The Fuzzy Mortality Model based on Quaternion Theory

Andrzej Szymanski and Agnieszka Rossa
Institute of Statistics and Demography, University of Lodz, 41 Rewolucji 1905, Lodz, Poland

Keywords: Mortality Models, Fuzzy Numbers, Membership Functions, Quaternion Theory.

Abstract: The mortality models are of fundamental importance in many areas, such as the pension plans, the care of the elderly, the provision of health service, etc. In the paper, we propose a new class of mortality models based on a fuzzy version of the well-known Lee–Carter model (1992). Theoretical backgrounds are based on the algebraic approach to fuzzy numbers (Ishikawa, 1997, Kosiński, Prokopowicz, Ślęzak, 2003, Rossa, Socha, Szymański, 2015, Szymański, Rossa, 2014). The essential idea in our approach focuses on representing a membership function of a fuzzy number as an element of quaternion 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. Thus, the membership function $\mu(z)$ can be represented by means of a complex-valued function $f(u) + ig(u)$, where i is an imaginary unit. Then the pair (f, g) is a quaternion. The quaternion-valued, square integrable functions form a tool for constructing the new class of mortality models.

1 INTRODUCTION

Mortality models presented in the literature can be classified into two categories consisting of the so-called static or stationary models and dynamic models, respectively. Among models in the first group three main approaches can be identified: extrapolation (the well-known Lee-Carter model can serve as an example), expectation (e.g. some scenario models adopted by the United Nations) and explanation (e.g. the bit-string Penna model among others). The most common approaches are extrapolative methods which use a real or fuzzy variable functions of age and time to describe patterns and trends in death probabilities, specific mortality rates or other life-table measures. Models in the second group express the force of mortality in terms of stochastic differential equations.

The widely used Lee–Carter model (1992) 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 mortality models based on the fuzzy version of the standard Lee–Carter model.

The paper is organized as follows. In sections 2 and 3 both the Lee–Carter and Koissi–Shapiro models are presented. The theoretical backgrounds for a new class of mortality models are introduced in section 4. Two next sections describe model estimation and data fuzzification procedures. Section 7 is focused on the model evaluation. Especially, the *ex post* prediction accuracy based on the real data is studied and compared with analogous results obtained for the standard Lee–Carter model. The last section contains final remarks.

2 THE LEE–CARTER MODEL

One of the most popular mortality models is the Lee–Carter model (1992). Let $m_x(t)$ denote an age-specific death rate for the subset of a population that is between ages x and x+1 years

$$m_x(t) = \frac{D_x(t)}{L_x(t)}, \tag{2.1}$$

where:

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

$L_x(t)$ – the mid-year population alive at the age x in the year t,

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

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

Szymanski, A. and Rossa, A.
The Fuzzy Mortality Model based on Quaternion Theory.
DOI: 10.5220/0006813201570165
In *Proceedings of the 3rd International Conference on Complexity, Future Information Systems and Risk (COMPLEXIS 2018)*, pages 157-165
ISBN: 978-989-758-297-4

The measure $m_x(t)$ is the ratio of deaths between ages x and $x + 1$ years to the mid-year population alive at age x in the year t, also referred to as the mean population in the year t. Therefore, $m_x(t)$ is often described as the central rate because the mid-year population is used in the denominator.

The Lee–Carter model (LC) can be written as

$$\ln m_x(t) = \alpha_x + \beta_x \kappa_t + \epsilon_{xt}, \quad (2.2)$$

or, equivalently, as

$$m_x(t) = \exp\{\alpha_x + \beta_x \kappa_t + \epsilon_{xt}\}, \quad (2.3)$$

where $m_x(t)$ are age-specific mortality rates, α_x, β_x, κ_t are some model parameters, of which α_x, β_x $(x = 0,1,\ldots,X)$ represent age-related effects and κ_t $(t = 1,2,\ldots,T)$ – time-related effects. The double-indexed terms ϵ_{xt} are random components, assumed to be independent and have the same normal distribution with an expected value of 0 and constant variance.

The system of equations (2.2) or (2.3) cannot be explicitly solved unless normalizing constraints are imposed. For full model identification, it is assumed that the sum of parameters β_x is 1 and the sum of parameters κ_t is equal to 0, i.e.

