

# A Comparison of Generalized Stochastic Milevsky-Promislov Mortality Models with continuous non-Gaussian Filters

Śliwka Piotr<sup>1</sup>[0000-0001-6226-9580] and Socha Lesław<sup>1</sup>[0000-0003-0701-9700]

Department of Mathematics and Natural Sciences, Cardinal Stefan Wyszyński University, Woycickiego 1/3, 01-938 Warsaw, Poland, email: [p.sliwka@uksw.edu.pl](mailto:p.sliwka@uksw.edu.pl) and [l.socha@uksw.edu.pl](mailto:l.socha@uksw.edu.pl)

**Abstract.** The ability to precisely model mortality rates  $\mu_{x,t}$  plays an important role from the economic point of view in healthcare. The aim of this article is to propose a comparison of the estimation of the mortality rates based on a class of stochastic Milevsky-Promislov mortality models. We assume that excitations are modeled by second, fourth and sixth order polynomials of outputs from a linear non-Gaussian filter. To estimate the model parameters we use the first and second moments of  $\mu_{x,t}$ . The theoretical values obtained in both cases were compared with theoretical  $\widehat{\mu_{x,t}}$  based on a classical Lee-Carter model. The obtained results confirm the usefulness of the switched model based on the continuous non-Gaussian processes used for modeling  $\mu_{x,t}$ .

**Keywords:** forecasting of mortality rates · hybrid mortality models · switched models · Lee-Carter model · Ito stochastic differential equations.

## 1 Introduction

The determination of the mortality models is one of the basic problems not only in the field of life insurance but recently particularly in the economics of healthcare ([1],[3]-[5]). Currently, the most frequently used model is Lee-Carter ([9], [11], [12], [16], [17]) not only the change in mortality associated with age  $x$  and calendar year  $t$  but also takes into account the influence of belonging to a particular generation (cohort effect) and takes the form:  $\ln(\mu_{x,t}) = \alpha_x + \beta_x k_t + \varepsilon_{x,t}$ . The assumption that the estimated  $a_x$  and  $b_x$  are fixed at time  $t$  causes a wave of criticism, especially from the point of view of forecasting. Therefore, there is a need to look for other methods predicting mortality rates that take into account the variability of parameters over time. One of these propositions may be the approach recently proposed in Rossa and Socha [18], Rossa et al. [19], Sliwka and Socha [20], Sliwka ([21],[22]), and based on the Milevsky-Promislov family of models [15] with extensions ([2],[7],[13]). Methods of modeling  $\mu_{x,t}$  taking into account the causes of death have been characterized in article [23].

Works on switchings model, which consist of several subsystems with the same structure and different parameters and which can switch over time according to an unknown switching rule, have been taken, among others in Sliwka

and Socha [20]. In the mentioned paper it was shown that modeling of empirical mortality coefficients  $\mu_{x,t}$  using the non-Gaussian linear scalar filters model second order with switchings (nGLSFo2s) allows a more precise estimate of  $\mu_{x,t}$  than using the Gaussian linear scalar filters model with switchings (GLSFs) and Lee-Carter with switchings (LCs) model for some fixed ages  $x$ .

In this paper we first propose three extended Milevsky and Promislov models with continuous non-Gaussian filters. We assume that excitations are modeled not only by the second but also by the fourth (nGLSFo4) and the sixth order polynomials (nGLSFo6) of outputs from a linear nGLSF. To estimate the model parameters we use the first and second moments of mortality rates. We show that in considered models some of the parameters can be estimated. Next, we use these models to create hybrid models, where submodels have the same structure and possible different parameters. To estimate the model parameters we use the first and second moments of mortality rates. According to our knowledge, the mortality models proposed above, their hybrid versions and methods for estimating their parameters and switching are new in the field of life insurance.

The paper is organized as follows. In Section 2 basic notations and definitions of stochastic hybrid systems are entered. Three new basic models represented by even-order polynomials of outputs from linear Gaussian filter are introduced and the non-stationary solutions of corresponding moment equations are presented in Section 3. The derivation of these non-stationary solutions are derived in Appendix. In Section 4 the procedure of the parameters estimation and determination of switching points is presented. Based on the adapted numerical algorithm of a nonlinear minimization problem, parameter estimation is performed. In Section 5 we have compared empirical mortality rates with theoretical ones obtained from proposed models as well as from standard LC model in two versions with switchings and without switchings. The last Section summarizes the obtained results.

## 2 Mathematical preliminaries

Throughout this paper we use the following notation. Let  $|\cdot|$  and  $\langle \cdot \rangle$  be the Euclidean norm and the inner product in  $\mathbb{R}^n$ , respectively. We mark  $\mathbb{R}_+ = [0, \infty)$ ,  $\mathbb{T} = [t_0, \infty)$ ,  $t_0 \geq 0$ . Let  $\Xi = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$  be a complete probability space with a filtration  $\{\mathcal{F}_t\}_{t \geq 0}$  satisfying usual conditions. Let  $\sigma(t) : \mathbb{R}_+ \rightarrow \mathbb{S}$  be the switching rule, where  $\mathbb{S} = \{1, \dots, N\}$  is the set of states. We denote switching times as  $\tau_1, \tau_2, \dots$  and assume that there is a finite number of switches on every finite time interval. Let  $W_k(t)$  be the independent Brownian motions. We assume that processes  $W_k(t)$  and  $\sigma(t)$  are both  $\{\mathcal{F}_t\}_{t \geq 0}$  adapted.

