Bardoutsos *et al. Genus* (2018) 74:17
https://doi.org/10.1186/s41118-018-0039-5

Genus

ORIGINAL ARTICLE Open Access

Projecting delay and compression of mortality

Anastasios Bardoutsos[1*], Joop de Beer[2] and Fanny Janssen[1,2]

* Correspondence:
a.bardoutsos@rug.nl
[1]Population Research Centre, University of Groningen, Groningen, The Netherlands
Full list of author information is available at the end of the article

Abstract

Background: Although mortality delay (the shift of the age-at-death distribution to older ages) and mortality compression (less variability in the age at death) are the key dynamics that drove past mortality trends, they have seldom been included in mortality projections.

Objective: We compare the projections of a new parametric mortality model that captures delay and compression of mortality (CoDe) with projections based on the well-known Lee-Carter (LC) model.

Data and methods: We compare the two models' properties and in-sample and out-of-sample performance using data from 1960 to 2014 for French, Japanese, and American women and men.

Results: The CoDe model has less parameters to describe the shape of the age pattern, but more parameters to describe the changes in the age pattern, provides extrapolation to higher ages, allows to estimate the modal age at death, does not assume the exponential decline of rates across all ages, decomposes the delay and compression effect, and can serve as a diagnostic tool. While the LC model provides a better fit at younger ages, the CoDe model provides a better fit at older ages. The LC model consistently projects a slowdown of mortality delay and thus of the increase in life expectancy at birth, whereas the CoDe model can project a continuation of delay and thus a steady increase in life expectancy.

Conclusion: Projecting mortality by including mortality delay and compression can result in better forecast performance than using the LC model, especially when the modal age at death increases linearly.

Keywords: Mortality projections, Delay, Compression, Age-at-death distribution

Introduction

Over the past three decades, demographers, actuaries, and epidemiologists have made considerable efforts to produce accurate mortality projections. While the unforeseen extension of longevity is beneficial for society, it comes with financial consequences (International Monetary Fund 2016). Accurate mortality projections are essential to ensure the solvency of government and private pension providers.

Mortality projections are usually based on the analysis of past trends in age-specific mortality rates (see Booth and Tickle 2008; Cairns et al. 2009; Dowd et al. 2010a, 2010b; Cairns et al. 2011; Stoeldraijer et al. 2013, for an extensive review and a comparison of mortality projection models). Currently, the best-known mortality projection

model is the Lee-Carter model (Lee and Carter 1992), which explains the age and period patterns of central death rates using a single period-dependent parameter and two age-dependent parameters. The Lee-Carter model is the "forefather" of many of the well-known mortality models described in the literature, such as the models proposed by Lee and Miller 2001, Booth et al. 2002, Renshaw and Haberman 2006, Plat 2009, and Kleinow 2015. The Lee-Carter model is regarded as a benchmark model, despite having some shortcomings (Booth et al. 2002).

Recently, mortality researchers have shown increased interest in studying trends in the age-at-death distribution instead of age-specific mortality rates (Wilmoth and Horiuchi 1999; Cheung et al. 2005; Canudas-Romo 2008; Canudas-Romo 2010; Ediev 2013; de Beer and Janssen 2016; Basellini et al. 2016; Bergeron-Boucher et al. 2015).

One of the main reasons for analyzing the age-at-death distribution is that doing so allows researchers to distinguish between delay and compression of mortality (Bergeron-Boucher et al. 2015; de Beer and Janssen 2016; Basellini et al. 2016). Delay is defined as the shift of the age-at-death distribution to the right (Vaupel 2010), which is reflected in an increase in the modal age at death (Canudas-Romo 2008). Compression is defined as a change in the shape of the age-at-death distribution resulting from a decline in the variability of the age at dying (Fries 1980; Wilmoth and Horiuchi 1999; Kannisto 2000). Until the 1970s, increases in life expectancy in low-mortality countries were largely attributable to a decline in infant and child mortality and consequently compression. But in recent years, delay of mortality has been the main cause of increases in life expectancy in these countries (de Beer and Janssen 2016).

However, to the best of our knowledge, very few of the existing mortality projection models have made use of the two dynamics that drive the changes in the age-at-death distribution. Bongaarts (2005) projected mortality for adult Swedish women based on the assumption that improvements in (senescent) mortality rates are the result of mortality delay. However, he did not consider changes in the shape of the age-at-death distribution, i.e., the compression effect. Two recently developed projection methods that include both compression and delay can only be applied to the adult population. Terblanche (2016) projected mortality for Australian men and women aged 50–100 based on the linear extrapolation of both the modal age and the concentration of deaths around the modal age. Basellini et al. (2016) recently proposed a methodology for modeling and forecasting adult mortality that is based entirely on the modal age and variance in the age-at-death distribution. Moreover, two mortality models have been proposed that capture mortality delay and mortality compression, i.e., the adapted Siler model by Bergeron-Boucher et al. (2015) and the compression and delay (CoDe) model by de Beer and Janssen (2016). Although the CoDe model has been used to project mortality for Japanese women (de Beer et al. 2017), the method has not yet been applied to other countries.

Our objective is to project mortality across all ages by examining mortality delay and compression using the compression and delay (CoDe) model. In addition, to assess the advantages and disadvantages of this new projection model, we compare it qualitatively and quantitatively with the well-known Lee-Carter model.

Methods and data

In this section, we introduce and discuss the CoDe model for projecting mortality using the concepts of delay and compression. We then briefly describe the Lee-Carter (LC)

model, and outline the qualitative and quantitative methods we use to compare the CoDe model with the Lee-Carter (LC) model.

Description of the CoDe model

To project mortality using delay and compression, we will use the compression and delay (CoDe) model. The CoDe model is a non-linear parametric mortality model for modeling mortality and forecasting the full age range that distinguishes between mortality delay and mortality compression (de Beer and Janssen 2016). We refer to our model as CoDe 2.1, as it is a slight adaptation of the CoDe 2.0 version (de Beer et al. 2017)—which was in turn a slight adaptation of the original CoDe model (de Beer and Janssen 2016). The main difference between CoDe 2.0 and CoDe 2.1 is that we followed the principle of parsimony and limited the number of time-varying parameters as much as possible, as this may be expected to result in more robust projections.

The CoDe 2.1 model is formulated as follows:

$$q_{x,t} = \frac{A_t}{xB+1} + \frac{e^{(a_t+x-z)}}{1+e^{(x-z)}} + \frac{b_{x,t}e^{b_{x,t}(x-M_t+h_1)}}{1+\frac{b_{x,t}}{g}e^{b_{x,t}(x-M_t+h_1)}} \tag{1}$$

with

$$b_{x,t} = \beta_1 + \left(\beta_{3,t}-\beta_1\right)\frac{e^{\beta_2(x-M_t+h_2)}}{1+e^{\beta_2(x-M_t+h_2)}}, \tag{2}$$

where $q_{x,t}$ denotes the death probability at age x in year t. The CoDe 2.1 model includes 10 parameters, of which four are assumed to change across time, i.e., $A_t \in (0,1)$, $a_t < 0$, $M_t > 0$, and $\beta_{3,t} > 0$; six are assumed to be constant across time, i.e., $\beta_1 > 0$, $\beta_2 > 0$, $B > 0$, $g \in (0,1)$, $h_2 > 0$ and $z > 0$ and $h_1 > 0$ is a constant, that improves the fit of parameter M_t on the modal age. For more details for the calibration of the model and the bounds necessary for the calibration of the model see Appendix 1.

The first term of Eq. 1 describes the sharp decline in mortality at young ages from a relatively high value at age zero to a very low value in young childhood. Parameter $A_t \in (0,1)$ reflects the level of infant mortality in year t, and $B > 0$ affects the rate of decrease in mortality in young childhood.

The so-called accident hump of mortality is described by the second term of Eq. 1, which is a logistic function. The value of $\exp a_t$ reflects the so-called background mortality (Bongaarts 2005), i.e., the part of mortality that does not vary with age in adulthood. Makeham (1860) was the first to introduce a term that does not change with age. For ages above 25, the first term of the CoDe 2.1 model is zero and the second term converges to a constant value, i.e., $\exp a_t$. Thus, the second term describes the level of background mortality. Adulthood mortality starts to increase strongly around age z.

The third term of Eq. 1 is a logistic function describing adult and old-age mortality. The logistic assumption ensures that the death probabilities will eventually slow down (de Beer et al. 2017, extended data Fig. 1). Thatcher (1999) and Kannisto (1994) showed that—in general—the logistic model provides a better fit at the oldest ages than the Gompertz model (Gompertz 1825), even though for the US there is some debate whether this conclusion is true (Gavrilov and Gavrilova 2011; de Beer et al. 2017, supplementary information).

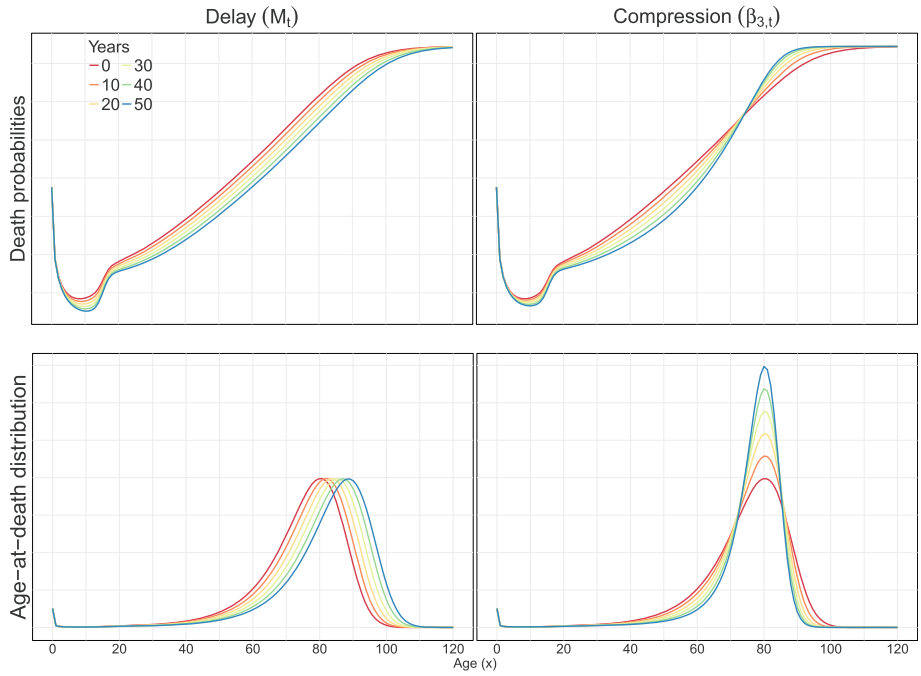

Fig. 1 The impact of increases in the modal age at death M_t (i.e., delay) and of increases in $\beta_{3,t}$ (i.e., compression) on the death probabilities and the age-at-death distribution. (We assume an increase of modal age by 2 months per year and a 5% increase of the compression parameter per 10 years. The values for year zero are chosen such that we observe the well-known mortality pattern by age)

In our logistic function, parameter M_t approximates the modal age at death; whereas the slope of the logistic function, i.e., $b_{x,t}$, is an age- and time-dependent function with four parameters. An increase in parameter M_t reflects the delay effect, whereas an increase in parameter $\beta_{3,t}$ reflects the compression effect. In Fig. 1, we present the impact of increases in M_t (i.e., delay) and of increases in $\beta_{3,t}$ (i.e., compression) on the death probabilities and on the age-at-death distribution. We observe that the age-at-death distribution shifts to the right if M_t increases and the shape remains the same. If we change the value of the compression parameter $\beta_{3,t}$, we observe that the age-at-death distribution does not shift to the right but the shape changes. At age $M_t - h_2$ we observe an increase in the density, indicating compression. At progressively higher ages, the death probabilities level off to a maximum value of g. For a collection of countries with reliable data, Gampe (2010) estimated that the death probability above age 110 is equal to 0.5. The Human Mortality Database assumes that the death rates have an upper bound equal to one (Wilmoth et al. 2017), which translates into an upper bound for death probabilities that are equal to 0.63 (based on Eq. 4). On the other hand, Rau et al. (2017) showed that the death probabilities at very old ages vary by gender and country. Thus, we assume that parameter g is not time dependent, but we allow it to vary by country and gender. Since the death probabilities level off at very old ages, and we use mortality data up to maximum age of 100 in the current application, our assumption for g is expected to minimally affect our results.

