 

# Application of Stochastic Lognormal Diffusion Model with Polynomial Exogenous Factors to Energy Consumption in Ghana

Godfred Kwame Abledu(PhD)

School of Applied Science and Technology, Koforidua Polytechnic, PO Box 981, Koforidua, Ghana

* E-mail of the corresponding author: godfredabledu@gmail.com

**Abstract.**

The main objective of this paper was the application of maximum likelihood ratio tests in lognormal diffusions with polynomial exogenous factors. The model described an innovation diffusion process considering at the same time disturbances coming from the environment of the system. Finally, the model was applied to energy consumption data in Ghana from 1999 to 2010.

Maximum likelihood estimators (MLEs) were obtained for the drift and diffusion coefficients characterizing lognormal diffusion models involving exogenous factors affecting the drift term. The present paper provides the distribution of these MLEs, the Fisher information matrix, and the solution to some likelihood ratio tests of interest for hypotheses on the parameters weighting the relative effect of the exogenous factors.

The results show that the total consumption of primary energy presents structural characteristics. The endogenous consumption pattern in Ghana, in absolute terms, also presents a clear upward trend.

**Key works:** lognormal diffusions model, maximum likelihood estimators, endogenous actors, energy consumption

.

## 1 Introduction

The use of diffusion processes with exogenous factors and their trend is common in many fields. The reason of its application is the usual presence of deviations of the observed data with respect to the trend of some known homogenous diffusion process, in some time intervals. These factors are time dependent functions that allow, on one hand, a best fit to the data and, on the other hand, an external control to the process behaviour. The factors must be totally or partially known, that is, their functional form or some aspects about their time evolution must be available. The problem of estimating the parameters of the drift coefficient in these models has received considerable attention recently, especially in situations in which the process is observed continuously. The statistical inference is usually based on approximating maximum likelihood methodology. An extensive review of this theory an be found in Prakasa(1999), and related new work has been done by Kloeden et al. (1999),

The usefulness of diffusion random fields in describing, for example, economic or environmental phenomena, has led to significant developments, particularly regarding inferential aspects. In that respect, from the contribution to theoretical foundations for diffusions given in Nualart (1983) and Ricciardi (1976), the lognormal diffusions involving exogenous factors affecting the drift term is considered. The maximum likelihood estimators (MLEs) for the drift and diffusion coefficients is obtained, which characterize these diffusions under certain conditions. Using these MLEs, techniques for estimation, prediction and conditional simulation of lognormal diffusions are developed.

The study of variables that model dynamical systems has undergone a great development over the last decades, and a variety of statistical and probabilistic techniques has been worked out for this purpose. Among these, stochastic processes, and in particular diffusion processes, have been systematically employed.

## 2. Lognormal Diffusion Process Model (LNDP)

The lognormal diffusion process with exogenous factors is defined as $X(t) : t_0 \leq T$ with infinitesimal moments $A_1(x,t) = xh(t)$ and $A_1(x,t) = \sigma^2 x^2$, where $\sigma > 0$ and $h$ is continuous function in $t_0, T$ containing the external information sources. So it is usual to take h as a linear combination of continuous functions. A class of two-parameter random fields which are diffusions on each coordinate and satisfy a particular Markov property related to partial ordering in $\mathbb{R}_+^2$ are considered by Nualart(1983). Using this theory, Skiadas(2007) introduced a 2D lognormal diffusion random field as follows.

 

Let $X(z):z=(s,t)\in I=[0,S]\times[0,T]\subset \mathbb{R}_+^2$ be a positive-valued Markov random field, defined on a probability space $(\Omega, A, P)$, where $X(0,0)$ is assumed to be constant or a lognormal random variable with $E[\ln X(0,0)]=\phi_0$ and $Var[\ln X(0,0)]=\sigma_0^2$. The distribution of the random field is determined by the following transition probabilities:

$$P(B,(s+h,t+k)\mid(x_1,x,x_2),z)=P(X(s+h,t+k)\in B\mid X(s,t+k)=x_1, X(z))$$

where $z=(s,t)\in I, h,k>0$, $(x_1,x,x_2)\in \mathbb{R}_+^2$ and $B$ is a Borel subset. It is assumed that the transition densities exist and are given by

$$g(y,(s+h,t+k)\mid(x_1,x,x_2),z)$$
$$=\frac{1}{y\sqrt{2\pi\sigma^2_{z;h,k}}}\exp\left\{-\frac{1}{2}\left(\frac{\ln\frac{yx}{x_1x_2}-m_{z;h,k}}{\sigma_{z;h,k}}\right)^2\right\}$$

