International Actuarial Association
Association Actuarielle Internationale

RESEARCH ARTICLE

A calendar year mortality model in continuous time

Donatien Hainaut

UCLouvain, LIDAM-ISBA, Voie du Roman Pays 20, 1348 Louvain-la-Neuve, Belgium
E-mail: donatien.hainaut@uclouvain.be

Abstract

This article proposes a continuous time mortality model based on calendar years. Mortality rates belong to a mean-reverting random field indexed by time and age. In order to explain the improvement of life expectancies, the reversion level of mortality rates is the product of a deterministic function of age and of a decreasing jump-diffusion process driving the evolution of longevity. We provide a general closed-form expression for survival probabilities and develop it when the mean reversion level of mortality rates is proportional to a Gompertz–Makeham law. We develop an econometric estimation method and validate the model on the Belgian population.

1. Introduction

Actuarial valuation involves to consider the uncertainty arising from the evolution of longevity. The statistical and insurance literature has produced a variety of statistical mortality forecasting models, such as the Lee and Carter (LC, 1992) model and its variants (e.g., Wilmoth, 1993; Booth *et al.,* 2002, and Cairns *et al.,* 2006a; Renshaw and Haberman, 2003). To summarize, the log-force of mortality is a linear combination of one or several age-specific functions and longevity random processes. Among the existing extensions of the LC model, Renshaw and Haberman (2006) propose to include a cohort effect for explaining the UK mortality. Whereas Li *et al.* (2009) introduce in the LC model an age-specific random variable that accounts for heterogeneity of individuals. We refer the interested reader to the work of Cairns *et al.* (2011) for a comparison of six mortality models extending the LC framework.

Statistical approaches and in particular the LC model became a standard in the insurance industry due to their robustness and reliability. Based on a discrete time framework, statistical models generate yearly simulated sample paths of mortality rates. Their main drawback is that survival probabilities do not have a closed-form expression. Premium or solvency capital requirement (SCR) calculations rely then on computationally intensive Monte-Carlo simulations.

Affine diffusion processes offer an interesting alternative to statistical approaches for modeling the mortality of a cohort. Directly inspired from the literature about the term structure of interest rates, an affine model is a trade-off between complexity and computational tractability of pricing and estimation. In this respect, Milevsky and Promislow (2001) were among the firsts to consider mean-reverting stochastic affine processes for modeling mortality. Luciano and Vigna (2005) develop a jump-diffusion affine model for modeling rates and show that constant mean-reverting process are not adapted for describing the mortality. Biffis (2005) exploits the analytical tractability of affine processes to evaluate life insurance contracts in a jump-diffusion setup. Cairns *et al.* (2006b) examine how to model mortality risks and price mortality-related instruments using adaptations of the arbitrage-free pricing frameworks for interest rate derivatives. Schrager (2006) develop a multivariate affine model and fit it to Dutch mortality rates. Luciano and Vigna (2008) calibrate various time homogeneous affine models to different generations in the UK population and investigate their empirical appropriateness. Hainaut

and Devolder ([2008](#)) propose a cohort model for mortality rates based on a mean-reverting Lévy process. Jevtic *et al.* ([2013](#)) study and calibrate a cohort-based model which captures the characteristics of a mortality surface with a continuous-time factor approach. Chen *et al.* ([2018](#)) adopt a bivariate affine Gaussian factors model for calculating SCRs of mixed gender portfolios of annuities. Zeddouk and Devolder ([2020](#)) model the force of mortality with mean-reverting affine processes in which the level of reversion is time-dependent. Xu *et al.* ([2020](#)) develop a multi-cohort mortality model for age-cohort mortality rates with common factors across cohorts as well as cohort-specific factors. Zhu and Bauer ([2022](#)) propose a Gaussian framework for describing the stochastic evolution of mortality projections rather than realized mortality rates.

This article proposes a calendar year model in the sense that we associate with each age x, a mortality process indexed by the calendar time. This contributes to the literature in several directions. Previously cited papers rely on a single mortality process per cohort of individuals. Such an approach presents a high level of analytical tractability. Nevertheless, their calibration requires either to observe the mortality of cohorts till their extinction or to fit them to prospective tables built with a different statistical method. Another difficulty arises for the joint-modeling of multiple cohorts. The calendar year model remedies to these issues. In the same manner as a statistical model, parameters are indeed estimated with data collected on a smaller time window than a cohort model. By construction, the calendar year model also allows for correlation between cohorts and survival probabilities admit a closed-form expression. Finally, the calendar year model remains analytically tractable. Our model may be seen as a continuous extension of the model of Wu and Wang ([2018](#)) who propose a Gaussian process regression for mortality modeling.

The structure of the paper is as follows. Section [2](#) presents the structure of the calendar year model. Mortality rates are stochastic processes reverting to a mean level that is the product of an age-specific function and of process driving the evolution of longevity. The third section focuses on the calculation of survival probabilities. We discuss the conditions of existence of an equivalent pricing measure and develop survival probabilities when mortality rates revert to a Gompertz–Makeham function scaled by the longevity process. In Section [4](#), we exploit the Gaussian feature of our model to estimate its parameters. The model is estimated in Section [5](#) for the Belgian male and female population over the period 1950–2010. Its predictive capacity is benchmarked to LC and Renshaw–Haberman forecasts from 2011 to 2020. In Section [6](#), we provide parameter estimates for Belgium, UK, Italy, male and female populations fitted to data from 1950 up to 2020. Cross-sectional and prospective life expectancies are compared. In the last section, we test the capacity of our model to evaluate a term life annuity and the related longevity risk.

2. A calendar year model

Let us denote by $\tau_{t,x}$, the random remaining lifetime of a x years old individual at time t. We assume that the death occurs when then integral of mortality rates reaches the realization of an exponential random variable ϵ, of parameter 1. Mortality processes are denoted by $\left(\mu_{t,x}\right)_{t\geq0,x\geq0}$ where $t \geq 0$ is the calendar time and $x \in [0, \omega]$ is the age. They are defined on a probability space Ω, endowed with their filtration $\left(\mathcal{G}_t\right)_{t\geq0}$ and a probability measure P. Contrary to existing affine frameworks, we consider a continuum of mortality processes for each age x instead of a single mortality process per cohort. The set of mortality rates is then a random field indexed by time and age, as illustrated in Figure [1](#).

In this model, the probability that the individual survives up to time s, conditionally to the knowledge of $\left(\mu_{t+u,x+u}\right)_{u\in[0,s-t]}$, is then equal to

$$P\left(\tau_{t,x} \geq s - t \mid \mathcal{G}_s\right) = P\left(\int_0^{s-t} \mu_{t+u,x+u}du \geq \epsilon\right) \tag{2.1}$$

$$= \exp\left(-\int_0^{s-t} \mu_{t+u,x+u}du\right).$$
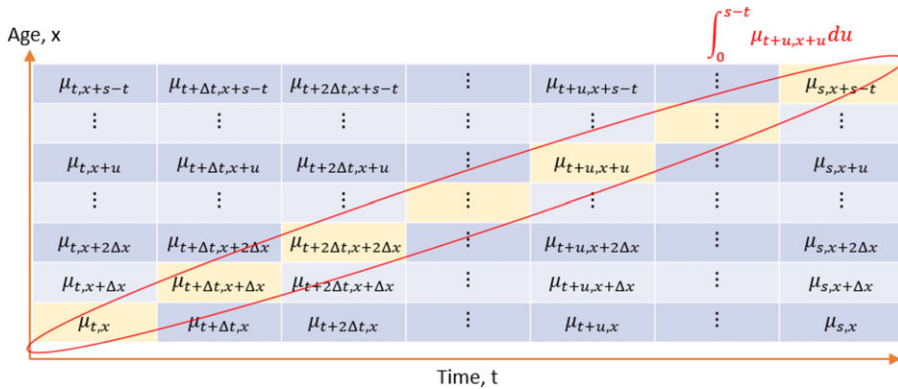
Figure 1. *Mortality rates belong to a random field indexed by time and age. The survival probability $P(\tau_{t,x} \geq s \mid \mathcal{G}_s)$ depends then on a continuum of random variables $\left(\mu_{t+u,x+u}\right)_{u \in [0, s-t]}$.*
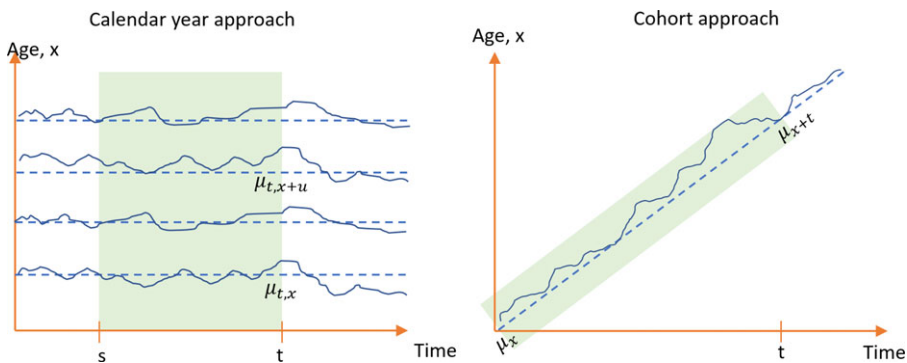
Figure 2. *Comparison of calendar year and cohort approaches. Shaded areas represent the time window of data needed for statistical inference.*

The integral in this last expression is an integral of an infinity of processes indexed by $x + u$ for $u \in [0, s - t]$ as illustrated in Figure 1.

Our model is a calendar year approach in the sense that we associate with each age x, a mortality process indexed by the calendar time. In an affine framework such as the one of Luciano and Vigna (2005), mortality rates are defined by cohort and the survival probability depends on a single mortality process indexed by the cohort age. Figure 2 compares the calendar year and the cohort approaches. Cohort models usually present a high level of analytical tractability and offer closed-form expression for survival probabilities. Nevertheless, their calibration requires to observe the mortality of cohorts till their extinction. For this exercise, we have then to consider mortality rates dating from the first half of the 19th century. This introduces a bias in the estimation due to the considerable progress made in the healthcare sector during the last 50 years. An alternative to estimate affine cohort models consists in fitting them to prospective tables computed with a statistical method. Nevertheless, this approach lacks of consistency and accumulates biases from the cohort model and from the statistical method used for the construction of prospective tables. Another issue is the joint-modeling of multiple cohorts. A life insurer is exposed to mortality/longevity risks of customers belonging to different cohorts. Introducing correlations between cohorts is a challenging task which reduces the analytical tractability of cohort models. From a statistical inference point of view, a calendar year approach is more reliable as it is based on multiple time series of mortality rates. These time series are stationary, and the calendar year

model can be estimated with data on a smaller time window than a cohort model. In the remainder of this article, we show that it is still possible to infer analytical expression for survival probabilities subject to some assumptions on the evolution of mortality. By construction, the correlation between cohorts is included in the calendar year approach.

In our framework, the death occurs when then integral of mortality processes reaches the realization of an exponential random variable of parameter 1. An alternative approach consists in assuming that the death occurs at the first jump of a point process. But the alternative definition (2.1) allows us to consider mortality processes $(\mu_{t,x})_{t \geq 0}$ defined on $\mathbb{R}$ instead of $\mathbb{R}^+$. This feature makes then possible to adopt a Gaussian model for $(\mu_{t,x})_{t \geq 0}$ leading to closed-form expression of survival probabilities. Notice that Gaussian frameworks are also used for approximating the dynamic of mortality rates in various affine cohort models, as in Luciano and Vigna (2005) or Zeddouk and Devolder (2020).

