Scientific
Research
Publishing

Applications of Dynamic-Equilibrium Continuous Markov Stochastic Processes to Elements of Survival Analysis

Eugen Mamontov[1,2]*, Ziad Taib[1]

[1]Early Clinical Biometrics, Early Clinical Development, IMED Biotech Unit, AstraZeneca, Gothenburg, Sweden
[2]Versati AB, Askim, Sweden
Email: *eugen.mamontov@astrazeneca.com, ziad.taib@astrazeneca.com, eugmav@gmail.com

Abstract

In this article, we summarize some results on invariant non-homogeneous and dynamic-equilibrium (DE) continuous Markov stochastic processes. Moreover, we discuss a few examples and consider a new application of DE processes to elements of survival analysis. These elements concern the stochastic quadratic-hazard-rate model, for which our work 1) generalizes the reading of its Itô stochastic ordinary differential equation (ISODE) for the hazard-rate-driving independent (HRDI) variables, 2) specifies key properties of the hazard-rate function, and in particular, reveals that the baseline value of the HRDI variables is the expectation of the DE solution of the ISODE, 3) suggests practical settings for obtaining multi-dimensional probability densities necessary for consistent and systematic reconstruction of missing data by Gibbs sampling and 4) further develops the corresponding line of modeling. The resulting advantages are emphasized in connection with the framework of clinical trials of chronic obstructive pulmonary disease (COPD) where we propose the use of an endpoint reflecting the narrowing of airways. This endpoint is based on a fairly compact geometric model that quantifies the course of the obstruction, shows how it is associated with the hazard rate, and clarifies why it is life-threatening. The work also suggests a few directions for future research.

Keywords

1. Introduction

Various non-stationary phenomena, which are studied in the natural/life sciences or engineering and described with solutions of deterministic (ordinary or partial) differential equations, develop on the entire time axis. Examples are a steady-state mode of a non-living system or the dynamic-equilibrium (DE) mode of a living system [1] provided that the latter is mature, *i.e.* considered in a time interval sufficiently far away from the birth of the system. However, the transition from deterministic differential equations to their stochastic counterparts, which are often associated with multi-dimensional non-homogeneous continuous Markov stochastic processes, is not free of problems. In particular, in the most common formulation, any of these processes can be solely defined at the time points to the right from the initial time point. The present work focuses on the versions of the processes that are defined on the entire time axis.

These processes are in focus in Section 2. Section 3 considers the stochastic quadratic-hazard-rate (SQHR) model and deals with applications of DE continuous Markov stochastic processes to survival analyses based on this model. Specifications of the SQHR model for problems in chronic obstructive lung disease (COPD) are discussed in Section 4. Section 5 concludes the work and suggests a few directions for future research.

2. Invariant Non-Homogeneous and Dynamic-Equilibrium Continuous Markov Stochastic Processes

We denote a stochastic process with $\chi(\xi, t)$ where $\xi \in \Xi$ is a simple event, $t \in \mathbb{R} = (-\infty, \infty)$ denotes time, and Ξ is the space of simple events. We further let M_ρ be the set of non-homogeneous continuous Markov stochastic processes on the Euclidean space $\mathbb{R}^n$ ($n \geq 1$), that satisfy the following two properties.

- All processes in M_ρ have the same transition probability distribution.
- This distribution is defined at all time points $s, t \in \mathbb{R}$ such that $s < t$, and has the density, $\rho(s, x, t, y)$. (The fact that the transition density ρ is the same for all processes in the set is emphasized with the subscript in the notation "M_ρ").

As is well known, the transition density $\rho(s, x, t, y)$, as a function of $y \in \mathbb{R}^n$ at any fixed *s*, *t*, and $x \in \mathbb{R}^n$, is the conditional probability density of a random variable $\chi(\cdot, t)$ under the condition that $\chi(\xi, s) = x$, and is such that

$$\lim_{t \downarrow s} \rho(s, x, t, y) = \delta(x - y) \tag{1}$$

where $\delta(\cdot)$ is the *n*-dimensional Dirac delta-function.

Definition 1 ([2], Definition 1.11). A stochastic process $\chi \in M_\rho$ specified by its marginal probability density $\rho_{inv}(t, \cdot)$ such that ([2], (1.7.3))

$$\rho_{inv}(t, y) = \int_{\mathbb{R}^n} \rho_{inv}(s, x) \rho(s, x, t, y) \mathrm{d}x, \quad s < t \tag{2}$$

is called invariant. The density $\rho_{inv}(t, \cdot)$ in (2) is called the invariant probability

density of the process. □

If the transition density ρ is stationary, *i.e.* depends on $t-s$ only, rather than on both s and t, all processes in M_ρ are homogeneous (e.g. [3], (2.2.8)). Stationary invariant processes are well known since long ago (e.g. [3], (2.2.11), (9.2.14)).

Proposition 1. Any process $\chi \in M_\rho$ is defined at all t if and only if it is invariant.

Proof. By the definition of the set M_ρ, the marginal probability density of the process χ at time t is determined by the marginal density of the random variable $\chi(\cdot,s)$ where $s < t$ and the transition density $\rho(s,x,t,y)$ by means of the well-known integral representation, which is similar to (2). The mentioned marginal density and marginal density $\rho_{inv}(t,\cdot)$ need not be the same. If they are different, then the process is only defined at $s < t$. However, it is obvious that a process χ is defined at all $t \in \mathbb{R}$ if and only if both marginal-density functions coincide, as is shown in relation (2). This proves the proposition. □

The above consideration generalizes the notion of an invariant process for the case where the process does not need be stationary. This generalization goes back to [4]. Thus, invariant processes are, in general, non-homogeneous.

The example below is probably the simplest example of invariant processes.

