Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER Applied Mathematics and Computation 175 (2006) 628–644

APPLIED MATHEMATICS AND COMPUTATION

www.elsevier.com/locate/amc

The Stochastic Rayleigh diffusion model: Statistical inference and computational aspects. Applications to modelling of real cases

R. Gutiérrez *, R. Gutiérrez-Sánchez, A. Nafidi

Department of Statistics and Operational Research, University of Granada, Facultad de Ciencias, Campus de Fuentenueva, 18071 Granada, Spain

Abstract

In this paper, we consider a Stochastic System modelling by the Stochastic Rayleigh Diffusion Process and we discuss theoretical aspects of the latter and establish a statistical methodology to adjust it to real cases, particulary, in the field of biometry and related areas. The Rayleigh process, according to the definition of [C. Giorno, A. Nobile, L. Ricciardi, L. Sacerdote, Some remarks on the Rayleigh process, Journal of Applied Probability 23 (1986) 398–408], is examined from the perspective of the corresponding nonlinear stochastic differential equation, and from its associated probability density function we obtain the corresponding mean functions (trend function and conditional trend function), which depend of Kummer functions. The drift parameters are estimated by maximum likelihood on the basis of continuous sampling of the process and they are calculated by computational methods. We propose numerical approximations for the diffusion coefficient, from an extension of the [M. Chesney, J. Elliot, Estimating the instantaneous volatility and covariance of risky assets, Applied Stochastic Models and Data Analysis 11 (1995) 51–58] procedure to the case of nonlinear

* Corresponding author.
E-mail address: rgjaimez@ugr.es (R. Gutiérrez).

doi:10.1016/j.amc.2005.07.047

stochastic differential equations and we establish also computational procedures and simulation algorithm, that are applied to obtened simulated paths of the fitted process. The proposed methodology is applied to two studies carried out in Andalusia (Spain) on females and males life expectancy at birth, between 1944 and 2001.

Keywords: Stochastic Rayleigh diffusion model; Kummer function; Computation of the estimation of drift parameters; Diffusion parameter; Simulation algorithm for stochastic differential equation; Life expectancy at birth

1. Introduction

The stochastic Rayleigh diffusion process (SRDP), from its first formulation Rayleigh [27], has been widely used in physics. For example, the radial Ornstein–Uhlenbeck process, a special case of the general Rayleigh process, is quite naturally related with the description of "optical field" by means of a pair of Ornstein–Uhlenbeck processes, which in turn describe the behaviour, both real and imaginary, of an electrical field, and thus the Ito stochastic differential equation (SDE) of the corresponding amplitude coincides with the SDE of a Rayleigh diffusion process (see, for example, [9]).

In the last few decades, the Rayleigh process, or one of its particular cases, has been discussed in conjunction with important theoretical and practical problems in various aspects of Stochastic Modelling. For example, in the theory of Point Processes, and in particular in modelling based on the Cox process [4], a stationary radial Ornstein–Uhlenbeck process (a particular Rayleigh process) plays an important role, as discussed by Clifford and Wei [3], who showed that if a Cox process has an intensity that is the square of a stationary radial Ornstein–Uhlenbeck process, then it is equivalent to a death process within a "simple stationary immigration, birth and death process", with all the advantages implicit in this, for example with respect to its simulation.

Recently, see for example Davidov and Linetsky [7], the Rayleigh diffusion process has also been considered in the context of the path-dependent options models used in economics and stochastic finance studies. Specifically, the Rayleigh process (radial Ornstein–Uhlenbeck process) is used in relation with important models formulated under the hypothesis that volatility is not constant, but is rather a function of the underlying asset price. On the basis of constant elasticity of variance (CEV) stochastic diffusions, or Cox processes [5], which include as particular cases the lognormal diffusion of the Black–Scholes model [1], the diffusion of the Merton model [24], the diffusion of the Cox–Ross model [6] and others, it can be shown that there exists a specific functional transformation between these CEV diffusion processes and the Rayleigh process and its particular cases. The CEV models, and especially the Rayleigh model, are today recognized to be very appropriate in real-life situations, under which the Black–Scholes model and the others named above have been found unable to achieve good fits. The article by Davidov and Linetsky [7] describes various simulations of the above-cited diffusions and some fits to real applications in the Theory of Options and Risk. In other fields of science, and in particular in Biomedical Science and related areas (such as demographic statistics and the growth of cell populations), we are unaware of applications of the Rayleigh process to theoretical modelling or to statistical fits to observed real data.

Giorno et al. [10], taking as their starting point a definition of the Rayleigh process, in its broad sense, (see Section 2.1 below) established the basic probabilistic theory of this process, obtaining their own transition density function and first-passage time density for arbitrary constant barriers, on the basis of the corresponding Kolmogorov equations (see, for example, [28]). As regards the statistical inference of the Rayleigh process, especially concerning the estimation of its drift and diffusion parameters, it is noteworthy that from the published literature, it does not seem to have been dealt with in a complete, integrated way (i.e. dealing with the estimation of the two types of parameter). With respect to the drift parameters, and in the case of a Rayleigh (radial Ornstein–Uhlenbeck) process with a diffusion coefficient equal to one, which is thus a very special Rayleigh model with a constant volatility equal to one, Prakasa Rao [25] considers the estimation to be based on eigenfunctions.

For other diffusions, and in particular Gompertz-type diffusions, the problem of the statistical inference of their drift and diffusion parameters has been addressed. Drift parameters are estimated by continuous sampling and an approximation of the diffusion coefficient is made by means of formulas based on the quadratic variation of the process. This method represents an alternative to that presented in this paper for the Rayleigh process. In this Gompertzian case, the statistical methodology is applied to real problems of tumor growth [8] and energy consumptions [15].

