STATISTICS IN TRANSITION new series, December 2017
Vol. 18, No. 4, pp. 701–724, DOI 10.21307/stattrans-2017-008

IMPROVEMENT OF FUZZY MORTALITY MODELS BY MEANS OF ALGEBRAIC METHODS

Andrzej Szymański, Agnieszka Rossa[1]

ABSTRACT

The forecasting of mortality is of fundamental importance in many areas, such as the funding of public and private pensions, the care of the elderly, and the provision of health service. The first studies on mortality models date back to the 19th century, but it was only in the last 30 years that the methodology started to develop at a fast rate. Mortality models presented in the literature form two categories (see, e.g. Tabeau *et al*., 2001, Booth, 2006) consisting of the so-called static or stationary models and dynamic models, respectively. Models contained in the first, bigger group contains models use a real or fuzzy variable function with some estimated parameters to represent death probabilities or specific mortality rates. The dynamic models in the second group express death probabilities or mortality rates by means of the solutions of stochastic differential equations, etc.

The well-known Lee-Carter model (1992), which is widely used today, is considered to belong to the first group, similarly as its fuzzy version published by Koissi and Shapiro (2006). In the paper we propose a new class of fuzzy mortality models based on a fuzzy version of the Lee-Carter model. Theoretical backgrounds are based on the algebraic approach to fuzzy numbers (Ishikawa, 1997a, Kosiński, Prokopowicz and Ślęzak, 2003, Rossa, Socha and Szymański, 2015, Szymański and Rossa, 2014). The essential idea in our approach focuses on representing a membership function of a fuzzy number as an element of C^*-Banach algebra. If the membership function $\mu(z)$ of a fuzzy number is strictly monotonic on two disjoint intervals, then it can be decomposed into strictly decreasing and strictly increasing functions $\Phi(z)$, $\Psi(z)$, and the inverse functions $f(u)=\Phi^{-1}(u)$ and $g(u)=\Psi^{-1}(u)$, $u \in [0, 1]$ can be found.

Ishikawa (1997a) proposed foundations of the fuzzy measurement theory, which is a general measurement theory for classical and quantum systems. We have applied this approach, termed C^*-measurement, as the theoretical foundation of the mortality model. Ishikawa (1997b) introduced also the notions of objective and subjective C^*-measurement called real and imaginary C^*-measurements. In our proposal of the mortality model the function f is treated as an objective C^*-measurement and the function g as an subjective C^*-measurement, and the

[1] Institute of Statistics and Demography, University of Łódź, e-mail: agrossa@uni.lodz.pl

membership function $\mu(z)$ is represented by means of a complex-valued function $f(u) + ig(u)$, where i is the imaginary unit. We use the Hilbert space of quaternion algebra as an introduction to the mortality models.

Key words: C^*-Banach algebra, non-commutative C^*-algebra, quaternion algebra, fuzzy mortality model.

1. Introduction

Long-lasting observations of mortality rates or death probabilities lead to the conclusion that in developed countries they decline for most age groups, whereas the upper limit of human lifetime is moving upwards. Other life-table parameters also change in time. The mortality trends and patterns observed in developed countries in the second half of the 20th century can be summed up as follows (see also Wilmoth and Horiuchi, 1999):

- the normal lifetime drifts toward older ages,
- ages at deaths are concentrating around the normal lifetime,
- the survival curve is undergoing rectangularization (because of the aforementioned trends),
- the life expectancy is increasing,
- in the young population (especially among young males aged 20+), the number and percentage of deaths from external causes (injuries, accidents, poisoning) is rising.

These measures are therefore not constant in time. They are rather functions of time or, in broader terms, stochastic processes showing some variability. Past works on this subject have used, for instance, time-series analysis tools to examine the stochastic nature of these processes. One of the most popular is the Lee-Carter mortality model (Lee and Carter, 1992).

2. The Lee-Carter mortality model

Let $m_x(t)$ denote an age-specific (central) death rate for the subset of a population that is between exact ages x and x+1

$$m_x(t) = \frac{D_x(t)}{L_x(t)}, \quad x=0,1,2,\ldots,X, \quad t=1,2,\ldots,T, \tag{2.1}$$

where

$D_x(t)$ – the number of deaths at age x in the year t,

$L_x(t)$ – the midyear population at the age x in the year t,

$x=0,1,\ldots,X$ – index of one-year age groups,

$t=1,2,\ldots,T$ – years of observation period.

The measure $m_x(t)$ is the ratio of deaths between ages x and $x + 1$ to the midyear population alive at age x in the given year t, also referred to as the mean population in the year t. The measure is described as the central rate because the midyear population is used in the denominator.
The Lee-Carter model can be written as

$$\ln m_x(t) = \alpha_x + \beta_x \kappa_t + \epsilon_{xt}, \qquad x=0,1,\dots,X,\ t=1,2,\dots,T \tag{2.2}$$

or, equivalently, as

$$m_x(t) = \exp\{\alpha_x + \beta_x \kappa_t + \epsilon_{xt}\}, \qquad x=0,1,\dots,X,\ t=1,2,\dots,T, \tag{2.3}$$

where $m_x(t)$, $t \in \mathbf{N}$ are age-specific mortality rates, α_x, β_x and κ_t are the model parameters, of which α_x, β_x depend on age x and κ_t on time t. The double-indexed terms $\varepsilon_{x,t}$ are error terms, which are assumed to be independent and to have the same normal distributions with an expected value of 0 and constant variance.

The parameters α_x, $x=0,1,\dots,X$ indicate the general shape of the mortality schedule, the time-varying parameters κ_t, $t=1,2,\dots,T$ represent the time-trend indices of the general mortality level, whereas β_x indicate the pattern of deviations from the age profile when the general level of mortality κ_t changes. In general, β_x could be negative at some ages, indicating that mortality rates at those ages tend to rise when falling at other ages. In other words, the shape of β_x profile tells which rates decline rapidly and which slowly over time in response to change of κ_t.

Because of the form of (2.2), the Lee-Carter model is called a bilinear model. The system of equations (2.2) or (2.3) cannot be explicitly solved unless additional restrictions are imposed. Let us assume, for instance, that for a set of parameters $\{\alpha_x\}$,$\{\beta_x\}$, and $\{\kappa_t\}$ the model (2.2) is valid. It is easy to see that the model holds true also for any constant c and parameters $\{\alpha_x - c\beta_x\}$, $\{\beta_x\}$, $\{\kappa_t + c\}$ or $\{\alpha_x\}$, $\{c\beta_x\}$, $\{\kappa_t/c\}$.

To make sure that an unambiguous solution is obtained, some additional restrictions must be defined. To this end, it is assumed that the sum of parameters β_x over age index x is 1 and the sum of parameters κ_t over time index t is equal to 0, i.e.

$$\sum_{x=0}^{X} \beta_x = 1, \ \sum_{t=1}^{T} \kappa_t = 0. \tag{2.4}$$

Parameters α_x and β_x do not depend on time t, which means that once they have been established they can also be used for the future period, i.e. $t > T$. The time-varying rates are κ_t. They can be further modelled using, for instance, the time series analysis methods.

Lee and Carter (1992) proposed a random walk model, but the range of proposals discussed in the literature is wider. A random walk process with a drift is given by the formula

$$\kappa_t = \delta + \kappa_{t-1} + \xi_t, \tag{2.5}$$

where δ is a constant (a drift), and ξ_t is a random term.

Parameter δ in (2.5) mostly takes negative values that point to declining mortality. Random fluctuations around this trend are represented by independent random terms ξ_t, each having a normal distribution with the expected value of 0 and finite variance.

With the values of κ_t predicted from (2.5) and the estimations of α_x and β_x the partial death rates can be forecasted, as well as other life-table mortality rates.