By the stochastic hybrid system we call the vector Itô stochastic differential equations with a switching rule described by

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), t, \sigma(t))dt + \mathbf{g}(\mathbf{x}(t), t, \sigma(t))dW(t), \quad (\sigma(t_0), x(t_0)) = (\sigma_0, x_0), \quad (1)$$

where  $x \in R^n$  is the state vector,  $(\sigma_0, x_0)$  is an initial condition,  $t \in T$  and  $M$  is a number of Brownian motions.  $f(x(t), t, \sigma(t))$  and  $g(x(t), t, \sigma(t))$  are defined by sets of  $f(x(t), t, l)$  and  $g(x(t), t, l)$ , respectively i.e.

$$f(x(t), t, \sigma(t)) = f(x(t), t, l), g(x(t), t, \sigma(t)) = g(x(t), t, l) \text{ for } \sigma(t) = l.$$

Functions  $f : R^n \times T \times S \rightarrow R^n$  and  $g : R^n \times T \times S \rightarrow R^n$  are locally Lipschitz and such that  $\forall l \in S, t \in T, f(\mathbf{0}, t, l) = g(\mathbf{0}, t, l) = \mathbf{0}, k = 1, \dots, M$ . These conditions together with these enforced on the switching rule  $\sigma$  ensure that there exists a unique solution to the hybrid system (1).

Hence it follows that equation (1) can be treated as a family (set) of subsystems defined by

$$d\mathbf{x}(t, l) = \mathbf{f}(\mathbf{x}(t), t, l)dt + \sum_{k=1}^M \mathbf{g}_k(\mathbf{x}(t), t, l)dW_k(t), \quad l \in \mathbb{S} \quad (2)$$

where  $\mathbf{x}(t, l) \in \mathbb{R}^n$  is the state vector of  $l$ - subsystem.

We assume additionally that the trajectories of the hybrid system are continuous. It means, when the stochastic system is switched from  $l_1$  subsystem to  $l_2$  subsystem in the moment  $\tau_j$ , then

$$\mathbf{x}(\tau_j, l_1) = \mathbf{x}(\tau_j, l_2), \quad l_1, l_2 \in \mathbb{S}. \quad (3)$$

### 3 Models with continuous non-Gaussian linear scalar filters

We consider a family of mortality models with a continuous nGLSF described by

$$\mu_x(t, l) = \mu_{x_0}^l \exp\{\alpha_x^l t + \sum_{i=1}^m q_{x_i}^l y^i(t, l)\}, \quad (4)$$

$$dy(t, l) = -\beta_{x_1}^l y(t, l)dt + \gamma_{x_1}^l dW(t), \quad (5)$$

where  $\mu_x(t, l)$  is a stochastic process representing a mortality rate for a person aged  $x$  ( $x \in X = 0, 1, \dots, \omega$ ) at time  $t$ ;  $\alpha_x^l, \beta_{x_1}^l, q_{x_i}^l, i = 1, \dots, m, \mu_{x_0}^l, \gamma_{x_1}^l$  are constant parameters,  $l \in \mathbb{S}$ ;  $W(t)$  is a standard Wiener process.

We will show that the proposed model (4), (5) can be transformed to the formula (2) for all  $l \in \mathbb{S}$ .

Introducing new variables  $y_1(t, l) = y(t, l), \quad y_i(t, l) = y^i(t, l), i = 1, \dots, m, l \in \mathbb{S}$  and applying Ito formula we obtain

$$dy_1(t, l) = -\beta_{x_1}^l y_1(t, l)dt + \gamma_{x_1}^l dW(t), \quad (6)$$

$$dy_2(t, l) = [-2\beta_{x_1}^l y_2(t, l) + (\gamma_{x_1}^l)^2]dt + 2\gamma_{x_1}^l y_1(t, l)dW(t), \quad (7)$$

⋮

$$dy_m(t, l) = \left[ -m\beta_{x_1}^l y_m(t, l) + \frac{m(m-1)}{2} (\gamma_{x_1}^l)^2 y_{m-2}(t, l) \right] dt + m\gamma_{x_1}^l y_{m-1}(t, l) dW(t).$$

Taking natural logarithm of both sides of equation (4) and applying Ito formula for all  $l \in \mathbb{S}$  we find

$$d \ln \mu_x(t, l) = \alpha + \sum_{i=1}^m q_{x_i}^l dy_i(t, l) = \alpha_x^l - \sum_{i=1}^m [i\beta_{x_1}^l q_{x_i}^l y_i(t, l) \quad (8)$$

$$+ \frac{i(i-1)}{2} q_{x_i}^l (\gamma_{x_1}^l)^2 y_i(t, l)^{i-2}] dt + \sum_{i=1}^m i\gamma_{x_1}^l q_{x_i}^l y_{i-1}(t, l) dW(t) \quad (9)$$

Now we consider in details three cases of model (4) and (6)-(7), namely for  $m = 2, 4$  and  $6$ .

### 3.1 Model with six order output of a scalar linear filter

Equations (4) and (6)-(7) for  $m = 6$  take the form

$$\mu_x(t, l) = \mu_{x_0}^l \exp\{\alpha_x^l t + \sum_{i=1}^6 q_{x_i}^l y^i(t, l)\}, \quad (10)$$

$$dy_1(t, l) = -\beta_{x_1}^l y_1(t, l) dt + \gamma_{x_1}^l dW(t), \quad (11)$$

$$dy_2(t, l) = [-2\beta_{x_1}^l y_2(t, l) + (\gamma_{x_1}^l)^2] dt + 2\gamma_{x_1}^l y_1(t, l) dW(t), \quad (12)$$

$$dy_3(t, l) = [-3\beta_{x_1}^l y_3(t, l) + 3(\gamma_{x_1}^l)^2 y_1(t, l)] dt + 3\gamma_{x_1}^l y_2(t, l) dW(t), \quad (13)$$

$$dy_4(t, l) = [-4\beta_{x_1}^l y_4(t, l) + 6(\gamma_{x_1}^l)^2 y_2(t, l)] dt + 4\gamma_{x_1}^l y_3(t, l) dW(t), \quad (14)$$

$$dy_5(t, l) = [-5\beta_{x_1}^l y_5(t, l) + 10(\gamma_{x_1}^l)^2 y_3(t, l)] dt + 5\gamma_{x_1}^l y_4(t, l) dW(t), \quad (15)$$

$$dy_6(t, l) = [-6\beta_{x_1}^l y_6(t, l) + 15(\gamma_{x_1}^l)^2 y_4(t, l)] dt + 6\gamma_{x_1}^l y_5(t, l) dW(t). \quad (16)$$

