جامعة الحسن الأول
UNIVERSITÉ HASSAN 1ER

Université Hassan 1er
Centre d'Études Doctorales en Sciences et Techniques & Sciences Médicales

كلية العلوم و التقنيات سطات
FACULTÉ DES SCIENCES ET TECHNIQUES SETTAT

Faculté des Sciences et Techniques
Settat

THÈSE DE DOCTORAT

Pour l'obtention de grade de Docteur en ***Sciences et Techniques***

Formation Doctorale: Mathématiques, Informatique et Applications

Spécialité: **Mathématiques Appliquées**

Sous le thème

Stochastic diffusion processes related to some special growth curves: Statistical inference, simulation aspects and applications

Présentée par :

Abdenbi EL AZRI

Soutenue le: *02/11/2022.*

À l'Ecole National des Sciences Appliquées de Berrechid devant le jury composé de :

Pr. Boujemâa ACHCHAB	**PES**	**ENSA de Berrechid**	Président
Pr. Karam ALLALI	**PES**	**FST de Mohammedia**	Rapporteur
Pr. Mohamed NAIMI	**PH**	**ENSA de Berrechid**	Rapporteur
Pr. Abdeljalil SAKAT	**PH**	**ENSA de Berrechid**	Examinateur
Pr. Ahmed NAFIDI	**PES**	**ENSA de Berrechid**	Directeur de thèse

Année Universitaire: 2021/2022

Dedication

To my parents,
especially to my dear mother, there is no dedication quite eloquent to express my great love, my high regard and what you deserve for your prayers and your great sacrifices, which you have never ceased to provide me from my birth up to now.

To my wife,
thank you for giving meaning to my life. Thank you for your love, support and encouragement It has always been a great comfort to me. Thank you for your kindness and sense of sacrifice.

To my close family,
to communicate my best feelings of gratitude, my high regard and my great love as recognition of your valuable tips, your encouragement and your huge support.

To my whole family.

To all my friends.

To all my professors,
from primary level to higher education, who taught and helped me to arrive here.

I dedicate this modest work.

Acknowledgments

Firstly, I would like to address my sincere gratitude to my supervisors Mr Ahmed NAFIDI for their support, their encouragements and their valuable guidance. I am very grateful for their valuable time spent in guiding my research and on sharing with me their expertise and their know-how. Finally, I would also like to say how much i appreciate their great availability and their professional and personal qualities.

I would also like to express my gratitude and appreciation to my origin laboratory, Laboratory of Analysis and Modeling System for Decision Support of Berrechid, for the great working conditions, the numerous activities and the different training which it provides. I address my gratitude to all the teaching and administrative staff of the National School of Applied Mathematics of Berrechid, in particular, Mr Boujemâa ACHCHAB, Mr Mohamed NAIMI and Mr Hamid HRIMECH for their encouragements.

I would like to send my most sincere thanks to Mr Ramón Gutiérrez Sanchez, who has always been present for all our collaborations in scientific articles.

My thanks and appreciation for the doctoral schools of sciences and techniques of Hassan First University, which has significantly contributed in my scientific training through numerous training programs during my thesis.

I acknowledge all my professors, from the primary level up to now, who taught and educated me to arrive here.

I am grateful too, to all my lab mates and my friends PhD students at Hassan First University. I wish you all happiness and full success in your personal and professional life.

Last, a very special thanks to all my family. I am very happy to have them by my sides and i am very grateful to each one of them.

Résumé

Les processus stochastiques de diffusion (PSDs) ont un rôle important dans la modélisation stochastique. Ces processus sont utilisés dans de nombreuses disciplines, telles que: la physique, les phénomènes biologiques, la croissance tumorale, l'économie, la finance, la consommation d'énergie et les phénomènes environnementaux, etc. Pendant ces dernières années, l'estimation des paramètres pour les PSDs a reçu une attention considérable, dans les deux cas où le processus est observé de manière continue ou d'une manière discrète. En général, une telle estimation n'est pas directe, sauf dans des cas simples. Afin de remédier à ce problème, on effectue une approche méthodologique basée sur l'approximation de la fonction du maximum de vraisemblance. Par conséquent, plusieurs auteurs ont étudié et ont développé plusieurs méthodes pour traiter ce problème. Ainsi, l'approche adoptée lors de nos applications couvre les étapes suivantes: la première étape consiste à définir des nouveaux PSDs qui correspondent aux courbes de croissance spéciales ou modifiées. La seconde étape concerne l'étude de l'inférence statistique. Puis, dans la troisième étape, on passe au stade de la modélisation pour justifier le choix du PSD en l'appliquant à des données qui peuvent être réelles ou simulées. Finalement, on élabore des comparaisons avec d'autres PSDs usuels.

Dans cette thèse, nous introduisons des nouveaux PSDs basés sur des courbes de croissance spécifiques, notamment la courbe de croissance de Schumacher, de Lundqvist-Korf, de Lundqvist-Korf modifiée, de logistique et du log-logistique. Notre première contribution porte sur une méthodologie statistique numérique effectuée sur un nouveau PSD dont la fonction moyenne est proportionnelle à la courbe de croissance de Schumacher. Notre deuxième contribution s'articule autour de la définition d'un nouveau PSD avec une fonction moyenne proportionnelle à la courbe de croissance de Lundqvist-Korf. Puis, à l'élaboration des nouveaux PSDs comme la puissance du processus de Lundqvist-Korf. Notre troisième contribution souligne la possibilité de modéliser l'évolution du CO_2 au Maroc à l'aide d'un PSD non homogène (sans facteurs exogènes). Concrètement, nous utilisons un nouveau processus, où la fonction moyenne est proportionnelle à la courbe de croissance modifiée de Lundqvist-korf. Dans notre dernière contribution, nous présentons un nouveau PSD lié à une reformulation de la courbe de croissance logistique. Nous étudions ensuite l'application de l'algorithme de recuit simulé pour l'estimation des paramètres en utilisant la méthode du maximum de vraisemblance. Finalement, nous présentons un nouveau PSD basé sur la théorie des PSDs dont la fonction de densité est proportionnelle à la courbe de croissance log-logistique. Ensuite, nous comparons le comportement de ce processus à celui du PSD logistique obtenu en paramétrant la courbe logistique.

Mots-clés: Processus stochastique de diffusion; La courbe de croissance de Lundqvist-

Korf; La courbe de croissance de Schumacher; La courbe de croissance logistique; Inférence statistique; Estimation du maximum de vraisemblance; Simulation computationnelle; Algorithme de recuit simulé; Fonctions moyenne.

Abstract

The stochastic diffusion processes (SDPs) play an outstanding role in stochastic modeling. We used these processes in many disciplines, such as physics, biological phenomena, tumor growth, economy, finance, energy consumption, environmental science, etc. In most cases, the estimation of the parameters in the SDPs has received considerable attention in recent years. Both when we observe the process continuously and discretely. The problem of the estimation of the parameters is not direct, except in simple cases. We based one possible methodological approach on approximating the maximum likelihood function. Therefore, various authors studied and developed several methods to deal with this problem. Then, the approach adopted during our applications of the SDPs covers the following stages. The first step is to define new SDPs corresponding to specific or modified growth curves. The second step concerns the study of statistical inference. Then, in the third step, we go to the modeling stage to justify the choice of the SDP and highlight the SDP adopted for the data that can be real or simulated. Finally, on proceeds to make comparisons with other usual dissemination processes.

In this thesis, we introduce some new SDPs based on specific growth curves, such as the Schumacher, the Lundqvist-Korf, the modified Lundqvist-Korf, the logistic, and the log-logistic. Our first contribution concerns a computational statistical methodology for a new SDP whose mean function (MF) is proportional to the Schumacher growth curve. Our second contribution makes a point to define a new SDP whose MF is proportional to the growth curve of the Lundqvist-Korf and the elaboration of new stochastic diffusion processes such as the power of the Lundqvist-Korf process. Our third contribution emphasizes the possibility of using a stochastic nonhomogeneous (without exogenous factors) diffusion process to model the evolution of CO_2 in Morocco. Concretely, we use a new model in which the MF is proportional to the modified Lundqvist-Korf growth curve. In the last contribution, we present a new SDP related to a reformulation of the logistic growth curve and explore the application of the simulated annealing algorithm for the maximum likelihood estimation of the parameters of this process. Then, we introduce a new SDP based on the theory of diffusion processes whose MF is proportional to the log-logistic growth curve. Finally, we compare the behavior of this process with that of the logistic diffusion process obtained by the parameterizing logistic curve.

Keywords: Stochastic diffusion process; Lundqvist-Korf growth curve; Schumacher growth curve; Logistic growth curve; Statistical inference; Maximum likelihood estimation; Computational simulation; Simulated annealing algorithm; Mean functions.

Publications

Journal publications

1. A. Nafidi and A. El Azri, A stochastic diffusion process based on the Lundqvist-Korf growth: Computational aspects and simulation. Mathematics and Computers in Simulation, 182: 25-38, 2021, doi: https://doi.org/10.1016/j.matcom.2020.10.022.

2. A. Nafidi, A. El Azri, R. Gutiérrez-Sánchez, The Stochastic Modified Lundqvist-Korf Diffusion Process: Statistical and Computational Aspects and Application to Modeling of the CO_2 Emission in Morocco. Stochastic Environmental Research and Risk Assessment, 36(4): 1163-1176, 2022, doi: https://doi.org/10.1007/s00477-021-02089-8.

3. A. El Azri and A. Nafidi, A γ-power stochastic Lundqvist-Korf diffusion process: Computational aspects and simulation. Moroccan Journal of Pure and Applied Analysis, 8(3): 364–374, 2022, doi: https://doi.org/10.2478/mjpaa-2022-0025.

4. A. Nafidi and A. El Azri, R. Gutiérrez-Sánchez, A stochastic Schumacher diffusion process: Probability characteristics computation and statistical analysis. Preprint submitted to Methodology and Computing in Applied Probability, 2022 (Under Revision).

5. A. Nafidi and A. El Azri, A stochastic log-logistic diffusion process: Statistical computational aspects and application to modelling real example. Preprint submitted to Stochastic Models, 2022 (Under Revision).

6. A. Nafidi and A. El Azri, Stochastic logistic-type diffusion process with application to random growth: Estimating the parameters by means of simulated annealing. Preprint submitted to Stochastics, 2022.

Conference Proceedings

1. A. El Azri and A. Nafidi, "Stochastic diffusion process based on the Schumacher curve: Simulation and application", Conference talk in the 6^{th} International Congress of Moroccan Society of Applied Mathematics (SM2A'6), FST Beni Mellal, Morocco, November $7-9^{th}$, 2019.

2. A. El Azri and A. Nafidi, “Statistical inference in a stochastic diffusion process related to the log-logistic curve: Simulation", Conference talk in the 1^{st} International Congress on Complex Systems and Artificial Intelligence (COSAI) ENSA Berrechid, Morocco, June $29-30^{th}$, 2021.

3. A. El Azri and A. Nafidi, “Trend analysis and computational statistical estimation in a stochastic Lundqvist-Korf diffusion process: Simulation", Conference talk in the 2^{nd} International Conference on Big data, Modeling and Machine Learning (BML'21) ENSA Kenitra, Morocco, July $15-16^{th}$, 2021.

4. A. El Azri and A. Nafidi, “Trend analysis and statistical computation in a stochastic Modified Lundqvist-Korf diffusion process: Simulation", Conference talk in the 2^{nd} edition of the International Conference on Mathematics & Data Science (ICMDS), ENSA Khouribga, Morocco, October $28-30^{th}$, 2021.

5. A. El Azri and A. Nafidi, “A γ-power stochastic Lundqvist-Korf diffusion process: Computational aspects and simulation", Conference talk in the 4^{th} edition of the International Conference on Research in Applied Mathematics & Computer Science (ICRAMCS), FS Ben M'Sick Casablanca, Morocco, March $24-26^{th}$, 2022.

6. A. El Azri and A. Nafidi, “Estimating the parameters of a stochastic Lundqvist-Korf diffusion process by means of simulated annealing", Conference talk in the 1^{st} International Conference on New Trends in Applied Mathematics (ICNTAM'22), FST Béni Mellal, Morocco, May $19-21^{th}$, 2022.

Other activities

1. Poster: A. El Azri and A. Nafidi, “Trend analysis and computational statistical in a stochastic logistic diffusion process: Simulation", presented at the 8^{th} edition of “La journée du doctorant", FST Settat, Morocco, July 8^{th}, 2021.

2. Poster: A. El Azri, A. Nafidi, R. Gutiérrez Sánchez, E. M. Ramos Ábalos, “Stochastic square root of the Lundqvist-Korf diffusion process: Computational statistical inference and simulation aspects", presented at the 39^{th} National Congress of Statistics and Operations Research and the 13^{th} Conference on Public Statistics of the Society of Statistics and Operations Research (SEIO2022), University of Granada, Spain, June $7-10^{th}$, 2022.

Contents

List of Figures

List of Tables

General notations

Notation		Signification
CB	:=	Confidence Bound.
CCB	:=	Conditional Confidence Bound.
CIR	:=	Cox-Ingersoll-Ross.
CMF	:=	Conditional Mean Function.
CV	:=	Coefficient of Variation.
ECMF	:=	Estimated Conditional Mean Function.
EMF	:=	Estimated Mean Function.
ECB	:=	Estimated Confidence Bound.
ECCB	:=	Estimated Conditional Confidence Bound.
LF	:=	Likelihood Function.
LLF	:=	Log-Likelihood Function.
ln	:=	natural logarithm.
MAE	:=	one-step-ahead Mean Absolute Error.
MAPE	:=	Mean Absolute Percentage Error.
MF	:=	Mean Function.
ML	:=	Maximum Likelihood.
MLM	:=	Maximum Likelihood Method.
ODE	:=	Ordinary Differential Equation.
RMSE	:=	Root Mean Square Error.
SA	:=	Simulated Annealing
SDE	:=	Stochastic Differential Equation.
SDP	:=	Stochastic Diffusion Processe
SLDP	:=	Stochastic Logistic Diffusion Process.
SLLDP	:=	Stochastic Log-Logistic Diffusion Process.
SLKDP	:=	Stochastic Lundqvist-Korf Diffusion Process.
γ-SLKDP	:=	The γ power of the Stochastic Lundqvist-Korf Diffusion Process
SMLKDP	:=	Stochastic Modified Lundqvist-Korf Diffusion Process.
SSDP	:=	Stochastic Schumacher Diffusion Process
std	:=	standard deviation.
TF	:=	Trend Function
TPDF	:=	Transition Probability Density Function.

CHAPTER 1

General Introduction

Contents

1.1 Goals and motivations

Stochastic diffusion processes (SDPs) based on deterministic models are currently one of the most frequently used mathematical tools for modeling and describing growth phenomena. In this context, various authors added white noise (Gaussian process) to a deterministic model based on growth curves or modified growth curves. In the same perspective, many types of stochastic diffusion processes have been obtained by this mechanism, such as stochastic Gompertz diffusion process [1], stochastic logistic diffusion process [2,3], stochastic Gamma diffusion process [4], stochastic Richards-type diffusion process [5] and stochastic Hubbert diffusion process [6], etc. From there, we identified our first objective consists in proposing and studying a new stochastic non-homogeneous diffusion process under its probabilistic and statistical aspects. In other words, we introduce a new PSD with a mean function proportional to the specific growth curve, such as the Schumacher growth curve, the Lundqvist-Korf, the modified Lundqvist-Korf, the logistic, or the log-logistic, that allows us to study the real phenomena. Due to this functionality, we use the model in situations where data are available from several individuals, each presenting the same growth curve. Our second objective is to answer the question of statistical inference and the problem of estimating parameters. The maximum likelihood (ML) estimators pose some difficulties because the resulting system of likelihood equations is exceedingly complex. It does not have an explicit solution and requires numerical methods for their resolution. To address this problem, we propose the simulated annealing algorithm to maximize the likelihood function. We establish a logarithm and a strategy for bounding the parametric space. Also, we include the results obtained from several examples of simulation to validate this methodology. Finally, we apply the process and the methodology established to real data. Then, we determine the conditional and unconditional mean functions to compare the real data to those estimated and assess the quality of fit and prediction.

1.2 Generalities

This section presents the general introduction to SDPs, statistical inference for SDPs, and simulated annealing algorithm. Then, the definitions and some properties that we needed in this thesis.

1.2.1 Introduction to stochastic diffusion processes

The SDPs is a family of random variables $\{x(t), t \in \tau\}$ defined on a probability space $(\Omega, \mathcal{A}, P)$ and indexed by a parameter t where t varies over a set τ. If the set τ is discrete, the stochastic process is called discrete. If the set τ is continuous, the stochastic process is called continuous. The parameter t usually plays the role of time and the random variables can be discrete valued or continuous-valued at each value of t. For example, a continuous stochastic process can be discrete-valued. For modeling purposes, it is useful to understand both continuous and discrete stochastic processes and how they are related. Indeed, the solutions of stochastic differential equations are stochastic processes. However, a continuous time parameter stochastic process which possesses the Markov property and for which the sample paths $x(t)$ are continuous functions of t is called a diffusion process (for more details see, Karlin and Taylor [7]).

On the one hand, we define SDPs as a solution of a stochastic differential equation (SDE). It is a differential equation in which one or more terms are a stochastic process. Many physical, biological, finance, radiotherapy, chemotherapy, energy consumption, and economic phenomena are either well approximated or reasonably modeled by SDE. Typically, SDEs contain a variable that represents random white noise calculated as the derivative of Brownian motion or the Wiener process. In general, SDE has the following form:

$$dx(t) = A_1(t, x(t))dt + (A_2(t, x(t)))^{\frac{1}{2}}\, dw(t); \quad x(t_1) = x_1, \tag{1.2.1}$$

where $\{w(t) : t \in [t_1; T], \ t_1 \geq 0\}$ is a one-dimensional Wiener process, with an independent increment $w(t) - w(s)$ normally distributed with mean $\mathbb{E}(w(t) - w(s)) = 0$ and variance $\text{Var}(w(t) - w(s)) = t - s$, for $t \geq s$, and x_1 is a positive random variable, independent on $w(t)$ for $t \geq t_1$, and $A_1(t, x(t))$ is drift coefficient and $A_2(t, x(t))$ is diffusion coefficient. The infinitesimal moments $A_1(t, x(t))$ and $A_2(t, x(t))$ satisfy the Lipschitz and the growth conditions for the existence and uniqueness of the solution to the SDEs (see, Kloeden and Platen [8]). Thus, there exists a non negative constant C, such that for all x and y of $\mathbb{R}^+$ and t of $[t_1, T]$, we have:

$$\begin{aligned}
&|A_1(t,x) - A_2(t,y)| + \left|(A_2(t,x))^{\frac{1}{2}} - (A_2(t,y))^{\frac{1}{2}}\right| \leq C \mid x - y \mid, \\
&|A_1(t,x)|^2 + \left|(A_2(t,x))^{\frac{1}{2}}\right|^2 \leq C^2 \left(1 + \mid x \mid^2\right).
\end{aligned}$$

Then, the SDE in Eq. (1.2.1) has a unique solution $\{x(t) : t \in [t_1, T], \ t_1 \geq 0\}$ continuous with probability 1, and satisfies the initial condition $x(t_1) = x_1$.

On the other hand, SDPs can be defined by a partial differential equations (PDEs). So, we consider the Fokker-Plank equation of the non-homogeneous diffusion process $\{x(t) : t \in [t_1, T], \quad t_1 > 0\}$ with infinitesimal moments $A_1(t, x)$ and $A_2(t, x)$. The transition probability density function (TPDF) of this process $f(x, t|y, s)$ is the solution of the following partial derivative equations

$$\frac{\partial f(t,x)}{\partial t} = -\frac{\partial [A_1(t,x) f(t,x)]}{\partial x} + \frac{1}{2}\frac{\partial^2 [A_2(t,x) f(t,x)]}{\partial x^2},$$

and the Kolmogorov or bakward equation

$$\frac{\partial f(s,y)}{\partial s} = -A_1(s,y)\frac{\partial f(s,y)}{\partial y} - \frac{1}{2}A_2(s,y)\frac{\partial^2 f(s,y)}{\partial y^2}; \quad y > 0, \ s \geq t_1,$$

with the initial condition $\lim_{t \to t_1} f(x, t|y, s) = \delta(x - y)$ and $\delta(.)$ is the Dirac delta function.

Various diffusion-type stochastic models have been developed and successfully applied to the fitting and prediction of real phenomena, such as in physics, biological phenomena, tumor Growth, economy and finance, life expectancy at birth, energy consumption, etc. These models include the homogeneous case such as the log-normal diffusion process [9–12], Cox-Ingersoll-Ross process [13], Gompertz diffusion process [14–18], Rayleigh diffusion process [19–21], Hyperbolic processes [22], Jacobi diffusion process [23], and Pearson diffusion [24]. In the case of the non-homogeneous case (with an exogenous factor), we can quote, for example, the gamma diffusion process [4], the stochastic Bertalanffy diffusion process [25], the stochastic Richards-type diffusion process [5], the stochastic Hubbert diffusion process [6], Brennan-Schwartz process [26], the Weibull diffusion process [27] and the Lundqvist-Korf diffusion process [28].

Then, Ricciardi [29] and Capocelli et al. [1] performed the first studies associated with the exponential curve Malthusian model, gave rise to the lognormal diffusion process, and applied it in ecology. In addition, Ricciardi [29] defined the Gompertz diffusion that is the most suitable for applications in biology thanks to its predictive capacity through its sigmoid type tendency. The non-homogeneous version of this process was considered by Ferrante et al. [30, 31] and Giorno et al. [21].

Various SDPs have been applied to describe and forecast the evolution of CO_2 emissions. In this perspective, Gutiérrez et al. [32] carried out the study of the SDP with cubic drift with application to a modeling of the global CO_2 emissions in Spain. Then, Gutiérrez et al. [33] in the case of the non-homogeneous stochastic Vasicek diffusion process in the case of CO_2 emission in Morocco. Moreover, Gutiérrez et al. [34] proposed the bivariate stochastic Gompertz diffusion model as the solution for a system of two Itô's stochastic differential equations (SDEs). The drift and diffusion coefficients are similar to those considered in the univariate stochastic Gompertz diffusion process. We applied this SDP to model gross domestic product and CO_2 emissions in Spain. Moreover, Abbass et al. [35] presented a systematic review of two decades of research from 1995 to 2017 on CO_2 emissions and economic growth. Magazzino and Cerulli [36] examined the relationship between CO_2 emissions, Gross domestic product (GDP), and energy in the Middle East and North African countries using a responsiveness scores approach. Then, Solaymani [37] treated the CO_2 emissions patterns in seven top carbon emitter economies in the case of the transport sector.

As a result, many works of SDPs related to logistic growth models. Such as, Capocelli and Ricciardi [38] derive a diffusion process from a reparameterization of the logistic growth model. Giovanis and Skaidas [2] proposed a stochastic version of the well-known logistic model that is solved analytically using reducible stochastic differential equation. They applied this model to study electricity consumption in Greece and the United States. Heydari et al. [39] introduced two new Vrst order linear noise approximations of a stochastic logistic diffusion process, one with multiplicative and one with additive intrinsic noise. Tang and Heron [40] have used Markov chain Monte Carlo techniques to carry out Bayesian inference for piecewise stochastic logistic growth models using discretely observed data sets. It allows us to fit models for time series data, including data on fish productions and yields, with structural changes.

1.2.2 Inference for stochastic diffusion processes

The statistical inference for SDPs has been an active research area in recent years, both when the process is observed continuously and when it is observed discretely. In general, the estimation of the parameters in stochastic models, is not direct, except in simple cases, and one possible methodological approach is based on approximating the maximum likelihood function. In this context, estimation methods addressing this problem have been developed, and many works have been published on this subject, focusing on several variants of approximate likelihood methodology. The general case of this methodology can be consulted in Prakasa-Rao [41], Bibby and al. [22], Ait-Sahalia [42], and Egorov et al. [43]; and in the case of particular diffusions, the following can be seen, for example, Gutierrez and al. [15, 17, 19]. Also, other works studying the hypothesis testing on diffusion process, for example, Dadgar et al. [44], who provides a long list of references on the subject.

Various researchers approached the problem of Maximum likelihood estimation by equating partial derivatives of the log-likelihood function to zero. Looking for a stationary point of the local maximum likelihood equations by iterative methods (see, for example, Wilson and Worcester [45], Cohen [46], Lambert [47], Harter et al. [48], and Calitz [49]), etc.

Moreover, we must not forget the other estimation methods, such as the method of moments (see Chan et al. [50]), the non-parametric methods (see, Arapis and Gao. [51] and Jiang and Knight [52]), and the Bayesian methods (see, Elerian and al. [53]).

Recently, several researchers have used the simulated annealing algorithm for estimating the parameters in the SDP (see, for instance, Nafidi and El-Azri [28], Nafidi et al. [54], Istoni et al. [6], and Román-Román et al. [5,55]). Other works have used the global simulated annealing heuristic for the three parameters log-normal maximum likelihood estimation (see, Fernando et al. [56]). Finally, Pedersen [57] suggested a new approach to maximum likelihood estimation for SDEs based on discrete observations when the likelihood function is unknown.

1.2.3 Simulated annealing algorithm

The simulated annealing (SA) algorithm is a stochastic global optimization method for problems of the type $\min_{\theta\in\Theta} g(\theta)$. It was developed by Kirkpatrick et al. [58] and Cerny [59]. This algorithm performs an iterative exploration of solution space Θ searching for improvements on the value of the objective function, say g, and intends to avoid an attraction towards local minima. Concretely, in each iteration, let x be a current solution, x' be a new value selected in a neighborhood of x in the next iteration, and the objective difference is $\delta = g\left(x'\right) - g(x)$. If $\delta \leq 0$, then x' is selected as the new solution. Otherwise, it could be accepted with probability $p = \exp\left(\frac{-\delta}{T}\right)$, where T is a scale factor called temperature.

The general application of the SA algorithm depends on the definition of the several parameters:

1. Initializing the parameters of the algorithm such as the initial solution θ_0, the initial temperature T_0, the final temperature T_F, and the chain length for each application of the Metropolis algorithm L and the cooling procedure and the stopping condition.
2. Apply the selection procedure for a new solution L times.
3. Verifying the stopping condition. If it is not verified, decrease the temperature and return to the previous step.

1.2.4 Itô's formula

1.2.4.1 Lemma

Let $x(t)$ be an Itô process given by

$$dx(t) = A_1(t, x(t))dt + (A_2(t, x(t)))^{\frac{1}{2}}\, dw(t).$$

Let $g(t,x) \in \mathcal{C}^2\left([0,\infty) \times \mathbb{R}\right)$ (i.e., g is twice continuously differentiable on $[0,\infty) \times \mathbb{R}$). Then $y(t) = g\left(t, x(t)\right)$ is again an Itô process, and

$$\begin{aligned} dy(t) = &\left[\frac{\partial g\left(t, x(t)\right)}{\partial t} + A_1\left(t, x(t)\right)\frac{\partial g\left(t, x(t)\right)}{\partial x} + \frac{A_2(t, x(t))}{2}\frac{\partial^2 g\left(t, x(t)\right)}{\partial x^2}\right] dt \\ &+ \left[(A_2(t, x(t)))^{\frac{1}{2}}\frac{\partial g\left(t, x(t)\right)}{\partial x}\right] dw(t). \end{aligned} \tag{1.2.2}$$

1.2.5 Existence and uniqueness solutions to a SDE

1.2.5.1 Theorem

Let $A_1(.,.) : [t_1, T] \times \mathbb{R} \longrightarrow \mathbb{R}$ and $A_2(.,.) : [t_1, T] \times \mathbb{R} \longrightarrow \mathbb{R}$ be measurable functions such that for all x and y of $\mathbb{R}$, $t \in [t_1, T]$, and for some constant K, we have

$$|A_1(t,x) - A_1(t,y)| + \left|(A_2(t,x))^{\frac{1}{2}} - (A_2(t,y))^{\frac{1}{2}}\right| \leq K \mid x - y \mid,$$

and

$$|A_1(t,x)|^2 + \left|(A_2(t,x))^{\frac{1}{2}}\right|^2 \leq K^2 \left(1+ \mid x \mid^2\right),$$

where $| \, . \, |$ denotes the absolute value in $\mathbb{R}$ and $0 \leq t_1 < T < \infty$.

Let x_1 be a random variable which is independent of the σ-algebra $\mathcal{F}$ generated by $w_s(.)$, $s \geq 0$ and such that $\mathbb{E}\left(|x_1|^2\right) < \infty$. Then, the stochastic differential equation

$$dx(t) = A_1\left(t, x(t)\right) dt + \left(A_2(t, x(t))\right)^{\frac{1}{2}} dw(t); \quad t_1 \leq t \leq T, \quad x\left(t_1\right) = x_1,$$

has a unique t-continuous solution $x_t(w)$ with the property that $x_t(w)$ is adapted to the filtration $\mathcal{F}^{x_1}$ generated by x_1 and $w_s(.);\ s \leq t$ and $\mathbb{E}\left[\int_{t_1}^{T} |x_t|^2\right] < \infty$.

1.2.6 Log-normal distribution

1.2.6.1 Definition

In probability theory, a log-normal (or lognormal) distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. Thus, if the random variable X is log-normally distributed, then $Y = \ln(X)$ has a normal distribution. Equivalently, if Y has a normal distribution, then $X = \exp(Y)$ has a log-normal distribution.

1.2.6.2 Probability density function

A positive random variable X is log-normally distributed, if the natural logarithm of X is normally distributed with mean μ and variance σ^2:

$$\ln(X) \sim \mathcal{N}\left(\mu, \sigma^2\right).$$

Let Φ and φ be respectively the cumulative probability distribution function and the probability density function of the $\mathcal{N}\left(\mu, \sigma^2\right)$ distribution, then we have that

$$\begin{aligned} f_X(x) &= \frac{\mathrm{d}}{\mathrm{d}x} \Pr\left(X \leq x\right) = \frac{\mathrm{d}}{\mathrm{d}x} \Pr\left(\ln(X) \leq \ln(x)\right) = \frac{\mathrm{d}}{\mathrm{d}x} \Phi\left(\ln(x)\right) \\ &= \varphi\left(\ln(x)\right) \frac{\mathrm{d}}{\mathrm{d}x}\left(\ln(x)\right) = \varphi\left(\ln(x)\right) \frac{1}{x} \\ &= \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{\left(\ln(x) - \mu\right)^2}{2\sigma^2}\right). \end{aligned}$$

1.2.6.3 Arithmetic moments

For any real or complex number n, the n-th moment of a log-normally distributed variable X is given by

$$\mathrm{E}\left[X^n\right] = \exp\left(n\mu + \frac{n^2\sigma^2}{2}\right).$$

Specifically, the arithmetic mean, expected square, arithmetic variance, and arithmetic standard deviation of a log-normally distributed variable X are respectively given by

$$\begin{aligned}
\mathrm{E}[X] &= \exp\left(\mu + \tfrac{1}{2}\sigma^2\right), \\
\mathrm{E}\left[X^2\right] &= \exp\left(2\mu + 2\sigma^2\right), \\
\mathrm{Var}\left[X\right] &= \mathrm{E}\left[X^2\right] - \mathrm{E}\left[X\right]^2 = \mathrm{E}\left[X\right]^2\left(\exp\left(\sigma^2\right) - 1\right) = \exp\left(2\mu + \sigma^2\right)\left(\exp\left(\sigma^2\right) - 1\right), \\
\mathrm{SD}\left[X\right] &= \sqrt{\mathrm{Var}\left[X\right]} = \mathrm{E}\left[X\right]\sqrt{\exp\left(\sigma^2\right) - 1} = \exp\left(\mu + \tfrac{1}{2}\sigma^2\right)\sqrt{\exp\left(\sigma^2\right) - 1}.
\end{aligned}$$

The arithmetic coefficient of variation $\mathrm{CV}[X]$ is the ratio $\dfrac{\mathrm{SD}[X]}{\mathrm{E}[X]}$. For a log-normal distribution it is equal to

$$\mathrm{CV}[X] = \sqrt{\exp\left(\sigma^2\right) - 1}.$$

1.3 Contributions and structure of the thesis

The research work conducted in this thesis is the definition of new stochastic diffusion processes. Its purpose is to introduce a new diffusion model related to many specific growth curves. We present the stochastic Schumacher diffusion process based on the Schumacher growth curve, the stochastic Lundqvist-Korf diffusion process based on the Lundqvist-Korf growth curve and its extended diffusion processes, the stochastic modified Lundqvist-Korf diffusion process based on the modified Lundqvist-Korf growth curve, the stochastic logistic type diffusion process based on a reformulation of the logistic curve, and the stochastic log-logistic diffusion process based on the log-logistic growth curve. Below, we summarize the contents of the five chapters with a focus on the original contributions contained in each of them.

In Chapter 2, we introduce a new stochastic diffusion process, in which his mean function is proportional to the Schumacher growth curve. We used this model for forest plantations, animal science, and artificial neural networks. Firstly, the main features probabilistic of the process are analyzed. Including the transition probability density function, the moment functions, and the mean conditioned and unconditioned functions. Then, the parameters are estimated by considering discrete sampling of the sample path of the process and by using the maximum likelihood methodology. Finally, to provide the performance of the proposed process, we will apply this process with its associated statistical methodology to simulated data based on a discretization of the exact solution of the stochastic differential equation of the model.

In Chapter 3, we present a new stochastic diffusion process in which the mean function is proportional to the growth curve of the Lundqvist-Korf. We analyze the main features of the process, including the transition probability density function and mean functions. We estimate the parameters of the model by the maximum likelihood method using discrete sampling. We propose the simulated annealing algorithm to solve the problem of maximum likelihood estimation of the parameters. We then define the stochastic square root and the order power of the basic Lundqvist-Korf diffusion process. Then, we obtain all the probabilistic characteristics of these processes and the inference study. Finally, we apply the proposed process and statistical results to simulated data. It was the subject of our contribution to the Mathematics and Computers in Simulation [28].

In Chapter 4, we introduce a new non-homogeneous stochastic diffusion process in which the mean function is proportional to the modified Lundqvist-Korf growth curve. First, we study the main characteristics of the process. Moreover, we establish a computational statistical methodology based on the maximum likelihood estimation method and the trend functions. Otherwise, when we estimate the parameters of the model, we obtain a nonlinear equation. Hence, we propose the simulated annealing method to solve it after bounding the parametric space by a stagewise procedure. Also, we include the results obtained from several examples of simulation to validate this methodology. Finally, we apply the process and the methodology established before to real data corresponding to the evolution of CO_2 emissions in Morocco. It was the subject of our contribution to the Stochastic Environmental Research and Risk Assessment [60].

In Chapter 5, we present two new stochastic diffusion processes. One is related to a reformulation of the logistic growth curve and explores the application of the simulated annealing algorithm for the maximum likelihood estimation of the parameters of this process. The main characteristics of the process are analyzed, including the transition probability density function and mean functions. We estimate the parameters of the process by the maximum likelihood method using discrete sampling. To this end, we apply the simulated annealing algorithm after bounding the parametric space by a stagewise procedure. In the end, to validate this methodology, we include the results obtained from several examples of simulation. Then, we give an application for the growth of a microorganism culture. The other is based on the theory of diffusion processes, whose mean function is proportional to the log-logistic growth curve. The main characteristics of both processes are analyzed, including the transition probability density function and mean functions. We estimate the parameters of the process by the maximum likelihood method using discrete sampling. We apply the simulated annealing algorithm after bounding the parametric space by a strategy procedure to solve the likelihood equations. The behavior of the diffusion process here derived is finally compared with that of the well-known diffusion process obtained by parameterizing logistic curve.

CHAPTER 2

Stochastic Schumacher diffusion process: Statistical inference and simulation aspects

Contents

This chapter is the complete version of the paper [61] preprint submitted to Methodology and Computing in Applied Probability (Under Revision).

2.1 Introduction

Stochastic processes (SDPs) in general and those of diffusion in particular have an extremely important role in the modeling of various phenomena in several disciplines, such as in health, environmental phenomena, economic and financial study, consumption of energy and others.

In recent years, statistical inference in diffusion processes and especially parameter estimation have received considerable attention from researchers, one where the process is continuously observed and the other where it is discretely observed. In general, such an estimation is not always direct, apart from simple cases and a possible methodology is founded on the approximation of the maximum likelihood function. Several methods answering this problematic have been elaborated, and numerous publications have been produced on this topic, based on the different versions of the approximate likelihood strategy, the generalized case of previous strategy is detailed in Prakasa-Rao [41], Bibby et al. [22], Egorov and al. [43], and Ait-Sahalia [42], and we can see, for instant, Gutierrez and al. [17] for the case of particular diffusions. Also, other works studying the hypothesis testing on diffusion process, like for example, Dadgar et al. [44], who provide a long list of references on the subject. Finally, we must not forget the other estimation methods, for example the moments method described by Chan et al. [50], the nonparametric methods, see for example,Arapis and Gao [51] and Jiang and Knight [52] and Bayesian methods, we recommend to see, Elerian and al. [53].

