Computational Features and Applications of Inference in the Stochastic of Inhomogeneous Gompertz Diffusion Process with Discrete Sampling

Nadia Makhlouki, Ahmed Nafidi and Boujemaâ Achchab

July 8, 2023

Computational features and applications of inference in the stochastic of inhomogeneous Gompertz diffusion process with discrete sampling*

Nadia Makhlouki[1] [0009−0005−3946−0092], Ahmed Nafidi[2,3] [0000−0001−9371−945X], and Achchab Boujemâa[3] [0000−0001−6365−0877]

Hassan First University of Settat, Berrechid National School of Applied Sciences, Mathematics and Informatics Department, LAMSAD, B.P. 218, 26103 Berrechid, Morocco. n.makhlouki@uhp.ac.ma

Abstract. In this study, We consider the Gompertz diffusion process-based stochastic inhomogeneous model. We begin by obtaining the analytical formulation for the process's probabilistic properties, the mean functions (conditional and non-conditional). Then, using the maximum likelihood technique and discrete sampling, we estimate the model's parameters. Finally, we used the stochastic inhomogeneous Gompertz diffusion process to analyze the development of the electric power consumption in Morocco in order to assess this method's capacity for modeling actual data.

Keywords: Inhomogeneous Gompertz diffusion model · Stochastic differential equation · Statistical inference in diffusion process · Mean function · Application to electric power consumption in Morocco.

1 Introduction

Our daily use of electricity results from the transformation of primary energy sources like coal, natural gas, nuclear energy, solar energy, and wind energy into electrical power, making it a secondary energy source. Since electricity may be transformed into other types of energy, such as mechanical energy or heat, it is sometimes referred to as an energy carrier. Although the power we consume is neither renewable nor nonrenewable, the primary energy sources are both.

The energy industry in Morocco is largely reliant on imported hydrocarbons. Currently, the nation imports around 90% of its energy requirements. Since 2004, the total primary energy consumption has grown by roughly 5% annually.
The development of the renewable energy industry is a top priority for the Moroccan government. The General Secretariat of the Government has received an amendment to Laws 13-09 on Renewable Energy and 16-08 on Self-Generation

* ENSAB

from the Ministry of Energy, Mines, and the Environment. While ensuring the security and viability of the national electrical grid, these revisions seek to strengthen the legal and regulatory framework governing renewable energy projects undertaken by the private sector. Per the state-owned power utility ONE, Moroccos electricity demand increased at an average annual rate of 6.7% between 2003 and 2013 as a result of population and economic expansion, resulting in an energy consumption of 32,015 GWh at the end of that year. From 483 kWh in 2002 to 843 kWh (approximate, estimate) in 2013, annual consumption per person has continuously climbed. Therefore, we draw the conclusion that modeling the evolution of energy consumption in general, and total electrical energy consumption in particular, as well as obtaining short- and medium-term forecasts, are very helpful in better understanding the historical development of the Moroccan economy and predicting its future development, as well as evaluating the impacts of this consumption on the global energy market. Determining short- and medium-term demand projections was the goal of this study, which served as the foundation for a more thorough investigation of the Moroccan energy market. This Fig. 1 shows the total renewable electricity net consumption and total electricity net consumption in Morocco (can be consulted at https://morocco.opendataforafrica.org/).