In order to preserve the analytical tractability of further developments, we assume that mortality processes are Gaussian with a mean-reverting dynamic defined by:

$$d\mu_{t,x} = \kappa \left(\theta_t \, \mu(x) - \mu_{t,x} \right) dt + \Sigma(x)^\top dW_t^\mu \,, \tag{2.2}$$

where $\kappa \in \mathbb{R}^+$ and $\left(W_t^\mu \right)_{t \geq 0} \in \mathbb{R}^d$ is a Brownian motion of dimension d, common to all $\left(\mu_{t,x} \right)_{x \in [0, \omega]}$. $(\theta_t)_{t \geq 0}$ is a stochastic process driving the evolution of longevity. $\Sigma(x) = (\Sigma_1(x), ..., \Sigma_d(x))$ is a volatility vector where, $\Sigma_k(x) : [0, \omega] \to \mathbb{R}^+$ is a continuous function of age for $k = 1, ..., d$. The mean reversion level of $\mu_{t,x}$ is the product of the longevity process $(\theta_t)_{t \geq 0}$ and of a continuous function $\mu(x) : [0, \omega] \to \mathbb{R}^+$, defining the baseline mortality. This function can be the Gompertz–Makeham law or any other parametric function. The longevity process is a mean-reverting jump-diffusion defined by:

$$d\theta_t = \alpha (\beta(t) - \theta_t) \, dt + \nu dW_t^\theta + dL_t \,, \tag{2.3}$$

where $\alpha \in \mathbb{R}^+$ and $\left(W_t^\theta \right)_{t \geq 0}$ is a Brownian motion independent of $\left(W_t^\mu \right)_{t \geq 0}$. The function $\beta(t) : \mathbb{R}^+ \to \mathbb{R}^+$ is strictly positive and decreasing. In later developments, we assume that $\beta(t)$ is equal to

$$\beta(t) = \theta_0 e^{-\gamma t} + \bar{\theta} \left(1 - e^{-\gamma t} \right) \,,$$

where θ_0, $\bar{\theta}$, and $\gamma \in \mathbb{R}^+$. For this choice of $\beta(t)$, θ_t converges on average to $\bar{\theta}$ that is lower than θ_0. In order to replicate jumps in mortality caused by pandemics such as the one of COVID-19, we assume that the longevity risk process is exposed to positive jumps. The jump size, denoted by J, is distributed according to an exponential distribution of parameter ρ. The moment generating function of jumps is denoted by:

$$\psi(\omega) = \mathbb{E}\left(e^{\omega J} \right) = \frac{\rho}{\rho - \omega}, \omega < \rho.$$

Shocks hitting the longevity process, occur according to a Poisson process, denoted by $(N_t)_{t \geq 0}$, with an intensity λ and such that

$$L_t = \sum_{k=1}^{N_t} J_k \,.$$

The filtration of $(\theta_t)_{t \geq 0}$ is denoted by $\mathcal{H}_t$. The union of mortality and longevity filtrations is $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{H}_t$. Notice that we could think to introduce correlation between longevity and mortality processes. Nevertheless, given the additive structure of $\mu_{t,x}$ and as $(\theta_t)_{t \geq 0}$ is not directly observed, this correlation cannot be estimated. Contrary to an affine framework in which there is one single mortality rate process per cohort, the calculation of survival probabilities involves an infinity of mortality processes. Nevertheless, the next section shows that it is still possible to infer closed-form expression for these probabilities.
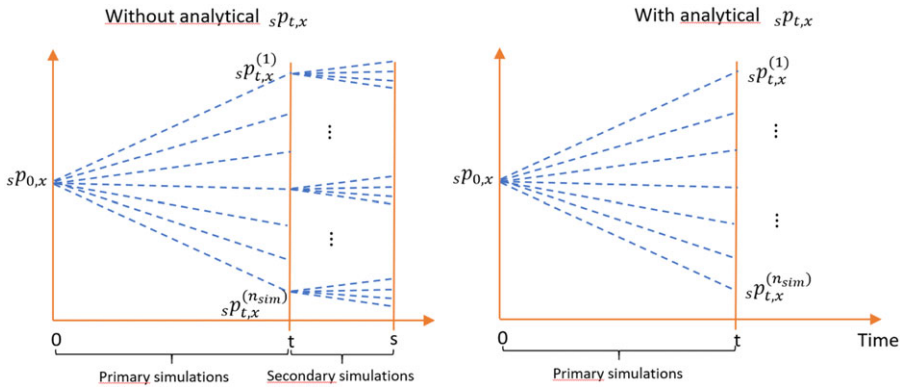
Figure 3. *Analytical framework versus simulations in simulations.*

3. Survival probabilities

Over the last 20 years, statistical frameworks based on the LC model (1992) became a standard in the insurance industry. This success is explained by several features such as their statistical robustness and their ability to explain the evolution of longevity. Their main disadvantage is that survival probabilities do not admit analytical solution and need to be computed by Monte-Carlo simulations.

In the context of Solvency II, estimating the SCR and the risk margin (RM) for covering adverse evolution of longevity is therefore computationally intensive. For instance, to forecast future SCR's covering the longevity risk of a portfolio of annuities, we have to consider simulations in simulations (eventually approached by the Longstaff–Schwartz method) to calculate survival probabilities. Figure 3 illustrates this point. Let us imagine that we generate n_{sim} primary scenarios of mortality and evaluate survival probabilities, noted $\left({}_sp^{(k)}{}_{t,x}\right)_{k=1,\dots,n_{sim}}$ in each simulation. In the absence of analytical formula, we have to simulate for each primary scenario several secondary mortality paths up to time s to estimate ${}_sp^{(k)}{}_{t,x}$. This is done by computing by the following average:

$$
{}_sp^{(k)}_{t,x} = \mathbb{E}\left(e^{-\int_0^{s-t}\mu_{t+u,x+u}du} \mid \mathcal{F}_t\right)
$$

$$
\approx \frac{1}{n_{sec.\,sim.}} \sum_{j=1}^{n_{sec.\,sim.}}\left(\exp\left(-\sum_{u=0}^{s-t-1}\mu^{(k,j)}_{t+u,x+u}\right)\right),
$$

where $n_{sec.\,sim.}$ is the number of secondary simulations and $\mu^{(k,j)}_{t+u,x+u}$ is the simulated mortality force in the jth secondary scenario of the kth primary scenario (i.e., $\mu^{(k,j)}_{t,x} = \mu^{(k)}_{t,x}$). We will see that the calendar year model studied in this article admits analytical expressions for survival probabilities and then remedies to this drawback. By construction, conditionally to the sample paths of $\left(\mu_{t+u,x+u}\right)_{u\in[s,t]}$, the survival probability is provided by Equation (2.1). The unconditional survival probability of a x-year-old individual at time t up to time s noted ${}_sp_{t,x}$ is therefore equal to an expectation:

$$
{}_sp_{t,x} = P\left(\tau_{t,x} \geq s\right)
$$

$$
= \mathbb{E}\left(\exp\left(-\int_0^{s-t}\mu_{t+u,x+u}du\right) \mid \mathcal{F}_t\right).
$$

In order to calculate this expectation, we first develop the analytical expression of $\mu_{t+u,x+u}$ conditioned by $\mathcal{F}_t$, as stated in the next proposition.

Proposition 3.1. *For $u \in [0, \omega - x]$, the mortality process $\mu_{t+u,x+u}$ may be developed as follows:*

$$\mu_{t+u,x+u} = e^{-\kappa u}\mu_{t,x+u} + \frac{\kappa\theta_t\mu(x+u)}{\kappa - \alpha}\left(e^{-\alpha u} - e^{-\kappa u}\right) + \Sigma(x+u)^\top \int_0^u e^{-\kappa(u-z)}d\tilde{\mathbf{W}}_z^\mu \tag{3.1}$$

$$+ \frac{\kappa\alpha\mu(x+u)}{\kappa - \alpha}\int_0^u \beta(t+z)\left(e^{-\alpha(u-z)} - e^{-\kappa(u-z)}\right)dz + \frac{\kappa\nu\mu(x+u)}{\kappa - \alpha}$$

$$\times \int_0^u \left(e^{-\alpha(u-z)} - e^{-\kappa(u-z)}\right)d\tilde{W}_z^\theta + \frac{\kappa\,\mu(x+u)}{\kappa - \alpha}\int_0^u \left(e^{-\alpha(u-z)} - e^{-\kappa(u-z)}\right)d\tilde{L}_z.$$

where $\tilde{\mathbf{W}}_z^\mu = \mathbf{W}_{t+z}^\mu$, $\tilde{W}_z^\theta = W_{t+z}^\theta$ are Brownian motions and $\tilde{L}_z = L_{t+z}$ is a compound Poisson process.

A key quantity for calculating the survival probability is the integral of $\mu_{t+u,x+u}$. The next proposition states that is may be rewritten as the sum of a deterministic drift, several Brownian integrals, and a jump component. This result will next be used to infer survival probabilities.

Proposition 3.2. *For any $\xi \geq 0$, the integral $\int_0^\xi \mu_{t+u,x+u}du$ may be developed as follows:*

$$\int_0^\xi \mu_{t+u,x+u}du = \int_0^\xi e^{-\kappa u}\mu_{t,x+u}du + \frac{\kappa\theta_t}{\kappa - \alpha}\int_0^\xi \mu(x+u)\left(e^{-\alpha u} - e^{-\kappa u}\right)du + \frac{\kappa\alpha\left(\theta_0 - \bar{\theta}\right)e^{-\gamma t}}{\kappa - \alpha}$$

$$\times \int_0^\xi \mu(x+u)\left(\frac{e^{-\gamma u} - e^{-\alpha u}}{\alpha - \gamma} - \frac{e^{-\gamma u} - e^{-\kappa u}}{\kappa - \gamma}\right)du + \int_0^\xi \int_z^\xi \Sigma(x+u)^\top e^{-\kappa(u-z)}du\,d\tilde{\mathbf{W}}_z^\mu$$

$$+ \frac{\kappa\alpha\bar{\theta}}{\kappa - \alpha}\int_0^\xi \mu(x+u)\left(\frac{1}{\alpha}\left(1 - e^{-\alpha u}\right) - \frac{1}{\kappa}\left(1 - e^{-\kappa u}\right)\right)du + \frac{\kappa\nu}{\kappa - \alpha}\int_0^\xi \int_z^\xi \mu(x+u)$$

$$\times \left(e^{-\alpha(u-z)} - e^{-\kappa(u-z)}\right)du\,d\tilde{W}_z^\theta + \frac{\kappa}{\kappa - \alpha}\int_0^\xi \int_z^\xi \mu(x+u)\left(e^{-\alpha(u-z)} - e^{-\kappa(u-z)}\right)du\,d\tilde{L}_z. \tag{3.2}$$

We dispose of all the elements to propose a closed-form expression of the survival probability.

Proposition 3.3. *The survival probability of a x-year-old individual at time t up to time s, noted $_sp_{t,x}$, is equal to*

$$_sp_{t,x} = \exp\left(-\,_{s-t}m_{t,x} + \frac{1}{2}\,_{s-t}v_{t,x} + A(0, s-t)\right) \tag{3.3}$$

where $_\xi m_{t,x}$ is drift of the integral $\int_0^\xi \mu_{t+u,x+u}du$, that is

$$_\xi m_{t,x} = \int_0^\xi e^{-\kappa u}\mu_{t,x+u}du + \frac{\kappa\theta_t}{\kappa - \alpha}\int_0^\xi \mu(x+u)\left(e^{-\alpha u} - e^{-\kappa u}\right)du \tag{3.4}$$

$$+ \frac{\kappa\alpha\left(\theta_0 - \bar{\theta}\right)e^{-\gamma t}}{\kappa - \alpha}\int_0^\xi \mu(x+u)\left(\frac{e^{-\gamma u} - e^{-\alpha u}}{\alpha - \gamma} - \frac{e^{-\gamma u} - e^{-\kappa u}}{\kappa - \gamma}\right)du$$

$$+ \frac{\kappa\alpha\bar{\theta}}{\kappa - \alpha}\int_0^\xi \mu(x+u)\left(\frac{1}{\alpha}\left(1 - e^{-\alpha u}\right) - \frac{1}{\kappa}\left(1 - e^{-\kappa u}\right)\right)du.$$

and $_\xi v_{t,x}$ is the variance of the Brownian part of $\int_0^\xi \mu_{t+u,x+u}du$:

$$_\xi v_{t,x} = \left(\frac{\kappa\nu}{\kappa - \alpha}\right)^2 \int_0^\xi \left(\int_z^\xi \mu(x+u)\left(e^{-\alpha(u-z)} - e^{-\kappa(u-z)}\right)du\right)^2 dz \tag{3.5}$$

$$+ \int_0^\xi \left(\int_z^\xi \Sigma(x+u)e^{-\kappa(u-z)}du\right)^\top \left(\int_z^\xi \Sigma(x+u)e^{-\kappa(u-z)}du\right)dz.$$

Whereas the function $A(\xi, s-t)$ solves the following ordinary differential equation (ODE):

$$\partial_\xi A(\xi, s-t) = -\lambda \left(\psi \left(-\frac{\kappa}{\kappa-\alpha} \int_\xi^{s-t} \mu(x+v) \left(e^{-\alpha(v-\xi)} - e^{-\kappa(v-\xi)} \right) dv \right) - 1 \right), \tag{3.6}$$

with the terminal condition $A(s-t, s-t) = 0$.

For general expressions of $\mu(x)$ and $\Sigma(x)$, the integrals in Equations (3.4), (3.5), and (3.6) are numerically computable. The ODE (3.6) is also solved numerically by finite difference approximation. For some model specifications, they can be fully developed. For instance, this is the case when $\mu(x)$ follows a Gompertz–Makeham distribution:

$$\mu(x) = a + bc^x, \tag{3.7}$$

where $a \in \mathbb{R}^+$ is the constant death rate caused by accidents and bc^x with $b, c \in \mathbb{R}^+$ is proportional to age. We will see in numerical illustrations that this assumption provides a good fit to mortality tables with a parsimonious model.