The method of parameter estimation proposed by Lee and Carter is based on the method of Singular Value Decomposition (SVD), which decomposes a data matrix **M** = [ln $m_x(t) - a_x$] into a matrix of singular values **D** and two matrices **W** and **V** of left and right singular vectors.

Let a_x, b_x, k_t represent the estimators of parameters α_x, β_x, κ_t. Assuming that random terms ϵ_{xt} in model (2.2) have an expected value of 0, we have

$$\mathrm{E}(\epsilon_{xt}) = 0. \tag{2.6}$$

This property will be used to find a_x. To this end, we will determine the analogous first row moment from the sample, i.e. from time series {ln $m_x(t)$, t = 1,2,...,T} for x = 0,1,2,...,X we calculate the sum

$$\sum_{t=1}^{T}[\ln m_x(t) - (a_x + b_x k_t)], \tag{2.7}$$

then by comparing (2.7) with 0

$$\sum_{t=1}^{T}[\ln m_x(t) - (a_x + b_x k_t)] = 0, \tag{2.8}$$

we obtain the following equality

$$T a_x + b_x \sum_{t=1}^{T} k_t = \sum_{t=1}^{T} \ln m_x(t). \tag{2.9}$$

By allowing additionally for condition $\sum_{t=1}^{T} k_t = 0$, we arrive at

$$a_x = \frac{1}{T}\sum_{t=1}^{T} \ln m_x(t). \tag{2.10}$$

To estimate β_x, κ_t the first singular value and the first vector of matrices **W** and **V** are used. For a general case, all singular values and singular vectors can be employed, which gives the following extension of the model (2.2)

$$\ln m_x(t) = \alpha_x + \sum_{i=1}^{r} \beta_x^{(i)} \kappa_t^{(i)}, \qquad x=0,1,\dots,X,\ t=1,2,\dots,T, \tag{2.11}$$

where r is the number of non-zero singular values.

3. The Koissi-Shapiro model

One of the most interesting generalisations of the Lee-Carter model, referring to the algebra of fuzzy numbers, was proposed by Koissi and Shapiro (2006). Their version of the Lee-Carter model (FLC model) assumes a fuzzy representation of the central death rates. It allows taking account of uncertainty involved in mortality rates and entering a random term into the fuzzy structure of the model.

Their approach builds on the assumption that the real rates of mortality are not exactly $m_x(t)$, but rather around $m_x(t)$, thus the role of the explanatory variable is played by fuzzified mortality rates.

It is well-known that death statistics are subject to reporting errors of several kinds. They may be reported for incorrect year, area, or assigned statistics that are incorrect, e.g. age. Moreover, the midyear population data that serve as the denominators of mortality rates are also the subject of errors. It is regarded as the population at July 1 and is assumed to be the point at which half of the deaths in the population during the year have occurred. Such an estimate can be actually underestimated or overestimated. For these reasons, fuzzy representation of the central death rates seems to be justified.

Koissi and Shapiro proposed fuzzy representation of the logarithms of age-specific mortality rates ln $m_x(t)$, by converting them into symmetric, triangular fuzzy numbers (basic notions of the fuzzy numbers are given in Rossa, Socha, Szymański (2015, appendix) presented as

$$Y_{xt} = (y_{xy}, e_{xt}), \quad x = 0,1,...,X, \quad t = 1,2,...,T, \tag{3.1}$$

where y_{xt} = ln $m_x(t)$ and e_{xt} are the spreads of the membership functions of triangular fuzzy numbers.

In fuzzification approach, a fuzzy least-squares regression based on minimum fuzziness criterion was employed, and – for simplicity – triangular symmetric fuzzy numbers were considered.

Given the log-central death rates y_{xt}= ln $m_x(t)$ for age x in year t, the task is to find symmetric triangular fuzzy numbers $A_0 = (c_{0x}, s_{0x})$, $A_1 = (c_{1x}, s_{1x})$ and $Y_{xt} = (y_{xt}, e_{xt})$ with centers c_{0x}, c_{1x}, y_{xt} and spreads s_{0x}, s_{1x}, e_{xt} such that

$$(y_{xt}, e_{xt}) = (c_{0x}, s_{0x})+(c_{1x}, s_{1x})\times t\,. \tag{3.2}$$

To find the fuzzy numbers A_0 and A_1, the approach is as follows:

1. First, ordinary least-squares (OLS) regression is used to find the center values c_{0x} and c_{1x} such that

$$y_{xt} = c_{ox} + c_{1x}t + \varepsilon_{xt}, \quad \text{for each } x, \tag{3.3}$$

where y_{xt} = ln $m_x(t)$ are the observed log-central death rates, t is time variable, and ε_{xt} represent random terms.

2. The spreads (s_{0x} and s_{1x}) are obtained by using the minimum fuzziness criterion. This consists in minimizing the following optimization problem, which can be solved through standard optimization software, i.e. minimize

$$Ts_{0x} + s_{1x}\sum_{t=1}^{T} t \tag{3.4}$$

subject to

$$\forall_t \ s_{0x}, s_{1x} \geq 0$$

$$c_{0x} + c_{1x}t + (s_{0x}+s_{1x}t) \geq y_{xt}, \quad \text{and} \quad c_{0x} + c_{1x}t - (s_{0x}+s_{1x}t) \leq y_{xt}\,.$$

Once the log-central death rates are fuzzified, the FLC model can be defined as

$$Y_{xt} = A_x \oplus_{Tw} (B_x \otimes_{Tw} K_t), \qquad x = 0,1,...,X, \ t = 1,2,...,T, \tag{3.5}$$

where Y_{xt} are known fuzzy log-central mortality rates, A_x, B_x, K_t are unknown parameters, and $\oplus_{Tw}$, $\otimes_{Tw}$ are the addition and multiplication operators of fuzzy numbers in the norm T_w, respectively. For the definition of the norm T_w see Koissi and Shapiro (2006).

The authors assumed that the model parameters can be estimated by minimizing the criterion function based on the Diamond distance measure between fuzzy variables. The criterion can be expressed as the following sum

$$\sum_{x=0}^{X}\sum_{t=1}^{T}[3a_x^2 + 3b_x^2k_t^2 + 3y_{xt}^2 + 6a_xb_xk_t - 4y_{xt}(a_x + b_xk_t) + 2e_{xt}^2] + \tag{3.6}$$

$$+2\sum_{x=0}^{X}\sum_{t=1}^{T}\left[\left(\max\{s_{A_x}, |b_x|s_{K_t}, |k_t|s_{B_x}\}\right)^2 - 2e_{xt}\max\{s_{A_x}, |b_x|s_{K_t}, |k_t|s_{B_x}\}\right].$$

However, the FLC model poses major problems in the estimation algorithm, because expression $\max\{s_{A_x}, |b_x|s_{K_t}, |k_t|s_{B_x}\}$ in the criterion (3.6) prevents the standard use of non-linear optimization methods.

In the rest of the paper, modification to the fuzzy mortality model based on fuzzified mortality rates with exponential membership functions will be proposed. The model simplifies both operations on fuzzy numbers and the model estimation. The essential idea in this approach is representing the membership functions of fuzzy numbers as elements of C*-Banach algebra.

4. A new class of mortality models based on algebraic approach to fuzzy numbers

4.1. The theoretical background for the new mortality model

Fuzzification of data depends on the assumption about membership functions of fuzzy numbers. Koissi and Shapiro (2006) adopted triangular symmetric membership functions and used fuzzy least-squares regression. In our approach, we will assume exponential membership functions derived from relative frequencies of residuals in the least-squares regression model.