Introducing a new vector state

$$\mathbf{z}_x(t, l) = [z_{x_1}(t, l), z_{x_2}(t, l), \dots, z_{x_7}(t, l)]^T = \quad (17)$$

$$[\ln \mu_x(t, l), y_1(t, l), y_2(t, l), y_3(t, l), y_4(t, l), y_5(t, l), y_6(t, l)]^T, \quad (18)$$

equations (10) - (16) one can rewrite in a vector form

$$d\mathbf{z}_x(t, l) = [\mathbf{A}_x^6(l)\mathbf{z}_x(t, l) + \mathbf{b}_x^6(l)] dt + [\mathbf{C}_x^6(l)\mathbf{z}_x(t, l) + \mathbf{g}_x^6(l)] dW(t) \quad (19)$$

where

$$\mathbf{A}_x^6(l) = [a_{ij}^l], \mathbf{b}_x^6(l) = [b_i^l], \mathbf{C}_x^6(l) = [c_{ij}^l], \mathbf{g}_x^6(l) = [g_i^l]. \quad (20)$$

The elements of the matrices  $\mathbf{A}_x^6(l)$ ,  $\mathbf{C}_x^6(l)$  and vectors  $\mathbf{b}_x^6(l)$ ,  $\mathbf{g}_x^6(l)$  are defined by:

$$\begin{aligned}
a_{12}^l &= -\beta_{x_1}^l q_{x_1}^l + 3q_{x_3}^l (\gamma_{x_1}^l)^2, & a_{13}^l &= -2\beta_{x_1}^l q_{x_2}^l + 6q_{x_4}^l (\gamma_{x_1}^l)^2, \\
a_{14}^l &= -3\beta_{x_1}^l q_{x_3}^l + 10q_{x_5}^l (\gamma_{x_1}^l)^2, & a_{15}^l &= -4\beta_{x_1}^l q_{x_4}^l + 15q_{x_6}^l (\gamma_{x_1}^l)^2, \\
a_{16}^l &= -5\beta_{x_1}^l q_{x_5}^l, & a_{17}^l &= -6\beta_{x_1}^l q_{x_6}^l, & a_{22}^l &= -\beta_{x_1}^l, & a_{33}^l &= -2\beta_{x_1}^l, & a_{42}^l &= 3(\gamma_{x_1}^l)^2, \\
a_{44}^l &= -3\beta_{x_1}^l, & a_{53}^l &= 6(\gamma_{x_1}^l)^2, & a_{55}^l &= -4\beta_{x_1}^l, & a_{64}^l &= 10(\gamma_{x_1}^l)^2, & a_{66}^l &= -5\beta_{x_1}^l, \\
a_{75}^l &= 15(\gamma_{x_1}^l)^2, & a_{77}^l &= -6\beta_{x_1}^l, & b_1^l &= \alpha_x^l + q_{x_2}^l (\gamma_{x_1}^l)^2, & b_3^l &= (\gamma_{x_1}^l)^2, \\
c_{12}^l &= 2q_{x_2}^l \gamma_{x_1}^l, & c_{14}^l &= 4q_{x_4}^l \gamma_{x_1}^l, & c_{15}^l &= 5q_{x_5}^l \gamma_{x_1}^l, & c_{16}^l &= 6q_{x_6}^l \gamma_{x_1}^l, & c_{32}^l &= 2\gamma_{x_1}^l, \\
c_{43}^l &= 3\gamma_{x_1}^l, & c_{54}^l &= 4\gamma_{x_1}^l, & c_{65}^l &= 5\gamma_{x_1}^l, & c_{76}^l &= 6\gamma_{x_1}^l, \\
g_1^l &= q_{x_1}^l \gamma_{x_1}^l, & g_2^l &= \gamma_{x_1}^l.
\end{aligned}$$

We note that similarly to equation (2) we may treat equation (19) as a family (set) of subsystems. It means, we have obtained new mortality hybrid model. The unknown parameters in family of equations (19) are

$$\ln \mu_{x_0}^l (= \alpha_{0_x}^l), \alpha_x^l, \beta_{x_1}^l, q_{x_1}^l, q_{x_2}^l, q_{x_3}^l, q_{x_4}^l, q_{x_5}^l, q_{x_6}^l, \gamma_{x_1}^l. \quad (21)$$

### 3.2 Nonstationary solutions

Using linear vector stochastic differential equation (19) and Ito formula we derive differential equations for the first order moments  $E[z_{x_i}(l)]$  and second order moments  $E[z_{x_i}(l)z_{x_j}(l)]$ ,  $i, j = 1, \dots, 7$ . Next, we find the nonstationary solutions of the first moment of the processes  $z_{x_i}(t, l)$  for nGLSF of all order models, i.e. (nGLSFo1), (nGLSFo2), ... (nGLSFo6) models

$$E[z_{x_1}(t, l)] = \alpha_x^l t + \alpha_{0_x}^l, \quad l \in \mathbb{S} \quad (22)$$

In the case of second moment of the processes  $z_{x_i}(t, l)$  we find first the nonstationary solutions for nGLSF even order models. In the case of sixth order model it has the form

$$E[z_{x_1}^2(t, l)] = (\alpha_x^l)^2 t^2 - 2\alpha_x^l \left[ -\alpha_{0_x}^l + q_{x_2}^l \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right] \quad (23)$$

$$+ 3q_{x_4}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^2 + 15q_{x_6}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^3 \Big] t + c_{0_x}^l \quad (24)$$

where  $l \in \mathbb{S}$ ,  $q_{x_2}^l = q_{x_4}^l = q_{x_6}^l = 1$ ,  $q_{x_2}^l = q_{x_4}^l = q_{x_6}^l = 1$ , and  $\alpha_{0_x}^l, c_{0_x}^l$  are constants of integration (see section 7).