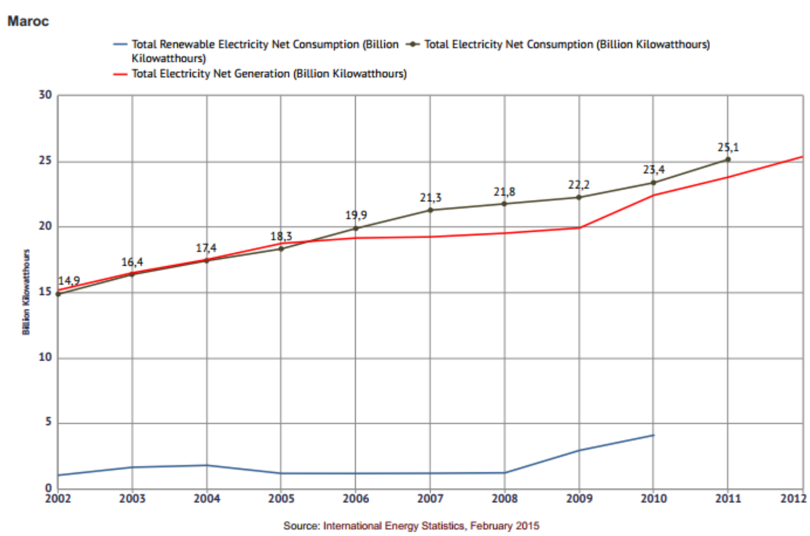


Fig. 1: Total renewable electricity net consumption and total electricity net consumption in Morocco.

The Stochastic Gompertz diffusion process (SGDP) is used to model stochastic phenomena in various fields of science. The homogenous case of this process was was introduced by Ricciardi (cf. [1]) in a theoretical form, and subsequently applied by Ferrante et al (cf. [2]) (growth of cancer cells) and by Gutiérrez et al (cf. [3]) (consumption of natural gas in Spain). and stock of motor vehicles in Spain (cf. [4]).However, the non-homogeneous case in which only the intrinsic growth rate in the drift is affected by exogenous factors (functions of time and some parameters) and with a constant deceleration coefficient, was applied, for example, in the price of new housing in Spain (cf. [5])and to the emission of CO2 (cf. [6]). Finally, Ferrante et al (cf. [7]) considered a non-homogeneous version in which the growth rate is the sum of two exponential functions that are exogenous factors. In the curent study, we define the stochastic inhomogeneous Gompertz diffusion process (SIGDP), which is used in various contexts. We first obtain the probabilistic characteristics of the process sush as the analytical expression, the transition probability density function (TPDF) , the mean functions (conditional and non-conditional). Then, we estimate the parameters by the maximum likelihood (ML) approach, with discrete sampling and getting the confidence bounds for the parameters. Finally, to evaluate the capability of this process for modeling real data, we applied the SIGDP to study the evolution of the electric power consumption in Morocco.

2 Basic probabilistic properties of the model

2.1 The suggested model

The following diffusion process provides the model's stochastic counterpart $\{x(t); t_0 \leq t \leq T\}$ taking values on $(0, \infty)$, $x(t)$ is a solution of the following stochastic differential equation (SDE)

$$dx(t) = \left(ax(t) - \frac{h'(t)}{h(t)} x(t) log(x(t)) \right) dt + \sigma x(t) dw(t). \tag{1}$$

Where $\sigma > 0$, $w(t)$ is a one-dimensional standard Wiener process, a represent the intrinsic growth rate and the function $h(t)$ is differentiable.

2.2 The process as analytically expressed

By the use of the Itô rule at the time-dependent transformation $y(t) = h(t)\, log(x(t))$, the SDE becomes

$$dy(t) = h(t) \left(a - \frac{\sigma^2}{2} \right) dt + \sigma h(t) dw(t),$$

By integrating, we have

$$y(t) = y(s) + (a - \frac{\sigma^2}{2}) \int_s^t h(\theta)\ d\theta + \sigma \int_s^t h(\theta) dw(\theta).$$

Finally, it follows that the explicit expression of the process

$$x(t) = exp\left\{ \frac{h(s)}{h(t)} log(x(s)) + \frac{(a - \frac{\sigma^2}{2})}{h(t)} \int_s^t h(\theta)\ d\theta + \frac{\sigma}{h(t)} \int_s^t h(\theta) dw(\theta) \right\} \tag{2}$$

2.3 Probability distribution of the

As the random variable $\int_s^t h(\theta)dw(\theta)$ has a one-dimensional normal distribution $\mathcal{N}_1(0, \int_s^t h^2(\theta)d\theta)$, we can deduce that the random variable $x(t)/x(s) = x_s \sim \Lambda_1(\mu(s,t,x_s), \sigma^2\nu^2(s,t))$, a one-dimensional log-normal distribution with

$$\mu(s,t,x_s) = \frac{h(s)}{h(t)}log(x(s)) + \frac{(a-\frac{\sigma^2}{2})}{h(t)}\int_s^t h(\theta)\ d\theta,$$
$$\nu^2(s,t) = \frac{1}{h^2(t)}\int_s^t h^2(\theta)d\theta.$$