In all the studies mentioned, and in other, similar ones, the Rayleigh process has been studied theoretically from various viewpoints and it has been applied to the theoretical structural modelling of the above-cited phenomena. However, no study has been carried out to obtain estimators of the parameters by general methods of parametric statistical inference, such that the Rayleigh model can be statistically fitted to real data obtained by time-continuous sampling (sample trajectories) or by discrete sampling (observations within a time discretization).

In the theoretical–practical context of the Rayleigh model as summarised above, the present study is intended to achieve the following objectives: firstly, taking into account the fitting and prediction methodology that will subsequently be used; Section 2.1 completes the probabilistic study of the Rayleigh model developed in Giorno et al. [10], obtaining the trend function and conditioned trend function, together with their asymptotic behaviour in time. Section 3 contains an integrated study of the estimation of the drift and diffusion parameters. The former are estimated by maximum likelihood methods, using continuous sampling; moreover, we propose an approximate calculation by means of numerical procedures applying the discretization of the Riemann integrals appearing in the corresponding expressions. As regards the coefficient of diffusion, we propose a methodology of approximate estimation, which is in fact a variation of the method described by Chesney and Elliot [2]. Finally, we propose the estimations of the conditioned and non-conditioned trend functions. Secondly, following this, Section 4 includes results on the simulation of the Rayleigh process and on the application of the proposed methodology to Females and Males Life Expectancy at Birth in Andalusia (1944–2001).

We note the versatility of the Rayleigh model, in that it is capable of describing, to a considerable degree of accuracy, phenomena with increasing trends and those with decreasing trends phenomena that in general are non-exponential.

2. The model and its trend functions

2.1. Rayleigh diffusion process model

The proposed model is a one dimensional diffusion process with values in $(0,\infty)$, and is defined by the process $\{X_t : t \in [t_0, T]\}$ solution of the following nonlinear stochastic differential equation (SDE) of the first order (see [10] for the general case and [9] for a particular case):

$$\mathrm{d}X_t = \left(\frac{a}{X_t} + bX_t\right)\mathrm{d}t + \sigma\,\mathrm{d}W_t; \quad X(t_0) = x_0, \tag{1}$$

where $\{W(t), t \in [t_0, T]\}$ is a one-dimensional standard Wiener process and $\sigma^2 > 0$, a and b $(b \neq 0)$ are real parameters (to be estimated).

2.2. Computation of the trend function

Taking into account the homogeneity of this process, and using the expression for the transition probability density function (t.p.d.f.) as obtained by Giorno et al. [10] for $a > -\frac{\sigma^2}{2}$ and with the zero-flux condition, the t.p.d.f. of the model is

$$f(x, t/y, s) = \frac{2by^{-\alpha}x^{\alpha+1}}{\sigma^2(e^{2b(t-s)} - 1)} \exp\left(-\frac{b(x^2 + y^2 e^{2b(t-s)})}{\sigma^2(e^{2b(t-s)} - 1)} - \alpha b(t-s)\right) \times \mathrm{I}_\alpha\left(\frac{2bxye^{b(t-s)}}{\sigma^2(e^{2b(t-s)} - 1)}\right),$$

where I_α denotes the modified Bessel function of the first kind and $\alpha = \frac{a}{\sigma^2} - \frac{1}{2}$.

The conditional trend function (CTF) of the process is

$$\mathbf{E}(X_t/X_s = x_s) = \int_0^\infty x f(x, t/x_s, s)\,\mathrm{d}x.$$

Then, we have

$$\begin{aligned}\mathbf{E}(X_t/X_s = x_s) &= \frac{2bx_s^{-\alpha}}{\sigma^2(e^{2b(t-s)}-1)}\exp\left(\frac{-bx_s^2e^{2b(t-s)}}{\sigma^2(e^{2b(t-s)}-1)} - \alpha b(t-s)\right)\\ &\quad\times\int_0^\infty x^{\alpha+2}\exp\left(\frac{-bx^2}{\sigma^2(e^{2b(t-s)}-1)}\right)I_\alpha\left(\frac{2bxx_se^{b(t-s)}}{\sigma^2(e^{2b(t-s)}-1)}\right)\mathrm{d}x.\end{aligned}$$

By applying the change of variable $y = x^2$ and by using the relations Gradshteyn and Ryzhik [11, 6.643],

$$\int_0^\infty e^{-\lambda y}y^{\mu-1/2}\mathrm{I}_{2v}(2\xi\sqrt{y})\,\mathrm{d}y = \frac{\Gamma(\mu+v+\frac{1}{2})}{\Gamma(2v+1)}\xi^{-1}\lambda^{-\mu}\exp\left(\frac{\xi^2}{2\lambda}\right)M_{-\mu,v}\left(\frac{\xi^2}{\lambda}\right),$$

where $\mu + v + 1/2 > 0$ and $M_{-\mu,v}$ is a Whittaker function Sepanier and Oldham [29, p. 477:48–13.1]

$$M_{v,\mu}(x) = x^{\mu+1/2}e^{-x/2}\Phi(\mu - v + 1/2, 2\mu+1, x),$$

where Φ is the confluent hypergeometric function (Kummer function), the conditional trend function of the proposed process leads to

$$\begin{aligned}\mathbf{E}(X_t/X_s = x_s) &= \frac{\Gamma(\alpha+3/2)}{\Gamma(\alpha+1)}\left(\frac{b}{\sigma^2(e^{2b(t-s)}-1)}\right)^{-1/2}\exp\left(\frac{-bx_s^2e^{2b(t-s)}}{\sigma^2(e^{2b(t-s)}-1)}\right)\\ &\quad\times\Phi\left(\alpha+3/2, \alpha+1, \frac{bx_s^2e^{2b(t-s)}}{\sigma^2(e^{2b(t-s)}-1)}\right).\end{aligned}$$