Example 1. Consider the case where $n=1$ and the stochastic process χ represents time, t, *i.e.* $\chi(\xi,t)=t$ for all t. Obviously, this process is defined and is continuous at all time points and its marginal density at time t is $\delta(t-y)$. Moreover, $\delta(t-y) = \int_{-\infty}^{\infty} \delta(s-x)\delta[s-x-(t-y)]\,\mathrm{d}x$ at all $s < t$. This, due to Proposition 1, means that the process under consideration is a Markovian one with transition probability density

$$\rho(s,x,t,y) = \delta[s-x-(t-y)] \tag{3}$$

and is an invariant process. Its invariant probability density is $\rho_{inv}(t,y) = \delta(t-y)$. Also note that the transition density in (3) satisfies property (1). □

One of the most important classes of continuous Markov stochastic processes is diffusion stochastic processes (DSPs). A survey of invariant processes of this type can be found in [2] (Chapter 3). However, no criterion for determination of invariant probability densities for non-homogeneous DSPs was known until recently [5]. Here is a brief summary.

We need the notations below.

- D_ρ is the set of all processes in M_ρ, each of which 1) is a DSP with drift n-vector $g(t,y)$ and diffusion $n \times n$-matrix $H(t,y)$ where functions g and H are sufficiently smooth in the entire space $\mathbb{R}^{n+1}$ and 2) corresponds to the transition probability density under consideration, *i.e.* $\rho(s,x,t,y)$, for instance, by means of the related Kolmogorov-forward/Fokker-Planck equation.

(All vectors used in the present work are assumed to be column vectors if not otherwise explicitly stated).

- $\nabla = \left(\partial/\partial y_1, \cdots, \partial/\partial y_n \right)^{\mathrm{T}}$ and ∇^{T} are the gradient and divergence differential operations with respect to the entries of vector x.

- $f(t, y)$ is the so-called Fichera drift. Its entries are

$$f_k(t, y) = g_k(t, y) - (1/2) \times \sum_{l=1}^{n} \partial H_{kl}(t, y)/\partial y_l , \quad k = 1, \cdots, n .$$

We also note that, as is well known (e.g. the paragraph above [2], Section 9.4), any process in D_ρ is a solution of an Itô stochastic ordinary differential equation (ISODE) of the form

$$\mathrm{d}y = g(t, y)\mathrm{d}t + G(t, y)\mathrm{d}w(\xi, t) \tag{4}$$

where $w(\xi, t)$ is the m-vector of the mutually stochastically independent Wiener stochastic processes ($m \geq 1$) and $G(t, y)$ is the $n \times m$-matrix coupled with diffusion matrix $H(t, y)$,

$$H(t, y) = G(t, y)\left[G(t, y) \right]^{\mathrm{T}} . \tag{5}$$

The above notations allow to summarize the aforementioned criterion of [5] as follows.

An invariant probability density of the processes in D_ρ, is the solution $\rho_{\mathrm{inv}}(t, x)$ of the partial-differential-equation system

$$\nabla^{\mathrm{T}} \left\{ f(t, y) - (1/2)H(t, y)\nabla\mu(t, y) \right\} = 0 , \tag{6}$$

$$\partial\mu(t, y)/\partial t + \left\{ f(t, y) - (1/2)H(t, y)\nabla\mu(t, y) \right\}^{\mathrm{T}} \nabla\mu(t, y) = 0 \tag{7}$$

under the condition

$$\int_{\mathbb{R}^n} \exp\left[\mu(t, y) \right]\mathrm{d}y = 1 \tag{8}$$

where $\mu(t, y) = \ln\left[\rho_{inv}(t, y) \right]$.

Example 2. A simple example of application of the above criterion deals with the particular case where $n = 1$, $g(t, y) \equiv -y$, and $H(t, y) \equiv 2$. In this case, the system described by (6)-(7) reduces to

$$\partial^2 \ln\left[\rho_{inv}(t, y) \right]/\partial y^2 = -1 , \tag{9}$$

$$\partial\rho_{inv}(t, y)/\partial t + \left\{ -y - \partial\left\{ \ln\left[\rho_{inv}(t, y) \right] \right\}/\partial y \right\}\partial\rho_{inv}(t, y)/\partial y = 0 . \tag{10}$$

Equation (9) has under condition (8) the solution

$$\rho_{inv}(t, y) = (2\pi)^{-1/2} \exp\left\{ -\left[y - e(t) \right]^2 \big/ 2 \right\} \tag{11}$$

where the function $e(t)$ is to be determined. In order to do that, one substitutes (11) into (10) resulting in the equation $\mathrm{d}e(t)/\mathrm{d}t + e(t) = 0$ having the one-dimensional manifold of trajectories

$$e(t) = e_o \exp\left(-(t - t_o) \right) \tag{12}$$

parameterized with two scalars, t_o and e_o. Application of (12) to (11) leads to the corresponding one-dimensional differentiable manifold of invariant probability densities

$$\rho_{inv}(t, y) = (2\pi)^{-1/2} \exp\left\{ -\left[y - e_o \exp\left(-(t - t_o) \right) \right]^2 \big/ 2 \right\} . \tag{13}$$

The members of manifold (13) at two values of e_o, $e_o = 0$ and $e_o = 1$, and one value of t_o, $t_o = 0$, are

$$\rho_{inv.0}(t, y) = (2\pi)^{-1/2} \exp(-y^2/2),$$
(14)

$$\rho_{inv.1}(t, y) = (2\pi)^{-1/2} \exp\left\{-\left[y - \exp(-t)\right]^2/2\right\}.$$
(15)

These are indicated in [4] without derivation (see also [5], (15). In contrast, expression (13) is derived and, thus, generalizes results (14) and (15) of [4].

