APPLIED STOCHASTIC MODELS AND DATA ANALYSIS, VOL. 11, 59–75 (1995)

ANALYTIC SOLUTION AND ESTIMATION OF PARAMETERS ON A STOCHASTIC EXPONENTIAL MODEL FOR A TECHNOLOGICAL DIFFUSION PROCESS

A. KATSAMAKI AND C. H. SKIADAS
Technical University of Crete, Department of Production Engineering and Management, 73100 Chania, Crete, Greece

SUMMARY

In this paper we examine the behaviour of a stochastic model that describes a technological diffusion process (continuously increasing process). Furthermore we obtain a solution for the proposed model through the estimation of the volatility using three different approximations. The adjustment of real data to the final stochastic model confirms its ability of describing and forecasting real cases.

KEY WORDS diffusion process; stochastic models; volatility

1. INTRODUCTION

In 1798 an anonymously published article became one of the most interesting works of the last centuries. Its title was 'An Essay on the Principle of Population as it Affects the Improvement of Society with Remarks on the Speculations of Mr. Godwin, M. Condorcet and Other Writers'. The writer was Thomas Malthus, a priest of the English Church. The first edition was published in a small number of copies since many disagreed with its content. But very soon the Essay of Malthus became the centre of social interest since many accepted its basic conclusion, which was that population increase obeys an arithmetic evolution when there are no restrictions.

Regardless of what we have considered correct or not, we acknowledge the fact that Malthus was the first to use the exponential distribution as a model of population growth. After him many others followed his ideas, such as Ludwig von Bertalanffy,[1] Wind *et al.*,[2] Mahajan and Muller,[3] Mahajan and Peterson[4] and Skiadas.[5]

The results we see in all cases where the exponential model has been used are good, although we have to be very careful since there is no real system that increases for ever. There is always an upper bound. So the best solution in this case is to have as many data (external and internal) as possible for the phenomenon studied and use in the final decision a method that takes into consideration all these data, allows the use of human judgement and constructs models that contain the minimum possible error (both systematic and random).

In our case, if f_t is the fraction of potential adopters who have adopted an innovation until time t, then the innovation growth model is described by an S-shaped f–t curve. This curve is always convex in the beginning, concave near the end and between these two situations has a point of inflection.

CCC 8755-0024/95/010059-17

The area we are going to study here is the convex area (before the inflection point) where the value of f_t is continuously increasing.

The paper is organized in the following way.

First we formulate a stochastic model for the area of the S-shaped f–t curve we are going to study.

Secondly, we estimate the values of the parameters of the proposed stochastic model using time series data by nonlinear or approximative estimation procedures.

Thirdly, we examine the prediction ability of the proposed model through the implementation of real data, using one estimate or a confidence interval for the value of the variable f_t at each time. Finally, we conclude with brief remarks.

2. THE STOCHASTIC MODEL: FORMULATION AND SOLUTION

One of the major problems in management nowadays is the quantification of the uncertainty regarding the future (Gihman,[6] Ito,[7] Bogolubov and Krylov[8]). A possible approach that has been proposed through many studies in the recent years is the use of stochastic models (Bretschneider and Bozeman,[9] Eliashberg, Tapiero and Wind,[10] El-Karoui,[11] Malliaris and Brock,[12] Oksendal[13] and Raman[14]). So from now on we will examine whether the stochastic model of the general form:

$$\mathrm{d}f_t = x(t, f_t)\,\mathrm{d}t + g(t, f_t)\,\mathrm{d}w_t \qquad t \geqslant t_0 \tag{1}$$

defined on the probability space (gkrW, F, P) is appropriate or not for the description of real cases of increasing processes.

First we can write the general model (1) for the case we study (technological diffusion process) in two possible forms:

$$\mathrm{d}f_t = \mu_t\,\mathrm{d}t + c_t\,\mathrm{d}w_t \tag{2}$$

$$\mathrm{d}f_t = \mu_t f_t\,\mathrm{d}t + c_t f_t\,\mathrm{d}w_t \tag{3}$$

where μ_t is a quantity that expresses the total influence (external and internal) on the whole population, and c_t is the coefficient of noise, which we will examine with more details in the next sections.

Model (2) can easily be solved. In that case the result we obtain is

$$f_t = f_0 + \int \mu_t\,\mathrm{d}t + \int c_t\,\mathrm{d}\omega_t$$

From now on, we will study model (2) no further. The model on which we are going to focus our interest is model (3).

We consider the noise for the diffusion process we are studying, as a one-dimensional variable of unknown behaviour (we do not know its value at every moment), but of known distribution function (Tanaka[15]). We will not study the multidimensional aspect because it does not offer anything more to the initial approximation of a diffusion process that we are trying to construct in this study.

A first approximation for the model (3) could be the form

$$\frac{\mathrm{d}f_t}{\mathrm{d}t} = \mu_t f_t + \text{noise}$$

where the noise has a simple form and is additive to the process.

The above equation is a model that could describe quite efficiently processes with constant measuring during their evolution or processes where the value of $\mathrm{d}f$ at each moment t depends on the value of f at that particular moment, and the noise is independent of the value of f at each moment t. However, it is obvious for the diffusion processes we are trying to describe (processes we meet in marketing, economy, etc), that we need a more complicated form than this equation.

The question is *how* we can count a process that is increased without the use of a constant meter.

For that reason we use the changeable fraction

$$\frac{\dot{f}}{f} \quad \text{or} \quad \frac{\Delta f}{f\,\Delta t} \rightarrow \frac{\mathrm{d}f}{f\,\mathrm{d}t} = \frac{\mathrm{d}\,(\ln f)}{\mathrm{d}t} = (\ln f)'$$

which expresses all the relative alterations (increases in our case) at each time t.

So from now on in the present study we consider that

$$\frac{\dot{f}}{f} = \mu + w(t)$$

where μ is a mean value function that counts locally the mean velocity of the noise we are modelling and $w(t)$ is the noise of the system at every moment.

