



HAL
open science

Inferring the effect of therapy on tumors showing stochastic Gompertzian growth

Giuseppina Albano, Virginia Giorno, Patricia Román-Román, Francisco Torres-Ruiz

► **To cite this version:**

Giuseppina Albano, Virginia Giorno, Patricia Román-Román, Francisco Torres-Ruiz. Inferring the effect of therapy on tumors showing stochastic Gompertzian growth. *Journal of Theoretical Biology*, 2011, 276 (1), pp.67. 10.1016/j.jtbi.2011.01.040 . hal-00682407

HAL Id: hal-00682407

<https://hal.science/hal-00682407>

Submitted on 26 Mar 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Author's Accepted Manuscript

Inferring the effect of therapy on tumors showing stochastic Gompertzian growth

Giuseppina Albano, Virginia Giorno, Patricia Román-Román, Francisco Torres-Ruiz

PII: S0022-5193(11)00066-X
DOI: doi:10.1016/j.jtbi.2011.01.040
Reference: YJTBI6352



www.elsevier.com/locate/jtbi

To appear in: *Journal of Theoretical Biology*

Received date: 10 September 2010
Revised date: 25 January 2011
Accepted date: 26 January 2011

Cite this article as: Giuseppina Albano, Virginia Giorno, Patricia Román-Román and Francisco Torres-Ruiz, Inferring the effect of therapy on tumors showing stochastic Gompertzian growth, *Journal of Theoretical Biology*, doi:[10.1016/j.jtbi.2011.01.040](https://doi.org/10.1016/j.jtbi.2011.01.040)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Inferring the effect of therapy on tumors showing stochastic Gompertzian growth

Giuseppina Albano ^a, Virginia Giorno ^b,
Patricia Román-Román ^{c*}, and Francisco Torres-Ruiz ^c

^a*Dipartimento di Scienze Economiche e Statistiche. Università di Salerno.
Fisciano(Sa)-Italy.*

^b*Dipartimento di Matematica e Informatica. Università di Salerno.
Fisciano(Sa)-Italy.*

^c*Departamento de Estadística e Investigación Operativa, Universidad de Granada.
Granada, Spain.*

Abstract

The present work deals with a Gompertz-type diffusion process, which includes in the drift term a time-dependent function $C(t)$ representing the effect of a therapy able to modify the dynamics of the underlying process. However, in experimental studies is not immediate to deduce the functional form of $C(t)$ from a treatment protocol. So a statistical approach is proposed in order to estimate this function, when a control group and one or more treated groups are observed. In order to validate the proposed strategy a simulation study for several interesting functional forms of $C(t)$ has been carried out. Finally, an application to infer the net effect of cisplatin and doxorubicin+ cyclophosphamide in actual murine models is presented.

Key words: Gompertz diffusion process; Therapy; Tumor growth.

1 Introduction

Models used to describe population growth currently play a relevant role in several fields such as economy, biology (evolution, demography, epidemiology,

* Corresponding author. Postal address: Departamento de Estadística e Investigación Operativa, Facultad de Ciencias, Universidad de Granada. Avda. Fuentenueva s/n 18071 Granada, Spain. Fax: +34 958243267. Tel.: +34 958240491. E-mail: proman@ugr.es

...), medicine, or ecology. In particular, the Gompertz curve seems particularly apt to capture the main features of several growth phenomena where resources are limited. This curve is associated with processes in which a slow initial growth is followed by faster growth and then by another period in which the limiting size, or limitation imposed by resources, is reached.

Traditionally, Gompertz-type models are solutions of certain ordinary differential equations. In this sense, they are deterministic, and do not take into account fluctuations or disturbances usually associated with the dynamics of real systems. Such fluctuations can be originated by a variety of factors, usually non-measurable or unknown. In order to include such environmental fluctuations in the model, the notion of “growth in random environment” was formulated (see, for instance, [12] and references therein) substituting the intrinsic fertility in the growth equation with a normal stochastic process whose mean is taken as representative of population fertility. Thus, growth is described by means of a stochastic differential equation (SDE) obtained from the deterministic model including a noise term in the associated ordinary differential equation. The resulting model is a process that can be treated with the stochastic methods of the diffusion equations (Lande et al. [8] popularized this methods for biologists).

In this context, Ricciardi et al. [13] used a homogeneous diffusion model associated to the Gompertz curve to model population growth and neuronal activity. Later, Gutiérrez et al. [6] introduced a non-homogeneous Gompertz-type diffusion process in order to include some exogenous factors. The main difference between the two models is the limit value of the process: in the former this value is independent on the initial distribution, whereas in the second one a dependence on this distribution exists. Such kind of dependence can be observed in some real contexts in which phenomena exhibit Gompertz-type growth and several sample-paths are available, each with a common growth pattern but with different initial values and a different limit value. This is the case, for example, of the dynamics of growth of individuals in the same species. It must be pointed out that the methodology used in Gutiérrez et al. [6] can be extended to other contexts in order to model several modes of population growth (see, for instance, Roman et al. [14]).

Regarding the use of this process in medicine, many of the models proposed in the literature to describe the dynamics of tumor growth are based on the Gompertz growth. In the 1960s, Laird [7] successfully used the Gompertz curve for the first time to fit data in tumor dynamics. Indeed, tumors are cellular populations growing in a confined space where the availability of nutrients is limited. In this context, Gompertzian-type models appear to be particularly consistent with the evidence of tumor growth (Sachs et al. [15], de Vladar and Gonzalez [4], Tabatai et al. [16], Castorina and Zappalà [3]). On the other side,

the abovementioned Gompertz diffusion process, due to be time homogeneous, can be unsatisfactory for the purpose of modeling the effects of an antitumoral treatment in such kind of dynamics. Such treatments, usually dependent on time, become suitable exogenous factors able to modify the growth of tumoral cells in the context of modeling.

This problem was studied by Albano and Giorno [1] by considering a homogeneous diffusion model described by the following SDE:

$$dX(t) = [\alpha X(t) - \beta X(t) \log X(t)] dt + \sigma X(t) dW(t)$$

in order to model the dynamics of an untreated tumor population $X(t)$. Here, α , β and σ are positive constants representing the growth, death rates and the width of random fluctuations, respectively, and $W(t)$ is a standard Brownian motion. Moreover, to include the effect of an antitumoral treatment, they introduced an exogenous term in the drift representing deterministic tumor regression rate due to the therapy. This term was modeled by means of a time-dependent function C , leading to the equation:

$$dX^C(t) = [(\alpha - C(t))X^C(t) - \beta X^C(t) \log X^C(t)] dt + \sigma X^C(t) dW(t). \quad (1)$$

This modified process was used in order to describe the evolution of tumoral cells in a parathyroid tumor by considering that the process is limited by two absorbing boundaries representing healing threshold and patient death (carrying capacity), respectively. The first exit time problem was analyzed for the process inside the boundary-restricted region by introducing a constant therapy, i.e. $C(t) = C$ as well as a logarithmic one, i.e. $C(t) = C_0 \log(e + \delta t)$.

Recently, Lo [9] generalized the model above in order to include both cell fission and mortality rates, and later, on the basis of this model, proposed another one in which the size of tumor cells is bounded (Lo [10]). In both cases the introduced rates are considered to be known.

It must be pointed out that in [1] the functional form of the therapy effect is assumed as known, as well as in the already mentioned research by Lo. Nevertheless, it seems reasonable to assume that in the first stages of a new medical treatment, the function C will be unknown; indeed, in experimental studies is not immediate to deduce the functional form of $C(t)$ from a treatment protocol. On the other hand, knowledge of such functional form is fundamental since it allows to introduce an external control to the system and to explain how the therapy acts. Moreover, the study of some problems related to the process (for example the first-passage-time) requires the functional form of $C(t)$. For these reasons, we provide a procedure for estimating the function C . The idea is to take the model $X(t)$ (homogeneous process) as a starting point and then to use the information provided by individuals in the treated group to

try approximate the function C . More precisely, the control group is assumed to be modeled by the homogeneous process $X(t)$, whereas the treated groups are modeled by means of the non-homogeneous one, i.e. $X^C(t)$. In this way, parameters of the homogeneous process, i.e. α , β and σ^2 , can be estimated by using trajectories of the control group and then, in a second step, $C(t)$ is estimated by the trajectories of the treated groups and suitable relations between the two models.

In section 2 we will deal with the non-homogeneous model, showing the distribution of the process as well as some interesting features that will be used later. Section 3 shows the strategy proposed in order to estimate function C , representing the therapy. To this end, the estimation of the parameters for the model associated to the control group is discussed, and some relations linking the two processes are shown. Section 4 discusses simulation results for different choices of the functional form of $C(t)$. Finally, in Section 5 an application to breast cancer in order to study the effect of two types of chemotherapy, concretely cisplatin and doxorubicin+cyclophosphamide, in actual murine models is presented.

