Application of the model with a non-Gaussian linear scalar filters to determine life expectancy, taking into account the cause of death

Piotr Sliwka[0000−0001−6226−9580]

Faculty of Mathematics and Natural Sciences. School of Exact Sciences, Cardinal S.Wyszynski University, Warsaw, Poland
p.sliwka@uksw.edu.pl

Abstract. It is well-known that civilization diseases shorten life expectancy. The most common causes of death in Poland, both for women and men, are cancer and cardiovascular disease. The aim of the article is to use the non-Gaussian scalar filter model to determine life expectancy based on death rates after eliminating one of the above causes of death. Based on the obtained results, it can be stated that depending on the sex and type of the cause of death, the life expectancy may extend to several years.

JEL classification: C32, C53, J11

Keywords: life tables · life expectancy · a cause of death · forecasting of mortality rates · Ito stochastic differential equations · hybrid mortality models.

1 Introduction

The creation of the first tables of life goes back to the 17th century. E. Halley used the death records available in the years 1687-91 of the inhabitants of Wrocław, on the basis of which he built the first life tables. The basis of the life table is the set of deaths at the age of x completed years (usually for one-year age ranges x from 0 to at least 100 years). Due to different length of life and gender diversity, separate life tables are built for women and for men. Existing expectancy life tables give the expected number of complete years remaining to live e_x for a person at age x without considering the cause of death. The following assumption can, therefore, be made: elimination of the cause of death extends e_x (mortality occurs as a result of natural death).

The purpose of this article is to try to estimate how many years life can last longer if mortality does not occur due to a specific disease, but because of natural death. The appointment of a precise e_x requires accurate data on the number of people in the cohort who died at the age of x due to the cause of y. Obtaining such data with the current restrictions of law in relation to The General Data Protection Regulation (GDPR-RODO) is very difficult. On the other hand, on the website of the Statistics Poland (GUS) there are data on the

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

number of deaths in particular age groups (0, 1, 2, 3, 4, 5-9, 10-14, ..., 90-95 and 95 and more years) and at the level defined by the International Statistical Classification of Diseases and Related Health Problems in Poland after revision since 1997. These data allow the percentage of deaths to be determined in the case of a selected cause in each calendar year for a fixed age group (for these data, linear interpolation was used for each age of age group x). On this basis, it is possible to correct the number of people in the cohort whose death will not occur due to a given cause, but due to natural death, and then set the corrected death rates. There are not many articles dedicated to modelling mortality rates and determining e_x with the assumption given above. According to the best knowledge of the author, the methods proposed in the literature for determining e_x including the cause of death are most often based on the Lee-Carter model and its mutations (e.g. [1], [14], [15], [17], [18], [19]). In some articles, modelling and forecasting changes in mortality due to the established cause of death, time series techniques (eg ARIMA(1,1,1), [7]) are more often used than stochastic processes (e.g. birth and death process [2]). In others, instead of e_x, the rate of mortality of the number of people susceptible to a given disease ending in death is indicated (e.g. [12]), which also allows life expectancy to be determined for the studied cohort. However, no one has used the scalar model where a stochastic process is a colored noise modeled by a scalar linear filter with white noise input described by a scalar linear stochastic differential equations with constant coefficients ([16], [9], [20]). The usefulness and advantages of this proposition in relation to the Lee-Carter model were shown, among others, in [22].

The paper is organized as follows. In Section 2 basic notations and definitions of stochastic hybrid systems are introduced. In Sections 3 materials and methods are presented: data set, the continuous non-Gaussian excitation model, the procedure of parameters estimation and the determination of submodels based on switching points to obtain hybrid model. In the case without restrictions on parameters, the standard estimation methods (such as: maximum likelihood or least squares) are used. Section 4 compares the empirical model with the theoretical model described in section 3 and discusses the obtained results. Section 5 with general conclusions ends the article.

2 Mathematical preliminaries

Throughout this paper we use the following notation. Let $|\cdot|$ and $<\cdot>$ 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 $\sigma(t) : \mathbb{R}_+ \to \mathbb{S}$ be the switching rule, where $\mathbb{S} = \{1, \ldots, N\}$ is the set of states. We denote switching times as $\tau_1, \tau_2, \ldots$ and assume that there is a finite number of switches on every finite time interval. Let $W_k(t)$ $(k = 1, \ldots, M)$ be the independent Brownian motions. We assume that processes $W_k(t)$ and $\sigma(t)$ are both $\{\mathcal{F}_t\}_{t\geq 0}$ adapted.