for $y\in\mathbb{R}_+^2$, with

$$m_{z;h,k}=\int_s^{s+k}\int_t^{t+k}\tilde{a}(\sigma,\tau)\,d\sigma d\tau,\ \sigma^2_{z;h,k}=\int_s^{s+k}\int_t^{t+k}\tilde{B}(\sigma,\tau)\,d\sigma d\tau$$

and $\tilde{a}$, $\tilde{B}$ being continuous functions on *I*. Under these conditions we can assert that $\{X(z):z\in I\}$ is a lognormal diffusion random field. The one parameter drift and diffusion coefficients associated are given by:

$$a_1(z)x:=\left(\tilde{a}_1(z)+\frac{1}{2}\tilde{B}_1(z)\right)x,\ B_1(z)x^2:=\tilde{B}_1(z)x^2$$
$$a_2(z)x:=\left(\tilde{a}_2(z)+\frac{1}{2}\tilde{B}_2(z)\right)x,\ B_2(z)x^2:=\tilde{B}_2(z)x^2$$

where

$$\tilde{a}_1(s,t)=\int_0^t\tilde{a}(s,r)\,dr,\ \tilde{B}_1(s,t)=\int_0^t\tilde{B}rdt$$
$$\tilde{a}_2(s,t)=\int_0^t\tilde{a}(\sigma,t)\,d\sigma,\ \tilde{B}_1(s,t)=\int_0^t\tilde{B}(\sigma,t)\,dt$$

for all $z=(s,t)\in I, x\in\mathbb{R}_+$.

The random field $\{Y(z):z\in I\}$ defined as $Y(z)=\ln X(z)$ is then a Gaussian diffusion random field with $\tilde{a}$ and $\tilde{B}$ being, respectively, the drift and diffusion coefficients, and $\tilde{a}_1$, $\tilde{a}_2$, $\tilde{B}_1$ and $\tilde{B}_2$ being the corresponding one parameter drift and diffusion coefficients. Furthermore, if $z,z'\in I, z=(s,t), z'=(s',t')$, then

$$m_Y(z):=E[Y(z)]=\phi_0+\int_0^s\int_0^t\tilde{a}(\sigma,\tau)\,d\sigma dr$$
$$\sigma_Y^2(z):=Var[Y(z)]=\sigma_0^2+\int_0^s\int_0^t\tilde{B}(\sigma,\tau)\,d\sigma dr$$
$$c_Y(z,z'):=Cov(Y(z),Y(z'))=\sigma_Y^2(z\wedge z')$$

It is also assumed that the conditions usually considered for estimation of the drift and diffusion coefficients in the one-parameter case hold, that is, $P[\ln X(0,0)=\phi_0]=1$ and $\sigma_Y^2(z)=\tilde{B}st$, $z=(s,t)\in I$.

**3. Inference in the Lognormal Diffusion Process Model (LNDP)**

Let $\{X(z): z \in I\}$ be a lognormal diffusion random field. Data $\mathrm{X}=(X(z_1),\ldots,X(z_n))^t$ are assumed to be observed at known spatial locations $z_1=(s_1,t_1)$, $z_2=(s_2,t_2)$, …, $z_n=(s_n,t_n) \in I$. Let $\mathrm{x}=(x_1,x_2,\ldots,x_n)^t$ be a sample. Let us consider the log-transformed n-dimensional random vector, $\mathrm{Y}=(Y(z_1),Y(z_2),\ldots,Y(z_n))^t=(\ln X(z_1),\ldots,\ln X(z_n))^t=\ln \mathrm{X}$, and the log-transformed sample $y=(y_1,y_2,\ldots,y_n)^t=\ln \mathrm{x}$. We denote $\mathrm{m}_Y=(m_Y(z_1),m_Y(z_2),\ldots,m_Y(z_n))^t$ and $\Sigma_Y=\left(\sigma_Y^2(z_i \Lambda z_j)\right)_{i,j=1,\ldots,n}$

In order to estimate the MLEs for the drift and diffusion coefficient using exogenous factors, it is supposed that the drift coefficient $\tilde{a}$ of $Y$ is a linear combination of several known functions, $\{h_1(z),\ldots,h_p(z): z \in I\}$, with real coefficients $\phi_1,\ldots,\phi_p$:

$$\tilde{a}(z)=\sum_{\alpha=1}^{p}\phi_\alpha h_\alpha(z), z \in I$$

Defining for $z=(s,t) \in I$,

$$f_0(z)=1,\ f_\alpha(z)=\int_0^s\int_o^t h_\alpha(\sigma,r)\,d\sigma dr,\ \alpha=1,\ldots,p$$

the mean of Y is given by

$$m_Y(s,t)=\phi_0+\sum_{\alpha=1}^{p}\phi_0\int_0^s\int_o^t h_\alpha(\sigma,r)\,d\sigma dr=\sum_{\alpha=0}^{p}\phi_\alpha f_\alpha(z)$$

