APPLIED STOCHASTIC MODELS IN BUSINESS AND INDUSTRY
Appl. Stochastic Models Bus. Ind. 2009; **25**:385–405
Published online 24 December 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asmb.754

Modelling and forecasting vehicle stocks using the trends of stochastic Gompertz diffusion models: The case of Spain

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

Department of Statistics and Operations Research, University of Granada, Facultad de Ciencias, Campus de Fuentenueva s/n, 18071 Granada, Spain

SUMMARY

In the present study, we treat the stochastic homogeneous Gompertz diffusion process (SHGDP) by the approach of the Kolmogorov equation. Firstly, using a transformation in diffusion processes, we show that the probability transition density function of this process has a lognormal time-dependent distribution, from which the trend and conditional trend functions and the stationary distribution are obtained. Second, the maximum likelihood approach is adapted to the problem of parameters estimation in the drift and the diffusion coefficient using discrete sampling of the process, then the approximated asymptotic confidence intervals of the parameter are obtained. Later, we obtain the corresponding inference of the stochastic homogeneous lognormal diffusion process as limit from the inference of SHGDP when the deceleration factor tends to zero. A statistical methodology, based on the above results, is proposed for trend analysis. Such a methodology is applied to modelling and forecasting vehicle stocks. Finally, an application is given to illustrate the methodology presented using real data, concretely the total vehicle stocks in Spain.

Received 22 November 2007; Revised 25 November 2008; Accepted 25 November 2008

KEY WORDS: stochastic homogeneous Gompertz diffusion model; statistical inference with discrete sampling; stationary distribution; trends analysis; vehicle stocks in Spain; car industry and climate change

*Correspondence to: R. Gutiérrez-Sánchez, Department of Statistics and Operations Research, University of Granada, Facultad de Ciencias, Campus de Fuentenueva s/n, 18071 Granada, Spain.
†E-mail: ramongs@ugr.es

Contract/grant sponsor: Ministerio de Educación y Ciencia; contract/grant numbers: MTM 2005-09209, MTM-2008-05785
Contract/grant sponsor: Junta de Andalucía; contract/grant number: P06-FQM-02271

1. INTRODUCTION. BACKGROUND AND STUDY AIMS

1.1. Controlling the emission of greenhouse gases and their effect on climate change: challenges facing the car industry

Studying the past, present and future stock of motor vehicles by vehicle types (cars, goods vehicles, motorcycles, etc.), fuel used (petrol, diesel, biofuels), motor size, CO_2 emissions, age, etc., in relation to specific geopolitical areas is a complex, significant task facing society today. It has implications, for example, for the emission of greenhouse gases (especially CO_2), for the demand for oil by-products and for the planning of communications and services networks, among other aspects.

In particular, there is an increasingly evident need to make such a study concerning the manufacture of all kinds of vehicles, as it is a matter of urgency to achieve models that could help reduce the effects of CO_2 emissions and thus alleviate the global problem of climate change.

An important example of this situation is the case of the EU. In February 2007, the European Commission approved a strategy aimed at producing a significant reduction in CO_2 emissions by cars in the EU by the year 2012, when the mean level of CO_2 emissions must be 120 g per km travelled, for all new cars driven within the EU, irrespective of their country of manufacture (some large companies manufacture cars in countries that have not signed the Kyoto protocol). This measure will have a very notable impact on the automobile industry in the next few years. At present, average levels of CO_2 emission are 160 g per km travelled.

The EU intends to give legal force to this strategy by the year 2008, in the light of the successive failures by European industry to observe the voluntary undertakings made; in 1998, for example, companies pledged to reduce average emissions to 140 g/km by 2008, a goal that will not be achieved.

The restructuring of European automobile industries, among others, to meet standards that take environmental considerations into account, has important repercussions for employment and for the economic sector in general. At the 'European Car Industry Restructuring Forum' (Brussels, October 2007), it was decided to create a European Observatory to monitor the effects of this restructuring on the 12 million workers in the car manufacturing industry (2 million directly employed and 10 million indirect workers), a reform that is intended to achieve 'socially responsible' restructuring. This sector currently accounts for 3% of the gross domestic product (GDP) of the entire EU and 4% (some 60 000 million euros) of the total export value of the goods manufactured there.

In coming years, it will also be necessary to take into account the burgeoning use of biofuels (biodiesel and bioethanol) as an alternative source of energy of particular importance to the worldwide sector of land transport; this will have knock-on repercussions of a technical nature and in other fields related to car manufacture. The U.S.A. recently signed agreements with Brazil to guarantee energy supplies, and the EU has set targets for 2010 and 2020 that 6% and 20%, respectively, of automotive fuel should be obtained from biodiesel.

In the last few decades, many studies have been made, using deterministic models, statistical-econometric models and, to a lesser degree, stochastic models, related to the stock of vehicles in certain geopolitical areas (countries, specific urban zones, the EU, etc.). For example, studies on the demand for fuels (petrol), by means of linear translog models, statistically adjusted, with a predictive capacity in terms of the evolution of regressors such as income, prices and car characteristics. One such technique, which is widely utilized, is Archibald and Gillingham's petrol demand model [1], which contains mixed terms that model interactions between regressors (see, for example, Kayser [2]). Many studies have also been carried out to relate the total stock of

DOI: 10.1002/asmb

vehicles and its growth pattern to policies such as technological control measures and territorial management. One such study was applied to the case of the Netherlands (see Priemus [3], Sbayti *et al.* [4]).

In addition, studies have been made on CO_2 emissions produced by the transport sector and the contribution of these to total greenhouse gas emissions in certain geographic areas. For example, Gutiérrez *et al.* [5] analysed the case of Spain, which is a paradigmatic example of a violation of the targets set out in the Kyoto Protocol (1992), see Gutiérrez *et al.* [6]. This study provides comparative data for the transport sector, with respect to the emissions of greenhouse gases, in the U.S.A., the EU and Spain, and models the evolution of these emissions in Spain. The above studies make use of stochastic models to model the effects of the current stock of vehicles within given geopolitical areas, and utilize various types of stochastic diffusion processes (Gompertz, CIR, logistic, Rayleigh, etc.). Technically speaking, the total stock of vehicles is considered an exogenous factor, or regressor, in the stochastic modelling of the evolution of other dynamic variables.

Many other studies published in recent years have considered CO_2 emissions, in various contexts. For example, Paravantis and Georgakellos [7] made an econometric analysis of emissions by passenger cars and buses in Greece. This study, moreover, provided an overview of diverse technical approaches, including those based on sigmoid-type deterministic growth curves (in particular, logistic and Gompertz curves), used to analyse the growth of total vehicle stocks.

In this context, however, few studies have addressed (even partially), using stochastic models, the question of modelling the past evolution and future trend of the total stock of vehicles (classified by types according to diverse criteria), in order to derive statistical predictions of future behaviour, in the medium and long term, based on the corresponding stochastic models statistically fitted to real trend data.

Greenman [8] proposed a random modelling of the car stock in Japan and the U.K.x, in terms of income and other effects, although this technique is not considered to be derived from stochastic modelling based on stochastic diffusion processes.

1.2. Modelling the stock of vehicles by means of stochastic diffusion models: background and goals

In recent decades, stochastic processes, and especially those of a diffusion type, have been subjected to in-depth theoretical study and used for the modelling and statistical prediction of stochastic variables that evolve in time. These studies have been applied to diverse areas such as economics and finance, physics, biology, industry and business. Lognormal, logistic, Bass, CIR and Gompertz diffusion processes, for example, have been used for this purpose.

For example, various real problems in the energy sector have been modelled by means of stochastic diffusion processes. Skiadas and Giovanis [9] and Giovanis and Skiadas [10] studied electricity consumption in Greece and the U.S.A., while the consumption of electricity and of natural gas has been examined in Gutiérrez *et al.* [11, 12].

One of the questions that has aroused greatest interest about these stochastic models, and one that has been the object of numerous studies in recent years, is that of statistical estimation and inference. If these are absent, then the corresponding diffusion process cannot be statistically fitted to the data series observed concerning the real phenomena under study. Such estimation, in general, is not direct, except in simple cases, and one possible methodological approach to deal with it would be based upon the maximum likelihood (ML) method. This, however, has the drawback that

DOI: 10.1002/asmb

the transition densities are rarely known, and it is not usually possible to express the likelihood function in an explicit way. Various methods addressing the question of statistical inference have been developed recently, and several papers have been published on the topic, including those by Bibby and Sorensen [13], Ait-Sahalia [14] and Singer [15], without overlooking the wide-ranging review of results presented by Prakasa Rao [16], who provides a lengthy list of references on the subject.