Paper [4] also draws attention to the non-uniqueness of the invariant density. Indeed, the one-dimensional manifold (13) of such densities shows that there can be a continuum of invariant processes that correspond to the same transition probability density. In other words, uniqueness is, in general, not fulfilled. More specifically, M_ρ can contain a continuum of invariant processes, each of which corresponds to its individual invariant probability density. □

Remark 1. If process $\chi \in M_\rho$ is invariant with one or another invariant density, say, $\rho_{inv}(t, y)$, then the process is completely described with this density. Consequently, there is no need to involve transition density ρ in the analysis of the process. This fact consistently and substantially simplifies the related theoretical and practical studies. □

It is also worth to notice that, at any $t \in \mathbb{R}$, the random variable $\chi(\cdot, t)$ in Remark 1 is described with probability density $\rho_{inv}(t, \cdot)$. This density depends on time t as a parameter. It can also depend on other parameters. If these do not depend on t (e.g. are similar parameters t_o and e_o in (13)), then temporal samples of the process χ can be treated by means of common statistical methods (e.g. the maximum-likelihood technique). If the parameters depend on t, then applicable statistical methods need to be indicated or, if necessary, developed.

A discussion on non-homogeneous invariant DSPs can be found in [5]. To our knowledge, in the general case of these processes, there is no method for derivation of exact invariant probability densities. One of the techniques that can provide approximate densities is based on the so-called detailed balance (DB) conditions (e.g. [2], (1.12.13) as well as Sections 3.5.2 and 3.5.3). If a DSP is stationary and the DB conditions are met, then the DB solution presents the exact invariant density. Otherwise, the DB solution does not exist but the DB approximation can be regarded as a quasi-stationary one for the corresponding solution. The aforementioned method is presented in ([2], Section 3.5.5; see also Section 3.5.4). It provides a simplified DB approximation for the invariant probability density. Theorem 3.2 in ([2], Section 3.5.5) proves the two-sided estimation for this density when both sides are Gaussian densities.

Apparently, the most important example of the invariant processes is the dynamic-equilibrium (DE) ones [1]. The DE probability density corresponding to transition density $\rho(s, x, t, y)$ is the one determined with limit relation (see [1], (10))

$$\rho_{DE}(t, y) = \lim_{s \to -\infty} \rho(s, x, t, y)$$
(16)

where the limit is uniform in (t, y). Relation (16) presumes that the limit function on the right-hand side does not depend on x. This is the very property that enables one to associate this function with an equilibrium. Since the function generally depends on t, the equilibrium is generally dynamic. Relation (16) is inspired by the well-known similar limit relation of R.Z. Has'minskii ([6], (9.12) on p. 139) in the case where the equilibrium is time-independent. One can easily see that the DE density determined with (16) is an invariant density. In order to realize that, it is sufficient to apply the Chapman—Kolmogorov equation $\rho(u, z, t, y) = \int_{\mathbb{R}^n} \rho(u, z, s, x) \rho(s, x, t, y) \, dx$, $u < s < t$, and pass in it to the limit as $u \to -\infty$. The invariant process $\chi \in M_\rho$ determined with the invariant density, which is the DE one, is termed the DE process.

Remark 2. According to definition (16) of the DE density, M_ρ can contain at most one DE process. Since it (if exists) is invariant, the advantage indicated in Remark 1, is equally applicable to it. □

Example 3. Passing in (3) to the limit as $s \to -\infty$ and taking into account (16), one obtains the density $\rho_{DE}(t, y) = \delta(t - y)$, which, as follows from Example 1, coincides with the invariant density. Thus, time is the DE process.

Notably, since the expectation corresponding to DE density $\delta(t - y)$, is t, the expectation of a DE process does not need be uniformly bounded in t.

Example 4. Due to the well-known result (e.g. [3], (9.4.8) and limit relation (16), stationary density (14) is a DE density. It corresponds to the DSPs considered in Example 2. □

If in M_ρ there exists the DE process, it is of special importance in connection with convergence of processes that belong to M_ρ. Convergence of stochastic processes in distribution is the most basic concept of stochastic convergence because it follows from other concepts of stochastic convergence. The well-known summary (e.g. [3], Point d) on p. 13)

$$\text{convergence in the } j\text{th mean} \Rightarrow \text{convergence in}$$
$$\text{the } i\text{th mean where } i \leq j \Rightarrow \text{almost certain convergence} \Rightarrow$$
$$\text{stochastic convergence} \Rightarrow \text{convergence in distribution} \quad (17)$$

outlines the relationships between different concepts of stochastic convergence. One can show, under rather mild conditions, that if the above DE process exists, then processes that belong to M_ρ converge in distribution, *i.e.* the marginal probability density of any of them at time t converges to the DE density $\rho_{DE}(t, \cdot)$ as $t \to \infty$. The t-dependent relation (16) and the above mentioned property are in line with the idea, which is formulated by R.Z. Has'minskii ([6], Remark 2 on pp. 140-141) and goes back to the result of A.M. Il'in and R.Z. Has'minskii ([4], Theorem 5).

In connection with the concept of dynamic equilibrium, one can note the following.

A physical equilibrium is independent of t, *i.e.* static. A DE need not be static. Thus, it need not be a physical equilibrium.

A non-living system may have physical equilibrium(s). It may also have modes,

which can be interpreted as DEs. These modes are known as the steady-state ones.

A living system cannot have static, t-independent equilibrium(s) but, as rule, has a DE.

3. Application of Dynamic-Equilibrium Processes to the Survival Analyses Based on the Stochastic Quadratic-Hazard-Rate Model

The remaining of the present work discusses a new application of the notion of DE DSP.

In statistics, there are different quantitative approaches to survival analysis (e.g. [7]). Recently, interest has been in joint modeling, *i.e.* modeling longitudinal independent variables, or covariates, on top of time to event, whereas the covariates are modeled using a mixed-effects-model approach. The main idea of joint modeling is to complement conditional probability distributions, which are common in evolutions of multi-component populations and are conditioned with fixed values of deterministic parameters of the populations, with probability distributions of the component-individual parameters, thereby generalizing the deterministic parameters to their stochastic versions and taking into account their stochastic variability. Parameters that are not component-individual, remain deterministic.

