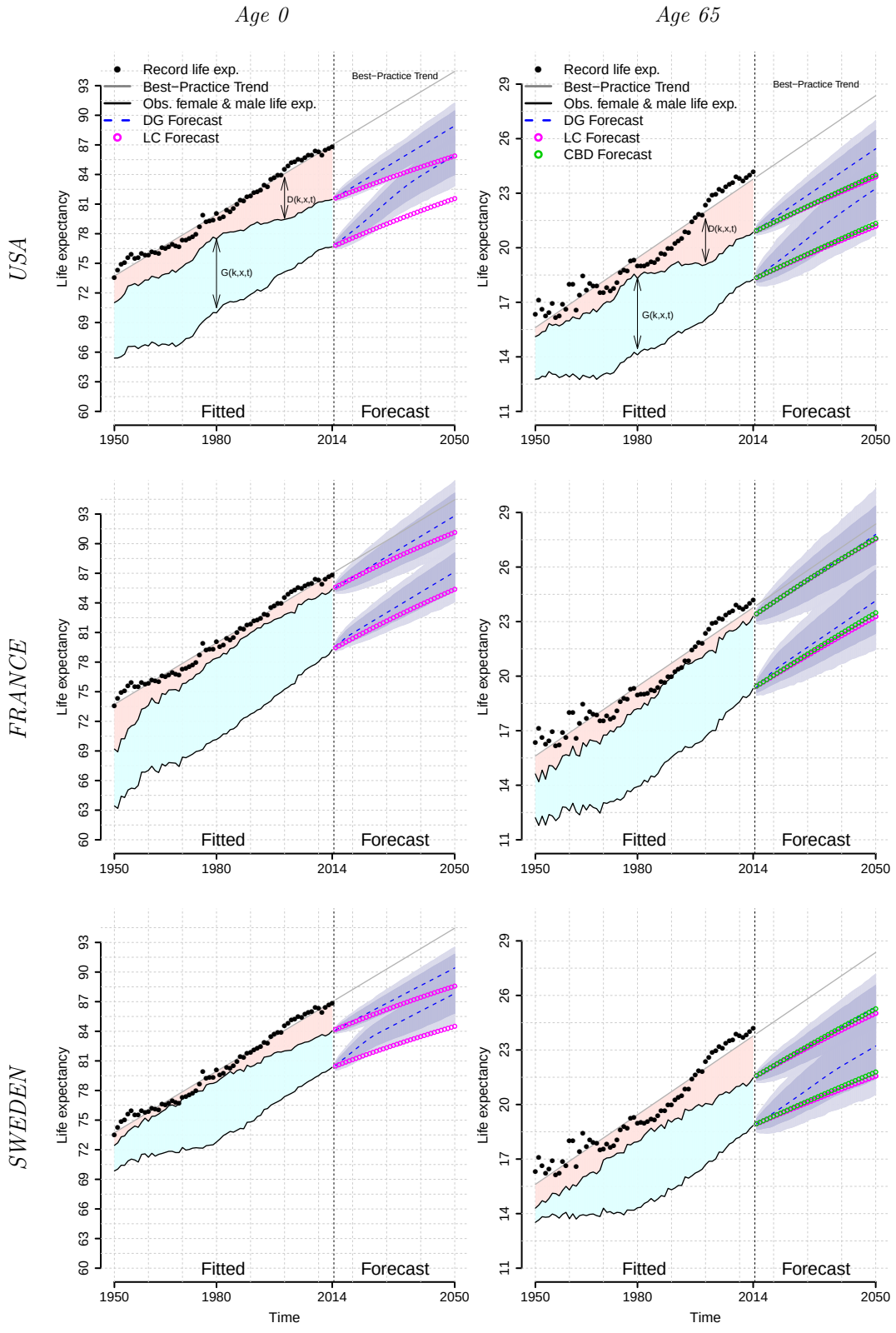


Figure 2.4: *Actual and forecast life expectancy at birth and at age 65 generated by the DG, LC and CBD models for females and males, 1950-2050. Prediction interval at 80% and 95% level are shown only for the DG model.*

Table 2.6: *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.*

		FEMALE POPULATION				**MALE POPULATION**			
		AGE 0		AGE 65		AGE 0		AGE 65	
COUNTRY	MODEL	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

More visual results for 18 countries are presented in Figure 2.5 and Figure 2.6 in the Appendix.

2.6 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 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 Figure 2.1) 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 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, 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.

2.7 Appendix

2.7.1 Out-of-sample forecasts for 18 countries, 1990–2014

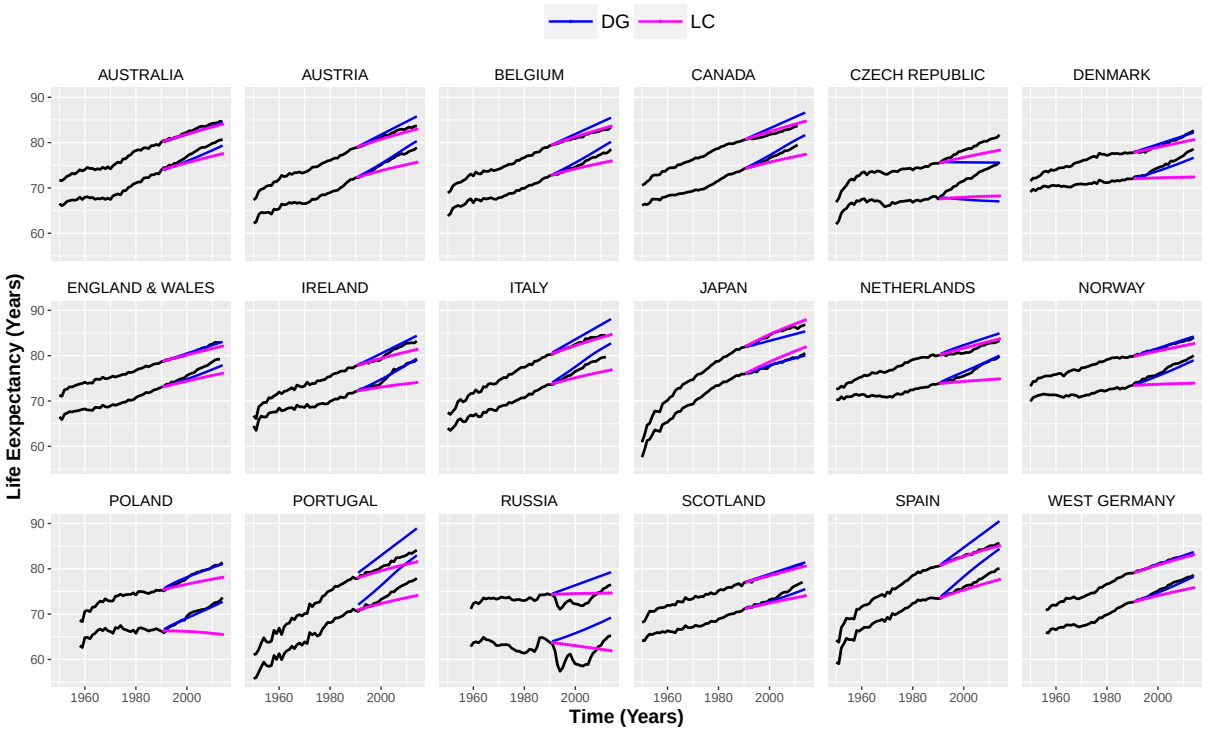


Figure 2.5: *Comparison of actual life expectancy at birth in 1990–2014 with forecasts generated by the Double–Gap and Lee–Carter models for 18 countries and regions.*

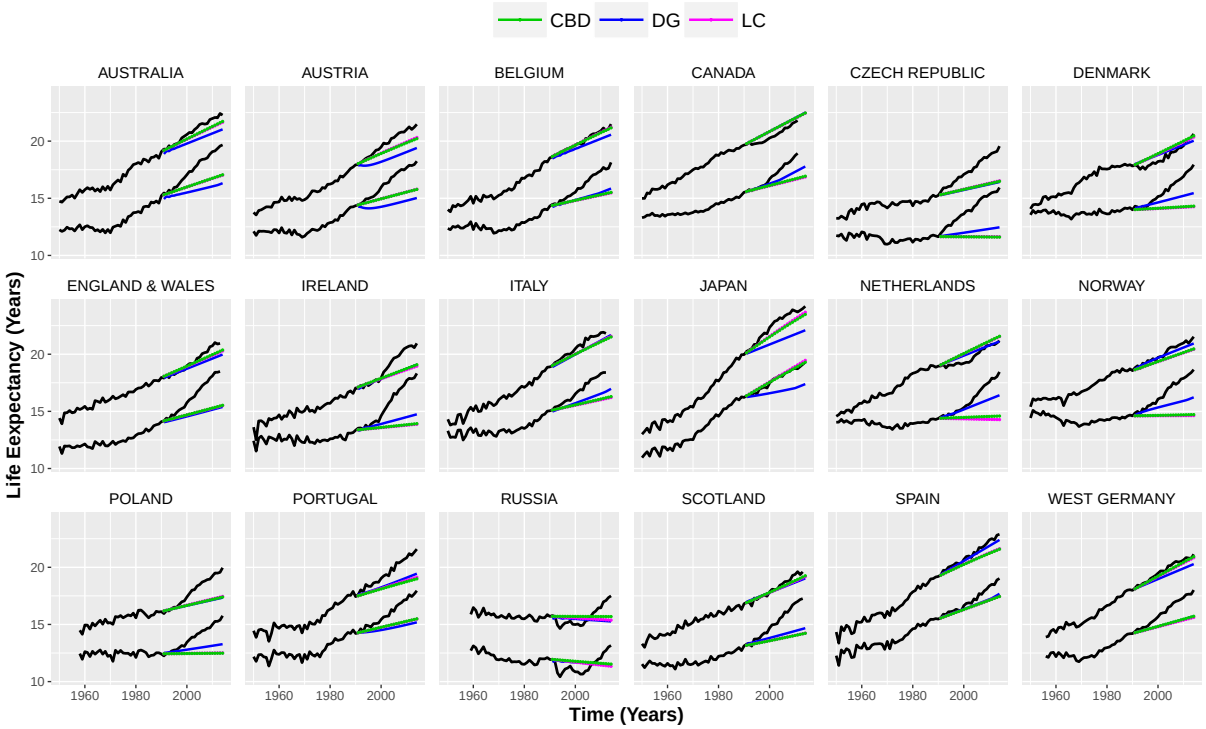


Figure 2.6: *Comparison of actual life expectancy at age 65 in 1990–2014 with forecasts generated by the Double–Gap, Lee–Carter and CBD models for 18 countries and regions.*

2.7.2 Forecasts for 18 countries and regions, 2015–2050

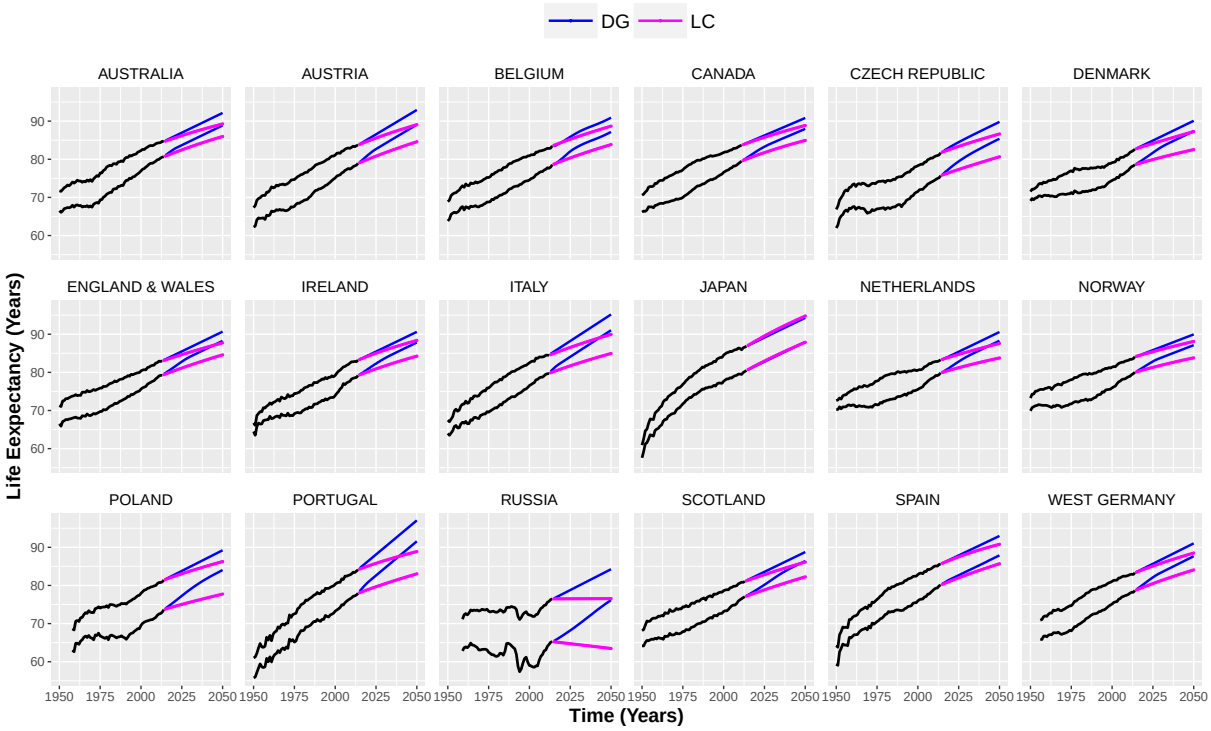


Figure 2.7: *Forecast life expectancy at birth in 2015–2050 for 18 countries and regions*

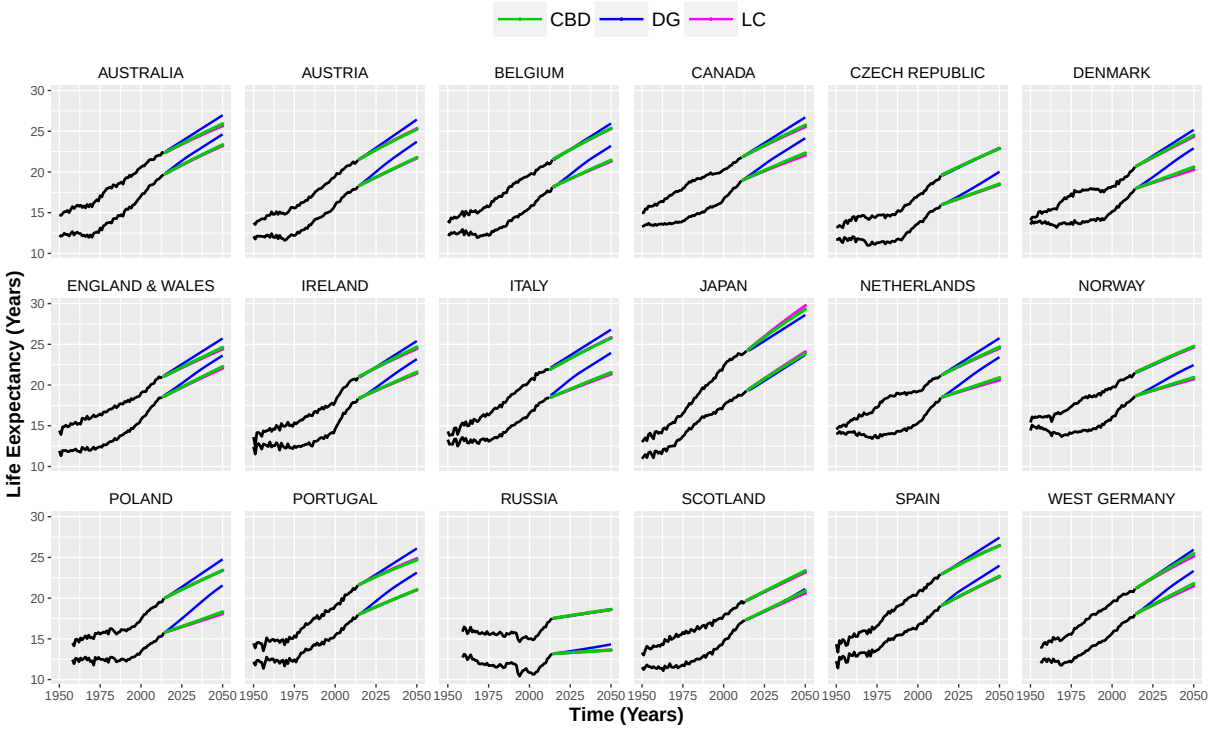


Figure 2.8: *Forecast life expectancy at age 65 in 2015–2050 for 18 countries and regions*

Chapter 3

The Linear Link: Deriving Age-Specific Death Rates from Life Expectancy

Marius Dan Pascariu
Ugofilippo Basellini
Jose Manuel Aburto
Vladimir Canudas-Romo

Corresponding Author: Marius D. Pascariu
Institute of Public Health
University of Southern Denmark
J.B. Winslows Vej 9B, 5000 Odense, Denmark
E-mail: mpascariu@health.sdu.dk

The Linear Link: Deriving Age-Specific Death Rates from Life Expectancy

Short title: The Linear-Link Model

Marius D. Pascariu

Institute of Public Health, University of Southern Denmark, Odense, Denmark

Ugofilippo Basellini

Institut national d'etudes démographiques (INED), Paris, France
Institute of Public Health, University of Southern Denmark, Odense, Denmark

José Manuel Aburto

Institute of Public Health, University of Southern Denmark, Odense, Denmark

Vladimir Canudas-Romo

School of Demography, The Australian National University, Canberra, Australia

Abstract

The prediction of human longevity levels in the future by direct forecasting of life expectancy offers numerous advantages, compared to methods based on extrapolation of age-specific death rates. However, the reconstruction of accurate life tables starting from a given level of life expectancy at birth, or any other age, is not straightforward. Model life tables have been extensively used for estimating age patterns of mortality in poor-data countries. We propose a new model inspired by indirect estimation techniques applied in demography, which can be used to estimate full life tables at any point in time, based on a given value of life expectancy at birth. The methods presented in this paper are implemented in a publicly available `R` package.

Keywords:

Indirect estimation; Life expectancy; Forecasting; Death rates; Age-patterns of mortality

3.1 Introduction

Understanding human mortality dynamics is of utmost importance in the context of rapid ageing process together with the increase in length of life experienced by most populations nowadays. The link between the pension systems sustainability and changes in life expectancy is more apparent than ever in light of the recent reforms that are taking place in Europe. In countries like Germany and Finland the level of retirement benefits are linked to life expectancy, in other countries like the U.K. and France the retirement age is set to increase from the current levels and implicitly the contribution period for pensions to be extended as people live longer (Stoeldraijer et al., 2013).

In predicting demographic processes, such as human mortality, methods involving extrapolation of mortality rates or probabilities are the most common approaches. Stochastic models, such as those proposed by Lee and Carter (1992) or Cairns et al. (2006) have gained significant popularity and have been extensively used in the last two decades. Ideas that focus only on life expectancy have given rise to a new approach. The models introduced by Torri and Vaupel (2012), Raftery et al. (2014) and Pascariu et al. (2018) are partially inspired by the linear time trends observed in life expectancy at birth in many developed countries, particularly in the second half of the twentieth century (Oeppen and Vaupel, 2002; White, 2002). These life expectancy models are very appealing because they offer the same, or higher, level of forecast accuracy in terms of life expectancy but with the advantage of being parsimonious, focusing on one variable rather than several. They rely on a measure that incorporates all the factors that influence longevity (lifestyle, access to healthcare, diet, economical status, etc.), namely life expectancy (Christensen and Vaupel, 1996). Furthermore, highly aggregated data by age provide valuable information that can be used to tackle the issue of mortality forecasting from a clearer perspective. The U.S. Census Bureau predicts the future mortality levels up to year 2100 based on projections of life expectancy at birth by sex and race, modelling an exponential decline of the gap to the observed upper asymptote of life expectancy. The period age-specific death rates are estimated in a subsequent step using these projections (United States Census Bureau, 2014).

Transformation of life expectancy into mortality rates at every age can be accomplished by exploiting the regularities of age patterns of mortality. In 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. Basically, actuarial methods combine the life table with functions related to an assumed rate of interest (Møller and Steffensen, 2007; Dickson et al., 2013). Based on the relevance of having a set of age-specific death rates, we propose a method to create such an array of values from one available life expectancy.

Our method extends the work initiated by the different systems of model life tables (Gabriel and Ronen, 1958; United Nations, 1955, 1967; Coale and Demeny, 1966, 1983; Ledermann, 1969; Sullivan, 1972); Brass' relational model (1968; 1971) and the recent extensions of techniques for estimating age patterns of mortality by Murray et al. (2003) and Wilmoth et al. (2012). Our model is also related to the work of Mayhew and Smith (2013) that uses the trends in life expectancy to establish a robust statistical relation between changes in life expectancy and survivorship. A further, similar approach to the one developed here is that of Ševčíková et al. (2016) which incorporates a method based on the Lee–Carter model for converting projected life expectancies at birth to age-specific death rates in the UN's 2014 probabilistic population projections.

Relational models were developed for estimation purposes in poor-data contexts. These models rely on parameters that depict the relationships between various measures of age-specific and overall mortality. The parameters in a relational model are estimated from an initial analysis of historical mortality data and become fixed thereafter. Once those values have been estimated, the model simplifies to a few initial inputs like: reported child survival, records of population growth, responses to questions about fertility and mortality and in our case life expectancy at birth or at any other age. Furthermore, in recent years the accessibility of historical mortality data, such as the Human Mortality Database (HMD), means that the necessary information to estimate the parameters of relational models is readily available. As shown here, an algorithm that derives a life table based only on life expectancy at birth can also be widely used in forecasting practice.

The remainder of the article is organized as follows. First, in Section 3.2 a new model to derive age-specific death rates is introduced and a description of the data used in testing is provided. Section 3.3 shows computed results and illustrations of life expectancy decomposition into death rates in several populations. The discussion and conclusion are presented in Section 3.4.

3.2 Data and Methods

3.2.1 Data

The data source used in this article is the Human Mortality Database (2017b), which contains historical mortality data for homogeneous populations in 43 different countries and territories. The HMD constitutes a reliable data source because it includes high quality data that were subject to a uniform set of procedures, thus maintaining the cross-national comparability of the information.

In order to test and illustrate the performance of the method, we fit the model using the death rates computed using death counts and population exposed to the risk of death in the calendar year for the female populations of the England & Wales, France, Sweden and USA available in the HMD. The reason for using these death rates is that old age mortality in the HMD is often subject to diverse correction procedures and modelling depending on the country (Wilmoth et al., 2007).

The reconstructive power of the method for a point forecast of life expectancy is demonstrated using the 1980–2014 mortality data between age 0 and 100. Data at higher ages might be unreliable or too sparse for different populations, which would make it difficult to differentiate between data related problems or modelling issues. To compute the accuracy measures and the estimation errors, the 1965–90 data is applied to the same age range.

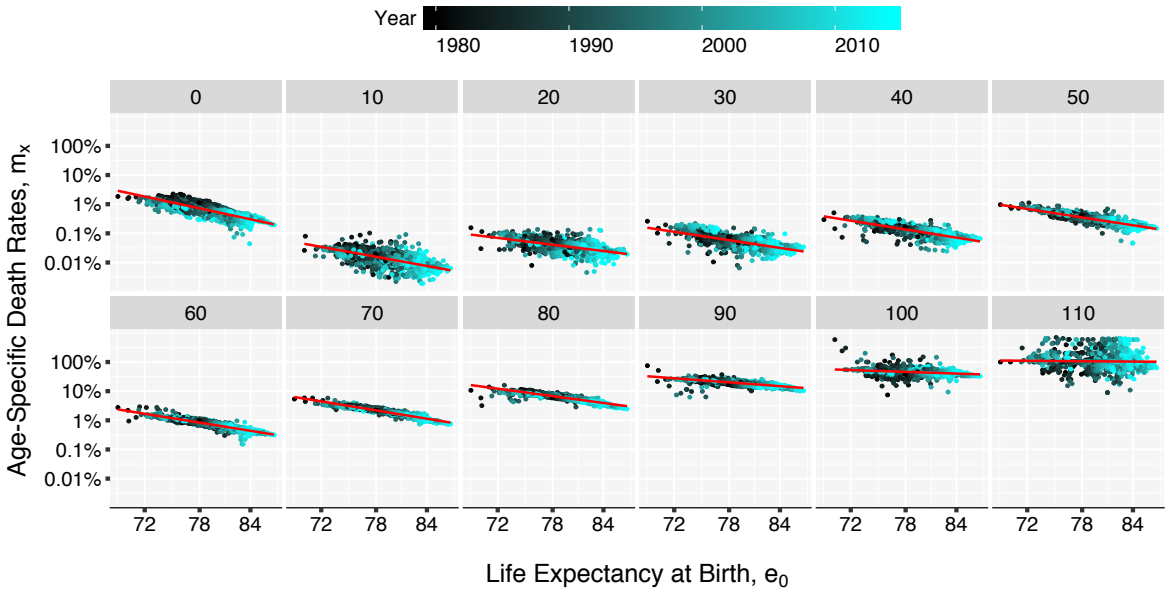


Figure 3.1: *Linear relation between life expectancy at birth and death-rates on a log-log scale, by age. The axis are labelled in normal scale for better interpretability. Based on HMD mortality data starting from 1980 for 43 countries and territories.*

3.2.2 The Model

Given a predicted level of life expectancy the age pattern of mortality can be derived using a linear relation. The logarithm age-specific death rate at time t, denoted $m_{x,t}$, can be expressed as a linear function of the logarithm of life expectancy at a given age θ, denoted $e_{\theta,t}$. Formally:

$$\log m_{x,t} = \beta_x \log e_{\theta,t} + \varepsilon_{x,t} \quad \text{for} \quad x \geq \theta, \tag{3.1}$$

where x can take values between 0 and ω, the highest attainable age, and β_x can be regarded as an age-specific parameter. Parameter $\varepsilon_{x,t}$ denotes a set of normally distributed errors with mean zero and variance σ^2. For example, when θ equals zero, we estimate an entire mortality curve based on life expectancy at birth using this equation; when $\theta > 0$, we estimate the mortality curve starting from age θ.

