ENVIRONMETRICS
Environmetrics 2008; **19**: 137–161
Published online 5 February 2008 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/env.862

Emissions of greenhouse gases attributable to the activities of the land transport: modelling and analysis using I-CIR stochastic diffusion—The case of Spain

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

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

SUMMARY

In this study, carried out on the basis of the conclusions and methodological recommendations of the Fourth Assessment Report (2007) of the International Panel on Climate Change (IPCC), we consider the emissions of greenhouse gases (GHG), and particularly those of CO_2, attributable to the activities of land transport, for all sectors of the economy, as these constitute a significant proportion of total GHG emissions. In particular, the case of Spain is an example of a worrying situation in this respect, both in itself and in the context of the European Union. To analyse the evolution, in this case, of such emissions, to enable medium-term forecasts to be made and to obtain a model that will enable us to analyse the effects of possible corrector mechanisms, we have statistically fitted a inverse Cox–Ingersoll–Ross (I-CIR) type nonlinear stochastic diffusion process, on the basis of the real data measured for the period 1990–2004, during which the Kyoto protocol has been applicable. We have studied the evolution of the trend of these emissions using estimated trend functions, for which purpose probabilistic complements such as trend functions and stationary distribution are incorporated, and a statistical methodology (estimation and asymptotic inference) for this diffusion, these tools being necessary for the application of the analytical methodology proposed.

KEY WORDS: climate change; Cox–Ingersoll–Ross model; emission of greenhouse gases; simulation of SDE; statistical inference for a Inverse-CIR diffusion

1. INTRODUCTION AND AIMS OF THE STUDY

In order to identify the aims of this study and the results to be obtained, let us now consider some general aspects of climate change, in the context of previous studies and of the present situation, thus constituting the scenario in which appropriate action must be taken, in the near future, to mitigate the negative consequences of this phenomenon.

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

Received 9 March 2007
Accepted 4 June 2007

1.1. Climate change according to the Fourth Assessment Report of the International Panel on Climate Change, as part of the UN Environment Programme

The IV Assessment Report of the International Panel on Climate Change (IPCC), as part of the UN Environment Programme, was recently presented at the international conference 'Citizens of the Earth' (Paris, February 2007), as a continuation of the Reports published in 1990, 1996 and 2001. The latest report has aroused enormous interest in the world's media, throughout society and at all political levels, and it has had an immediate effect on social and governmental awareness of the far-reaching changes and consequences that will be caused by climate change during the 21st century. Among other political and governmental repercussions, it has inspired the proposal that the work carried out by the UN through the IPCC should be transferred to a UN body for the environment with greater power and institutional weight than the current programme.

The most startling conclusion of this Fourth Report is that, at a worldwide scale, the earth's temperature will rise by between 1.8 and 4°C during the present century, and that in the most unfavourable scenario, this rise could be as much as 6°C.

With respect to the causes of this change, the report concludes that 'human responsibility' exceeds 90%, a significant increase on the 66% reported in the Third Report IPCC (2001). This conclusion, in the words of the IPCC President, R. Pachauri, is to some extent now irrelevant, because 'the debate on the link between human actions and climate change should really be over. Henceforth, the debate should only consider the measures that must be taken'. Furthermore, the final text of the report, agreed upon by 500 scientists from 113 countries, sets out its most significant conclusions in terms of what is 'very likely', what is 'virtually certain' and what is 'probable'. On this scale, the IPCC assumes, respectively, a likelihood of over 90%, a degree of certainty of 99% and a probability of 66%. Thus, the final conclusion referred to above would mean that it is 'very likely' (>90% certainty) that the increase in worldwide mean temperatures observed since the mid-20th century is the result of the increase in concentrations of greenhouse gases (GHG), due to mankind's use of fossil fuels. According to the IV Report, the global warming caused by human activities will continue for centuries, even if concentrations of GHG in the atmosphere become stabilised. Past and future emissions of CO_2 (and of GHG in general) will continue contributing to global warming and their consequences will last for over a millennium. Thus, it becomes a matter of urgent priority to seek to control the greenhouse effect, reducing CO_2 emissions by at least 50%.

Nevertheless, not all governments and scientific debates around the world are in agreement in identifying human activities as the basic cause of climate change. There are also disagreements concerning the methodology to be adopted by the IPCC Work Group (I Summary, II Impacts, III Response Strategies) and the terminology used by experts, depending on whether they work in Experimental Sciences or in Social and Human Sciences. In some scientific circles, while acknowledging the important role played by the reports in stimulating interest in climate change and in research into Environmental Sciences in general, questions have been raised that will contribute to a greater rigor and precision in the conclusions reached concerning causes, impacts and the measures to be adopted in response to climate change. The paper-editorial by Ha-Duong *et al.* (2007) provides a general perspective, and a very lucid one from the standpoint of experimental science, on the improvements that have been made to the methodology used by the Working Groups, in the respective reports. Moreover, a classification is made of the terminology used by each Working Group in their application of the concepts underlying the different types of probability (classical, frequentist, Popper's, views, e.g. Fine (1973)) or degrees of belief in subjective (Bayesian) theories. Also discussed are issues concerning uncertainly in natural systems and in human and social systems, in contrast to the concept of risk (using precise probabilities).

DOI: 10.1002/env

In view of the IPCC Reports, in the coming years it is foreseeable that more studies will be made to objectively determine reality and to control local, national and transnational situations, as far as possible, taking into account the feasibility of joint actions, especially in sectors such as energy, transport, house construction and urban development in general. Moreover, the IPCC methodology will be particularised to specific geographic systems. For example to cite just a few recent studies in the above fields, taking as a starting point four emission scenarios defined and recommended by the IPCC in 2001 (Special Report on Emissions Scenarios—AOGCM), together with five Atmosphere Ocean General Circulation Models, Nogues-Bravo *et al.* (2007) studied the global climate warming of mountain systems during the 21st century. Such systems are highly sensitive to climate change. This study provides results that are coherent with those expressed in the Conclusions of the IV IPCC Report (2007), referred to above. For example, in the A1F1 IPCC scenario (global economics), characterised by intensive use of fossil fuels (high rate of emission of GHG), rapid economic growth, low population growth and the rapid introduction of new and more efficient technologies, the most pessimistic of the possible situations, the authors estimate projected average temperature changes of $+3.2°C$ for 2055 and of $5.3°C$ for 2085. The methodology by which these projected temperature increases are obtained for the different scenarios proposed in this paper, followed by their statistical analysis (fundamentally based on ANOVA), is paradigmatic of the way global studies of climate change are structured and applied.

On the subject of studies focused on control, also paradigmatic is a recent paper by Den Elzen *et al.* (2007), which addresses general aspects concerning the control and the costs of emissions of GHG (the Kyoto Protocol), from the quantitative standpoint and within the framework of the IPCC methodologies, taking into consideration probabilistic aspects of the question, with particular reference to the lognormal distribution, among others, which has been shown to play an influential role in the corresponding statistical analysis.

1.2. The emission of greenhouse gases attributable to land transport in the EU and in Spain

In order to be able to make an objective analysis of given policies concerning the control and reduction of CO_2 emissions, we require statistical models that are appropriate to the observed trends, enabling us to make short, medium and long-term predictions. It is also desirable that such models should have the capability to be affected, especially as regards the trend, by exogenous dynamic variables, so that we may study the effect of policies modelled by these variables on the response of the respective trend function.

In such a complex global economy, which is strongly interrelated with social, political and geographic aspects, etc., and in the context of which control policies are of necessity sectoral, what is required is models of CO_2 emission classified by emission sources or by geopolitical areas (such as the EU), taken as ambits for common policies for controlling the sources of CO_2 emission.

Such is the case, for example of CO_2 emissions resulting from land transport, which is one of the heaviest producers of these emissions. It would be very useful to have a model of growth trends of CO_2 in this sector, for specific countries or geopolitical areas, because exogenous control policies could then be applied to land transport. The effectiveness of such policies could be studied if a good model were available, especially if it enabled the inclusion of exogenous factors to model policies for restricting or reducing the growth of the endogenous variable (i.e. the emissions of CO_2).

The land transport constitutes an important source of total emissions of GHG (and of CO_2 in particular), at a world scale, in the EU and in Spain (this latter case is analysed in detail in Subsection 4.1). Let us now summarise the evolution of these emissions in the U.S.A., the EU and Spain, during the period 1990–2004. Table 1 presents data for the U.S.A. and the EU, for the above period, for the total

DOI: 10.1002/env

Table 1. Total GHG emission and GHG emission by land transport in EU and U.S.A.

Year	UE			U.S.A.		
	Total GHG emission	GHG transport emission	% GHG transport emission	Total GHG emission	GHG transport emission	% GHG transport emission
1990	4 252 461	701 677	16.50	6 103 283	1 459 958	23.92
1991	4 264 805	717 344	16.82	6 066 323	1 434 327	23.64
1992	4 176 880	743 302	17.80	6 140 485	1 483 212	24.15
1993	4 108 108	751 339	18.29	6 327 120	1 534 976	24.26
1994	4 106 695	754 993	18.38	6 370 859	1 575 141	24.72
1995	4 144 433	765 900	18.48	6 477 148	1 604 637	24.77
1996	4 231 297	783 983	18.53	6 678 309	1 641 105	24.57
1997	4 171 982	794 082	19.03	6 703 780	1 658 937	24.75
1998	4 184 732	819 371	19.58	6 767 132	1 686 776	24.93
1999	4 119 135	839 625	20.38	6 808 241	1 749 538	25.70
2000	4 129 317	841 976	20.39	6 975 929	1 794 114	25.72
2001	4 174 119	852 575	20.43	6 886 890	1 777 223	25.81
2002	4 155 328	863 585	20.78	6 909 407	1 820 865	26.35
2003	4 216 469	870 033	20.63	6 952 561	1 818 396	26.15
2004	4 228 006	884 432	20.92	7 067 570	1 869 641	26.45

Data in Gg CO_2 eq.