In the present study, the basis adopted to carry out the modelling is the Gompertz diffusion process.

Gompertz [17] proposed the Gompertz deterministic model (or Gompertz curve), which has been applied successfully in many fields, especially in the study of populational growth phenomena in general, such as cases of tumor growth and the diffusion of new technologies or of innovations. Taking the s-shaped Gompertz slope as a basis, various authors have created stochastic versions of the Gompertz deterministic model. Concretely, a homogeneous Gompertz diffusion process has been established, obtained by adding an infinitesimal variance (white noise fluctuations) to the Gompertz deterministic model. See, for example, Ricciardi [18] and Dennis and Patil [19].

The principal aim of the present study is to demonstrate the potential of the Gompertz homogeneous diffusion process in the stochastic modelling of real cases in the automobile industry. In consequence, we decided to present this model directly, explaining its probabilistic properties for forecasting and for analysing the trends present within applications. In fact, we made a direct examination of the Gompertz model (that is not included in any general class of diffusions), defined from the standpoint of the corresponding Kolmogorov equations and we obtain the probability transition density function (ptdf) and the trend functions (non-conditioned and conditioned) of the process, and on the other, that we obtain its stationary distribution and that of the asymptotic moments. Moreover, we study statistical estimation by means of the ML method, on the basis of discrete sampling of the process, to obtain the estimators of the process parameters. Following this, the statistical inference is completed by establishing the approximate and asymptotic confidence intervals for the drift parameters of the process, also producing ML estimates of the trend functions (both conditional and non-conditional). Alternatively, the homogeneous Gompertz process could be considered via a 'linearization' derived from the general theory of reducible SDEs (see, for example, Kloeden and Platen [20]). This theory is known to provide methods for reducing a nonlinear autonomous SDE in Y_t to a linear autonomous SDE in X_t by means of a time-dependent transformation (and in particular cases by a time-independent transformation). This methodology has been implemented in the case of many specific diffusions, such as the stochastic Bass innovation diffusion model [9], Gompertz homogeneous univariate diffusions, Ornstein–Uhlenbeck diffusions and Malthusian diffusions (see, for example, Skiadas *et al.* [21]), among others. In this respect, therefore, the basic probabilistic results for the Gompertz homogeneous univariate process examined in the present study, as well as those corresponding to certain more general versions of this process, can be obtained by particularizing the general results given in the above-cited theory of reducible SDEs.

With respect to previous approaches to stochastic homogeneous Gompertz diffusion process (SHGDP), for example, let us note that in Gutiérrez *et al.* [22] and in Ferrante *et al.* [23], this process was defined and studied as a solution to Ito's SDE, with the parameters of the drift coefficient being estimated by ML based on continuous sampling of the process, while the diffusion coefficient parameter was determined by approximation methods. Another version of the Gompertz process is the non-homogeneous Gompertz process, which was studied and implemented in Gutiérrez *et al.* [24], and where the Gompertz process was considered with exogenous factors (time-dependent

DOI: 10.1002/asmb

functions that affect the drift). Among other cases, this has been applied to data on new housing prices in Spain, taking as exogenous factors the (GDP), the retail price index and the long-term interest rate. Ferrante *et al.* [25] considered a non-homogeneous Gompertz process with an exogenous factor that was the sum of two exponential functions. Albano and Giorno [26] examined a Gompertz process with a logarithmic exogenous factor. Finally, Patanarapeelert *et al.* [27] studied a Gompertz model in relation to the study of Markov processes with delays, making reference to Gutiérrez *et al.* [24] with respect to statistical processing. Also Meade and Islam [28] cited the statistic methodology study in Gutiérrez *et al.* [11]. Finally, Gutiérrez *et al.* [29] introduced a Gompertz model with a threshold parameter. In all the above-referenced papers, the homogeneous process in question was studied from the standpoint of the corresponding Ito SDE, while the present study is fundamentally based on the Kolmogorov equations with their ptdf. Another difference is that none of the above studies tackles the problem of the stationary distribution of the process; nor do they propose confidence intervals or regions, which are necessary to be able to make a statistical analysis of the predictions obtained.

1.3. Aims and organization of the study

In our framework, as set out in Sections 1.1 and 1.2 above, the specific aim of this study is to model the evolution of diverse total stocks of vehicles (overall, cars, petrol-fuelled cars and diesel-fuelled cars), using stochastic diffusion models with a statistical method to fit the trends, which are subsequently employed as a statistical instrument for short- and medium-term predictions. This methodology provides an alternative to other, classical approaches (econometric models, or chronological series, among others). The proposed modelling procedure is a lognormal–Gompertz-type based on homogeneous Gompertz diffusion processes.

The paper is organized as follows: in Section 2 we consider the homogeneous Gompertz diffusion process in its probabilistic aspects, complementing the results of earlier studies. The process is examined from the perspective of its Kolmogorov equations, on the basis of which we are able to establish the ptdf, the trend functions (both non-conditional and conditional) of the process, and the stationary distribution. Subsequently, the parameters are estimated using the maximum likelihood method on the basis of a scheme of discrete sampling in time. We then determine the approximate and asymptotic confidence intervals of the drift parameters, and from these, deduce those of the trend functions. In Section 3, the study of stochastic homogeneous lognormal diffusion process (SHLDP) is derived from SHGDP. Finally, in Section 4, the methodology examined in this study is applied to real data, namely the evolution of the stock of vehicles in Spain, classified by type of vehicle and by the fuel used (petrol or diesel), statistically fitting Gompertz models on the basis of observations for the period 1978–2005.

2. THE MODEL AND ITS CHARACTERISTICS

2.1. The ptdf and moment of the SHGDP

SHGDP is defined as a diffusion process $\{X_t, t_0 \leqslant t \leqslant T\}$ on $(0,\infty)$ with infinitesimal moments given by (see Ricciardi [30])

$$A_1(x) = \alpha x - \beta x \log(x), \quad A_2(x) = \sigma^2 x^2 \tag{1}$$

DOI: 10.1002/asmb

where $\sigma > 0$, α, and β are real constants. In studies of population and cell growth, the parameter α is the intrinsic growth rate and β is the deceleration factor.

We denote the pdf of the process by $f(y,t\,|\,x,s)$, as both the boundaries 0 and ∞ are natural (for $\beta > 0$), and then f is the unique solution to the following equations, known as the forward Fokker Planck and the backward Kolmogorov expressions:

$$\frac{\partial f(y,t\,|\,x,s)}{\partial t} = -\frac{\partial}{\partial x}[A(y,t)f(y,t\,|\,x,s)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[A_2(y,t)f(y,t\,|\,x,s)]$$

$$\frac{\partial f(y,t\,|\,x,s)}{\partial s} = -A_1(x,t)\frac{\partial f(y,t\,|\,x,s)}{\partial x} - \frac{1}{2}A_2(x,t)\frac{\partial^2 f(y,t\,|\,x,s)}{\partial x^2}$$

with the delta type initial condition $\lim_{t\to s} f(y,t\,|\,x,s) = \delta(y-x)$.

Alternatively, the above-defined process can be considered as the solution of Itô's stochastic differential equation (SDE) (see for example Ferrante *et al.* [23]; Gutiérrez *et al.* [11])

$$\begin{aligned} \mathrm{d}X_t &= (\alpha X_t - \beta X_t \log(X_t))\,\mathrm{d}t + \sigma X_t\,\mathrm{d}W_t \\ X_{t_0} &= x_{t_0} > 0, \quad t \in [t_0, T] \end{aligned} \tag{2}$$

where W_t is a one-dimensional standard Wiener process.