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

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

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), t, \sigma)dt + \sum_{k=1}^{M} \mathbf{g}_k(\mathbf{x}(t), t, \sigma)dW_k(t), \quad (\sigma(t_0), \mathbf{x}(t_0)) = (\sigma_0, \mathbf{x}_0), \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state vector, $(\sigma_0, \mathbf{x}_0)$ is an initial condition, $t \in \mathbb{T}$ and M is a number of Brownian motions. $\mathbf{f}(\mathbf{x}(t), t, \sigma(t))$ and $\mathbf{g}_k(\mathbf{x}(t), t, \sigma(t))$ are defined by sets of $\{f(\mathbf{x}(t), t, l)\}$ and $\{\mathbf{g}_k(\mathbf{x}(t), t, l)\}$, respectively i.e. $\mathbf{f}(\mathbf{x}(t), t, \sigma(t)) = \mathbf{f}(\mathbf{x}(t), t, l)$, $\mathbf{g}_k(\mathbf{x}(t), t, \sigma(t)) = \mathbf{g}_k(\mathbf{x}(t), t, l)$ for $\sigma(t) = l$. Functions $\mathbf{f} : \mathbb{R}^n \times \mathbb{T} \times \mathbb{S} \to \mathbb{R}^n$ and $\mathbf{g}_k : \mathbb{R}^n \times \mathbb{T} \times \mathbb{S} \to \mathbb{R}^n$ are locally Lipschitz and such that $\forall l \in \mathbb{S}, t \in \mathbb{T}$ $\mathbf{f}(\mathbf{0}, t, l) = \mathbf{g}_k(\mathbf{0}, t, l) = \mathbf{0}, k = 1, \ldots, M$. These conditions together with these enforced on the switching rule $\sigma(t)$ 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 τ_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 Material and methods

3.1 Data

From the HMD database, both data describing death rates of men and women from 2002-2016 were taken, as well as for each year 1958-2016 and for each age X (X = 0, ..., 110) the number of people (l_x) surviving age X (eg for men from 2016: $l_0 = 100000, l_1 = 99541, \ldots, l_100 = 731, \ldots$). Based on l_x, q_x as the probability of death in the period up to 1 year was determined, and next μ_x as death rates computed. Using the data of Statistics Poland regarding the number of deaths due to cardiovascular disease (cause C) and cancer (cause I), the percentages of these deaths in the number of all deaths for each calendar year were separately determined. These percentages were used to correct (usually increase) l_x - the number of survivors aged x, and consequently, q_x and μ_x. The modelling of adjusted mortality ratios μ_x was based on the non-Gaussian scalar filter model, whose analysis and purposefulness of the application in the present study was included in the works of, among others, [22]-[24], while the general form is contained in subsection 3.2.

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

3.2 Model with a non-Gaussian linear scalar filters (non-Gaussian LSF)

We consider a family of mortality model with a continuous non-Gaussian scalar linear filter described by

$$\mu_x(t,l) = \mu^l_{x0} \exp\{\alpha^l_x t + \sum_{i=1}^{3} q^l_{x_i} y^i(t,l)\}, \tag{4}$$

$$dy(t,l) = -\beta^l_{x_1} y(t,l)dt + \gamma^l_{x_1} dW(t), \tag{5}$$

Introducing new variables $y_1(t,l) = y(t,l)$, $y_2(t,l) = y^2(t,l)$, $y_3(t,l) = y^3(t,l)$ and applying Ito formula we obtain

$$dy_2(t,l) = [-2\beta^l_{x_1} y_2(t,l) + (\gamma^l_{x_1})^2]dt + 2\gamma^l_{x_1} y_1(t,l)dW(t), \tag{6}$$

$$dy_3(t,l) = [-3\beta^l_{x_1} y_3(t,l) + 3(\gamma^l_{x_1})^2 y_1(t,l)]dt + 3\gamma^l_{x_1} y_2(t,l)dW(t), \tag{7}$$

where $\mu_x(t,l)$ is a stochastic process representing a mortality rate for a person aged x at time t, α^l_x, $\beta^l_{x_1}$, $q^l_{x_1}$, $q^l_{x_2}$, $q^l_{x_3}$, μ^l_{x0}, $\gamma^l_{x_1}$ are constant parameters, $l \in \mathbb{S}$; $W(t)$ is a standard Wiener process.

Taking natural logarithm of both sides of equation (4) and applying Ito formula we find

$$\begin{aligned} d\ln\mu_x(t,l) = \big[&\alpha^l_x - (\beta^l_{x_1} q^l_{x_1} - 3(\gamma^l_{x_2})^2) y_1(t,l) \\ &-(2\beta^l_{x_1} q^l_{x_1} - 6(\gamma^l_{x_2})^2) y_2(t,l) - (\gamma^l_{x_2})^2 - 3\beta^l_{x_1} q^l_{x_3} y_3(t,l)\big]\, dt \\ &+ \big[\gamma^l_{x_1} q^l_{x_1} + 2\gamma^l_{x_1} q^l_{x_2} y_2(t,l) + 3\gamma^l_{x_1} q^l_{x_3} y_3(t,l)\big]\, dW(t) \end{aligned} \tag{8}$$