GHG and the total GHG attributable to the transport sector. Table 2 shows the corresponding data for Spain, together with the total emission of CO_2 with respect to total emissions of GHG.

In this case, the variable considered (with respect to total emission of GHG attributable to the transport sector) is 'Emissions from the combustion and evaporation of fuel for all transport activity, regardless of the sector. Emissions from fuel sold to any air or marine vessel engaged in international transport (international bunker fuels) are not included'. The global emission value is taken to be 'the

Table 2. Total GHG emissions, total CO_2 emissions and GHG emissions by land transport in Spain

Year	Total emission GHG	CO_2 emissions	% CO_2 emissions	GHG emission transport	% GHG emission transport
1990	287 152	205 535	71.58	57 536	20.04
1991	293 134	211 251	72.07	59 849	20.42
1992	300 912	217 612	72.32	63 690	21.17
1993	289 550	207 015	71.50	63 011	21.76
1994	305 784	220 052	71.96	65 985	21.58
1995	317 941	230 977	72.65	67 028	21.08
1996	310 540	216 991	69.88	71 735	23.10
1997	331 324	235 215	70.99	72 498	21.88
1998	341 618	243 047	71.15	79 485	23.27
1999	369 927	267 744	72.38	84 273	22.78
2000	384 246	277 453	72.21	87 002	22.64
2001	384 552	279 792	72.76	91 277	23.74
2002	402 060	299 128	74.40	93 462	23.25
2003	408 169	303 602	74.38	98 045	24.02
2004	427 905	324 020	75.72	102 011	23.84

Data in Gg CO_2 eq.

DOI: 10.1002/env

global emission of GHG from all sources'. The data were obtained from the United Nations Framework Convention on Climate Change (http:www.ufccc.int).

From the data presented, we may draw the following conclusions:

1. The total emission of GHG in 2004, with respect to 2004, has increased in the U.S.A. by 15.8%, in the EU by −0.58% and in Spain by 49%.
2. The total emission of GHG attributable to the transport sector in 2004, with respect to 1990, has increased in the U.S.A. by 28%, in the EU by 26% and in Spain by 77.3%.
3. The proportion of emissions of GHG by the transport sector in 2004, with respect to total GHG emissions, was 26.4% in the U.S.A., 20.92% in the EU and 23.64% in Spain.
4. During the period 1990–2005, the above proportion ranged as follows: in the U.S.A. between 23.64 and 26.45%; in the EU, between 16.5 and 20.92%; and in Spain, between 20 and 24.62%. In all three cases, the trend was a continuously rising one.

1.3. Aims and background of this study

The scenario described in the preceding Subsections (1.1, 1.2) can be summarised as follows:

1. The emission of GHG is the cause of global warming, with a probability of 90% (according to the Fourth IPCC Assessment Report, 2007).
2. At a worldwide scale, especially in geopolitical spaces or countries classified as belonging to the A1F1 scenario of the IPCC Assessment Reports, in the EU and in Spain in particular, the emission of GHG attributable to the land transport sector contributes significantly to the total emission of GHG. This sectoral emission, moreover, tends to increase in many cases (e.g. Spain) and is, in the best of cases, only lightly regulated (e.g. the EU).
3. CO_2 constitutes the main component of the total emission of GHG within these countries and areas.
4. It is essential to adopt control measures so that, in the medium–long term, emission levels may be achieved that are compatible with sustainable socio-economic development. Moreover, we require mathematical–statistical models that will enable us to objectively study the effects of specific sectoral policies on the levels of emissions of different sources of GHG.
5. The case of Spain is particularly serious both in general and within the particular context of the EU. This country, as a signatory to the Kyoto Protocol in 1997, was assigned, on the basis of its total GHG emissions in 1990, an increase of 15% by the year 2010. In fact, however, from the trend observed during the period 1990–2005, it is estimated that the corresponding increase for this period will be 51.3%. Under the same circumstances, the total increase in GHG emissions in the EU as a whole is forecast to be −4.6%. This situation has been examined in greater detail in Gutiérrez *et al.* (2007a, 2007b). From the above analysis, it is evident that a fundamental cause of Spain's poor performance is the effect on the emission of GHG attributable to the land transport sector.
6. The latest studies carried out within the EU and in Spain in particular conclude that the annual mean temperature in Spain rose by 1.53°C between 1971 and 2000, and that in the last 100 years, it has risen by 1.5°C. The corresponding values for the EU and for the Earth as a whole are 0.95°C and 0.65°C, respectively (Source: Spanish Ministry of the Environment, and the United Nations).

We see, thus, from Scenarios 1–6, that the present study responds to well-founded concerns. It seeks to contribute to our knowledge of the patterns of GHG emissions attributable to the land transport sector in a country that is representative of Scenario A1F1 (see Subsection 1.1).

DOI: 10.1002/env

The main aim of this study, from the applied standpoint, is to model the secular trend that is present in the evolution of the emission of GHG attributable to the land transport sector in Spain, on the basis of observations made during the period 1990–2004, which covers a period of 15 years during which the Kyoto Protocol (1997) was in force. This study then seeks to predict the future development of this trend in the medium term (2005–2008). This goal is further discussed in Sections 4 and 5.

From a technical standpoint, we propose to carry out stochastic modelling, and to apply a statistical analysis methodology based on using the data observed to fit the trend and the conditioned trend functions corresponding to a (homogeneous) inverse Cox–Ingersoll–Ross (I-CIR) diffusion process. This stochastic model was selected from various possible stochastic diffusion processes (including Gompertz, Rayleigh, Lognormal and cubic processes) as being the most suitable for the situation to be studied. An original contribution of the present study is the explicit calculation, for the above purposes, of the trend functions, the EMV methodology, based on continuous sampling of its parameters (the drift and the coefficient of volatility of the diffusion coefficient), the stationary distribution and some results concerning asymptotic inference that enable us to establish confidence regions for the drift estimators of the process being considered, as well as the corresponding trend functions.

As concerns the immediate background to this study, let us cite the following: for the case of Spain, Gutiérrez *et al.* (2007a) examined the emission of CO_2, a fundamental component of GHG, in a one-dimensional way, modelling the evolution of this variable using a cubic-type homogeneous diffusion process. Gutiérrez *et al.* (2007b) studied the evolution of CO_2 emissions in relation to changes in the GDP. In this case, the most appropriate modelling instrument was found to be that based on a non-homogeneous Gompertz diffusion process, originally proposed by the authors (see too, e.g. Gutiérrez *et al.*, 2006b, in which the interrelation of CO_2 emissions and GDP was studied in order to determine the extent to which an increase in the latter variable affected these emissions). In both studies, the specific variable examined was 'CO_2 emissions from fossil-fuel burning, cement manufacture and gas flaring'. In both cases, the trend functions were analysed by means of a statistical methodology originally proposed by the authors (see, e.g. Gutiérrez *et al.*, 1999, 2001, 2005a, 2005b, 2006a, 2006b), subsequently cited in Meade and Islam (2006). Unlike these forerunners, the present study examines the variable emission of GHG (including CO_2) and, in addition, this variable is broken down in order to analyse the proportion of GHG emission that is caused by the activities of the land transport sector. Comments on this question are given in Section 5, in the Discussion and in the Conclusions.

2. PROBABILISTIC CHARACTERISTICS OF THE MODEL

The Feller diffusion process has constituted the basis for numerous others, both linear and nonlinear, and these have been the object of particular attention in the field of stochastic finance, for example interest rate models. Many of these models can be included within certain families of diffusions, such as that considered by Chan *et al.* (1992), which covers the diffusions corresponding to Ito's SDE, as follows:

$$\mathrm{d}X_t = (\alpha + \beta X_t)\mathrm{d}t + \sigma X_t^{\gamma}\mathrm{d}W_t \tag{1}$$

among which are diffusions such as those of Merton, Vasicek, Brenann–Schwartz, CIR-SR, CIR-VR and the CEV subfamily (see e.g. Davidov and Linetsky, 2001). The above-cited paper by Chan *et al.* (1992) considers the estimation of parameters belonging to the family Equation (1), using econometric methods based on the Generalised Method of Moments technique, and on the basis of

DOI: 10.1002/env

an appropriate example, empirically compares the particular features of each diffusion type in this family.