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

The age-related effects α_x indicate the age profile of mortality, the effects κ_t represent the general mortality trend, whereas β_x indicate the pattern of deviations from the age profile in response to change of the general trend κ_t. It is worth noting that 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, parameters β_x tell which age-specific rates $m_x(t)$ decline (or rise) fast and which slow in response to change of κ_t.

The method of parameter estimation proposed by Lee and Carter (1992) is based on the so-called SVD method (Singular Value Decomposition), further developed by Wilmoth (1993) as weighted SVD.

Parameters α_x and β_x do not depend on time t, which means that once derived can also be used for future periods $t > T$. The time-varying effects are κ_t. They can be predicted for $t > T$ using, for instance, the time series analysis.

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

$$\kappa_t = \delta + \kappa_{t-1} + \xi_t, \quad (2.5)$$

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

With predicted values of κ_t for $t > T$ and estimates of α_x and β_x the central death rates $m_x(t)$ can be easily forecasted from (2.2) or (2.3), and subsequently other life-table measures.

3 THE KOISSI–SHAPIRO MODEL

One of the most interesting modifications 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 assumes fuzzy representation of log-central death rates as well as model parameters. It allows taking account of uncertainty involved in mortality rates and including a random term into the fuzzy structure of the model.

The Koissi–Shapiro approach builds on the assumption that the observed mortality rates $m_x(t)$ are in fact not exactly known, as they are subject to reporting errors of several kinds. They may be reported for incorrect year, area, age or assigned statistics that are incorrect, etc. For these reasons, fuzzy representation of central death rates seems to be justified.

Koissi and Shapiro proposed the fuzzification procedure (see Koissi, Shapiro, 2006 for details) to convert the log-central age-specific mortality rates $\ln m_x(t)$ into symmetric, triangular fuzzy numbers

$$Y_{xt} = (y_{xy}, e_{xt}), \quad (3.1)$$

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

Next, they defined the fuzzy version of the Lee–Carter model in the following form

$$Y_{xt} = A_x \oplus (B_x \otimes K_t), \quad (3.2)$$

where Y_{xt} are fuzzified log-central mortality rates, A_x, B_x, K_t represent symmetric, triangular fuzzy numbers, playing an analogous role as parameters in the model (2.2), and $\oplus$, $\otimes$ are addition and multiplication operators of fuzzy numbers in the norm T_w (see Koissi, Shapiro, 2006 for the definition of T_w).

The authors assumed that the model parameters can be estimated by minimizing a criterion function based on the Diamond distance measure between fuzzy variables. However, this estimation method poses major problems in the optimization algorithm, because expression

$$\max\{s_{A_x}, |b_x| s_{K_t}, |k_t| s_{B_x}\}, \tag{3.3}$$

appearing in this criterion function prevents the use of standard non-linear optimization methods.

In the rest of the paper, a modification of the fuzzy mortality model using fuzzified mortality rates with exponential membership functions will be proposed. The essential idea in this approach is representing exponential membership functions of fuzzy numbers as elements of the quaternion algebra. It simplifies both operations on fuzzy numbers and model estimation.

4 THE NEW CLASS OF MORTALITY MODELS

4.1 Transformation of Exponential Membership Functions Into Complex-valued Functions

Data fuzzification 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 model in order to fuzzify the data. We will adopt exponential membership functions of fuzzy numbers adjusted to relative frequencies of residuals in the least-squares regression.

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, i.e. strictly increasing and strictly decreasing functions $\Psi(z)$ and $\Phi(z)$ of the form

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

Then there exist inverse functions

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

where $\psi(u)$ and $\varphi(u)$ for $u \in [0,1]$ are expressed as follows

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

Let us consider two complex functions

$$f(u) = c + i\psi(u), \ g(u) = c + i\varphi(u), \tag{4.5}$$

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

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

$$f(u) = c - i\tau(-\ln u)^{\frac{1}{2}}, \tag{4.6}$$
$$g(u) = c + i\nu(-\ln u)^{\frac{1}{2}}. \tag{4.7}$$

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

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

$$|f(u)|^2 = c^2 + \tau^2(-\ln u), \tag{4.8}$$
$$|g(u)|^2 = c^2 + \nu^2(-\ln u). \tag{4.9}$$

After integration both sides of (4.8) and (4.9) on [0,1] we find

$$\int_0^1 |f(u)|^2 du = c^2 + \tau^2 < \infty, \tag{4.10}$$
$$\int_0^1 |g(u)|^2 du = c^2 + \nu^2 < \infty. \tag{4.11}$$

4.2 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 (z, w) of complex functions $z = a + ib$, $w = c + id$ where $i = \sqrt{-1}$ is an imaginary unit.