Finally, by the Kummer transformation $\Phi(\beta,\gamma,z) = e^z\Phi(\gamma-\beta,\gamma,-z)$, we deduce that the final form of the conditional trend function of the model is:

$$\begin{aligned}\mathbf{E}(X_t|X_s = x_s) = \frac{\Gamma(\alpha+3/2)}{\Gamma(\alpha+1)}&\left(\frac{b}{\sigma^2(e^{2b(t-s)}-1)}\right)^{-1/2}\\ &\times\Phi\left(-\frac{1}{2}, \alpha+1, \frac{-bx_s^2}{\sigma^2(1-e^{-2b(t-s)})}\right).\end{aligned} \tag{2}$$

From (2) and by considering the initial distribution $\mathrm{P}(X(t_0) = x_0) = 1$, the trend function (TF) of the process is

$$\mathbf{E}(X_t) = \frac{\Gamma(\alpha+3/2)}{\Gamma(\alpha+1)}\left(\frac{b}{\sigma^2(e^{2b(t-t_0)}-1)}\right)^{-1/2} \times \Phi\left(-\frac{1}{2},\alpha+1,\frac{-bx_0^2}{\sigma^2(1-e^{-2b(t-t_0)})}\right). \quad (3)$$

These functions are utilized in the following Section to fit and predict the future evolution of the Stochastic Rayleigh Diffusion Process (SRDP).

Using the relation Sepanier and Oldham [29] p.467: 47–9.6, for very large, positive z and $a \neq 0, -1, -2, \ldots$,

$$\Phi(a,b,z) \sim \frac{\Gamma(b)}{\Gamma(a)} z^{a-b}\exp(z)$$

one can show the right t_0-continuity of (3), thus obtaining $\lim_{t\to t_0^+}\mathbf{E}(X_t) = x_0$.

We can also study the asymptotic behaviour in time of the trend function, thus obtaining, if $b < 0$: $\lim_{t\to+\infty}\mathbf{E}(X_t) = \frac{\Gamma(\alpha+3/2)}{\Gamma(\alpha+1)}\left(\frac{-b}{\sigma^2}\right)^{-1/2}$.

3. Parameter estimation. Computational aspects

We shall now estimate the parameters of the proposed model using two methods: firstly, estimating the drift parameters (the a and b parameters) by the maximum likelihood method, with continuous sampling of the process; and secondly, by estimating the coefficient of diffusion by an approach proposed in this paper, based on an extension of the procedure described by Chesney and Elliot [2] for the case of a linear EDS, to the nonlinear case.

3.1. Estimation of drift parameters

Let us consider the following SDE (scalar)

$$\mathrm{d}X_t = A_t(X_t)\theta\,\mathrm{d}t + B_t(X_t)\,\mathrm{d}W_t; \quad t_0 \leqslant t \leqslant T, \quad (4)$$

where the parameter θ is a $(k\times 1)$-vector, A_t is a $(1\times k)$-vector and B_t is $\mathbb{R}$-valued depending only on the sample path up to the given instant. We assume that Eq. (4) has a unique solution for every θ. The maximum likelihood estimator of the vector θ is given by (see, for example, [23,25])

$$\hat{\theta}_T = S_T^{-1}H_T, \quad (5)$$

where H_T is the following $(k\times 1)$-vector:

$$H_T = \int_{t_0}^{T} A_t^*(X_t)(B_t(X_t)B_t(X_t))^{-1}\,\mathrm{d}X_t \quad (6)$$

and S_T is the $k\times k$-matrix:

$$S_T = \int_{t_0}^{T} A_t^*(X_t)(B_t(X_t)B_t(X_t))^{-1}A_t(X_t)\,\mathrm{d}t \quad (7)$$

and the asterisk denotes the transpose.

The representative Eq. (1) of the model can be written in the vectorial form (4), with:

$$A(X_t) = \left(\frac{1}{X_t}, X_t\right); \quad \theta^* = (a, b) \text{ and } B(X_t) = \sigma.$$

The corresponding vector H_T in Eq. (6) in this case leads us to

$$H_T^* = \frac{1}{\sigma^2}\left(\int_{t_0}^{T} \frac{\mathrm{d}X_t}{X_t}, \int_{t_0}^{T} X_t \,\mathrm{d}X_t\right),$$

S_T is the following square matrix

$$S_T = \frac{1}{\sigma^2}\begin{pmatrix} \int_{t_0}^{T} \frac{\mathrm{d}t}{X_t^2} & (T - t_0) \\ (T - t_0) & \int_{t_0}^{T} X_t^2 \,\mathrm{d}t \end{pmatrix}.$$

