Article

Application of Continuous Non-Gaussian Mortality Models with Markov Switchings to Forecast Mortality Rates

Piotr Sliwka * and Leslaw Socha

Department of Mathematics and Natural Sciences, Cardinal Stefan Wyszyński University, Woycickiego 1/3, 01-938 Warsaw, Poland; l.socha@uksw.edu.pl
* Correspondence: p.sliwka@uksw.edu.pl

Featured Application: The proposed new methods of modelling and forecasting mortality rates are used, among others, to estimate life expectancy depending on the type of death as a fundamental life insurance factor.

Abstract: The ongoing pandemic has resulted in the development of models dealing with the rate of virus spread and the modelling of mortality rates $\mu_{x,t}$. A new method of modelling the mortality rates $\mu_{x,t}$ with different time intervals of higher and lower dispersion has been proposed. The modelling was based on the Milevski–Promislov class of stochastic mortality models with Markov switches, in which excitations are modelled by second-order polynomials of results from a linear non-Gaussian filter. In contrast to literature models where switches are deterministic, the Markov switches are proposed in this approach, which seems to be a new idea. The obtained results confirm that in the time intervals with a higher dispersion of $\mu_{x,t}$, the proposed method approximates the empirical data more accurately than the commonly used the Lee–Carter model.

Keywords: forecasting of mortality rates; hybrid mortality models; switched models; Lee–Carter model; Ito stochastic differential equations; Markov switchings

1. Introduction

The existing SARS-CoV-2 pandemic has resulted in the growth of epidemiological models that mainly describe the virus’s evolution and spread, affecting many public life spheres (social, economic, institutional, administrative, etc.). The most frequently used tools for building these models last year include: machine learning, taking into account various types of regression (e.g., multivariate regression, SVM, Cox multivariate regression, LASSO binary logistic regression [1], etc.). Other methods are based, e.g., on autoregressive models (ARiMA [2]), which also include the commonly used the Lee–Carter model (abbreviated as the LC model, [3–7]). Recently, mortality models described by stochastic differential equations have also been developed ([8–10]), including COVID-19 models ([11–13]).

This group of methods can be divided into two subgroups: the first subgroup consists of mortality models without jumps [14–18], while the second group consists of the ones with jumps [15,19–21].

Nowadays, the high dynamics of changes also force adjusting their parameters to the forecasting models’ situations. Therefore, it can be observed that there is still a need to search for models that consider the variability of parameters over time. Some authors have proposed using the methodology of stochastic dynamic hybrid (switched) systems ([22,23]). They considered the dynamic systems to consist of several subsystems described by deterministic or stochastic differential equations. These subsystems have the same structures and different parameters. The submodels can change over time according to a given switching rule, creating a hybrid system.

Recently, this idea was developed by [24–26], where the authors proposed measure changes (probability distributions) for LC-models. Another approach was developed...
by [27–32]. The authors proposed extended Milevsky and Promislov models with the non-Gaussian linear scalar filters for modelling of empirical mortality coefficients ($\mu_{x,t}$). Compared to the Gaussian linear scalar filters model with switchings (GLSFs) and Lee–Carter with switchings (LCs), the model proposed above allows for a more precise estimate of $\mu_{x,t}$ for some fixed age groups of $X$. The proposed models with switchings seem to represent the natural way the epidemic oscillates. Unfortunately, the switching times were assumed to be deterministic in all mentioned approaches ([33]).

In this paper, we first propose a special case when the empirical data of $\mu_{x,t}$ can be divided into three basic parts. The observed data are “regular” (lower dispersion) in their first and third part. The second part is “nonregular”, chaotically distributed, and with higher dispersion. The mortality models corresponding to the first and third part of the observed data are assumed to be extended Milevsky and Promislov models with the non-Gaussian linear scalar filters. The Markov switched model is proposed for the second part of the empirical data, which is a consequence of the switching between the first and the third part of the observed data. We will use the first and second moments of mortality rates in an iterative procedure to estimate the model parameters and the switching points between three basic parts. To the author’s knowledge, the proposed approach is a new one.

The paper is organised as follows. In Section 2.1, basic notations and definitions of stochastic hybrid systems with Markov switchings are introduced. In Sections 2.2 and 2.3, two new basic models with Gaussian linear scalar filters and continuous non-Gaussian excitation are presented, respectively. The moment equations are derived, and the stationary solutions for part variables are found. It is shown that only some of the parameters of these models can be unambiguously estimated. In Section 2.4, parameter estimation is performed based on the adapted numerical algorithm of a nonlinear minimisation problem (random search). The hybrid models are obtained by: parameter estimation procedures and the determination of switching points based on the Chow test for both models. The following steps of the technical work in Sections 2.1–2.4 were performed:

S1. The family of extended Milevsky and Promislov mortality models with Gaussian linear scalar filters (GLSF) is studied. This family is described by stochastic processes representing a mortality rate $\mu_{x,t}$ for a person aged $X$ at time $t$. The solutions of the mentioned stochastic differential equations are considered with switches (Section 2.1).

S2. Considering the ln-function of $\mu_{x,t}$, and applying the Ito formula, a new vector state with unknown parameters is introduced (Section 2.2).

S3. Using moment equations for GLSF with Markov switches, the first- and the second-order moments of equations for a particular case of two subsystem models (stationary and nonstationary solutions) are obtained, and approximate solutions are analysed (Sections 2.2.1–2.2.3).

S4. A similar analysis to step S3 for the non-Gaussian linear scalar filters (nGLSF) model with Markov switches is repeated (Section 2.3).

S5. The estimation procedure of the parameters (introduced in step S2) is applied (Section 2.4).

In Section 3, we have compared empirical mortality rates with theoretical ones obtained from the standard LC model and the models proposed in Section 2.1. Section 4 contains the discussion. Conclusions research are then given in Section 5.

2. Materials and Methods

2.1. 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 $\mathbb{R}_+ = [0, \infty)$, $T = [t_0, \infty)$, $t_0 \geq 0$. Let $\mathbb{S} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ 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 $W_k(t)$ 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

\[ dx(t) = f(x(t), t, \sigma)dt + \sum_{k=1}^{M} g_k(x(t), t, \sigma)dW_k(t), \quad (\sigma(t_0), x(t_0)) = (\sigma_0, x_0), \]

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

\[ f(x(t), t, \sigma(t)) = f(x(t), t, l), \quad g_k(x(t), t, \sigma(t)) = g_k(x(t), t, l) \text{ for } \sigma(t) = l. \]

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

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

\[ dx(t, l) = f(x(t), t, l)dt + \sum_{k=1}^{M} g_k(x(t), t, l)dW_k(t), \quad l \in \mathbb{S} \]

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

We assume that processes \( W_k(t) \) and \( \sigma(t) \) are mutually independent and both are \( \{\mathcal{F}_t\}_{t \geq 0} \) adapted. Let \( r(t), t \geq 0 \), be a right-continuous Markov chain on the probability space taking values in a finite state space \( \mathbb{S} \) with the generator \( \Gamma = [\gamma_{ij}]_{N \times N} \), i.e.,

\[
\mathbb{P}\{r(t + \Delta) = j | r(t) = i\} = \begin{cases} \gamma_{ij}\Delta + o(\Delta) & \text{if } i \neq j \\ 1 + \gamma_{ii}\Delta + o(\Delta) & \text{if } i = j, \end{cases}
\]

where \( \Delta > 0, \gamma_{ij} \geq 0 \) is the transition rate from \( i \) to \( j \) if \( i \neq j \), \( \gamma_{ii} = -\sum_{i\neq j} \gamma_{ij} \). We assume that the Markov chain is irreducible, i.e., rank(\( \Gamma \)) = \( N - 1 \), and has a unique stationary distribution \( \mathcal{P} = [p_1, p_2, \ldots, p_N]^T \in \mathbb{R}^N \), which can be determined by solving

\[
\begin{align*}
\mathcal{P}\Gamma &= 0 \\
\text{subject to } \sum_{i=1}^{N} p_i &= 1 \text{ and } p_i > 0 \text{ for all } i \in \mathbb{S}.
\end{align*}
\]