To obtain the moment equations for nGLSF second and fourth order models and the corresponding stationary and nonstationary solutions we assume that:

- in the case of second order model the parameters  $q_{x_2}^l = 1$ , and  $q_{x_4}^l = q_{x_6}^l = 0$ ,

- in the case of fourth order model the parameters  $q_{x_2}^l = q_{x_4}^l = 1$ , and  $q_{x_6}^l = 0$ .

The corresponding nonstationary solution for the second moment of the process  $z_{x_1}(t, l)$  takes the form:

$$E[z_{x_1}^2(t, l)] = (\alpha_x^l)^2 t^2 + 2\alpha_x^l \alpha_{0_x}^l t - 2\alpha_x^l q_{x_2}^l p t + c_{0_x}^l \quad (25)$$

for nGLSF second order model, where  $c_{0_x}^l$  is an integration constant, and

$$E[z_{x_1}^2(t, l)] = (\alpha_x^l)^2 t^2 - 2\alpha_x^l [-\alpha_{0_x}^l + q_{x_2}^l p + 3q_{x_4}^l p^2] t + c_{0_x}^l \quad (26)$$

for nGLSF fourth order model, where  $c_{0_x}^l$  is an integration constant and  $p = \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}$ .

It can be proved that in the case odd order models the nonstationary solutions have similar forms, i.e. in the case of the first order (nGLSFo1) model

$$E[z_{x_1}^2(t, l)] = (\alpha_x^l)^2 t^2 + 2\alpha_x^l \alpha_{0_x}^l t + c_{0_x}^l \quad (27)$$

and in the case of other odd order models, i.e. (nGLSFo3), (nGLSFo5), (nGLSFo7) models the nonstationary solutions are the same as the nonstationary solutions for nGLSF even order models, i.e. (nGLSFo2), (nGLSFo4), (nGLSFo6) models, respectively.

## 4 The procedure of the parameters estimation and determination of submodels (based on switching points)

### 4.1 The procedure

Simultaneous estimation of parameters:  $\alpha_0^l, \alpha_x^l, \beta_x^l, \gamma_x^l, c_{0_x}^l, q_x^l$  (where  $l \in \mathbb{S}$ ) nGLSF models of 2, 4 or 6 order given by formulas (22)-(26) using traditional methods does not provide unambiguous results (this problem has already been considered in [20] part 4.1.1, in particular by considering the analytical formula for estimating parameters of GLSFo2 model). Therefore, in this case, a two-step procedure was used to estimate the parameters. In the first step, the  $\alpha_{0_x}$  and  $\alpha_x$  of the first moment  $E[z_{x_1}(t)]$  of the process  $z_{x_1}(t)$  were estimated. In the second step,  $c_{0_x}$  and  $p_x$  of the second moment  $E[z_{x_1}^2(t)]$  were estimated based on the already known  $\widehat{\alpha_{0_x}^l}, \widehat{\alpha_x^l}$ , where  $p_x$  was defined as follows:  $p_x = \frac{\gamma_{x_1}^2}{2\beta_{x_1}}$ . The applied procedure allows to obtain unambiguous estimates of all parameters assuming that  $q_{x_i} = 1, \forall i=1, \dots, 6$ .

One of the fundamental problems in the field of switching models is to find the set of switching points. This problem is closely related to the problem of segmentation of a time series discussed in many papers (see for instance [10],[14]).

In our considerations we propose a procedure which is a combination of a statistical test (based on [6]) and so called Top-Down algorithm. It has the following form.

First we introduce some notations. We assume that an extracted time series (Input) consists of  $n$  empirical values  $y_{emp_1}, y_{emp_2}, \dots, y_{emp_n}$  defined in time points  $t_1, t_2, \dots, t_n$ , respectively. By  $\langle t_1, t_2 \rangle$  we denote an interval that begins at  $t_1$  and ends in  $t_2$ . We define three sets

- $\mathcal{P}$  - the set of non-verified intervals,
- $\mathcal{R}$  - the set of intervals without switching points,
- $\mathcal{T}$  - the set of switching points.

Then the initial conditions have the form

$$\langle t_1, t_n \rangle \in \mathcal{P}, \quad \mathcal{R} = \phi, \quad \mathcal{T} = \phi.$$

**Step 1.**

We calculate the values of function  $L(*)$  given by formula (28)

$$\begin{aligned} L(\alpha_{0_x}^l, \alpha_x^l, \sigma^2, \tau) = & -\frac{\tau}{2} \ln(2\pi) - \frac{\tau}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{\tau} (y_{emp_i} - E[z_{x_1}(i, l)])^2 \\ & - \frac{n-\tau}{2} \ln(2\pi) - \frac{n-\tau}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=\tau+1}^n (y_{emp_i} - E[z_{x_1}(i, l)])^2 \end{aligned} \quad (28)$$

for all points from an interval  $\langle t_1, t_n \rangle$  and assuming the random component  $\epsilon_t \sim N(\mu, \sigma^2)$ .