We can show that analytic expression of the process is

$$\begin{aligned} x(t) &= \exp\left(\mathrm{e}^{-\beta(t-t_0)}\log(x_{t_0}) + \frac{\alpha-\sigma^2/2}{\beta}(1-\mathrm{e}^{-\beta(t-t_0)})\right) \\ &\quad \times \exp\left(\left[\frac{\sigma^2}{2\beta}(1-\mathrm{e}^{-2\beta(t-t_0)})\right]^{-1/2} z\right) \quad \text{with } z \rightsquigarrow \mathcal{N}(0,1) \end{aligned} \tag{3}$$

Note that when $\beta = 0$, the SHLDP is obtain as a particular case of SHGDP.

The common solution to the Fokker Planck and Kolmogorov equations can be obtained using Ricciardi's theorem (see Ricciardi [18]) for the transformation of the diffusion process into the Wiener process. The infinitesimal moments equation (1) verifies the conditions of the cited theorem; therefore such a transform exists and has the following form:

$$\Psi(x,t) = \frac{\mathrm{e}^{\beta t}}{\sigma}\log(x) - \frac{\alpha-\sigma^2/2}{\sigma}\int^t \mathrm{e}^{\beta\tau}\,\mathrm{d}\tau$$

$$\Phi(t) = \int^t \mathrm{e}^{2\beta\tau}\,\mathrm{d}\tau$$

From the above, the pdf for the considered process is

$$f(y,t\,|\,x,s) = \frac{1}{y}[2\pi\sigma^2\lambda^2(s,t)]^{-1/2}\exp\left(-\frac{[\log(y)-\mu(s,t,x)]^2}{2\sigma^2\lambda^2(s,t)}\right) \tag{4}$$

DOI: 10.1002/asmb

This function is the density function of the one-dimensional lognormal distribution: $\Lambda_1(\mu(s,t,x),\sigma^2\lambda^2(s,t))$ where $\mu(s,t,x)$ and $\lambda(s,t)$ are given respectively by

$$\mu(s,t,x)=\mathrm{e}^{-\beta(t-s)}\log(x)+\frac{\alpha-\sigma^2/2}{\beta}(1-\mathrm{e}^{-\beta(t-s)})$$

$$\lambda^2(s,t)=\frac{1}{2\beta}(1-\mathrm{e}^{-2\beta(t-s)})$$

Taking into account the random variable $X_t\,|\,X_s=x_s$ has lognormal distribution $\Lambda_1(\mu(s,t,x_s),\sigma^2\lambda^2(s,t))$, and bearing in mind the properties of this distribution, the rth conditional moment of the SHGDP is given by

$$E(X_t^r/X_s=x_s)=\exp\{r\mathrm{e}^{-\beta(t-s)}\log(x_s)\}\exp\left\{\frac{r\alpha}{\beta}(1-\mathrm{e}^{-\beta(t-s)})\right\}$$
$$\times\exp\left\{\frac{r\sigma^2}{4\beta}(1-\mathrm{e}^{-\beta(t-s)})[r(1+\mathrm{e}^{-\beta(t-s)})-2]\right\}$$

from which, the conditional trend function leads us to

$$E(X_t/X_s=x_s)=\exp\left\{\mathrm{e}^{-\beta(t-s)}\log(x_s)+\frac{\alpha-\sigma^2/2}{\beta}(1-\mathrm{e}^{-\beta(t-s)})+\frac{\sigma^2}{4\beta}(1-\mathrm{e}^{-2\beta(t-s)})\right\} \quad (5)$$

Assuming the initial condition $P(X_{t_1}=x_{t_1})=1$, the trend function of the process is

$$E(X_t)=\exp\left\{\mathrm{e}^{-\beta(t-t_1)}\log(x_{t_1})+\frac{\alpha-\sigma^2/2}{\beta}(1-\mathrm{e}^{-\beta(t-t_1)})+\frac{\sigma^2}{4\beta}(1-\mathrm{e}^{-2\beta(t-t_1)})\right\} \quad (6)$$

2.2. Stationary distribution

We shall now determine the stationary distribution of the process, the density function and the asymptotic moments. In general (see Nobile and Ricciardi [31]) the density function of a stationary distribution, $f_{\mathrm{S}}(x)$, in a diffusion can be expressed, under given conditions that satisfy the process, as follows:

$$f_{\mathrm{S}}(x)=\frac{c}{A_2(x)}\exp\left[2\int_z^x\frac{A_1(y)}{A_2(y)}\,\mathrm{d}y\right]$$

where z is an arbitrary point in the interval $]0,+\infty[$, and c is a constant to be determined by the following normalization condition:

$$c=\left[\int_0^{+\infty}\frac{1}{A_2(x)}\exp\left(2\int_z^x\frac{A_1(y)}{A_2(y)}\,\mathrm{d}y\right)\mathrm{d}x\right]^{-1}$$

By applying these results, we can deduce that for $\beta>0$, the density function of the stationary distribution of the process exists and has the following form:

$$f_{\mathrm{S}}(x)=\left[\frac{\pi\sigma^2}{\beta}\right]^{-1/2}x^{-1}\exp\left(-\frac{\beta}{\sigma^2}\left[\log(x)-\frac{\alpha-\sigma^2/2}{\beta}\right]^2\right) \quad (7)$$

DOI: 10.1002/asmb

and that this is also the density of a lognormal distribution $\Lambda_1((\alpha-\sigma^2/2)/\beta;\sigma^2/2\beta)$. Therefore, the k order asymptotic moment of the process (for $\beta>0$) is given by

$$E[X^k(\infty)]=\exp\left(\frac{k}{\beta}\left[\alpha-\frac{\sigma^2(2-k)}{4}\right]\right)$$

The asymptotic trend function of the process is, for $\beta>0$

$$E[X(\infty)]=\exp\left(\frac{\alpha}{\beta}-\frac{\sigma^2}{4\beta}\right) \tag{8}$$

It can be seen that the limit of the trend function in Equation (5) (when t tends to ∞) coincides with the asymptotic trend function in Equation (8).

2.3. Statistical inference

The problem of estimating the parameters of this process has been studied by Ferrante *et al.* [23] and by Gutiérrez *et al.* [11], applying the ML method on the basis of continuous sampling to estimate the drift parameters (α and β), while the parameter of the diffusion coefficient σ is obtained by means of the quadratic variation associated with the process in Ferrante *et al.* [23] and by an extension of the procedure reported by Chesney and Elliot [32], in turn based on the SDE approach that characterizes the process in Gutiérrez *et al.* [11].

In the present study, as the densities of the ptdf of the process are known (the distribution is lognormal), we can constitute the likelihood function corresponding to the process (the product of the ptdf) via discrete sampling of the process and estimating all its parameters by ML, expressing the latter in vector form in order to facilitate the computation when the model is to be applied to real data.

The ML method for parameter estimation was chosen for two main reasons: first, because of its good computational behaviour in the Gompertz model, particularly when this is based on equally-spaced discrete sampling (as in the real case analysed in Section 4). Second, and very importantly, because of its well-known good behaviour with respect to asymptotic situations (for example, asymptotic distributions of estimators, with both discrete and continuous sampling).

2.3.1. Likelihood estimation. The drift parameter α and β and the diffusion coefficient σ of the process are estimated by means of the ML method using discrete sampling. Thus, we consider a discrete sampling of the process $x_{t_1},x_{t_2},\ldots,x_{t_n}$ in times $t_1,t_2,\ldots,t_n$. Furthermore, we assume that the time lag between consecutive observations is one (i.e. $t_i-t_{i-1}=1$, for $i=2,\ldots,n$). Henceforth, the abbreviation $x_i\equiv x_{t_i}$ will be used. In addition by assuming an initial distribution of $P[X(t_1)=x_1]=1$, the associated likelihood function can be obtained from Equation (4) by the following expression:

$$\mathbb{L}(x_1,\ldots,x_n,\alpha,\beta,\sigma^2)=\prod_{j=2}^{n} f(x_j,t_j\,|\,x_{j-1},t_{j-1})$$

As mentioned above, in order to facilitate the computation of the ML estimators and to express them in a simplified form, we shall state the likelihood function in a vector form, considering the following transformation of the sample used: $v_1=x_1$, $v_{i,\beta}=\lambda_\beta^{-1}(\log(x_i)-\mathrm{e}^{-\beta}\log(x_{i-1}))$, for

DOI: 10.1002/asmb

$i=2,\ldots,n$, denoting $\mathbf{V}_\beta=(v_{2,\beta},\ldots,v_{n,\beta})'$. Thus, in terms of $\mathbf{V}_\beta$, the likelihood function is expressed as follows:

$$\mathbb{L}_{\mathbf{V}_\beta}(\mathbf{a},\beta,\sigma^2)=[2\pi\sigma^2\lambda_\beta^2]^{-(n-1)/2}\exp\left(-\frac{1}{2\sigma^2}(\mathbf{V}_\beta-\gamma_\beta\mathbf{aU})'(\mathbf{V}_\beta-\gamma_\beta\mathbf{aU})\right)$$

where $\mathbf{a}=\alpha-\sigma^2/2$, $\gamma_\beta=\lambda_\beta^{-1}(1-\mathrm{e}^{-\beta})/\beta$, $\lambda_\beta=\lambda(t_{i-1},t_i)$ and $\mathbf{U}=(1,\ldots,1)'$ is a vector of the order $(n-1)$.