The method presented here combines the strong linear relations found when comparing life expectancies and age-specific death rates on a log-log scale. Figures 3.1 and 3.2 show those relations, although the slopes and intercepts vary between ages, in all cases there

is a strong linear concordance between the level of overall mortality, as depicted by life expectancy, and the individual age-specific death rates. These relations have been key in much of the work on model life tables (Gabriel and Ronen, 1958; United Nations, 1955, 1967; Coale and Demeny, 1966, 1983; Ledermann, 1969). Inspired by the two-dimensional system (age and time) of the Log-quadratic model Wilmoth et al. (2012) and the strong linear trends in Figures 3.1 and 3.2, we derive the age pattern of mortality based on a given value of life expectancy, e.g. forecasted life expectancy value, and a matrix of age-specific death rates from the past.

This model can be seen as a method that links the life expectancy at age θ at any point in time to a mortality curve estimated from the death rates m_x's that return a life expectancy level of e_θ. Therefore we will refer to it as the linear-link (LL) model. To gain precision in the fitting of the death rates the LL model can be extended by including additional parameters:

$$\begin{aligned} \log m_{x,t} &= \beta_x \log e_{\theta,t} + \nu_x k + \varepsilon_{x,t} \quad \text{for} \quad x \geq \theta, \\ \sum_{x=\theta}^{\omega} \nu_x &= 1, \quad \text{and} \quad \nu_x \geq 0, \end{aligned} \tag{3.2}$$

where ν_x is the speed of mortality improvement over time at age x, k is an estimated correction factor independent of time and $\varepsilon_{x,t}$ are independent and identically distributed random variables normally distributed with mean zero and variance σ^2. Different than the Log-quadratic model that has a fixed set of parameters for any input value, here the parameters β_x, ν_x and k can be calculated for each set of age-specific death rates and future life expectancy. Thus, it can be seen as an extension of the log-quadratic model for countries that have good quality data, where an entire life table is completed from one target value of life expectancy.

In addition, the LL model is closely related to the LC model. Indeed, if one sets the parameters $\beta_x log e_{\theta,t} = \alpha_x$, $\nu_x = \beta_x$, and $k = k_t$, we obtain the LC model. Despite their similarities, there are two important differences between the two models. First, while the shape of the mortality pattern α_x is constant in the LC model, the first term of the

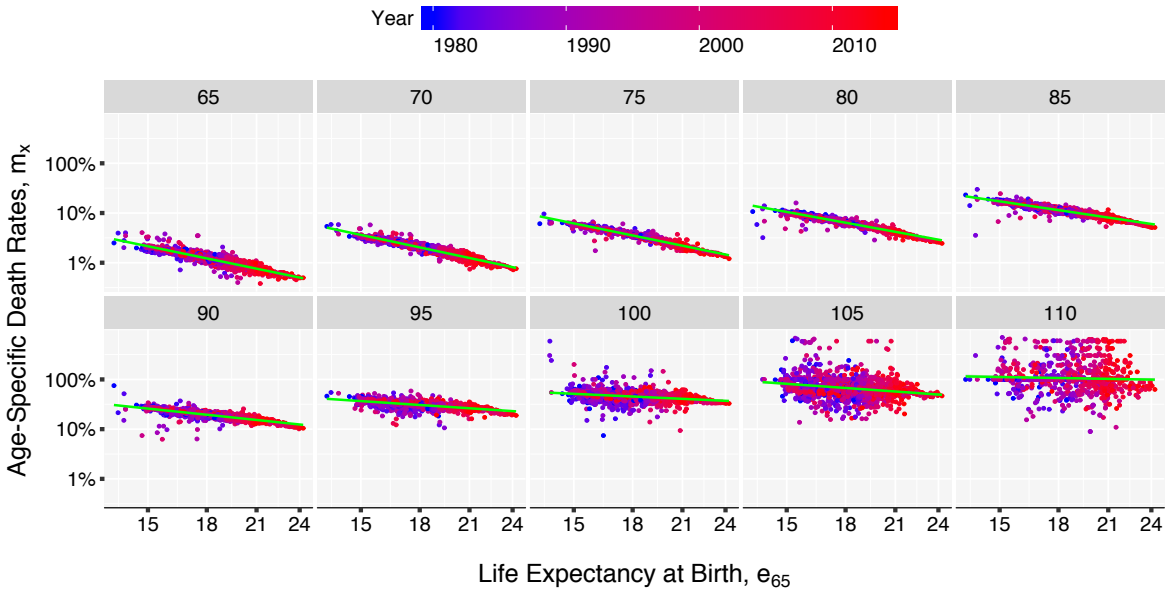


Figure 3.2: *Linear relation between life expectancy at age 65 and death-rates on a log-log scale, by age. The axis are labelled in normal scale for better interpretability. Based on HMD mortality data starting from 1980 for 43 countries and territories.*

LL changes with the level of life expectancy considered; as such, there exists a range of different baseline mortality curves of the LL model depending on the particular level of e_θ. Second, the k parameter is not modelled as a function of time, instead the parameter is used as an optimization variable affecting the shape of the age pattern of mortality to achieve the desired target life expectancy e_θ. Thus, the k parameter enhances the flexibility of the method and the accuracy of the results.

3.2.3 Algorithm

Let t be an observed unit of time in the interval $\{1, ..., T\}$ and τ be a an unobserved point in time $T+n$ e.g. a date in the future. The objective is to convert a value of life expectancy, $e^*_{\theta,\tau}$, into a schedule of age-specific death rates $m^*_{x,\tau}$. The level of life expectancy can be a predicted value given by certain extrapolation method or the target values resulted following a subjective judgement. Input data will be a collection of observed death rates $m_{x,t}$ and a level of life expectancy $e^*_{\theta,\tau}$. The steps involved in the algorithm to obtain the desired death rates are the following:

1. Using the Kannisto mortality model (see Appendix) extend $m_{x,t}$ to higher age groups up to age ω for all times t. The highest attainable age, ω, can be set for example to 120.

2. Estimate the slope of the linear relation between life expectancy and the death-rates, β_x, over the observation time t. This is done by using the method of the least squares approach, by minimizing the sum of squared residuals:

$$\sum_x \left[\log m_{x,t} - \beta_x \log e_{\theta,t}\right]^2 = \sum_x \left[\varepsilon_{x,t}\right]^2. \tag{3.3}$$

   Alternatively, the parameters of the model can be estimated by assuming that deaths follow a Poisson distribution (Brouhns et al., 2002), $D_x \backsim Poisson(E_x \cdot m_{x,t})$, with $m_{x,t} = \exp(\beta_x \log e_\theta + \nu_x k)$. In order to use this approach death counts $(D_{x,t})$ and exposure data $(E_{x,t})$ are needed. Sensitivity analysis shows that the difference between the two fitting procedure return minor discrepancies (see section 3.5.2 in the Appendix for more details).

3. Estimate the parameter ν_x by computing the singular value decomposition (SVD) of the matrix of regression residuals, $\mathbf{R}$, obtained in the previous step,

$$SVD\left[\mathbf{R}\right] = \mathbf{DPQ}^T = d_1 p_1 q_1^T + ..., \tag{3.4}$$

   where

$$\mathbf{R} = \begin{bmatrix} \varepsilon_{0,1} & \varepsilon_{0,2} & \cdots & \varepsilon_{0,T} \\ \varepsilon_{1,1} & \varepsilon_{1,2} & \cdots & \varepsilon_{1,T} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{\omega,1} & \varepsilon_{\omega,2} & \cdots & \varepsilon_{\omega,T} \end{bmatrix},$$

   $\mathbf{P} = [p_1, p_2, ...]$ and $\mathbf{Q} = [q_1, q_2, ...]$ are matrices of left and right singular vectors, and $\mathbf{D}$ is a diagonal matrix with singular values along the diagonal. The first term of the SVD, $d_1 p_1 q_1^T$, is used for obtaining the estimates of ν_x. Parameter ν_x can be interpreted as the rate of mortality improvement over age.

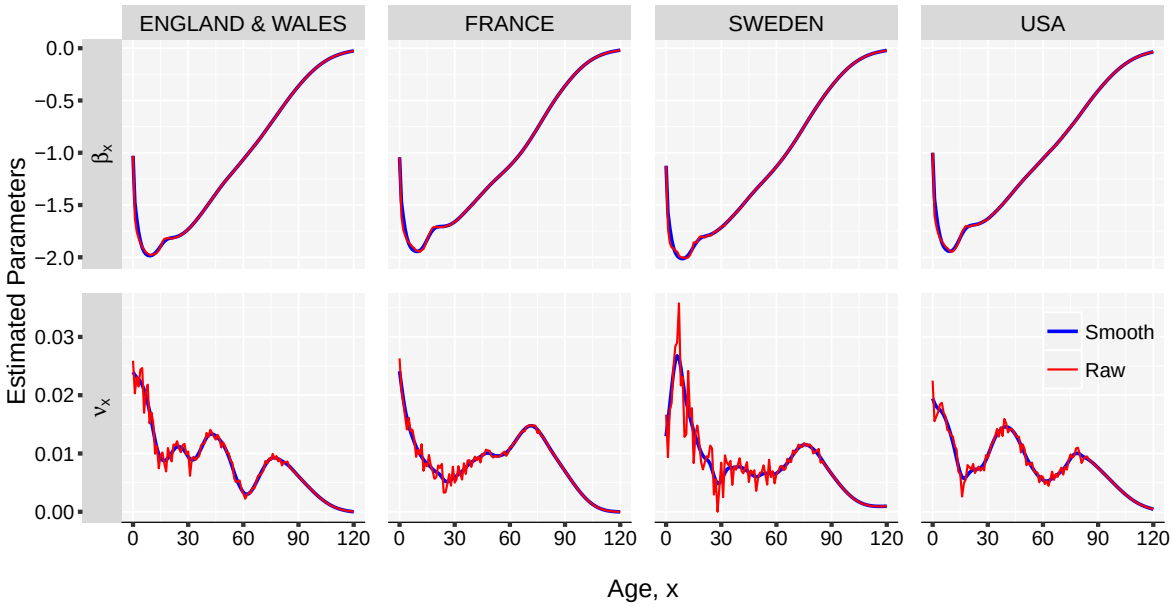


Figure 3.3: *Estimated parameters of the Linear–link model, using HMD data from 1980 to 2014 and life expectancy at birth* ($\theta = 0$).

4. Smooth the β_x and ν_x parameters using splines. This step is important to obtain graduated mortality curves and avoid projecting age-specific noise in the jump-off life table. However, if the graduation is not of interest or if the input data-set is large enough, this step can be skipped.

5. Compute the initial mortality rates[1] by $m^*_{x,\tau} = \exp\{\beta_x \log e^*_{\theta,\tau} + \nu_x k\}$, where $k = 0$.

6. Optimize the mortality curve given in the previous step by finding the value of k where the difference between target life expectancy $e^*_{\theta,\tau}$ and an estimated life expectancy $e_{\theta,\tau}$ is below a tolerance level, for example 0.001. Where $e_{\theta,\tau}$ represents the level of life expectancy at birth computed based on the mortality rates obtained in step (5). Usually k will be in the range of $(-150, +150)$ depending on the length of the forecast window.

The estimated β parameters for the female populations in England & Wales, France, Sweden and USA, exhibit minor differences between the countries, and capture well the important stages of human mortality: the decreasing infant mortality, the accidental hump, the adult mortality characterized by an exponential increase with age and, finally,

[1]The change in age-specific death rates can be assumed to be constant over time, in which case the fitted ν_x is used in computing m_x. Or, a shift in the speed of improvement can be imposed by "rotating" the ν_x coefficients. For more details see Section 3.5.3 in the Appendix.

a mortality plateau above the age of 100 years. As shown in Figure 3.3, the ν_x pattern differs from population to population. In the case of Sweden, a larger variance is observed over ages due to a smaller population size and more significant changes at younger ages in the period analysed.

3.3 Results and Illustration

We perform back-testing against the observed mortality for the female populations living in England & Wales, France, Sweden and USA. We take the period of 1965–90 as reference and use the death-rates and life expectancies at birth in this time interval to fit our model. Based on single values of life expectancy at birth observed in the subsequent years we

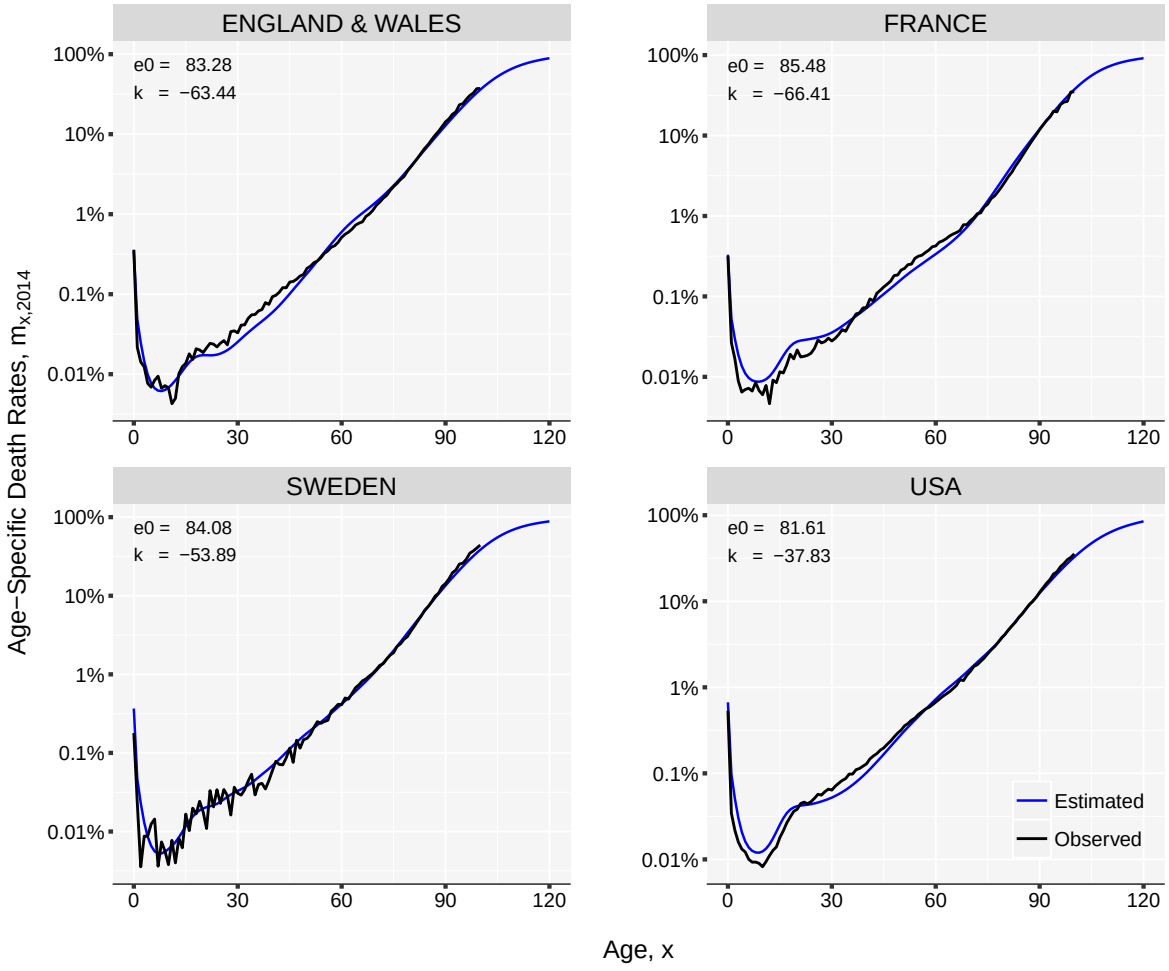


Figure 3.4: *Observed and estimated death rates for female populations in 2014. Computed based on mortality data in the period 1965–90.*

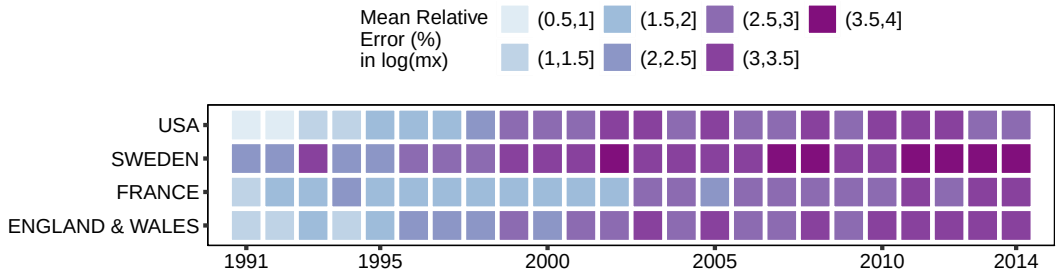


Figure 3.5: *Mean absolute errors (%) of the estimated log-death rates against the actual log-death rates between 1991 and 2014. Computed based on female mortality data in the period 1965–90.*

derive complete mortality curves. For example, the estimation of the age-specific death rates in 2014 is demonstrated in Figure 3.4. The reconstructed mortality curves are in general smoother than the observed data; this is more evident in the case of Sweden, where the population is smaller compared with the other three countries.

Figure 3.5 shows that the average relative error of the estimated log-death rates, compared to the actual rates between 1991 and 2014, is between 0.9% and 3.8%. It can be also noted that the longer the prediction interval, the larger the errors. In the case of female populations living in England & Wales, France, Sweden and USA, the largest errors occurred in 2014; nonetheless these are smaller than 3.8% of the actual log-death rate.

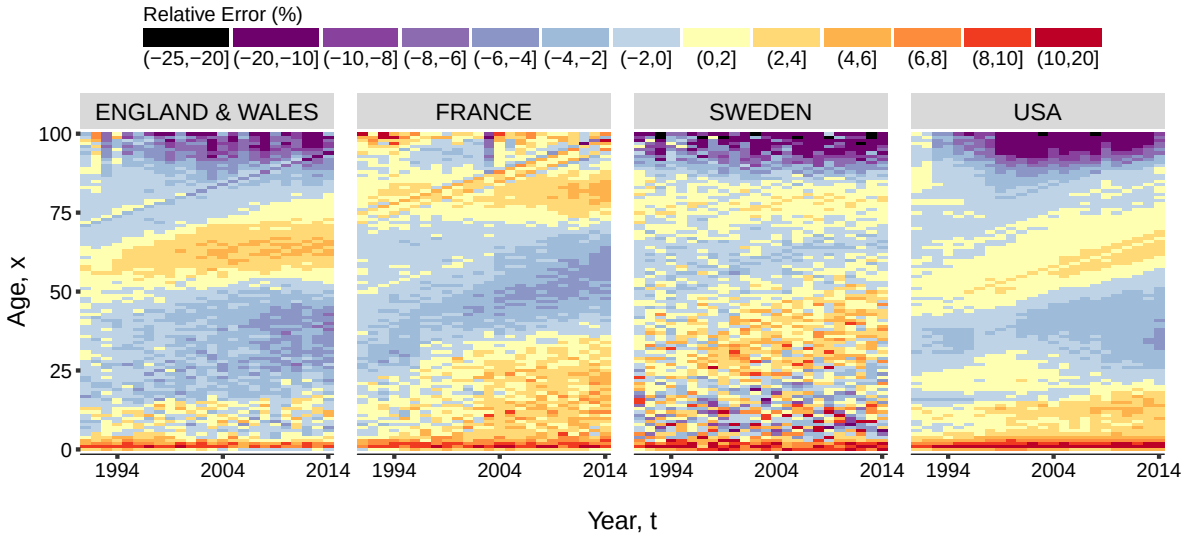


Figure 3.6: *Relative errors (%) of the estimated log-death rates against the actual log-death rates between 1991 and 2014. Computed based on female mortality data in the period 1965–90.*

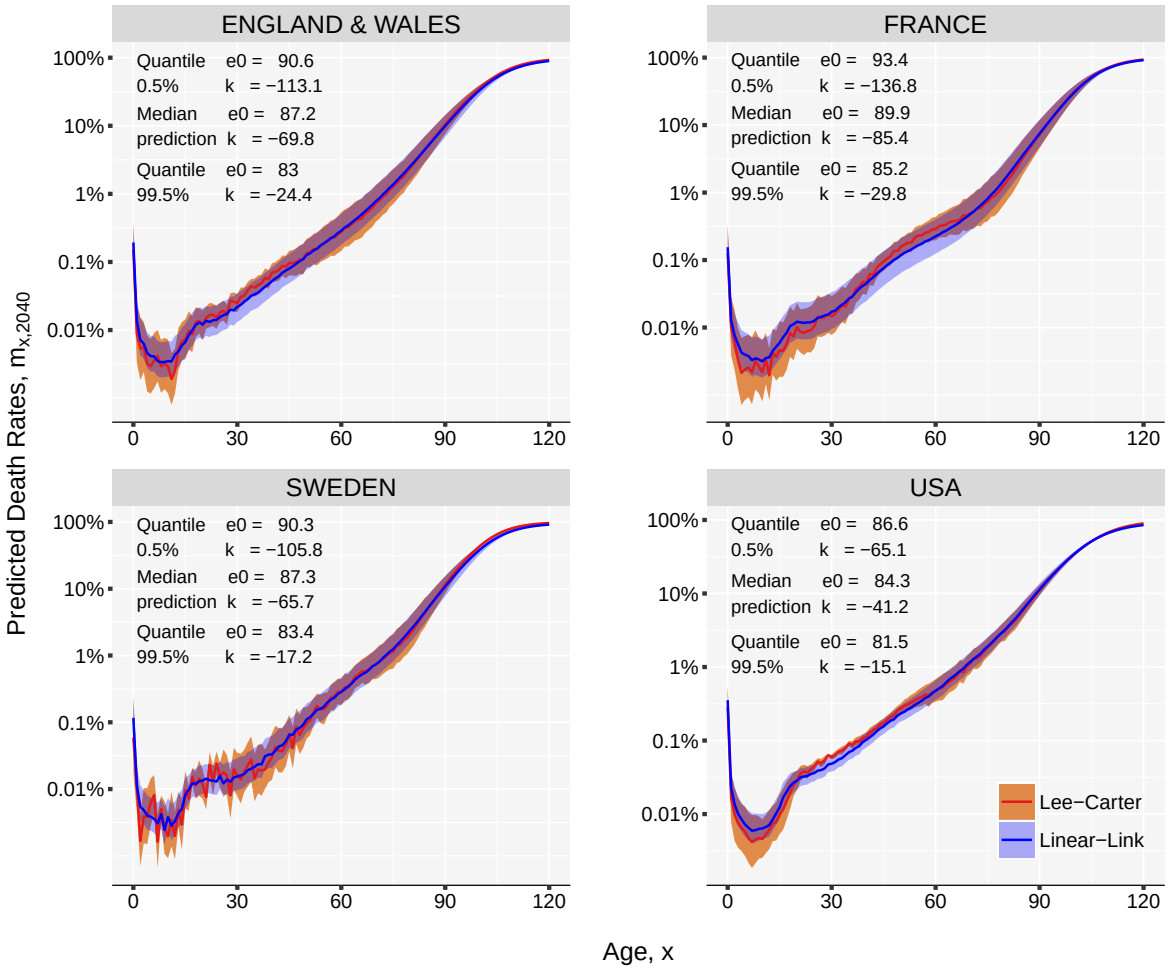


Figure 3.7: *Comparison of the mortality curves predicted by Lee–Carter and Linear-Link models in 2040 from female populations. The models are fitted on the 1980–2014 historical period.*

This value is an average over the entire comparable age range (0–100). The largest impact on the overall accuracy occurs at advanced ages, where the level of uncertainty is higher. Figure 3.6 offers a view of the error distribution by age and time. However, the life expectancy at birth computed based on the estimated death rates matches exactly the actual life expectancy in the respective year and country.

In order to test the conversion reliability of a forecast value of life expectancy, we compare the results generated by the LL model against the predicted mortality from the Lee–Carter model (1992).

The LC model is fitted over the 0–95 age-range using the historical data from 1980 to 2014, and used to forecast death rates 26 years in the future, until 2040. The estimated

matrix of predicted death rates between age 0 and age 95 is extended up to age 120 using the Kannisto model (see equation (3.5) in the Appendix). If multiple projections are simulated for the same forecast point, the LC would produce a range of outcomes that can be translated into life expectancies using standard life table calculations. Any predicted life expectancy given by LC is used as an input value in the LL model to derive the mortality curve, thus obtaining two comparable curves. For every simulated trajectory, the LL method can produce a mortality curve, generating the uncertainty around the median prediction. Figure 3.7 shows that the reconstruction method employed by the LL model gives an almost coincident mortality curve when compared with the LC curve for female populations in 2040. The 99% prediction intervals are computed based on ten thousand Monte-Carlo simulations. Besides the LL model showing a more smooth age-pattern when compared with the LC results, it can perfectly estimate the predetermined life expectancy.

3.4 Discussion

We have introduced a simple method, the Linear-link model, to derive the entire schedule of age-specific death rates, based on a single value of life expectancy and prior knowledge of human mortality patterns. The model is based on the observed linearity between age-specific death rates, m_x, and life expectancy at a certain age, e_θ. The model can be regarded as a decomposition approach of the human mortality curve between the general age pattern, β_x, and an age-specific speed of improvement, ν_x. The method is inspired by: (1) the Log-quadratic model (Wilmoth et al., 2012) in the sense of using a leading indicator in determining the age pattern of mortality; (2) the model introduced by Ševčíková et al. (2016) by adopting an inverse approach to death rates estimation starting from life expectancy; (3) the Lee–Carter model (1992) using the same interpretation of mortality improvement over time and age; and finally (4) the Li et al. (2013) method to model the rotation of age patterns of mortality decline for long-term projections.