2 The model

Let $\{X^C(t); t_0 \leq t \leq T\}$ be a diffusion process taking values on \mathbb{R}^+ and with infinitesimal mean and variance

$$\begin{aligned} A_1^C(x, t) &= [\alpha - C(t)]x - \beta x \log x \\ A_2^C(x, t) &= \sigma^2 x^2, \end{aligned} \tag{2}$$

where C is a continuous function in $[t_0, T]$ and α , β and σ are positive constants.

This process was introduced by Albano and Giorno [1] and studied from a transformation that leads to an Ornstein Uhlenbeck diffusion process. Now we present a general study of the process. Firstly, the distribution is obtained from the typical approaches: Kolmogorov's partial differential equations and SDE's. Secondly some of the main characteristics of the process, concretely the mean, mode and quantile functions, are obtained. Finally, some relations between the processes $X(t)$ and $X^C(t)$ are discussed in order to fit the function C from real data.

2.1 Distribution of the process

In this section we will present the distribution of the process $\{X^C(t); t_0 \leq t \leq T\}$ by means of the two approaches mentioned above. For both cases, finite dimensional distributions (for tumoral growth, the distribution of tumor size at all times) will be calculated.

- **Distribution of the process from stochastic differential equations** (see Appendix A for details)

The solution to Ito's SDE (1) with initial condition $X^C(t_0) = x_0$, being x_0 a positive random variable independent of $W(t)$, is

$$X^C(t) = \exp\left(\sigma e^{-\beta t} Z(t)\right) \quad (3)$$

where

$$Z(t) = z_0 + \left(\alpha - \frac{\sigma^2}{2}\right) \frac{e^{\beta t} - e^{\beta t_0}}{\sigma\beta} - \frac{1}{\sigma} \int_{t_0}^t C(s) e^{\beta s} ds + \int_{t_0}^t e^{\beta s} dW(s),$$

with $z_0 = (e^{\beta t_0} \log x_0) / \sigma$.

The finite dimensional distributions of the process $X^C(t)$ are lognormal $\Lambda_n(\eta; \Delta)$, where the components of the vector η and matrix Δ are

$$\eta_i = \sigma e^{-\beta t_i} m_Z(t_i) \quad \text{and} \quad \delta_{ij} = \sigma^2 e^{-\beta(t_i+t_j)} R_Z(t_i, t_j),$$

respectively, where

$$m_Z(t) = E[z_0] + \left(\alpha - \frac{\sigma^2}{2}\right) \frac{e^{\beta t} - e^{\beta t_0}}{\sigma\beta} - \frac{1}{\sigma} \int_{t_0}^t C(s) e^{\beta s} ds$$

and

$$R_Z(t, s) = \text{Var}[z_0] + \frac{e^{2\beta(t \wedge s)} - e^{2\beta t_0}}{2\beta},$$

with $t \wedge s = \min(t, s)$.

The transition probability density function (p.d.f.) can then be calculated from the two-dimensional distributions.

- **Distribution of the process from partial differential equations.**

The solution to the Fokker-Plank and Kolmogorov equations, with initial conditions $\lim_{t \downarrow \tau} f^C(x, t|y, \tau) = \delta(x - y)$ and $\lim_{\tau \uparrow t} f^C(x, t|y, \tau) = \delta(x - y)$ (see Appendix B for details) provides, for $t > \tau$,

$$f^C(x, t|y, \tau) = \frac{1}{x\sqrt{2\pi\sigma_{X^C}^2(t|\tau)}} \exp\left(-\frac{[\log x - M_{X^C}(t|\log y, \tau)]^2}{2\sigma_{X^C}^2(t|\tau)}\right), \quad (4)$$

where

$$M_{X^C}(t|y, \tau) = e^{-\beta(t-\tau)} y + \left(\alpha - \frac{\sigma^2}{2} \right) \frac{1 - e^{-\beta(t-\tau)}}{\beta} - e^{-\beta t} \int_{\tau}^t C(\theta) e^{\beta\theta} d\theta$$

and

$$\sigma_{X^C}^2(t|\tau) = \frac{\sigma^2}{2\beta} \left(1 - e^{-2\beta(t-\tau)} \right),$$

corresponding to a lognormal distribution.

Since $X^C(t)$ is Markovian, its finite-dimensional distributions depend on the initial distribution and on the transition p.d.f. In our case, the transition p.d.f. is lognormal, given in (4), so we only need to specify the initial distribution. Accordingly, we consider two initial distributions: a degenerate one in $x_0 > 0$, i.e. $P[X^C(t_0) = x_0] = 1$, and a lognormal one, i.e. $X^C(t_0) \sim \Lambda_1(\mu_0; \sigma_0^2)$. These choices ensure that finite dimensional distributions are lognormal (in accordance with what was established in the previous approach for the distribution of the process). It must be pointed out that the first case is a particular case of the second one with $\sigma_0 = 0$ and $\mu_0 = \log x_0$. Moreover, in real cases the degenerate initial distribution is able to describe situations in which only one sample path is available whereas the lognormal case requires several trajectories.

In particular, since $[X^C(t)|X^C(t_0) = x_0] \sim \Lambda_1(M_{X^C}(t|\log x_0, t_0); \sigma_{X^C}^2(t|t_0))$, taking $X^C(t_0) \sim \Lambda_1(\mu_0; \sigma_0^2)$, it is possible to calculate the joint distribution of $(X^C(t_0), X^C(t))'$, $t > t_0$, resulting in a two-dimensional lognormal distribution $\Lambda_2(\mu_{t_0,t}; \Sigma_{t_0,t})$ with

$$\mu_{t_0,t} = \begin{pmatrix} \mu_0 \\ M_{X^C}(t|\mu_0, t_0) \end{pmatrix} \quad \text{and} \quad \Sigma_{t_0,t} = \mathbf{A}_{t_0,t} \Gamma_{t_0,t} \mathbf{A}_{t_0,t}'$$

being

$$\mathbf{A}_{t_0,t} = \sigma \begin{pmatrix} e^{-\beta t_0} & 0 \\ 0 & e^{-\beta t} \end{pmatrix} \quad \text{and} \quad \Gamma_{t_0,t} = \frac{\sigma_0^2 e^{2\beta t_0}}{\sigma^2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{e^{2\beta t}}{\sigma^2} \begin{pmatrix} 0 & 0 \\ 0 & \sigma_{X^C}^2(t|t_0) \end{pmatrix}.$$

Thus, $X^C(t) \sim \Lambda_1(M_{X^C}(t|\mu_0, t_0), B_t + \sigma_{X^C}^2(t|t_0))$, where $B_t = \sigma_0^2 e^{-2\beta(t-t_0)}$, $t > t_0$.

In the same way, the distribution of $(X^C(t), X^C(s))$, $0 \leq t_0 < s < t$, and the finite dimensional distributions of $X^C(t)$ can be obtained using Markov's property.

2.2 Some characteristics

Some of the main characteristics of the growth can be obtained through the unidimensional distributions of the process. We will now proceed to describe some of them, focusing on three of the most commonly employed in practice, in particular for fitting and forecasting purposes: the n -th moment, the mode and the quantile functions (as well as their conditional versions), whose expressions can be formulated jointly for the two aforementioned initial distributions.

Assuming $X^C(t_0) \sim \Lambda_1(\mu_0; \sigma_0^2)$ and taking into account the distributions of $X^C(t)$ and $X^C(t)|X^C(\tau) = y$ ($t > \tau$), the mean, mode and quantile functions can be jointly expressed in the form

$$G_{X^C}^\lambda(t|y, \tau) = \exp\left(\lambda_1 M_{X^C}(t|y, \tau) + \lambda_2 \left(\lambda_3 B_t + \sigma_{X^C}^2(t|\tau)\right)^{\lambda_4}\right), \quad (5)$$

with $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)'$. Table 1 summarizes the different functions according to the values of λ , τ and y , where z_α is the α -quantile of a standard normal distribution.

Table 1

Values of λ , t , τ and y used to obtain the n -th moment, the mode and quantile functions from $G_{X^C}^\lambda(t|y, \tau)$.

