

StMoMo: An R Package for Stochastic Mortality Modelling

Andrés M. Villegas*, Vladimir Kaishev, Pietro Millossovich

Cass Business School, City University London, United Kingdom

Abstract

In this paper we mirror the framework of generalised linear models to define the family of generalised Age-Period-Cohort stochastic mortality models which encompasses the vast majority of stochastic mortality projection models proposed to date. In addition, we introduce the **R** package **StMoMo** which exploits the unifying framework of the generalised Age-Period-Cohort family to provide tools for fitting stochastic mortality models, analysing their goodness of fit and performing mortality projections with them. We illustrate some of the capabilities of the package by doing a comparison of several stochastic mortality models applied to the England and Wales population.

Keywords: Mortality modelling; mortality forecasting; generalised linear models

1. Introduction

During the last two centuries the developed world experienced a persistent increase in life expectancy. For instance, Oeppen and Vaupel (2002) estimate that during the last 160 years the world record in female life expectancy at birth has increased at an approximate steady pace of 3 month per year. This increase in life expectancy, though a sign of social progress, poses a challenge to governments, private pension plans and life insurers as it can have a potential impact on pension and health costs. Actuaries and demographers have recognised the issues of an ageing population and rising longevity and have thus devoted a fair amount of effort to the development of statistical techniques for the modelling and projection of mortality rates.

One of the most influential approaches for the stochastic modelling of mortality rates is the parsimonious mortality model proposed by Lee and Carter (1992). This model uses principal component analysis to decompose the evolution of age-specific mortality rates into a bilinear combination of age and period parameters, with these period parameters being treated as time series to produce mortality projections. The Lee-Carter model has inspired numerous variants and extensions. For instance, Lee and Miller (2001), Booth et al. (2002), and Brouhns et al. (2002) have proposed alternative estimation approaches in order to improve the goodness-of-fit and forecasting properties of the model. In particular, Brouhns

*Corresponding author

Email addresses: `Andres.Villegas.1@cass.city.ac.uk` (Andrés M. Villegas),
`Vladimir.Kaishev.1@city.ac.uk` (Vladimir Kaishev), `Pietro.Millossovich.1@city.ac.uk` (Pietro Millossovich)

et al. (2002) propose a more formal statistical approach to estimating the parameters by embedding the Lee-Carter model into a Poisson regression setting. Other authors have extended the Lee-Carter model by including additional terms, such as more than one bilinear age-period term (Renshaw and Haberman, 2003; Hyndman and Ullah, 2007), or a cohort effect (Renshaw and Haberman, 2006; Currie, 2006).

The two factor model introduced by Cairns et al. (2006) is one of the most prominent extensions to the Lee-Carter method. This model relies on the linearity of the logit of one-year death probability at older ages. Specifically, the Cairns-Blake-Dowd (CBD) model assumes that for a fix year the logit of one-year death probability is a linear function of age, and treats the intercept and slope parameters across years as stochastic processes. Cairns et al. (2009) consider three extensions to the original CBD model by incorporating combinations of a quadratic age term and a cohort effect. Plat (2009) has combined the CBD method with some features of the Lee-Carter method to produce a model that is suitable for full age ranges and captures the cohort effect.

Given the abundance and rapid increase in the number of stochastic mortality models proposed in the literature, there have been some recent attempts to find the commonalities among these models. Hunt and Blake (2014b) review holistically the structure of mortality models and describe an Age-Period-Cohort model structure which encompasses the vast majority of stochastic mortality models. Currie (2014) shows how many common mortality models can be expressed in the terms of generalised linear models or generalised non-linear models.

In this paper, we build on the work of Hunt and Blake (2014b) and Currie (2014) to define the family of Generalised Age-Period-Cohort stochastic mortality models by mirroring the terminology of generalised linear models. In addition, we introduce the **R** package **StMoMo**¹ which exploits the unifying framework of the Generalised Age-Period-Cohort family to provide computational tools for implementing many of the stochastic mortality models proposed to date. The **StMoMo** package is available at <https://github.com/amvillegas/StMoMo>. Version 0.2 has been used for this paper.

Several packages for mortality modelling are available in the **R** environment (R Core Team, 2014). The package **demography** (Hyndman, 2014), whose usage is explained in detail in Booth et al. (2014), implements the original Lee-Carter model along with the Lee and Miller (2001), Booth et al. (2002), and Hyndman and Ullah (2007) variants. The **ilc** package (Butt et al., 2014) implements the Renshaw and Haberman (2006) cohort extension of the Lee-Carter model together with the Lee-Carter model under a Poisson regression framework. The **LifeMetrics** **R** functions implement the original CBD model and the three extended CBD models considered in Cairns et al. (2009), along with the Lee-Carter model (using Poisson maximum likelihood), the Age-Period-Cohort model of Currie (2006) and the

¹The acronym StMoMo, pronounced Saint Momo, stands for Stochastic Mortality Modelling. Momo is the king of Carnivals in numerous Latin American festivities (Wikipedia, 2014).

Renshaw and Haberman (2006) model. This package, which is not on CRAN, is available at <http://www.macs.hw.ac.uk/~andrewc/lifemetrics/>.

There are however several areas where the previous packages fall short and which our package **StMoMo** seeks to overcome. First, the existing packages are based on model-specific fitting algorithms limiting the models available to those already predefined in the packages. By contrast, **StMoMo** is general enough to allow users to easily expand the numbers of models available. In addition, whilst **StMoMo** provides forecasting and simulation functions for any model within the Generalised Age-Period-Cohort family, the existing packages only provide such facilities for a limited number of models. For instance, package **ilc** only includes forecasting functions for the Lee-Carter model. Similarly, the simulation facilities of package **Lifemetrics** are limited to the Lee-Carter and the standard CBD model. Finally, **StMoMo** provides facilities which are not available in existing packages, such as tools for analysing the goodness-of-fit and evaluating the impact of parameter uncertainty using bootstrapping techniques.

In the rest of this paper we describe the statistical framework underlying **StMoMo** and illustrate the usage of our package. To aid illustration, we use as a running example a comparison of several stochastic mortality models fitted to the England and Wales population. This example is in the same spirit of the comparison exercises of Cairns et al. (2009, 2011), Haberman and Renshaw (2011) and Lovász (2011) and as such we show how several of the analysis performed in those papers can easily be replicated using **StMoMo**. The structure of this paper is as follows. In Section 2 we make precise our notation. In Section 3, we mirror the framework of generalised linear models to define the family of Generalised Age-Period-Cohort (GAPC) stochastic mortality models and describe how many of the mortality models discussed in the literature can be framed within this family. Section 4 discusses how **StMoMo** implements the GAPC family of models. In Section 5, we describe the fitting of GAPC mortality models and illustrate how this can be accomplished using **StMoMo**. Section 5 discusses the evaluation of the goodness-of-fit of GAPC models. In Section 7 we discuss the forecasting and simulation of GAPC models using time series techniques. Section 8 describes the use of bootstrapping techniques to incorporate parameter uncertainty in the estimation and forecasting of GAPC mortality models. Finally, Section 9 concludes and discusses possible extensions of the **StMoMo** package.

2. Notation and Data

Let the random variable D_{xt} denote the number of deaths in a population at age x last birthday during calendar year t . Also let d_{xt} denote the observed number of deaths, E_{xt}^c the central exposed to risk at age x in year t , and E_{xt}^i the corresponding initial exposed to risk. The one-year mortality rate for an individual aged x last birthday and in calendar year t , denoted q_{xt} , can be estimated as $\hat{q}_{xt} = d_{xt}/E_{xt}^i$. The force of mortality and central death rates are denoted by μ_{xt} and m_{xt} , respectively, with the empirical estimate of the latter being $\hat{m}_{xt} = d_{xt}/E_{xt}^c$. Under the assumption that the force of mortality is constant over each year of age, i.e. from age x to age $x + 1$ and t to $t + 1$, then the force

of mortality μ_{xt} and the death rate m_{xt} coincide. We assume that this is the case throughout.

In **StMoMo** and throughout this paper we assume that deaths, d_{xt} , and either central exposures, E_{xt}^c , or initial exposures, E_{xt}^i , are available in a rectangular array format comprising ages (on the rows) $x = x_1, x_2, \dots, x_k$, and calendar years (on the columns) $t = t_1, t_2, \dots, t_n$. When only central exposures are available and initial exposures are required (or vice-versa), one can approximate the initial exposures by adding half the matching reported numbers of deaths to the central exposures, i.e., $E_{xt}^i \approx E_{xt}^c + \frac{1}{2}d_{xt}$. When the context is clear, we may write E_{xt} to refer to E_{xt}^i or E_{xt}^c .

3. Generalised Age-Period-Cohort stochastic mortality models

Some authors have recently sought to identify the similarities amongst stochastic mortality models. For instance, Hunt and Blake (2014b) describe an Age-Period-Cohort model structure which encompasses the vast majority of stochastic mortality models. In another interesting contribution, Currie (2014) shows how many common mortality models can be expressed in the standard terminology of either generalised linear models or generalised non-linear models. In this section, we build on the aforementioned papers to define the family of Generalised Age-Period-Cohort (GAPC) stochastic mortality models.

Akin to generalised linear models (see e.g. McCullagh and Nelder (1989)), a GAPC stochastic mortality model is comprised of four components:

- i. The *random component*: the number of deaths D_{xt} are assumed to follow a conditional Poisson distribution or a conditional Binomial distribution so that

$$D_{xt} \sim \text{Poisson}(E_{xt}^c \mu_{xt}) \quad (1)$$

or

$$D_{xt} \sim \text{Binomial}(E_{xt}^i, q_{xt}) \quad (2)$$

with $\mathbb{E}(D_{xt}/E_{xt}^c) = \mu_{xt}$ and $\mathbb{E}(D_{xt}/E_{xt}^i) = q_{xt}$, respectively.

- ii. The *systematic component*: following Hunt and Blake (2014b) the effects of age x , calendar year t and year-of-birth (cohort) $c = t - x$ are captured through a *predictor* η_{xt} given by:

$$\eta_{xt} = \alpha_x + \sum_{i=1}^N \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x} \quad (3)$$

Here:

- The term α_x is a static age function capturing the general shape of mortality by age.
- N is an integer indicating the number of age-period terms describing the mortality trends, with each time index $\kappa_t^{(i)}$, $i = 1, \dots, N$, contributing in specifying the mortality trend and $\beta_x^{(i)}$ modulating its effect across ages.

- The term γ_{t-x} accounts for the cohort effect with $\beta_x^{(0)}$ modulating its effect across ages.

The age modulating terms $\beta_x^{(i)}$, $i = 0, 1, \dots, N$, can be either pre-specified functions of age, i.e. $\beta_x^{(i)} \equiv f^i(x)$, as in CBD type models, or non-parametric terms without any prior structure and which need to be estimated as in the Lee-Carter model. In the GAPC family we assume that the period indexes $\kappa_t^{(i)}$, $i = 1, \dots, N$, and the cohort index γ_{t-x} are stochastic processes rather than smooth functions of time or cohort. This is the key feature that allows the stochastic projection of GAPC models and thus the generation of probabilistic forecasts of future mortality rates.

- iii. The *link function* $g(\cdot)$ associating the random component and the systematic component so that

$$g\left(\mathbb{E}\left(\frac{D_{xt}}{E_{xt}}\right)\right) = \eta_{xt} \quad (4)$$

Although a number of link functions would be possible it is convenient to use the so-called canonical link and pair the Poisson distribution with the log link function and the Binomial distribution with the logit link function (see e.g. Currie (2014) for a discussion of this in the context of mortality models and McCullagh and Nelder (1989) in the wider context of GLMs).

- iv. The *set of parameter constraints*: most stochastic mortality models are only identifiable up to a transformation and thus require parameter constraints to ensure unique parameter estimates. These parameter constraints are applied using the *constraint function* $h : \theta \rightarrow \tilde{\theta}$ which maps an arbitrary set of parameter estimates

$$\theta := \{\alpha_{xt}, \beta_x^{(1)}, \dots, \beta_x^{(N)}, \kappa_t^{(1)}, \dots, \kappa_t^{(N)}, \beta_x^{(0)}, \gamma_{t-x}\}$$

to a set of parameter estimates

$$\tilde{\theta} := \{\tilde{\alpha}_{xt}, \tilde{\beta}_x^{(1)}, \dots, \tilde{\beta}_x^{(N)}, \tilde{\kappa}_t^{(1)}, \dots, \tilde{\kappa}_t^{(N)}, \tilde{\beta}_x^{(0)}, \tilde{\gamma}_{t-x}\}$$

satisfying the model constraints.

Most stochastic mortality models proposed in the literature belong to the GAPC family. This includes the original Lee-Carter model, the extensions of the Lee-Carter proposed in Renshaw and Haberman (2003, 2006), the original CBD model, and the extended CBD models of Cairns et al. (2009). In addition, all the model structures considered in Haberman and Renshaw (2011), Lovász (2011) and van Berkum et al. (2014), as well as the models of Plat (2009), Aro and Pennanen (2011), OHare and Li (2012) and Börger et al. (2013), are part of the GAPC class of models.