The algebra of quaternions is often denoted as $\boldsymbol{H}$. Quaternions were first described by 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

$$(z, w) + (u, x) = (z + u, w + x). \tag{4.12}$$

The multiplication of an element of $\boldsymbol{H}$ by a real number α is defined as the product of both components and the scalar α

$$\alpha(z, w) = (\alpha z, \alpha w). \tag{4.13}$$

To define the product of two elements in $\boldsymbol{H}$ a choice of a basis for $\boldsymbol{R}^4$ is needed. The elements of this basis are usually 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 multiplying by 1 does not change the value and any element of $\boldsymbol{H}$ can be uniquely written as

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

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

Multiplication of quaternions could be defined in the form

$$(z, w)(u, x) = (zu - w\bar{x}, zx + w\bar{u}), \qquad (4.15)$$

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. We have for example

$$(i, 0)(0,1) = (0, i), \qquad (4.16)$$

but

$$(0,1)(i, 0) = (0, -i). \qquad (4.17)$$

The norm of a quaternion q is denoted as $\|q\|$ and may be expressed as follows

$$\|q\| = \sqrt{a^2 + b^2 + c^2 + d^2}. \qquad (4.18)$$

The norm 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 q by a real number α scales its norm by the absolute value of this number

$$\|\alpha q\| = |\alpha| \|q\|. \qquad (4.19)$$

The norm (4.18) allows to define a distance $d(p, q)$ between p and q as the norm of their difference

$$d(p, q) = \|p - q\|. \qquad (4.20)$$

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

According to (4.5) we have

$$f(u) = c + i\psi(u), \quad u \in [0,1] \qquad (4.21)$$

and

$$g(u) = c + i\varphi(u), \quad u \in [0,1], \qquad (4.22)$$

where ψ, φ are defined in (4.4).

Let us denote

$$P(u) = \big(f(u), g(u)\big), \quad u \in [0,1]. \qquad (4.23)$$

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

$$\|P(u)\|^2 = |f(u)|^2 + |g(u)|^2. \qquad (4.24)$$

By integrating both sides of (4.24) and taking into account results (4.10) and (4.11), we receive

$$\int_0^1 \|P(u)\|^2 du < \infty. \qquad (4.25)$$

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

4.3 A Mortality Model Based on Quaternion-valued Functions

Let us assume that $\tilde{Y}_{x,t} = \big(f_{Y_{x,t}}, g_{Y_{x,t}}\big)$ are quaternions with complex functions $f_{Y_{x,t}}$, $g_{Y_{x,t}}$ defined as follows

$$f_{Y_{x,t}}(u) = y_{xt} - i\tau_x(-\ln u)^{\frac{1}{2}}, \qquad (4.26)$$

$$g_{Y_{x,t}}(u) = y_{xt} + i\nu_x(-\ln u)^{\frac{1}{2}}, \qquad (4.27)$$

where $u \in [0,1]$, $y_{xt} = \ln m_x(t)$, $i = \sqrt{-1}$ is an imaginary unit, and τ_x, ν_x are some known parameters determined by means of the Nasibov–Peker method (see section 6 for more details).

We will also assume that $\tilde{A}_x = \big(f_{A_x}, g_{A_x}\big)$, $\tilde{K}_t = \big(f_{K_t}, g_{K_t}\big)$ are quaternions determined by the following complex functions defined for $u \in [0,1]$