Corollary 3.4. *If the function $\mu(x)$ defining the level of mean reversion of mortality rates is the Gompertz–Makeham law (3.7), the drift $_\xi m_{t,x}$ of $\int_0^\xi \mu_{t+u,x+u} du$, is*

$$_\xi m_{t,x} = \int_0^\xi e^{-\kappa u} \mu_{t,x+u} du + \frac{\kappa \theta_t}{\kappa-\alpha} \left(\frac{a}{\alpha} \left(1 - e^{-\alpha\xi} \right) - \frac{a}{\kappa} \left(1 - e^{-\kappa\xi} \right) \right) \tag{3.8}$$

$$+ \frac{\kappa \theta_t}{\kappa-\alpha} \left(\frac{bc^x \left(e^{(\ln c - \alpha)\xi} - 1 \right)}{\ln c - \alpha} - \frac{bc^x \left(e^{(\ln c - \kappa)\xi} - 1 \right)}{\ln c - \kappa} \right)$$

$$+ \frac{\kappa \alpha \left(\theta_0 - \bar\theta \right) e^{-\gamma t} \left(\frac{1}{\alpha-\gamma} - \frac{1}{\kappa-\gamma} \right)}{\kappa-\alpha} \left(\frac{a}{\gamma} \left(1 - e^{-\epsilon\gamma} \right) - \frac{a}{\alpha} \left(1 - e^{-\alpha\xi} \right) \right)$$

$$+ \frac{\kappa \alpha \left(\theta_0 - \bar\theta \right) e^{-\gamma t} \left(\frac{1}{\alpha-\gamma} - \frac{1}{\kappa-\gamma} \right)}{\kappa-\alpha} \left(\frac{bc^x \left(e^{(\ln c - \gamma)\xi} - 1 \right)}{\ln c - \gamma} - \frac{bc^x \left(e^{(\ln c - \alpha)\xi} - 1 \right)}{\ln c - \alpha} \right)$$

$$+ \frac{\kappa \alpha \bar\theta}{\kappa-\alpha} \left(\frac{1}{\alpha} - \frac{1}{\kappa} \right) \left(a\xi + \frac{bc^x \left(e^{\xi \ln c} - 1 \right)}{\ln c} \right)$$

$$+ \frac{\kappa \alpha \bar\theta}{\kappa-\alpha} \left(\frac{a}{\kappa^2} \left(1 - e^{-\kappa\xi} \right) + \frac{bc^x \left(e^{(\ln c - \kappa)\xi} - 1 \right)}{\kappa \ln c - \kappa^2} \right)$$

$$- \frac{\kappa \alpha \bar\theta}{\kappa-\alpha} \left(\frac{a}{\alpha^2} \left(1 - e^{-\alpha\xi} \right) + \frac{bc^x \left(e^{(\ln c - \alpha)\xi} - 1 \right)}{\alpha \ln c - \alpha^2} \right).$$

Let us introduce the function $h(\xi|\beta_1, \beta_2)$ for $\xi \in \mathbb{R}^+$ and $\beta_1, \beta_2 \in \mathbb{R}$:

$$h(\xi|\beta_1, \beta_2) = \xi + \frac{1}{\beta_1+\beta_2} \left(1 - e^{-(\beta_1+\beta_2)\xi} \right)$$

$$- \frac{1}{\beta_2} \left(1 - e^{-\beta_2\xi} \right) - \frac{1}{\beta_1} \left(1 - e^{-\beta_1\xi} \right).$$

The variance of $\int_0^\xi \mu_{t+u,x+u}du$ simplifies as follows:

$$_\xi v_{t,x} = \left(\frac{\kappa v}{\kappa - \alpha}\right)^2 \left[\frac{a^2}{\alpha^2}h(\xi|\alpha,\alpha) + \frac{a^2}{\kappa^2}h(\xi|\kappa,\kappa) - 2\frac{a^2}{\alpha\kappa}h(\xi|\alpha,\kappa)\right. \tag{3.9}$$

$$+ \left(\frac{bc^x}{\ln c - \alpha}\right)^2 h(\xi|\alpha - \ln c, \alpha - \ln c) + \left(\frac{bc^x}{\ln c - \kappa}\right)^2 h(\xi|\kappa - \ln c, \kappa - \ln c)$$

$$-2\frac{abc^x}{\alpha(\ln c - \alpha)}h(\xi|\alpha, \alpha - \ln c) + 2\frac{abc^x}{\alpha(\ln c - \kappa)}h(\xi|\alpha, \kappa - \ln c)$$

$$+2\frac{abc^x}{\kappa(\ln c - \alpha)}h(\xi|\kappa, \alpha - \ln c) - 2\frac{abc^x}{\kappa(\ln c - \kappa)}h(\xi|\kappa, \kappa - \ln c)$$

$$\left. -2\frac{(bc^x)^2}{(\ln c - \alpha)(\ln c - \kappa)}h(\xi|\alpha - \ln c, \kappa - \ln c)\right]$$

$$+ \int_0^\xi \left(\int_z^\xi \Sigma(x+u)e^{-\kappa(u-z)}du\right)^\top \left(\int_z^\xi \Sigma(x+u)e^{-\kappa(u-z)}du\right) dz.$$

The function $A(\xi, s-t)$ related to the jump parts of θ_t is solution of

$$0 = \partial_\xi A(\xi, s-t) + \lambda \tag{3.10}$$

$$\times \left(\psi\left(\begin{array}{c} -\frac{\kappa}{\kappa-\alpha}\left[\frac{a}{\alpha}\left(1 - e^{-\alpha(s-t-\xi)}\right) + \frac{bc^{x+\alpha\xi}}{\ln(c)-\alpha}\left(e^{(s-t)(\ln(c)-\alpha)} - e^{\xi(\ln(c)-\alpha)}\right)\right. \\ \left. -\frac{a}{\kappa}\left(1 - e^{-\kappa(s-t-\xi)}\right) - \frac{bc^{x+\kappa\xi}}{\ln(c)-\kappa}\left(e^{(s-t)(\ln(c)-\kappa)} - e^{\xi(\ln(c)-\kappa)}\right)\right] \end{array}\right) - 1\right).$$

In Corollary 3.4, the ODE (3.10) and the last term of Equation (3.9) can be numerically computed without particular difficulty. Notice that in the absence of jumps and for a simple function $\Sigma(x)$, the variance admits a full analytical solution. For instance, if we consider a two factors model ($d = 2$) with

$$\Sigma(x)^\top = \left(\sigma_0, \ \sigma_1 e^{\sigma_2 x}\right),$$

where $\sigma_0, \sigma_1, \sigma_2 \in \mathbb{R}^+$, the double integral in the variance (3.9) becomes

$$\int_0^\xi \left(\int_z^\xi \Sigma(x+u)e^{-\kappa(u-z)}du\right)^\top \left(\int_z^\xi \Sigma(x+u)e^{-\kappa(u-z)}du\right) dz$$

$$= \frac{\sigma_0^2}{\kappa^2}\left(\xi - \frac{2}{\kappa}\left(1 - e^{-\kappa\xi}\right) + \frac{1}{2\kappa}\left(1 - e^{-2\kappa\xi}\right)\right) + \left(\frac{\sigma_1}{\sigma_2 - \kappa}\right)^2 e^{2\sigma_2 x}$$

$$\times \left(\frac{e^{2\sigma_2\xi}}{2\kappa}\left(1 - e^{-2\kappa\xi}\right) - \frac{2e^{\sigma_2\xi}}{\kappa + \sigma_2}\left(e^{\sigma_2\xi} - e^{-\kappa\xi}\right) + \frac{1}{2\sigma_2}\left(e^{2\sigma_2\xi} - 1\right)\right).$$

In this particular case, the $_s p_{t,x}$ are fully analytical and parameters may be estimated by minimizing the mean square errors between observed and modeled survival probabilities. We do not explore this way for calibrating the model. Instead, we opt in the next section for a more complex expression of $\Sigma(x)$ and a higher number of Brownian factors. The calibration will be done by log-likelihood maximization. But before, we conclude this section by a short discussion about the calculation of survival probabilities for pricing. Insurers evaluate policies under an equivalent probability measure to P, called the pricing or risk-neutral measure and denoted here by Q. Roughly speaking, this measure allows to include a safety loading in insurance premiums, to face adverse evolution of mortality. Our model is defined under the P-measure, and we need therefore to specify the condition that Q must fulfill to keep similar dynamics of mortality rates. Given two equivalent measures Q and P, there exists a $\mathcal{F}_t-$ measurable random variable,

$\frac{dQ}{dP}$ called the Radon–Nikodym derivative which is strictly positive and such that $\mathbb{E}\left(\frac{dQ}{dP}|\mathcal{F}_0\right) = 1$. The Radon–Nikodym density process is defined as follows:

$$\eta_t = \mathbb{E}\left(\frac{dQ}{dP}|\mathcal{F}_t\right).$$

The next proposition reminds us the general structure of this Radon–Nikodym density process and the dynamic of processes under the equivalent measure Q.

Proposition 3.5. *Let $\boldsymbol{\phi}_t^\mu = \left(\phi_{1,t}^\mu, ..., \phi_{d,t}^\mu\right)$, ϕ_t^θ and ϕ_t^N be $\mathcal{F}_t$-adapted square-integrable processes. We also define*

$$\Upsilon(x, y) = \ln\left(y\frac{f_J^Q(x)}{f_J(x)}\right)$$

where $f_J(x)$ and $f_J^Q(x)$ are, respectively, the probability density functions of jumps under the P and Q measures, such that their ratio is well defined. An equivalent measure Q to P is defined by the following Radon–Nikodym derivative:

$$\eta_t = \eta_t^\mu \eta_t^\theta \eta_t^L, \tag{3.11}$$

where

$$\eta_t^\mu = \exp\left(-\int_0^t \left(\boldsymbol{\phi}_u^\mu\right)^\top dW_u^\mu - \frac{1}{2}\int_0^t \left(\boldsymbol{\phi}_u^\mu\right)^\top \boldsymbol{\phi}_u^\mu du\right),$$

$$\eta_t^\theta = \exp\left(-\int_0^t \phi_u^\theta dW_u^\theta - \frac{1}{2}\int_0^t \left(\phi_u^\theta\right)^2 du\right),$$

$$\eta_t^L = \exp\left(\int_0^t \Upsilon\left(J_u, \phi_u^N\right) dN_u - \int_0^t \lambda(\phi_u^N - 1)du\right).$$

Under the Q measure, $W_t^{\mu,Q}$ and $W_t^{\theta,Q}$ defined as:

$$W_t^{\mu,Q} = W_t^\mu + \int_0^t \boldsymbol{\phi}_u^\mu du$$

$$W_t^{\theta,Q} = W_t^\theta + \int_0^t \phi_u^\theta du$$

are Brownian motion. Whereas $(L_t)_{t\geq 0}$ is point process with an intensity $\lambda_t^Q = \lambda\phi_t^N$ and jumps distributed as $f_J^Q(x)$.

We see from the previous proposition that the dynamics of mortality and longevity processes under the pricing measure are

$$d\mu_{t,x} = \kappa\left(\theta_t\,\mu(x) - \mu_{t,x}\right) dt - \Sigma(x)^\top \boldsymbol{\phi}_t^\mu dt + \Sigma(x)^\top dW_t^{\mu,Q},$$

$$d\theta_t = \alpha(\beta(t) - \theta_t) dt - \nu\phi_t^\theta dt + \nu dW_t^{\theta,Q} + dL_t,$$

and L_t is a point process with an intensity λ_t^Q and marks distributed as $f_J^Q(x)$. The structure of the model can then widely differ under Q from the one under P. This observation motivates us to focus on changes of measure preserving the features of processes under the pricing measure. From previous results, we immediately obtain the following corollary:

Corollary 3.6. *Let us define $\mu^Q(x) : \mathbb{R}^+ \to \mathbb{R}^+$ and $\beta^Q(t) : \mathbb{R}^+ \to \mathbb{R}^+$ the baseline mortality function and the mean reversion level of θ_t under Q. If processes $\boldsymbol{\phi}_t^\mu$ and ϕ_t^θ satisfies the equalities:*

$$\kappa\theta_t\mu^Q(x) = \Sigma(x)^\top\boldsymbol{\phi}_t^\mu,$$

$$\alpha\beta^Q(t) = \nu\phi_t^\theta.$$

and ϕ_t^N is constant, then mortality rates have dynamics under Q similar to those under P.
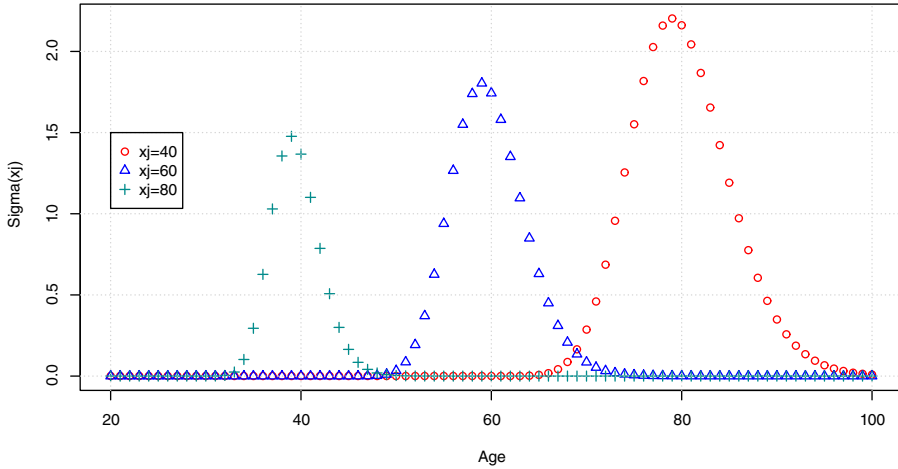
Figure 4. $\Sigma(x_j)$ *for* $x_j = 40, 60, 80$ *with* $\sigma_0 = 1$, $\sigma_1 = 0.01$, *and* $\sigma_2 = 0.09$.