Next, we describe in detail some of these models highlighting how they can be framed within the GAPC family.

3.1. Lee-Carter model under a Poisson setting

The Lee Carter model as implemented by Brouhns et al. (2002) assumes a Poisson distribution of the deaths (*random component*) using a log *link function* to target the force of mortality μ_{xt} . The predictor structure (*systematic component*) proposed by Lee and Carter (1992) assumes that there is a static age function, α_x , a unique non-parametric age-period term ($N = 1$), and no cohort effect. Thus, the predictor is given by:

$$\eta_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} \quad (5)$$

The Lee-Carter model is only identifiable up to a transformation as for arbitrary real constants c_1 and $c_2 \neq 0$ the parameters in Equation (5) can be transformed in the following way

$$\left\{ \alpha_x, \beta_x^{(1)}, \kappa_t^{(1)} \right\} \rightarrow \left\{ \alpha_x + c_1 \beta_x^{(1)}, \frac{1}{c_2} \beta_x^{(1)}, c_2 (\kappa_t^{(1)} - c_1) \right\}, \quad (6)$$

leaving η_{xt} unchanged. To ensure identifiability of the model Lee and Carter (1992) suggest the following *set of parameter constraints*

$$\sum_x \beta_x^{(1)} = 1, \quad \sum_t \kappa_t^{(1)} = 0, \quad (7)$$

which can be imposed by choosing

$$c_1 = \frac{1}{n} \sum_t \kappa_t^{(1)}, \quad c_2 = \sum_x \beta_x^{(1)} \quad (8)$$

in transformation (6).

3.2. Renshaw and Haberman model: Lee-Carter with cohort effects

Renshaw and Haberman (2006) generalise the Lee-Carter model by incorporating a cohort effect to obtain the predictor (*systematic component*):

$$\eta_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t + \beta_x^{(0)} \gamma_{t-x} \quad (9)$$

To estimate the model Renshaw and Haberman (2006) assume a Poisson distribution of deaths (*random component*) and use a log *link function* targeting the force of mortality μ_{xt} . As with the Lee-Carter model, the predictor η_{xt} is invariant to the transformation:

$$\left\{ \alpha_x, \beta_x^{(1)}, \kappa_t^{(1)}, \beta_x^{(0)}, \gamma_{t-x} \right\} \rightarrow \left\{ \alpha_x + c_1 \beta_x^{(1)} + c_2 \beta_x^{(1)}, \frac{1}{c_3} \beta_x^{(1)}, c_3 (\kappa_t^{(1)} - c_1), \frac{1}{c_4} \beta_x^{(0)}, c_4 (\gamma_{t-x} - c_2) \right\}, \quad (10)$$

where $c_1, c_2, c_3 \neq 0$ and $c_4 \neq 0$ are real constants. Identifiability of the model can be ensured using the following *set of parameter constraints*:

$$\sum_x \beta_x^{(1)} = 1, \quad \sum_t \kappa_t^{(1)} = 0, \quad \sum_x \beta_x^{(0)} = 1, \quad \sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c = 0, \quad (11)$$

which can be imposed by setting

$$c_1 = \frac{1}{n} \sum_t \kappa_t^{(1)}, \quad c_2 = \frac{1}{n+k-1} \sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c, \quad c_i = \sum_x \beta_x^{(i)}, \quad i = 3, 4, \quad (12)$$

in transformation (10).

Renshaw and Haberman (2006) also consider several substructures of predictor (9) obtained by setting to a constant one or both of the age modulating terms. Of particular interest is the substructure

$$\eta_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} + \gamma_{t-x} \quad (\beta_x^{(0)} = 1) \quad (13)$$

which has been suggested by Haberman and Renshaw (2011) as a simpler structure that resolves some stability issues of the original model.

3.3. Age-Period-Cohort model

Another commonly used substructure of the Renshaw and Haberman model is the so-called Age-Period-Cohort (APC) model

$$\eta_{xt} = \alpha_x + \kappa_t^{(1)} + \gamma_{t-x} \quad (\beta_x^{(1)} = 1, \beta_x^{(0)} = 1), \quad (14)$$

which has a long-standing tradition in the fields of medicine and demography (see, e.g., Clayton and Schifflers (1987) and Hobcraft et al. (1982)), but has not been widely used in the actuarial literature until considered by Currie (2006). The APC model is known to be unidentified under the following two transformations

$$\left\{ \alpha_x, \kappa_t^{(1)}, \gamma_{t-x} \right\} \rightarrow \left\{ \alpha_x + \phi_1 - \phi_2 x, \kappa_t^{(1)} + \phi_2 t, \gamma_{t-x} - \phi_1 - \phi_2(t-x) \right\} \quad (15)$$

$$\left\{ \alpha_x, \kappa_t^{(1)}, \gamma_{t-x} \right\} \rightarrow \left\{ \alpha_x + c_1, \kappa_t^{(1)} - c_1, \gamma_{t-x} \right\}, \quad (16)$$

where c_1 , ϕ_1 , and ϕ_2 are real constants. However, we can ensure identifiability of the model by imposing the *set of constraints*:

$$\sum_t \kappa_t^{(1)} = 0, \quad \sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c = 0, \quad \sum_{c=t_1-x_k}^{t_n-x_1} c \gamma_c = 0, \quad (17)$$

where the last two constraints ensure that the cohort effect fluctuates around zero with no discernible linear trend. Following Haberman and Renshaw (2011, Appendix A), the constraints on the cohort effect can be imposed by applying transformation (15) with constants ϕ_1 and ϕ_2 obtained by regressing γ_{t-x} on $t-x$, so that

$$\gamma_{t-x} = \phi_1 + \phi_2(t-x) + \epsilon_{t-x}, \quad \epsilon_{t-x} \sim N(0, \sigma^2) \quad \text{i.i.d.} \quad (18)$$

The constraint on the period index can then be imposed by applying transformation (16) with

$$c_1 = \frac{1}{n} \sum_t \kappa_t^{(1)} \quad (19)$$

3.4. CBD model

Cairns et al. (2006) propose a predictor structure (*systematic component*) with two age-period terms ($N = 2$) with pre-specified age-modulating parameters $\beta_x^{(1)} = 1$ and $\beta_x^{(2)} = x - \bar{x}$, no static age function and no cohort effect. Thus, the predictor of the CBD model is given by:

$$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)}, \quad (20)$$

where \bar{x} is the average age in the data. The CBD model does not have identifiability issues and hence the *set of parameter constraints* is empty. In order to estimate the parameter of the CBD model we can follow Haberman and Renshaw (2011) and assume a Binomial distribution of deaths (*random component*) using a logit *link function* targeting the one-year death probabilities q_{xt} .

3.5. M7: Quadratic CBD model with cohort effects

Cairns et al. (2009) extend the original CBD model by adding a cohort effect and a quadratic age effect to obtain the predictor (*systematic component*):

$$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + \kappa_t^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2) + \gamma_{t-x}, \quad (21)$$

where $\hat{\sigma}_x^2$ is the average value of $(x - \bar{x})^2$. This model, usually referred to as model M7, is not identifiable under the transformation

$$\begin{aligned} \left\{ \kappa_t^{(1)}, \kappa_t^{(3)}, \kappa_t^{(3)}, \gamma_{t-x} \right\} &\rightarrow \left\{ \kappa_t^{(1)} + \phi_1 + \phi_2(t - \bar{x}) + \phi_3((t - \bar{x})^2 + \hat{\sigma}^2), \kappa_t^{(2)} - \phi_2 - 2\phi_3(t - \bar{x}), \right. \\ &\quad \left. \kappa_t^{(3)} + \phi_3, \gamma_{t-x} - \phi_1 - \phi_2(t - x) - \phi_3(t - x)^2 \right\}, \end{aligned} \quad (22)$$

for real constants ϕ_1 , ϕ_2 and ϕ_3 . To identify the model Cairns et al. (2009) impose the *set of constraints*:

$$\sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c = 0, \quad \sum_{c=t_1-x_k}^{t_n-x_1} c\gamma_c = 0, \quad \sum_{c=t_1-x_k}^{t_n-x_1} c^2\gamma_c = 0, \quad (23)$$

which ensure that the cohort effect fluctuates around zero and has no discernible linear or quadratic trend. Following Haberman and Renshaw (2011, Appendix A), these constraints can be imposed by applying transformation (22) with constants ϕ_1 , ϕ_2 and ϕ_3 obtained by regressing γ_{t-x} on $t - x$ and $(t - x)^2$, so that

$$\gamma_{t-x} = \phi_1 + \phi_2(t - x) + \phi_3(t - x)^2 + \epsilon_{t-x}, \quad \epsilon_{t-x} \sim N(0, \sigma^2) \quad \text{i.i.d.} \quad (24)$$

Cairns et al. (2009) also consider the simpler predictor structures

$$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + \gamma_{t-x}, \quad (25)$$

$$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + (x_c - x)\gamma_{t-x}, \quad (26)$$

where x_c is a constant parameter to be estimated. These structures are typically referred to as models M6 and M8, respectively.

3.6. Plat model

Plat (2009) combines the CBD method with some features of the Lee-Carter method to produce a model that is suitable for full age ranges and captures the cohort effect. The proposed predictor structure (*systematic component*) assumes that there is a static age function, α_x , three age-period terms ($N = 3$) with pre-specified age-modulating parameters $\beta_x^{(1)} = 1$, $\beta_x^{(2)} = \bar{x} - x$, $\beta_x^{(3)} = (\bar{x} - x)^+ = \max(0, \bar{x} - x)$, and a cohort effect with pre-specified age-modulating parameters $\beta_x^{(0)} = 1$. Thus, the predictor is given by:

$$\eta_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} + \gamma_{t-x} \quad (27)$$

Plat (2009) targets the force of mortality μ_{xt} with the log *link* and estimates the parameters of the model by assuming a Poisson distribution of the deaths (*random component*). The following parameter transformations leave the predictor in (27) unchanged:

$$\left\{ \alpha_x, \kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \gamma_{t-x} \right\} \rightarrow \left\{ \alpha_x + \phi_1 - \phi_2 x + \phi_3 x^2, \kappa_t^{(1)} + \phi_2 t + \phi_3 (t^2 - 2\bar{x}t), \right. \\ \left. \kappa_t^{(2)} + 2\phi_3 t, \kappa_t^{(3)}, \gamma_{t-x} - \phi_1 - \phi_2(t - x) - \phi_3(t - x)^2 \right\} \quad (28)$$

$$\left\{ \alpha_x, \kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \gamma_{t-x} \right\} \rightarrow \left\{ \alpha_x + c_1 + c_2(\bar{x} - x) + c_3(\bar{x} - x)^+, \right. \\ \left. \kappa_t^{(1)} - c_1, \kappa_t^{(2)} - c_2, \kappa_t^{(3)} - c_3, \gamma_{t-x} \right\}, \quad (29)$$

where $c_1, c_2, c_3, \phi_1, \phi_2$, and ϕ_3 are any real constants. The following *set of parameter constraints* can be imposed to ensure identifiability:

$$\sum_t \kappa_t^{(1)} = 0, \sum_t \kappa_t^{(2)} = 0, \sum_t \kappa_t^{(3)} = 0, \sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c = 0, \sum_{c=t_1-x_k}^{t_n-x_1} c\gamma_c = 0, \sum_{c=t_1-x_k}^{t_n-x_1} c^2\gamma_c = 0 \quad (30)$$

The first three constraints ensure that the period indexes are centred around zero, while the last three constraints ensure that the cohort effect fluctuates around zero and has no linear or quadratic trend. Following Haberman and Renshaw (2011, Appendix A), the constraints on the cohort effect can be imposed by applying transformation (28) with constants ϕ_1, ϕ_2 , and ϕ_3 obtained by regressing γ_{t-x} on $t - x$ and $(t - x)^2$, so that

$$\gamma_{t-x} = \phi_1 + \phi_2(t - x) + \phi_3(t - x)^2 + \epsilon_{t-x}, \quad \epsilon_{t-x} \sim N(0, \sigma^2) \quad \text{i.i.d.}, \quad (31)$$

The constraints on the period indexes can then be imposed by applying transformation (29) with

$$c_i = \frac{1}{n} \sum_t \kappa_t^{(i)}, \quad i = 1, 2, 3. \quad (32)$$

In the cases where only older ages are of interests, Plat (2009) suggests to drop the third period term from predictor (27) to have²:

$$\eta_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + \gamma_{t-x}. \quad (33)$$

We note that this reduced Plat model has the same identifiability issues as the complete model with the omission of the transformations and constraints involving $\kappa_t^{(3)}$ and c_3 .

²Note that the reduced Plat model is essentially the M6 model of Cairns et al. (2009) with an added static age term α_x .

Table 1: Model structures considered in this paper.