Using Eq. (5) and after some calculation (not shown), we obtain the expressions of the estimators

$$\hat{a} = \frac{\left(\int_{t_0}^{T} X_t^2 \,\mathrm{d}t\right)\left(\int_{t_0}^{T} \frac{\mathrm{d}X_t}{X_t}\right) - (T - t_0)\int_{t_0}^{T} X_t \,\mathrm{d}X_t}{\left(\int_{t_0}^{T} \frac{\mathrm{d}t}{X_t^2}\right)\left(\int_{t_0}^{T} X_t^2 \,\mathrm{d}t\right) - (T - t_0)^2},$$

$$\hat{b} = \frac{\left(\int_{t_0}^{T} \frac{\mathrm{d}t}{X_t^2}\right)\left(\int_{t_0}^{T} X_t \,\mathrm{d}X_t\right) - (T - t_0)\left(\int_{t_0}^{T} \frac{\mathrm{d}X_t}{X_t}\right)}{\left(\int_{t_0}^{T} \frac{\mathrm{d}t}{X_t^2}\right)\left(\int_{t_0}^{T} X_t^2 \,\mathrm{d}t\right) - (T - t_0)^2}.$$

The stochastic integrals in the latter expressions can be transformed into Riemann–Stieljes integrals by using the Itô formula, hence

$$\int_{t_0}^{T} \frac{\mathrm{d}X_t}{X_t} = \log(X_T) - \log(x_0) + \frac{\sigma^2}{2}\int_{t_0}^{T} \frac{\mathrm{d}t}{X_t^2},$$

$$\int_{t_0}^{T} X_t \,\mathrm{d}X_t = \frac{1}{2}\left(X_T^2 - x_0^2\right) - \frac{\sigma^2}{2}(T - t_0).$$

Therefore, the resulting maximum likelihood estimators are

$$\hat{a} = \frac{\left(\int_{t_0}^{T} X_t^2 \,\mathrm{d}t\right)\left(\log\left(\frac{X_T}{x_0}\right) + \frac{\sigma^2}{2}\int_{t_0}^{T} \frac{\mathrm{d}t}{X_t^2}\right) - \frac{T - t_0}{2}\left(X_T^2 - x_0^2 - \sigma^2(T - t_0)\right)}{\left(\int_{t_0}^{T} \frac{\mathrm{d}t}{X_t^2}\right)\left(\int_{t_0}^{T} X_t^2 \,\mathrm{d}t\right) - (T - t_0)^2}, \tag{8}$$

$$\hat{b} = \frac{\frac{1}{2}\left(\int_{t_0}^{T} \frac{\mathrm{d}t}{X_t^2}\right)\left(X_T^2 - x_0^2 - \sigma^2(T - t_0)\right) - (T - t_0)\left(\log\left(\frac{X_T}{x_0}\right) + \frac{\sigma^2}{2}\int_{t_0}^{T} \frac{\mathrm{d}t}{X_t^2}\right)}{\left(\int_{t_0}^{T} \frac{dt}{X_t^2}\right)\left(\int_{t_0}^{T} X_t^2 \,\mathrm{d}t\right) - (T - t_0)^2}. \tag{9}$$

The use of these expressions in estimating the parameters requires continuous observations of the process, which in practice cannot be made. In such a situation, we consider a finite number of discrete observations at the instants $t_0 < t_1 < \cdots < t_n = T$, and so the corresponding likelihood function is the product of the transition density. The latter is normally difficult to obtain (in our case it would be the product of very complicated Bessel functions). An alternative method, and one that is frequently used (see, for example, [30,15]), consists in using the above expressions for the estimators and numerically calculating the Riemann integrals appearing in them, for example, by means of the trapezium method.

3.2. Estimation of the diffusion coefficient

In general, no direct estimation methods exist to estimate the coefficient of the diffusion parameter σ. Various expressions have been proposed to obtain approximate estimations, such as the formula proposed by Guerra and Stefanini [12], based on the quadratic variation associated with the process, while others, such as those of Skiadas and Giovani [30] and Katsamaki and Skiadas [21] approximate this coefficient in the case of a linear SDE with multiplicative white noise. In this paper, we propose a formula based on an extension to the case of a nonlinear SDE from the procedure described by Chesney and Elliot [2], as follows: By applying the Itô formula, we have

$$\mathrm{d}\left(\frac{1}{X_t}\right) = -\frac{\mathrm{d}X_t}{X_t^2} + \frac{\sigma^2\,\mathrm{d}t}{X_t^3}.$$

Using the following approximation in the interval $[t-1, t]$:

$$\mathrm{d}\left(\frac{1}{X_t}\right) \simeq \frac{1}{X_t} - \frac{1}{X_{t-1}} \text{ y } \mathrm{d}(X_t) \simeq X_t - X_{t-1}$$

then

$$X_t^3\left(\frac{1}{X_t} - \frac{1}{X_{t-1}}\right) + X_t(X_t - X_{t-1}) = \sigma^2.$$

An estimator for σ is, therefore

$$\hat{\sigma} = \sqrt{X_t/X_{t-1}}\,|X_t - X_{t-1}|.$$

For $n+1$ observations of one trajectory of the process, the resulting estimator has the following expression:

$$\hat{\sigma} = \frac{1}{n}\sum_{t=1}^{n}\sqrt{X_t/X_{t-1}}\,|X_t - X_{t-1}|. \tag{10}$$

Remark. By using Zehna's theorem, the estimated conditional trend function (ECTF) of the SRDP is obtained by replacing the parameters in expression (2) by Eqs. (8)–(10), and thus the ECTF is given by the following expression:

$$\widehat{\mathbf{E}}(X_t|X_s = x_s) = \frac{\Gamma(\hat{\alpha}+3/2)}{\Gamma(\hat{\alpha}+1)}\left(\frac{\hat{b}}{\hat{\sigma}^2(e^{2\hat{b}(t-s)}-1)}\right)^{-1/2} \times \Phi\left(-\frac{1}{2}, \hat{\alpha}+1, \frac{-\hat{b}x_s^2}{\hat{\sigma}^2(1-e^{-2\hat{b}(t-s)})}\right), \tag{11}$$

where $\hat{\alpha} = \frac{\hat{a}}{\hat{\sigma}^2} - \frac{1}{2}$.