The TPDF of this process $f(x,t|y,s)$ takes the form

$$f(x,t|y,s) = \frac{1}{x\sqrt{2\pi\sigma^2\nu^2(s,t)}}exp\left(-\frac{[log(x)-\mu(s,t,x)]^2}{2\sigma^2\nu^2(s,t)}\right). \tag{3}$$

2.4 Computation of the mean function

From the properties of the Lognormal distribution, the r-th conditional moment of the process is

$$\mathbb{E}(x^r/x(s) = x_s) = exp\left(r\mu(s,t,x_s) + \frac{r^2\sigma^2\nu^2(s,t)}{2}\right).$$

For $r = 1$, the conditional mean function (CMF) of the process is:

$$\mathbb{E}(x(t)/x(s) = x_s) = exp\left\{\frac{h(s)}{h(t)}log(x(s)) + \frac{\left(a-\frac{\sigma^2}{2}\right)}{h(t)}\int_s^t h(\theta)d\theta + \frac{\sigma^2}{2h^2(t)}\int_s^t h^2(\theta)d\theta\right\}. \tag{4}$$

Assuming the initial condition $P(x(t_1) = x_1) = 1$, the mean function (MF) of the process is

$$\mathbb{E}(x(t)) = exp\left\{\frac{h(t_1)}{h(t)}log(x_{t_1}) + \frac{\left(a-\frac{\sigma^2}{2}\right)}{h(t)}\int_{t_1}^t h(\theta)d\theta + \frac{\sigma^2}{2h^2(t)}\int_{t_1}^t h^2(\theta)d\theta.\right\} \tag{5}$$

3 Inference on the model

Let us then examine in this section the ML estimation of the parameters of the model from which we can obtain, by virtue of Zehnas theorem [8], the corresponding for the aforementioned parametric functions.

3.1 Parameter estimation

We consider a discrete sampling of the process $x_1, x_2, \cdots, x_n$ for times $t_1 < t_2 < \cdots < t_n$. The likelihood function depends on the choice of the initial distribution. If $P(x(t_1) = x_1) = 1$, the associated likelihood function can be written as

$$\mathbb{L}(x_1, \cdots, x_n, \alpha, \sigma^2) = \prod_{i=2}^{n} f(x_i, t_i | x_{i-1}, t_{i-1}),$$

which is written as

$$\mathbb{L} = \prod_{i=2}^{n} \frac{1}{x_i\sqrt{2\pi\sigma^2\nu^2(s,t)}} exp\left(-\frac{\left\{ log(x_i) - \frac{h(t_{i-1})}{h(t_i)} log(x_{i-1}) - \frac{(a - \frac{\sigma^2}{2})}{h(t_i)} \int_{ti-1}^{t_i} h(\theta)\ d\theta \right\}^2}{2\sigma^2\nu^2(s,t)} \right).$$

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 discrete sampling of the process: $v_1 = x_1$, and $v_i = \nu_i^{-1} \left(log(x_i) - \frac{h(t_{i-1})}{h(t_i)} log(x_{i-1}) \right)$ for $i = 2, \cdots, n$ with thus, the likelihood function can be obtained from Equation (3) by the following expression

$$\mathbb{L}(\mathbf{v}, \mathbf{a}, \sigma^2) = [2\pi\sigma^2]^{-(n-1)/2} exp\left\{ -\frac{1}{2\sigma^2} (\mathbf{v} - \mathbf{U}'\mathbf{a})' (\mathbf{v} - \mathbf{U}'\mathbf{a}) \right\}$$

where

$$\mathbf{a} = a - \sigma^2/2,\ \mathbf{v} = (v_2, \cdots, v_n)',$$

$$\nu_i = \nu(t_{i-1}, t_i),$$

$$u_i = \frac{\nu_i^{-1}}{h(t_i)} \int_{t_{i-1}}^{t_i} h(\theta) d\theta$$

and $\mathbf{U}$ is the $1 \times (n-1)$ matrix, whose rank is assumed to be 1, given by $\mathbf{U} = (\mathbf{u}_2, \cdots, \mathbf{u}_n)$.