Function	Expression	y	τ	λ
n -th moment	$E[(X^C(t))^n]$	μ_0	t_0	$(n, n^2/2, 1, 1)'$
n -th conditional moment	$E[(X^C(t))^n X^C(s) = x_s]$	$\log x_s$	s	$(n, n^2/2, 0, 1)'$
mode	$Mode[X^C(t)]$	μ_0	t_0	$(1, -1, 1, 1)'$
conditional mode	$Mode[X^C(t) X^C(s) = x_s]$	$\log x_s$	s	$(1, -1, 0, 1)'$
α -quantile	$C_\alpha[X^C(t)]$	μ_0	t_0	$(1, z_\alpha, 1, 1/2)'$
α - conditional quantile	$C_\alpha[X^C(t) X^C(s) = x_s]$	$\log x_s$	s	$(1, z_\alpha, 0, 1/2)'$

3 Fitting the effect of the therapy

The aim of this section is to approximate function C included in the infinitesimal mean of the Gompertz diffusion process (2).

Experimental designs to test the effectiveness of a therapy treatment include untreated (control) and treated groups, and the progress of the tumor size is observed in both.

Clearly, the growth of the control group is modeled by the process without including the function C . The idea is to take this model as a starting point and then use the information provided by individuals in the treated group to try and approximate the function C . More precisely, the proposed methodology for estimating C is based on a two-step procedure: first, to develop the inference on the original Gompertz diffusion process and second, to establish a relationship between the characteristics of the two processes (with and without the effect of the therapy).

3.1 Estimation of the parameters of the process without the effect of a therapy

We will consider the Gompertz diffusion process $\{X(t); t_0 \leq t \leq T\}$, taking values on \mathbb{R}^+ , with infinitesimal moments

$$\begin{aligned} A_1(x, t) &= \alpha x - \beta x \log x \\ A_2(x, t) &= \sigma^2 x^2, \end{aligned} \tag{6}$$

whose transition p.d.f. can be obtained from (4) by choosing $C(t) = 0$.

By denoting

$$a = \alpha - \frac{\sigma^2}{2}, \quad T_\beta = \frac{1 - e^{-\beta(t-\tau)}}{\sqrt{\beta(1 - e^{-2\beta(t-\tau)})}}$$

and considering the transformation in the variable $X^C(t)|X^C(\tau) = y$ ($t > \tau$),

$$x_\beta = \frac{\sqrt{\beta} [\log x - e^{-\beta(t-\tau)} \log y]}{\sqrt{1 - e^{-2\beta(t-\tau)}}},$$

the transition p.d.f. can be expressed as

$$f(x_\beta, t|y, \tau) = \frac{1}{\sqrt{\pi\sigma^2}} \exp\left(-\frac{[x_\beta - aT_\beta]^2}{\sigma^2}\right). \tag{7}$$

Let us consider a discrete sampling of the process, based on d sample paths, for times t_{ij} , ($i = 1, \dots, d$, $j = 1, \dots, n_i$) with $t_{i1} = t_1$, $i = 1, \dots, d$. That is, we observe the variables $X(t_{ij})$, the values of which, $\{x_{ij}\}_{i=1, \dots, d; j=1, \dots, n_i}$, constitute the sample for the inferential study.

The likelihood function depends on the choice of the initial distribution. When

$$P[X(t_1) = x_1] = 1, \text{ this function is } L_{x_{ij}}(a, \beta, \sigma^2) = \prod_{i=1}^d \prod_{j=2}^{n_i} f(x_{ij}, t_{ij}|x_{i,j-1}, t_{i,j-1}),$$

where a , β and σ^2 are the parameters to be estimated. If $X(t_1) \sim \Lambda_1(\mu_1, \sigma_1^2)$ the likelihood is $L_{x_{ij}}(\mu_1, \sigma_1^2, a, \beta, \sigma^2) = \prod_{i=1}^d f_{X(t_{i1})}(x_{i1}) \prod_{j=2}^{n_i} f(x_{ij}, t_{ij} | x_{i,j-1}, t_{i,j-1})$.

In the second case, there are two additional parameters that must be included in the estimation procedure. Nevertheless, it can be immediately confirmed that the estimations of μ_1 and σ_1^2 depend only on the initial values and do not influence the estimation of the other parameters. Hence, the maximum likelihood (ML) estimators of a , β and σ^2 are the same in both cases.

Henceforth, we will consider the general case in which the initial distribution is lognormal. In order to simplify calculations, we transform the x_{ij} values by means of $v_{1j} = x_{1j}$ and

$$v_{\beta}^{ij} = \frac{\sqrt{\beta} \left[\log x_{ij} - e^{-\beta(t_{ij}-t_{i,j-1})} \log x_{i,j-1} \right]}{\sqrt{1 - e^{-2\beta(t_{ij}-t_{i,j-1})}}}, \quad i = 1, \dots, d; j = 1, \dots, n_i.$$

From (7), and denoting $n = \sum_{i=1}^d n_i$ and \mathbf{v}_{β} , the vector containing the v_{β}^{ij} values, the log-likelihood function for the transformed sample is

$$\begin{aligned} \log L_{\mathbf{v}_{\beta}}(\mu_1, \sigma_1^2, a, \beta, \sigma^2) &= -\frac{n}{2} \log(2\pi) - \frac{d}{2} \log \sigma_1^2 - \frac{n-d}{2} \log \sigma^2 \\ &\quad - \frac{1}{\sigma_1^2} \sum_{i=1}^d [\log x_{i1} - \mu_1]^2 - \frac{1}{\sigma^2} \sum_{i=1}^d \sum_{j=2}^{n_i} [v_{\beta}^{ij} - T_{\beta}^{ij}]^2 \end{aligned}$$

from which the ML estimations of μ_1 and σ_1^2 are

$$\hat{\mu}_1 = \frac{1}{d} \sum_{i=1}^d \log x_{i1}, \quad \hat{\sigma}_1^2 = \frac{1}{d} \sum_{i=1}^d (\log x_{i1} - \hat{\mu}_1)^2.$$

Deriving the log-likelihood function with respect to a , σ^2 and β , and making the derivatives equal to zero, we obtain, respectively, the following system of equations:

$$\sum_{i=1}^d \sum_{j=2}^{n_i} [v_{\beta}^{ij} - aT_{\beta}^{ij}] v_{\beta}^{ij} = 0, \quad (8)$$

$$\sum_{i=1}^d \sum_{j=2}^{n_i} [v_{\beta}^{ij} - aT_{\beta}^{ij}]^2 = \frac{\sigma^2(n-d)}{2}, \quad (9)$$

$$\sum_{i=1}^d \sum_{j=2}^{n_i} \left[v_{\beta}^{ij} - a T_{\beta}^{ij} \right] \frac{d}{d\beta} \left\{ v_{\beta}^{ij} - a T_{\beta}^{ij} \right\} = 0. \quad (10)$$

Denoting

$$X_{1,\beta} = \sum_{i=1}^d \sum_{j=2}^{n_i} v_{\beta}^{ij} T_{\beta}^{ij}, \quad X_{2,\beta} = \sum_{i=1}^d \sum_{j=2}^{n_i} \left(T_{\beta}^{ij} \right)^2, \quad X_{3,\beta} = \sum_{i=1}^d \sum_{j=2}^{n_i} \left(v_{\beta}^{ij} \right)^2,$$

$$Y_{1,\beta} = \sum_{i=1}^d \sum_{j=2}^{n_i} v_{\beta}^{ij} \frac{d}{d\beta} v_{\beta}^{ij}, \quad Y_{2,\beta} = \sum_{i=1}^d \sum_{j=2}^{n_i} v_{\beta}^{ij} \frac{d}{d\beta} T_{\beta}^{ij},$$

$$Y_{3,\beta} = \sum_{i=1}^d \sum_{j=2}^{n_i} T_{\beta}^{ij} \frac{d}{d\beta} v_{\beta}^{ij}, \quad Y_{4,\beta} = \sum_{i=1}^d \sum_{j=2}^{n_i} T_{\beta}^{ij} \frac{d}{d\beta} T_{\beta}^{ij},$$

and, after some algebra, from equations (8) and (9), we obtain

$$a_{\beta} = \frac{X_{1,\beta}}{X_{2,\beta}} \quad \text{and} \quad \sigma_{\beta}^2 = \frac{2}{n-d} \left[X_{3,\beta} - \frac{X_{1,\beta}^2}{X_{2,\beta}} \right]$$

whereas (10) becomes

$$Y_{1,\beta} - a_{\beta} Y_{2,\beta} - a_{\beta} Y_{3,\beta} + a_{\beta}^2 Y_{4,\beta} = 0.$$

This last equation does not have an explicit solution, so the estimation of β can be found by numerical methods. Estimations of a and σ^2 are then found by $\hat{a} = a_{\hat{\beta}}$ and $\hat{\sigma}^2 = \sigma_{\hat{\beta}}^2$, respectively, whereas the ML estimation of α is $\hat{\alpha} = \hat{a} + \hat{\sigma}^2/2$. Finally, the ML estimation of any parametric function expressed in (5) can be obtained by applying Zehna's theorem.