Joint modeling attracts growing attention in statistics. However, its main idea was developed in mathematical physics rather long ago (e.g. [8] [9]). It was motivated by needs in analysis of multi-component populations with a large number of components. This development treats the multi-componentness of the population as multi-modality of the probability density of stochastic parameters. It describes the parameters with the McKean—ISODE ([8], Section 6.2), which is even more general than ISODE (4) (not to mention linear Equation (18) below). Another example is a mean-field generalization [10] of the classical, Bogolyubov—Born—Green—Kirkwood—Yvon statistical mechanics. This generalization is free from the thermodynamic-limit assumption, and is more compact and flexible than the classical counterpart.

One of the aforementioned approaches is proposed in [11]. It, in particular, presumes that the hazard-rate-driving independent (HRDI) variables (also known as "risk factors" or "covariates"; e.g. see [12]) are the entries of some vector $y \in \mathbb{R}^n$ described with a particular case of ISODE (4) where $g(t, y) \equiv A(t) y + a(t)$ and $G(t, y) \equiv B(t)$, *i.e.* (see [11], (1)), namely

$$\mathrm{d}y = \left[A(t) y + a(t) \right] \mathrm{d}t + B(t) \mathrm{d}w(\xi, t). \tag{18}$$

In this equation, the functions *a*, *A*, and *B* are defined on the entire axis $\mathbb{R}$, and t is interpreted as the age of a person whose survival is described with the HRDI vector *y*. Notice that ISODE (18) is linear in the narrow sense ([3], Section 8.2) and that it was suggested in [11] four years later the above McKean-ISODE of [8]. According to [11], the corresponding hazard rate is (see [11], (2))

$$\lambda(t,y) = \lambda_0(t) + (1/2)\big[y - \zeta(t)\big]^{\mathrm{T}} F(t)\big[y - \zeta(t)\big] \qquad (19)$$

where $\lambda_0(t) \geq 0$ is the baseline hazard rate and $\zeta(t)$ is described as follows

$\zeta(t)$ is the so-called "optimal" trajectory of solutions of (18)
(which is, however, not endowed with any modeling representation). (20)

$F(t)$ being a $n \times n$-matrix is symmetric and positive-definite uniformly in t. The advantages of model (19) under condition (18) compared to the well-known Cox proportional hazard-rate model are discussed in detail in ([11], p. 539).

Remark 3. Reference [11] emphasizes that $\zeta(t)$ need not coincide with $-\big[A(t)\big]^{-1}/a(t)$, which is in fact the zero-drift approximation for y. (It is denoted with "$f_1(t)$" in [11]). Thus, in general, these vectors are different. Moreover ([11], the text below (2)), any stochastic solution y of (18) can follow a deterministic trajectory $\zeta(t)$. In more precise terms, this behavior means that any solution y converges to $\zeta(t)$ as time tends to infinity. □

The discussion in ([11], p. 539) stresses that typical dependences of hazard rates on the HRDI variables are J- or U-shaped. A particular, exponential case of the J-shaped dependences can be exemplified with the Cox hazard model. Along with this, the U-shaped dependences are meaningful for many HRDI variables, for instance, human body temperature, weight, volume, and surface, blood pressure, serum potassium concentration, serum calcidiol (or calcifediol) concentration and other characteristics. Each of them is in a system-relevant bounded interval. Values in the middle part of the interval correspond to the lowest hazard rates. However, if a value approaches any of the two bounds, the hazard rate rapidly increases. The simplest version of U-shaped hazard rates is quadratic. Also note that both J- and U-shaped dependences are convex.

Still, it may seem that the quadratic expression (19) is nothing but a modeling assumption. However, it has a consistent ground. Indeed, it follows from the exact representation, which is valid under rather mild conditions ([13], Section 3.3.11), that

$$\begin{aligned}\lambda(t,y) = &\lambda(t,\zeta(t)) + \nabla\lambda(t,\zeta(t)) \\ &+ (1/2)\big[y - \zeta(t)\big]^{\mathrm{T}} G(t,\zeta(t),y - \zeta(t))\big[y - \zeta(t)\big]\end{aligned} \qquad (21)$$

where the term ∇ is described in the second bullet above Equation (4) and

$$G(t,\zeta(t),y - \zeta(t)) = 2\int_0^1 (1-u)\nabla\nabla^{\mathrm{T}}\lambda\big[t,\zeta(t) + u(y - \zeta(t))\big]\mathrm{d}u. \qquad (22)$$

Notice that the second multiplier in the integrand in (22) is the Hesse matrix. Comparing (19) and (21), one obtains the relations

$$\lambda_0(t) = \lambda(t,\zeta(t)), \qquad (23)$$

$$\nabla\lambda(t,\zeta(t)) = 0, \qquad (24)$$

$$F(t) = G(t,\zeta(t),y - \zeta(t))\big|_{y=\zeta(t)} = \nabla\nabla^{\mathrm{T}}\lambda(t,\zeta(t)). \qquad (25)$$

These elucidate a number of topics.

First, equalities (23) and (25) express the vector $\lambda_0(t)$ and matrix $F(t)$ in the model described by (19) in terms of the hazard-rate function $\lambda(t, y)$. Moreover, since $\lambda_0(t)$ is the baseline hazard rate, relation (23) shows that the function $\zeta(t)$ (cp., (20)) can be interpreted as the baseline value of the variable y. Thus, it appeared that the "optimal" trajectory of solutions of Equation (18) in the SQHR model is nothing but the baseline value of the variable of this equation.

Next, as follows from (24), the baseline value $\zeta(t)$ of y can be determined as a solution of the equation

$$\nabla \lambda(t, y) = 0 . \tag{26}$$

Finally, in view of ([13], Section 3.4.4) and equality (21), the scalar $\lambda(t, y)$, as a function of y, is strictly convex if and only if the matrix $G(t, \zeta(t), y - \zeta(t))$ is positive definite. In that case, Equation (26) has as unique solution, namely the baseline value $\zeta(t)$ of y, which corresponds to the only local minimum of $\lambda(t, y)$ in y. This minimum is also the global one.