The method can be useful in three different situations: future target life expectancy, life tables for countries with deficient data and historical life table construction. The

former is the one explored in the present manuscript, while the latter two are only briefly discussed since their development goes beyond the scope of the paper of presenting the LL model.

First the model can be used in forecasting practice when the level of life expectancy is forecast first. We showed that this model can accurately reconstruct a Lee–Carter forecast starting from a single value of life expectancy at birth. This is important, because the Linear-link model offers the possibility of taking advantage of the more regular pattern of the life expectancy evolution. It is much easier and parsimonious, from a technical perspective to forecast one time series of expectation of life than to extrapolate 100 or 110 series of death probabilities corresponding to each age group. In the same manner adult mortality can be estimated based on a value of life expectancy at an advanced age, say age 65. In Figure 3.2, we showed that the linearity between death rates at advance ages and life expectancy at age 65 on a log scale is maintained. A greater variation is observed only at advanced ages, above 100, where data is sparse in general.

Second, the method can be used to build model life tables and estimate the current age patterns of mortality in poor-data countries or regions, like Sub-Saharan Africa. In this case the parameters of the model are estimated based on a collection of historical life tables from several regions or populations. Once the parameters have been estimated, and implicitly the model life table, they remain fixed. The relevant mortality curve is simply calibrated in accordance with a single value of life expectancy at birth or any other age instead of child mortality like in the case of Wilmoth et al. (2012). In our analysis we show examples using high quality data from developed countries in order to demonstrate the efficiency of the model, and to be able to asses the accuracy of the mortality curve reconstruction. However, the estimation procedure and the steps of the algorithm are the same for this case too.

Third, the LL model can be a useful tool in a variety of research contexts of historical demography like backward projections and estimation of mortality levels in historical populations. Due to the existence of scarce non-standardized population data in the past and population census only for the more recent times, the very possibility of projecting mortality backward is of theoretical interest (Ediev, 2011).

According to our analysis, the optimal number of years to be used in the fitting of the model is between 30 and 35 years. If a longer time interval was used, the parameter estimates would lose their relevance. For example, the present rate of improvement in the death rates is different from that experienced 50 years ago, because of fundamental changes in society and scientific advances during this period (Bengtsson, 2006; Rau et al., 2008). In the same manner over a longer period of time the linearity between life expectancy and death rates might be challenged, however this should be investigated from case to case.

The speed of improvement in age-specific death rates over ages changes over time. For example, in recent decades, a faster pace of improvement was observed at ages 65 and above (Vaupel, 1997; Shkolnikov et al., 2011). We address the possibility of experiencing accelerating or decelerating speeds of mortality improvements over different age ranges by assigning different weights to the estimated ν_x curve when the life expectancy at birth continues to advance over age 75. The effect of this method can be best observed in Figure 3.7 in the case of France. Under the implicit assumption of constant mortality improvements, the LC forecast generates a second mortality hump around age 50. The estimated mortality curve given by the LL model has a less pronounced effect due to the rotated ν_x parameter. See a detailed description of the method in Appendix 3.5.3.

The evolution of human mortality is a complex process that is driven by a large number of factors and can not be explained by a single statistical model. The Linear-link method offers an alternative approach to deriving the unknown levels of mortality in the future. In contrast with methods like the Lee–Carter model that extrapolate age-specific rates or probabilities directly the method presented here recognizes life expectancy as the main driver of mortality at any given age and employs an indirect estimation algorithm. These methods can complement each other and help us understand better the future longevity experienced by populations.

3.5 Appendix

3.5.1 The Kannisto Model

Normally, mortality data is available in tables that contain detailed information up to age 85, 100 or 110, with last age group being open. In order to extend the mortality rates up to age 120, the Kannisto method (Thatcher et al., 1998) for old-age mortality with an asymptote equal to one can be employed:

$$m_x = \frac{\alpha e^{\beta x}}{1 + \alpha e^{\beta x}}, \tag{3.5}$$

which can also be written as a linear function of age

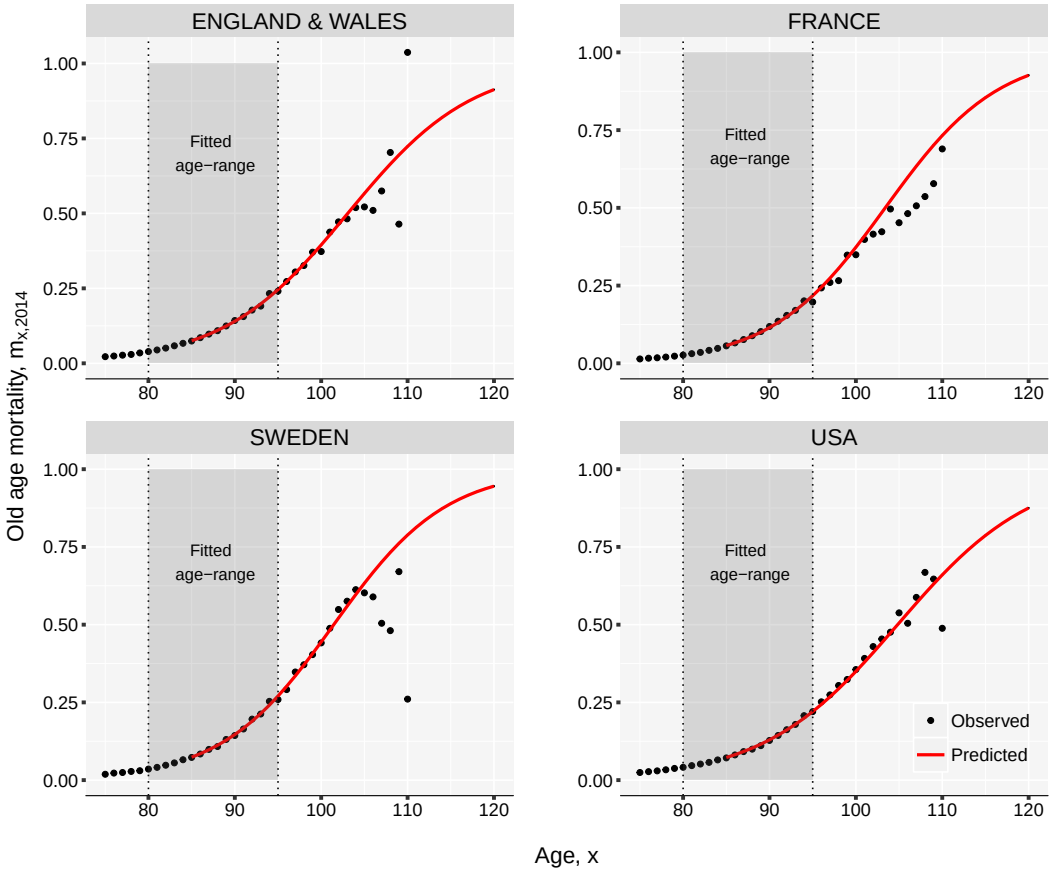


Figure 3.8: *Extension of female mortality rates using the Kannisto model in 2014.*

$$logit(m_x) = ln(\alpha) + \beta\chi + \epsilon_\chi, \tag{3.6}$$

where ϵ is a normally distributed variable with mean zero, $\chi = x - 80$, and parameters α and β are positive real numbers. The model is usually fitted between age 80 and 95.

Assuming that $D_x \backsim Poisson(E_x \cdot m_x(\alpha, \beta))$ the parameters α and β can be derived by maximizing the log-likelihood function:

$$logL(\alpha, \beta) = \sum_{x=80}^{95} \Big\{ D_x log\big[m_x(\alpha, \beta)\big] - E_x m_x(\alpha, \beta) \Big\} + constant, \tag{3.7}$$

where D_x denotes the number of deaths that occurred at age x, E_x represents the population exposed to risk at the same age, and m_x is the age-specific death rate.

The Kannisto model is not only useful to obtain values for oldest-old mortality but also to smooth the rates computed on smaller sample sizes. The case of Sweden presented in Figure 3.8 can be relevant here, where in 2014 the number of females aged 100 and above was less than 1,700. A small sample size can create difficulties in obtaining reliable mortality estimates based only on empirical observations. Outliers are expected to show up from year to year.

3.5.2 Maximum likelihood estimation

Assuming that deaths are Poisson distributed, the LL model can be fitted by maximising the log-likelihood given by

$$logL(a, \nu, k) = \sum_{x,t} \Big\{ D_{x,t}(a_x + \nu_x k) - E^c_{x,t} \exp(a_x + \nu_x k) \Big\} + constant, \tag{3.8}$$

where $a_x = \beta_x \log e_{\theta,t}$. The parameters are estimated following an updating scheme proposed by Brouhns et al. (2002) based on the Newton-Raphson algorithm. The updating procedure, with initial values $\hat{a}_x(0) = 0, \hat{\nu}_x(0) = 1$, and $\hat{k}(0) = 0$, is as follows:

$$
\begin{aligned}
\hat{a}_x(w+1) &= \hat{a}_x(w) - \frac{\sum_t (D_{x,t} - \hat{D}_{x,t}(w))}{-\sum_t \hat{D}_{x,t}(w)}, \\
\hat{\nu}_x(w+1) &= \hat{\nu}_x(w), \\
\hat{k}(w+1) &= \hat{k}(w), \\
\hat{k}(w+2) &= \hat{k}(w+1) - \frac{\sum_x (D_{x,t} - \hat{D}_{x,t}(w+1))\hat{\nu}_x(w+1)}{-\sum_x \hat{D}_{x,t}(w)(\hat{\nu}_x(w+1))^2}, \\
\hat{a}_x(w+2) &= \hat{a}_x(w+1), \\
\hat{\nu}_x(w+2) &= \hat{\nu}_x(w+1), \\
\hat{\nu}_x(w+3) &= \hat{\nu}_x(w+2) - \frac{\sum_t (D_{x,t} - \hat{D}_{x,t}(w+2))\hat{k}(w+2)}{-\sum_t \hat{D}_{x,t}(w+2)(\hat{k}(w+2))^2}, \\
\hat{a}_x(w+3) &= \hat{a}_x(w+2), \\
\hat{k}(w+3) &= \hat{k}(w+2),
\end{aligned}
$$

where $\hat{D}_{x,t}(w) = E^c_{x,t} \exp(\hat{a}_x(w) + \hat{\nu}_x(w)\hat{k}(w))$, is the estimated number of deaths after iteration w. After the parameters are estimated, the parameter a_x is transformed using the LL model, $a_x = \beta_x \log e_{\theta,t}$.

The maximum likelihood estimation (MLE) has several advantages over least squares (OLS) and SVD methods or even weighted least squares (WLS) used in Wilmoth et al. (2007). Several reasons have been given in the literature (Brouhns et al., 2002; Alho, 2000). One example would be the increasing confidence intervals by age. This is because in the OLS estimation via SVD the errors are assumed to be homoskedastic and normally distributed, which is quite a heavy assumption. The logarithm of the observed force of mortality is much more variable at older ages than at younger ages because of the much smaller absolute number of deaths at older ages. Therefore, since the number of deaths is a counting variable, the Poisson assumption seems more reasonable (Brillinger, 1986).

However, in order to use this approach we need death counts $D_{x,t}$ and exposures $E_{x,t}$, which are not always available. This being the reason why the model described by equation (3.1) is chosen in the article. Our methodology is targeting populations with deficient data as well as populations whose estimates of mortality rates are provided without dis-

aggregation by deaths and exposures. Although conceptually a Poisson setting would be better, the SVD approach is a pragmatic decision for practical reasons. As shown in figure 3.9 the difference between the two estimation methods for the case of England & Wales females is very small but the data requirements is higher in the case of MLE.

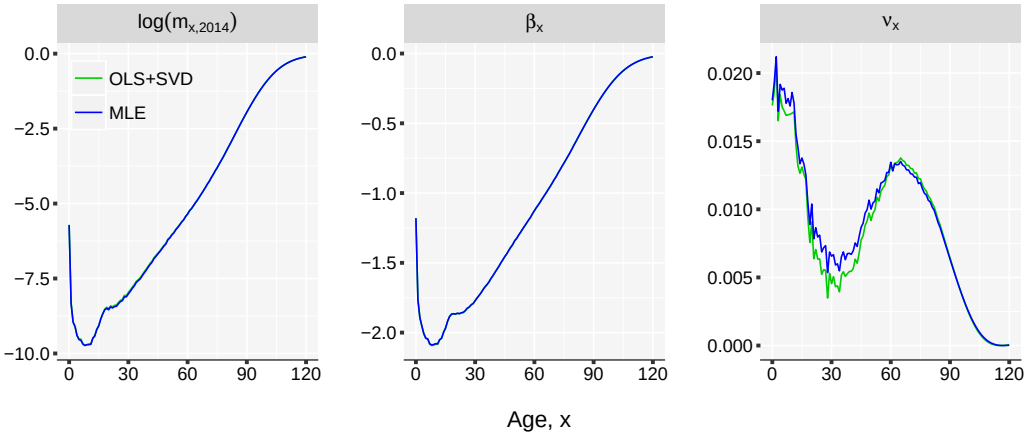


Figure 3.9: *Comparison of the fitted mortality curves and parameter estimates of the Linear-Link model using the OLS+SVD and MLE fitting procedures. England & Wales female data for 1980–2014 period is used.*

3.5.3 Rotation of mortality improvements

One of the main limitations of the LC model (1992) is the central assumption of constant rates of mortality declines at different ages, resulting from the time-invariant b_x coefficient of age-specific mortality improvements (Bongaarts, 2005). The assumption has been violated in several low-mortality countries in recent decades, because rates of mortality improvements have tended to decline over time at younger ages, and they have risen at older ages (Kannisto et al., 1994; Vaupel et al., 1998; Wilmoth and Horiuchi, 1999).

It is important to take into consideration the changing age pattern of mortality improvements to produce more accurate mortality forecasts, and projection methodologies that ignore such rotation will lead to errors, particularly in the projected age patterns of future death rates (Li et al., 2013). Li et al. proposed an extension of the LC method to incorporate the rotation of the age patterns of mortality decline for long-term projections.

Here, we propose a modification of the original Li et al. (2013) methodology that ensures the rotation of the rate of mortality improvement over age in the LL model, ν_x. The methodology is composed of two different steps.

First, we derive an ultimate schedule of mortality improvements, ν_x^u, from the estimated coefficient ν_x. In particular, the ultimate rates of improvement between ages 0 and 65 are set equal to the average improvement at adolescent and adult ages (15–65); from age 65

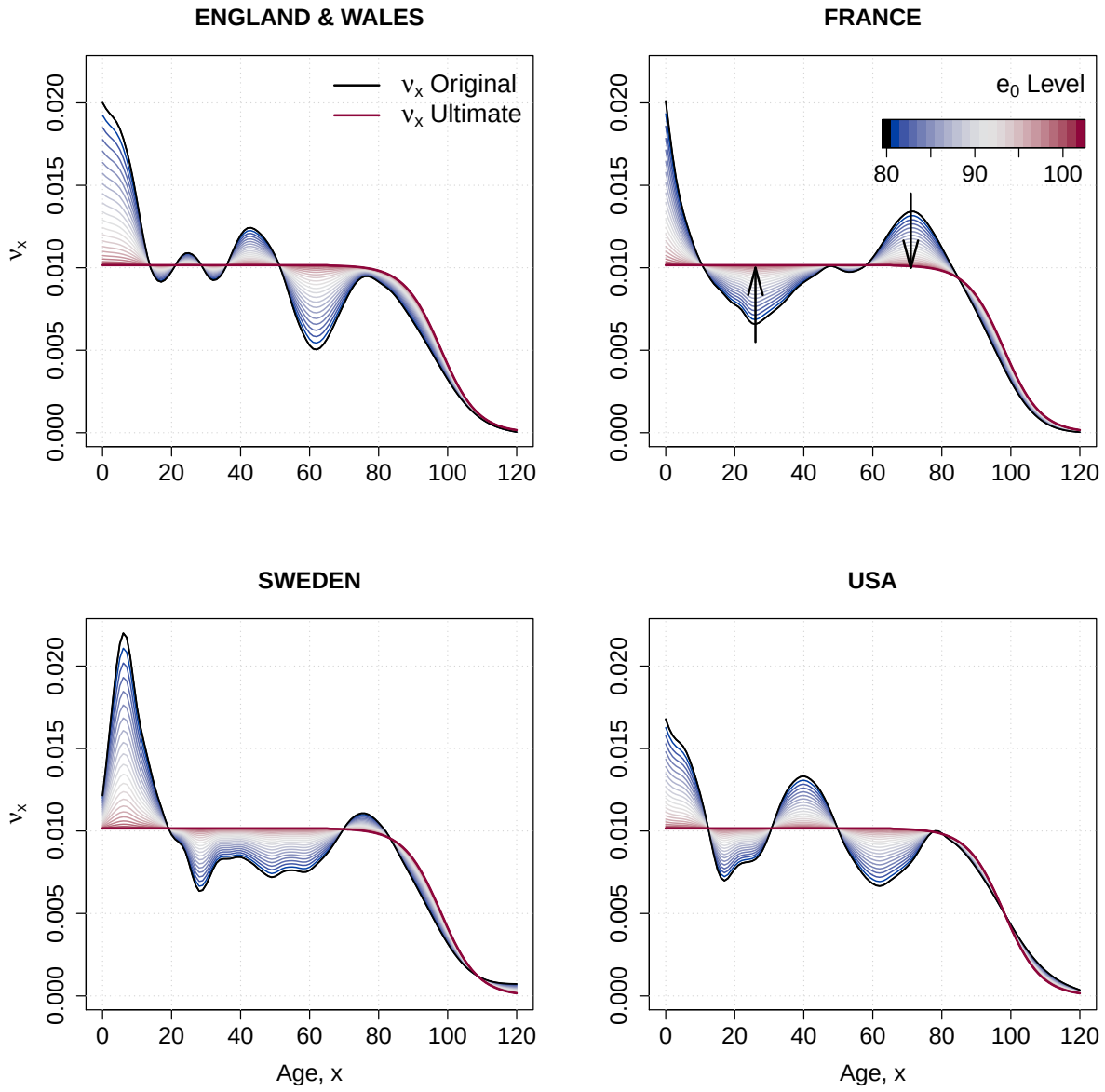


Figure 3.10: *Assumption of the change in ν_x pattern following the increase in life expectancy at birth from 75 to 102 years.*

onwards, improvements decrease following a logistic shape, and they converge to zero at age 130.

Second, we smooth the transition from ν_x to ν_x^u using the weight function proposed by Li et al. (2013). The transition, and therefore the degree of rotation of ν_x, is dependent on $e_0^*(\tau)$, the predicted value of life expectancy at birth (an input in our LL model). Formally, the weight function w_s can be expressed as:

$$w_s(\tau) = \left\{ \frac{1}{2} \left[1 + \sin \left[\frac{\pi}{2} \left(2w(\tau) - 1 \right) \right] \right] \right\}^p \quad \text{with} \quad w(t) = \frac{e_0^*(\tau) - 80}{e_0^u - 80} \,. \tag{3.9}$$

Because ν_x parameter is scaled in order to take values between 0 and 1 the ultimate pattern of mortality improvement by age will be the same for all countries. If the scaling process is ignored, the same pattern is obtained in all cases with a different level of ν_x between age 0 and 65. The estimated death rates would be the same in both case because of the adjustment provided by k parameter.

The power of the smooth-weight function, p, regulates the speed of the rotation. It varies between 0 and 1, and lower values correspond to faster rotations for levels of $e_0^*(\tau)$ closer to 80. The level of life expectancy at which the rotation finishes, e_0^u, is also arbitrary; here, we follow the recommendations of Li et al. (2013) and set the intermediate value of 0.5 for p and the age 102 for e_0^u.

The rotated coefficient of mortality improvement over age, denoted $N_x^r(\tau)$, can thus be written as:

$$N_x^r(\tau) = \begin{cases} \nu_x \,, & e_0^*(\tau) < 80 \,, \\ \left[1 - w_s(\tau) \right] \nu_x + w_s(\tau) \nu_x^u \,, & 80 \leq e_0^*(\tau) < e_0^u \,, \\ \nu_x^u \,, & e_0^*(\tau) \geq e_0^u \,. \end{cases} \tag{3.10}$$

Chapter 4

Forecasting the Age-at-Death Distribution

Marius D. Pascariu
Adam Lenart
Vladimir Canudas-Romo

Corresponding Author: Marius D. Pascariu
Institute of Public Health
University of Southern Denmark
J.B. Winslows Vej 9B, 5000 Odense, Denmark
E-mail: mpascariu@health.sdu.dk

The Maximum Entropy Mortality Model

Forecasting the Age-at-Death Distribution

Marius D. Pascariu

Institute of Public Health, University of Southern Denmark, Odense, Denmark

Adam Lenart

Novo Nordisk, Copenhagen, Denmark
Institute of Public Health, University of Southern Denmark, Odense, Denmark

Vladimir Canudas-Romo

School of Demography, The Australian National University, Canberra, Australia

Abstract

The age-at-death distribution is a representation of the mortality experience in a population. Although it proves to be highly informative it is often neglected when it comes to the practice of past or future mortality assessment. We propose an innovative method to mortality modelling and forecasting by making use of the location and shape measures of a density function i.e. statistical moments. Time series methods for extrapolating a limited number of moments are used and then the reconstruction of the future age-at-death distribution is performed. The accuracy of the estimated distributions proves to be very good and the predictive power of the method seems to be net superior when compared to the results obtained using classical approaches to extrapolating age specific death rates. The method is tested using data from the Human Mortality Database and implemented in a publicly available `R` package.

Keywords:

Mortality forecasting; Density estimation; Statistical moments; Maximum entropy

4.1 Introduction

The mortality experience of a population is well described by its hazard rates, a property that has been well exploited in numerous methods of mortality modelling (Gompertz, 1825; Makeham, 1867; Siler, 1983; Heligman and Pollard, 1980) and forecasting (Lee and Carter, 1992; Li and Lee, 2005; Haberman et al., 2014). The main reasons why hazard rates have been used predominantly in modelling and forecasting is that they readily represent the change in the risk of death over age and time. In addition the five components of the pattern of human mortality (infant, child, youth, adult, and old-age mortality) can be identified clearly in a graphical representation of the hazard curve; and a variety of time series model can be employed to extrapolate the identified trends over time.

Surprisingly few mortality methods acknowledge that the probability density function of the distribution of deaths can be equally informative when compared to the hazard indices: and more than that it can give immediate indication on key measure of longevity like how long a population lives on average and the degree of variability of ages at death. By employing statistical knowledge about the shape of the distribution, and how the level of mortality at a certain age is fully dependent on the levels of mortality at all the other ages, brings enormous advantages. When hazard rates are investigated no conclusion can be made about mean age-at death, or the inequality experienced by the population when it comes to death or what is the prevalence of extraordinary long life-spans (the outliers in old-age mortality) or even summary statistics related to the mortality experience. For example, Figure 4.1 illustrates the transition in the age-at-death distribution for men living in England & Wales. We can learn that in 1960 the population experienced a pronounced infant mortality level up to the age of 5, and the fact that it takes about 53 years in order for the first 10% of men to die. In 1960, only 10% of the male population had the chance of living beyond the age of 85. Analysing the same distribution in 2016 we can see that infant mortality dropped to almost insignificant levels (an inspection on the logarithmic scale would still show differences over ages), and that it takes about 63 years to eliminate the first 10% of the male population through death. We can also learn

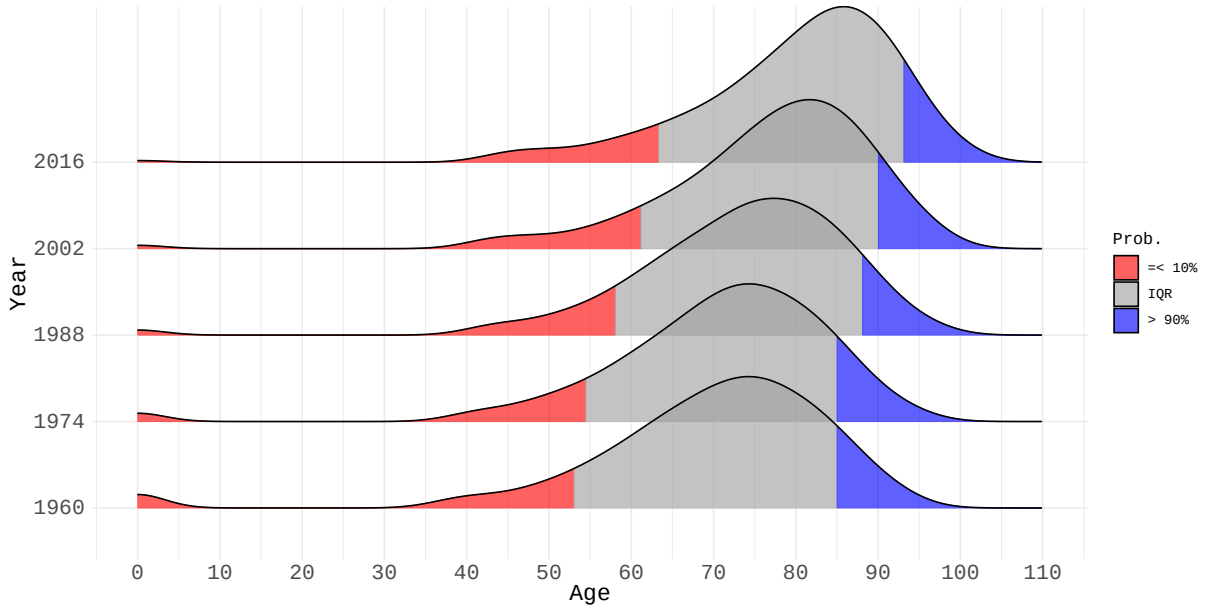


Figure 4.1: *Convergence of the age–at–death distribution (England & Wales, Male population)*

that more that 50% of men are surviving to age 85, and more than 10% of the population will surpass the age of 93.