3.2 Relationship between the main characteristics of the processes

Since the process in absence of therapy is a particular case of (2) for $C(t) = 0$, we can find the relations between the two processes. On one hand, and using (3), the trajectories of the processes are related by

$$X^C(t) = X(t) \exp \left(- \int_{t_0}^t C(s) e^{-\beta(t-s)} ds \right).$$

On the other hand, denoting by $G_X^\lambda(t|y, \tau)$ the function defined in (5) for process $X(t)$, and noting that

$$M_X(t|y, \tau) = M_{X^C}(t|y, \tau) - e^{-\beta t} \int_{\tau}^t C(\theta) e^{\beta\theta} d\theta \text{ and } \sigma_X^2(t|\tau) = \sigma_{X^C}^2(t|\tau),$$

we obtain the following relation

$$\begin{aligned} \frac{G_{X^C}^\lambda(t|y, \tau)}{G_X^\lambda(t|y, \tau)} &= \frac{\exp\left(\lambda_1 M_{X^C}(t|y, \tau) + \lambda_2 (\lambda_3 B_t + \sigma_X^2(t|\tau))^{\lambda_4}\right)}{\exp\left(\lambda_1 M_X(t|y, \tau) + \lambda_2 (\lambda_3 B_t + \sigma_X^2(t|\tau))^{\lambda_4}\right)} \\ &= \exp\left(-\lambda_1 e^{-\beta t} \int_{\tau}^t C(\theta) e^{\beta\theta} d\theta\right). \end{aligned}$$

In particular, taking $\lambda = (1, 1/2, 1, 1)'$, $t = t_0$ and $y = \mu_0$, the mean functions of both processes verify

$$\frac{E[X^C(t)]}{E[X(t)]} = \exp\left(-e^{-\beta t} \int_{t_0}^t C(\theta) e^{\beta\theta} d\theta\right)$$

from which

$$C(t) = -e^{-\beta t} \frac{d}{dt} \left\{ e^{\beta t} \log \left(\frac{E[X^C(t)]}{E[X(t)]} \right) \right\}. \quad (11)$$

3.3 Proposed methodology

The data required for the proposed strategy are the values of tumor size on d_1 individuals of a non treated (control) group (x_{ij} , $i = 1, \dots, d_1$, $j = 1, \dots, n$) and d_2 individuals of a treated group (x_{ij}^C , $i = 1, \dots, d_2$, $j = 1, \dots, n$), observed in the same time instants t_1, \dots, t_n .

By using equation (11), we are able to find an approximation of the therapy function C , and thus of the modified growth rate of the process. We are then able to implement the following procedure:

- From the data of the control group, estimate the parameters of process $X(t)$, i.e. without the effect of a therapy. From this first step, we obtain ML estimations $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\sigma}^2$.

- Calling x_j and x_j^C the mean tumor size at time t_j in the control group and in the treated group respectively (that is, $x_j = \sum_{i=1}^{d_1} x_{ij}/d_1$ and $x_j^C = \sum_{i=1}^{d_2} x_{ij}^C/d_2$), obtain function $m(t)$ by the interpolation of values

$$m_j = e^{\hat{\beta}t_j} \log\left(\frac{x_j^C}{x_j}\right). \quad (12)$$

Note that if $t_1 = t_0$, the starting time of therapy, the distribution of $X(t_1)$ and $X^C(t_1)$ must match, so x_1 and x_1^C should be approximately equal and thus should be considered for the procedure leading to $C(t_1) = 0$.

- Finally, consider the following function as an approximation of $C(t)$:

$$\hat{C}(t) = -m'(t) e^{-\hat{\beta}t}. \quad (13)$$

We point out that the proposed methodology consists of a two-step procedure: (i) estimation of parameters α , β and σ for the control process; (ii) interpolation of the points m_i in (12). Hence, the consistence of the proposed estimator (13) derives from the consistence of the maximum likelihood estimators in (i) and from the uniform convergence of the interpolation method in (ii) (for example cubic spline interpolation). So, when the number of data points increases the proposed estimate becomes better and better.

4 Simulation study

In order to validate the strategy proposed above, a simulation study has been carried out for several interesting functional forms of function C . Specifically, we have considered the case of constant, linear and logarithmic functions.

In all cases, for the process modeling the control group we have chosen the following values for the parameters: $\alpha = 0.3$, $\beta = 0.1$ and $\sigma = 0.01$, so the original homogeneous diffusion model has infinitesimal moments

$$\begin{aligned} A_1(x) &= 0.3x - 0.1x \log x \\ A_2(x) &= 0.01^2 x^2. \end{aligned} \quad (14)$$

For this model, 100 sample paths (see Figure 1) have been simulated in the $[0, 50]$ time interval, taking $t_i = i - 1$, $i = 1, \dots, 51$, and considering a log-normal initial distribution $\Lambda_1[1; 0.2]$. The estimation procedure developed in Section 3.1 provides

$$\hat{\alpha} = 0.299239, \quad \hat{\beta} = 0.0997214, \quad \hat{\sigma} = .00991149.$$

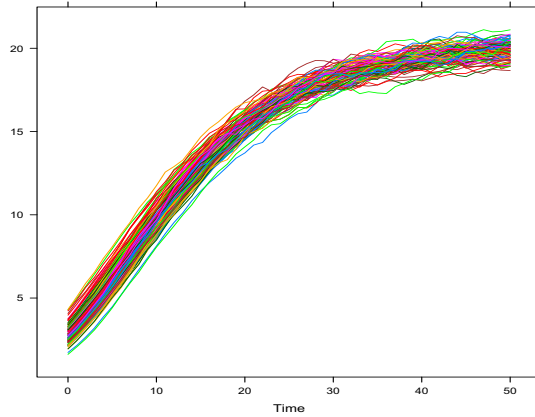


Figure 1. Simulated sample paths of the process $X(t)$ with $\alpha = 0.3$, $\beta = 0.1$, $\sigma = 0.01$ and initial distribution $\Lambda_1[1;0.2]$

In the following subsections we will consider the three types of functions mentioned before. In each case, the mean function of the involved processes is plotted for different $C(t)$ functions in order to better understand the effect of their introduction in the original model. Then, and for a particular $C(t)$ function, 100 sample paths have been simulated for the modified process, taking the same initial distribution and time instants as those chosen for the control group. Finally, the estimation procedure for the therapy introduced in section 3.3. is applied, showing its validity in this context. Moreover, a study of how the number of data points (the number of sample paths and the number of data in each of them) affects to the mean relative absolute error of the estimated therapies has been performed.

4.1 Constant function $C(t) = C$

Choosing $C(t) = C$, process $X^C(t)$ in (2) is of the same type of the original one, with a different value of the growth rate α , i.e. $\alpha \rightarrow \alpha - C$. Thus, studying this model is equivalent to evaluate the influence of a change in the parameter α of the original model (6). Figure 2 shows the influence of a constant function in the mean function of $X^C(t)$.

In Figure 3, 100 simulated sample paths of the modified version of process (14) are plotted, choosing $C(t) = 0.03$, $\forall t \geq t_1 = 0$.

Using the proposed methodology, the functional form of function $C(t)$ has been fitted. In Figures 4a and 4b the approximation of $C(t)$, i.e. $\hat{C}(t)$, and the difference $C(t) - \hat{C}(t)$ are plotted, showing the proposed procedure is appropriate.

Table 2 shows the values (expressed as percentages) of the mean relative ab-

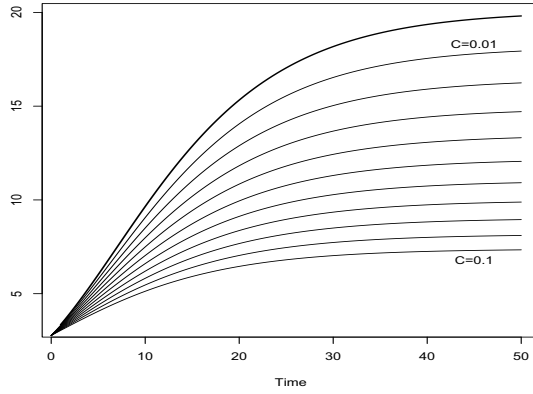


Figure 2. Mean function for processes $X(t)$ (upper line) and $X^C(t)$ with $\alpha = 0.3$, $\beta = 0.1$, $\sigma = 0.01$ and $C(t) = C$ and C from 0.01 to 0.1 by 0.01.