Description of the Lee-Carter model

The Lee-Carter (LC) model is a log-bilinear model developed to fit the death rates surface. For each age and year, it has two age-dependent parameters and a 1-year-dependent parameter. The LC model has the following simple form

$$\log\left(\mu_{x,t}\right) = \alpha_x + \beta_x \kappa_t, \tag{3}$$

where $\mu_{x,t}$ is the death rate at age x in year t. For the identifiability of the parameters of the LC model, two non-unique constraints are required. Based on the constraints proposed by Lee and Carter (1992), α_x reflects the average age-specific death rates. The changes in the standardized age-specific death rates log $\mu_{x,t}$ are decomposed into a time effect κ_t, which captures the average improvement over time across all ages; and an age effect β_x, which adjusts the rate of mortality decline for each age.

We select the LC model as a benchmark model, foremost because it is a well-known model in the literature. In addition, however, the CoDe 2.1 and the LC model have a fairly equal number of parameters (see Table 1), which enables to focus the comparison on the most important aspect on which the two models differ: the ability to project delay and compression of mortality.

Qualitative comparison

The analysis and the comparison of mortality models is usually based on quantitative criteria, such as measures of goodness of fit or the ex-post facto comparison of demographic indices, such as life expectancy (Cairns et al. 2009; Bohk-Ewald et al. 2017).

Table 1 Comparison of the underlying assumptions and properties of the CoDe 2.1 model and the LC model

Properties	Model	
	LC	CoDe 2.1
Mathematical form	Models death rates and is linear	Models death probabilities and is non-linear
Total number of parameters	Two parameters for each age, one parameter for each year, minus two parameters for identification	Six parameters to describe the age pattern plus four parameters for each year
Modeling of the age pattern	More flexible to capture age-specific effects due to many age-dependent parameters	Fixed mathematical form to describe the age pattern
Extrapolation to high ages	Additional assumptions are required, such as additional parameters and a two-step approach	Is part of the model. Death probabilities to higher ages are asymptotically equal to parameter g
Estimating modal age	Limited, only integer values	Possible
Modeling of the mortality trend	One parameter explains the mortality trends across all ages	For each age group, there is a parameter to explain the mortality trends
Ability to distinguish between compression and delay	Cannot decompose	Decomposes the two effects and projects them
Speed of mortality decline	Death rates decrease exponentially for all ages	Non-exponential decay for all age groups
Incorporation of exogenous variables and expert opinion	Very limited	Possible
Ability to use as diagnostic tool	Not possible	Possible

Nevertheless, performing a qualitative comparison is also very important, as a comparison of this kind is needed to understand and interpret the causes of differences across mortality projections. For the qualitative comparison of the CoDe 2.1 model and the LC model, we studied the models' underlying assumptions and properties, i.e., the mathematical form, the number of parameters, the modeling of the age pattern, the extrapolation of mortality to high ages, the modeling of mortality trends, the ability to distinguish between delay and compression, the speed of the mortality decline, the incorporation of exogenous variables and expert opinion in the projections, and the extent to which the model can be used as an ex-post diagnostic tool.

Quantitative comparison

The quantitative comparison of the two models is based on an empirical application in which we compare the in-sample fit and the out-of-sample and the in-sample projections of life expectancy at birth and the modal age at death using data on recent mortality trends for three low-mortality countries.

Empirical application

For our empirical analysis, we obtained annual sex- and age-specific deaths and exposures from the Human Mortality Database (HMD) for France, Japan, and the USA for the 1960–2014 period and for ages 0–100 (Human Mortality Database n.d.). We then used these data to calculate unsmoothed death probabilities and death rates. The reason for selecting these three low-mortality countries is that they show clearly different trends in the modal age at death (de Beer and Janssen 2016). The modal age of French women is the highest among European populations and increased linearly from 1960 onward (Ouellette and Bourbeau 2011), whereas the modal age of American women was high in 1960 but decelerated in the 1980s and the 1990s. Japanese women have displayed the highest modal age since the beginning of the twenty-first century, along with the highest (non-linear) improvement rate in the world (de Beer et al. 2017). Furthermore, for these countries, the CoDe model has been applied before (de Beer and Janssen 2016 and de Beer et al. 2017).

Note that we have used death rates to calibrate the LC model and the death probabilities to fit the CoDe 2.1 model. We assume that the death probabilities can be estimated from death rates by

$$q_{x,t} = 1 - e^{-\mu_{x,t}}. \tag{4}$$

Comparison of in-sample fit

For the comparison of the models' in-sample fit, we have made sure that the calibration methodologies for the two models are similar (see Appendix 1). The calibration period is from 1960 until 2014 and for ages zero to 100.

For the calibration of the CoDe 2.1 model, we follow de Beer and Janssen (2016) and de Beer et al. (2017). We use total weighted least squares: namely, we minimize

$$\frac{1}{3(t_{\max} - t_{\min} + 1)} \sum_{t=t_{\min}}^{t_{\max}} \left[\frac{\mathrm{MSE}(\log q_{x,t})}{\mathrm{Var}(\log q_{x,t})} + \frac{\mathrm{MSE}(q_{x,t})}{\mathrm{Var}(q_{x,t})} + \frac{\mathrm{MSE}(d_{x,t})}{\mathrm{Var}(d_{x,t})} \right], \tag{5}$$

where $d_{(x,t)}$ denotes the age-at-death distribution at age x and calendar year t.

Consequently, we maximize the average coefficient of determination, $R^2 := 1 - \left(\frac{\text{MSE}}{\text{Var}}\right)$, of the log-death probabilities log $q_{x,t}$, the death probabilities $q_{x,t}$, and the life table age-at-death $d_{x,t}$.

We calibrated the LC model using singular value decomposition (SVD), as proposed by Lee and Carter (1992), which results in total least squares minimization. As it is commonly employed to compare the in-sample fit of the models, we use the root mean square error (RMSE).

Time series framework for projections

To project the four time-dependent parameters (A_t, α_t, M_t, $\beta_{3,t}$) of the CoDe 2.1 model and to estimate the future level of the death probabilities, we used autoregressive integrated moving average (ARIMA) models (Box and Jenkins 1970). For each population and parameter, we employ a different ARIMA model. We use a variation of the Hyndman-Khandakar algorithm for automatic ARIMA selection, which can be found in the *forecast* package (Hyndman and Khandakar 2008) in **R** (R Core Team 2017). The selection of the best time series is based on the corrected Akaike's Information Criterion (AICc) (Hurvich and Tsai 1989).

For the projections of the time-dependent parameter κ_t of the LC model, we assumed a random walk (RW) with drift, i.e.,

$$\kappa_{t+1} = \kappa_t + \delta + \varepsilon_t, \tag{6}$$

where δ is the drift and ε_t is the error term that follows a normal distribution with zero mean and variance σ^2. The RW—which represents a special case of an ARIMA model, i.e., ARIMA (0,1,0)—is the most common assumption for projecting κ_t that appears in the literature (Lee and Carter 1992; Renshaw and Haberman 2006). Using a more general ARIMA model, as in the case of the CoDe 2.1 model, is expected to minimally influence our results due to the linear decline of the time-dependent parameter κ_t in the LC model.

In-sample projections (backtesting)

Using these time series frameworks, we benchmark the forecast performance of the CoDe 2.1 model against the LC model using in-sample projections of life expectancy at birth and modal age. Life expectancy at birth, e_0, is a key measure of mortality trends that is used in life tables. But while future levels of life expectancy at birth are commonly derived from mortality projections, it is not an efficient measure of the reliability of age-specific projections because different kinds of mortality improvements can result in similar estimations of life expectancy at birth (Bohk-Ewald et al. 2017). For instance, an underestimation in the projections of adult mortality can be compensated for by an overestimation of premature mortality. We therefore use the modal age at death as a complementary measure to evaluate the forecast performance. Improvements in the modal age indicate mortality delay, which has been the main cause of life expectancy improvement in recent decades.

For our illustration, we fit both models for the 1960–2004 period (45 years in total) and project the time-dependent parameters from 2005 to 2014 (10 years in total). We visually compare the projections with the estimations of life expectancy at birth and modal age for the 2004–2014 period. In choosing our backtest strategy, we followed

recent literature which also often forecasts up to 10 years (e.g., Li and O'Hare 2017). Using a longer period for the backtest or a different strategy is not expected to alter the backtest results of the modal age given the clear linear past trends of the modal age for women and for men after 1970 (Janssen and de Beer: The timing of the transition from mortality compression to mortality delay in Europe, Japan, and the United States, submitted).

Out-of-sample projections

For both models, we project up to 2044 the death rates, the age-at-death distribution, life expectancy at birth, and the modal age at death. We then compare the point estimates and forecast intervals. We did not choose a longer out-of-sample projection period, since the differences between the two models in the projections of the modal age and life expectancy are already evident with the current out-of-sample projection period.

As the LC model cannot produce estimations and projections of death rates beyond the age sample, it is not possible to estimate life expectancy based only on the death rate estimates of the LC model. To extend the age range up to 130, we used the Kannisto model (Thatcher et al. 1998), following previous literature (Ševčíková et al. 2016; Antonio et al. 2017; Bohk-Ewald et al. 2017).

Results

Qualitative comparison

Table 1 compares the underlying assumptions and properties of the CoDe 2.1 model and the LC model.

There are fundamental differences in the mathematical formulations of the two models. The CoDe 2.1 model is a non-linear model that describes the death probabilities. The LC model is a bilinear model (age and period dimensions) that describes the logarithmic transformation of death rates. Compared to the CoDe 2.1 model, the LC model allows for easier and faster calibration of the parameters due to its linear form.

To avoid model complexity, it is desirable to use models with a limited number of parameters. The LC model has two parameters for each age and one parameter for each year. However, two parameters are excluded due to the two constraints that are required for the identifiability of the parameters (Lee and Carter 1992). The CoDe 2.1 model has six parameters to describe the age pattern, plus four parameters for each year. In most cases, thus, the LC model will end up with more parameters than the CoDe 2.1 model. For instance, if we fit the two models using deaths rates from 1960 to 2014 (55 years) and for ages 0–100 (101 ages), the total number of parameters is 255 for the LC model, and 225 for the CoDe 2.1 model. While the LC model would have fewer parameters than the CoDe 2.1 model if we used data covering a period of more than 70 years and for ages 0–100, the difference would not be substantial. For any additional year in the sample, the CoDe 2.1 model requires four additional parameters, while the LC model requires one additional parameter. For any additional age in the sample, the LC model requires two parameters, whereas the CoDe 2.1 model requires no parameters.

These differences in the number of parameters also influenced how the age pattern is captured by the two models. The CoDe 2.1 model uses a combination of three age-dependent functions to describe the age pattern of the death probabilities, whereas

the LC model requires two age-specific parameters to capture the shape of the log-deaths rates. Because of its many age-dependent parameters, the LC model is more flexible than the CoDe 2.1 model in capturing the age pattern.