If conditions of the previous corollary are fulfilled, we can apply formulas of Proposition 3.3 and Corollary 3.4 to calculate survival probabilities under the pricing measure. In the next section, we explain how to calibrate the model under P.

4. Filtering of θ_t and parameters estimation

If we remember Equation (2.2), mortality rates oscillate around a level of reversion proportional to the longevity process, $(\theta_t)_{t \geq 0}$. For this process being hidden in practice, we need to guess its most likely value based on observed mortality rates. Under the assumption of absence of jumps in the dynamics of $(\theta_t)_{t \geq 0}$, we exploit the properties of the multivariate normal distribution to filter the longevity process. Jumps are next estimated by a simple but robust "peak over threshold" method applied to the filtered longevity process. Let us now detail each step of the estimation procedure.

The data set contains sampled mortality rates at $n + 1$ equispaced times $\{0, 1, ..., n\}$ for p ages $(x_j)_{j=1,...,p}$. The interval between two successive sampling times is equal to 1 year. We assume that we have as much Brownian motions as observed ages, that is, $d = p$. We consider a step-wise volatility vector. For $(x_j)_{j=1,...,p}$, $\Sigma(.)$ is assumed equal to

$$\Sigma(x_j) = \left(\sigma_0 e^{\sigma_1 x_j} \times e^{-(j-k)^2 / (\sigma_2 x_j)^2} \right)_{k=1,...,p} \qquad j = 1, ..., p \tag{4.1}$$

where σ_0, σ_1, and $\sigma_2 \in \mathbb{R}^+$. With such a specification, the variance of mortality rates increases exponentially with age, whereas the covariance between μ_{t,x_j} and μ_{t,x_k} is proportional to a normal kernel valued at $|j - k|$. For this choice of $\Sigma(x_j)$, the covariance of variations of mortality rates for close (resp. distant) ages is high (resp. low). Figure 4 illustrates the structure of correlation between mortality rates, generated with a vector $\Sigma(.)$ such as defined by Equation (4.1). In this framework, mortality rates are locally correlated to their close neighborhood which size is determined by the parameter σ_2. The covariance function $\Sigma(.)$ is similar to the squared exponential (SE) kernel used by Wu and Wang (2018) who propose a Gaussian process regression for mortality rates. SE kernels are also commonly used in spatial statistics, and we refer the reader to Adler (1981) for details.

For non-integer ages, $\Sigma(x)$ is set to closest $\Sigma(x_j)$-value :

$$\Sigma(x) = \left\{ \Sigma(x_j) \, | \, j = \arg \min_{k \in \{1,...,p\}} |x - x_k| \right\} .$$

In numerical illustrations, we assume that the baseline mortality is a Gompertz–Makeham function. Nevertheless, we draw the attention of the reader that the estimation procedure is not constrained by the choice of the volatility and baseline mortality functions. Any other specification may be considered under the condition that $d = p$. The next proposition is a key result to filter the longevity process.

Proposition 4.1. *Let us denote the vector of mortality rates by* $\boldsymbol{\mu}_t = (\mu_{t,x})_{x=x_1,\dots,x_p}$. *In the absence of shock on the longevity process* ($L_t = 0$, *for all* $t \geq 0$), $(\boldsymbol{\mu}_{t+1}, \theta_{t+1} \mid \boldsymbol{\mu}_t, \theta_t)$ *is a multivariate normal distribution:*

$$
\left(\begin{pmatrix} \theta_{t+1} \\ \boldsymbol{\mu}_{t+1} \end{pmatrix} \middle| \begin{pmatrix} \theta_t \\ \boldsymbol{\mu}_t \end{pmatrix} \right) \sim N \left(\begin{pmatrix} m_\theta(t, \boldsymbol{\mu}_t, \theta_t) \\ m_{\boldsymbol{\mu}}(t, \boldsymbol{\mu}_t, \theta_t) \end{pmatrix}, \begin{pmatrix} \sigma_\theta^2 & \sigma_{\boldsymbol{\mu},\theta}^\top \\ \sigma_{\boldsymbol{\mu},\theta} & \sigma_{\boldsymbol{\mu}}^2 \end{pmatrix} \right), \tag{4.2}
$$

where $m_\theta(t, \boldsymbol{\mu}_t, \theta_t)$ *and* $m_{\boldsymbol{\mu}}(t, \boldsymbol{\mu}_t, \theta_t) = \left(m_{\boldsymbol{\mu}}(t, \mu_{t,x_j}, \theta_t) \right)_{j=1,\dots,p}$ *are respectively a scalar and a p-vector equal to*

$$
m_\theta(t, \boldsymbol{\mu}_t, \theta_t) = e^{-\alpha} \theta_t + \bar{\theta} \left(1 - e^{-\alpha} \right) + \frac{\alpha(\theta_0 - \bar{\theta})}{\alpha - \gamma} e^{-\gamma t} \left(e^{-\gamma} - e^{-\alpha} \right), \tag{4.3}
$$

$$
m_{\boldsymbol{\mu}}(t, \mu_{t,x_j}, \theta_t) = e^{-\kappa} \mu_{t,x_j} + \frac{\kappa \theta_t \mu(x_j) \left(e^{-\alpha} - e^{-\kappa} \right)}{\kappa - \alpha} \tag{4.4}
$$

$$
+ \mu(x_j) \frac{\kappa \alpha \bar{\theta}}{\kappa - \alpha} \left(\frac{1}{\alpha} \left(1 - e^{-\alpha} \right) - \frac{1}{\kappa} \left(1 - e^{-\kappa} \right) \right)
$$

$$
+ \mu(x_j) \frac{\kappa \alpha \left(\theta_0 - \bar{\theta} \right) e^{-\gamma t}}{\kappa - \alpha} \left(\frac{e^{-\gamma} - e^{-\alpha}}{\alpha - \gamma} - \frac{e^{-\gamma} - e^{-\kappa}}{\kappa - \gamma} \right).
$$

The variance of $\theta_{t+1} \mid \boldsymbol{\mu}_t, \theta_t$ *is equal to* $\sigma_\theta^2 = \frac{v^2}{2\alpha} \left(1 - e^{-2\alpha} \right)$. *The covariance vector* $\sigma_{\boldsymbol{\mu},\theta} = \left(\sigma_{\boldsymbol{\mu},\theta}(x_j) \right)_{j=1,\dots,p}$ *of dimension* $p - 1$ *is given by 4.*

$$
\sigma_{\boldsymbol{\mu},\theta}(x_j) = v^2 \frac{\kappa \mu(x_j)}{\kappa - \alpha} \left(\frac{1}{2\alpha} \left(1 - e^{-2\alpha} \right) - \frac{1}{\alpha + \kappa} \left(1 - e^{-(\alpha+\kappa)} \right) \right). \tag{4.5}
$$

The $(p-1) \times (p-1)$ *covariance matrix* $\sigma_{\boldsymbol{\mu}}^2 = \left(\sigma_{\boldsymbol{\mu}}^2(x_j, x_k) \right)_{j,k=1,\dots,p}$ *contains the following elements:*

$$
\sigma_{\boldsymbol{\mu}}^2(x_j, x_k) = \frac{\Sigma(x_j)^\top \Sigma(x_k)}{2\kappa} \left(1 - e^{-2\kappa} \right) + \mu(x_j)\mu(x_k) \left(\frac{\kappa v}{\kappa - \alpha} \right)^2 \tag{4.6}
$$

$$
\times \left(\frac{1}{2\alpha} \left(1 - e^{-2\alpha} \right) + \frac{1}{2\kappa} \left(1 - e^{-2\kappa} \right) - \frac{2}{\alpha + \kappa} \left(1 - e^{-(\alpha+\kappa)} \right) \right).
$$

If $\boldsymbol{\mu}_{t+1}$ and θ_{t+1} would be both observable, we could estimate the model by maximizing the log-likelihood of the distribution (4.2). In practice, θ_{t+1} is not directly visible and we need to filter it. We exploit the properties of the multivariate normal distribution to find an estimate, noted $\widehat{\theta}_{t+1}$ of θ_{t+1}. In the absence of jump, we know that that the conditional expectation of θ_{t+1} is related to parameters of the multivariate normal (4.2) as follows:

$$
\mathbb{E} \left(\theta_{t+1} | \boldsymbol{\mu}_{t+1}, \boldsymbol{\mu}_t, \theta_t \right) = m_\theta(t, \boldsymbol{\mu}_t, \theta_t) + \sigma_{\boldsymbol{\mu},\theta}^\top \left(\sigma_{\boldsymbol{\mu}}^2 \right)^{-1} \left(\boldsymbol{\mu}_{t+1} - m_{\boldsymbol{\mu}}(t, \boldsymbol{\mu}_t, \theta_t) \right),
$$

and its conditional variance is equal to

$$
\mathbb{V} \left(\theta_{t+1} | \boldsymbol{\mu}_{t+1} \right) = \sigma_\theta^2 - \sigma_{\boldsymbol{\mu},\theta}^\top \left(\sigma_{\boldsymbol{\mu}}^2 \right)^{-1} \sigma_{\boldsymbol{\mu},\theta}.
$$

Therefore, a natural recursive estimator of θ_{t+1}, based on the observation of mortality rates and previous filtered values of the longevity process is provided by:

$$\widehat{\theta}_{t+1} = \mathbb{E}\left(\theta_{t+1} | \boldsymbol{\mu}_{t+1}, \boldsymbol{\mu}_t, \widehat{\theta}_t\right) \tag{4.7}$$

$$= m_\theta(t, \boldsymbol{\mu}_t, \widehat{\theta}_t) + \sigma_{\boldsymbol{\mu},\theta}^\top \left(\sigma_{\boldsymbol{\mu}}^2\right)^{-1} \left(\boldsymbol{\mu}_{t+1} - m_{\boldsymbol{\mu}}(t, \boldsymbol{\mu}_t, \widehat{\theta}_t)\right) \ .$$

For a given set of parameters, we filter recursively $(\theta_t)_{t \geq 0}$ with Equation (4.7). The matrix $\sigma_{\boldsymbol{\mu}}^2$ is inverted by singular value decomposition (SVD). We first compute the $d \times d$ matrix U, Λ, and V where Λ is diagonal and such that $\sigma_{\boldsymbol{\mu}}^2 = U\Lambda V^\top$. The inverse is $\left(\sigma_{\boldsymbol{\mu}}^2\right)^{-1} = U\Lambda^{-1}V^\top$. After the filtration of $\widehat{\theta}_t$, the log-likelihood of the sample of observations, denoted by $\ln L$, is approached by the following sum:

$$\ln L = \sum_{t=0}^{n-1} \ln\left(f_{\boldsymbol{\mu}}\left(\boldsymbol{\mu}_{t+1} | m_{\boldsymbol{\mu}}(t, \boldsymbol{\mu}_t, \widehat{\theta}_t), \ \sigma_{\boldsymbol{\mu}}^2\right)\right) \tag{4.8}$$

where $f_{\boldsymbol{\mu}}\left(. \,|\, m_{\boldsymbol{\mu}}, \sigma_{\boldsymbol{\mu}}^2\right)$ is the probability density function of a multivariate normal with mean $m_{\boldsymbol{\mu}}$ and covariance matrix $\sigma_{\boldsymbol{\mu}}^2$. Algorithm 1 summarizes the procedure to compute the log-likelihood. For a model without jump, the parameter estimates are next found by maximizing the log-likelihood (4.8) in which $(\theta_t)_{t \geq 0}$ is replaced by its filtered values $\left(\widehat{\theta}_t\right)_{t \geq 0}$.

Algorithm 1 Procedure to compute the log-likelihood of a sample of observations, model with no jump.

Initialization :

      Set $\widehat{\theta}_0$ to θ_0 and $\ln L = 0$

Main procedure :

      For $t = 0$ to $n-1$

          1. Filter the longevity process

$$\widehat{\theta}_{t+1} = m_\theta(t, \boldsymbol{\mu}_t, \widehat{\theta}_t) + \sigma_{\boldsymbol{\mu},\theta}^\top \left(\sigma_{\boldsymbol{\mu}}^2\right)^{-1} \left(\boldsymbol{\mu}_{t+1} - m_{\boldsymbol{\mu}}(t, \boldsymbol{\mu}_t, \widehat{\theta}_t)\right) \ .$$

          4. Update the log-likelihood

$$\ln L = \ln L + \ln\left(f_{\boldsymbol{\mu}}\left(\boldsymbol{\mu}_{t+1} | m_{\boldsymbol{\mu}}(t, \boldsymbol{\mu}_t, \widehat{\theta}_t), \ \sigma_{\boldsymbol{\mu}}^2\right)\right)$$

      End loop on epochs

Proposition 4.1 and Algorithm 1 allow us to fit a model without jump. The estimation by log-likelihood maximization of the model with jump risk is a challenging task because the multivariate distribution of $(\boldsymbol{\mu}_t, \theta_t)$ and the conditional distribution $\theta_{t+1} | \boldsymbol{\mu}_{t+1}, \boldsymbol{\mu}_t, \theta_t$ are unknown in this case. To remedy this issue, we employ a peak-over-threshold approach, similar to Hainaut (2022, Chapter 4, Section 3). This approach is robust and easy to implement. The first stage consists in fitting a Gaussian model without jumps by log-likelihood maximization and in filtering $\{\widehat{\theta}_0, \widehat{\theta}_1, ..., \widehat{\theta}_n\}$. We next consider the discrete record of n variations $\{\Delta\widehat{\theta}_1, ..., \Delta\widehat{\theta}_n\}$ where $\Delta\widehat{\theta}_k = \widehat{\theta}_k - \widehat{\theta}_{k-1}$ for $k = 1, ..., n$. A mortality jump is believed to occur when one of these variations is above a threshold noted $g(\alpha)$. In practice, $g(\alpha)$ is the α-percentile of the empirical distribution of $\left(\Delta\widehat{\theta}_k\right)_{k=1,...,n}$. The time of the kth jump is therefore:

$$t_k = \min\{t_j \in \{1, ..., n\} \,|\, \Delta\widehat{\theta}_j \geq g(\alpha), \, j \geq k\}\ .$$

and the sample path of $(N_t)_{t \geq 0}$ is approached by the following time series:

$$N(t_j) = \max\{k \in \mathbb{N} \,|\, t_k \leq t_j\}\ .$$

The level of confidence α is chosen such that years during which a jump is detected, correspond to a year during which an abnormal over-mortality is identified (e.g., such as in 2020 due to the COVID 19 pandemics). In numerical illustrations, we set α to 94%. The jump size parameter, ρ, is estimated by

Table 1. *Parameter estimates for the* 20–100 *years old female and male populations,* 1950–2010.