The log-likelihood for Equation (3) has the following form

$$Log(\mathbb{L}(\mathbf{v}, \mathbf{a}, \sigma^2)) = -\frac{n-1}{2} log(2\pi) - \frac{n-1}{2} log(\sigma^2) - \frac{1}{2\sigma^2} (\mathbf{v} - \mathbf{U}'\mathbf{a})' (\mathbf{v} - \mathbf{U}'\mathbf{a})$$

By deriving the log-likelihood function with respect to σ^2 and a we obtain

$$\frac{\partial Log(\mathbb{L})}{\partial \sigma^2} = -\frac{n-1}{2\sigma^2} + \frac{1}{2\sigma^4} (\mathbf{v} - \mathbf{U}'\mathbf{a})' (\mathbf{v} - \mathbf{U}'\mathbf{a}) \qquad (6)$$

$$\frac{\partial Log(\mathbb{L})}{\partial \mathbf{a}} = -\frac{1}{2\sigma^2}\frac{\partial[(\mathbf{v}-\mathbf{U}^{'}\mathbf{a})^{'}(\mathbf{v}-\mathbf{U}^{'}\mathbf{a})]}{\partial \mathbf{a}} = \frac{1}{\sigma^2}\mathbf{U}(\mathbf{v}-\mathbf{U}^{'}\mathbf{a}) \tag{7}$$

Making the derivatives (6) and (7) equal to zero, we obtain the following equations

$$-(n-1)\sigma^2 + (\mathbf{v}-\mathbf{U}^{'}\mathbf{a})^{'}(\mathbf{v}-\mathbf{U}^{'}\mathbf{a}) = 0 \tag{8}$$

$$\mathbf{U}\mathbf{v} - \mathbf{U}\mathbf{U}^{'}\mathbf{a} = 0 \tag{9}$$

The equations (8) and (9) becomes

$$\mathbf{U}\mathbf{v} = \mathbf{U}\mathbf{U}^{'}\mathbf{a} \tag{10}$$

$$(n-1)\sigma^2 = (\mathbf{v}-\mathbf{U}^{'}\mathbf{a})^{'}(\mathbf{v}-\mathbf{U}^{'}\mathbf{a}) \tag{11}$$

The ML estimators of $\mathbf{a}$ and σ^2 yield

$$\hat{\mathbf{a}} = (\mathbf{U}\mathbf{U}^{'})^{-1}\mathbf{U}\mathbf{v} \tag{12}$$

$$(n-1)\hat{\sigma}^2 = \mathbf{v}^{'}\mathbf{H_u}\mathbf{v} \tag{13}$$

where the matrix $\mathbf{H_u}$ is the symmetric and idempotent matrix given by

$$\mathbf{H_u} = \mathbf{I}_{n-1} - \mathbf{U}^{'}(\mathbf{U}\mathbf{U}^{'})^{-1}\mathbf{U}.$$

3.2 Estimated mean functions

By using Zehna's theorem [8], the Estimated Mean Function (EMF) and Estimated Conditional Mean Function (ECMF) of the proposed model are obtained by replacing the parameters in Equations (4) and (5) by their estimators given in Equations (12) and (13). Then, the ECMF has the following expressions:

$$\mathbb{E}(x(t)) = exp\left\{\frac{h(s)}{h(t)}log(x_s) + \frac{\left(\hat{\mathbf{a}} - \frac{\hat{\sigma}^2}{2}\right)}{h(t)}\int_s^t h(\theta)d\theta + \frac{\sigma^2}{2h^2(t)}\int_s^t h^2(\theta)d\theta.\right\} \tag{14}$$

Under the initial condition $P(x(t_1) = x_1) = 1$, the EMF of the process is:

$$\mathbb{E}(x(t)) = exp\left\{\frac{h(t_1)}{h(t)}log(x_{t_1}) + \frac{\left(\hat{\mathbf{a}} - \frac{\hat{\sigma}^2}{2}\right)}{h(t)}\int_{t_1}^t h(\theta)d\theta + \frac{\sigma^2}{2h^2(t)}\int_{t_1}^t h^2(\theta)d\theta.\right\} \tag{15}$$