The estimation of life expectancy requires the estimation of death rates beyond the age sample. By design, the LC model cannot be used to estimate death rates outside of the age sample. A potential solution to this problem is to use a parametric mortality model, such as the Kannisto law model (Kannisto 1996), and to combine it with the LC estimations of death rates (Ševčíková et al. 2016; Antonio et al. 2017; Bohk-Ewald et al. 2017). When using the CoDe 2.1 model, we do not encounter this problem because the estimates of the death probabilities at higher ages stem from Eqs. 1 and 2. Moreover, in the CoDe 2.1 model, the death probabilities level off at higher ages and asymptotically approach the value of parameter g. In the LC model, by contrast, we cannot estimate such a level a priori.

Another important lifespan measure is the modal age at death (e.g., Canudas-Romo 2008). The LC model can only produce integer values of the modal age at death, as a result of the fact that the LC model estimates an age-discrete age-at-death distribution. The CoDe model, however, can be used to estimate the modal age at death. Due to the non-linear and complex structure of the CoDe 2.1 model, we unfortunately cannot prove mathematically that the parameter M_t is equal to modal age. Nevertheless, our empirical analysis shows that M_t is a good estimator of the modal age, see Fig. 3.

Capturing the mortality trends is essential when making projections. In the LC model, there is a single time parameter, i.e., κ_t, that explains the improvements over time across all ages. Combined with the log-linear form assumption, the LC model assumes a continuation of the changes in the death rates observed in the reference period. Indeed, from Eqs. 3 and 6, we can conclude that the relative changes in the death rates are constant over time (Haberman and Renshaw 2012); i.e.,

$$\frac{\mu_{x,t+1}}{\mu_{x,t}} - 1 = \frac{\mu_{x,t}}{\mu_{x,t-1}} - 1. \tag{7}$$

For instance, if in the past there were greater relative improvements in infant mortality than in old-age mortality, the LC model will project a continuation of this pattern. Thus, Seligman et al. (2016) concluded that forecasts based on the LC model significantly underestimated the increase in male life expectancy after 1990 in the G7 countries. This limitation is well known, and solutions have been suggested, see for instance Li et al. (2013) that extended the Lee-Carter to take into account rotation of the age pattern (i.e., the changes in the age pattern of mortality decline). The CoDe 2.1 model has four parameters (A_t, a_t, M_t, $\beta_{3,t}$) to explain the improvements in the death probabilities over time and across three age groups, i.e., infant (A_t), adolescent (a_t), and adult (M_t, $\beta_{3,t}$) mortality. In contrast, an important assumption of the CoDe 2.1 model is that relative changes in the death rates are non-constant over time and different across ages.

Because capturing mortality delay and compression can provide insight into adult mortality trends, it is desirable for a mortality model to have this ability (de Beer and Janssen 2016; Basellini et al. 2016; Bergeron-Boucher et al. 2015). As the LC model cannot decompose the compression and delay effect, it cannot identify the influence of compression and delay in life expectancy improvements. The CoDe 2.1 model, on the

other hand, has one parameter for each of the two effects, i.e., M_t for delay and $\beta_{3,t}$ for compression in adult and old ages.

The LC model assumes that for all ages, the death rates decrease exponentially over time. Due to the RW assumption (see Eq. 6) and the exponential decline, the future death rates will follow a log-normal distribution. As a result, when the variance of the RW is not very small, we consider the possibility of an increase of mortality rates, which results in non-symmetric projections intervals. The CoDe 2.1 model incorporates different functions for each age group (infant, adolescent, and adult ages) due to its non-linear form. For instance, for background mortality, we assume an exponential decline over time, whereas for adult ages we assume that death rates decline as a logistic function of the modal age. For parameter A_t (infant mortality) we use a logarithm transformation in our projections to ensure that $A_t > 0$. Nevertheless, the variance is very small, so we simulate no scenarios of increasing infant mortality.

When making long-run projections, some mortality forecasters incorporate exogenous variables such as gross domestic product (GDP), weather conditions, or expert opinion (Lutz and Scherbov 1998; Niu and Melenberg 2014; Villegas and Haberman 2014; Seklecka et al. 2017). Our projection model provides such flexibility. Consider a scenario in which experts believe there will be a swift and unprecedented decline in young adult mortality because the introduction of effective interventions is expected to result in a sharp decline in traffic accidents. Using the CoDe 2.1 model, we can take this information into account by projecting a more rapid decline in background mortality, i.e., parameter a_t. But if when using the LC model, we assume a faster decline of parameter κ_t, that assumption will affect the death rates across all ages. In the literature, methods for including exogenous variables in the LC model have been proposed (Niu and Melenberg 2014; Toczydlowska et al. 2017). These methods require additional parameters and use exogenous determinants as covariates in the model calibration. Consequently, the trend of the time-dependent parameter κ_t is altered depending on the exogenous variables. Another example is that the values of the death probabilities at very old ages can be determined on the basis of expert opinion by fixing parameter g (Gampe 2010).

To explain why projections failed (or did not fail) to capture mortality trends is important. The CoDe 2.1 model can be used as a diagnostic tool for addressing such questions. Using the CoDe 2.1 model, we can perform a backtest and compare the time-dependent parameters. For instance, if we underestimated the improvement in the death rates at adult ages, we can examine whether this was due to an underestimation of delay or of compression. By contrast, the LC model is inadequate as a diagnostic tool because the parameter κ_t provides information only for the average improvement over all ages.

Quantitative comparison

In-sample fit

The CoDe 2.1 model provides a good fit for the death probabilities and the age-at-death distribution for French, Japanese, and American women and men for the 1960–2014 period and ages 0–100 (Table 2). The LC model provides a better fit for the log-death probabilities. The root mean square error (RMSE) of $\log q_{x,t}$ gives a relatively

Table 2. Root mean square error (RMSE) of the CoDe 2.1 model and the LC model. Reference period 1960–2014 and ages 0–100. We use green color for the model with better fit and red color for the model with poorer fit

Model	Gender	Country	$q_{x,t}$	$\log q_{x,t}$	$d_{x,t}$
CoDe 2.1	Female	France	0.00487	0.0935	0.000458
		USA	0.00200	0.0690	0.000349
		Japan	0.00468	0.0721	0.000353
	Male	France	0.00894	0.120	0.000584
		USA	0.00329	0.0983	0.000491
		Japan	0.00665	0.0957	0.000491
Lee-Carter	Female	France	0.00407	0.0879	0.000602
		USA	0.00382	0.0589	0.000454
		Japan	0.00783	0.0878	0.00118
	Male	France	0.00710	0.0882	0.000646
		USA	0.00419	0.0621	0.000561
		Japan	0.00799	0.0734	0.000537

large amount of weight to errors at young ages, RMSE of $q_{x,t}$ to errors at old ages, and RMSE of $d_{x,t}$ to errors around the modal age. Hence, we conclude that the LC model provides a better fit for young ages, with the exception of Japanese females, whereas the CoDe 2.1 model provides a better fit for ages around the modal age and for old ages.

In Fig. 2, we show the in-sample estimations of the death probabilities (top) (logarithmic transformation) and the age-at-death distribution (bottom) for the CoDe 2.1 model (left) and the LC model (right) for Japanese females for the years 1960 and 2014. From the death probabilities, we observe that for 1960, the LC model does not fit very well around the modal age, whereas the CoDe 2.1 model provides a good fit for both years. For other populations, see Figures 5–9 in Appendix 2.

In Fig. 3, we show the estimates of life expectancy at birth and the modal age at death for French, Japanese, and American women. We observe that the CoDe 2.1 model provides a better estimation of life expectancy and smoothed estimates of the modal age, whereas the LC model does not provide smooth estimations of the modal age. We obtained similar results for men, see Figure 10 in Appendix 2.

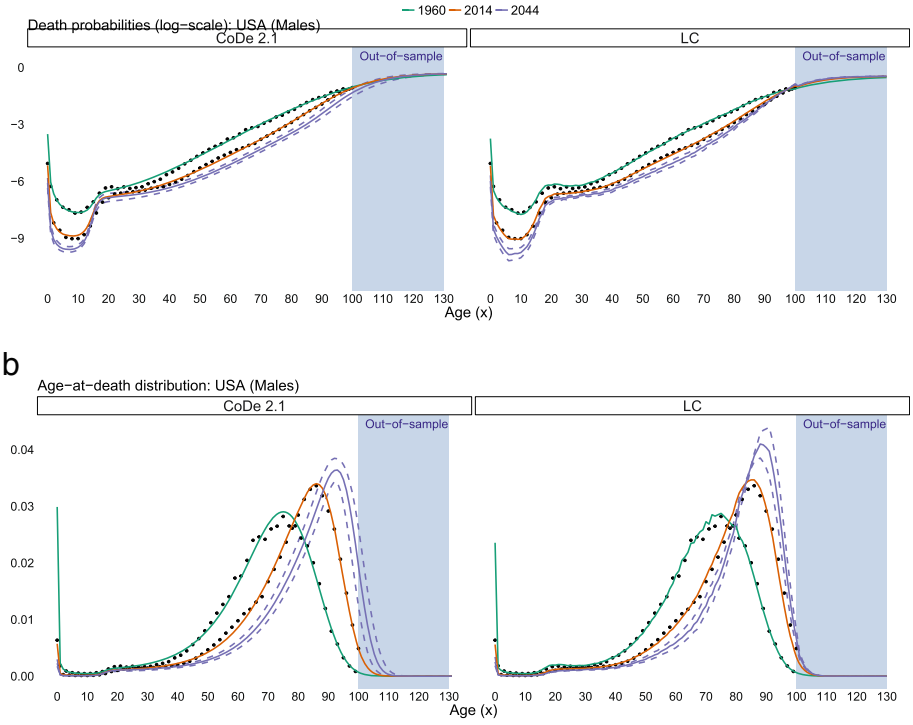

Fig. 2 Estimations and projections of the death probabilities and the age-at-death distribution for Japanese females for the years 1960, 2014, and 2044 and for ages 0–130. CoDe 2.1 (left) and LC (right). The dashed lines are 95% projection intervals. The reference period is 1960–2014 for ages 0–100

In-sample projections

Figure 4 shows the in-sample projections of life expectancy at birth and the modal age at death for the 2004–2014 period based on data for the 1960–2004 period for French, Japanese, and American women. The LC model fails to capture the observed linear increase in the modal age among French women, whereas it accurately projects the deceleration of the increase in the modal age among Japanese women. The ARIMA model used to project the delay parameter of the CoDe 2.1 model projects a linear continuation of the modal age. Thus, this model captures the linear increase in the modal age among French women, but fails to capture the slowdown in the increase in the modal age among Japanese women. We can therefore conclude that the CoDe 2.1 model performs better when estimating French life expectancy, whereas the LC model performs better when estimating Japanese life expectancy. We observe that the two models project similar (short-term) trajectories for life expectancy and the modal age among American women.