In particular, Cox–Ingersoll–Ross (CIR) investigated and applied the CIR-SR and CIR-VR models with SDE given, respectively, by

$$\mathrm{d}X_t = (\alpha + \beta X_t)\mathrm{d}t + \sigma X_t^{1/2}\mathrm{d}W_t \tag{2}$$

$$\mathrm{d}X_t = \sigma X_t^{3/2}\mathrm{d}W_t \tag{3}$$

Subsequently, Ahn and Gao (1998) and Ait-Sahalia (1999), among others, considered, together with other diffusions, a type of nonlinear diffusion with an SDE expressed as

$$\mathrm{d}X_t = X_t(\kappa - (\sigma^2 - \kappa\alpha)X_t)\mathrm{d}t + \sigma X_t^{3/2}\mathrm{d}W_t \tag{4}$$

which is empirically compared, using a financial example, with other diffusions, both linear and nonlinear.

This diffusion, Equation (4), has a diffusion coefficient which is similar to that of CIR-VR diffusion except that unlike the latter, which has no drift, it possesses a drift coefficient, and moreover, one that is nonlinear. With respect to CIR-SR, the above diffusion is more general, in two ways: it contains a nonlinear drift, rather than the linear one found in CIR-SR, and its coefficient of diffusion is different ($\gamma = 3/2$ and Equation (1)). This diffusion, suitably reparametrised, is the one we consider in the present paper for modelling GHG emissions. In the first place, we studied probabilistic and statistical questions concerning the diffusion process we denominate I-CIR. These issues are fundamental in the trend analysis methodology proposed in this paper for the case of GHG emissions attributable to the land transport.

2.1. *The proposed model*

Let $\{X_t; t \in [0, T]\}$ be the one-dimensional diffusion process taking values on $(0, \infty)$ and with infinitesimal moments

$$A_1(x) = ax - bx^2, \quad A_2(x) = \sigma^2 x^3 \tag{5}$$

where $\sigma > 0$, a and b are real parameters.

Alternatively, the above process can be defined by the following Itô's SDE

$$\mathrm{d}X_t = \left(aX_t - bX_t^2\right)\mathrm{d}t + \sigma X_t^{3/2}\mathrm{d}W_t, \quad X_0 = x_0 \tag{6}$$

where W_t is a standard Wiener process and x_0 is fixed in $\mathbb{R}_+^*$.

2.2. *The TPDF of the model*

The Transition Probability Density Function (TPDF) of the model can be obtained by using the Theorem 2.1 of Karlin and Taylor (see Karlin and Taylor, 1981; p. 173). In this result, we consider the function

DOI: 10.1002/env

$g(x) = 1/x$; then, the process proposed in Equation (5) can be transformed into the diffusion process $Y_t = g(X_t)$ whose infinitesimal moments are given by

$$\tilde{A}_1(y) = (\sigma^2 + b) - ay, \quad \tilde{A}_2(x) = \sigma^2 y$$

Given the form of these infinitesimal moments, the Y_t process is of the CIR type (see e.g. Chan *et al.*, 1992; Eugen and Manfredi, 1999), and if we denote the TPDF of the X_t and Y_t processes by f and f^*, respectively, then these two TPDF are related by

$$f(y, t \mid x, s) = |\varphi'(y)| \, f^*\,(\varphi(y), t \mid \varphi(x), s)$$

Using the closed form of TPDF of the process Y_t (see e.g. Going-Jaeschke, 1998) and the homogeneity of this (i.e. $f^*(y, t \mid x, s) = f^*(y, t - s \mid x, 0)$) and after some calculations (omitted here), we can deduce that the TPDF of our original process is for $\sigma^2 \geq -b$

$$\begin{aligned} f(y, t \mid x, s) = {} & \frac{2ay^{-\frac{\alpha}{2}-2}\left(xe^{a(t-s)}\right)^{\frac{\alpha}{2}}}{\sigma^2\left(1 - \mathrm{e}^{-a(t-s)}\right)} \exp\left(\frac{-2a\left(x^{-1}\mathrm{e}^{-a(t-s)} + y^{-1}\right)}{\sigma^2\left(1 - \mathrm{e}^{-a(t-s)}\right)}\right) \\ & \times \mathrm{I}_\alpha\left(\frac{4a\mathrm{e}^{-\frac{\alpha}{2}(t-s)}}{\sigma^2\left(1 - \mathrm{e}^{-a(t-s)}\right)\sqrt{xy}}\right) \end{aligned} \tag{7}$$

In this expression I_α denotes the modified Bessel function of the first kind and $\alpha = \frac{2b}{\sigma^2} + 1$.

2.3. *The conditional moments*

The k-order of conditional moment of the model is given by:

$$m_k(t \mid s) = E\left(X_t^k \mid X_s = x_s\right) = \int_0^\infty y^k f(y, t \mid x_s, s)\mathrm{d}y$$

To simplify the calculations involved in the above, the following notation is used:

$$\lambda(s, t) = \lambda = 2a/\sigma^2\left(1 - \mathrm{e}^{-a(t-s)}\right) \quad \text{and} \quad \xi(s, t) = \xi = \lambda\left(x_s\mathrm{e}^{a(t-s)}\right)^{-\frac{1}{2}}.$$

Then, in terms of λ and ξ, this conditional moment is

$$m_k(t \mid s) = \lambda^{\alpha+1}\xi^{-\alpha}\mathrm{e}^{-\frac{\xi^2}{\lambda}} \int_0^\infty y^{k-\frac{\alpha}{2}+1}\mathrm{e}^{-\lambda/y}\mathrm{I}_\alpha\left(2\xi/\sqrt{y}\right)\mathrm{d}y$$

DOI: 10.1002/env

With the change variable $z = 1/y$, we obtain

$$m_k(t \mid s) = \lambda^{\alpha+1}\xi^{-\alpha}\mathrm{e}^{-\frac{\xi^2}{\lambda}} \int_0^\infty z^{\frac{\alpha}{2}-k}\mathrm{e}^{-\lambda z}\mathrm{I}_\alpha\left(2\xi\sqrt{z}\right)\mathrm{d}z$$

then, by the relation (see Gradshteyn and Ryzhik, 1979; p. 720: 6.643): for $\mathcal{R}e(\mu + \nu + 1/2) > 0$

$$\int_0^\infty y^{\mu-1/2}\mathrm{e}^{-\beta y}\mathrm{I}_{2\nu}(2\gamma\sqrt{y})\mathrm{d}y = \frac{\Gamma\left(\mu+\nu+\frac{1}{2}\right)}{\Gamma(2\nu+1)}\gamma^{-1}\beta^{-\mu}\exp\left(\frac{\gamma^2}{2\beta}\right)\mathrm{M}_{-\mu,\nu}\left(\frac{\gamma^2}{\beta}\right)$$

where $\mathrm{M}_{\cdot,\cdot}$ is a Whittaker function; from this, we have for $\alpha > k - 1$

$$m_k(t \mid s) = \frac{\Gamma(\alpha-k+1)}{\Gamma(\alpha+1)}\lambda^{\frac{\alpha}{2}+k+\frac{1}{2}}\xi^{-\alpha-1}\mathrm{e}^{-\frac{\xi^2}{2\lambda}}\mathrm{M}_{k-\frac{\alpha}{2}-\frac{1}{2},\frac{\alpha}{2}}\left(\frac{\xi^2}{\lambda}\right)$$

by means of the relation (see Gradshteyn and Ryzhik, 1979; p. 1059: 9.220) with the confluent hypergeometric function Φ (Kummer function):

$$\mathrm{M}_{\nu,\mu}(x) = x^{\mu+1/2}\mathrm{e}^{-x/2}\Phi\left(\mu-\nu+1/2, 2\mu+1, x\right)$$

We have

$$m_k(t|s) = \frac{\Gamma(\alpha-k+1)}{\Gamma(\alpha+1)}\lambda^k\mathrm{e}^{-\frac{\xi^2}{\lambda}}\Phi\left(\alpha-k+1, \alpha+1, \frac{\xi^2}{\lambda}\right)$$

Finally, by the Kummer transformation (see Gradshteyn and Ryzhik, 1979; p. 1058: 9.212) $\Phi(\beta, \gamma, z) = \mathrm{e}^z\Phi(\gamma-\beta, \gamma, -z)$, we deduce that, for $\alpha > k - 1$

$$m_k(t \mid s) = \frac{\Gamma(\alpha-k+1)}{\Gamma(\alpha+1)}\lambda^k\Phi\left(k, \alpha+1, -\frac{\xi^2}{\lambda}\right)$$

Then, by substituting λ and ξ, the final form of the conditional moment of the kth order of the model takes the following form:

$$E\left(X_t^k \mid X_s = x_s\right) = \frac{\Gamma(\alpha-k+1)}{\Gamma(\alpha+1)}\left(\frac{2a}{\sigma^2\left(1-\mathrm{e}^{-a(t-s)}\right)}\right)^k \times \Phi\left(k, \alpha+1, \frac{-2a\mathrm{e}^{-a(t-s)}/x_s}{\sigma^2\left(1-\mathrm{e}^{-a(t-s)}\right)}\right)$$

for $\alpha > k - 1$.