3.3 Properties of maximum likelihood estimators

Distribution of maximum likelihood estimators The likelihood function can be rewritten in the following form

$$\mathbb{L}(\mathbf{v},\mathbf{a},\sigma^2) = [2\pi]^{-\frac{(n-1)}{2}}|\sigma^2 I_{n-1}|^{-1/2}exp\left\{-\frac{1}{2}(\mathbf{v}-\mathbf{U}^{'}\mathbf{a})^{'}(\sigma^2 I_{n-1})^{-1}(\mathbf{v}-\mathbf{U}^{'}\mathbf{a})\right\}$$

From which, we deduce that

$$\mathbf{v} \sim \mathcal{N}_{n-1}(\mathbf{U}^{'}\mathbf{a}, \sigma^2 I_{n-1})$$

The rank of $\mathbf{U}$ is supposed equal 2. Then, $(\mathbf{U}\mathbf{U}^{'})^{-1}\mathbf{U}$ has the same rank, and we have

$$(\mathbf{U}\mathbf{U}^{'})^{-1}\mathbf{U}\mathbf{v} \sim \mathcal{N}_2\left((\mathbf{U}\mathbf{U}^{'})^{-1}\mathbf{U}\mathbf{U}^{'}\mathbf{a}, \sigma^2(\mathbf{U}\mathbf{U}^{'})^{-1}(\mathbf{U}\mathbf{U}^{'})(\mathbf{U}\mathbf{U}^{'})^{-1}\right)$$

and therefore, we have

$$\hat{\mathbf{a}} \sim \mathcal{N}_1\left(\mathbf{a}, \sigma^2(\mathbf{U}\mathbf{U}^{'})^{-1}\right)$$

In order to obtain the distribution of $\hat{\sigma}^2$, we make use the following result (see for example [9], corollary 2.11.2):

Corollary 1. *If $Z \sim \mathcal{N}_p[\mu, \Sigma]$, Σ non singular and $\mathbf{B}_{p\times p}$ symmetric, then, $Z'\mathbf{B}Z \sim \chi^2_k(\delta)$, where $k = rank(\mathbf{B})$ and $\delta = \mu'\mathbf{B}\mu$ if and only if $\mathbf{B}\Sigma$ is idempotent.*

As $\mathbf{H}_U$ is symmetric and idempotent, then,

$$rank(\mathbf{H}_U) \;=\; tr(\mathbf{H}_U) \;=\; n-2,$$

then using the last result in the particular case: $Z = \sigma^{-1}\mathbf{v},\ \Sigma = I_{n-1},\ \mathbf{B} = \mathbf{H}_U$ and $\mu = \mathbf{U}'\mathbf{a}$, we have

$$\frac{\mathbf{v}'}{\sigma}\mathbf{H}_U\frac{\mathbf{v}}{\sigma} \sim \chi^2_{n-2}(\delta),\ with\ \delta = \mathbf{a}'\mathbf{U}\mathbf{H}_U\mathbf{U}'\mathbf{a} = 0$$

and therefore

$$\frac{(n-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-2}$$

The independence between $\hat{\mathbf{a}}$ and $\hat{\sigma}^2$ can be proved by using the following result (see Ref. [9] corollary 2.11.4, p.66).

Corollary 2. *Let $Z \sim \mathcal{N}_p[\mu, \Sigma]$, with $\Sigma > 0$. Then, $y_j = Z'A_jZ + 2b_j'Z + c_j,\ j = 1,2$ are independently distributed if and only if $A_1\Sigma A_2 = 0,\ A_2\Sigma b_1 = 0,\ A_1\Sigma b_2 = 0$, and $b_1'\Sigma b_2 = 0$.*

If we choose $Z = \mathbf{v} \sim \mathcal{N}_{n-1}(\mathbf{U}'\mathbf{a}, \sigma^2 I_{n-1});\ A_1 = \mathbf{H}_U;\ b_1 = 0;\ c_1 = 0$ and $A_2 = 0;\ b_2 = (\mathbf{U}\mathbf{U}')^{-1}\mathbf{U}$ and $c_2 = 0$ then the necessary and sufficient conditions of corollary 3 are satisfied and therefore $(\mathbf{U}\mathbf{U}')^{-1}\mathbf{U}\mathbf{v}$ and $\mathbf{v}'\mathbf{H}_U\mathbf{v}$ are independently distributed, which means that $\hat{a}$ and $\hat{\sigma}^2$ are independently distributed.