Introducing a new vector state

$$\begin{aligned} \mathbf{z}_x(t,l) &= [z_{x_1}(t,l), z_{x_2}(t,l), z_{x_3}(t,l), z_{x_4}(t,l)]^T \\ &= [\ln\mu_x(t,l), y_1(t,l), y_2(t,l), y_3(t,l)]^T, \end{aligned} \tag{9}$$

equations (8) and (5) - (7) one can rewrite in a vector form

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

$$d\mathbf{z}_x(t,l)=\begin{bmatrix} 0 & -\beta^l_{x_1}q^l_{x_1}+3(\gamma^l_{x_1})^2 & -2\beta^l_{x_1}q^l_{x_2}+6(\gamma^l_{x_1})^2 & -3\beta^l_{x_1}q^l_{x_3} \\ 0 & -\beta^l_{x_1} & 0 & 0 \\ 0 & 0 & -2\beta^l_{x_1} & 0 \\ 0 & 3(\gamma^l_{x_1})^2 & 0 & -3\beta^l_{x_1} \end{bmatrix}\mathbf{z}_x(t,l)dt$$

$$+\begin{bmatrix} \alpha^l_x+q^l_{x_2}(\gamma^l_{x_1})^2 \\ 0 \\ (\gamma^l_{x_1})^2 \\ 0 \end{bmatrix}dt \tag{10}$$

$$+\begin{bmatrix} \gamma^l_{x_1}q^l_{x_1}+2\gamma^l_{x_1}q^l_{x_2}y_1(t,l)+3\gamma^l_{x_1}q^l_{x_3}y_2(t,l) \\ \gamma^l_{x_1} \\ +2\gamma^l_{x_1}y_1(t,l) \\ +3\gamma^l_{x_1}y_2(t,l) \end{bmatrix}dW(t)$$

The unknown parameters are

$$\ln\mu^l_0, \alpha^l_x, \beta^l_{x_1}, q^l_{x_1}, q^l_{x_2}, q^l_{x_3}, \gamma^l_{x_1}.$$

Using the method of the moment equations (see Appendix 3 in [22]) we find the nonstationary solutions of the first and second moment of the process $z_{x_1}(t,l), l\mathbb{S}$ (see Appendix 4 in [22])

$$E[z_{x_1}(t,l)] = \alpha^l_x t + \alpha^l_{0_x}, \tag{11}$$

$$E[z^2_{x_1}(t,l)] = (\alpha^l_x)^2t^2 + 2\alpha^l_x\alpha^l_{0_x}t - 2\alpha^l_x\frac{(\gamma^l_{x_1})^2}{2\beta^l_{x_1}}t + c^l_{0_x} \tag{12}$$

where $\alpha^l_{0_x}$ and $c^l_{0_x}$ are constants of integration.

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

To find the parameters estimation and the determination of the switching points for non-Gaussian linear scalar filters (non-Gaussian LSF) we use similar procedure to the one described in [22].

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

Due to the limited number of observations (from 2002 to 2016) the parameter estimation procedure was performed for two types of models, namely:

- for the moment model with non-Gaussian LSF without switchings (for only one l)
- for the moment model with non-Gaussian LSFs with switchings, i.e. for $l \in \mathbb{S}$ using the estimation methods for each subsystem (next subsection).

Parameters estimation We note that the first and second moments of $z_{x_1}(t,l) = \ln \mu_x(t,l)$ depend on only six parameters α_x^l, $\alpha_{0_x}^l = \ln \mu_{x0}^l(t)$, $c_{0_x}^l$, $q_{x_2}^l$, $\beta_{x_1}^l$, $\frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}$ and does not depend on the others, namely $q_{x_1}^l$, $q_{x_2}^l$, $q_{x_3}^l$. As it was shown in [22], only two parameters: α_x^l and $\alpha_{0_x}^l = \ln \mu_{x_0}^l(t), l \in \mathbb{S}$ are used and are found separately from minimization of the following square criterion

$$I_1 = \left(E[z_{x_1}(t,l)] - \alpha_x^l t - \alpha_{0_x}^l\right)^2. \tag{13}$$

Next, we assume for simplicity that $q_{x_1}^l = q_{x_2}^l = q_{x_3}^l = 1$. Then from the second moments of $z_{x_1}^2(t,l)$, i.e. $E[z_{x_1}^2(t,l)]$ we find the two parameters p_1^l and p_2^l, where $p_1^l = \frac{(\gamma_{x_1}^l)^2}{2\beta_{x_1}^l}$ and $p_2^l = c_{0_x}^l$, the relationship of which is nonlinear, namely

$$E[z_{x_1}^2(t,l)] = (\alpha_x^l)^2 t^2 + 2\alpha_x^l \alpha_{0_x}^l t - 2\alpha_x^l p_1^l t + p_2^l \tag{14}$$

Hence, the square criterion has the form

$$I_2 = \left(E[z_{x_1}^2(t,l)] - (\alpha_x^l)^2 t^2 - 2\alpha_x^l \alpha_{0_x}^l t + 2\alpha_x^l p_1^l t - p_2^l\right)^2 \tag{15}$$

In this case, all parameters ($\alpha_x^l, \alpha_{0_x}^l, p_1^l$ and p_2^l) in the formula (14) - (15) based on the numerical algorithm of nonlinear minimization with additional conditions of $\alpha_{0_x}^l$ parameters ($\forall x\ \alpha_{0_x}^l < 0$) 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.