According to this equation, the process we examine receives random influences from its environment that are increasing continuously and analogically to the value of the process f_t, which is also an increasing function and has the form

$$\mathrm{d}f_t = \mu_t f_t\,\mathrm{d}t + c_t f_t\,\mathrm{d}w_t \tag{4}$$

If we assume that μ_t and c_t in the proposed model $\mathrm{d}f_t = \mu_t f_t\,\mathrm{d}t + c_t f_t\,\mathrm{d}w_t$ have constant values, using Ito's lemma we obtain (Gihman and Skorohod[16]):

$$\begin{aligned}
&\mathrm{d}(\ln f_t) = \frac{1}{f_t}\,\mu f_t\,\mathrm{d}t - \frac{1}{2}\frac{1}{f_t^2}\,c^2 f_t^2\,\mathrm{d}t + \frac{1}{f_t}\,c f_t\,\mathrm{d}w_t \Leftrightarrow \\
&\mathrm{d}(\ln f_t) = \left[\mu - \tfrac{1}{2}c^2\right]\mathrm{d}t + c\,\mathrm{d}w_t \Leftrightarrow \ln f_t = \ln f_0 + \left(\mu - \tfrac{1}{2}c^2\right)t + cw_t \Leftrightarrow \\
&f_t = f_0 \exp\left\{\left(\mu - \tfrac{1}{2}c^2\right)t + cw_t\right\}
\end{aligned} \tag{5}$$

From equation (5) it is obvious that

for $\mu > \frac{1}{2}c^2$ $\quad f_t \rightarrow \infty$ as $t \rightarrow \infty$

for $\mu < \frac{1}{2}c^2$ $\quad f_t \rightarrow \infty$ as $t \rightarrow \infty$

for $\mu = \frac{1}{2}c^2$ $\quad f_t$ takes randomly big and small values as $t \rightarrow \infty$.

3. ESTIMATION OF THE PARAMETERS IN THE MODEL

The application of any model involves estimating its parameters. The greater the number of parameters, the better the data fit will be. However, estimation in that case becomes more complex.

For the model we have already formulated (see equation (4)), the estimation of parameters μ_t and c_t includes the steps given in Sections 3.1 and 3.2 as follows.

3.1. Estimation of the coefficient μ_t

We consider that μ_t is the solution of the following generalized exponential differential equation:

$$\frac{d\mu_t}{dt} = a + b\mu_t^\delta \tag{6}$$

where a, b and δ are parameters.

The differential equation (6) always has the solution

$$\frac{d\mu_t}{dt} = a + b\mu_t^\delta \Leftrightarrow \int_{\mu_0}^{\mu_t} \frac{d\mu_t}{a + b\mu_t^\delta} = \int_0^t ds \Leftrightarrow \frac{1}{a} \int_{\mu_0}^{\mu_t} \frac{d\mu_t}{1 + \frac{b}{a}\mu_t^\delta} = t, \quad a \neq 0 \tag{7}$$

Assuming that

$$\frac{b}{a}\mu_t^\delta = m_t^\delta$$

$d\mu_t$ and μ_t can be replaced in equation (7) so that we have

$$\frac{1}{a}\int_{(a/b)^{1/\delta}m_0}^{(a/b)^{1/\delta}m_t} \frac{(a/b)^{1/\delta}}{1+m_t^\delta}\, dm_t = t, \quad b \neq 0 \quad \text{and} \quad ba > 0 \Leftrightarrow \frac{1}{a}(a/b)^{1/\delta} \int_{(a/b)^{1/\delta}m_0}^{(a/b)^{1/\delta}m_t} \frac{dm_t}{1+m_t^\delta} = t \tag{8}$$

Based on complex number theory, the indefinite integral $\int dm_t/(1 + m_t^\delta)$ has the solution

$$\int \frac{dm_t}{1+m_t^\delta} = -\frac{2}{\delta}\sum_{k=0}^{\delta/2-1} P_k \cos\frac{2k+1}{\delta}\pi + \frac{2}{\delta}\sum_{k=0}^{\delta/2-1} Q_k \sin\frac{2k+1}{\delta}\pi$$

for $\delta = 2k$, $k \in N^*$ and

$$\int \frac{dm_t}{1+m_t^\delta} = \frac{1}{\delta}\ln(1+m_t) - \frac{2}{\delta}\sum_{k=0}^{(\delta-3)/2} P_k \cos\frac{2k+1}{\delta}\pi + \frac{2}{\delta}\sum_{k=0}^{(\delta-3)/2} Q_k \sin\frac{2k+1}{\delta}\pi \tag{9}$$

for $\delta = 2k+1$, $k \in N^*$ where

$$P_k = \frac{1}{2}\ln\left(m_t^2 - 2m_t\cos\frac{2k+1}{\delta}\pi + 1\right)$$

and

$$Q_k = \arctan\frac{m_t \sin\frac{2k+1}{\delta}\pi}{1 - m_t\cos\frac{2k+1}{\delta}\pi} = \arctan\frac{m_t - \cos\frac{2k+1}{\delta}\pi}{\sin\frac{2k+1}{\delta}\pi}$$

So the equation that estimates m_t which ends to μ_t is

For $\delta = 2k,\ k \in N^*$:

$$\frac{d\mu_t}{dt} = a + b\mu_t^\delta \Leftrightarrow \frac{1}{a}(a/b)^{1/\delta}\int_{(a/b)^{1/\delta}m_0}^{(a/b)^{1/\delta}m_t}\frac{dm_t}{1+m_t^\delta} = t$$

$$\frac{1}{a}(a/b)^{1/\delta}\left[-\frac{2}{\delta}\sum_{k=0}^{\delta/2-1}P_k\cos\frac{2k+1}{\delta}\pi + \frac{2}{\delta}\sum_{k=0}^{\delta/2-1}Q_k\sin\frac{2k+1}{\delta}\pi\Big|_{\mu_0}^{\mu_t}\right] = t \tag{10}$$