### 2.2. Model with Gaussian Linear Scalar Filter (GLSF)

We consider a family of extended Milevsky and Promislow mortality models ([34]) with Gaussian linear scalar filters described by

\[ \mu_x(t, l) = \mu_{x,0}^l \exp\{a_x^l t + q_x^l y_1(t, l)\}, \]

\[ dy_1(t, l) = -\beta_{x_1}^l y_1(t, l)dt + \gamma_{x_1}^l dW(t), \]

where \( \mu_x(t, l) \) is a stochastic process representing a mortality rate for a person aged \( x \) (x-fixed) at time \( t \), \( a_x^l, \beta_{x_1}^l, q_x^l, \mu_{x,0}^l, \gamma_{x_1}^l \), are constant parameters, \( l \in \mathbb{S} \); \( W(t) \) is a standard Wiener process, and \( y_1(t) \) is an output of a linear filter with an input process \( W(t) \).

Taking the natural logarithm of both sides of Equation (5) and applying the Ito formula we find

\[ d\ln \mu_x(t, l) = [a_x^l - \beta_{x_1}^l q_x^l y_1(t, l)]dt + \gamma_{x_1}^l q_x^l dW(t). \]

Introducing a new vector state: \( z_x(t, l) = [z_1(t, l), z_2(t, l)]^T = [\ln \mu_x(t, l), y_1(t, l)]^T \)

Equations (6) and (7) can be rewritten in a vector form

\[ dz_x(t, l) = \begin{bmatrix} 2 & -\beta_{x_1}^l q_x^l \\ 0 & -\beta_{x_1}^l \end{bmatrix} z_x(t, l) + \begin{bmatrix} a_x^l \\ 0 \end{bmatrix} dt + \begin{bmatrix} \gamma_{x_1}^l q_x^l \\ \gamma_{x_1}^l \end{bmatrix} dW(t), \quad z_x(0, l) = \begin{bmatrix} \ln \mu_{x,0}^l \\ 0 \end{bmatrix}. \]
The unknown parameters of this model are: \( \mu'_0, a'_{s1}, \beta_{s1}, q'_{s1}, \gamma_{s1}, l \in S \) (further in the text, the parameters \( q'_{s1} \) are denoted by \( q'_l \), \( l = 1, 2, 3 \)).

2.2.1. Moment Equations for GLFS Model with Markov Switchings

Using the methodology proposed in [28,29,31,32,35,36], and notation \( \rho_1 = -\gamma_{11} \), \( \rho_1 = \gamma_{12}, \rho_2 = -\gamma_{22}, \rho_2 = \gamma_{21} \) one can derive first and second order moment equations for a particular case of two subsystem models with Markov switchings

\[
\frac{dE[z_1(t, 1)]}{dt} = a'_1 - \beta_{s1} q'_1 E[z_2(t, 1)] - \rho_1 E[z_1(t, 1)] + \rho_1 E[z_1(t, 2)],
\]

\[
\frac{dE[z_1(t, 2)]}{dt} = a'_2 - \beta_{s1} q'_1 E[z_2(t, 2)] - \rho_2 E[z_1(t, 1)] + \rho_2 E[z_1(t, 2)],
\]

\[
\frac{dE[z_2(t, 1)]}{dt} = -\beta_{s1} E[z_2(t, 1)] - \rho_1 E[z_2(t, 1)] + \rho_1 E[z_2(t, 2)],
\]

\[
\frac{dE[z_2(t, 2)]}{dt} = -\beta_{s1} E[z_2(t, 2)] + \rho_2 E[z_2(t, 1)] - \rho_2 E[z_2(t, 2)],
\]

\[
\frac{dE[z_1^2(t, 1)]}{dt} = 2a'_1 E[z_1(t, 1)] - 2\beta_{s1} q'_1 E[z_1(t, 1)z_2(t, 1)] + (\gamma_{s1})^2 + \rho_1 E[z_1^2(t, 1)] + \rho_1 E[z_1^2(t, 2)],
\]

\[
\frac{dE[z_1^2(t, 2)]}{dt} = 2a'_2 E[z_1(t, 2)] - 2\beta_{s1} q'_1 E[z_1(t, 2)z_2(t, 2)] + (\gamma_{s1})^2 + \rho_2 E[z_1^2(t, 1)] - \rho_2 E[z_1^2(t, 2)],
\]

\[
\frac{dE[z_1(t, 1)z_2(t, 1)]}{dt} = a'_1 E[z_1(t, 1)] - \beta_{s1} q'_1 E[z_2(t, 1)] - \beta_{s1} E[z_1(t, 1)z_2(t, 1)]
\]

\[
+ (\gamma_{s1})^2 q'_1 - \rho_1 E[z_1(t, 1)z_2(t, 1)] + \rho_1 E[z_1(t, 2)z_2(t, 2)],
\]

\[
\frac{dE[z_1(t, 2)z_2(t, 2)]}{dt} = a'_2 E[z_2(t, 2)] - \beta_{s1} q'_1 E[z_2(t, 2)] - \beta_{s1} E[z_1(t, 2)z_2(t, 2)]
\]

\[
+ (\gamma_{s1})^2 q'_1 + \rho_2 E[z_1(t, 1)z_2(t, 1)] - \rho_2 E[z_1(t, 2)z_2(t, 2)],
\]

\[
\frac{dE[z_2^2(t, 1)]}{dt} = -2\beta_{s1} E[z_2^2(t, 1)] + (\gamma_{s1})^2 - \rho_1 E[z_2^2(t, 1)] + \rho_1 E[z_2^2(t, 2)],
\]

\[
\frac{dE[z_2^2(t, 2)]}{dt} = -2\beta_{s1} E[z_2^2(t, 2)] + (\gamma_{s1})^2 + \rho_2 E[z_2^2(t, 1)] - \rho_2 E[z_2^2(t, 2)].
\]

2.2.2. Analysis of First Order Moments

Equating the derivatives in Equations (11) and (12) to zero, we find

\[
(-\rho_1 - \beta_{s1}) E[z_2(t, 1)] + \rho_1 E[z_2(t, 2)] = 0,
\]

\[
\rho_2 E[z_2(t, 1)] + (-\rho_2 - \beta_{s1}) E[z_2(t, 2)] = 0.
\]

From Equations (19) and (20), it follows that

\[
E[z_2(t, 1)] = 0, \quad E[z_2(t, 2)] = 0.
\]

We denote

\[
u_1(t) = E[z_1(t, 1)], \quad \nu_2(t) = E[z_1(t, 2)].
\]

Substituting (21) to Equations (9) and (10), we obtain

\[
\frac{d}{dt} u(t) = Au(t) + v(t), \quad u(0) = u_0
\]
If we denote depending functions then we obtain approximate formulas for first-order moments in the form of linear time.