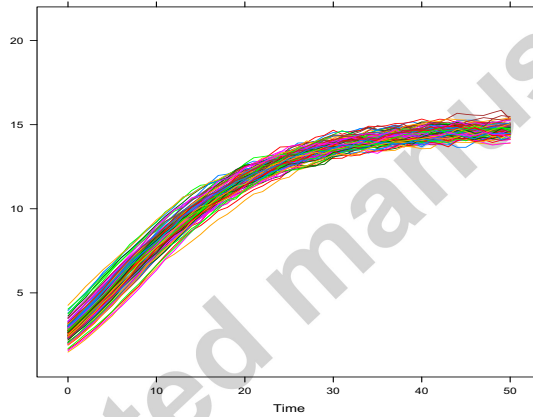


Figure 3. 100 simulated sample paths of $X^C(t)$ with $\alpha = 0.3$, $\beta = 0.1$, $\sigma = 0.01$, initial distribution $\Lambda_1[1; 0.2]$ and $C(t) = 0.03$.

solute error of the estimated therapy, i.e. the mean of the quotients of the absolute differences between estimated and actual values of the therapy and the actual ones, for several values of number of data points (number of sample paths and number of data in each of them).

Table 2

Mean relative absolute errors (expressed as percentages) for the estimation of the constant therapy $C = 0.03$.

Number of data in each path	Number of sample paths					
	1	5	10	25	50	100
5	22.51	22.14	22.08	22.03	21.96	21.90
7	5.74	5.13	5.06	5.06	5.05	5.05
10	2.67	1.65	1.49	1.37	1.36	1.35
25	1.83	0.85	0.59	0.39	0.28	0.21
50	1.83	0.85	0.59	0.37	0.28	0.21

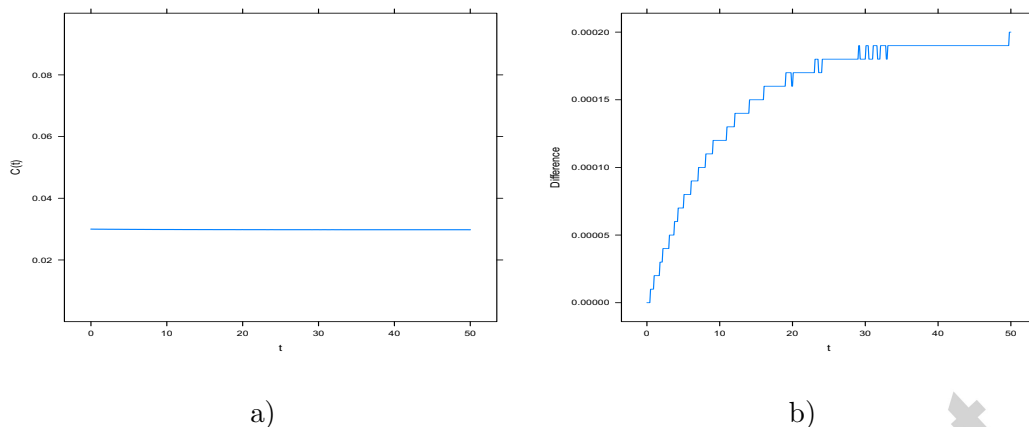


Figure 4. a) Fitted $C(t)$ function. b) Difference between approximate and real $C(t)$ functions (in the case of a constant therapy).

Note that the errors decrease with increasing the number of sample paths and the number of data in each one of them, being more marked the improvement in the latter case for fixed values of the number of sample paths. In addition, the values of errors are acceptable for a number of data not too high even in the case of a single sample path.

4.2 Linear function $C(t) = Ct$

We will focus on the case $C(t) = Ct$ since the effect of the independent term has been studied in the previous subsection. Figure 5 shows the influence of this function in the mean function of $X^C(t)$, whereas Figure 6 exhibits 100 sample paths of the modified version of process (14) choosing $C(t) = 0.002t$, $\forall t \geq t_1 = 0$.

Using the proposed procedure, $C(t)$ has been fitted. Figures 7a and 7b show $\hat{C}(t)$ and the difference $C(t) - \hat{C}(t)$.

Table 3 shows the values (expressed as percentages) of the mean relative absolute error of the estimated therapy for several values of number of data points (number of sample paths and number of data in each of them). The interpretation is analogous to that of Table 2.

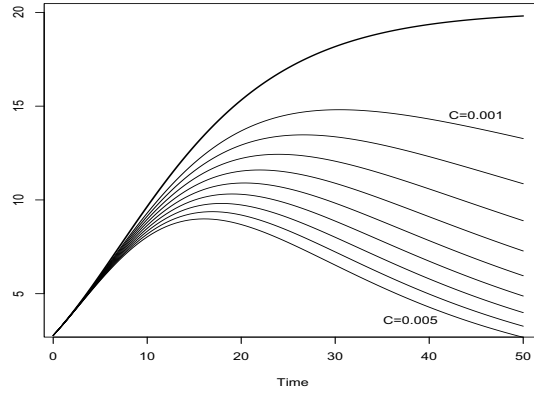


Figure 5. Mean function of $X(t)$ (upper line) and $X^C(t)$ with $\alpha = 0.3$, $\beta = 0.1$, $\sigma = 0.01$ and $C(t) = Ct$, taking C from 0.001 to 0.005 by 0.0005.

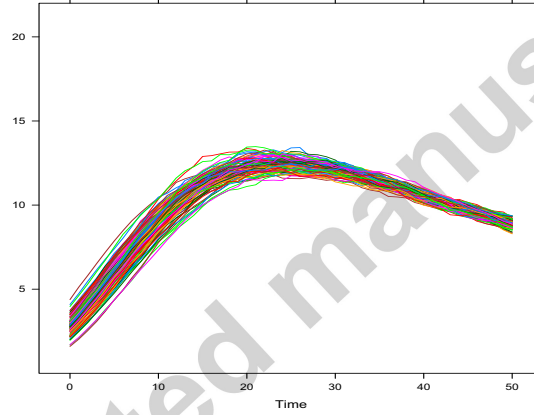


Figure 6. 100 simulated sample paths of $X^C(t)$ with $\alpha = 0.3$, $\beta = 0.1$, $\sigma = 0.01$, initial distribution $\Lambda_1[1; 0.2]$ and $C(t) = 0.002t$.

Table 3

Mean relative absolute errors (expressed as percentages) for the estimation of the linear therapy $C(t) = 0.002t$.

Number of data in each path	Number of sample paths					
	1	5	10	25	50	100
5	18.61	18.40	18.27	18.27	18.26	18.26
7	7.03	6.87	6.84	6.83	6.83	6.83
10	2.98	2.69	2.64	2.64	2.64	2.64
25	1.70	1.10	0.93	0.82	0.78	0.75
50	1.70	1.10	0.93	0.82	0.78	0.75

4.3 Logarithmic function $C(t) = C_0 \log(e + Ct)$

A logarithmic $C(t)$ function was considered by Albano et al. [1] and de Vladar et al. [5] in the study of antitumoral therapies since it seems that a logarithmic

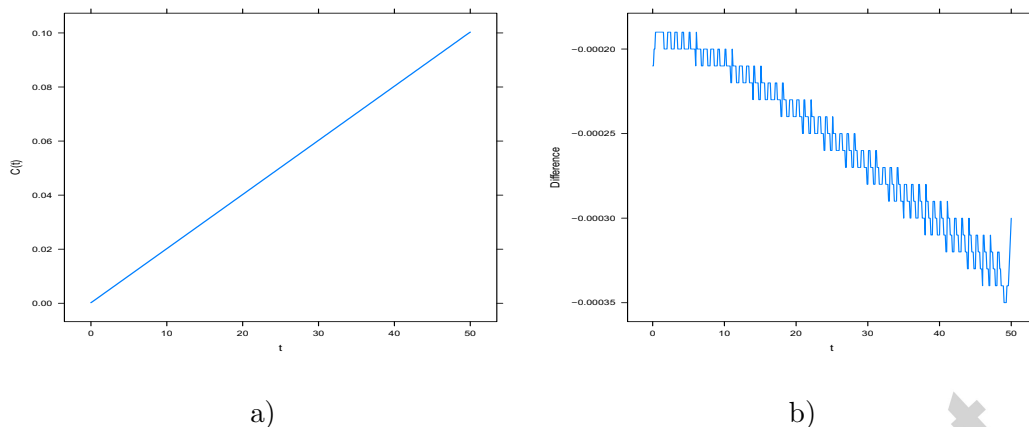


Figure 7. a) Fitted $C(t)$ function. b) Difference between approximate and real $C(t)$ functions (in the case of a linear therapy).

therapy induces a reduction of tumor population larger than the one experienced in a constant therapy and it is more tolerable than one that linearly increases in time.

Figure 8 shows the influence of a logarithmic function in the mean function of process $X^C(t)$.