For $\delta = 2k+1,\ k \in N^*$:

$$\frac{d\mu_t}{dt} = a + b\mu_t^\delta \Leftrightarrow \frac{1}{a}(a/b)^{1/\delta}\int_{(a/b)^{1/\delta}m_0}^{(a/b)^{1/\delta}m_t}\frac{dm_t}{1+m_t^\delta} = t$$

$$\frac{1}{a}(a/b)^{1/\delta}\left[\frac{1}{\delta}\ln(1+(b/a)^{1/\delta})\mu_t) - \frac{2}{\delta}\sum_{k=0}^{(\delta-3)/2}P_k\cos\frac{2k+1}{\delta}\pi + \frac{2}{\delta}\sum_{k=0}^{(\delta-3)/2}Q_k\sin\frac{2k+1}{\delta}\pi\Big|_{\mu_0}^{\mu_t}\right] = t \tag{11}$$

where

$$P_k = \frac{1}{2}\ln\left((b/a)^{2/\delta}\mu_t^2 - 2(b/a)^{1/\delta}\mu_t\cos\frac{2k+1}{\delta}\pi + 1\right)$$

and

$$Q_k = \arctan\frac{(b/a)^{1/\delta}\mu_t - \cos\frac{2k+1}{\delta}\pi}{\sin\frac{2k+1}{\delta}\pi}$$

under the considerations $ab > 0$ and $\delta \neq 0, 1$.

As we can see, the final equations for μ_t (equations (10) and (11)), are so complicated that they are quite impossible to apply in real situations. Secondly, trying to estimate the final solutions we assumed some restrictions for the parameters a, b and δ that reduce the usefulness of the final results. So we decided to study the initial model (equation (3)), for some specific values of a, b and δ.

Case A: $a = 0$

For $\delta = 0$:

$$\frac{d\mu_t}{dt} = b \Leftrightarrow \mu_t = \mu_0 + b(t - t_0)$$

For $\delta = 1$:

$$\frac{d\mu_t}{dt} = b\mu_t \Leftrightarrow \mu_t = \mu_0\,e^{bt} \qquad \text{(exponential model)}$$

For $\delta = 2$:

$$\mu_t = \frac{\mu_0}{1 - b\mu_0 t}$$

For $\delta = 3$:

$$\mu_t^2 = \frac{\mu_0^2}{1 - 2b\mu_0^2 t}$$

$$\vdots$$

For $\delta = \delta$:

$$\mu_t^{\delta-1} = \frac{\mu_0^{\delta-1}}{1 - (\delta - 1)b\mu_0^{\delta-1}t} \Leftrightarrow \mu_t = \mu_0\left(\frac{1}{1 - (\delta - 1)b\mu_0^{\delta-1}t}\right)^{1/(\delta-1)}$$

So finally we have for $a = 0$:

For $\delta = 0$ $\mu_t = \mu_0 + b(t - t_0)$ (12)

For $\delta = 1$ $\mu_t = \mu_0 \, e^{bt}$ (13)

For $\delta \in N - \{0, 1\}$:

$$\mu_t = \mu_0\left(\frac{1}{1 - (\delta - 1)b\mu_0^{\delta-1}t}\right)^{1/(\delta-1)} \quad (14)$$

Case B: $b = 0$

$$\frac{d\mu_t}{dt} = a \Leftrightarrow \mu_t = \mu_0 + a(t - t_0) \quad (15)$$

Case C: $\delta = 0$

$$\frac{d\mu_t}{dt} = a + b \Leftrightarrow \mu_t = \mu_0 - (a + b)t_0 + (a + b)t \quad (16)$$

Case D: $\delta = 1$

$$\frac{d\mu_t}{dt} = a + b\mu_t \Leftrightarrow \mu_t = (\mu_0 + (a/b))e^{bt} - (a/b), \quad b \neq 0 \quad (17)$$

Case E: $\delta = 2$

$$\frac{d\mu_t}{dt} = a + b\mu_t^2 \Leftrightarrow \mu_t = \frac{\mu_0 + (a/b)^{1/2}\tan((ab)^{1/2}t)}{1 - (b/a)^{1/2}\mu_0 \tan((ab)^{1/2}t)}, \quad ab > 0 \quad (18)$$

Case F: $\delta = 3$

$$\frac{d\mu_t}{dt} = a + b\mu_t^3 \Leftrightarrow \frac{1}{3c^2}\ln|c+\mu_t| + \frac{4}{3\sqrt{3}c^2}\arctan\left(\frac{2\mu_t - c}{\sqrt{3}c}\right) - \frac{1}{6c^2}\ln(\mu_t^2 - c\mu_t + c^2) + \frac{\sqrt{3}}{9c^2}\arctan\left(\frac{2\mu_t - c}{\sqrt{3}c}\right) = bt + \lambda \quad (19)$$

Where $c = (a/b)^{1/3}$ and λ is the integration constant.

Case G: $\delta = 4$

$$\frac{d\mu_t}{dt} = a + b\mu_t^4 \Leftrightarrow \frac{1}{2}\ln(\mu_t^2 + \sqrt{2}c\mu_t + c^2) - \arctan\left(\frac{\sqrt{2}}{c}\mu_t + 1\right)\left(\frac{\sqrt{2}-1}{\sqrt{2}}\right) - \frac{1}{2}\ln(\mu_t^2 - \sqrt{2}c\mu_t + c^2) - \arctan\left(\frac{\sqrt{2}}{c}\mu_t - 1\right)\left(\frac{\sqrt{2}-1}{\sqrt{2}}\right) = (a^3b)^{1/4}t + \lambda \quad (20)$$

where $c = (a/b)^{1/4}$. Note that the calculations for $\delta = 3$ are presented in the Appendix. The analytical solution of the proposed deterministic model is available upon request to the authors.