Then Equations (13) and (14) can be represented in the form

\[
u(t) = e^{At}u_0 + \int_0^t e^{A(t-\tau)}v(\tau)d\tau,
\]

(25)
after algebraic manipulation, one can find the solutions for both coordinations of the vector \(u(t)\) in the form

\[
E[z_1(t,1)] = u_1(t) = \left[ e^{-\beta t} + \frac{\beta_2}{\beta} (1 - e^{-\beta t}) \right] u_{10} + \left[ \frac{\beta_1}{\beta} (1 - e^{-\beta t}) \right] u_{20} + \frac{a_1^1 \beta_2 + a_2^2 \beta_1}{\beta} t - \frac{\beta_1}{\beta^2} e^{\beta t} (\alpha_2^2 - \alpha_1^1) (1 - e^{-\beta t}),
\]

(26)

\[
E[z_1(t,2)] = u_2(t) = \left[ \frac{\beta_2}{\beta} (1 - e^{-\beta t}) \right] u_{10} + \left[ e^{-\beta t} + \frac{\beta_1}{\beta} (1 - e^{-\beta t}) \right] u_{20} + \frac{a_1^1 \beta_2 + a_2^2 \beta_1}{\beta} t + \frac{\beta_2}{\beta^2} e^{\beta t} (\alpha_2^2 - \alpha_1^1) (1 - e^{-\beta t}).
\]

(27)

If we assume \(u_{10} = a_1 x_0\) and \(u_{20} = a_2^2 x_0\) are constant parameters, and we use an approximation

\[e^{-\beta t} \approx 1 - \beta t,\]

(28)

then we obtain approximate formulas for first-order moments in the form of linear time depending functions

\[
E[z_1(t,1)] = a_1^1 x_0 + \left(a_1^1 + \beta_1 (a_2^2 - a_1^1)\right) t,
\]

(29)

\[
E[z_1(t,2)] = a_2^2 x_0 + \left(a_2^2 + \beta_2 (a_1^1 - a_2^2)\right) t.
\]

(30)

2.2.3. Analysis of Second-Order Moments

Similar analysis may be derived for second-order moments \(E[z_1^2(t,1)]\) and \(E[z_1^2(t,2)]\). If we denote

\[
\phi_1(t) = E[z_1^2(t,1)],\quad \phi_2(t) = E[z_1^2(t,2)],
\]

(31)

then Equations (13) and (14) can be represented in the form

\[
\frac{d}{dt}\phi(t) = A\phi(t) + \psi(t), \quad \phi(0) = \phi_0
\]

(32)

where

\[
\phi(t) = \begin{bmatrix} \phi_1(t) \\ \phi_2(t) \end{bmatrix}, \quad A = \begin{bmatrix} -\rho_1 & \rho_1 \\ \rho_2 & -\rho_2 \end{bmatrix}, \quad \psi(t) = \begin{bmatrix} \delta(t,1) \\ \delta(t,2) \end{bmatrix}, \quad \phi_0 = \begin{bmatrix} \phi_{10} \\ \phi_{20} \end{bmatrix},
\]

(33)

\[
\delta(t,1) = 2a_1^1 E[z_1(t,1)] - 2\beta_1^2 q_1^2 [E[z_1(t,1)z_2(t,1)] + (\gamma_1^1 q_1^1)^2],
\]

(34)

\[
\delta(t,2) = 2a_2^2 E[z_1(t,2)] - 2\beta_2^2 q_1^2 [E[z_1(t,2)z_2(t,2)] + (\gamma_1^1 q_1^1)^2].
\]

(35)

Using Equalities (29) and (30) we find

\[
\delta(t,1) = a_1 t + b_1, \quad \delta(t,2) = a_2 t + b_2
\]

(36)
where
\[ a_1 = 2\alpha_x^1 \left[ \alpha_x^1 + \rho_1 (\alpha_x^2 - \alpha_x^0) \right], \quad a_2 = 2\alpha_x^2 \left[ \alpha_x^2 + \rho_2 (\alpha_x^1 - \alpha_x^2) \right], \]  
\[ b_1 = 2\alpha_x^1 \alpha_x^1 - 2\beta_x^2 q_1^1 E[z_1(t,1)z_2(t,1)] + (\gamma_x q_1^1)^2, \]  
\[ b_2 = 2\alpha_x^2 \alpha_x^2 - 2\beta_x^2 q_1^2 E[z_1(t,2)z_2(t,2)] + (\gamma_x q_1^2)^2. \]  

The general solution of Equation (23) has the form
\[ \phi(t) = e^{At} \phi_0 + \int_0^t e^{A(t-\tau)} \psi(\tau) d\tau. \]  

If we assume \( \phi_{10} = c_0^1 \) and \( \phi_{20} = c_0^2 \) are constant parameters, and we use approximation (28), then after algebraic manipulation, we obtain approximate formulas for second-order moments in the form of quadratic linear time-dependent functions
\[ E[z_1^2(t,1)] = c_{0_1}^1 + \left( b_1 + \rho_1 (c_{0_1}^2 - c_0^1) \right) t + \frac{\rho_1 \rho_2 - \rho_2^2}{2\rho} t^2, \]  
\[ E[z_1^2(t,2)] = c_{0_1}^2 + \left( b_2 + \rho_2 (c_{0_1}^1 - c_0^2) \right) t + \frac{\rho_1 \rho_2 - \rho_2^2}{2\rho} t^2. \]  

We note that in the case of nonstationary solutions of the first and second moment of the process \( z_{1_i}(t,l) \) without Markov switchings \( (\rho_1 = \rho_2 = 0) \), we obtain equalities
\[ E[z_{1_1}(t,l)] = a_{1_1}^1 t + a_{0_1}^1, \]  
\[ E[z_{1_2}(t,l)] = (a_{1_2}^1)^2 t^2 + 2a_{1_2}^1 a_{0_2}^1 t + c_{0_2}^1, \]  
where \( a_{1_0}^1 \) and \( c_{0_0}^1 \) are constants of integration and \( l = 1, 2 \).

2.3. Model with a Non-Gaussian Linear Scalar Filters (nGLSF)

We consider a family of mortality model with a continuous non-Gaussian scalar linear filter described by
\[ \mu_x(t,l) = \mu_{x_0}^l \exp \left[ \sum_{i=1}^{3} q_{1_i} y_i(t,l) \right], \]  
\[ dy(t,l) = -\beta_{x_1}^l y(t,l) dt + \gamma_{x_1}^l dW(t). \]  

Introducing new variables: \( y_1(t,l) = y(t,l), \quad y_2(t,l) = y^2(t,l), \quad y_3(t,l) = y^3(t,l) \) and applying the Ito formula, we obtain
\[ dy_2(t,l) = \left[ -2\beta_{x_1}^l y_2(t,l) + (\gamma_{x_1}^l)^2 \right] dt + 2\gamma_{x_1}^l y_1(t,l) dW(t), \]  
\[ dy_3(t,l) = \left[ -3\beta_{x_1}^l y_3(t,l) + 3(\gamma_{x_1}^l)^2 y_1(t,l) \right] dt + 3\gamma_{x_1}^l y_2(t,l) dW(t), \]  
where \( \mu_{x}(t,l) \) is a stochastic process representing a mortality rate for a person aged \( x \) at time \( t \), \( a_{1_0}^l, \beta_{x_1}^l, q_{1_0}, q_{1_1}, q_{1_2}, \mu_{x_0}^l, \gamma_{x_1}^l \) are constant parameters, \( l \in \mathbb{S} \); \( W(t) \) is a standard Wiener process.