4. Applications and simulation. The life expectancy at birth in Spain

4.1. Applications

The homogeneous SRDP model with the statistical methodology proposed above has been used for us to study the evolution of biometric-type variables in Spain, such as "life expectancy at birth", "the rate of infant mortality" and "the number of deaths by cancer". The present paper considers, particularly, the females life expectancy at birth and the males life expectancy at birth which are, in this moment, of particular socio-demographic and economic interest in Spain. Concretely we consider the Spanish region of Andalusia, which has a large degree of self-government and a population of seven and half million. The respective time-dependent stochastic variables (stochastic processes) that are considered are:

(1) $X_1(t)$, the value of female life expectancy at birth corresponding to year, t of birth, in Andalusia (Spain).
(2) $X_2(t)$, the value of male life expectancy at birth corresponding to year, t of birth, in Andalusia (Spain).

In each case, we took the time period of 1944–1999 to fit the corresponding SRDP models $X_1(t)$ and $X_2(t)$. The values observed were the life expectancy at birth calculated for each year of birth. Tables 1 and 2 show the sample values observed for both processes. Note that these values correspond to observations of the two stochastic processes in a time discretisation at equal-amplitude intervals of one year. In both cases, the source for the data was The INE (National Statistics Institute of Spain).

In applying the statistical methodology, the following steps were applied:

(1) Take the values observed for the period 1944–1999 for the two processes, reserving the 2000–2001 values to compare these values with the corresponding ones forecasted by the adjusted models.

Table 1
Observed values and predicted values of $Z_1(t)$

Years	X_1	ETF X_1	ECTF X_1
1944	57.900	57.9000	57.9000
1945	58.618	59.0799	59.0799
1946	59.081	60.1923	59.7565
1947	60.243	61.2426	60.1932
1948	61.779	62.2355	61.2901
1949	62.430	63.1753	62.7425
1950	62.666	64.0656	63.3588
1951	64.229	64.9098	63.5823
1952	65.869	65.7111	65.0641
1953	67.594	66.4722	66.6214
1954	68.595	67.1957	68.2619
1955	69.004	67.8838	69.2149
1956	68.980	68.5387	69.6046
1957	69.581	69.1623	69.5817
1958	70.329	69.7564	70.1545
1959	71.432	70.3228	70.8678
1960	71.775	70.8629	71.9203
1961	72.002	71.3781	72.2477
1962	72.136	71.8699	72.4645
1963	72.546	72.3394	72.5925
1964	72.928	72.7879	72.9841
1965	73.317	73.2163	73.3491
1966	73.486	73.6257	73.7209
1967	73.690	74.0171	73.8824
1968	73.621	74.3913	74.0775
1969	73.839	74.7493	74.0115
1970	73.893	75.0917	74.2199
1971	74.454	75.4194	74.2716
1972	74.625	75.7329	74.8081
1973	74.947	76.0331	74.9717
1974	75.074	76.3205	75.2798
1975	75.535	76.5958	75.4013
1976	75.615	76.8594	75.8429
1977	76.454	77.1119	75.9190
1978	76.433	77.3538	76.7228
1979	77.048	77.5856	76.7026
1980	77.273	77.8077	77.2919
1981	77.595	78.0207	77.5073
1982	78.100	78.2248	77.8155
1983	78.080	78.4204	78.3001
1984	78.632	78.6080	78.2804
1985	78.702	78.7879	78.8099
1986	78.824	78.9605	78.8771
1987	79.029	79.1259	78.9935
1988	79.058	79.2846	79.1904
1989	79.453	79.4368	79.2186

(*continued on next page*)

Table 1 (*continued*)

Years	X_1	ETF X_1	ECTF X_1
1990	79.424	79.5829	79.5977
1991	79.653	79.7230	79.5692
1992	80.136	79.8574	79.7892
1993	80.038	79.9865	80.2525
1994	80.461	80.1102	80.1583
1995	80.650	80.2290	80.5648
1996	80.724	80.3431	80.7459
1997	80.899	80.4525	80.8171
1998	80.976	80.5575	80.9854
1999	81.050	80.6584	81.0590
Predictions			
2000	81.421	80.8481	81.1304
2001	81.710	80.8481	80.4860

Table 2
Observed values and predicted values of $Z_2(t)$

Years	X_2	ETF X_2	ECTF X_2
1944	51.609	51.6099	51.6100
1945	51.701	52.9232	52.9232
1946	52.472	54.1471	53.0079
1947	53.710	55.2902	53.7261
1948	55.904	56.3599	54.8812
1949	56.522	57.3623	56.9340
1950	56.860	58.3031	57.5134
1951	58.604	59.1872	57.8305
1952	60.428	60.0189	59.4690
1953	62.504	60.8022	61.1864
1954	63.517	61.5405	63.1455
1955	63.776	62.2371	64.1030
1956	63.753	62.8948	64.3480
1957	64.256	63.5161	64.3262
1958	65.014	64.1036	64.8021
1959	66.000	64.6594	65.5198
1960	66.326	65.1854	66.4540
1961	66.510	65.6835	66.7630
1962	66.676	66.1555	66.9375
1963	66.977	66.6028	67.0949
1964	67.389	67.0270	67.3805
1965	67.642	67.4294	67.7714
1966	67.671	67.8112	68.0115
1967	67.788	68.1736	68.0391
1968	67.635	68.5178	68.1501
1969	67.793	68.8446	68.0049
1970	67.773	69.1551	68.1549
1971	68.420	69.4502	68.1359