Various diffusion stochastic models have been developed and successfully applied to the fitting and prediction of real phenomena. These models include homogeneous case such as Log-normal diffusion process [9, 10], CIR process [13], Gompertz diffusion process [16], Rayleigh diffusion process [19], Hyperbolic processes [22], Jacobi diffusion process [23] and Pearson diffusion [24]. In the case of the inhomogeneous case (with exogeneous factor), we can quote for example, the gamma diffusion process [4], Brennan-Schwartz process [26], the Weibull diffusion process [27] and the Lundqvist-Korf diffusion process [28].

The Schumacher growth curve was first used to model stem volume growth in forest plantations (see, Schumacher [62]). This curve has been applied previously in animal science (see, Schulin-Zeuthen et al. [63]) and to express productive capacity in a young plantation of Tectona grandis (see, Silva et al. [64]). Silva et al. [65] suggested an adjustment of the Schumacher for application of artificial neural networks to estimate volume of eucalypt trees. Liang and al. [66] introduced a uncertain Johnson–Schumacher growth model with imprecise observations and k-fold cross-validation test. In the context of this research, we define a new type non-homogeneous extension of the log-normal diffusion, based on the Schumacher growth curve.

The main aim of this chapter is to define a new SDP termed Stochastic Schumacher Diffusion Process (SSDP). The term adopted for this process can be proved by the relationship between the mean function of this process and the Schumacher growth curve. The rest of this paper is organized as follows: in the section 2.2, we presented an overview of the Schumacher growth curve and we define the model in terms of stochastic differential equation (SDE). Then we determine the main probabilistic characteristics of the model, such as the solution to the stochastic differential equation, transition probability density function (TPDF), the moment functions, and in particular the mean conditioned and unconditioned functions. In the section 2.3, discusses parameter estimation by maximum likelihood (ML), using discrete sampling. In this case, the ML estimators can be given in explicit form because the system of likelihood equations have an explicit solution. In the section 2.4, presents the results obtained from the simulations. We then illustrate the fit and prediction possibilities of the process. Finally, the briefly summarizes and concludes from this study.

2.2 The model and its characteristics

2.2.1 An overview of the Schumacher curve

The growth curve proposed by Schumacher [62] is based on the hypothesis that the relative growth rate has a linear relationship with the inverse time squared

$$\frac{1}{x(t)}\frac{dx(t)}{dt} = \frac{\beta}{t^2}; \quad t > 0,\ \beta > 0. \tag{2.2.1}$$

By integrating (2.2.1) with respect to t and imposing that $x\left(t_1\right) = x_1$, thus, we have

$$x(t) = x_1 \exp\left(\frac{\beta}{t_1}\right) \exp\left(\frac{-\beta}{t}\right). \tag{2.2.2}$$

The curve represented by equation (2.2.2), in the following referred by Schumacher-type growth curve, verifies the following properties:

1. It is strictly increasing.
2. $\lim_{t\to\infty} x(t) = x_1 \exp\left(\frac{\beta}{t_1}\right)$, hence the line $y = x_1 \exp\left(\frac{\beta}{t_1}\right)$ is the horizontal asymptote of the curve represented by equation (2.2.2). Its asymptote is dependent on the initial value.
3. It shows an inflection point at $t_I = \frac{\beta}{2}$, verifying $x\left(t_I\right) = x_1 \exp\left(\frac{\beta}{t_1}\right) \exp\left(-2\right)$.
4. In addition, this point verify $t_I > t_1$. Then, $t_I > t_1$ if and only if $\beta > 2t_1 > t_1$.

Figure 2.1 shows the Schumacher curve and inflection point for several values of β.

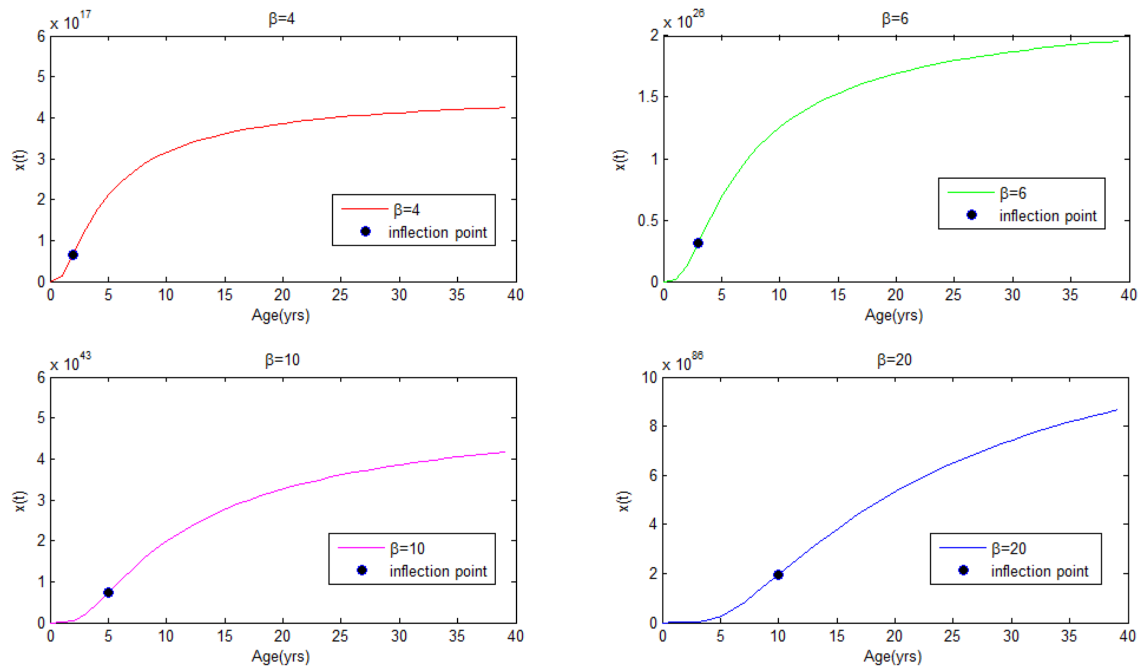


Figure 2.1: The Schumacher curve for several values of β $(x_1 = 2,\ t_1 = 0.1,\ T = 40)$.

2.2.2 The SSDP model

In order to obtain the stochastic version of the Schumacher growth curve, our contribution is to consider a diffusion process whose mean function given by equation (2.2.2). Now, starting from (2.2.2), one is lead to considering the ordinary differential equation (ODE)

$$\frac{dx(t)}{dt} = h(t)x(t); \quad x(t_1) = x_1, \tag{2.2.3}$$

where $h(t) = \dfrac{\beta}{t^2}$. Hence (2.2.3) can be viewed as a generalisation of the Malthusian growth model with time dependent fertility rate $h(t)$. Note that $h(t)$ is a decreasing continuous positive function and has a horizontal asymptote at $y = 0$ and a vertical asymptote at $t = 0$.

A stochastic version of the model is given by the diffusion process $\{x(t) : t \geq t_1\}$, taking values on $(0, \infty)$ and with infinitesimal moments

$$\begin{aligned} A_1(t, x) &= \frac{\beta}{t^2}x, \\ A_2(t, x) &= \sigma^2 x^2, \end{aligned} \tag{2.2.4}$$

where $\beta > 0$ and $\sigma^2 > 0$ are real parameters.

Alternatively, the above process can be defined by the following Itô's stochastic differential equation:

$$dx(t) = A_1(t, x)dt + (A_2(t, x))^{\frac{1}{2}}\, dw(t); \quad x(t_1) = x_1, \tag{2.2.5}$$

where $w(t)$ is a standard Wiener process and x_1 is fixed in $\mathbb{R}_+^*$.

2.2.3 Analytic solution of the SSDP

The infinitesimal moments $A_1(t,x)$ and $A_2(t,x)$ specified in Eq. (2.2.4) satisfy the Lipschitz and the growth conditions for the existence and unicity of the solution to the SDEs (see, Kloeden and Platen [8]). In fact, there exists a non negative constant $C = \frac{\beta}{t_1^2} + \sigma$, such that for all $x, y \in \mathbb{R}^+$ and $t \in [t_1, T]$, we have:

$$\begin{aligned}
&|A_1(t,x) - A_1(t,y)| + \left|(A_2(t,x))^{\frac{1}{2}} - (A_2(t,y))^{\frac{1}{2}}\right| \leq C \mid x - y \mid, \\
&|A_1(t,x)|^2 + \left|(A_2(t,x))^{\frac{1}{2}}\right|^2 \leq C^2 \left(1 + \mid x \mid^2\right).
\end{aligned}$$

Then, the SDE Eq. (2.2.5) has a unique solution $\{x(t) : t \in [t_1, T], \ t_1 > 0\}$ continuous with probability 1, and satisfies the initial condition $x(t_1) = x_1$.

The analytical expression of this solution is obtained by applying Itô's formula to the transform $y(t) = \ln(x(t))$, then we have

$$dy(t) = \left(\frac{\beta}{t^2} - \frac{\sigma^2}{2}\right) dt + \sigma dw(t); \quad y(t_1) = \ln(x_1),$$

by integrating both sides yields,

$$y(t) = y_1 - \beta\left(\frac{1}{t} - \frac{1}{t_1}\right) - \frac{\sigma^2}{2}(t - t_1) + \sigma\left(w(t) - w(t_1)\right).$$

Finally, we have:

$$x(t) = x_1 \exp\left[-\beta\left(\frac{1}{t} - \frac{1}{t_1}\right) - \frac{\sigma^2}{2}(t - t_1) + \sigma\left(w(t) - w(t_1)\right)\right]. \tag{2.2.6}$$

2.2.4 Probability distribution

Taking to account that the random variable $w(t) - w(s)$ has the one-dimensional normal distribution $\mathcal{N}_1(0, t-s)$, it can be deduced that $(x(t)|x(s) = x_s)$ is distributed as one-dimensional lognormal distribution $\Lambda_1\left(\mu(s,t,x_s); \sigma^2(t-s)\right)$, where $\mu(s,t,x_s)$ is given by

$$\mu(s,t,x_s) = \ln(x_s) - \beta\left(\frac{1}{t} - \frac{1}{s}\right) - \frac{\sigma^2}{2}(t-s)$$

and therefore the TPDF of the process has the following form:

$$f(y,t|x,s) = \frac{1}{x\sqrt{2\pi\sigma^2(t-s)}} \exp\left(-\frac{[\ln\left(\frac{y}{x}\right) + \beta\left(\frac{1}{t} - \frac{1}{s}\right) + \frac{\sigma^2}{2}(t-s)]^2}{2\sigma^2(t-s)}\right). \tag{2.2.7}$$

2.2.5 The MFs of the SSDP

By the properties of the log-normal distribution, we obtain the conditional moments of order r of this process:

$$\begin{aligned} E\left(x^r(t)|x(s)=x_s\right) &= \exp\left(r\mu\left(s,t,x_s\right)+\frac{r^2\sigma^2}{2}(t-s)\right) \\ &= \exp\left(\ln\left(x_s^r\right)-\beta r\left(\frac{1}{t}-\frac{1}{s}\right)-\frac{r\sigma^2}{2}(t-s)+\frac{r^2\sigma^2}{2}(t-s)\right) \\ &= x_s^r\exp\left(-\beta r\left(\frac{1}{t}-\frac{1}{s}\right)\right)\exp\left(\frac{r(r-1)\sigma^2}{2}(t-s)\right). \end{aligned}$$

Then, for $r=1$, the conditional mean function (CMF) of the process is

$$E\left(x(t)|x(s)=x_s\right)=x_s\exp\left(-\beta\left(\frac{1}{t}-\frac{1}{s}\right)\right). \tag{2.2.8}$$

In addition, taking into the initial condition $P\left(x\left(t_1\right)=x_1\right)=1$, the mean function (MF) of the process is given by

$$E(x(t))=x_1\exp\left(\frac{\beta}{t_1}\right)\exp\left(\frac{-\beta}{t}\right). \tag{2.2.9}$$

2.3 Statistical inference on the model

2.3.1 Parameters estimation

Let's take a discrete sampling of the model, i.e. for the time points $t_1, t_2, \cdots, t_n$ with $n>2$, we observe the values of the variables $x\left(t_1\right)$, $x\left(t_2\right)$, $\cdots$, $x\left(t_n\right)$ and their values constitute the basic sample for the inference procedure. Then, we suppose $\mathrm{P}\left[x\left(t_1\right)=x_1\right]=1$ and $t_i-t_{i-1}=h$, for $i=2,\cdots,n$. We note by x_1, x_2, $\cdots$, x_n the actual values of the sampling. Then, the corresponding likelihood function (LF) is

$$\begin{aligned} \mathbf{L} &= L_{x_1,x_2,\cdots,x_n}\left(\beta,\sigma^2\right)=\prod_{i=2}^{n} f\left(x_i,t_i|x_{i-1},t_{i-1}\right) \\ &= \prod_{i=2}^{n}\frac{1}{x_i\sqrt{2\pi\sigma^2 h}}\exp\left(-\frac{\left[\ln\left(\frac{x_i}{x_{i-1}}\right)+\beta\left(\frac{1}{t_i}-\frac{1}{t_{i-1}}\right)+\frac{\sigma^2}{2}h\right]^2}{2\sigma^2 h}\right). \end{aligned}$$

Finally, the expression of the LF is

$$\mathbf{L}=\prod_{i=2}^{n}\frac{1}{x_i\sqrt{2\pi\sigma^2 h}}\exp\left(-\frac{\left[\ln\left(\frac{x_i}{x_{i-1}}\right)-\frac{\beta h}{t_i t_{i-1}}+\frac{\sigma^2}{2}h\right]^2}{2\sigma^2 h}\right). \tag{2.3.1}$$

The log-likelihood function (LLF) of the sample is

$$\begin{aligned}\ln(\mathbf{L}) &= -\frac{(n-1)}{2}\ln(2\pi h) - \frac{(n-1)}{2}\ln\left(\sigma^2\right) - \sum_{i=2}^{n}\ln\left(x_i\right) \\ &- \frac{1}{2\sigma^2 h}\sum_{i=2}^{n}\left[\ln\left(\frac{x_i}{x_{i-1}}\right) - \frac{\beta h}{t_i t_{i-1}} + \frac{\sigma^2}{2}h\right]^2.\end{aligned}$$

Deriving the LLF with respect to β and σ^2, we obtain

$$\begin{aligned}\frac{\partial\ln(\mathbf{L})}{\partial\sigma^2} = -\frac{(n-1)}{2\sigma^2} + \frac{1}{2\sigma^4 h}\sum_{i=2}^{n}\left[\ln\left(\frac{x_i}{x_{i-1}}\right) - \frac{\beta h}{t_i t_{i-1}} + \frac{\sigma^2}{2}h\right]^2 \\ - \frac{1}{2\sigma^2}\sum_{i=2}^{n}\left[\ln\left(\frac{x_i}{x_{i-1}}\right) - \frac{\beta h}{t_i t_{i-1}} + \frac{\sigma^2}{2}h\right].\end{aligned} \tag{2.3.2}$$

$$\frac{\partial\ln(\mathbf{L})}{\partial\beta} = \frac{1}{\sigma^2}\sum_{i=2}^{n}\left[\frac{1}{t_i t_{i-1}}\left(\ln\left(\frac{x_i}{x_{i-1}}\right) - \frac{\beta h}{t_i t_{i-1}} + \frac{\sigma^2}{2}h\right)\right]. \tag{2.3.3}$$

Making the derivatives (2.3.2) and (2.3.3) equal to zero, we obtain the following set of equations

$$\frac{\sigma^4 h^2}{4} + \sigma^2 h - \frac{1}{n-1}\sum_{i=2}^{n}\left(\ln\left(\frac{x_i}{x_{i-1}}\right) - \frac{\beta h}{t_i t_{i-1}}\right)^2 = 0 \tag{2.3.4}$$

$$\sum_{i=2}^{n}\left[\frac{1}{t_i t_{i-1}}\left(\ln\left(\frac{x_i}{x_{i-1}}\right) - \frac{\beta h}{t_i t_{i-1}} + \frac{\sigma^2}{2}h\right)\right] = 0. \tag{2.3.5}$$

By the Eq. (2.3.5), the estimator of β is given by

$$\beta = \frac{1}{h\sum\limits_{i=2}^{n}\left(\frac{1}{t_i t_{i-1}}\right)^2}\left[\sum_{i=2}^{n}\frac{1}{t_i t_{i-1}}\left\{\ln\left(\frac{x_i}{x_{i-1}}\right) + \frac{\sigma^2}{2}h\right\}\right]. \tag{2.3.6}$$

From Eq. (2.3.4) and Eq. (2.3.6), we can deduce

$$\frac{ah^2}{4}\sigma^4 + (n-1)\sum_{i=2}^{n}\left(\frac{1}{t_i t_{i-1}}\right)^2 h\sigma^2 - b = 0, \tag{2.3.7}$$

where $a = (n-1)\sum\limits_{i=2}^{n}\left(\frac{1}{t_i t_{i-1}}\right)^2 - \left(\sum\limits_{i=2}^{n}\frac{1}{t_i t_{i-1}}\right)^2$ and

$$b = \sum_{i=2}^{n}\left(\frac{1}{t_i t_{i-1}}\right)^2\sum_{i=2}^{n}\ln^2\left(\frac{x_i}{x_{i-1}}\right) - \left[\sum_{i=2}^{n}\frac{1}{t_i t_{i-1}}\ln\left(\frac{x_i}{x_{i-1}}\right)\right]^2.$$

The Eq. (2.3.7) is a second-degree equation in σ^2, as the discriminant of this equation is positive, then, the non-negative solution corresponding to σ^2 is

$$\sigma^2 = \frac{2}{ah}\left[\left\{\left[(n-1)\sum_{i=2}^{n}\left(\frac{1}{t_i t_{i-1}}\right)^2\right]^2 + ab\right\}^{1/2} - (n-1)\sum_{i=2}^{n}\left(\frac{1}{t_i t_{i-1}}\right)^2\right]. \tag{2.3.8}$$

2.3.2 A confidence bounds of the SSDP

It's possible to obtain a confidence bounds (CBs) of the process via the same technique used in [67]. As $w(t) - w(t_1)$ is Gaussian with mean zero and variance $t - t_1$ (for $t \geq t_1$), therefore, the random variable z given by

$$z = \frac{\ln(x(t)) - \mu(t_1, t, x_1)}{\sigma\sqrt{(t - t_1)}} \sim N_1(0, 1).$$

A $100(1 - \alpha)\%$ CB for z is given by $P(-\xi \leq z \leq \xi) = 1 - \alpha$. From this, we can obtain a CB of $x(t)$ with following form $(x_{lower}(t), x_{upper}(t))$ where,

$$x_{lower}(t) = \exp\left(\mu(t_1, t, x_1) - \xi\sigma\sqrt{(t - t_1)}\right), \tag{2.3.9}$$

and

$$x_{upper}(t) = \exp\left(\mu(t_1, t, x_1) + \xi\sigma\sqrt{(t - t_1)}\right), \tag{2.3.10}$$

with $\xi = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ and where Φ^{-1} is the inverse cumulative normal standard distribution.

2.3.3 Estimated MFs and CB

By using Zehna's theorem [68], On the one hand, the estimated MF (EMF) and the estimated CMF (ECMF) are obtained by replacing the parameters by their estimators in the equations Eq. (2.2.8) and Eq. (2.2.9). Thus the ECMF is:

$$\hat{E}\left(x(t)|x(s) = x_s\right) = x_s \exp\left(-\hat{\beta}\left(\frac{1}{t} - \frac{1}{s}\right)\right). \tag{2.3.11}$$

In addition, taking into the initial condition, the EMF of the process is given by

$$\hat{E}(x(t)) = x_1 \exp\left(\frac{-\hat{\beta}}{t_1}\right) \exp\left(\frac{-\hat{\beta}}{t}\right). \tag{2.3.12}$$

On the other hand, the estimated lower bound (ELB) $\hat{x}_{lower}(t)$ and an estimated upper bound (EUB) $\hat{x}_{upper}(t)$ for $x(t)$ can be obtained by substituting the parameters by theirs estimators in the equations Eq. (2.3.9) and Eq. (2.3.10), the estimated CB are given by:

$$\hat{x}_{lower}(t) = \exp\left(\hat{\mu}(t_1, t, x_1) - \xi\hat{\sigma}\sqrt{(t - t_1)}\right), \tag{2.3.13}$$

and

$$\hat{x}_{upper}(t) = \exp\left(\hat{\mu}(t_1, t, x_1) + \xi\hat{\sigma}\sqrt{(t - t_1)}\right), \tag{2.3.14}$$

where $\hat{\mu}(t_1, t, x_1) = \ln(x_1) - \hat{\beta}\left(\frac{1}{t} - \frac{1}{t_1}\right) - \frac{\hat{\sigma}^2}{2}(t - t_1)$.

2.3.4 Goodness of Fit

The results according to the one-step-ahead mean absolute error (MAE), the root mean square error (RMSE) and the mean absolute percentage error (MAPE), given by the Table 2.1

Table 2.1: The expressions of MAE, RMSE, and MAPE.

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |x(t_i) - \hat{x}(t_i)|$$

$$\text{RMSE} = \left(\frac{1}{N}\sum_{i=1}^{N} (x(t_i) - \hat{x}(t_i))^2\right)^{\frac{1}{2}}$$

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N} \frac{|x(t_i) - \hat{x}(t_i)|}{x(t_i)} \times 100$$

With $x(t_i)$ is the observed values, $\hat{x}(t_i)$ is the predicted values and N is the number of predictions.

From Lewis [69], we deduce the accuracy of the forecast can be judged from the MAPE result Table 2.2. The absolute mean error in percentage MAPE is the average of the deviations in absolute value compared to the observed values. It is a practical indicator of comparison, it makes it possible to evaluate the forecasts obtained from the models. We denote by $x(t_i)$, $\hat{x}(t_i)$ and N respectively the real values, the values predicted by the model and the number of predictions, so we have: MAPE$= 0.5092\%$.

Table 2.2: Interpretation of typical MAPE values.

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

2.4 Simulation study

2.4.1 Simulated sample paths of the SSDP

In this section we present some simulated sample paths for the SSDP. To illustrate Eq. (2.2.6), we consider the equidistant time discretization of the interval $[t_1, T]$, with time points $t_i = t_{i-1} + (i-1)h$; for $i = 2, \cdots, N$ and step size $h = \dfrac{T_N - t_1}{N}$ for an integer N (N is the sample size). Therefore, 50 sample paths of the SSDP are simulated with $t_1 = 0.05$, $T = 40$, and $N = 100$. This, Figure 2.2 shows the some simulated sample paths for the SSDP for several values of σ.

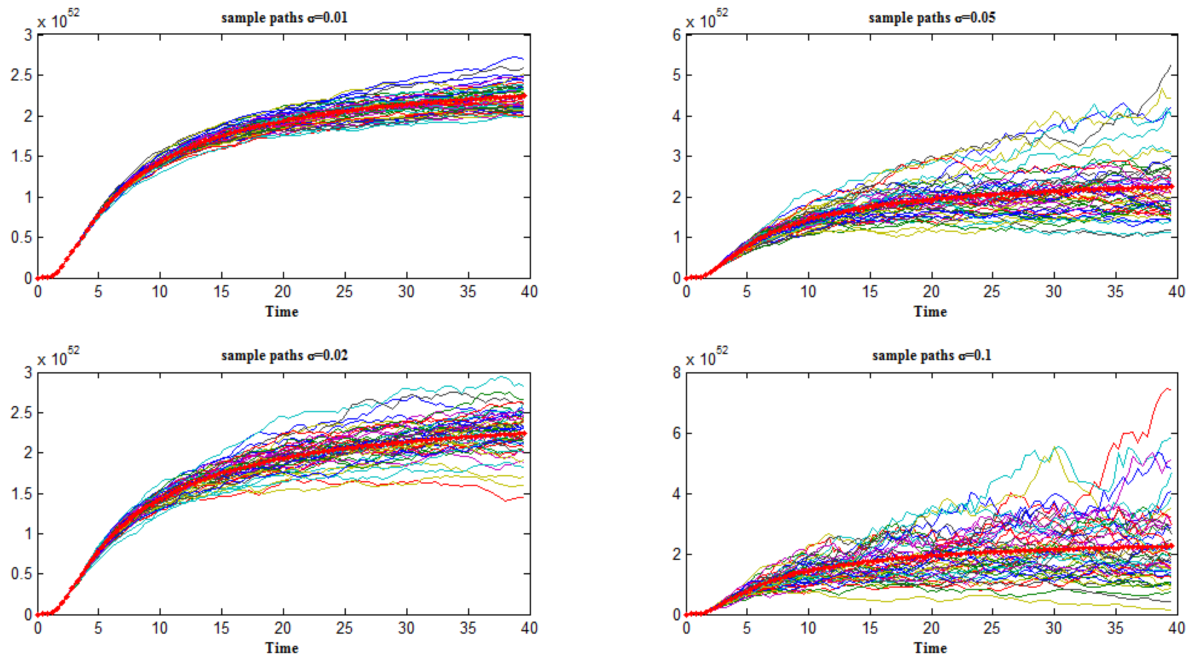


Figure 2.2: Some simulated sample paths and MF for the SSDP for several values of σ ($x_1 = 2,\ \beta = 6$).

2.4.2 Parameters estimation

This section shows different examples to validate the methodology of estimation developed earlier. To do so, we have considered four cases in which $N = 50,\ 100,\ 250,\ 500$ sample paths have been simulated. Each trajectory has been simulated with $M = 30$ data starting at $t_1 = 0.05$, $T = 40$ and $x_1 = 2$ with the step size $h = \dfrac{T - t_1}{N}$. So as to realize the subsequent inference, in every case, we consider 30 data with $t_i = t_{i-1} + (i-1)\,h$, for $i = 2, \cdots, N$.

Table 2.4 shows the results for the empirical mean of the parameters, along with the std, and the CV obtained for β and σ, given by Table 2.3.

Table 2.3: The empirical mean, std and CV expressions of the parameters β and σ.

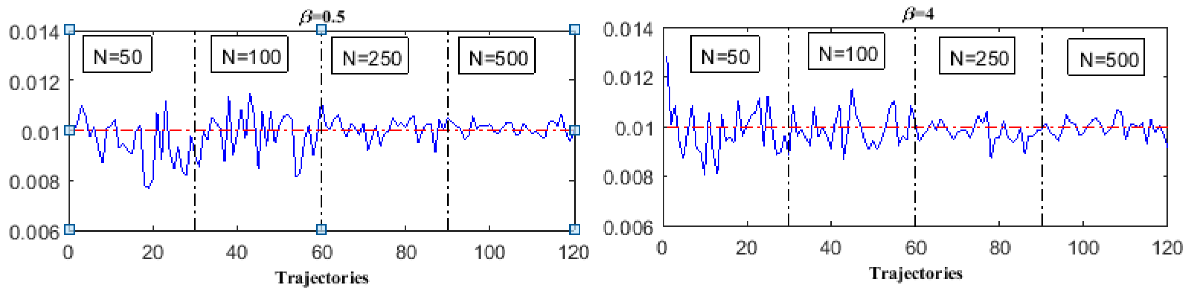
Empirical mean	std	CV
$\overline{\beta} = \frac{1}{M}\sum\limits_{i=1}^{M}\beta_i$	$\text{std}(\beta) = \left(\frac{1}{M-1}\sum\limits_{i=1}^{M}\left(\beta_i - \overline{\beta}\right)^2\right)^{\frac{1}{2}}$	$\text{CV}(\beta) = \frac{\text{std}(\beta)}{\overline{\beta}}$
$\overline{\sigma} = \frac{1}{M}\sum\limits_{i=1}^{M}\sigma_i$	$\text{std}(\sigma) = \left(\frac{1}{M-1}\sum\limits_{i=1}^{M}\left(\sigma_i - \overline{\sigma}\right)^2\right)^{\frac{1}{2}}$	$\text{CV}(\sigma) = \frac{\text{std}(\sigma)}{\overline{\sigma}}$

Table 2.4: Data estimation, std, and CV of the parameters β and σ for different values of σ.

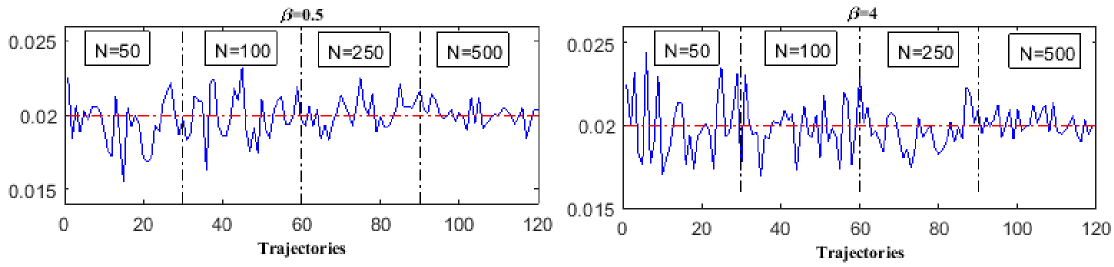
σ	N	$\overline{\beta}$	$\overline{\sigma}$	std(β)	std(σ)	CV(β)	CV(σ)
$\beta = 0.5$							
0.01	50	0.50002	0.00946	4.5207×10^{-4}	9.4215×10^{-4}	0.00090	0.09957
	100	0.499877	0.00991	3.1194×10^{-4}	9.1600×10^{-4}	0.00062	0.09242
	250	0.50003	0.01006	2.4639×10^{-4}	3.9334×10^{-4}	0.00049	0.03907
	500	0.50001	0.01007	2.1700×10^{-4}	2.6847×10^{-4}	0.00043	0.02665
0.02	50	0.50022	0.01939	9.1465×10^{-4}	16.466×10^{-4}	0.00182	0.08488
	100	0.50005	0.01990	6.6178×10^{-4}	16.331×10^{-4}	0.00132	0.08205
	250	0.49997	0.02016	4.5609×10^{-4}	10.263×10^{-4}	0.00091	0.05089
	500	0.49999	0.02001	4.3199×10^{-4}	6.6969×10^{-4}	0.00086	0.03346
0.05	50	0.49983	0.04918	23.782×10^{-4}	43.861×10^{-4}	0.00476	0.08918
	100	0.49976	0.04999	15.771×10^{-4}	34.237×10^{-4}	0.00315	0.06848
	250	0.49986	0.05015	11.246×10^{-4}	21.876×10^{-4}	0.00225	0.04362
	500	0.50014	0.04948	11.395×10^{-4}	15.702×10^{-4}	0.00228	0.03173
0.1	50	0.50119	0.09658	43.287×10^{-4}	110.89×10^{-4}	0.00863	0.11480
	100	0.49955	0.09932	38.344×10^{-4}	61.907×10^{-4}	0.00767	0.06233
	250	0.50007	0.10025	30.369×10^{-4}	51.549×10^{-4}	0.00607	0.05142
	500	0.49981	0.10026	28.915×10^{-4}	31.016×10^{-4}	0.00578	0.03093
$\beta = 4$							
0.01	50	3.99988	0.00982	4.4977×10^{-4}	10.338×10^{-4}	0.00011	0.10521
	100	4.00005	0.00990	3.8429×10^{-4}	7.0099×10^{-4}	0.00009	0.07077
	250	4.00003	0.00976	3.4307×10^{-4}	4.0839×10^{-4}	0.00008	0.04184
	500	3.99998	0.00994	2.6908×10^{-4}	3.6431×10^{-4}	0.00006	0.03664
0.02	50	4.00006	0.01994	12.166×10^{-4}	21.326×10^{-4}	0.00031	0.10693
	100	3.99982	0.019875	7.3735×10^{-4}	15.245×10^{-4}	0.00018	0.076709
	250	4.00009	0.01962	4.6337×10^{-4}	11.344×10^{-4}	0.00011	0.05782
	500	3.99993	0.02011	4.4082×10^{-4}	6.7713×10^{-4}	0.00011	0.03366
0.05	50	4.00032	0.04943	26.337×10^{-4}	46.066×10^{-4}	0.00065	0.09318
	100	4.00057	0.05055	16.196×10^{-4}	31.104×10^{-4}	0.00040	0.06152
	250	3.99974	0.04976	13.230×10^{-4}	22.499×10^{-4}	0.00033	0.04521
	500	3.99995	0.04946	10.856×10^{-4}	16.170×10^{-4}	0.00027	0.03269
0.1	50	3.99821	0.09881	33.551×10^{-4}	103.35×10^{-4}	0.00084	0.10460
	100	3.99997	0.10104	43.799×10^{-4}	84.672×10^{-4}	0.00109	0.08379
	250	3.99998	0.09800	20.167×10^{-4}	43.620×10^{-4}	0.00050	0.04451
	500	3.99975	0.10107	21.812×10^{-4}	27.656×10^{-4}	0.00054	0.02736

Figure 2.3 shows the evolution of $\sigma = 0.01, 0.02, 0.05$ and 0.1 computed for the sample sizes $N = 50$, 100, 250, and 500 for several values of β. Results show, as could be predicted, that the coefficient of variation decreases as σ does and the sample size increases.

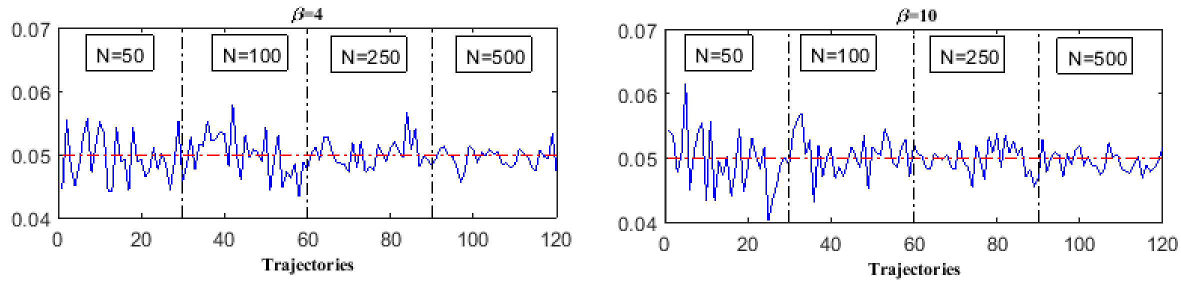
Figure 2.4 shows the evolution of $\beta = 0.5$ and $\beta = 4$ computed for the sample sizes $N = 50$, 100, 250, and 500 for several values of σ.



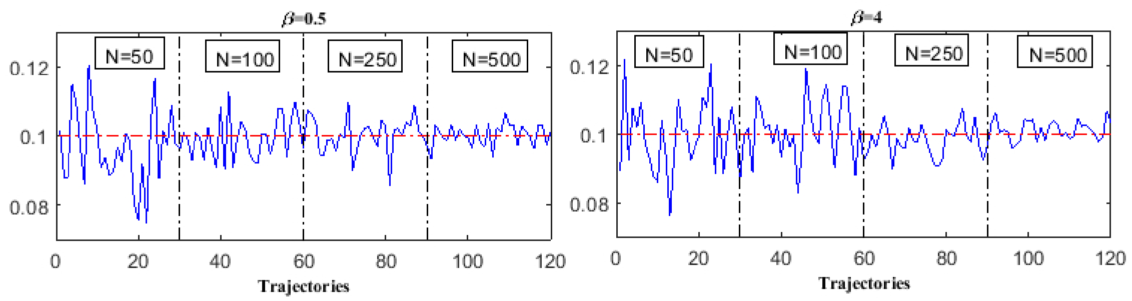
(a) The case $\sigma = 0.01$



(b) The case $\sigma = 0.02$

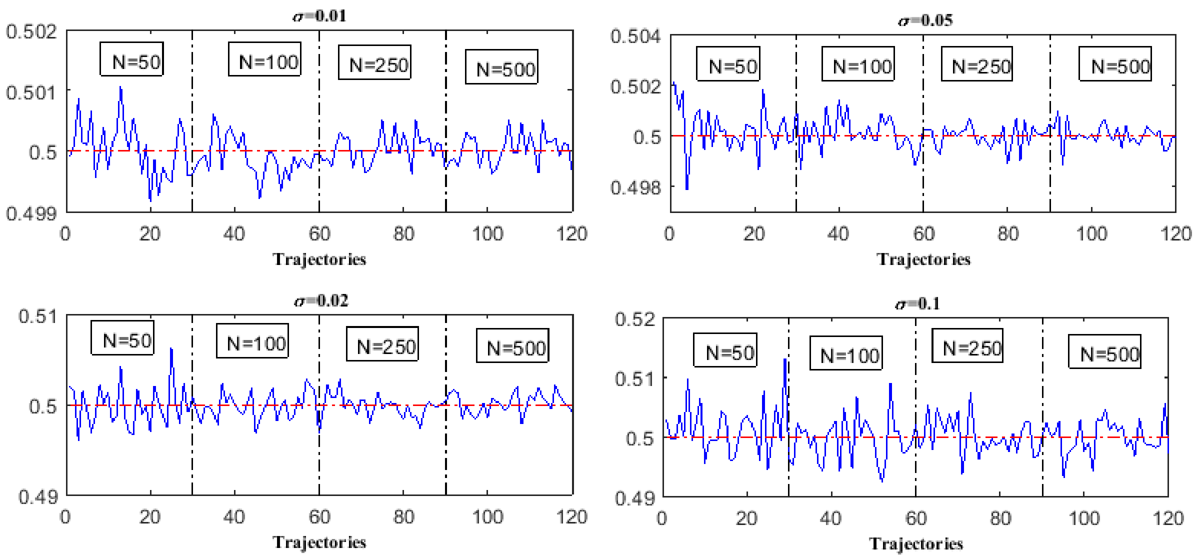


(c) The case $\sigma = 0.05$

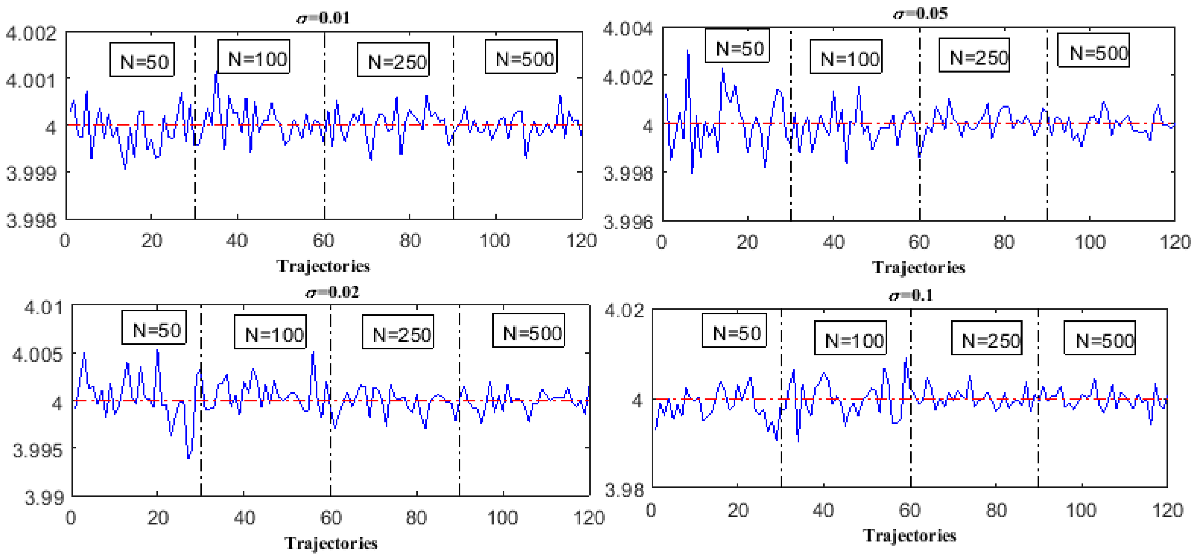


(d) The case $\sigma = 0.1$

Figure 2.3: Evolution of σ computed for the sample sizes N =50, 100, 250, and 500 for several values of β.