We present the in-sample projection for men in Figure 11 in Appendix 2. For French men, we observe that the LC model projects a slowdown and fails to capture the increase in life expectancy at birth for the 2004–2014 period, whereas the CoDe 2.1 model projects a continuation of the linear increase in life expectancy at birth. Moreover, both models project a linear continuation of the modal age for the shorter 2004–2014 period. For Japanese men, the LC model provides a more accurate estimation of the average lifespan trend than the CoDe 2.1 model for the years 2004–2010, but the models provide similar estimations for the 2010–2014 period. For American men, the

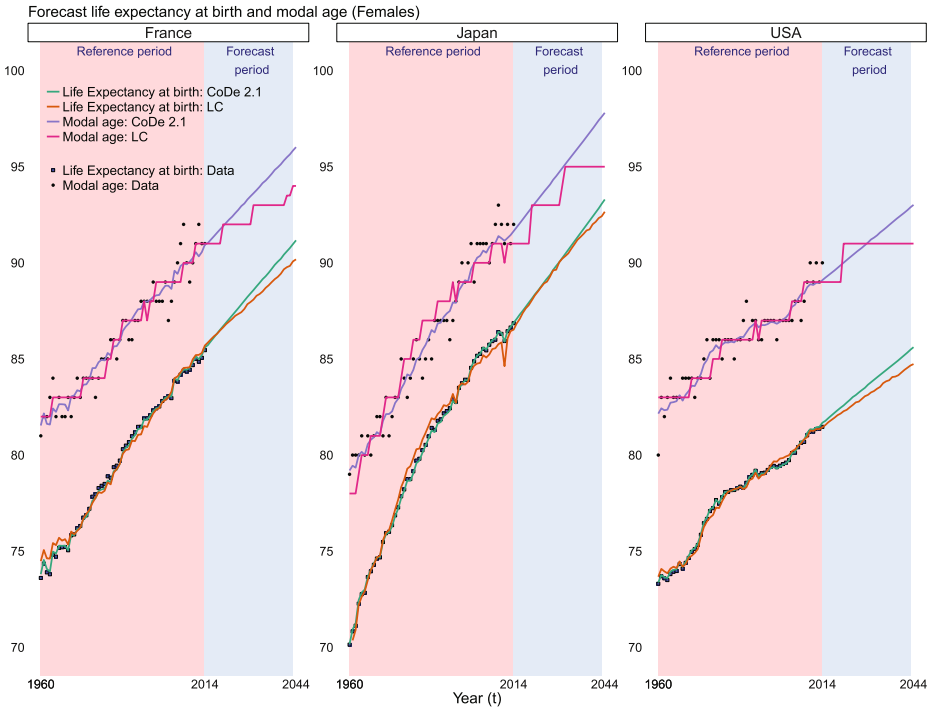

Fig. 3 Estimation and out-of-sample projections of life expectancy at birth and of the modal age of females using the CoDe 2.1 model and the LC model. The reference period is 1960–2014 (pink); the projection period is 2015–2044 (blue)

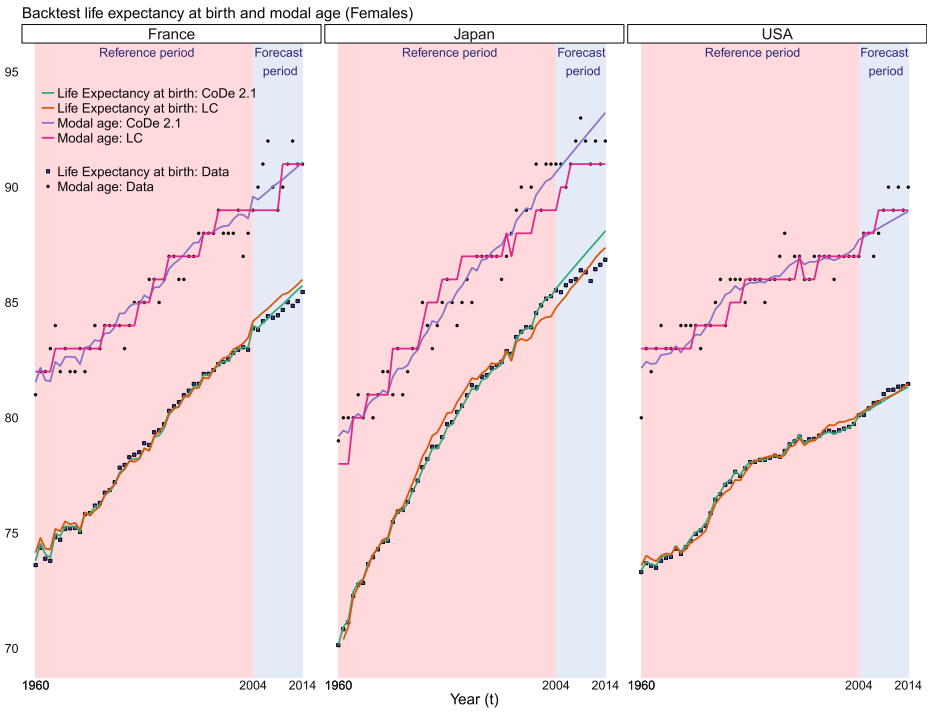

Fig. 4 Estimation and in-sample projections of life expectancy at birth and the modal age of females using the CoDe 2.1 model and the LC model. The reference period is 1960–2004 (pink); the projection period is 2005–2014 (blue)

CoDe 2.1 model captures the linear increase in the modal age and provides a better estimation of life expectancy at birth. The LC model fails to project the increases in the modal age and the average lifespan.

Out-of-sample projections and extrapolations

In Fig. 2, we also present the out-of-sample projections of the death probabilities and the age-at-death distribution for the year 2044 among Japanese women. For the results of analyses of other populations, see Figures 5-9 in Appendix 2. We conclude that the CoDe 2.1 model has smooth point estimates and projections of intervals of the death probabilities and the age-at-death distribution, whereas the LC model produces unsmoothed prediction intervals around the modal age. Looking at the LC model, we observe that the prediction intervals of the death probabilities are uniform for all ages because they stem from the same time-dependent parameter, κ_t. In contrast, the CoDe 2.1 model has smaller projection intervals that are bigger for younger ages because parameters A_t and a_t have higher levels of uncertainty than M_t and $\beta_{3,t}$.

In Fig. 3, we show the projections of life expectancy at birth and the modal age for French, Japanese, and American women from 2015 to 2044. We observe that the CoDe 2.1 model projects a continuation of the increase in the modal age and in life expectancy at birth, whereas the LC model consistently projects a slowdown of the increase in the modal age and in life expectancy. Using the LC model, life expectancy is projected to move in the direction of the modal age, whereas using the CoDe 2.1 model, life expectancy is projected to move almost parallel to the modal age. Furthermore, the LC model produces unsmoothed projection intervals of the modal age that are outside of the 95% (smoothed) projection intervals of the CoDe 2.1 model (see Figure 12 in Appendix 2). We obtained similar results for men, see Figures 10 and 12 in Appendix 2.

Conclusion and discussion

Summary of results

In this paper, we proposed a method, the CoDe 2.1 model, for projecting mortality by capturing the delay and the compression of mortality. We compared our method with the well-known Lee-Carter (LC) model.

Our qualitative comparison showed that the CoDe 2.1 model has more attractive and useful properties for making mortality estimates and projections than the LC model. The CoDe 2.1 model has a non-linear and fixed form for describing the age pattern of mortality, it uses less parameters to describe the shape in the age pattern of mortality, but more parameters to describe the changes in the age pattern, it provides extrapolation of death probabilities to higher ages, it allows to estimate the modal age at death, it does not assume an exponential decline in rates across all ages, it decomposes the delay and compression effect, it can be combined with exogenous variables and expert opinion, and it can be used as a diagnostic tool.

Our quantitative comparison based on data for France, Japan, and the USA for the 1960–2014 period showed that the CoDe 2.1 model provides a better fit around the modal age and at older ages, whereas the LC model provides a better fit at younger ages. Our in-sample projections revealed that for the years 2005 to 2014, the CoDe 2.1 model captures life expectancy and the modal age as well as or better than the LC

model, except among Japanese women and men. Our out-of-sample projections showed that the LC model consistently projects a slowdown of mortality delay and thus of life expectancy at birth, whereas the CoDe 2.1 model projected a continuation of the increase in the modal age at death and thus a steady increase in life expectancy at birth (see Ediev 2011). Obviously, there may be a bound to the increase of the modal age at death in the long run. If we would project a linear increase of the modal age up to 2100, this would result in projecting a modal age beyond age 100. The question whether this would be a plausible projection is beyond the scope of this paper. In Dong et al. 2016 published a paper in Nature suggesting that the limit to human lifespan is 115 years. If this would be true, one may question whether a modal age above 100 is plausible, as this would imply much more compression in old age than observed so far. However, the claim by Dong et al. (2016) was questioned by five other papers in Nature (http://www.nature.com/articles/nature19793). Thus, there is not yet a general agreement about the limit to human lifespan. Hence, it is not obvious what would be the ceiling of the increase in the modal age. Consequently, a longer projection horizon, e.g., until 2100, is realistic only under logical assumptions regarding the possible limit of human lifespan and the future age pattern. The CoDe 2.1 model is flexible enough (see the "Qualitative comparison" section and Table 1) to take into account such assumptions during the projection process.

Explanation of the results

The CoDe 2.1 model provides a better fit around the modal age and for old-age mortality because it uses a logistic function (the third term in Eq. 1) with an age-dependent slope (see Eq. 2) to capture adult mortality. The LC model provides a better fit at younger ages because it has more age-dependent parameters to explain the non-linear pattern observed at young ages.

Using backtesting, we observed that the CoDe 2.1 model captures life expectancy and the modal age for the 2005–2014 period as well as or better than the LC model, except among Japanese women and men. For the LC model, we cannot relate the errors observed in the average lifespan projections to errors in the mortality forecasts of specific age groups for each population (see Table 1, Ability to use as diagnostic tool). It is interesting to examine the reason why the LC model performs better than the CoDe 2.1 model in the backtesting of Japanese women. The LC model projects the slowdown in the increase in the modal age after 2009 (see Fig. 4), even though during the 1960–2004 observation period, the modal age increased linearly. Thus, the LC model did not project a slowdown because that was a logical assumption given the development during the observation period but because the LC model consistently projects a slowdown. The suitability of the CoDe 2.1 model for use as a diagnostic tool can help to explain the observed results. By evaluating the in-sample projections of the parameters (A_t, α_t, M_t, $\beta_{3,t}$), we can relate the errors we observed in the projections of life expectancy for each population to the errors in the mortality forecasts of specific age groups. Especially in the years 2010–2014, we underestimated background mortality for French women and we underestimated compression for French men (see Figures 13 and 14 in Appendix 2). For Japanese women, the CoDe 2.1 model overestimated delay because the modal age had been increasing linearly until 2004, but the increase slowed down

after 2004 (see Figure 13 in Appendix 2). For Japanese men, we underestimated compression at old ages, as compression was steady over the 1970–2004 period (see Figure 14 in Appendix 2). Finally, for American women and men, the errors stem mainly from the overestimation of compression for the 2010–2014 period. Among American women and men, the compression parameter increased steadily from 1990 to 2005, stagnated from 2005 to 2010, and declined from 2010 onward (see Figures 13 and 14 in Appendix 2). The ARIMA model therefore projected a continuation of compression as observed in the reference period, which resulted in the overestimation of compression for American women and men in the projection period.

From the out-of-sample projection (long-run) (see Fig. 4 and Figure 11 in Appendix 2), we observed that the LC model consistently projects the deceleration of the increase in the modal age and in life expectancy at birth. The inability of the LC model to project a linear increase in the modal age is caused by the erroneous assumption that the relative changes in the death rates are constant over time (see Table 1, modeling of the mortality trend). In the LC model, the relative changes in the death rates are captured by a single age-dependent parameter (i.e., β_x) that is a (weighted) average of the relative changes observed in the calibration period. Hence, the estimation of the relative changes in the death rates is highly dependent on the calibration. Furthermore, the LC model assumes a continuation of the estimated relative changes in the age pattern rather than a continuation of the most recent changes in the age pattern. As a result, when mortality continues to be delayed, as in the case of French women, the LC model tends to underestimate the increase in the modal age at death, i.e., the model projects a deceleration of the increase.

The CoDe 2.1 model projected the steady continuation of trends in mortality delay and life expectancy at birth. By design, the CoDe 2.1 model has a parameter for approximating modal age. Thus, the modal age was not estimated in a separate step in the CoDe 2.1 model as it was in the LC model. It is important to note that the time series framework we used to project the modal age (i.e., M_t) influences the future levels of the modal age. We used AICc to select ARIMA models to project the modal age. Results show a steady increase of modal age and life expectancy for all the populations.