	Female	Male		Female	Male
κ	1.320e-01	1.574e-01	α	2.827e+01	1.739e+01
a	2.299e-04	5.890e-04	ν	8.092	3.749
b	1.758e-05	7.133e-05	θ_0	1.000	1.000
c	1.113	1.098	$\bar{\theta}$	5.320e-04	1.602e-03
σ_0	2.020e-05	2.865e-05	γ	1.157e-02	8.544e-03
σ_1	6.601e-02	6.831e-02	λ	6.667e-02	6.667e-02
σ_2	1.016e-02	8.904e-03	ρ	1.650	2.523
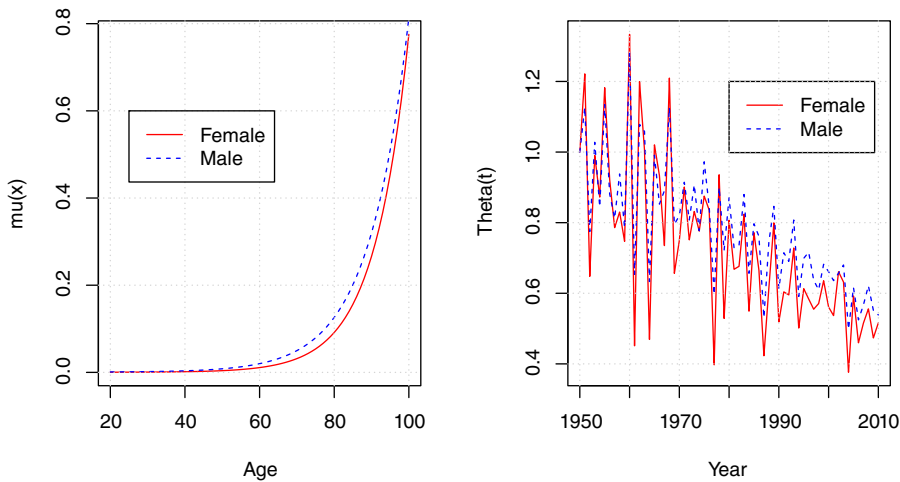
Figure 5. *Belgian female and male populations,* 1950–2010, 20 *to* 100 *years old. Left plot: estimated* $\mu(x)$ *function. Right plot: filtered* $\left(\widehat{\theta}_t\right)_{t \geq 0}$.

log-likelihood maximization under the assumption that the entire variation $\Delta\widehat{\theta}_k$ is caused by the jump. This choice is motivated by the difficulty to decompose the yearly variation into Brownian and jump components.

## 5. Estimation and validation on the Belgian population

We fit the model to the Belgian female and male populations. The data set[1] contains mortality rates from 1950 to 2010 for the range of ages 20–10 years. The number of Brownian motion is then equal to $d = p = 81$. The baseline mortality, $\mu(x)$, is the Gompertz–Makeham function. Notice that we can choose any other baseline function without affecting the performance of the estimation procedure. Indeed, the Algorithm 1 depends solely on $\mu(x)$ and not of its integral. Parameter estimates found by log-likelihood maximization are reported in Table 1. To avoid identification issues, the parameter θ_0 is set to 1. The instantaneous volatility ν of θ_t is high, but the speed of mean reversion, α, is also high and drives quickly back θ_t to $\beta(t)$. The mean reversion level $\beta(t)$ of $\theta(t)$ slowly decays from θ_0 to a lower value, $\bar{\theta}$.

Figures 5 presents calibration results. The left plots shows the estimated baseline mortality functions, $\mu(x)$. Male baseline rates are slightly higher than female ones. The right plots present filtered values of

---

[1]Source: Human Mortality Database, https://www.mortality.org/.
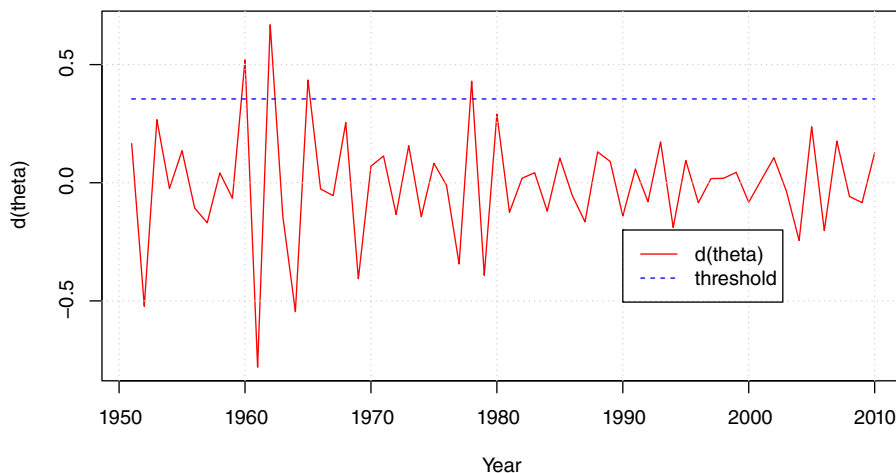
Figure 6. *Belgian female population, jumps detection by peaks over threshold. The red line is $\Delta\widehat{\theta}_t = \widehat{\theta}_{t+1} - \widehat{\theta}_t$, and the threshold is the 95% percentile of $\Delta\widehat{\theta}_t$.*
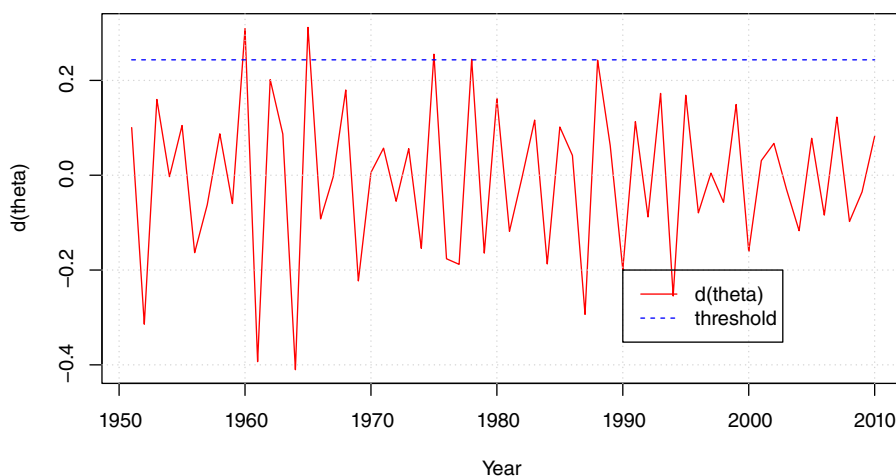

Figure 7. *Belgian male population, jumps detection by peaks over threshold. The red line is $\Delta\widehat{\theta}_t = \widehat{\theta}_{t+1} - \widehat{\theta}_t$, and the threshold is the 95% percentile of $\Delta\widehat{\theta}_t$.*

longevity process, $\widehat{\theta}_t$ for $t \in \{1950, ..., 2010\}$. Female and male longevity processes decline on average over time, but the decrease is slightly more pronounced for women than for men. This is explained by the fact that the female life expectancy has risen quicker than the one of the male population. Figures 6 and 7 show the variation of female and male longevity processes, $\Delta\widehat{\theta}_k$ and the percentile above which a jump is supposed to occur. For the female and male populations and a confidence level of $\alpha = 94\%$, we identify four shocks.

In order to evaluate the predictive power of our model, we generate 1000 mortality scenarios over the period 2011 up to 2020 for ages from 0 up to 105 years. Processes $\mu_{t,x}$ and θ_t are simulated by discretizing Equations (2.2) and (2.3) with a time step equal to $1/300$. Figures 8 and 9 display statistics about male and female mortality forecasts in 2015 and 2019. For men, the gap between average simulated and observed mortality rates is small. This gap is slightly wider for women, but observed rates are located in the 5–95% confidence interval of simulated rates. This confirms the ability of our model to generate
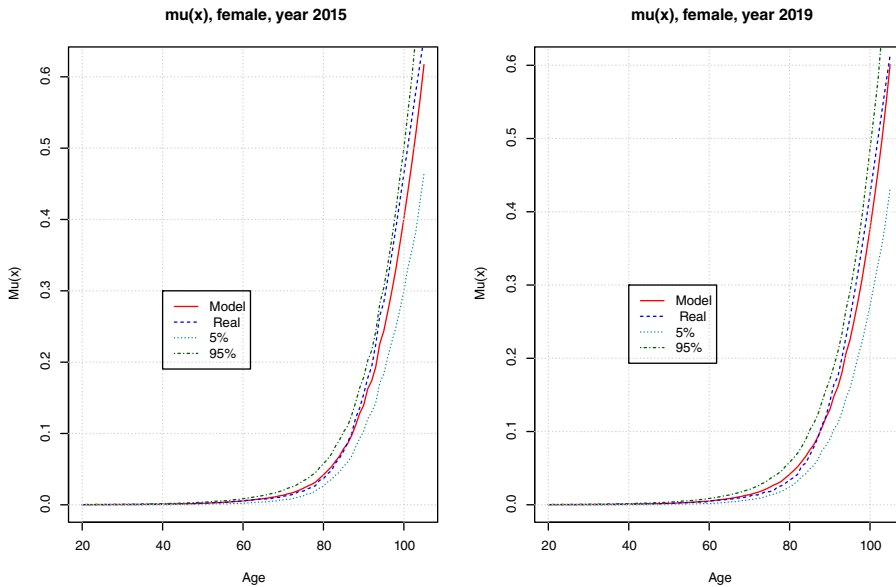
Figure 8. *Female population, forecast mortality rates with parameter estimates fitted to the sample* 1950–2010, *ages from to* 20 *to* 105.
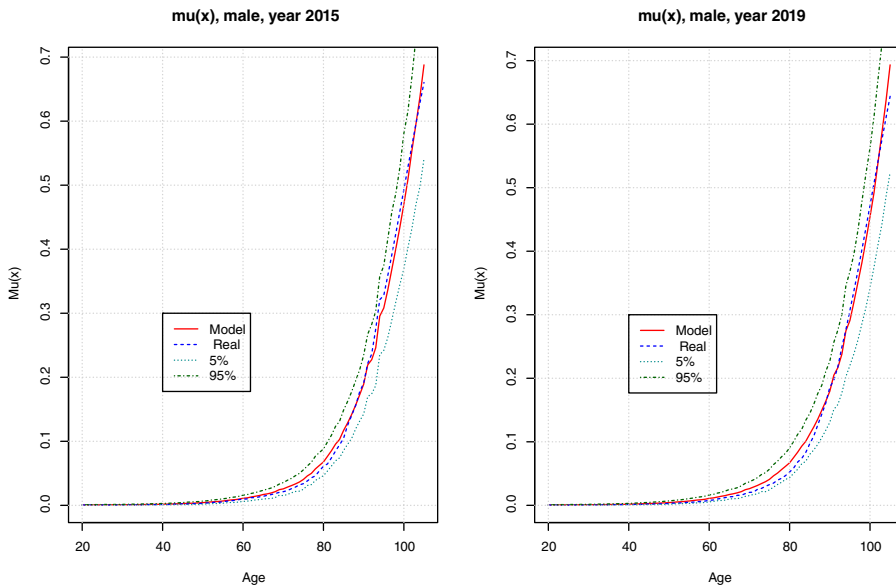
Figure 9. *Male population, forecast mortality rates with parameter estimates fitted to the sample* 1950–2010, *ages from to* 20 *to* 105.

realistic mortality scenarios. We also notice that 5% and 95% percentiles of mortality rates are positive. This observation justifies the Gaussian assumption for the dynamics of $\mu_{t,x}$ and θ_t.