Suppose that the membership function $\mu(z)$ of a fuzzy number is strictly monotonic on two disjoint intervals. Following Nasibov and Peker (2011), we will consider an exponential membership function of the form

$$\mu(z) = \begin{cases} \exp\left\{-\left(\frac{c-z}{\tau}\right)^2\right\}, & \text{for } z \le c, \\ \exp\left\{-\left(\frac{z-c}{\nu}\right)^2\right\}, & \text{for } z > c, \end{cases} \tag{4.1}$$

where c, τ, ν are scalars.

Note that we can decompose $\mu(z)$ into two parts – strictly increasing and strictly decreasing functions $\Psi(z)$ and $\Phi(z)$ of the form

$$\begin{aligned} \Psi(z) &= \exp\left\{-\left(\frac{c-z}{\tau}\right)^2\right\}, \quad \text{for } z \le c, \\ \Phi(z) &= \exp\left\{-\left(\frac{z-c}{\nu}\right)^2\right\}, \quad \text{for } z > c. \end{aligned} \tag{4.2}$$

Then, there exist inverse functions

$$\Psi^{-1}(u) = c + \psi(u), \quad \Phi^{-1}(u) = c + \varphi(u), \quad u \in [0,1], \tag{4.3}$$

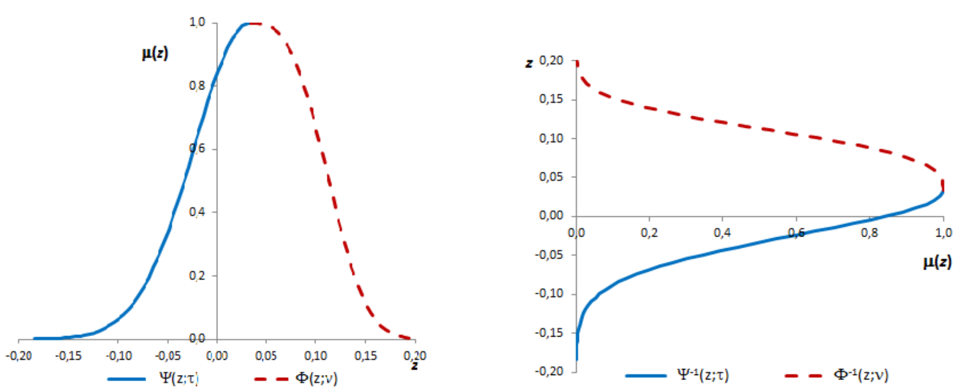
where $\psi(u)$ and $\varphi(u)$ are expressed as follows

$$\psi(u) = -\tau(-\ln u)^{\frac{1}{2}}, \quad \varphi(u) = \nu(-\ln u)^{\frac{1}{2}}, \quad u \in [0,1]. \tag{4.4}$$

Example 1. Figure 1(a) illustrates an exponential functions (4.2) for fixed values of parameters c=0.03, $\tau = 0.08, \nu = 0.09$, Figure 1(b) presents respective inverse functions (4.3).

Figure 1. An example of an exponential membership function, c=0.03, τ =0.08, ν =0.09

(a) **(b)**



Source: developed by the authors.

4.2. Transformation of membership functions into complex-valued functions

Let us consider the complex functions

$$f(u) = c + i\psi(u), \quad \text{and} \quad g(u) = c + i\varphi(u), \quad u \in [0,1], \tag{4.5}$$

where $i = \sqrt{-1}$ is an imaginary unit.

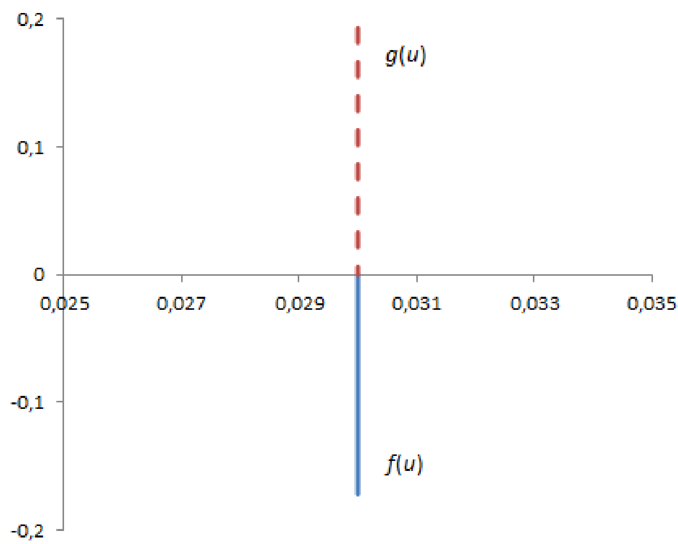
Assuming that functions $\psi(u)$ and $\varphi(u)$ are expressed as in (4.4) we get

$$f(u) = c - i\tau(-\ln u)^{\frac{1}{2}}, \quad \text{and} \quad g(u) = c + i\nu(-\ln u)^{\frac{1}{2}}, \quad u \in [0,1]. \tag{4.6}$$

The pair of two complex functions $(f(u), g(u))$ is called a quaternion.

An illustration of a quaternion $(f(u), g(u))$ on the complex plane for c=0.03, τ=0.08, ν=0.09 is presented in Figure 2.

Figure 2. A quaternion $(f(u), g(u))$, with $f(u)$ and $g(u)$ defined in (4.4) with $\tau = 0.08, \nu = 0.09$



Source: developed by the authors.

The modules of $f(u)$ and $g(u)$ are as follows

$$|f(u)|^2 = c^2 + \tau^2(-\ln u), \; u \in [0,1], \tag{4.7}$$

$$|g(u)|^2 = c^2 + \nu^2(-\ln u), \; u \in [0,1]. \tag{4.8}$$

After integrating both sides of (4.7) and (4.8) on the interval [0,1] we obtain

$$\int_0^1 |f(u)|^2 du = c^2 + \tau^2 \int_0^1 (-\ln u)\, du = c^2 + \tau^2, \tag{4.9}$$

$$\int_0^1 |g(u)|^2 du = c^2 + \nu^2 \int_0^1 (-\ln u)\, du = c^2 + \nu^2. \tag{4.10}$$

4.3. Basic properties of quaternions

It is well known that the complex numbers could be viewed as ordered pairs of real numbers. By analogy, the quaternions can be treated as ordered pairs of complex functions

$$(z, w), \quad \text{where} \quad z = a + ib, \quad w = c + id \quad \text{and} \quad i = \sqrt{-1}. \qquad (4.11)$$

The algebra of quaternions is often denoted by $\boldsymbol{H}$. Quaternions were first described by Irish mathematician William Hamilton in 1843. The space $\boldsymbol{H}$ is equipped with three operations: addition, scalar multiplication and quaternion multiplication.

The sum of two elements of $\boldsymbol{H}$ is defined as the sum of their components. Therefore, we have

$$(z, w) + (u, x) = (z + u, w + x). \qquad (4.12)$$

The product of an element of $\boldsymbol{H}$ by a real number $\alpha \in R$ is defined to be the same as the product by scalar of both components

$$\alpha(z, w) = (\alpha z, \alpha w). \qquad (4.13)$$

To define the product of two elements in $\boldsymbol{H}$ a choice of the basis for $\boldsymbol{R}^4$ is needed. The elements of this basis are customarily denoted as $1, i, j$ and k. Each element of $\boldsymbol{H}$ can be uniquely denoted as a linear combination $a \cdot 1 + bi + cj + dk$, where a, b, c, d are real numbers.

The basis element 1 could be viewed as the identity element of $\boldsymbol{H}$. It means that multiplication by 1 does not change the value, and elements of $\boldsymbol{H}$ can be uniquely denoted as

$$(z, w) = a + bi + cj + dk, \qquad (4.14)$$

where a, b, c, d are real numbers. Therefore, each element of $\boldsymbol{H}$ is determined by four numbers and hence the term "quaternion".