The procedure of the determination of switching time points To identify the switching time points s_t the procedure based on the Chow test [8] (which allows to assess whether the respective regression coefficients are different for split data sets) due to limited series of time observations only on three- and six-years intervals was used.

Step 1. Split the 2002-2016 mortality data (source: [11]) into two groups of intervals. The first group consists of six-years intervals e.g.:

$$\tilde{\tau}_6(1) = \{2002, 2003, ..., 2007\}, \quad \ldots, \quad \tilde{\tau}_6(10) = \{2011, 2012, ..., 2016\}$$

The second group consists of three-years intervals e.g.:

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

$$\tilde{\tau}_3(1) = \{2002, 2003, 2004\}, \quad \ldots, \quad \tilde{\tau}_3(13) = \{2010, 2011, ..., 2015\}$$

Note: $\tilde{\tau}_3(1) \cup \tilde{\tau}_3(4) = \tilde{\tau}_6(1), \ldots$ and so on. In the next steps of the algorithm, the following sets of indices will be considered: $\tilde{\tau}_3(i), \tilde{\tau}_3(i+3)$ and $\tilde{\tau}_6(i)$ for i=1,...,10.

Step 2. For $i = 1, l = 1$:
Using estimated parameters $\widehat{\alpha_0^l}$ and $\widehat{\alpha_1^l}$ of the regression model: $\mu_{x,t} = \alpha_0^l + \alpha_1^l t + \varepsilon_t$ for the years 2002-2016 ($t \in 1, \ldots, 15$) we determine three types of sums of residual squares (based on the above regression):
- the first one for the 3-element subinterval $S_{\tilde{\tau}_3(i)} = \sum_{k \in \tilde{\tau}_3(i)} e_k^2$,
- the second one for the 3-element subinterval $S_{\tilde{\tau}_3(i+3)} = \sum_{k \in \tilde{\tau}_3(i+3)} e_k^2$,
- the third one for $S_{\tilde{\tau}_6(i)} = \sum_{k \in \tilde{\tau}_6(i)} e_k^2$ for $i = 1, \ldots, 10$.

Step 3. To test the existence of a switching point, we propose to apply the Chow test of statistic F_{emp} [8] based on the Fisher-Snedecor distribution F:

$$F_{emp,i} = \frac{\frac{S_{\tilde{\tau}_6(i)} - S_{\tilde{\tau}_3(i)} - S_{\tilde{\tau}_3(i+3)}}{m}}{\frac{S_{\tilde{\tau}_3(i)} + S_{\tilde{\tau}_3(i+3)}}{n_1 + n_2 - 2m}} \tag{16}$$

where: m is the number of the estimated parameters (with intercept), $n_1 = n_2 = 3$ are numbers of observations in two neighbor rolling subintervals.

If $F_{emp,i} > F_{r_1,r_2,\alpha}$ (alternatively: $p - value \leq \alpha$, where α is the level of significance; usually $\alpha = 0.05$) then reject null hypothesis H_0 with the set of statistical hypotheses as follows:

$$H_0 : \alpha_{0_x}^l = \alpha_{0_x}^{l+1} \wedge \alpha_x^l = \alpha_x^{l+1} \wedge c_{0_x}^l = c_{0_x}^{l+1} \quad \text{against the alternative} \quad H_1 : \neg H_0$$

and accept as the switching point the first element of the set $\tilde{\tau}_3(i+3)$, where: $F_{r_1,r_2,\alpha}$ is a value of theoretical Fisher-Snedecor distribution F with $r_1 = m$ and $r_2 = n_1 + n_2 - 2m$ at significance level α.

If we have rejected H_0 then we have found a switching point s_l between subsystem l and subsystem $l+1$ and we add it to the set of switching points, $l = l + 1$.

Step 4 Go back to Step 2, $i = i + 1$, repeat Step 2 and Step 3 until i=13.

Step 5 Finally, we have created the set of switching points s_j, $j = 1, ..., N - 1$ and the corresponding N intervals of the mortality data.

4 Results

Based on empirical central death rates $\mu_{x,t}$ for all ages x (x=0,...,100) the parameters of the models non-Gaussian SLF without and with switchings (in the case that at least one switching point has appeared) were evaluated and two sets of theoretical mortality rates $\widehat{\mu_{x,t}}$ were determined.