DOI: 10.1002/env

From this expression, we deduce the conditional trend function of the process the expression of which is for $\alpha > 0$

$$E(X_t \mid X_s = x_s) = \frac{2a/\alpha}{\sigma^2(1 - \mathrm{e}^{-a(t-s)})} \Phi\left(1, \alpha + 1, \frac{-2a\mathrm{e}^{-a(t-s)}/x_s}{\sigma^2(1 - \mathrm{e}^{-a(t-s)})}\right) \tag{8}$$

From Equation (8) and under the initial distribution $P(X_0 = x_0) = 1$, the trend function of the process takes the following form (for $\alpha > 0$)

$$E(X_t) = \frac{2a/\alpha}{\sigma^2(1 - \mathrm{e}^{-at})} \Phi\left(1, \alpha + 1, \frac{-2a\mathrm{e}^{-at}/x_0}{\sigma^2(1 - \mathrm{e}^{-at})}\right) \tag{9}$$

Equations (8) and (9) play a fundamental role in constituting the estimated trend functions (see following section), which enable us to fit and to predict the future evolution of the stochastic variable under consideration.

2.4. *Stationary distribution of the process*

Let us now analyse the existence of the stationary distribution and obtain the explicit expression for its density. In general (see Nobile and Ricciardi, 1984; Ricciardi, 1977) the density function of stationary distribution, $S(x)$, in a diffusion can be expressed, under specific conditions that satisfy the inverse CIR, as

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

where z is an arbitrary point in the interval $]0, +\infty[$, and c is a constant to be determined by the following normalisation 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 the above results, we can deduce that for $a > 0$ and $\alpha > -1$ (i.e. $a > 0$ and $b > -\sigma^2$), the density function of the stationary distribution of the process exists, and takes the form

$$S(x) = \frac{\left(\frac{2a}{\sigma^2}\right)^{\alpha+1} x^{-\alpha-2} \mathrm{e}^{-2a/\sigma^2 x}}{\Gamma(\alpha + 1)} \tag{10}$$

S is the density function of inverse Gamma distribution.

From the above expression we can calculate the asymptotic moment of order k, and then we have for $a > 0$ and $\alpha > k - 1$

$$E[X^k(\infty)] = \int_0^\infty x^k S(x)\mathrm{d}x = \frac{\Gamma(\alpha - k + 1)}{\Gamma(\alpha + 1)} \left(\frac{2a}{\sigma^2}\right)^k$$

The asymptotic trend function of the process is, for $a > 0$ and $\alpha > 0$

$$E[X(\infty)] = \frac{2a}{\alpha\sigma^2} = \frac{a}{b + \sigma^2/2} \tag{11}$$

It can be seen that the limit of the trend function in Equation (9) (when t tends to ∞) coincides with the asymptotic trend function in Equation (11).

The asymptotic variance function of the process is, for $a > 0$ and $\alpha > 1$

$$\mathrm{Var}[X(\infty)] = \frac{4a^2}{\sigma^4\alpha^2(\alpha - 1)}$$

3. STATISTICAL METHODOLOGY

For the drift parameters, since the PDFT of the process is known (the inverse of non-central chi-square distribution with non-integer parameters), likelihood inference with discrete sampling is, in principle, possible, but it is extremely complicated to calculate. As an alternative, one might apply maximum likelihood using continuous sampling, after transforming the stochastic integrals that relate the likelihood estimators to Riemann integrals, the latter being approximated by the trapezoid method. On the other hand, the parameter in coefficient diffusion can be approximated using an extension of the procedure described by Chesney and Elliot (1995); this extension has been considered by Skiadas and Giovani (1997), Gutiérrez *et al.* (2006a) and Gutiérrez *et al.* (2007a).

3.1. Likelihood estimation of drift parameters

Let us consider the one-dimensional diffusion process defined by the following SDE

$$\mathrm{d}X_t = \mathbf{A}(X_t)\cdot \boldsymbol{\theta}\, \mathrm{d}t + B(X_t)\,\mathrm{d}W_t; \quad 0 \le t \le T \tag{12}$$

where the parameter $\theta \in R^k$, $\mathbf{A}$ is a k-dimensional vector and B is R-valued depending only on the sample path up to the given instant. We assume that Equation (12) has a unique solution for every θ. The maximum likelihood estimator of the vector $\boldsymbol{\theta}$ is given by (see, e.g. Kloeden *et al.*, 1992; Prakasa Rao, 1999)

$$\hat{\theta}_T = \mathbf{S}_T{}^{-1}\mathbf{H}_T \tag{13}$$

DOI: 10.1002/env

where $\mathbf{H}_T$ is the following k-dimensional vector:

$$\mathbf{H}_T = \int_0^T \mathbf{A}_t^*(X_t)\big(B_t^*(X_t)B_t(X_t)\big)^{-1}\mathrm{d}X_t \tag{14}$$

$\mathbf{S}_T$ is the $k \times k$ matrix:

$$\mathbf{S}_T = \int_0^T \mathbf{A}_t^*(X_t)\big(B_t^*(X_t)B_t(X_t)\big)^{-1}\mathbf{A}_t(X_t)\mathrm{d}t \tag{15}$$

and the asterisk denotes the transpose.

The SDE of the proposed model Equation (6) can be written in the vector form [Equation (12)], with:

$$\mathbf{A}(X_t) = \left(X_t, -X_t^2\right);\ \theta^* = (a, b) \text{ and } B(X_t) = \sigma X_t^{3/2}$$

The corresponding vector $\mathbf{H}_T$ in Equation (14) in this case leads us to

$$\mathbf{H}_T^* = \frac{1}{\sigma^2}\left(\int_0^T \frac{\mathrm{d}X_t}{X_t^2}, -\int_0^T \frac{\mathrm{d}X_t}{X_t}\right)$$

$\mathbf{S}_T$ is the following square matrix

$$\mathbf{S}_T = \frac{1}{\sigma^2}\begin{pmatrix} \int_0^T \frac{\mathrm{d}t}{X_t} & -T \\ -T & \int_0^T X_t\mathrm{d}t \end{pmatrix} \tag{16}$$

Using Equation (13) and after some calculation (not shown), we obtain the expressions of the estimators

$$\hat{a} = \frac{\left(\int_0^T X_t\mathrm{d}t\right)\left(\int_0^T \frac{\mathrm{d}X_t}{X_t^2}\right) - T\int_0^T \frac{\mathrm{d}X_t}{X_t}}{\left(\int_0^T \frac{\mathrm{d}t}{X_t}\right)\left(\int_0^T X_t\mathrm{d}t\right) - T^2}$$

$$\hat{b} = \frac{T\left(\int_0^T \frac{\mathrm{d}X_t}{X_t^2}\right) - \left(\int_0^T \frac{\mathrm{d}t}{X_t}\right)\left(\int_0^T \frac{\mathrm{d}X_t}{X_t}\right)}{\left(\int_0^T \frac{\mathrm{d}t}{X_t}\right)\left(\int_0^T X_t\mathrm{d}t\right) - T^2}$$

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

$$\int_0^T \frac{\mathrm{d}X_t}{X_t} = \log(X_T) - \log(x_0) + \frac{\sigma^2}{2}\int_0^T X_t \mathrm{d}t$$

$$\int_0^T \frac{\mathrm{d}X_t}{X_t^2} = \frac{1}{x_0} - \frac{1}{X_T} + \sigma^2 T$$

Therefore, the resulting maximum likelihood estimators are

$$\hat{a} = \frac{\int_0^T X_t \mathrm{d}t \left(\frac{X_T - X_0}{x_0 X_T} + T\sigma^2\right) - T\left(\log(X_T/x_0) + \frac{\sigma^2}{2}\int_0^T X_t \mathrm{d}t\right)}{\left(\int_0^T \frac{\mathrm{d}t}{X_t}\right)\left(\int_0^T X_t \mathrm{d}t\right) - T^2} \tag{17}$$

$$\hat{b} = \frac{T\left(\frac{X_T - X_0}{X_T X_0} + T\sigma^2\right) - \int_0^T \frac{\mathrm{d}t}{X_t}\left(\log(X_T/x_0) + \frac{\sigma^2}{2}\int_0^T X_t \mathrm{d}t\right)}{\left(\int_0^T \frac{\mathrm{d}t}{X_t}\right)\left(\int_0^T X_t \mathrm{d}t\right) - T^2} \tag{18}$$

3.2. Estimation of the noise coefficient

As mentioned above, the coefficient of diffusion is estimated using an approximative method similar to that described by Chesney and Elliot (1995) and by Skiadas and Giovani (1997). This method can be summarised in the following steps:

By applying the Itô formula, we have

$$\mathrm{d}\left(\frac{1}{X_t}\right) = -\frac{\mathrm{d}X_t}{X_t^2} + \sigma^2 \mathrm{d}t \tag{19}$$

The differentials shown in Equation (19) can be approximated by consecutive observations of a sample path of the process in $t-1$ and t, as follows:

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

DOI: 10.1002/env

By inserting these approximations in Equation (19), an approximate estimator of the σ parameter between the latter observations is found to be

$$\hat{\sigma}_{(t-1,t)} = \frac{|X_t - X_{t-1}|}{X_t\sqrt{X_{t-1}}}$$

For n observations of a sample path of the process, the resulting approximate estimator has the following expression:

$$\hat{\sigma} = \frac{1}{n-1}\sum_{t=1}^{n}\frac{|X_t - X_{t-1}|}{X_t\sqrt{X_{t-1}}} \tag{20}$$

Remarks.