With the aim of validating the proposed strategy, we have simulated 100 sample paths of the modified version of process (14), choosing $C(t) = 0.03 \log(e + 0.2t)$, $\forall t \geq t_1 = 0$. The values of the parameters have been chosen arbitrarily but so that the therapy is effective, i.e. the condition $C_0 > -\beta \log(x_0)$ is verified. Figure 9 shows the trajectories whereas Figures 10a and 10b show $\hat{C}(t)$ and the difference $C(t) - \hat{C}(t)$.

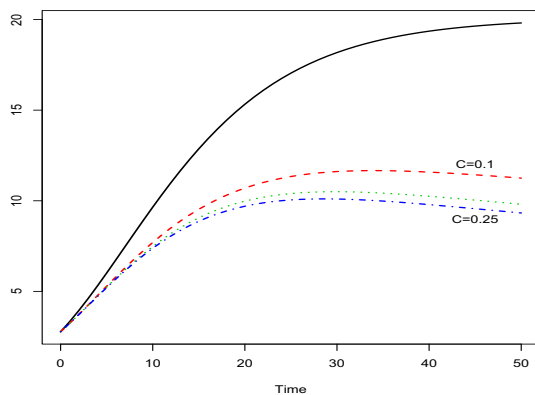


Figure 8. Mean function of $X(t)$ (upper line) and $X^C(t)$ with $\alpha = 0.3$, $\beta = 0.1$, $\sigma = 0.01$ and $C(t) = C_0 \log(e + Ct)$, taking $C_0 = 0.03$ and $C = 0.1, 0.2$ and 0.25 .

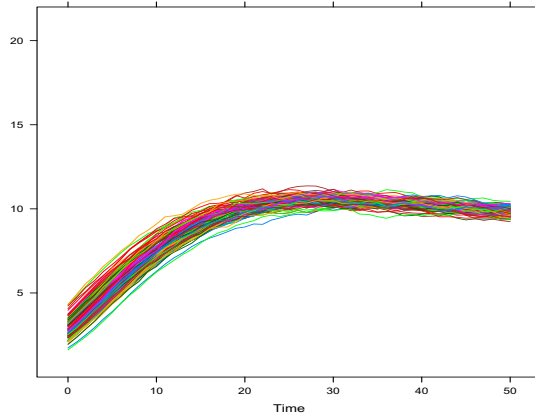


Figure 9. 100 simulated sample paths of $X^C(t)$ with $\alpha = 0.3$, $\beta = 0.1$, $\sigma = 0.01$, initial distribution $\Lambda_1[1; 0.2]$ and $C(t) = 0.03 \log(e + 0.2t)$.

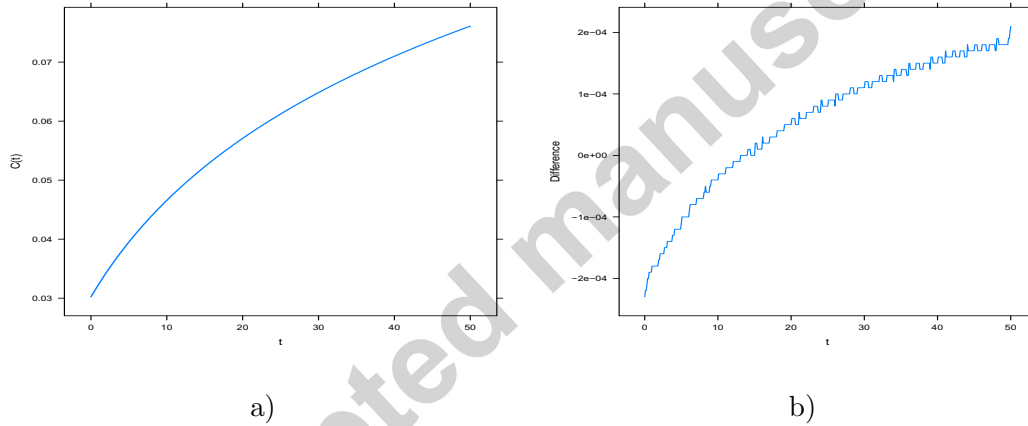


Figure 10. a) Fitted $C(t)$ function. b) Difference between approximate and real $C(t)$ functions (in the case of a logarithmic therapy).

Table 4 shows the values (expressed as percentages) of the mean relative absolute error of the estimated therapy for several values of number of data points (number of sample paths and number of data in each of them). The interpretation is analogous to Tables 2 and 3.

5 Application to real data of tumor growth

In order to study the effect of two types of chemotherapy in breast cancer xenografts, we consider data about the growth of *BC297MONp5* from three experimental groups of 7, 8 and 8 mice, respectively. The first one was a control group (untreated), whereas two treatments, Doxorubicin (a dose of

Table 4

Mean relative absolute errors (expressed as percentages) for the estimation of the logarithmic therapy $C(t) = 0.003\log(e + 0.2t)$.

Number of data in each path	Number of sample paths					
	1	5	10	25	50	100
5	49.01	47.87	47.87	47.62	47.59	47.57
7	3.70	3.39	3.33	3.33	3.33	3.33
10	3.44	2.74	2.57	2.55	2.53	2.51
25	1.70	0.85	0.61	0.42	0.34	0.28
50	1.66	0.76	0.57	0.40	0.29	0.24

2mg/Kg at 3-week intervals) + Cyclophosphamide (a dose of 100mg/Kg at 3-week intervals) and Cisplatin (a dose of 6mg/Kg at 3-week intervals) was applied for the second and the third group, respectively.

Mice of the three groups were treated at days 1, 21 and 42 and the tumor volume was measured at days 1, 7, 14, 21, 28, 35, 42 and 49. Subsequently, the relative volume of tumor with respect to the initial volume was calculated.

Figure 11 shows the mean of the relative tumor volume for the three experimental groups as a function of the days after starting the treatment.

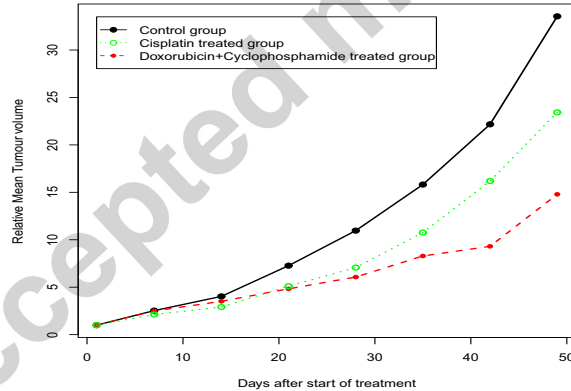


Figure 11. Tumor growth for the three experimental groups.

From the data of the control group we have estimated the parameters of the homogeneous process, and obtained the following values

$$\hat{\alpha} = 0.112784, \hat{\beta} = 0.0184158, \hat{\sigma} = 0.104842.$$

With these estimations, we have considered a process $X^C(t)$ for the two treatments, and the functional forms of functions $C(t)$ have been fitted by using the proposed methodology. The approximate functions are shown in Figure 12.

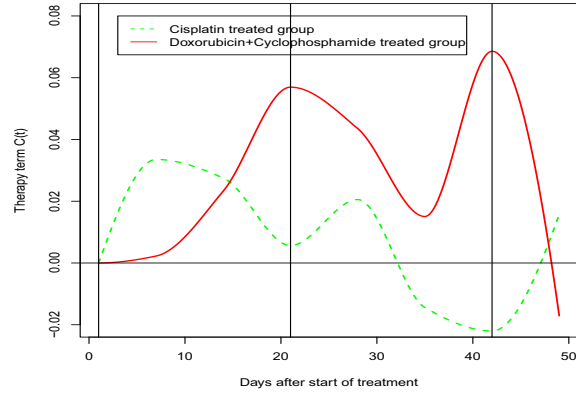


Figure 12. $C(t)$ functions fitted for the two treated groups

Once the models $X^C(t)$ have been estimated, we have fitted the observed data by using the mean function of the process (see figures 13a and 13b), the conditional mean function (figures 14a and 14b) and the percentile functions (figures 15a and 15b).