The possible products of basic elements i, j, k can be described as follows

$$i^2 = j^2 = k^2 = ijk = -1, \qquad (4.15)$$

$$ij = k, \ ji = -k, \qquad (4.16)$$

$$jk = i, \ kj = -i, \qquad (4.17)$$

$$ki = j, \ ik = -j. \qquad (4.18)$$

Quaternions can be represented as pairs of complex numbers as a generalization of the construction of the complex numbers being pairs of real numbers.

Let C be a two-dimensional vector space over the complex numbers. Let us choose a basis consisting of two elements 1 and j. For $z, w \in C$ of the form $z = a + bi$ and $w = c + di$, we can write

$$q = z + wj = (a + bi) + (c + di)j = a + bi + cj + dij. \tag{4.19}$$

If we denote $k = ij$ then

$$q = z + wj = a + bi + cj + dk. \tag{4.20}$$

Thus, the vector (z, w) corresponds to a quaternion $q = a + bi + cj + dk$. Then, each quaternion $q \in \boldsymbol{H}$ is uniquely represented by

$$q = z + wj. \tag{4.21}$$

Multiplication of quaternions could be defined in the form

$$(z, w)(u, x) = (zu - w\bar{x}, zx + w\bar{u}), \tag{4.22}$$

where $\bar{x}, \bar{u}$ denote conjugations of x and u.

Multiplication of quaternions is associative and distributive with respect to addition, however it is not commutative, since, for example, we have

$$(i, 0)(0,1) = (0, i), \tag{4.23}$$

but

$$(0,1)(i, 0) = (0, -i). \tag{4.24}$$

Let us denote

$$q^* = z - wj \tag{4.25}$$

as the conjugate of q.

Conjugation is an involution. It means that for $p, q \in \boldsymbol{H}$ we have

$$(q^*)^* = q, \quad (pq)^* = q^*p^*, \quad (p + q)^* = p^* + q^*. \tag{4.26}$$

The square root of the product of a quaternion with its conjugate is called a norm, and is denoted $\|q\|$. This is expressed as follows

$$\|q\| = \sqrt{qq^*} = \sqrt{q^*q} = \sqrt{a^2 + b^2 + c^2 + d^2}. \tag{4.27}$$

It is always a non-negative real number, and it is the same as the Euclidean norm on $\boldsymbol{H}$ considered as the vector space $\boldsymbol{R}^4$. Multiplying a quaternion by a real number scales its norm by the absolute value of this number

$$\|\alpha q\| = |\alpha|\|q\|. \tag{4.28}$$

This is a special case of the following property

$$\|pq\| = \|p\|\|q\| \tag{4.29}$$

for any two quaternions p and q.

The norm (4.27) allows us to define the distance $d(p,q)$ between p and q as the norm of their difference

$$d(p,q) = \|p - q\|. \tag{4.30}$$

This defines $\boldsymbol{H}$ as a metric space.

According to (4.6) we have

$$f(u) = c + i\psi(u), \quad u \in [0,1],$$

and

$$g(u) = c + i\varphi(u), \quad u \in [0,1],$$

where ψ, φ are defined in (4.4).

Hence,

$$|f(u)|^2 = c^2 + \psi^2(u), \quad \text{and} \quad |g(u)|^2 = c^2 + \varphi^2(u).$$

Let us denote

$$P(u) = \big(f(u), g(u)\big), \quad u \in [0,1]. \tag{4.31}$$

The function P is a quaternion-valued function. The norm of $P(u)$ could be expressed as follows

$$\|P(u)\|^2 = |f(u)|^2 + |g(u)|^2 = c^2 + \psi^2(u) + c^2 + \varphi^2(u), \tag{4.32}$$

and from (4.9) and (4.10) we have

$$\int_0^1 |f(u)|^2 du < \infty \quad and \quad \int_0^1 |g(u)|^2 du < \infty.$$

Integrating both sides in (4.32) we receive also

$$\int_0^1 \|P(u)\|^2 du = \int_0^1 |f(u)|^2 du + \int_0^1 |g(u)|^2 \, du < \infty. \tag{4.33}$$

Thus, the functions f and g are the elements of the Hilbert space $L_2[0,1]$, and the quaternion-valued function P is integrable with squared norm on the interval [0,1]. Let us denote the space of such functions as $L_2(\boldsymbol{H})$.

5. A mortality model based on quaternion-valued functions

5.1. Formulation of the model

We will assume that $\tilde{Y}_{x,t} = \left(f_{Y_{x,t}}, g_{Y_{x,t}}\right)$ are quaternions with complex functions $f_{Y_{x,t}}, g_{Y_{x,t}}$ of the form

$$f_{Y_{x,t}}(u) = y_{xt} - i\tau_x(-\ln u)^{\frac{1}{2}}, \qquad g_{Y_{x,t}}(u) = y_{xt} + i\nu_x(-\ln u)^{\frac{1}{2}}, \qquad u \in [0,1],$$

where i is an imaginary unit, $y_{xt} = \ln m_x(t)$, and τ_x, υ_x are known parameters evaluated by means of Nasibov-Peker method (see section 5.3 for more details).

Similarly, we will assume that $\tilde{A}_x = (f_{A_x}, g_{A_x})$, $\widetilde{K}_t = (f_{K_t}, g_{K_t})$ are quaternions determined by complex functions

$$f_{A_x}(u) = a_x - i(-\ln u)^{\frac{1}{2}} s_{A_x}^L, \quad g_{A_x}(u) = a_x + i(-\ln u)^{\frac{1}{2}} s_{A_x}^R, \quad u \in [0,1] \tag{5.1}$$

$$f_{K_t}(u) = k_t - i(-\ln u)^{\frac{1}{2}} s_{K_t}, \quad g_{K_t}(u) = k_t + i(-\ln u)^{\frac{1}{2}} s_{K_t}, \quad u \in [0,1]. \tag{5.2}$$

As in other models based on functional analysis, we postulate the following mortality model based on quaternion-valued functions

$$\widetilde{Y}_{x,t} = \tilde{A}_x + b_x \widetilde{K}_t, \qquad x = 0,1,\dots,X,\ t = 1,2,\dots,T, \tag{5.3}$$

where $Y_{x,t}$ are fuzzified log-central mortality rates expressed in terms of quaternion-valued functions in the Hilbert space $L_2(\boldsymbol{H})$, $b_x \in \boldsymbol{R}, x = 0,1,\dots,X,$ is a set of unknown scalar parameters, and quaternions $\tilde{A}_x, \widetilde{K}_t$ represent unknown parameters in $L_2(\boldsymbol{H})$ determined by the complex functions (5.1) and (5.2). The proposed model (5.3) will be termed Complex Number Mortality Model (CNMM).