The above properties present the modeling settings for the function $\zeta(t)$ and clarify its role.

Hazard rate (19) and the HRDI variables (*i.e.* entries of vector y in (18)) are key ingredients of the survival function (e.g. [7]) and, thus, important elements of survival analysis. Relations (18) and (19) constitute the stochastic quadratic-hazard-rate (SQHR) model of [11] [14] [15]. The entire approach of [11] is developed for aging-related changes in a human organism. The subsequent work [14] generalizes the approach of [11] to a joint-modeling paradigm (e.g. see ([14], (3) and the text on "Z" below it)) in connection with human aging, health, and longevity. In these settings, the population and its components correspond to a group of persons and the persons in this group, respectively. Also, the component-individual parameters are represented with parameters that are individual to the persons, whereas the group parameters are the same for all persons. This approach is discussed in-depth in connection with predicting health and survival in [15].

Remark 4. In connection with clinical-trial applications, one should note an important advantage of ISODE models. The form of the probability densities of solutions of ISODE (18) is well known (e.g. [3], (8.2.9), (8.2.10), (9.2.12), and (9.4.8)). In statistical problems, ISODEs with known distributions enable consistent and systematic reconstruction of missing data by means of the Gibbs sampling (e.g. [16]). Missing data are common in many areas, in particular, in clinical trials due to attrition, participants drop out, and various types of censoring. □

No matter what the area of application of the SQHR model is, the key principles remain valid. In particular, the approach emphasizes the role of the property associated with a special solution of ISODE (18). This property is known as the "mean-reverting" one (mentioned in [15], pp. 228/3-228/4). Strictly speaking,

it is only applicable to Equation (18) in the case where functions *a*, *A*, and *B* are independent of *t*, *i.e.* solutions of this equation are the Ornstein—Uhlenbeck processes (e.g. [3], Section 8.3). However, it is possible to generalize the "mean-reverting" property to the present settings where *a*, *A*, and *B* are *t*-dependent with the help of well-known results and the above concept of a DE solution. This can be accomplished in the following way.

Let there exist numbers $\gamma > 0$ and $\Gamma > 0$ such that

$$\left\| C(t,s) \right\| \le \Gamma \exp\left[-\gamma(t-s) \right]$$

for all $s < t$ where $C(t,s)$ is the Cauchy matrix of the ordinary differential equation (ODE) $\mathrm{d}y/\mathrm{d}t = A(t)y$. Then one can, on the basis of ([3], Point c) in Section 1.4, (11.2.19), and (8.2.4)), show that the "mean-revering" property of equation (18) is the property that any solution of this equation converges to its DE solution, y_{DE}, in quadratic mean as $t \to \infty$. This solution presents the DSP with transition probability density $\rho(s,x,t,y)$ and marginal probability density $\rho_{DE}(t,y)$. In view of ([3], (8.2.9)), the density $\rho_{DE}(t,y)$ is Gaussian, with expectation vector

$$e_{DE}(t) = \int_{-\infty}^{t} C(t,s)a(s)\mathrm{d}s = \int_{\mathbb{R}^n} y \rho_{DE}(t,y)\mathrm{d}y \tag{27}$$

and variance matrix

$$\begin{aligned} V_{DE}(t) &= \int_{-\infty}^{t} C(t,s)\left[B(s) \right]\left[B(s) \right]^{\mathrm{T}}\left[C(t,s) \right]^{\mathrm{T}}\mathrm{d}s \\ &= \int_{\mathbb{R}^n} \left[y - e_{DE}(t) \right]\left[y - e_{DE}(t) \right]^{\mathrm{T}} \rho_{DE}(t,y)\mathrm{d}y \end{aligned} \tag{28}$$

Notably, according to (17), the above mentioned convergence in quadratic mean (see the text above (27)) implies convergence in distribution. One of the outcomes of the latter convergence is density $\rho_{DE}(t,y)$.

Model (18), (19) is not sufficiently complete. For example, the lack of modeling representations for the "optimal" trajectory $\zeta(t)$ (see the parentheses in (20)), is not resolved in [11] [14] [15]. This gap is partly filled with the results discussed in the text on (21)-(26). In addition to that, comparison of the aforementioned convergence in quadratic mean and the convergence noted in Remark 3 shows that

$$\zeta(t) = e_{DE}(t), \tag{29}$$

i.e., the baseline value $\zeta(t)$, or the "optimal" trajectory, of solutions of ISODE (18) is provided by the expectation $e_{DE}(t)$ of the DE solution (see (27)). Relation (29) agrees with the difference of $\zeta(t)$ from the zero-drift approximation for *y* as is emphasized in Remark 3. One can also show that the expectation of the second term on the right-hand side of (19) under condition (29) is, in the infinite-time limit, $(1/2)\mathrm{tr}\left[F(t)V_{DE}(t) \right]$ (or, equivalently, $(1/2)\mathrm{tr}\left[V_{DE}(t)F(t) \right]$) where $\mathrm{tr}(\cdot)$ is the trace of a (square) matrix.

4. Specifications of the Stochastic Quadratic-Hazard-Rate Model for Problems in Obstructive Lung Diseases

The SQHR model of [11] [14] [15] can be applied to various areas, whereas we

choose to apply it to clinical trials of treatments against obstructive lung diseases (OLDs), *i.e.* diseases that cause lower airway obstruction (e.g. [17]). These include asthma, bronchiectasis, bronchitis, and COPD (e.g. [18]). Airway obstruction is a blockage of respiration in the airway (e.g. [17]). For example, the citation below is outlined in ([19], p. 1342 and Figure 1) and explained in ([20], pp. 448-449 and Figure 7) in more detail.