- In order to use the above expressions, Equations (17) and (18), to estimate the parameters, we must have continuous observations. In practice, continuous sample path processes cannot be monitored in continuous time, but only at sequences of discrete instants ($0 = t_0 < t_1 < \cdots < t_n = T$). By the Markov properties, the likelihood function corresponding to such data is the product of transition densities (in the present case, it has a complicated form, see Equation (7)) and it is very difficult to find the estimators explicitly. An alternative estimation procedure that is frequently utilised (see e.g. Giovanis and Skiadas, 1999; Gutiérrez *et al.*, 2007a) for such data is to use the continuous time maximum likelihood estimators with suitable approximations of the integrals that appear in the Equations (17) and (18); specifically, the Riemann–Stieltjes integrals are approximated by means of the trapezoidal formula.
- An approximation of the standard error of the estimator of $\hat{\sigma}$ is given by

$$\mathrm{es}(\hat{\sigma}) = \frac{1}{n-1}\sum_{t=1}^{n}\left(\hat{\sigma}_{(t-1,t)} - \hat{\sigma}\right)^2 \tag{21}$$

- By using Zehna's theorem, the estimated trend function (ETF) and estimated conditional trend function (ECTF) of the process are obtained by replacing the parameters in Equations (8) and (9) by their estimators given in Equations (17), (18) and (20). Then the ETF and ECTF are given by the following expression:

$$\hat{E}(X_t) = \frac{2\hat{a}/\hat{\alpha}}{\hat{\sigma}^2(1-\mathrm{e}^{-\hat{a}t})}\Phi\left(1, \hat{\alpha}+1, \frac{-2\hat{a}\mathrm{e}^{-\hat{a}t}/x_0}{\hat{\sigma}^2(1-\mathrm{e}^{-\hat{a}t})}\right) \tag{22}$$

$$\hat{E}(X_t/X_s = x_s) = \frac{2\hat{a}/\hat{\alpha}}{\hat{\sigma}^2(1-\mathrm{e}^{-\hat{a}(t-s)})}\Phi\left(1, \hat{\alpha}+1, \frac{-2\hat{a}\mathrm{e}^{-\hat{a}(t-s)}/x_s}{\hat{\sigma}^2(1-\mathrm{e}^{-\hat{a}(t-s)})}\right) \tag{23}$$

  where $\hat{\alpha} = 2\hat{b}/\hat{\sigma}^2 + 1$.

3.3. Asymptotic normality of likelihood estimators

For $a > 0$ and $\alpha > 0$, we confirm the conditions of ergodicity (see Kutoyants Yu, 2004) and that the process has ergodic properties. If we denote by X the random variable with density function S in

DOI: 10.1002/env

Equation (10), then $1/X$ has a Gamma($\alpha+1, \sigma^2/2a$) distribution and we have, for a known σ and for $\theta \in (a_1, a_2) \times (b_1, b_2)$, with $a_1 > 0$ and $b_1 > -\sigma^2/2$,

$$\mathcal{L}_\theta\left(\sqrt{T}(\hat{\theta}-\theta)\right) \to \mathcal{N}_2\left(0, \mathbb{I}^{-1}(\theta)\right); \quad \text{when} \quad T \to \infty \tag{24}$$

where

$$\mathbb{I}(\theta) = \mathbb{E}_\theta\left(\frac{\dot{A}_1(X)\dot{A}_1^*(X)}{A_2(X)}\right) \quad \text{and} \quad \dot{A}_1(x) = \left(\frac{\partial A_1(x,\theta)}{\partial a}; \frac{\partial A_1(x,\theta)}{\partial b}\right)^*$$

Then, by calculation, we obtain

$$\mathbb{I}(\theta) = \frac{1}{\sigma^2}\mathbb{E}_\theta\begin{pmatrix} \frac{1}{X} & -1 \\ -1 & X \end{pmatrix} = \begin{pmatrix} \frac{\alpha+1}{2a} & \frac{-1}{\sigma^2} \\ \frac{-1}{\sigma^2} & \frac{2a}{\alpha\sigma^4} \end{pmatrix} \tag{25}$$

and their inverse is

$$\mathbb{I}^{-1}(\theta) = \begin{pmatrix} 2a & \alpha\sigma^2 \\ \alpha\sigma^2 & \frac{\alpha(\alpha+1)\sigma^4}{2a} \end{pmatrix} \tag{26}$$

An approximated and asymptotic confidence region of θ and an approximated and asymptotic marginal confidence intervals of $\hat{a}$ and $\hat{b}$ can be obtained by substitution of Equation (25) in (24). The above mentioned region is given, for a large T, by

$$P\left[T\left(\theta-\hat{\theta}\right)^* \hat{\mathbb{I}}(\theta)\left(\theta-\hat{\theta}\right) \le \chi^2_{2,\gamma}\right] = 1-\gamma \tag{27}$$

where $\hat{\mathbb{I}}(\theta)$ is obtained by replacing the parameters by their estimators in the Equation (25) and $\chi^2_{2,\gamma}$ is the upper 100γ per cent points of the chi-squared distribution with two degrees of freedom.

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

$$P\left(a \in \left[\hat{a} \pm \lambda_\gamma\sqrt{2\hat{a}/T}\right]\right) = 1-\gamma \tag{28}$$

$$P\left(b \in \left[\hat{b} \pm \lambda_\gamma\sigma^2\left(\frac{\hat{\alpha}(\hat{\alpha}+1)}{2T\hat{a}}\right)^{1/2}\right]\right) = 1-\gamma \tag{29}$$

where λ_γ is the 100γ per cent points of the normal standard distribution.

In Equations (27)–(29) we have supposed that σ is known with value $\sigma = \hat{\sigma}$.

By using those results, we could also obtain an approximated and asymptotic confidence interval of the ETF and ECTF.

DOI: 10.1002/env

4. APPLICATION AND SIMULATION

4.1. Application: land transport as a source of emission of greenhouse gases in Spain

In recent years, policies to reduce GHG (particulary, CO_2) emissions from cars have been contradictory: on the one hand, significant efforts have been made in technological innovation in the sector, with more efficient engines producing lower levels of CO_2, in relation to the power provided (by means of diesel and petrol engines with direct injection, and by the better use of compressors and turbo compressors, among other advances).

However, there has been a growth in the market for high-powered vehicles, which are very large and heavy, consume more fuel and in some cases produce over 250 g of CO_2 per km. In 1998, the European Automobile Manufacturers Association (ACEA) committed its members to reducing car emissions of CO_2 to 140 g/km by 2008, a target that is in fact impossible for most to meet.

The situation varies widely throughout Europe, with Germany being the most affected country. Manufacturers, for their part, demand that a more global policy be established, involving other factors that influence fuel consumption, such as improvements in road infrastructures and changes in drivers' behaviour.

At present, the maximum permitted emission rate for cars is 160 g of CO_2 per km travelled, and the European Commissioner for the Environment has proposed that this limit should be reduced to 120 g/km for all new cars from the year 2012.

All the above should be considered in the context of the relative failure of the Kyoto Protocol; in general, the targets established have not been met, with this non-compliance being particularly severe in the case of Spain (see Gutiérrez *et al.*, 2007a, 2007b).

We shall now summarise the situation in the land transport sector in Spain, for the period 1996–2005, in accordance with the following criteria:

1. Type of vehicle (cars; lorries, vans and buses; industrial tractors; other types).
2. Type of fuel (petrol; diesel) used.
3. Age of vehicle.

The data are classified using the above factors (type of vehicle; type of fuel; age of vehicle) because these are crucial in determining levels of emission of CO_2 and of other GHG. The data thus classified are of interest with respect to the application of certain, more restrictive, policies for the control and reduction of vehicle emissions.

Table 3 describes the evolution of the total stock of vehicles in Spain; in 2005, there were a total of 27 657 276 vehicles, which can be classified by vehicle and fuel type. This figure is 41.5% higher than the corresponding one for 1996. In 2005, the total stock of petrol-driven cars was 0.24% higher than in 1996, with a relatively stable trend being observed for the period 1996–2005. In contrast, the stock of diesel-driven cars in 2005 was 168.5% higher than in 1996. For cars, these increases were—4.43 and 252.7%, respectively. These cars made up 76.5% of the total vehicle stock in 1996, and 73.2% in 2005, with a relatively stable evolution over the period in question.

Table 4 shows the age of the stock of vehicles, by type and by fuel, taking as a baseline the total stock recorded in 2005. Thus we see that 38.4% of the total stock of vehicles is aged over 10 years. In the case of cars, the corresponding figure is 36.7%. The cars that were 10 or more years old in 2005 constituted 70% of the total stock of vehicles of this age.