By differentiating the log-likelihood function with respect to $\mathbf{a}$ and σ^2, we obtain the following equations:

$$\hat{\mathbf{a}}\gamma_\beta\mathbf{U}'\mathbf{U}=\mathbf{U}'\mathbf{V}_\beta$$

$$(n-1)\hat{\sigma}^2=(\mathbf{V}_\beta-\hat{\mathbf{a}}\gamma_\beta\mathbf{U})'(\mathbf{V}_\beta-\hat{\mathbf{a}}\gamma_\beta\mathbf{U})$$

The third likelihood equation is obtained by differentiating the log-likelihood function with respect to β and by using the effect that $\mathbf{V}_\beta=\lambda_\beta^{-1}(J_x-\mathrm{e}^{-\beta}I_x)$ with $J_x=(\log(x_2),\ldots,\log(x_n))'$ and $I_x=(\log(x_1),\ldots,\log(x_{n-1}))'$. After various operations, we have

$$I_x'(\mathbf{V}_\beta-\hat{\mathbf{a}}\gamma_\beta\mathbf{U})=0$$

Taking into account that $\mathbf{U}'\mathbf{U}=n-1$ and after algebraic rearrangement (not shown), the ML estimators of $\mathbf{a}$ and σ^2 are

$$(n-1)\hat{\mathbf{a}}=\gamma_\beta^{-1}\mathbf{U}'\mathbf{V}_\beta \tag{9}$$

$$(n-1)\hat{\sigma}^2=\mathbf{V}_\beta'\mathrm{H}_\mathbf{U}\mathbf{V}_\beta \tag{10}$$

The ML estimator of β is given by

$$\hat{\beta}=\log\left(\frac{I_x'H_\mathbf{U}I_x}{I_x'H_\mathbf{U}J_x}\right) \tag{11}$$

where $H_\mathbf{U}=I_{n-1}-(1/(n-1))\mathbf{UU}'$ is idempotent and a symmetric matrix.

2.3.2. Asymptotic normality of drift parameters. Let X be the random variable with a density function $f_\mathrm{S}(x)$ in Equation (7); then $\log(X)$ is distributed as a normal distribution $N_1((\alpha-\sigma^2/2)/\beta;\sigma^2/2\beta)$. If $\beta>0$, the process under consideration has ergodic properties, and for $\theta^*=(\alpha,\beta)\in(\alpha_1,\alpha_2)\times(\beta_1,\beta_2)$, with $\beta_1>0$, we have

$$\mathscr{L}_\theta(\sqrt{T}(\hat{\theta}-\theta))\rightarrow\mathscr{N}_2(0,\mathbb{I}^{-1}(\theta))\quad\text{when } T\rightarrow\infty \tag{12}$$

where

$$\mathbb{I}(\theta)=\mathbb{E}_\theta\left(\frac{\dot{A}_1(X)\dot{A}_1^*(X)}{A_2(X)}\right)$$

DOI: 10.1002/asmb

and

$$\dot{A}_1(x) = \left(\frac{\partial A_1(x,\theta)}{\partial \alpha}, \frac{\partial A_1(x,\theta)}{\partial \beta} \right)^*$$

Then, we have

$$\mathbb{I}(\theta) = \frac{1}{\sigma^2} \mathbb{E}_\theta \begin{pmatrix} 1 & -\log(X) \\ -\log(X) & \log^2(X) \end{pmatrix} = \frac{1}{\sigma^2} \begin{pmatrix} 1 & -\dfrac{\alpha-\sigma^2/2}{\beta} \\ -\dfrac{\alpha-\sigma^2/2}{\beta} & \dfrac{\sigma^2}{2\beta} + \dfrac{(\alpha-\sigma^2/2)^2}{\beta^2} \end{pmatrix} \tag{13}$$

and the inverse is

$$\mathbb{I}^{-1}(\theta) = \begin{pmatrix} \sigma^2 + \dfrac{2}{\beta}\left(\alpha - \dfrac{\sigma^2}{2}\right)^2 & 2\alpha - \sigma^2 \\ 2\alpha - \sigma^2 & 2\beta \end{pmatrix} \tag{14}$$

An approximated and asymptotic confidence region of θ and an approximated and asymptotic marginal confidence interval of α and β can be obtained by substitution of Equations (12) and (14). The above mentioned region is given, for a large T, by

$$P[T(\theta - \hat{\theta})^* \hat{\mathbb{I}}(\theta)(\theta - \hat{\theta}) \leqslant \chi^2_{2,\gamma}] = 1 - \gamma \tag{15}$$

where $\widehat{\mathbb{I}}(\theta)$ is obtained by replacing the parameters by their estimators in the expression Equation (13) and $\chi^2_{2,\gamma}$ is the upper $100\gamma\%$ points of the chi-squared distribution with two degrees of freedom.

The $\gamma\%$ confidence (marginal) intervals for the parameters α and β are given, for a large T, by

$$P\left(\alpha \in \left[\hat{\alpha} \pm \lambda_\gamma \left(\frac{\hat{\beta}\hat{\sigma}^2 + 2(\hat{\alpha} - \hat{\sigma}^2/2)^2}{\hat{\beta} T} \right)^{1/2} \right] \right) = 1 - \gamma \tag{16}$$

$$P(\beta \in [\hat{\beta} \pm \lambda_\gamma \sqrt{2\hat{\beta}/T}]) = 1 - \gamma \tag{17}$$

where λ_γ is the $100\gamma\%$ points of the normal standard distribution.

In Equations (15), (16) and (17) we have assumed that σ is known with a value $\sigma = \hat{\sigma}$.

Remark 1

By Zehna's theorem, the estimated conditional trend (ECT) and the estimated trend (ET) functions can be obtained from Equations (5) and (6) in the case of SHGDP by replacing the parameters by their estimators. Furthermore, we can obtain an approximated and asymptotic confidence interval

DOI: 10.1002/asmb

of the ETF and ECTF of the SHGDP by means of the approximated and asymptotic confidence interval of the parameters given by Equations (16) and (17).

3. STOCHASTIC HOMOGENEOUS LOGNORMAL DIFFUSION PROCESS

The SHGDP discussed in this paper, together with the non-homogeneous version discussed by Gutiérrez *et al.* [33], is an extension of the lognormal diffusion models that have been studied in depth and applied, see Gutiérrez *et al.* [34–37]. Specifically (see, for example, paragraph 2.1), the SHGDP contains as a particular limit case that of the one-dimensional homogeneous lognormal process. In using the expression 'limit case', we understand the succession of Gompertz processes, this succession being given in terms of a succession of values of the parameter β tending to 0, to converge to the lognormal process. Indeed, it is possible to prove the convergence, for example in distribution, of a succession of Gompertz processes (a succession in β, with $\beta \to 0$)) to the lognormal process. Nevertheless, this question goes beyond the aims of the present study, which is why its proof has not been included. Our aim, by including in Section 3 the lognormal case, which in fact is, technically coincident with that of a Gompertz process with $\beta = 0$ is to demonstrate the possibility of fitting a lognormal model to the variables being studied in the real case analysed in Section 4. Furthermore, we wish to show that the corresponding fits can be discarded in favour of the Gompertz process $\beta \neq 0$. In addition, it is evident that Gompertz transition densities, as deterministic functions that are dependent on β when $\beta \to 0$, tend to the transition density of the lognormal case. It is interesting to observe that although this process is, technically, a special case, it presents certain peculiarities that differ greatly from its Gompertz extension. For example, it does not possess a stationary distribution, but rather presents an increasing exponential trend function that does not weaken.