$$f_{A_x}(u) = a_x - i(-\ln u)^{\frac{1}{2}} s_{A_x}^L, \qquad (4.28)$$

$$g_{A_x}(u) = a_x + i(-\ln u)^{\frac{1}{2}} s_{A_x}^R, \qquad (4.29)$$

$$f_{K_t}(u) = k_t - i(-\ln u)^{\frac{1}{2}} s_{K_t}, \qquad (4.30)$$

$$g_{K_t}(u) = k_t + i(-\ln u)^{\frac{1}{2}} s_{K_t}. \qquad (4.31)$$

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

$$\tilde{Y}_{x,t} = \tilde{A}_x + b_x \tilde{K}_t, \qquad (4.32)$$

where $Y_{x,t}$ are fuzzified log-central mortality rates expressed in terms of quaternion-valued functions in the space $L_2(\boldsymbol{H})$, $b_x \in \boldsymbol{R}$ for $x = 0,1, \dots, X$ is a set of scalar parameters, $\tilde{A}_x$, $\tilde{K}_t$ are quaternions in $L_2(\boldsymbol{H})$ determined by the complex functions (4.28) – (4.31). The proposed model (4.32) will be termed Complex Number Mortality Model (CNMM).

Note that the quaternions $\tilde{A}_x = \big(f_{A_x}, g_{A_x}\big)$, $\tilde{K}_t = \big(f_{K_t}, g_{K_t}\big)$ on the right-hand side of (4.32) reflect also some fuzzy numbers A_x, K_t with exponential membership functions $\mu_{A_x}(z)$ and $\mu_{K_t}(z)$ as follows

$$\mu_{A_x}(z) = \begin{cases} \exp\left\{-\left(\frac{a_x - z}{s_{A_x}^L}\right)^2\right\} & \text{for } z \le a_x, \\ \exp\left\{-\left(\frac{z - a_x}{s_{A_x}^R}\right)^2\right\} & \text{for } z > a_x, \end{cases} \qquad (4.33)$$

$$\mu_{K_t}(z)=\begin{cases}\exp\left\{-\left(\frac{k_t-z}{s_{K_t}}\right)^2\right\} & \text{for } z\le k_t,\\ \exp\left\{-\left(\frac{z-k_t}{s_{K_t}}\right)^2\right\} & \text{for } z> k_t.\end{cases}\qquad(4.34)$$

Using properties (4.12) and (4.13) the complex functions defining the quaternion $A_x+b_xK_t$ on the right-hand side of (4.32) are as follows

$$f_{A_x+b_xK_t}(u)=a_x+b_xk_t-i(-\ln u)^{\frac{1}{2}}\left(s_{A_x}^L+b_xs_{K_t}\right),$$

$$g_{A_x+b_xK_t}(u)=a_x+b_xk_t+i(-\ln u)^{\frac{1}{2}}\left(s_{A_x}^R+b_xs_{K_t}\right).$$

It means that $\tilde{A}_x+b_x\tilde{K}_t$ reflects a fuzzy number W_{xt} with an exponential membership function $\mu(z)$ equal to

$$\mu(z)=\begin{cases}\exp\left\{-\left(\frac{a_x+b_xk_t-z}{s_{A_x}^L+b_xs_{K_t}}\right)^2\right\}, & z\le a_x+b_xk_t\\ \exp\left\{-\left(\frac{z-a_x-b_xk_t}{s_{A_x}^R+b_xs_{K_t}}\right)^2\right\}, & z> a_x+b_xk_t\end{cases}\qquad(4.35)$$

5 ESTIMATION OF THE MODEL PARAMETERS

In order to estimate the parameters a_x, b_x, k_t and $s_{A_x}^L, s_{A_x}^R, s_{K_t}$ appearing in (4.35) we will use the norm (4.24) defined in the space of quaternion-valued functions. Thus, the following distance between left- and right-hand sides of the model (4.32) will be defined for some fixed x and t

$$d_{x,t}=\int_0^1\left\|\tilde{Y}_{x,t}(u)-\left(\tilde{A}_x(u)+b_x\tilde{K}_t(u)\right)\right\|^2du.\qquad(5.1)$$

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

$$f_{Y_{x,t}-(A_x+b_xK_t)}(u)=y_{x,t}-(a_x+b_xk_t)-i(-\ln u)^{\frac{1}{2}}\left(\tau_x-s_{A_x}^L-b_xs_{K_t}\right),\qquad(5.2)$$

$$g_{Y_{x,t}-(A_x+b_xK_t)}(u)=y_{x,t}-(a_x+b_xk_t)+i(-\ln u)^{\frac{1}{2}}\left(\nu_x-s_{A_x}^R-b_xs_{K_t}\right).\qquad(5.3)$$