The first attempt to describe the mortality pattern by analysing the age-at-distribution was made by Pearson (1897), building on Lexis' work (1879), where different functions were being used to capture the components of the pattern of human mortality. Pearson used a skewed function, arguing that the skewness of old-age mortality depends on the incidence of premature mortality. More recently Dellaportas et al. (2001) uses death counts to fit the Heligman-Pollard model with Bayesian methods and Mazzuco et al. (2018) models mortality by fitting a half-normal and a skew-bimodal-normal distribution to the observed empirical age-at-death density function.

Despite being well suited to portray the dynamics of mortality patterns and to study longevity and lifespan variability, age-at-death distributions have generally been neglected in forecasting practice. The life table distribution of deaths, or $f(x)$[1], is constrained ensuring that its elements add to the radix of the population, usually $l_0 = 1$, thus having $\sum_x f(x) = 1$. Time-series extrapolations of its trends are likely to violate this assumption

[1]The commonly accepted life-table notation for the age-at-death distribution is $d(x)$. However we will use the $f(x)$ notation in order to maintain consistency with the notation used in statistics for densities.

making burdensome the use of the same extrapolative methods as in the case of hazard rates. Notable attempts at forecasting the age-at-death distribution were made by Oeppen (2008) and Bergeron-Boucher et al. (2017) by adapting the Lee–Carter model to the Compositional Data Analysis framework (Aitchison, 1982). While this approach solves the problem of respecting the unit sum constraint, it also requires changing the coordinate system from a Euclidean space to an Aitchison simplex (1986) which might hinder the interpretation of the results. Another novel idea is introduced by Basellini and Camarda (2018) by modelling the shifting and compression dynamics of the adult mortality distribution around the modal age at death with a 3-parameter function, parameters that can be extrapolated using standard time series models.

Inspire by the idea of forecasting mortality given the information offered by the age-at-death distribution, we propose a novel approach to forecasting age-specific mortality levels by making use of statistical moments i.e. the shape measures of a density. By employing the knowledge about the shape of the distribution mentioned earlier, and how the level of mortality at a certain age is fully dependent on the levels of mortality at all the other ages, brings an enormous advantage to the extrapolation methods based on death frequencies over the methods based on hazard or mortality rates extrapolation. One statistical moment describes one characteristic of the distribution it belongs to. For example, in the case of the distribution of deaths of a population the first moments correspond to: i) the mean age at death or life expectancy; ii) the variance offers information about the inequality of the age at death; iii) the skewness states how concentrated the mortality is around young or old ages, and iv) the kurtosis indicates the weight of the tails or the presence of outliers (extreme old age) in the distribution. Beyond moments higher than the fourth the interpretation is limited, however these are relevant in the case of more complex distribution (multi-modal densities) helping in fine-tuning the observed irregularities. Thus, for a distribution, the collection of all the moments uniquely determines its density function. In order to gain a perfect understanding of the underlying density function one needs to have information about all the moments up to infinity. However a limited number of moments like mean, variance, skewness and kurtosis can already offer a good approximation of the shape of the probability density function of the underly-

ing distribution of deaths, and therefore a good understanding of the levels of mortality experienced by a population at different ages.

The method proposed here considers the *finite moment problem* where a positive density, $f(x)$, is sought from knowledge of a limited number of its power moments. We assess the evolution of several observed moments of the age-at-death distribution in order to forecast them by employing multivariate time series models. And we reconstruct the forecast distribution using the maximum entropy approach (Mead and Papanicolaou, 1984) that relies on the average rate at which information is produced by a stochastic source of data or density function i.e. *information entropy*. Reconstructing the density function from a set of predicted moments has the advantage of allowing accelerating/decelerating rates of mortality improvement over age and time. It also eliminates the necessity of altering the coordinate system as in Oeppen (2008). We will refer to this method as the *maximum entropy mortality* model (MEM).

The purpose of this study is:

- to demonstrate that accurate forecasts of age-specific mortality levels can be obtained using statistical moments and the information provided by the age-at-death distribution;
- to compare the MEM against other well established mortality models and determine its newly added value.
- to validate the MEM predictions against a benchmark, that is, a simplistic trend extrapolation of the age specific death-rates (naïve model) in order to justify the increase in complexity of the proposed method. This objective is justified and inspired by Bohk-Ewald et al. (2018), a study where 20 major fertility forecasting methods are evaluated. The main findings show that across multiple measures of fertility forecast accuracy only 4 methods consistently outperform the naïve model.

4.2 Methods

In the current section we introduce our method of modelling and forecasting mortality and the main concepts required to understand its estimation procedure.

4.2.1 Statistical Moments

The statistical moments are defined as the expected value of the n-th power of a random variable x. The n-th moment, μ_n, for a continuous density function, $f(x)$, about a value c can be defined as:

$$\mu_n = \int_a^{\omega} (x-c)^n f(x) dx, \quad \text{where} \quad n = 0, 1, 2 \ldots \tag{4.1}$$

If variable c denotes the mean it is said that μ_n is the n-th moment about the mean or the n-th *central moment*. The moments can be computed about zero as well, in which case the moment is called a *raw* or *crude moment*. The *normalized moment* of a probability distribution is a central moment that is standardized. The normalization is typically a division by an expression of the standard deviation, σ, which renders the moment scale invariant. This has the advantage that such normalized moments differ only in other properties than variability, facilitating e.g. comparison of the shape of different probability distributions. Then the normalized moment of degree n of a central moment is

$$\tilde{\mu}_n = \mu_n / \sigma^n. \tag{4.2}$$

Both, the raw and the normalized moments are used in this study. While the normalized moments are a good choice to serve as indices in time series extrapolation (as we will see in section 4.2.4), the raw moments are more efficient in the estimation of the underlying density functions (section 4.2.3). If one type of moments is known, all the others can be derived from these without losing their properties.

4.2.2 Information Entropy

The concept of information entropy was introduced by Shannon (1948) in an effort to mathematically formalize the process of communication in computer science (between two devices). The word *information*, here, is used in a special sense and must not be confused with *meaning*. The information is seen as a measure of the numbers of the possible outcomes generate by a probabilistic process and it is defined as the logarithm of this value. Warren (1949) reveals that the quantity which uniquely meets the natural requirements that one sets up for *information* turns out to be exactly that which is known in thermodynamics as *entropy* and which is a measure of the degree of randomness.

To explore the numerical values of the entropy measure H, let's assume that the individuals in a population can die in one of the following two states: childhood (C) and adulthood (A). The associate probabilities would be $f(C)$ for the first state and $f(A) = 1 - f(C)$ for the second. It follows that:

$$H = -\left[f(C)\log f(C) + f(A)\log f(A)\right] \quad \text{or} \quad H = -\sum_{i=C}^{A} f(i)\log f(i) \tag{4.3}$$

Since the logarithm of a number smaller that 1 is a negative value, the minus sign in the equation is added for convenience only, so that H is always positive. It turns out that the information entropy has its largest value, when the events of dying in childhood and adulthood are equally probable; that is when $f(C) = f(A) = 0.5$. Just as soon as one outcome becomes more probable than the other (e.g. adult mortality greater than childhood mortality) the value of H decreases. And when one outcome is very probable ($f(A)$ almost one and $f(C)$ almost zero, say) the value of H is approaching zero. A situation in which the stochastic process of dying in different states becomes less random and the outcome almost certain. Therefore, information entropy is also a measure that can show the level of inequality experienced by individuals represented in a distribution such as age-at-death. A small value of the entropy indicates that everyone dies around the same ages, i.e. the population is characterized by a small degree of inequality in terms of age-at-death. If H is large one can say that the inequality is pronounced.

To generalize, the entropy of a random variable x with probability distribution function $f(x)$ is the negative logarithm of the density function for the value, and can be written as:

$$H = -\int f(x) \log_b f(x) dx, \tag{4.4}$$

and

$$H = E\left[I(x)\right] = E\left[-\log_b f(x)\right], \tag{4.5}$$

where E is the expected value operator, I is the information content of x and b is the base of the logarithms used. The entropy can be measured in binary units (*bits*) natural units (*nats*) or decimal units (*bann*) for b equal to 2, e and 10 respectively. In the case of $f(x_i) = 0$ for some i, the value of $log_b(0)$ is taken to be 0, which is consistent with the limit:

$$\lim_{p \to 0+} p \log p = 0. \tag{4.6}$$

Various entropy measures have been proposed in the scientific literature including the entropy of a life table (Keyfitz, 1977) commonly used in demography. We note that Keyfitz's measure can not be an alternative to the Shannon entropy in this study. Keyfitz's entropy is defined as measure of elasticity of the life expectancy with respect to a uniform change in age-specific death rates, and it does not represent a true measure of entropy in the probability sense (Hill, 1993).

4.2.3 The finite moment problem

The problem of reconstructing a distribution from a given number of moments is not straightforward. It is known in the mathematical literature as the *finite moment problem*. The method has been extensively studied from a theoretical perspective and has practical applications in thermodynamics and quantum-physics. It can be regarded as a finite dimensional version of the Hausdorff moment problem (Hausdorff, 1921; Shohat

and Tamarkin, 1943). Various methods for solving this problem have been proposed in the last decades, by making use of orthogonal polynomials (Chihara, 2011), splines (John et al., 2007), or other numerical strategies (Frontini et al., 1990). All the procedures aim at constructing specific sequences of functions $f_N(x)$ which eventually converge to the true distribution $f(x)$ as the number of moments N, used in estimation, approaches infinity

$$\mu_n = \int_a^\omega x^n f_N(x) dx, \quad n = 0, 1, 2 \ldots, N. \tag{4.7}$$

Equation (4.7) should be seen as a system of $N + 1$ equations, where the moments $\mu_0, \ldots, \mu_N$ come from the $f_N(x)$ density.

Taking advantage of the regularity of human mortality the reconstruction of a density function can be realised using a small number of moments, usually 3 to 6. And, a good fit of the true density is achieved by imposing a prior restriction of the class on functions where the solution is sought.

Here we follow the maximum entropy reconstruction (*MaxEnt*) and the algorithm developed by Mead and Papanicolaou (1984) as a definite procedure for the construction of a sequence of approximations to the true density. This method is based on the information entropy given by the density function. As a strategy for finding the local maxima of the entropy functional $\mathcal{L} = \mathcal{L}(f)$, we employ the method of Lagrange multipliers, λ_n for the n-th moment:

$$\mathcal{L} = H + \sum_{n=0}^{N} \lambda_n \left[\hat{\mu}_n - \mu_n\right]. \tag{4.8}$$

The entropy is maximized under the condition that the first $N+1$ moments, $\hat{\mu}_n$, are equal to the true moments μ_n, where n takes values between 0 and N. Functional variation of $\mathcal{L}$ with respect to the unknown density $f(x)$ yields

$$\frac{\delta \mathcal{L}}{\delta f(x)} = 0 \Longrightarrow f = f_N(x) = \exp\left(-\lambda_0 - \sum_{n=1}^{N} \lambda_n x^n\right), \tag{4.9}$$

and the $n-$th raw moment

$$\mu_n = \int_a^\omega x^n \exp\left(-\lambda_n - \sum_{n=1}^{N} \lambda_n x^n\right) dx. \tag{4.10}$$

Considering the availability of the first $N+1$ moments, the equations (4.9) and (4.10) (that are closely related to equation 4.7) should be viewed as a non-linear system of $N+1$ equations for the unknown Lagrange multipliers $\lambda_0, \lambda_1, \ldots, \lambda_N$. If we assume that the density $f(x)$ is normalized such that the first moment is always equal to 1 ($\mu_0 = 1$, i.e. respecting the unit sum constraint) the first equation in (4.10) then reads

$$\mu_0 = \int_a^\omega x^0 f_N(x) dx = \int_a^\omega \exp\left(-\lambda_0 - \sum_{n=1}^{N} \lambda_n x^n\right) dx = 1 \tag{4.11}$$

and results in the first Lagrange multiplier, λ_0, being expressed in terms of the remaining Lagrange multipliers:

$$\int_a^\omega \exp\left(-\sum_{n=1}^{N} \lambda_n x^n\right) = e^{\lambda_0}. \tag{4.12}$$

The system of equations then reduces to

$$\mu_n = \frac{\int_a^\omega x^n \exp\left(-\sum_{n=1}^{N} \lambda_n x^n\right)}{\int_a^\omega \exp\left(-\sum_{n=1}^{N} \lambda_n x^n\right)}, \quad n = 0, 1, 2 \ldots, N. \tag{4.13}$$

For a numerical solution, one introduces $\Gamma = \Gamma(\lambda_1, \lambda_2, \ldots, \lambda_N)$ through the Legendre transformation

$$\Gamma = \ln(e^{\lambda_0}) + \sum_{n=1}^{N} \mu_n \lambda_n, \tag{4.14}$$

where the μ_n's are the observed numerical values of the known moments. The stationary points of the potential Γ are solutions to the equations

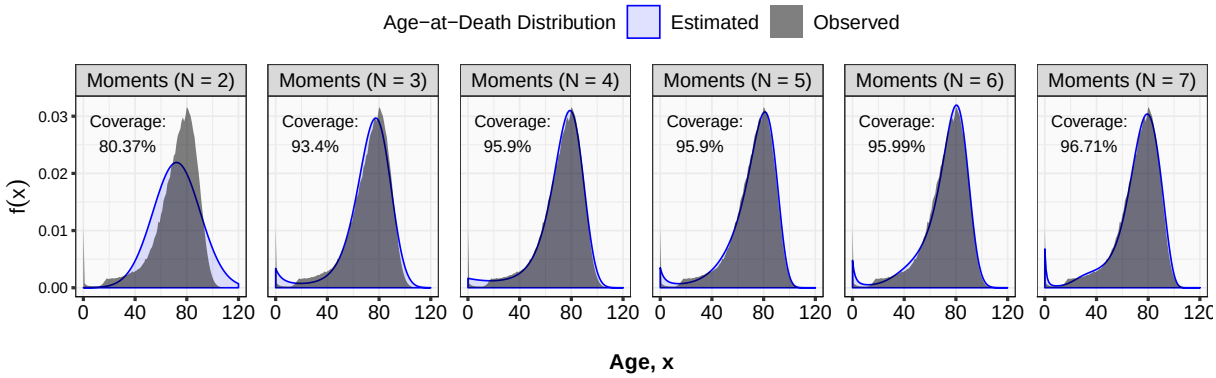


Figure 4.2: *Observed and estimated empirical density functions (USA, 1990, Male population)*

$$\frac{\delta\Gamma}{\delta\lambda_n} = 0 \Longrightarrow \mu_n, \quad n = 0, 1, 2 \ldots, N, \tag{4.15}$$

which is the solution to the finite moment problem. The convexity of Γ guarantees that if a stationary point is found for some finite values of $\lambda_1, \lambda_2, \ldots, \lambda_N$ it must be a unique absolute minimum. A more detailed description and analytic demonstration of the method and also alternative algorithms for solving the finite moment problem can be found in Mead and Papanicolaou (1984).

Figure 4.2 shows the observed and reconstructed distribution of deaths using different numbers of observed statistical moments for the male population living in the USA in 1990. This distribution of deaths is a good study case because it is characterized by a pronounced level of mortality in the first year of life, an accident hump around age 20 and an exponential increase in adult and old age mortality. More recent distributions exhibit less pronounced local maxima or modes, therefore making them easier to estimate. Knowing only the first two moments, mean and variance, is not sufficient to obtain a good reconstruction of the underlying distribution. The obtained coverage, i.e. the common surface of the observed and estimated distributions, would be around 80%. However, the more moments we employ in the estimation procedure the bigger the coverage becomes. The results indicate that of 6 moments are enough to obtain a coverage above 96% where the infant and adult mortality is captured adequately. We note here that above a large enough coverage level, the measure does not necessarily indicate a more accurate ap-

proximation of the true distribution, but better identification of the main body of the distribution.

4.2.4 The Maximum Entropy Mortality model

The idea behind our forecasting method is simple. The future age-specific levels of mortality for a population are determined by extrapolating a limited number of statistical moments given by the available life-table age-at-death distributions. The extrapolation is realised with multivariate time series models. The age-at-death distribution is estimate at any point in time from the predicted moments using the $MaxEnt$ algorithm (introduced in section 4.2.3).

Prior to generating future realisations, the moments of ordinal 3 and higher are normalized, and the logarithmic transformation is applied to the absolute values of all observed moments. This ensures that the relevant shape measures remain positive on any forecasting horizon (e.g. the mean and the variance of the distribution). The period index of interest to be used in forecasting, with first order differences, can be defined as follows:

$$y_{n,t} = \log|\tilde{\mu}_{n,t}| - \log|\tilde{\mu}_{n,t-1}|. \tag{4.16}$$

We are using a multivariate random–walk, with a vector of drift parameters θ_n (see Appendix 4.5.1 for a detailed description), to drive the dynamics of the multiple period indices, so that

$$y_{n,t} = \theta_n + y_{n,t-1} + \varepsilon_{n,t} \quad \text{with} \quad t = 1, 2, ..., \tau \quad \text{and} \quad n = 1, 2, ..., N, \tag{4.17}$$

where $\varepsilon_{n,t} \sim \mathbf{N}(0, \Omega)$, with $\Omega = CC'$. C represents the Cholesky factorisation matrix of the variance-covariance matrix Ω. The parameters θ_n and Ω are estimated by ordinary least squares (OLS). A similar model is used by Haberman and Renshaw (2011) to generate trajectories of the multiple period indices in various mortality models.

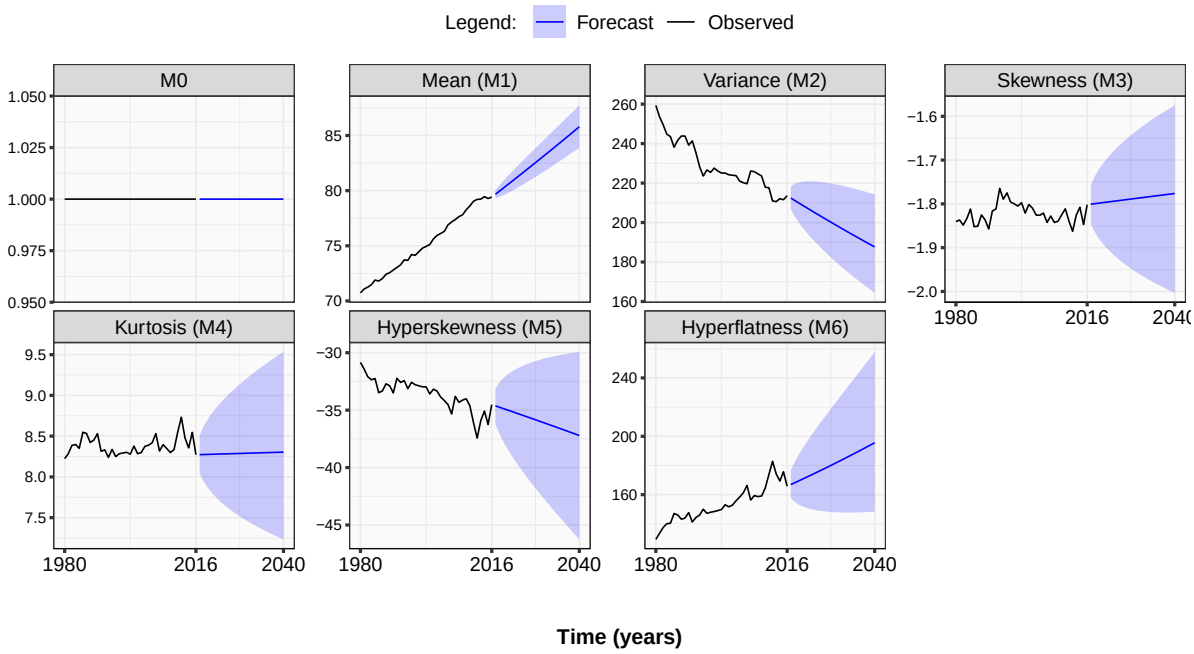


Figure 4.3: *Forecast of statistical moments of the life table distribution of deaths together with 95% prediction intervals, using a multivariate random–walk model (England &Wales, Male population, 1980–2040)*

Once the $y_{n,t}$ forecasts are obtained one can compute the statistical moments, estimate the distribution of deaths at time t using $MaxEnt$, and derive any other life table indicator by applying standard life table calculations (Preston et al., 2000).

4.2.5 Prediction intervals

We simulate prediction intervals for the indices of interest using an algorithm, which makes full allowance for the forecast error generated by the multivariate random–walk model.

Algorithm:

M simulations are performed on a forecasting horizon J.

For simulation $m = 1, 2, \ldots, M$

1. randomly sample a variable z_m^* from the multivariate normal distribution $\mathbf{N}(0, \mathbf{I})$;

For $j = 1, 2, \ldots, J$

2. compute $y_{n,t+j}^* = y_{n,t+j} + j\hat{\theta}_n + \sqrt{j}\hat{C}z_m^*$

3. compute statistical moments $\mu^*_{t+j,n,m}$
4. estimate the density using the $MaxEnt$ algorithm and determine $d^*_{x,t+j,m}$.

4.3 Case study: England and Wales 1960–2016 male mortality experience, ages 0-95

4.3.1 The data

The data source used in this article is the Human Mortality Database (2018), which contains historical mortality data for 43 different countries and territories. 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.

In order to test and illustrate the performance of the method, we fit the model using life table death counts for male population in England & Wales between 1960 and 2016. Additional results for various countries are presented in Appendix 4.5.2.

4.3.2 Model Comparison

In addition to the MEM model the following mortality models are evaluated and used to forecast mortality:

- The multivariate random-walk with drift model:

$$\log(m_{x,t}) = \theta_x + \log(m_{x,t-1}) + \varepsilon_{x,t}. \tag{4.18}$$

  This model represents a simple linear extrapolation of the logarithm of the age-specific death rates, m_x, based on the first and the last observed values in the multivariate time series.

- The Lee and Carter (1992) mortality model:

$$\log(m_{x,t}) = \alpha_x + \beta_x k_t + \varepsilon_{x,t}, \tag{4.19}$$

  which is a numerical algorithm to estimate the age-specific effects α_x and β_x and employs the singular value decomposition (SVD) to derive a univariate time series vector k_t, that becomes the main leading indicator of future mortality.

- The Hyndman and Ullah (2007) – functional mortality model:

$$\log(m_{x,t}) = \alpha_x + \sum_{k=1}^{K} \beta_{x,k}\phi_{t,k} + e_{x,t} + \sigma_{t,x}\varepsilon_{x,t} \tag{4.20}$$

  This model is an extension of the Lee–Carter model where the sum term allows for smooth functions of age and $\sigma_{t,x}$ allows the amount of noise to become age-specific.

- The Oeppen (2008) – compositional-data mortality model:

$$clr(f_{x,t}) = \alpha_x + \beta_x k_t + \varepsilon_{x,t} \tag{4.21}$$

  Again, a variant of the Lee–Carter model where the index of interest to be modelled is the life-table age-at-death distribution $f(x)$, subject to a *clr* transformation (instead of a logarithmic one).

  For all models $\varepsilon_{x,t}$ are independent and identically distributed random variables (`iid`) normally distributed with mean zero.

Thus, five models are evaluated. Three of them are targeting the log–transformed death rates, $\log m_x$, and the other two (Oeppen and MEM) are focusing on modelling the age-specific frequencies of the age-at-death distribution or mortality data in compositional format. We mention that the random-walk model is chosen because of its simplicity, the Lee–Carter and Hydman-Ullah methods are included in comparison because of their popularity and acceptance in the demographic and actuarial literature, and finally the Oeppen model is selected because of its similarity with the MEM model (mainly, the $f(x)$ focus). Discussing the advantages and the features of the models used in comparison

is beyond the scope of this paper. For more details about the models please refer to the original articles.

The analysis is performed using the `R` programming language (R Core Team, 2018). The Lee–Carter and the Hyndman–Ullah models are fitted and forecasted using the `demography` R package (2017). The source code of the other three models can be downloaded and installed in form of an `R` software package from authors' GitHub repository.

4.3.3 Evaluation and predictive power measurements

We assess the performance of the proposed MEM method and the other four mortality models based on forecasts of life expectancy at all ages[2] as compared to the observed life expectancies. All the models are fitted and evaluated over the 0–95 age-range. Since a perfect fit of the data can always be obtained by using a model with enough parameters and due to the fact that a good fit does not guarantee good forecasting performance (Hyndman and Athanasopoulos, 2018), we will not evaluate the models based on their ability to fit the historical data. The models are evaluated based on the out-of-sample forecasting performance over the observed data, where we refer to *sample* as the dataset used in fitting. The predictive power is our ultimate goal, translated into a high degree of accuracy of forecast trajectories.

Many of accuracy measures have been published. See Hyndman and Koehler (2006) for a comprehensive review of the most common accuracy measures used in forecasting literature. Only six of them are considered here:

- ME – Mean Error;
- MAE – Mean Absolute Error;
- MAPE – Mean Absolute Percentage Error;

[2]Because the five mortality extrapolation methods are modelling different life table indicators (m_x vs. f_x), in order to perform a fair comparison one needs to make sure that the life table computation guarantees the transitivity between the indicators. That is, given a certain mortality level the same values of life expectancy (indicator evaluated in this article) are obtained regardless of whether the life table construction starts from m_x, q_x, l_x or f_x. In this article the goal is achieved by using the life table methods implemented in the `MortalityLaws` R package (2018).

- sMAPE – Symmetric Mean Absolute Percentage Error;
- sMRAE – Symmetric Mean Absolute Relative Error;
- MASE – Mean Absolute Scaled Error.