Thus, denoting $\mathrm{F}=(\mathrm{f}_0,\mathrm{f}_1,\ldots,\mathrm{f}_p)$, with $\mathrm{f}_\alpha=(f_\alpha(z_1),f_\alpha(z_2),\ldots,f_\alpha(z_n))^t$ for $\alpha=0,1,\ldots,p$, and $\phi=(\phi_0,\ldots,\phi_p)^t$ the following is obtained:

$$m_Y=(\phi_0\mathrm{f}_0+\phi_1\mathrm{f}_1+\ldots+\phi_p\mathrm{f}_p)=\mathrm{F}\phi$$

Let us write

$$\Sigma_Y=\hat{B}\mathrm{M}:=\hat{B}\begin{pmatrix} s_1t_1 & (s_1\Lambda s_2)(t_1\Lambda t_2) & \ldots & (s_1\Lambda s_n)(t_1\Lambda t_n) \\ (s_1\Lambda s_2)(t_1\Lambda t_2) & s_2t_2 & \ldots & (s_2\Lambda s_n)(t_2\Lambda t_n) \\ \ldots & \ldots & \ldots & \ldots \\ (s_1\Lambda s_n)(t_1\Lambda t_n) & (s_2\Lambda s_n)(t_2\Lambda t_n) & \ldots & s_nt_n \end{pmatrix}$$

With this notation, the MLEs for the drift and diffusion coefficients are, respectively:

$$\phi^*=(\phi_0^*,\phi_1^*,\ldots,\phi_p^*)^t=(\mathrm{F}^t\mathrm{M}^{-1}\mathrm{F})^{-1}\mathrm{F}^t\mathrm{M}^{-1}\ln \mathrm{x} \tag{1}$$

and

$$\tilde{B}^*=\frac{1}{2}(\ln \mathrm{x}-m_Y^*)^t\mathrm{M}^{-1}(\ln \mathrm{x}\text{-}\mathrm{m}_Y^*) \tag{2}$$

where $m_Y^*=\mathrm{F}\phi^*$ (Gutierrez, et. al., 2005)

The expected value of the MLE $\phi^*$ is

$$E\left[\phi^*\right]=(\mathrm{F}^t\mathrm{M}^{-1}\mathrm{F})^{-1}\mathrm{F}^t\mathrm{M}^{-1}\mathrm{E}(\ln \mathrm{X})=(\mathrm{F}^t\mathrm{M}^{-1}\mathrm{F})^{-1}\mathrm{F}^t\mathrm{M}^{-1}E(Y)$$
$$=(\mathrm{F}^t\mathrm{M}^{-1}\mathrm{F})^{-1}\mathrm{F}^t\mathrm{M}^{-1}\mathrm{F}\phi=\phi$$

 

and then, $\phi^*$ is unbiased. Taking into account that $\mathrm{Y} \sim \mathrm{N}\ \phi, \tilde{B}M$ , it is clear that the distribution of the estimators are given by

$$\phi^* \sim \mathrm{N}\ \phi, \tilde{B}\ \mathrm{F^t M^{-1} F}^{\ -1} \tag{3}$$

and

$$\mathrm{n}\tilde{B}^* \sim W_1\ \text{n-p+1}, \tilde{B} \tag{4}$$

The last estimator is biased. Therefore, the following transformation is considered:

$$\tilde{B}^{**} = \frac{n}{n-p-1} = \frac{1}{n-p-1}\ \mathrm{In\ X - F}\phi^{*\ t}\ \mathrm{M}^{-1}\ \mathrm{In\ X - F}\phi^* \tag{5}$$

Which is unbiased estimator of $\tilde{B}$ . In addition, $\tilde{B}^*$ and $\phi^*$ are independent and

$$Var\ \tilde{B}^{**} = Var\left(\frac{n}{n-p+1}\tilde{B}^*\right) = \frac{1}{n-p+1} Var\ n\tilde{B}^* = \frac{2\tilde{B}^2}{n-p+1}$$

Therefore, the covariance matrix of $\phi^*$ and $\tilde{B}^{**}$ is

$$\tilde{B}^* \begin{pmatrix} \tilde{B}\ \mathrm{F^t M^{-1} F}^{\ -1} & 0 \\ 0^t & \frac{2B^2}{n-p-1} \end{pmatrix} \tag{6}$$

with $0^t = \overbrace{0,...0}^{p+1}$ . The point estimation of the $\phi_k$ function(maximum-likelihood and minimum variance unbiased estimation) was developed in Gutierrez et. al.(2001) for h =2 . The mean and mode functions and their conditional versions can be written in the form $\exp\ \mu_k\ t,s\ + \phi\sigma_k^2\ t,s$ with mean, conditional mean, mode and conditional mode as in Table 1 being the problem of building confidence bands for them solved in Gutierrez et al.( 2003).