Table 2 (*continued*)

Years	X_2	ETF X_2	ECTF X_2
1972	68.681	69.7307	68.7503
1973	69.121	69.9973	68.9982
1974	69.341	70.2509	69.4163
1975	69.638	70.4920	69.6254
1976	69.720	70.7214	69.9079
1977	70.296	70.9396	69.9854
1978	70.286	71.1473	70.5336
1979	70.882	71.3450	70.5242
1980	71.499	71.5331	71.0909
1981	71.342	71.7122	71.6781
1982	71.891	71.8827	71.5291
1983	72.215	72.0451	72.0515
1984	72.178	72.1998	72.3603
1985	72.210	72.3470	72.3248
1986	72.470	72.4873	72.3550
1987	72.681	72.6210	72.6029
1988	72.672	72.7483	72.8039
1989	72.583	72.8696	72.7953
1990	72.520	72.9852	72.7103
1991	72.587	73.0954	72.6503
1992	73.095	73.2004	72.7140
1993	73.199	73.3005	73.1985
1994	73.590	73.3958	73.2979
1995	73.547	73.4868	73.6704
1996	73.714	73.5734	73.6297
1997	74.310	73.6560	73.7887
1998	74.063	73.7348	74.3572
1999	74.175	73.8099	74.1212
Predictions			
2000	74.846	73.8815	74.2287
2001	75.106	73.9498	74.8687

(2) Calculate the estimators of the parameters of each SRDP model, using expressions Eqs. (8) and (9) for the drift parameters and expression (10) for the respective coefficients of diffusion (volatilities). Expressions Eqs. (8) and (9) are numerically approximated as described in this paper, Section 3.1, using Mathematica 5.1. The parameters obtained are shown in Table 3.

(3) After estimating the parameters, obtain the conditioned trend functions (CTF), using expression (11). The respective trend functions (TF) are obtained from expression (3), replacing the parameters by their corresponding estimators. Tables 1 and 2 show the observed and adjusted values by the conditioned trends of the SRDP models for each of the examples.

Table 3
Estimation of parameters

Variable	$\hat{a}$	$\hat{b}$	$\hat{\sigma}^2$
Z_1	136.83102	−0.01982	0.60859
Z_2	132.35918	−0.02332	0.67293

Finally, the years 2001 and 2002, which were not used for the statistical fit, were compared with the values forecast by the respective ECTF for these years (Tables 1 and 2). Figs. 1 and 2 show, respectively, the observed and adjusted values by ETF and ECTF for each variables.

4.2. Simulation

For the simulation of the sample paths, we have used the procedure proposed by Rao et al. [26], see also, for example, Kloeden and Platen [22]. The derivation of this algorithm involves approximate discretization of the Itô integral equation in time intervals of length h.

In the case of the Rayleigh diffusion process, the algorithm becomes

$$x_{n+1} = x_n + h\left(\frac{a}{x_n} + bx_n\right) + \sigma Z_{1n} + \frac{h^2}{2}\left(b^2 x_n - \frac{a^2}{x_n^3}\right) + \left(b - \frac{a}{x_n^2}\right)\sigma Z_{2n} + \frac{a\sigma^2}{x_n^3}(Z_{1n}Z_{2n} - Z_{3n}); \quad x_0 = x_{t_0},$$

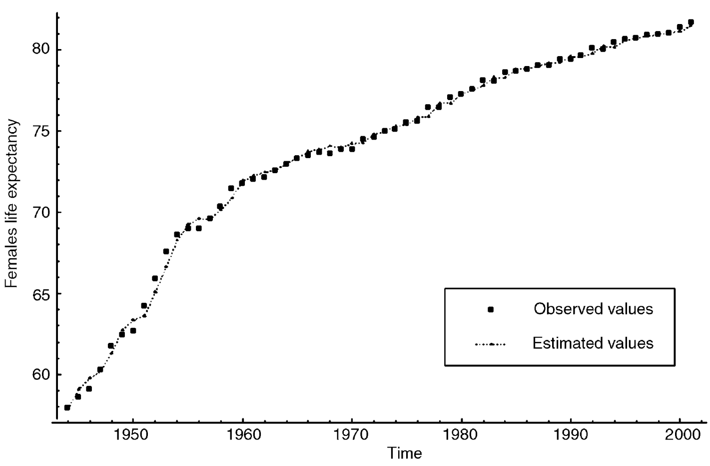

Fig. 1. Fit and forecast of $X_1(t)$, using ECTF.