Table 3. Vehicles by category and fuel consumption in Spain

	Cars		Lorries, vans and buses		Industrial tractors	
	Petrol	Diesel	Petrol	Diesel	Petrol	Diesel
1996	12 362 457	2 391 352	906 848	2 198 904	1694	92 863
1997	12 490 612	2 806 754	891 999	2 364 010	1845	102 276
1998	12 681 210	3 368 847	879 072	2 566 179	2050	114 255
1999	12 802 978	4 044 419	859 384	2 799 128	2233	127 983
2000	12 746 971	4 702 264	832 348	3 002 605	2367	140 588
2001	12 795 735	5 355 145	813 409	3 191 738	2453	153 504
2002	12 728 713	6 003 919	790 179	3 358 649	2486	164 528
2003	12 095 876	6 592 444	750 494	3 494 409	2410	172 097
2004	12 035 097	7 506 821	738 030	3 736 966	2426	182 953
2005	11 815 652	8 434 725	716 421	3 997 240	2396	191 810

	Other types		Total		Total
	Petrol	Diesel	Petrol	Diesel	
1996	1 478 386	109 600	14 749 385	4 792 719	19 542 104
1997	1 511 961	116 951	14 896 417	5 389 991	20 286 408
1998	1 566 510	128 370	15 128 842	6 177 651	21 306 493
1999	1 631 559	143 510	15 296 154	7 115 040	22 411 194
2000	1 699 018	158 054	15 280 704	8 003 511	23 284 215
2001	1 763 857	174 030	15 375 454	8 874 417	24 249 871
2002	1 825 903	191 355	15 347 281	9 718 451	25 065 732
2003	1 854 709	207 013	14 703 489	10 465 963	25 169 452
2004	2 003 188	227 160	14 778 741	11 653 900	26 432 641
2005	2 250 348	248 684	14 784 817	12 872 459	27 657 276

Table 4. Vehicles by category, fuel consumption and age. Spain, 2005

	Cars		Lorries, vans and buses		Industrial tractors	
	Petrol	Diesel	Petrol	Diesel	Petrol	Diesel
0–5	2 781 162	4 433 287	85 782	1 491 588	768	95 154
6–10	2 838 566	2 756 807	94 536	1 141 784	849	57 683
11–15	2 470 138	761 269	171 283	610 557	276	15 764
>15	3 725 786	483 362	364 820	753 311	503	23 209
Total	11 815 652	8 434 725	716 421	3 997 240	2396	191 810

	Other types		Total		Total
	Petrol	Diesel	Petrol	Diesel	
0–5	747 825	106 412	3 615 537	6 126 441	9 741 978
6–10	345 522	58 546	3 279 473	4 014 820	7 294 293
11–15	330 232	29 110	2 971 929	1 416 700	4 388 629
>15	826 769	54 616	4 917 878	1 314 498	232 376
Total	2 250 348	248 684	14 784 817	12 872 459	27 657 276

DOI: 10.1002/env

Table 5. Car engine size, by year of first registration

	<1199	1200–1599	1600–1999	>2000	Total
1996	76 399	282 688	363 152	77 844	800 083
1997	89 430	323 409	437 187	97 120	947 146
1998	88 227	370 593	573 490	115 259	1 147 569
1999	103 021	413 859	737 776	111 247	1 365 903
2000	79 170	399 552	739 377	116 573	1 334 672
2001	87 114	423 999	724 897	138 503	1 374 513
2002	80 533	394 871	664 206	137 835	1 277 445
2003	81 888	447 566	692 129	157 667	1 379 250
2004	81 764	547 021	752 571	173 333	1 554 689
2005	69 827	669 972	713 183	175 570	1 628 552
Total	837 373	4 273 530	6 397 968	1 300 951	12 809 822

Table 5 describes the variation of the stock of vehicles, by engine size, between 1996 and 2005. In 2005, 10.8% of all new cars registered had an engine size of over 2000 cm^3, while in 1996 the corresponding figure was 9.72%. In 2005, 43.8% of the new cars registered had an engine size of 1600–1999 cm^3, while in 1996 the corresponding figure was 45.4%; thus, there has been a change in motorists' preferences, with larger, more powerful cars being bought than 10 years ago. Of all cars registered during the period 1996–2005, 9.84% were 'up-market' models (engine size over 2000 cm^3).

In summary, the stock of vehicles in Spain is as follows:

(i) The tendency is for high, sustained rates of growth.
(ii) A large proportion of the vehicles are significantly old, which contributes to a high degree of global emission of GHG, and in particular of CO_2.
(iii) There is a growing trend towards more powerful cars, with the corresponding increase in fuel consumption.
(iv) There has been a marked growth in the number of cars using diesel fuel, versus petrol-driven vehicles, especially as concerns cars; this also contributes to the increased emission of GHG.

In summary, the stock of vehicles in Spain constitutes an important source of the emission of GHG and the trend is for this emission to increase.

4.2. Application to analysis of the trend of emissions of GHG attributable to transport activities in Spain

In this application, we consider 'Emissions from the combustion and evaporation of fuel for all transport activity, regardless of the sector', with the qualifications and sources mentioned in Subsection 1.2. We examine the variable X_t that at each instant of time t corresponds to the total annual amount of GHG emitted during the year ending at t. For instants coinciding with the end of each natural year, we possess data provided by the above-mentioned sources; these, therefore, constitute discrete time sampling at equally spaced intervals.

In the present study, we shall analyse the results obtained from modelling the stochastic process X_t by the SDE (Equation (6)) corresponding to I-CIR diffusion and by applying the statistical methodology described in Section 3. Specifically, we fit the trend functions (ETF and ECTF) given by Equations (22) and (23), respectively. Prior to this, we calculate the estimators of the drift coefficient parameters for

DOI: 10.1002/env

Table 6. Fit (1990–2003) and forecasts (2004)

Years	Real Data	ETF	ECTF
1990	57.536	57.536	57.536
1991	59.849	60.242	60.242
1992	63.691	63.032	62.628
1993	63.011	65.905	66.586
1994	65.986	68.859	65.886
1995	67.028	71.890	68.947
1996	71.735	74.996	70.017
1997	72.498	78.175	74.848
1998	79.485	81.421	75.630
1999	84.274	84.731	82.777
2000	87.003	88.101	87.662
2001	91.278	91.525	90.441
2002	93.463	94.997	94.786
2003	98.045	98.512	97.003
2004	102.011	102.064	101.646

Data in Tg CO_2 eq.

the diffusion, given by Equations (17) and (18), and the estimator of the σ parameter (volatility) of the diffusion coefficient, given by Equation (20). To perform these calculations, programs were developed in Mathematica 5.1, including the numerical calculus subroutines necessary to approximate the integers in Equations (17) and (18), and to calculate the confluent hypergeometric functions involved in the utilisation of I-CIR diffusion.

Different fits were achieved, corresponding to the following situations: (i) As baseline values for the fit, the data observed for the entire period considered, 1990–2004; (ii) Data for the subperiod 1990–2003 were used for the fit, and the forecast for 2004 was obtained by means of the ETF and ECTF and then compared with the real data for this year; (iii) The same process was performed for the subperiod 1990–2002 and with forecasts for 2003 and 2004; (iv) Finally, the entire period 1990–2004 was considered and medium-term forecasts (2005–2008) made.

In every case, the values fitted in accordance with the ETF and ECTF of the I-CIR model present a high degree of accuracy with respect to the observed values. Below, we describe just the results corresponding to the above situations (ii) and (iv).

Table 6 shows the values observed for the period 1990–2004, together with the fitted values, taking the subperiod 1990–2003 as the basis for fitting, by application of the ETF and the ECTF. The datum corresponding to 2004 is the forecast made by applying these trend functions. Figure 1 shows the real data observed and the respective values fitted using the ECTF. The estimated values for the three parameters of the I-CIR Model are, in this case, as follows: $a = 0.060176$, $b = 0.000241$, $\sigma = 0.005292$.

Table 7 shows the values that correspond to the case in which the entire period considered, 1990–2004, is taken as the basis for fitting the data, and in which forecasts are made for 2005–2008. We show the values fitted both by the ETF and by the ECTF. In this case, the estimations of the I-CIR parameters are: $a = 0.058573$, $b = 0.000216$, $\sigma = 0.005187$.

Finally, Figure 2 shows the values fitted by the ECTF, together with the forecasts made; in this case, the latter are not strictly comparable with the respective real values observed, as they are, in fact, unknown.

DOI: 10.1002/env

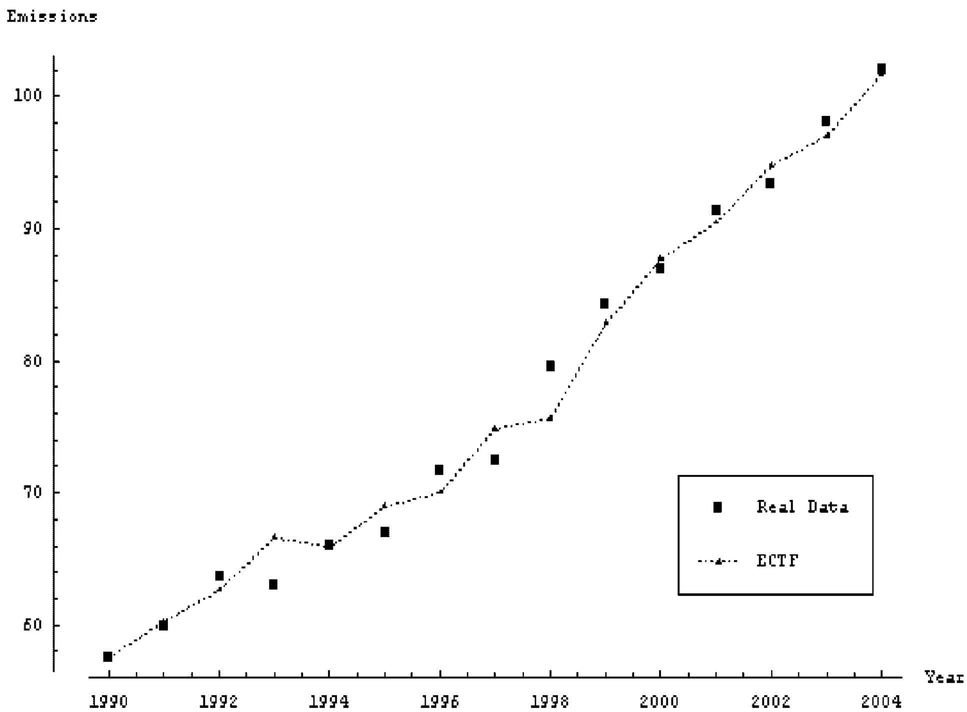

Figure 1. Estimated conditioned trend function (ECTF) versus data

Table 7. Fit (1990–2004) and forecasts (2005–2008)