## 4. Fisher Information Matrix

The Fisher information is a way of measuring the amount of information that an observable random variable *X* carries about an unknown parameter $\phi$ upon which the probability of *X* depends. The probability function for *X*, which is also the likelihood function for $\phi$ , is a function *f*(*X*; $\phi$ ); it is the probability mass (or probability density) of the random variable *X* conditional on the value of $\phi$ . The partial derivative with respect to $\phi$ of the natural logarithm of the likelihood function is called the score.

The Fisher information matrix is determined by first calculating the following:

$$\frac{\partial \mathrm{In\ L}}{\partial \phi} = \tilde{B}^{-1}\ \mathrm{In\ X - F}\phi^{\ t}\ \mathrm{M^{-1}F}$$

$$\frac{\partial^2 \mathrm{In\ L}}{\partial \phi^2} = \tilde{B}^{-1}\mathrm{F}^t\mathrm{M^{-1}F}$$

$$\frac{\partial^2 \mathrm{In\ L}}{\partial \phi \partial \tilde{B}} = \tilde{B}^{-2}\ \mathrm{In\ X - F}\phi^{\ t}\ \mathrm{M^{-1}F}$$

$$\frac{\partial \mathrm{In\ L}}{\partial \phi} = \frac{n}{2\tilde{B}} + \frac{1}{2\tilde{B}^2}\ \mathrm{In\ X - F}\phi^{\ t}\ \mathrm{M^{-1}F}\ \ \mathrm{In\ X - F}\phi$$

$$\frac{\partial^2 \mathrm{In\ L}}{\partial \tilde{B}^2} = \frac{n}{2\tilde{B}^2} - \frac{n}{\tilde{B}^3}\ \mathrm{In\ X - F}\phi^{\ t}\ \mathrm{M^{-1}F}\ \ \mathrm{In\ X - F}\phi$$

and

$$E\left[\frac{\partial^2 \text{In L}}{\partial \phi^2}\right] = -\tilde{B}^{-1}\text{F}^t\text{M}^{-1}\text{F}\,,\; E\left[\frac{\partial^2 \text{In L}}{\partial \phi \partial \tilde{B}}\right] = 0$$

$$E\left[\frac{\partial^2 \text{In L}}{\partial \tilde{B}^2}\right] = \frac{n}{2\tilde{B}^2} - \frac{2\ n+p-1}{2\tilde{B}^2} = \frac{2n+2p-2}{2\tilde{B}^2}$$

Therefore, the Fisher information matrix is

$$\text{I} = \begin{pmatrix} \tilde{B}^{-1}\text{F}^t\text{M}^{-1}\text{F} & 0 \\ 0^t & \frac{2n+2p-2}{2\tilde{B}^2} \end{pmatrix} \quad (7)$$

**5. Hypotheses Testing**

In order to test the hypothesis, the vector $\phi = \ \phi_0, \phi_1, ..., \phi_p\ ^t$ is split as follows (Skiadas, 2007; Anderson, 2003):

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

where $\phi_1$ is $p_1 \times 1$ and $\phi_2$ is $p_2 \times 1$, with $p_1 + p_2 = p+1$. The hypothesis of interest is

$$H_0 : \phi_1 = \overline{\phi}_1$$
$$H_1 : \phi_1 \neq \overline{\phi}_1$$

where $\overline{\phi}_1$ is $p_1 \times 1$ fixed vector. The total region and the region associated with the null hypothesis are, respectively,

$$\Omega = \ \phi, \tilde{B}\ : \tilde{B} > 0\ \subset \text{R}^{p+2}$$

$$\omega = \ \phi, \tilde{B}\ : \phi_1 = \tilde{\phi}_1 : B > 0\ \subset \text{R}^{p_2+1}$$

Under these hypotheses,

$$\max_{\Omega} L\ \ \text{x};\phi,\tilde{B}\ = \ 2\pi\ ^{n/2}\ \tilde{B}^*_{\Omega}\ ^{-n/2} |\,\text{M}|^{1/2} \prod_{i=1}^{n} x_i^{-1} \exp\left\{-\frac{n}{2}\right\}$$

$$\max_{\omega} L\ \ \text{x};\phi,\tilde{B}\ = \ 2\pi\ ^{n/2}\ \tilde{B}^*_{\omega}\ ^{-n/2} |\,\text{M}|^{-1/2} \prod_{i=1}^{n} x_i^{-1} \exp\left\{-\frac{n}{2}\right\}$$

and the likelihood ratio statistic for testing $H_0$ is

$$\Lambda = \frac{\max_{\omega} L}{\max_{\Omega} L} = \left[\frac{\tilde{B}^*_{\omega}}{\tilde{B}^*_{\Omega}}\right]^{-\frac{n}{2}} = \left[\frac{\tilde{B}^*_{\omega}}{\tilde{B}^*_{\Omega}}\right]^{\frac{n}{2}}$$