If  $\tilde{L}(\tau_1) = \max L(*)$  is found at the beginning or at the end of the considered interval, then there is not a switching point in this interval. Then we receive

$$\mathcal{P} = \phi, \quad \langle t_1, t_n \rangle \in \mathcal{R}, \quad \mathcal{T} = \phi,$$

If  $\tilde{L}(\tau_1) = \max L(*)$  is found inside the interval for  $\tau_1 = t_k$ , then  $\langle t_1, t_k \rangle, \langle t_{k+1}, t_n \rangle \in \mathcal{P}$ ,  $\mathcal{R} = \phi$ ,  $\tau_1 \in \mathcal{T}$ .

**Step 2.**

Choose an interval from the set  $\mathcal{P}$  and check if its length is greater than 2.

**Step 3.**

If "no", then transfer this interval from the set  $\mathcal{P}$  to the set  $\mathcal{R}$  and go back to **Step 2**, if "yes", go back to **Step 1**.

**Step 4.**

The procedure is ended when

$$\mathcal{P} = \phi,$$

$\mathcal{R}$  consists only with subintervals without switching points,

$\mathcal{T}$  consists of all switching points that can be sorted from the smallest to the greatest one.

## 4.2 The determination of submodels

In subsection 4.1 we have established the switching points set, which allow to define submodels. From (21) and further considerations we find that unknown parameters in family of (19) are

$$\alpha_{0_x}^l, \alpha_x^l, p^l, c_{0_x}^l \quad (29)$$

where  $p = \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}$ , and parameters  $q_{x_2}, q_{x_4}, q_{x_6}$  are equal 0 or 1.

Based on the numerical algorithm of nonlinear minimization with additional conditions of  $\alpha_{0_x}^l$  ( $\forall x \alpha_{0_x}^l < 0$ ) parameters (29) given in the formula (23) - (26) were assessed. The algorithm works by generating a population of random starting points and next uses a local optimization method from each of the starting points to converge to a local minimum. As the solution, the best local minimum was chosen.

For a fixed sex, fixed age  $x$ , and knowing the switching points (designated in accordance with the procedure described above) two sets of time series of  $\widehat{\mu_{x,t}}$  values were created. In the first case, the estimation of  $\widehat{\mu_{x,t}}$  was based on empirical data from 1958-2010 (using the next 6 years for ex-post error evaluation). Similar estimation based on the years 1958-2016 was done in the second case. In both cases the choice of the theoretical value  $\widehat{\mu_{x,t}}$  at a fixed moment  $t$  from the theoretical values of the models (nGLSFo2), (nGLSFo4) and (nGLSFo6) was based on minimization of the absolute error (AE), i.e.

$$\min_{i=2,4,6} |\widehat{\mu_{x,t}}^{nGLSFo_i} - \mu_{x,t}|.$$

In addition, point forecasts for the period 2017-2025 have been determined. The parameters for the Lee-Carter model with switchings were estimated based on the formulas given in the literature [11] and using the same set of switches as in the case of the nGLSF model.

We note that the hybrid model (19) is continuous. However, the moment equations of the first and second-order defined by (22) - (27) are not continuous in switching points because the empirical data of mortality rates we have used were discrete, and these moments are determined separately for every submodel.

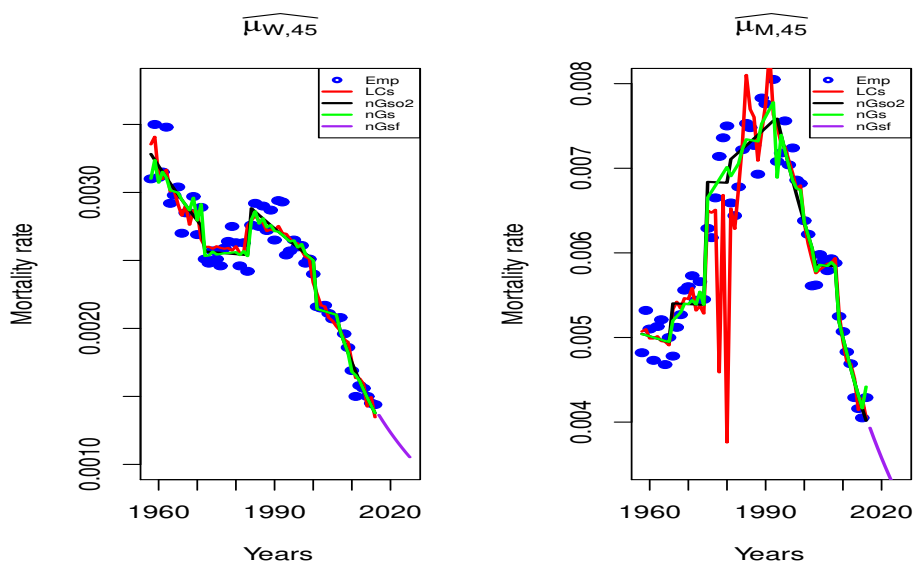
## 5 Results

Selected results for a 45-year old and a 60-year old woman and man presented in Figures 1-2 (source of empirical data: [8]). In figures 1 and 2, blue circular points indicate empirical data, red, black and green solid lines indicate the theoretical values of the models: Lee-Carter (LCs), nGLSF order 2 (nGso2) and nGLSF mixed order 2,4, and 6 (nGs) with switchings respectively, while the solid purple line indicates the forecast of the nG (nGsf) for the next five years.