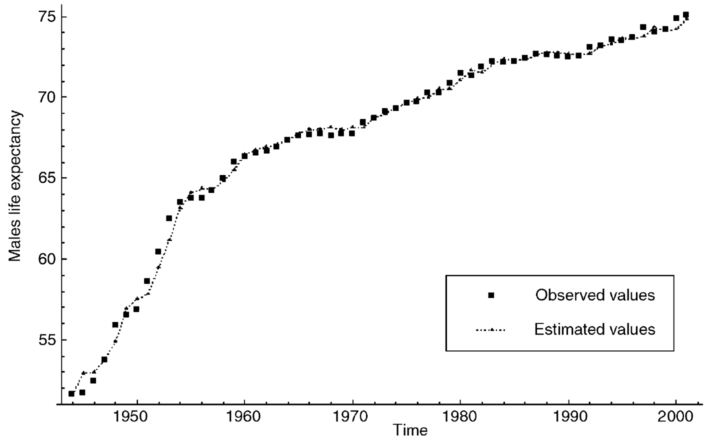

Fig. 2. Fit and forecast of $X_2(t)$, using ECTF.

where $Z_{1n} \rightsquigarrow N[0,h]$, $Z_{2n} \rightsquigarrow N\left[0,\frac{h^3}{3}\right]$, $E[Z_{1n}Z_{2n}] = \frac{h^2}{2}$ and Z_{3n} is not normal, but for small values of h, one can approximate it by a normal variable with zero mean and variance $\frac{h^4}{12}$ verifying $E[Z_{1n}Z_{3n}] = E[Z_{2n}Z_{3n}] = 0$.

The simulation was carried out by taking values of a, b, σ^2 and x_0 that were close to the values estimated for these parameters in the two real-life applications for which this study was developed. In each case we generated 100 trajectories with 601 values each, and considered the time instants $t_i = (i-1)h$, $i = 1,\ldots,101$. Figs. 3 and 4 show, for the particular cases of $a = 136$,

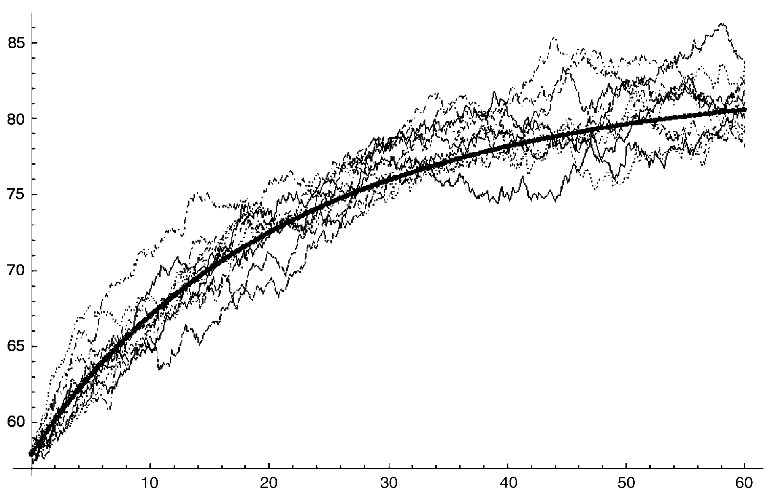

Fig. 3. A simulation of the trajectory of SRDP.

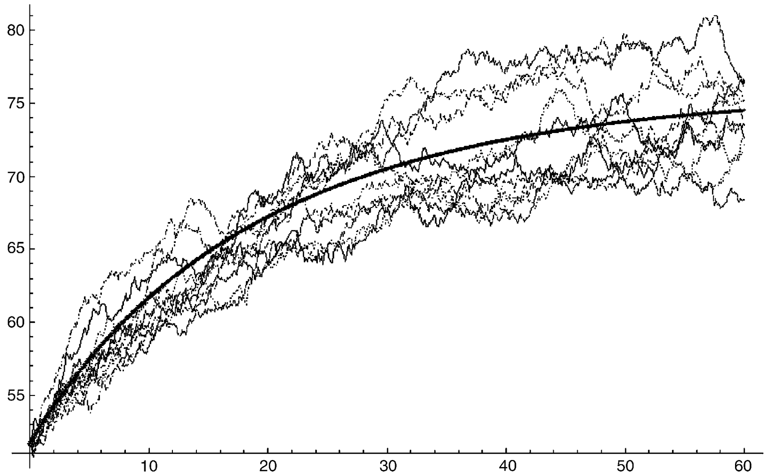

Fig. 4. A simulation of the trajectory of SRDP.

$b = -0.02$, $\sigma^2 = 0.608$, $h = 0.1$ and $x_0 = 57.9$, and of $a = 132$, $b = -0.023$, $\sigma^2 = 0.673$, $h = 0.1$ and $x_0 = 51.6$, which correspond, respectively to values close to those obtained in the study of $X_1(t)$ and $X_2(t)$, ten of the simulated trajectories, together with the theoretical trend function for the values of these parameters. The Figs. 3 and 4 show the simulated trajectories of the Rayleigh process, with the dark line representing the trend.

5. Conclusion and discussion

The main conclusion of this study is that the homogeneous Rayleigh model, under the proposed statistical methodology, provides an accurate fit for the Life expectancy by females and males. Furthermore, the model enables us to accurately predict, in the medium term, the behaviour of the two dynamic variables.

An alternative approach was to fit the two cases described in this paper using other diffusion models, such as those based on lognormal or on Gompertz homogeneous stochastic diffusions (see, for example, [18,14]). For this purpose, a statistical-fit methodology was applied, but in both cases the fits achieved were unsatisfactory. On the other hand, the SRDP model was found to be ideal for interpreting the internal dynamic of the evolution of the two variables considered in this paper, by means of the estimated Ito stochastic equation (1).

As a possible area for future research, a study could be made of a non-homogeneous version of the SRDP model, defined by introducing time-dependent exogenous variables into the trend function, analogously to the work carried out on lognormal and Gompertz diffusions (see, for example, [17,13,19,20,16]). This would enable us to study, for example, the effect of certain preventive health policies on the behaviour of the endogenous variables considered. The Life expectancy at birth is likely that there will be a ‘‘rebound’’ in the next few years as a result of the influence of an exogenous variable related to immigration, which is currently increasing at a fast rate in Spain. Thus, the SRDP process described in this paper, with its associated statistical methodology, is technically ready for the construction of the above-mentioned nonhomogeneous version.