Appraisal of the CoDe model and recommendations

A prominent functional attribute of the CoDe 2.1 model is that it analyzes mortality trends in three age groups (infancy, young adulthood, and adult ages), which are modeled by three sets of time-dependent parameters. The LC model, by contrast, considers only one period effect. For making projections, the CoDe 2.1 model has four advantages relative to the LC model. First, because it has more time-dependent parameters, the CoDe 2.1 model accounts for the diversity of the mortality decline with age more effectively than the LC model. Thus, the CoDe 2.1 model is more flexible in capturing changes in the age pattern. The second advantage is that the CoDe 2.1 model by construction does not project monotonic age-profiles of death rates, which is a disadvantage of the Lee-Carter model (Li et al. 2013). The third advantage is that one of the parameters of the CoDe 2.1 model, i.e., $\beta_{3,t}$, measures and quantifies mortality compression. The LC model cannot provide any information about the compression of deaths, even though this information is valuable in assessing mortality (Bohk-Ewald

et al. 2017). The fourth advantage is that because it studies the past trends in the time-dependent parameters, the CoDe 2.1 model can provide insights into forecasting failures. By contrast, the single period parameter of the LC model is an estimation of average improvement across all ages that provides no explicit information on a specific age group.

In addition, the CoDe 2.1 model is flexible. Indeed, the estimation of the linear continuation of delay in the case of Japanese women, and of compression in the case of American women, can always be improved by adapting the time series framework used to project the parameters of the CoDe 2.1 model. As the AICc is an in-sample measure, it selects the best ARIMA models based on the calibration period of the time series. Looking at Figure 13 in Appendix 2, we observe that many of the parameters have non-linear patterns. Choosing different calibration periods for the time series may result in non-linear patterns, and thus in different trajectories for the modal age and for compression. Furthermore, we observe that the modal age for Japanese females and males has a concave shape. Thus, we can use Box-Cox transformation to project the concave shape of the modal age.

In this paper, we propose a methodology to project mortality rates by projecting mortality delay and compression and as far as we know CoDe 2.1 is the first model able to project these two effects. In presenting our methodology, we add a comparison with the benchmark Lee-Carter method. Although various extensions and variations of the LC model deal with certain—and different—shortcomings of the LC model (e.g., Booth et al. 2006; Li et al. 2013), they do not disentangle delay and compression effects, nor separately project them. Note that we do not claim that our model produces more accurate projections than all other existing mortality projection models. Rather we aim to show that accounting for delay and compression is necessary for more robust and accurate projections of the age pattern of mortality.

We consider parameter M_t of the CoDe 2.1 model a good estimator of the modal age. Although due to mathematical complexities we cannot provide a mathematical justification why this is the case, we consider our empirical illustration as sufficient evidence. The empirical analysis furthermore showed that the trend over time in the modal age at death is very well captured by a change in parameter M_t (see Figure 15 in Appendix 2). An accurate estimation of this trend over time was a very important component of the current research. For other analyses with the modal age at death by means of the parameter M_t, further improvement of the model is required, such that parameter M_t will always coincide with the modal age.

Currently, we applied the CoDe 2.1 mortality projection to three low-mortality countries: France, Japan, and the USA. Indeed, our model is particularly useful to project mortality for countries which have experienced a shift in the age-at-death distribution: most non-Eastern European countries plus other low-mortality countries like Japan and the USA (Janssen and de Beer: The timing of the transition from mortality compression to mortality delay in Europe, Japan, and the United States, submitted). Eastern European countries experienced a stagnation in the increase in life expectancy from 1975 onwards due to the public health crisis in these countries (McKee and Shkolnikov 2001; Vallin and Meslé 2004; Leon 2011). As a result, no shift in the age-at-death distribution occurred in these countries up to 2000 (Janssen and de Beer: The timing of the transition from mortality compression to mortality delay in Europe, Japan, and the United States, submitted). Similar to other mortality models and projection methods, the CoDe 2.1 model is therefore not suitable to adequately forecast mortality in Eastern European countries.

In this paper, we projected mortality for each country and for men and women separately. The CoDe 2.1 model can, however, be adapted to produce coherent forecasts simultaneously for both women and men for several countries by coherently projecting the time-dependent parameters. We can further improve projections by, for example, investigating alternative projection scenarios for the modal age and compression, taking into account the effects of the smoking epidemic (Janssen et al. 2013), or by including socioeconomic differences (Villegas and Haberman 2014; Cairns et al. 2016).

Overall conclusion and implications

Projecting mortality by taking mortality delay and compression into account can result in more accurate forecasts than using the LC model, especially when the modal age increases linearly. Projecting mortality by including mortality delay and compression is therefore essential to making sound mortality projections.

Because the LC model is currently the standard tool used in making mortality projections, the tendency of the model to underestimate mortality delay—and thus life expectancy at birth—is worrisome. The underestimation of mortality delay and life expectancy has important implications for society and policymakers, as it can lead to inaccurate assessments of the solvency of government and private pension funds, and of longevity risk hedgers. The results of our CoDe 2.1 model, which projects mortality by taking mortality delay and compression into account, indicate that future life expectancy values may be higher than has previously been predicted. Thus, we estimate that the capital requirements of pension providers will be higher than the LC model has estimated.

The CoDe 2.1 model we propose for projecting mortality by estimating mortality delay and compression can be considered an important new tool for projecting mortality. The method is flexible and can be further extended to take into account non-linear past trends, the mortality experiences of other populations, and lifestyle factors.

Appendix 1

The algorithm to calibrate the CoDe 2.1 model

Step 1: Define necessary bounds for the parameters. For our analysis, we used the same bounds for all the countries, genders, and years, see Table 3.

Step 2: For each year, we calibrate the CoDe 2.1 assuming that all the parameters are time dependent and h_1 is equal to zero, by minimizing for each year t

$$\frac{\mathrm{MSE}(\log q_{x,t})}{\mathrm{Var}(\log q_{x,t})}+\frac{\mathrm{MSE}(q_{x,t})}{\mathrm{Var}(q_{x,t})}+\frac{\mathrm{MSE}(d_{x,t})}{\mathrm{Var}(d_{x,t})}.$$

For this optimization, we use the **DEoptim** package in **R** that performs global optimization by differential evolution and does not require initial values only the bounds from Step 1.

Step 3: In this step, we ensure that parameter M_t is indeed a good estimator of the modal age. To do so, we compare the parameter M_t with the smoothed modal age. We estimate the smooth modal age based on the methodology of Ouellette and Bourbeau (2011). If we observe a level difference between parameter M_t and the smoothed modal age M_t^s, we estimate

Table 3 Bounds for the parameters required for the calibration of the model

	A_t	B	a_t	z	M_t	$\beta_{3,t}$	g	h_2	b_1	b_2
Lower	e^{-10}	e^{-10}	$\log(e^{-10})$	12	70	0	0.4	0	0	0
Upper	0.05	5	log(0.01)	18	100	0.5	1	100	0.5	1

$$h_1 := \sum_t \left(M_t - M_t^s\right).$$

Step 4: This is the last step of the calibration, where we minimize Eq. 5. Due to the large number of parameters, we use package **rgenoud** in **R** that performs genetic optimization using derivatives. For the time-dependent parameters, we use as starting values the estimations from Step 1 and for the constant across time parameters the average across years from Step 1. The bounds are the same as in Steps 1 and 2. We set h_1 equal to the value we estimate in Step 2.

An **R** script for the calibration of the CoDe 2.1 model is available to the GENUS readers upon request. Please address your request to the corresponding author.

The algorithm to calibrate the Lee-Carter model

For the calibration of the Lee-Carter model, we use the algorithm which is described by Pitacco et al. (2009) in their Section 4.2. We do not adjust the parameters to the deaths count or to the life expectancy as was originally suggested by Lee and Carter, for optimal comparison with the CoDe 2.0 model.

Appendix 2

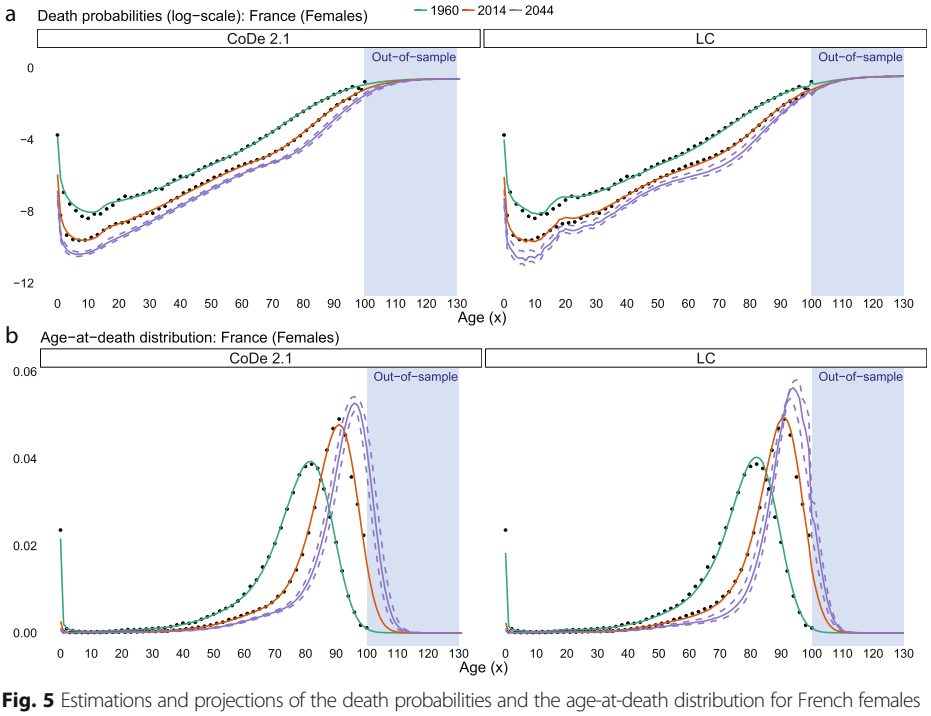

Fig. 5 Estimations and projections of the death probabilities and the age-at-death distribution for French females for the years 1960, 2014, and 2044 and for ages 0–130. CoDe 2.1 (left) and LC (right). The dotted lines are observations and the dashed lines are 95% projection intervals. The reference period is 1960–2014 for ages 0–100

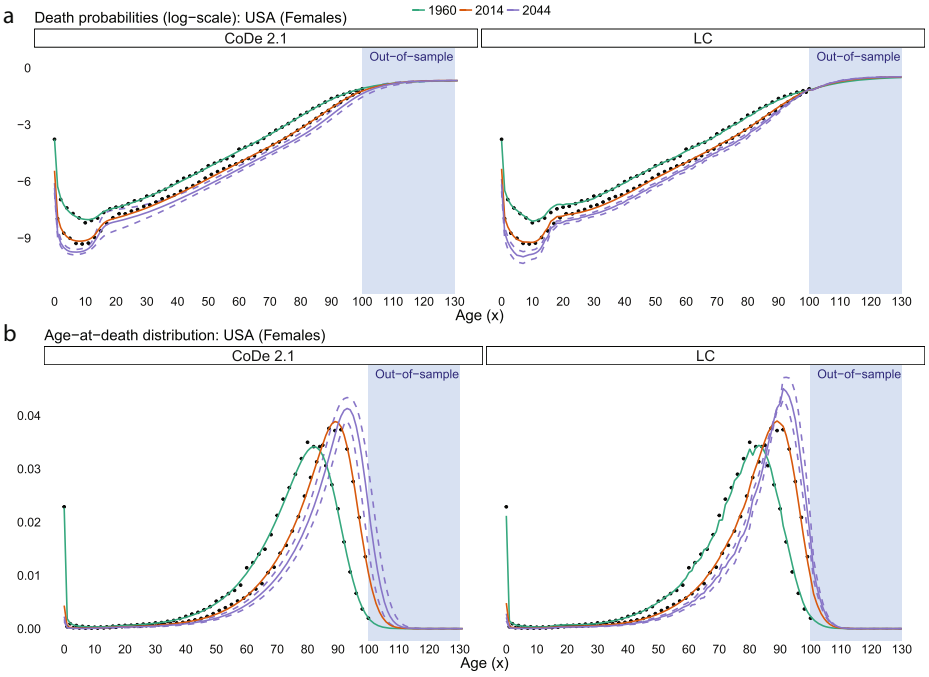

Fig. 6 Estimations and projections of the death probabilities and the age-at-death distribution for American females for the years 1960, 2014, and 2044 and for ages 0–130. CoDe 2.1 (left) and LC (right). The dotted lines are observations and the dashed lines are 95% projection intervals. The reference period is 1960–2014 for ages 0–100