Hence,

$$\left|f_{Y_{x,t}-(A_x+b_xK_t)}(u)\right|^2=\left(y_{x,t}-(a_x+b_xk_t)\right)^2+(-\ln u)\left(\tau_x-s_{A_x}^L-b_xs_{K_t}\right)^2,\qquad(5.4)$$

$$\left|g_{Y_{x,t}-(A_x+b_xK_t)}(u)\right|^2=\left(y_{x,t}-(a_x+b_xk_t)\right)^2+(-\ln u)\left(\nu_x-s_{A_x}^R-b_xs_{K_t}\right)^2.\qquad(5.5)$$

After integration both sides of (5.4) – (5.5) on the interval [0,1] we receive

$$d_{x,t}=2\left(y_{x,t}-(a_x+b_xk_t)\right)^2+\left(\tau_x-s_{A_x}^L+-b_xs_{K_t}\right)^2+\left(\nu_x-s_{A_x}^R-b_xs_{K_t}\right)^2.\qquad(5.6)$$

By analogy to the Lee–Carter model and constraints (2.4) we will assume that

$$\sum_{x=0}^X b_x=1,\quad \sum_{t=1}^T k_t=0.\qquad(5.7)$$

Moreover, the additional restriction will be also imposed

$$\sum_{t=1}^T s_{K_t}=(X+1)\sqrt{\sum_{t=1}^T(\bar{y}_t-\bar{y})^2},\qquad(5.8)$$

where

$$\bar{y}_t=\frac{1}{X+1}\sum_{x=0}^X y_{xt},\ \bar{y}=\frac{1}{T(X+1)}\sum_{t=1}^T\sum_{x=0}^X y_{xt}.\qquad(5.9)$$

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

$$F=\sum_{x=0}^X\sum_{t=1}^T d_{x,t}+\lambda_1\left(\sum_{x=0}^X b_x-1\right)+\lambda_2\sum_{t=1}^T k_t+\lambda_3\left(\sum_{t=1}^T s_{K_t}-(X+1)\sqrt{\sum_{t=1}^T(\bar{y}_t-\bar{y})^2}\right),\qquad(5.10)$$

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

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

$$\begin{cases}a_x=\frac{1}{T}\sum_{t=1}^T y_{xt}=\bar{y}_x\\ b_x=\frac{\sum_{t=1}^T\left[2k_t(y_{xt}-a_x)+s_{K_t}\left(\tau_x+\nu_x-s_{A_x}^L-s_{A_x}^R\right)\right]-\frac{\lambda_1}{2}}{2\sum_{t=1}^T\left(k_t^2+s_{K_t}^2\right)}\\ k_t=\frac{\sum_{x=0}^X b_x(y_{xt}-a_x)-\frac{\lambda_2}{4}}{\sum_{x=0}^X b_x^2}\\ s_{A_x}^L=\tau_x-\frac{1}{T}b_x\sum_{t=1}^T s_{K_t}\\ s_{A_x}^R=\nu_x-\frac{1}{T}b_x\sum_{t=1}^T s_{K_t}\\ s_{K_t}=\frac{\sum_{x=0}^X b_x\left(\tau_x+\nu_x-s_{A_x}^L-s_{A_x}^R\right)-\frac{\lambda_3}{2}}{2\sum_{x=0}^X b_x^2}\\ \sum_{x=1}^X b_x=1\\ \sum_{t=1}^T k_t=0\\ \sum_{t=1}^T s_{K_t}-(X+1)\sqrt{\sum_{t=1}^T(\bar{y}_t-\bar{y})^2}=0\end{cases}\qquad(5.11)$$

Note, that the last three equations satisfy restrictions (5.7) and (5.8).

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 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 parameters k_t for future periods $t > T$.

Next, functions (5.2) and (5.3) should be determined using estimates of $a_x, b_x, s_{A_x}^L, s_{A_x}^R, s_{K_t}$ and the sequence of predicted time-related parameters k_t for $t > T$. Note, that the functions (5.2) and (5.3) define the right-hand side of the mortality model (4.32) for $t > T$, i.e. they determine quaternions $\tilde{A}_x + b_x\tilde{K}_t$ for future periods. Finally, the quaternions $\tilde{A}_x + b_x\tilde{K}_t$ can be transformed into fuzzy numbers W_{xt} using exponential membership function $\mu(z)$ defined in (4.35). Moreover, W_{xt} can be defuzzified into crisp numbers w_{xt}, if necessary, using i.e. the centroid defuzzification formula

$$w_{xt} = \frac{\sum_{z=\epsilon}^{1} z \cdot \mu_{W_{xt}}(z)}{\sum_{z=\epsilon}^{1} \mu_{W_{xt}}(z)}, \qquad (5.12)$$