For obtaining the distridution of this statistic, let us denote $\text{A}=\text{F}^t\text{M}^{-1}\text{F}$ and $\text{C}=\text{F}^t\text{M}^{-1}\text{In X}$ and consider the following partitions:

$$\text{A}=\begin{pmatrix} \text{A}_{11} & \text{A}_{21} \\ \text{A}_{21} & \text{A}_{22} \end{pmatrix} = \begin{pmatrix} \text{F}_1^t \\ \text{F}_2^t \end{pmatrix} \text{M}^{-1}\ \ \text{F}_1 \,|\, \text{F}_2\ = \begin{pmatrix} \text{F}_1^t\text{M}^{-1}\text{F}_1 & \text{F}_1^t\text{M}^{-1}\text{F}_2 \\ \text{F}_2^t\text{M}^{-1}\text{F}_1 & \text{F}_2^t\text{M}^{-1}\text{F}_2 \end{pmatrix}$$

$$\text{C}=\text{A}\phi^*_{\Omega} \Rightarrow\ \ \text{C}_1 \,|\, \text{C}_2\ = \begin{pmatrix} \text{A}_{11} & \text{A}_{21} \\ \text{A}_{21} & \text{A}_{22} \end{pmatrix} \begin{pmatrix} \phi^*_{1\Omega} \\ \phi^*_{2\Omega} \end{pmatrix} \Rightarrow \begin{matrix} \text{C}_1 = \text{A}_{11}\phi^*_{1\Omega} + \text{A}_{12}\phi^*_{2\Omega} \\ \text{C}_2 = \text{A}_{12}\phi^*_{1\Omega} + \text{A}_{22}\phi^*_{2\Omega} \end{matrix}$$

where $\text{A}_{11}$ is $p_1 \times p_1$ and $\text{C}_1$ is $p_1 \times 1$. Using the last two expression, the following is obtained:

$$\phi^*_{2\Omega} = \text{A}_{22}^{-1}\text{C}_2 - \text{A}_{22}^{-1}\text{A}_{21}\phi^*_{1\Omega} = \text{A}_{22}^{-1}\ \ \text{C}_2 - \text{A}_{11}\phi^*_{1\Omega}$$

Subtracting $\text{F}_2\ \ \phi^*_{2w} - \phi_2$ in both sides of this equation, taking into account that

$$\ln x - F\phi - F_2\left(\phi^*_{2w} - \phi_2\right) = \ln x - F_1\phi_1 - F_2\phi_2 - F_2\left(\phi^*_{2w} - \phi_2\right)$$
$$= \ln x - F_1\overline{\phi}_1 - F_1\phi^*_{2w}$$

we obtain

$$\ln x - F_1\overline{\phi}_1 - F_2\phi^*_{2w} = \ln x - F\phi^*_\Omega + \left(F_1 - F_2A^{-1}_{22}A_{21}\right)\left(\phi^*_{1\Omega} - \overline{\phi}_1\right)$$

Since $\left(\ln x - F_1\phi^*_{2w}\right)^t M^{-1}F = 0$, it is clear that

$$\left(\ln x - F_1\phi^*_{2w}\right)^t M^{-1}F - F_2A^{-1}_{22}A_{21} = 0$$

Using the previous notation,

$$\left(F_1 - F_2A^{-1}_{22}A_{21}\right)^t M^{-1}\left(F_1 - F_2A^{-1}_{22}A_{21}\right) = A_{12}A^{-1}_{22}A_{21} = A_{11,2}$$

It can be established that

$$n\tilde{B}^*_w = \left(\ln x - F_1\overline{\phi}_1 - F_2\phi^*_{2w}\right)^t M^{-1}\left(\ln x - F_1\overline{\phi}_1 - F_2\phi^*_{2w}\right)$$
$$= n\tilde{B}^*_\Omega + \left(\phi^*_{1\Omega} - \overline{\phi}_1\right)^t A_{11,2}\left(\phi^*_{1\Omega} - \overline{\phi}_1\right)$$

The likelihood ratio statistics can now be written as

$$\Lambda = \left(n\tilde{B}^*_\Omega\left(n\tilde{B}^*_\Omega + \left(\phi^*_{1\Omega} - \overline{\phi}_1\right)^t A_{11,2}\left(\phi^*_{1\Omega} - \overline{\phi}_1\right)\right)^{-1}\right)^{\frac{n}{2}}$$

where $n\tilde{B}^*_\Omega + \left(\phi^*_{1\Omega} - \overline{\phi}_1\right)^t A_{11,2}\left(\phi^*_{1\Omega} - \overline{\phi}_1\right)$ is $W_1\left(p_1, \tilde{B}\right)$ and distributes independently of $n\tilde{B}^*_\Omega$ (Anderson, 2003). This means that the distribution of $\Lambda^{2/n}$ is the same as $U\left(U+V\right)^{-1}$ where $U$ and $V$ are independent random variables with distribution given by $W_1\left(n-p-1, \tilde{B}\right)$ and $W_2\left(p_1, \tilde{B}\right)$ respectively.