(a) The case $\beta = 0.5$



(b) The case $\beta = 4$

Figure 2.4: Evolution of β computed for different sample sizes N =50, 100, 250, and 500 for several values of σ.

2.4.3 Prediction

In this section, we have considered two examples in which $N = 20$ sample paths have been simulated. Each trajectory has been simulated with the discritization times $t_i = t_{i-1} + (i-1)\,h$, for $i = 2, \cdots, N$ beginning with $t_1 = 0.05$, $x_1 = 2$ and $h = 0.2$. The procedure is applied as follows: we first use the first 17 data for estimating β and σ^2, using the expressions (2.3.6) and (2.3.8). Then, we find the associated EMF and ECMF expressed by the expressions (2.3.11) and (2.3.12). The associated values of the 3 last data are predicted by the EMF and ECMF. In addition, the results obtained are associated an estimated CB (ECB) of 95% and an estimated conditional CB (ECCB) of 95% (for more details, see, section 2.3.2). To highlight the performance of the method, the results are attached to the MAE, RMSE and MAPE shown in the Tabale 2.1.

From Lewis [69], we deduce the accuracy of the forecast can be judged from the MAPE result Table 2.2.

Table 2.5 presents the achieved results for the estimated MF, CMF, CB and CCB of the model.

Table 2.6 shows the estimation of the parameters of the process.

Table 2.7 shows the goodness of fit of the process. The accuracy of the forecast can be judged from the MAPE result is less than 10%, showing the forecast to be highly accurate.

The performance of the SSDP for the forecasting using the trend function and the conditional trend function for the data is illustrated in Figure 2.5 and Figure 2.6.

Table 2.5: Simulated data, estimated MF, CMF, CB and CCB results.

i	Simulated data $x(t_i)$	EMF	ECMF	ECB	ECCB
1	2.0000000	2.0000000	2.0000000	[2.0000000 2.0000000]	[2.0000000 2.0000000]
2	5968.9190	5969.8432	5969.8432	[5902.7241 6023.7604]	[5902.7241 6023.7604]
3	14508.920	14523.343	14521.094	[14283.269 14699.216]	[14357.833 14652.242]
4	20505.799	20444.100	20423.797	[20018.212 20734.504]	[20194.172 20608.256]
5	24655.497	24501.145	24575.088	[23897.904 24888.013]	[24298.789 24797.040]
6	27655.915	27406.584	27579.239	[26636.982 27873.814]	[27269.165 27828.324]
7	29758.227	29576.686	29845.759	[28650.756 30111.254]	[29510.202 30115.313]
8	31200.534	31254.645	31446.485	[30180.975 31846.088]	[31092.931 31730.496]
9	32435.022	32588.916	3.2532.495	[31374.774 33228.731]	[32166.731 32826.314]
10	33661.894	33674.393	33515.373	[32325.958 34355.558]	[33138.559 33818.070]
11	34870.550	34574.275	34561.442	[33096.846 35290.899]	[34172.866 34873.586]
12	35474.898	35332.165	35634.934	[33730.333 36079.237]	[35234.289 35956.774]
13	36231.504	35979.068	36124.414	[34256.851 36752.272]	[36380.293 36450.675]
14	37121.920	36537.615	36793.969	[34698.558 37333.189]	[36389.494 37126.277]
15	37233.678	37024.695	37616.789	[35071.955 37839.330]	[37193.862 37956.528]
16	37732.066	37453.165	37664.566	[35389.574 38283.925]	[37241.102 38004.736]
17	38140.877	37832.978	38114.707	[35661.094 38677.239]	[37686.183 38458.943]
Prediction					
18	38230.634	38171.962	38482.619	[35894.108 39027.360]	[38049.958 38830.178]
19	38670.220	38476.351	38535.492	[36094.653 39340.747]	[38102.236 38883.528]
20	39008.842	38751.178	38946.431	[36267.585 39622.622]	[38508.555 39298.179]

Table 2.6: Estimation of the parameters of the SSDP.

Parameters	Estimated value
β	0.500083048495758
σ	0.011578353626586

Table 2.7: Goodness of fit of the SSDP.

MAE	RMSE	MAPE
173.2253	222.1143	0.5092%

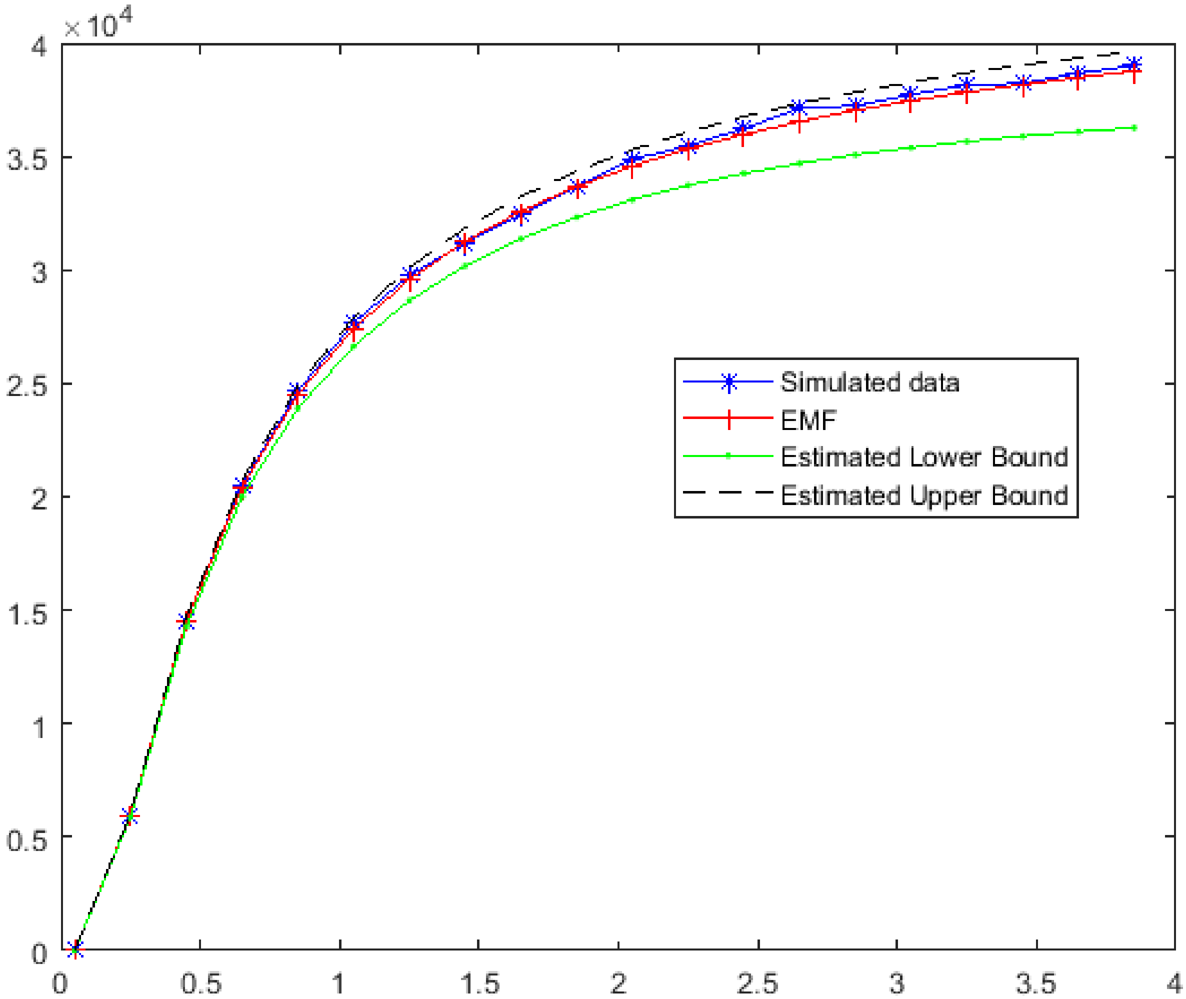


Figure 2.5: Simulated data, EMF, ELB and EUB.

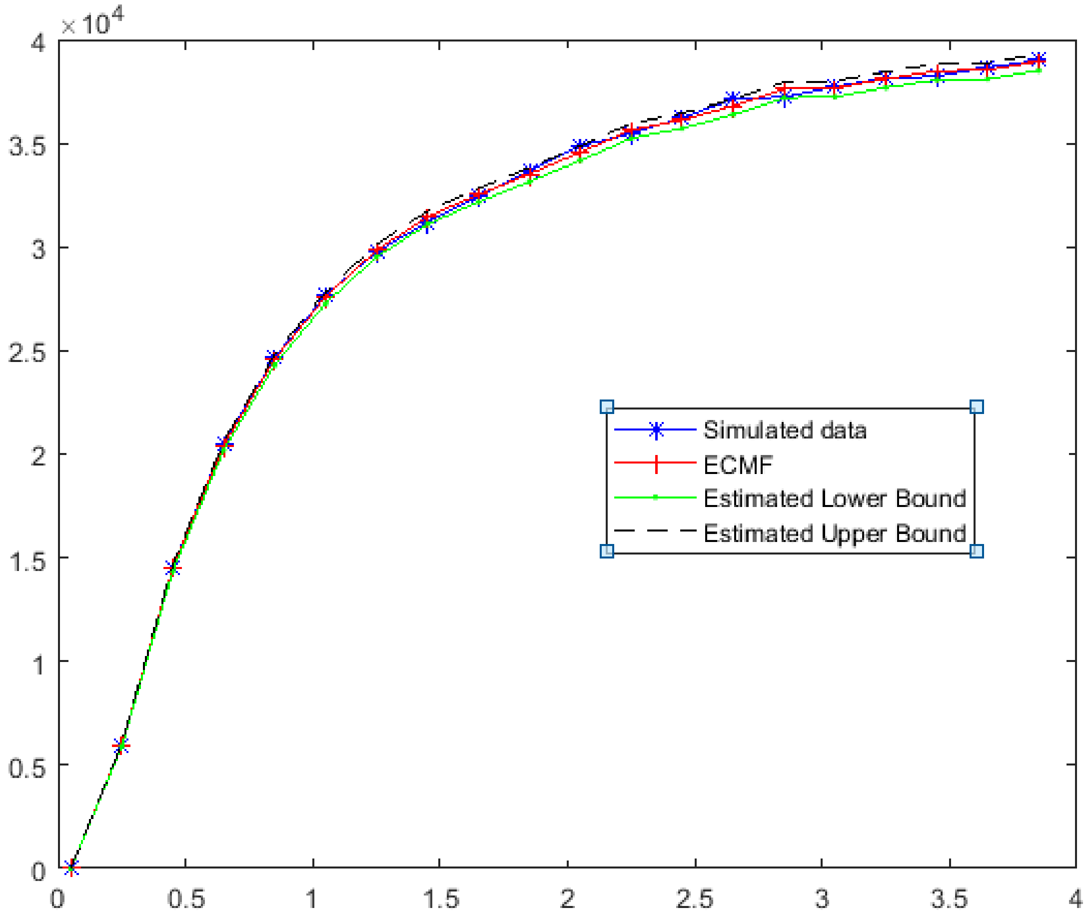


Figure 2.6: Simulated data, ECMF, ECLB and ECUB.

2.5 Conclusions

The resulting stochastic Schumacher diffusion process is advantageous compared to the deterministic Schumacher curve, which is widely used in several research areas. So, this diffusion process makes it possible to introduce all the information coming from the data into the model, in the same way as the random factors that need to be taken into consideration for clarify the various phenomena.

This article introduces a new diffusion process linked to the Schumacher curve. Its main probability characteristics were examined, and its unconditional and conditional mean functions was proved to be Schumacher curve.

The inferential approach is performed on the discrete sampling using the maximum likelihood principle based on a discretization of the exact analytical solution of the process.

Finally, an application of the proposed model to simulated data has shown its effectiveness in practice, and the ability of the process to predict and predict is demonstrated. Consequently, the variability of this study, could be generalized in the upcoming studies such that the application to forestry and in other areas.

CHAPTER 3

The stochastic Lundqvist-Korf diffusion process and its extended diffusion processes

Contents

This chapter is the complete version of the paper [28] published in Mathematics and Computers in Simulation, vol. 182, 25-38, 2021 and the paper [70] published in Moroccan Journal of Pure and Applied Analysis, vol. 8(3), 364-374, 2022.

3.1 Introduction

Stochastic diffusion models have extensive areas of applications. They have been the object of particular attention in diverse fields of sciencesuch as biological phenomena, energy consumption, economy and finance, and environmental phenomena, see for example, Gutiérrez et al. [17,20,71], Román-Román et al. [3], and Nafidi et al. [60], the process generally being defined by means of stochastic differential equations.

The Korf growth function was initially proposed by Václav Korf [72] and is considered as a particular mathematical formulation of trees and forest populations over time. It was belonged to the power decline type according to Zeide [73]. Zarnovican [74] introduced a briefly analysed of Korf's mathematical formula and applied it to forest mensuration such the relation between height and age for three black spruce stands. Sghaier et al. [75] introduced a six generalized algebraic difference equations derived from the base models of log-logistic, Bertalanffy-Richards, and Lundqvist-Korf were used to develop site index model for Pinus pinea plantations in north-west of Tunisia and concludes a generalized algebraic difference equation derived from the base model of Lundqvist-Korf realized the best compromise between biological and statistical constraints, producing the most adequate site index curves. Amaro et al. [76] suggested the difference forms of the Richards and Lundqvist-Korf growth equations. Sánchez-González et al. [77] tested and fitted using the generalized least squares regression method the difference forms of the Lundqvist-Korf, McDill-Amateis and Richards growth functions. Martins et al. [78] established that the Lundqvist–Korf model provided the most accurate estimates for diameter and height growth, in comparison with the other models, providing better statistical values, greater proximity to observed values and better distribution of residual percentages. Krisnawati et al. [79] established that the Gompertz, Chapman-Richards, Weibull, modified logistic, Lundqvist-Korf and exponential models provided very similar results in term of the resulting fit and the prediction error (bias) with the Lundqvist-Korf model being slightly better than the other. Crescenzo and Spina [80] presented an analysis of a growth model inspired by Gompertz and Korf laws, and an analogous birth-death process. This model may be used when only diameter measurements are available, although the error of prediction can be relatively large.

In literature, there are many types of diffusion process models, for example, the stochastic Lognormal diffusion process [11, 12], the stochastic logistic diffusion process [3], the stochastic Gompertz diffusion process [14–18], the stochastic Rayleigh diffusion process ([20, 21]), the stochastic Bertalanfly diffusion process [25], the stochastic gamma diffusion process [4], the stochastic Richards-type diffusion process [5], the stochastic Hubbert diffusion process [6] and the stochastic Weibull diffusion process [54].

Various authors, approached the problem of Maximum likelihood estimation by equating partial derivatives of the log-likelihood function to zero, looking for a stationary point of the local maximum likelihood equations by iterative methods (see for example, [45–49]). Vera and Díaz-García [56] proposed the global simulated annealing heuristic for the three-parameter log-normal maximum likelihood estimation. Román-Román et al. [55] used the simulated annealing algorithm to estimate the parameters of a Gompertz-type diffusion process. Pedersen [57] suggested a new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations when the likelihood function is unknown.

The approach of maximum likelihood estimation of the parameters using likelihood equations can be problematic, which is why we propose the use of the simulated annealing algorithm (SA). This algorithm is a method for solving unconstrained and bound-constrained optimization problems developed by Kirkpatrick et al. [58]. The method models the physical process of heating a material and then slowly lowering the temperature to decrease defects, thus minimizing the system energy. It has been successfully used for optimization in continuous spaces (see, Duflo [81]). In the last years many works have used the simulated annealing algorithm (see, for example, Nafidi et al. [54], Istoni et al. [6] and Román-Román et al. [5]).

In the present chapter, we propose two new stochastic diffusion processes. One is the stochastic lundqvist-korf diffusion process (SLKDP), which presents a mean function that is proportional to the Lundqvist-Korf growth curve. The other is the γ power of the Lundqvist-Korf diffusion process (γ-SLKDP). The rest of this paper is organized as follows: In the section 3.2, we present an overview of the Lundqvist-Korf growth curve and we define the proposed model in terms of stochastic differential equation (SDE). We then determine the explicit expression of the solution to the SDEs, the transition probability density function (TPDF) and the mean functions. The estimation of parameters are discussed using the maximum likelihood (ML) method on the basis of discrete sampling of the process. Since the closed form of the ML estimators cannot be given because the system of likelihood equations does not have an explicit form, numerical methods are needed. In fact, the SA algorithm is proposed, either carried through to convergence or terminated after reaching a given stop criterion, to calculate or approximate the resulting ML estimator. The results obtained are presented from the examples of simulation and we illustrate the predictive study by fitting the diffusion process to simulated data. In the section 3.3, we study a new family non-homogeneous stochastic γ-power Lundqvist-Korf diffusion process, defined from a non-homogeneous Lundqvist-Korf diffusion process. First, we determine the probabilistic characteristics of the process, such as its analytic expression, the transition probability density function from the corresponding Itô stochastic differential equation and obtain the conditional and non-conditional mean functions. We then study the statistical inference in this process. The parameters of this process are estimated by using the maximum likelihood estimation method with discrete sampling, thus we obtain a nonlinear equation, which is achieved via the simulated annealing algorithm. Finally, the results of the paper are applied to simulated data. In the last section, we summarise the main conclusions.

3.2 The SLKDP model

3.2.1 An overview of the Lundqvist-Korf growth curve

As is well known, the lundqvist-Korf curve represent a generalization of the Schumacher curve that has been proposed by Schumacher [62]. In fact, the curve mentioned above is based on the hypothesis that the relative growth rate has a linear relationship with the inverse square in the time

$$\frac{1}{y(t)}\frac{dy(t)}{dt} = \frac{\beta}{t^2}; \quad t > 0,\ \beta > 0. \tag{3.2.1}$$

By integrating equation (3.2.1) with respect to t and assuming that $y(t_1) = y_1$, thus, we have

$$y(t) = y_1 \exp\left(\frac{\beta}{t_1}\right) \exp\left(\frac{-\beta}{t}\right); \quad t > 0,\ \beta > 0, \tag{3.2.2}$$

where β is scale parameter. Then, the expression of Lundqvist-Korf growth curve is

$$x(t) = x_\infty \exp\left(\frac{-\beta}{t^\alpha}\right); \quad t > 0,\ \alpha > 0,\ \beta > 0,$$

where x_∞ is the upper bound for the the studied variable, that can only be reached after infinity time and β is the scale parameter.

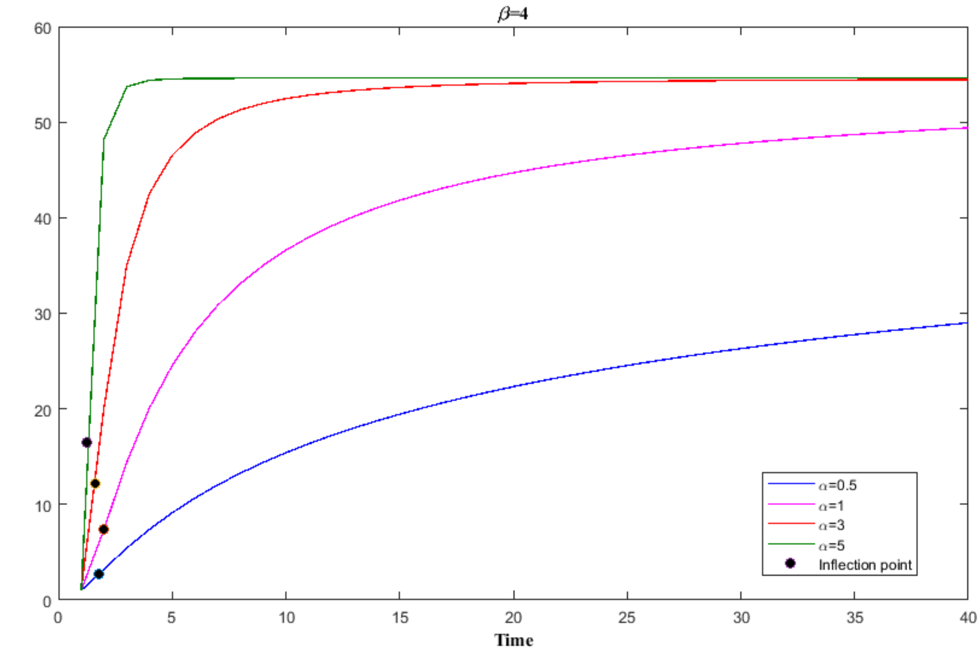
We impose that $x(t_1) = x_1 > 0$, hence $x_\infty = x_1 \exp\left(\frac{\beta}{t_1^\alpha}\right)$. Thus, we reach

$$x(t) = x_1 \exp\left(\frac{\beta}{t_1^\alpha}\right) \exp\left(\frac{-\beta}{t^\alpha}\right); \quad t \geq t_1,\ \alpha > 0,\ \beta > 0, \tag{3.2.3}$$

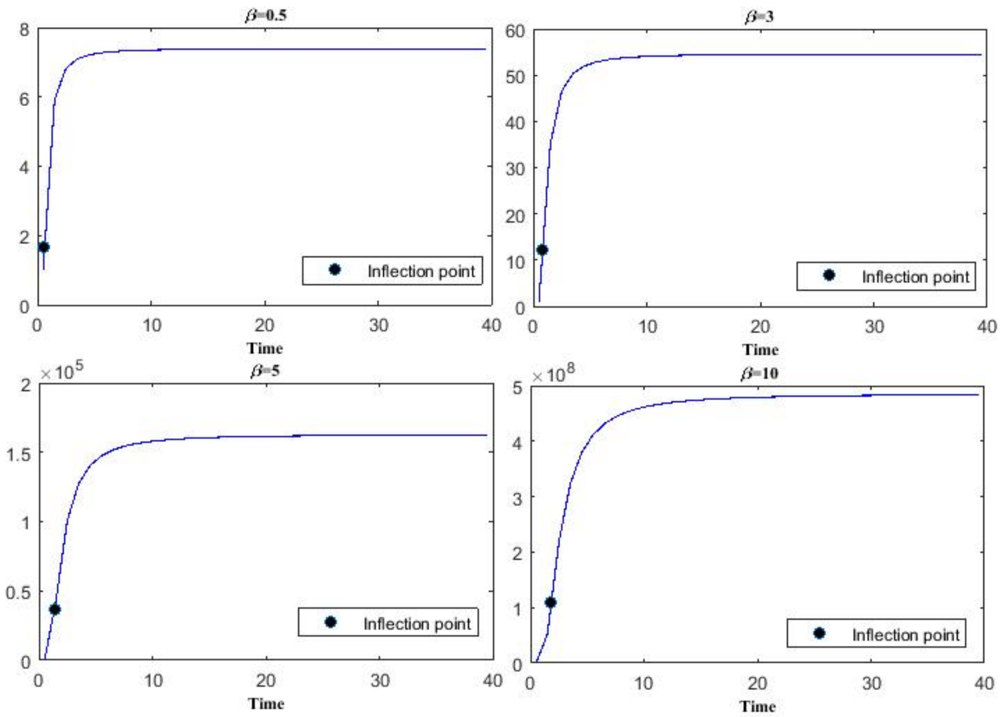
Moreover, the Lundqvist-Korf growth curve (3.2.3), verifies the following properties:

1. It is strictly increasing.
2. We have $\lim_{t\to\infty} x(t) = x_1 \exp\left(\frac{\beta}{t_1^\alpha}\right)$, so the line $y = x_1 \exp\left(\frac{\beta}{t_1^\alpha}\right)$ is a horizontal asymptote of the curve (3.2.3) when t tend to ∞.
3. It shows an inflection point at $t_I = \left(\frac{\alpha\beta}{\alpha+1}\right)^{\frac{1}{\alpha}}$, and the value of $x(t)$ at t_I verifying $x(t_I) = x_1 \exp\left(\frac{\beta}{t_1^\alpha}\right) \exp\left(-\frac{\alpha+1}{\alpha}\right)$.
4. In addition, this point verify $t_I > t_1$. Furthermore, we have $t_I > t_1$ if and only if $\beta > \left(1 + \frac{1}{\alpha}\right) t_1^\alpha > t_1^\alpha$.

Figure 3.1 show the Lundqvist-Korf curve and its inflection point for several values of α and β.



(a) The case $\beta = 4$.



(b) The case $\alpha = 2$.

Figure 3.1: The Lundqvist-Korf curve for several values of α (a) and β (b) for $x_1 = 1$, $t_1 = 1$, and $T = 40$.

3.2.2 The SLKDP model

In order to model the Lundqvist-Korf type behaviors from a stochastic point of view, our contribution is to consider a diffusion process whose mean function has the expression given in (3.2.3). Now, starting from (3.2.3), one is lead to considering the ODE

$$\frac{dx(t)}{dt} = r(t)x(t); \quad x(t_1) = x_1, \tag{3.2.4}$$

where $r(t) = \frac{\alpha\beta}{t^{\alpha+1}}$. Hence (3.2.4) can be viewed as a generalisation of the Malthusian growth model with time dependent fertility rate $r(t)$. Note that $r(t)$ is a decreasing continuous positive function and has a horizontal asymptote at $y = 0$ and a vertical asymptote at $t = 0$.

The stochastic version of the model is given by the following diffusion process $\{x(t) : t \in [t_1, T],\ t_1 > 0\}$, taking values on $(0, \infty)$ and characterized by infinitesimal moments

$$\begin{aligned} A_1(t,x) &= \left(\frac{\alpha\beta}{t^{\alpha+1}}\right) x, \\ A_2(t,x) &= \sigma^2 x^2, \end{aligned} \tag{3.2.5}$$

with $\alpha > 0$, $\beta > 0$, $\sigma > 0$ and initial distribution $x(t_1)$.

Alternatively, the above process can be defined as the unique solution of the following Itô's SDE:

$$dx(t) = A_1(t,x)dt + (A_2(t,x))^{\frac{1}{2}}\, dw(t); \quad x(t_1) = x_1, \tag{3.2.6}$$

where $w(t)$ is a standard Wiener process and x_1 is a positive random variable, independent on $w(t)$ for $t \geq t_1$.

In fact, the infinitesimal moments $A_1(t,x)$ and $A_2(t,x)$ specified in Eq. (3.2.5) satisfy the Lipschitz and the growth conditions for the existence and unicity of the solution to the SDEs (see, Kloeden et al. [8]). Thus, there exists a non negative constant $C = \frac{\alpha\beta}{t_1^{\alpha+1}} + \sigma$, such that for all $x, y \in \mathbb{R}^+$ and $t \in [t_1, T]$, we have:

$$\begin{aligned} &|A_1(t,x) - A_1(t,y)| + \left|(A_2(t,x))^{\frac{1}{2}} - (A_2(t,y))^{\frac{1}{2}}\right| \leq C \mid x - y \mid, \\ &|A_1(t,x)|^2 + \left|(A_2(t,x))^{\frac{1}{2}}\right|^2 \leq C^2 \left(1 + \mid x \mid^2\right). \end{aligned}$$

Then, the SDE Eq. (3.2.6) has a unique solution $\{x(t) : t \in [t_1, T],\ t_1 > 0\}$ continuous with probability 1, and satisfies the initial condition $x(t_1) = x_1$.

3.2.3 Basic probabilistic characteristics of the SLKDP

3.2.3.1 Probability distribution from SDE

By means of the appropriate transformation of the form $y(t) = \ln(x(t))$, and by using the Itô rule, the SDE Eq. (3.2.6) becomes

$$dy(t) = \left(\frac{\alpha\beta}{t^{\alpha+1}} - \frac{\sigma^2}{2}\right) dt + \sigma dw(t); \quad y_1 = \ln(x_1),$$

by integrating both sides yields,

$$y(t) = y_1 - \beta\left(\frac{1}{t^\alpha} - \frac{1}{t_1^\alpha}\right) - \frac{\sigma^2}{2}(t - t_1) + \sigma\left(w(t) - w(t_1)\right).$$

Finally, we have:

$$x(t) = x_1 \exp\left[-\beta\left(\frac{1}{t^\alpha} - \frac{1}{t_1^\alpha}\right) - \frac{\sigma^2}{2}(t - t_1) + \sigma\left(w(t) - w(t_1)\right)\right]. \tag{3.2.7}$$

The $y(t)$ is a gaussian process if and only if $\ln(x_1)$ is constant or normally distributed (see, Arnold [82]). In such a case, the $x(t)$ is a Lognormal process. That is, the TPDF

$$f(x, t|y, s) = \frac{1}{x\sqrt{2\pi\sigma^2(t-s)}} \exp\left(-\frac{\left[\ln\left(\frac{x}{y}\right) + \beta\left(\frac{1}{t^\alpha} - \frac{1}{s^\alpha}\right) + \frac{\sigma^2}{2}(t-s)\right]^2}{2\sigma^2(t-s)}\right). \tag{3.2.8}$$

3.2.3.2 Probability distribution from partial differential equations

We consider the Fokker-Plank equation of the non-homogeneous diffusion process with infinitesimal moments $A_1(t, x)$ and $A_2(t, x)$ in Eq. (3.2.5)

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial x}\left[A_1(t, x) f\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[A_2(t, x) f\right], \tag{3.2.9}$$

hence

$$\frac{\partial f}{\partial t} = -\left(\frac{\beta\alpha}{t^{\alpha+1}}\right)\frac{\partial}{\partial x}[xf] + \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}\left[x^2 f\right]; \quad x > 0,\ t \geq t_1, \tag{3.2.10}$$

and the Kolmogorov or bakward equation

$$\frac{\partial f}{\partial s} + \left(\frac{\beta\alpha}{s^{\alpha+1}} y\right)\frac{\partial f}{\partial y} + \left(\frac{\sigma^2}{2} y^2\right)\frac{\partial^2 f}{\partial y^2}; \quad y > 0,\ s \geq t_1, \tag{3.2.11}$$

where $f = f(x, t|y, s)$ is the transition probability density of this process and with the initial condition $\lim_{t \to t_1} f(x, t|y, s) = \delta(x - y)$ and $\delta(.)$ is the Dirac delta function.

We change the bakward Eq. (3.2.11) into that of the Wiener process (see, Ricciardi [83]) can be obtained by means of the following transformation

$$x' = \psi(x, t),$$

$$t' = \phi(t).$$

The infinitesimal moments (3.2.5) verify the conditions of the theorem proposed by Ricciardi [83], and so this transformation exists

$$\psi(x, t) = \frac{1}{\sigma}\ln(x) + \frac{1}{\sigma}\frac{\beta}{t^\alpha} + \frac{\sigma}{2}t,$$

$$\phi(t) = t.$$

This transformation allows to obtain the TPDF

$$f(x,t|y,s)=\frac{1}{x\sqrt{2\pi\sigma^2(t-s)}}\exp\left(-\frac{[\ln\left(\frac{x}{y}\right)+\beta\left(\frac{1}{t^\alpha}-\frac{1}{s^\alpha}\right)+\frac{\sigma^2}{2}(t-s)]^2}{2\sigma^2(t-s)}\right), \tag{3.2.12}$$

which corresponds to a lognormal distribution, that is

$$[x(t)|x(s)=x_s]\sim\Lambda_1\left(\mu\left(s,t,x_s\right);\sigma^2(t-s)\right), \tag{3.2.13}$$

where $\mu\left(s,t,x_s\right)=\ln\left(x_s\right)-\beta\left(\frac{1}{t^\alpha}-\frac{1}{s^\alpha}\right)-\frac{\sigma^2}{2}(t-s)$.

3.2.3.3 Computation of the MFs

By the properties of the log-normal distribution, we obtain the conditional moments of order r of this process:

$$\begin{aligned}E\left(x^r(t)|x(s)=x_s\right) &= \exp\left(r\mu\left(s,t,x_s\right)+\frac{r^2\sigma^2}{2}(t-s)\right)\\ &= \exp\left(\ln\left(x_s^r\right)-\beta r\left(\frac{1}{t^\alpha}-\frac{1}{s^\alpha}\right)-\frac{r\sigma^2}{2}(t-s)+\frac{r^2\sigma^2}{2}(t-s)\right)\\ &= x_s^r\exp\left(-\beta r\left(\frac{1}{t^\alpha}-\frac{1}{s^\alpha}\right)\right)\times\exp\left(\frac{r(r-1)\sigma^2}{2}(t-s)\right).\end{aligned}$$

Then, for $r=1$, the conditional mean function (CMF) of the process is

$$E\left(x(t)|x(s)=x_s\right)=x_s\exp\left(-\beta\left(\frac{1}{t^\alpha}-\frac{1}{s^\alpha}\right)\right). \tag{3.2.14}$$

In addition, taking into the initial condition $\mathrm{P}\left(x\left(t_1\right)=x_1\right)=1$, the mean function (MF) of the process is given by

$$E(x(t))=x_1\exp\left(\frac{\beta}{t_1^\alpha}\right)\exp\left(\frac{-\beta}{t^\alpha}\right). \tag{3.2.15}$$

3.2.4 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 Zehna's theorem [68], the corresponding for the aforementioned parametric functions.