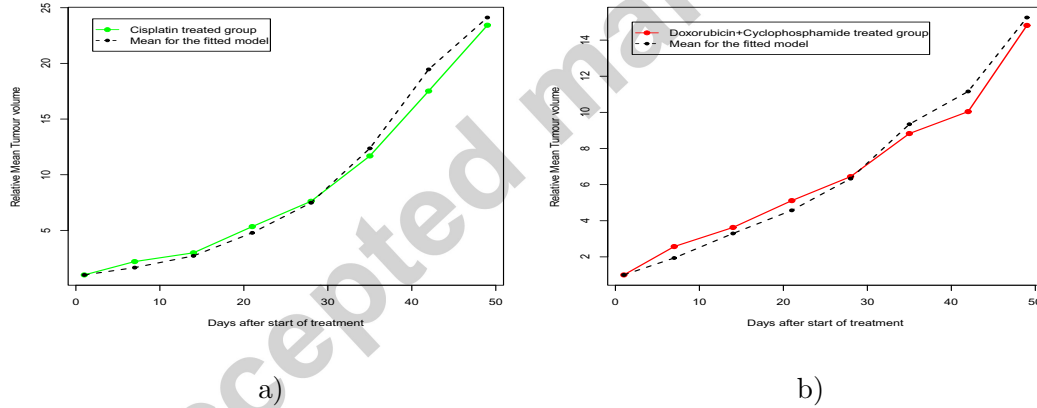


Figure 13. Fit (by the mean function) for the group treated with (a) Cisplatin and (b) Doxorubicin+Cyclophosphamide.

Once the therapies have been adjusted we can interpret their influence on tumor growth and draw conclusions about the effectiveness of treatment. Recall that the effect of therapy on the rate of tumor growth in an time instant t is expressed in terms of

$$\int_{t_0}^t C(\theta)e^{\beta\theta}d\theta. \quad (15)$$

As depicted in Figure 12, changes in $C(t)$ determines the trend in tumor

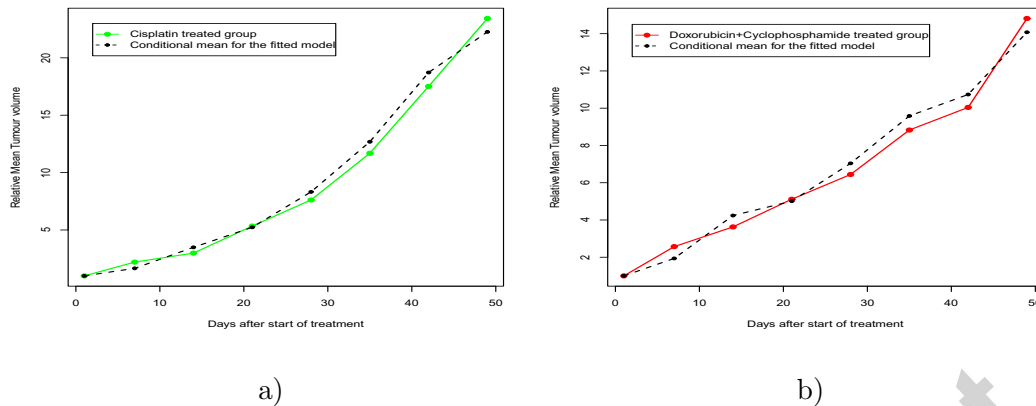


Figure 14. Fit (by the conditional mean function) for the group treated with (a) Cisplatin and (b) Doxorubicin+Cyclophosphamide.

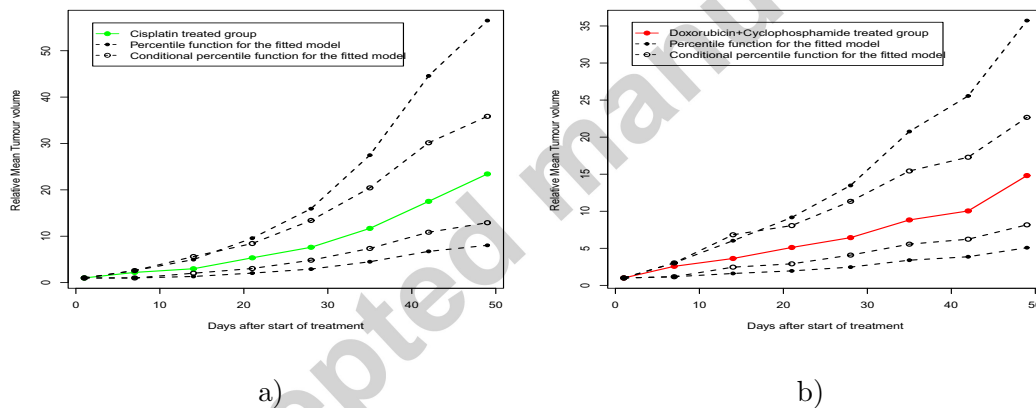


Figure 15. Fit (by the percentile functions) for the group treated with (a) Cisplatin and (b) Doxorubicin+Cyclophosphamide.

growth. When (15) is positive the rate of tumor growth is negatively affected and viceversa.

Note that the $C(t)$ patterns displayed by the two therapeutic schedules examined in this study are different. Cisplatin therapy seems to affect tumor growth very rapidly with a peak of activity at about 7 days. After day 7 the efficacy of cisplatin on tumor growth declines and the situation is reversed after 21 days, at the time of a second application of the drug. A second peak of activity, lower than the previous one, is found a week after treatment. Importantly after 30 days $C(t)$ becomes negative, indicating the absence of therapeutic effect of cisplatin at this period. In addition, between the days 32 and 46, where the growth curves of the control group and that of the group treated with Cisplatin

are practically parallel (similar growth rate), the $C(\cdot)$ function is negative, a reason for which its global effect, in terms of its cumulative integral, vanishes. The third application of cisplatin results at day 42 in a significant reduction of tumor growth. The association of cyclophosphamide/doxorubicin reduces tumor growth slower than cisplatin but it is more effective reaching a peak at 21 days. Importantly, the second administration of this therapeutic association does not result in a $C(t)$ increase and the inversion in the pattern of the curve is only observed at about 36 days, two weeks after the second administration. The same phenomenon seems to take place after the third administration of drugs with a dramatic and rapid decrease in $C(t)$. These results can explain why cyclophosphamide/doxorubicin affect tumor growth rate only until day 42. After the third administration treated tumors display growth rate kinetics similar to control untreated tumors.

The experimental data consists of tumor volume throughout the time. Changes in tumor volume during treatments can be largely affected by the pharmacodynamics and pharmacokinetics parameters of the drugs. It is therefore difficult to attribute observed $C(t)$ patterns to the availability of the drugs at the tumor site, to tumor response itself or both. Moreover, tumors can become resistant to the drugs after an initial response. The observed patterns of $C(t)$ resulting from the analyses suggest that the drug schedules applied to the mice, are not optimal. This can be explained by potential differences in pharmacokinetic parameters between humans and mice. It is worthy to point out that the schedules used in the mouse experiments analyzed in this work were chosen by mimicking the schedules applied to humans. Inferring $C(t)$ for new drugs in preclinical setting can be a valuable method to find the optimal administration schedule without conducting complex pharmacokinetic studies.

6 Conclusions

In this paper, a modification of a Gompertz diffusion process is considered. This modified process includes, in its infinitesimal mean a function of time C , in order to model the effect of a therapy.

We propose a strategy that allows to approximate this function from data of the evolution over time of the variable under study for a control group and for one or more treated groups. For example, for the study of tumor growth, from data on the evolution of tumor growth in individuals of a control group and the individuals subjected to a chemotherapy treatment.

Our simulation studies show how the proposed strategy detects the functional form of such function. In addition, an application has been made in order to explain the behavior of two types of chemotherapy in breast cancer. This

study has allowed us to interpret the influence of the function modeling the therapy in the tumor growth rate. In future applications, the knowledge of the functional form of this function should allow to know and to explain the effectiveness of a treatment throughout time, and also to establish protocols for the application of treatments if they do not exist, or to modify those previously established.

7 Acknowledgements

The authors are very grateful to Dr. Didier Decaudin (Institut Curie, Paris) for providing the data used in this research, to Dr. Sergio Román-Román (Institut Curie, Paris) for the interpretation of the results and to the anonymous referees whose comments and suggestions have contributed to improve this paper. This work was supported in part by the Ministerio de Educación y Ciencia, Spain, under Grants MTM2008-05785 and HI2007-0034 and by MIUR (PRIN 2008 and Azione Integrata Italia Spagna).

A Appendix A: Distribution of the process from stochastic differential equations

We consider Ito's SDE (1) with initial condition $X^C(t_0) = x_0$, being x_0 a positive random variable independent on $W(t)$. Considering the analytical properties of $C(t)$, it follows (see, for example, Arnold [2]) that this equation has an unique solution that will be the \mathbb{R}^+ -valued diffusion process with initial distribution x_0 and infinitesimal moments given by (2).