Taking natural logarithm of both sides of Equation (45) and applying the Ito formula, we find
\[ d \ln \mu_{x}(t,l) = \left[ a_{1_0}^l - (\beta_{x_1}^l q_{1_0} - 3(\gamma_{x_1}^l)^2) y_1(t,l) \right. \]  
\[ -2\beta_{x_1}^l q_{1_1} - 6(\gamma_{x_1}^l)^2 y_2(t,l) - (\gamma_{x_1}^l)^2 - 3\beta_{x_1}^l q_{1_2} y_3(t,l) \]  
\[ \left. + \left[ \gamma_{x_1}^l q_{1_1} + 2\gamma_{x_1}^l q_{1_2} y_2(t,l) + 3\gamma_{x_1}^l q_{1_3} y_3(t,l) \right] dW(t). \]
Introducing a new vector state

\[
\mathbf{z}_x(t) = [z_{x_1}(t), z_{x_2}(t), z_{x_3}(t), z_{x_4}(t)]^T = [\ln \mu_x(t), y_1(t), y_2(t), y_3(t)]^T,
\]

Equations (46)-(49) can be rewritten in a vector form

\[
d\mathbf{z}_x(t) = \begin{bmatrix}
0 & -\beta_1 y_1^2 + 3(\gamma_1')^2 & -2\beta_1 y_1 q_0' + 6(\gamma_1')^2 & -3\beta_1 q_0'
0 & -\beta_2 & 0 & 0
0 & 0 & -2\beta_1 & 0
0 & 3(\gamma_1')^2 & 0 & -3\beta_1
\end{bmatrix} \mathbf{z}_x(t) dt + \begin{bmatrix}
0 \\
\gamma_1 y_1(t) + 2\gamma_1 y_1 q_0 y_1(t) + 3\gamma_1 y_1 q_2 y_2(t)
\gamma_1 y_1(t) + 2\gamma_1 y_1 q_0 y_1(t)
0
\gamma_1 y_1(t)
\end{bmatrix} dW(t),
\]

and \( \mathbf{z}_x(0) = [\ln \mu_0, 0, 0, 0]^T \). The unknown parameters are: \( \ln \mu_0, \alpha_1, \beta_1, \gamma_1, q_0 \).

Similar to Section 2.2, using the method of the moment equations (see, e.g., Appendix B.1), we find the nonstationary solutions of the first and second moment of the process \( z_{x_1}(t, l) \) (see, e.g., Appendix B.2)

\[
E[z_{x_1}(t, l)] = \alpha_{x_1}' t + \alpha_{0_1}'
\]

\[
E[z_{x_1}^2(t, l)] = (\alpha_{x_1}')^2 t^2 + 2\alpha_{x_1}' \alpha_{0_1}' t - 2\alpha_{x_1}' \frac{(\gamma_{x_1})^2}{2\beta_{x_1}} t + c_{0_1}'
\]

where \( \alpha_{0_1}' \) and \( c_{0_1}' \) are constants of integration.

To obtain the differential equations for first- and second-order moments with Markov switchings, we differentiate Equations (52) and (53) and denote in the forms (22)-(24) for first-order moments and (31)-(39) for second-order moments.

In the case of first-order moments with Markov switchings, we obtain exactly the same equations and solutions as in the previous chapter. In the case of second-order moments with Markov switchings, we find

\[
\frac{d}{dt} \phi(t) = A \phi(t) + \psi(t), \quad \phi(0) = \phi_0
\]

where

\[
\phi_1(t) = E[z_{x_1}^2(t, l)], \quad \phi_2(t) = E[z_{x_1}^2(t, l)],
\]

\[
\phi(t) = \begin{bmatrix}
\phi_1(t) \\
\phi_2(t)
\end{bmatrix}, \quad A = \begin{bmatrix}
-\rho_1 & \rho_1 \\
\rho_2 & -\rho_2
\end{bmatrix}, \quad \psi(t) = \begin{bmatrix}
\delta(t, 1) \\
\delta(t, 2)
\end{bmatrix}, \quad \phi_0 = \begin{bmatrix}
\phi_{10} \\
\phi_{20}
\end{bmatrix},
\]

\[
\delta(t, 1) = a_1 t + b_1, \quad \delta(t, 2) = a_2 t + b_2
\]

where

\[
a_1 = 2(\alpha_1')^2, \quad a_2 = 2(\alpha_2')^2,
\]

\[
b_1 = -2\alpha_1' \left[ -\alpha_{x_0}' + q_2' \frac{(\gamma_{x_1}')^2}{2\beta_{x_1}} \right], \quad b_2 = -2\alpha_2' \left[ -\alpha_{x_0}' + q_2' \frac{(\gamma_{x_1}')^2}{2\beta_{x_1}} \right].
\]
The general solution of Equation (23) has the form
\begin{equation}
\phi(t) = e^{At} \phi_0 + \int_0^t e^{A(t-\tau)} \psi(\tau) d\tau.
\end{equation}

If we assume \( \phi_{t_0} = c_{t_0}^1 \) and \( \phi_{t_0} = c_{t_0}^2 \) are constant parameters, and we use approximation (28), then, after algebraic manipulation, we obtain approximate formulas for second-order moments in the form of quadratic linear time-dependent functions
\begin{equation}
E[z_1^2(t, 1)] = c_{t_0}^1 + \left( b_1 + \rho_1 (c_{t_0}^3 - c_{t_0}^1) \right) t + \frac{\rho_1 a_2 + \rho_2 a_1}{2\rho} t^2
\end{equation}
\begin{equation}
E[z_1^2(t, 2)] = c_{t_0}^2 + \left( b_2 + \rho_2 (c_{t_0}^3 - c_{t_0}^1) \right) t + \frac{\rho_1 a_2 + \rho_2 a_1}{2\rho} t^2.
\end{equation}

### 2.4. Procedure

Based on the following procedure (steps S1–S3), the time intervals with significantly higher and lower dispersion of the empirical \( \mu_{x,t} \) were determined:

**S1:** The 10-year segments \( y_1 = \{1958, 1959, \ldots, 1967\} \), \( y_2 = \{1959, 1960, ..., 1968\} \), ..., \( y_{47} = \{2007, 2008, \ldots, 2016\} \) were determined, and for each \( y_{k,(k=1,...,47)} \), a dispersion measure \( \sigma_k \) was computed based on the empirical \( \mu_{x,t} \) mortality data.

**S2.** The statistical hypothesis was verified: \( H_0 : \sigma_k^2 = \sigma_{k+1}^2 \) (versus \( H_1 : \sigma_k^2 \neq \sigma_{k+1}^2 \)), which assumes no statistically significant difference in dispersion between \( y_k \) and \( y_{k+1} \) segments, based on the F statistic and the Fisher–Snedecor distribution. Rejection of \( H_0 \) indicates a significant change in dispersion between \( y_k \) and \( y_{k+1} \) segments (selection based on the lowest p-value < 0.05, [37]).

**S3.** The non-Gaussian Linear Scalar Filter model with Markov switchings (nGMs), given by Formulas (61)–(62), was proposed in the range with the higher dispersion. In the range with the lower dispersion, the nGs model (with switchings), given by (52)–(53), was applied (see: [29,32]).

The random search algorithm was used to estimate the parameters of the models mentioned in S3 and described in more detail, for example, in [29].

Due to the simultaneous occurrence of parameters \( a_l, b_l, c_{t_0}^l \) and \( \rho_l \) in Formulas (61)–(62), the parameter estimation was performed iteratively.

The empirical data of mortality rates \( \mu_{x,t} \) used in the article are contained in the Statistics Poland and The Human Mortality Database (source: [38,39]).

### 3. Results

Based on the observations \( \mu_{x,t} \) and the procedure described in Section 2.4, the variability of dispersion throughout the intervals seemed to occur mainly among women and men between 30 and 40 years of age (with some exceptions for both women and men). Apart from a woman/man 30–40 years old, the Fisher F-test for variance did not show significant differences in terms of lower and higher dispersion periods based on the F statistic. The obtained results are divided into two subgroups: one of them with significant differences between lower and higher dispersion and the second one in another case. The first group includes, for example, a 37-year-old woman (\( F_{37} \)), a 32-year-old man (\( M_{32} \)), a 36-year-old man (\( M_{36} \)), and a 37-year-old man (\( M_{37} \)). The second group, for example, consists of a 35-year-old woman (\( F_{35} \)) and a 22-year-old man (\( M_{22} \)). The above results are presented in Tables 1 and 2 and Figures 1–7 (source of empirical data: [38,39]).
Figure 1. The values of the empirical mortality rates: \( F_{35}, F_{37}, M_{22}, M_{32}, M_{36}, M_{37} \).