3.2. Estimation of the coefficient c_t using three different methods of approximation

The three different ways we used to approximate coefficient c_t if we assume that it has always a positive value, are as follows:

1st method. According to Chesney and Elliot,[17] coefficient c_t is called the volatility coefficient. Suppose that the value of coefficient μ is already known, c_t is estimated using Ito's lemma as follows:

$$d\left(\frac{1}{f_t}\right) = -\frac{1}{f_t^2}df_t + \frac{1}{2}\left(\frac{2}{f_t^3}\right)(df_t)^2 = -\frac{df_t}{f_t^2} + \frac{c^2\,dt}{f_t} \Leftrightarrow f_t\,d\left(\frac{1}{f_t}\right) = -\frac{df_t}{f_t} + c^2\,dt$$

Substituting

$$df_t \cong f_t - f_{t-1} \quad \text{and} \quad d\left(\frac{1}{f_t}\right) \cong \frac{1}{f_t} - \frac{1}{f_{t-1}}$$

an approximate value for c_t is

$$f_t\left(\frac{1}{f_t} - \frac{1}{f_{t-1}}\right) = \frac{f_t - f_{t-1}}{f_t} + c_t^2 t \Leftrightarrow \frac{f_{t-1} - f_t}{f_{t-1}} + \frac{f_t - f_{t-1}}{f_t} = c_t^2 t$$

$$\Leftrightarrow (f_t - f_{t-1})\left[\frac{1}{f_t} - \frac{1}{f_{t-1}}\right] = c_t^2 t \Leftrightarrow -\frac{(f_t - f_{t-1})^2}{f_t f_{t-1}} = c_t^2 t$$

$$\Leftrightarrow c_t = \left| -\frac{f_t - f_{t-1}}{(f_t f_{t-1} t)^{1/2}} \right|$$

and the constant value of c_t when we apply the model

$$f_t = f_0 \exp\left\{\left(\mu - \tfrac{1}{2} c^2\right)t + cw_t\right\}$$

on a set of real data, length n, is

$$c = \sum_{i=1}^{n} \left| \frac{f_t - f_{t-1}}{(f_t f_{t-1} t)^{1/2}} \right| \tag{21}$$

2nd method. Suppose again that the value of coefficient μ is already known. Then we have the following.

Using Ito's lemma

$$\mathrm{d}(\ln f_t) = \frac{\mathrm{d}f_t}{f_t} - \frac{1}{2}\frac{(\mathrm{d}f_t)^2}{f_t^2} \tag{22}$$

From the stochastic differential equation of the variable f_t we see that

$$\mathrm{d}f_t = \mu f_t\,\mathrm{d}t + c_t f_t\,\mathrm{d}w_t \Leftrightarrow \frac{\mathrm{d}f_t}{f_t} = \mu\,\mathrm{d}t + c_t\,\mathrm{d}w_t \tag{23}$$

and

$$(\mathrm{d}f_t)^2 = c_t^2 f_t^2\,\mathrm{d}t \tag{24}$$

Substituting (23) and (24) into (22) we get

$$\mathrm{d}(\ln f_t) = \mu\,\mathrm{d}t + c_t\,\mathrm{d}w_t - \tfrac{1}{2} c_t^2\,\mathrm{d}t \Leftrightarrow (\mathrm{d}(\ln f_t))^2 = c_t^2\,\mathrm{d}t$$

Considering that $\mathrm{d}(\ln f_t) \cong \ln(f_t) - \ln(f_{t-1})$, an approximate value for c_t is

$$c_t = \left| \frac{\ln f_t - \ln f_{t-1}}{t^{1/2}} \right|$$

So the constant value of c is obtained as in the first method by the equation

$$c = \sum_{i=1}^{n} \left| \frac{\ln f_t - \ln f_{t-1}}{t^{1/2}} \right| \tag{25}$$

3rd method. Assuming again that μ is already known from the stochastic differential equation of the variable f_t we get

$$\mathrm{d}f_t = \mu f_t\,\mathrm{d}t + c_t f_t\,\mathrm{d}w_t \Leftrightarrow \frac{\mathrm{d}f_t}{f_t} = \mu\,\mathrm{d}t + c_t\,\mathrm{d}w_t \Leftrightarrow \left(\frac{\mathrm{d}f_t}{f_t}\right)^2 = c_t^2\,\mathrm{d}t \Leftrightarrow c_t(\mathrm{d}t)^{1/2} = \frac{\mathrm{d}f_t}{f_t}$$

Considering that $\mathrm{d}f_t \cong f_t - f_{t-1}$ a third approximation for c_t is:

$$c = \left| \frac{f_t - f_{t-1}}{f_t t^{1/2}} \right|$$

So the third approximated constant value for c is

$$c = \sum_{i=1}^{n} \left| \frac{f_t - f_{t-1}}{f_t t^{1/2}} \right| \tag{26}$$

4. APPLICATION

Most cost data which are available for immediate use in a preliminary or predesign estimate are based on conditions at some time in the past. Because prices may change considerably with time due to changes in economic conditions, some method must be used for updating cost data applicable at a past date to costs that are representative of conditions at a later time. This can be done by the use of cost indexes.

A cost index is merely an index value for a given point in time showing the cost at that time relative to a certain basetime. If the cost at some time in the past is known, the equivalent cost at the present time can be determined by multiplying the original cost by the ratio of the present index value to the index value applicable when the original cost was obtained. Cost indexes can

Table I.

	Index				
Year	MS 1926 = 100	ENR 1913 = 100	Nelson 1946 = 100	Chem.En. 1957 = 100	STP 1957–59 = 100
1948	163	460	133	70	65
1949	161	477	140	71	66
1950	168	510	146	74	69
1951	180	543	157	80	74
1952	181	569	164	81	77
1953	183	600	174	85	81
1954	185	628	180	86	83
1955	191	660	184	88	87
1956	209	690	195	94	92
1957	225	724	206	99	98
1958	229	759	214	100	102
1959	235	797	222	102	104
1960	238	824	228	102	105
1961	237	847	233	102	106
1962	239	872	238	102	107
1963	239	901	244	102	109
1964	242	936	252	103	110
1965	245	971	261	104	112
1966	252	1021	273	107	116
1967	263	1070	288	110	119
1968	273	1165	304	114	124
1969	285	1272	329	119	133
1970	303	1418	365	126	144
1971	321	1620	406	132	160
1972	332	1670	429	137	172

be used to give a general estimate but no index can take into account all factors such as special technological advancements or local conditions. Many different types of cost indexes are published regularly. Some of these can be used for estimating equipment costs; others apply specifically to labour, construction, materials or other specialized fields.