By using the transformation $Z(t) = e^{\beta t} \log X^C(t)/\sigma$, and by Ito's lemma, equation (1) becomes

$$\begin{aligned} dZ(t) &= \frac{e^{\beta t}}{\sigma} \left(\alpha - C(t) - \frac{\sigma^2}{2} \right) dt + e^{\beta t} dW(t) \\ Z(t_0) &= z_0 \end{aligned}$$

whose solution is

$$Z(t) = z_0 + \left(\alpha - \frac{\sigma^2}{2} \right) \frac{e^{\beta t} - e^{\beta t_0}}{\sigma \beta} - \frac{1}{\sigma} \int_{t_0}^t C(s) e^{\beta s} ds + \int_{t_0}^t e^{\beta s} dW(s),$$

with $z_0 = (e^{\beta t_0} \log x_0) / \sigma$.

Furthermore, and according to Arnold [2], $Z(t)$ is a Gaussian process if and only if z_0 is constant or normally distributed. In such case, the mean and covariance function of $Z(t)$ is given by

$$m_Z(t) = E[z_0] + \left(\alpha - \frac{\sigma^2}{2} \right) \frac{e^{\beta t} - e^{\beta t_0}}{\sigma \beta} - \frac{1}{\sigma} \int_{t_0}^t C(s) e^{\beta s} ds$$

and

$$R_Z(t, s) = Var[z_0] + \frac{e^{2\beta(t \wedge s)} - e^{2\beta t_0}}{2\beta},$$

respectively, where $t \wedge s = \min(t, s)$. Hence, the finite dimensional distributions of process $Z(t)$ are normal, that is, $\forall n \in \mathbb{N}$

$$(Z(t_1), Z(t_2), \dots, Z(t_n))' \sim \mathcal{N}_n(\mu; \Sigma)$$

where the i -th component of the vector μ is $m_Z(t_i)$, $i = 1, \dots, n$, whereas Σ is a definite positive matrix with components $R_Z(t_i, t_j)$, $i, j = 1, \dots, n$.

Finally, the solution of (1) is

$$X^C(t) = \exp\left(\sigma e^{-\beta t} Z(t)\right)$$

and, therefore, the finite dimensional distributions of the process $X^C(t)$ are lognormal $\Lambda_n(\eta; \Delta)$, where the components of the vector η and matrix Δ are

$$\eta_i = \sigma e^{-\beta t_i} m_Z(t_i) \quad \text{and} \quad \delta_{ij} = \sigma^2 e^{-\beta(t_i+t_j)} R_Z(t_i, t_j),$$

respectively. The transition probability density function (p.d.f.) can then be calculated from the two-dimensional distributions.

B Appendix B: Transition p.d.f. through diffusions equations

The transition p.d.f. of $X^C(t)$, $f^C(x, t|y, \tau)$, verifies the forward (or Fokker-Planck) equation:

$$\begin{aligned} \frac{\partial f^C(x, t|y, \tau)}{\partial t} = & - \frac{\partial [([\alpha - C(t)]x - \beta x \log x) f^C(x, t|y, \tau)]}{\partial x} \\ & + \frac{\sigma^2}{2} \frac{\partial^2 [x^2 f^C(x, t|y, \tau)]}{\partial x^2} \end{aligned} \quad (\text{B.1})$$

and the backward (or Kolmogorov) equation:

$$0 = \frac{\partial f^C(x, t|y, \tau)}{\partial \tau} + ([\alpha - C(s)]y - \beta y \log y) \frac{\partial f^C(x, t|y, \tau)}{\partial y} + \frac{\sigma^2}{2} y^2 \frac{\partial^2 f^C(x, t|y, \tau)}{\partial y^2}. \quad (\text{B.2})$$

Equations (B.1) and (B.2) verify the conditions for the existence and uniqueness of the solution with respect to the initial conditions $\lim_{t \downarrow \tau} f^C(x, t|y, \tau) = \delta(x - y)$ and $\lim_{\tau \uparrow t} f^C(x, t|y, \tau) = \delta(x - y)$, where $\delta(\cdot)$ is the Dirac's delta function. Observe that for this process the end points of the diffusion interval are natural non-attracting boundaries.

In this case, the transition p.d.f. of $X^C(t)$ can be found by looking for a transformation

$$\begin{aligned} t' &= \phi(t) \\ x' &= \psi(x, t) \end{aligned}$$

which changes Kolmogorov equation (B.2) into that of a standard Wiener process. Indeed, infinitesimal moments (2) verify the conditions of Theorem 1 in Ricciardi [11], so such transformation exists. Specifically,

$$\begin{aligned} \psi(x, t) &= \frac{k_1^{1/2}}{\sigma e^{\beta t_0}} \left(e^{\beta t} \log x - e^{\beta t_2} \log z - \left(\alpha - \frac{\sigma^2}{2} \right) \frac{e^{\beta t} - e^{\beta t_2}}{\beta} \right. \\ &\quad \left. + \int_{t_2}^t C(\theta) e^{\beta \theta} d\theta \right) + k_2 \\ \phi(t) &= \frac{k_1}{2\beta} \left(e^{2\beta(t-t_0)} - e^{2\beta(t_1-t_0)} \right) + k_3 \end{aligned}$$

where $z \in \mathbb{R}^+$, $t_i > 0$ and the k_i 's are arbitrary constants with $k_1 > 0$. In this way, we obtain, for $t > \tau$,

$$f^C(x, t|y, \tau) = \frac{1}{x \sqrt{2\pi\sigma_{X^C}^2(t|\tau)}} \exp \left(-\frac{[\log x - M_{X^C}(t|\log y, \tau)]^2}{2\sigma_{X^C}^2(t|\tau)} \right),$$

where

$$M_{X^C}(t|y, \tau) = e^{-\beta(t-\tau)} y + \left(\alpha - \frac{\sigma^2}{2} \right) \frac{1 - e^{-\beta(t-\tau)}}{\beta} - e^{-\beta t} \int_{\tau}^t C(\theta) e^{\beta \theta} d\theta$$

and

$$\sigma_{X^c}^2(t|\tau) = \frac{\sigma^2}{2\beta} \left(1 - e^{-2\beta(t-\tau)}\right),$$

corresponding to a lognormal distribution.

References

- [1] Albano, G., Giorno, V., 2006. A stochastic model in tumor growth. *Journal of Theoretical Biology* 242, 329–336.
- [2] Arnold, L., 1974. *Stochastic differential equations*. Wiley, New York.
- [3] Castorina, P., Zappalà, D., 2006. Tumor Gompertzian growth by cellular energetic balance. *Physica A: Statistical Mechanics and its Applications*, 365(2), 473–480
- [4] de Vladar, H.P., Gonzalez, J.A., 2004. Dynamics response of cancer under the influence of immunological activity and therapy. *Journal of Theoretical Biology* 227, 335–348.
- [5] de Vladar, H.P., Gonzalez, J.A., Rebolledo, M., 2003. New-late intensification schedules for cancer treatments. *Acta Cientifica Venezolana* 54, 263–276.
- [6] Gutiérrez, R., Román, P., Romero, D., Serrano, J.J., Torres, F., 2007. A new gompertz-type diffusion process with application to random growth. *Mathematical Biosciences* 208, 147–165.
- [7] Laird, A. K., 1964. Dynamics of tumor growth. *British Journal of Cancer* 18, 490-502.
- [8] Lande, R., Engen, S., Saether, B.E., 2003. *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press, Oxford.
- [9] Lo, C.F., 2007. Stochastic Gompertz model of tumour cell growth. *Journal of Theoretical Biology* 248, 317–321.
- [10] Lo, C.F., 2010. A modified stochastic Gompertz model for tumour cell growth. *Computational and Mathematical Methods in Medicine*, 11(1), 3–11.
- [11] Ricciardi, L.M., 1976. On the transformation of diffusion processes into the Wiener processes. *Journal of Mathematical Analysis and Applications* 54, 185–199.
- [12] Ricciardi, L.M. 1979. On the conjecture concerning population growth in random environment. *Biological Cybernetics* 32, 95–99.

- [13] Ricciardi, L.M., Sacerdote, L., Sato, S. (1983). Diffusion approximation and first-passage-time problem for a model neuron II. Outline of a computation method. *Mathematical Biosciences*, 64, 29–44.
- [14] Román-Román, P., Romero, D. Torres-Ruiz, F., 2010. A diffusion process to model generalized von Bertalanffy growth patterns: Fitting to real data. *Journal of Theoretical Biology* 263, 59–69.
- [15] Sachs, R.K., Hlatky, L.R. Hahnfeldt, P. 2001. Simple ODE models of tumor growth and anti-angiogenic or radiation treatments. *Mathematical and Computer Modelling*, 33, 1297–1305.
- [16] Tabatai, M., Williams, D.K., Bursac, Z., 2005. Hyperbolastic growth models: theory and applications. *Theoretical Biology and Medical Modelling*, 2(1), 1–14.

Accepted manuscript