**6. Simulation Studies**

The stochastic differential equation $dx(t) = \mu\left(X(t) - \phi\right)dt + \sigma\left(X(t) - \phi\right)dw(t)$, where W(t ) represents the Wiener process with independent increments W(t ) - W(s) distributed according to $N\left(0, t-s\right)$ for t > 0, has a single continuous solution in the interval [$t_0$ , T]. This corresponds to the parameter of the lognormal diffusion process, the explicit expression of which can be obtained by means of Itô's formula, applied to the transform $\ln\left(X(t) - \phi\right)$, and which has the following form

$$X(t) = \phi + \left(x_0 - \phi\right)\exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)\left(t - t_0\right) + \sigma\left(W(t) - W(t_0)\right)\right\}$$

From this explicit solution, the simulated trajectories of the process can be obtained by discretizing the time interval [$t_0$, T], with the initial condition W($t_0$) = 0. The Wiener process is obtained as the sum of the distributions N(*0, h*), where h = $t_i$ $h = t_i - t_{i-1}$.

From this simulated process sample, the parameters can be estimated by ML, first using the Newton–Raphson (NR) nonlinear approach to approximate the value of $\phi$. Secondly, the problems that occur in estimating the parameters of the lognormal diffusion process are discussed. The SA optimization to the estimation of the parameters is used in order to perform a compression of the range of values over which the conditioned log-likelihood function must be maximized to find $\phi$. The parameters of the process are estimated by applying the method to the simulated data set described previously, which enable the effectiveness of the method to be tested.

Table 2 shows the values used in the simulation and the results obtained by estimating the parameters, using the methods described above, implemented using the mathematical packages by considering h = 1, n = 30 and an initial value $x_o$ = 1.12149. These results clearly show that the SA algorithm was a good estimation method and that it enabled the elimination of many of the difficulties encountered with ML estimation.

 

## 7. Empirical Results

The LNDP is applied to the data of total natural-petroleum products consumption in Ghana from 1999 to 2010. These data were provided by the Ministry of Economic Planning. Data of the above time series were used to estimate the parameters of the process using the methods described in Section 3. Gutierrez et al. (1999; 2005; 2006) proposed a methodology for building a theoretical model of lognormal diffusion process with exogenous factors that fit the data, that is, a method for searching for the h function.

The goodness-of-fit to the data was one criterion to compare various models for petroleum consumption in Ghana. The statistical results from the models such as $R^2$, MSE, MAPE, MAD and d values were calculated. The performance of the SGIDP for the forecasting period using the trend and conditional trend function is illustrated in Figure 1. Finally, in order to evaluate the results obtained using the SGIDP in studying the data series, the model was compared with two alternative models; the first being the stochastic logistic innovation process and the second is the stochastic lognormal innovation process (Skiadas and Giovani, 1997). A Matlab program was implemented to carry out the calculations required for this study. A Matlab program was implemented to carry out the calculations required for this study. The methodology is summarised as follows:

i. Use the first 50 data set in the series of observations to estimate the parameters of the model, using expressions (5) and (6). Then, determine the corresponding confidence intervals using equations (7) and (8).

ii. For the years 2000, 2001 and 2002, predict the corresponding values for electricity consumption in Morocco using the estimated trend function (ETF) and the estimated conditional trend function (ECTF), obtained by replacing the parameters with their estimators in expressions (3) and (4), and compare the results with the corresponding observed data for the same years.

The data from 1999 to 2010 are used to make forecasts of the future values of the process, with the trend and conditional trend functions given by expressions (2) and (3) and the confidence interval (given a 95%) in the expressions (6) and (7). The results are summarized in Table 3. Comparing the parameter estimation results for the demand for oil in Ghana, it can be seen that the maximum energy consumption level (*F*) of the process resulting from the stochastic model is larger than that resulting from the deterministic model. The estimators of the stochastic model seem more reasonable since the forecasting values of the deterministic model underestimate the real values. The approximate distribution function and cumulative distribution function for a random point are also provided. These distribution functions are not symmetric due to the nonlinearity of the stochastic model and are in accordance with the assumption of a multiplicative noise. As it is shown, the model behaves well since in both cases the real data are included in the lower and upper limits.