3.2.4.1 Parameters estimation

We consider a discrete sampling of the process, that is, for fixed times $t_1,t_2,\cdots,t_n$, $(n>2)$, we observe the variables $x\left(t_1\right)$, $x\left(t_2\right)$, $\cdots$, $x\left(t_n\right)$ whose values provide the basic sample for the inference process. In addition, we assume $t_i-t_{i-1}=h$, for $i=2,\cdots,n$. Let $x_1,x_2,\cdots,x_n$

be the observed values of the sampling. The likelihood function (LF) depends on the choice of the initial distribution. If $\mathrm{P}\left(x\left(t_1\right)=x_1\right)=1$, the associated LF is

$$\mathbf{L}=L_{x_1,x_2,\cdots,x_n}\left(\alpha,\beta,\sigma^2\right)=\prod_{i=2}^{n} f\left(x_i,t_i|x_{i-1},t_{i-1}\right),$$

which is written as

$$\mathbf{L}=\prod_{i=2}^{n}\frac{1}{x_i\sqrt{2\pi\sigma^2 h}}\exp\left(-\frac{\left\{\ln\left(\frac{x_i}{x_{i-1}}\right)+\beta\left(\frac{1}{t_i^\alpha}-\frac{1}{t_{i-1}^\alpha}\right)+\frac{\sigma^2}{2}h\right\}^2}{2\sigma^2 h}\right).$$

where α, β, and σ^2 are the parameters to be estimated. If $x\left(t_1\right)\sim\Lambda_1\left(\mu_1,\sigma_1^2\right)$, the associated LF is

$$\mathbf{L}=L_{x_1,x_2,\cdots,x_n}\left(\mu_1,\sigma_1^2,\alpha,\beta,\sigma^2\right)=f_{x(t_1)}\left(x_1\right)\prod_{i=2}^{n} f\left(x_i,t_i|x_{i-1},t_{i-1}\right),$$

which is written as

$$\begin{aligned}\mathbf{L}=&\frac{1}{x_1\sqrt{2\pi\sigma_1^2}}\exp\left(-\frac{\left(\ln\left(x_1\right)-\mu_1\right)^2}{2\sigma_1^2}\right)\\&\times\prod_{i=2}^{n}\frac{1}{x_i\sqrt{2\pi\sigma^2 h}}\exp\left(-\frac{\left\{\ln\left(\frac{x_i}{x_{i-1}}\right)+\beta\left(\frac{1}{t_i^\alpha}-\frac{1}{t_{i-1}^\alpha}\right)+\frac{\sigma^2}{2}h\right\}^2}{2\sigma^2 h}\right).\end{aligned}\tag{3.2.16}$$

In the following, we will consider the case when the initial distribution is lognormal. From (3.2.16), the log-likelihood function (LLF) of the sample is

$$\begin{aligned}\ln(\mathbf{L})=&-\frac{n}{2}\ln(2\pi)-\frac{1}{2}\ln\left(\sigma_1^2\right)-\frac{(n-1)}{2}\ln\left(\sigma^2\right)-\sum_{i=1}^{n}\ln\left(x_i\right)-\frac{1}{2\sigma_1^2}\left(\ln\left(x_1\right)-\mu_1\right)^2\\&-\frac{(n-1)}{2}\ln\left(h\right)-\frac{1}{2\sigma^2 h}\sum_{i=2}^{n}\left\{\ln\left(\frac{x_i}{x_{i-1}}\right)+\beta\left(\frac{1}{t_i^\alpha}-\frac{1}{t_{i-1}^\alpha}\right)+\frac{\sigma^2}{2}h\right\}^2.\end{aligned}\tag{3.2.17}$$

By deriving the LLF (3.2.17) with respect to μ_1, σ_1^2, σ^2, β and α we obtain

$$\frac{\partial\ln(\mathbf{L})}{\partial\mu_1}=\frac{1}{\sigma_1^2}\left(\ln\left(x_1\right)-\mu_1\right);\qquad\frac{\partial\ln(\mathbf{L})}{\partial\sigma_1^2}=-\frac{1}{2\sigma_1^2}+\frac{1}{2\sigma_1^4}\left(\ln\left(x_1\right)-\mu_1\right)^2.\tag{3.2.18}$$

$$\begin{aligned}\frac{\partial\ln(\mathbf{L})}{\partial\sigma^2}=&-\frac{(n-1)}{2\sigma^2}+\frac{1}{2\sigma^4 h}\sum_{i=2}^{n}\left[\ln\left(\frac{x_i}{x_{i-1}}\right)+\beta\left(\frac{1}{t_i^\alpha}-\frac{1}{t_{i-1}^\alpha}\right)+\frac{\sigma^2}{2}h\right]^2\\&-\frac{1}{2\sigma^2}\sum_{i=2}^{n}\left[\ln\left(\frac{x_i}{x_{i-1}}\right)+\beta\left(\frac{1}{t_i^\alpha}-\frac{1}{t_{i-1}^\alpha}\right)+\frac{\sigma^2}{2}h\right].\end{aligned}\tag{3.2.19}$$

$$\frac{\partial\ln(\mathbf{L})}{\partial\beta}=\frac{1}{\sigma^2 h}\sum_{i=2}^{n}\left(\frac{1}{t_i^\alpha}-\frac{1}{t_{i-1}^\alpha}\right)\left[\ln\left(\frac{x_i}{x_{i-1}}\right)+\beta\left(\frac{1}{t_i^\alpha}-\frac{1}{t_{i-1}^\alpha}\right)+\frac{\sigma^2}{2}h\right].\tag{3.2.20}$$

$$\frac{\partial \ln(\mathbf{L})}{\partial \alpha} = \frac{\alpha\beta}{\sigma^2 h} \sum_{i=2}^{n} \left(\frac{\ln(t_i)}{t_i^\alpha} - \frac{\ln(t_{i-1})}{t_{i-1}^\alpha} \right) \left[\ln\left(\frac{x_i}{x_{i-1}}\right) + \beta\left(\frac{1}{t_i^\alpha} - \frac{1}{t_{i-1}^\alpha}\right) + \frac{\sigma^2}{2} h \right]. \quad (3.2.21)$$

Making the derivatives (3.2.18), (3.2.19), (3.2.20) and (3.2.21) equal to zero, we obtain the following set of equations

$$\ln(x_1) - \mu_1 = 0; \qquad \sigma_1^2 - (\ln(x_1) - \mu_1)^2 = 0. \quad (3.2.22)$$

$$\frac{\sigma^4 h^2}{4} + \sigma^2 h - \frac{1}{n-1} \sum_{i=2}^{n} \left(\ln\left(\frac{x_i}{x_{i-1}}\right) + \beta\left(\frac{1}{t_i^\alpha} - \frac{1}{t_{i-1}^\alpha}\right) \right)^2 = 0. \quad (3.2.23)$$

$$\sum_{i=2}^{n} \left(\frac{1}{t_i^\alpha} - \frac{1}{t_{i-1}^\alpha}\right) \left[\ln\left(\frac{x_i}{x_{i-1}}\right) + \beta\left(\frac{1}{t_i^\alpha} - \frac{1}{t_{i-1}^\alpha}\right) + \frac{\sigma^2}{2} h \right] = 0. \quad (3.2.24)$$

$$\sum_{i=2}^{n} \left(\frac{\ln(t_i)}{t_i^\alpha} - \frac{\ln(t_{i-1})}{t_{i-1}^\alpha} \right) \left[\ln\left(\frac{x_i}{x_{i-1}}\right) + \beta\left(\frac{1}{t_i^\alpha} - \frac{1}{t_{i-1}^\alpha}\right) + \frac{\sigma^2}{2} h \right] = 0. \quad (3.2.25)$$

On the one hand, from (3.2.22), the ML estimators of μ_1 and σ_1^2 are

$$\hat{\mu}_1 = \ln(x_1) \quad \text{and} \quad \hat{\sigma}_1^2 = (\ln(x_1) - \hat{\mu}_1)^2.$$

On the other hand, from the equations (3.2.23), (3.2.24) and (3.2.25), we can deduce the expression of the estimation of α, β and σ^2. Thus, from the latter equations, we can deduce the expression of the estimation of α, β and σ^2: In our case, from (3.2.23), we can deduce (as a positive solution) the expression of the estimation of σ^2, which is obtained from

$$\sigma_{\alpha\beta}^2 = \frac{2}{h} \left[\left\{ 1 + \frac{1}{n-1} \sum_{i=2}^{n} \left(\ln^2\left(\frac{x_i}{x_{i-1}}\right) + 2\beta B_{i,\alpha} \ln\left(\frac{x_i}{x_{i-1}}\right) + \beta^2 B_{i,\alpha}^2 \right) \right\}^{1/2} - 1 \right]. \quad (3.2.26)$$

Thus, from (3.2.24) and (3.2.26), we can deduce (as a positive solution) the expression of the estimation of β, which is obtained from

$$\beta_\alpha = \frac{(n-1)\left(\sum_{i=2}^{n} B_{i,\alpha}\right)\left(\sum_{i=2}^{n} B_{i,\alpha}^2\right) - D_\alpha \left(\sum_{i=2}^{n} B_{i,\alpha} \ln\left(\frac{x_i}{x_{i-1}}\right)\right) - \Delta^{1/2} \sum_{i=2}^{n} B_{i,\alpha}}{D_\alpha \left(\sum_{i=2}^{n} B_{i,\alpha}^2\right)}, \quad (3.2.27)$$

where $B_{i,\alpha} = \frac{1}{t_i^\alpha} - \frac{1}{t_{i-1}^\alpha}$, $D_\alpha = \left[(n-1)\sum_{i=2}^{n} B_{i,\alpha}^2 - \left(\sum_{i=2}^{n} B_{i,\alpha}\right)^2 \right]$, and

$$\Delta = D_\alpha \left\{ \sum_{i=2}^{n} B_{i,\alpha}^2 \sum_{i=2}^{n} \ln^2\left(\frac{x_i}{x_{i-1}}\right) - \left(\sum_{i=2}^{n} B_{i,\alpha} \ln\left(\frac{x_i}{x_{i-1}}\right) \right)^2 \right\} + \left((n-1) \sum_{i=2}^{n} B_{i,\alpha}^2 \right)^2.$$

However, estimating the parameters α, β and σ^2 remains difficult. In fact, the resulting system of equations is exceedingly complex and does not have an explicit solution, therefore numerical procedures must be employed. As an alternative we can use Newton-Raphson method. This algorithm is designed to solve problems of the type $g(\alpha) = 0$ where g is a well-behaved function. In our case, we consider the following function

$$g(\alpha) = \sum_{i=2}^{n} \left(\frac{\ln(t_i)}{t_i^\alpha} - \frac{\ln(t_{i-1})}{t_{i-1}^\alpha} \right) \left[\ln\left(\frac{x_i}{x_{i-1}}\right) + \beta\left(\frac{1}{t_i^\alpha} - \frac{1}{t_{i-1}^\alpha}\right) + \frac{\sigma^2}{2} h \right]. \quad (3.2.28)$$

Nevertheless, these kind of methods may not be suitable for random data. The reason is that the function g depends on observations, so the conditions to guarantee the convergence of the Newton-Raphson method may not be met, for example, when there exists a change of sign in the second derivative of g (see, for example, Román-Román et al. [55]). In this sense, we propose the use of the metaheuristic SA algorithm in order to maximise the likelihood function. This algorithm is designed to solve problems of the type $\min_{\theta\in\Theta} h(\theta)$, where h is the objective function to be optimized. This method is often more appropriate than classical numerical methods since it imposes fewer restrictions on the space of solutions Θ and on the analytical properties of h. In our case, the problem becomes maximizing the function $\ln\left(L_{x_1,x_2,\cdots,x_n}\left(\alpha,\beta,\sigma^2\right)\right)$. Since the algorithm aforementioned is usually formulated for minimization problems, then, from (3.2.17), the objective function is a function of the parameters α, β and σ^2 and has the following form:

$$h\left(\alpha,\beta,\sigma^2\right)=\frac{(n-1)}{2}\ln\left(\sigma^2\right)+\frac{1}{2\sigma^2 h}\sum_{i=2}^{n}\left\{\ln\left(\frac{x_i}{x_{i-1}}\right)+\beta\left(\frac{1}{t_i^\alpha}-\frac{1}{t_{i-1}^\alpha}\right)+\frac{\sigma^2}{2}h\right\}^2. \quad (3.2.29)$$

3.2.4.2 Estimated MFs and CBs

3.2.4.2.1 Estimated MFs

By using Zehna's theorem [68], the estimated MF (EMF) of the process is obtained by replacing the parameters by replacing the parameters in equations Eq. (3.2.14) and Eq. (3.2.15) by their estimators given in equations Eq. (3.2.26), Eq. (3.2.27) and Eq. (3.2.28), then the estimated CMF (ECMF) is given by the following expression

$$\hat{E}\left(x(t)|x(s)=x_s\right)=x_s\exp\left(-\hat{\beta}\left(\frac{1}{t^{\hat{\alpha}}}-\frac{1}{s^{\hat{\alpha}}}\right)\right), \quad (3.2.30)$$

Taking into account the initial condition that is $\mathrm{P}\left(x\left(t_1\right)=x_1\right)=1$, the EMF of the process is given by

$$\hat{E}(x(t))=x_1\exp\left(\frac{\hat{\beta}}{t_1^{\hat{\alpha}}}\right)\exp\left(\frac{-\hat{\beta}}{t^{\hat{\alpha}}}\right). \quad (3.2.31)$$

3.2.4.2.2 Confidence bounds

The confidence bounds (CBs) of the process are obtained by using the procedure described in [67]. Let $v(s,t)=x(t)|x(s)=x_s$. Since the variable $w(t)-w(s)$ is Gaussian with the mean equal to zero and the variance $t-s$ for $t\geq s$. Therefore, the random variable z is given by

$$z=\frac{\ln(v(s,t))-\mu\left(s,t,x_s\right)}{\sigma\sqrt{(t-s)}}\sim N_1(0,1),$$

where $\mu\left(s,t,x_s\right)=\ln\left(x_s\right)-\beta\left(\frac{1}{t^\alpha}-\frac{1}{s^\alpha}\right)-\frac{\sigma^2}{2}(t-s)$.

A $100(1-\kappa)\%$ conditional CBs for z is given by $P(-\xi \leq z \leq \xi) = 1-\kappa$. From this, we can obtain a CB of $v(s,t)$ with following form $(v_{lower}(s,t), v_{upper}(s,t))$, where

$$v_{lower}(s,t) = \exp\left(\mu\left(s,t,x_s\right) - \xi\sigma\sqrt{(t-s)}\right), \tag{3.2.32}$$

and

$$v_{upper}(s,t) = \exp\left(\mu\left(s,t,x_s\right) + \xi\sigma\sqrt{(t-s)}\right), \tag{3.2.33}$$

with $\xi = F^{-1}_{N(0,1)}\left(1-\dfrac{\kappa}{2}\right)$ and where $F^{-1}_{N(0,1)}$ is the inverse cumulative normal standard distribution and $\mu\left(s,t,x_s\right) = \ln\left(x_s\right) - \beta\left(\dfrac{1}{t^{\alpha}} - \dfrac{1}{s^{\alpha}}\right) - \dfrac{\sigma^2}{2}(t-s)$.

On the other hand, the estimated lower bound $\hat{v}_{lower}(t)$ and an estimated upper bound $\hat{v}_{upper}(t)$ can be obtained by substituting the parameters by theirs estimators in the equations Eq. (3.2.32) and Eq. (3.2.33), the estimated CBs (ECBs) are given by:

$$\hat{v}_{lower}(s,t) = \exp\left(\hat{\mu}\left(s,t,x_s\right) - \xi\hat{\sigma}\sqrt{(t-s)}\right), \tag{3.2.34}$$

and

$$\hat{v}_{upper}(s,t) = \exp\left(\hat{\mu}\left(s,t,x_s\right) + \xi\hat{\sigma}\sqrt{(t-s)}\right), \tag{3.2.35}$$

where $\hat{\mu}\left(s,t,x_s\right) = \ln\left(x_s\right) - \hat{\beta}\left(\dfrac{1}{t^{\hat{\alpha}}} - \dfrac{1}{s^{\hat{\alpha}}}\right) - \dfrac{\hat{\sigma}^2}{2}(t-s)$.

3.2.5 Application of the SA algorithm

3.2.5.1 The algorithm

The SA algorithm is a metaheuristic algorithm for problems of the type $\min\limits_{\theta\in\Theta} h\left(\theta\right)$, where h is the objective function to be optimized and Θ is the solution space.

In our case, the parametric space Θ linked to the objective function (3.2.29), on which the selected algorithms must operate, is continuous and unbounded. Consequently

$$\Theta = \{(\alpha,\beta,\sigma): \quad \alpha > 0,\ \beta > 0,\ \sigma > 0\}.$$

The drawback is that the solution space might not be explored with enough depth. This requires us to find arguments for bounding said space. The following are some strategies to this end.

3.2.5.2 Bounding the solution space

Regarding the parameter σ, when it has high values it leads to sample paths with great variability around the mean of the process. Thus, excessive variability in available paths would make an Lundqvist-Korf-type modeling inadvisable (see, Figure 3.2). Some simulations performed for several values of σ have led us to consider that $0 < \sigma < 0.1$, so that we may have paths compatible with an Lundqvist-Korf-type growth. On the other hand, there does not seem to be an upper bound for α and β. To this end, we will use the following reformulation: Setting $b = \exp(-\beta)$ and $a = \dfrac{1}{\alpha}$. This reformulation leads to the condition $0 < b < 1$ and does not seem to be an upper bound for a. Nevertheless, the Lundqvist-Korf curve is sigmoidal and has an inflection point at $t_I = \left(\dfrac{\ln\left(\frac{1}{b}\right)}{a+1}\right)^a$, which is higher than t_1 if and only if $b < \exp\left(-t_1^{\frac{1}{a}}\right)$. The parameter a can be bounded taking into account the information provided by the sample paths and the asymptote of the curve verifying $k = x_1\left(\dfrac{1}{b}\right)^{\left(\frac{1}{t_1}\right)^{\frac{1}{a}}}$ and the curve has an inflection point at $t_I = \left(\dfrac{\ln\left(\frac{1}{b}\right)}{a+1}\right)^a$. From these expressions, and if we denote by $k_{\max}$ the maximum value of the sample path, the following expression holds $0 < a < \ln\left(\dfrac{k_{\max}}{x_1}\right) - 1$, where x_1 is the initial value of the sample path. Thus, the solution space, which is obtained numerically for a, b and σ is bounded and takes the form

$$(0, \ln(\eta) - 1) \times \left(0, \exp\left(-t_1^{\frac{1}{\ln(\eta)-1}}\right)\right) \times (0, 0.1)\,,$$

where $\eta = \dfrac{k_{\max}}{x_1}$.

Once the solution space has been bounded, we specify the choice of the initial parameters of the algorithms and the stopping conditions. Let consider:

1. The initial solution is chosen randomly in the bounded subspace

$$\Theta' = (0, \ln(\eta) - 1) \times \left(0, \exp\left(-t_1^{\frac{1}{\ln(\eta)-1}}\right)\right) \times (0, 0.1)\,.$$

2. The initial temperature should be high enough such that in the first iteration of the algorithm, the probability of accepting a worse solution is at least of 80% (see, Kirkpatrick et al. [58]). For this, we assume the initial temperature of 10.

3. For the cooling process we have considered a geometric scheme in which the current temperature is multiplied by a constant γ ($0 < \gamma < 1$), i.e. $T_i = \gamma T_{i-1}$, $i \geq 1$. The usual values for γ are between 0.80 and 0.99. For this we have set $\gamma = 0.95$.

4. The length of each temperature level L determines the number of solutions generated at a certain temperature T. For this we have set $L = 50$.

5. The stopping criterion defines when the system has reached a desired energy level (freezing temperature). Equivalently it defines the total number of solutions generated, or when an acceptance ratio (ratio between the number of solutions accepted and the number of solutions generated) is reached. The application of the algorithm will be limited to 1000 iterations.

The coding is performed using Matlab computer software.

3.2.6 Simulation study

3.2.6.1 Simulated sample paths of the SLKDP

In this section, we present some simulated sample paths of the SLKDP. The trajectory of the model is obtained by simulating the exact solution of the SDE Eq. (3.2.7). We obtain the simulated trajectories of the model by considering the time discretization of the interval $[t_1, T]$, with time points $t_i = t_{i-1} + (i-1)h$; for $i = 2, \cdots, N$ and $h = \dfrac{T - t_1}{N}$ is the descritization step size for an integer N (N is the sample size).

The random variable $\sigma\left(w(t) - w\left(t_1\right)\right)$ in the equation (3.2.7) is distributed as one-dimensional normal distribution $N_1\left(0, \sigma^2(t - t_1)\right)$. Therefore, in this simulation, 50 sample paths are simulated with $t_1 = 1$, $T = 40$, and $N = 100$.

Figure 3.2 shows some simulated sample paths of the SLKDP for several values of α, β and σ.

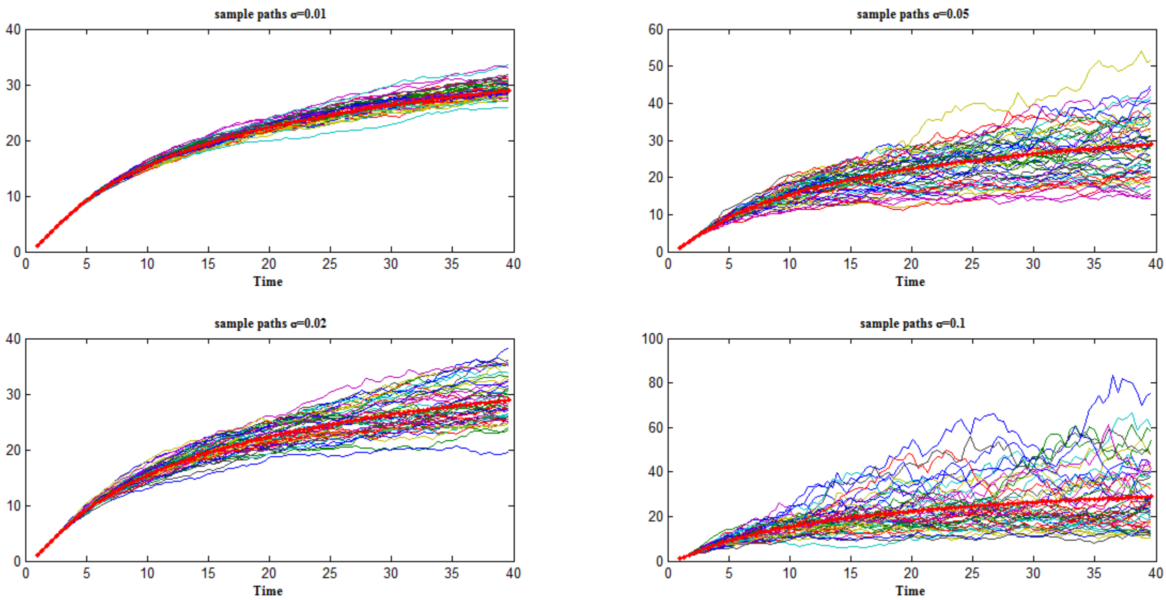


Figure 3.2: Simulated trajectories of the SLKDP and its MF for several values of σ with $(x_1 = 1,\ \alpha = 0.5,\ \beta = 4)$.

3.2.6.2 Parameters estimation

In this section, we present several examples in order to validate the estimation procedure previously developed. To this end, equation (3.2.7) was simulated 30 times under the following assumptions $t_1 = 0.1$, taking the step size $h = 0.2$ and $x_1 \sim \Lambda_1(1, 0.16)$ and $N = 50$, 100, 250 and 500 respectively. In order to make the subsequent inference we have considered, in each case, 30 sample paths with $t_i = t_{i-1} + (i-1)h$; for $i = 2, \cdots, N$.

The empirical mean, the std, and the CV for a, b and σ (functions of parameters α, β and σ which are the arguments of the objective function) are defined in Table 3.1. Then, Table

3.2 shows the results obtained for calculating the latter measures. The results obtained show the performance of the methodology.

Table 3.1: The empirical mean, the std and the CV for a, b and σ (M is the sample paths)

Empirical mean	Std	CV
$\overline{a} = \frac{1}{M}\sum_{i=1}^{M} a_i$	$\text{std}(a) = \left(\frac{1}{M-1}\sum_{i=1}^{M}(a_i - \overline{a})^2\right)^{\frac{1}{2}}$	$\text{CV}(a) = \frac{\text{std}(a)}{\overline{a}}$
$\overline{b} = \frac{1}{M}\sum_{i=1}^{M} b_i$	$\text{std}(b) = \left(\frac{1}{M-1}\sum_{i=1}^{M}\left(b_i - \overline{b}\right)^2\right)^{\frac{1}{2}}$	$\text{CV}(b) = \frac{\text{std}(b)}{\overline{b}}$
$\overline{\sigma} = \frac{1}{M}\sum_{i=1}^{M} \sigma_i$	$\text{std}\,(\sigma) = \left(\frac{1}{M-1}\sum_{i=1}^{M}(\sigma_i - \overline{\sigma})^2\right)^{\frac{1}{2}}$	$\text{CV}\,(\sigma) = \frac{\text{std}(\sigma)}{\overline{\sigma}}$

Table 3.2: Estimation values, the std and the CV of a, b and σ for several values of σ.

$a = 2,\ b = 0.02$									
$\sigma = 0.01$									
N	$\overline{a}$	$\overline{b}$	$\overline{\sigma}$	$\text{std}(a)$ $.10^{-4}$	$\text{std}(b)$ $.10^{-4}$	$\text{std}(\sigma)$ $.10^{-4}$	$\text{CV}(a)$ $.10^{-4}$	$\text{CV}(b)$ $.10^{-4}$	$\text{CV}(\sigma)$ $.10^{-4}$
50	1.99657	0.02027	0.01074	127.9	9.041	11.43	64.07	446.1	106.4
100	1.98886	0.02082	0.00987	72.79	5.194	5.757	36.60	249.4	583.2
250	1.99025	0.02070	0.01007	60.13	4.500	5.283	30.21	217.4	524.3
500	1.98826	0.02083	0.01001	28.36	1.964	2.479	14.27	94.29	247.7
$\sigma = 0.02$									
50	2.00358	0.01976	0.01936	62.96	4.373	10.96	31.43	221.3	566.3
100	2.00558	0.01961	0.01991	38.12	2.861	14.05	19.01	145.9	706.1
250	2.00529	0.01962	0.01994	24.94	2.413	9.925	12.44	122.9	497.8
500	2.00402	0.01969	0.01986	38.97	2.904	7.244	19.44	147.5	364.7
$\sigma = 0.05$									
50	1.99025	0.02091	0.04994	133.9	19.17	49.63	116.6	916.5	993.7
100	1.98917	0.02089	0.04998	164.7	11.74	34.65	82.75	562.2	693.2
250	1.99189	0.0205	0.04996	220.3	15.86	22.37	110.6	772.2	447.7
500	1.98858	0.02088	0.04992	162.1	12.14	17.59	81.47	581.1	352.4

3.2.6.3 Prediction

In this section, we have considered an example of the predictive study in which $N = 25$ and $t_i = t_{i-1} + (i-1)h$; for $i = 2, \cdots, N$ starting at $t_1 = 0.1$, taking the step size $h = 0.2$ and $x_1 = 2$. The methodology can be summarised in the following steps: First, we use the first 22 data to estimate the parameters a, b and σ^2 of the process by SA. Moreover, we obtain the corresponding EMF and ECMF values given by the expressions (3.2.14) and (3.2.15). For the three last data, we predict the corresponding values using the EMF and ECMF. Also, we give the results attached to a 95% ECB and a 95% estimated conditional CB (ECCB) of the process (see, the expressions (3.2.34) and (3.2.35)). Finally, to illustrate the performance of procedure, the results according to the MAE, the RMSE and MAPE, given by Table 2.1.

From Lewis [69], we deduce that the accuracy of the forecast can be judged from the MAPE result Table 2.2.

Table 3.3 shows the results of the simulated values and the estimated MF, CMF, CB and CCB of the process.

Table 3.3: Simulated and predicted data, showing EMF, ECMF, ECB and ECCB using the model.

i	Simulated data $x(t_i)$	EMF	ECMF	ECB $(.10^4)$	ECCB $(.10^4)$
1	2.0000000	2.0000000	2.0000000	[0.00020 0.00020]	[0.00020 0.00020]
2	422.18230	422.52508	422.52508	[0.04164 0.04287]	[0.04164 0.04287]
3	2198.0757	2203.0824	2201.2951	[0.21581 0.22486]	[0.21695 0.22334]
4	5278.2411	5309.1673	5297.1016	[0.51769 0.54439]	[0.52206 0.53744]
5	9429.1186	9362.3703	9307.8339	[0.90934 0.96371]	[0.91734 0.94437]
6	14199.186	14032.305	14132.347	[1.35822 1.44933]	[1.39282 1.43386]
7	19268.748	19081.794	19308.726	[1.84120 1.97693]	[1.90298 1.95905]
8	24703.936	19081.793	24586.690	[2.34260 2.52967]	[2.42316 2.49455]
9	30291.324	29720.825	30155.132	[2.85187 3.09599]	[2.97196 3.05953]
10	36300.858	35124.807	35799.036	[3.36194 3.66794]	[3.52820 3.63215]
11	41306.230	40509.112	41865.440	[3.86806 4.24007]	[4.12608 4.24765]
12	47004.079	45839.212	46741.213	[4.36709 4.80860]	[4.60662 4.74234]
13	52442.472	51091.924	52390.273	[4.85698 5.37096]	[5.16336 5.31550]
14	57230.651	56251.966	57738.913	[5.33639 5.92539]	[5.69051 5.85817]
15	61415.223	61309.602	62376.281	[5.80455 6.47075]	[6.14755 6.32868]
16	65804.131	66259.014	66373.162	[6.26102 7.00628]	[6.54146 6.73420]
17	70506.561	71097.165	70609.066	[6.70563 7.53154]	[6.95894 7.16397]
18	75426.385	75822.989	75193128	[7.13838 8.04630]	[7.41072 7.62907]
19	80457.836	80436.821	80016.084	[7.55942 8.55051]	[7.88605 8.11840]
20	84679.801	84939.980	84962.172	[7.96894 9.04421]	[8.37352 8.62023]
21	88835.694	89334.471	89060.832	[8.36723 9.52752]	[8.77746 9.03608]
22	93622.933	93622.766	93100.046	[8.75458 10.0006]	[9.17555 9.44590]
Prediction					
23	98496.745	97807.643	97807.817	[9.13135 10.4637]	[9.63953 9.92355]
24	103678.20	101892.07	102609.95	[9.49786 10.9170]	[10.1128 10.4107]
25	105881.59	105879.12	107735.14	[9.85448 11.3609]	[10.6179 10.9307]

Table 3.4 shows the values obtained from the estimation of the parameters of the process. Then the accuracy of the forecast can be judged from the MAPE, in other words if the value of the MAPE is less than 10%, the forecast is highly accurate.

Table 3.5 shows the values obtained after the calculation of the three measures of the goodness of fit.

Table 3.4: Estimation of the parameters of the SLKDP.

Parameters	Estimated value
a	2.008971
b	0.017608
σ^2	0.000274

Table 3.5: Goodness of fit of the SLKDP.

MAE	RMSE	MAPE
466.239	674.796	0.982%

Figure 3.3 and Figure 3.4 illustrate the performance of the SLKDP for forecasting using the mean function and the conditional mean function.

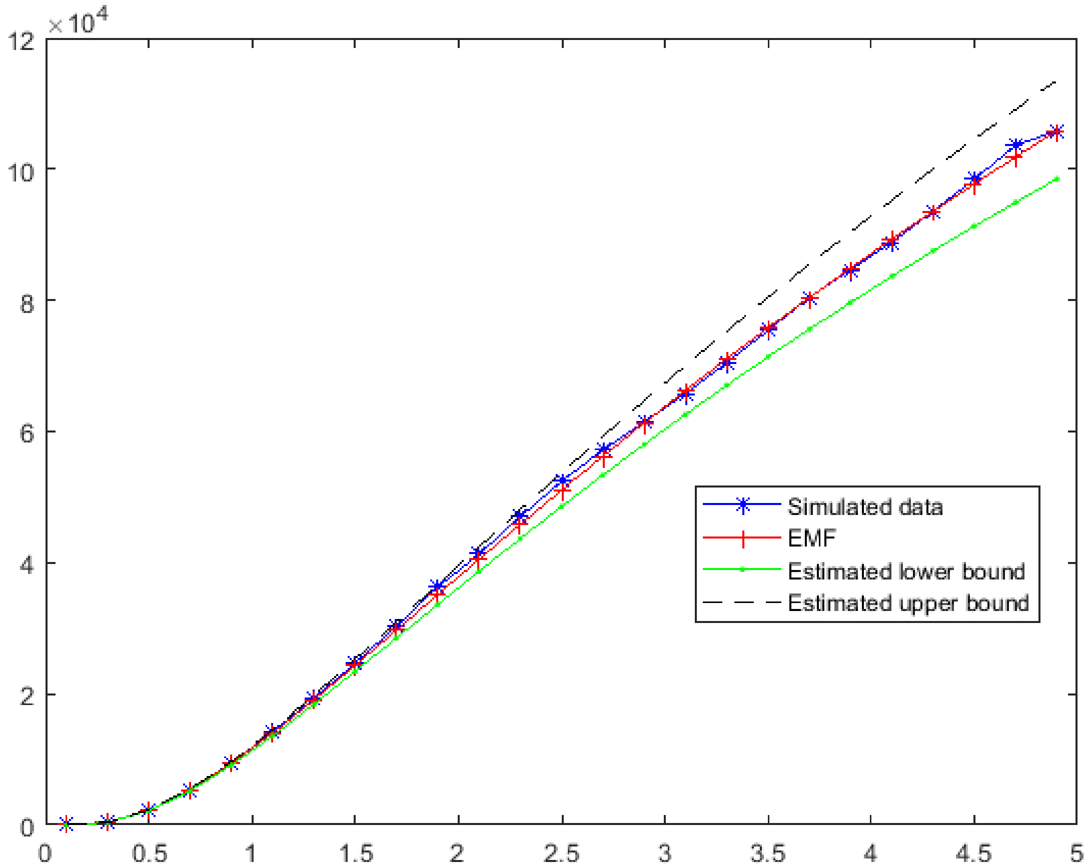


Figure 3.3: Simulated data, EMF, ECB and the predicted values.

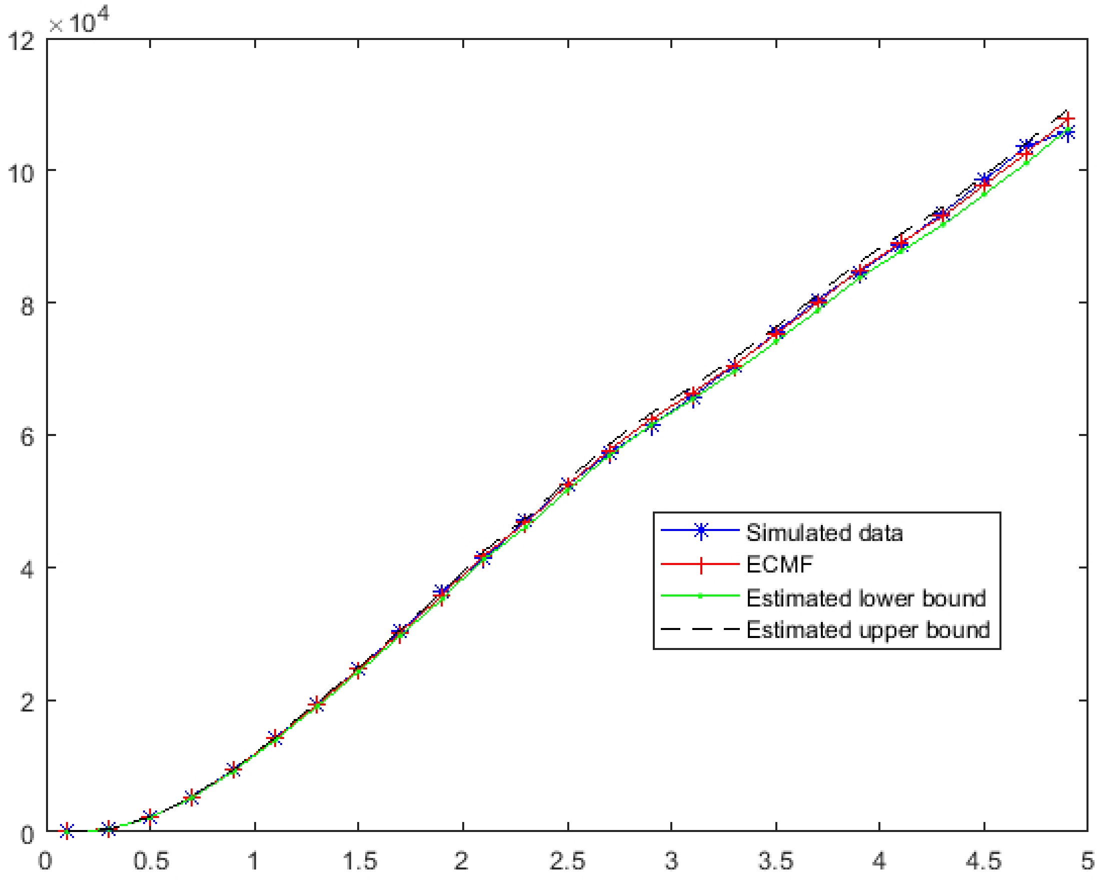


Figure 3.4: Simulated data, ECMF, ECCB and the conditional predicted values.

3.3 The γ-SLKDP

3.3.1 The proposed model

Let $\{x(t);\, t \in [t_1; T];\, t_1 > 0\}$ be a stochastic Lundqvist-Korf diffusion process. Then, the γ-power of the stochastic Lundqvist-Korf diffusion process $x(t)$ (γ-SLKDP) is defined by

$$x_\gamma(t) = x^\gamma(t); \qquad \gamma \in \mathbb{R}^*. \tag{3.3.1}$$

The process $\{x_\gamma(t) : t \in [t_1; T];\; t_1 > 0\}$ is also a diffusion process with values in $(0; \infty)$, and has the drift and diffusion coefficients are shown below.