As stated above, when $\beta \to 0$, we obtain the SHLDP, and then the trend functions and the estimators of the parameters are deduced from those of the Gompertz function, via the limit $\beta \to 0$. Thus, the conditional trend function obtained from Equation (5) has the following form:

$$E(X_t/X_s = x_s) = x_s \exp(\alpha(t-s)) \tag{18}$$

Taking the limit when $\beta \to 0$ in Equation (6), with the initial condition $P(X_{t_1} = x_{t_1}) = 1$, the trend function of the SHLDP is

$$E(X_t) = x_{t_1} \exp(\alpha(t-t_1)) \tag{19}$$

Then, from Equations (9) and (10), we obtain the estimators of the parameters of the SHLDP, and then by denoting $\mathbf{V} = J_x - I_x$, these estimators are expressed as follows:

$$(n-1)\hat{\mathbf{a}} = \mathbf{U}'\mathbf{V} \tag{20}$$

$$(n-1)\hat{\sigma}^2 = \mathbf{V}' H_{\mathbf{U}} \mathbf{V} \tag{21}$$

On the contrary, the stationary distribution and the approximated confidence interval cannot be deduced from the Gompertz functions, but in this case it is possible to determine the exact

DOI: 10.1002/asmb

confidence interval of the parameter, as the distribution of the estimators are known; in the present case, they are

$$\hat{\mathbf{a}} \sim \mathcal{N}_1(\mathbf{a}; \sigma^2/n-1) \quad \text{and} \quad (n-1)\hat{\sigma}^2/\sigma^2 \sim \chi^2_{n-2}$$

$(1-\alpha)\%$ confidence intervals for $\mathbf{a}$ and σ^2 are given respectively by

$$[\hat{\mathbf{a}} - \hat{\sigma} \cdot t_{\gamma/2,n-1}/\sqrt{n-1}, \hat{\mathbf{a}} + \hat{\sigma} \cdot t_{\gamma/2,n-1}/\sqrt{n-1}] \tag{22}$$

$$[(n-1)\hat{\sigma}^2/\chi^2_{\gamma/2,n-1}, (n-1)\hat{\sigma}^2/\chi^2_{1-\gamma/2,n-1}] \tag{23}$$

where $\chi^2_{\gamma,n}$ and $t_{\gamma,n}$ are the upper $100\gamma\%$ points of the chi-squared distribution and student distribution respectively with n degrees of freedom.

Remark 2
By Zehna's theorem, the ECT and the ET functions of the SHLDP can be obtained from Equations (18) and (19) by replacing the parameters by their estimators. In addition the exact confidence interval of the ETF and ECTF of the SHLDP by means of the confidence interval of the parameters is given by Equations (22) and (23).

4. APPLICATION TO REAL DATA: THE CASE OF SPAIN

The challenge to be faced, within the EU as a whole, by the automobile industry, as set out in the Introduction to this paper, is not of the same dimension for all countries or for all manufacturers of certain types of car. Thus, Spain is a paradigmatic case, characterized by its having conspicuously failed to fulfill the commitments made on signing the Kyoto protocol, and by its specialization in the manufacture of mid-upper range vehicles. In this context, an interesting question, among others, is that of modelling the future trends of various 'total vehicle stocks' on the road, classified by type of fuel utilized (petrol or diesel). This question comprises one of the aims of the application discussed below.

The benefits of using diesel fuel rather than petrol, in terms of CO_2 emissions, seem clearcut today. Zervas and Lazarou [38], for example, have analysed this question hypothesizing various future scenarios of the future evolution of the stock of diesel passenger cars in Sweden and in the EU; a point of technical interest is that they did not use time-series stochastic or econometric modelling.

In October 2007, the Spanish Government adopted fiscal measures concerning newly-registered motor vehicles (whether or not manufactured in Spain) concerning the CO_2 emissions they produce. Under the new provisions, those emitting up to 120 g of CO_2 would be exempted from the registration tax; the rate applicable to those producing 121–160 g would be reduced from 7 to 4.75%; for those producing 161–200 g, it would fall from 12 to 9.75%; and finally, for those in the highest category, emitting over 200 g of CO_2, the tax would rise from 12 to 14.75%.

On other hand, these political economic strategies are proposed in Spain, which currently (2007) has 501 cars per 1000 inhabitants (in comparison to 565, 596 and 501 in Germany, Italy and France, respectively). This value is expected to rise to 510 by the year 2012.

In this Spanish context, we shall develop the application of the Gompertz modelling described in Section 2, using the statistical methodology proposed in paragraph 2.3, in order to model evolution

DOI: 10.1002/asmb

Table I. Real data ($\times 10^7$).

Years	$X_1(t)$	$X_2(t)$	$X_3(t)$	$X_4(t)$
1978	0.8952628	0.653042	0.0152199	0.637822
1979	0.9586752	0.705752	0.0184651	0.687287
1980	1.0192748	0.755651	0.0223368	0.733314
1981	1.0666714	0.794332	0.0273418	0.766990
1982	1.1170404	0.835405	0.0328692	0.802535
1983	1.1628151	0.871407	0.0422526	0.829155
1984	1.1390564	0.887444	0.0534743	0.833969
1985	1.1716339	0.927371	0.0646153	0.862755
1986	1.2284071	0.964344	0.0758191	0.888525
1987	1.3072840	1.021852	0.0871487	0.934703
1988	1.3881323	1.078742	0.0987094	0.980033
1989	1.4870484	1.146772	0.1107089	1.036063
1990	1.5696715	1.199564	0.1220746	1.077489
1991	1.6528396	1.253709	0.1317475	1.121962
1992	1.7347203	1.310228	0.1461618	1.164066
1993	1.7809897	1.344069	0.1602062	1.183863
1994	1.8218924	1.373379	0.1806248	1.192754
1995	1.8847245	1.421225	0.2059126	1.215313
1996	1.9542104	1.475380	0.2391352	1.236245
1997	2.0286408	1.529736	0.2806754	1.249061
1998	2.1306493	1.605005	0.3368847	1.268121
1999	2.2411194	1.684739	0.4044419	1.280297
2000	2.3284215	1.744923	0.4702264	1.274697
2001	2.4249871	1.815088	0.5355145	1.279573
2002	2.5065732	1.873263	0.6003919	1.272871
2003	2.5169452	1.868832	0.6592444	1.209587
2004	2.6432641	1.954191	0.7506821	1.203509
2005	2.7657276	2.025037	0.8434725	1.181565

Table II. Estimation of the parameters and the limits of the 95% confidence intervals.

Variable	β lower	β	β upper	α lower	α	α upper	σ^2
$X_1(t)$	0.01080	0.01175	0.01270	0.21712	0.23713	0.25558	3.21626e−004
$X_2(t)$	0.02304	0.02441	0.02578	0.41552	0.44098	0.46506	1.36022e−004
$X_3(t)$	0.02193	0.02327	0.02460	0.44809	0.47632	0.50298	0.00145
$X_4(t)$	0.07694	0.07941	0.08188	1.26836	1.30976	1.34988	1.95418e−004

trends and to obtain statistical forecasts for diverse 'stocks of vehicles'. Specifically, in this study we shall examine the dynamic stochastic variables $X_1(t), X_2(t), X_3(t)$ and $X_4(t)$ that represent, respectively, the 'total stock of vehicles', 'total stock of private cars', 'total stock of private cars—diesel' and 'total stock of private cars—petrol'. These variables are considered to be defined for the continuous time variable t, in which $t \geqslant 0$, such that any of them will provide, for each t, the respective value for total stocks in the year ending at time t. This variable is measured in 'years'.

DOI: 10.1002/asmb

Table III. Fits by using Gompertz ETF ($\times 10^7$).