The sMRAE and MASE are accuracy measure computed relative to a benchmark model. In the case of sMRAE the reference model in our study is the multivariate random-walk with drift, because we consider this model to be the simplest reliable method to extrapolate age specific death-rates. The MASE measure assesses the accuracy of a forecast with reference to a simple 1-step random-walk model (without drift). For all the presented accuracy measure, except ME, a smaller value is preferred over a larger one. A model performs better in terms of ME, compared to another model, if the obtained value is closer to zero i.e. the smallest value in absolute terms. Because the six measures are evaluating the accuracy by analysing different aspects of the realised forecasts, it is possible but not mandatory to obtain a different classification of the model performance, depending on the considered measure. We compute a general classification (GC) of the resulted accuracy performance of the analysed models by considering the median classification over the six measures for each model. The best performing model is marked with: rank (1).

4.3.4 Out–of–sample forecasting strategy

The period between 1960–2016 is long enough and relevant at the same time for assessing the predictive power of the estimated models. A typical testing scenario would use sub-periods of 20 years of data to fit/train the models, and subsequent periods of 20 years of data to forecast and validate the results. Multiple testing scenarios are defined by rolling forward the training/validation windows in steps of 1 year. This strategy will be called: *20–20–1*. Therefore, considering our timeline, in the first scenario we will use data from 1960 to 1979 for fitting the models. The resulting models will be used to predict mortality for the period between 1980 and 1999, and validate the forecasts against the observed values in the same period. We will refer to this as the *1960–1979–1999* scenario. By moving the evaluation windows 1 year forward, the second scenario would be *1961–1980–2000*. And so on until the last scenario: *1977–1996–2016*. In total 18 scenarios have

been defined, containing equal fitting and forecasting period lengths, making possible in this way the aggregation (by averaging) and comparison of the accuracy results over all scenarios in addition to the specific scenario results.

4.3.5 Results

Across multiple measures of forecast accuracy computed based on the mortality experience of the male population living in England and Wales, we find that the predictive power of the MEM method is net superior to the other models. Table 4.1 displays a summary of the aggregated measures over 18 defined scenarios. We also learn that the differences between the multivariate random-walk with drift, Lee–Carter and Hydman-Ullah models are insignificant in the case of this population. And that the Oeppen method consistently offers better results among the Lee–Carter type models, obtaining the second position among the best-performing models in this study.

Table 4.1: *Forecast accuracy measures aggregated over 18 scenarios in the 1960–2016 period*

Model	ME	MAE	MAPE	sMAPE	sMRAE	MASE	GC
M.Random-Walk w Drift	0.72 (3)	0.73 (3)	2.98 (3)	3.06 (3)	100.00 (3)	4.06 (3)	(3)
Lee–Carter	0.78 (5)	0.78 (5)	3.15 (5)	3.25 (5)	103.93 (5)	4.30 (5)	(5)
Hyndman–Ullah	0.76 (4)	0.77 (4)	3.10 (4)	3.19 (4)	102.69 (4)	4.22 (4)	(4)
Oeppen	0.61 (2)	0.62 (2)	2.61 (2)	2.68 (2)	91.98 (2)	3.47 (2)	(2)
MEM–6	0.42 (1)	0.45 (1)	2.11 (1)	2.15 (1)	79.59 (1)	2.66 (1)	(1)

When the models are tested over the mortality experience of multiple populations using the same strategy we discovered a similar patter, namely, in the majority of the cases the methods based of age-at-death distribution (MEM and Oeppen) would be ranked as first and second in the general classification.

The projected trajectories given by these methods can be inspected across different life table indicators, which are equivalent in the sense that they represent the same level of mortality. In Figure 4.4 we show the resulting mean trends in life expectancy at age 0, 25, 45, 65, 75 and 85; and in Figure 4.5 the trends in central death rate at the same ages are represented. It is noticeable that the MEM can cope with different levels of mortality

improvement over age and time. This is the main reason why the model is able to return significantly better forecasts.

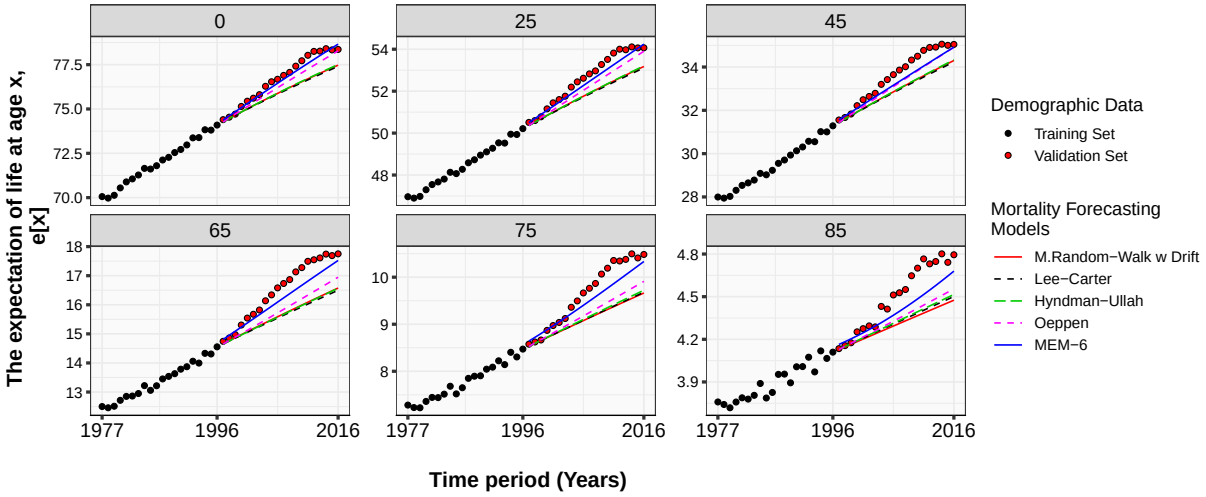


Figure 4.4: *Out–of–sample forecast of the remaining life expectancy at various ages using the five mortality models (England & Wales, Male population, Scenario 18: 1977–1996–2016)*

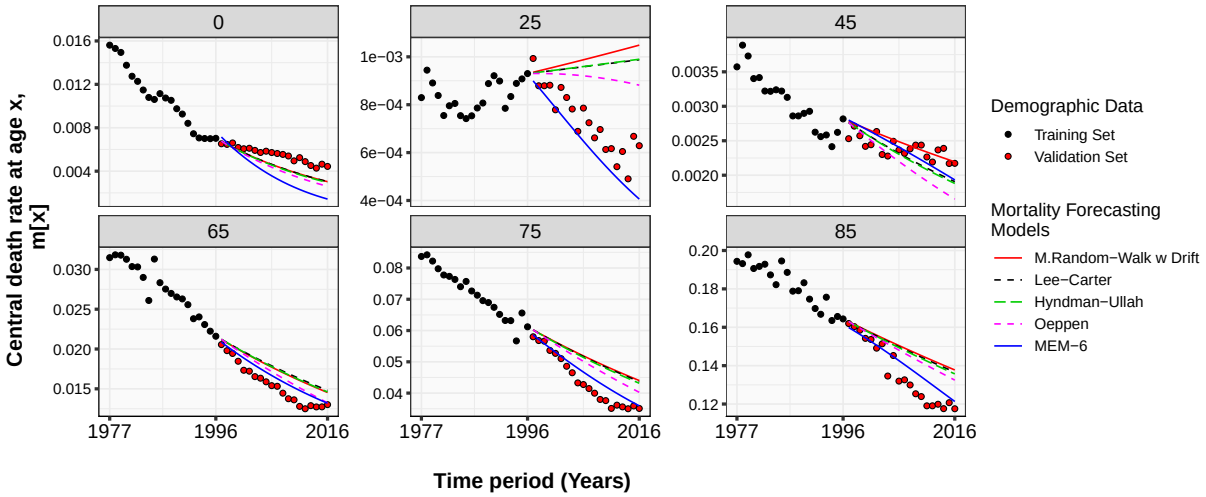


Figure 4.5: *Out–of–sample forecast of the age-specific death rates using the five mortality models (England & Wales, Male population, Scenario 18: 1977–1996–2016)*

4.3.6 How many moments to use in MEM forecasting?

In general, the number of statistical moments to be considered in the MEM model depends on the regularity of the age-at-death distribution in the population of interest. The more moments employed, the more accurate the estimation of the underlying distribution becomes. However, the cost of using a lager number of moments is paid in processing

speed and the likelihood of convergence of the MaxEnt algorithm. For 7 or more moments a more complex time series model for moment extrapolation might be required.

We tested the MEM models of order 2 to 6, that is models that are estimated based on the first 2, 3 until 6 moments (plus μ_0). The testing was carried out in the same manner and over the same scenarios as in section 4.3.5. From the results in table 4.2 we learn that only the MEM–2 is disqualified by the benchmark model, all the other variants of the MEM return significantly better results.

Table 4.2: *Forecast accuracy measures aggregated over 18 scenarios in the 1960–2016 period*

Model	ME	MAE	MAPE	sMAPE	sMRAE	MASE	GC
M.Random-Walk w Drift	0.72 (5)	0.73 (5)	2.98 (5)	3.06 (5)	100.00 (5)	4.06 (5)	(5)
MEM–2	1.00 (6)	1.01 (6)	4.02 (6)	4.16 (6)	113.12 (6)	5.44 (6)	(6)
MEM–3	0.56 (4)	0.58 (4)	2.72 (4)	2.80 (4)	88.14 (4)	3.34 (4)	(4)
MEM–4	0.47 (3)	0.49 (3)	2.41 (3)	2.47 (3)	83.10 (3)	2.94 (3)	(3)
MEM–5	0.43 (2)	0.46 (2)	2.17 (2)	2.22 (2)	80.55 (2)	2.73 (2)	(2)
MEM–6	0.42 (1)	0.45 (1)	2.11 (1)	2.15 (1)	79.59 (1)	2.66 (1)	(1)

4.4 Conclusion & Discussion

The maximum entropy mortality model represents a new approach to modelling age-specific mortality levels. Nevertheless, its novelty refers only to the authors' idea of defining an algorithm to forecast mortality using well established methods and concepts like "statistical moments", "information entropy", "the finite moment problem" and "the multivariate random-walk with drift" introduced decades or centuries ago. All these individually have important application in different scientific fields.

The main advantage of the MEM is that the possible forecast age-specific trends are no longer based on the assumption of constant changes in mortality as in the Lee–Carter model. A different speed of improvement can be predicted across ages by taking into account the observed dynamics of the distribution of deaths and the change in its shape and location. This makes possible the identification of the "location" of the longevity risk across the x-axis. The model has the required features to predict the different rates of change in life expectancy at different ages, e.g. by maintaining a linear increase in life

expectancy at birth and at the same time inducing accelerating rates for ages above 65. This result is consistent with the observed trends in the past and across many developed countries.

Even if not shown here, the model introduced in this article is flexible enough in the sense that different covariates, influencing the mortality dynamics, can be considered in order to further improve the predictions. Examples of such covariates are information on the prevalence of smoking or obesity but also the trends in life expectancy at birth and modal age at death. This can be done by extending the time series model used in extrapolation (the multivariate random-walk with drift) by attaching extra cause-specific parameters. Similarly, a wide range of multivariate autoregressive time series models can be used to capture the coherent trends given by the observed empirical statistical moments, however this is subject to the requirements imposed by the available data.

Including higher order moments in the prediction can sometimes increase the importance of relatively small effects seen in the age-at-death distribution such as the accident hump for males or infant mortality. In countries with a low level of mortality, four moments can return pertinent results but in countries that still exhibit a pronounced multi-modal distribution of deaths a larger number of moments might be required. The coverage proportion, introduced in the article, is a very good measure to study the model's ability to estimate the shape of the true distribution, but it has an important drawback: two errors of the same magnitude at different ends of the distribution will be assigned the same weight in the measure. This is a drawback because underestimating or overestimating the force of mortality at younger ages has a higher impact on the general level of mortality of a population than underestimating or overestimating the force of mortality at an older age. If we investigate the evolution of longevity by looking at life expectancy we can see that relying solely on information given by the mean age at death and life-span disparity in order to generate forecasts is an endeavour doomed to failure (as showed in table 4.2). The model would mostly return pronounced pessimistic or overoptimistic results across ages without a correspondence to reality. A real improvement can be noticed if 4 to 6 moments are used.

Other possible extensions of the MEM method worth exploring in the future are the use of an optimal weighting scheme which decreases the importance of higher order moments or the use of various smoothing methods that can be applied to the data prior to computing the observed moments and fitting the model.

An important finding revealed in this study is the superiority of the age-at-death distribution based extrapolation methods like the MEM and the Oeppen over the death-rate based extrapolation methods.

4.5 Appendix

4.5.1 The Multivariate random–walk with drift model

Denote a multivariate time series $m_{x,t}$ with $t = \{0, 1, 2, \ldots, \tau\}$, $x = \{1, 2, \ldots, \omega\}$ and first order differences

$$y_{x,t} = m_{x,t} - m_{x,t-1}, \quad t = 1, 2, \ldots, \tau.$$

For the random–walk with drift

$$y_{x,t} = \theta_x + y_{x,t-1} + \varepsilon_{x,t}.$$

Refer to the multivariate Gaussian model

$$Y = GA + \varepsilon,$$

for which

$$Y = \begin{bmatrix} y_{1,1} & y_{2,1} & \cdots & y_{\omega,1} \\ y_{1,2} & y_{2,2} & \cdots & y_{\omega,2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{1,\tau} & y_{2,\tau} & \cdots & y_{\omega,\tau} \end{bmatrix}, \quad G = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad A = \big[\theta_1, \theta_2, \ldots, \theta_\omega\big], \quad \text{and} \quad \varepsilon \sim \mathbf{N}(0, \Omega),$$

so that

$$G'G = \tau, \quad G'Y = \big[y_{1+}, y_{2+}, \ldots, y_{\tau+}\big], \quad \text{where} \quad y_{i+} = \sum_{t=1}^{\tau} y_{i,t} = m_{i,\tau} - m_{i,0}.$$

Then the OLS estimates

$$\hat{A} = (G'G)^{-1}(G'Y) = \left[\frac{m_{i,\tau} - m_{i,0}}{\tau}\right], \quad \text{with} \quad \hat{\Omega} = \left[\frac{\hat{\varepsilon}'\hat{\varepsilon}}{\tau - 1}\right].$$

The matrix of residuals

$$\hat{\varepsilon} = Y - G\hat{A} = \left[y_{i,t} - \hat{\theta}_i\right] = \left[r_{i,t}\right],$$

so that

$$\hat{\varepsilon}'\hat{\varepsilon} = \begin{bmatrix} \sum r_{1,t}^2 & \sum r_{1,t}r_{2,t} & \cdots & \sum r_{1,t}r_{\omega,t} \\ \sum r_{2,t}r_{1,t} & \sum r_{2,t}^2 & \cdots & \sum r_{2,t}r_{\omega,t} \\ \vdots & \vdots & \ddots & \vdots \\ \sum r_{\omega,t}r_{1,t} & \sum r_{\omega,t}r_{2,t} & \cdots & \sum r_{\omega,t}^2 \end{bmatrix}.$$

Forecasting: successive substitution gives

$$m_{x,t+j} = m_{x,t} + j\theta_x + \varepsilon_{x,t+j} + \varepsilon_{x,t+j-1} + \cdots + \varepsilon_{x,t+1}.$$

Then, taking expected values, the j-step ahead forecast, from $t(= \tau)$, is

$$\hat{m}_{x,t+j|t} = m_{x,t} + j\theta_x.$$

4.5.2 Out–of–sample forecasts in various countries

Out–of–sample forecasts: Australia, Male population, 1960–2014

Table 4.3: *Forecast accuracy measures aggregated over 16 scenarios in the 1960–2014 period*

Model	ME	MAE	MAPE	sMAPE	sMRAE	MASE	GC
M.Random-Walk w Drift	0.59 (4)	0.61 (4)	2.46 (5)	2.52 (5)	100.00 (4)	3.12 (4)	(4)
Lee–Carter	0.60 (5)	0.62 (5)	2.41 (4)	2.45 (4)	101.68 (5)	3.13 (5)	(5)
Hyndman–Ullah	0.48 (3)	0.51 (3)	2.11 (2)	2.14 (2)	94.54 (3)	2.64 (3)	(3)
Oeppen	0.36 (2)	0.51 (2)	2.15 (3)	2.19 (3)	84.95 (1)	2.62 (2)	(2)
MEM–6	0.11 (1)	0.41 (1)	1.75 (1)	1.76 (1)	85.70 (2)	2.15 (1)	(1)

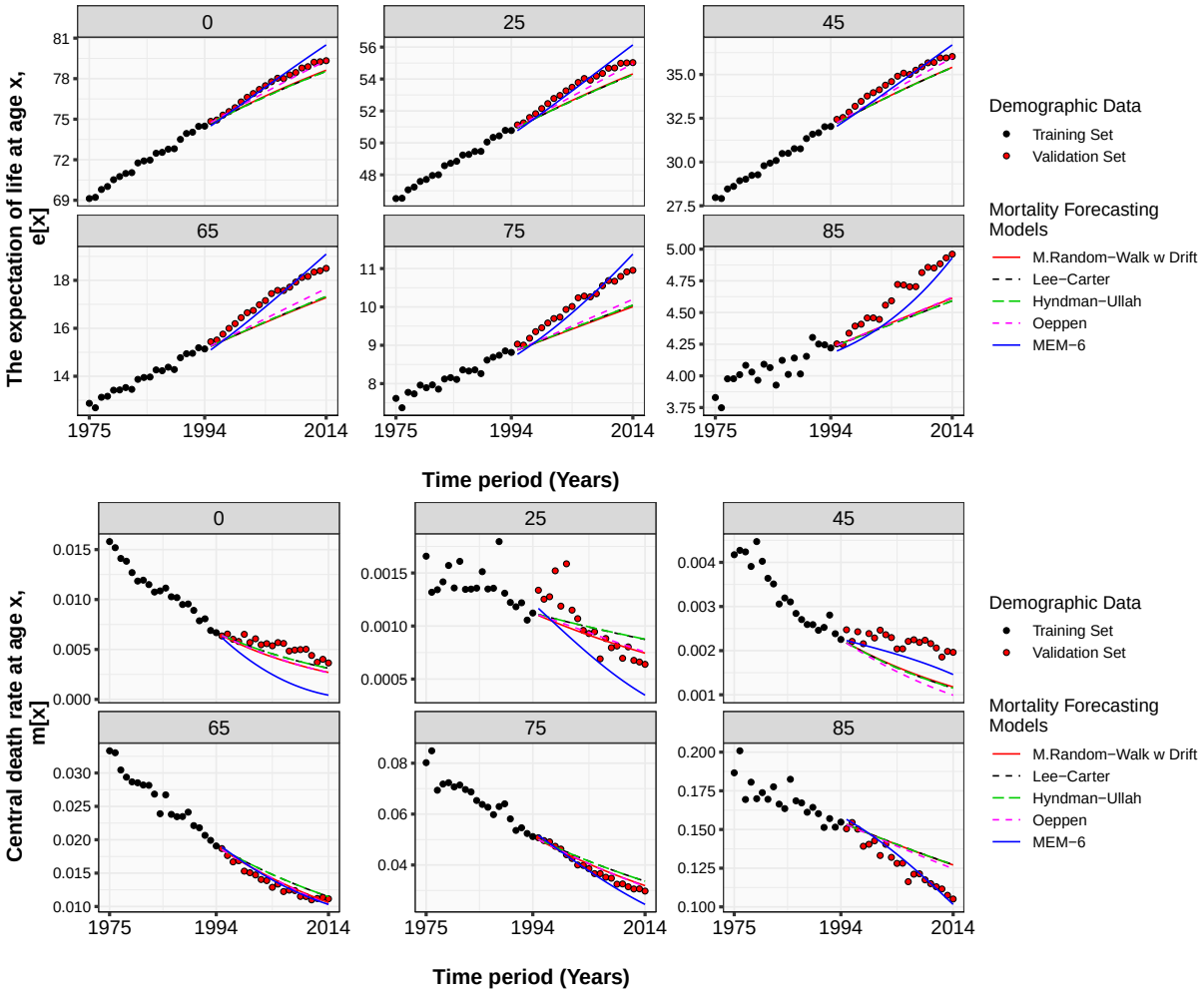


Figure 4.6: *Out–of–sample forecast of the remaining life expectancy and central death rates at various ages using the five mortality models (Australia, Male population, Scenario 16: 1975–1994–2014)*

Out–of–sample forecasts: Canada, Male population, 1960–2011

Table 4.4: *Forecast accuracy measures aggregated over 13 scenarios in the 1960–2011 period*

Model	ME	MAE	MAPE	sMAPE	sMRAE	MASE	GC
M.Random-Walk w Drift	0.58 (4)	0.59 (4)	2.19 (4)	2.23 (4)	100.00 (4)	3.77 (4)	(4)
Lee–Carter	0.63 (5)	0.65 (5)	2.32 (5)	2.37 (5)	103.35 (5)	4.03 (5)	(5)
Hyndman–Ullah	0.44 (3)	0.50 (3)	1.98 (3)	2.01 (3)	90.19 (3)	3.27 (3)	(3)
Oeppen	0.35 (2)	0.45 (2)	1.88 (2)	1.89 (2)	88.39 (2)	3.05 (2)	(2)
MEM–6	0.15 (1)	0.25 (1)	1.06 (1)	1.07 (1)	68.07 (1)	1.72 (1)	(1)

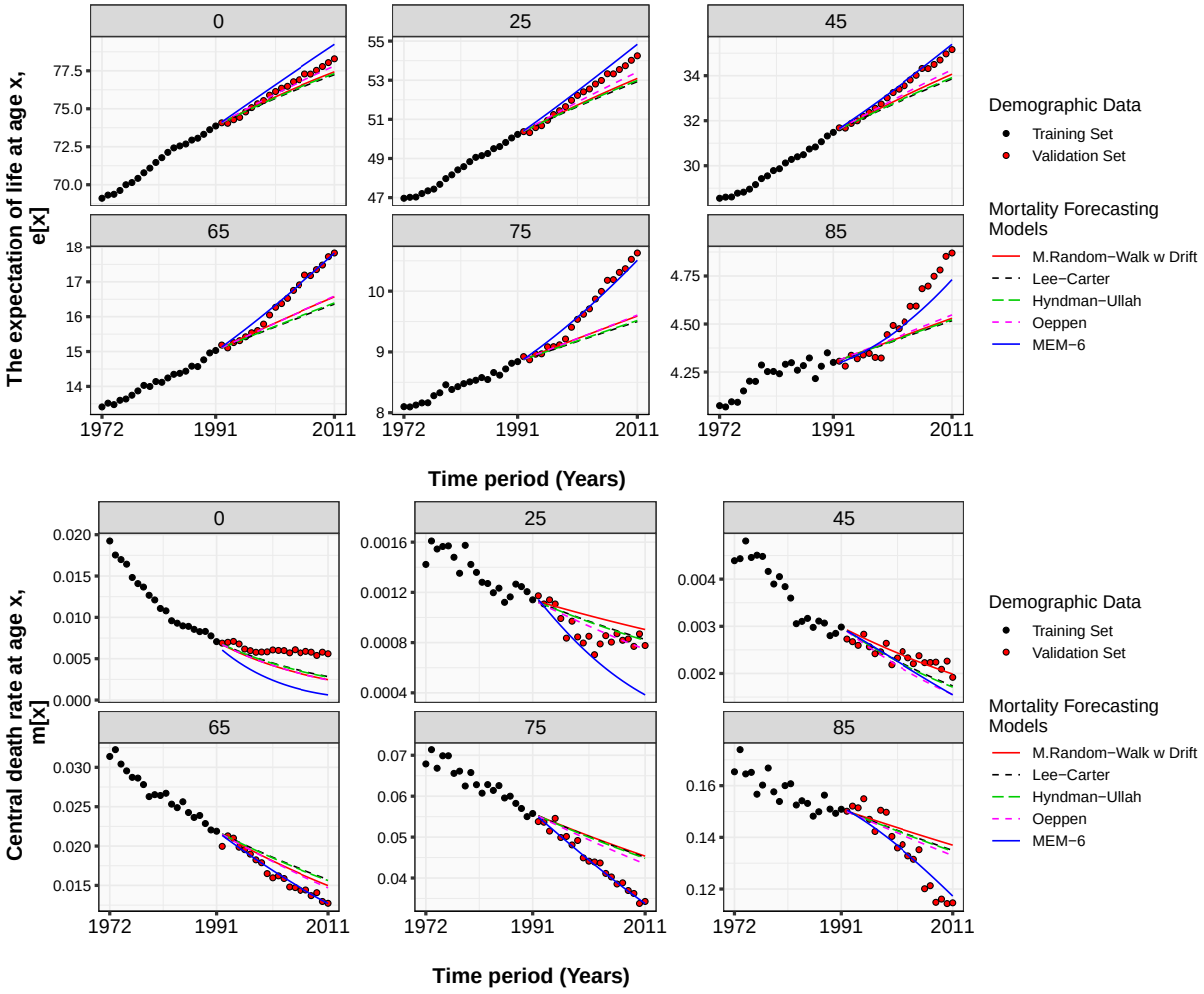


Figure 4.7: *Out–of–sample forecast of the remaining life expectancy and central death rates at various ages using the five mortality models (Canada, Male population, Scenario 13: 1972–1991–2011)*

Out–of–sample forecasts: France, Male population, 1960–2016

Table 4.5: *Forecast accuracy measures aggregated over 18 scenarios in the 1960–2016 period*

Model	ME	MAE	MAPE	sMAPE	sMRAE	MASE	GC
M.Random-Walk w Drift	0.46 (3)	0.50 (3)	1.85 (3)	1.88 (3)	100.00 (4)	3.13 (3)	(3)
Lee–Carter	0.53 (5)	0.55 (5)	2.02 (5)	2.06 (5)	106.19 (5)	3.44 (5)	(5)
Hyndman–Ullah	0.47 (4)	0.51 (4)	1.90 (4)	1.93 (4)	99.63 (3)	3.21 (4)	(4)
Oeppen	0.22 (1)	0.40 (2)	1.61 (2)	1.63 (2)	92.73 (2)	2.60 (2)	(2)
MEM–6	0.23 (2)	0.35 (1)	1.56 (1)	1.58 (1)	91.36 (1)	2.43 (1)	(1)