The most common of these indexes in chemical engineering are the Marshall and Swift all-industry equipment indexes (MS), the Engineering News-Record construction index (ENR),

Table II.

Time series	Index μ	$c=\sum_{t=1}^{15}\left\|\frac{f_t-f_{t-1}}{(f_tf_{t-1}t)^{1/2}}\right\|$	$c=\sum_{t=1}^{15}\left\|\frac{\ln f_t-\ln f_{t-1}}{t^{1/2}}\right\|$	$c=\sum_{t=1}^{15}\left\|\frac{f_t-f_{t-1}}{f_t t^{1/2}}\right\|$	f_0
MS	0·0320	0·010875	0·010873	0·011165	1·546
ENR	0·0458	0·017034	0·017033	0·016615	4·515
Nelson	0·0413	0·015904	0·015903	0·016301	1·327
Chem.En.	0·0287	0·010522	0·010521	0·010788	0·702
STP	0·0381	0·013310	0·013309	0·013631	0·638

Table III.

	MS		ENR		Nel		CE		STP	
Time	Real value	Pred. value	Real value	Pred. value	Real value	Pred. value	Real value	Pred. value	Real value	Pred. value
1	1·63	1·60	4·60	4·73	1·33	1·38	0·70	0·72	0·65	0·66
2	1·61	1·65	4·77	4·95	1·40	1·44	0·71	0·74	0·66	0·69
3	1·68	1·70	5·10	5·18	1·46	1·50	0·74	0·77	0·69	0·72
4	1·80	1·76	5·43	5·42	1·57	1·57	0·80	0·79	0·74	0·74
5	1·81	1·82	5·69	5·68	1·64	1·63	0·81	0·81	0·77	0·77
6	1·83	1·87	6·00	5·94	1·74	1·70	0·85	0·84	0·81	0·80
7	1·85	1·93	6·28	6·22	1·80	1·77	0·86	0·86	0·83	0·83
8	1·91	2·00	6·60	6·51	1·84	1·85	0·88	0·88	0·87	0·87
9	2·09	2·06	6·90	6·82	1·95	1·92	0·94	0·91	0·92	0·90
10	2·25	2·13	7·24	7·16	2·06	2·09	0·99	0·93	0·98	0·93
11	2·29	2·20	7·59	7·47	2·14	2·09	1·00	0·96	1·02	0·97
12	2·35	2·27	7·97	7·82	2·22	2·18	1·02	0·99	1·04	1·01
13	2·38	2·34	8·24	8·19	2·28	2·27	1·02	1·02	1·05	1·05
14	2·37	2·42	8·47	8·57	2·33	2·37	1·02	1·05	1·06	1·09
15	2·39	2·50	8·72	8·97	2·38	2·47	1·02	1·08	1·07	1·13
16	2·39	2·58	9·01	9·40	2·44	2·57	1·02	1·11	1·09	1·18
17	2·42	2·66	9·36	9·84	2·52	2·68	1·03	1·15	1·10	1·22
18	2·45	2·75	9·71	10·26	2·61	2·79	1·04	1·17	1·12	1·27
19	2·52	2·84	10·21	10·78	2·73	2·91	1·07	1·21	1·16	1·32
20	2·63	2·93	10·70	11·29	2·88	3·03	1·10	1·24	1·19	1·37
21	2·73	3·03	11·65	11·82	3·04	3·16	1·14	1·28	1·24	1·42
22	2·85	3·12	12·72	12·37	3·29	3·29	1·19	1·32	1·33	1·48
23	3·03	3·23	14·18	12·95	3·65	3·43	1·26	1·36	1·44	1·53
24	3·21	3·33	16·20	13·56	4·06	3·58	1·32	1·34	1·60	1·59
25	3·32	3·44	16·70	14·19	4·29	3·73	1·37	1·44	1·72	1·66

the Nelson refinery construction index (Nelson), the Chemical Engineering plant cost index (Chem.En.) and the Materials and Labour cost index (STP).

Table I presents the annual mean values of these five indexes from 1948 to 1972. We are going to use these for our implementation. It is worth mentioning that all of them are used mostly in industry to estimate the cost of an investment. So it is obvious that their values are continuously increasing in time.

The stochastic model we used for this application was the specific form of the general model $df_t = \mu_t f_t \, dt + c_t f_t \, dw_t$ when μ_t and c_t have constant values (equation (5)).

Parameters μ and c were estimated as described before for 15-year time series data (1948–1962). Predictions for 10 years ahead on each case were made. For simplicity, all the time series have been divided by 100. Table II presents the results of the parameters estimation.

The results of the prediction are shown in Table III and Figures 1–5. For the predicted values of $f(t)$ we used the approximated values of c we estimated from the first method. Finally, the values of w_t are obtained using an appropriate random generator. In all the figures, we consider for $t = 1$ to 15 the real values (f_0 has been estimated using nonlinear regression analysis).

Here, through the illustration of the above diagrams it is evident that the yield of the proposed model is quite satisfactory. However, a weakness of this approach is that we predict what is going to happen in the future with one value at each moment. So we don't have the flexibility of predicting in case some external changes affect the value of f_t so its estimated value is useless. Therefore in order to improve the prediction ability of the proposed stochastic model even in