Years	ETF of $X_1(t)$	ETF of $X_2(t)$	ETF of $X_3(t)$	ETF of $X_4(t)$
1978	0.8952628	0.653042	0.0152199	0.637822
1979	0.9356658	0.688806	0.0183134	0.677636
1980	0.9774269	0.725645	0.0219472	0.716710
1981	1.0205714	0.763546	0.0261984	0.754892
1982	1.0651243	0.802495	0.0311529	0.792054
1983	1.1111101	0.842473	0.0369048	0.828091
1984	1.1585534	0.883463	0.0435578	0.862916
1985	1.2074780	0.925444	0.0512250	0.896463
1986	1.2579078	0.968394	0.0600296	0.928686
1987	1.3098659	1.012290	0.0701051	0.959553
1988	1.3633753	1.057106	0.0815956	0.989046
1989	1.4184584	1.102815	0.0946563	1.017161
1990	1.4751371	1.149391	0.1094534	1.043904
1991	1.5334329	1.196805	0.1261642	1.069292
1992	1.5933668	1.245026	0.1449774	1.093350
1993	1.6549594	1.294025	0.1660930	1.116107
1994	1.7182304	1.343768	0.1897219	1.137600
1995	1.7831993	1.394223	0.2160862	1.157869
1996	1.8498848	1.445358	0.2454185	1.176959
1997	1.9183051	1.497137	0.2779618	1.194915
1998	1.9884777	1.549528	0.3139689	1.211785
1999	2.0604195	1.602495	0.3537021	1.227618
2000	2.1341468	1.656003	0.3974323	1.242462
2001	2.2096751	1.710016	0.4454384	1.256366
2002	2.2870194	1.764499	0.4980068	1.2693797

When we refer to 'total stocks', this is understood as the total number of vehicles, of the class in question, that 'are in circulation' during the annual period under consideration, irrespective of the distance travelled, as these are all potential emitters of CO_2.

For these variables, and as is usually the case at the Department of Economics and Industry, data on real observations are available for calendar years. Specifically, in this study we make use of observations for the period 1978–2005, these data being obtained from the official statistics of the Spanish Ministry of Economics (National Transport Yearbook), and can be consulted at *http//www.ine.es*. On the basis of these real data, we modelled the above-mentioned dynamic variables, using SHGDP, which have been described in Section 2. The methodology consisted of the following steps:

1. We took the data observed for the period 1978–2002, reserving those for 2003–2005 for later confirmation with the values forecasted using the fitted SHGDP. These data are shown in Table I. Using these data, which obviously comprise an 'equally spaced', 'discrete time' sample with a one-year unit interval, we fitted the SHGDP model with the estimation methodology described in paragraph 2.3. In particular, following this methodology, and as described in the Remark 1, we obtained ML estimations for the drift parameters of each diffusion (i.e. the estimators of the individual parameters α and β) and for the coefficient of diffusion (volatility) σ and for the estimated trend functions (ETF and ECTF). The parameters estimated are shown in Table II.

Table IV. Fits by using Gompertz ECTF ($\times 10^7$).

Years	ECTF of $X_1(t)$	ECTF of $X_2(t)$	ECTF of $X_3(t)$	ECTF of $X_4(t)$
1978	0.8952628	0.6530428	0.0152199	0.637822
1979	0.9356658	0.6888063	0.0183134	0.677636
1980	1.0012036	0.7430869	0.0221253	0.726160
1981	1.0637908	0.7943882	0.0266541	0.771080
1982	1.1127139	0.8341033	0.0324837	0.803814
1983	1.1646796	0.8762252	0.0388949	0.838249
1984	1.2118835	0.9131088	0.0497268	0.863962
1985	1.1873856	0.9295265	0.0626128	0.868607
1986	1.2209754	0.9703732	0.0753478	0.896333
1987	1.2794889	1.0081631	0.0881064	0.921096
1988	1.3607357	1.0668750	0.1009665	0.965338
1989	1.4439585	1.1248820	0.1140516	1.008609
1990	1.5457093	1.1941570	0.1275981	1.061893
1991	1.6306441	1.2478494	0.1403997	1.101149
1992	1.7160907	1.3028637	0.1512742	1.143170
1993	1.8001695	1.3602310	0.1674473	1.182838
1994	1.8476621	1.3945529	0.1831720	1.201453
1995	1.8896350	1.4242636	0.2059808	1.209806
1996	1.9540914	1.4727333	0.2341519	1.230978
1997	2.0253468	1.5275489	0.2710497	1.250597
1998	2.1016425	1.5825215	0.3170302	1.262597
1999	2.2061587	1.6585718	0.3790161	1.280426
2000	2.3192840	1.7390449	0.4532214	1.291807
2001	2.4086419	1.7997288	0.5252200	1.286573
2002	2.5074396	1.8704163	0.5964588	1.291130

2. After having obtained the ETs, we determined the fitted values for the period 1978–2002. The results obtained from the ETF are shown in Table III and those from ECTF, in Table IV.
3. We then made forecasts for the year 2003–2005, on the basis of the ETF and the ECTF. First, we applied the asymptotic statistical inference proposed in paragraph 2.3.2, with the approximations described therein. By these means, we obtained the upper and lower limits of the 95% confidence intervals for the parameters of the drift parameters of the process (see Table II). Then, in accordance with the description given in Remark 1, we obtained the predictions for 2003–2005; for the four models under consideration, see Table V, which gives the results for the case in which the ETF was used. Note that in the first three cases, the observed values lie within the 95% confidence region of the ETF, which confirms the high degree of accuracy of the predictions made. For $X_4(t)$, see discussion in Section 5.
4. It is possible to construct a confidence region of the trajectories of the Gompertz process, based on the explicit expression of its random variables $X(t)$ and following a known methodology (see, for example, Katsamaki and Skiadas [39]). In particular, for the fitted Gompertz process (with the technical parameters replaced by their estimators), the following lines can be calculated for a confidence of $(1-\alpha)\%$. This estimated region of confidence includes the one corresponding to the trend function (conditioned or otherwise). With these elements, the validation methodology used in the present study, based on trend functions, can be completed.

DOI: 10.1002/asmb

Table V. Forecasting (2003–2005) using ETF, the limits of confidence regions for ETF and ECTF.

Years	Real data	ETF lower	ETF	ETF upper	ECTF
$X_1(t)$					
2003	2.516945	2.366193	2.596996	2.750818	2.600707
2004	2.643264	2.447211	2.693317	2.857050	2.611343
2005	2.765727	2.530085	2.792019	2.965963	2.740831
$X_2(t)$					
2003	1.868832	1.819417	1.929278	1.990181	1.933765
2004	1.954191	1.874733	1.990129	2.053510	1.929301
2005	2.025037	1.930413	2.051360	2.117164	2.015250
$X_3(t)$					
2003	0.6592444	0.5554299	0.6639763	0.7662196	0.6718881
2004	0.7506821	0.6180062	0.7411469	0.8570027	0.7361637
2005	0.8434725	0.6860381	0.8251928	0.9559318	0.8357693
$X_4(t)$					
2003	1.209587	1.281549	1.300179	1.300457	1.286066
2004	1.203509	1.292921	1.311475	1.311294	1.226894
2005	1.181565	1.303540	1.321995	1.321358	1.221198

Figure 5 shows the behaviour of the estimated trend function with respect to its own region of confidence and to that of the adjusted process, in the case of $X_1(t)$. The same figure also shows 10 simulated paths of the adjusted process.

5. An important complementary question is that of studying the consistency of the estimators obtained for the Gompertz homogeneous univariate model, specifically, the behaviour of its biases and variates when the time is increased, and especially at times ‘previous’ to the inflection of this Gompertz model. By following the theoretical–practical model (see Biby and Sorensen [13]), which has also been applied by Gutiérrez *et al.* [40] in the case of a gamma diffusion process, it is possible to analyse the above-mentioned question for the adjusted Gompertz process in the present case. The bias and variance behaviour, in the time interval considered in the application (1985–2005), which is well before the point of inflection (see Figure 5), is in accordance with the hypothesis that it decreases with increasing time.
6. We have carried out several studies similar to the one described in this paper, considering, in fact, the final 10 and 8 data items, but we concluded that the first and forecasts are less accurate because the sample, in years, is a small one. For this reason we adopted the situation described in the present paper (i.e. taking the last 3 years).
7. For the variable $X_4(t)$, the Gompertz fit was discarded (see Comment 3, Section 5). A fundamental reason for this rejection is, precisely, the behaviour of the 2004–2005 series, the values of which decreased after 2003. This real decrease is poorly modelled by the Gompertz process, which provides slightly rising average ETF values. On the other hand, the ECTF provides clearly falling values. This clearly demonstrates the utility of the ECTF in our methodology; in this it is capable of highlighting a poor fit.
8. The lognormal process has been fitted to the variables considered, discarding the validity of this fit in favour of Gompertz or other models.