Note that the quaternions $\tilde{A}_x = (f_{A_x}, g_{A_x})$, $\widetilde{K}_t = (f_{K_t}, g_{K_t})$ on the right-hand side of (5.3) reflect fuzzy numbers A_x, K_t with exponential membership functions $\mu_{A_x}(z)$ and $\mu_{K_t}(z)$ (see sections 4.1 and 4.2)

$$\mu_{A_x}(z) = \begin{cases} \exp\left\{-\left(\frac{a_x - z}{s_{A_x}^L}\right)^2\right\}, & \text{for } z \le a_x, \\ \exp\left\{-\left(\frac{z - a_x}{s_{A_x}^R}\right)^2\right\}, & \text{for } z > a_x, \end{cases} \tag{5.4}$$

$$\mu_{K_t}(z) = \begin{cases} \exp\left\{-\left(\frac{k_t - z}{s_{K_t}}\right)^2\right\}, & \text{for } z \le k_t, \\ \exp\left\{-\left(\frac{z - k_t}{s_{K_t}}\right)^2\right\}, & \text{for } z > k_t. \end{cases} \tag{5.5}$$

Using the properties (4.12) and (4.13) the complex functions defining the quaternion $\tilde{A}_x + b_x \widetilde{K}_t$ on the right-hand side of (5.3) are as follows

$$f_{A_x + b_x K_t}(u) = a_x + b_x k_t - i(-\ln u)^{\frac{1}{2}} \left(s_{A_x}^L + b_x s_{K_t}\right),\ u \in [0,1], \tag{5.6}$$

$$g_{A_x + b_x K_t}(u) = a_x + b_x k_t + i(-\ln u)^{\frac{1}{2}} \left(s_{A_x}^R + b_x s_{K_t}\right),\ u \in [0,1]. \tag{5.7}$$

It means that $\tilde{A}_x + b_x\widetilde{K}_t$ reflects a fuzzy number W_{xt} with an exponential membership function

$$\mu_{W_{xt}}(z) = \begin{cases} \exp\left\{-\left(\frac{a_x+b_xk_t-z}{s^L_{A_x}+b_xs_{K_t}}\right)^2\right\}, & \text{for } z \le a_x + b_xk_t, \\ \exp\left\{-\left(\frac{z-a_x-b_xk_t}{s^R_{A_x}+b_xs_{K_t}}\right)^2\right\}, & \text{for } z > a_x + b_xk_t. \end{cases} \quad (5.8)$$

5.2. Estimation of the model parameters

In order to estimate the parameters $a_x, b_x, k_t, s^L_{A_x}, s^R_{A_x}, s_{K_t}$ we will use the notion of the norm (4.32) defined in the space of quaternion-valued functions. Thus, the following distance between left- and right-hand sides of the model (5.3) will be defined for fixed x and t

$$d_{x,t} = \int_0^1 \left\|\tilde{Y}_{x,t} - \left(\tilde{A}_x + b_x\widetilde{K}_t\right)\right\|^2 du$$
$$= \int_0^1 \left|f_{Y_{x,t}-(A_x+b_xK_t)}(u)\right|^2 du + \int_0^1 \left|g_{Y_{x,t}-(A_x+b_xK_t)}(u)\right|^2 du.$$

Let us find functions $f_{Y_{x,t}-(A_x+b_xK_t)}(u)$ and $g_{Y_{x,t}-(A_x+b_xK_t)}(u)$ determining the difference of quaternions $\tilde{Y}_{x,t} - \left(\tilde{A}_x + b_x\widetilde{K}_t\right)$. We have

$$f_{Y_{x,t}-(A_x+b_xK_t)}(u) = y_{x,t} - (a_x + b_xk_t) - i(-\ln u)^{\frac{1}{2}}\left(\tau_x - s^L_{A_x} - b_xs_{K_t}\right), \quad (5.9)$$

$$g_{Y_{x,t}-(A_x+b_xK_t)}(u) = y_{x,t} - (a_x + b_xk_t) + i(-\ln u)^{\frac{1}{2}}\left(\nu_x - s^R_{A_x} - b_xs_{K_t}\right). \quad (5.10)$$

Hence,

$$\left|f_{Y_{x,t}-(A_x+b_xK_t)}(u)\right|^2 = \left(y_{x,t} - (a_x + b_xk_t)\right)^2 + (-\ln u)\left(\tau_x - s^L_{A_x} - b_xs_{K_t}\right)^2, \quad (5.11)$$

$$\left|g_{Y_{x,t}-(A_x+b_xK_t)}(u)\right|^2 = \left(y_{x,t} - (a_x + b_xk_t)\right)^2 + (-\ln u)\left(\nu_x - s^R_{A_x} - b_xs_{K_t}\right)^2. \quad (5.12)$$

Integrating (5.11) and (5.12) on the interval [0,1] we receive

$$d_{x,t} = \int_0^1 \left\|\tilde{Y}_{x,t} - \left(\tilde{A}_x + b_x\widetilde{K}_t\right)\right\|^2 du =$$
$$= 2\left(y_{x,t} - (a_x + b_xk_t)\right)^2 + \left(\tau_x - s^L_{A_x} - b_xs_{K_t}\right)^2 + \left(\nu_x - s^R_{A_x} - b_xs_{K_t}\right)^2. \quad (5.13)$$

By the analogy to the Lee-Carter model and restrictions (2.4) we will assume that

$$\sum_{t=1}^{T} k_t = 0,\ \ \sum_{x=0}^{X} b_x = 1. \tag{5.14}$$

An additional restriction will be also imposed on the sum of s_{K_t}

$$\sum_{t=1}^{T} s_{K_t} = (X+1)\sqrt{\sum_{t=1}^{T}(\bar{y}_t - \bar{y})^2}, \tag{5.15}$$

where $\bar{y}_t = \frac{1}{X+1}\sum_{x=0}^{X} y_{xt}$ and $\bar{y} = \frac{1}{T(X+1)}\sum_{t=1}^{T}\sum_{x=0}^{X} y_{xt}$.

Thus, the criterion used to estimate model parameters takes the form

$$F\left(a_x, b_x, k_t, s_{A_x}^L, s_{A_x}^R, s_{K_t}, \lambda_1, \lambda_2, \lambda_3\right) = \sum_{x=0}^{X}\sum_{t=1}^{T} d_{x,t} + \lambda_1\left(\sum_{x=0}^{X} b_x - 1\right) + \lambda_2 \sum_{t=1}^{T} k_t + \lambda_3\left(\sum_{t=1}^{T} s_{K_t} - (X+1)\sqrt{\sum_{t=1}^{T}(\bar{y}_t - \bar{y})^2}\right), \tag{5.16}$$

where $\lambda_1, \lambda_2, \lambda_3$ represent Lagrange multipliers.

To minimize (5.16) it is necessary to compute its first derivatives with respect to $a_x, b_x, k_t, s_{A_x}^L, s_{A_x}^R, s_{K_t}, \lambda_1, \lambda_2, \lambda_3$. We have

$$\begin{cases} \frac{\partial F}{\partial a_x} = -4\sum_{t=1}^{T}(y_{xt} - a_x - b_x k_t), \\ \frac{\partial F}{\partial b_x} = -2\sum_{t=1}^{T}\left[2k_t(y_{xt} - a_x - b_x k_t) + s_{K_t}\left(\tau_x + \nu_x - s_{A_x}^L - s_{A_x}^R - 2b_x s_{K_t}\right)\right] + \lambda_1 \\ \frac{\partial F}{\partial k_t} = -4\sum_{x=0}^{X} b_x(y_{xt} - a_x - b_x k_t) + \lambda_2 \\ \frac{\partial F}{\partial s_{A_x}^L} = -2\sum_{t=1}^{T}\left(\tau_x - s_{A_x}^L - b_x s_{K_t}\right) \\ \frac{\partial F}{\partial s_{A_x}^R} = -2\sum_{t=1}^{T}\left(\nu_x - s_{A_x}^R - b_x s_{K_t}\right) \\ \frac{\partial F}{\partial s_{K_t}} = -2\sum_{x=0}^{X} b_x\left(\tau_x + \nu_x - s_{A_x}^L - s_{A_x}^R - 2b_x s_{K_t}\right) + \lambda_3 \\ \frac{\partial F}{\partial \lambda_1} = \sum_{x=1}^{X} b_x - 1 \\ \frac{\partial F}{\partial \lambda_2} = \sum_{t=1}^{T} k_t \\ \frac{\partial F}{\partial \lambda_3} = \sum_{t=1}^{T} s_{K_t} - (X+1)\sqrt{\sum_{t=1}^{T}(\bar{y}_t - \bar{y})^2} \end{cases} \tag{5.17}$$