We benchmark our framework to the LC and the Renshaw–Haberman (RH) models, under the assumption that the number of deaths follows a Poisson distribution. For this purpose, we use the R package STMoMo to fit the two first models to Belgian mortality rates from 1950 to 2010 and ages

Table 2. *Mean Absolute Errors (MAE) between average simulated and observed mortality rates from* 2011 *to* 2020 *and for ages from* 20 *to* 105 *years.*

	MAE, female			MAE, male		
	LC	RH	Model	LC	RH	Model
2011	3.643e-03	3.090e-03	1.458e-03	7.001e-03	4.970e-03	2.823e-03
2012	6.736e-03	7.079e-03	3.303e-03	5.206e-03	5.446e-03	4.197e-03
2013	5.924e-03	6.774e-03	3.281e-03	5.220e-03	4.960e-03	4.519e-03
2014	4.674e-03	4.550e-03	2.589e-03	8.554e-03	4.435e-03	4.648e-03
2015	8.477e-03	9.897e-03	5.642e-03	6.106e-03	4.197e-03	4.296e-03
2016	5.582e-03	6.898e-03	3.945e-03	7.353e-03	4.075e-03	4.404e-03
2017	5.968e-03	8.205e-03	4.671e-03	6.658e-03	6.089e-03	5.528e-03
2018	7.079e-03	9.647e-03	5.613e-03	7.554e-03	3.837e-03	5.155e-03
2019	5.126e-03	7.651e-03	4.494e-03	8.874e-03	4.987e-03	5.512e-03
2020	1.533e-02	1.942e-02	1.113e-02	9.655e-03	1.544e-02	1.023e-02

between 20 and 105 years. Let us recall that in the LC and RH models, the mortality rates are ruled by the following equations:

$$\ln \mu_x(t) = \alpha_x + \beta_x \kappa_t$$

$$\ln \mu_x(t) = \alpha_x + \rho_{t-x} + \beta_x^1 \kappa_t$$

where κ_t explains the global evolution of mortality and ρ_{t-x} represents the cohort effect. Both models are estimated with a Poisson law for the number of deaths. Next, we forecast 1000 mortality scenarios over the period 2011–2020 with standard auto-regressive models AR(1) for κ_t and ρ_{t-x}. Table 2 reports the mean absolute errors (MAE) between observed and average simulated mortality rates. According to this criterion, our model outperforms the LC and RH approaches from 2011 up to 2019 for the female population. Knowing that the year 2020 is particular due to the over-mortality caused by the COVID-19, this is a remarkable performance. For men, our model outperforms the LC framework up to 2019 and has smaller or similar MAE to the RH model. It is worth to mention that such a good performance of the calendar year model is obtained with only 14 parameters, whereas the LC and RH models count more than 100 coefficients to estimate.

Finally, we compare our model to a Gaussian cohort framework in which the mortality rate at time t, of a $x + t$ years old individual is

$$\mu_{x+t} = a_x + \lambda_{x+t} \,,$$

$$d\lambda_{x+t} = \kappa_x \lambda_{x+t} dt + \sigma_x dW_t^x \,,$$

where a_x, κ_x, and $\sigma_x \in \mathbb{R}^+$. The survival probability in this model, $_t p_x = \mathbb{E}\left(e^{-\int_0^t \mu_{x+s} ds}\right)$ is equal to

$$_t p_x = \exp\left(-\left(a_x - \frac{\sigma_x^2}{2\kappa_x^2}\right) t - \frac{1}{\kappa_x}\left(e^{\kappa_x t} - 1\right)\left(\lambda_x + \frac{\sigma_x^2}{\kappa_x^2} - \frac{\sigma_x^2}{4\kappa_x^2}\left(e^{\kappa_x t} + 1\right)\right)\right).$$

We fit this model to cohorts of ages 30–90 years in 2010. We use the same window of time as for the LC, RH, and calendar year model: Belgian mortality rates from 1950 to 2010. The 60 sets of parameters $(a_x, \kappa_x, \sigma_x, \lambda_x)$ are estimated by least square minimization between observed and modeled survival of probabilities. The model for the youngest cohort is estimated with 20 observations, but we wanted to use data prior to the second world war as sanitary conditions were significantly different from those of the post-war period. We have also sufficient to have a good fit and forecasts at short term. We next use the 60 models to compute mortality rates over the period 2011 (ages 31–91 years) to 2020 (ages 40–100 years). The corresponding MAE's are reported and compared to these obtained with the calendar year

Table 3. *Mean Absolute Errors (MAE) between average simulated and observed mortality rates from* 2011 *to* 2020 *and for* 30 *to* 90 *years old cohorts in* 2010.

	MAE, female		MAE, male	
	Cohort	Calendar	Cohort	Calendar
2011	1.949e-03	7.916e-04	1.643e-03	1.764e-03
2012	3.749e-03	1.302e-03	2.508e-03	1.963e-03
2013	4.630e-03	1.659e-03	3.476e-03	2.915e-03
2014	4.244e-03	2.114e-03	3.293e-03	4.724e-03
2015	8.044e-03	3.131e-03	5.443e-03	4.514e-03
2016	7.873e-03	2.831e-03	5.553e-03	5.363e-03
2017	1.059e-02	3.781e-03	8.298e-03	6.160e-03
2018	1.357e-02	5.125e-03	8.944e-03	6.350e-03
2019	1.482e-02	4.892e-03	1.003e-02	6.971e-03
2020	2.865e-02	1.314e-02	2.326e-02	1.281e-02

model in Table 3. For the female population, the calendar year model leads to a lower MAE from 2011 to 2020. For the male population, cohort and calendar model have a similar predictive power from 2011 to 2013. At longer term, the MAE computed with the calendar year are the longer. Globally, the calendar year model outperforms the cohort approach. It is also more parsimonious when we need to forecast the mortality of multiple cohorts (in our illustration, 14 instead of 60×4 parameters).

6. Estimation over 1950–2020 and life expectancy

Table 4 presents parameter estimates for Belgium, UK, and Italy. The model is fitted to mortality rates of female and male populations from 1950 up to 2020 (2019 for Italy because the 2020 data is not available yet on HMD) and for ages between 20 and 105 years. As previously, the longevity process is filtered using a subsample of $q = 15$ equispaced mortality rates. We observe many similarities between parameters for UK and Belgium. Some parameters for Italy move away from their Belgian and UK equivalents, but this is explained by the longer life expectancy in this country.

Tables 5, 6, and 7 report the prospective life expectancy in 2020 (2019 for Italy) at 0, 20, 40, 60, and 80 years old. The columns "Diffusion" and "Jump-Diffusion," respectively, report expectancies computed within a pure Brownian framework and with the jump-diffusion model. Whatever the considered population, these life expectancies are very close (less than 1/10 of year). This emphasizes the small impact of punctual mortality shocks on the evolution of life expectancy. For the year 2019 in Belgium, the (cross-sectional) female and male life expectancies were, respectively, equal to 84.0 years and 79.6 years. In comparison, the prospective life expectancies at birth calculated increase of 7.39 years for women and of 4.48 years for men. The life expectancy at birth in the UK in 2018 to 2020 was 79.0 years for males and 82.9 years for females. Our model forecast a higher prospective life expectancy of 4.79 years for men and of 4.27 years for women. In Italy, the male and female life expectancies were equal to 81.40 and 85.70 years in 2019. In comparison, female and male prospective life expectancies, respectively, rise by 9.06 and by 6.52 years.

Tables 5, 6, and 7 also report the prospective life expectancies computed with the LC model (column LC). The expectation at birth predicted by calendar year and LC models are relatively close to each other. A gap appears for life expectancies at older ages. We explain this gap by the trend of the model to overestimate mortality rates between 60 and 85 years of age. This overestimation is clearly visible in Figures 8 and 9 and is partly due to shape of the Makeham–Gompertz law.

Table 4. *Parameter estimates for Belgium, UK, and Italy.*

| | BE | | UK | | Italy | |
	Female	Male	Female	Male	Female	Male
κ	9.651e-02	1.277e-01	8.224e-02	3.182e-02	6.242e-02	7.036e-02
a	1.948e-04	5.691e-04	2.857e-04	4.736e-04	5.744e-06	2.295e-04
b	1.234e-05	6.361e-05	2.936e-05	5.520e-05	1.701e-05	1.709e-04
c	1.118	1.100	1.109	1.108	1.114	1.096
σ_0	1.905e-05	2.757e-05	7.837e-06	1.008e-05	6.769e-06	1.268e-05
σ_1	6.590e-02	6.790e-02	6.990e-02	7.238e-02	7.402e-02	7.100e-02
σ_2	9.844e-03	8.654e-03	1.295e-02	1.290e-02	1.421e-02	1.249e-02
α	1.730e+01	2.127e+01	1.007e+01	1.560e+01	3.430e+01	1.544e+01
ν	7.233	5.428	4.940	9.443	1.826e+01	3.004
θ_0	1.000	1.000	1.000	1.000	1.000	1.000
$\bar{\theta}$	2.402e-04	3.541e-04	1.096e-03	1.808e-02	3.897e-02	1.384e-02
γ	1.174e-02	1.018e-02	1.088e-02	1.335e-02	1.661e-02	1.961e-02
λ	5.714e-02	5.714e-02	7.143e-02	7.143e-02	7.246e-02	7.246e-02
ρ	1.164	2.086	1.819	1.172	1.178	2.525

Table 5. *Prospective life expectancies, e_x, in 2020 at 0, 20, 40, 60, and 80 years old.*

| | Female, Belgium | | | Male, Belgium | | |
| | e_x, 2020 | | | e_x, 2020 | | |
Age	Diffusion	Jump-diffusion	LC	Diffusion	Jump-diffusion	LC
0	91.39	91.39	90.13	84.08	84.08	85.54
20	69.04	69.04	70.51	62.17	62.17	65.88
40	46.79	46.78	50.80	40.9	40.89	46.57
60	26.00	25.98	31.82	21.63	21.62	27.83
80	9.15	9.14	13.63	7.36	7.36	11.00

7. Pricing and risk management

In this last section, we test the capacity of our model to evaluate a term life annuity and the related longevity risk. We consider annuities with three maturities 10, 15, and 20 years. They are valued with parameter estimates of Table 4 for a 60-year-old Belgian woman and man. The valuation is then done under the real measure P. The interest rate r is set to 0%. Prices are reported in the columns "0" of Tables 8 and 9. The female longevity being higher that the male one, annuity prices are slightly more expensive for women than for men.

As emphasized in Section 3, we do not need to perform simulations in simulations for determining the distribution of future values as survival probabilities admit a closed-form expression. To illustrate this, we simulate a sample of 1000 mortality rates over a period of 5 years. Annuities are re-evaluated in each scenario at the end of each year. Figure 10 shows the histogram of 20 years annuity values after 5 years. Both for male and female, we observe a small left asymmetry.

The average values, the 1% and 99% percentile of annuity values, are reported in columns "1" up to "5" of Tables 8 and 9. The exposure to the longevity risk is measured by the 99% relative Value at Risk (VaR), computed as the spread between average and percentile values divided by the average price. This exposure is higher for men than for women, whatever the duration of the annuity. This is explainable by

Table 6. *Prospective life expectancies, e_x, in 2020 at 0, 20, 40, 60, and 80 years old.*

	Female, UK			Male, UK		
	e_x, 2020			e_x, 2020		
Age	Diffusion	Jump-diffusion	LC	Diffusion	Jump-diffusion	LC
0	87.17	87.16	89.03	83.79	83.79	85.51
20	64.99	64.97	69.44	61.73	61.73	65.98
40	43.31	43.29	49.72	40.33	40.33	46.76
60	23.84	23.81	30.64	21.48	21.47	28.04
80	8.57	8.55	13.00	7.43	7.43	11.08

**Table 7.** *Prospective life expectancies, e_x, in 2019 at 0, 20, 40, 60, and 80 years old.*

	Female, Italy			Male, Italy		
	e_x, 2019			e_x, 2019		
Age	Diffusion	Jump-diffusion	LC	Diffusion	Jump-diffusion	LC
0	94.76	94.75	93.54	87.92	87.91	88.93
20	71.83	71.82	73.76	64.09	64.07	69.09
40	49.1	49.08	53.88	41.61	41.58	49.48
60	27.93	27.91	34.36	22.19	22.16	30.12
80	10.16	10.15	15.42	7.79	7.78	12.30

Table 8. *Annuity values and percentiles over the first 5 years, male, Belgium.*

Annuity 10 years	0	1	2	3	4	5
1% percentile		8.42	7.56	6.68	5.78	4.85
Price	9.31	8.50	7.65	6.76	5.84	4.90
99% percentile		8.58	7.74	6.84	5.91	4.95
99% relative VaR		0.95%	1.23%	1.28%	1.14%	1.03%

Annuity 15 years	0	1	2	3	4	5
1% percentile		12.49	11.78	11.04	10.25	9.44
Price	13.28	12.66	11.97	11.22	10.42	9.58
99% percentile		12.84	12.18	11.42	10.60	9.73
99% relative VaR		1.41%	1.77%	1.76%	1.67%	1.54%

Annuity 20 years	0	1	2	3	4	5
1% percentile		15.92	15.42	14.89	14.29	13.66
Price	16.56	16.22	15.77	15.23	14.61	13.92
99% percentile		16.52	16.13	15.58	14.97	14.22
99% relative VaR		1.88%	2.32%	2.35%	2.44%	2.16%

Table 9. *Annuity values and percentiles over the first 5 years, female, Belgium.*