References

[1] F. Black, M. Scholes, The pricing of options and corporate liabilities, Political Economics 81 (1973) 637–659.

[2] M. Chesney, J. Elliot, Estimating the instantaneous volatility and covariance of risky assets, Applied Stochastic Models and Data Analysis 11 (1995) 51–58.

[3] P. Clifford, G. Wei, The equivalence of the cox process with squared radial Ornstein–Uhlembeck intensity and the death process in a simple population model, Annals of Applied Probability 3 (3) (1993) 863–873.

[4] D. Cox, Some statistical methods connected with series of events (with discussion), Journal of the Royal Statistical Society Series B 27 (1955) 129–164.

[5] J. Cox, Notes on option pricing i: constant elasticity of variance diffusions, Stanford University (1955). Reprinted in Journal of Portfolio Management 1996;22:15–17.

[6] J. Cox, S. Ross, The valuation of options for alternative stochastic processes, Journal of Financial Economics 3 (1976) 145–166.

[7] D. Davidov, V. Linetsky, Pricing and hedging path-dependent options under cev processes, Management Sciences 47 (7) (2001) 949–965.

[8] L. Ferrante, S. Bompade, L. Possati, L. Leone, Parameter estimation in a Gompertzian stochastic-model for tumor growth, Biometrics 56 (2000) 1076–1081.

[9] C. Gardiner, Handbook of stochastic methods for physics chemistry and natural sciences, second ed., Springer Verlag, Berlin Germany, 1990.

[10] V. Giorno, A. Nobile, L. Ricciardi, L. Sacerdote, Some remarks on the Rayleigh process, Journal of Applied Probability 23 (1986) 398–408.

[11] L. Gradshteyn, I. Ryzhik, Table of integrals, series and products, Academic Press, 1979.

[12] M. Guerra, L. Stefanini, A comparative simulation study for estimating diffusion coefficient, Mathematical and Computers in Simulation 53 (2000) 193–203.

[13] R. Gutiérrez, A. González, F. Torres, Estimation in multivariate lognormal diffusion process with exogenous factors, Applied Statistics 4 (1) (1997) 140–146.

[14] R. Gutiérrez, R. Gutiérrez-Sánchez, A. Nafidi, Stochastic Gompertz diffusion model statistical inference and applications, in: Proceedings of the 17th Symposium on E.M.C.S.R., Austrian Society for Cybernetic Studies, Vienna, 2004, pp. 146–150.

[15] R. Gutiérrez, A. Nafidi, R. Gutiérrez-Sánchez, Forecasting total natural-gas consumption in spain by using the stochastic gompertz innovation diffusion model, Applied Energy 80 (2) (2005) 115–124.

[16] R. Gutiérrez, A. Nafidi, R. Gutiérrez-Sánchez, P. Román, F. Torres, Inference in Gompertz type nonhomogeneous stochastic systems by means of discrete sampling, Cybernetics and Systems 36 (2005) 203–216.

[17] R. Gutiérrez, L.M. Ricciardi, P. Román, F. Torres, First passage-time densities for time-non-homogeneous diffusion processes, Journal of Applied Probability 34 (3) (1997) 623–631.
[18] R. Gutiérrez, P. Román, F. Torres, A note on the volterra integral equation for the first-passage-time density, Journal of Applied Probability 53 (3) (1995) 635–648.
[19] R. Gutiérrez, P. Román, F. Torres, Inference and first-passage-time for the lognormal diffusion process with exogenous factors: application to modelling in economics, Applied Stochastic Models in Business and Industry 15 (4) (1999) 325–332.
[20] R. Gutiérrez, P. Román, F. Torres, Inference on some parametric functions in the univariate lognormal diffusion process with exogenous factors, Test 10 (2) (2001) 357–373.
[21] A. Katsamaki, C. Skiadas, Analytic solution and estimation of parameters on a stochastic exponential model for a technological diffusion process application, Applied Stochastic Models and Data Analysis 11 (1995) 59–75.
[22] P. Kloeden, E. Platen, The numerical solution of stochastic differential equations, Springer, Berlin, Germany, 1992.
[23] P. Kloeden, E. Platen, H. Schurz, M. Sorensen, On effects of discretization on estimators of drift parameters for diffusion processes, Journal of Applied Probability 33 (1996) 1061–1071.
[24] R. Merton, Theory of rational options pricing, Bell Journal of Economics and Management Science 4 (1973) 141–183.
[25] B. Prakasa Rao, Statistical inference for diffusion type process, Arnold (Ed.), 1999.
[26] N. Rao, J. Borwankar, D. Ramkrishma, Numerical solution of Itô integral equations, SIAM Journal of Control 12 (1) (1974) 124–139.
[27] L. Rayleigh, Philosophical Magazine Letters 32 (1902) 473.
[28] L. Ricciardi, Diffusion processes and related topics in biology, Lecture Notes in Biomathematics, Springer Verlag, Berlin Germany, 1977.
[29] J. Sepanier, K. Oldham, An Atlas of functions, Springer Verlag, Berlin, Germany, 1977.
[30] C. Skiadas, A. Giovani, A stochastic bass innovation diffusion model for studying the growth of electricity consumption in Greece, Applied Stochastic Models and Data Analysis 13 (1997) 85–101.