Then, setting each derivative in (5.17) equal to zero and solving for required parameters yields the set of normal equations

$$\begin{cases} a_x = \frac{1}{T}\sum_{t=1}^{T} y_{xt} = \bar{y}_x \\ b_x = \frac{\sum_{t=1}^{T}\left[2k_t(y_{xt}-a_x)+s_{K_t}\left(\tau_x+\nu_x-s_{A_x}^L-s_{A_x}^R\right)\right]-\frac{\lambda_1}{2}}{2\sum_{t=1}^{T}\left(k_t^2+s_{K_t}^2\right)} \\ k_t = \frac{\sum_{x=0}^{X} b_x(y_{xt}-a_x)-\frac{\lambda_2}{4}}{\sum_{x=0}^{X} b_x^2} \\ s_{A_x}^L = \tau_x - \frac{1}{T} b_x \sum_{t=1}^{T} s_{K_t} \\ s_{A_x}^R = \nu_x - \frac{1}{T} b_x \sum_{t=1}^{T} s_{K_t} \\ s_{K_t} = \frac{\sum_{x=0}^{X} b_x\left(\tau_x+\nu_x-s_{A_x}^L-s_{A_x}^R\right)-\frac{\lambda_3}{2}}{2\sum_{x=0}^{X} b_x^2} \\ \sum_{x=1}^{X} b_x = 1 \\ \sum_{t=1}^{T} k_t = 0 \\ \sum_{t=1}^{T} s_{K_t} - (X+1)\sqrt{\sum_{t=1}^{T}(\bar{y}_t-\bar{y})^2} = 0 \end{cases} \tag{5.18}$$

Note that the last three equations in (5.18) satisfy restrictions (5.14) and (5.15).

This set of normal equations can be solved numerically by means of an iterative procedure. After choosing a set of starting values, equations are computed sequentially using the most recent set of parameter estimates obtained from the right-hand side of each equation. In addition to numerical solution of the normal equations, there are also other minimizing algorithms, e.g. computer routines available in several mathematical packages (e.g. quasi-Newton or simplex methods).

Prediction of the log-central death rates with the CNMM can be performed in three steps. First, the random-walk model with a drift (2.5) should be used to predict time parameters k_t for future periods *t*>*T*. Next, functions (5.6) and (5.7) should be determined using estimated parameters $a_x, b_x, s_{A_x}^L, s_{A_x}^R, s_{K_t}$ and the sequence of predicted time indices $k_t, t > T$. Note that the functions (5.6) and (5.7) define the right-hand side of the mortality model (5.3) for $t > T$, i.e. they define quaternions $\tilde{A}_x + b_x\tilde{K}_t$ for future periods. Finally, these quaternions $\tilde{A}_x + b_x\tilde{K}_t$ can be transformed into fuzzy numbers W_{xt} using exponential membership function $\mu_{W_{xt}}(z)$ given in (5.8). They also can be further defuzzified into crisp numbers w_{xt}, if necessary, i.e. by means of the centroid defuzzification method

$$w_{xt} = \frac{\sum_{z=\epsilon}^{1} z\mu_{W_{xt}}(z)}{\sum_{z=\epsilon}^{1} \mu_{W_{xt}}(z)}, \tag{5.19}$$

where $\epsilon > 0$ denotes a small positive number.

The crisp values w_{xt} represent predicted fuzzy log-central death rates for $t>T$, whereas W_{xt} are their fuzzy counterparts.

5.3. Fuzzification of log-central death rates

Fuzzification of the log-central death rates $y_{xt} = \ln m_x(t)$ for x=0,1,...,X, t = 1,2,...,T by means of exponential membership functions (4.1) will be based on the method proposed by Nasibov and Peker (2011), which allows us to determine parameters τ_x, ν_x for a fixed x based on an empirical distribution of a sequence of data. The main results of their work are introduced in this section.

Assume that $\{r_t, t = 1,2,\ldots,T\}$ is a sequence of T observations in a data set. Assume that observation are grouped into a frequency table with k mutually exclusive class intervals (Table 1).

Table 1. Frequency table

Class intervals	Midpoints z_i	Frequencies f_i	Relative frequencies p_i
r_1- r_2	$z_1=(r_1+r_2)/2$	f_1	$p_1=f_1/T$
r_2- r_3	$z_2=(r_2+r_3)/2$	f_2	$p_2=f_2/T$
…	…	…	…
r_{K-1}- r_K	$r_k = (r_{K-1}+ r_K)/2$	f_k	$p_k=f_k/T$

Source: developed by the authors.

Let us consider the exponential membership function (4.1). To find estimates of parameters $\tau \equiv \tau_x$, $\nu \equiv \nu_x$ the following criterion will be used

$$\sum_{i=1}^{m-1}\left(\ln(-\ln \tilde{p}_i) - 2\ln\left(\frac{c-z_i}{\tau}\right)\right)^2 + \sum_{i=m+1}^{k}\left(\ln(-\ln \tilde{p}_i) - 2\ln\left(\frac{z_i-c}{\nu}\right)\right)^2, \tag{5.19}$$

where c denotes the midpoint of m-th class interval with maximum relative frequency $p_m = \max(p_1, p_2, \ldots, p_k)$, and $\tilde{p}_i, i = 1,2,\ldots,k$ are normalized frequencies for separate class intervals

$$\tilde{p}_i = \frac{p_i}{p_m},\ i = 1,2,\ldots,k. \tag{5.20}$$

It is worth noting that normalized frequencies (5.20) are included in the criterion (5.19) in order to find an exponential membership function of a fuzzy number similar to an empirical histogram.

The expressions (5.21) and (5.22) give the minimum of (5.19) with respect to the unknown parameters τ, ν (see Nasibov and Peker (2011) for more details). Thus, we have

$$\hat{\tau} = \exp\left(\frac{2\sum_{i=1}^{m-1}\ln(c-z_i)-\sum_{i=1}^{m-1}\ln(-\ln \tilde{p}_i)}{2(m-1)}\right), \tag{5.21}$$

$$\hat{\nu} = \exp\left(\frac{2\sum_{i=1}^{m-1}\ln(z_i-c)-\sum_{i=1}^{m-1}\ln(-\ln \tilde{p}_i)}{2(k-m)}\right). \tag{5.22}$$

Example 3. Let us consider the data aggregated in the frequency Table 2.

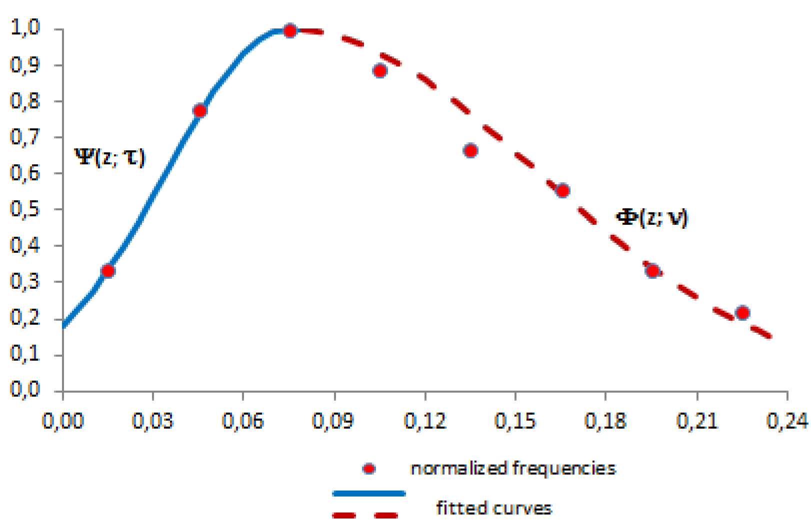
Table 2. Frequency table

Class intervals	Midpoints z_i	Frequencies f_i	Relative frequencies p_i	Normalized frequencies $\tilde{p}_i$
0.00 – 0.03	0.015	3	0.0698	0.3333
0.03 – 0.06	0.045	7	0.1628	0.7778
0.06 – 0.09	0.075	9	0.2093	1.0000
0.09 – 0.12	0.105	8	0.1860	0.8889
0.12 – 0.15	0.135	6	0.1395	0.6667
0.15 – 0.18	0.165	5	0.1163	0.5556
0.18 – 0.21	0.195	3	0.0698	0.3333
0.21 – 0.24	0.225	2	0.0465	0.2222

Source: developed by the authors.