Model	Predictor
LC	$\eta_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)}$
CBD	$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x}) \kappa_t^{(2)}$
APC	$\eta_{xt} = \alpha_x + \kappa_t^{(1)} + \gamma_{t-x}$
RH	$\eta_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} + \gamma_{t-x}$
M7	$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x}) \kappa_t^{(2)} + ((x - \bar{x})^2 - \hat{\sigma}_x^2) \kappa_t^{(3)} + \gamma_{t-x}$
PLAT	$\eta_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x) \kappa_t^{(2)} + \gamma_{t-x}$

4. GAPC stochastic mortality models with StMoMo

The **StMoMo** package implements GAPC stochastic mortality models via the **StMoMo** function. This function takes as input information on the *link function* (and the associated distributional assumption), the *predictor structure*, and the *set of parameter constraints* to create an object of type "StMoMo" representing a GAPC mortality model:

- The argument **link** defines the *link function* and error distribution associated with the mortality model. Setting **link** = "log" assumes that deaths follow a Poisson distribution and uses a log link targeting the force of mortality μ_{xt} , while setting **link** = "logit" assumes that deaths follow a Binomial distribution and uses a logit link targeting one-year mortality rates q_{xt} .
- The *predictor* of the model is defined via arguments **staticAgeFun**, **periodAgeFun** and **cohortAgeFun**. Argument **staticAgeFun** is a logical value indicating whether the model has a static age function α_x or not. Argument **periodAgeFun** is a list of length N containing the definitions of the period age-modulating parameters $\beta_x^{(i)}$, $i = 1, \dots, N$, with each entry being "NP" for non-parametric age terms, "1" for $\beta_x^{(i)} = 1$, or a predefined parametric function of age. Argument **cohortAgeFun** defines the cohort age modulating parameter $\beta_x^{(0)}$ and can take values "NP" for non-parametric age terms, "1" for $\beta_x^{(0)} = 1$, a predefined parametric function of age, or NULL if the model does not have a cohort effect.
- The *set of parameter constraints* are defined via argument **constFun** which is a user-defined function mapping an arbitrary set of parameter estimates to a set of parameter estimates satisfying the model constraints.

In order to illustrate the creation of particular GAPC mortality models and other facilities of **StMoMo**, in the rest of this paper we will focus on the models summarised in Table 1. From now onwards, LC stands for the Lee-Carter model; CBD for the original Cairns-Blake-Dowd model; APC for the Age-Period-Cohort model of Currie (2006); RH for the cohort extension of the Lee-Carter model defined in equation (13) and proposed by Renshaw and

Haberman (2006); M7 for the quadratic CBD model defined in equation (21); and PLAT for the reduced Plat model defined previously in equation (33). For the sake of comparability, in all cases we will assume a Binomial distribution of deaths and use the logit function to link q_{xt} to the predictor structure n_{xt} .

Below, we show how to define with package **StMoMo** each of the models in Table 1.

Lee-Carter model. The LC model under a Binomial setting can be defined using the following code:

```
library(StMoMo)
constLC <- function(ax, bx, kt, b0x, gc, wxt, ages){
  c1 <- mean(kt, na.rm = TRUE)
  c2 <- sum(bx, na.rm = TRUE)
  list(ax = ax + c1 * bx, bx = bx / c2, kt = c2 * (kt - c1))
}
LC <- StMoMo(link = "logit", staticAgeFun = TRUE, periodAgeFun = "NP",
             constFun = constLC)
```

Recalling Section 3.1, we note that the constraint function `constLC` is the R implementation of transformation (6) with constants c_1 and c_2 calculated using Equation (8) to impose the constraints defined in equation (7). The **StMoMo** package also contains the function `lc` to facilitate the definition of Lee-Carter models. Hence, we could alternatively define the LC model using the much simpler predefined command:

```
LC <- lc(link = "logit")
```

CBD model. To define the CBD model we use the following commands:

```
f2 <- function(x, ages) x - mean(ages)
CBD <- StMoMo(link = "logit", staticAgeFun = FALSE,
              periodAgeFun = c("1", f2))
```

Here, we note that function `f2` defines the second age-modulating parameter $\beta_x^{(2)} = x - \bar{x}$ and that a `constFUN` argument needs not be provided since the CBD model does not have identifiability issues. Alternatively, we can define the CBD model using the predefined function `cbd`:

```
CBD <- cbd()
```

APC model, RH model and model M7. These three models could be defined by implementing explicitly the discussions in Sections 3.2, 3.3 and 3.5. However, package **StMoMo**

includes predefined functions `apc`, `rh`, `m7` that facilitate the definition of the APC model, the RH model and model M7, respectively³. Thus, these models are defined with the code:

```
RH <- rh(link = "logit", cohortAgeFun = "1")
APC <- apc(link = "logit")
M7 <- m7()
```

PLAT model. Package **StMoMo** does not include a predefined function for the Plat model. Nevertheless, recalling Section 3.6, we can define the reduced Plat model using the code:

```
f2 <- function(x, ages) mean(ages) - x
constPlat <- function(ax, bx, kt, b0x, gc, wxt, ages){
  nYears <- dim(wxt)[2]
  x <- ages
  t <- 1:nYears
  c <- (1 - tail(ages, 1)):(nYears - ages[1])
  xbar <- mean(x)
  #\sum g(c)=0, \sum cg(c)=0, \sum c^2g(c)=0
  indC <- !is.na(gc)
  phiReg <- lm(gc ~ 1 + c + I(c^2), data.frame(gc = gc[indC], c = c[indC]))
  phi <- coef(phiReg)
  gc[indC] <- residuals(phiReg)
  kt[2, ] <- kt[2, ] + 2 * phi[3] * t
  kt[1, ] <- kt[1, ] + phi[2] * t + phi[3] * (t^2 - 2 * xbar * t)
  ax <- ax + phi[1] - phi[2] * x + phi[3] * x^2
  #\sum kt[i, ] = 0
  ci <- rowMeans(kt, na.rm = TRUE)
  ax <- ax + ci[1] + ci[2] * (xbar - x)
  kt[1, ] <- kt[1, ] - ci[1]
  kt[2, ] <- kt[2, ] - ci[2]
  list(ax = ax, bx = bx, kt = kt, b0x = b0x, gc = gc)
}
PLAT <- StMoMo(link = "logit", staticAgeFun = TRUE,
               periodAgeFun = c("1", f2), cohortAgeFun = "1",
               constFun = constPlat)
```

We note that the constraint function `constPlat` is the R implementation of transformations (28) and (29) omitting the terms involving $\kappa^{(3)}$ and c_3 , and with constants ϕ_1, ϕ_2, ϕ_3 obtained via the linear regression defined in (31) and constant c_1 and c_2 as in Equation (32). Function `constPlat` imposes the constraints in Equation (30).

³The **StMoMo** package also includes functions `m6` and `m8` for defining models M6 and M8. We also note that the Renshaw and Haberman (2006) cohort extension of the Lee-Carter in Equation (9) can be defined using function `rh` with argument `cohortAgeFun = "NP"`.

5. Model fitting

Parameter estimates of GAPC stochastic mortality models can be obtained by maximising the model log-likelihood, which is given by

$$\mathcal{L}(d_{xt}, \hat{d}_{xt}) = \sum_x \sum_t \omega_{xt} \left\{ d_{xt} \log \hat{d}_{xt} - \hat{d}_{xt} - \log d_{xt}! \right\} \quad (34)$$

in the case of a Poisson distribution of deaths, and by

$$\mathcal{L}(d_{xt}, \hat{d}_{xt}) = \sum_x \sum_t \omega_{xt} \left\{ d_{xt} \log \left(\frac{\hat{d}_{xt}}{E_{xt}^i} \right) + (E_{xt}^i - d_{xt}) \log \left(\frac{E_{xt}^i - \hat{d}_{xt}}{E_{xt}^i} \right) + \binom{E_{xt}^i}{d_{xt}} \right\} \quad (35)$$

in the case of a Binomial distribution of deaths. In both cases, ω_{xt} are 0-1 weights indicating empty or omitted data cells and

$$\hat{d}_{xt} = E_{xt} g^{-1} \left(\alpha_x + \sum_{i=1}^N \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x} \right)$$

is the expected number of deaths predicted by the model, with $g^{-1}(\cdot)$ denoting the inverse of the *link function* $g(\cdot)$.

In the mortality literature, maximisation of the log-likelihood is typically performed using the Newton-Raphson iterative procedure tailored for each model (see e.g. Brouhns et al. (2002), Renshaw and Haberman (2006) and Cairns et al. (2009)). This is in fact the approach implemented in packages **ilc** and **Lifemetrics**. Nonetheless, as discussed extensively by Currie (2014), many stochastic mortality models are examples of generalised linear models or generalised non-linear-model, which facilitates their fitting using standard statistical software⁴. Currie (2014) exemplifies this fact by fitting several stochastic mortality models in R using the standard function **glm** or the function **gnm** of package **gnm** (Turner and Firth, 2012).

StMoMo provides the generic function **fit** for estimating the parameters of GAPC mortality models. In line with the remarks of Currie (2014), the corresponding **S3** method for objects of class "**StMoMo**" relies on function **gnm** of package **gnm** to estimate the parameters of a GAPC model⁵. Internally, this is accomplished by constructing the equivalent **gnm** formulation of the GAPC mortality model⁶. For instance, the **gnm** formula of the Binomial LC model created before is

⁴Haberman and Renshaw (2011) have also noticed this fact and profit from GLM facilities in standard statistical packages when fitting CBD type models.

⁵The current version of **StMoMo** only includes the **S3** method **fit.StMoMo** for objects of class "**StMoMo**", but future versions may provide tailored **fit** methods for particular stochastic mortality models.

⁶We note that when all the $\beta_x^{(i)}$ are parametric functions of age, the model is a GLM and therefore **gnm** by default resorts to the **glm** function of R when fitting the parameter of the model.

```
LC$gnmFormula
```

```
## [1] "D/E ~ -1 + offset(o) + factor(x) + Mult(factor(x), factor(t),  
inst = 1)"
```

while the **gnm** formula of the Binomial CBD model defined before is

```
CBD$gnmFormula
```

```
## [1] "D/E ~ -1 + offset(o) + factor(t) + B2:factor(t)"
```

We now illustrate the usage of the function `fit` of package **StMoMo** by fitting the six models defined before to England and Wales mortality data. The object `EWMaleData` included in package **StMoMo** contains data on deaths and central exposures for England and Wales males for the period 1961-2011 and for ages 0-100 obtained from the Human Mortality Database (2014). In our examples we concentrate however on ages 55 to 89 as the CBD model and the M7 model have been particularly designed to fit high ages. Additionally, since some models include cohort effects and in agreement with the usual practice (see e.g. Cairns et al. (2009) and Haberman and Renshaw (2011)), we exclude (by setting $\omega_{xt} = 0$) all cohorts that have fewer than three observations.

Models LC, APC, CBD, M7 and PLAT can be fitted to England and Wales male mortality data for ages 55 to 89 using the code:

```
Dxt <- EWMaleData$Dxt  
Ext <- EWMaleData$Ext + 0.5 * EWMaleData$Dxt  
ages <- EWMaleData$ages  
years <- EWMaleData$years  
ages.fit <- 55:89  
wxt <- genWeightMat(ages = ages.fit, years = years, clip = 3)  
LCfit <- fit(LC, Dxt = Dxt, Ext = Ext, ages = ages, years = years,  
            ages.fit = ages.fit, wxt = wxt)  
APCfit <- fit(APC, Dxt = Dxt, Ext = Ext, ages = ages, years = years,  
            ages.fit = ages.fit, wxt = wxt)  
CBDfit <- fit(CBD, Dxt = Dxt, Ext = Ext, ages = ages, years = years,  
            ages.fit = ages.fit, wxt = wxt)  
M7fit <- fit(M7, Dxt = Dxt, Ext = Ext, ages = ages, years = years,  
            ages.fit = ages.fit, wxt = wxt)  
PLATfit <- fit(PLAT, Dxt = Dxt, Ext = Ext, ages = ages, years = years,  
            ages.fit = ages.fit, wxt = wxt)
```

From this code we note the following:

- In order to match the logit-Binomial setting used before in the definition of the mortality models, initial exposures are approximated by transforming the available central exposures.

- The first and last three cohorts years are excluded from the fitting via argument `wxt`. The appropriate 0-1 weighting matrix, `wxt`, is constructed using the utility function `genWeightMat` of package **StMoMo**.