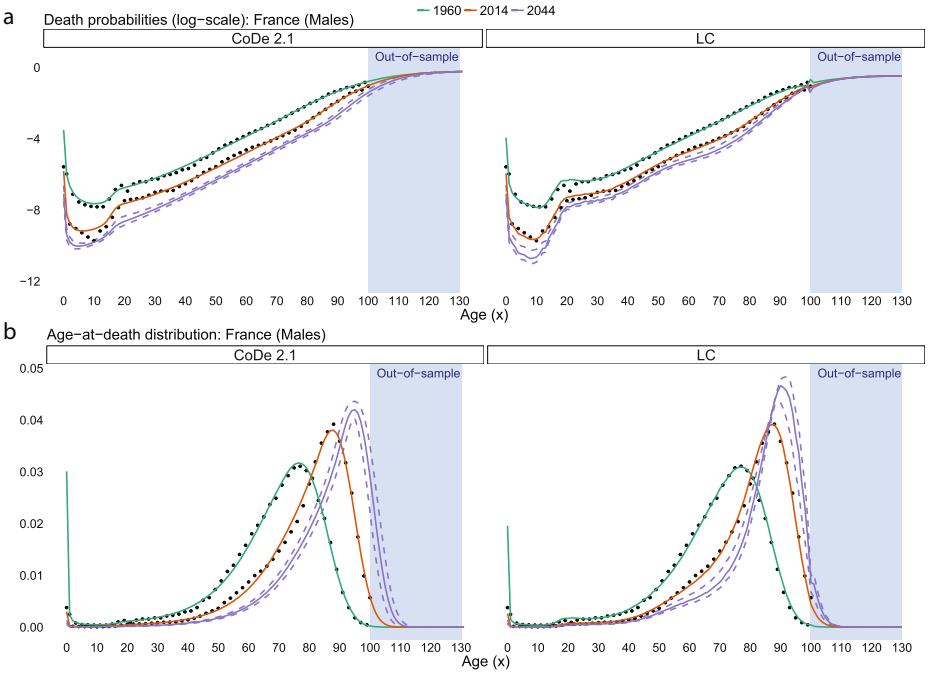

Fig. 7 Estimations and projections of the death probabilities and the age-at-death distribution for France males for the years 1960, 2014, and 2044 and for ages 0–130. CoDe 2.1 (left) and LC (right). The dotted lines are observations and the dashed lines are 95% projection intervals. The reference period is 1960–2014 for ages 0–100

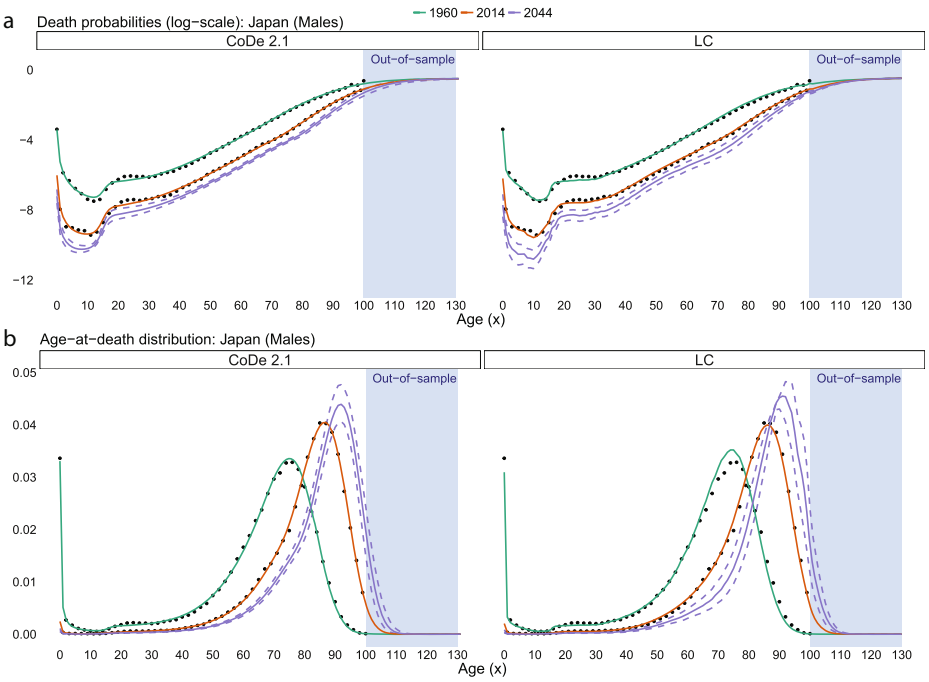

Fig. 8 Estimations and projections of the death probabilities and the age-at-death distribution for Japanese males for the years 1960, 2014, and 2044 and for ages 0–130. CoDe 2.1 (left) and LC (right). The dotted lines are observations and the dashed lines are 95% projection intervals. The reference period is 1960–2014 for ages 0–100

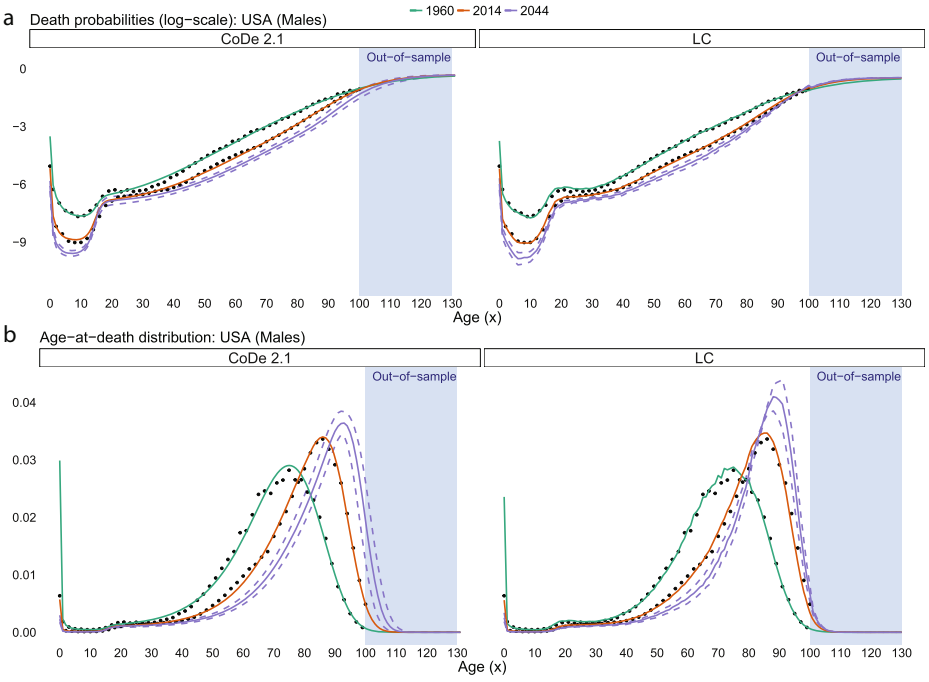

Fig. 9 Estimations and projections of the death probabilities and the age-at-death distribution for American males for the years 1960, 2014, and 2044 and for ages 0–130. CoDe 2.1 (left) and LC (right). The dotted lines are observations and the dashed lines are 95% projection intervals. The reference period is 1960–2014 for ages 0–100

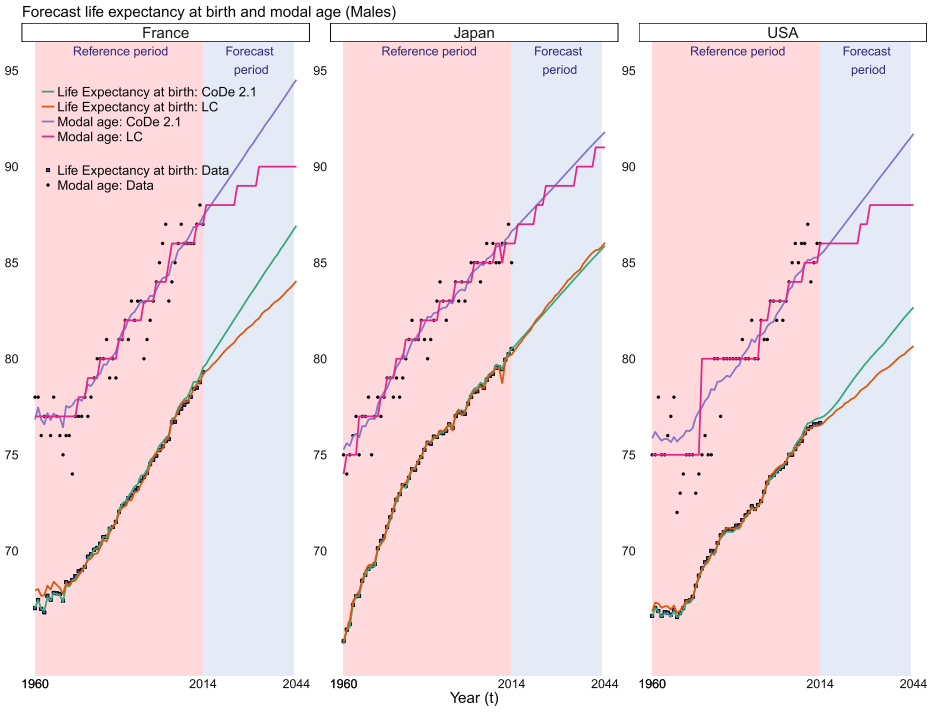

Fig. 10 Estimation and out-of-sample projections of life expectancy at birth and of the modal age of males using the CoDe 2.1 model and the LC model. The reference period is 1960–2014 (pink); the projection period is 2015–2044 (blue)

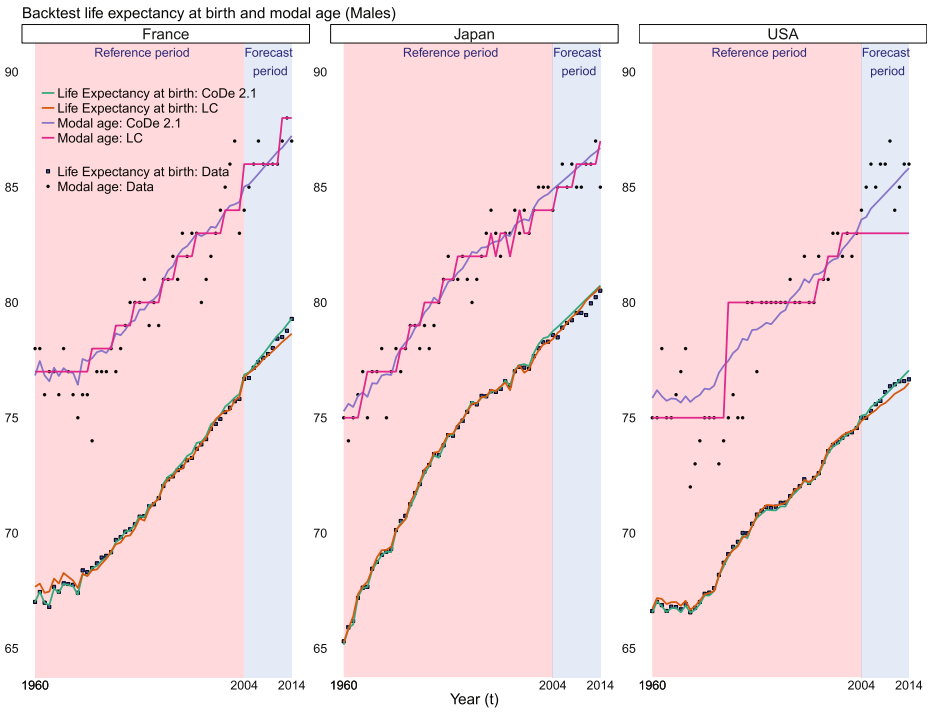

Fig. 11 Estimation and in-sample projections of life expectancy at birth and the modal age of males using the CoDe 2.1 model and the LC model. The reference period is 1960–2004 (pink); the projection period is 2005–2014 (blue)