"*Although the measurements of FEV$_1$ and FEV$_1$/FVC (forced expiratory vital capacity, or volume of air expired between full inspiration and residual lung volume) provide a reliable way of diagnosing airflow limitation and classifying COPD severity, they cannot separate the precise contribution of either small-airway obstruction or emphysematous destruction to the airflow limitation in individuals with COPD. However, direct measurements of pressures and flows within the lung indicate that the smaller bronchi and bronchioles less than 2 mm in diameter are the major sites of airway obstruction in COPD. Moreover, the reduced expiratory flow that defines COPD results from reduction of the lumen by peribronchiolar fibrosis, thickening of the small-airway walls, and occlusion of the lumen of the small conducting airways by exudate containing mucus* [20]."

The above histopathological results on the major sites enable one to reveal the physiological and biological meaning of the hazard rate in the case of OLDs (in particular, COPD) in the compact form. This form, following the definition of the hazard rate, includes relations

$$\lambda_k(t) = -\left[\mathrm{d}\Phi_k(t)/\Phi_k(t)\right]/\mathrm{d}t, \quad k = 1, 2, \cdots, m \tag{30}$$

where t is, as in model (18), the age, $m \geq 1$ is the (integer) number of the terminal or respiratory (also known as transitional) bronchioles in the lung (e.g. [21], Section II and Figure 3), $\Phi_k(t)$ is the area of the cross-section of the lumen (non-occluded or occluded) of the kth bronchiole (e.g. [22], Section "Results" and Figure 1), $\lambda_k(t)$ is the kth-bronchiole hazard rate (assumed to be independent of $\Phi_k(t)$), and $\mathrm{d}\Phi_k(t)/\Phi_k(t)$ is the infinitesimal relative change of $\Phi_k(t)$. Linear ODEs (30) are complemented with expressions for the survival functions of the bronchioles, $S_k(t)$, and the survival function of the entire bronchiole system, $S(t)$ (e.g. [7])

$$S_k(t) = \Phi_k(t)/\Phi_k(0) = \exp\left[-\int_0^t \lambda_k(s)\mathrm{d}s\right], \quad k = 1, 2, \cdots, m \tag{31}$$

$$S(t) = 1 - \prod_{k=1}^m \left[1 - S_k(t)\right]. \tag{32}$$

In the particular case where all λ_k are the same and independent of *t*, one can readily check that survival function (32) corresponds to the exponentiated exponential distribution with the parameters independent of *t*. Also note that expression (32) is well known in reliability theory where it is associated with the so-called parallel systems, and the survival functions are termed the reliability functions. One can also show that this expression corresponds to *m* stochastically independent random variables. They are associated with the component-individual hazard rates by means of (31).

The role of the bronchiole-specific hazard rates (30) is illustrated in the following remark.

Remark 5. In the particular case where, firstly, the inner surface of the kth bronchiole is circular cylindrical of the non-occluded radius $R_k(t)$ and, secondly, this surface is covered by the occlusive layer of highly viscous mucus (e.g. [21], Figure 10 and p. 543) of the thickness $T_k(t)$, the relation $\Phi_k(t) = -\pi \left[R_k(t) - T_k(t) \right]^2$ holds and one can readily show that a linear ODE in system (30) applied to the above bronchiole reduces to a linear ODE $\mathrm{d}\left[R_k(t) - T_k(t) \right]/\mathrm{d}t = -\left[\lambda_k(t)/2 \right]\left[R_k(t) - T_k(t) \right]$ for $T_k(t)$. The corresponding initial condition is $T_k(0) = 0$. The solution of this initial-value problem is

$$T_k(t) = R_k(t) - R_k(0)\exp\left[-(1/2)\int_0^t \lambda_k(s)\,\mathrm{d}s \right]. \tag{33}$$

Expression (33) explicitly shows the following. Firstly, the bronchiole-individual hazard is the occlusive-layer thickness. Secondly, this hazard is obstructive because it, in the course of the age, grows. And, thirdly, this growth is life-threatening because the thickness, in the infinite-age limit, tends to the radius $R_k(t)$ of the non-occluded lumen, thereby completely obstructing the lumen with the occlusive-layer mucus. Moreover, relation (33) is one of apparently few results that discover informal meaning of the concept of a hazard rate (which is the term that can, however, be formally determined for any probability density). Indeed, (33) explicitly indicates the informal, geometric meaning inherent in the specific biophysical, bronchiole/occlusive-layer system.

The model described by (30)-(32) and Remark 5 present the biostatistical reading of a single component of OLDs, the narrowing of airways. This component is a necessary phenomenon in all OLDs including COPD. With respect to COPD, the model corresponds to the core measurement results of [23]. The range of the related measurement data is contributed with advances in COPD imaging [24]. These outcomes can help to better focus research on further modeling of hazard rate $\lambda(t)$ (see (30)) for OLDs and interpretation of the results in a less arbitrary way.

The quantity $\Phi(t)$ is a biophysical characteristic. Along with this, clinical studies often focus on exacerbations of COPD, which are presumably coupled with the occlusion-caused lumen narrowing (e.g. [20], Section "Acute exacerbation") and are usually formulated in terms of patient symptoms and medical signs (e.g. test results). In this case, modeling of $\lambda(t)$ can be based on (19) but the HRDI-variable vector, say, v, can include not only variables on the entire axis $\mathbb{R}$ but also variables that are non-negative or represent membership functions (MFs) (e.g. [25]). This means that v is in a bounded set $R \subseteq \mathbb{R}^n$ rather than in the entire Euclidean space $\mathbb{R}^n$. However, the latter two types of variables can be represented with entries in $\mathbb{R}$ of vector y. Examples of these representations, which presume that a scalar v_* is an entry of a vector v and a scalar y_* is the corresponding entry of a vector y, are the following:

$$v_* = y_*, \text{ if } v_* \text{ is in } \mathbb{R}, \tag{34}$$

$$v_* = (1/2)\left\{\sqrt{1+\left[s_*\left(y_*-p_*\right)\right]^2}+s_*\left(y_*-p_*\right)\right\}, \text{ if } v_* \text{ is non-negative}, \quad (35)$$

$$v_* = (1/2)\left\{1+s_*\left(y_*-p_*\right)\Big/\sqrt{1+\left[s_*\left(y_*-p_*\right)\right]^2}\right\}, \text{ if } v_* \text{ is an MF variable} \quad (36)$$

where $s_* > 0$ and p_* are the scaling and positioning coefficient. In the latter case, v_* can also represent a categorical variable (a particular case of an MF one) such as any of the following three variables (e.g. [26]):