In the actual situation, it is suppose that they are not additional information but only values $x_1,\ldots,x_n$ of the endogenous variable in times $t_1,\ldots,t_n$. Suppose $P\left[X\ t_1\ = x_1\right]=1$, it is known that $\text{In}\left(\frac{E\left[X\ t\ \right]}{x_1}\right)=\int_{t_1}^{t} h\ s\ ds = H\ t$ and the values $f_i = \text{In}\left(\frac{x_i}{x_1}\right)$, $i=1,\ldots,n$ **are** considered as approximation to $H\ t_i$. So, with these values we fit $k+1$ - degree polynomial $P\ t\ =\sum_i^{k+1} a_i\phi_i\ t$ and we can approach the lognormal diffusion process $X\ t\ : t_0 \le t \le T$ by the lognormal diffusion process with polynomial exogenous factors $X^k\ t\ : t_0 \le t \le T$ with infinitesimal moments as

$$B_1^k\ x,t\ = x\left[\sum_{j=0}^{k}\phi_j^k P_j^k\ t\ \right] \text{ and } B_2^k\ x,t\ =\sigma_k^2 x^2$$

taking, in this case, $P_1^k\ t\ = a_j + 1\phi_j^{'} + 1\ t$, $j=1,\ldots,k$ and $P_0^k\ t\ = 1$. It is assumed that $\sum_{j=0}^{k}\phi_j^k P_j^k\ t$ has more than one factors, because it is possible that after a posterior study and analysis. Some values of $P_j^k$ may not be relevant.

Considering, for example, the function $f(t) = t^{-4}$, the values of the corresponding estimators are: $\hat{\phi} = 0.0209$, $\hat{h} = 0.0279$ and $\hat{B} = -0.2109$, with confidence intervals (0.015215;0.098076) and (1.011172;3.812135).10−4. Table 4 summarises the prediction results, that is, the observed data, the values predicted by ETF and ECTF and the lower and upper limits of these functions,

## 7. Conclusions

The main objective of this paper was the application of maximum likelihood ratio tests in lognormal diffusions model. The model described an innovation diffusion process considering at the same time disturbances coming from the environment of the system. Finally, the model was applied to energy consumption data in Ghana and showed sufficiently good results.

Considering a lognormal diffusion model, in this paper we have calculated the distribution of the MLEs of the drift and diffusion coefficients, the Fisher information matrix, and solved some likelihood ratio tests for hypotheses on the parameters weighting the relative effect of exogenous factors affecting the drift. The results obtained are important for real applications; in particular, for prediction and conditional simulation.

The endogenous consumption pattern in Ghana, in absolute terms, presents a clear upward trend. Between 1973 and 2010, the consumption rose from 763 to 12292 barrels (thousand metric tons of oil equivalent), while between 1990 and 2010, from 4531 to 12292 metric tons (an increase of 171.3%). With respect to the total consumption of primary energy derived from natural gas, the increase between 1990 and 2010 was even greater, at 204.46%. Finally, the separation, within total demand for petroleum products (final energy), of domestic-commercial use from industrial use (including electricity generation and cogeneration), reveals values of 18% and 82%, respectively.

The energy market in Ghana has been characterized in recent decades by very important quantitative and structural changes, especially concerning natural petroleum products as a source of energy. Moreover, this has taken place in a context of an expanding phase of the economic cycle and significant social changes. The energy market in Ghana has been characterized in recent decades by very important quantitative and structural changes, especially concerning petroleum as a source of energy. Moreover, this has taken place in a context of an expanding phase of the economic cycle and significant social changes.

## References

Anderson, T.W. (2003): *An Introduction to Multivariate Statistical Analysis(*3rd edition). Wiley & Sons, New Jersey.

Gutiérrez, R., Gutiérrez-Sánchez, R,. Nafidi, A and Ramos, E.(2007): Three-parameter stochastic lognormal diffusion model: statistical computation and simulating annealing – application to real case. *Journal of Statistical Computation and Simulation.* Vol. 2, pp. 1–14.

Gutiérrez ,R., Rom´an P., Romero, D., and Torres, F. (2003): Application of the univariate lognormal diffusion process with exogenous factors in forecasting. Cybernetics & Systems, Vol. 34, no. 8, pp. 709– 724.

Gutiérrez, R., Román, P. and 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*, Vol. 15, no. 4, pp. 325–332.

Gutiérrez, R., Román, P. and Torres, F., 2001, Inference on some parametric functions in the univariate lognormal diffusion process with exogenous factors. *Test*, Vol. 10, no. 2, pp. 357–373.