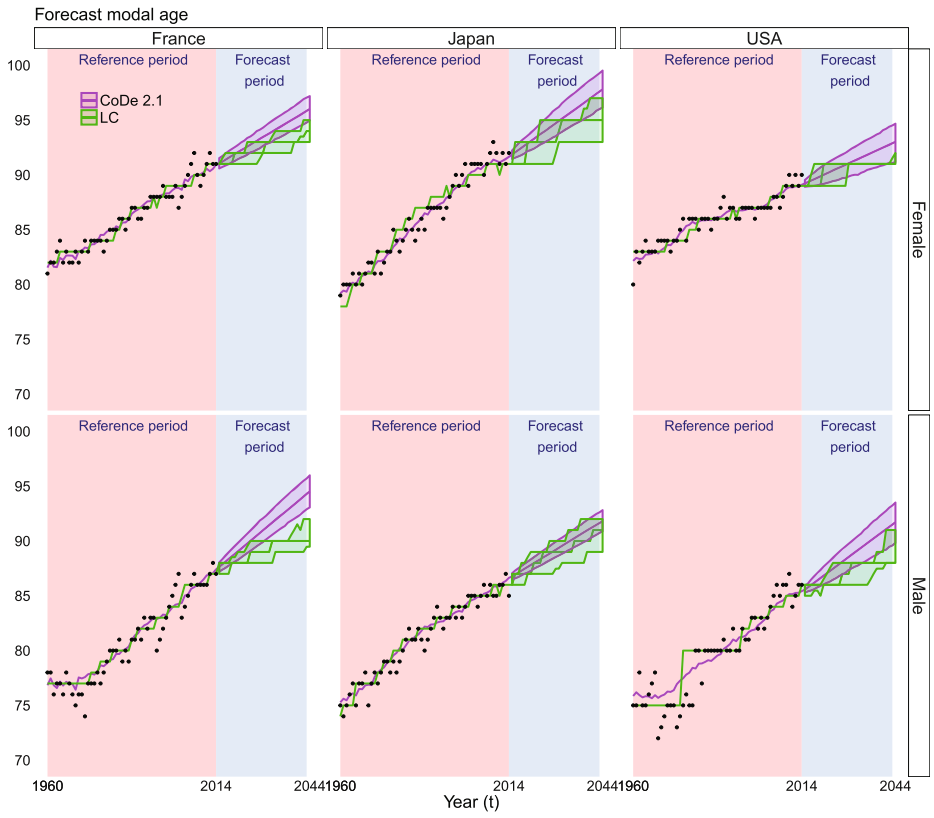

Fig. 12 Estimation and out-of-sample projections of the modal age using the CoDe 2.1 model and the LC model. The reference period is 1960–2014 (pink); the projection period is 2015–2044 (blue)

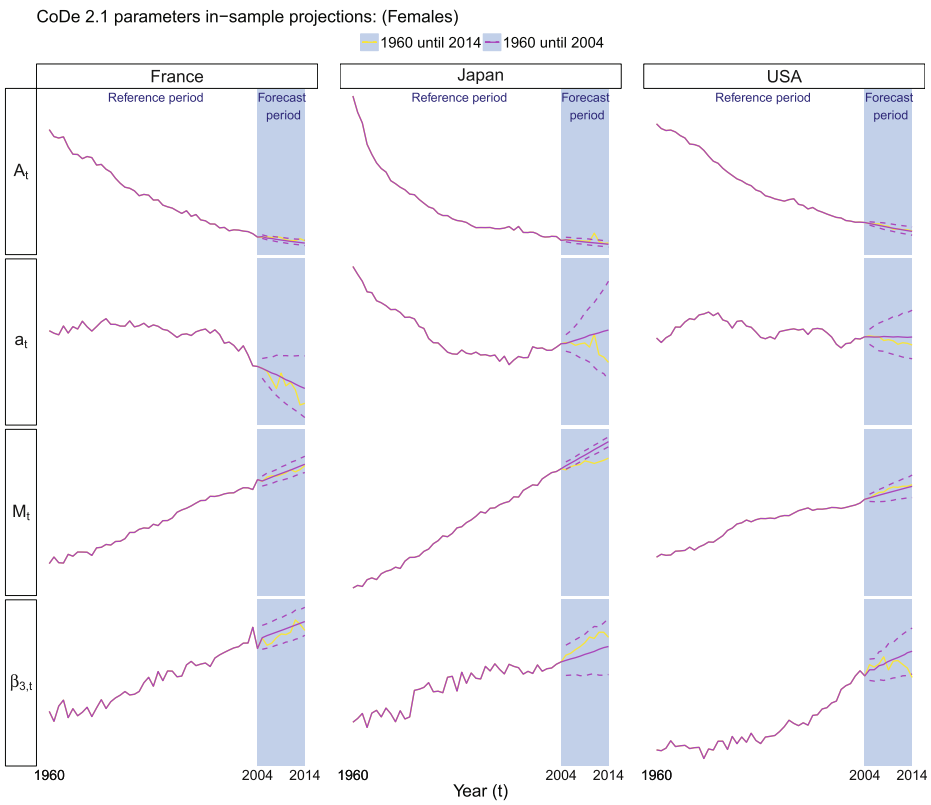

Fig. 13 Estimation and in-sample projections of CoDe 2.1 model parameters of females. The reference period is 1960–2004 (white); the projection period is 2005–2014 (blue)

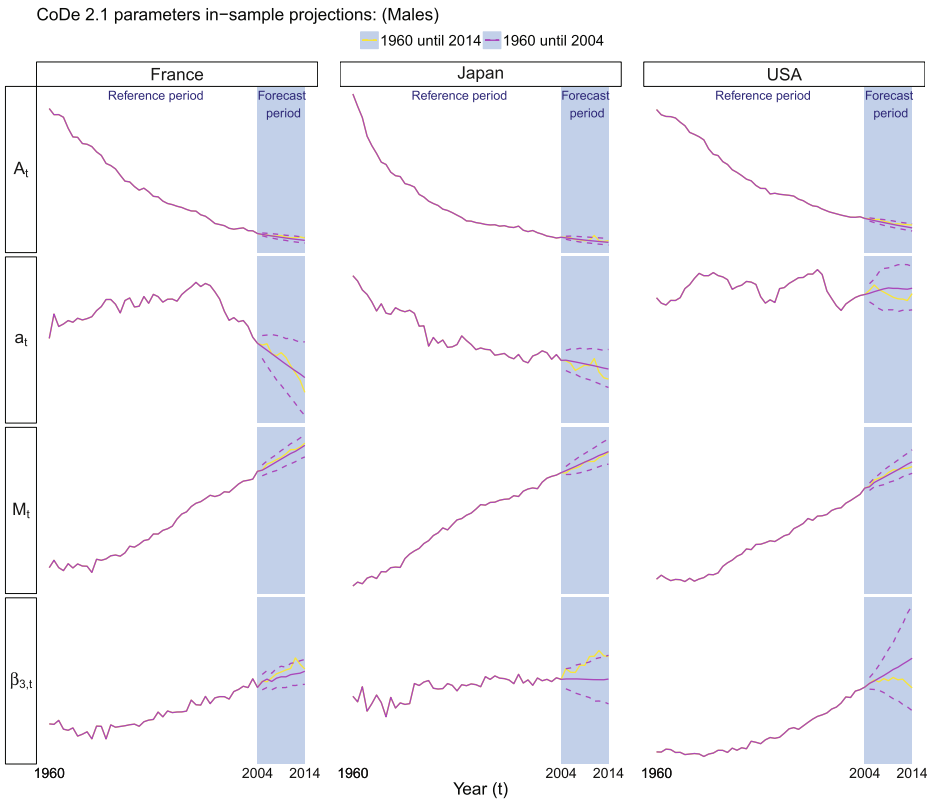

Fig. 14 Estimation and in-sample projections of CoDe 2.1 model parameters of males. The reference period is 1960–2004 (white); the projection period is 2005–2014 (blue)

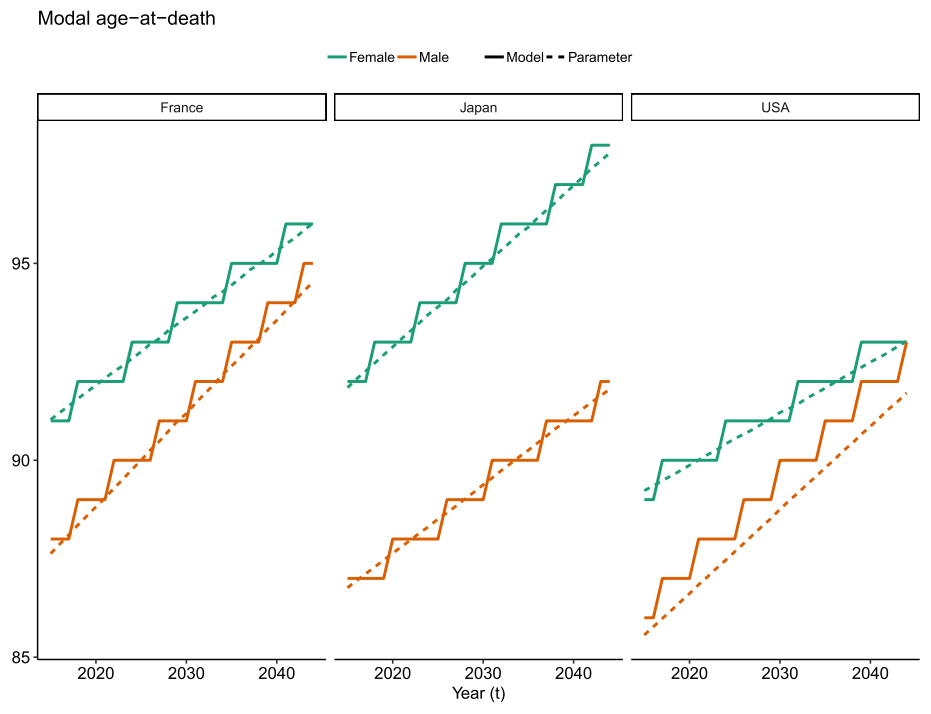

Fig. 15 Comparison of the parameter M_t with the modal age based on the estimated age-at-death distribution using CoDe 2.1 model for the years 2015–2044

Abbreviations
AICc: Corrected Akaike's Information Criterion; ARIMA: Autoregressive integrated moving average; CoDe model: Compression and delay mortality model; HMD: Human Mortality Database; RMSE: Root mean squared error; RW: Random walk

Acknowledgements
We highly appreciate the comments received from the anonymous referees which helped improving the quality of the paper.