Selected results for a 40, 65, 67 and 70 year old woman and man are shown in Tables 1-2

Due to the small number of observations (only 15 years), there was no switching point in every age group. Based on the results of the Chow test in table 1

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

Table 1. Chow test values, 3- and 6-year intervals 2002-2016 (woman-W, man-M).

Sex	Age	02-07	03-08	04-09	05-10	06-11	07-12	08-13	09-14	10-15	11-16
W	40	0.65	1.42	**141.32**	0.30	1.25	4.35	2.37	11.49	8.88	3.31
	65	2.86	10.92	3.05	0.14	6.26	1.96	0.03	2.90	1.81	0.09
	67	4.57	5.10	3.32	6.38	5.41	1.11	1.04	0.23	1.72	1.93
	70	3.69	0.70	0.03	2.70	1.42	0.03	2.09	7.47	3.97	0.30
M	40	1.36	1.05	**127.50**	9.97	2.54	**28.52**	4.31	5.37	17.02	2.64
	65	**93.61**	12.62	0.62	0.11	3.61	0.98	0.06	1.04	0.14	1.57
	67	0.23	0.27	1.15	4.38	0.59	1.32	1.10	0.69	0.07	6.41
	70	2.20	0.12	0.96	0.23	1.43	0.13	0.49	2.36	0.27	0.46

Table 2. Life expectancy e_x: all causes of death, after removing cause C ($e_{x,C}$) and I ($e_{x,I}$) separately - selected years of life for women (e_{x_W}) and men (e_{x_M}).

age x	e_{x_W}	$e_{x_W,C}$	$e_{x_W,I}$	e_{x_M}	$e_{x_M,C}$	$e_{x_M,I}$
40	42.56	48.31	50.27	35.57	42.88	44.80
65	20.13	24.91	27.91	15.86	23.10	25.10
67	18.56	23.17	26.34	14.65	21.79	23.87
70	16.25	20.60	24.03	12.89	19.87	22.08

for selected age groups, it can be seen that there is only one switching point for women aged 40 and men aged 65 years, while for men aged 40 years, there are two switching points. Removal of the cause of death as expected generally extends the life expectancy. If the cause of death C is removed, the average life expectancy will increase, depending on the age group, from 4.35 to 5.75 years for women and from 6.98 to 7.31 for men. If the cause of I is removed, the average life expectancy will be extended by approx. 7.8 years for women and approx. 9.2 for men (see table 2). In addition, it can be seen that in all cases empirical mortality rates without taking into account the cause of death decrease over time, which means an increase in life expectancy.

Value of empirical, theoretical death rates $\widehat{\mu_x}$, $\widehat{\mu_{x,C}}$ (without the cause of death C) and $\widehat{\mu_{x,I}}$ (without the cause of death I) determined by the nGLSF model (nGLSFC - without the cause of death C, nGLSFI - without the cause of death I) and forecasts from 2017 to 2025 (denoted by an additional letter f, i.e. nGLSFf) for women (W) and men (M) are included in Figures 1-8.

The following conclusions can be drawn from the figures 1-8:

1. For a 40-year-old woman, the difference between $\widehat{\mu_{40}}$ and $\widehat{\mu_{40,C}}$, as well as between $\widehat{\mu_{40}}$ and $\widehat{\mu_{40,I}}$ is more or less stable, while for men it is decreasing, which means that C and I definitely increase their share in the total number of deaths and they are definitely the dominant causes of deaths in this age group.
2. For a 65-year-old woman, the difference between $\widehat{\mu_{65}}$ and $\widehat{\mu_{65,C}}$ is more or less stable, whereas between $\widehat{\mu_{65}}$ and $\widehat{\mu_{65,I}}$ it decreases, thus becoming the dominant cause of death in time, while in the case of men, the trend is slightly decreasing.

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

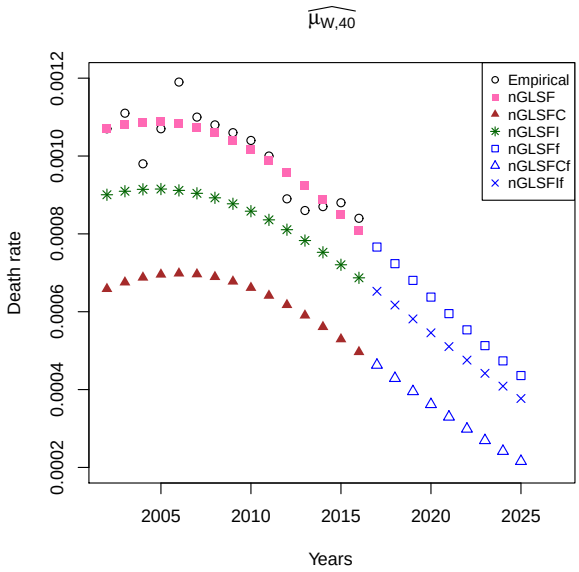

Fig. 1. Values of death rates for women aged 40 - empirical and theoretical values as well as forecasts based on the nGLSF model respectively