DOI: 10.1002/asmb

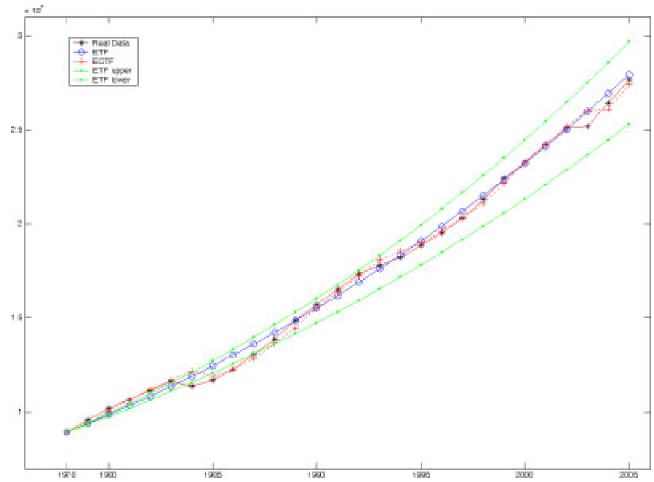

Figure 1. Total stock of vehicles.

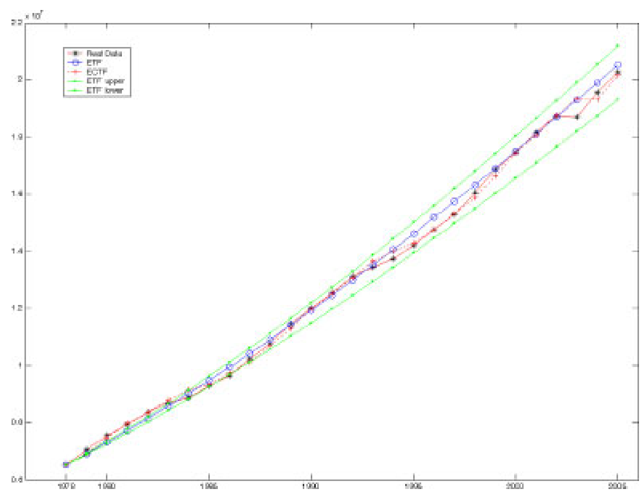

Figure 2. Total stock of private car.

Figures 1, 2, 3 and 4 show, respectively for $X_1(t)$, $X_2(t)$, $X_3(t)$ and $X_4(t)$ the real data, the ECT and the ECTF and the confidence ranges calculated for the respective ETF. These graphs also show the forecasts for 2003–2005.

All the calculations necessary for this application were implemented using MATLAB7.0.1.

5. DISCUSSION AND CONCLUSIONS

- The proposed SHGDP provides a significantly good fit, in statistical terms, to the real evolution of the total stock of motor vehicles in Spain $X_1(t)$, in accordance with real data for the period

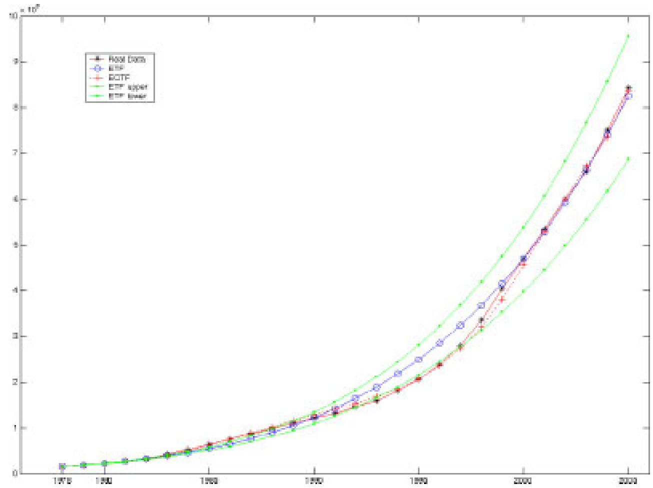

Figure 3. Total stock of private car—diesel.

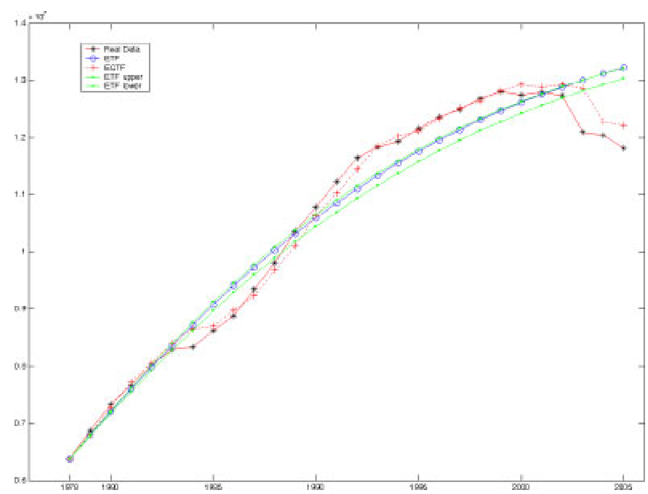

Figure 4. Total stock of private car—petrol.

1978–2002. Figure 1 and Tables II and III show that the ETF and ECTF, estimated using the methodology described in paragraph 2.3, provides a good description of the evolution of vehicles in Spain. The forecasts for 2003, 2004 and 2005, based on the extrapolation of the ETF fitted to the data for the period 1978–2002 (see Table V) are particularly good statistically, as the real data for the period 2003–2005 lie within the confidence interval $(1-\alpha=95\%)$ of the ETF.

- This conclusion is also valid, in every respect, for the total stock of private cars $X_2(t)$ and the subsector of diesel-fuelled cars $X_3(t)$. The fitted results obtained using the ETF and the

DOI: 10.1002/asmb

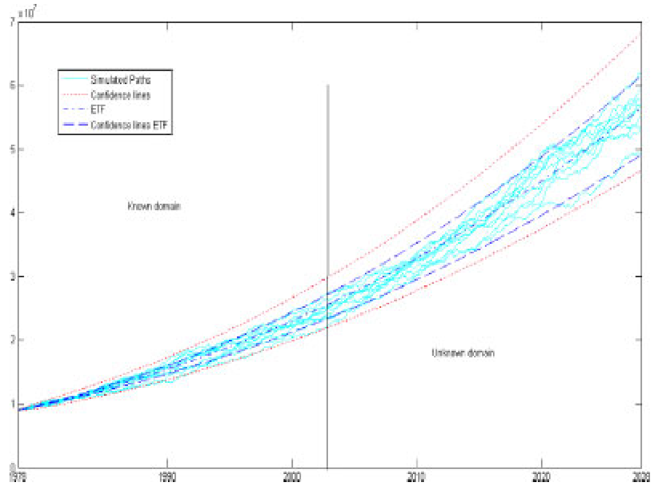

Figure 5. ETF, simulated paths and confidence lines of $X_1(t)$.

ECTF (Figures 2, 3 and Tables II and III), reveal a rising Gompertz trend. The estimate values of β coefficients (Equation 1) in the Gompertz model are 0.02441 and 0.02327, respectively, both values being moderately high. Note that as these coefficients rise or fall (as β tends to 0), the Gompertz model ceases to be valid. In the second case, as it tends to 0 (see Section 3), the Gompertz model becomes a lognormal model. Therefore, β values that are excessively high or low are indicative of the non-validity of the Gompertz trend.

- With respect to the subsector of petrol-fuelled cars in Spain (variable $X_4(t)$), the conclusion we reach is that the Gompertz fit is inadequate, mainly due to a significant variation in the trend beginning in 2002. Thus, a new stochastic model based on a diffusion process of a different type from the Gompertz–lognormal model should be considered. This would probably lead to a gamma-type diffusion model such as that recently described by Gutiérrez *et al.* [40].
- In this case, the β coefficient is very high in relative terms. The β for petrol-driven cars is 0.07941 and that for diesel is 0.02327; the first named, thus, is 3.41 times greater than the second. In fact, the behaviour of the β in the SDE (Equation (2)) reveals significant trend changes in the evolution of the processes in question. Indirectly, this is a further advantage to the proposed method for studying trends versus those based on econometric models or on time series.
- The Gompertz model that we successfully implement in this study can be extended, moreover, to a non-homogeneous version, in cases 1 and 2 above, including exogenous factors within the trend (functioning as regressors). This would make it possible to improve, in statistical terms, the fits obtained in the present study. To do so, we follow the methodology previously described by the authors and implemented it to address real problems (see Gutiérrez *et al.* [40]). This non-homogeneous method would require us to identify the exogenous variables that affect the stock of motor vehicles in Spain, such as the GDP, retail sales prices, interest rates, etc.).