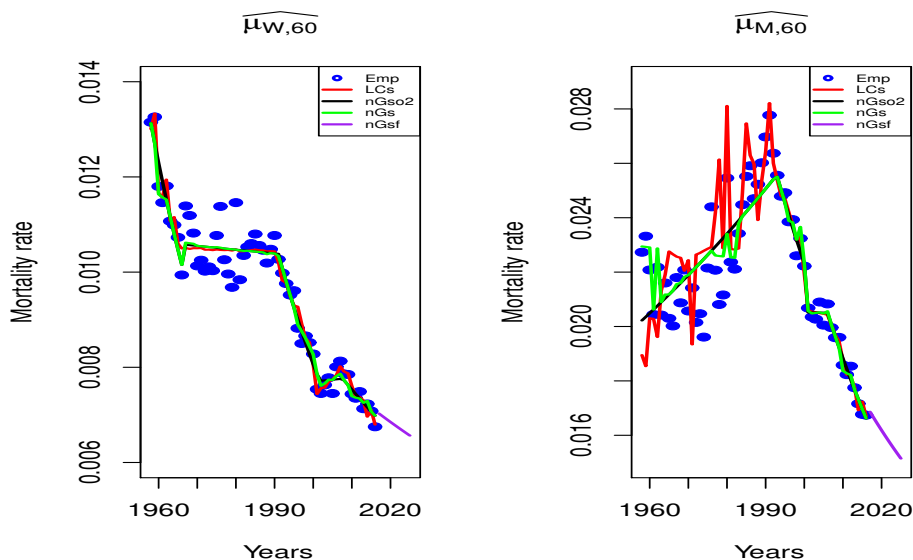
To verify the goodness of fit of the proposed nGs models with switchings to the empirical mortality rates and compared with Lee-Carter model the mean squared errors (MSE) between empirical mortality  $\mu_{x,t}$  and theoretical values  $\widehat{\mu_{x,t}}$  in the years 1958-2010 ('10) and 1958-2016 ('16) as well as the 95% confidence interval for MSE has been calculated. Selected results (45 and 60-year old female and male) are presented in Table 1 (where:  $CI_L$ -lower -,  $CI_U$ -upper confidence interval,  $\{W, M\}_{X, MSE}$  - MSE value for {female, male} aged X). The results in column 5th illustrate the model (nGso2) considered in [20].

MSE values calculated on the basis of empirical and theoretical data from 1958-2016 and included in Table 1 and Figures 1-2 provide the following conclusions:





**Fig. 1.** Mortality rates for women (left side) and men (right side) aged 45 and empirical, theoretical values based on the following models: LCs, nGs, and forecasts



**Fig. 2.** Mortality rates for women (left side) and men (right side) aged 60 and empirical, theoretical values based on the following models: LCs, nGs, and forecasts

**Table 1.** Goodness of fit measures (woman-W, man-M) based on MSE.

Sex & Age	EMP-LCs <sup>10</sup>	EMP-nGs <sup>10</sup>	EMP-LCs <sup>16</sup>	EMP-nGso2 <sup>16</sup>	EMP-nGs <sup>16</sup>
$W_{45,CI_L}$	7.145E-09	8.373E-09	9.468E-09	9.296E-09	7.380E-09
$W_{45,MSE}$	1.010E-08	1.190E-08	1.320E-08	1.290E-08	1.030E-08
$W_{45,CI_U}$	1.541E-08	1.806E-08	1.960E-08	1.925E-08	1.528E-08
$W_{60,CI_L}$	8.639E-08	8.983E-08	8.351E-08	9.293E-08	8.190E-08
$W_{60,MSE}$	1.220E-07	1.270E-07	1.160E-07	1.290E-07	1.140E-07
$W_{60,CI_U}$	1.863E-07	1.937E-07	1.729E-07	1.924E-07	1.696E-07
$M_{45,CI_L}$	3.176E-07	5.247E-08	2.882E-07	7.714E-08	4.546E-08
$M_{45,MSE}$	4.490E-07	7.430E-08	4.010E-07	1.070E-07	6.330E-08
$M_{45,CI_U}$	6.851E-07	1.132E-07	5.967E-07	1.597E-07	9.412E-08
$M_{60,CI_L}$	1.995E-06	8.113E-07	1.827E-06	9.966E-07	6.578E-07
$M_{60,MSE}$	2.820E-06	1.148E-06	2.540E-06	1.390E-06	9.160E-07
$M_{60,CI_U}$	4.303E-06	1.750E-06	3.784E-06	2.063E-06	1.362E-06

- the theoretical values of the mortality rate  $\widehat{\mu}_{x,t}^{nGs}$  based on the non-Gaussian linear scalar filters with switching provide closer estimates to empirical values than  $\widehat{\mu}_{x,t}^{LCs}$  based on LC model and  $\widehat{\mu}_{x,t}^{nGso2}$  with switching for both a 45-year-old and a 60-year-old woman and man,
- the range confidence interval is the smallest for the nGs model compared to all other models given in Table 1, which means greater precision of the proposed nGs for forecasting than the other models presented here,
- the empirical mortality rates for women are more accurately fitted using the proposed nGs model than for men (lower MSE value),
- based on graphical results (Figure 1 - Figure 2), it can be seen that the proposed method of modeling  $\mu_{x,t}$  using nGs more precisely adapts to empirical data, especially for data with a large variance than the LC model (e.g. see empirical data from 1980-1990 for a 60-year-old man on Figure 2, right side).

Moreover, taking into account all results for people aged  $x = 0, \dots, 100$  years (also partly included in Table 1) it can be seen that the proposed nGs model fits more accurately to the empirical data for younger than older (lower MSE for 45 years old than for 60 years old man and woman).

## 6 Conclusions

In this paper, three extended Milevsky and Promislov models with excitations modeled by the second, the fourth and the sixth order polynomials of outputs from a linear non-Gaussian filter are proposed and adopted to Polish mortality data. To obtain hybrid models the procedures of parameters estimation and the determination of switching points were proposed. Based on the theoretical values obtained from these three models, one series of theoretical values based on the AE criterion was constructed and compared with the theoretical mortality rates based on classical the Lee–Carter model. In addition, a point forecast was

computed. The obtained results confirm the usefulness of the switched model based on the continuous non-Gaussian process for modeling mortality rates.