By applying Itô's formula to the transform given in Eq. (3.3.1), we have

$$dx_\gamma(t) = \left(\frac{\gamma\alpha\beta}{t^{\alpha+1}} + \frac{\gamma(\gamma-1)\sigma^2}{2}\right) x_\gamma(t)\, dt + \gamma\sigma x_\gamma(t)\, dw(t). \tag{3.3.2}$$

The last SDE Eq. (3.3.2) has a unique solution (see, Kloeden and Platen [8]) can be obtained using the relation $x_\gamma(t) = x^\gamma(t)$ and Eq. (3.3.2), from which we have

$$x_\gamma(t) = x_\gamma(t_1) \exp\left[-\gamma\beta\left(\frac{1}{t^\alpha} - \frac{1}{t_1^\alpha}\right) - \frac{\gamma\sigma^2}{2}(t - t_1) + \gamma\sigma(w(t) - w(t_1))\right]. \tag{3.3.3}$$

From this, we deduce that the process $(x_\gamma(t) \mid x_\gamma(s) = x_s)$ is distributed as the following one-dimensional lognormal distribution $\Lambda_1\left[m(s,t,x_s); \gamma^2\sigma^2(t-s)\right]$ with
$m(s,t,x_s) = \ln(x_s) - \gamma\beta\left(\frac{1}{t^\alpha} - \frac{1}{s^\alpha}\right) - \frac{\gamma\sigma^2}{2}(t-s)$.

3.3.2 The PTDF of the γ-SLKDP

When $(x_\gamma(t) \mid x_\gamma(s) = x_s)$ has the log-normal distribution $\Lambda_1\left[m(s,t,x_s); \gamma^2\sigma^2(t-s)\right]$, then the TPDF of the γ-SLKDP is

$$f(x,t \mid y,s) = \frac{1}{x\sqrt{2\pi\gamma^2\sigma^2(t-s)}} \exp\left(-\frac{[\ln(x) - m(s,t,y)]^2}{2\gamma^2\sigma^2(t-s)}\right). \tag{3.3.4}$$

3.3.3 The MFs of the γ-SLKDP

From the properties of the log-normal distribution, the $r-$th conditional moment of the process is

$$\begin{aligned} E\left(x_\gamma^r(t) \mid x_\gamma(s) = x_s\right) &= \exp\left(rm(s,t,x_s) + \frac{r^2\gamma^2\sigma^2}{2}(t-s)\right) \\ &= x_s^r \exp\left(-r\gamma\beta\left(\frac{1}{t^\alpha} - \frac{1}{s^\alpha}\right) + \frac{r\gamma(\gamma r - 1)\sigma^2}{2}(t-s)\right). \end{aligned}$$

Then, for $r = 1$, the conditional mean function (CMF) of the process is

$$E(x_\gamma(t) \mid x_\gamma(s) = x_s) = x_s \exp\left(-\gamma\beta\left(\frac{1}{t^\alpha} - \frac{1}{s^\alpha}\right) + \frac{\gamma(\gamma-1)\sigma^2}{2}(t-s)\right).$$

In addition, taking into the initial condition $P(x_\gamma(t_1) = x_1) = 1$, the mean function (MF) of the process is given by

$$E(x_\gamma(t)) = x_1 \exp\left(-\gamma\beta\left(\frac{1}{t^\alpha} - \frac{1}{t_1^\alpha}\right) + \frac{\gamma(\gamma-1)\sigma^2}{2}(t-t_1)\right).$$

Remark. If $\gamma = 1$, note that $E(x_\gamma(t) \mid x_\gamma(s) = x_s) = x_s \exp\left(-\beta\left(\frac{1}{t^\alpha} - \frac{1}{s^\alpha}\right)\right)$ is the conditional mean function of the stochastic Lundqvist-Korf diffusion process.

3.3.4 Statistical inference on the model

3.3.4.1 Parameters estimation

Let us then examine in this section the ML estimation of the parameters α, β and σ^2 of the model from which we can obtain, by virtue of Zehna's theorem [52], we obtain the corresponding estimated trend functions of the process.

We consider a discrete sampling of the process, that is, for fixed times $t_1, t_2, \cdots, t_n$, where $n > 2$, with the initial condition $P\left[x_\gamma\left(t_1\right) = x_1\right] = 1$, we observe the variables $x_\gamma\left(t_1\right)$, $x_\gamma\left(t_2\right)$, $\cdots$, $x_\gamma\left(t_n\right)$ whose values provide the basic sample for the inference process. In addition, we assume $t_i - t_{i-1} = h$, for $i = 2, \cdots, n$. Let x_1, x_2, $\cdots$, x_n be the observed values of the sampling. The associated likelihood function is thus

$$\mathbf{L} = L_{x_1, x_2, \cdots, x_n}\left(\alpha, \beta, \sigma^2\right) = \prod_{i=2}^{n} f\left(x_i, t_i \mid x_{i-1}, t_{i-1}\right),$$

which is written as

$$\mathbf{L} = \prod_{i=2}^{n} \frac{1}{x_i\sqrt{2\pi\gamma^2\sigma^2 h}} \exp\left(-\frac{\left[\ln\left(\frac{x_i}{x_{i-1}}\right) + \gamma\beta\left(\frac{1}{t_i^\alpha} - \frac{1}{t_{i-1}^\alpha}\right) + \frac{\gamma\sigma^2}{2}h\right]^2}{2\gamma^2\sigma^2 h}\right). \tag{3.3.5}$$

From Eq. (3.3.5), the log-likelihood function of the sample is

$$\begin{aligned} \ln(\mathbf{L}) = &- \frac{(n-1)}{2}\ln\left(2\pi\gamma^2 h\right) - \frac{(n-1)}{2}\ln\left(\sigma^2\right) - \sum_{i=2}^{n}\ln(x_i) \\ &- \frac{1}{2\gamma^2\sigma^2 h}\sum_{i=2}^{n}\left\{\ln\left(\frac{x_i}{x_{i-1}}\right) + \gamma\beta\left(\frac{1}{t_i^\alpha} - \frac{1}{t_{i-1}^\alpha}\right) + \frac{\gamma\sigma^2}{2}h\right\}^2. \end{aligned} \tag{3.3.6}$$

However, estimating α, β and σ^2 poses some difficulties. In fact, the resulting system of equations is exceedingly complex and does not have an explicit solution, and numerical procedures must be employed.

3.3.4.2 Computational aspects

In this section we propose the use of the SA algorithm to maximize the likelihood function or, equivalently, its logarithm. Hereafter is the description of the method.

SA algorithm is a metaheuristic algorithm to approximating the solution of optimization problems of the type $\min_{x\in\Theta} g(x)$, where g is the objective function to be optimized and Θ is the solution space.

In our case, the problem becomes maximizing the log-likelihood function $\ln(\mathbf{L})$. Since the algorithm aforementioned is usually formulated for minimization problems,then, from

Eq. (3.3.6), the objective function is a function of the parameters α, β and σ^2 and has the following form:

$$g\left(\alpha,\beta,\sigma^2\right)=\frac{(n-1)}{2}\ln\left(\sigma^2\right)+\frac{1}{2\gamma^2\sigma^2h}\sum_{i=2}^{n}\left\{\ln\left(\frac{x_i}{x_{i-1}}\right)+\gamma\beta\left(\frac{1}{t_i^{\alpha}}-\frac{1}{t_{i-1}^{\alpha}}\right)+\frac{\gamma\sigma^2}{2}h\right\}^2. \tag{3.3.7}$$

The parametric space Θ linked to the objective function (3.3.7), on which the selected algorithms must operate, is continuous and unbounded. Consequently

$$\Theta = \{(\alpha,\beta,\sigma):\alpha>0,\ \beta>0,\ \sigma>0\}.$$

The drawback is that the solution space might not be explored with enough depth. This requires us to find arguments for bounding said space. For this we use the strategies developed by Nafidi et al. [28]. Thus, for a known γ, the solution space, which is obtained numerically for $a=\frac{1}{\alpha}$, $b=\exp(-\beta)$ and σ is bounded and takes the form

$$\left(0,\ \frac{\lambda}{\gamma}\right)\times\left(0,\ \exp\left(-t_1^{\frac{\gamma}{\lambda}}\right)\right)\times(0,\ 0.1)\,,$$

where $\lambda=\ln\left(\frac{k_{\max}}{x_1^{\gamma}}\right)-1$ and k_{max} is the maximum value of the sample path.

3.3.5 Simulation study

3.3.5.1 Simulated sample paths of the γ-SLKDP

This section presents some simulated sample paths of the γ-SLKDP. We obtain the trajectory of the model by simulating the exact solution of the SDE Eq. (3.3.3). We obtain the simulated trajectories of the model by considering the time discretization of the interval $[t_1,T]$, with time points $t_i=t_{i-1}+(i-1)h$, $i=2,\cdots,N$ and step size $h=\frac{T-t_1}{N}$ for the sample size N.

The random variable $\gamma\sigma\left(w(t)-w\left(t_1\right)\right)$ in the Eq. (3.3.3) is distributed as one-dimensional normal distribution $\mathrm{N}_1\left(0,\gamma^2\sigma^2(t-t_1)\right)$. This simulation for $x_\gamma(t)$ ($\gamma=1$, $\gamma=1.5$ and $\gamma=2$) is based on 20 sample paths with $t_1=1$, $T=40$, $x_1=2$, and $N=50$.

Figures 3.5, 3.6 and 3.7 shows some simulated sample paths of the γ-SLKDP for $(\gamma=1,\gamma=1.5,\gamma=2)$ and its mean function for several values of σ with $\alpha=0.5$ and $\beta=4$.

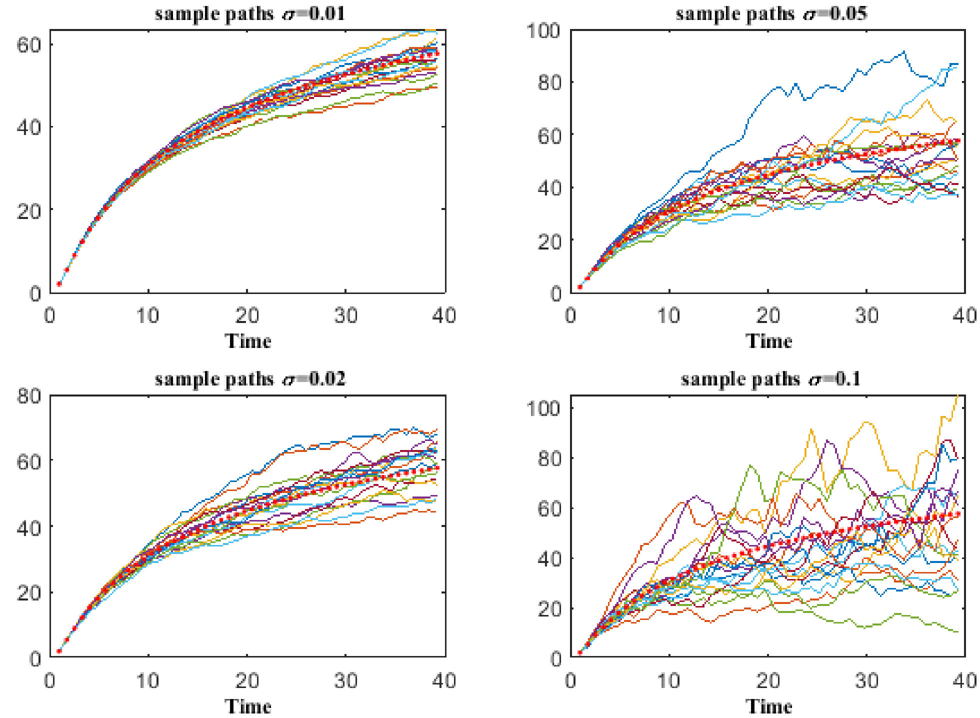


Figure 3.5: Simulated trajectories of the γ-SLKDP and its MF for several values of σ with $\alpha = 0.5$, $\beta = 4$ and $\gamma = 1$.

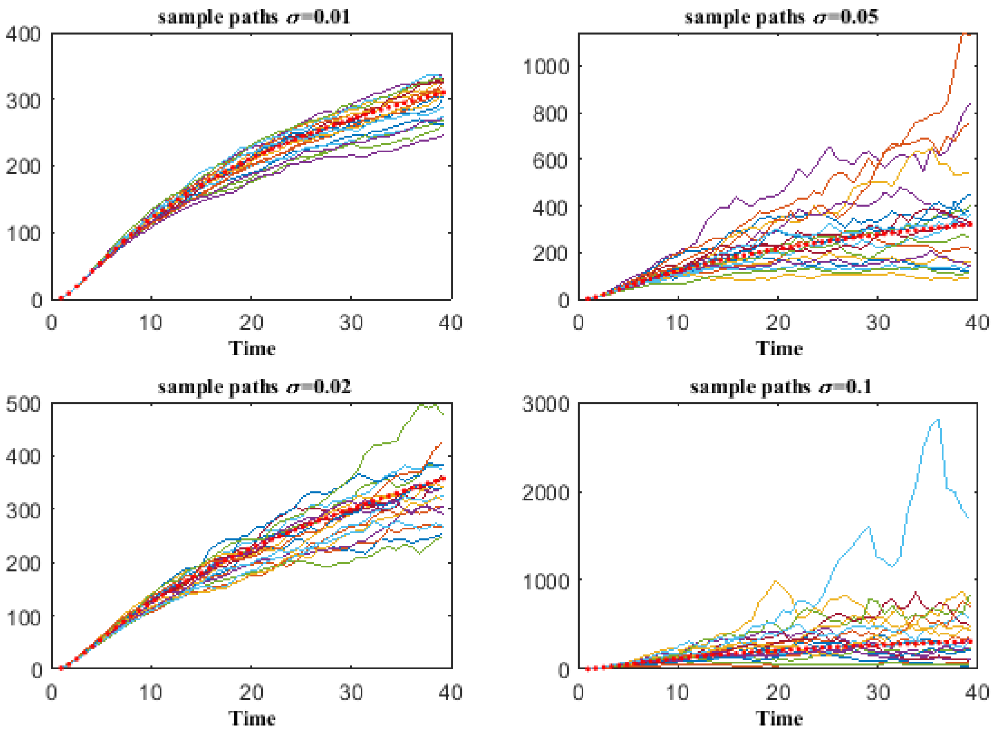


Figure 3.6: Simulated trajectories of the γ-SLKDP and its MF for several values of σ with $\alpha = 0.5$, $\beta = 4$ and $\gamma = 1.5$.

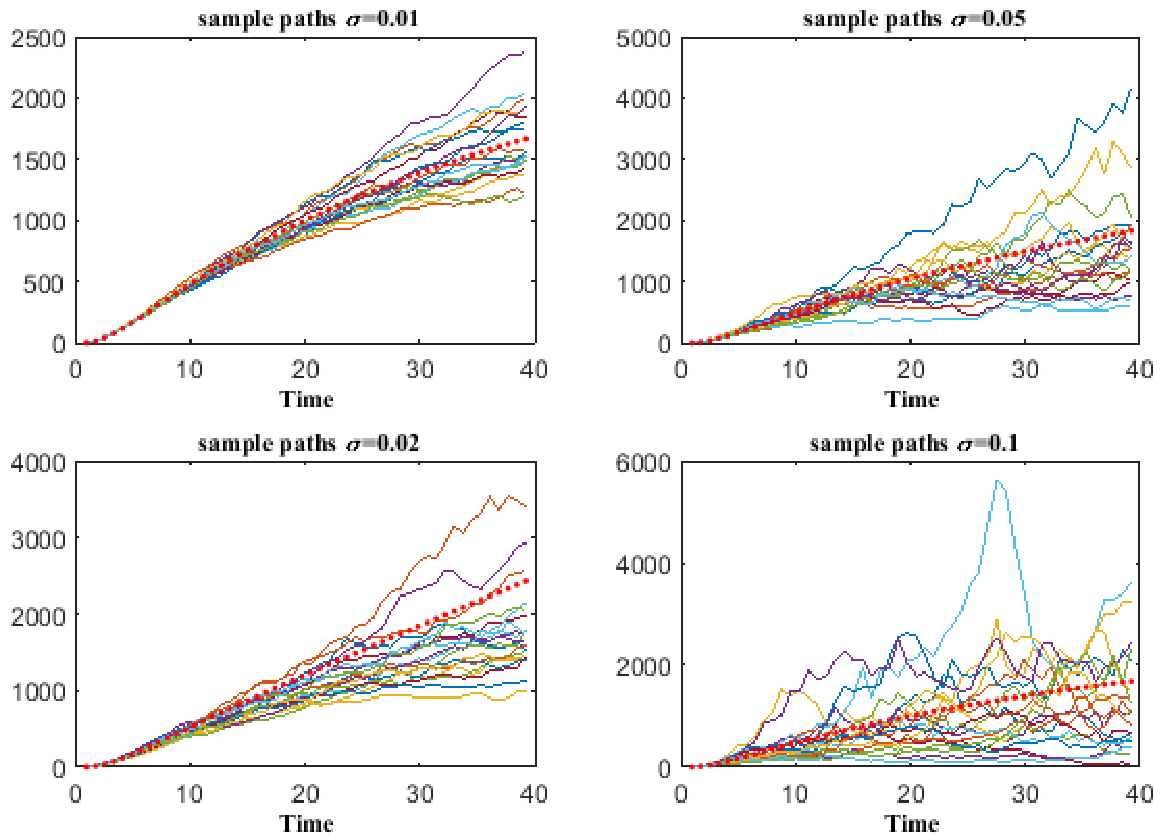


Figure 3.7: Simulated trajectories of the γ-SLKDP and its MF for several values of σ with $\alpha = 0.5$, $\beta = 4$ and $\gamma = 2$.

3.3.5.2 Parameter estimation

This section will present several examples in order to validate the estimation procedure previously developed. To this end, Eq. (3.3.3) was simulated 25 times under the following assumptions, for each one, we consider the equidistant time discretization of the interval $[t_1, T]$ with time points $t_i = t_{i-1} + (i-1)h$; for $i = 2, \cdots, N$ where $T = 50$ and step size $h = \dfrac{T - t_1}{N}$, starting from instant $t_1 = 0.1$ and $x_1 = 2$. As for the sample size, values 100, 250 and 500 have been considered for N. The SA algorithm has been applied for estimating the parameters of the process.

The results obtained are summarised in Table 3.6 which shows for each of the above data sets the empirical mean, the std, and the CV obtained for $a = \dfrac{1}{\alpha}$, $b = \exp(-\beta)$ and σ given by Table 3.1.

The results obtained show the performance as well as the usefulness and importance of the methodology studied.

Table 3.6: Estimation values, the std and the CV of a, b and σ for several values of γ.

$a = 2, b = 0.02, \sigma = 0.01$									
N	$\overline{a}$	$\overline{b}$	$\overline{\sigma}$	std(a) $.10^{-4}$	std(b) $.10^{-4}$	std(σ) $.10^{-4}$	CV(a) $.10^{-4}$	CV(b) $.10^{-4}$	CV(σ) $.10^{-4}$
$\gamma = 1$									
100	1.98886	0.02082	0.00987	72.79	5.194	5.757	36.60	249.4	583.2
250	1.99025	0.02070	0.01007	60.13	4.500	5.283	30.21	217.4	524.3
500	1.98826	0.02083	0.01001	28.36	1.964	2.479	14.27	94.29	247.7
$\gamma = 1.5$									
100	2.01439	0.01964	0.01187	733.8	51.04	32.54	364.3	2599	2742
250	2.04887	0.01713	0.01141	592.7	34.11	24.50	289.3	1992	2147
500	2.00738	0.02031	0.01129	812.0	64.45	14.16	404.5	3173	1254
$\gamma = 2$									
100	1.97972	0.02169	0.01108	503.4	37.79	17.57	254.3	1742	1586
250	1.99655	0.02082	0.01033	372.1	32.47	8.482	186.4	1559	821.0
500	1.99498	0.02059	0.01027	411.2	34.31	8.762	206.1	1665	853.0
$a = 0.02, b = 0.7, \sigma = 0.02$									
$\gamma = 1$									
100	2.00558	0.01961	0.01991	38.12	2.861	14.05	19.01	145.9	706.1
250	2.00529	0.01962	0.01994	24.94	2.413	9.925	12.44	122.9	497.8
500	2.00402	0.01969	0.01986	38.97	2.904	7.244	19.44	147.5	364.7
$\gamma = 1.5$									
100	1.99800	0.02039	0.02005	466.1	38.64	16.65	233.3	1894	830.3
250	2.00177	0.01996	0.01978	308.2	22.88	10.02	153.9	1147	506.2
500	1.98623	0.02128	0.02001	429.1	35.48	8.020	216.1	1667	400.6
$\gamma = 2$									
100	1.96747	0.02273	0.02071	619.0	50.83	18.01	314.6	2236	869.7
250	1.99415	0.02073	0.02023	462.6	39.00	14.69	231.9	1881	726.2
500	1.99418	0.02075	0.02019	487.2	43.67	9.208	244.3	2104	455.9

3.4 Conclusions

In this chapter, first, we introduced a new stochastic diffusion process associated with the Lundqvist-Korf growth curve. Its distribution and main characteristics were analyzed, and its mean function as well as its conditional mean function was found to be proportional to the Lundqvist-Korf growth curve. In this work we have developed the theoretical base for the practical use of Lundqvist-Korf-type diffusion processes as a particular case of the stochastic log-normal diffusion process, and applying the simulated annealing algorithm in order to solve inference problems. In the end, an application to simulated data of the proposed model to an example showed its usefulness in practice then, the capability of the model for forecasting and predicting is shown.

Secondly, we define and examine a new extension of the stochastic Lundqvist-Korf diffusion process by the powers of the stochastic Lundqvist-Korf diffusion process. Them we define the proposed model as the solution to a stochastic differential equation. From this, we obtain the explicit expression of the process, the probability transition density function, the moments of different orders and, in particular, the conditioned and unconditioned mean functions of the process. The third section, discusses parameter estimation by ML method, using discrete sampling. In this case, ML estimators cannot be given in closed form because the system of likelihood equations does not have an explicit solution. The simulated annealing method is proposed to calculate or approximate the resulting ML estimator. Then we validate this methodologies by the results obtained from the simulation examples.

Therefore, the studied variable can be applied in several fields in the future studies.

CHAPTER 4

The stochastic modified Lundqvist-Korf diffusion process: Statistical inference and application

Contents

This chapter is the complete version of the paper [60] published in Stochastic Environmental Research and Risk Assessment, vol. 36, 1163-1176, 2021. https://doi.org/10.1007/s00477-021-02089-8.

4.1 Introduction

Stochastic diffusion processes (SDPs) play an efficient role in modeling several phenomena in various disciplines. We will quote, for example, physics, biology, economy and finance, radiotherapy, chemotherapy, energy consumption, and others. In the same perspective, many types of stochastic diffusion processes have been defined, such as the stochastic Gompertz diffusion process [15, 16], the stochastic Gamma diffusion process [4, 84], the stochastic Weibull diffusion process [54].

The Lundqvist-Korf growth curve belongs to the smooth sigmoidal functions arising from tree growth and forest populations. It was developed by Korf [72]. In this respect, Zarnovican [74] introduced a brief analysis of Korf's mathematical formula. It was applied to forest mensuration, such as the relation between height and age for three black spruce stands. Nafidi and El Azri [28] proposed a new non-homogeneous stochastic diffusion process in which the trend function is proportional to the growth curve of the Lundqvist-Korf. Crescenzo and Spina [80] and Crescenzo and Paraggio [85] introduced a new deterministic growth model which captures certain features of both the Gompertz and Korf laws. We also perform a comparison between the Logistic growth process and other models, such as the Gompertz model, the Korf model, and the modified Lundqvist-Korf model.

The evolution of CO_2 emissions is currently one of the most significant subjects in environmental science, climate change and health (see, for instance Amit et al. [86, 87], Balram et al. [88–90], Balram and Shrikanta [91], Susanta and Balram [92]). In this context, various SDPs have been applied to describe and forecast this issue. In this perspective, see, for example, Gutiérrez et al. [32] in the study of the SDP with cubic drift with application to a modeling of the global CO_2 emissions in Spain. Then, Gutiérrez et al. [33] in the case of the non-homogeneous (with exogenous factors) stochastic Vasicek diffusion process in the case of CO_2 emission in Morocco. Note that the non-homogeneity in this process is obtained by adding exogenous factors such as external variables that affect the drift of the homogeneous Vasicek model. Moreover, Gutiérrez et al. [34] proposed the bivariate stochastic Gompertz diffusion model as the solution for a system of two Itô's stochastic differential equations (SDEs). The drift and diffusion coefficients are similar to those considered in the univariate stochastic Gompertz diffusion process. This stochastic model is applied to the modeling of the gross domestic product and CO_2 emissions in Spain. Moreover, Abbass et al. [35] presented a systematic review of two decades of research from 1995 to 2017 for CO_2 emissions and economic growth. Magazzino and Cerulli [36] examined the relationship among CO_2 emissions, GDP, and energy in the Middle East and North Africa countries using a responsiveness scores approach. Then, Solaymani [37] treated the CO_2 emissions patterns in seven top carbon emitter economies in the case of the transport sector. Furthermore, several techniques, different from the one we proposed in this work, have been used to study CO_2 emissions in several countries and different geographic regions of the world, we will cite, for example, regression models (see, Wang et al. [93], Lin and Xu [94], Hosseini et al. [95]), and temporal series models (see, Nguyen and Le [96]).

The question of statistical inference in SDPs has received considerable attention in recent years, both when the process is observed continuously and when it is observed discretely. The estimation of the parameters in stochastic models, in general, is not direct, except in simple cases and one possible methodological approach is based on approximating the maximum likelihood function. In the same context, various methods addressing this question have been developed, and many papers have been published on this subject, focusing on several variants of approximate likelihood methodology. The general case of this methodology can be consulted in Prakasa-Rao [41], Bibby et al. [22], Ait-Sahalia [42] and Egorov et al. [43], and in the case of particular diffusions, the following can be seen, for example, Gutiérrez et al. [15, 17]. However, the approach of maximum likelihood estimation of the parameters using likelihood equations can be problematic, which is why we propose the use of the simulated annealing (SA) method. This algorithm is a method for solving unconstrained and bound-constrained optimization problems developed by Kirkpatrick et al. [58]. Duflo [81] used for optimization in continuous spaces the method models the physical process of heating a material and then slowly lowering the temperature to decrease defects, thus minimizing the system energy. Recently, many works have used the SA algorithm for estimating the parameters in the stochastic diffusion process (see, for instance, Nafidi et al. [28, 54]).

The main objective of this chapter is to introduce a new non-homogeneous stochastic modified Lundqvist-Korf diffusion process (SMLKDP). In our work, the non-homogeneity in the process is of nature. That is to say, that in the definition of the process its drift is dependent on time. Then in the future, it is possible to introduce exogenous factors in this process to obtain a version of double non-homogeneity which presents a trend function (TF) that is proportional to the modified Lundqvist-Korf growth curve in Eq. (4.2.2). Also, we can apply this process to fit and forecast real data. The rest of this paper is organized as follows: In section 4.2, we present an overview of the Lundqvist–Korf growth curve and we define the proposed model in terms of the stochastic differential equation. We then determine the explicit expression of the solution to the SDE, the transition probability density function (TPDF) and the trend functions. In section 4.3, we discuss the parameter estimation using the maximum likelihood method (MLM) on the basis of discrete sampling of the process. Since the closed form of the MLM estimators cannot be given because the system of likelihood equations does not have an explicit form, numerical methods are needed. We deal with this problem through the application of SA method: First, a brief summary of the algorithm is provided, and then its adaptation to the problem at hand is presented. Some strategies are suggested for bounding the space of solutions, and a description is provided for the application of the algorithms selected. In Section 4.4, we present the results obtained from the examples of simulation, and we illustrate the predictive study by fitting the SDP to simulated data. In Section 4.5, we present an application to real data by considering the evolution of global CO_2 emissions in Morocco. This subject is the primary driver of global climate change. It is widely recognized that to avoid the worst impacts of climate change. Finally, in the last section, we recapitulate the main conclusions from this study.

4.2 The model and its characteristics

In this section, we provide an overview of the modified growth of the Lundqvist-Korf curve and defines the proposed process in terms of SDE. We then determine the explicit expression of the solution to the SDE, the TPDF, and the trend functions.

4.2.1 The modified Lundqvist-Korf curve

The most commonly used expression of the Lundqvist-Korf curve is:

$$y(t) = K \exp\left(-\frac{\beta}{t^{\alpha}}\right); \quad t > 0,\ \alpha > 0,\ \beta > 0, \tag{4.2.1}$$

where K is the upper bound for the studied variable, that can only be reached after infinite time and is the scale parameter. If in the Lundqvist-Korf curve Eq. (4.2.1), we replace t by $1+t$, then we obtain the modified Lundqvist-Korf growth curve:

$$x(t) = K \exp\left(-\frac{\beta}{(1+t)^{\alpha}}\right); \quad t \geq 0,\ \alpha > 0,\ \beta > 0.$$

We impose that $x(t_1) = x_1 > 0$, hence $K = x_1 \exp\left(\frac{\beta}{(1+t_1)^{\alpha}}\right)$. Thus, we reach:

$$x(t) = x_1 \exp\left(\frac{\beta}{(1+t_1)^{\alpha}}\right) \exp\left(-\frac{\beta}{(1+t)^{\alpha}}\right); \quad t \geq t_1,\ \alpha > 0,\ \beta > 0. \tag{4.2.2}$$

This curve is a sigmoidal strictly increasing curve showing an inflection point at
$t_I = \left(\frac{\alpha\beta}{\alpha+1}\right)^{1/\alpha} - 1$, its value is

$$x(t_I) = x_1 \exp\left(\frac{\beta}{(1+t_1)^{\alpha}}\right) \exp\left(-1 - \frac{1}{\alpha}\right).$$

In addition, $t_I > t_1$ (that is, the inflection can be visualized) if and only if

$$\beta > \left(1 + \frac{1}{\alpha}\right) t_1^{\alpha} > t_1^{\alpha}.$$

Whereas $\lim_{t\to\infty} x(t) = x_1 \exp\left(\frac{\beta}{(1+t_1)^{\alpha}}\right)$, that is the line $y = x_1 \exp\left(\frac{\beta}{(1+t_1)^{\alpha}}\right)$ is a horizontal asymptote of the curve in Eq. (4.2.2) when t tends to ∞. Its asymptote is dependent on the initial value.

4.2.2 The SMLKDP model

To model the modified Lundqvist-Korf type behaviors from a stochastic point of view, our contribution is to consider a SDP whose trend function has the expression given in the curve in Eq. (4.2.2). Now, starting from the curve in Eq. (4.2.2), one is lead to considering the ODE

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = h(t)x(t); \quad x(t_1) = x_1, \tag{4.2.3}$$

where $h(t) = \dfrac{\alpha\beta}{(1+t)^{\alpha+1}}$ and the parameters α, β and σ are all positive. Hence, Eq. (4.2.3) can be viewed as a generalisation of the Malthusian growth model with time dependent fertility rate $h(t)$. Note that $h(t)$ is a decreasing continuous positive and bounded function and has a horizontal asymptote at $y = 0$.

The form of the proposed one-dimensional SMLKDP is defined as a diffusion process $\{x(t) : t \in [t_1, T],\ t_1 \geq 0\}$, taking values on $(0, \infty)$ and characterized by these infinitesimal moments (drift and diffusion coefficient):

$$\begin{aligned} A_1(t,x) &= h(t)x, \\ A_2(t,x) &= \sigma^2 x^2, \end{aligned} \tag{4.2.4}$$

with initial distribution $x(t_1) = x_1$.

Alternatively, the above process can be defined as the unique solution of the following Itô's SDE:

$$\mathrm{d}x(t) = A_1(t,x)\mathrm{d}t + (A_2(t,x))^{\frac{1}{2}}\,\mathrm{d}w(t); \quad x(t_1) = x_1, \tag{4.2.5}$$

where $w(t)$ is a standard Wiener process and x_1 is a positive random variable, independent on $w(t)$ for $t \geq t_1$.

The infinitesimal moments $A_1(t,x)$ and $A_2(t,x)$ specified in Eq. (4.2.4) satisfy the Lipschitz and the growth conditions for the existence and uniqueness of the solution to the SDEs (see, Kloeden and Platen [8]). Thus, there exists a non negative constant $C = \dfrac{\alpha\beta}{(1+t_1)^{\alpha+1}} + \sigma$, such that for all x and y of $\mathbb{R}^+$ and t of $[t_1, T]$, we have:

$$\begin{aligned} &|A_1(t,x) - A_1(t,y)| + \left|(A_2(t,x))^{\frac{1}{2}} - (A_2(t,y))^{\frac{1}{2}}\right| \leq C \mid x - y \mid, \\ &|A_1(t,x)|^2 + \left|(A_2(t,x))^{\frac{1}{2}}\right|^2 \leq C^2 \left(1 + \mid x \mid^2\right). \end{aligned}$$

Then, the SDE in Eq. (4.2.5) has a unique solution $\{x(t) : t \in [t_1, T], \quad t_1 > 0\}$ continuous with probability 1, and satisfies the initial condition $x(t_1) = x_1$.

4.2.3 Probability distribution of the SMLKDP

By means of the appropriate transformation of the form $y(t) = \ln(x(t))$, and by using the Itô rule, the SDE Eq. (4.2.5) becomes:

$$\mathrm{d}y(t) = \left(\frac{\alpha\beta}{(1+t)^{\alpha+1}} - \frac{\sigma^2}{2}\right)\mathrm{d}t + \sigma \mathrm{d}w(t); \quad y_1 = \ln(x_1),$$

by integrating both sides yields,

$$y(t) = y_1 - \beta\left(\frac{1}{(1+t)^{\alpha}} - \frac{1}{(1+t_1)^{\alpha}}\right) - \frac{\sigma^2}{2}(t-t_1) + \sigma\left(w(t) - w(t_1)\right).$$

Finally, we have:

$$x(t) = x_1 \exp\left[-\beta\left(\frac{1}{(1+t)^{\alpha}} - \frac{1}{(1+t_1)^{\alpha}}\right) - \frac{\sigma^2}{2}(t-t_1) + \sigma\left(w(t) - w(t_1)\right)\right]. \quad (4.2.6)$$

The process $y(t)$ is a gaussian process if and only if $\ln(x_1)$ is constant or normally distributed (see, for instance, Arnold [82]). In such a case, the process $x(t)$ is a Lognormal process. That is, the TPDF:

$$f(x,t|x_s,s) = \frac{1}{x\sqrt{2\pi\sigma^2(t-s)}} \exp\left(-\frac{\left(\ln(x) - \mu(s,t,x_s)\right)^2}{2\sigma^2(t-s)}\right), \quad (4.2.7)$$

where $\mu(s,t,x_s) = \ln(x_s) - \beta\left(\frac{1}{(1+t)^{\alpha}} - \frac{1}{(1+s)^{\alpha}}\right) - \frac{\sigma^2}{2}(t-s)$.

4.2.4 The moments of the SMLKDP

By using the properties of the log-normal distribution, we obtain the conditional moments of order r of this process:

$$\begin{aligned} E\left(x^r(t)|x(s)=x_s\right) &= \exp\left(r\mu(s,t,x_s) + \frac{r^2\sigma^2}{2}(t-s)\right) \\ &= \exp\left\{\ln(x_s^r) - \beta r\left(\frac{1}{(1+t)^{\alpha}} - \frac{1}{(1+s)^{\alpha}}\right) + \frac{r(r-1)\sigma^2}{2}(t-s)\right\} \\ &= x_s^r \exp\left(-\beta r\left(\frac{1}{(1+t)^{\alpha}} - \frac{1}{(1+s)^{\alpha}}\right)\right) \exp\left(\frac{r(r-1)\sigma^2}{2}(t-s)\right). \end{aligned}$$

Then, for $r = 1$, the conditional TF (CTF) of the process is:

$$E\left(x(t)|x(s)=x_s\right) = x_s \exp\left(-\beta\left(\frac{1}{(1+t)^{\alpha}} - \frac{1}{(1+s)^{\alpha}}\right)\right). \quad (4.2.8)$$

In addition, considering the initial condition $\mathrm{P}\left(x(t_1) = x_1\right) = 1$, the TF of the process is given by:

$$E(x(t)) = x_1 \exp\left(\frac{\beta}{(1+t_1)^{\alpha}}\right) \exp\left(-\frac{\beta}{(1+t)^{\alpha}}\right). \quad (4.2.9)$$

Note that the TF (4.2.9) satisfies the deterministic counterpart

$$dE\left(x(t)\right) = h(t)E\left(x(t)\right)dt.$$

This connects the stochastic and the deterministic models.

4.3 Statistical inference on the model

This section examines the MLM estimators of the parameters α, β and σ^2 of the model using discrete sampling. Then, by Zehna's theorem [68], we obtain the corresponding estimated TFs (ETFs) of the process.

4.3.1 Parameters estimation

Let us consider a discrete sampling of the process, based on d sample paths, for fixed times t_{ij}, $(i = 1, \cdots, d, \quad j = 1, \cdots, n_i)$ with $t_{i1} = t_1, \quad i = 1, \cdots, d$. That is, we observe the variables $x(t_{ij})$ whose values provide the basic sample for the inference process. In addition, we assume $t_{ij} - t_{i,j-1} = h$, for $i = 1, \cdots, d; j = 2, \cdots, n_i$. Let $x = \{x_{ij}\}_{i=1,\cdots,d; j=1,\cdots,n_i}$, be the observed values of the sampling. The likelihood function (LF) depends on the choice of the initial distribution.