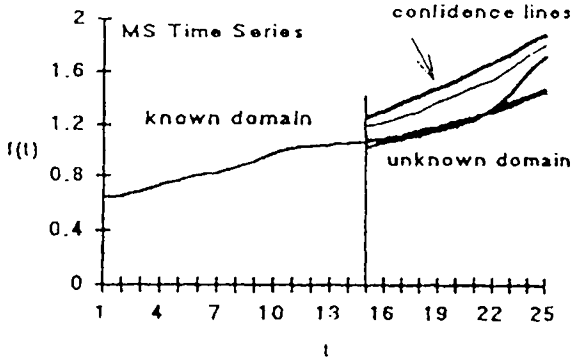


Figure 1

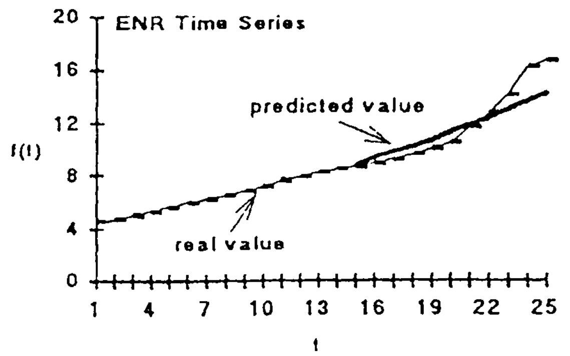


Figure 2

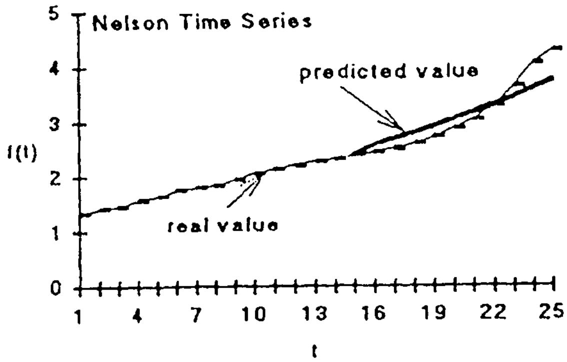


Figure 3

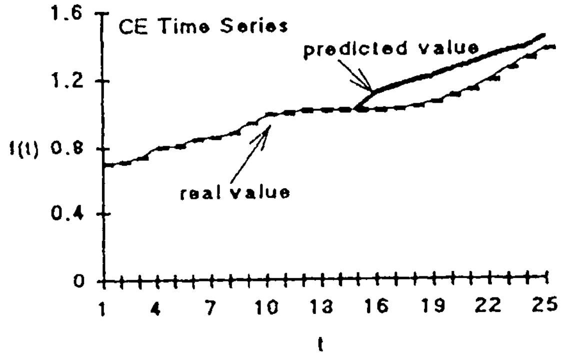


Figure 4

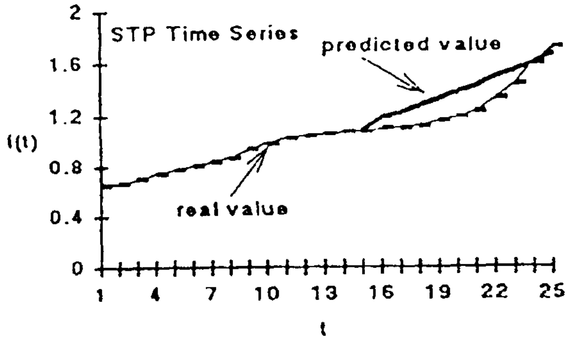


Figure 5

cases when unexpected events take place, we will study the prediction ability for the same five time series through confidence intervals. The procedure we followed to estimate the confidence interval of the variable f_t was

$$f_t = f_0 \exp\left[\left(\mu - \tfrac{1}{2} c^2\right)t\right]\exp[cw_t] \Leftrightarrow \exp[cw_t] = \frac{f_t}{f_0 \exp\left[\left(\mu - \frac{1}{2} c^2\right)t\right]} \Leftrightarrow cw_t$$

$$= \ln\left[\frac{f_t}{f_0 \exp\left[\left(\mu - \frac{1}{2} c^2\right)t\right]}\right] \Leftrightarrow cw_t = \ln f_t - \ln\left[f_0 \exp\left[\left(\mu - \tfrac{1}{2} c^2\right)t\right]\right] \Leftrightarrow cw_t = \ln f_t - \ln f_0 - \left(\mu - \frac{c^2}{2}\right)t$$

According to the properties of Wiener processes we know that $cw_t \sim N(0, c^2 t)$.

Therefore, an estimation for the variable $x = cw_t$ (where w_t follows the normal distribution $N(0, 1)$) is

$$Z = \frac{X - EX}{[\mathrm{Var}(X)]^{1/2}} = \frac{\ln f_t - \ln f_0 - \left(\mu - \frac{c^2}{2}\right)t - 0}{ct^{1/2}} \sim N(0, 1)$$

So the 95% confidence interval for the variable f_t is:

$$P\left[-\gamma \leqslant \frac{\ln f_t - \ln f_0 - \left(\mu - \frac{c^2}{2}\right)t}{ct^{1/2}} \leqslant \gamma\right] = 0{\cdot}95 \Leftrightarrow -1{\cdot}96 \leqslant \frac{\ln f_t - \ln f_0 - \left(\mu - \frac{c^2}{2}\right)t}{ct^{1/2}} \leqslant 1{\cdot}96$$

$$\Leftrightarrow -1{\cdot}96ct^{1/2} + \ln f_0 + \left(\mu - \frac{c^2}{2}\right)t \leqslant \ln f_t \leqslant 1{\cdot}96ct^{1/2} + \ln f_0 + \left(\mu - \frac{c^2}{2}\right)t$$

$$\Leftrightarrow \exp\left[\ln f_0 + \left(\mu - \frac{c^2}{2}\right)t - 1{\cdot}96ct^{1/2}\right] \leqslant f_t \leqslant \exp\left[\ln f_0 + \left(\mu - \frac{c^2}{2}\right)t + 1{\cdot}96ct^{1/2}\right]$$

The values of μ, c_i $(i = 1, 2, 3)$ and f_0 remain the same we estimated for the 15-year time series data (see Table II). In the following diagrams we can see whether the calculated confidence intervals are a satisfactory area or not for the predicted and the real values of f_t. The predicted values are related to the Brownian motion which is expressed by the term w_t. It is

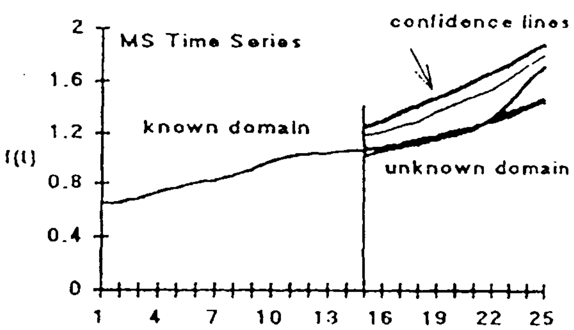


Figure 6

obvious that we can have innumerable values for f_t, one for each w_t. In each of the following diagrams (Figures 6 to 10), and for each set of data, are presented four of these possible values so as immediately to be able to see whether the confidence interval is a good tool for prediction or not.