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

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

6 DATA FUZZIFICATION

We propose fuzzification of the log-central death rates $\ln m_x(t)$ for each $x = 0,1,\dots,X,\ t = 1,2,\dots,T$ using the method proposed by Nasibov and Peker (2011) which leads to determine parameters τ_x, ν_x for a fixed x based on an empirical distribution of a sequence of data. Main results of their work are introduced in this section.

Let us assume that $\{r_t, t = 1,2,\dots,T\}$ 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.

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 function will be used

$$Q = \sum_{i=1}^{m-1}\left(\ln(-\ln\tilde{p}_i) - 2\ln\left(\tfrac{c-z_i}{\tau}\right)\right)^2 + \sum_{i=m+1}^{k}\left(\ln(-\ln\tilde{p}_i) - 2\ln\left(\tfrac{z_i-c}{\nu}\right)\right)^2, \qquad (6.1)$$

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

$$\tilde{p}_i = \frac{p_i}{p_m},\ i = 1,2,\dots,k. \qquad (6.2)$$

It is worth noting that normalized frequencies (6.2) are included in the criterion (6.1) in order to find an exponential membership function of a fuzzy number by analogy to an empirical histogram.

The next two expressions (6.3) and (6.4) give the minimum of (6.1) with respect to the unknown parameters τ, ν (see Nasibov, Peker, 2011 for more details). We have

$$\tau = \exp\left(\frac{2\sum_{i=1}^{m-1}\ln(c-z_i) - \sum_{i=1}^{m-1}\ln(-\ln\tilde{p}_i)}{2(m-1)}\right), \qquad (6.3)$$

$$\nu = \exp\left(\frac{2\sum_{i=1}^{m-1}\ln(z_i-c) - \sum_{i=1}^{m-1}\ln(-\ln\tilde{p}_i)}{2(k-m)}\right). \qquad (6.4)$$

7 MODEL EVALUATION

To illustrate theoretical discussions presented in the previous sectors dealing with the mortality model CNMM based on quaternion-valued functions the estimates of model parameters will be derived using the real mortality data set 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 mortality rates served the purpose of evaluating the *ex-post* forecasting accuracy, therefore were not used in estimations.

Estimates of the parameters a_x, b_x, k_t in the mortality model (4.32) were obtained with the log-central age-specific mortality rates registered 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 described in previous section, with $\{r_t, t = 1,2,\dots,T\}$ represented by standardized residuals from the ordinary least squares regression. To ensure the clarity of data presentation, the parameter estimates are plotted as Figures 1–3.

Interpretation of the model parameters' estimates is similar as in the standard Lee–Carter approach, meaning that estimates of a_x, $x = 0,1,\dots,X$ indicate the overall shape of the mortality schedule, the estimates of the time-varying parameters k_t, $t = 1,2,\dots,T$ approximate the general mortality trend and b_x, $x = 0,1,\dots,X$ indicate the pattern of deviations from the general age profile.

The conclusion that can be drawn by comparing two curves plotted in Figure 1 is that average mortality in almost all age groups was higher for men than for women. Despite of 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 years, rising rapidly in the older age groups.

The arrangement of curves in Figure 2 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 changes of k_t for males than those noted for females. What is more, some negative values of b_x are found, i.e. for males at age group (34, 67) years. They indicate that male mortality rates at those ages grew in some years when declining at other ages.

Figure 3 shows that the overall mortality trend was generally declining, but at a varying pace. It is also worth noting that this general mortality trend (expressed by k_t) was faster in the population of females.

The *ex-post* errors for the CNMM model were determined using crisp forecasts of log-central death rates (5.12). Two types of prediction accuracy measures were used, i.e. a mean squared error (MSE) and a mean absolute deviation (MAD). The results obtained indicate that the CNMM model generates markedly smaller *ex-post* errors in terms of MSE or MAD measures than the LC model. For instance, for the prediction period 2010-2014 the *ex-post* errors obtained with the CNMM model were less than half of what was obtained with the LC model.