The fitting of the RH model requires some care as it is well known that fitting cohort extensions of the Lee-Carter models is problematic (Hunt and Villegas, 2014). In particular, Currie (2014) has encountered convergence issues when using package **gnm** to fit the RH model. As a possible way to circumvent these issues, Currie (2014) suggests the use of appropriate starting values when fitting model RH. This can be achieved in function `fit` via input arguments `start.ax`, `start.bx`, `start.kt`, `start.b0x`, and `start.gc`. Using the parameters of the Lee-Carter model as starting values, model RH can be fitted with the code:

```
RHfit <- fit(RH, Dxt = Dxt, Ext = Ext, ages = ages, years = years,
            ages.fit = ages.fit, wxt = wxt, start.ax = LCfit$ax,
            start.bx = LCfit$bx, start.kt = LCfit$kt)
```

The output from the function `fit` is an object of class "fitStMoMo" including, among others things, the following information:

- `model`: the "StMoMo" object defining the underlying GAPC stochastic mortality model;
- `ax`, `bx`, `kt`, `b0x`, `gc`: the estimated parameters;
- `loglik`: the log-likelihood of the model;
- `deviance`: the model deviance;
- `nobs`: the number of observations in the data;
- `npar`: the effective number of parameters of the model;
- `fittingModel`: the output of the `gnm` call used to fit the model.

There are `print`, `plot`, `summary`, `fitted`, and `residuals` methods for the "fitStMoMo" class. For instance, Figures 1, 2 and 3 depicting the fitted parameters of the LC model, the CBD model and the APC model, respectively, were produced with the code:

```
plot(LCfit)
plot(CBDfit, parametricbx = FALSE)
plot(APCfit, parametricbx = FALSE)
```

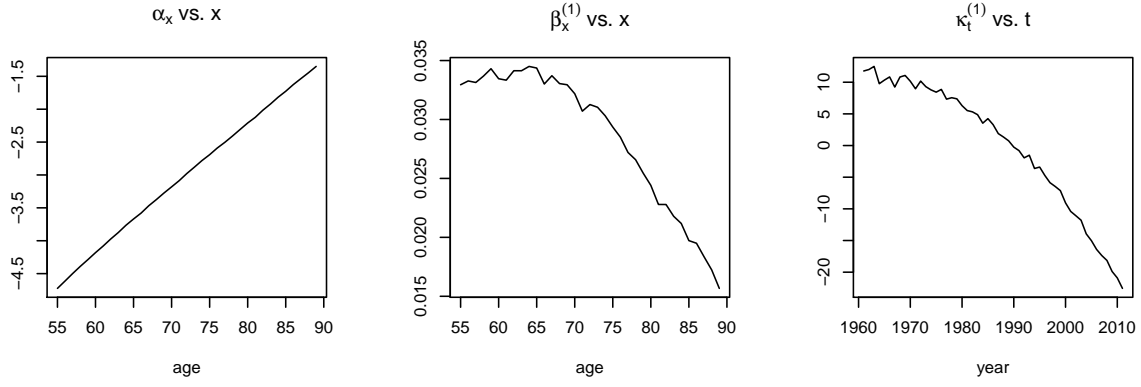


Figure 1: Parameters for the Lee-Carter (LC) model fitted to the England and Wales male population for ages 55-89 and the period 1961-2011.

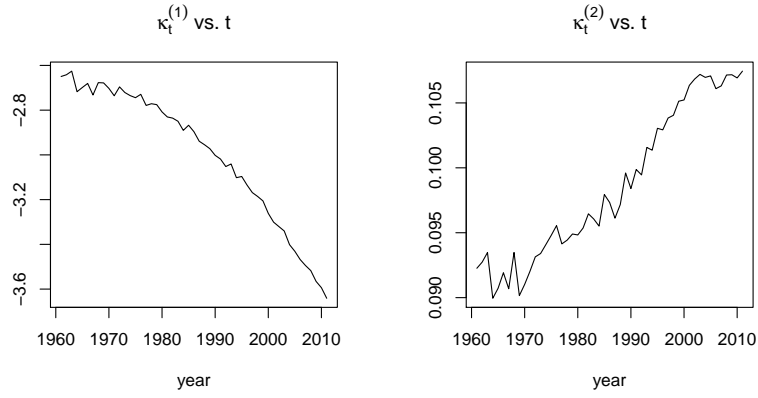


Figure 2: Parameters for the CBD model fitted to the England and Wales male population for ages 55-89 and the period 1961-2011.

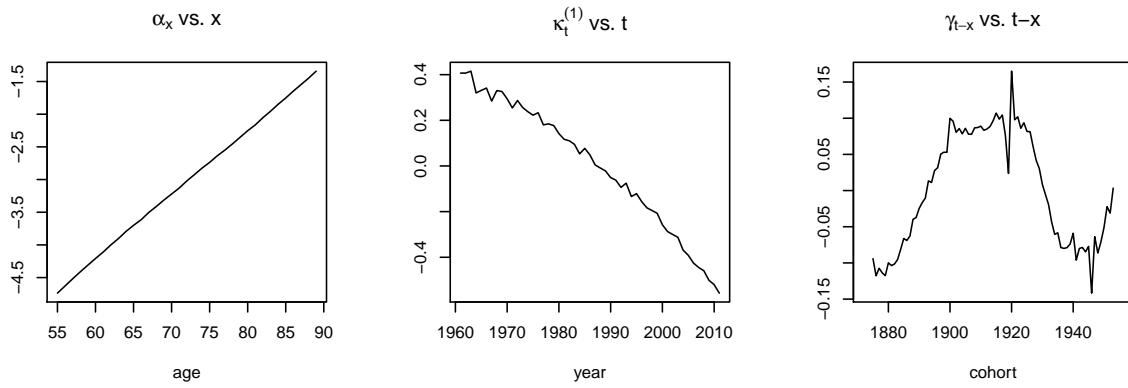


Figure 3: Parameters for the APC model fitted to the England and Wales male population for ages 55-89 and the period 1961-2011.

6. Goodness-of-fit analysis

A typical way of analysing the goodness-of-fit of mortality models is by inspecting the residuals of the fitted model. Regular patterns in the residuals indicate the inability of the model to describe all the features of the data appropriately. With a Poisson or Binomial random component, it is appropriate to look at the scaled deviance residuals defined as:

$$r_{xt} = \text{sign}(d_{xt} - \hat{d}_{xt}) \sqrt{\frac{\text{dev}(x, t)}{\hat{\phi}}}, \quad \hat{\phi} = \frac{D(d_{xt}, \hat{d}_{xt})}{K - \nu}, \quad (36)$$

where

$$\text{dev}(x, t) = 2d_{xt} \log \left(\frac{d_{xt}}{\hat{d}_{xt}} \right) - (d_{xt} - \hat{d}_{xt})$$

for a Poisson random component, or

$$\text{dev}(x, t) = 2d_{xt} \log \left(\frac{d_{xt}}{\hat{d}_{xt}} \right) + (E_{xt}^i - d_{xt}) \log \left(\frac{E_{xt}^i - d_{xt}}{E_{xt}^i - \hat{d}_{xt}} \right)$$

for a Binomial random component;

$$D(d_{xt}, \hat{d}_{xt}) = \sum_x \sum_t \omega_{xt} \text{dev}(x, t)$$

is the total deviance of the model; $K = \sum_x \sum_t \omega_{xt}$ is the number of observations in the data; and ν is the effective number of parameters in the model.

In **StMoMo** standardised deviance residuals can be obtained with the generic function `residuals` applied to a fitted stochastic mortality model of class `"fitStMomo"`. For example, to obtain the residuals of the LC model and of the CBD model fitted before we use the commands:

```
LCres <- residuals(LCfit)
CBDres <- residuals(CBDfit)
```

Graphs of these residuals can be produced using the generic function `plot`. This function supports, via argument `type`, three types of plots:

- scatter plots of residuals by age, period and cohort of such as those extensively used in Haberman and Renshaw (2011);
- black and white sign-plots of the residuals such as those used in Cairns et al. (2009) and Lovász (2011); and
- colour maps (heat-maps) of the residuals.

Figure 4 presents heat-maps of the deviance residuals for the six models fitted to the England and Wales male mortality experience. These charts were produced using function `plot` with option `type = "colourmap"`. For instance, Figure 4b was obtained with the code:

```
plot(CBDres, type = "colourmap", reslim = c(-3.5, 3.5))
```

From Figure 4 we see that models LC, CBD and APC display strong residual patterns while the residuals of models RH, M7 and PLAT look reasonably random. The APC model shows a strong clustering of residuals due to its inability to allow for varying improvement rates with age. The LC and CBD models, which do not incorporate a cohort effect, show very marked diagonal patterns indicating the inability of these models to capture the well-known cohort effect observed in the England and Wales population (Willets, 2004). The issues with the fit of the LC and CBD models become more evident when looking at scatter plots of the residuals by age, period and cohort. Such plots for the LC model (Figure 5a) and the CBD model (Figure 5b) can be produced using argument `type = "scatter"` of function `plot` via the commands:

```
plot(LCres, type = "scatter", reslim = c(-3.5, 3.5))
plot(CBDres, type = "scatter", reslim = c(-3.5, 3.5))
```

The right panels in Figure 5 clearly show that the LC and CBD models are unable to capture the cohort effect. In addition, the centre panel in Figure 5b reveals some strong patterns by age, reflecting the the lack of a quadratic age term in the CBD which may be necessary to capture the commonly observed curvature of the mortality rates in a logit scale.

When evaluating the goodness-of-fit of different models, it is generally anticipated that models with more parameters provide a better fit to the data. To rule out the possibility that the better fit observed in a model is the result of over-parametrisation and compare the relative performance of several models, it has become common in the mortality literature to use information criteria which modify the maximum likelihood criterion by penalising models with more parameters⁷. Two of these criteria are the Akaike Information Criteria (AIC) and the Bayesian Information Criteria (BIC), defined as $AIC = 2\nu - 2\mathcal{L}$ and $BIC = \nu \log K - 2\mathcal{L}$, respectively, with a lower value of AIC and BIC being preferable. In **R** these information criteria can be computed using the generic functions `AIC` and `BIC`. For example, we can get the AIC and BIC of the CBD model as follows:

```
AIC(CBDfit)

## [1] 34697.82

BIC(CBDfit)

## [1] 35256.83
```

Table 2 presents AIC and BIC values for the six models fitted to the England and Wales male data. We note that both criteria lead to the same ranking of models with M7, PLAT, and

⁷For examples of such analysis see Cairns et al. (2009, Section 6.1.1), Haberman and Renshaw (2011, Section 3.3), and Lovász (2011, Section 4.1).

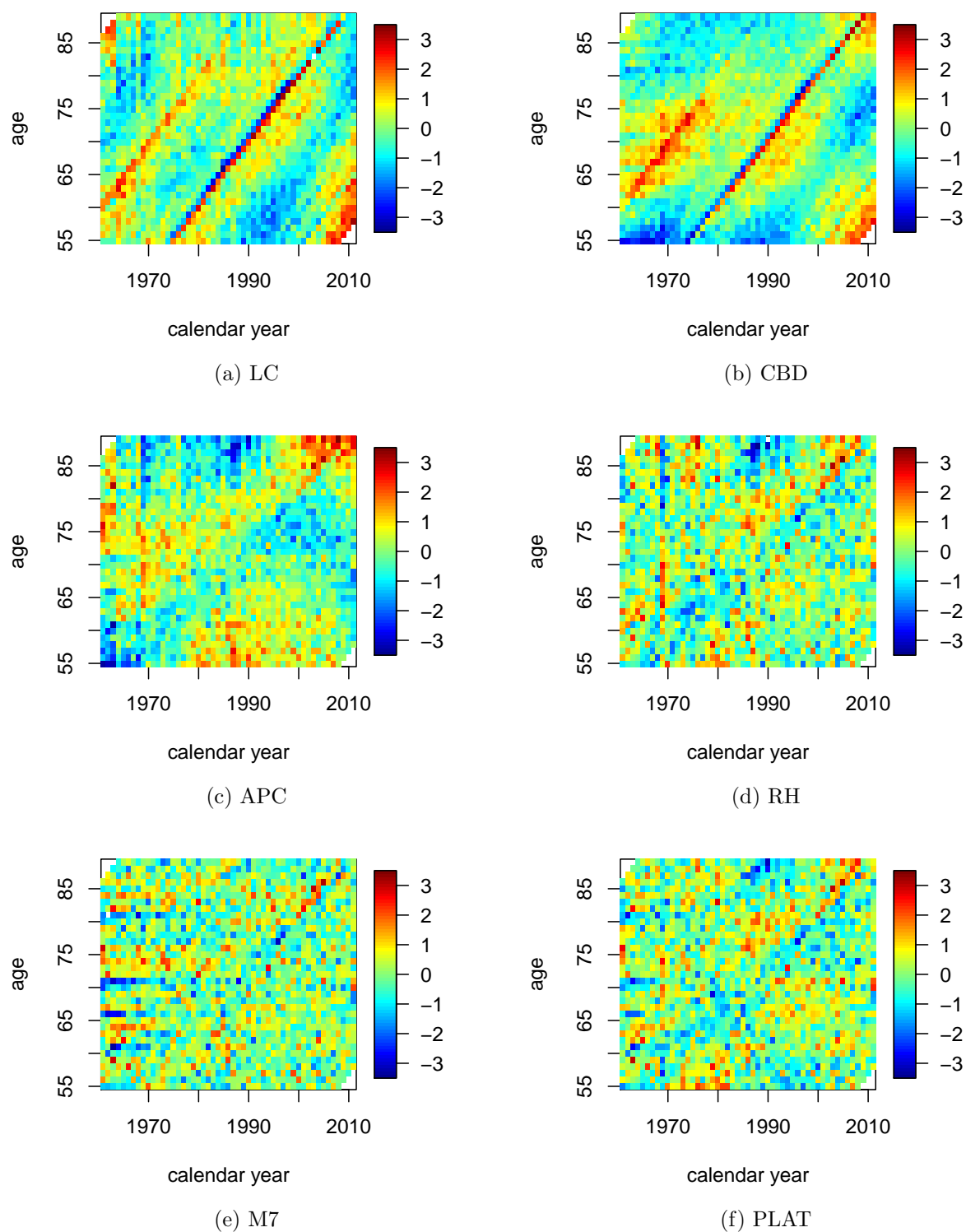


Figure 4: Heat-maps of deviance residuals for different model fitted to the England and Wales males population for ages 55-89 and the period 1961-2011.