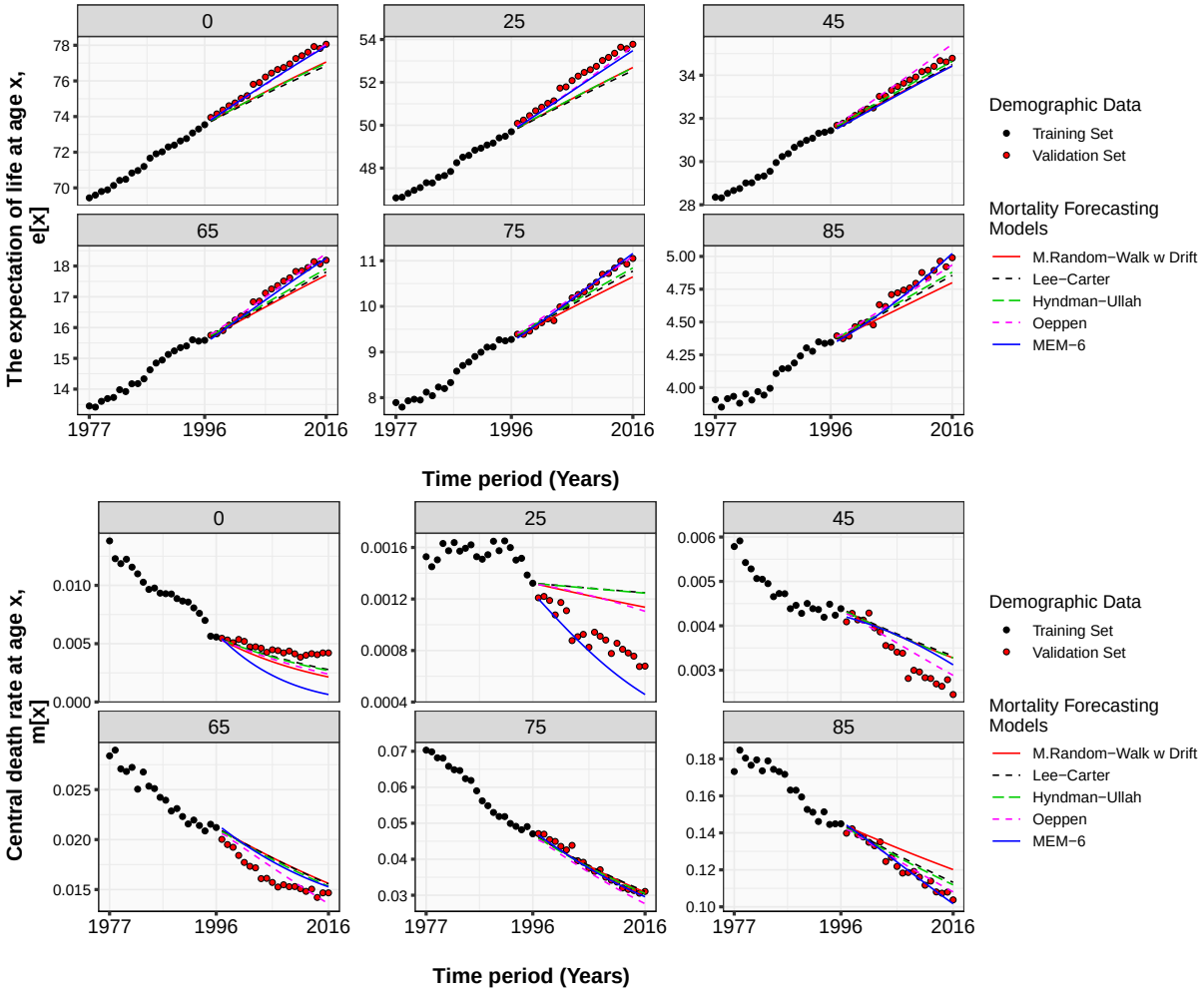


Figure 4.8: *Out–of–sample forecast of the remaining life expectancy and central death rates at various ages using the five mortality models (France, Male population, Scenario 18: 1977–1996–2016)*

Out–of–sample forecasts: Italy, Male population, 1960–2014

Table 4.6: *Forecast accuracy measures aggregated over 16 scenarios in the 1960–2014 period*

Model	ME	MAE	MAPE	sMAPE	sMRAE	MASE	GC
M.Random-Walk w Drift	0.88 (3)	0.88 (3)	3.27 (3)	3.37 (3)	100.00 (3)	4.82 (3)	(3)
Lee–Carter	0.94 (5)	0.94 (5)	3.46 (5)	3.56 (5)	103.71 (5)	5.10 (5)	(5)
Hyndman–Ullah	0.92 (4)	0.93 (4)	3.41 (4)	3.51 (4)	102.27 (4)	5.03 (4)	(4)
Oeppen	0.79 (2)	0.80 (2)	3.01 (2)	3.10 (2)	93.97 (2)	4.38 (2)	(2)
MEM–5	0.34 (1)	0.41 (1)	2.05 (1)	2.10 (1)	72.98 (1)	2.63 (1)	(1)

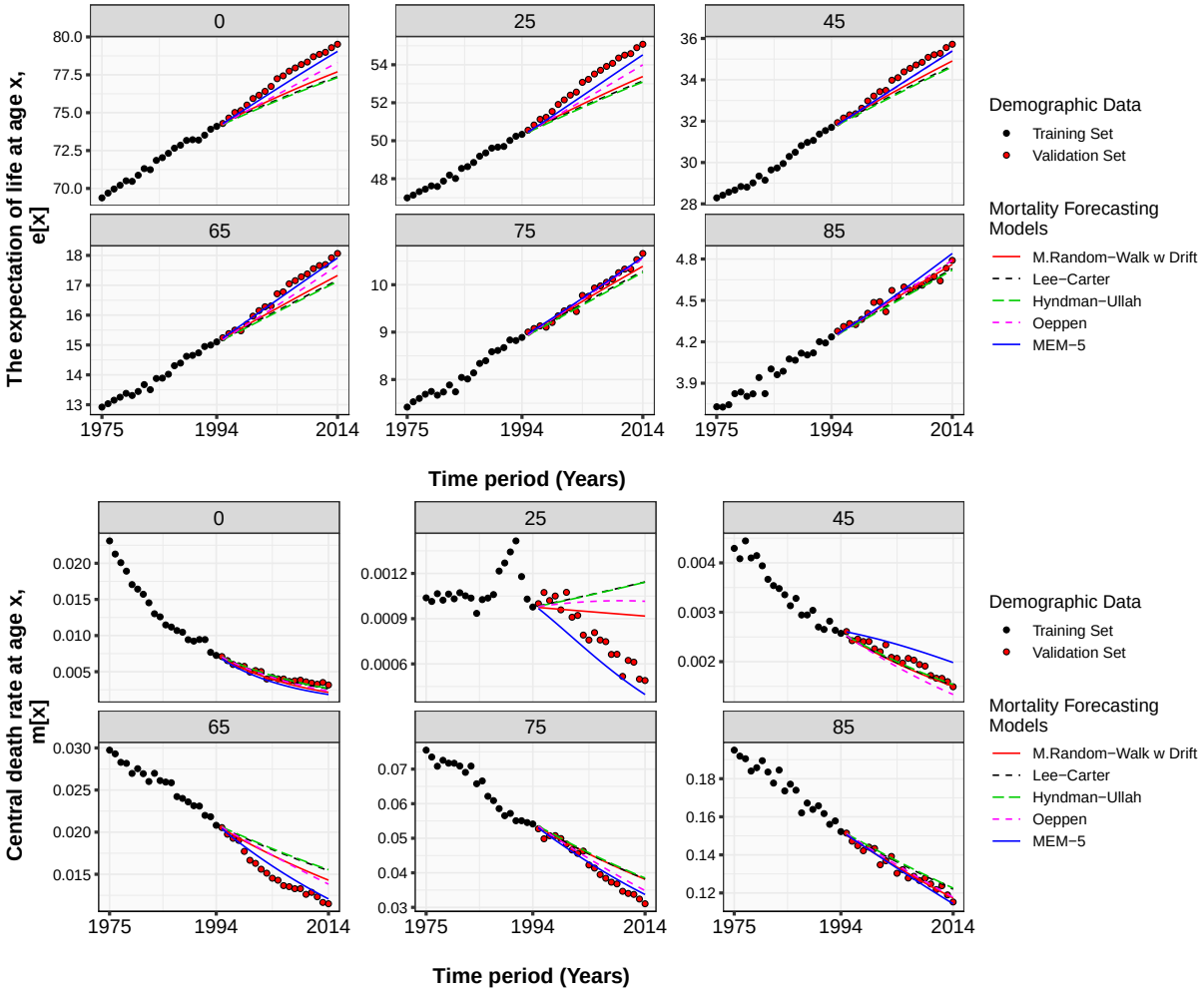


Figure 4.9: *Out–of–sample forecast of the remaining life expectancy and central death rates at various ages using the five mortality models (Italy, Male population, Scenario 16: 1975–1994–2014)*

Out–of–sample forecasts: The Netherlands, Male population, 1960–2016

Table 4.7: *Forecast accuracy measures aggregated over 18 scenarios in the 1960–2016 period*

Model	ME	MAE	MAPE	sMAPE	sMRAE	MASE	GC
M.Random-Walk w Drift	0.81 (5)	0.84 (5)	3.36 (5)	3.46 (5)	100.00 (5)	5.10 (5)	(5)
Lee–Carter	0.80 (4)	0.82 (4)	3.25 (4)	3.34 (4)	98.56 (4)	4.92 (4)	(4)
Hyndman–Ullah	0.79 (3)	0.81 (3)	3.24 (3)	3.33 (3)	98.33 (3)	4.88 (3)	(3)
Oeppen	0.76 (2)	0.79 (2)	3.17 (2)	3.26 (2)	96.94 (2)	4.76 (2)	(2)
MEM–5	0.54 (1)	0.59 (1)	2.54 (1)	2.60 (1)	84.66 (1)	3.70 (1)	(1)

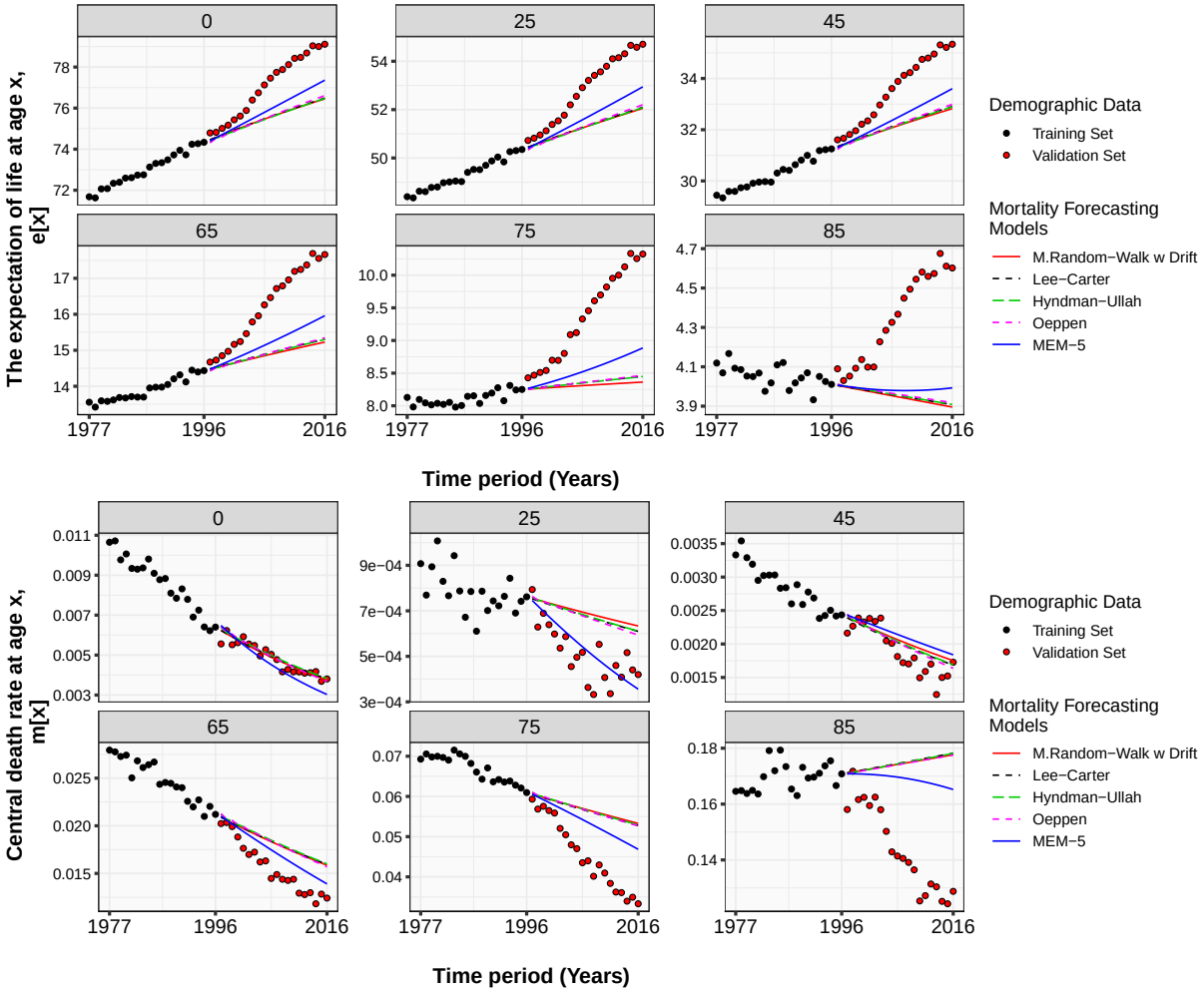


Figure 4.10: *Out–of–sample forecast of the remaining life expectancy and central death rates at various ages using the five mortality models (The Netherlands, Male population, Scenario 18: 1977–1996–2016)*

Out–of–sample forecasts: Spain, Male population, 1960–2014

Table 4.8: *Forecast accuracy measures aggregated over 16 scenarios in the 1960–2014 period*

Model	ME	MAE	MAPE	sMAPE	sMRAE	MASE	GC
M.Random-Walk w Drift	0.10 (1)	0.33 (1)	1.22 (3)	1.22 (3)	100.00 (3)	2.18 (2)	(2)
Lee–Carter	0.11 (2)	0.33 (2)	1.15 (1)	1.15 (1)	97.74 (1)	2.15 (1)	(1)
Hyndman–Ullah	0.13 (3)	0.34 (3)	1.18 (2)	1.18 (2)	98.76 (2)	2.21 (3)	(2)
Oeppen	-0.15 (4)	0.38 (4)	1.38 (4)	1.37 (4)	108.45 (4)	2.56 (4)	(4)
MEM–5	-0.29 (5)	0.50 (5)	1.60 (5)	1.59 (5)	112.71 (5)	3.40 (5)	(5)

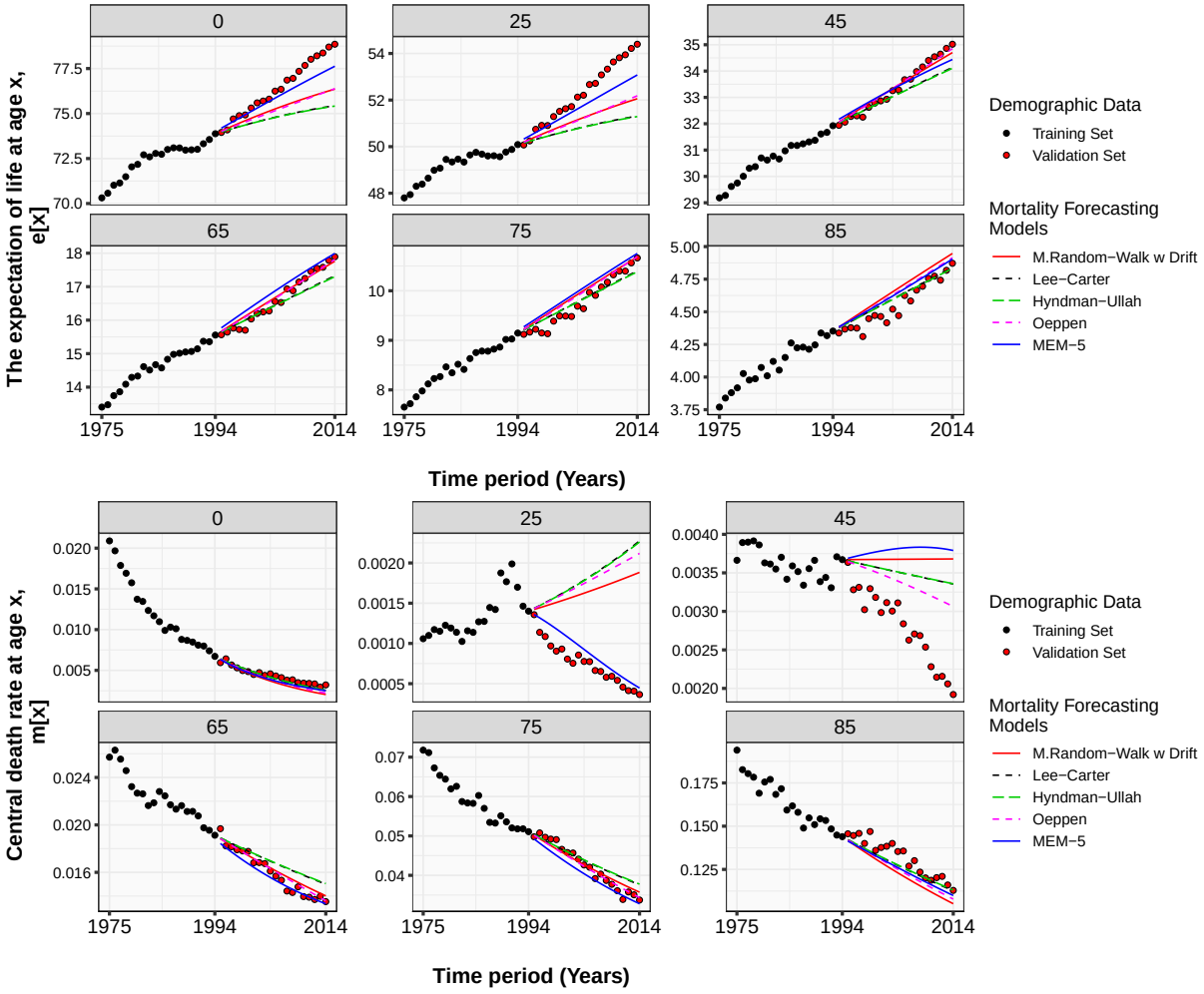


Figure 4.11: *Out–of–sample forecast of the remaining life expectancy and central death rates at various ages using the five mortality models (Spain, Male population, Scenario 16: 1975–1994–2014)*

Out–of–sample forecasts: Switzerland, Male population, 1960–2016

Table 4.9: *Forecast accuracy measures aggregated over 18 scenarios in the 1960–2016 period*

Model	ME	MAE	MAPE	sMAPE	sMRAE	MASE	GC
M.Random-Walk w Drift	0.58 (3)	0.61 (3)	2.12 (3)	2.15 (3)	100.00 (3)	3.29 (3)	(3)
Lee–Carter	0.64 (4)	0.66 (4)	2.24 (4)	2.27 (4)	104.34 (4)	3.55 (4)	(4)
Hyndman–Ullah	0.66 (5)	0.68 (5)	2.30 (5)	2.34 (5)	107.12 (5)	3.65 (5)	(5)
Oeppen	0.44 (2)	0.48 (2)	1.67 (2)	1.69 (2)	90.50 (2)	2.58 (2)	(2)
MEM–5	0.32 (1)	0.41 (1)	1.50 (1)	1.51 (1)	85.97 (1)	2.25 (1)	(1)

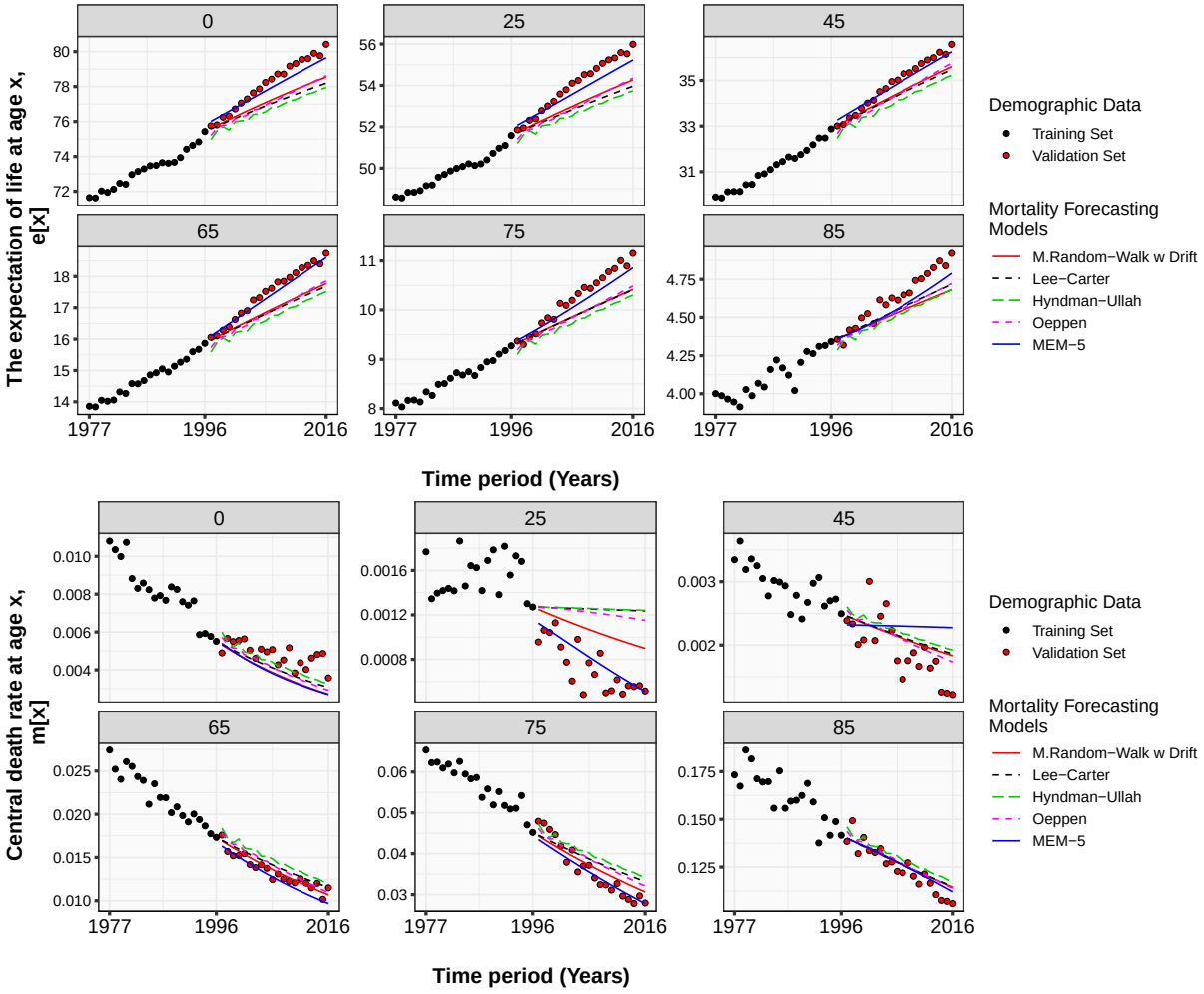


Figure 4.12: *Out–of–sample forecast of the remaining life expectancy and central death rates at various ages using the five mortality models (Switzerland, Male population, Scenario 18: 1977–1996–2016)*

Out–of–sample forecasts: Sweden, Male population, 1960–2016

Table 4.10: *Forecast accuracy measures aggregated over 18 scenarios in the 1960–2016 period*

Model	ME	MAE	MAPE	sMAPE	sMRAE	MASE	GC
M.Random-Walk w Drift	0.85 (3)	0.86 (3)	3.03 (3)	3.10 (3)	100.00 (3)	5.26 (3)	(3)
Lee–Carter	0.89 (4)	0.90 (4)	3.12 (4)	3.21 (4)	101.10 (4)	5.46 (4)	(4)
Hyndman–Ullah	0.92 (5)	0.92 (5)	3.19 (5)	3.28 (5)	103.79 (5)	5.60 (5)	(5)
Oeppen	0.84 (2)	0.84 (2)	2.94 (2)	3.01 (2)	97.04 (2)	5.12 (2)	(2)
MEM–5	0.61 (1)	0.65 (1)	2.53 (1)	2.59 (1)	86.42 (1)	4.08 (1)	(1)