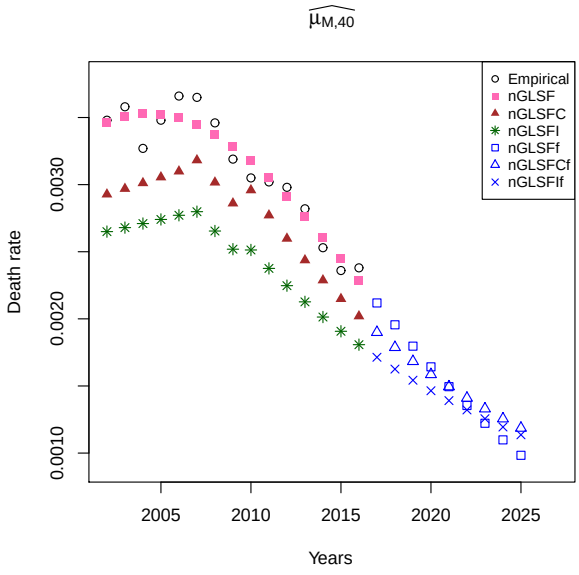

Fig. 2. Values of death rates for men aged 40 - empirical and theoretical values as well as forecasts based on the nGLSF model respectively

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

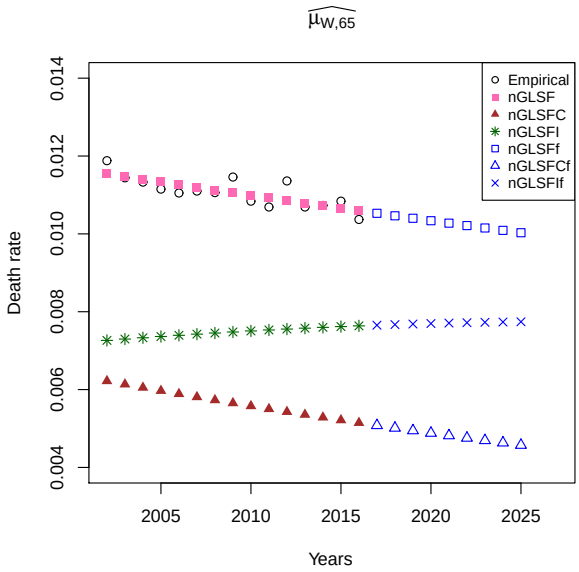

Fig. 3. Values of death rates for women aged 65 - empirical and theoretical values as well as forecasts based on the nGLSF model respectively

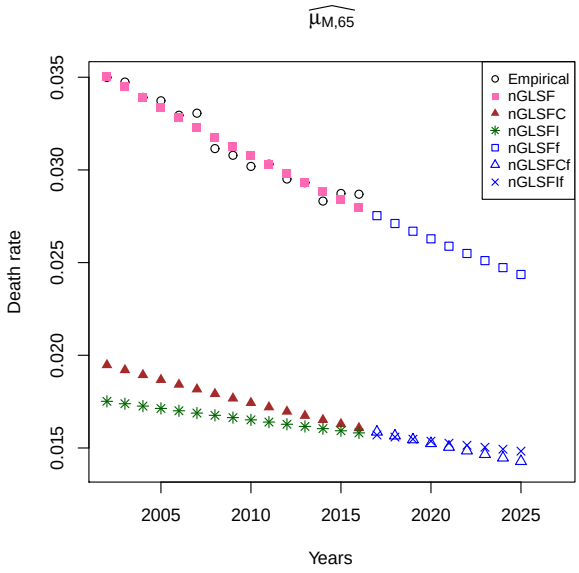

Fig. 4. Values of death rates for men aged 65 - empirical and theoretical values as well as forecasts based on the nGLSF model respectively

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

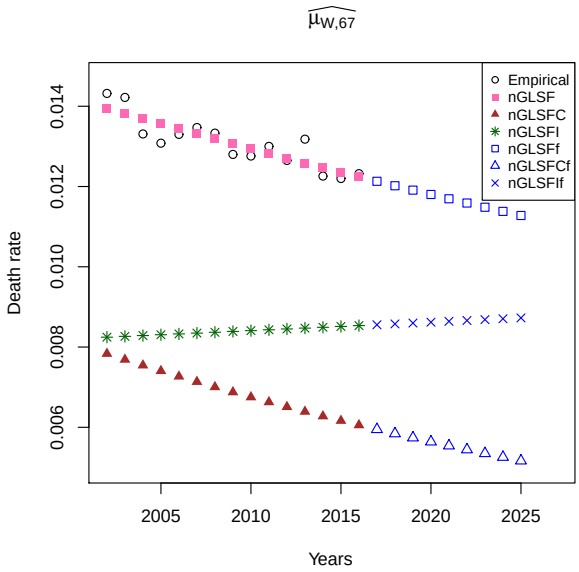

Fig. 5. Values of death rates for women aged 67 - empirical and theoretical values as well as forecasts based on the nGLSF model respectively