Gutiérrez, R., González, A. and Torres, F., 1997, Estimation in multivariate lognormal diffusion process with exogenous factors. *Applied Statistics (JRSC)*, Vol. 46, no. 1, pp. 140–146.

Kloeden, P., Platen E., Schurz H. , Sorensen H. , (1999): On effects of discretization on estimators of drift parameters for diffusion processes, *Journal of Applied Probability.*Vol. 33, pp. 1061–1076.

Prakasa B.L., (1999): S*tatistical Inference for Diffusion Type Processes*, Arnold, London and University Press, New York,

Ricciardi, L. M. On the transformation of diffusion processes into the Wiener processes.
*Journal of Mathematics Analysis and Appication. Vol. 54, pp.*185–99.

 

Skiadas, C. H, Giovani, A.N(1997): A stochastic Bass innovation diffusion model for studying the growth of electricity consumption in Greece. *Applied Stochastic Models Data and Analysis. Vol. 13, pp.85–101.*

Skiadas, C. H..; (2007): *Recent Advances in Stochastic Modeling and Data Analysis*. Chania, Greece

Nualart, l. D. (1983): Stochastic Processes and their Applications. *Recent Advances in Stochastic Modeling and Data Analysis* : Chania, Greece,

**Table 1: Gutierrez point estimation of the $\phi_k$ function**

| | **Mean** | **Conditional Mean** | **Mode** | **Conditional Mode** |
|---|---|---|---|---|
| $\mu_k \; t,s$ | $\text{In}x_0 + \bar{\mu}_k' \; t \; a_k$ | $\text{In}x_s + \bar{\mu}_k' \; t,s \; a_k$ | $\text{In}x_0 + \bar{\mu}_k' \; t \; a_k$ | $\text{In}x_s + \bar{\mu}_k' \; t,s \; a_k$ |
| $\sigma_k^2 \; t,s$ | $t - t_0 \; \sigma_k^2$ | $t - s \; \sigma_k^2$ | $t - t_0 \; \sigma_k^2$ | $t - s \; \sigma_k^2$ |
| $\phi$ | $1/2$ | $1/2$ | $-1$ | $-1$ |

**Table 2: Simulation and Estimation of the Parameters**

| | $\phi$ | $\mu$ | $\sigma$ |
|---|---|---|---|
| Simulation | 1 | 0.32 | 0.00024 |
| Estimation NR | 2.00005 | 0.400009 | 0.00013 |
| Estimation SA | 2.00475 | 0.432842 | 0.00086 |

**Table 3: Forecasting based on ETF and ECTF**

| Times | Data | EET | LL-ETF | UL-ETF | ECTF | LL- ECTF | UL- ECTF |
|---|---|---|---|---|---|---|---|
| 2010 | 15.275 | 15.478 | 14.776 | 16.578 | 14.274 | 14.377 | 15.985 |
| 2011 | 17.446 | 17.847 | 17.346 | 16.248 | 15.762 | 16.978 | 17.934 |
| 2012 | 18.274 | 18.845 | 18.274 | 17.679 | 18.978 | 18.367 | 18.709 |

**Table 4: Confidence intervals for parameter estimates**

| | Forecasting period | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 80% Lower Limit | 13.34 | 13.55 | 14.83 | 15.17 | 15.54 |
| Real Values | 24.65 | 25.37 | 26.37 | 29.39 | 29.99 |
| 80% Upper Limit | 325.50 | 44.46 | 45.23 | 46.75 | 46.90 |

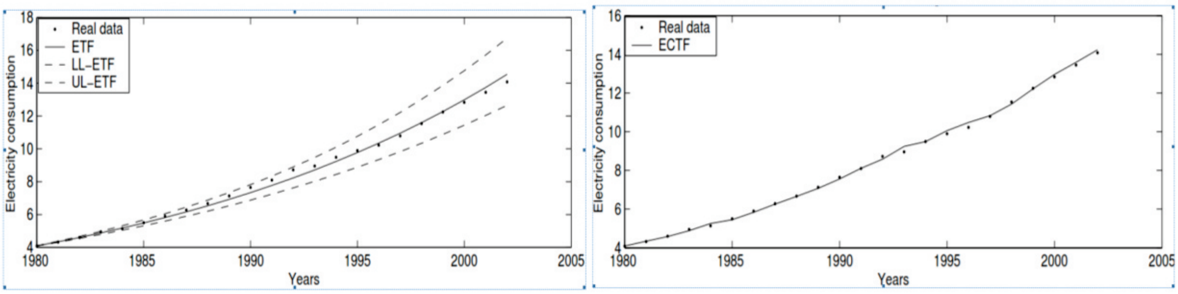


**Figure 1: Fits and predictions made using the ETF and the ECTF**