DOI: 10.1002/asmb

ACKNOWLEDGEMENTS

The authors are very grateful to the editor and referees for constructive comments and suggestions. This work was supported partially by research projects MTM 2005-09209 and MTM-2008-05785, Ministerio de Educación y Ciencia and project P06-FQM-02271, Junta de Andalucía, Spain.

REFERENCES

1. Archibald R, Gillingham R. An analysis of the short-run consumer demand for gasoline using household survey data. *The Review of Economics and Statistics* 1980; **62**(4):622–628.
2. Kayser HA. Gasoline demand and car choice: estimating gasoline demand using household information. *Energy Economics* 2000; **22**:331–348.
3. Priemus H. Reduction of car use: instruments of national and local policies. A Dutch perspective. *Environment and Planning and Design* 1995; **22**(6):721–737.
4. Sbayti H, El-Fadet M, Kaysi I, Baaj H. Automotive emissions in developing countries: traffic management and technological control measures. *Environmental Engineering Science* 2001; **18**(6):347–358.
5. Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A. Emissions of the greenhouse gases attributable to the activities of the land transport: modelling and analysis using I-CIR stochastic diffusion. The case of Spain. *Environmetrics* DOI: 10.1002/env.862.
6. Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A, Ramos E. A diffusion model with cubic drift: statistical and computational aspects and application to modelling the global CO_2 emission in Spain. *Environmetrics* 2007; **18**:55–69.
7. Paravantis JA, Georgakellos DA. Trends in energy consumption and carbon dioxide emissions of passenger cars and buses. *Technological Forecasting and Social Change* 2007; **74**:682–707.
8. Greenman JV. The car park: diffusion models revisited. *Energy Economics* 1996; **18**:107–128.
9. Skiadas CH, Giovanis AN. A stochastic Bass innovation diffusion model for studying the growth of electricity consumption in Greece. *Applied Stochastics Model and Data Analysis* 1997; **13**:85–101.
10. Giovanis AN, Skiadas CH. A stochastic logistic innovation diffusion model for studying the electricity consumption in Greece and United States. *Technological Forecasting and Social Change* 1999; **61**:235–246.
11. Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A. Forecasting total natural-gas consumption in Spain by using the stochastic Gompertz innovation diffusion model. *Applied Energy* 2005; **80**(2):115–124.
12. Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A. Electricity consumption in Morocco: stochastic Gompertz diffusion analysis with exogenous factors. *Applied Energy* 2006; **83**:1139–1151.
13. Bibby BM, Sorensen M. Martingale estimation functions for discretely observed diffusion processes. *Bernoulli* 1995; **1**(1/2):17–39.
14. Ait-Sahalia Y. Maximum-likelihood estimation of discretely-sampled diffusion: a closed-form approximation approach. *Econometrica* 2002; **70**:223–262.
15. Singer H. Parameter estimation of nonlinear stochastic differential equations: simulated maximum likelihood versus extended Kalman filter and Itô–Taylor expansion. *Journal of Computational and Graphical Statistics* 2002; **11**(4):972–995.
16. Prakasa Rao BLS. *Statistical Inference for Diffusion Type Process*. Arnold: New York, 1999.
17. Gompertz B. On nature of the function expressive of the law of human mortality and on an new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London* 1825; **115**:513–583.
18. Ricciardi LM. On the transformation of diffusion processes into the Wiener processes. *Journal of Mathematical Analysis and Applications* 1976; **54**:185–199.
19. Dennis B, Patil GP. Application in ecology. In *Lognormal Distributions*: *Theory and Applications*, Crow EL, Shimizu K (eds), Chapter 12. Marcel Dekker: New York, 1988.
20. Kloeden P, Platen E. *The Numerical Solution of Stochastic Differential Equations*. Springer: Berlin, Germany, 1992.
21. Skiadas CH, Giovanis AN, Dimoticalis J. Investigation of stochastic differential models: the Gompertzian case. In *Selected Topics on Stochastics Modeling*, Gutiérrez R, Valderrama M (eds). World Scientific: Singapore, 1994; 296–310.
22. Gutiérrez R, Nafidi A, Gutiérrez-Sánchez R. Inference in the stochastic Gompertz diffusion model with continuous sampling. *Monografias del Seminario Garcia de Galdeano* 2004; **31**:247–253.

23. Ferrante L, Bompade S, Possati L, Leone L. Parameter estimation in a Gompertzian stochastic model for tumor growth. *Biometrics* 2000; **56**:1076–1081.
24. Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A, Román P, Torres F. Inference in Gompertz type non homogeneous stochastic systems by means of discrete sampling. *Cybernetics and Systems* 2005; **36**:203–216.
25. Ferrante L, Bompade S, Possati L, Leone L, Montanari MP. A stochastic formulation of the Gompertzian growth model for *in vitro* bactericidad kinetics: parameter estimation and extinction probability. *Biometrical Journal* 2005; **47**(3):309–318.
26. Albano G, Giorno V. A stochastic model in tumor growth. *Journal of Theoretical Biology* 2006; **242**(2):329–336.
27. Patanarapeelert K, Frank TD, Friedrich R, Tang IM. On reducible nonlinear time-delayed stochastic systems: fluctuation–dissipation relations, transitions to bistability, and secondary transitions to non-stationarity. *Journal of Physics A*: *Mathematical and General* 2005; **38**:10069–10083.
28. Meade N, Islam T. Modelling and forecasting the diffusion of innovation—a 25-year review. *International Journal of Forecasting* 2006; **22**(3):519–545.
29. Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A, Ramos E. A new stochastic Gompertzian process: computational inference and application in modelling. *Applied Mathematics and Computation* 2006; **183**:738–747.
30. Ricciardi LM. *Diffusion Process and Related Topic in Biology*. Lectures Notes in Biomathematics, vol. 14. Springer: Berlin, 1977.
31. Nobile AG, Ricciardi LM. Growth with regulation in fluctuating environments. *Biology Cybernetics* 1984; **49**:179–188.
32. Chesney M, Elliot RJ. Estimating the volatility of an exchange rate. In *Applied Stochastic Model and Data Analysis*, Janssen J, Skiadas CH (eds). World Scientific: Singapore, 1993; 131–135.
33. Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A. Trend analysis using nonhomogeneous stochastic diffusion processes. Emission of CO_2; Kyoto protocol in Spain. *Stochastics Environmental Research Risk Assessment* 2008; **22**:57–66.
34. Gutiérrez R, Angulo JM, González A, Pérez R. Inference in lognormal multidimensional diffusion process with exogenous factors: application to modelling in economics. *Applied Stochastic Models and Data Analysis* 1991; **7**:295–316.
35. Gutiérrez R, González A, Torres F. Estimation in multivariate lognormal diffusion process with exogenous factors. *Applied Statistics* 1997; **46**(1):140–146.
36. Gutiérrez R, Román P, Torres F. 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* 1999; **15**(4):325–332.
37. Gutiérrez R, Román P, Torres F. Inference on some parametric functions in the univariate lognormal diffusion process with exogenous factors. *Test* 2001; **10**(2):357–373.
38. Zervas E, Lazarou C. CO_2 benefit from the increasing percentage of diesel passenger cars in Sweden. *International Journal of Energy Research* 2007; **31**(2):192–203.
39. Katsamaki A, Skiadas CH. Analytic solution and estimation of parameters on stochastics exponential model for a technological diffusion process. *Applied Stochastics Model and Data Analysis* 1995; **11**:59–75.
40. Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A, Merbouha A. Stochastic diffusion process based on the Gamma density function: statistical inference. *Monografías del Seminario García de Galdeano* 2008; **34**:117–125.

DOI: 10.1002/asmb