8 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 function to the empirical distribution of the observed data, or – as in our case – to the normalized frequencies of residuals in the regression model.

The theoretical advantage of applying such membership functions are lying also in the desirable theoretical properties, because exponential functions can be transformed into the Hilbert spaces of quaternion valued functions.

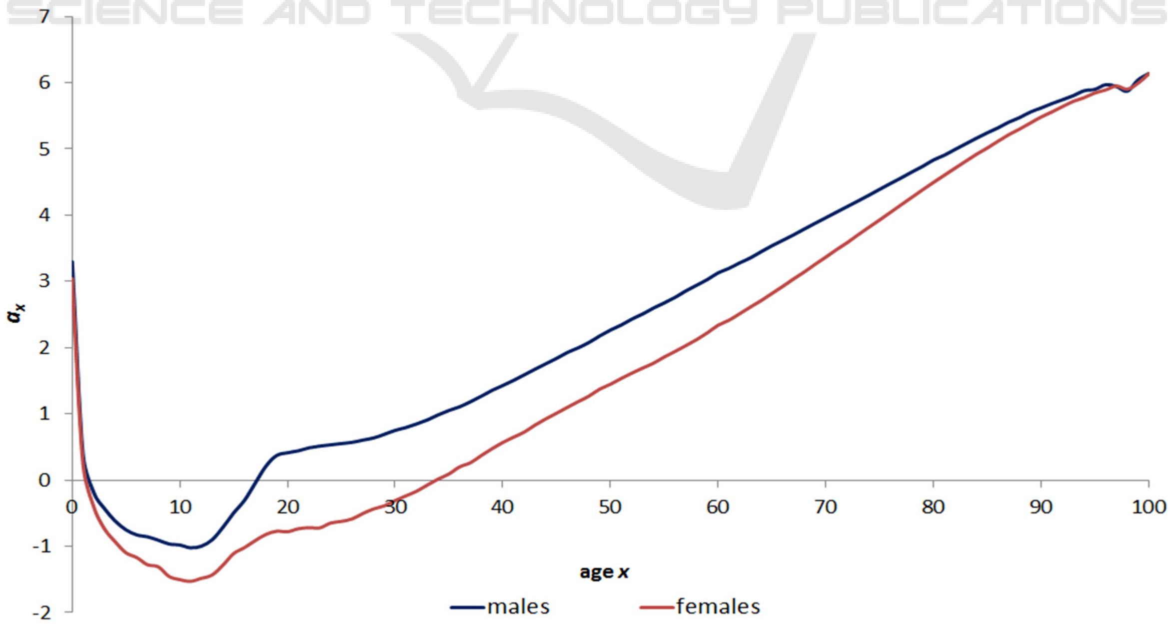

Figure 1: Parameters a_x, x = 0,1,...,100 estimated with the CNMM model (males and females).