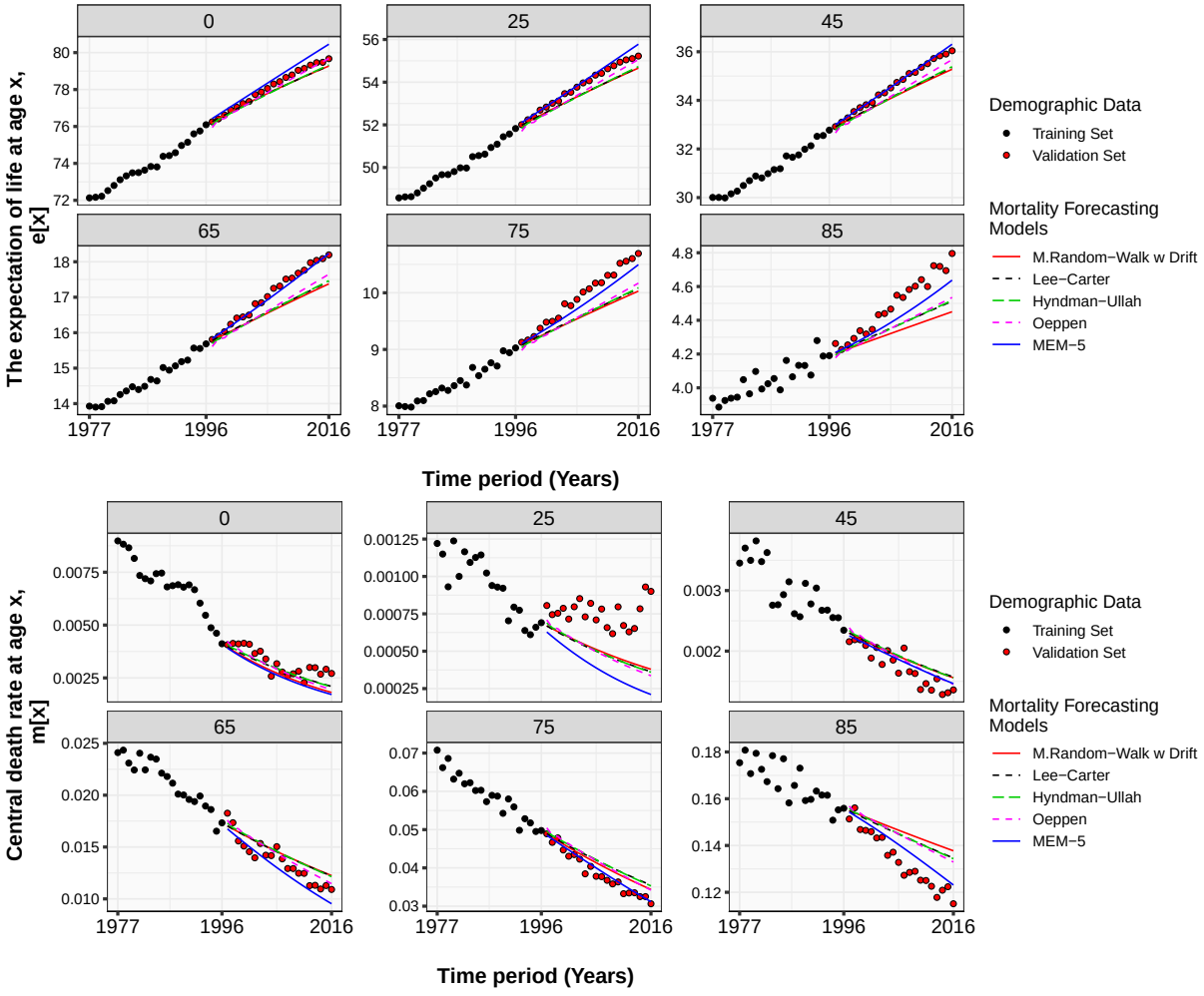


Figure 4.13: *Out–of–sample forecast of the remaining life expectancy and central death rates at various ages using the five mortality models (Sweden, Male population, Scenario 18: 1977–1996–2016)*

Out–of–sample forecasts: USA, Male population, 1960–2016

Table 4.11: *Forecast accuracy measures aggregated over 18 scenarios in the 1960–2016 period*

Model	ME	MAE	MAPE	sMAPE	sMRAE	MASE	GC
M.Random-Walk w Drift	0.21 (4)	0.28 (3)	1.23 (3)	1.24 (3)	100.00 (3)	2.39 (3)	(3)
Lee–Carter	0.25 (5)	0.31 (4)	1.33 (4)	1.34 (4)	104.51 (4)	2.59 (4)	(4)
Hyndman–Ullah	0.15 (3)	0.42 (5)	1.76 (5)	1.78 (5)	115.63 (5)	3.49 (5)	(5)
Oeppen	0.02 (1)	0.24 (1)	1.16 (2)	1.16 (2)	94.18 (1)	2.21 (2)	(1)
MEM–6	-0.09 (2)	0.24 (2)	0.95 (1)	0.95 (1)	94.82 (2)	1.92 (1)	(1)

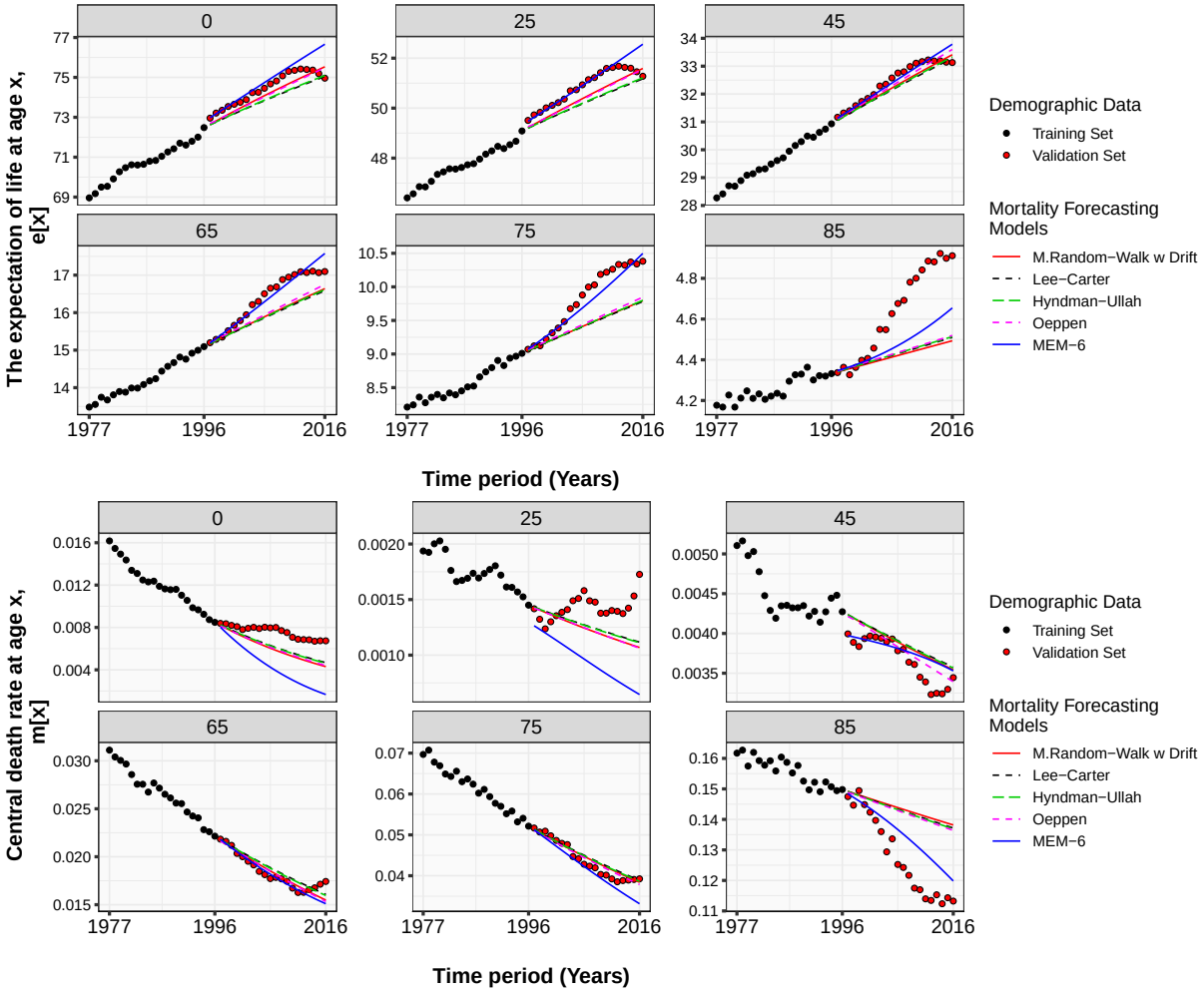


Figure 4.14: *Out–of–sample forecast of the remaining life expectancy and central death rates at various ages using the five mortality models (USA, Male population, Scenario 18: 1977–1996–2016)*

Chapter 5

MortalityLaws: Parametric Mortality Models, Life Tables and HMD

Marius D. Pascariu

Published in the *Comprehensive R Archive Network*

Package 'MortalityLaws'

August 31, 2018

Type Package

Title Parametric Mortality Models, Life Tables and HMD

Version 1.6.0

Date 2018-08-31

Maintainer Marius D. Pascariu <mpascariu@outlook.com>

Description Fit the most popular human mortality 'laws', and construct full and abridge life tables given various input indices. A mortality law is a parametric function that describes the dying-out process of individuals in a population during a significant portion of their life spans. For a comprehensive review of the most important mortality laws see Tabeau (2001) <doi:10.1007/0-306-47562-6_1>. An elegant function for downloading data from Human Mortality Database <https://www.mortality.org> is provided as well.

License GPL-3

LazyData TRUE

Depends R (>= 3.0.0)

Imports minpack.lm (>= 1.2), RCurl (>= 1.95), pbapply (>= 1.3-4), tidyr (>= 0.8.1)

Suggests testthat, knitr, rmarkdown

URL https://github.com/mpascariu/MortalityLaws

BugReports https://github.com/mpascariu/MortalityLaws/issues

RoxygenNote 6.1.0

VignetteBuilder knitr

NeedsCompilation no

Author Marius D. Pascariu [aut, cre, cph] (<https://orcid.org/0000-0002-2568-6489>), Vladimir Canudas-Romo [ctb]

Repository CRAN

Date/Publication 2018-08-31 12:00:03 UTC

R topics documented:

---

ahmd *MortalityLaws Test Data*

Description

Dataset containing altered death rates (mx), death counts (Dx) and exposures (Ex) for the female population living in England & Wales in four different years: 1850, 1900, 1950 and 2010. The data-set is provided for testing purposes only. Download the actual data free of charge from https://www.mortality.org. Once a username and a password are created on the website the function ReadHMD can be used for downloading.

Usage

```
ahmd
```

Format

An object of class `list` of length 3.

Source

Human Mortality Database

See Also

ReadHMD

Examples

```
head(ahmd$mx)
```

availableHMD *Check Data Availability in HMD*

Description

The function returns information about available data in HMD (period life tables etc.), with the range of years covered by the life tables.

Usage

```
availableHMD(username, password, ...)
```

Arguments

username	Your HMD username. If you don't have one you can sign up for free on the Human Mortality Database website.
password	Your HMD password.
...	Other parameters to be passed in ReadHMD function.

Value

An availableHMD object.

Author(s)

Marius D. Pascariu

See Also

ReadHMD

Examples

```
## Not run:
# This will take few seconds...
datainfo <- availableHMD(username = "your_username",
                         password = "your_password")
datainfo

## End(Not run)
```

`availableLaws` *Check Available Mortality Laws*

Description

The function returns information about the parametric models that can be called and fitted in the `MortalityLaw` function. For a comprehensive review of the most important mortality laws, Tabeau (2001) is a good starting point.

Usage

```
availableLaws(law = NULL)
```

Arguments

`law`	Optional. Default: `NULL`. One can extract details about a certain model by specifying its codename.

Value

The output is of the `"availableLaws"` class with the components:

`table`	Table with mortality models and codes to be used in `MortalityLaw`
`legend`	Table with details about the section of the mortality curve

Author(s)

Marius D. Pascariu

References

1. Gompertz, B. (1825). On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. Philosophical Transactions of the Royal Society of London, 115, 513-583.
2. Makeham, W. (1860). On the Law of Mortality and Construction of Annuity Tables. The Assurance Magazine and Journal of the Institute of Actuaries, 8(6), 301-310.
3. Thiele, T. (1871). On a Mathematical Formula to express the Rate of Mortality throughout the whole of Life, tested by a Series of Observations made use of by the Danish Life Insurance Company of 1871. Journal of the Institute of Actuaries and Assurance Magazine, 16(5), 313-329.
4. Oppermann, L. H. F. (1870). On the graduation of life tables, with special application to the rate of mortality in infancy and childhood. The Insurance Record Minutes from a meeting in the Institute of Actuaries., 42.
5. Wittstein, T. and D. Bumsted. (1883). The Mathematical Law of Mortality. Journal of the Institute of Actuaries and Assurance Magazine, 24(3), 153-173.

6. Steffensen, J. (1930). Infantile mortality from an actuarial point of view. Skandinavisk Aktuarietidskrift 13, 272-286.

7. Perks, W. (1932). On Some Experiments in the Graduation of Mortality Statistics. Journal of the Institute of Actuaries, 63(1), 12-57.

8. Harper, F. S. (1936). An actuarial study of infant mortality. Scandinavian Actuarial Journal 1936 (3-4), 234-270.

9. Weibull, W. (1951). A statistical distribution function of wide applicability. Journal of applied mechanics 103, 293-297.

10. Beard, R. E. (1971). Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. Biological aspects of demography 999, 57-68.

11. Vaupel, J., Manton, K.G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16(3): 439-454.

12. Siler, W. (1979), A Competing-Risk Model for Animal Mortality. Ecology, 60: 750-757.

13. Heligman, L., & Pollard, J. (1980). The age pattern of mortality. Journal of the Institute of Actuaries, 107(1), 49-80.

14. Rogers A and Planck F (1983). MODEL: A General Program for Estimating Parametrized Model Schedules of Fertility, Mortality, Migration, and Marital and Labor Force Status Transitions. IIASA Working Paper. IIASA, Laxenburg, Austria: WP-83-102

15. Martinelle S. (1987). A generalized Perks formula for old-age mortality. Stockholm, Sweden, Statistiska Centralbyran, 1987. 55 p. (R&D Report, Research-Methods-Development, U/STM No. 38)

16. Carriere J.F. (1992). Parametric models for life tables. Transactions of the Society of Actuaries. Vol.44

17. Kostaki A. (1992). A nine-parameter version of the Heligman-Pollard formula. Mathematical Population Studies. Vol. 3 277-288

18. Thatcher AR, Kannisto V and Vaupel JW (1998). The force of mortality at ages 80 to 120. Odense Monographs on Population Aging Vol. 5 Odense University Press, 1998. 104, 20 p. Odense, Denmark

19. Tabeau E. (2001) A Review of Demographic Forecasting Models for Mortality. In: Tabeau E., van den Berg Jeths A., Heathcote C. (eds) Forecasting Mortality in Developed Countries. European Studies of Population, vol 9. Springer, Dordrecht

See Also

`MortalityLaw`

Examples

```
availableLaws()
```

availableLF *Check Available Loss Function*

Description

The function returns information about the implemented loss function used by the optimization procedure in the MortalityLaw function.

Usage

```
availableLF()
```

Value

A list of class availableLF with the components:

table	Table with loss functions and codes to be used in MortalityLaw.
legend	Table with details about the abbreviation used.

Author(s)

Marius D. Pascariu

See Also

MortalityLaw

Examples

```
availableLF()
```

convertFx *Convert Life Table Indicators*

Description

Easy conversion between the life table indicators. This function is based on the LifeTable function and methods behind it.

Usage

```
convertFx(x, data, from, to, ...)
```

Arguments

x	Vector of ages at the beginning of the age interval.
data	Vector or data.frame/matrix containing the mortality indicators.
from	Specify the life table indicator in the input data. Character. Options: mx, qx, dx, lx.
to	What indicator would you like to obtain? Character. Options: mx, qx, dx, lx, Lx, Tx, ex.
...	Further arguments to be passed to the LifeTable function with impact on the results to be produced.

Author(s)

Marius D. Pascariu

See Also

LifeTable

Examples

```
# Data ---
x  <- 0:110
mx <- ahmd$mx

# mx to qx
qx <- convertFx(x, data = mx, from = "mx", to = "qx")
# mx to dx
dx <- convertFx(x, data = mx, from = "mx", to = "dx")
# mx to lx
lx <- convertFx(x, data = mx, from = "mx", to = "lx")


# There are 28 possible combinations --------------------------------
# Let generate all of them.
from <- c("mx", "qx", "dx", "lx")
to   <- c("mx", "qx", "dx", "lx", "Lx", "Tx", "ex")
K    <- expand.grid(from = from, to = to) # all possible cases/combinations

for (i in 1:nrow(K)) {
  In  <- as.character(K[i, "from"])
  Out <- as.character(K[i, "to"])
  N <- paste0(Out, "_from_", In)
  cat(i, " Create", N, "\n")
  # Create the 28 sets of results
  assign(N, convertFx(x = x, data = get(In), from = In, to = Out))
}
```

LawTable *Compute Life Tables from Parameters of a Mortality Law*

Description

Compute Life Tables from Parameters of a Mortality Law

Usage

```
LawTable(x, par, law, sex = NULL, lx0 = 1e+05, ax = NULL)
```

Arguments

x	Vector of ages at the beginning of the age interval.
par	The parameters of the mortality model.
law	The name of the mortality law/model to be used. e.g. gompertz, makeham, ... To investigate all the possible options, see availableLaws function.
sex	Sex of the population considered here. Default: NULL. This argument affects the first two values in the life table ax column. If sex is specified the values are computed based on the Coale-Demeny method and are slightly different for males than for females. Options: NULL, male, female, total.
lx0	Radix. Default: 100 000.
ax	Numeric scalar. Subject-time alive in age-interval for those who die in the same interval. If NULL this will be estimated. A common assumption is ax = 0.5, i.e. the deaths occur in the middle of the interval. Default: NULL.

Details

The "life table" is also called "mortality table" or "actuarial table". This shows, for each age, what the probability is that a person of that age will die before his or her next birthday, the expectation of life across different age ranges or the survivorship of people from a certain population.

Value

The output is of the "LifeTable" class with the components:

lt	Computed life table;
call	Call in which all of the specified arguments are specified by their full names;
process_date	Time stamp.

Author(s)

Marius D. Pascariu

See Also

LifeTable MortalityLaw

Examples

```
# Example 1 --- Makeham --- 4 tables ----------
x1 = 45:100
L1 = "makeham"
C1 = matrix(c(0.00717, 0.07789, 0.00363,
              0.01018, 0.07229, 0.00001,
              0.00298, 0.09585, 0.00002,
              0.00067, 0.11572, 0.00078),
            nrow = 4, dimnames = list(1:4, c("A", "B", "C")))

LawTable(x = 45:100, par = C1, law = L1)

# WARNING!!!

# It is important to know how the coefficients have been estimated. If the
# fitting of the model was done over the [x, x+) age-range, the LawTable
# function can be used to create a life table only for age x onward.

# What can go wrong?

# ** Example 1B - is OK.
LawTable(x = 45:100, par = c(0.00717, 0.07789, 0.00363), law = L1)

# ** Example 1C - Not OK, because the life expectancy at age 25 is
# equal with life expectancy at age 45 in the previous example.
LawTable(x = 25:100, par = c(0.00717, 0.07789, 0.00363), law = L1)

# Why is this happening?

# If we have a model that covers only a part of the human mortality curve
# (e.g. adult mortality), in fitting the x vector is scaled down, meaning age (x) becomes
# (x - min(x) + 1). And, the coefficients are estimated on a scaled x in ordered
# to obtain meaningful estimates. Otherwise the optimization process might
# not converge.

# What can we do about it?

# a). Know which mortality laws are rescaling the x vector in the fitting process.
# If these models are fitted with the MortalityLaw() function, you can find out
# like so:
A <- availableLaws()$table
A[, c("CODE", "SCALE_X")]

# b). If you are using one of the models that are applying scaling,
# be aware over what age-range the coefficients have been estimated. If they
# have been estimated using, say, ages 50 to 80, you can use the
# LawTable() to build a life tables from age 50 onwards.


# Example 2 --- Heligman-Pollard -- 1 table ----
x2 = 0:110
L2 = "HP"
```

```
C2 = c(0.00223, 0.01461, 0.12292, 0.00091,
       2.75201, 29.01877, 0.00002, 1.11411)

LawTable(x = x2, par = C2, law = L2)

# Because "HP" is not scaling down the x vector, the output is not affected by
# the problem described above.

# Check
LawTable(x = 3:110, par = C2, law = L2)
# Note the e3 = 70.31 in both tables
```

`LifeTable` *Compute Life Tables from Mortality Data*

Description

Construct either a full or abridged life table with various input choices like: death counts and mid-interval population estimates (`Dx, Ex`) or age-specific death rates (`mx`) or death probabilities (`qx`) or survivorship curve (`lx`) or a distribution of deaths (`dx`). If one of these options is specified, the other can be ignored. The input data can be an object of class: numerical `vector`, `matrix` or `data.frame`.

Usage

```
LifeTable(x, Dx = NULL, Ex = NULL,
             mx = NULL,
             qx = NULL,
             lx = NULL,
             dx = NULL,
             sex = NULL,
             lx0 = 1e5,
             ax  = NULL)
```

Arguments

`x`	Vector of ages at the beginning of the age interval.
`Dx`	Object containing death counts. An element of the `Dx` object represents the number of deaths during the year to persons aged x to x+n.
`Ex`	Exposure in the period. `Ex` can be approximated by the mid-year population aged x to x+n.
`mx`	Death rate in age interval [x, x+n).
`qx`	Probability of dying in age interval [x, x+n).
`lx`	Probability of survival up until age x.
`dx`	Deaths by life-table population in the age interval [x, x+n).

sex	Sex of the population considered here. Default: NULL. This argument affects the first two values in the life table ax column. If sex is specified the values are computed based on the Coale-Demeny method and are slightly different for males than for females. Options: NULL, male, female, total.
lx0	Radix. Default: 100 000.
ax	Numeric scalar. Subject-time alive in age-interval for those who die in the same interval. If NULL this will be estimated. A common assumption is ax = 0.5, i.e. the deaths occur in the middle of the interval. Default: NULL.

Details

The "life table" is also called "mortality table" or "actuarial table". This shows, for each age, what the probability is that a person of that age will die before his or her next birthday, the expectation of life across different age ranges or the survivorship of people from a certain population.

Value

The output is of the "LifeTable" class with the components:

lt	Computed life table;
call	Call in which all of the specified arguments are specified by their full names;
process_date	Time stamp.

Author(s)

Marius D. Pascariu

See Also

LawTable convertFx

Examples

```
# Example 1 --- Full life tables with different inputs ---

y  <- 1900
x  <- as.numeric(rownames(ahmd$mx))
Dx <- ahmd$Dx[, paste(y)]
Ex <- ahmd$Ex[, paste(y)]

LT1 <- LifeTable(x, Dx = Dx, Ex = Ex)
LT2 <- LifeTable(x, mx = LT1$lt$mx)
LT3 <- LifeTable(x, qx = LT1$lt$qx)
LT4 <- LifeTable(x, lx = LT1$lt$lx)
LT5 <- LifeTable(x, dx = LT1$lt$dx)

LT1
LT5
ls(LT5)
```

```
# Example 2 --- Compute multiple life tables at once ---

LTs = LifeTable(x, mx = ahmd$mx)
LTs
# A warning is printed if the input contains missing values.
# Some of the missing values can be handled by the function.

# Example 3 --- Abridged life table ------------

x  = c(0, 1, seq(5, 110, by = 5))
mx = c(.053, .005, .001, .0012, .0018, .002, .003, .004,
       .004, .005, .006, .0093, .0129, .019, .031, .049,
       .084, .129, .180, .2354, .3085, .390, .478, .551)
lt = LifeTable(x, mx = mx, sex = "female")
lt
```

`MortalityLaw` *Fit Mortality Laws*

Description

Fit parametric mortality models given a set of input data which can be represented by death counts and mid-interval population estimates (`Dx`, `Ex`) or age-specific death rates (`mx`) or death probabilities (`qx`). Using the argument `law` one can specify the model to be fitted. So far more than 27 parametric models have been implemented; check the `availableLaws` function to learn about the available options. The models can be fitted under the maximum likelihood methodology or by selecting a loss function to be optimised. See the implemented loss function by running the `availableLF` function.

Usage

```
MortalityLaw(x, Dx = NULL, Ex = NULL, mx = NULL, qx = NULL,
                law = NULL,
                opt.method = "LF2",
                parS = NULL,
                fit.this.x = x,
                custom.law = NULL,
                show = FALSE, ...)
```

Arguments

`x`	Vector of ages at the beginning of the age interval.
`Dx`	Object containing death counts. An element of the `Dx` object represents the number of deaths during the year to persons aged x to x+n.
`Ex`	Exposure in the period. `Ex` can be approximated by the mid-year population aged x to x+n.
`mx`	Death rate in age interval [x, x+n).

qx	Probability of dying in age interval [x, x+n).
law	The name of the mortality law/model to be used. e.g. gompertz, makeham, ... To investigate all the possible options, see availableLaws function.
opt.method	How would you like to find the parameters? Specify the function to be optimize. Available options: the Poisson likelihood function poissonL; the Binomial likelihood function -binomialL; and 6 other loss functions. For more details, check the availableLF function.
parS	Starting parameters used in the optimization process (optional).
fit.this.x	Select the ages to be considered in model fitting. By default fit.this.x = x. One may want to exclude from the fitting procedure, say, the advanced ages where the data is sparse.
custom.law	Allows you to fit a model that is not defined in the package. Accepts as input a function.
show	Choose whether to display a progress bar during the fitting process. Logical. Default: FALSE.
...	Arguments to be passed to or from other methods.

Details

Depending on the complexity of the model, one of following optimization strategies is employed:

1. Nelder-Mead method: approximates a local optimum of a problem with n variables when the objective function varies smoothly and is unimodal. For details see optim
2. PORT routines: provides unconstrained optimization and optimization subject to box constraints for complicated functions. For details check nlminb
3. Levenberg-Marquardt algorithm: damped least-squares method. For details check nls.lm

Value

The output is of the "MortalityLaw" class with the components:

input	List with arguments provided in input. Saved for convenience.
info	Brief information about the model.
coefficients	Estimated coefficients.
fitted.values	Fitted values of the selected model.
residuals	Deviance residuals.
goodness.of.fit	List containing goodness of fit measures like AIC, BIC and log-Likelihood.
opt.diagnosis	Resultant optimization object useful for checking the convergence etc.
stats	List containing statistical measures like: parameter correlation, standard errors, degrees of freedom, deviance, gradient matrix, QR decomposition, covariance matrix etc.