- scores 0, 1, ..., 40 on the COPD assessment test;
- grades 1, 2, 3, and 4 according to the GOLD;
- grades 1, 2, 3, 4, and 5 on the British Medical Research Council short-ness-of-breath test;

after a proper transformation of the actual values into values in interval $[0,1]$. In general, scoring/grading systems should also take into account weight loss and muscle weakness, as well as the presence of other diseases. In agreement with this, a use of a number of non-negative and MF variables (including the GOLD grading) is in the core of clinical studies reported in [27] and other works. Also note that Remark 5 draws attention to the question on how specifically each entry of the HRDI-variable vector v can affect the bronchiole-individual hazard rate in its role indicated in (33).

Vector $v \in R \subseteq \mathbb{R}^n$ can formally be described with an equation similar to ISODE (4), namely

$$\mathrm{d}v = q\left(t,v\right)\mathrm{d}t + Q\left(t,v\right)\mathrm{d}w\left(\xi,t\right). \quad (37)$$

Since vector-function $q\left(t,v\right)$ and matrix-function $Q\left(t,v\right)$ in (37) allow for variables of three different types (see (34)-(36)), Equation (37) can hardly be linear in v. In general, models and methods for ISODEs cannot be applied to (37) directly because ISODEs are usually considered in the entire Euclidean space $\mathbb{R}^n$ rather than in bounded set $R \subseteq \mathbb{R}^n$ (e.g. [3], Section 6). In order to resolve this matter, one can, in Equation (37), pass from vector v to vector $y \in \mathbb{R}^n$ by means of changes of variables (34)-(36) and Itô's theorem (e.g. [3], (5.3.8)-(5.3.10)). The resulting equation is an ISODE for vector y in the form (4).

Remark 6. The generalized SQHR model suitable for methods of DSP analysis comprises nonlinear ISODE (4) for vector y of the HRDI-variable representatives (such as terms y_* in (34)-(36)), expressions (21)-(25) under specification (29) and the three properties, which are formulated in the text below (25) and, because of (23)-(25), admit (19) as a particular case. Nonlinear Equation (4), in contrast to its predecessor the linear Equation (18), takes into account not only the HRDI variables in the entire axis but also the ones that are non-negative or MF/categorical. Note, however, that, in the present case, *i.e.* the case of nonlinear ISODE (4), the term $e_{DE}\left(t\right)$ in (29) is the expectation of the DE solution of the mentioned nonlinear ISODE. □

As is above in Remark 6, Equation (4) is generally nonlinear in y. It is desirable that Equation (4) inherits the important advantage of linear Equation (18)

emphasized in Remark 4. However, in the problems of interest, the number of variables (or scalar equations) in system (4), n, is of the order of a few units or tens, or greater. How can one efficiently obtain probability distributions of such high dimensions from nonlinear ISODE (4)? Work [2] focuses on high-dimensional DSPs and ISODEs, and includes a few approaches to the problem. The simplest one provides an approximate model that comprises:

- a non-autonomous nonlinear ODE system of the second order ([2], (2.3.9), (2.3.10), and Section 2.3.2) with the initial conditions for entries of the expectation vector $e(t)$ ([2], (A.6) and (2.3.12)); the unique advantage of this system is that it includes the influence of diffusion matrix (5) on the expectation;

- a non-autonomous linear ODE system of the first order ([2], (2.5.8), (2.5.13), (2.5.14), and Section 2.5.2) with the initial conditions for entries of variance matrix $V(t)$ ([2], (1.6.14), (1.6.8)).

The DE versions of the expectation and variance are provided by numerical integration of both systems at time intervals, which is sufficiently far away from the initial time point. One can use a marginal probability density for y, which is the DE density and Gaussian, with the corresponding DE expectation and variance. These outcomes are approximate representations. Reference [2] includes other approaches that are more precise, such as the ones based on the Schauder bases of function Banach spaces and differential-quadrature/stochastic-adaptive-interpolation method.

5. Concluding Remarks

The above analysis of the SQHR model of [11] [14] [15] generalizes the reading of its Itô ISODE for the HRDI variables. Moreover, it specifies key properties of the hazard-rate function. In particular, it reveals that the baseline value of the HRDI variables is the same as the so-called "optimal" trajectory in the SQHR model and is the expectation of the DE solution of the ISODE. The work also suggests practical settings for obtaining multi-dimensional probability densities necessary for consistent and systematic reconstruction of missing data by the Gibbs sampling, and further develops the corresponding line of modeling. The resulting advantages are emphasized in connection with general survival analysis and the statistical framework for clinical trials of new treatments. The present work also proposes a use of endpoints reflecting the narrowing of airways, which is a major component of obstructive lung diseases (including COPD). This endpoint is based on a fairly compact geometric model that quantifies the course of the obstruction, shows how it is associated with the hazard rate, and clarifies why it is life-threatening.

Directions for Future Research can Include

- Problem-specific derivations and specifications of the hazard-rate functions and ISODEs for the HRDI variables in various applications;
- Implementation and improvement of the computational scenario formulated

at the end of Section 4;

- Development of a practical statistical framework that would enable an aggregated, fairly compact treatment of the suggested bronchiole-system analysis for clinical trials of new treatments of COPD and other OLDs.