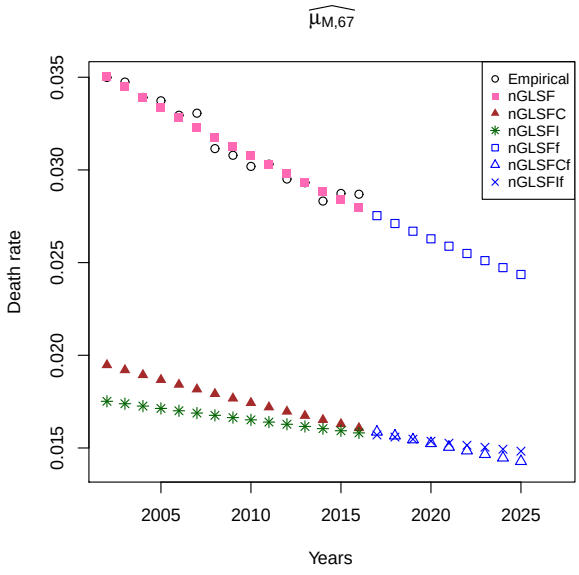

Fig. 6. Values of death rates for men aged 67 - empirical and theoretical values as well as forecasts based on the nGLSF model respectively

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

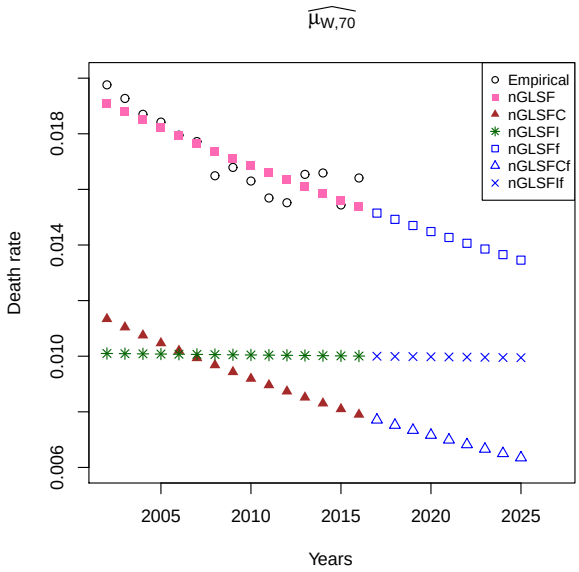

Fig. 7. Values of death rates for women aged 70 - empirical and theoretical values as well as forecasts based on the nGLSF model respectively

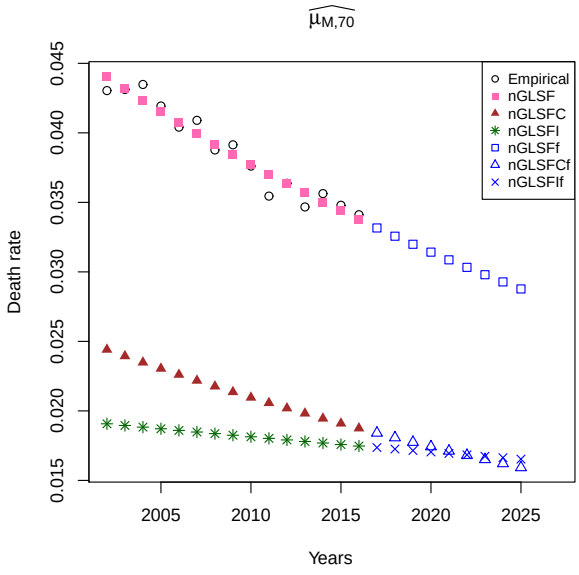

Fig. 8. Values of death rates for men aged 70 - empirical and theoretical values as well as forecasts based on the nGLSF model respectively

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

3. For a 67-year-old woman and man, the situation is similar to the case of a 65-year-old.
4. In the case of a 70-year-old woman the difference between $\widehat{\mu_{70}}$ and $\widehat{\mu_{70,C}}$ to 2008 is slightly decreasing, and since 2008 it is slightly growing, while the difference between $\widehat{\mu_{70}}$ and $\widehat{\mu_{70,I}}$ is decreasing until 2012, then more or less stable; in the case of men, the trend is slightly diminishing with a decreasing disparity between $\widehat{\mu_{70,C}}$ and $\widehat{\mu_{70,I}}$

where
$\widehat{\mu_{x,C}}$ - death rate without a cardiological cause,
$\widehat{\mu_{x,I}}$ - death rate without a cancer cause.