Years	Real Data	ETF	ECTF
1990	57.536	57.536	57.536
1991	59.849	60.234	60.234
1992	63.691	63.020	62.624
1993	63.011	65.894	66.5879
1994	65.986	68.853	65.886
1995	67.028	71.896	68.952
1996	71.735	75.021	70.025
1997	72.498	78.226	74.865
1998	79.485	81.507	75.648
1999	84.274	84.862	82.812
2000	87.003	88.286	87.710
2001	91.278	91.775	90.496
2002	93.463	95.325	94.854
2003	98.045	98.930	97.078
2004	102.011	102.585	101.735
2005		106.284	100.753
2006		110.022	105.479
2007		113.791	104.484
2008		117.586	109.272

Data in Tg CO_2 eq.

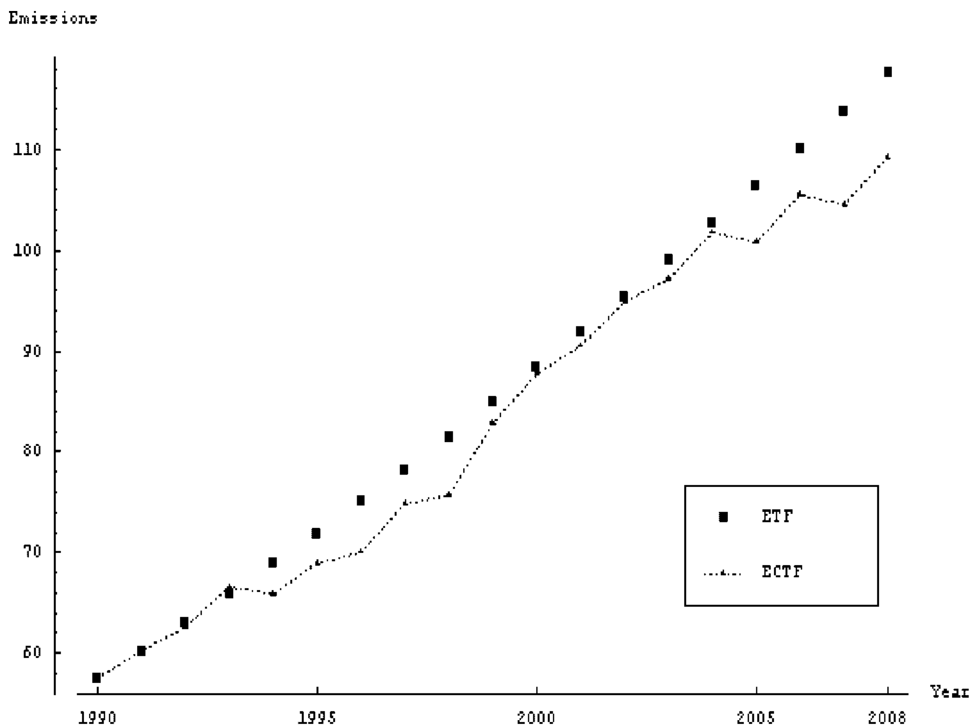

Figure 2. Estimated trend function (ETF) and ECTF

In every case, we may apply the approximate, asymptotic confidence intervals that can be calculated from Equations (28) and (29).

4.3. Simulation

For the simulation of the sample paths, we have used the procedure proposed by Kloeden and Platen (1992). Derivation of this algorithm involves the approximate discretisation of the Itô integral equation in time intervals of length h. The algorithm is given by

$$\begin{aligned} x_{n+1} = {} & x_n \left\{1 + \left[a - \left(b + \frac{3\sigma^2}{4}\right)\right] h\right\} + x_n \left\{a^2 - 3abx_n + 2b\left(b - \frac{\sigma^2}{2}\right) x_n^2\right\} \frac{h^2}{2} \\ & - \frac{\sigma}{2} x_n^{3/2} \left\{a + \left(b + \frac{3\sigma^2}{4}\right) x_n\right\} \Delta Z + \frac{3}{2}\sigma x_n^{3/2} \left\{a - \left(b + \frac{3\sigma^2}{4}\right) x_n\right\} h.\Delta W \\ & + \sigma x_n^{3/2} \Delta W \left\{1 + \frac{3}{4}\sigma x_n^{1/2} \Delta W + \frac{\sigma^2}{2} x_n (\Delta W)^2\right\} \end{aligned}$$

where $\Delta W = \sqrt{h} U_1$ and $\Delta Z = \frac{h^{3/2}}{2}(U_1 + U_2/\sqrt{3})$, with U_1 and U_2 being two standard normal independent random variables, and h is the discretisation step.

DOI: 10.1002/env

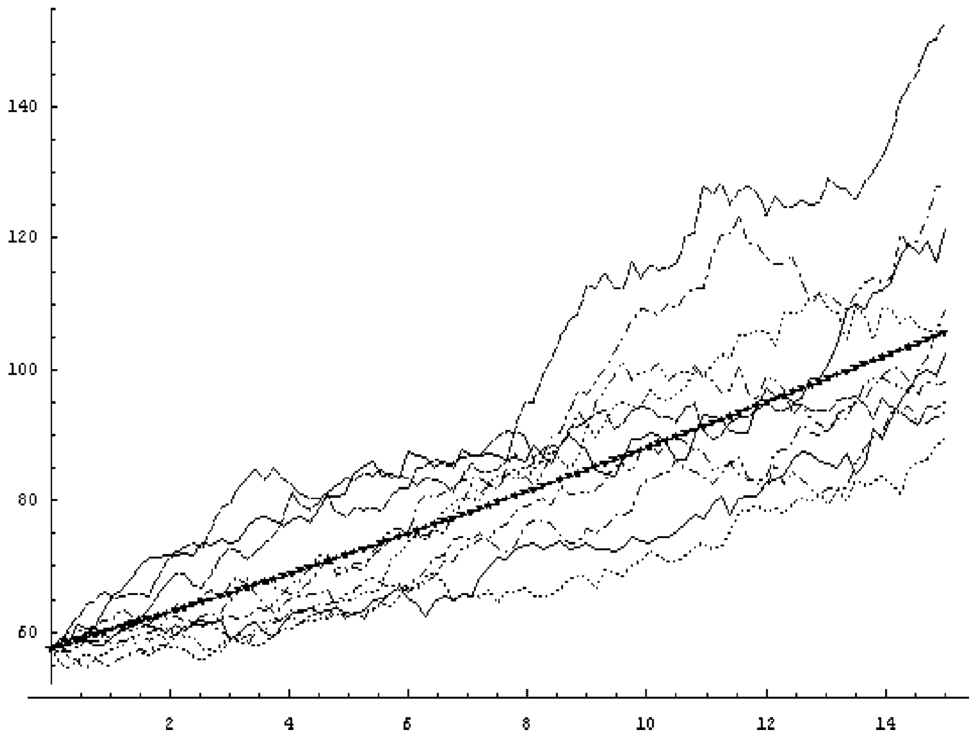

Figure 3. Simulated sample paths

Figure 3, shows the simulation of 10 sample paths of the I-CIR with parameters taken from the neighbourhood of the estimators obtained for real case studied, that is $a = 0.0601762$, $b = 0.0002411$, $\sigma = 0.0052926$ and with an initial value $x_0 = 57.536$ and using a discretisation step of $h = 0.1$. This simulation was calculated according to the above mentioned algorithm.

Also, Figure 3 shows the estimated conditioned trend function (ECTF) fitted to the considered data for GHG emission in Spain.

5. CONCLUSIONS AND DISCUSSION

One of the conclusions reached in the present study is that, whilst the Gompertz diffusion model performs optimally when total CO_2 emissions in all sectors are considered (see Gutiérrez *et al.*, 2007a), the inverse-CIR diffusion provides a better statistical fitting for the case of emissions of greenhouse gasses attributable to the land transport sector. As the different sources of CO_2 emissions are studied individually, we find that some of them, and specifically that of land transport, perform better with respect to a diffusion type such as inverse CIR than a diffusion of the same type as that of the global emissions. This is somewhat surprising, because the different sources making up global emissions have not behaved in the same way over recent decades, that is the CO_2 corresponding to land transport, for example does not present the same diffusion as that emitted by electricity generating plants.