Figure 2. The values of mortality rates—female, age 37: empirical (Emp), theoretical (LC, nGMs), forecasts (nGMsf), and c. interval based on the LC, and nGM model with Markov switch.
Figure 3. The values of mortality rates—male, age 32: empirical (Emp), theoretical (LC, nGMs), forecasts (nGMsf), and c. interval based on the LC, and nGM model with Markov switch.

Figure 4. The values of mortality rates—male, age 36: empirical (Emp), theoretical (LC, nGs), forecasts (nGsf), and c. interval based on the LC, and nGs model with switch.
Figure 5. The values of mortality rates—male, age 37: empirical (Emp), theoretical (LC, nGs), forecasts (nGsf), and c. interval based on the LC, and nGs model with switch.

Figure 6. The values of mortality rates—female, age 35: empirical (Emp), theoretical (LC, nGs), forecasts (nGsf), and c. interval based on the LC, and nGs model with switch.
Figure 7. The values of mortality rates—male, age 22: empirical (Emp), theoretical (LC, nGs), forecasts (nGsf), and c. interval based on the LC, and nGs model with switch.

Table 1. MSE (goodness of fit), F-statistics, p-value (pv).

| Age  | $MSE_{EM\text{-}LC}$ | $MSE_{EM\text{-}nG}$ | $MSE_{EM\text{-}nGMs}$ | $F_1 (pv)$ | $F_2 (pv)$ |
|------|---------------------|----------------------|------------------------|------------|------------|
| $F_{37}$ | $5.265 \times 10^{-9}$ | $1.720 \times 10^{-8}$ | $4.849 \times 10^{-9}$ | 0.195 (0.02) | 8.69 (0.004) |
| $M_{32}$ | $2.122 \times 10^{-8}$ | $7.639 \times 10^{-8}$ | $1.762 \times 10^{-8}$ | 0.22 (0.03) | 0.18 (0.02) |
| $M_{36}$ | $9.591 \times 10^{-8}$ | $1.873 \times 10^{-7}$ | $3.701 \times 10^{-8}$ | 0.12 (0.01) | 5.15 (0.02) |
| $M_{37}$ | $4.818 \times 10^{-8}$ | $2.481 \times 10^{-8}$ | $–$ | 0.27 (0.0197) | 192 (2.7 $\times 10^{-23}$) |
| $F_{35}$ | $3.74 \times 10^{-9}$ | $2.471 \times 10^{-9}$ | $–$ | – | – |
| $M_{22}$ | $1.728 \times 10^{-8}$ | $1.319 \times 10^{-8}$ | $–$ | – | – |

Table 2. The 90% and 95% CI for the forecast of the mortality rates (2020–2027; L: left, R: right).

| CI   | $F_{37}$  | $M_{32}$  | $M_{36}$  | $M_{37}$  | $F_{35}$  | $M_{22}$  |
|------|-----------|-----------|-----------|-----------|-----------|-----------|
| 90% L | $4.57 \times 10^{-4}$ | $1.05 \times 10^{-3}$ | $1.45 \times 10^{-3}$ | $1.47 \times 10^{-3}$ | $3.38 \times 10^{-4}$ | $7.98 \times 10^{-4}$ |
| 90% R | $5.66 \times 10^{-4}$ | $1.27 \times 10^{-3}$ | $1.73 \times 10^{-3}$ | $1.84 \times 10^{-3}$ | $4.35 \times 10^{-4}$ | $8.97 \times 10^{-4}$ |
| 95% L | $4.47 \times 10^{-4}$ | $1.03 \times 10^{-3}$ | $1.42 \times 10^{-3}$ | $1.44 \times 10^{-3}$ | $3.29 \times 10^{-4}$ | $7.89 \times 10^{-4}$ |
| 95% R | $5.77 \times 10^{-4}$ | $1.30 \times 10^{-3}$ | $1.76 \times 10^{-3}$ | $1.88 \times 10^{-3}$ | $4.44 \times 10^{-4}$ | $9.06 \times 10^{-4}$ |

Figure 1 shows an example of an empirical $\mu_{x,t}$ from 1958–2019, regarding periods with lower and higher dispersion in the case of $F_{37}$, $M_{32}$, $M_{36}$, and $M_{37}$. Additionally, the example points to a lack of statistical differences between periods in cases $F_{35}$ and $M_{22}$.

In Figures 2–7, the observed (blue circular points) and theoretical values of the $\mu_{x,t}$ (using the Formulas (61) and (62)) mortality rates based on the Lee–Carter (red squares), nGs or nGMs (green diamonds) models from 1958-2019 (HMD) are presented. Moreover, the forecast (purple diamonds) and the yellow and sienna lines define 95% (CI95) and 90% (CI90) confidence intervals. In the case of $F_{37}$, $M_{32}$, and $M_{36}$, the data were split into three periods of lower, higher, and again lower dispersion. In the case of $M_{37}$, it is possible to distinguish an initial period with a higher dispersion (in the years 1958–1963), but the number of observations is not sufficient to estimate all parameters of the model (61)–(62).
(Section 2.3). In this case, the $nG_2m$ model with switches in 1963 and 1991 was used [28]. For the $F_{35}, M_{22}$, the F-test showed no significant change in variance concerning all 10-year time intervals described in Section 2.4 ($p$-value > 0.05). Therefore, in this case, the sequence of observations in the period 1958–2019 cannot be divided into areas with significantly lower and higher dispersion. In this case, the nGs model with one switch (Chow test: $F = 11.54, p$-value $= 7.88 \times 10^{-4}$ in the case of $F_{35}$ and $F = 9.99, p$-value $= 0.0016$ in the case of $M_{22}$) was applied to the complete set of observations (see Figures 6 and 7). In other cases, the nGMs model was better suited to the empirical data than the LC model or the nG model without switches. The LC model is generally accepted and treated as a benchmark model (see the Mean Squared Error (MSE) in Table 1). The 90% and 95% confidence intervals (CI) for the 2020–2027 forecast of the mortality rates based on nGMs model ($F_{37}, M_{32}, M_{36}$) and nGs model ($F_{35}, M_{22}, M_{37}$) are included in Table 2.

4. Discussion

Based on Figures 1–7, the values of MSE contained in Table 1, the CI values contained in Table 2, and the results of the research contained in papers [28–32], the following observations can be made:

- All the results obtained in the article regarding the proposed model were compared with the benchmark, i.e., the LC model,
- In all the studied cases, the MSE value for the nGMs model, which took into account the divisions of periods into higher and lower dispersion, is lower than for the LC model,
- As a consequence of the above observation, the respective confidence intervals are narrower, thereby resulting in more accurate forecasts,
- In cases that neither contained higher or lower periods of dispersion nor switchings, the LC model works as a better model, one better suited to the empirical data, the LC model usually fits the empirical data better ($F_{35}, M_{22}$),
- In cases that neither contained higher or lower periods of dispersion nor switchings, the nGs model fits the empirical data better ($M_{37}$),
- In cases of determining periods with significantly smaller and higher dispersion, the proposed method of modelling $\mu_{x,t}$ reflects the shape of the empirical data better than the LC model and the nGs model,
- The weakness of the proposed nGMs model can be found in the minimum number of observations that uniquely estimate all model parameters in the period with higher dispersion (in the case of the model (52)–(53) and (61)–(62), a minimum of seven observations, this condition is not met by $M_{37}$).

Based on the above observations, the nGMs model is recommended for modelling and forecasting mortality rates in time intervals with a higher dispersion of $\mu_{x,t}$. The model proposed in this paper provides closer estimates of empirical values than the model based on the Lee–Carter model. The proposed model can be applied, among other things, to estimate life expectancy by age group, taking into account the type of death. This fundamental life insurance factor allows the insurance company to evaluate risks more accurately and thus increase competitiveness in the financial market.