Annuity 10 years	0	1	2	3	4	5
1% percentile		8.64	7.73	6.80	5.86	4.90
Price	9.583	8.69	7.79	6.85	5.90	4.94
99% percentile		8.75	7.85	6.91	5.94	4.97
99% relative VaR		0.61%	0.81%	0.79%	0.69%	0.63%
Annuity 15 years	**0**	**1**	**2**	**3**	**4**	**5**
1% percentile		13.05	12.24	11.40	10.532	9.65
Price	13.91	13.16	12.36	11.51	10.64	9.73
99% percentile		13.27	12.48	11.63	10.75	9.82
99% relative VaR		0.81%	1.03%	1.04%	1.04%	0.93%
Annuity 20 years	**0**	**1**	**2**	**3**	**4**	**5**
1% percentile		17.00	16.36	15.66	14.92	14.13
Price	17.73	17.18	16.55	15.86	15.10	14.31
99% percentile		17.35	16.76	16.07	15.30	14.48
99% relative VaR		1.00%	1.24%	1.32%	1.32%	1.19%
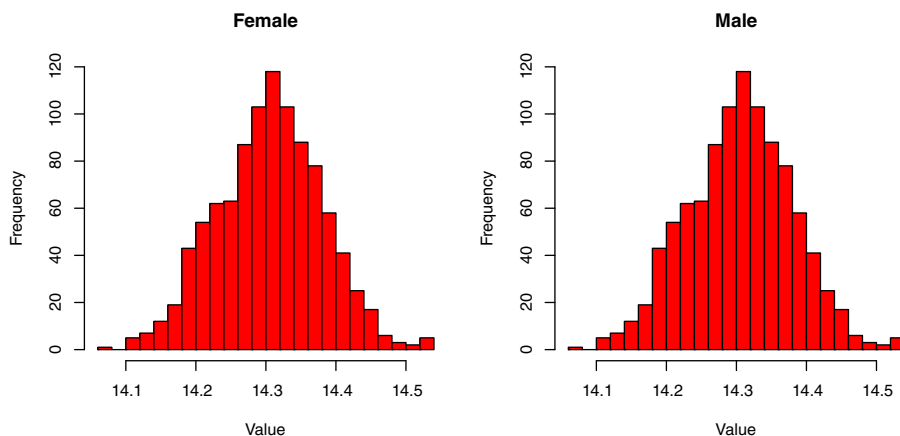
Figure 10. *Simulated values in 5 years of an annuity with an initial maturity of* 20 *years.* 1000 *simulations.*

the difference between male and female life expectancies. Women live on average longer than men and therefore cash flows paid by a term life annuity are less volatile. This drives up prices of female annuity and decreases the uncertainty related to premature deaths. The VaR for male and female, respectively, climbs up to 2.44% and 1.32% for a 20-year annuity.

To conclude this section, we analyze the probability of generating negative mortality rates. Tables 10 and 11 report the probability of generating negative mortality rates by ages and time horizons, for the Belgian female and male populations. We observe that this probability is the higher for younger ages (ages at which the average mortality is quasi null) but becomes null or neglible after 35 years old. This confirms the reliability of our model for modeling longevity for the usual annuitants range of ages.

Table 10. *Probability of generating negative mortality rates by ages and time horizons. Belgian female population.*

Age	Time horizon (years)				
	1	2	3	4	5
5	0	0.007	0.050	0.091	0.112
20	0	0.050	0.096	0.139	0.163
35	0	0.003	0.016	0.039	0.056
50	0	0	0.001	0.002	0.006
65	0	0	0	0	0
80	0	0	0	0	0
95	0	0	0	0	0

Table 11. *Probability of generating negative mortality rates by ages and time horizons. Belgian male population.*

Age	Time horizon (years)				
	1	2	3	4	5
5	0	0.165	0.17	0.168	0.126
20	0	0	0.008	0.013	0.01
35	0	0	0.001	0.006	0.007
50	0	0	0	0	0
65	0	0	0	0	0
80	0	0	0	0	0
95	0	0	0	0	0

8. Conclusions

This article proposes a calendar year model in which mortality rates revert to a long-term level that is the product of an age-specific function and of a mean-reverting process driving the evolution of longevity. This presents several interesting features. Firstly, the estimation is based on past mortality rates observed on an adjustable time window, in a similar manner to a statistical discrete approach such as the LC model. Secondly, the model manages correlation between different cohorts. Thirdly, survival probabilities admit a closed-form expression reducing the computational time. The degree of analytical tractability depends on assumptions done on age- and volatility-specific functions.

The empirical illustrations emphasize that our model is capable to explain the evolution of Belgian death rates, and that its predictive power competes with LC, Renshaw–Haberman, and cohort approaches. In Section 6, the estimation of the model to Belgian, UK, and Italian populations allows us to compare cross-sectional with prospective life expectancies. These last ones are higher of 4 up to 5 years, depending on the country. We also see that jumps have a limited impact on prospective life expectancies. Finally, our model allows us to calculate the forward longevity VaR of 10-, 15-, and 20-year annuities.

References

Adler, R.J. (1981) *The Geometry of Random Fields*. Chichester: John Wiley.
Biffis, E. (2005) Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics*, **37**(3), 443–468.

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

Cairns, A.J.G., Blake, D. and Dowd, K. (2006a) 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.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D. and Khalaf-Allah, M. (2011) Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, **48**(3), 355–367.

Cairns, A.J.G., Blake, D. and Dowd, K. (2006b) Pricing death: Frameworks for the valuation and securitization of Mortality risk. *ASTIN Bulletin*, **36**(1), 79–120.

Chen, A., Guillen, M. and Vigna, E. (2018) Solvency requirement in a unisex mortality model. *ASTIN Bulletin*, **48**(3), 1219–1243.

Hainaut, D. and Devolder, P. (2008) Mortality modelling with Lévy processes. *Insurance: Mathematics and Economics*, **42**(1), 409–418.

Hainaut, D. (2022) *Continuous Time Processes for Finance. Switching, Self-Exciting, Fractional and Other Recent Dynamics.* Springer Nature Switzerland: Bocconi & Springer Series.

Jevtic, P., Luciano, E. and Vigna, E. (2013) Mortality surface by means of continuous time cohort models. *Insurance: Mathematics and Economics*, **53**(1), 122–133.

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

Li, J.S-H., Hardy, M.R. and Tan, K.S. (2009) Uncertainty in mortality forecasting: An extension to the classical Lee-Carter approach. *ASTIN Bulletin*, **39**(1), 137–164.

Luciano, E. and Vigna, E. (2005) *Non mean reverting affine processes for stochastic mortality*. ICER working paper 4/2005.

Luciano, E. and Vigna, E. (2008) *Mortality risk via affine stochastic intensities: Calibration and empirical relevance*. MPRA Paper No. 59627.

Milevsky, M. and Promislow, S. (2001) Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics*, **29**(3), 299–318.

Protter, P.E. (2004) *Stochastic Integration and Differential Equations*, 2nd Edition. Berlin: Springer.

Renshaw, A.E. and Haberman, S. (2003) Lee-Carter mortality forecasting with age-specic enhancement. *Insurance: Mathematics and Economics*, **33**, 255–272.

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

Schrager, D. (2006) Affine stochastic mortality. *Insurance: Mathematics and Economics*, **38**, 81–97.

Wilmoth, J.R. (1993) *Computational methods for fitting and extrapolating the Lee-Carter model of mortality change*. Report, University of California, Berkeley, CA.

Wu, R. and Wang, B. (2018) Gaussian process regression method for forecasting of mortality rates. *Neurocomputing*, **316**, 232–239.

Xu, Y., Sherris, M. and Ziveyi, J. (2020) Continuous-time multi-cohort mortality modelling with affine processes. *Scandinavian Actuarial Journal*, **20**(6), 526–552.

Zeddouk, F. and Devolder, P. (2020) Mean reversion in stochastic mortality: Why and how? *European Actuarial Journal*, **10**(2), 499–525.

Zhu, N. and Bauer, D. (2022) Modeling the Risk in Mortality Projections. *Operation Researh*. Online https://doi.org/10.1287/opre.2021.2255.

Appendix

Proof of Proposition **3.2**. We can check by direct differentiation that the solution to Equation (2.2) is given by the next expression:

$$\mu_{s,x} = e^{-\kappa(s-t)}\mu_{t,x} + \kappa\mu(x)\int_t^s \theta_u\, e^{-\kappa(s-u)}du + \Sigma(x)^\top \int_t^s e^{-\kappa(s-u)}dW_u^\mu. \tag{A1}$$

In a similar manner, the longevity process $(\theta_t)_{t\geq 0}$ is equal to the sum:

$$\theta_u = e^{-\alpha(u-t)}\theta_t + \alpha\int_t^u \beta(v)\, e^{-\alpha(u-v)}dv + \nu\int_t^u e^{-\alpha(u-v)}dW_v^\theta + \int_t^u e^{-\alpha(u-v)}dL_v. \tag{A2}$$

Using this last equation, the integral of $e^{-\kappa(s-u)}\theta_u$ that is involved in Equation (A1) of $\mu_{s,x}$ is developed as follows:

$$\int_t^s \theta_u\, e^{-\kappa(s-u)}du = \theta_t \int_t^s e^{-\alpha(u-t)-\kappa(s-u)}\, du + \alpha \int_t^s \int_t^u \beta(v)\, e^{-\alpha(u-v)-\kappa(s-u)}\, dv\, du$$

$$+ v \int_t^s \int_t^u e^{-\alpha(u-v)-\kappa(s-u)}dW_v^\theta\, du + \int_t^s \int_t^u e^{-\alpha(u-v)-\kappa(s-u)}dL_v\, du\,.$$

Changing the order of integration allows us to rewrite this integral as:

$$\int_t^s \theta_u\, e^{-\kappa(s-u)}du = \frac{\theta_t\left(e^{-\alpha(s-t)}-e^{-\kappa(s-t)}\right)}{\kappa-\alpha} + \int_t^s \frac{\alpha\beta(v)\left(e^{-\alpha(s-v)}-e^{-\kappa(s-v)}\right)}{\kappa-\alpha}\, dv$$

$$+ \int_t^s \frac{v\left(e^{-\alpha(s-v)}-e^{-\kappa(s-v)}\right)}{\kappa-\alpha}dW_v^\theta + \int_t^s \frac{\left(e^{-\alpha(s-v)}-e^{-\kappa(s-v)}\right)}{\kappa-\alpha}dL_v \qquad \text{(A3)}$$

Combining this last expression with Equation (A1) leads to

$$\mu_{s,x} = e^{-\kappa(s-t)}\mu_{t,x} + \Sigma(x)^\top \int_t^s e^{-\kappa(s-u)}d\mathbf{W}_u^\mu \qquad \text{(A4)}$$

$$+\frac{\kappa\theta_t\mu(x)\left(e^{-\alpha(s-t)}-e^{-\kappa(s-t)}\right)}{\kappa-\alpha} + \int_t^s \frac{\kappa\alpha\beta(v)\mu(x)\left(e^{-\alpha(s-v)}-e^{-\kappa(s-v)}\right)}{\kappa-\alpha}\, dv$$

$$+\int_t^s \frac{\kappa v\mu(x)\left(e^{-\alpha(s-v)}-e^{-\kappa(s-v)}\right)}{\kappa-\alpha}dW_v^\theta + \int_t^s \frac{\kappa\,\mu(x)\left(e^{-\alpha(s-v)}-e^{-\kappa(s-v)}\right)}{\kappa-\alpha}dL_v$$

Finally, we perform the change of variable $v = t + z$ and obtain Equation (3.1). $\qquad \square$

Proof of Proposition **3.3**. The survival probability is the expectation of $\exp\left(-\int_0^{s-t} \mu_{t+u,x+u}du\right)$, conditionally to $\mathcal{F}_t$. From Equation (3.2), we know that $\int_0^{s-t} \mu_{t+u,x+u}du$ is the sum of a normal $N(_{s-t}m_{t,x},\ _{s-t}v_{t,x})$ and of an inhomogeneous Poisson random variable. The survival probability is then the product of the expectation of a log-normal and of the moment generating function of an inhomogeneous Poisson distribution:

$$_sp_{t,x} = \exp\left(-\ _{s-t}m_{t,x} + \frac{1}{2}\ _{s-t}v_{t,x}\right) \times \mathbb{E}\left(e^{-\frac{\kappa}{\kappa-\alpha}\int_0^{s-t}\int_z^{s-t}\mu(x+u)\left(e^{-\alpha(u-z)}-e^{-\kappa(u-z)}\right)du\, d\tilde{L}_z}\Big|\mathcal{F}_t\right). \qquad \text{(A5)}$$

If we remember that $\tilde{L}_z = L_{t+z}$, the change of variables $v = t + z$ and $w = t + u$ allows us to rewrite the double integral in this expectation as:

$$\int_0^{s-t}\int_z^{s-t} \mu(x+u)\left(e^{-\alpha(u-z)}-e^{-\kappa(u-z)}\right)du\, d\tilde{L}_z$$

$$= \int_t^s \int_v^s \mu(x+w-t)\left(e^{-\alpha(w-v)}-e^{-\kappa(w-v)}\right)\, dw\, dL_v\,.$$