Author(s)

Marius D. Pascariu

See Also

availableLaws availableLF LifeTable ReadHMD

Examples

```
# Example 1: --------------------------
# Fit Makeham Model for Year of 1950.

x  <- 45:75
Dx <- ahmd$Dx[paste(x), "1950"]
Ex <- ahmd$Ex[paste(x), "1950"]

M1 <- MortalityLaw(x = x, Dx = Dx, Ex = Ex, law = 'makeham')

M1
ls(M1)
coef(M1)
summary(M1)
fitted(M1)
predict(M1, x = 45:95)
plot(M1)


# Example 2: --------------------------
# We can fit the same model using a different data format
# and a different optimization method.
x  <- 45:75
mx <- ahmd$mx[paste(x), ]
M2 <- MortalityLaw(x = x, mx = mx, law = 'makeham', opt.method = 'LF1')
M2
fitted(M2)
predict(M2, x = 55:90)

# Example 3: --------------------------
# Now let's fit a mortality law that is not defined
# in the package, say a reparameterized Gompertz in
# terms of modal age at death
# hx = b*exp(b*(x-m)) (here b and m are the parameters to be estimated)

# A function with 'x' and 'par' as input has to be defined, which returns at least
# an object called 'hx' (hazard rate).
my_gompertz <- function(x, par = c(b = 0.13, M = 45)){
  hx  <- with(as.list(par), b*exp(b*(x - M)) )
  return(as.list(environment()))
}

M3 <- MortalityLaw(x = x, Dx = Dx, Ex = Ex, custom.law = my_gompertz)
summary(M3)
# predict M3 for different ages
predict(M3, x = 85:130)
```

```
# Example 4: -------------------------
# Fit Heligman-Pollard model for a single
# year in the dataset between age 0 and 100 and build a life table.

x  <- 0:100
mx <- ahmd$mx[paste(x), "1950"] # select data
M4 <- MortalityLaw(x = x, mx = mx, law = 'HP', opt.method = 'LF2')
M4
plot(M4)

LifeTable(x = x, qx = fitted(M4))
```

`MortalityLaws` *MortalityLaws: Parametric Mortality Models, Life Tables and HMD*

Description

Fit the most popular human mortality 'laws', and construct full and abridge life tables given various input indices. A mortality law is a parametric function that describes the dying-out process of individuals in a population during a significant portion of their life spans. For a comprehensive review of the most important mortality laws see Tabeau (2001) <doi:10.1007/0-306-47562-6_1>. An elegant function for downloading data from Human Mortality Database <https://www.mortality.org> is provided as well.

Details

To learn more about the package, start with the vignettes: `browseVignettes(package = "MortalityLaws")`

Author(s)

Maintainer: Marius D. Pascariu <mpascariu@outlook.com> (0000-0002-2568-6489) [copyright holder]

Other contributors:

- Vladimir Canudas-Romo [contributor]

See Also

Useful links:

- https://github.com/mpascariu/MortalityLaws
- Report bugs at https://github.com/mpascariu/MortalityLaws/issues

plot.MortalityLaw *Plot Function for MortalityLaw*

Description

Plot Function for MortalityLaw

Usage

```
## S3 method for class 'MortalityLaw'
plot(x, ...)
```

Arguments

x	An object of class MortalityLaw
...	Arguments to be passed to methods, such as graphical parameters (see par).

Author(s)

Marius D. Pascariu

See Also

MortalityLaw

Examples

```
# See complete example in MortalityLaw help page
```

predict.MortalityLaw *Predict function for MortalityLaw*

Description

Predict function for MortalityLaw

Usage

```
## S3 method for class 'MortalityLaw'
predict(object, x, ...)
```

Arguments

object	An object of class "MortalityLaw"
x	Vector of ages to be considered in prediction
...	Additional arguments affecting the predictions produced.

Author(s)

Marius D. Pascariu

See Also

MortalityLaw

Examples

```
# Extrapolate old-age mortality with the Kannisto model
# Fit ages 80-94 and extrapolate up to 120.

Mx <- ahmd$mx[paste(80:94), "1950"]
M1 <- MortalityLaw(x = 80:94, mx  = Mx, law = 'kannisto')
fitted(M1)
predict(M1, x = 80:120)

# See more examples in MortalityLaw function help page.
```

ReadHMD *Download Mortality and Population Data (HMD)*

Description

Download detailed mortality and population data for different countries and regions in a single object from the Human Mortality Database.

Usage

```
ReadHMD(what, countries = NULL, interval = "1x1", username, password,
  save = TRUE, show = TRUE)
```

Arguments

what	What type of data are you looking for? There are available: birth records "births", death counts "Dx", deaths by Lexis triangles "lexis", population size "population", exposure-to-risk "Ex", death-rates "mx", life tables for females "LT_f", life tables for males "LT_m", life tables both sexes combined "LT_t", life expectancy at birth "e0", cohort death-rates "mxc" and cohort exposures "Exc".
countries	HMD country codes.
interval	HMD data format: (age interval x year interval). Interval options: 1x1, 1x5, 1x10, 5x1, 5x5, 5x10.
username	Your HMD username. If you don't have one you can sign up for free on the Human Mortality Database website.
password	Your HMD password.

save	Do you want to save a copy of the dataset on your local machine? Logical. Default: FALSE.
show	Choose whether to display a progress bar. Logical. Default: TRUE.

Value

A ReadHMD object that contains:

input	List with the input values (except the password).
data	Data downloaded from HMD.
download.date	Time stamp.
years	Numerical vector with the years covered in the data.
ages	Numerical vector with ages covered in the data.

Author(s)

Marius D. Pascariu

Examples

```
## Not run:
# Download demographic data for 3 countries in 1x1 format
age_int  <- 1  # age interval: 1,5
year_int <- 1  # year interval: 1,5,10
interval <- paste0(age_int, "x", year_int)  # --> 1x1
# And the 3 countries: Sweden Denmark and USA. We have to use the HMD codes
cntr  <- c('SWE', 'DNK', 'USA')

# Download death counts. We don't want to export data outside R.
HMD_Dx <- ReadHMD(what = "Dx",
                  countries = cntr,
                  interval  = interval,
                  username  = "user@email.com",
                  password  = "password",
                  save = FALSE)
ls(HMD_Dx)
HMD_Dx

# Download life tables for female population and export data.
LTF <- ReadHMD(what = "LT_f",
               countries = cntr,
               interval  = interval,
               username  = "user@email.com",
               password  = "password",
               save = TRUE)
LTF

## End(Not run)
```

Index

Bibliography

Aitchison, J. 1982. The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 139–177.

Aitchison, J. 1986. A concise guide to compositional data analysis. *London: Chapman and Hall. 2015*.

Alho, J. M. 1990. Stochastic methods in population forecasting. *International Journal of Forecasting 6*(4), 521–530.

Alho, J. M. 2000. The Lee–Carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal 4*(1), 91–93.

Alho, J. M. and B. D. Spencer. 1991. A population forecast as a database: implementing the stochastic propagation of error. *Journal of Official Statistics 7*(3), 295.

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

Basellini, U. and C. G. Camarda. 2018. Modelling and forecasting adult age-at-death distributions. *Population Studies. Forthcoming*.

Beard, R. E. 1971. Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. *Biological Aspects of Demography 999*, 57–68.

Bell, W. and B. Monsell. 1991. Using principal components in time series modeling and forecasting of age-specific mortality rates. *Alexandria Virginia American Statistical Association*.

Bengtsson, T. 2006. Linear increase in life expectancy: past and present. *Perspective on Mortality Forecasting. III. The Linear Rise in Life Expectancy: History and Present 3*, 83–99.

Bergeron-Boucher, M.-P., V. Canudas-Romo, J. Oeppen, and J. W. Vaupel. 2017. Coherent forecasts of mortality with compositional data analysis. *Demographic Research 37*, 527–566.

Bohk-Ewald, C., P. Li, and M. Myrskylä. 2018. Forecast accuracy hardly improves with method complexity when completing cohort fertility. *Proceedings of the National Academy of Sciences 115*(37), 9187–9192.

Bongaarts, J. 2005. Long-range trends in adult mortality: Models and projection methods. *Demography 42*(1), 23–49.

Booth, H., R. J. Hyndman, L. Tickle, and P. De Jong. 2006. Lee-carter mortality forecasting: a multi-country comparison of variants and extensions. *Demographic Research 15*, 289–310.

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

Box, G. E. and G. M. Jenkins. 1976. Time series analysis, control, and forecasting. *San Francisco, CA: Holden Day 3226*(3228), 10.

Box, G. E., G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. 1970. *Time series analysis: forecasting and control*. John Wiley & Sons 2015.

Brass, W. 1971. *Biological aspects of demography*. Society for the Study of Human Biology Symposium Series. Taylor & Francis.

Brass, W., A. Coale, P. Demeny, D. Heisel, A. Lorimer, A. Romaniuk, and E. Van de Walle. 1968. *The Demography of Tropical Africa*. Princeton University Press.

Brillinger, D. R. 1986. A biometrics invited paper with discussion: the natural variability of vital rates and associated statistics. *Biometrics*, 693–734.

Brouhns, N., M. Denuit, and J. K. Vermunt. 2002. A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics 31*(3), 373–393.

Cairns, A. J., D. Blake, and K. Dowd. 2006. A Two Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration. *Journal of Risk and Insurance 73*(4), 687–718.

Cairns, A. J., D. Blake, K. Dowd, G. D. Coughlan, and M. Khalaf-Allah. 2011. Bayesian stochastic mortality modelling for two populations. *ASTIN Bulletin: The Journal of the IAA 41*(1), 29–59.

Cannan, E. 1895. The probability of a cessation of the growth of population in England and Wales during the next century. *The Economic Journal 5*(20), 505–515.

Carriere, J. F. 1992. Parametric models for life tables. *Transactions of the Society of Actuaries 44*, 77–99.

Chen, H. and S. H. Cox. 2009. Modeling mortality with jumps: Applications to mortality securitization. *Journal of Risk and Insurance 76*(3), 727–751.

Chihara, T. S. 2011. *An introduction to orthogonal polynomials*. Courier Corporation.

Christensen, K. and J. W. Vaupel. 1996. Determinants of longevity: Genetic, environmental and medical factors. *Journal of internal medicine 240*(6), 333–341.

Coale, A. and P. Demeny. 1966. *Regional model life tables and stable populations*. Princeton New Jersey. Princeton Univ. Press 1966.

Coale, A. J., P. G. Demeny, and B. Vaughan. 1983. Models of mortality and age composition. *Regional model life tables and stable population. 2nd ed. New York: Academic Press 193*, 3–7.

De Gans, H. 1999. *Population forecasting 1895–1945: The transition to modernity*, Volume 5. Springer Science & Business Media.

De Jong, P. and L. Tickle. 2006. Extending Lee–Carter mortality forecasting. *Mathematical Population Studies 13*(1), 1–18.

Dellaportas, P., A. F. Smith, and P. Stavropoulos. 2001. Bayesian analysis of mortality data. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 164*(2), 275–291.

DeMoivre, A. 1725. Annuities on lives: Or, the valuation of annuities upon any number of lives as also of reversions. *William Person, London*.

Deng, Y., P. L. Brockett, and R. D. MacMinn. 2012. Longevity/mortality risk modeling and securities pricing. *Journal of Risk and Insurance 79*(3), 697–721.

Dickson, D. C., M. R. Hardy, and H. R. Waters. 2013. *Actuarial mathematics for life contingent risks*. Cambridge University Press.

Doob, J. L. and J. L. Doob. 1953. *Stochastic processes*, Volume 7. Wiley New York.

Dowd, K., A. J. Cairns, D. Blake, G. D. Coughlan, and M. Khalaf-Allah. 2011. A gravity model of mortality rates for two related populations. *North American Actuarial Journal 15*(2), 334–356.

Ediev, D. M. 2011. Robust backward population projections made possible. *International Journal of Forecasting 27*(4), 1241–1247.

Frontini, M., G. Rodriguez, and S. Seatzu. 1990. An algorithm for computing minimum norm solutions of finite moment problem. *Algorithms for Approximation II*, 361–368.

Gabriel, K. and I. Ronen. 1958. Estimates of mortality from infant mortality rates. *Population Studies 12*(2), 164–169.

Gompertz, B. 1825. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical transactions of the Royal Society of London 115*, 513–583.

Graunt, J. 1662. Natural and political observations mentioned in a following index, and made upon the bills of mortality. *London: John Martyn and James Allestry, 1973*.

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

Haberman, S. and A. Renshaw. 2011. A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics 48*(1), 35–55.

Halley, E. 1693. An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the price of annuities upon lives, pt nr. 196.

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

Harper, F. S. 1936. An actuarial study of infant mortality. *Scandinavian Actuarial Journal 1936*(3-4), 234–270.

Hausdorff, F. 1921. Summationsmethoden und momentfolgen. i. *Mathematische Zeitschrift 9*(1-2), 74–109.

Heligman, L. and J. H. Pollard. 1980. The age pattern of mortality. *Journal of the Institute of Actuaries 107*(01), 49–80.

Hill, G. 1993. The entropy of the survival curve: An alternative measure. *Canadian Studies in Population 20*(1), 43–57.

Hoem, J. M. 1983. The reticent trio: some little-known early discoveries in life insurance mathematics by LHF Oppermann, TN Thiele and JP Gram. *International Statistical Review/Revue Internationale de Statistique*, 213–221.

Hollmann, F. W., T. J. Mulder, and J. E. Kallan. 1999. *Methodology & Assumptions for the Population Projections of the United States: 1999 to 2010*. US Department of Commerce, Bureau of the Census, Population Division, Population Projections Branch.

Human Mortality Database 2017a. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Data downloaded on 10/06/2017. URL: <https://www.mortality.org>.

Human Mortality Database 2017b. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Data downloaded on 07/12/2017. URL: <https://www.mortality.org>.

Human Mortality Database 2018. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Data downloaded on 21/08/2018. URL: <https://www.mortality.org>.

Hyndman, R. J. 2017. *demography: Forecasting Mortality, Fertility, Migration and Population Data*. R package version 1.20.

Hyndman, R. J. and G. Athanasopoulos. 2018. *Forecasting: principles and practice*. OTexts. Available online at https://otexts.org/fpp2/.

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

Hyndman, R. J. and A. B. Koehler. 2006, October. Another look at measures of forecast accuracy. *International Journal of Forecasting 22*(4), 679–688.

Hyndman, R. J. and M. S. Ullah. 2007. Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics & Data Analysis 51*(10), 4942–4956.

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

John, V., I. Angelov, A. Öncül, and D. Thévenin. 2007. Techniques for the reconstruction of a distribution from a finite number of its moments. *Chemical Engineering Science 62*(11), 2890–2904.

Kannisto, V., J. Lauritsen, A. R. Thatcher, and J. W. Vaupel. 1994. Reductions in mortality at advanced ages: several decades of evidence from 27 countries. *Population and Development Review*, 793–810.

Keyfitz, N. 1977. *Applied mathematical demography*. John Wiley and Sons.

Kirkwood, T. B. 2015. Deciphering death: a commentary on Gompertz (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Phil. Trans. R. Soc. B 370*(1666), 20140379.

Kostaki, A. 1992. A nine-parameter version of the Heligman–Pollard formula. *Mathematical Population Studies 3*(4), 277–288.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, and Y. Shin. 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? *Journal of Econometrics 54*(1-3), 159–178.

Laplace, P.-S. 1825. *Pierre-Simon Laplace Philosophical Essay on Probabilities: Translated from the fifth French edition of 1825 with notes by the Translator*, Volume 13. Springer Science & Business Media (2012).

Ledermann, S. 1969. *Nouvelles tables-types de mortalité*. Travaux et Documents - Institut National d'Études Démographiques. Presses Universitaires de France.

Lee, R. D. and L. R. Carter. 1992. Modeling and forecasting US mortality. *Journal of the American statistical association 87*(419), 659–671.

Lexis, W. H. R. A. 1879. *Sur la durée normale de la vie humaine et sur la théorie de la stabilité des rapports statistiques*. Vve. F. Henry.

Li, J. S.-H., W.-S. Chan, and S.-H. Cheung. 2011. Structural changes in the Lee–Carter mortality indexes: detection and implications. *North American Actuarial Journal 15*(1), 13–31.

Li, J. S.-H. and M. R. Hardy. 2011. Measuring basis risk in longevity hedges. *North American Actuarial Journal 15*(2), 177–200.

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

Li, N., R. Lee, and P. Gerland. 2013. Extending the Lee-Carter method to model the rotation of age patterns of mortality decline for long-term projections. *Demography 50*(6), 2037–2051.

Makeham, W. M. 1867. On the law of mortality. *Journal of the Institute of Actuaries (1866-1867) 13*(6), 325–358.

Malthus, T. R. 1798. *An Essay on the Principle of Population*. London. J. Johnson.

Martinelle, S. 1987. *A generalized Perks formula for old-age mortality.* Stockholm Sweden Statistiska Centralbyran 1987.

Mayhew, L. and D. Smith. 2013. A new method of projecting populations based on trends in life expectancy and survival. *Population Studies 67*(2), 157–170.

Mazzuco, S., B. Scarpa, and L. Zanotto. 2018. A mortality model based on a mixture distribution function. *Population Studies*, 1–10.

McNown, R. and A. Rogers. 1989. Forecasting mortality: A parameterized time series approach. *Demography 26*(4), 645–660.

McNown, R. and A. Rogers. 1992. Forecasting cause-specific mortality using time series methods. *International Journal of Forecasting 8*(3), 413–432.

Mead, L. R. and N. Papanicolaou. 1984. Maximum entropy in the problem of moments. *Journal of Mathematical Physics 25*(8), 2404–2417.

Messner, J. and A. Zeileis. 2015. Censored Regression with Conditional Heteroscedasticity, Version: 0.9-1. *CRAN - electronic resource*.

Mitchell, D., P. Brockett, R. Mendoza-Arriaga, and K. Muthuraman. 2013. Modeling and forecasting mortality rates. *Insurance: mathematics and economics 52*(2), 275–285.

Møller, T. and M. Steffensen. 2007. *Market-valuation methods in life and pension insurance*. Cambridge University Press.

Murray, C. J., B. D. Ferguson, A. D. Lopez, M. Guillot, J. A. Salomon, and O. Ahmad. 2003. Modified logit life table system: principles, empirical validation, and application. *Population Studies 57*(2), 165–182.

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

Oeppen, J. 2008. Coherent forecasting of multiple-decrement life tables: a test using japanese cause of death data. *Paper presented at the European Population Conference 2008, Barcelona, Spain, July 9-12*.

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

Olshansky, S. J. and B. A. Carnes. 1997. Ever since Gompertz. *Demography 34*(1), 1–15.

Oppermann, L. H. F. 1870. On the graduation of life tables, with special application to the rate of mortality in infancy and childhood. *The Insurance Record Minutes from a meeting in the Institute of Actuaries.*, 42.

Pascariu, M. D. 2018. *MortalityLaws: Parametric Mortality Models, Life Tables and HMD*. R package version 1.6.0.

Pascariu, M. D., V. Canudas-Romo, and W. J. Vaupel. 2018. The double-gap life expectancy forecasting model. *Insurance: Mathematics and Economics*.

Pearson, K. 1897. The chances of death, and other studies in evolution. *Folklore Enterprises, Ltd. 8*(4), 370–372.

Perks, W. 1932. On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries (1886-1994) 63*(1), 12–57.

Perls, T. T. and R. C. Fretts. 1998. Why Women Live Longer than Men - What gives women the extra years? *Scientific American* (2), 100–103.

Pollard, J. H. 1987. Projection of age-specific mortality rates. *Population Bulletin of the United Nations* (21/22), 55–69.

Preston, S., P. Heuveline, and M. Guillot. 2000. Demography: measuring and modeling population processes.

R Core Team 2018. R: A language and environment for statistical computing [Internet]. Vienna, Austria. R version 3.5.0 (2018-04-23).

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

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

Rau, R., E. Soroko, D. Jasilionis, and J. W. Vaupel. 2008. Continued reductions in mortality at advanced ages. *Population and Development Review 34*(4), 747–768.

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

Rogers, A. and F. Planck. 1984. Parameterized multistage population projections. *Working Paper for presentation at the Annual Meeting of the Population Association of America, Minnesota, May 3-5.*.

Saboia, J. L. 1974. Modeling and forecasting populations by time series: The Swedish case. *Demography 11*(3), 483–492.

Saboia, J. L. M. 1977. Autoregressive integrated moving average (ARIMA) models for birth forecasting. *Journal of the American Statistical Association 72*(358), 264–270.

Sacher, G. A. 1977. Life table modification and life prolongation. *Handbook of the Biology of Aging 582*.

Ševčíková, H., N. Li, V. Kantorová, P. Gerland, and A. E. Raftery. 2016. Age-specific mortality and fertility rates for probabilistic population projections. pp. 285–310.

Shannon, C. E. 1948. A mathematical theory of communication. *Bell System Technical Journal 27 (3)*, 379–423.

Shkolnikov, V. M., D. A. Jdanov, E. M. Andreev, and J. W. Vaupel. 2011. Steep increase in best-practice cohort life expectancy. *Population and Development Review 37*(3), 419–434.

Shohat, J. A. and J. D. Tamarkin. 1943. *The problem of moments*. Number 1. American Mathematical Soc.

Siler, W. 1983. Parameters of mortality in human populations with widely varying life spans. *Statistics in medicine 2*(3), 373–380.

Steffensen, J. 1930. Infantile mortality from an actuarial point of view. *Skandinavisk Aktuarietidskrift 13*, 272–286.

Stoeldraijer, L., C. van Duin, L. van Wissen, and F. Janssen. 2013. Impact of different mortality forecasting methods and explicit assumptions on projected future life expectancy: The case of the Netherlands. *Demographic Research 29*, 323–353.

Sullivan, J. M. 1972. Models for the estimation of the probability of dying between birth and exact ages of early childhood. *Population Studies 26*(1), 79–97. PMID: 22070174.

Tabeau, E., A. van den Berg Jeths, and C. Heathcote. 2001. *Forecasting mortality in developed countries: Insights from a statistical, demographic and epidemiological perspective*, Volume 9. Springer Science & Business Media.

Thatcher, A. R., V. Kannisto, and J. W. Vaupel. 1998. The force of mortality at ages 80 to 120. *Odense Denmark: Odense University Press 1998.*.

Thiele, T. N. 1871. On a mathematical formula to express the rate of mortality throughout the whole of life. *Journal of the Institute of Actuaries and Assurance Magazine. Translated by T.B. Sprague in 1872. 16*(5), 313–329.

Torri, T. and J. W. Vaupel. 2012. Forecasting life expectancy in an international context. *International Journal of Forecasting 28*(2), 519–531.

United Nations 1955. *Age and Sex Patterns of Mortality: Model Life Tables for Under-Developed Countries*. Population Studies, No. 22. Department of Social Affairs.

United Nations 1967. *Methods for Estimating Basic Demographic Measures from Incomplete Data: Manuals and Methods of Estimating Populations*. Manual IV. Population Studies, No. 42. Department of Social Affairs.

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

United States Census Bureau 2014. *Methodology, Assumptions, and Inputs for the 2014 National Projections*. U.S. Census Bureau.

Vallin, J. and F. Meslé. 2009. The segmented trend line of highest life expectancies. *Population and Development Review 35*(1), 159–187.

Vaupel, J. W. 1997. The remarkable improvements in survival at older ages. *Philosophical Transactions of the Royal Society of London B: Biological Sciences 352*(1363), 1799–1804.

Vaupel, J. W., J. R. Carey, K. Christensen, T. E. Johnson, A. I. Yashin, N. V. Holm, I. A. Iachine, V. Kannisto, A. A. Khazaeli, P. Liedo, et al. 1998. Biodemographic trajectories of longevity. *Science 280*(5365), 855–860.

Vaupel, J. W., K. G. Manton, and E. Stallard. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography 16*(3), 439–454.

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

Warren, W. 1949. Recent contributions to the mathematical theory of communication. *The Mathematical Theory of Communication. Urbana. IL: University of Illinois Press*.

Weibull, W. 1951. A statistical distribution function of wide applicability. *Journal of applied mechanics 103*, 293–297.

Whelpton, P. K. 1928. Population of the United States, 1925 to 1975. *American Journal of Sociology 34*(2), 253–270.

Whelpton, P. K., H. T. Eldridge, and J. S. Seigel. 1948. *Forecasts of the Population of the United States, 1945-1975.* U.S. Government Publishing Office.

White, K. M. 2002. Longevity advances in high-income countries, 1955–96. *Population and Development Review 28*(1), 59–76.

Wicksell, S. D. 1926. Sveriges framtida befolkning under olika förutsättningar. *Ekonomisk tidskrift*, 91–123.

Wiebols, G. A. H. 1925. *De toekomstige bevolkingsgrootte in Nederland. PhD thesis.* NV Dorsman Odé's bookshop and printing company.

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

Wilmoth, J. R., K. Andreev, D. Jdanov, D. A. Glei, C. Boe, M. Bubenheim, D. Philipov, V. Shkolnikov, and P. Vachon. 2007. Methods protocol for the human mortality database. *University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock. 9*, 10–11.

Wilmoth, J. R. and S. Horiuchi. 1999. Rectangularization revisited: Variability of age at death within human populations. *Demography 36*(4), 475–495.

Wittstein, T. and D. Bumsted. 1883. The mathematical law of mortality. *Journal of the Institute of Actuaries and Assurance Magazine 24*(3), 153–173.

Yang, S. S., J. C. Yue, and H.-C. Huang. 2010. Modeling longevity risks using a principal component approach: A comparison with existing stochastic mortality models. *Insurance: Mathematics and Economics 46*(1), 254–270.