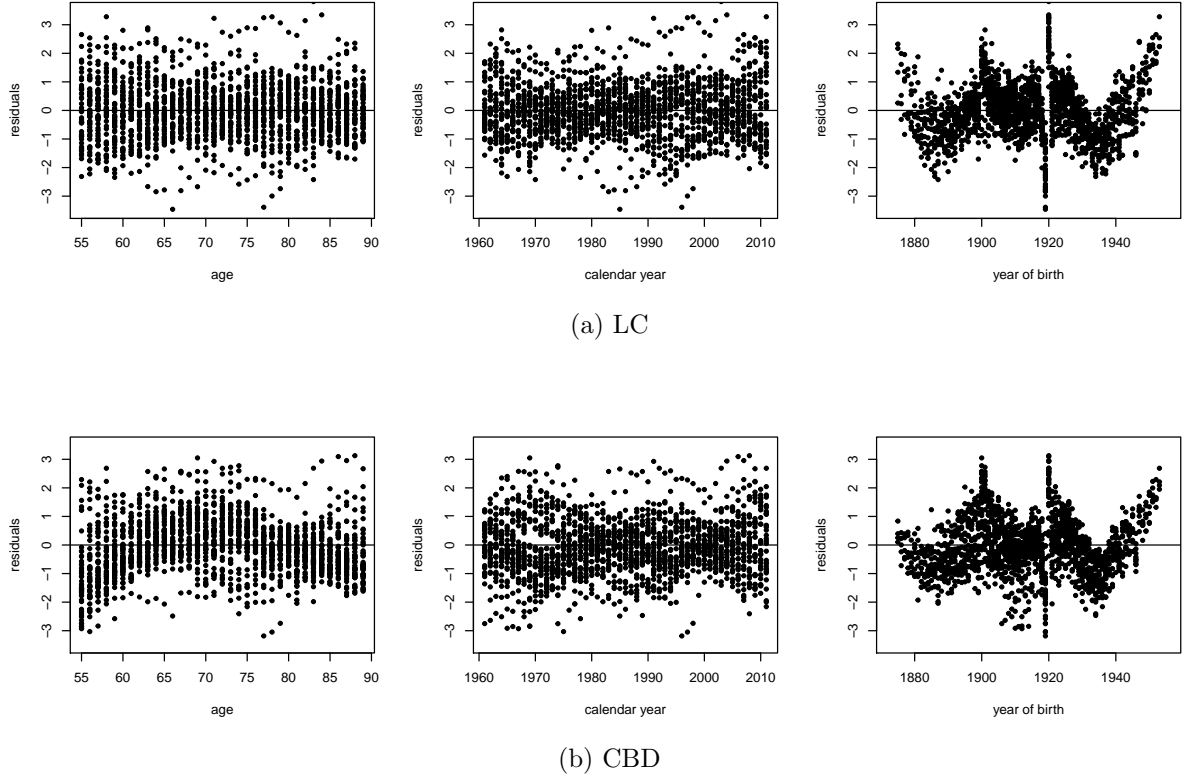


Figure 5: Scatter plots of deviance residuals for models LC and CBD fitted to the England and Wales males population for ages 55-89 and the period 1961-2011.

	LC	CBD	APC	RH	M7	PLAT
Number of Parameters	119	102	162	197	229	211
AIC	29866	34698	24469	21779	21406	21624
BIC	30518	35257	25357	22859	22661	22780

Table 2: Number of parameters, AIC and BIC values for different model fitted to the England and Wales males population for ages 55-89 and the period 1961-2011

RH being the best performing models. Overall, these results are consistent with the existing literature comparing single population models, where the Renshaw-Haberman extension of the Lee-Carter model and the M7 model have been identified as good candidates for modelling mortality in the England and Wales population (Cairns et al., 2009; Haberman and Renshaw, 2011).

7. Forecasting and simulation with stochastic mortality models

In the family of GAPC stochastic mortality models the dynamics of mortality are driven by the period indexes $\kappa_t^{(i)}$, $i = 1, \dots, N$, and the cohort index γ_{t-x} . Therefore, the forecasting and simulation of mortality rates requires the modelling of these indexes using time series techniques.

For the period indexes we follow the standard approach (Cairns et al., 2006, 2011; Haberman and Renshaw, 2011; Lovász, 2011) and use a multivariate random walk with drift. Specifically, we assume that

$$\boldsymbol{\kappa}_t = \boldsymbol{\theta} + \boldsymbol{\kappa}_{t-1} + \boldsymbol{\xi}_t, \quad \boldsymbol{\kappa}_t = \begin{bmatrix} \kappa_t^{(1)} \\ \vdots \\ \kappa_t^{(N)} \end{bmatrix}, \quad \boldsymbol{\xi}_t \sim N(\mathbf{0}, \Sigma), \quad (37)$$

where $\boldsymbol{\theta}$ is an N -dimension vector of drift parameters and Σ is the $N \times N$ variance-covariance matrix of the multivariate white noise $\boldsymbol{\xi}_t$.

As pointed out by Currie (2014), the main challenge when forecasting stochastic mortality models is specifying the dynamics of the cohort effect. To have a simple starting point, we follow previous studies (Renshaw and Haberman, 2006; Cairns et al., 2011; Lovász, 2011) and assume that the cohort index, γ_{t-x} , follows an univariate ARIMA process which is independent to the dynamics of the period index, $\boldsymbol{\kappa}_t$. In general, we assume that $\gamma_c \equiv \gamma_{t-x}$ follows an ARIMA(p, q, d) with drift, so that

$$\Delta^d \gamma_c = \theta_0 + \phi_1 \Delta^d \gamma_{c-1} + \cdots + \phi_p \Delta^d \gamma_{c-p} + \xi_c + \theta_1 \xi_{c-1} + \cdots + \theta_q \xi_{c-q}, \quad (38)$$

where Δ is the difference operator, $\Delta \gamma_c = \gamma_c - \gamma_{c-1}$, $\Delta^2 \gamma_c = \gamma_c - 2\gamma_{c-1} + \gamma_{c-2}, \dots$; θ_0 is the drift parameter; ϕ_1, \dots, ϕ_p are the autoregressive coefficients with $\phi_p \neq 0$; $\theta_1, \dots, \theta_q$ are the moving average coefficients with $\theta_q \neq 0$; and ξ_c is a Gaussian white noise process with variance σ_ξ .

The time series models in (37) and (38) can be used to obtain projected (simulated) values of the period index $\dot{\boldsymbol{\kappa}}_{t_n+s} := \left(\dot{\kappa}_{t_n+s}^{(1)}, \dots, \dot{\kappa}_{t_n+s}^{(N)} \right)'$ and cohort index $\dot{\gamma}_{t_n+s-x_1}$, $s = 1, \dots, h$, respectively, to derive forecasted (simulated) values of the predictor

$$\dot{\eta}_{x,t_n+s} = \alpha_x + \sum_{i=1}^N \beta_x^{(i)} \dot{\kappa}_{t_n+s}^{(i)} + \beta_x^{(0)} \dot{\gamma}_{t_n+s-x}, \quad (39)$$

which can in turn be used to obtain forecasted (simulated) age-specific central mortality rates, $\dot{\mu}_{x,t_n+s}$ or age-specific one-year mortality rates, \dot{q}_{x,t_n+s} .

In package **StMoMo** the forecasting of GAPC stochastic mortality models is implemented via the generic method `forecast`. This function estimates and forecasts the multivariate random walk with drift in Equation (37) using the approach described in Haberman and Renshaw (2011, Appendix B)⁸ and uses function `Arima` of package **forecast** (Hyndman and Khandakar, 2008) to estimate and forecast the ARIMA process of Equation (38). For instance, if we assume that the cohort indexes of the APC, RH, M7, and PLAT model follow the ARIMA processes specified in Table 3, central projections of the period indexes, cohort

⁸This is implemented in function `mrwd` of package **StMoMo**.

Table 3: ARIMA Models for the cohort effect for models APC, RH, M7 and PLAT.

Mortality Model	Model for γ_{t-x}
APC	ARIMA(1, 1, 0) with drift
RH	ARIMA(1, 1, 0) with drift
M7	ARIMA(2, 0, 0) with non-zero intercept
PLAT	ARIMA(2, 0, 0) with non-zero intercept

index, and one-year death probabilities for the England and Wales mortality experience can be obtain with the code:

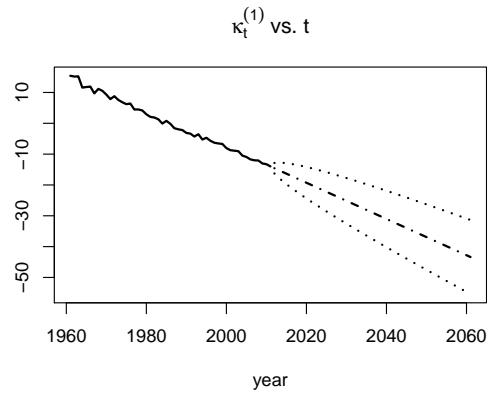
```
LCfor <- forecast(LCfit)
CBDfor <- forecast(CBDfit)
APCfor <- forecast(APCfit, gc.order = c(1, 1, 0))
RHfor <- forecast(RHfit, gc.order = c(1, 1, 0))
M7for <- forecast(M7fit, gc.order = c(2, 0, 0))
PLATfor <- forecast(PLATfit, gc.order = c(2, 0, 0))
```

The output from the function `forecast` is an object of class "forStMoMo" including, among others things, the following information:

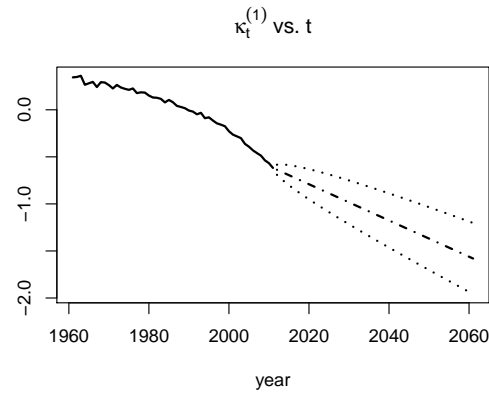
- **rates**: a matrix with the central projection of the mortality rates, $\dot{\mu}_{x,t_n+s}$ or \dot{q}_{x,t_n+s} , $s = 1, \dots, h$;
- **kt.f**: a list containing information on the multivariate random walk with drift fitted to the period index κ_t ; and
- **gc.f**: a list containing information on the ARIMA(p, q, d) model fitted to the cohort index γ_{t-x} .