The resulting outcomes will further enrich the field of continuous stochastic approaches to survival analysis.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Mamontov, E. (2008) Dynamic-Equilibrium Solutions of Ordinary Differential Equations and Their Role in Applied Problems. *Applied Mathematics Letters*, **21**, 320-325. https://doi.org/10.1016/j.aml.2007.02.031

[2] Mamontov, Y.V. and Willander, M. (2001) High-Dimensional Nonlinear Diffusion Stochastic Processes. World Scientific, River Edge, NJ. https://doi.org/10.1142/4494

[3] Arnold, L. (1974) Stochastic Differential Equations: Theory and Applications. John Wiley Sons, New York.

[4] Il'in, A.M. and Has'minskiĭ, R.Z. (1965) Asymptotic Behavior of Solutions of Parabolic Equations and an Ergodic Property of Nonhomogeneous Diffusion Processes. *American Mathematical Society Translations: Series* 2, **49**, 241-268.

[5] Mamontov, E. (2005) Nonstationary Invariant Distributions and the Hydrodynamic-Style Generalization of the Kolmogorov-Forward/Fokker-Planck Equation. *Applied Mathematics Letters*, **18**, 976-982. https://doi.org/10.1016/j.aml.2004.06.027

[6] Has'minskiĭ, R.Z. (1980) Stochastic Stability of Differential Equations. Sijthoff Noordhoff, Alphen aan den Rijn. https://doi.org/10.1007/978-94-009-9121-7

[7] https://en.wikipedia.org/wiki/Survival_analysis

[8] Bellomo, N., Mamontov, E. and Willander, M. (2003) The Generalized Kinetic Modelling of a Multicomponent "Real-Life" Fluid by Means of a Single Distribution Function. *Mathematical and Computer Modelling*, **38**, 637-659. https://doi.org/10.1016/S0895-7177(03)90033-1

[9] Willander, M., Mamontov, E. and Chiragwandi, Z. (2004) Modelling Living Fluids with the Subdivision into the Components in Terms of Probability Distributions. *Mathematical Models and Methods in Applied Sciences*, **14**, 1495-1520. https://doi.org/10.1142/S0218202504003702

[10] Mamontov, E. (2009) Ordinary Differential Equation System for Population of Individuals and the Corresponding Probabilistic Model. *Mathematical and Computer Modelling*, **49**, 1551-1562. https://doi.org/10.1016/j.mcm.2008.09.010

[11] Yashin, A.I., *et al.* (2007) Stochastic Model for Analysis of Longitudinal Data on Aging and Mortality. *Mathematical Biosciences*, **208**, 538-551. https://doi.org/10.1016/j.mbs.2006.11.006

[12] https://en.wikipedia.org/wiki/Dependent_and_independent_variables#Statistics_synonyms

[13] Ortega, J.M. and Rheinboldt, W.C. (1970) Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York.

[14] Yashin, A.I., *et al.* (2012) The Quadratic Hazard Model for Analyzing Longitudinal

Data on Aging, Health, and the Life Span. *Physics of Life Reviews*, **9**, 177-188. https://doi.org/10.1016/j.plrev.2012.05.002

[15] Arbeev, K.G., *et al.* (2014) Joint Analyses of Longitudinal and Time-to-Event Data in Research on Aging: Implications for Predicting Health and Survival. *Frontiers in Public Health*, **2**, 228/1-228/12. https://doi.org/10.3389/fpubh.2014.00228

[16] Struthers, C.A. and McLeish, D.L. (2011) A Particular Diffusion Model for Incomplete Longitudinal Data: Application to the Multicenter AIDS Cohort Study. *Biostatistics*, **12**, 493-505. https://doi.org/10.1093/biostatistics/kxq079

[17] https://en.wikipedia.org/wiki/Airway_obstruction

[18] https://en.wikipedia.org/wiki/Obstructive_lung_disease

[19] Decramer, M., Janssens, W. and Miravitlles, M. (2012) Chronic Obstructive Pulmonary Disease. *The Lancet*, **379**, 1341-1351.
https://doi.org/10.1016/S0140-6736(11)60968-9

[20] Hogg, J.C. and Timens, W. (2009) The Pathology of Chronic Obstructive Pulmonary Disease. *Annual Review of Pathology*, **4**, 435-459.
https://doi.org/10.1146/annurev.pathol.4.110807.092145

[21] Hogg, J.C., Paré, P.D. and Hackett, T.-L. (2017) The Contribution of Small Airway Obstruction to the Pathogenesis of Chronic Obstructive Pulmonary Disease. *Physiological Reviews*, **97**, 529-552. https://doi.org/10.1152/physrev.00025.2015

[22] Hogg, J.C., *et al.* (2004) The Nature of Small-Airway Obstruction in Chronic Obstructive Pulmonary Disease. *The New England Journal of Medicine*, **350**, 2645-2653. https://doi.org/10.1056/NEJMoa032158

[23] Koo, H.K., *et al.* (2018) Small Airways Disease in Mild and Moderate Chronic Obstructive Pulmonary Disease: A Cross-Sectional Study. *The Lancet Respiratory Medicine*, **6**, 591-602. https://doi.org/10.1016/S2213-2600(18)30196-6

[24] Thiboutot, J., *et al.* (2018) Current Advances in COPD Imaging. *Academic Radiology*. https://doi.org/10.1016/j.acra.2018.05.023

[25] https://en.wikipedia.org/wiki/Membership_function_(mathematics)

[26] https://en.wikipedia.org/wiki/Chronic_obstructive_pulmonary_disease#Severity

[27] Liu, D., *et al.* (2015) Prediction of Short Term Re-Exacerbation in Patients with Acute Exacerbation of Chronic Obstructive Pulmonary Disease. *International Journal of COPD*, **10**, 1265-1273.

Abbreviations

COPD: chronic obstructive pulmonary disease
DB: detailed balance
DE: dynamic equilibrium
DSP: diffusion stochastic process
FEV_1: forced expiratory volume in 1 second
FVC: forced vital capacity
HRDI: hazard-rate-driving independent
ISODE: Itô stochastic ordinary differential equation
MF: membership function
ODE: ordinary differential equation
OLD: obstructive lung diseases
SQHR: stochastic quadratic-hazard-rate