The maximum relative frequency refers to the third class interval, thus we obtain m=3, $p_m = 0.2093$, and c=0.075. The membership function with $\hat{\nu}$, $\hat{\tau}$ derived from (5.21)–(5.22) is illustrated on Figure 3.

Figure 3. Normalized frequencies and a fitted membership function for $\hat{\tau} = 0.059$, $\hat{\nu} = 0.106$



Source: developed by the authors.

5.4. Evaluation of the proposed mortality model based on real data

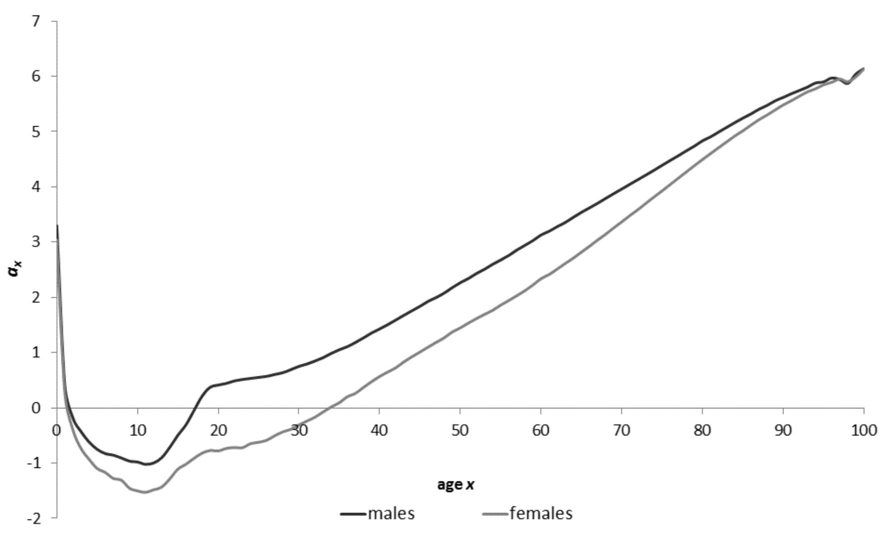
To illustrate theoretical discussions presented in the previous chapters dealing with the proposed mortality model based on quaternion-valued functions the estimates of model parameters will be calculated using real data to compare the *ex-post* forecasting errors with errors yielded by the standard Lee–Carter model (LC).

The analysis is based on the log-central death rates for males and females in Poland from the years 1958–2014. The necessary data were sourced from the Human Mortality Database (www.mortality.org) and from the GUS database (stat.gov.pl). The 2001–2014 death rates served the purpose of evaluating the models' forecasting properties and were not used in estimations.

Estimates a_x, b_x, k_t of the parameters of the quaternion mortality model (5.3) were obtained with the log-central death rates for males and females from the years 1958–2000. Parameters τ_x, ν_x were derived for each separate x using the Nasibov-Peker method, with $\{r_t, t = 1,2,\dots,T\}$ represented by standardized residuals from the ordinary least regression (3.3).

To ensure the clarity of data presentation, the parameter estimates are plotted as shown in Figures 4-6.

Figure 4. Parameters a_x, x = 0,1,...,100 estimated with the CNMM model for males and females

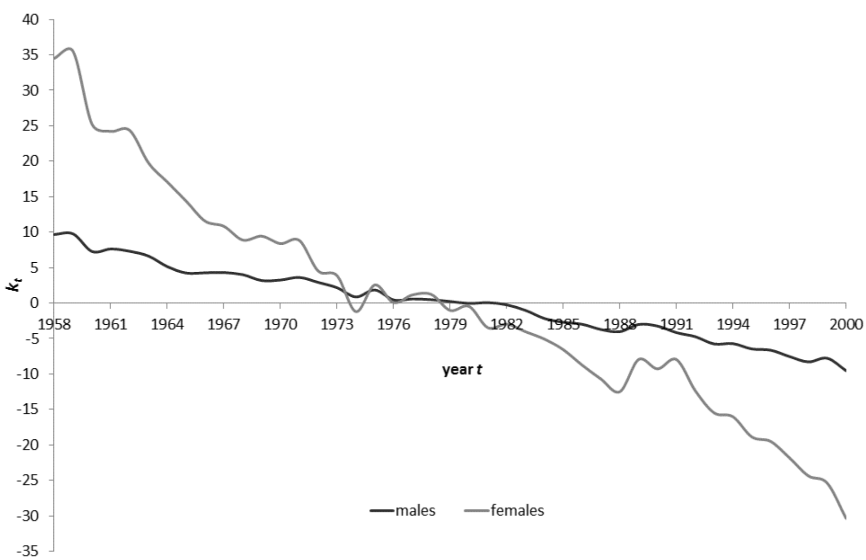


Source: developed by the authors.

Figure 5. Parameters b_x, x = 0.1,...,100 estimated with the CNMM model for males and females

Source: developed by the authors.

Figure 6. Parameters k_t, t = 1958,...,2000 estimated with model CNMM (males and females)



Source: developed by the authors.

The interpretation of the model parameters' estimates a_x, b_x, k_t is similar as in the standard Lee-Carter approach, meaning that a_x, x=0,1,...,X indicate the general shape of the mortality schedule, the time-varying parameters k_t, t=1,2,…,T represent the general mortality level, and b_x, x=0,1,…,X indicate the pattern of deviations from the age profile when the general level of mortality k_t changes.

The conclusion that can be drawn by comparing two curves plotted in Figure 4 is that average mortality in almost all age groups was higher for men than for women. Despite this fact the shapes of mortality profiles for both sexes seem rather similar, i.e. with a high mortality among children under two years of age, relatively low mortality for children aged 8–12, rising rapidly in the older age groups.

The arrangement of curves in Figure 5 shows that in some age groups the absolute values of b_x are higher for males than for females (i.e. for young or middle ages). It means that the log-central death rates clearly are more sensitive to the temporal changes in mortality for males than those noted for females. What is more, some negative values of b_x are estimated, i.e. for males at age group (34, 67) years. They indicate that male log-central mortality rates at those ages grew in some years of the period under consideration when declining at other ages in response to change of k_t. Figure 6 also shows that the overall mortality trend was generally declining, but at a varying rate. It is also worth noting that this general mortality trend (expressed by k_t) was faster in the subpopulation of women.

The forecasting properties of LC and CNMM models were compared based on the *ex-post* errors measured for each year in the period 2001–2014, i.e. the period which was omitted from parameter estimation. The *ex-post* errors were determined using crisp forecasts of log-central death rates (5.23). Two types of prediction accuracy measures will be used, i.e. a mean squared error (*MSE*) and a mean absolute deviation (*MAD*). The results are summarized in Table 3.