There are `print`, `plot`, and `summary` methods for the "forStMoMo" class. For instance, plots of the forecast of the period indexes of models RH, M7 and PLAT (see Figure 6) can be produced using the code:

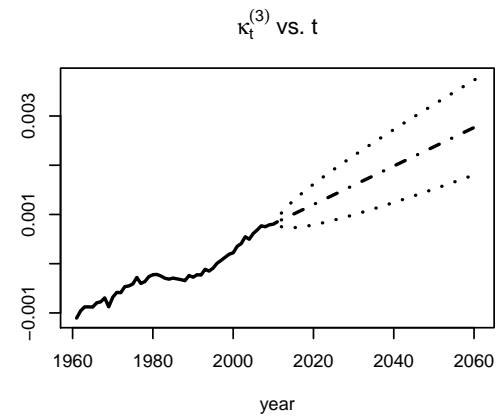
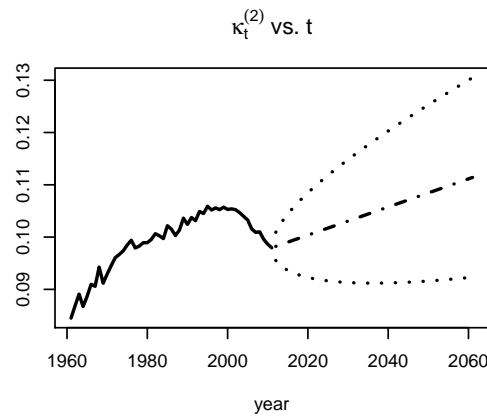
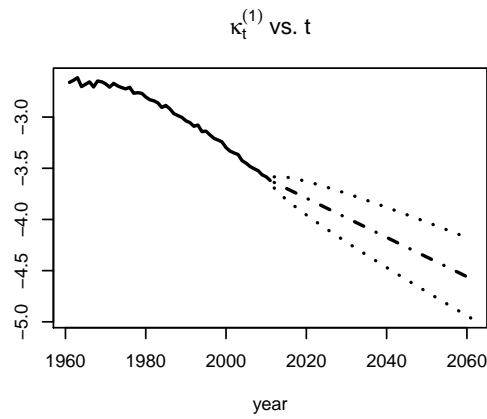
```
plot(RHfor, only.kt = TRUE)
plot(M7for, only.kt = TRUE, nCol = 3)
plot(PLATfor, only.kt = TRUE)
```



(a) RH



(b) PLAT



(c) M7

Figure 6: Forecast of the period indexes of the RH, M7 and PLAT models applied to the England and Wales males population for ages 55-89 and the period 1961-2011. Dashed lines represent central forecast and dotted lines represent 95% prediction intervals.

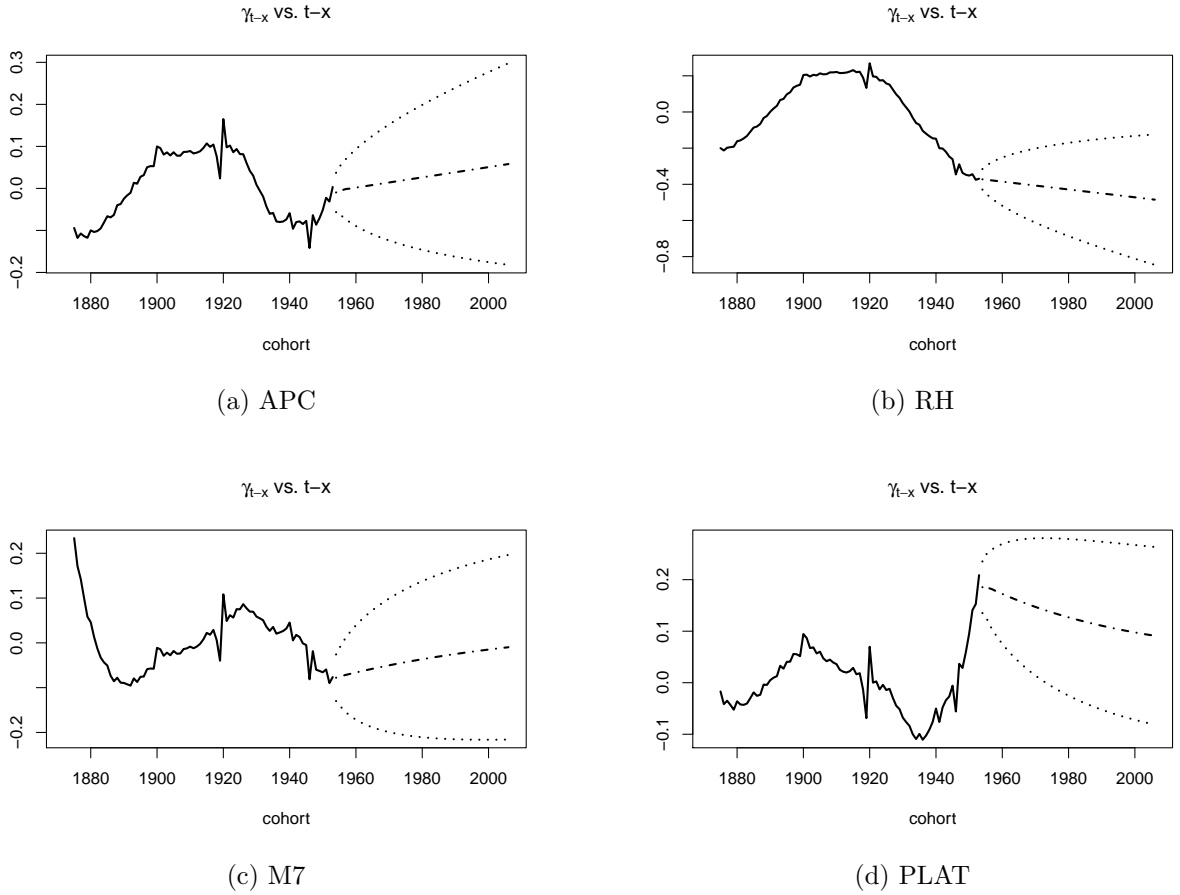


Figure 7: Forecast of the cohort indexes of the APC, RH, M7 and PLAT models applied to the England and Wales males population for ages 55-89 and the period 1961-2011. Dashed lines represent central forecast and dotted lines represent 95% prediction intervals.

Similarly, plots of the forecast of the cohort indexes of the APC, RH, M7 and PLAT models (see Figure 7) can be obtained with the commands:

```
plot(APCfor, only.gc = TRUE)
plot(RHfor, only.gc = TRUE)
plot(M7for, only.gc = TRUE)
plot(PLATfor, only.gc = TRUE)
```

Package **StMoMo** also provides the function **simulate** for simulating trajectories from GAPC stochastic mortality models. To simulate the period index, κ_t , **StMoMo** implements a multivariate adaptation of Algorithm 2 in Haberman and Renshaw (2009) without provision for parameter error⁹, while to simulate the cohort index, γ_{t-x} , function **simulate** uses the equivalent S3 method for objects of class "Arima" provided by package **forecast**. For

⁹We note that Algorithm 2 in Haberman and Renshaw (2009) is itself an adaptation of the prediction interval approach of Cairns et al. (2006).

example, the code below produces 5000 simulated trajectories of the six stochastic mortality models fitted previously to the England and Wales male mortality experience:

```
set.seed(1234)
nsim <- 5000
LCsim <- simulate(LCfit, nsim = nsim)
CBDsim <- simulate(CBDfit, nsim = nsim)
APCsim <- simulate(APCfit, nsim = nsim, gc.order= c(1, 1, 0))
RHsim <- simulate(RHfit, nsim = nsim, gc.order= c(1, 1, 0))
M7sim <- simulate(M7fit, nsim = nsim, gc.order= c(2, 0, 0))
PLATsim <- simulate(PLATfit, nsim = nsim, gc.order= c(2, 0, 0))
```

The output from the function `simulate` is an object of class `"simStMoMo"` including, among others things, the following information:

- **rates**: a three dimensional array with the future simulated mortality rates;
- **kt.s**: a list containing information on the simulated paths of the period index κ_t ; and
- **gc.s**: a list containing information on the simulated paths of the cohort index γ_{t-x} .

This output can be used to extract sample trajectories from a model. For instance, Figure 8, which depicts 20 trajectories of the period index, cohort index and one-year death rates at age 65 from model RH, was produced with the code:

```
#Plot period index
par(mfrow=c(1, 3))
tlim <- range(RHfit$years, RHsim$kt.s$years)
plot(RHfit$years, RHfit$kt[1, ], xlim = tlim,
     ylim = range(RHfit$kt, RHsim$kt.s$sim[1, , 1:20]), type = "l",
     xlab = "year", ylab = "kt", main = "Period index")
matlines(RHsim$kt.s$years, RHsim$kt.s$sim[1, , 1:20], type = "l", lty = 1)
#Plot cohort index
plot(RHfit$cohorts, RHfit$gc,
     xlim = range(RHfit$cohorts, RHsim$gc.s$cohorts),
     ylim = range(RHfit$gc, RHsim$gc.s$sim[, 1:20], na.rm = TRUE),
     type = "l", xlab = "year", ylab = "kt",
     main = "Cohort index (ARIMA(1,1,0) with drift)")
matlines(RHsim$gc.s$cohorts, RHsim$gc.s$sim[, 1:20], type = "l", lty = 1)
#Plot death probabilities at age 65
qxt <- Dxt / Ext
plot(RHfit$years, qxt["65", ], xlim = tlim,
     ylim = range(qxt["65", ], RHsim$rates["65", , 1:20]),
     type = "l", xlab = "year", ylab = "rate",
     main = "Mortality rates at age 65")
matlines(RHsim$years, RHsim$rates["65", , 1:20], type = "l", lty = 1)
```

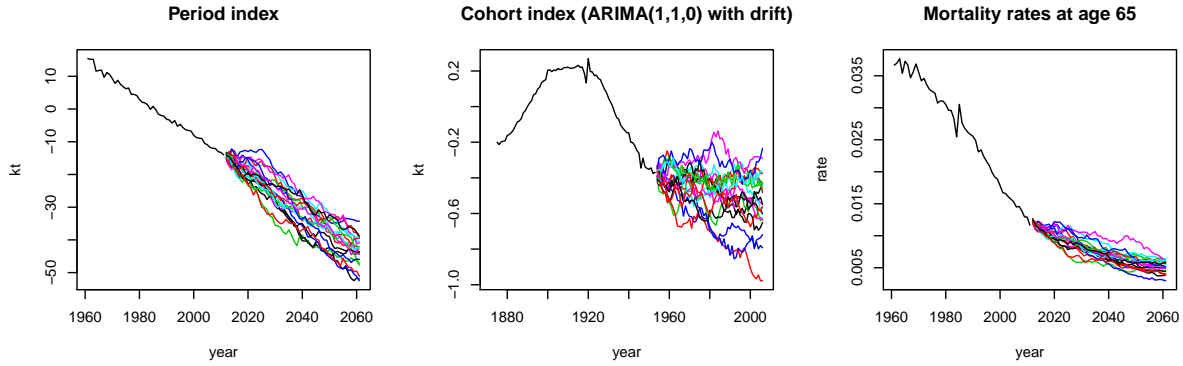


Figure 8: 20 simulated trajectories of the period index $\kappa_t^{(1)}$, cohort index γ_{t-x} and one-year death rates at age 65 q_{xt} for model RH fitted to the England and Wales males population for ages 55-89 and the period 1961-2011.

We can also use the output from function `simulate` to produce fan charts depicting the uncertainty associated with a model forecast. According to Cairns et al. (2011), such plots are central to the analysis of the plausibility of the forecast from a model, and can be used as a criterion when deciding upon what is the most appropriate model amongst a group of possible stochastic mortality models. Figure 9 shows fan charts depicting 50%, 80% and 95% prediction intervals for mortality rates at ages 65, 75 and 85 for each of the six models fitted to the England and Wales experience. Figure 9b, for instance, was produced with the code:

```
library(fanplot)
probs = c(2.5, 10, 25, 50, 75, 90, 97.5)
qxt <- Dxt / Ext
#age 65
plot(years, qxt["65", ], xlim = c(1960, 2061), ylim = c(0.0025, 0.2),
     ylab = "mortality rate (log scale)", pch = 20, log = "y")
fan(t(CBDsim$rates["65", , ]), start = 2012, probs = probs, n.fan = 4,
    fan.col = colorRampPalette(c("black", "white")), ln = NULL)
#age 75
points(years, qxt["75", ], pch = 20)
fan(t(CBDsim$rates["75", , ]), start = 2012, probs = probs, n.fan = 4,
    fan.col = colorRampPalette(c("red", "white")), ln = NULL)
#age 85
points(years, qxt["85", ], pch = 20)
fan(t(CBDsim$rates["85", , ]), start = 2012, probs = probs, n.fan = 4,
    fan.col = colorRampPalette(c("blue", "white")), ln = NULL)
#labels
text(1965, qxt[c("65", "75", "85"), "1990"],
     labels = c("x=65", "x=75", "x=85"))
```

From Figure 9 we note the following:

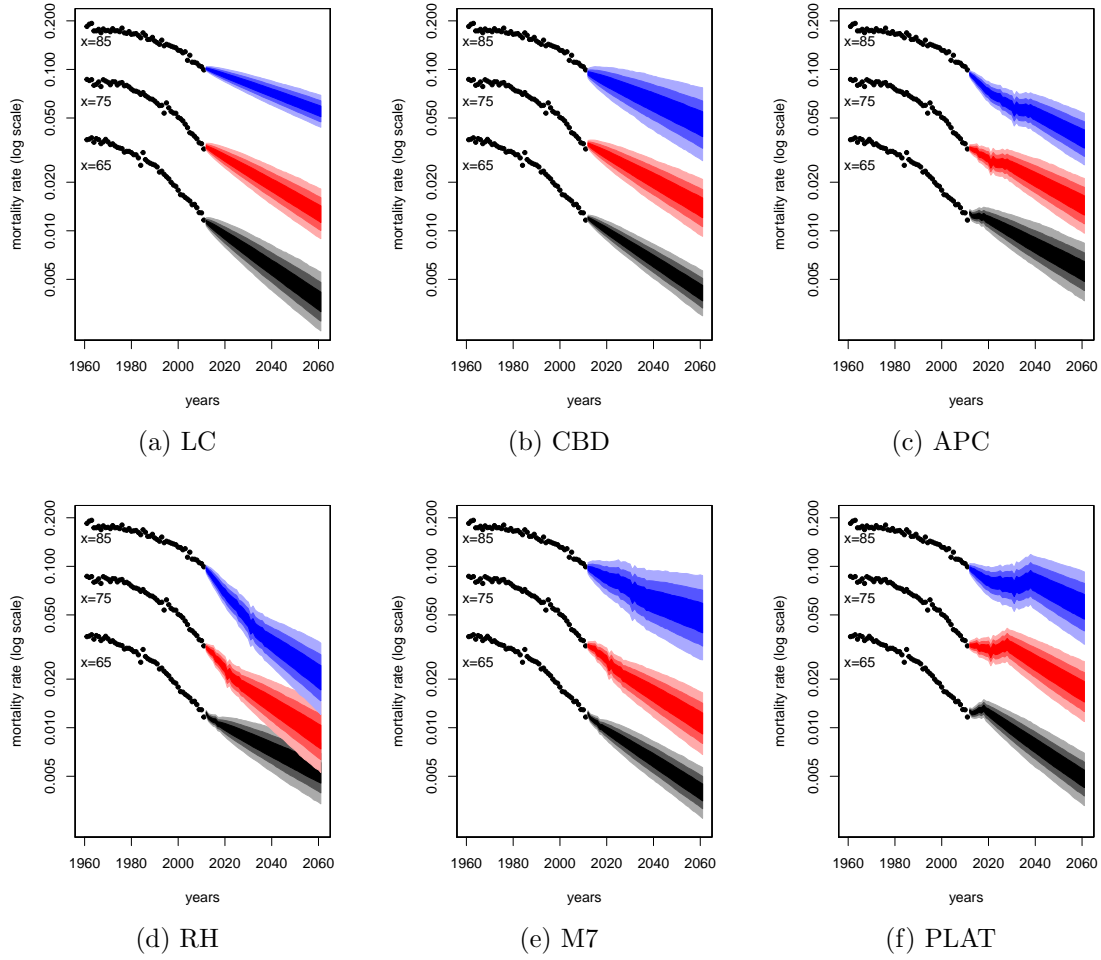


Figure 9: Fan charts for mortality rates q_{xt} at ages $x = 65$ (bottom fan), $x = 75$ (middle fan) and $x = 85$ (top fan) from the six models fitted to the England and Wales males population for ages 55-89 and the period 1961-2011. The dots show historical mortality rates for 1961-2011. Shades in the fan represent prediction intervals at the 50%, 80% and 95% level.

- Whilst for models CBD, RH, M7 and PLAT the fans at age 85 are wider than at age 65 in accordance with historical evidence (see Cairns et al. (2011, Appendix B)), for models LC and APC the fans at age 85 are narrower than at age 65. This suggests that forecasts from models LC and APC are not plausible for the dataset used in this paper.
- Forecasts for the PLAT model show an implausible increase of mortality rates. This is because the central trend is linked to the estimated cohort effect γ_{t-x} for the PLAT model (see Figure 7d) which shows a steep upward trend between 1935 and 1955.
- The central trend and uncertainty levels produced by each of the models have noticeable differences. This highlights the importance of recognising model risk as a significant issue when modelling mortality (Cairns et al., 2011).

8. Parameter uncertainty and bootstrapping

When analysing the uncertainty in mortality projections in an actuarial context it is important to consider all sources of risk. However, the prediction intervals (fan charts) obtained in the previous section only account for the uncertainty arising from the error in the forecast of the period and cohort indexes and ignore the uncertainty arising from the estimation of the parameters of the GAPC model.

Due to the analytical intractability of many stochastic mortality models, parameter uncertainty is typically accounted for using the bootstrap procedure. This procedure yields B samples $\alpha_{xt}^b, \beta_x^{(1),b}, \dots, \beta_x^{(N),b}, \kappa_t^{(1),b}, \dots, \kappa_t^{(N),b}, \beta_x^{(0),b}, \gamma_{t-x}^b$, $b = 1, \dots, B$, of the parameters of the GAPC model which can then be used to produce confidence and prediction intervals of demographic and actuarial quantities. To sample the parameters of GAPC models we consider the semiparametric bootstrap proposed by Brouhns et al. (2005) and the residual bootstrap first considered in Koissi et al. (2006):

- **Semiparametric bootstrap:** Brouhns et al. (2005) propose a semiparametric bootstrap where first B samples of the number of deaths d_{xt}^b , $b = 1, \dots, B$, are generated by sampling from the Poisson Distribution with mean d_{xt} . Each bootstrapped sample d_{xt}^b , $b = 1, \dots, B$, is then used to re-estimate the model to obtain B bootstrapped parameter estimates $\alpha_{xt}^b, \beta_x^{(1),b}, \dots, \beta_x^{(N),b}, \kappa_t^{(1),b}, \dots, \kappa_t^{(N),b}, \beta_x^{(0),b}, \gamma_{t-x}^b$, $b = 1, \dots, B$. Renshaw and Haberman (2008) use the fitted number of death \hat{d}_{xt} (instead of the observed deaths d_{xt}) to perform the sampling from the Poisson distribution. Wang and Lu (2005) consider a similar semiparametric approach using a Binomial distribution of deaths.
- **Residual bootstrap:** Another possibility is to bootstrap the residuals of the model as suggested by Koissi et al. (2006). Under this approach the deviance residuals r_{xt} are resampled with replacement to generate B replications r_{xt}^b , $b = 1, \dots, B$, which are mapped to the corresponding resampled death counts d_{xt}^b , $b = 1, \dots, B$, using the inverse formula. The model is then re-fitted using these resample number of deaths to produce B set of estimated parameters $\alpha_{xt}^b, \beta_x^{(1),b}, \dots, \beta_x^{(N),b}, \kappa_t^{(1),b}, \dots, \kappa_t^{(N),b}, \beta_x^{(0),b}, \gamma_{t-x}^b$, $b = 1, \dots, B$. We refer to Koissi et al. (2006) and Renshaw and Haberman (2008) for details on the inverse formula under a Poisson distribution of deaths and to Debón et al. (2010) for the inverse formula under a Binomial framework. Finally, we note that in implementing the residual bootstrap we follow Renshaw and Haberman (2008) and apply the inverse formula to produce samples of the observed number of deaths rather than samples of the fitted number of deaths as originally done by Koissi et al. (2006).

In what follows we illustrate the assessment of parameter uncertainty using package **StMoMo**. In doing so we deviate from the England and Wales example we have used so far and use instead New Zealand mortality data. This new example follows closely the work of Li (2014) who uses New Zealand mortality data to compare several simulation strategies for assessing the risk in mortality projections with a Poisson Lee-Carter model. The main rationale for the change of dataset is that parameter uncertainty is particularly important when analysing the mortality of smaller populations such as smaller countries or pension

plans. In addition, this new example also serves as a means for illustrating the use of package **StMoMo** in other datasets.

Mortality data for New Zealand can be extracted from the Human Mortality Database (2014) using function `hmd.mx` of package **demography** with the code:

```
library(demography)
NZdata <- hmd.mx(country = "NZL_NP", username = username,
                 password = password)
```

We note that the `username` and `password` above are for the Human Mortality Database and should be replaced appropriately. Following Li (2014), we fit a Poisson Lee-Carter model to New Zealand male data for ages 0 to 89 and for the period 1985 to 2008. This can be carried out with the commands:

```
Ext <- NZdata$pop$male
Dxt <- NZdata$rate$male * Ext
LCfit <- fit(lc(), Dxt = Dxt, Ext = Ext, ages = NZdata$age,
            years = NZdata$year, ages.fit = 0:89, years.fit = 1985:2008)
```

In package **StMoMo** the bootstrap of GAPC stochastic mortality models is implemented with generic function `bootstrap`. This functions supports both the semiparametric bootstrap and the residual bootstrap. For instance, 5000 semiparametric bootstrap samples of the Lee-Carter model can be obtained with the code¹⁰:

```
LCboot <- bootstrap(LCfit, nBoot = 5000, type = "semiparametric")
```

The output from function `bootstrap` is an object of class `"bootStMoMo"` in which the component `bootParameters` contains the `nBoot` replications of the bootstrap parameters. A fan chart depicting the 50%, 80% and 95% confidence intervals of the bootstrapped Lee-Carter model (Figure 10) can be obtained with the command:

```
plot(LCboot)
```

From Figure 10 we note that whilst the parameter uncertainty in the static age function α_x and the period index $\kappa_t^{(1)}$ is modest, the uncertainty in the age-modulating parameters $\beta_x^{(1)}$ is more significant.

Once a stochastic mortality model has been bootstrapped we can simulate it forward to obtain simulated trajectories which account for both the forecast error in the period and cohort indexes and the error in the model fitting. In **StMoMo** we can accomplish this using function `simulate` applied to an object of class `"bootStMoMo"`. Thus, to obtain 5000 simulated trajectories of the Lee-Carter model which take into account parameter uncertainty we use the instruction:

¹⁰We note that the bootstrap is a computer intensive procedure and can take a couple of hours to run.

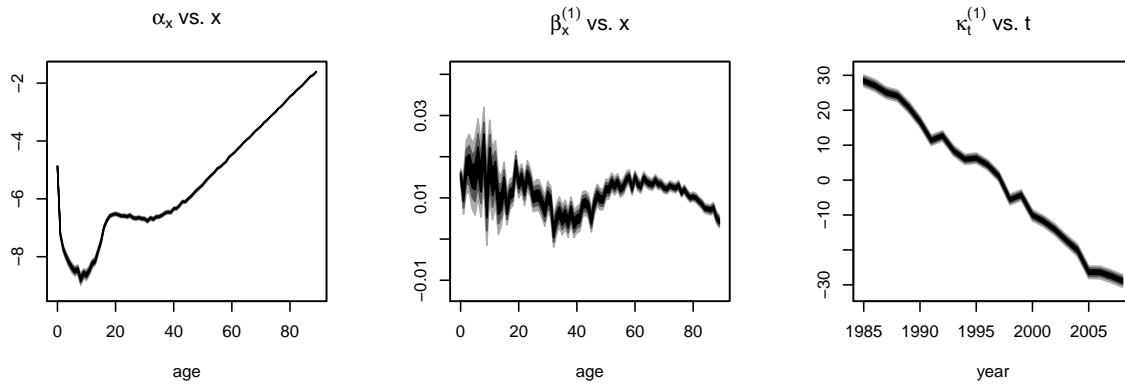


Figure 10: Bootstrapped parameters for the Poisson Lee-Carter model fitted to the New Zealand male population for ages 0-89 and the period 1985-2008. Shades in the fan represent confidence intervals at the 50%, 80% and 95% level.

```
h <- 24
LCsimPU <- simulate(LCboot, h = h)
```

To highlight the impact of parameter risk on mortality rate projections it is instructive to compare prediction intervals with and without the allowance of parameter uncertainty. Central forecast together with 5000 trajectories of the Lee-Carter model allowing only for forecast error in the random walk with drift and ignoring the model fitting error can be obtained with the code:

```
LCfor <- forecast(LCfit, h = h)
LCsim <- simulate(LCfit, nsim = 5000, h = h)
```

Figure 11 depicts 95% prediction intervals for mortality rates at age 40, 60 and 80 with and without allowance for parameter uncertainty. This graph was produced with the code:

```
#Observed, fitted and central forecasts
mxt <- LCfit$Dxt / LCfit$Ext
mxtHat <- fitted(LCfit, type = "rates")
mxtCentral <- LCfor$rates
#95% Prediction intervals without parameter uncertainty
mxtPred2.5 <- apply(LCsim$rates, c(1, 2), quantile, probs = 0.025)
mxtPred97.5 <- apply(LCsim$rates, c(1, 2), quantile, probs = 0.975)
#95% intervals with parameter uncertainty (in sample, and predictions)
mxtHatPU2.5 <- apply(LCsimPU$fitted, c(1, 2), quantile, probs = 0.025)
mxtHatPU97.5 <- apply(LCsimPU$fitted, c(1, 2), quantile, probs = 0.975)
mxtPredPU2.5 <- apply(LCsimPU$rates, c(1, 2), quantile, probs = 0.025)
mxtPredPU97.5 <- apply(LCsimPU$rates, c(1, 2), quantile, probs = 0.975)
#Plot
x <- c("40", "60", "80")
```