Of course, the results we have got up to this point are only empirical and cannot prove with certainty the high contribution in prediction using confidence intervals.

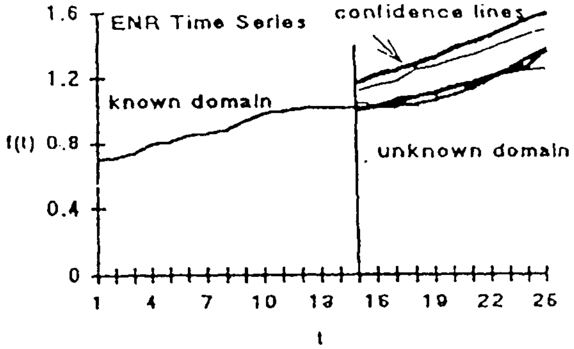


Figure 7.

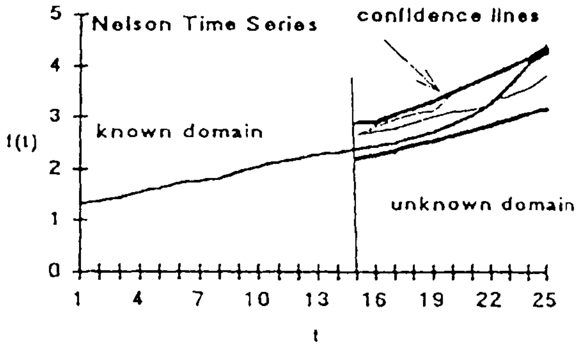


Figure 8.

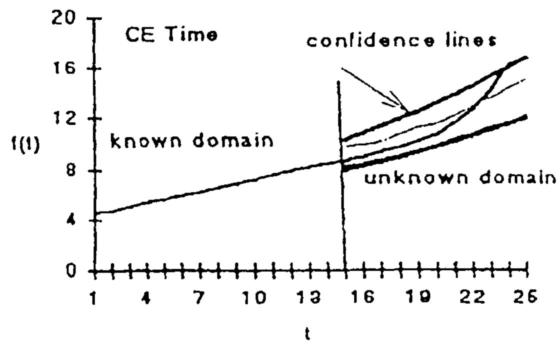


Figure 9.

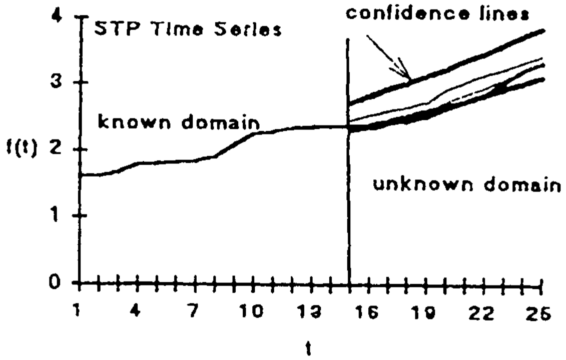
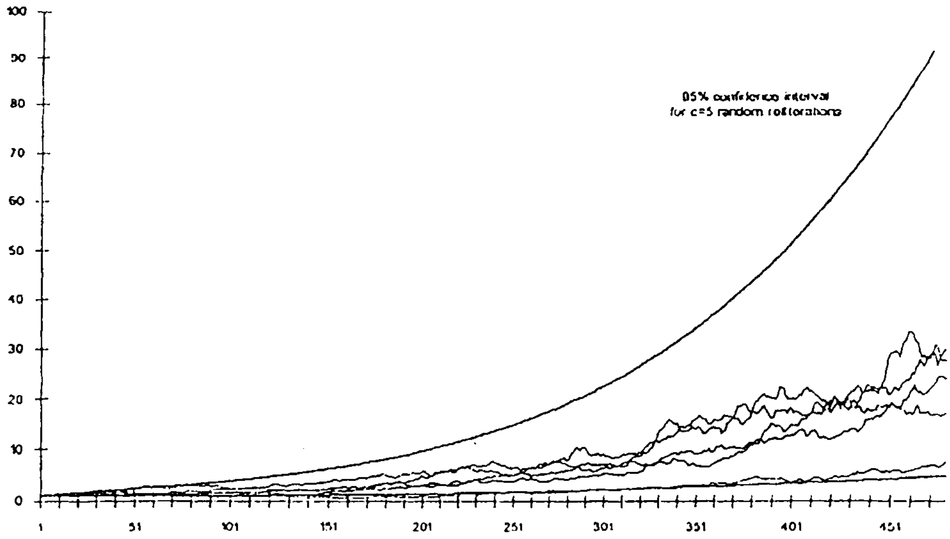


Figure 10.

In order to get a measure of success in predictions using confidence intervals, we performed a final experiment. For each of the five time series and each combination of the known parameters (μ, c_i, f_0), $i = 1, 2, 3$ for $t = 1$ to 720, we examined 2000 simulations of random time series using a computer program and calculated how many values where found out of the 95% confidence interval we had already estimated using the known parameter values. In all cases



sum<L/c = 15.360 sum<L+sum>U=170.6 REAL CONFIDENCE
sum>U/c = 18.760 INTERVAL FOR c=5
REITERATIONS(%) = 95.261

Sum < L is the total sum of points that have lower value than the lower bound of the confidence interval.
Sum > U is the total sum of points that exceed the upper bound of the confidence interval.
Note that we don't present on the diagram the first experiments we made because it is obvious that the graphical presentation of 2000 lines would not give us more information than we already have.

Figure 11

($5\times 3=15$ different cases) the results showed that the 95% confidence intervals contain the predicted values of f_t from 95·35 to 95·40%.

For a better understanding, we present Figure 11 showing graphically an implementation of the above experiment for $f_0=1$, $\mu=0{\cdot}009$, $c=0{\cdot}02$ and five random reiterations of the predicted values of f_t.