If $\mathrm{P}(x(t_1) = x_1) = 1$, then the LF is

$$\mathrm{L}_x\left(\alpha, \beta, \sigma^2\right) = \prod_{i=1}^{d} \prod_{j=2}^{n_i} f\left(x_{ij}, t_{ij} | x_{i,j-1}, t_{i,j-1}\right).$$

If $x(t_1)$ is distributed as one-dimensional lognormal distribution $\Lambda_1\left(\mu_1, \sigma_1^2\right)$, then the LF is:

$$\mathrm{L}_x\left(\mu_1, \sigma_1^2, \alpha, \beta, \sigma^2\right) = \prod_{i=1}^{d} f_{x(t_1)}\left(x_{i1}\right) \prod_{j=2}^{n_i} f\left(x_{ij}, t_{ij} | x_{i,j-1}, t_{i,j-1}\right),$$

where μ_1, σ_1^2, α, β and σ^2 are the parameters to be estimated.

Henceforth, we will consider the case when the initial distribution is lognormal. Denoting $n = \sum_{i=1}^{d} n_i$, from Eq. (4.2.7), the log-likelihood function (LLF) of the sample is:

$$\begin{aligned}
\ln \mathrm{L}_x\left(\mu_1, \sigma_1^2, \alpha, \beta, \sigma^2\right) &= -\frac{n}{2}\ln(2\pi) - \frac{d}{2}\ln\left(\sigma_1^2\right) - \frac{(n-d)}{2}\ln\left(\sigma^2\right) \\
&- \sum_{i=1}^{d} \ln\left(x_{i1}\right) - \frac{1}{2\sigma_1^2}\sum_{i=1}^{d}\left(\ln\left(x_{i1}\right) - \mu_1\right)^2 - \frac{(n-d)}{2}\ln(h) \\
&- \frac{1}{2\sigma^2 h}\sum_{i=1}^{d}\sum_{j=2}^{n_i}\left\{\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right) + \beta\left(\frac{1}{(1+t_{ij})^{\alpha}} - \frac{1}{(1+t_{i,j-1})^{\alpha}}\right) + \frac{\sigma^2}{2}h\right\}^2.
\end{aligned} \tag{4.3.1}$$

By differentiating the LLF in Eq. (4.3.1) with respect to the parameters μ_1 and σ_1^2, we obtain the ML estimators of μ_1 and σ_1^2 are

$$\hat{\mu}_1 = \frac{1}{d}\sum_{i=1}^{d}\ln\left(x_{i1}\right) \quad \text{and} \quad \hat{\sigma}_1^2 = \frac{1}{d}\sum_{i=1}^{d}\left(\ln\left(x_{i1}\right) - \hat{\mu}_1\right)^2.$$

The estimation of the parameters α, β and σ^2 poses some difficulties. The resulting system of equations is exceedingly complex and does not have an explicit solution. Therefore numerical procedures must be employed. To address this problem, we propose the use of the simulated annealing algorithm to maximize the likelihood function or, equivalently, its logarithm.

4.3.2 Estimated TFs and CB

Using the invariance property of the ML estimator (see, Zehna [68]), we determine the estimated TFs (ETFs) and the confidence bounds (CB). Therefore, if $\widehat{\lambda}$ is a ML estimator for λ, then $u\left(\widehat{\lambda}\right)$ is a ML estimator for $u\left(\lambda\right)$ where u is some function of λ.

4.3.2.1 Estimated TFs

By using Zehna's theorem [68], we obtain the ETFs. Then, the estimated CTF (ECTF) is given by the following expression:

$$\hat{E}\left(x(t)|x(s)=x_s\right)=x_s\exp\left(-\hat{\beta}\left(\frac{1}{(1+t)^{\hat{\alpha}}}-\frac{1}{(1+s)^{\hat{\alpha}}}\right)\right). \tag{4.3.2}$$

Taking into account the initial condition that is $\mathrm{P}\left(x\left(t_1\right)=x_1\right)=1$, the ETF of the process is given by:

$$\hat{E}\left(x(t)\right)=x_1\exp\left(\frac{\hat{\beta}}{(1+t_1)^{\hat{\alpha}}}\right)\exp\left(-\frac{\hat{\beta}}{(1+t)^{\hat{\alpha}}}\right). \tag{4.3.3}$$

4.3.2.2 Confidence bounds

The CB of the process are obtained by using the procedure described in Katsamaki and Skiadas [67]. Let $v(s,t)$ be the random variable distributed with the probability law $x(t)|x(s)=x_s$. Since the variable $w(t)-w(s)$ is Gaussian with the mean equal to zero and the variance $t-s$ for $t\geq s$. Therefore, the random variable z is given by

$$z=\frac{\ln(v(s,t))-\mu\left(s,t,x_s\right)}{\sigma\sqrt{t-s}}\sim N_1\left(0,1\right).$$

A $100(1-\kappa)\%$ conditional CB for z is given by $P(-\xi\leq z\leq\xi)=1-\kappa$. From this, we can obtain a CB of $v(s,t)$ with following form $(v_{\text{lower}}(s,t),v_{\text{upper}}(s,t))$, where

$$v_{\text{lower}}(s,t)=\exp\left(\mu\left(s,t,x_s\right)-\xi\sigma\sqrt{(t-s)}\right), \tag{4.3.4}$$

and

$$v_{\text{upper}}(s,t)=\exp\left(\mu\left(s,t,x_s\right)+\xi\sigma\sqrt{(t-s)}\right), \tag{4.3.5}$$

with $\xi=F^{-1}_{\mathrm{N}(0,1)}\left(1-\kappa/2\right)$ and where $F^{-1}_{\mathrm{N}(0,1)}$ is the inverse cumulative normal standard distribution.

By substituting the parameters by theirs estimators in the equations Eq. (4.3.4) and Eq. (4.3.5), we obtain the estimated lower bound $v_{\text{lower}}(t)$ and the estimated upper bound $\hat{v}_{\text{upper}}(t)$:

$$\hat{v}_{\text{lower}}(s,t) = \exp\left(\hat{\mu}\left(s,t,x_s\right) - \xi\hat{\sigma}\sqrt{(t-s)}\right), \tag{4.3.6}$$

and

$$\hat{v}_{\text{upper}}(s,t) = \exp\left(\hat{\mu}\left(s,t,x_s\right) + \xi\hat{\sigma}\sqrt{(t-s)}\right), \tag{4.3.7}$$

where $\hat{\mu}\left(s,t,x_s\right) = \ln\left(x_s\right) - \hat{\beta}\left(\frac{1}{(1+t)^{\hat{\alpha}}} - \frac{1}{(1+s)^{-\hat{\alpha}}}\right) - \frac{\hat{\sigma}^2}{2}\left(t-s\right)$.

4.3.3 Application of the SA method

4.3.3.1 The algorithm

SA algorithm is a metaheuristic algorithm to approximating the solution of optimization problems of the type $\min_{\theta\in\Theta} g(\theta)$, where g is the objective function to be optimized and Θ is the solution space.

In our case, the problem becomes maximizing the function $\ln \mathrm{L}_x\left(\mu_1, \sigma_1^2, \alpha, \beta, \sigma^2\right)$. Since the algorithm aforementioned is usually formulated for minimization problems, then, from Eq. (4.3.1), the objective function is a function of the parameters α, β and σ^2 and has the following form:

$$\begin{aligned} g\left(\alpha,\beta,\sigma^2\right) &= \frac{(n-d)}{2}\ln\left(\sigma^2\right) + \frac{1}{2\sigma^2 h} \\ &\times \sum_{i=1}^{d}\sum_{j=2}^{n_i}\left\{\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right) + \beta\left(\frac{1}{(1+t_{ij})^{\alpha}} - \frac{1}{(1+t_{i,j-1})^{\alpha}}\right) + \frac{\sigma^2}{2}h\right\}^2, \end{aligned} \tag{4.3.8}$$

and the parametric space Θ linked to the objective function Eq. (4.3.8), on which the selected algorithms must operate, is continuous and unbounded. Consequently

$$\Theta = \{(\alpha,\beta,\sigma):\quad \alpha>0,\ \beta>0,\ \sigma>0\}.$$

The drawback is that the solution space might not be explored with enough depth. This requires us to find arguments for bounding said space. The following are some strategies to this end.

4.3.3.2 Bounding the solution space

Regarding the parameter σ, when it has high values it leads to sample paths with great variability around the mean of the process. Thus, excessive variability in available paths would make modified Lundqvist-Korf-type modeling inadvisable (see Figure 4.1). Some simulations performed for several values of σ have led us to consider that $0<\sigma<0.1$, so that we may have paths compatible with the modified Lundqvist-Korf-type growth. On the other hand, there does not seem to be an upper bound for α and β. To this end, we will use the following reformulation: Setting $b = \exp(-\beta)$ and $a = \dfrac{1}{\alpha}$. This reformulation leads to the condition

$$0 < b < 1$$

and does not seem to be an upper bound for a. Nevertheless, the modified Lundqvist-Korf curve is sigmoidal and has an inflection point at $t_I = \left(-\dfrac{\ln(b)}{a+1}\right)^a - 1$, which is higher than t_1 if and only if

$$b < \exp\left(-t_1^{\frac{1}{a}}\right).$$

The parameter a can be bounded taking into account the information provided by the sample paths and the asymptote of the curve verifying

$$k = x_1 \left(\frac{1}{b}\right)^{(1+t_1)^{-\frac{1}{a}}}$$

and the inflection point of the curve at

$$t_I = \left(\frac{-\ln(b)}{a+1}\right)^a - 1.$$

From these expressions, and if we denote by k_i the maximum value of the $i-$th sample path, the following expression holds

$$0 < a < \ln(\eta) - 1,$$

where $\eta = \max\limits_{i=1,\cdots,d} \left(\dfrac{k_i}{x_{i1}}\right)$ and x_1 is the initial value of the sample path. Thus, the solution space, which is obtained numerically for a, b and σ is bounded and takes the form

$$(0, \ln(\eta) - 1) \times \left(0, \exp\left(-t_1^{\frac{1}{\ln(\eta)-1}}\right)\right) \times (0, 0.1).$$

Once the solution space has been bounded, we specify the choice of the initial parameters of the algorithms and the stopping conditions. Let consider:

1. The initial solution is chosen randomly in the bounded subspace

$$(0, \ln(\eta) - 1) \times \left(0, \exp\left(-t_1^{\frac{1}{\ln(\eta)-1}}\right)\right) \times (0, 0.1).$$

2. The initial temperature should be high enough such that in the first iteration of the algorithm, the probability of accepting a worse solution is at least of 80% (see, Kirkpatrick et al. [58]). For this, we assume the initial temperature of 10.

3. For the cooling process we have considered a geometric scheme in which the current temperature is multiplied by a constant γ ($0 < \gamma < 1$), i.e. $T_i = \gamma T_{i-1}$, $i \geq 1$. The usual values for γ are between 0.80 and 0.99. For this we have set $\gamma = 0.95$.

4. The length of each temperature level L determines the number of solutions generated at a certain temperature, T. For this we have set $L = 50$.

5. The stopping criterion defines when the system has reached a desired energy level (freezing temperature). Equivalently it defines the total number of solutions generated, or when an acceptance ratio (ratio between the number of solutions accepted and the number of solutions generated) is reached. The application of the algorithm will be limited to 1000 iterations.

The coding is performed using Matlab computer software.

4.4 Simulation study

In this section, we present some simulated sample paths of the SMLKDP and demonstrate the performance of the proposed procedure using simulation examples. Then, we validate the predictive study by fitting this process to simulated data.

4.4.1 Simulated sample paths of the SMLKDP

This section presents some simulated sample paths of the SMLKDP. We obtain the trajectory of the model by simulating the exact solution of the SDE Eq. (4.2.6). We obtain the simulated trajectories of the model by considering the time discretization of the interval $[t_1, T]$, with time points $t_i = t_{i-1} + (i-1)h$, $i = 2, \cdots, N$ and $h = \dfrac{T - t_1}{N}$ is the discretization step size for the sample size N. The random variable $\sigma\left(w(t) - w\left(t_1\right)\right)$ in the Eq. (4.2.6) is distributed as one-dimensional normal distribution $N_1\left(0, \sigma^2\left(t - t_1\right)\right)$. This simulation is based on 25 sample paths with $t_1 = 0$, $T = 40$, $x_1 \sim \Lambda_1(1, 0.16)$, and $N = 250$.

Fig. 4.1 shows some simulated sample paths of the SMLKDP and its trend function for several values of α, β and σ.

4.4.2 Estimation of drift parameters and the diffusion coefficient

In this section, we present several examples to evaluate the performance of the estimation procedure previously developed. To this end, Eq. (4.2.6) was simulated 25 times under the following assumptions $t_1 = 0$, $h = 1$, $x_1 \sim \Lambda_1(1, 0.16)$ and the samples of sizes $N = 100$, 250 and 500 are used to investigate the effects of the sample size on the performances of the estimation procedure. In order to make the subsequent inference we have considered, in each example the 25 sample paths with $t_i = t_{i-1} + (i-1)h$, $i = 2, \cdots, N$.

The empirical mean, the std, and the CV for a, b and σ (functions of parameters α, β and σ which are the arguments of the objective function) are defined in Table 3.1. Then, Table 4.1 shows the results obtained for calculating the latter measures. The results obtained show the performance of the methodology.

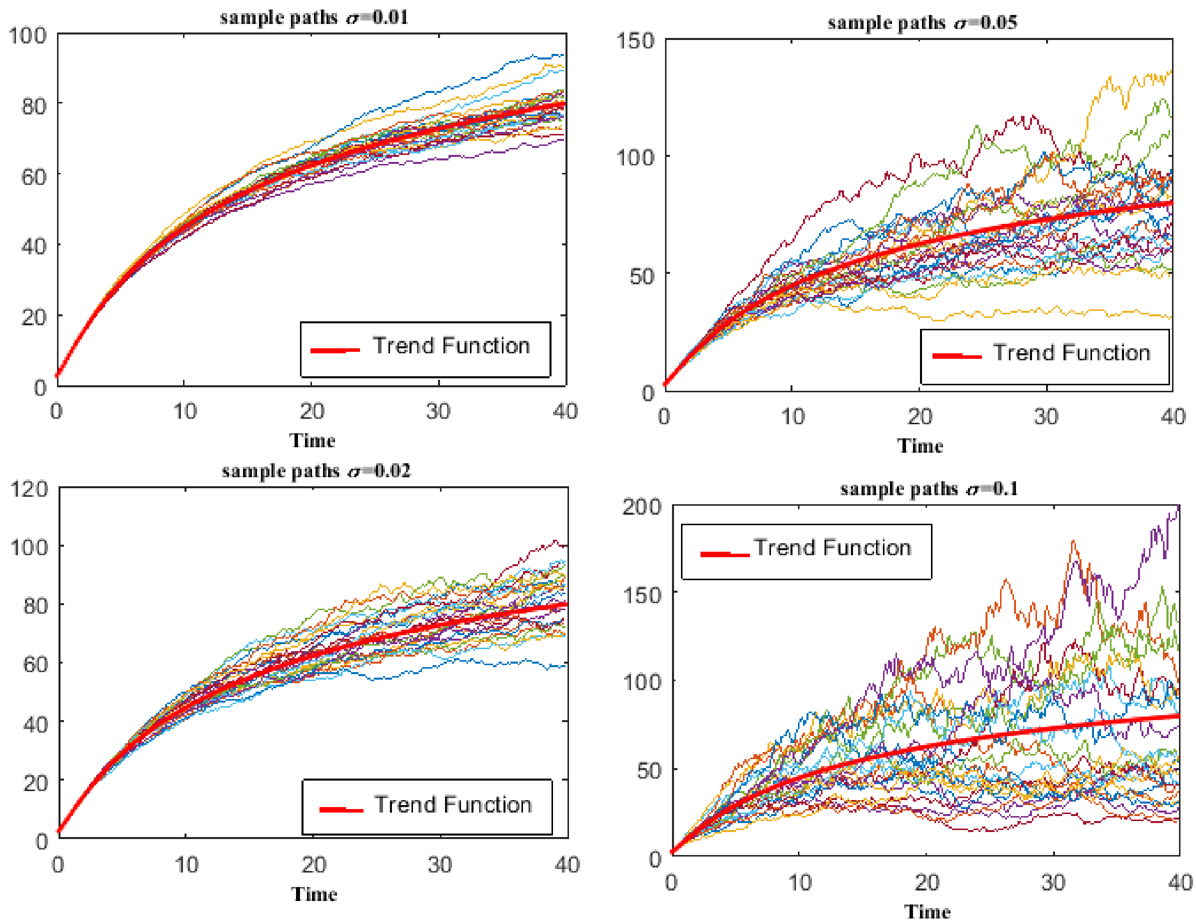


Figure 4.1: Simulated trajectories of the SMLKDP and its TF for several values of σ with $(\alpha = 0.5, \beta = 4)$.

Table 4.1: Estimation values, the std and the CV of a, b and σ for several values of σ.

$a = 2, b = 0.02$									
$\sigma = 0.01$									
N	$\overline{a}$	$\overline{b}$	$\overline{\sigma}$	std(a)	std(b)	std(σ)	CV(a)	CV(b)	CV(σ)
100	1.99539	0.02031	0.00965	0.06374	0.00223	0.00150	0.03195	0.11004	0.15546
250	2.00543	0.02006	0.00972	0.07726	0.00195	0.00084	0.03853	0.09709	0.08686
500	2.00019	0.02003	0.00961	0.07988	0.00266	0.00087	0.03994	0.13298	0.09129
$\sigma = 0.02$									
100	2.11445	0.01764	0.01926	0.19080	0.00448	0.00248	0.09023	0.25418	0.12905
250	2.03829	0.02006	0.01911	0.11803	0.00315	0.00219	0.05791	0.15706	0.11468
500	2.01817	0.02047	0.01856	0.04587	0.00285	0.00101	0.02273	0.13961	0.05450
$\sigma = 0.025$									
100	2.09409	0.01894	0.02277	0.22070	0.00524	0.00260	0.10539	0.27683	0.11421
250	2.03920	0.02106	0.02329	0.13176	0.00419	0.00257	0.06461	0.19923	0.11051
500	2.11080	0.01807	0.02346	0.19919	0.00410	0.00219	0.09436	0.22719	0.09347

4.4.3 Predicted data using ETF and ECTF

In this section, we have considered the predictive study based on fitting the diffusion process to simulated data in which $N = 25$ and $t_i = t_{i-1} + (i-1)h$, $i = 2, \cdots, N$ starting at $t_1 = 1$, taking the step size $h = 0.05$, and $x_1 = 2.6574$. First, we use the first 22 data to estimate the parameters a, b and σ^2 of the process by SA. Moreover, we obtain the corresponding ETF and ECTF values given by the expressions (4.3.3) and (4.3.2). For the three last data, we predict the corresponding values using the ETF and ECTF. Also, we give the results attached to a 95% estimated CB (ECB) and a 95% estimated conditional CB (ECCB) of the process (see, the expressions (4.3.6) and (4.3.7)). Finally, to illustrate the performance of the procedure, the results according to the MAE, the RMSE, and the MAPE, given by Table 2.1. From Lewis [69], we conclude the accuracy of the forecast from the MAPE result in Table 2.2.

Table 4.2 shows the results obtained of the simulated values with the ETF, ECTF, ECB, and ECCB.

Table 4.2: Simulated and predicted data, showing ETF, ECTF, ECB and ECCB using the SM-LKDP.

i	Simulated data $x(t_i)$	ETF	ECTF	ECB	ECCB
1	2.65740	2.65740	2.65740	[2.65740 2.65740]	[2.65740 2.65740]
2	3.12870	3.09158	3.09158	[3.05598 3.12748]	[3.05598 3.12748]
3	3.61360	3.55041	3.59303	[3.49269 3.60882]	[3.55166 3.63476]
4	4.10860	4.03040	4.10213	[3.95027 4.11173]	[4.05490 4.14978]
5	4.61090	4.52808	4.61593	[4.42424 4.63371]	[4.56278 4.66954]
6	5.11820	5.04005	5.13223	[4.91096 5.17163]	[5.07313 5.19183]
7	5.62840	5.56305	5.64930	[5.40711 5.72229]	[5.58425 5.71492]
8	6.14000	6.09402	6.16560	[5.90966 6.28260]	[6.09460 6.23721]
9	6.65170	6.63012	6.68014	[6.41586 6.84963]	[6.60322 6.75773]
10	7.16250	7.16877	7.19210	[6.92323 7.42069]	[7.10928 7.27563]
11	7.67140	7.70762	7.70088	[7.42955 7.99333]	[7.61220 7.79032]
12	8.17770	8.24460	8.20584	[7.93284 8.56533]	[8.11135 8.30115]
13	8.68080	8.77785	8.70662	[8.43139 9.13474]	[8.60636 8.80774]
14	9.18020	9.30578	9.20289	[8.92371 9.69982]	[9.09691 9.30977]
15	9.67550	9.82700	9.69438	[9.40854 10.2590]	[9.58274 9.80697]
16	10.1664	10.3403	10.1809	[9.88482 10.8111]	[10.0636 10.2991]
17	10.6527	10.8447	10.6623	[10.3516 11.3550]	[10.5395 10.7862]
18	11.1340	11.3395	11.1387	[10.8083 11.8898]	[11.0104 11.2680]
19	11.6104	11.8239	11.6096	[11.2542 12.4146]	[11.4759 11.7444]
20	12.0817	12.2974	12.0753	[11.6890 12.9290]	[11.9363 12.2156]
21	12.5477	12.7597	12.5358	[12.1123 13.4323]	[12.3914 12.6814]
22	13.0085	13.2104	12.9909	[12.5239 13.9243]	[12.8412 13.1417]
Prediction					
23	13.4639	13.6493	13.4407	[12.9237 14.4046]	[13.2859 13.5968]
24	13.9140	14.0763	13.8851	[13.3115 14.8731]	[13.7252 14.0464]
25	14.3588	14.4914	14.3243	[13.6875 15.3298]	[14.1593 14.4907]

Table 4.3 shows the values obtained from the estimation of the parameters of the SMLKDP.

Table 4.4 shows the values obtained after the calculation of the three measures of the goodness of fit. Then, we evaluated the accuracy of the forecast from the MAPE. In other words, if the value of the MAPE is less than 10%, the forecast is highly accurate.

Fig. 4.2 and Fig. 4.3 illustrate the performance of the SMLKDP for forecasting using the trend function and the conditional trend function.

Table 4.3: Estimation of the parameters of the SMLKDP.

Parameters	a	b	σ
Estimated value	0.374785121021031	0.000000280539236	0.026387464094434

Table 4.4: Goodness of fit of the SMLKDP.

MAE	RMSE	MAPE
0.11408	0.13433	1.26105%

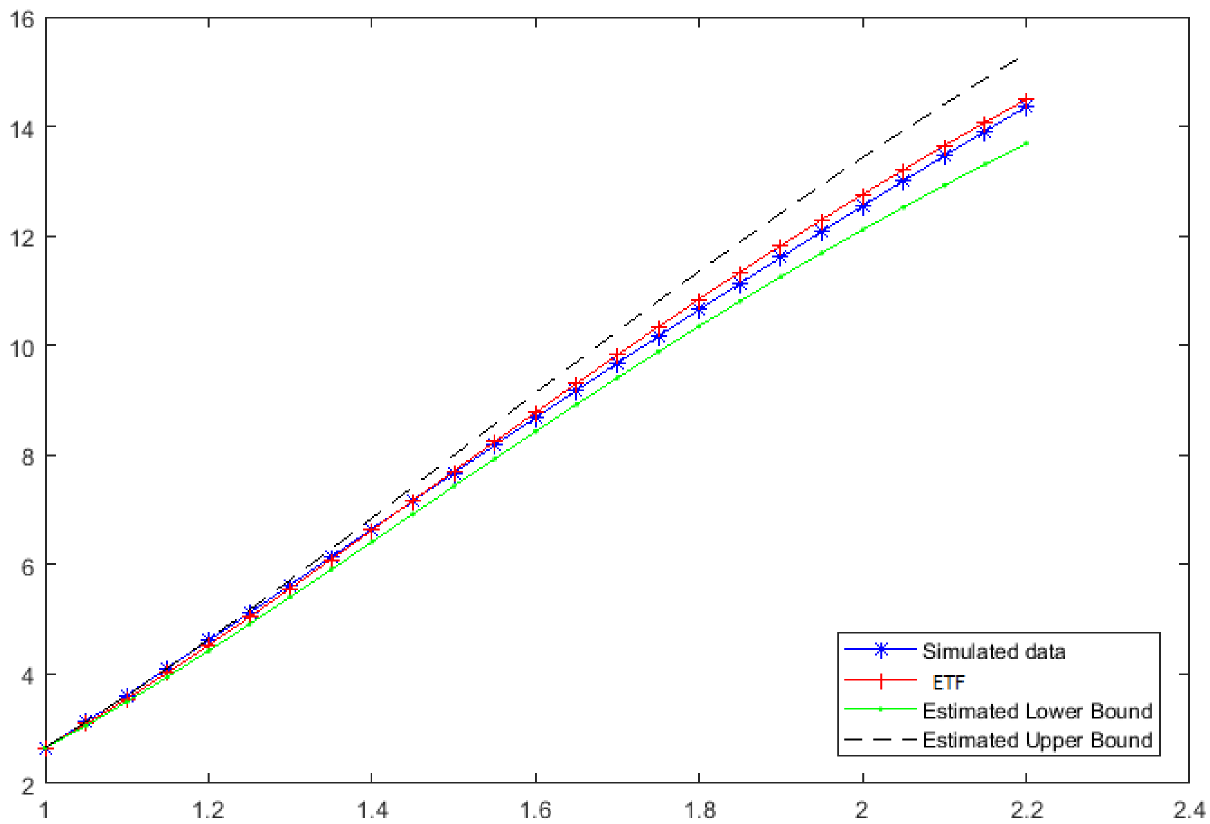


Figure 4.2: Simulated data, ETF, ECB and the predicted values.

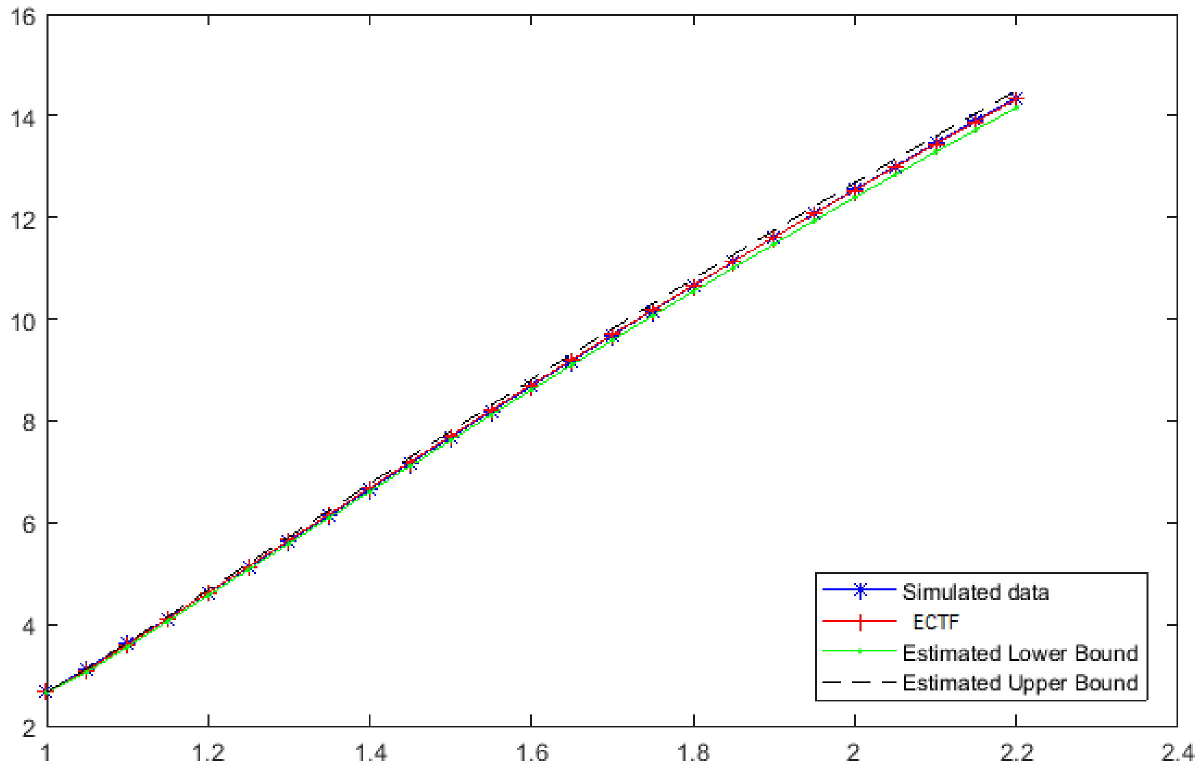


Figure 4.3: Simulated data, ECTF, ECCB and the predicted values.

4.5 Application to real data

In this section, we present an application of the proposed process and the computational statistical methodology described above to fits and forecast the global CO_2 emissions in Morocco using the ETF and ECTF. The variable $x(t)$ represents CO_2 total emissions from the Consumption of Energy and $t_i = t_{i-1} + (i-1)h$; for $i = 2, \cdots, N$ starting at $t_1 = 1$, taking the step size $h = 0.01$, $N = 32$ and $x_1 = 19.2$.

The real data are annual values and correspond to the period 1987-2018. CO_2 emissions are expressed in Million Metric Tons of Carbon Dioxide. These data are published by Atlas Mondial de Données (Maroc-Environnement) at https://knoema.fr/atlas/Maroc/%C3%89mission-de-CO2-kt.

We used the series of observations considered from 1987 to 2018 to estimate the parameters of the model by SA method. The values of the estimators of the parameters are

$$\hat{a} = 0.258755, \quad \hat{b} = 2.214039 \times 10^{-19} \quad \text{and} \quad \hat{\sigma} = 0.099976.$$

For the years 2019 and 2020, we predict the corresponding values using the ETF and ECTF. Also, we give the results attached to a 95% ECB and a 95% ECCB of the process. Finally, to illustrate the performance of the procedure, we calculate the one-step-ahead MAE, the RMSE, and the MAPE. Then, the values of the MAE and RMSE are respectively 0.95229 and 1.28012. The MAPE is 2.48211%, so we conclude the forecast is highly accurate.

Table 4.5 shows the results of the observed values, the ETF, ECTF, ECB and ECCB.

Fig. 4.4 and Fig. 4.5 illustrate the fits and forecast of the SMLKDP using the ETF and ECTF.

Table 4.5: Observed data, fits and forecast using the ETF and ECTF.

Year	CO_2 emissions	ETF	ECTF	ECB	ECCB
1987	19.2	19.2000	19.2000	[19.2000 19.2000]	[19.2000 19.2000]
1988	20.2	20.3118	20.3118	[19.9166 20.7127]	[19.9166 20.7127]
1989	21.8	21.4590	21.3409	[20.8704 22.0598]	[20.9258 21.7621]
1990	22.4	22.6415	23.0012	[21.8827 23.4197]	[22.5537 23.4552]
1991	23.5	23.8588	23.6043	[22.9373 24.8075]	[23.1451 24.0702]
1992	25.2	25.1105	24.7329	[24.0280 26.2287]	[24.2517 25.2210]
1993	25.9	26.3963	26.4903	[25.1517 27.6859]	[25.9749 27.0131]
1994	28.1	27.7155	27.1944	[26.3060 29.1801]	[26.6654 27.7312]
1995	29.2	29.0678	29.4709	[27.4895 30.7120]	[28.8976 30.0526]
1996	28.9	30.4524	30.5909	[28.7009 32.2818]	[29.9958 31.1947]
1997	30.4	31.8689	30.2442	[29.9391 33.8892]	[29.6559 30.8412]
1998	31.1	33.3166	31.7809	[31.2030 35.5342]	[31.1626 32.4082]
1999	32.7	34.7948	32.4798	[32.4918 37.2163]	[31.8479 33.1209]
2000	33.4	36.3028	34.1172	[33.8044 38.9351]	[33.4534 34.7905]
2001	36.9	37.8398	34.8141	[35.1401 40.6900]	[34.1369 35.5013]
2002	38.1	39.4052	38.4265	[36.4980 42.4804]	[37.6789 39.1849]
2003	37.9	40.9982	39.6402	[37.8771 44.3056]	[38.8690 40.4226]
2004	41.8	42.6180	39.3973	[39.2767 46.1649]	[38.6309 40.1749]
2005	44.4	44.2637	43.4141	[40.6959 48.0576]	[42.5695 44.2710]
2006	45.5	45.9345	46.0759	[42.1339 49.9829]	[45.1796 46.9854]
2007	46.1	47.6296	47.1790	[43.5897 51.9400]	[46.2612 48.1103]
2008	49.0	49.3481	47.7633	[45.0626 53.9279]	[46.8341 48.7060]
2009	48.9	51.0893	50.7288	[46.5517 55.9459]	[49.7419 51.7300]
2010	52.0	52.8521	50.5872	[48.0562 57.9931]	[49.6031 51.5857]
2011	56.2	54.6357	53.7548	[49.5752 60.0685]	[52.7091 54.8158]
2012	58.5	56.4392	58.0551	[51.1078 62.1712]	[56.9258 59.2010]
2013	58.1	58.2618	60.3891	[52.6534 64.3003]	[59.2143 61.5810]
2014	59.7	60.1026	59.9356	[54.2109 66.4549]	[58.7696 61.1186]
2015	61.5	61.9606	61.5455	[55.7797 68.6341]	[60.3483 62.7603]
2016	63.6	63.8351	63.3605	[57.3590 70.8368]	[62.1279 64.6111]
2017	66.6	65.7250	65.4830	[58.9478 73.0621]	[64.2091 66.7755]
2018	68.3	67.62969	68.5299	[60.5455 75.3090]	[67.1968 69.8826]
Prediction					
2019	-	69.5480	70.2374	[62.1512 77.5767]	[68.8710 71.6237]
2020	-	71.4793	72.1878	[63.7642 879.864]	[69.4065 75.0655]

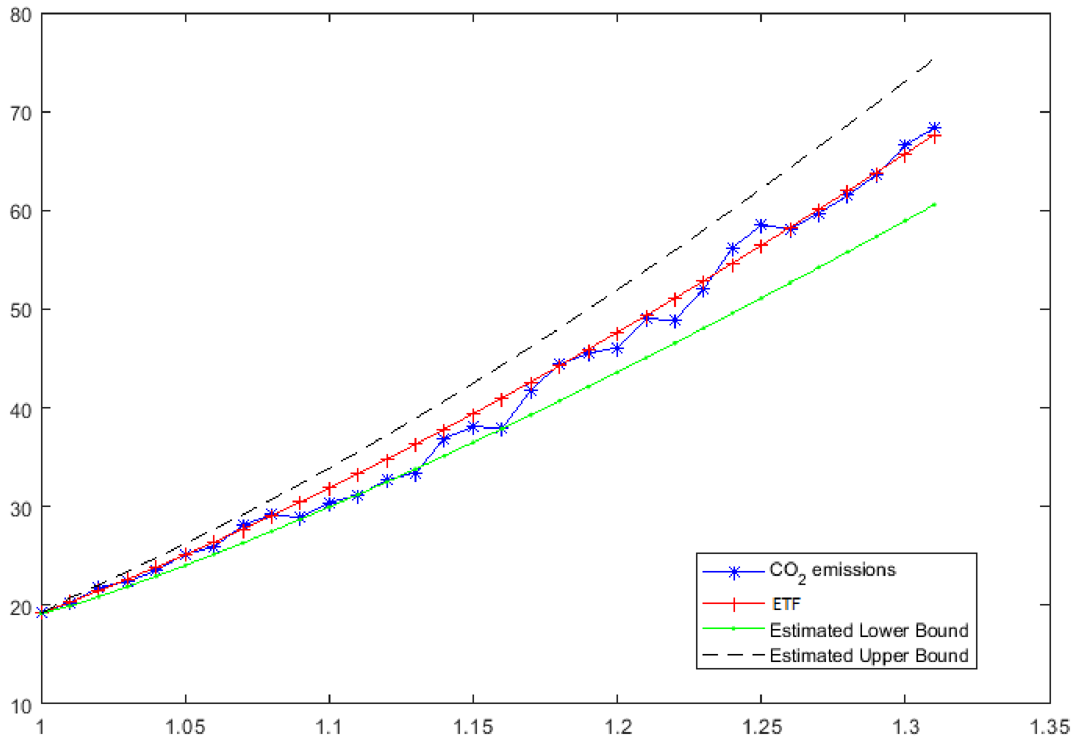


Figure 4.4: Fits and forecast using the ETF and ECB.