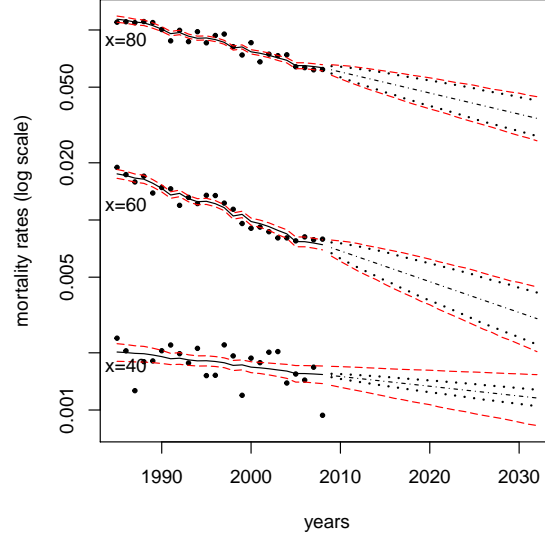


Figure 11: 95% Prediction intervals for mortality rates q_{xt} at ages $x = 40$ (bottom lines), $x = 60$ (middle lines) and $x = 80$ (top lines) for the Poisson Lee-Carter model fitted to the New Zealand male population for ages 0-89 and the period 1985-2008. Dots show historical mortality rates for 1985-2008 and solid black lines show the corresponding fitted rates. Dashed lines represent central forecast, black dotted lines represent 95% prediction intervals excluding parameter uncertainty and dot-dashed red lines depict 95% confidence and prediction intervals including parameter uncertainty.

```
matplot(LCfit$years, t(mxt[x, ]), xlim = range(LCfit$years, LCfor$years),
        ylim = range(mxtHatPU97.5[x, ], mxtPredPU2.5[x, ], mxt[x, ]),
        type = "p", xlab = "years", ylab = "mortality rates (log scale)",
        log = "y", pch = 20, col = "black")
matlines(LCfit$years, t(mxtHat[x, ]), lty = 1, col = "black")
matlines(LCfit$years, t(mxtHatPU2.5[x, ]), lty = 5, col = "red")
matlines(LCfit$years, t(mxtHatPU97.5[x, ]), lty = 5, col = "red")
matlines(LCfor$years, t(mxtCentral[x, ]), lty = 4, col = "black")
matlines(LCsim$years, t(mxtPred2.5[x, ]), lty = 3, col = "black")
matlines(LCsim$years, t(mxtPred97.5[x, ]), lty = 3, col = "black")
matlines(LCsimPU$years, t(mxtPredPU2.5[x, ]), lty = 5, col = "red")
matlines(LCsimPU$years, t(mxtPredPU97.5[x, ]), lty = 5, col = "red")
text(1986, mxtHatPU2.5[x, "1995"], labels = c("x=40", "x=60", "x=80"))
```

In figure 11 we can clearly see that parameter uncertainty has an important impact on the prediction intervals. This is particularly notable at age 40 where in year 2030, for instance, the width of the prediction interval with parameter uncertainty is around 3 times bigger than that without parameter uncertainty. These results are in line with the results obtained by Li (2014) using the same dataset.¹¹

¹¹We note that our prediction intervals without parameter uncertainty correspond to method (I) in Li (2014) while our prediction intervals with parameter uncertainty correspond to method (IVa). We also note

9. Conclusion

In this paper we have introduced the family of Generalised Age-Period-Cohort stochastic mortality models by paralleling the standard framework of generalised linear models. In addition, we have presented the **R** package **StMoMo** which takes advantage of the unifying framework of the GAPC family to provide tools for fitting a diverse number of stochastic mortality models, analysing their goodness of fit and performing mortality projections with them. A key feature of the generalised APC family and of **StMoMo** is that they not only encompass models from the Lee-Carter and CBD families, but can also accommodate possible new models. Furthermore, model risk is a prevalent issue when forecasting mortality and we thus believe that the possibility of easily implementing and comparing a wide range of models makes **StMoMo** a valuable addition to the toolkit for measuring and managing longevity risk.

Our package can be expanded in several directions. The current version of **StMoMo** only allows the use of a log link with a Poisson distribution of deaths or a logit link with a Binomial distribution of deaths. However, we plan to expand the possible combinations of error distribution and link function to include, for instance, Binomial errors with a complementary log-log link as suggested by Currie (2014).

In agreement with the standard in the literature, package **StMoMo** assumes a multivariate random walk with drift for the dynamics of the period index κ_t . Nevertheless, the package could be expanded to accommodate more general time series models which might be more appropriate for some datasets or some models. For instance, Renshaw and Haberman (2006) use an ARIMA(2, 1, 0) for $\kappa_t^{(1)}$ when forecasting mortality for England and Wales males with the Lee-Carter model with cohorts, while Plat (2009) uses general univariate ARIMA(p, d, q) for each $\kappa_t^{(i)}$, $i = 1, \dots, N$, fitted using seemingly unrelated regressions to allow for correlations between the different period indexes.

In the generalised APC family it is assumed that the age-modulating terms $\beta_x^{(i)}$, $i = 0, \dots, N$, are either non-parametric functions of age which need to be estimated or parametric functions of age with a pre-specified functional form $f^{(i)}(x)$. This former case could be extended to include the more general case of a pre-specified functional form with a set of parameters that need to be estimated, that is, $\beta_x^{(i)} = f^{(i)}(x; \theta_i)$ with θ_i being some model parameters. This generalisation would allow the implementation of the family of model considered in Hunt and Blake (2014a).

Finally, the increasing attention that multiple population mortality models are receiving in the mortality forecasting literature provides an avenue for a wealth of extensions of our package. In particular, we are working on a two-population version of **StMoMo** which implements the class of relative two-population models considered in Villegas et al. (2015).

that our mortality rate projections use as jump-off rates the fitted rates in the final year of observation while Li (2014) appears to use the observed rates instead.

References

- Aro, H., Pennanen, T., 2011. A user-friendly approach to stochastic mortality modelling. *European Actuarial Journal* 1, 151–167.
- Booth, H., Hyndman, R. J., Tickle, L., 2014. Prospective life tables. In: Charpentier, A. (Ed.), *Computational Actuarial Science with R*. CRC Press, Boca Raton, FL, pp. 319–344.
- Booth, H., Maindonald, J., Smith, L., Nov. 2002. Applying Lee-Carter under conditions of variable mortality decline. *Population studies* 56 (3), 325–36.
- Börger, M., Fleischer, D., Kuksin, N., Oct. 2013. Modeling the Mortality Trend Under Modern Solvency Regimes. *ASTIN Bulletin* 44 (01), 1–38.
URL http://www.journals.cambridge.org/abstract_S051503611300024X
- Brouhns, N., Denuit, M., Van Keilegom, I., May 2005. Bootstrapping the Poisson log-bilinear model for mortality forecasting. *Scandinavian Actuarial Journal* 2005 (3), 212–224.
URL <http://www.tandfonline.com/doi/abs/10.1080/03461230510009754>
- Brouhns, N., Denuit, M., Vermunt, J., Dec. 2002. A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics* 31 (3), 373–393.
- Butt, Z., Haberman, S., Shang, H. L., 2014. *ilc: Lee-Carter Mortality Models using Iterative Fitting Algorithms*.
URL <http://cran.r-project.org/package=ilc>
- Cairns, A. J., Blake, D., Dowd, K., Dec. 2006. A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration. *Journal of Risk and Insurance* 73 (4), 687–718.
- Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Khalaf-Allah, M., May 2011. Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics* 48 (3), 355–367.
- Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., 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.
- Clayton, D., Schifflers, E., Jun. 1987. Models for temporal variation in cancer rates. II: Age-period-cohort models. *Statistics in medicine* 6 (4), 469–81.
URL <http://www.ncbi.nlm.nih.gov/pubmed/3629048>
- Currie, I. D., 2006. Smoothing and Forecasting Mortality Rates with P-splines.
URL <http://www.macs.hw.ac.uk/~iain/research/talks/Mortality.pdf>

- Currie, I. D., 2014. On fitting generalized linear and non-linear models of mortality. *Scandinavian Actuarial Journal* 0 (0), 1–28.
URL <http://dx.doi.org/10.1080/03461238.2014.928230><http://www.tandfonline.com/doi/abs/10.1080/03461238.2014.928230>
- Debón, A., Martínez-Ruiz, F., Montes, F., 2010. A geostatistical approach for dynamic life tables: The effect of mortality on remaining lifetime and annuities. *Insurance: Mathematics and Economics* 47 (3), 327–336.
URL <http://dx.doi.org/10.1016/j.insmatheco.2010.07.007><http://www.sciencedirect.com/science/article/pii/S0167668710000843>
- Haberman, S., Renshaw, A., Oct. 2009. On age-period-cohort parametric mortality rate projections. *Insurance: Mathematics and Economics* 45 (2), 255–270.
URL <http://linkinghub.elsevier.com/retrieve/pii/S016766870900081X>
- Haberman, S., Renshaw, A., Jan. 2011. A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics* 48 (1), 35–55.
- Hobcraft, J., Menken, J., Preston, S., 1982. Age, period, and cohort effects in demography: a review. *Population index* 48 (1), 4–43.
- Human Mortality Database, 2014. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). www.mortality.org or www.humanmortality.de.
- Hunt, A., Blake, D., Feb. 2014a. A General Procedure for Constructing Mortality Models. *North American Actuarial Journal* (February), 140210105231006.
URL <http://www.tandfonline.com/doi/abs/10.1080/10920277.2013.852963>
- Hunt, A., Blake, D., 2014b. On the Structure and Classification of Mortality Models.
- Hunt, A., Villegas, A. M., 2014. Robustness and convergence in the Lee-Carter model with cohorts. Working paper.
- Hyndman, R. J., 2014. demography: Forecasting mortality, fertility, migration and population data.
URL <http://cran.r-project.org/package=demography>
- Hyndman, R. J., Khandakar, Y., 2008. Automatic time series forecasting : the forecast package for R. *Journal Of Statistical Software* 27 (3), 1–22.
- Hyndman, R. J., Ullah, S., Jun. 2007. Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics & Data Analysis* 51 (10), 4942–4956.
- Koissi, M.-c., Shapiro, a., Hognas, G., Feb. 2006. Evaluating and extending the LeeCarter model for mortality forecasting: Bootstrap confidence interval. *Insurance: Mathematics and Economics* 38 (1), 1–20.
URL <http://linkinghub.elsevier.com/retrieve/pii/S0167668705000909>

- Lee, R., Miller, T., Nov. 2001. Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography* 38 (4), 537–49.
- Lee, R. D., Carter, L. R., Sep. 1992. Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* 87 (419), 659–671.
- Li, J., Aug. 2014. A quantitative comparison of simulation strategies for mortality projection. *Annals of Actuarial Science* 8 (02), 281–297.
URL http://www.journals.cambridge.org/abstract_S1748499514000153
- Lovász, E., Oct. 2011. Analysis of Finnish and Swedish mortality data with stochastic mortality models. Vol. 1.
URL <http://www.springerlink.com/index/10.1007/s13385-011-0039-8>
- McCullagh, P., Nelder, J., 1989. *Generalized Linear Models*, second edi Edition. Chapman & Hal, London.
- Oeppen, J., Vaupel, J. W., Jul. 2002. Broken limits to life expectancy. *Science* 296 (5570), 1029–1031.
- OHare, C., Li, Y., Jan. 2012. Explaining young mortality. *Insurance: Mathematics and Economics* 50 (1), 12–25.
URL <http://linkinghub.elsevier.com/retrieve/pii/S0167668711000990>
- Plat, R., Dec. 2009. On stochastic mortality modeling. *Insurance: Mathematics and Economics* 45 (3), 393–404.
- R Core Team, 2014. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
URL <http://www.r-project.org/>
- Renshaw, A., Haberman, S., 2003. Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics* 33 (2), 255–272.
- Renshaw, A., Haberman, S., Jun. 2006. A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics* 38 (3), 556–570.
- Renshaw, A., Haberman, S., Apr. 2008. On simulation-based approaches to risk measurement in mortality with specific reference to Poisson LeeCarter modelling. *Insurance: Mathematics and Economics* 42 (2), 797–816.
URL <http://linkinghub.elsevier.com/retrieve/pii/S0167668707000960>
- Turner, H., Firth, D., 2012. *Generalized nonlinear models in R: An overview of the gnm package*.
URL <http://cran.r-project.org/package=gnm>
- van Berkum, F., Antonio, K., Vellekoop, M., Dec. 2014. The impact of multiple structural changes on mortality predictions. *Scandinavian Actuarial Journal* (January 2015), 1–23.

- Villegas, M., Haberman, S., Kaishev, V., Millossovich, P., 2015. A comparative study of two-population models for the assessment of basis risk in longevity hedges. Working paper, 1–29.
- Wang, D., Lu, P., 2005. Modelling and forecasting mortality distributions in England and Wales using the LeeCarter model. *Journal of Applied Statistics* 32 (February 2015), 873–885.
- Wikipedia, 2014. King Momo, in Wikipedia, The Free Encyclopedia.
URL http://en.wikipedia.org/w/index.php?title=King_Momo&oldid=635124783
- Willets, R., 2004. The cohort effect: Insights and explanations. *British Actuarial Journal* 10 (4), 833–877.