A natural extension of the research contained in this article is the Markov chain application (homogeneous or heterogeneous), which will be used to describe the space of states built on extended Milevsky and Promislov models with excitations modeled by the second, the fourth and the sixth order polynomials. The issues discussed above will be examined in the next article.

## 7 Appendix

### The derivation of stationary and nonstationary solutions of moment equations in nGLSF six order model

Using linear vector stochastic differential equation (19) and Ito formula we derive differential equations for the first order moments  $E[z_{x_i}(l)]$  and second order moments  $E[z_{x_i}(l)z_{x_j}(l)]$ ,  $i, j = 1, \dots, 7$ .

Next we find the stationary solutions for the first order moments  $E[z_{x_i}(l)]$ ,  $i = 2, 3, \dots, 7$  and for second order moments  $E[z_{x_i}(l)z_{x_j}(l)]$ ,  $i, j = 1, \dots, 7$ ,  $(i, j) \neq (1, 1)$  equating to zero the corresponding time derivatives, i.e.

$$\frac{dE[z_{x_i}(t, l)]}{dt} = 0, \quad i = 2, 3, \dots, 7 \quad (30)$$

$$\frac{dE[z_{x_i}(t, l)z_{x_j}(t, l)]}{dt} = 0, \quad i, j = 1, \dots, 7, (i, j) \neq (1, 1) \quad (31)$$

Then we obtain

$$E[z_{x_2}(t, l)] = E[z_{x_4}(t, l)] = E[z_{x_6}(t, l)] = 0, E[z_{x_3}(t, l)] = \frac{\gamma_{x_1}^2}{2\beta_{x_1}}, \quad (32)$$

$$E[z_{x_3}(t, l)] = \frac{\gamma_{x_1}^2}{2\beta_{x_1}}, E[z_{x_5}(t, l)] = 3 \left( \frac{\gamma_{x_1}^2}{2\beta_{x_1}} \right)^2, E[z_{x_7}(t, l)] = 15 \left( \frac{\gamma_{x_1}^2}{2\beta_{x_1}} \right)^3 \quad (33)$$

Hence, from conditions (32)-(33) and equality

$$\frac{E[z_{x_1}(t, l)]}{dt} = \alpha_x^l \quad (34)$$

we find the nonstationary solution for the first moment of the process  $z_{x_1}(t, l)$

$$E[z_{x_1}(t, l)] = \alpha_x^l t + \alpha_0^l, \quad (35)$$

where  $\alpha_0^l$  is an integration constant.

Next, taking into account conditions (30)-(31), (32)-(33) and (35) we obtain

$$E[z_{x_2}^2(l)] = \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}, E[z_{x_3}^2(l)] = 3 \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^2, E[z_{x_2}z_{x_3}(l)] = 0, \quad (36)$$

$$E[z_{x_2}z_{x_4}(l)] = 3 \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^2, E[z_{x_4}^2(l)] = 15 \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^3, E[z_{x_4}z_{x_5}(l)] = 0, \quad (37)$$

$$E[z_{x_3}z_{x_4}(l)] = 0, E[z_{x_2}(l)z_{x_5}(l)] = 0, E[z_{x_3}z_{x_5}(l)] = 15 \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^3, \quad (38)$$

$$E[z_{x_5}^2(l)] = 105 \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^4, E[z_{x_2}z_{x_6}(l)] = 15 \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^3, E[z_{x_3}z_{x_6}(l)] = 0 \quad (39)$$

$$E[z_{x_4}z_{x_6}(l)] = 105 \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^4, E[z_{x_5}z_{x_6}(l)] = 0, E[z_{x_6}^2(l)] = 945 \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^5 \quad (40)$$

$$E[z_{x_2}(l)z_{x_7}(l)] = 0, E[z_{x_3}z_{x_7}(l)] = 105 \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^4, E[z_{x_4}z_{x_7}(l)] = 0, \quad (41)$$

$$E[z_{x_5}(l)z_{x_7}(l)] = 945 \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^5, E[z_{x_6}z_{x_7}(l)] = 0, E[z_{x_7}^2(l)] = 10395 \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^6 \quad (42)$$

$$E[z_{x_1}(l)z_{x_2}(l)] = q_{x_1}^l \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} + 3q_{x_3}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^2 + 15q_{x_5}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^3, \quad (43)$$

$$E[z_{x_1}(l)z_{x_3}(l)] = \frac{1}{2\beta_{x_1}^l} \left[ (\gamma_{x_1}^l)^2 E[z_{x_1}(l)] + \alpha_x^l \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} + 4\beta_{x_1}^l q_{x_2}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^2 \right] \quad (44)$$

$$+ 24\beta_{x_1}^l q_{x_4}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^3 + 180\beta_{x_1}^l q_{x_6}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^4 \Big],$$

$$E[z_{x_1}(l)z_{x_4}(l)] = 3q_{x_1}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^2 + 15q_{x_3}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^3 + 105q_{x_5}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^4 \quad (45)$$

$$E[z_{x_1}(l)z_{x_5}(l)] = \frac{1}{4\beta_{x_1}^l} \left[ 6(\gamma_{x_1}^l)^2 \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} E[z_{x_1}(l)] + 9\alpha_x^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^2 + \right] \quad (46)$$

$$+ 384\beta_{x_1}^l q_{x_4}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^4 + 3600\beta_{x_1}^l q_{x_6}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^5 \Big],$$

$$E[z_{x_1}z_{x_6}(l)] = 15q_{x_1}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^3 + 105q_{x_3}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^4 + 945q_{x_5}^l \left( \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right)^5, \quad (47)$$

$$\begin{aligned}
E[z_{x_1}(l)z_{x_7}(l)] &= 15\left(\frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}\right)^3 E[z_{x_1}(l)] + \frac{55}{4\beta_1} \alpha_x^l \left(\frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}\right)^3 + 90q_{x_2}^l \left(\frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}\right)^4 \\
&+ 900q_{x_4}^l \left(\frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}\right)^5 + 10170q_{x_6}^l \left(\frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}\right)^6.
\end{aligned} \tag{48}$$