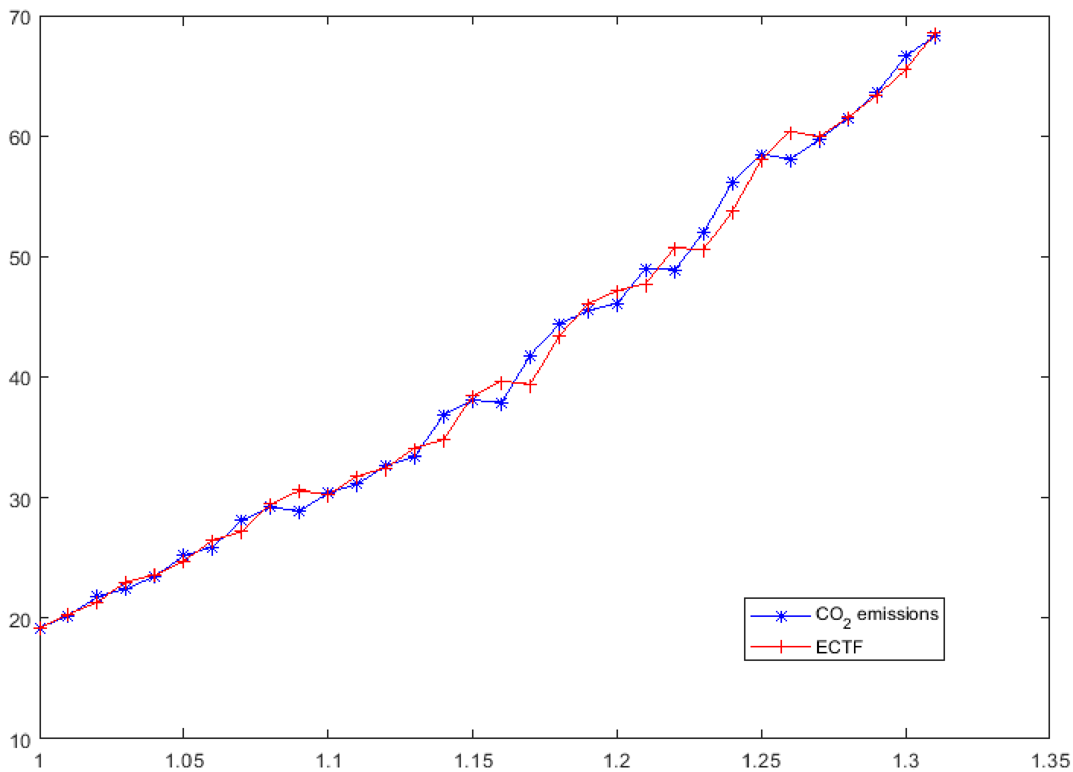


Figure 4.5: Fits and forecast using the ECTF.

4.6 Conclusions

In this chapter, we defined a new non-homogeneous SDP related to the modified Lundqvist-Korf growth curve. Then, we analyzed its distribution and main characteristics as its trend function and its conditional trend function, which were found to be proportional to the modified Lundqvist-Korf growth curve. This process is advantageous over the deterministic modified Lundqvist-Korf growth curve.

In this study, we developed the theoretical and the practical aspects of modified Lundqvist-Korf-type diffusion processes as a particular case of the stochastic log-normal diffusion process. Then, we applied the simulated annealing algorithm to solve inference problems.

Finally, an application to simulated data of the proposed model showed its usefulness in practice and demonstrated that the strategy used for bounding the parametric space behaves well. Finally, we applied the process to study the total emission of CO_2 in Morocco. By fitting the SMLKDP to the real data from the period corresponding to 1987 to 2018, we obtained a good description of the series and good short-medium term forecasts 2019–2020.

The description and forecast using the conditioned trend function are considerably better than those based on the trend function alone, although they are only optima in the short term.

CHAPTER 5

New stochastic logistic and log-logistic diffusion processes: Simulation and application

Contents

This chapter is the complete version of the paper [97] preprint submitted to Stochastic Models (Under Revision) and the paper [98] preprint submitted to Stochastics: An International Journal of Probability and Stochastic Processes.

5.1 Introduction

In 1838, Pierre François Verhulst [99, 100] proposed a particular mathematical formulation for studying population growth. This formulation is the logistic curve. It describes an asymptotic sigmoidal time response and has been widely used in ecological investigations (we recommend to see, for example, Solomon [101], May [102], Tang and Chen [103] and Xu et al. [104]), the growth of animals (see, Bordy [105]), the growth of plants (see, Charles-Edwards [106]), and elsewhere.

The stochastic diffusion processes plays an outstanding role in stochastic modelling, these are used in many disciplines, such as in physics, biological phenomena, tumor Growth, economy and finance, life expectancy at birth, energy consumption, and others (see, for example, Gutiérrez et al. [17, 52, 71], Albano and Giorno [107] and Román-Román et al. [3]). Stochastic process related to a reformulation of the logistic growth models have been applied to a variety of scientific areas (see, for example, Capocelli and Ricciardi [38], Román-Román et al. [3] and Giovanis and Skaidas [2]).

In literature, there are many works of stochastic diffusion process related to logistic growth models, for example, Capocelli and Ricciardi [38] derive a diffusion process from a reparameterization of the logistic growth model. Giovanis and Skaidas [2] proposed a stochastic version of the well-known logistic model is solved analytically using the theory of reducible stochastic differential equation (SDE) and applied this model to studying the electricity consumption in Greece and the United States. Heydari et al. [39] introduced two new first order linear noise approximations of stochastic logistic diffusion process, one with multiplicative and one with additive intrinsic noise. Tang and Heron [40] have used Markov chain Monte Carlo techniques to carry out Bayesian inference for piecewise stochastic logistic growth models using discretely observed data sets, which allows us to fit models for time series data, including data on fish productions and yields, with structural changes.

The log-logistic curve belongs to the important class of smooth sigmoidal functions arising from dose-response studies and cell growth models. It was initially developed by Jerne and Wood [108] and Berkso [109], whose applied it to pharmaceutics. The log-logistic curve is an appropriate method for analyzing most dose-response studies. This curve has been used widely and successfully in weed science and possesses several clear advantages over other analysis methods and the authors suggest that it should be widely adopted as a standard herbicide dose-response analysis method. For a more statistically oriented coverage of the log-logistic and other models, we refer to see, Streibig and Kudsk [110] and Seefeldt et al. [111].

The estimation of the parameters in the diffusion processes have received considerable attention in recent years, both when the process is observed continuously and when discretely. Such estimation, in general, is not direct, except in simple cases and one possible methodological approach is based on approximating the maximum likelihood function. Various methods addressing this question have been developed, and many papers have been published on this subject, focusing on several variant of approximate likelihood methodology, the general case of this methodology can be consulted in [41–43] and others. Moreover, we must not forget the other estimation methods, such as method of moments (see, Chan et al. [50]), the non-parametric method (see, Arapis and Gao. [51] and Jiang and Knight [52]) and Bayesian methods (see, for example, [39, 40]). Also, we find various authors approached the problem of Maximum likelihood (ML) estimation by equating partial derivatives of the log-likelihood function (LLF) to zero, looking for a stationary point of the local ML equations by iterative methods (see, for example, Wilson and Worcester [45], Cohen [46], Lambert [47], Harter and Moore [48] and Calitz [49]). Fernando et al. [56] proposed the global simulated annealing (SA) heuristic for the three-parameter log-normal ML estimation. Pedersen [57] suggested a new approach to ML estimation for SDEs based on discrete observations when the likelihood function (LF) is unknown.

The approach of ML estimation of the parameters using likelihood equations can be problematic, which is why we propose the use of the SA algorithm. This algorithm is a method for solving unconstrained and bound-constrained optimization problems developed by Kirkpatrick et al. [58]. The method models the physical process of heating a material and then slowly lowering the temperature to decrease defects, thus minimizing the system energy. It has been successfully used for optimization in continuous spaces (see, Duflo [81]). In the last years many works have used the SA algorithm (see, for example, Nafidi et al. [28, 54] and Román-Román et al. [112]).

In this chapter, we propose a new stochastic logistic diffusion process related to a reformulation of the logistic curve (SLDP), which presents a mean function that is the logistic growth curve and we introduce a stochastic process whose mean function (MF) is a log-logistic growth curve and to compare the prediction and forecasting using this process with that the logistic from two examples: the first over simulated data, and the second over real data. The rest of this paper is organized as follows: In section 5.2, we present a reformulation of the logistic curve and we define the proposed model in terms of SDE. We then determine the explicit expression of the solution to the SDEs, the transition probability density function (TPDF) and the MFs, and we discuss the parameter estimation using the ML method on the basis of discrete sampling of the process. Since the closed form of the ML estimators cannot be given because the system of likelihood equations does not have an explicit form, numerical methods are needed. In fact, the SA algorithm is proposed. A summary of the algorithm and its properties is then shown, as well as the modifications required in our context. In this sense, a procedure for bounding the parametric space is proposed. The results obtained are presented from the examples of simulation and we illustrate the predictive study by fitting the diffusion process to study the growth of a microorganism culture. In section 5.3, we define a new stochastic log-logistic diffusion process (SLLDP) in terms of SDE. We then determine the explicit expression of the solution to the SDEs, the TPDF and the MFs, and we discuss the parameters estimation using the ML method on the basis of discrete sampling of the process. Since the closed form of the ML estimators cannot be given because the system of likelihood equations does not have an explicit form, numerical methods are needed. In fact, the SA algorithm is proposed. A summary of the algorithm and its properties is then shown, as well as the modifications required in our context. In this sense, a procedure for bounding the parametric space is proposed. Finally, we present the results obtained from the simulations, comparing the parameters estimation of this process with that the SLDP using an example based on the simulated data. We then illustrate the predictive study by fitting the two diffusion processes to study the growth of a microorganism culture. In the last section, we summarize the main conclusions.

5.2 The new SLDP

5.2.1 The proposed model and its characteristics

5.2.1.1 A reformulation of the logistic curve

As is well known that the logistic model describes the self-limiting growth of a population of size $x(t)$ at time t, this model is formalized by the ODE

$$\frac{dx(t)}{dt} = \beta x(t)\left(1 - \frac{x(t)}{\gamma}\right); \quad t \geq t_1, \tag{5.2.1}$$

where the parameters β is the growth rate parameter and γ is the carrying capacity. The ODE Eq. (5.2.1) has the following analytic solution

$$x(t) = \frac{\gamma}{1+\alpha e^{-\beta t}}; \quad t \geq t_1,$$

where α is a positive constant. At $t = t_1$, this curve starts out at $x(t_1) = \dfrac{\gamma}{1+\alpha e^{-\beta t_1}}$ and as $t \to \infty$, $x(t) \to \gamma$ i.e. the line $y = \gamma$ is the horizontal asymptote of the logistic growth curve. We impose that $x(t_1) = x_1 > 0$, hence $\gamma = x_1\left(1+\alpha e^{-\beta t_1}\right)$. Thus, we reach

$$x(t) = x_1\left(\frac{1+\alpha e^{-\beta t_1}}{1+\alpha e^{-\beta t}}\right); \quad t \geq t_1,\ \alpha > 0,\ \beta > 0. \tag{5.2.2}$$

Note that $\lim_{t\to\infty} x(t) = x_1\left(1+\alpha e^{-\beta t_1}\right)$, so the line $y = x_1\left(1+\alpha e^{-\beta t_1}\right)$ is a horizontal asymptote of the curve (5.2.2) when t tend to ∞. Finally, by setting $a = \dfrac{1}{\alpha}$ and $b = e^{-\beta}$, we find the new reformulation of the logistic growth curve

$$x(t) = x_1 \left(\frac{a + \xi(t_1)}{a + \xi(t)} \right), \tag{5.2.3}$$

where $\xi(t) = b^t$. This reformulation leads to condition $0 < b < 1$. The curve (5.2.3) shows an inflection point at $t_I = \dfrac{\ln(a)}{\ln(b)}$. In addition, this point verify $t_I > t_1$. Furthermore, $t_I > t_1$ if and only if $a < \xi(t_1)$.

5.2.1.2 The new SLDP

In this section, we will introduce a diffusion process associated to a reformulation of the curve (5.2.3). Now, starting from the curve (5.2.3), one is lead to considering the ODE

$$\frac{dx(t)}{dt} = \varphi(t)x(t); \quad x(t_1) = x_1, \tag{5.2.4}$$

where $\varphi(t) = \dfrac{-d\xi(t)/dt}{a + \xi(t)}$. Hence Eq. (5.2.4) can be viewed as a generalisation of the Malthusian growth model with time dependent fertility rate $\varphi(t)$.

A new stochastic version of the SLDP is given by the diffusion process $\{x(t) : t \geq t_1\}$, taking values on $(0, \infty)$ and characterized by infinitesimal moments

$$\begin{aligned} A_1(x,t) &= \varphi(t)x, \\ A_2(x,t) &= \sigma^2 x^2, \end{aligned} \tag{5.2.5}$$

with $\sigma > 0$ and initial distribution $x(t_1)$. This is a solution of the Itô's SDE

$$dx(t) = A_1(x,t)dt + (A_2(x,t))^{\frac{1}{2}} dw(t); \quad x(t_1) = x_1, \tag{5.2.6}$$

where $w(t)$ is a standard Wiener process and x_1 is a positive random variable, independent on $w(t)$ for $t \geq t_1$. The infinitesimal moments $A_1(x,t)$ and $A_2(x,t)$ specified in Eq. (5.2.5) satisfy the Lipschitz and the growth conditions for the existence and unicity of the solution to the SDE Eq. (5.2.6) (see, Kloeden and Platen [8]). In fact, there exists a non negative constant $C = \dfrac{-\ln(b)\xi(t_1)}{a + \xi(t_1)} + \sigma$, such that for all $x, y \in \mathbb{R}^+$ and $t \geq t_1$ we have:

$$\begin{aligned} &|A_1(t,x) - A_1(t,y)| + \left|(A_2(t,x))^{\frac{1}{2}} - (A_2(t,y))^{\frac{1}{2}}\right| \leq C \mid x - y \mid, \\ &|A_1(t,x)|^2 + \left|(A_2(t,x))^{\frac{1}{2}}\right|^2 \leq C^2 \left(1 + \mid x \mid^2\right). \end{aligned}$$

Then, the SDE Eq. (5.2.6) has a unique solution $\{x(t) : t \geq t_1\}$ continuous with probability 1, and satisfies the initial condition $x(t_1) = x_1$.

5.2.1.3 Basic probabilistic characteristics

5.2.1.3.1 Probability distribution of the SLDP

By means of the appropriate transformation of the form $y(t) = \ln(x(t))$, and by using the Itô's rule, the SDE Eq. (5.2.6) becomes

$$dy(t) = \left(\varphi(t) - \frac{\sigma^2}{2}\right) dt + \sigma dw(t); \quad y_1 = \ln(x_1),$$

by integrating both sides yields,

$$y(t) = y_1 + \ln\left(\frac{a + \xi(t_1)}{a + \xi(t)}\right) - \frac{\sigma^2}{2}(t - t_1) + \sigma(w(t) - w(t_1)).$$

Finally, we have:

$$x(t) = x_1 \left(\frac{a + \xi(t_1)}{a + \xi(t)}\right) \exp\left[-\frac{\sigma^2}{2}(t - t_1) + \sigma(w(t) - w(t_1))\right]. \quad (5.2.7)$$

The $y(t)$ is a gaussian process if and only if $\ln(x_1)$ is constant or normally distributed (see, Arnold [82]). In such a case, the $x(t)$ is a log-normal process. That is, the TPDF

$$f(x, t|y, s) = \frac{1}{x\sqrt{2\pi\sigma^2(t - s)}} \exp\left(-\frac{[\ln\left(\frac{x}{y}\right) - \ln\left(\frac{a+\xi(s)}{a+\xi(t)}\right) + \frac{\sigma^2}{2}(t - s)]^2}{2\sigma^2(t - s)}\right). \quad (5.2.8)$$

5.2.1.3.2 The MFs

By using the properties of the log-normal distribution, the r-th conditional moment of the process is

$$\begin{aligned} E(x^r(t)|x(s) = x_s) &= \exp\left(\ln(x_s^r) + r\ln\left(\frac{a + \xi(s)}{a + \xi(t)}\right) - \frac{r\sigma^2}{2}(t - s) + \frac{r^2\sigma^2}{2}(t - s)\right) \\ &= x_s^r \left(\frac{a + \xi(s)}{a + \xi(t)}\right)^r \times \exp\left(\frac{r(r-1)\sigma^2}{2}(t - s)\right). \end{aligned}$$

Then, for $r = 1$, the conditional MF (CMF) of the process is

$$E(x(t)|x(s) = x_s) = x_s \left(\frac{a + \xi(s)}{a + \xi(t)}\right). \quad (5.2.9)$$

In addition, taking into the initial condition $\mathrm{P}(x(t_1) = x_1) = 1$, the MF of the process is given by

$$E(x(t)) = x_1 \left(\frac{a + \xi(t_1)}{a + \xi(t)}\right). \quad (5.2.10)$$

5.2.2 Statistical 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 Zehna's theorem (see, Zehna [68]), the corresponding MFs.

5.2.2.1 Estimation of the parameters

We consider a discrete sampling of the process, that is, for fixed times $t_1, t_2, \cdots, t_n$, $(n > 2)$, we observe the variables $x(t_1)$, $x(t_2)$, $\cdots$, $x(t_n)$ whose values provide the basic sample for the inference process. In addition, we assume $t_i - t_{i-1} = h$, for $i = 2, \cdots, n$. Let x_1, x_2, $\cdots$, x_n be the observed values of the sampling. The LF depends on the choice of the initial distribution.

If $\mathrm{P}(x(t_1) = x_1) = 1$, the associated LF is

$$\mathbf{L} = L_{x_1,x_2,\cdots,x_n}\left(a, b, \sigma^2\right) = \prod_{i=2}^{n} f\left(x_i, t_i | x_{i-1}, t_{i-1}\right),$$

which is written as

$$\mathbf{L} = \prod_{i=2}^{n} \frac{1}{x_i\sqrt{2\pi\sigma^2 h}} \exp\left(-\frac{\left\{\ln\left(\frac{x_i}{x_{i-1}}\right) - \ln\left(\frac{a+\xi(t_{i-1})}{a+\xi(t_i)}\right) + \frac{\sigma^2}{2}h\right\}^2}{2\sigma^2 h}\right).$$

where a, b and σ^2 are the parameters to be estimated.

If $x(t_1)$ is distributed as one-dimensional log-normal distribution $\Lambda_1\left(\mu_1, \sigma_1^2\right)$, the associated LF is

$$\mathbf{L} = L_{x_1,x_2,\cdots,x_n}\left(\mu_1, \sigma_1^2, a, b, \sigma^2\right) = f_{x(t_1)}\left(x_1\right) \prod_{i=2}^{n} f\left(x_i, t_i | x_{i-1}, t_{i-1}\right),$$

which is written as

$$\begin{aligned} \mathbf{L} = & \frac{1}{x_1\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{\left(\ln\left(x_1\right) - \mu_1\right)^2}{2\sigma_1^2}\right) \\ & \times \prod_{i=2}^{n} \frac{1}{x_i\sqrt{2\pi\sigma^2 h}} \exp\left(-\frac{\left\{\ln\left(\frac{x_i}{x_{i-1}}\right) - \ln\left(\frac{a+\xi(t_{i-1})}{a+\xi(t_i)}\right) + \frac{\sigma^2}{2}h\right\}^2}{2\sigma^2 h}\right). \end{aligned} \tag{5.2.11}$$

In the following, we will consider the case when the initial distribution is log-normal. From Eq. (5.2.11), the LLF of the sample is

$$\begin{aligned} \ln(\mathbf{L}) = & -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\sigma_1^2\right) - \frac{(n-1)}{2}\ln\left(\sigma^2\right) - \sum_{i=1}^{n}\ln\left(x_i\right) - \frac{1}{2\sigma_1^2}\left(\ln\left(x_1\right) - \mu_1\right)^2 \\ & - \frac{(n-1)}{2}\ln(h) - \frac{1}{2\sigma^2 h}\sum_{i=2}^{n}\left\{\ln\left(\frac{x_i}{x_{i-1}}\right) - \ln\left(\frac{a+\xi\left(t_{i-1}\right)}{a+\xi\left(t_i\right)}\right) + \frac{\sigma^2}{2}h\right\}^2. \end{aligned} \tag{5.2.12}$$

Deriving the LLF with respect to μ_1 and σ_1^2 we obtain

$$\frac{\partial \ln(\mathbf{L})}{\partial \mu_1} = \frac{1}{\sigma_1^2}\left(\ln\left(x_1\right)-\mu_1\right); \qquad \frac{\partial \ln(\mathbf{L})}{\partial \sigma_1^2} = -\frac{1}{2\sigma_1^2}+\frac{1}{2\sigma_1^4}\left(\ln\left(x_1\right)-\mu_1\right)^2. \quad (5.2.13)$$

Making the derivatives Eqs. (5.2.13) equal to zero, we obtain the following set of equations

$$\ln\left(x_1\right)-\mu_1=0; \qquad \sigma_1^2-\left(\ln\left(x_1\right)-\mu_1\right)^2=0. \quad (5.2.14)$$

Firstly, from Eqs. (5.2.14), the ML estimates of μ_1 and σ_1^2 are

$$\hat{\mu}_1=\ln\left(x_1\right) \quad \text{and} \quad \hat{\sigma}_1^2=\left(\ln\left(x_1\right)-\hat{\mu}_1\right)^2.$$

However, estimating a, b and σ^2 poses some difficulties. In fact, the resulting system of equations is exceedingly complex and does not have an explicit solution, and numerical procedures must be employed. In this sense, we propose the use of the metaheuristic simulated annealing algorithm in order to maximise the likelihood function. This algorithm designed to solve problems of the type $\min_{\theta\in\Theta} g(\theta)$, g being the objective function to be optimized, and are often more appropriate than classical numerical methods since they impose fewer restrictions on the space of solutions Θ and on the analytical properties of g. In our case, the problem becomes maximizing function $\ln L_{x_1,x_2,\cdots,x_n}\left(a,b,\sigma^2\right)$. Since the algorithm aforementioned are usually formulated for minimization problems, from Eq. (5.2.12), the objective function is then a function of parameters a, b and σ^2 is consider as follows

$$g\left(a,b,\sigma^2\right)=\frac{(n-1)}{2}\ln\left(\sigma^2\right)+\frac{1}{2\sigma^2 h}\sum_{i=2}^{n}\left\{\ln\left(\frac{x_i}{x_{i-1}}\right)-\ln\left(\frac{a+\xi\left(t_{i-1}\right)}{a+\xi\left(t_i\right)}\right)+\frac{\sigma^2}{2}h\right\}^2. \quad (5.2.15)$$

5.2.2.2 Estimated MFs

By using Zehna's theorem given by Zehna [68], the estimated mean functions (MFs) of the process is obtained by replacing the parameters by replacing the parameters in Eq. (5.2.9) and Eq. (5.2.10) by their estimators. Then the estimated conditional MF (ECMF) is given by the following expression

$$\hat{E}\left(x(t)|x(s)=x_s\right)=x_s\left(\frac{\hat{a}+\hat{\xi}(s)}{\hat{a}+\hat{\xi}(t)}\right). \quad (5.2.16)$$

Taking into account the initial condition that is $\mathrm{P}\left(x\left(t_1\right)=x_1\right)=1$, the estimated MF (EMF) of the process is given by

$$\hat{E}(x(t))=x_1\left(\frac{\hat{a}+\hat{\xi}\left(t_1\right)}{\hat{a}+\hat{\xi}(t)}\right). \quad (5.2.17)$$

5.2.2.3 Confidence bounds

The confidence bounds (CB) of the process are obtained by using the procedure described by Katsamaki and Skiadas [67]. Let $v(s,t) = x(t)|x(s) = x_s$. Since the variable $w(t) - w(s)$ is Gaussian with the mean equal to zero and the variance $t - s$ for $t \geq s$. Therefore, the random variable z is given by

$$z = \frac{\ln(v(s,t)) - \mu\left(s,t,x_s\right)}{\sigma\sqrt{(t-s)}} \sim N_1(0,1),$$

where $\mu\left(s,t,x_s\right) = \ln\left(x_s\right) + \ln\left(\frac{a+\xi(s)}{a+\xi(t)}\right) - \frac{\sigma^2}{2}\left(t-s\right)$.

A $100(1-\kappa)\%$ conditional CB for z is given by $P(-\lambda \leq z \leq \lambda) = 1 - \kappa$. From this, we can obtain a CB of $v(s,t)$ with following form $(v_{lower}(s,t), v_{upper}(s,t))$, where

$$v_{lower}(s,t) = \exp\left(\mu\left(s,t,x_s\right) - \lambda\sigma\sqrt{(t-s)}\right), \tag{5.2.18}$$

and

$$v_{upper}(s,t) = \exp\left(\mu\left(s,t,x_s\right) + \lambda\sigma\sqrt{(t-s)}\right), \tag{5.2.19}$$

with $\lambda = F^{-1}_{N_1(0,1)}\left(1-\frac{\kappa}{2}\right)$ and where $F^{-1}_{N_1(0,1)}$ is the inverse cumulative normal standard distribution and $\mu\left(s,t,x_s\right) = \ln\left(x_s\right) + \ln\left(\frac{a+\xi(s)}{a+\xi(t)}\right) - \frac{\sigma^2}{2}\left(t-s\right)$.

On the other hand, the estimated lower bound $\hat{v}_{lower}(t)$ and an estimated upper bound $\hat{v}_{upper}(t)$ can be obtained by substituting the parameters by theirs estimators in the equations Eq. (5.2.18) and Eq. (5.2.19), the estimated CB are given by:

$$\hat{v}_{lower}(s,t) = \exp\left(\hat{\mu}\left(s,t,x_s\right) - \lambda\hat{\sigma}\sqrt{(t-s)}\right), \tag{5.2.20}$$

and

$$\hat{v}_{upper}(s,t) = \exp\left(\hat{\mu}\left(s,t,x_s\right) + \lambda\hat{\sigma}\sqrt{(t-s)}\right), \tag{5.2.21}$$

where $\hat{\mu}\left(s,t,x_s\right) = \ln\left(x_s\right) + \ln\left(\frac{\hat{a}+\hat{\xi}(s)}{\hat{a}+\hat{\xi}(t)}\right) - \frac{\hat{\sigma}^2}{2}\left(t-s\right)$.

5.2.2.4 Application of the SA algorithm

5.2.2.4.1 The algorithm

The SA algorithm is a metaheuristic algorithm for problems of the type $\min_{\theta\in\Theta} g(\theta)$, where Θ is the solution space and g is the objective function.

In our case,the parametric space Θ linked to the objective function (5.2.15) ,on which the selected algorithms must operate,is continuous and unbounded. Consequently

$$\Theta = \{(a, b, \sigma) : \quad a > 0,\ 0 < b < 1,\ \sigma > 0\}.$$

The drawback is that the solution space might not be explored with enough depth. This requires us to find arguments for bounding said space. The following are some strategies to this end.

5.2.2.4.2 Bounding the solution space

Regarding parameter σ, when it has high values it leads to sample paths with great variability around the mean of the process. Thus, excessive variability in available paths would make an logistic-type modeling inadvisable (see, Figure 5.1). Some simulations performed for several values of σ have led us to consider that $0 < \sigma < 0.1$, so that we may have paths compatible with an logistic-type growth curve. On the other hand, there does not seem to be an upper bound for a. Nevertheless, the logistic curve is sigmoidal and has an inflection point at $t_I = \dfrac{\ln(a)}{\ln(b)}$, which is higher than t_1 if and only if $a < b^{t_1}$. So the parameter a is bounded $0 < a < b^{t_1} < 1$. Thus, the solution space, which is obtained numerically for a, b and σ is bounded and takes the form $(0,1) \times (0,1) \times (0, 0.1)$.

Once the solution space has been bounded, we specify the choice of the initial parameters of the algorithms and the stopping conditions. We consider the following:

1. The initial solution is chosen randomly in the bounded subspace

$$\Theta' = (0,1) \times (0,1) \times (0, 0.1) .$$

2. The initial temperature should be high enough such that in the first iteration of the algorithm, the probability of accepting a worse solution is at least of 80% (see, Kirkpatrick and al. [58]). For this we assume the initial temperature of 10.

3. For the cooling process we have considered a geometric scheme in which the current temperature is multiplied by a constant γ $(0 < \gamma < 1)$, i.e. $T_i = \gamma T_{i-1}$, $i \geq 1$. The usual values for γ are between 0.80 and 0.99. For this we have set $\gamma = 0.95$.

4. The length of each temperature level (L) determines the number of solutions generated at a certain temperature, T. For this we have set $L = 50$.

5. The stopping criterion defines when the system has reached a desired energy level (freezing temperature). Equivalently it defines the total number of solutions generated, or when an acceptance ratio (ratio between the number of solutions accepted and the number of solutions generated) is reached. The application of the algorithm will be limited to 1000 iterations.

The coding is performed using Matlab computer software.

5.2.3 Simulation study

This section will analyze the application of the SA algorithm to the obtainment of maximum likelihood estimations for parameters a, b and σ (functions of parameters α, β and σ) in a new stochastic logistic diffusion process with infinitesimal moments given by Eq. (5.2.5).

5.2.3.1 Simulated sample paths of the SLDP

This section will present some simulated sample paths of the SLDP. The trajectory of the model is obtained by simulating the exact solution of the SDE Eq. (5.2.7). We obtain the simulated trajectories of the model by considering the equally spaced time discretization of the interval $[t_1, T]$, with time points $t_i = t_{i-1}+(i-1)h$; for $i = 2, \cdots, N$ and the discretization step size $h = \dfrac{T - t_1}{N}$ for the sample size N. The random variable $\sigma\left(w(t) - w\left(t_1\right)\right)$ in the Eq. (5.2.7) is distributed as one-dimensional normal distribution $N_1\left(0, \sigma^2(t-t_1)\right)$. Therefore, in this simulation, 50 sample paths are simulated with $t_1 = 0$, $T = 50$, $x_1 \sim \Lambda_1(1, 0.16)$ and 250 observations of the process. Figure 5.1 shows some simulated sample paths of the SLDP for several values of a, b and σ.

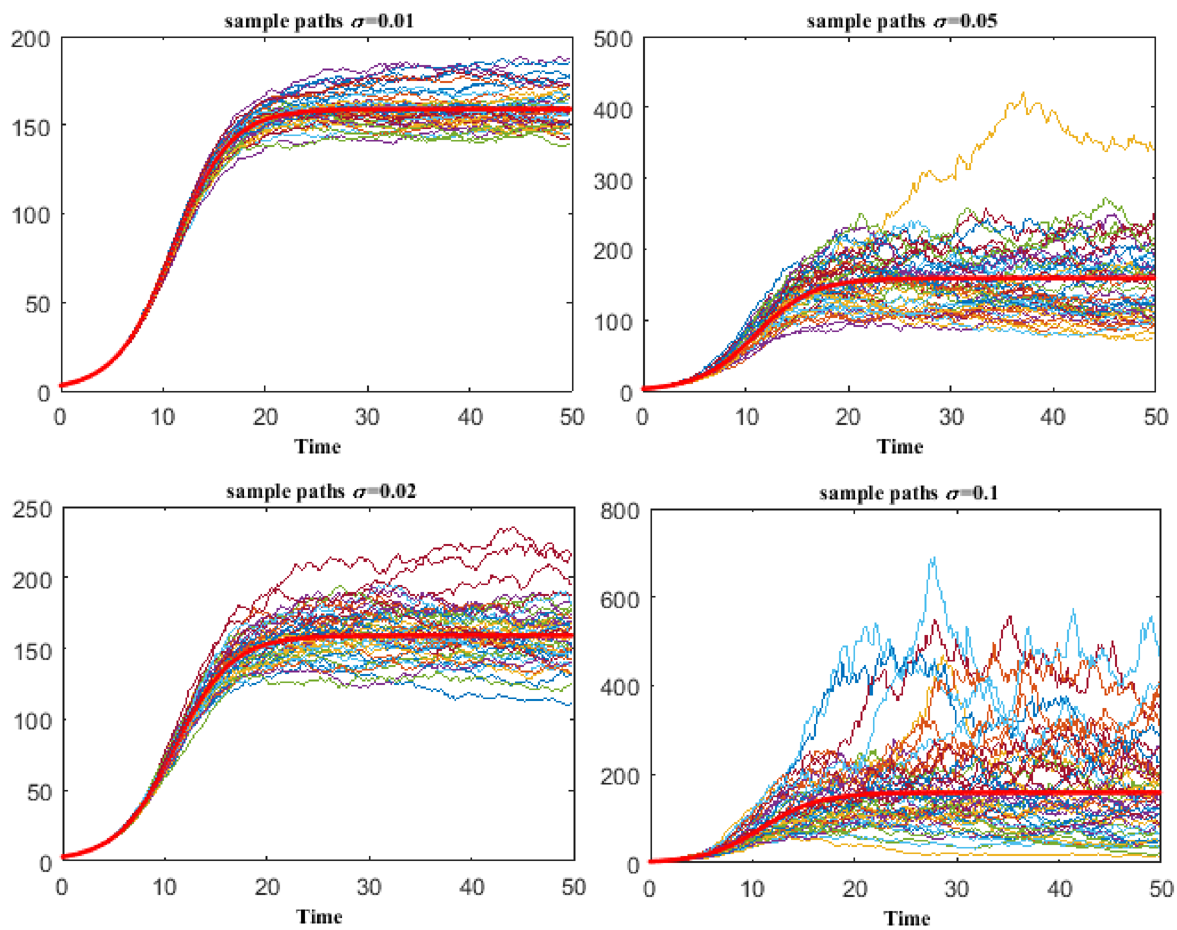


Figure 5.1: Simulated trajectories of the SLDP and its MF for several values of σ with ($a = 0.02, b = 0.7$).

5.2.3.2 Parameters estimation

This section will present several examples in order to validate the estimation procedure previously developed. To this end, Eq. (5.2.7) was simulated 25 times under the following assumptions, for each one, equally spaced time instants in the interval $[t_1, T]$ with step $h = 0.1$, starting from instant $t_1 = 0$ and $x_1 \sim \Lambda_1(1, 0.16)$. As for the sample size, values 50, 100, 250 and 500 have been considered for N.

In order to make the subsequent inference we have considered, in each case, 25 sample paths with $t_i = t_{i-1} + (i-1)h$; for $i = 2, \cdots, N$. The SA algorithm has been applied for estimating the parameters of the process with the specifications detailed in section 5.2.2.4.

The empirical mean, the std, and the CV for a, b and σ are defined in Table 3.1. Then Table 5.1 shows the results obtained for calculating the latter measures. The results obtained show the performance of the methodology.

Table 5.1: Estimation values, the std and the CV of a, b and σ for several values of σ.

$a = 0.02,\quad b = 0.7$									
$\sigma = 0.01$									
N	$\overline{a}$	$\overline{b}$	$\overline{\sigma}$	std(a) $.10^{-4}$	std(b) $.10^{-4}$	std(σ) $.10^{-4}$	CV(a) $.10^{-4}$	CV(b) $.10^{-4}$	CV(σ) $.10^{-4}$
50	0.016835	0.701094	0.009657	62.105	58.744	11.305	3688.9	83.789	1170.7
100	0.020283	0.700044	0.009792	11.106	38.196	6.5462	547.58	54.563	668.53
250	0.019902	0.700102	0.010065	8.9535	31.767	5.6171	449.87	45.376	558.08
500	0.020101	0.699292	0.009972	7.3911	27.878	2.2226	367.69	39.866	222.87
$\sigma = 0.02$									
50	0.027197	0.695666	0.019512	207.27	146.91	20.239	7621.02	211.18	1037.2
100	0.020328	0.700350	0.019893	292.94	77.411	14.094	1441.02	110.53	708.50
250	0.020243	0.700268	0.019634	16.049	59.968	8.2236	792.86	85.637	418.84
500	0.020034	0.699662	0.019898	12.851	57.377	7.6907	641.47	82.007	386.49
$\sigma = 0.05$									
50	0.035738	0.6821633	0.049341	382.17	284.92	52.750	10693.4	417.68	1069.1
100	0.019419	0.699931	0.049655	69.028	182.37	31.032	3554.6	260.55	624.96
250	0.021215	0.706099	0.049351	33.797	112.24	27.370	1593.1	158.96	554.60
500	0.021399	0.702548	0.049807	50.068	111.10	15.221	2339.7	158.14	305.61

5.2.3.3 Application to biological systems

The following example based on the studies developed by Román-Román et al. [3] on some aspects related to the growth in cultures of some microorganisms in the context of the logistic-type process. The growth of a culture for which it is known that the intrinsic growth rate is 0.25 per day and the equilibrium density is 1000 individuals per millilitre. There is a total of 50 containers, in which cultures are placed at the beginning of the study $t_1 = 0$, with a density of five individuals per millilitre. The experiment is then carried out for 50 days.

We will now proceed to linking the values specified for the experiment to the parameters of the model. The parameter $\beta = \ln\left(\frac{1}{b}\right)$, identifying the intrinsic growth rate, is 0.25 days^{-1} i.e.,

$$b = e^{-0.25}.$$

The initial distribution is degenerate for value $x_1 = 5$. The equilibrium density determines the limit value of the logistic curve, in this case $\gamma = 1000$. From these values, and taking into consideration that

$$a = \frac{x_1 b^{t_1}}{\gamma - x_1},$$

it is deduced that a equals $\frac{1}{199}$. Having 50 containers for the experiment implies simulating 50 paths for the process, taking place from $t_1 = 0$ to $T = 50$.

The simulation of the sample paths include 401 observations of the process starting from instant $t_1 = 0$, with $h = \frac{T - t_1}{400}$ and the stochastic variability term $\sigma = 0.01$.