5. Conclusions

The aim of the article was to verify the validity of the non-Gaussian linear scalar filter model with homogeneous Markov switchings concerning empirical $\mu_{x,t}$. The obtained MSE results from Section 3 confirm the usefulness of the proposed model, by dividing the data into periods of lower and higher dispersion. To determine the time intervals with the significantly different variances of $\mu_{x,t}$, the Fisher test (F) was used.

Based on the lowest $p$-value, the limits of these intervals with statistically significantly higher and lower dispersion have been determined, according to the procedure described in Section 2.4. The article focuses mainly on the range between 30 and 40 years of age because in this interval, based on historical data, periods of lower ($\pm 1958$–1970 and 1995–2019)
and higher (1971–1994) dispersion, regardless of gender, were observed. Currently, in the case of the increased number of deaths from the end of 2019 to nowadays, also caused by different types of the virus, the modelling approach proposed in the article should find wide applications. However, the current epidemic time series is too short to model annual mortality rates. The paper opens possibilities of research on the heterogeneity of the Markov chain in the case of the nGLSF model with Markov switchings.

Additionally, one unresolved issue remains the selection of appropriate methods for estimating the parameters of models. While the parameters have been proposed already earlier in this article, further research could fine-tune the methods that would aid in estimating parameters. One of the proposals could be using MCMC methods, while another could be the use of machine learning techniques. Finally, the use of fractional differential equations for \( \mu_{x,t} \) modelling remains an area of inquiry for researchers.

**Author Contributions:** Conceptualisation, P.S. and L.S.; methodology, P.S. and L.S.; software, P.S.; validation, P.S. and L.S.; formal analysis, P.S. and L.S.; investigation, P.S. and L.S.; resources, P.S. and L.S.; data curation, P.S.; writing—original draft preparation, P.S. and L.S.; writing—review and editing, P.S. and L.S.; visualisation, P.S.; supervision, P.S. and L.S. All authors have read and agreed to the published version of the manuscript.

**Acknowledgments:** The authors would like to thank the anonymous reviewers for their valuable comments.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** [http://demografia.stat.gov.pl/bazademografia/TrwanieZycia.aspx](http://demografia.stat.gov.pl/bazademografia/TrwanieZycia.aspx), Statistical Office, Poland, 2020 (accessed on 23 May 2022). [https://mortality.org/hmd/STATS/Deaths-1x1.txt](https://mortality.org/hmd/STATS/Deaths-1x1.txt), The Human Mortality Database, 2020 (accessed on 23 May 2022).

**Appendix A. Model with Gaussian Linear Scalar Filter**

**Appendix A.1. Moment Equations for GLSF Model**

The moment equations in this model for all \( l \in S \) are

\[
\frac{dE[z_{x_1}(l,t)]}{dt} = \alpha_{x_1}^l - \beta_{x_1}^l q_{x_1}^l E[z_{x_2}(l,t)] - \beta_{x_2}^l q_{x_2}^l E[z_{x_3}(l,t)], \tag{A1}
\]

\[
\frac{dE[z_{x_2}(l,t)]}{dt} = -\beta_{x_1}^l E[z_{x_2}(l,t)], \tag{A2}
\]

\[
\frac{dE[z_{x_3}(l,t)]}{dt} = -\beta_{x_2}^l E[z_{x_3}(l,t)], \tag{A3}
\]

\[
\frac{dE[z_{x_1}^2(l,t)]}{dt} = 2\alpha_{x_1}^l E[z_{x_1}(l,t)] - 2\beta_{x_1}^l q_{x_1}^l E[z_{x_1}(l,t)z_{x_2}(l,t)] - 2\beta_{x_2}^l q_{x_2}^l E[z_{x_2}(l,t)z_{x_1}(l,t)] - (\gamma_{x_1}^l q_{x_1}^l + \gamma_{x_2}^l q_{x_2}^l)^2, \tag{A4}
\]

\[
\frac{dE[z_{x_1}(l,t)z_{x_2}(l,t)]}{dt} = \alpha_{x_1}^l E[z_{x_2}(l,t)] - \beta_{x_1}^l q_{x_1}^l E[z_{x_2}^2(l,t)] - \beta_{x_2}^l q_{x_2}^l E[z_{x_1}(l,t)z_{x_3}(l,t)] + (\gamma_{x_1}^l q_{x_1}^l + \gamma_{x_2}^l q_{x_2}^l)\gamma_{x_1}^l, \tag{A5}
\]
where the second moment of the process $z$ is given by

$$E[z^2_{x_1}(t,l)] = a^i_{x_1} t + b_{x_1}^i + l$$

(A10)

Hence, and from Equation (A1), we find the nonstationary solution for the first moment of the process $z_{x_1}(t,l)$:

$$E[z_{x_1}(t,l)] = a^i_{x_1} t + a^0_{x_1}, \quad l \in S$$

(A11)

where $a^0_{x_1}$ is a constant in the integration of Equation (A1).

Next, equating the derivatives in Equations (A5)–(A9) to zero and taking into account conditions (A10), we obtain

$$E[z^2_{x_2}(l)] = \frac{(\gamma_{x_2})^2}{2\beta_{x_1}}, \quad E[z^2_{x_3}(l)] = \frac{\gamma_{x_3}^2}{2\beta_{x_2}}, \quad E[z_{x_2}(l)z_{x_3}(l)] = \frac{\gamma_{x_2} \gamma_{x_3}}{\beta_{x_1} + \beta_{x_2}}$$

(A12)

$$E[z_{x_1}(l)z_{x_2}(l)] = \frac{1}{\beta_{x_1}} \left( - \frac{q_{x_1}(\gamma_{x_1})^2}{2} - \beta_{x_2} q_{x_2} \gamma_{x_2}^j \gamma_{x_2}^l \beta_{x_1} + \beta_{x_2} \right)$$

(A13)

Substituting quantities (A12)–(A14) to Equation (A4), we obtain

$$\frac{dE[z^2_{x_1}(t,l)]}{dt} = 2a^i_{x_1} E[z_{x_1}(t,l)]$$

(A15)

Hence, and from Equations (A11) and (A15), we find the nonstationary solution for the second moment of the process $z_{x_1}(t,l)$.
where $c^l_{0_{x}}$ is a constant of integration.

**Appendix B. Model with Non-Gaussian Linear Scalar Filter**

**Appendix B.1. Moment Equations and Stationary Solutions for Non-GLSF Model**

The moment equations in the considered model are

$$
\frac{dE[z_{x_1}(t,I)]}{dt} = (a^l_x + q^l_{x_2} (\gamma^l_1)^2 - (\beta^l_{x_1} q^l_{x_3} - 3q^l_{x_3} \gamma^l_1 (t))E[z_{x_2}(t,I)]]
$$

(A17)

- $2\beta^l_{x_1} q^l_{x_2} E[z_{x_3}(t,I)] - 3\beta^l_{x_1} q^l_{x_3} E[z_{x_4}(t,I)],$

$$
\frac{dE[z_{x_2}(t,I)]}{dt} = -2\beta^l_{x_1} E[z_{x_3}(t,I)] + (\gamma^l_1)^2,
$$

(A18)

$$
\frac{dE[z_{x_3}(t,I)]}{dt} = -3\beta^l_{x_1} E[z_{x_4}(t,I)] + 3(\gamma^l_1)^2 E[z_{x_4}(t,I)],
$$