Let us define the process $(X_u)_{t\le u\le s}$ as a double integral:

$$X_u = \int_t^u \int_v^s \mu(x+w-t)\left(e^{-\alpha(w-v)}-e^{-\kappa(w-v)}\right)\, dw\, dL_v\,,$$

which has the following infinitesimal dynamic at time $u \ge t$:

$$dX_u = \left(\int_u^s \mu(x+w-t)\left(e^{-\alpha(w-u)}-e^{-\kappa(w-u)}\right)\, dw\right)dL_u\,.$$

Let us find the moment generating function (mgf) of this process. We momentarily denote it by $g(u,X_u,L_u) = \mathbb{E}\left(e^{\omega X_s}|\mathcal{F}_u\right)$ for $u \ge t$. As dX_u has no deterministic drift and no Brownian part, the function

$g(.)$ solves the following Itô's equation:

$$0 = \partial_u g + \lambda \times \left(\mathbb{E}g\left(u, X_u + J \int_u^s \mu(x + w - t) \left(e^{-\alpha(w-u)} - e^{-\kappa(w-u)}\right) \, dw, L_u + J\right) - g \right).$$

with the terminal condition $g(s, X_s, L_s) = e^{\omega X_s}$. We do the ansatz that the mgf is an exponential affine function of state variables:

$$g(u, X_u, L_u) = \exp\left(A(u, s) + B(u, s)X_u + C(u, s)L_u\right)$$

and we infer that $B(u, s) = \omega$, $C(u, s) = 0$. We find that $A(u, s)$ solves for $u \geq t$ the following differential equation:

$$0 = \partial_u A(u, s) + \lambda \left(\psi \left(\omega \int_u^s \mu(x + w - t) \left(e^{-\alpha(w-u)} - e^{-\kappa(w-u)}\right) \, dw \right) - 1 \right)$$

where $\psi(z) = \mathbb{E}\left(e^{zJ}\right)$ is the mgf of jumps. If we perform the changes of variable $v = w - t$ and $\xi = u - t$, then $\partial_u A(u, s) = \partial_\xi A(\xi, s - t)$ and $A(\xi, s - t)$ solves Equation (3.6). By definition and given that $X_t = 0$, the expectation in Equation (A5) is rewritten as:

$$\mathbb{E}\left(e^{-\frac{\kappa}{\kappa-\alpha} \int_0^{s-t} \int_z^{s-t} \mu(x+u)\left(e^{-\alpha(u-z)} - e^{-\kappa(u-z)}\right) du \, d\bar{L}_z} \mid \mathcal{F}_t\right)$$

$$= \mathbb{E}\left(e^{-\frac{\kappa}{\kappa-\alpha} X_s} \mid \mathcal{F}_t\right)$$

$$= e^{A(0, s-t)}$$

where $A(\xi, s - t)$ solves Equation (3.6) with the terminal condition $A(s - t, s - t) = 0$.   □

Proof of Corollary **3.4**. The proof does not present any technical difficulties. We just develop integrals of Equation (3.4) in which we replace $\mu(x)$ by the Gompertz–Makeham function (3.7). We can show that for any $\epsilon \in \mathbb{R}^+$

$$\int_0^\xi \mu(x + u)e^{-\epsilon u} du = a \int_0^\xi e^{-\epsilon u} du + bc^x \int_0^\xi e^{(\ln c - \epsilon)u} du$$

$$= \frac{a}{\epsilon}\left(1 - e^{-\epsilon\xi}\right) + \frac{bc^x}{\ln c - \epsilon}\left(e^{(\ln c - \epsilon)\xi} - 1\right).$$

Given that

$$\int_0^\xi \mu(x + u) du = a\xi + bc^x \int_0^\xi e^{u \ln c} du$$

$$= a\xi + \frac{bc^x}{\ln c}\left(e^{\xi \ln c} - 1\right)$$

we retrieve Equation (3.8). Combining the following relation in a similar manner, we can show that

$$\int_\xi^{s-t} \mu(x + v)\left(e^{-\alpha(v-\xi)} - e^{-\kappa(v-\xi)}\right) \, dv$$

$$= \frac{a}{\alpha}\left(1 - e^{-\alpha(s-t-\xi)}\right) + \frac{bc^{x+\alpha\xi}}{\ln(c) - \alpha}\left(e^{(s-t)(\ln(c)-\alpha)} - e^{\xi(\ln(c)-\alpha)}\right)$$

$$- \frac{a}{\kappa}\left(1 - e^{-\kappa(s-t-\xi)}\right) - \frac{bc^{x+\kappa\xi}}{\ln(c) - \kappa}\left(e^{(s-t)(\ln(c)-\kappa)} - e^{\xi(\ln(c)-\kappa)}\right)$$

and infer Equation (3.10). If we insert the expression of $\mu(x)$ into the variance (3.5), we obtain

$$
{}_\xi v_{t,x} = \left(\frac{\kappa \nu}{\kappa - \alpha}\right)^2 \int_0^\xi \left(\frac{a}{\alpha}\left(1 - e^{-\alpha(\xi - z)}\right) + \frac{bc^x\left(e^{(\ln c - \alpha)(\xi - z)} - 1\right)}{\ln c - \alpha}\right.
\tag{A6}
$$

$$
\left. -\frac{a}{\kappa}\left(1 - e^{-\kappa(\xi - z)}\right) - \frac{bc^x\left(e^{(\ln c - \kappa)(\xi - z)} - 1\right)}{\ln c - \kappa}\right)^2 dz
$$

$$
+ \int_0^\xi \left(\int_z^\xi \Sigma(x + u)e^{-\kappa(u - z)}du\right)^\top \left(\int_z^\xi \Sigma(x + u)e^{-\kappa(u - z)}du\right) dz .
$$

Rewriting the first integral in the right-hand term as a function of

$$
h(\xi | \beta_1, \beta_2) = \int_0^\xi \left(1 - e^{-\beta_1(\xi - z)}\right)\left(1 - e^{-\beta_2(\xi - z)}\right) dz ,
$$

leads to Equation (3.9). □

Proof of Proposition **3.5**. We just sketch the proof as it is standard in the literature. We first use the Itô's lemma to infer that $d\eta_t^\mu = -\eta_{t-}^\mu \left(\boldsymbol{\phi}_t^\mu\right)^\top dW_t^\mu$ and $d\eta_t^\theta = -\phi_u^\theta dW_u^\theta$. Let $(M_t)_{t \geq 0}$ be a jump martingale defined by:

$$
M_t = \int_0^t e^{\Upsilon\left(J_u, \phi_u^N\right)} - 1 \, dN_u - \int_0^t \lambda \mathbb{E}\left(e^{\Upsilon\left(J_u, \phi_u^N\right)} - 1\right) du
$$

$$
= \int_0^t e^{\Upsilon\left(J_u, \phi_u^N\right)} - 1 \, dN_u - \int_0^t \lambda\left(\phi_u^N - 1\right) du ,
$$

given that $\mathbb{E}\left(e^{\Upsilon\left(J_u, \phi_u^N\right)} - 1\right) = \phi_u^N - 1$. We can check that η_t^L is solution of $d\eta_t^L = \eta_{t-}^L \, dM_t$ that emphasizes that η_{t-}^L is also martingale.

We can then rewritten η_t as a sum

$$
\eta_t = 1 - \int_0^t \eta_{s-} \left(\boldsymbol{\phi}_u^\mu\right)^\top dW_u^\mu + \int_0^t \eta_{s-}\phi_u^\theta dW_u^\theta + \int_0^t \eta_{s-}dM_u
\tag{A7}
$$

revealing that η_t is well a martingale such that $\eta_0 = 1$. The dynamics of is inferred from the Girsanov theorem (see Protter, 2004, p. 134). □

Proof of Proposition **4.1**. From Equation (A4), we immediately infer that in the absence of jumps, $\mu_{t+1,x}$ is the following function of $\mu_{t,x}$:

$$
\mu_{t+1,x} = e^{-\kappa}\mu_{t,x} + \frac{\kappa\theta_t\mu(x)\left(e^{-\alpha} - e^{-\kappa}\right)}{\kappa - \alpha} + \int_t^{t+1} \frac{\kappa\alpha\beta(v)\mu(x)\left(e^{-\alpha(t+1-v)} - e^{-\kappa(t+1-v)}\right)}{\kappa - \alpha} dv
$$

$$
+ \int_t^{t+1} \frac{\kappa\nu\mu(x)\left(e^{-\alpha(t+1-v)} - e^{-\kappa(t+1-v)}\right)}{\kappa - \alpha} dW_v^\theta + \Sigma(x)^\top \int_t^{t+1} e^{-\kappa(t+1-u)}dW_u^\mu .
$$

The components of the vector $m_{\boldsymbol{\mu}}(t, \boldsymbol{\mu}_t, \theta_t)$ are then equal to

$$
m_{\boldsymbol{\mu}}(t, \mu_{t,x_j}, \theta_t) = e^{-\kappa}\mu_{t,x_j} + \frac{\kappa\theta_t\mu(x_j)\left(e^{-\alpha} - e^{-\kappa}\right)}{\kappa - \alpha}
$$

$$
+ \frac{\kappa\alpha\mu(x)}{\kappa - \alpha}\int_t^{t+1} \beta(v)\left(e^{-\alpha(t+1-v)} - e^{-\kappa(t+1-v)}\right) dv .
$$

As $\beta(v) = \bar{\theta} + (\theta_0 - \bar{\theta})e^{-\gamma v}$, a direct calculation allows us to develop the integral:

$$\int_t^{t+1} \beta(v) \left(e^{-\alpha(t+1-v)} - e^{-\kappa(t+1-v)}\right) dv$$

$$= (\theta_0 - \bar{\theta}) e^{-\gamma t} \left(\frac{e^{-\gamma} - e^{-\alpha}}{\alpha - \gamma} - \frac{e^{-\gamma} - e^{-\kappa}}{\kappa - \gamma}\right)$$

$$+ \bar{\theta} \left(\frac{1}{\alpha} \left(1 - e^{-\alpha}\right) - \frac{1}{\kappa} \left(1 - e^{-\kappa}\right)\right).$$

and to obtain Equation (4.4). From Equation (A2) and as $L_t = 0$, we deduce that

$$\theta_{t+1} = e^{-\alpha}\theta_t + \alpha \int_t^{t+1} \beta(v) e^{-\alpha(t+1-v)}dv + v \int_t^{t+1} e^{-\alpha(t+1-v)}dW_v^\theta$$

and the expectation of θ_{t+1}, conditionally to the information available up to t is equal to.

$$m_\theta(t, \boldsymbol{\mu}_t, \theta_t) = e^{-\alpha}\theta_t + \alpha \int_t^{t+1} \beta(v) e^{-\alpha(t+1-v)}dv.$$

As the integral $\int_t^{t+1} \beta(v) e^{-\alpha(t+1-v)}dv$ is equal to

$$\int_t^{t+1} \beta(v) e^{-\alpha(t+1-v)}dv = \frac{\bar{\theta}}{\alpha} \left(1 - e^{-\alpha}\right) + \frac{(\theta_0 - \bar{\theta})}{\alpha - \gamma} e^{-\gamma t} \left(e^{-\gamma} - e^{-\alpha}\right),$$

then we obtain Equation (4.3). The variance of theta is

$$\sigma_\theta^2 = v^2 \int_t^{t+1} e^{-2\alpha(t+1-v)}dv = \frac{v^2}{2\alpha} \left(1 - e^{-2\alpha}\right).$$

The covariance is a vector $\boldsymbol{\sigma}_{\mu,\theta} = \left(\sigma_{\mu,\theta}(x)\right)_{x=x_1...x_u}$ that is given by:

$$\sigma_{\mu,\theta}(x) = v^2 \frac{\kappa\mu(x)}{\kappa - \alpha} \int_t^{t+1} \left(e^{-2\alpha(t+1-v)} - e^{-(\alpha+\kappa)(t+1-v)}\right) dv$$

$$= v^2 \frac{\kappa\mu(x)}{\kappa - \alpha} \left(\frac{1}{2\alpha} \left(1 - e^{-2\alpha}\right) - \frac{1}{\alpha + \kappa} \left(1 - e^{-(\alpha+\kappa)}\right)\right).$$

The covariance matrix of mortality rates contains the following elements:

$$\sigma_\mu^2(x_j, x_k) = \mu(x_j)\mu(x_k) \left(\frac{\kappa v}{\kappa - \alpha}\right)^2 \int_t^{t+1} \left(e^{-\alpha(t+1-v)} - e^{-\kappa(t+1-v)}\right)^2 dv$$

$$+ \Sigma(x_j)^\top \Sigma(x_k) \int_t^{t+1} e^{-2\kappa(t+1-v)}dv$$

$$= \mu(x_j)\mu(x_k) \left(\frac{\kappa v}{\kappa - \alpha}\right)^2 \left(\frac{1}{2\alpha} \left(1 - e^{-2\alpha}\right) + \frac{1}{2\kappa} \left(1 - e^{-2\kappa}\right) - \frac{2}{\alpha + \kappa} \left(1 - e^{-(\alpha+\kappa)}\right)\right)$$

$$+ \frac{\Sigma(x_j)^\top \Sigma(x_k)}{2\kappa} \left(1 - e^{-2\kappa}\right)$$

$\square$