Sufficiency and completeness of the estimators By subtracting and adding $\mathbf{U}'\hat{\mathbf{a}}$ to $\mathbf{v} - \mathbf{U}'\mathbf{a}$, expression of likelihood function becomes

$$\mathbb{L}(\mathbf{v}, \mathbf{a}, \sigma^2) = \frac{1}{[2\pi\sigma^2]^{\frac{(n-1)}{2}}} exp\left\{-\frac{1}{2\sigma^2}[(n-1)\hat{\sigma}^2 + (\hat{\mathbf{a}}-\mathbf{a})'\mathbf{U}\mathbf{U}'(\hat{\mathbf{a}}-\mathbf{a})]\right\}$$

which means that $(\hat{\mathbf{a}}, \hat{\sigma}^2)$ is conjointly sufficient for $(\mathbf{a}, \sigma^2)$.
The completeness follows by means of a similar reasoning to that established for the maximum likelihood estimators of the parameters of the multivariate normal distribution (see, for example, Anderson [10]).
And so the estimators $\hat{\mathbf{a}}$ and $\frac{(n-1)\hat{\sigma}^2}{(n-2)\sigma^2}$ are the UMVUE for $\mathbf{a}$, σ^2, respectively.

confidence bounds The $(1-\alpha)\%$ confidence bound for the parameter σ^2 is given, by

$$\left[\hat{\mathbf{a}} - \hat{\sigma}.t_{\alpha/2,n-2}/\sqrt{n-1}, \hat{\mathbf{a}} + \hat{\sigma}.t_{\alpha/2,n-2}/\sqrt{n-1}\right] \quad (16)$$

$$\left[(n-1)\hat{\sigma}^2/\chi^2_{\alpha/2,n-2}, (n-1)\hat{\sigma}^2/\chi^2_{1-\alpha/2,n-2}\right] \quad (17)$$

where $\chi^2_{\alpha,n}$ and $t_{\alpha,n}$ are the upper 100α per cent points of the chi squared distribution and the Student distribution, respectively, with n degrees of freedom.

4 Application

The model used in this study was applied to actual data for Morocco's total electricity usage (reported in billion kilowathours) from 1800 to 2012. These statistics, which relate to sales by ONE, the Moroccan authority, are accessible at https://morocco.opendataforafrica.org/. Two steps that make up the methodology are as follows::
• The first step: To estimate the model's parameters, start with the first 30 data in the sequence of observations being analyzed, using expressions (12) and (13). Then establish the relevant confidencebounds using equations (16) and (17).
• The second step:Predict the corresponding values for Morocco's electricity consumption for the years 2011 and 2012 using the estimated mean function (EMF) and estimated conditional mean function (ECMF), which were obtained by swapping the parameters in expressions (14) and (15), with their respective estimators, and then contrast the results with the corresponding observed data for the same years.

For the computations needed for the present study, a Matlab application was used. Think about it, for example, the function $h(t) = \dfrac{1-t^2-t^4}{t+1}$, The corresponding estimators' values, and the confidence bounds, are $\hat{\mathbf{a}} = 0.060651$ and $\hat{\sigma} = 1.094854.10^{-3}$ with confidence bounds $(0.048313; 0.072988)$ and $(0.699151; 1.956171).10^{-3}$.

Table 1 explains the fit and forecast made possible by the EMF with its EMF_l and EMF_u.
Table 2 explains the fit and forecast made possible by the $ECMF$ with its $ECMF_l$ and $ECMF_u$.
Fig. 1 illustrates the relationship between the real data and the fit and forecast made using the EMF with EMF_l and EMF_u.
Fig. 2 illustrates the fit and forecasting made with the $ECMF$ of the model

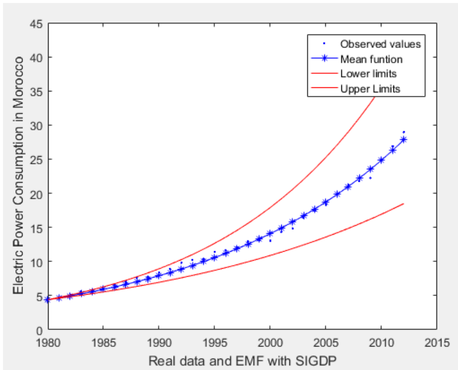
with respect to the real data.
MATLAB was used to do all computations.