5 Conclusions

The purpose of this article was to try to estimate how many years life can last longer if death does not occur because of a specific disease, but because of natural death using a non-Gaussian linear scalar filter model. Determining the exact life expectancy e_x of individual people requires accurate data on the number of people in the cohort who died in the age of x due to the cause of C or I. Obtaining such data (e.g. from a hospital) is very difficult due to the current restrictions of law and in connection with the Act on protection of personal data GDPR. Nevertheless, the attempt to determine life expectancy with the exclusion of death due to C or I illness using the proposed model and the method of estimation seems realistic. Thus, the results obtained in the article should be treated as an approximation of the real life expectancy e_x. Determining the "real" e_x after exclusion of the cause of death would occur if, in fact, one observes a cohort for nearly 100 years. Currently, however, the implementation of such an experience seems unrealistic and therefore the methods of stochastic simulation should be further developed.

References

1. Booth, H., Tickle, L.: Mortality modelling and forecasting: a review of methods. ADSRI Working Paper **3** (2008)
2. Boumezoued, A., Hardy, H.L., El Karoui, N., Arnold, S.: Cause-of-Death Mortality: What Can Be Learned From Population Dynamics? INSUR MATH ECON **78**, 301–315 (2018)
3. Boukas, E.K.: Stochastic Hybrid Systems: Analysis and Design. Birkhauser, Boston, (2005)
4. Cairns, A.J.G., Blake, D., Dowd, K.: Modelling and management of mortality risk: a review. SCAND ACTUAR J **2- 3**, 79–113 (2008)
5. Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A., Balevich, I.: A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. N AM ACTUAR J **13**, 1–35 (2009)

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34

6. Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D.: Mortality density forecasts: An analysis of six stochastic mortality models. INSUR MATH ECON **48**, 355–367 (2011)
7. 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: 10.1136/bmjopen-2017-015941
8. Chow, G.C.: Tests of Equality Between Sets of Coefficinets in Two Linear Regressions. Econometrica **28**, 591–605 (1960)
9. Giacometti, R., Ortobelli, S., Bertocchi, M.: A stochastic model for mortality rate on Italian Data. J OPTIMIZ THEORY APP. **149**, 216–228 (2011)
10. Gompertz, B.: On the nature of the function expressive of the law of human mortality and on a new mode of determining life contingencies. R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci. Ser. A **CXV**, 513–580 (1825)
11. GUS: http://demografia.stat.gov.pl/bazademografia/TrwanieZycia.aspx, Central Statistical Office of Poland (2015)
12. Jahangiri, K., Aghamohamadi, S., Khosravi, A., Kazemi, E.: Trend forecasting of main groups of causes-of-death in Iran using the Lee-Carter model. Med J Islam Repub Iran (2018)
13. Lee, R.D., Carter, L.: Modeling and forecasting the time series of U.S. mortality. J AM STAT ASSOC **87**, 659–671 (1992)
14. Lee, R.D., Miller, T.: Evaluating the performance of the Lee-Carter method for forecasting mortality. Demography **38**, 537–549 (2001)
15. Maccheroni, C., Nocito, S.: Backtesting the Lee-Carter and the Cairns-Blake-Dowd Stochastic Mortality Models on Italian Death Rates. CeRP Working Papers **166**, (2017)
16. Milevsky, M.A., Promislov, S.D.: Mortality derivatives and the option to annuitise. INSUR MATH ECON **29**, 299–318 (2001)
17. Pollard, J.H.: Projection of age-specific mortality rates. Population bulletin of the United Nations **21-22**, 55–69 (1987)
18. Renshaw, A., Haberman, S.: Lee - Carter mortality forecasting with age-specific enhancement. INSUR MATH ECON **33**, 255–272 (2003)
19. Renshaw, A., Haberman, S.: A cohort-based extension to the Lee-Carter model for mortality reduction factor. INSUR MATH ECON **38**, 556–570 (2006)
20. Rossa, A., Socha, L.: Proposition of a hybrid stochastic Lee-Carter mortality model. Metodoloski zvezki (Advances in Methodology and Statistics) **10**, 1–17 (2013)
21. Rossa, A., Socha, L., Szymanski, A.: Hybrid Dynamic and Fuzzy Models of Mortality 1st edn. WUL, Lodz (2018)
22. Sliwka, P., Socha, L.: A proposition of generalized stochastic Milevsky-Promislov mortality models. SCAND ACTUAR J **8**, 706–726 (2018)
23. 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. Oficyna Wydawnicza SGH, Warszawa (2018)
24. Sliwka, P.: Application of the Markov Chains in the Prediction of the Mortality Rates in the Generalized Stochastic Milevsky-Promislov Model. In: BIOMAT Consortium Lectures. Springer (2019).

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_34