5. CONCLUDING REMARKS

A stochastic exponential model has been presented as a tool for the description of many continuously increasing processes. Application of the proposed model to five financial indexes showed its usefulness in practice. However, most effort in this study has been centred on the estimation of the unknown parameters of the model.

Based on the assumption that the external and internal influence on the diffusion process is independent of the whole system's noise, the influence of parameter μ_t was estimated as the result of a generalized exponential deterministic differential equation. Furthermore, the noise parameter c_t was estimated through the stochastic model with a known value for μ_t.

The results we obtained showed that more work in this area must be carried out. According to the characteristics and the dynamic behaviour of any diffusion process, we have to consider the changes of the system that can be represented by the model, through more study on the unknown parameters.

APPENDIX

The procedure followed to solve equation 19 was:

$$\frac{\mathrm{d}\mu_t}{\mathrm{d}t}=a+b\mu_t^3 \Leftrightarrow \int \frac{\mathrm{d}\mu_t}{a+b\mu_t^3}=t \Leftrightarrow \int \frac{\mathrm{d}\mu_t}{([a/b]^{1/3})^3+\mu_t^3}=bt$$

The denominator $([a/b]^{1/3})^3+\mu_t^3$ can be written in the form

$$([a/b]^{1/3})^3+\mu_t^3=([a/b]^{1/3}+\mu_t)([a/b]^{2/3}-\mu_t[a/b]^{1/3}+\mu_t^2)$$

If we substitute $(a/b)^{1/3}=c$ then $([a/b]^{1/3})^3+\mu_t^3=(c+\mu_t)(c^2-c\mu_t+\mu_t^2)$. The integral we have to solve becomes

$$\int \frac{\mathrm{d}\mu_t}{(c+\mu_t)(c^2-c\mu_t+\mu_t^2)}=bt \Leftrightarrow \frac{\mathrm{d}\mu_t}{(c+\mu_t)(c^2-c\mu_t+\mu_t^2)}=b\,\mathrm{d}t$$

where after some calculations the final solution of the initial differential equation is

$$\frac{1}{3c^2}\ln|c+\mu_t|+\frac{2}{3c}\left(\frac{2}{\sqrt{3}c}\arctan\left(\frac{2\mu_t-c}{\sqrt{3}c}\right)\right)-\frac{1}{3c^2}\left(\frac{1}{2}\ln(\mu_t^2-c\mu_t+c^2)+\frac{\sqrt{3}}{3}\arctan\left(\frac{2\mu_t-c}{c\sqrt{3}}\right)\right)$$

$$=bt+\lambda \Leftrightarrow \frac{1}{3c^2}\ln|c+\mu_t|+\frac{4}{3\sqrt{3}c^2}\arctan\left(\frac{2\mu_t-c}{\sqrt{3}c}\right)-\frac{1}{6c^2}\ln(\mu_t^2-c\mu_t+c^2)$$

$$+\frac{\sqrt{3}}{9c^2}\arctan\left(\frac{2\mu_t-c}{c\sqrt{3}}\right)=bt+\lambda$$

In a similar way, we solve equation (20) with $\delta=4$.

REFERENCES

1. L. von Bertalanffy, 'Untersuchungen über die Geselzlichkeit des Wachtums. Allgemeine Grundlagen der theorie mathematische und physiologische geselzlichkeiten des Wachtums bei Wassertieren', *Arch. Entwick. Mech. Org.*, **131**, 613–653 (1945).
2. Y. Wind, T. S. Robertson and C. Fraser, 'Industrial product diffusion by market segment', *Industrial Marketing Mgmt*, **II**, 1–8 (1982).
3. V. Mahajan and E. Muller, 'Innovation diffusion and new product growth models in marketing', *J. Marketing*, **43**, 55–68 (1979).
4. V. Mahajan and R. A. Peterson, 'First purchase diffusion models of new product acceptance', *Technological Forecasting & Social Change*, **15**, 127–146 (1979).
5. C. H. Skiadas, 'Two generalized rational models for forecasting innovation diffusion', *Technological Forecasting & Social Change*, **27**, 39–61 (1985).
6. I. I. Gihman, *Dokl. Acad. Nauk SSSR*, **58**, 961–964 (1947), in Russian.
7. K. Ito, 'Lectures on stochastic processes', Tata Institute of Fundamental Research, Bombay, India, 1961.
8. N. N. Bogolubov and N. M. Krylov, *Zap. Kaf. Mat. Fiz. Acad. Nauk UK.SSR*, **4**, 5–158 (1939), in Russian.
9. S. Bretschneider and B. Bozeman, 'Adaptive diffusion models for the growth of robotics in New York State industry', *Technological Forecasting & Social Change*, **30**, 111–121 (1986).
10. C. Eliashberg, C. Tapiero and Y. Wind, 'Innovation diffusion models with stochastic parameters', The Wharton School University of Pennsylvania, Working Paper, 1992.
11. L. El-Karoui, 'Representation des processus ponctuels multivaries a l'aide d'un processus de Poisson', *Z. Wahrsch. Verw. Gebiete*, **39**, 111–133 (1977).
12. A. G. Malliaris and W. A. Brock, *Stochastic Methods in Economics and Finance*, Advanced Textbooks in Economics 17, North Holland, 1982.
13. B. Oksendal, *Stochastic Differential Equations: an Introduction with Applications*, 2nd Edn, Springer-Verlag, 1989.
14. K. Raman, 'Dynamic optimization of the marketing mix in a stochastic environment', Working Paper, Department of Marketing, Auburn University, 1985.
15. H. Tanaka, 'Probabilistic treatment of the Boltzmann equation for Maxwellian molecules', *Z. Wahtsch. Verw. Gebiete*, **46**, 67–105 (1978).
16. I. I. Gihman and A. V. Skorohod, *Stochastic Differential Equations*, Springer-Verlag, German, 1972.
17. M. Chesney and R. J. Elliot, 'Estimating the volatility of an exchange rate'. Department de France, Groupe H.E.C., Jony en Josas, France and Department of Statistics and Applied Probability, University of Alberta, Edmonton, Alberta, Canada TGG 261, 1992.