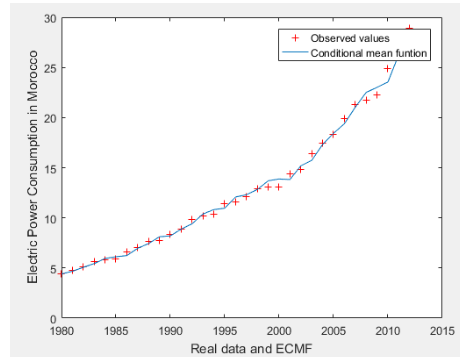
Table 1: Real data, EMF, EMF_l and EMF_u

Year	**Real data**	**EMF**	**EMF**$_l$	**EMF**$_u$
1980	4.409	4.409	4.409	4.409
1981	4.774	4.676	4.615	4.734
1982	5.130	4.959	4.832	5.082
1983	5.612	5.259	5.057	5.455
1984	5.776	5.577	5.294	5.856
1985	5.884	5.914	5.541	6.284
1986	6.568	6.269	5.798	6.744
1987	7.018	6.646	6.068	7.237
1988	7.656	7.045	6.349	7.764
1989	7.744	7.467	6.643	8.329
1990	8.370	7.914	6.951	8.935
1991	8.877	8.387	7.272	9.583
1992	9.804	8.888	7.607	10.277
1993	10.218	9.417	7.957	11.021
1994	10.350	9.977	8.323	11.817
1995	11.404	10.569	8.706	12.669
1996	11.617	11.196	9.105	13.581
1997	12.114	11.859	9.521	14.558
1998	12.935	12.560	9.956	15.603
1999	13.103	13.301	10.411	16.721
2000	13.050	14.085	10.885	17.917
2001	14.351	14.914	11.380	19.198
2002	14.856	15.790	11.897	20.567
2003	16.361	16.716	12.437	22.032
2004	17.411	17.695	13.001	23.599
2005	18.315	18.730	13.588	25.275
2006	19.872	19.824	14.202	27.067
2007	21.266	20.980	14.842	28.983
2008	21.751	22.201	15.511	31.032
2009	22.243	23.492	16.208	33.222
2010	24.844	24.855	16.935	35.563
2011	26.871	26.296	17.695	38.065
2012	28.946	27.818	18.486	40.739

Table 2: Real data, ECMF, $ECMF_l$ and $ECMF_u$

Year	**Real data**	**ECMF**	**ECMF**$_l$	**ECMF**$_u$
1980	4.409	4.409	4.409	4.409
1981	4.774	4.676	4.616	4.732
1982	5.130	5.063	4.998	5.123
1983	5.612	5.440	5.370	5.504
1984	5.776	5.950	5.874	6.021
1985	5.884	6.123	6.045	6.196
1986	6.568	6.238	6.158	6.312
1987	7.018	6.962	6.873	7.045
1988	7.656	7.438	7.343	7.527
1989	7.744	8.113	8.010	8.210
1990	8.370	8.206	8.102	8.304
1991	8.877	8.869	8.756	8.974
1992	9.804	9.405	9.285	9.517
1993	10.218	10.386	10.254	10.509
1994	10.350	10.824	10.686	10.953
1995	11.404	10.964	10.824	11.094
1996	11.617	12.078	11.924	12.222
1997	12.114	12.304	12.147	12.449
1998	12.935	12.829	12.666	12.982
1999	13.103	13.697	13.523	13.861
2000	13.050	13.875	13.698	14.041
2001	14.351	13.819	13.643	13.983
2002	14.856	15.195	15.001	15.375
2003	16.361	15.729	15.528	15.916
2004	17.411	17.319	17.098	17.525
2005	18.315	18.429	18.193	18.648
2006	19.872	19.385	19.137	19.6147
2007	21.266	21.031	20.762	21.281
2008	21.751	22.503	22.215	22.770
2009	22.243	23.016	22.722	23.289
2010	24.844	23.536	23.235	23.815
2011	26.871	26.284	25.948	26.596
2012	28.946	28.426	28.062	28.763