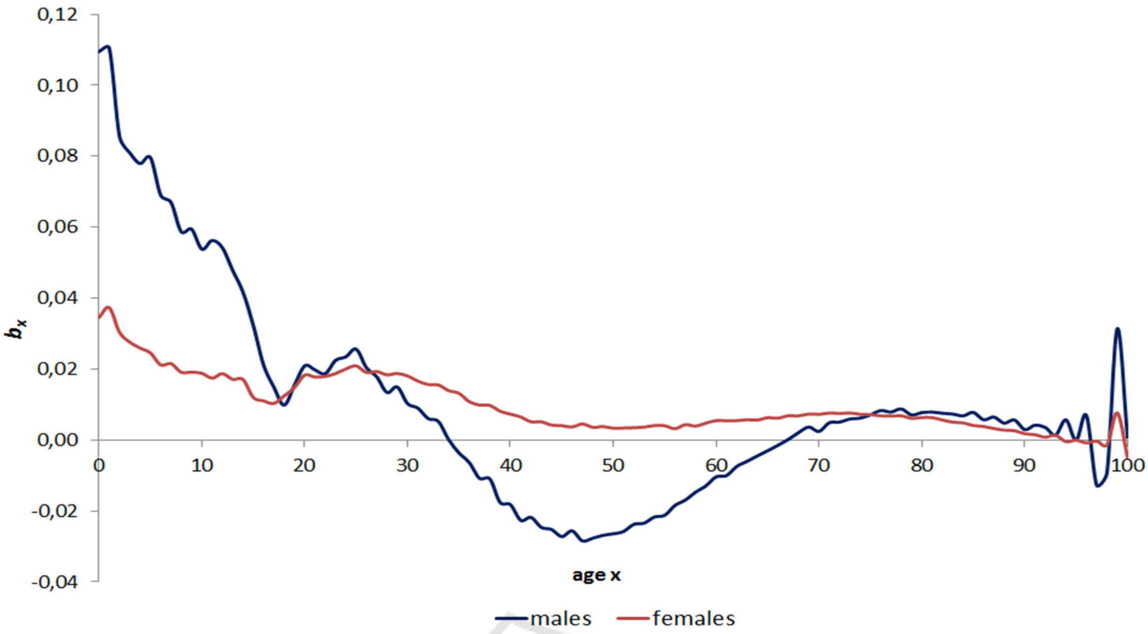

Figure 2: Parameters b_x, x = 0.1,...,100 estimated with the CNMM model (males and females).

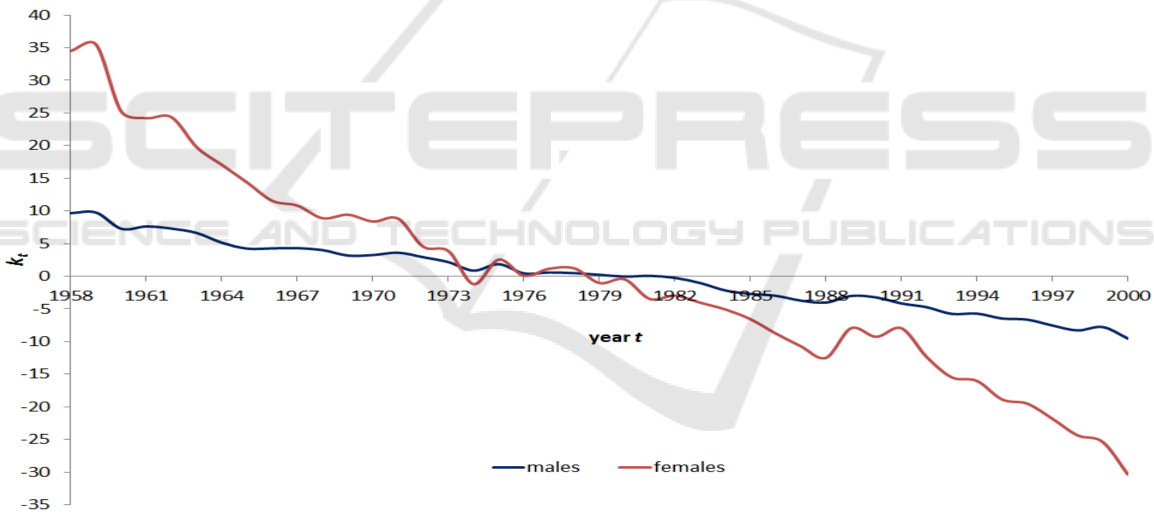

Figure 3: Parameters k_t, t = 1958,...,2000 estimated with the CNMM model (males and females).

Probably, other functions offer better fit to the observed data but this subject will be considered in our further research.

ACKNOWLEDGEMENTS

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

REFERENCES

Ishikawa, S., 1997. Fuzzy inferences by algebraic method. In *Fuzzy Sets and Systems*, 87, pp. 181–200.

Koissi, M.-C., Shapiro, A. F., 2006. Fuzzy formulation of the Lee-Carter model for mortality forecasting. In *Insurance: Mathematics and Economics*, 39, pp. 287–309.

Kosiński, W., Prokopowicz, P., Ślęzak, D., 2003. Ordered Fuzzy Numbers. In *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. In *Journal of the American Statistical Association*, 87, pp. 659-671.

Nasibov, E., Peker, S., 2011, Exponential Membership Function Evaluation based on Frequency. In *Asian Journal of Mathematics and Statistics*, 4(1), pp. 8-20.

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. In *International Journal of Intelligent Technologies and Applied Statistics*, 7, pp. 241–265.

Wilmoth, J. R., 1993. Computational methods for fitting and extrapolating the Lee–Carter model of mortality change. In *Technical Report*, University of California.