Table 3. Comparison of *ex-post* errors (*MSE* and *MAD*) for LC and CNMM models

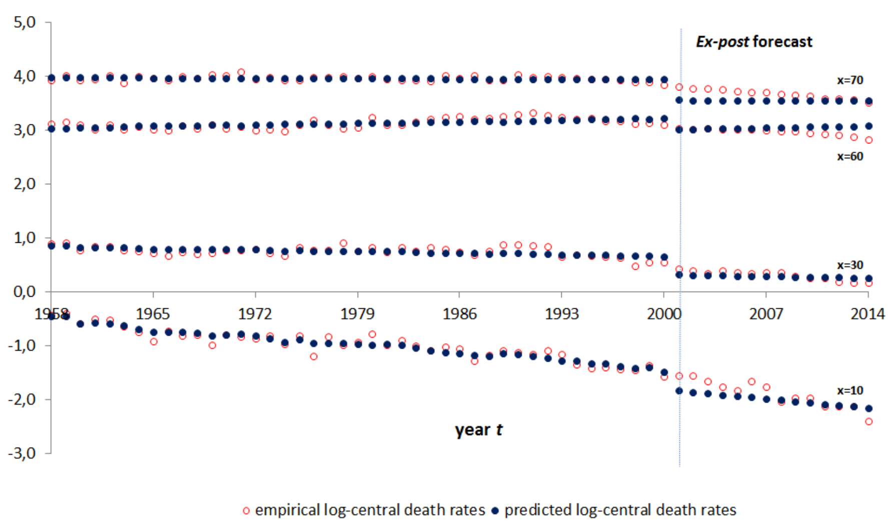
Year	Males				Females			
	MSE		*MAD*		*MSE*		*MAD*	
	LC	CNMM	LC	CNMM	LC	CNMM	LC	CNMM
2001	0.197	0.121	0.182	0.093	0.098	0.140	0.083	0.114
2002	0.204	0.119	0.185	0.091	0.122	0.120	0.107	0.096
2003	0.215	0.120	0.195	0.087	0.122	0.124	0.109	0.098
2004	0.223	0.111	0.206	0.081	0.132	0.113	0.117	0.089
2005	0.230	0.097	0.214	0.070	0.146	0.117	0.129	0.093
2006	0.232	0.110	0.214	0.081	0.152	0.105	0.130	0.083
2007	0.238	0.106	0.219	0.077	0.172	0.116	0.152	0.091
2008	0.257	0.107	0.234	0.083	0.174	0.111	0.156	0.086
2009	0.281	0.114	0.250	0.090	0.191	0.124	0.170	0.092
2010	0.330	0.137	0.302	0.110	0.190	0.095	0.167	0.072
2011	0.341	0.149	0.307	0.119	0.218	0.108	0.191	0.081
2012	0.373	0.174	0.335	0.137	0.215	0.105	0.185	0.081
2013	0.406	0.204	0.359	0.160	0.246	0.138	0.221	0.108
2014	0.469	0.257	0.430	0.212	0.273	0.148	0.245	0.117

Source: developed by the authors

It is worth noting that the CNMM model generates markedly smaller *ex-post* errors (in terms of *MSE* or *MAD* measures) than the LC model, which is visible especially for last years of prediction. For instance, for the prediction years 2010, 2011, 2012, 2013 and 2014 the *ex-post* errors obtained with the CNMM model are less than half of what was obtained with the LC model.

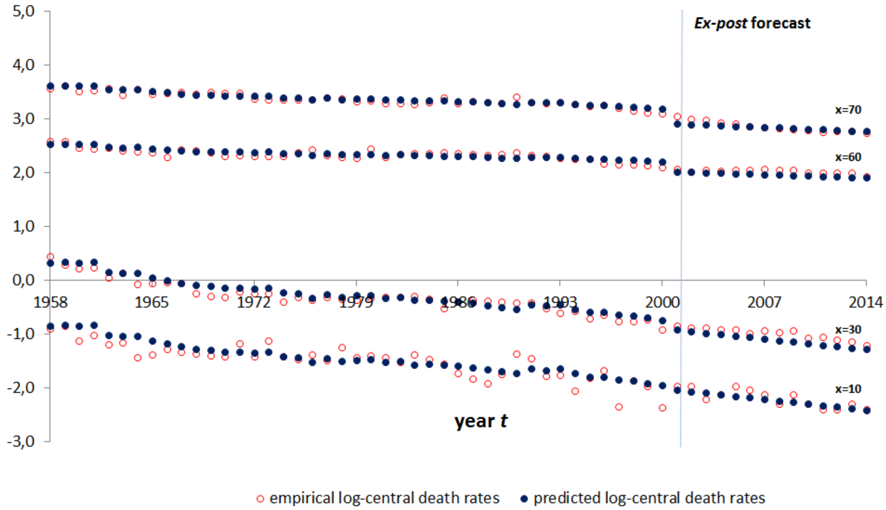
A comparison between empirical log-central death rates and those rates obtained from the CNMM model for some age groups is illustrated in Figures 7 and 8. It is worth noting that the models' parameters were estimated using the 1958–2000 data, therefore the log-central death rates estimated for the years 2001–2014 represent the *ex-post* forecasting.

Figure 7. Real and predicted log-central death rates for some age groups (males)



Source: developed by the authors.

Figure 8. Real and predicted log-central death rates for some age groups (females)



Source: developed by the authors.

6. Final remarks

We should explain to the reader why we have applied the exponential functions while building the theoretical function space as a basis of our new mortality model.

This approach has theoretical and practical advantages. Practical ones are delivered in the paper of Nasibov and Peker (2011), where an easy and useful fitting algorithm is proposed. Based on this algorithm it is possible to fit an exponential functions to the empirical distributions of the observed data, or – as in our case – to the normalized frequencies of residuals in the regression model.

The theoretical advantage of applying exponential membership functions lies in the desirable theoretical properties, because such functions can be transformed into the Hilbert spaces of quaternion valued functions. It is possible that other functions offer better fit to the observed data. This approach will be the subject of further research.

Acknowledgements

Both authors gratefully acknowledge that their research was supported by a grant from the National Science Centre under contract UMO-2015/17/B/HS4/00927.

REFERENCES

BOOTH, H., (2006). Demographic forecasting: 1980 to 2005 in review, International Journal of Forecasting, 22, pp. 547–581.

ISHIKAWA, S., (1997a). Fuzzy inferences by algebraic method, Fuzzy Sets and Systems, 87, pp. 181–200.

ISHIKAWA, S., (1997b). A quantum mechanical approach to a fuzzy theory, Fuzzy Sets and Systems, 90, pp. 277–306.

KOISSI, M.-C., SHAPIRO, A. F., (2006). Fuzzy formulation of the Lee-Carter model for mortality forecasting, Insurance: Mathematics and Economics, 39, pp. 287–309.

KOSIŃSKI, W., PROKOPOWICZ, P., ŚLĘZAK, D., (2003). Ordered Fuzzy Numbers, Bull. Polish Acad. Sci. Math., 51, pp. 327–338.

LEE R. D., CARTER, L., (1992). Modeling and forecasting the time series of U.S. mortality, Journal of the American Statistical Association, 87, pp. 659–671.

NASIBOV, E., PEKER, S., (2011), Exponential Membership Function Evaluation based on Frequency, Asian Journal of Mathematics and Statistics, 4 (1), pp. 8–20.

WILMOTH, J. R., HORIUCHI, S., (1999). Rectangularization revised: variability of age at death within human populations, Demography, 36 (4), pp. 475–495.

ROSSA, A., SOCHA, L., SZYMAŃSKI, A., (2015). Hybrydowe modelowanie umieralności za pomocą przełączających układów dynamicznych i modeli rozmytych, (in Polish), University of Lodz Press.

SZYMAŃSKI, A., ROSSA, A., (2014). Fuzzy mortality model based on Banach algebra, International Journal of Intelligent Technologies and Applied Statistics, 7, pp. 241–265.

TABEAU, E., BERG JETHS, A., HEATHCOTE, CH., (eds.), (2001). Forecasting of mortality in developed countries: insights form a statistical, demographic and epidemiological perspective, Kluwer Academic Publishers, London.