4.1 Fit quality

The following scale-dependent measurements and quantities are based on absolute errors, squared errors, and percentage errors:

$$Mean\ Absolute\ Error\ (MAE) = \frac{1}{N}\sum_{i=1}^{n} \mid x(t_i) - \hat{x}(t_i) \mid,$$

$$Root\ Mean\ Square\ Error\ (RMSE) = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(x(t_i) - \hat{x}(t_i))^2},$$

$$Mean\ Absolute\ Percentage\ Error\ (MAPE) = \frac{1}{N}\sum_{i=1}^{n}\frac{\mid x(t_i) - \hat{x}(t_i) \mid}{x(t_i)} \times 100.$$

with $\hat{x}(t)$ is obtained by substituting the parameters in Equation (2) by their estimators.
The values obtained for the above error measures are acceptably low, especially the MAPE according to Table (3). The statistical measures obtained are illustrated in the Table (4).

Table 3: Interpretation of typical Mean Absolute Percentage Error (MAPE) values.

MAPE	**Interpretation**
< 10	Highly accurate forecasting
$20 - 30$	Good forecasting
$30 - 50$	Reasonable forecasting
> 50	Inaccurate forecasting

5 Conclusions

This article presents a study of the non-homogeneous stochastic Gompertz diffusion process (NHSGDP), including all its probabilistic properties and the corre-

Table 4: Fit quality of the model.

Measures of Forecasting Accuracy Error	**Values of NHGDP**
MAE	0.361247990035216
RMSE	0.459024822910505
MAPE	2.759618319944835

sponding statistical inference. As a particular case in the limit comparison test, we also study the homogeneous stochastic Gompertz diffusion process (HSGDP). In the future, it will be possible to apply these models to fit real data and to obtain goodness of fit results between the processes and the data. We will also study the possibility of defining all these processes in their non-homogeneous form, by introducing exogenous factors, and considering the use of numerical methods to obtain the estimates.

References

1. L.M.Ricciardi, Diffusion processes and related topics in Biology. Lecture Notes in Biomathematics 14, (1977).
2. L.Ferrante, S.Bompade, L.Possati, and L.Leone, Parameter estimation in a Gompertzian stochastic-model for tumor growth. Biometrics 56 (2000), 107681.
3. R.Gutiérrez, A.Nafidi and R.Gutiérrez-Sánchez. Forecasting total natural gas consumption in Spain by using the stochastic Gompertz innovation diffusion model. Appl Energy 80, 2 (2005), 115124.
4. R.Gutiérrez, R.Gutiérrez-Sánchez and A.Nafidi. Modelling and forecasting vehicle stocks using the trends of stochastic Gompertz diffusion models. The case of Spain. Appl. stochastic models Bus. Ind. 25(2009) 385-405.
5. R.Gutiérrez, R.Gutiérrez-Sánchez, A.Nafidi, P.Román, F.Torres. Inference in Gompertz-type nonhomogeneous stochastic systems by means of discrete sampling. Cybernetics and Systems 36 (2005) 203-216.
6. R.Gutiérrez, R.Gutiérrez-Sánchez and A.Nafidi. Trend analysis using nonhomogeneous stochastic diffusion processes. Emission of CO2; Kyoto protocol in Spain. Stoch Environ Res Risk Assess. 22 (2008) 57-66.
7. L.Ferrante, S.Bompade, L.Possati, L.Leone, and M.P.Montanari. A stochastic formulation of the Gompertzian growth model for in vitro bactericidad kinetics: Parameter estimation and extinction probability. Biometrical Journal 47, 3(2005), 309318.
8. Zehna P.W. Invariance of maximum likelihood estimators. Ann. Math. Stat. 1966, 37, 744.
9. Srivastava MS, Khatri CG. An introduction to multivariate analysis. New York: North-Holland Scientific Publishers; 1979.
10. Anderson TW. An introduction to multivariate statistical analysis. 2nd ed. New York: Wiley; 1984.