(A19)

$$
\frac{dE[z_{x_4}(t,I)]}{dt} = (2a^l_x + 2q^l_2 (\gamma^l_1)^2)E[z_{x_1}(t,I)]
- 2(\beta^l_{x_1} q^l_{x_2} - 3(\gamma^l_1)^2 q^l_{x_3})E[z_{x_1}(t,I)z_{x_2}(t,I)]
- 4\beta^l_{x_1} q^l_{x_2} E[z_{x_1}(t,I)z_{x_3}(t,I)] - 6\beta^l_{x_1} q^l_{x_3} E[z_{x_1}(t,I)z_{x_4}(t,I)] + (\gamma^l_1)^2 (q^l_{x_2})^2
+ 4(\gamma^l_1)^2 (q^l_{x_2})^2 E[z_{x_2}^2(t,I)] + 9(\gamma^l_1)^2 (q^l_{x_3})^2 E[z_{x_3}^2(t,I)]
+ 4(\gamma^l_1)^2 q^l_{x_1} q^l_{x_2} E[z_{x_2}^2(t,I)] + 6(\gamma^l_1)^2 q^l_{x_1} q^l_{x_3} E[z_{x_3}(t,I)]
+ 12(\gamma^l_1)^2 q^l_{x_2} q^l_{x_3} E[z_{x_2}(t,I)z_{x_3}(t,I)]
$$

A21

- $2\beta^l_{x_1} q^l_{x_2} E[z_{x_2}(t,I)z_{x_3}(t,I)] - \beta^l_{x_1} E[z_{x_1}(t,I)z_{x_2}(t,I)]$

(A22)

- $3\beta^l_{x_1} q^l_{x_3} E[z_{x_2}(t,I)z_{x_4}(t,I)] + 2(\gamma^l_1)^2 (q^l_{x_2})^2 E[z_{x_2}(t,I)] +
+ 3(\gamma^l_1)^2 (q^l_{x_3})^2 E[z_{x_3}(t,I)] + (\gamma^l_1)^2 q^l_{x_1},$
\[
\frac{dE[z_{x_1}(t,l)z_{x_3}(t,l)]}{dt} = (a_1^l + q_{x_2}^l (\gamma_1^l)^2)E[z_{x_1}(t,l)] - 3(\gamma_1^l)^2 q_{x_3}^l E[z_{x_2}(t,l)z_{x_3}(t,l)] \\
- 2\beta_1^l q_{x_3}^l E[z_{x_3}(t,l)] - 2\beta_2^l E[z_{x_1}(t,l)z_{x_3}(t,l)] \\
- 3\beta_1^l q_{x_1}^l E[z_{x_1}(t,l)z_{x_3}(t,l)] + 2(\gamma_1^l)^2 q_{x_3}^l E[z_{x_2}(t,l)] + \\
+ 4(\gamma_1^l)^2 q_{x_2}^l E[z_{x_2}(t,l)] + (\gamma_1^l)^2 E[z_{x_1}(t,l)] + 6(\gamma_1^l)^2 q_{x_2}^l E[z_{x_2}(t,l)z_{x_3}(t,l)].
\]

\[
\frac{dE[z_{x_1}(t,l)z_{x_3}(t,l)]}{dt} = (a_1^l + q_{x_2}^l (\gamma_1^l)^2)E[z_{x_1}(t,l)] - 3(\gamma_1^l)^2 q_{x_3}^l E[z_{x_2}(t,l)z_{x_3}(t,l)] \\
- 2\beta_1^l q_{x_3}^l E[z_{x_3}(t,l)] - 3(\gamma_1^l)^2 q_{x_3}^l E[z_{x_1}(t,l)] \\
+ 6(\gamma_1^l)^2 q_{x_2}^l E[z_{x_2}(t,l)z_{x_3}(t,l)] + 9(\gamma_1^l)^2 q_{x_2}^l E[z_{x_2}(t,l)] - 9(\gamma_1^l)^2 q_{x_2}^l E[z_{x_2}(t,l)z_{x_3}(t,l)].
\]

\[
\frac{dE[z_{x_3}^2(t,l)]}{dt} = -2\beta_1^l E[z_{x_3}^2(t,l)] + \gamma_1^l q_{x_1}^l E[z_{x_3}(t,l)],
\]

\[
\frac{dE[z_{x_2}(t,l)z_{x_3}(t,l)]}{dt} = -2\beta_1^l E[z_{x_2}(t,l)z_{x_3}(t,l)] - 2\beta_2^l E[z_{x_2}(t,l)z_{x_3}(t,l)] \\
+ (\gamma_1^l)^2 E[z_{x_2}(t,l)] + 2(\gamma_1^l)^2 E[z_{x_2}(t,l)],
\]

\[
\frac{dE[z_{x_2}(t,l)z_{x_3}(t,l)]}{dt} = -2\beta_1^l E[z_{x_2}(t,l)z_{x_3}(t,l)] - 3\beta_1^l E[z_{x_2}(t,l)z_{x_4}(t,l)] \\
+ (\gamma_1^l)^2 E[z_{x_2}(t,l)] + 3(\gamma_1^l)^2 E[z_{x_2}(t,l)z_{x_3}(t,l)] \\
+ 6(\gamma_1^l)^2 E[z_{x_2}(t,l)z_{x_3}(t,l)].
\]

Appendix B.2. Partially Stationary Solutions for Non-GLSF Model

Equating the derivatives in Equations (A18)–(A20) to zero, we obtain

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

Hence, from conditions (A21) and equalities (A31), we find the nonstationary solution for the first moment of the process \( z_{x_1}(t,l) \)

\[
E[z_{x_1}(t,l)] = a_1^l t + a_0^l,
\]
where \( \alpha_0 \) is an integration constant.

Next, equating the derivatives in Equations (A22)–(A30) to zero and taking into account conditions (A31), (A32), we obtain

\[
E[z_{x_1}^2(l)] = \frac{(\gamma_{x_1})^2}{2\beta_{x_1}}, \quad E[z_{x_1}^2(l)] = 3\left(\frac{(\gamma_{x_1})^2}{2\beta_{x_1}}\right)^2, \quad E[z_{x_2}z_{x_3}(l)] = 0 \tag{A33}
\]

\[
E[z_{x_2}(l)z_{x_4}(l)] = 3\left(\frac{(\gamma_{x_1})^2}{2\beta_{x_1}}\right)^2, \quad E[z_{x_4}^2(l)] = 15\left(\frac{(\gamma_{x_1})^2}{2\beta_{x_1}}\right)^3, \quad E[z_{x_3}z_{x_4}(l)] = 0 \tag{A34}
\]

\[
E[z_{x_1}(l)z_{x_2}(l)] = q_{x_1}d\left(\frac{(\gamma_{x_1})^2}{2\beta_{x_1}}\right)^2 + 3q_{x_3}\left(\frac{(\gamma_{x_1})^2}{2\beta_{x_1}}\right)^2, \tag{A35}
\]

\[
E[z_{x_1}(l)z_{x_3}(l)] = \frac{1}{2\beta_{x_1}}\left[(\gamma_{x_1})^2E[z_{x_1}(l)] + a_{x_1}(\gamma_{x_1})^2 + 2q_{x_2}(\gamma_{x_2})^4\right] \tag{A36}
\]

\[
E[z_{x_1}(l)z_{x_4}(l)] = 3q_{x_1}\left(\frac{(\gamma_{x_1})^2}{2\beta_{x_1}}\right)^2 + 15q_{x_3}\left(\frac{(\gamma_{x_1})^2}{2\beta_{x_1}}\right)^3. \tag{A37}
\]

Substituting quantities (A31)–(A37) to Equation (A21), we obtain

\[
\frac{dE[z_{x_1}^2(t,l)]}{dt} = 2\alpha_{x_1}E[z_{x_1}(t,l)] - 2\alpha_{x_1}q_{x_2}\left(\frac{(\gamma_{x_1})^2}{2\beta_{x_1}}\right)^2 \tag{A38}
\]

Hence, from Equation (A38) and equality (A32), we find the nonstationary solution for the second moment of the process \( z_{x_1}(t,l) \)

\[
E[z_{x_1}^2(t,l)] = (\alpha_{x_1})^2t^2 + 2\alpha_{x_1}\alpha_{0_1}t - 2\alpha_{x_1}q_{x_2}\left(\frac{(\gamma_{x_1})^2}{2\beta_{x_1}}\right)t + c_{0_1} \tag{A39}
\]

where \( c_{0_1} \) is an integration constant.