DOI: 10.1002/env

The principal technical conclusion is that I-CIR diffusion constitutes a suitable model for the description and analysis of the evolution of the variable considered, in the case of Spain. This means that the diffusion coefficient fitted to the data is of the type $\sigma^2 X_t^{2\gamma}$; $\gamma > 1$, which according to Chan *et al.* (1992) enables us to better express the fact that the volatility of the process is sensitive to the values of the process itself. Moreover, the fitted trend corresponds to a quadratic-type drift coefficient in the variable. Note that in the study by Gutiérrez *et al.* (2007a), limited to the emission of CO_2 in Spain, attributable to various sources including transport, the most appropriate model was found to be that of cubic drift diffusion.

According to the I-CIR model, adjusted on the basis of real observations for the period 1990–2004, the emission of GHG in Spain attributable to the transport sector will increase during the period 2005–2008 by 7.41% with respect to 2004, and by 2008 will have increased by 89.9% with respect to 1990.

This emission is, thus, a fundamental part of the 52% by which total emissions of GHG in Spain are expected to increase for all activity sectors with respect to 1990.

By applying the results obtained in Subsection 3.3, the approximate, asymptotic interval for the 2008 forecast, using the ECTF given by Equation (23) is found to be [99.05; 117.56] (on the 1990 basal value) at 95% confidence.

The causes of this large increase are assumed to be those described in Subsection 4.1 of this paper, which details the profile of the stock of vehicles in Spain. Policies to control the variables that constitute this profile are essential if we are to reduce, or at least stabilise the sharp increase in the emissions in question. In this respect, we should consider with interest the proposal by the EU Council (20 February 2007, Brussels) to carry out studies to evaluate the potential impact on the transport sector and on the economy and industry, in general, of a reduction to a maximum of 130 g/km travelled.

Thus, future research could be done into the influence of the above factors on the pattern of vehicle emissions. For this purpose, the inverse-CIR model proposed in the present paper would have to be extended to non-homogeneous versions, as has been done for other diffusion types that have been studied and applied by the authors with respect to the energy sector, for example Gompertz diffusion (see Gutiérrez *et al.*, 2005a).

Like any type of activity that is analysed with respect to GHG emissions (and particularly, those of CO_2), the land transport sector (in any area of economic activity) presents special difficulties due to the fact that the level of emissions depends on a broad group of 'regressor' variables (factors that are exogenous to the variable, X_t, that is being modelled). From Subsection 4.1, we deduce some variables that are assumed to influence the evolution of emission levels. But the latter may also be affected by social variables or by regional ones, if the problem is being analysed within a local context. The modelling techniques proposed in this paper enable us to address the evolution of the endogenous variable, X_t, in itself, within a national or transnational macro-context, in which broad political decisions may be taken. Statistical means can be employed to perform analyses for more restricted contexts (micro-contexts), with particular socio-economic situations (demographic variables, unemployment rates, etc.). In such cases, see for example Paravantis and Georgakellos (2007), econometric models could be applied, with forecasts based on techniques of the Box-Jenkins type; these describe the evolution of variables like CO_2 emissions or the energy consumption of certain types of transport (like buses or private cars) by means of known, observed 'regressor' variables.

We believe the methodology proposed for fitting the diffusion process (continuous EMV for the drift parameters and by means of Chesney-Elliot for the diffusion coefficient, and especially as regards volatility) is more suitable for the situations that are discussed than that based on fitting by econometric approximation (General Linear Model), in view of the fitting results obtained. Nevertheless, further

DOI: 10.1002/env

empirical studies should be carried out, examining different subject areas and with real, observed data, fitted using the two methodologies.

ACKNOWLEDGEMENTS

The authors are very grateful to the editor and referees for constructive comments and suggestions. The authors are grateful to Dr Jose M. Angulo (University of Granada) for his careful reading and excellent comments. This work was supported partially by research project MTM 2005-09209, Ministerio de Educación y Ciencia, and project P06-FQM-02271, Junta de Andalucía, Spain.

REFERENCES

Ahn DH, Gao B. 1998. A parametric nonlinear model of term structure dynamics, Working paper, University of North CArolina and Chapel Hill.

Ait-Sahalia Y. 1999. Transition densities for interest rate and other nonlinear diffusions *The Journal of Finance* **54**: 1361–1396.

Chan KC, Karolyi GA, Longstaff FA, Sanders AB. 1992. An ampirical comparison of alternative models of the short-term interest rate *The Journal of Finance* **3**: 1209–1227.

Chesney M, Elliot J. 1995. Estimating the instantaneous volatility and covariance of risky assets. *Applied Stochastic Models and Data Analysis* **11**: 51–58.

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

Den Elzen M, Meinshausen M, van Vuuren D. 2007. Multigas emissions envelopes to meet greenhouse gas concentration targets: costs versus certainly of limiting temperature increase. *Global Environmental Change* 2007; **17**: 260–280. DOI: 10.1016/j.gloenvcha.2006.10.003

Eugene MC, Manfredi P. 1999. One dimensional SDE models, low order numerical methods and simulation based estimation: a comparison of alternative estimators. *Computational Economics* **13**: 177–197.

Fine TL. 1973. *Theories of Probability: An Examination of Foundations*. Academic Press: New York.

Giovanis AN, Skiadas CH. 1999. A stochastic logistic innovation diffusion-model studying the electricity consumption in Greece and the United States. *Technological Forecasting and Social Change* **61**: 253–264.

Going-Jaeschke A. 1998. Parameter estimation and Bessel processes in financial models and numerical analysis in Hamiltonian dynamics, Unpublished mathematics PhD dissertation, Diss. ETM N.12566, Swiss Federal Institute of Technology, Zurich.

Gradshteyn L, Ryzhik I. 1979. *Table of Integrals, Series and Products*. Academic Press: San Diego, CA.

Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A. 2005a. Forecasting total natural-gas consumption in Spain by using the stochastic Gompertz innovation diffusion model. *Applied Energy* **80**(2): 115–124.

Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A. 2006a. The stochastic Rayleigh diffusion model: statistical inference and computational aspects. Applications to modelling of real cases. *Applied Mathematics and Computation* **175**: 628–644.

Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A. 2006b. Electricity consumption in Morocco: stochastic Gompertz diffusion analysis with exogenous factors. *Applied Energy* **83**: 1139–1151.

Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A, Ramos E. 2007a. A diffusion model with cubic drift: statistical and computational aspects and application to modelling the global CO_2 emission in Spain. *Environmetrics* **18**: 55–69.

Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A. 2007b. Trend analysis using non-homogeneous stochastic diffusion processes. Emission of CO_2; Kyoto protocol in Spain. *Stochastic Environmental Research and Risk Assessment*. DOI 10.1007/s00477-006-0097-7

Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A, Ramos E. 2006c. A new stochastic Gompertzian process: computational inference and application in modelling. *Applied Mathematics and Computation* **183**: 738–747.

Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A, Román P, Torres F. 2005b. Inference in Gompertz type non homogeneous stochastic systems by means of discrete sampling. *Cybernetics and Systems* **36**: 203–216.

Gutiérrez R, Román P, Torres F. 1999. 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): 325–332.

Gutiérrez R, Román P, Torres F. 2001. Inference on some parametric functions in the univariate lognormal diffusion process with exogenous factors. *Test* **10**(2): 357–373.

Ha-Duonhd M, Swart R, Bernstein L. 2006. Uncertainty management in the IPCC: agreeing to disagree. *Global Environmental Change* 2007; **17**: 8–11. DOI: 10.1016/j.gloenvcha.2006.12.003

IPCC, Third Special Report. Climate Change 2001: The Scientific Basis. University Press, Cambridge.

Karlin S, Taylor HM. 1981. *A Second Course in Stochastic Processes*. Academic Press: New York.

Kloeden P, Platen E. 1992. *The Numerical Solution of Stochastic Differential Equations*. Springer: Berlin, Germany.

Kloeden P, Schurz H, Platten E, Sorensen M. 1992. *On Effects of Discretization on Estimators of Drift Parameters for Diffusion Processes.* Research Report no. 249, Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus, Aarhus, Denmark.

Kutoyants Yu. A. 2004. *Statistical Inference for Ergodic Diffusion Processes.* Springer Series in Statistics. Springer-Verlag London, Ltd: London.

Meade N, Islam T. 2006. Modelling and forecasting the diffusion of innovation. A 25-year review. *International Journal of Forecasting* **22**(3): 519–545.

Nobile AG, Ricciardi LM. 1984. Growth with regulation in fluctuating enviroments. *Biology Cybernetics* **49**: 179–188.

Nogues-Bravo D, Araújo MB, Errea MP, Martínez-Rica JD. 2007. Exposure of Global mountain systems to climate warning during the 21st Century. *Global Environmental Change*. DOI: 10.1016/j.gloenvcha.2006.11.007

Paravantis JA, Georgakellos DA. 2007. Trends in energy consumption and carbon dioxide emissions of passenger car and buses. *Technological Forecasting and Social Change* 2007; **74**: 682–707. DOI 10.1016/j.techfore.2006.05.005

Prakasa Rao B. 1999. *Statistical Inference for Diffusion Type Process*. Ed. Arnold: London.

Ricciardi LM. 1977. Diffusion process and related topic in biology. *Lectures Notes in Biomathematics* **14**: Springer Verlag.

Skiadas CH, Giovani AN. 1997. A stochastic Bass innovation diffusion model for studying the growth of electricity consumption in Greece. *Applied Stochastics Model and Data Analysis* **13**: 85–101.