Funding
This work is financed by the Netherlands Organisation for Scientific Research (NWO) in relation to the research program "Smoking, alcohol, and obesity, ingredients for improved and robust mortality projections", under grant no. 452-13-001. See www.futuremortality.com.

Availability of data and materials
The data were obtained from the Human Mortality Database http://www.mortality.org.

Authors' contributions
All authors participated in the design of the study and aided in interpreting the results. AB and JdB carried out the empirical analysis. AB and FJ drafted the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

Author details
[1]Population Research Centre, University of Groningen, Groningen, The Netherlands. [2]Netherlands Interdisciplinary Demographic Institute, The Hague, The Netherlands.

Received: 15 February 2018 Accepted: 7 September 2018
Published online: 25 October 2018

References
Antonio, K., Devriendt, S., de Boer, W., de Vries, R., De Waegenaere, A., Kan, H.-K., Kromme, E., Ouburg, W., Schulteis, T., Slagter, E., van der Winden, M., van Iersel, C., & Vellekoop, M. (2017). Producing the Dutch and Belgian mortality projections: a stochastic multi-population standard. *European Actuarial Journal, 7*(2), 297–336.

Basellini, U., Camarda, C.G., and Canudas-Romo, V. (2016) Modeling and forecasting age at death distributions: a nonparametric approach. https://iussp.confex.com/iussp/ipc2017/mediafile/ExtendedAbstract/Paper6042/ModelAndForecastDx_IPC2017.pdf

Bergeron-Boucher, M.-P., Ebeling, M., & Canudas-Romo, V. (2015). Decomposing changes in life expectancy: compression versus shifting mortality. *Demographic Research, 33*(14), 391–424.

Bohk-Ewald, C., Ebeling, M., & Rau, R. (2017). Lifespan disparity as an additional indicator for evaluating mortality forecasts. *Demography, 54*(4), 1559–1577.

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

Booth, H., Hyndman, R., Tickle, L., & de Jong, P. (2006). Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions. *Demographic Research, 15*(9), 289–310.

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

Booth, H., & Tickle, L. (2008). Mortality modelling and forecasting: a review of methods. *Annals of Actuarial Science, 3*(1–2), 3–43.

Box, G. E., & Jenkins, G. M. (1970). *Time-series analysis, forecasting and control*. San Francisco: Holden-Day (revised edn., 1976).

Cairns, A., Blake, D., Dowd, K., Coughlan, G., Epstein, D., & Khalaf-Allah, M. (2011). Mortality density forecasts: an analysis of six stochastic mortality models. *Insurance: Mathematics and Economics, 48*(3), 355–367.

Cairns, A., Blake, D., Dowd, K., Coughlan, G., Epstein, D., Ong, A., & Balevich, I. (2009). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal, 13*(1), 1–35.

Cairns, A. J. G., Kallestrup-Lamb, M., Rosenskjold, C. P. T., Blake, D., & Dowd, K. (2016). *Modelling socio-economic differences in the mortality of Danish males using a new affluence index*. Working paper, Heriot-Watt University http://www.macs.hw.ac.uk/~andrewc/papers/ajgc73.pdf.

Canudas-Romo, V. (2008). The modal age at death and the shifting mortality hypothesis. *Demographic Research, 19*(30), 1179–1204.

Canudas-Romo, V. (2010). Three measures of longevity: time trends and record values. *Demography, 47*(2), 299–312.

Cheung, S. L. K., Robine, J.-M., Tu, E. J.-C., & Caselli, G. (2005). Three dimensions of the survival curve: horizontalization, verticalization, and longevity extension. *Demography, 42*(2), 243–258.

de Beer, J., Bardoutsos, A., & Janssen, F. (2017). Maximum human lifespan may increase to 125 years. *Nature, 546*, E16–E17.

de Beer, J., & Janssen, F. (2016). A new parametric model to assess delay and compression of mortality. *Population Health Metrics, 14*(1), 46.

Dong, X., Milholland, B., & Vijg, J. (2016). Evidence for a limit to human lifespan. *Nature, 538*(7624), 257.

Dowd, K., Cairns, A., Blake, D., Coughlan, G., Epstein, D., & Khalaf-Allah, M. (2010a). Backtesting stochastic mortality models: an ex-post evaluation of multi-period-ahead density forecasts. *North American Actuarial Journal, 14*, 281–298.
Dowd, K., Cairns, A., Blake, D., Coughlan, G., Epstein, D., & Khalaf-Allah, M. (2010b). Evaluating the goodness of fit of stochastic mortality models. *Insurance: Mathematics and Economics, 47*, 255–265.
Ediev, D. M. (2011). Life expectancy in developed countries is higher than conventionally estimated. Implications from improved measurement of human longevity. *Population Ageing, 4*, 5 https://doi.org/10.1007/s12062-011-9040-x.
Ediev, D. M. (2013). Decompression of period old-age mortality: when adjusted for bias, the variance in the ages at death shows compression. *Mathematical Population Studies, 20*(3), 137–154.
Fries, J. F. (1980). Aging, natural death, and the compression of morbidity. *New England Journal of Medicine, 303*(3), 130–135.
Gampe, J. (2010). Human mortality beyond age 110. In: Maier H., Gampe J., Jeune B., Robine J.-M., & Vaupel J. W., editors. Supercentenarians: Demographic research monographs (pp. 219–230). Heidelberg: Springer-Verlag
Gavrilov, L. A., & Gavrilova, N. S. (2011). Mortality measurement at advanced ages: a study of the Social Security Administration Death Master File. *North American Actuarial Journal, 15*(3), 432–447.
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., 115*, 513–585.
Haberman, S., & Renshaw, A. (2012). Parametric mortality improvement rate modelling and projecting. *Insurance: Mathematics and Economics, 50*(3), 309–333.
Human Mortality Database. University of California, Berkeley (USA) and Max Planck Institute for Demographic Research, Rostock (Germany). Available at http://www.mortality.org. Data accessed December 2016.
Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika, 76*(2), 297–307.
Hyndman, R., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software, 27*(3), 1-22. https://doi.org/10.18637/jss.v027.i03.
International Monetary Fund. (2016). Chapter 3: The insurance sector—trends and systemic risk implications. In *Global financial stability report: potent policies for a successful normalization*. USA: IMF eLibrary. https://www.elibrary.imf.org/view/IMF082/23157-9781513506777/23157-9781513506777/ch03.xml.
Janssen, F., Van Wissen, L., & Kunst, A. (2013). Including the Smoking Epidemic in Internationally Coherent Mortality Projections. *Demography*, 50(4), 1341–1362.
Kannisto, V. (1994). *Development of the oldest-old mortality, 1950–1980: evidence from 28 developed countries (Odense monographs on population aging no. 1)*. Odense: Odense University Press.
Kannisto, V. (1996). *The advancing frontier of survival (Odense monographs on population aging no. 3)*. Odense: Odense University Press.
Kannisto, V. (2000). Measuring the compression of mortality. *Demographic Research, 3*(6), 1-24.
Kleinow, T. (2015). A common age effect model for the mortality of multiple populations. *Insurance: Mathematics and Economics, 63*, 147–152.
Lee, R., & Carter, L. (1992). Modeling and forecasting the time series of US mortality. *Journal of the American Statistical Association, 87*, 659–671.
Lee, R., & Miller, T. (2001). Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography, 38*(4), 537–549.
Leon, D. A. (2011). Trends in European life expectancy: a salutary view. *International Journal of Epidemiology, 40*(2), 271–277.
Li, H., & O'Hare, C. (2017). Semi-parametric extensions of the Cairns-Blake-Dowd model: a one-dimensional kernel smoothing approach. *Insurance: Mathematics and Economics, 77*, 166–176.
Li, N., Lee, R., & Gerland, P. (2013). Extending the Lee-Carter method to model the rotation of age patterns of mortality-decline for long-term projection. Demography, *50*(6), 2037–2051. https://doi.org/10.1007/s13524-013-0232-2.
Lutz, W., & Scherbov, S. (1998). An expert-based framework for probabilistic national population projections: the example of Austria. *European Journal of Population, 14*(1), 1–17.
Makeham, W. M. (1860). On the law of mortality and the construction of annuity tables. *Journal of the Institute of Actuaries and Assurance Magazine, 8*, 301–310.
McKee, M., & Shkolnikov, V. (2001). Understanding the toll of premature death among men in eastern Europe. *BMJ, 323*(7320), 1051–1055.
Niu, G., & Melenberg, B. (2014). Trends in mortality decrease and economic growth. *Demography, 51*(5), 1755–1773.
Ouellette, N., & Bourbeau, R. (2011). Changes in the age-at-death distribution in four low mortality countries: a nonparametric approach. *Demographic Research, 25*(19), 595–628.
Pitacco, E., Denuit, M., Haberman, S., & Olivieri, A. (2009). *Modeling longevity dynamics for pensions andannuity business*. London: Oxford University Press.
Plat, R. (2009). On stochastic mortality models. *Insurance: Mathematics and Economics, 45*, 393–404.
R Core Team. (2017). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing URL: https://www.R-project.org/.
Rau, R., Ebeling, M., Peters, F., Bohk-Ewald, C., Missov, T. I. (2017) Where is the level of the mortality plateau? Society of Actuaries (Ed.): 2017 Living to 100 monograph (2017) Living to 100 Schaumburg, IL: *Society of Actuaries*.
Renshaw, A., & Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics, 38*, 556–570.
Seligman, B., Greenberg, G., & Tuljapurkar, S. (2016). Convergence in male and female life expectancy: Direction, age pattern, and causes. *Demographic Research*, 34, 1063–1074.
Seklecka, M., Pantelous, A. A., & O'Hare, C. (2017). Mortality effects of temperature changes in the United Kingdom. *Journal of Forecasting, 36*(7), 824–841 for.2473.
Ševčíková, H., Li, N., Kantorová, V., Gerland, P., & Raftery, A. E. (2016). *Age-specific mortality and fertility rates for probabilistic population projections* (pp. 285–310). Cham: Springer International Publishing.
Stoeldraijer, L., van Duin, C., van Wissen, L., & Janssen, F. (2013). Impact of different mortality forecasting methods and explicit assumptions on projected future life expectancy: the case of the Netherlands. *Demographic Research, 29*(13), 323–354.
Terblanche, W. (2016). Retrospective testing of mortality forecasting methods for the projection of very elderly populations in Australia. *Journal of Forecasting, 35*(8), 703–717 FOR-15-0064.R2.

Thatcher, A. R. (1999). The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 162*(1), 5–43.
Thatcher, A. R., Kannisto, V., & Vaupel, J. W. (1998). *The force of mortality at ages 80 to 120*. Odense: Odense University Press.
Toczydlowska, D., Peters, G. W., Fung, M. C., & Shevchenko, P. V. (2017). Stochastic period and cohort effect state-space mortality models incorporating demographic factors via probabilistic robust principal components. *Risks, 5*, 42.
Vallin, J., & Meslé, F. (2004). Convergences and divergences in mortality. A new approach to health transition. *Demographic Research, Special Collection, 2*(2), 11–44.
Vaupel, J. W. (2010). Maximum human lifespan will increase. *Nature, 464*, 536–542.
Villegas, A. M., & Haberman, S. (2014). On the modeling and forecasting of socioeconomic mortality differentials: an application to deprivation and mortality in England. *North American Actuarial Journal, 18*(1), 168–193.
Wilmoth, J.R., Andreev, K., Jdanov, D., Glei, D.A., Riffe, T. (2017) Methods protocol for the human mortality database, http://www.mortality.org/Public/Docs/MethodsProtocol.pdf, accessed December 2017.
Wilmoth, J. R., & Horiuchi, S. (1999). Rectangularization revisited: variability of age at death within human populations. *Demography, 36*(4), 475–495.