References
1. Shang, Y.; Liu, T.; Wei, Y.; Li, J.; Shao, L.; Liu, M.; Zhang, Y.; Zhao, Z.; Xu, H.; Peng, Z.; et al. Scoring systems for predicting mortality for severe patients with COVID-19. *eClinicalMedicine* 2020, 24, 100426. [CrossRef] [PubMed]
2. Alaraj, M.; Majdalawieh, M.; Nizamuddin, N. Modeling and forecasting of COVID-19 using a hybrid dynamic model based on SEIRD with ARIMA corrections. *Infect. Dis. Model.* 2021, 6, 98–111. [CrossRef] [PubMed]
3. Booth, H.; Tickle, L. Mortality modelling and forecasting: A review of methods. *Ann. Actuar. Sci.* 2008, 3, 3–43. [CrossRef]
4. Cairns, A.J.G.; Blake, D.; Dowd, K. Modelling and management of mortality risk: A review. *Scand. Actuar. J.* 2008, 2–3, 79–113. [CrossRef]
5. 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, 32, 124. [CrossRef] [PubMed]
6. Lee, R.D.; Carter, L. Modeling and forecasting the time series of U.S. mortality. *J. Am. Stat. Assoc.* 1992, 87, 659–671.
7. Lee, R.D.; Miller, T. Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography* 2001, 38, 537–549. [CrossRef]
8. Agadi, R.P.; Talawar, A.S. Stochastic differential equation: An application to mortality data. *Int. J. Res. Granthaalayah* 2020, 8, 229–235. [CrossRef]
9. Christiansen, M.C. Gaussian and Affine Approximation of Stochastic Diffusion Models for Interest and Mortality Rates. *Risks* 2013, 1, 81–100. [CrossRef]
10. Gao, N.; Wang, X.; Liu, J. Dynamics of stochastic SIS epidemic model with nonlinear incidence rates. *Adv. Differ. Equ.* 2019, 2019, 41.
11. Kareem, A.M.; Al-Azzawi, S.N. A stochastic differential equations model for internal COVID-19 Dynamics. *J. Phys. Conf. Ser.* 2021, 1818, 012121. [CrossRef]
12. Zhang, X.; Liao, P.; Chen, X. The negative impact of COVID-19 on live insurers. *Front. Public Health* 2021, 9, 756977. [CrossRef] [PubMed]
13. Zhang, Z.; Zeb, A.; Hussain, S.; Alzahrani, E. Dynamics of COVID-19 mathematical model with stochastic perturbation. Adv. Differ. Equ. 2020, 2020, 451. [CrossRef] [PubMed]

14. Akushevich, I.; Akushevich, L.; Mantov, K.; Yashin, A. Stochastic process model of mortality and aging: Application to longitudinal data. Nonlinear Phenom. Complex Syst. 2003, 6, 515–523.

15. Biffis, E. Affine processes for dynamic mortality and actuarial valuations. Insur. Math. Econom. 2005, 37, 443–468. [CrossRef] [PubMed]

16. Giacometti, R.; Ortolotti, S.; Bertocchi, M. A stochastic model for mortality rate on Italian Data. J. Option. Theory Appl. 2011, 149, 216–228. [CrossRef]

17. Janssen, J.; Skiadas, C.H. Dynamic modelling of life table data. Appl. Stoch. Model. Data Anal. 1995, 11, 35–49. [CrossRef]

18. Russo, V.; Giacometti, R.; Ortolotti, S.; Rachev, S.; Fabozzi, F.J. Calibrating affine stochastic mortality models using term assurance premiums. Insur. Math. Econom. 2011, 49, 53–60. [CrossRef]

19. Hainaut, D.; Devolder, P. Mortality modelling with Lévy processes. Insur. Math. Econom. 2008, 42, 409–418. [CrossRef]

20. Bravo, J.M.; Braumann, C.A. The value of a random life: Modelling survival probabilities in a stochastic environment. Bull. Intern. Stat. Inst. 2007, LXII, 1–4.

21. Bravo, J.M. Modelling mortality using multiple stochastic latent factors. In Proceedings of the 7th International Workshop on Pension, Insurance and Saving, Paris, France, 28–29 May 2009; pp. 1–15.

22. Boukas, E.K. Stochastic Hybrid Systems: Analysis and Design; Birkhauser: Boston, MA, USA, 2005.

23. Liberzon, D. Switching in Systems and Control, Boston, Basel, Berlin; Birkhauser: Basel, Switzerland, 2003.

24. Biffis, E.; Denuit, M. Lee-Carter goes risk-neutral: An application to the Italian annuity market. Giornalle Dell’Istitutoitaliano Degli Attuari 2006, LXIX, 1–21.

25. Biffis, E.; Denuit, M.; Devolder, P. Stochastic mortality under measure changes. Scand. Actuar. J. 2010, 4, 284–311. [CrossRef]

26. Hainaut, D. Multi dimensions Lee-Carter model with switching mortality processes. Insur. Math. Econom. 2012, 47, 409–418.

27. Rossa, A.; Socha, L.; Szymanski, A. Hybrid Dynamic and Fuzzy Models of Mortality, 1st ed.; WUL: Lodz, Poland, 2018.

28. Sliwka, P.; Socha, L. A proposition of generalized stochastic Milevsky-Promislov mortality models. Scand. Actuar. J. 2018, 8, 706–726. [CrossRef]

29. Sliwka, P. Application of the Markov Chains in the Prediction of the Mortality Rates in the Generalized Stochastic Milevsky–Promislov Model. In Trends in Biomathematic 2018; Mondaini, R., Ed.; Springer: Cham, Switzerland, 2019; pp. 191–208. [CrossRef]

30. 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 Computational Science—ICCS 2019; Springer: Cham, Switzerland, 2019; Volume 11538, pp. 435–449. [CrossRef]

31. Sliwka, P.; Socha, L. A Comparison of Generalized Stochastic Milevsky-Promislov Mortality Models with continuous non-Gaussian Filters. In Computational Science—ICCS 2020; Krzhizhanovskaya, V., Ed.; Springer: Cham, Switzerland, 2020; Volume 12140, pp. 348–362. [CrossRef]

32. Sliwka, P. Markov (Set) chains application to predict mortality rates using extended Milevsky-Promislov generalized mortality models. J. Appl. Stat. 2021, 1–21. [CrossRef]

33. Lovrić, M.; Milanović, M.; Stamenković, M. Algorithmic methods for segmentation of time series: An overview. J. Contemp. Econ. Bus. Issues 2014, 1, 31–53.

34. Milevsky, M.A.; Promislov, S.D. Mortality derivatives and the option to annuitise. Insur. Math. Econom. 2001, 29, 299–318. [CrossRef]

35. Khasminski, R.Z.; Zhu, C.; Yin, G. Stability of regime-switching diffusions. Stoch. Proc. Appl. 2007, 117, 1037–1051. [CrossRef]

36. Mao, X. Stability of stochastic differential equations with Markovian switching. Stoch. Proc. Appl. 1999, 79, 45–67. [CrossRef]

37. Sliwka, P.; Swistowska, A. Economic Forecasting Methods with the R Package; UKSW: Warszawa, Poland, 2019; ISBN 978-83-8090-667-9.

38. Statistical Office, Poland. 2020. Available online: http://demografia.stat.gov.pl/bazademografia/TrwanieZycia.aspx (accessed on 23 May 2022).

39. The Human Mortality Database. 2020. Available online: https://mortality.org/hmd/POL/STATS/Deaths-1x1.txt (accessed on 23 May 2022).