Substituting quantities (36) - (48) to equation for the derivative  $\frac{dE[z_{x_1}^2(t,l)]}{dt}$  we obtain

$$\begin{aligned}
\frac{dE[z_{x_1}^2(t,l)]}{dt} &= 2\alpha_x^l E[z_{x_1}(t,l)] \\
&- 2\alpha_x^l \left[ q_{x_2}^l \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} + 3q_{x_4}^l \left(\frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}\right)^2 + 15q_{x_6}^l \left(\frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}\right)^3 \right]
\end{aligned} \tag{49}$$

Hence, from equation (49) and equality (35) we find the nonstationary solution for the second moment of the process  $z_{x_1}(t, l)$

$$E[z_{x_1}^2(t,l)] = (\alpha_x^l)^2 t^2 - 2\alpha_x^l \left[ -\alpha_{0_x}^l + q_{x_2}^l \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l} \right] \tag{50}$$

$$+ 3q_{x_4}^l \left(\frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}\right)^2 + 15q_{x_6}^l \left(\frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}\right)^3 \Big] t + c_{0_x}^l \tag{51}$$

where  $c_{0_x}^l$  is an integration constant.

## References

1. Booth, H., Tickle, L.: Mortality modelling and forecasting: a review of methods. ADSRI Working Paper **3** (2008)
2. Boukas, E.K.: Stochastic Hybrid Systems: Analysis and Design. Birkhauser, Boston, (2005)
3. Cairns, A.J.G., et al.: Modelling and management of mortality risk: a review. Scandinavian Actuarial Journal **2- 3**, 79–113 (2008)
4. Cairns, A.J.G., et al.: A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. North American Actuarial Journal **13**, 1–35 (2009)
5. Cao, H., Wang, J., et al.: Trend analysis of mortality rates and causes of death in children under 5 years old in Beijing, China from 1992 to 2015 and forecast of mortality into the future: an entire population-based epidemiological study. BMJ Open **7** (2017). doi.org/10.1136/bmjopen-2017-015941
6. Chow, G.C.: Tests of Equality Between Sets of Coefficients in Two Linear Regressions. Econometrica **28**, 591–605 (1960)
7. Giacometti, R., et al.: A stochastic model for mortality rate on Italian Data. Journal of Optimization Theory&Applications **149**, 216–228 (2011)

8. GUS:<http://demografia.stat.gov.pl/bazademografia/TrwanieZycia.aspx>, Central Statistical Office of Poland (2015)
9. Jahangiri, K., et al.: Trend forecasting of main groups of causes-of-death in Iran using the Lee-Carter model. *Med J Islam Repub Iran* **32**(1), 124 (2018)
10. Keogh, E., et al.: Segmenting Time Series: A Survey and Novel Approach. In: Last, M. (eds), Bunke, H. (eds), Kandel, A. (eds) *Data Mining In Time Series Databases*, vol. 83, pp. 1–22. World Scientific (2004)
11. Lee, R.D., Carter, L.: Modeling and forecasting the time series of U.S. mortality. *J AM STAT ASSOC* **87**, 659–671 (1992)
12. Lee, R.D., Miller, T.: Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography* **38**, 537–549 (2001)
13. Liberzon, D.: *Switching in Systems and Control*, Boston, Basel, Berlin: Birkhauser, (2003)
14. Lovrić, M., et al.: Algorithmic methods for segmentation of time series: an overview. *J. Contemp. Econ. Bus. Iss.* **1**, 31–53, (2014)
15. Milevsky, M.A., Promislov, S.D.: Mortality derivatives and the option to annuitise. *Insurance Mathematics&Economics* **29**(3), 299–318 (2001)
16. Renshaw, A., Haberman, S.: Lee-Carter mortality forecasting with age-specific enhancement. *Insurance Mathematics&Economics* **33**(2), 255–272 (2003)
17. Renshaw, A., Haberman, S.: A cohort-based extension to the Lee-Carter model for mortality reduction factor. *Insurance Mathematics&Economics* **38**(3), 556–570 (2006)
18. Rossa, A., Socha, L.: Proposition of a hybrid stochastic Lee-Carter mortality model. *Metodoloski zvezki (Advances in Methodology and Statistics)* **10**, 1–17 (2013)
19. Rossa, A., Socha, L., Szymanski, A.: *Hybrid Dynamic and Fuzzy Models of Mortality 1st edn.* WUL, Lodz (2018)
20. Sliwka, P., Socha, L.: A proposition of generalized stochastic Milevsky-Promislov mortality models. *Scandinavian Actuarial Journal*, 2018(8), 706–726, (2018). <https://doi.org/10.1080/03461238.2018.1431805>
21. Sliwka, P.: Proposed methods for modeling the mortgage and reverse mortgage installment. In: *Recent Trends In The Real Estate Market And Its Analysis*, pp. 189–206. SGH,Warszawa (2018)
22. Sliwka P.: Application of the Markov Chains in the Prediction of the Mortality Rates in the Generalized Stochastic Milevsky–Promislov Model. In: Mondaini, R. (eds) *Trends in Biomathematic 2018*, pp. 191–208. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-23433-1\\_14](https://doi.org/10.1007/978-3-030-23433-1_14)
23. Sliwka, P.: Application of the Model with a Non-Gaussian Linear Scalar Filters to Determine Life Expectancy,Taking into Account the Cause of Death. In: Rodrigues, J. et al. (eds) *Computational Science—ICCS 2019, LNCS*, vol. 11538, pp. 435–449. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-22744-9\\_34](https://doi.org/10.1007/978-3-030-22744-9_34)