The methodology can be summarised in the following steps: Firstly, we use the all data to estimate the parameters a, b and σ of the process by SA with the specifications detailed in section 5.2.2.4. Moreover, we obtain the corresponding EMF and ECMF values given by the expressions (5.2.9) and (5.2.10). Also, we give the results attached to a 95% estimated confidence bound (ECB) and a 95% estimated conditional confidence bound (ECCB) of the process in the expressions (5.2.20) and (5.2.21). Finally, to illustrate the performance of procedure, the results according to the MAE, the RMSE and the MAPE, given by Table 2.1. According to Lewis [69], we deduce that the accuracy of the forecast can be judged from the MAPE result Table 2.2.

Table 5.2 shows the values obtained from the estimation of the parameters of the process. Then the accuracy of the forecast can be judged from the MAPE, in other words if the value of the MAPE is less than 10%, the forecast is highly accurate for $\sigma = 0.01$. Table 5.3 shows the values obtained after the calculation of the three measures of the goodness of fit.

Figure 5.2 and Figure 5.3 illustrate the performance of the SLDP for forecasting using the mean function and the conditional mean function for $\sigma = 0.01$.

Table 5.2: Estimation of the parameters of the SLDP.

σ	$\hat{a}$	$\hat{b}$	$\hat{\sigma}$
0.01	0.004780	0.776117	0.009761

Table 5.3: Goodness of fit of the SLDP.

σ	MAE	RMSE	MAPE
0.01	8.690253	12.000116	1.523%

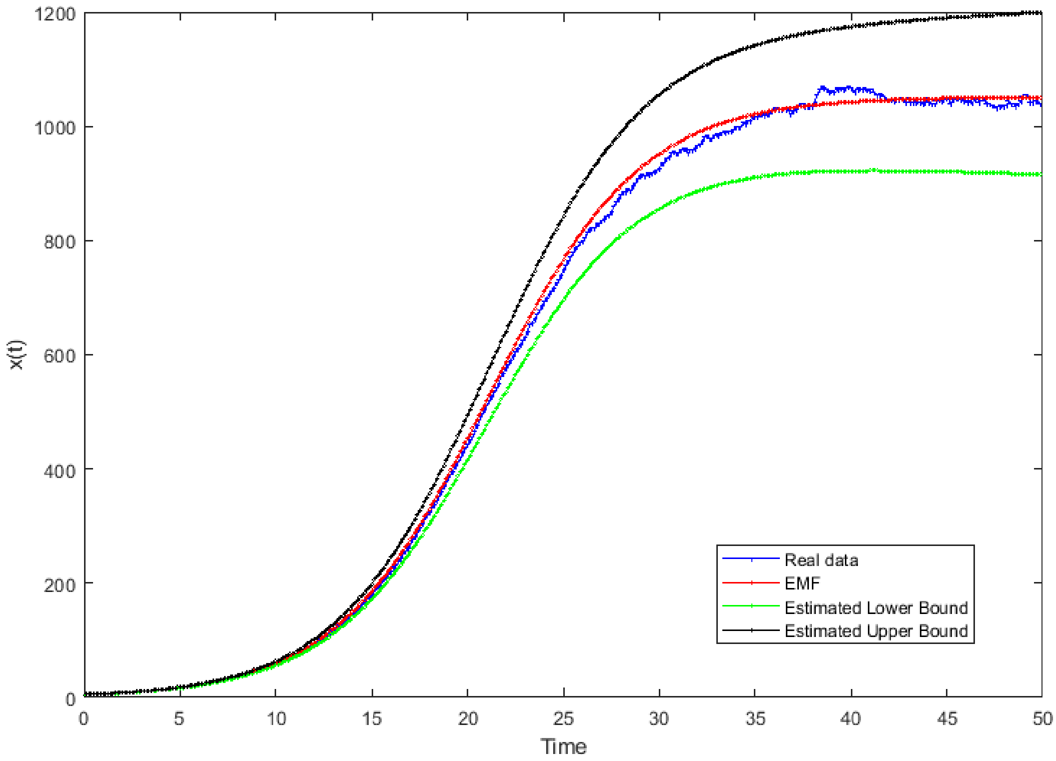


Figure 5.2: Real data versus EMF of the SLDP ($\sigma = 0.01$).

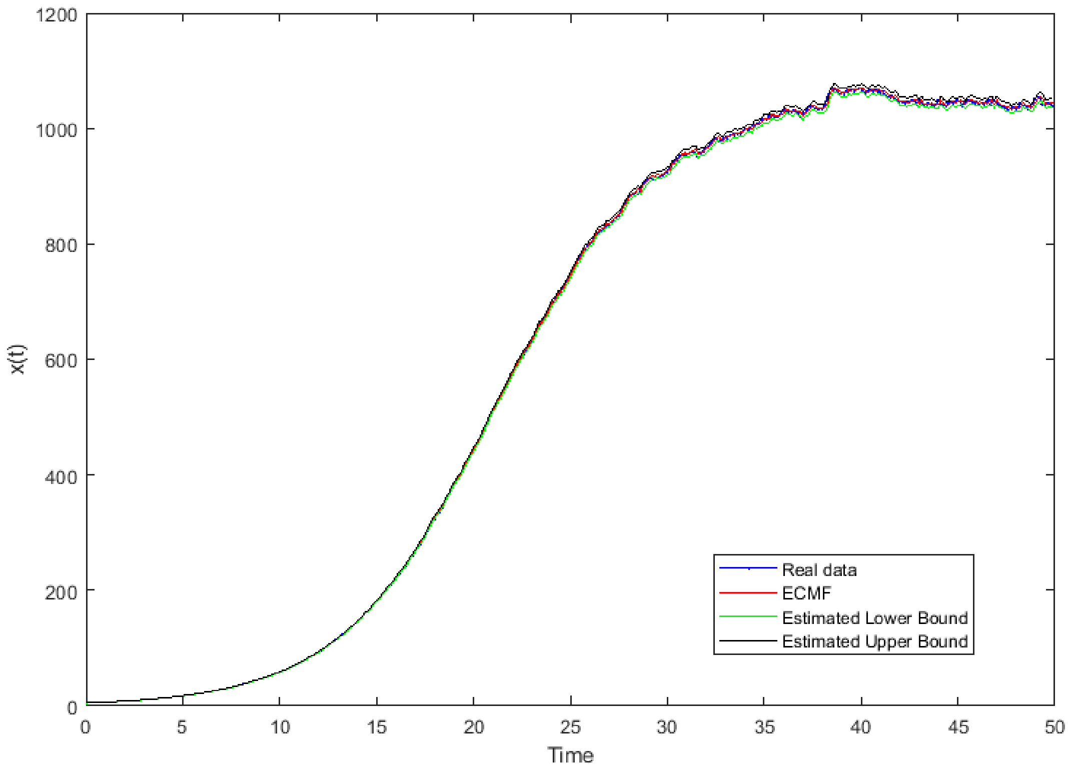


Figure 5.3: Real data versus ECMF of the SLDP ($\sigma = 0.01$).

5.3 The SLLDP

5.3.1 The model

This section will introduce a new diffusion process associated to a reformulation of the log-logistic curve Eq. (5.3.4). The most commonly used expression of the logistic curve is:

$$y(t) = \frac{\gamma}{1+\alpha e^{-\beta t}}; \quad t > 0. \tag{5.3.1}$$

If in the logistic curve Eq. (5.3.1), we replace t by $\ln(t)$, then we obtain the log-logistic curve

$$x(t) = \frac{\gamma}{1+\alpha e^{-\beta \ln(t)}}; \quad t > 0, \tag{5.3.2}$$

where the parameters α and β are growth rate parameters and γ are all positive. We impose that $x(t_1) = x_1 > 0$, hence $\gamma = x_1 \left(1+\alpha e^{-\beta \ln(t_1)}\right)$. Thus, we reach

$$x(t) = x_1 \left(\frac{1+\alpha e^{-\beta \ln(t_1)}}{1+\alpha e^{-\beta \ln(t)}}\right); \quad t \geq t_1,\ \alpha > 1,\ \beta > 0. \tag{5.3.3}$$

Whereas $\lim_{t\to\infty} x(t) = x_1 \left(1+\alpha e^{-\beta \ln(t_1)}\right)$, that is the line $k = x_1 \left(1+\alpha e^{-\beta t_1}\right)$ is a horizontal asymptote of the curve Eq. (5.3.3) when t tend to ∞. Its asymptote is dependent on the initial value. Finally, by setting $a = \frac{1}{\alpha}$ and $b = e^{-\beta}$, we find the new reformulation of the log-logistic curve

$$x(t) = x_1 \left(\frac{a+\varphi(t_1)}{a+\varphi(t)}\right); \quad t \geq t_1,\ 0 < a < 1,\ 0 < b < 1, \tag{5.3.4}$$

where $\varphi(t) = b^{\ln(t)}$. This reformulation leads to conditions $0 < b < 1$ and $0 < a < 1$.

Now, starting from Eq. (5.3.4), the one-dimensional SLLDP can be defined from the diffusion process $\{x(t) : t \geq t_1,\ t_1 > 0\}$, taking values on $(0, \infty)$ and with infinitesimal moments given by

$$\begin{aligned} A_1(x,t) &= h(t)x, \\ A_2(x,t) &= \sigma^2 x^2, \end{aligned} \tag{5.3.5}$$

with $\sigma > 0$, $h(t) = \frac{-d\varphi(t)/dt}{a+\varphi(t)}$ and initial distribution $x(t_1) = x_1$. This is a solution of the Itô SDE

$$dx(t) = A_1(x,t)dt + (A_2(x,t))^{\frac{1}{2}} dw(t); \quad x(t_1) = x_1, \tag{5.3.6}$$

where $w(t)$ is a standard Wiener process and x_1 is a positive random variable, independent on $w(t)$ for $t \geq t_1$. Taking into account that h is a continuous and bounded function, the infinitesimal moments $A_1(x,t)$ and $A_2(x,t)$ specified in Eq. (5.3.5) satisfy the Lipschitz and the growth conditions for the existence and uniqueness of the solution to the SDE Eq. (5.3.6) (see, for example, Kloeden and Platen [8]). In this sense, there exists a non negative constant $C = \dfrac{-\ln(b)\varphi(t_1)}{t_1(a+\varphi(t_1))} + \sigma$, such that for all x and y in $\mathbb{R}^+$ and $t \geq t_1$ we reach:

$$|A_1(t,x) - A_1(t,y)| + \left|(A_2(t,x))^{\frac{1}{2}} - (A_2(t,y))^{\frac{1}{2}}\right| \leq C \mid x - y \mid,$$
$$|A_1(t,x)|^2 + \left|(A_2(t,x))^{\frac{1}{2}}\right|^2 \leq C^2 (1 + \mid x \mid^2).$$

Then, the SDE Eq. (5.3.6) has a unique solution $\{x(t) : t \geq t_1;\ t_1 > 0\}$ continuous with probability 1, and satisfies the initial condition $x(t_1) = x_1$.

5.3.2 Probability distribution of the SLLDP

By means of the following transformation $y(t) = \ln(x(t))$, and by using the Itô lemma, the SDE Eq. (5.3.6) becomes

$$dy(t) = \left(h(t) - \frac{\sigma^2}{2}\right) dt + \sigma dw(t); \quad y_1 = \ln(x_1),$$

by integrating both sides yields,

$$y(t) = y_1 + \ln\left(\frac{a+\varphi(t_1)}{a+\varphi(t)}\right) - \frac{\sigma^2}{2}(t-t_1) + \sigma(w(t) - w(t_1)).$$

Finally, we have:

$$x(t) = x_1 \left(\frac{a+\varphi(t_1)}{a+\varphi(t)}\right) \exp\left[-\frac{\sigma^2}{2}(t-t_1) + \sigma(w(t) - w(t_1))\right]. \tag{5.3.7}$$

The $y(t)$ is a gaussian process if and only if $\ln(x_1)$ is constant or normally distributed (see, Arnold [82]). In such a case, the $x(t)$ is a log-normal process. That is, the TPDF

$$f(x,t|y,s) = \frac{1}{x\sqrt{2\pi\sigma^2(t-s)}} \exp\left(-\frac{\left[\ln\left(\frac{x}{y}\right) - \ln\left(\frac{a+\varphi(s)}{a+\varphi(t)}\right) + \frac{\sigma^2}{2}(t-s)\right]^2}{2\sigma^2(t-s)}\right). \tag{5.3.8}$$

5.3.3 The MFs of the SLLDP

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

$$\begin{aligned} E(x^r(t)|x(s) = x_s) =& \exp\left(\ln(x_s^r) + r\ln\left(\frac{a+\varphi(s)}{a+\varphi(t)}\right) - \frac{r\sigma^2}{2}(t-s) + \frac{r^2\sigma^2}{2}(t-s)\right) \\ =& x_s^r \left(\frac{a+\varphi(s)}{a+\varphi(t)}\right)^r \times \exp\left(\frac{r(r-1)\sigma^2}{2}(t-s)\right). \end{aligned} \tag{5.3.9}$$

Then, for $r=1$, the CMF of the process is

$$E\left(x(t)|x(s)=x_s\right)=x_s\left(\frac{a+\varphi(s)}{a+\varphi(t)}\right). \tag{5.3.10}$$

In addition, taking into the initial condition $\mathrm{P}\left(x\left(t_1\right)=x_1\right) = 1$, the MF of the process is given by

$$E(x(t))=x_1\left(\frac{a+\varphi\left(t_1\right)}{a+\varphi(t)}\right). \tag{5.3.11}$$

and we have

$$E[x(t)x(s)]=E\left[\exp\left\{\ln\left(x(t)x(s)\right)\right\}\right]=E\left[\exp\left\{y(t)+y(s)\right\}\right],$$

i.e.,

$$E[x(t)x(s)]=\exp\left\{\ln\left(x_1^2\right)+\ln\left[\left(\frac{a+\varphi\left(t_1\right)}{a+\varphi(t)}\right)\left(\frac{a+\varphi\left(t_1\right)}{a+\varphi(s)}\right)\right]+\sigma^2\left(t\wedge s-t_1\right)\right\},$$

which is written as

$$E[x(t)x(s)]=x_1^2\left(\frac{a+\varphi\left(t_1\right)}{a+\varphi(t)}\right)\left(\frac{a+\varphi\left(t_1\right)}{a+\varphi(s)}\right)\exp\left(\sigma^2\left(t\wedge s-t_1\right)\right), \tag{5.3.12}$$

where $t\wedge s=\min\left(t,s\right)$ (for further details see, [113]).

5.3.4 The covariance and autocorrelation functions of the SLLDP

From Eqs. (5.3.9) and (5.3.10), we can deduce that the variance function of the $x(t)$ is given by

$$\mathrm{Var}(x(t))=x_1^2\left(\frac{a+\varphi\left(t_1\right)}{a+\varphi(t)}\right)^2\left(\exp\left[\sigma^2\left(t-t_1\right)\right]-1\right). \tag{5.3.13}$$

The covariance function of $x(t)$ is given by

$$\mathrm{Cov}\left(x(t),x(s)\right)=E[x(t)x(s)]-E[x(t)]E[x(s)].$$

From Eqs. (5.3.11) and (5.3.12), we can deduce that the expression of $\mathrm{Cov}\left(x(t),x(s)\right)$ is

$$\mathrm{Cov}\left(x(t),x(s)\right)=x_1^2\left(\frac{a+\varphi\left(t_1\right)}{a+\varphi(t)}\right)\left(\frac{a+\varphi\left(t_1\right)}{a+\varphi(s)}\right)\left(\exp\left[\sigma^2(t\wedge s-t_1)\right]-1\right). \tag{5.3.14}$$

The autocorrelation function between time t and s of the process $x(t)$ is

$$R\left(t,s\right)=\frac{\mathrm{Cov}\left(x(t),x(s)\right)}{\left(\mathrm{Var}[x(t)]\mathrm{Var}[x(s)]\right)^{1/2}}.$$

From Eqs. (5.3.13) and (5.3.14), we can deduce that the expression of $R\left(t,s\right)$ is

$$\begin{aligned}R\left(t,s\right)=&\left(\exp\left[\sigma^2\left(t\wedge s-t_1\right)\right]-1\right)\left(\exp\left[\sigma^2\left(t-t_1\right)\right]-1\right)^{-1/2}\\&\times\left(\exp\left[\sigma^2\left(s-t_1\right)\right]-1\right)^{-1/2}.\end{aligned} \tag{5.3.15}$$

5.3.5 Statistical inference on the model

In this section, we examine the ML estimation of the parameters of the model from which and by virtue of Zehna's theorem [68], we can obtain the corresponding estimated mean functions of the process.

5.3.5.1 Estimation of the parameters

We consider a discrete sampling of the process, that is, for fixed times $t_1, t_2, \cdots, t_n$, $(n > 2)$, we observe the variables $x(t_1)$, $x(t_2)$, $\cdots$, $x(t_n)$ whose values provide the basic sample for the inference process. In addition, we assume $t_i - t_{i-1} = h$, for $i = 2, \cdots, n$. Let x_1, x_2, $\cdots$, x_n be the observed values of the sampling. The LF depends on the choice of the initial distribution.

If $\mathrm{P}\left(x\left(t_1\right) = x_1\right) = 1$, the associated LF is

$$\mathbf{L} = L_{x_1, x_2, \cdots, x_n}\left(a, b, \sigma^2\right) = \prod_{i=2}^{n} f\left(x_i, t_i | x_{i-1}, t_{i-1}\right),$$

which is written as

$$\mathbf{L} = \prod_{i=2}^{n} \frac{1}{x_i \sqrt{2\pi\sigma^2 h}} \exp\left(-\frac{\left\{\ln\left(\frac{x_i}{x_{i-1}}\right) - \ln\left(\frac{a+\varphi(t_{i-1})}{a+\varphi(t_i)}\right) + \frac{\sigma^2}{2} h\right\}^2}{2\sigma^2 h}\right),$$

where a, b and σ^2 are the parameters to be estimated.

If $x\left(t_1\right)$ is distributed as one-dimensional log-normal distribution $\Lambda_1\left(\mu_1, \sigma_1^2\right)$, the associated LF is

$$\mathbf{L} = L_{x_1, x_2, \cdots, x_n}\left(\mu_1, \sigma_1^2, a, b, \sigma^2\right) = f_{x(t_1)}\left(x_1\right) \prod_{i=2}^{n} f\left(x_i, t_i | x_{i-1}, t_{i-1}\right),$$

which is written as

$$\begin{aligned} \mathbf{L} = & \frac{1}{x_1 \sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{\left(\ln\left(x_1\right) - \mu_1\right)^2}{2\sigma_1^2}\right) \\ & \times \prod_{i=2}^{n} \frac{1}{x_i \sqrt{2\pi\sigma^2 h}} \exp\left(-\frac{\left\{\ln\left(\frac{x_i}{x_{i-1}}\right) - \ln\left(\frac{a+\varphi(t_{i-1})}{a+\varphi(t_i)}\right) + \frac{\sigma^2}{2} h\right\}^2}{2\sigma^2 h}\right). \end{aligned} \tag{5.3.16}$$

In the following, we will consider the case when the initial distribution is log-normal. From Eq. (5.3.16), the LLF of the sample is

$$\begin{aligned} \ln(\mathbf{L}) = & -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\sigma_1^2\right) - \frac{(n-1)}{2}\ln\left(\sigma^2\right) - \sum_{i=1}^{n}\ln\left(x_i\right) - \frac{1}{2\sigma_1^2}\left(\ln\left(x_1\right) - \mu_1\right)^2 \\ & - \frac{(n-1)}{2}\ln(h) - \frac{1}{2\sigma^2 h}\sum_{i=2}^{n}\left\{\ln\left(\frac{x_i}{x_{i-1}}\right) - \ln\left(\frac{a+\varphi\left(t_{i-1}\right)}{a+\varphi\left(t_i\right)}\right) + \frac{\sigma^2}{2}h\right\}^2. \end{aligned} \tag{5.3.17}$$

Deriving the LLF with respect to μ_1 and σ_1^2 we obtain

$$\begin{aligned}\frac{\partial \ln(\mathbf{L})}{\partial \mu_1} &= \frac{1}{\sigma_1^2}\left(\ln\left(x_1\right) - \mu_1\right),\\ \frac{\partial \ln(\mathbf{L})}{\partial \sigma_1^2} &= -\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_1^4}\left(\ln\left(x_1\right) - \mu_1\right)^2.\end{aligned} \tag{5.3.18}$$

Making the derivatives (5.3.18) equal to zero, we obtain the following set of equations

$$\begin{aligned}&\ln\left(x_1\right) - \mu_1 = 0,\\ &\sigma_1^2 - \left(\ln\left(x_1\right) - \mu_1\right)^2 = 0.\end{aligned} \tag{5.3.19}$$

Then, from (5.3.19), the ML estimates of μ_1 and σ_1^2 are

$$\hat{\mu}_1 = \ln\left(x_1\right) \quad \text{and} \quad \hat{\sigma^2}_1 = \left(\ln\left(x_1\right) - \hat{\mu}_1\right)^2.$$

Nevertheless, the estimation of the parameters a, b and σ^2 poses a few difficulties. In fact, the resulting likelihood equations system has no explicit solution and must be dealt with by numerical methods. To address this problem, we propose the use of the simulated annealing algorithm in order to maximise the likelihood function or, equivalently, its logarithm.

5.3.5.2 Estimated MFs

The EMFs of the process is obtained using the Zehna's theorem [68] by replacing the parameters in equation (5.3.11) by their estimators. Then the ECMF is given by the following expression

$$\hat{E}\left(x(t)|x(s) = x_s\right) = x_s\left(\frac{\hat{a} + \hat{\varphi}(s)}{\hat{a} + \hat{\varphi}(t)}\right). \tag{5.3.20}$$

Taking into account the initial condition that is $\mathrm{P}\left(x\left(t_1\right) = x_1\right) = 1$, the EMF of the process is given by

$$\hat{E}(x(t)) = x_1\left(\frac{\hat{a} + \hat{\varphi}\left(t_1\right)}{\hat{a} + \hat{\varphi}(t)}\right). \tag{5.3.21}$$

5.3.5.3 Confidence bounds

Using the procedure described by Katsamaki and Skiadas in [67], we can obtain the CB of the process. Let $v(s,t) = x(t)|x(s) = x_s$. Since the variable $w(t) - w(s)$ is Gaussian with the mean equal to zero and the variance $t - s$ for $t \geq s$. Therefore, the random variable z is given by

$$z = \frac{\ln(v(s,t)) - \mu\left(s,t,x_s\right)}{\sigma\sqrt{(t-s)}} \sim N_1(0,1),$$

where $\mu\left(s,t,x_s\right) = \ln\left(x_s\right) + \ln\left(\frac{a + \varphi(s)}{a + \varphi(t)}\right) - \frac{\sigma^2}{2}(t-s)$.

A $100(1-\kappa)\%$ conditional CB for z is given by $P(-\lambda \leq z \leq \lambda) = 1-\kappa$. From this, we can obtain a CB of $v(s,t)$ with following form $(v_{lower}(s,t), v_{upper}(s,t))$, where

$$v_{lower}(s,t) = \exp\left(\mu\left(s,t,x_s\right) - \lambda\sigma\sqrt{(t-s)}\right), \tag{5.3.22}$$

and

$$v_{upper}(s,t) = \exp\left(\mu\left(s,t,x_s\right) + \lambda\sigma\sqrt{(t-s)}\right), \tag{5.3.23}$$

with $\lambda = F^{-1}_{N_1(0,1)}\left(1-\frac{\kappa}{2}\right)$ and where $F^{-1}_{N_1(0,1)}$ is the inverse cumulative normal standard distribution and $\mu\left(s,t,x_s\right) = \ln\left(x_s\right) + \ln\left(\frac{a+\varphi(s)}{a+\varphi(t)}\right) - \frac{\sigma^2}{2}(t-s)$.

Besides, the estimated lower bound $\hat{v}_{lower}(t)$ and the estimated upper bound $\hat{v}_{upper}(t)$ can be obtained by substituting the parameters by theirs estimators in the equations Eq. (5.3.22) and Eq. (5.3.23), the estimated confidence bounds are given by:

$$\hat{v}_{lower}(s,t) = \exp\left(\hat{\mu}\left(s,t,x_s\right) - \lambda\hat{\sigma}\sqrt{(t-s)}\right), \tag{5.3.24}$$

and

$$\hat{v}_{upper}(s,t) = \exp\left(\hat{\mu}\left(s,t,x_s\right) + \lambda\hat{\sigma}\sqrt{(t-s)}\right), \tag{5.3.25}$$

where $\hat{\mu}\left(s,t,x_s\right) = \ln\left(x_s\right) + \ln\left(\frac{\hat{a}+\hat{\varphi}(s)}{\hat{a}+\hat{\varphi}(t)}\right) - \frac{\hat{\sigma}^2}{2}(t-s)$.

5.3.5.4 Application of the SA algorithm

5.3.5.5 The objective function for SLLDP

In the present case, the problem is to maximise the likelihood function or, equivalently, its logarithm Eq. (5.3.17). From Eq. (5.3.17), we get the target function (a function of parameters a, b and σ^2):

$$\phi\left(a,b,\sigma^2\right) = \frac{(n-1)}{2}\ln\left(\sigma^2\right) + \frac{1}{2\sigma^2 h}\sum_{i=2}^{n}\left\{\ln\left(\frac{x_i}{x_{i-1}}\right) - \ln\left(\frac{a+\varphi\left(t_{i-1}\right)}{a+\varphi\left(t_i\right)}\right) + \frac{\sigma^2}{2}h\right\}^2. \tag{5.3.26}$$

5.3.5.6 Bounding the solution space

Regarding parameter σ, when it has high values it leads to sample paths with great variability around the mean of the process. Thus, excessive variability in available paths would make an logistic-type modeling inadvisable (see, Figure 5.4). Some simulations performed for several values of σ have led us to consider that $0 < \sigma < 0.1$, so that we may have paths compatible with an log-logistic-type growth curve. On the other hand, the parameters a and b are bounded by the formulation developed in the section 5.3.1. Thus, the solution space, which is obtained numerically for a, b and σ is bounded and takes the form $(0,1)\times(0,1)\times(0,0.1)$.

Once the solution space has been bounded, we specify the choice of the initial parameters of the algorithms and the stopping conditions. We consider the following:

1. The initial solution is chosen randomly in the bounded subspace
$$\Theta = (0,1)\times(0,1)\times(0,0.1).$$
2. The initial temperature should be high enough such that in the first iteration of the algorithm, the probability of accepting a worse solution is at least of 80% (see, Kirkpatrick and al. [58]). For this we assume the initial temperature of 10.
3. For the cooling process we have considered a geometric scheme in which the current temperature is multiplied by a constant γ $(0<\gamma<1)$, i.e. $T_i=\gamma T_{i-1}$, $i\geq 1$. The usual values for γ are between 0.80 and 0.99. For this we have set $\gamma=0.95$.
4. The length of each temperature level (L) determines the number of solutions generated at a certain temperature, T. For this we have set $L=50$.
5. The stopping criterion defines when the system has reached a desired energy level. Equivalently it defines the total number of solutions generated, or when an acceptance ratio (ratio between the number of solutions accepted and the number of solutions generated) is reached. The application of the algorithm will be limited to 1000 iterations.

The coding is performed using Matlab computer software.

5.3.6 Simulation study

This section will analyze the application of the SA algorithm to the obtainment of maximum likelihood estimations for parameters a, b and σ (function of parameters α, β and σ) in a stochastic log-logistic diffusion process with infinitesimal moments given by Eq. (5.3.5), and are compared with those of a stochastic logistic diffusion process (SLDP) using a simulation example.

5.3.6.1 Simulated sample paths of the SLLDP

This section will present some simulated sample paths of the SLLDP. The trajectory of the model is obtained by simulating the exact solution of the SDE Eq. (5.3.7). We obtain the simulated trajectories of the model by considering the equally spaced time discretization of the interval $[t_1,T]$, with time points $t_i=t_{i-1}+(i-1)h$; for $i=2,\cdots,N$ and the discretization step size $h=\dfrac{T-t_1}{N}$ for the sample size N. The random variable $\sigma\left(w(t)-w\left(t_1\right)\right)$ in the Eq. (5.3.7) is distributed as one-dimensional normal distribution $N_1\left(0,\sigma^2\left(t-t_1\right)\right)$. Therefore, in this simulation, 50 sample paths are simulated with $t_1=0.01$, $T=50$, $x_1\sim\Lambda_1(1,0.16)$ and 250 observations of the process. Figure 5.4 shows some simulated sample paths of the SLLDP for several values of a, b and σ.

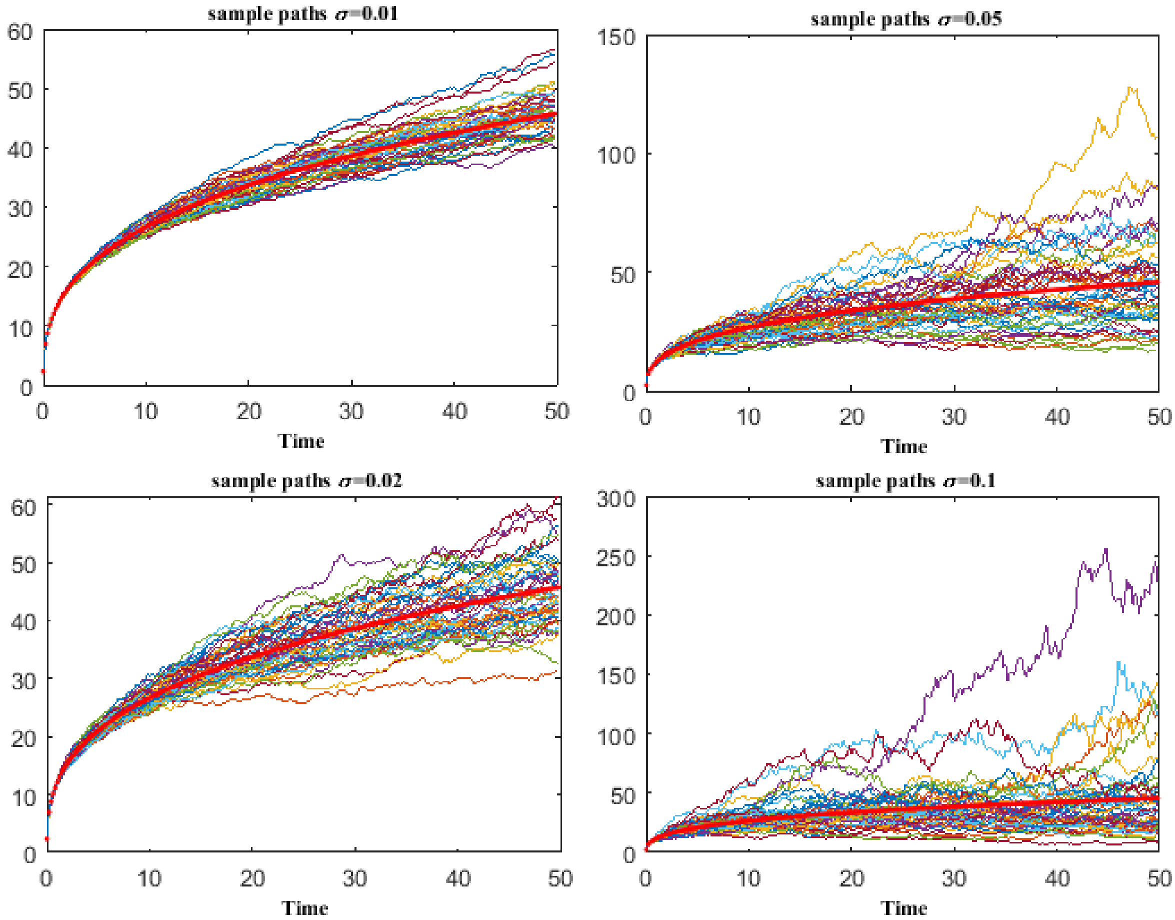


Figure 5.4: Simulated trajectories of the SLLDP and its MF for several values of σ with $(a = 0.02, b = 0.7)$.

5.3.6.2 Parameters estimation

This section will present several examples in order to validate the estimation procedure previously developed. To this end, Eq. (5.3.7) was simulated 25 times under the following assumptions, for each one, equally spaced time instants in the interval $[t_1, T]$ with step $h = 0.1$, starting from instant $t_1 = 0.01$ and $x_1 \sim \Lambda_1(1, 0.16)$. As for the sample size, values 50, 100, 250 and 500 have been considered for N.

In order to make the subsequent inference we have considered, in each case, 25 sample paths with $t_i = t_{i-1} + (i-1)h$; for $i = 2, \cdots, N$. The SA algorithm has been applied for estimating the parameters of the process with the specifications detailed in section 5.2.2.4.

The results obtained are compared with the results obtained for the SLDP and are summarised in Table 5.4, which shows for each of the above data sets the empirical mean, the std, and the CV obtained for a, b and σ are defined in Table 3.1.

Table 5.4: Comparison of a, b and σ obtained by SLDP and SLLDP.

		$a = 0.02,$		$b = 0.7,$		$\sigma = 0.01$				
N	Model	$\overline{a}$	$\overline{b}$	$\overline{\sigma}$	$\text{std}(a)$ $.10^{-4}$	$\text{std}(b)$ $.10^{-4}$	$\text{std}(\sigma)$ $.10^{-4}$	$\text{CV}(a)$ $.10^{-4}$	$\text{CV}(b)$ $.10^{-4}$	$\text{CV}(\sigma)$ $.10^{-4}$
50	SLDP	0.0211	0.6994	0.0093	99.552	70.142	8.5588	4719.8	100.29	914.98
	SLLDP	0.0183	0.6999	0.0102	67.302	10.139	10.807	3679.7	14.486	106.49
100	SLDP	0.0201	0.7003	0.0099	13.600	34.781	8.4067	678.04	49.663	841.81
	SLLDP	0.0196	0.7002	0.0097	21.902	8.2297	6.0292	111.78	11.754	619.20
250	SLDP	0.0199	0.6993	0.0099	7.5082	29.508	4.2502	377.16	42.196	428.63
	SLLDP	0.0204	0.7001	0.0101	13.721	11.288	3.1797	271.85	16.125	316.43
500	SLDP	0.0198	0.7001	0.0099	7.8697	21.927	3.1590	396.84	31.321	316.74
	SLLDP	0.0205	0.7002	0.0099	8.8669	8.5950	2.5593	231.75	12.275	257.06

The values obtained by SLDP and SLLDP are close to those expected. The CV obtained by SLLDP are usually smaller than those obtained by SLDP, from which we deduce that although the estimators obtained by these two methods differ only slightly, those obtained by SLLDP are more closely in line with the expected value.

5.3.7 Application

The following example based on the studies developed by Román-Román et al. [3] on some aspects related to the growth in cultures of some microorganisms in the context of the logistic-type process. The growth of a culture for which it is known that the intrinsic growth rate is 0.25 per day and the equilibrium density is 1000 individuals per millilitre. There is a total of 50 containers, in which cultures are placed at the beginning of the study $t_1 = 1$, with a density of five individuals per millilitre. The experiment is then carried out for 50 days.

We will now proceed to linking the values specified for the experiment to the parameters of the model. Parameter $\beta = \ln\left(\frac{1}{b}\right)$, identifying the intrinsic growth rate, is 0.25 days^{-1} i.e.,

$$b = e^{-0.25}.$$

The initial distribution is degenerate for value $x_1 = 5$. The value of parameter a is $\dfrac{1}{199}$. From these values, and taking into consideration that the equilibrium density determines the limit value of the logistic curve is

$$\gamma = \frac{x_1}{a}\left(a + b^{t_1}\right),$$

it is deduced that γ equals 779.90. In the case of the log-logistic curve, the equilibrium density determines the limit value is

$$\gamma = \frac{x_1}{a}\left(a + b^{\ln(t_1)}\right),$$

it is deduced that γ equals 1000. Having 50 containers for the experiment implies simulating 50 paths for the process, taking place from $t_1 = 1$ to $T = 50$.

The simulation of the sample paths include 401 observations of the process starting from instant $t_1 = 1$, with $h = \dfrac{T - t_1}{400}$ and the stochastic variability term $\sigma = 0.01$.

The methodology can be summarised in the following steps: Firstly, we use the all data to estimate the parameters a, b and σ of the process by SA with the specifications detailed in section 5.2.2.4. Moreover, we obtain the corresponding EMF value given by the expression (5.3.10). Also, we give the results attached to a 95% ECB of the process in the expressions (5.3.24) and (5.3.25). The results obtained are compared with the results obtained for the SLDP.

To illustrate the performance of SLDP and SLLDP, the results were compared according to the MAE, the RMSE and the MAPE, given by the Table 2.1.

Table 5.5 shows the values obtained from the estimation of the parameters of the SLDP and SLLDP.

Table 5.6 shows that the SLLDP performs better than SLDP.

The accuracy of the forecast can be judged from the MAPE result (see, Table 5.6), in other words if the value of the MAPE is less than 10%, the forecast is highly accurate, according to Lewis [69].

Figure 5.5 and Figure 5.6 illustrate the performance of the SLDP and SLLDP for forecasting using the MF.

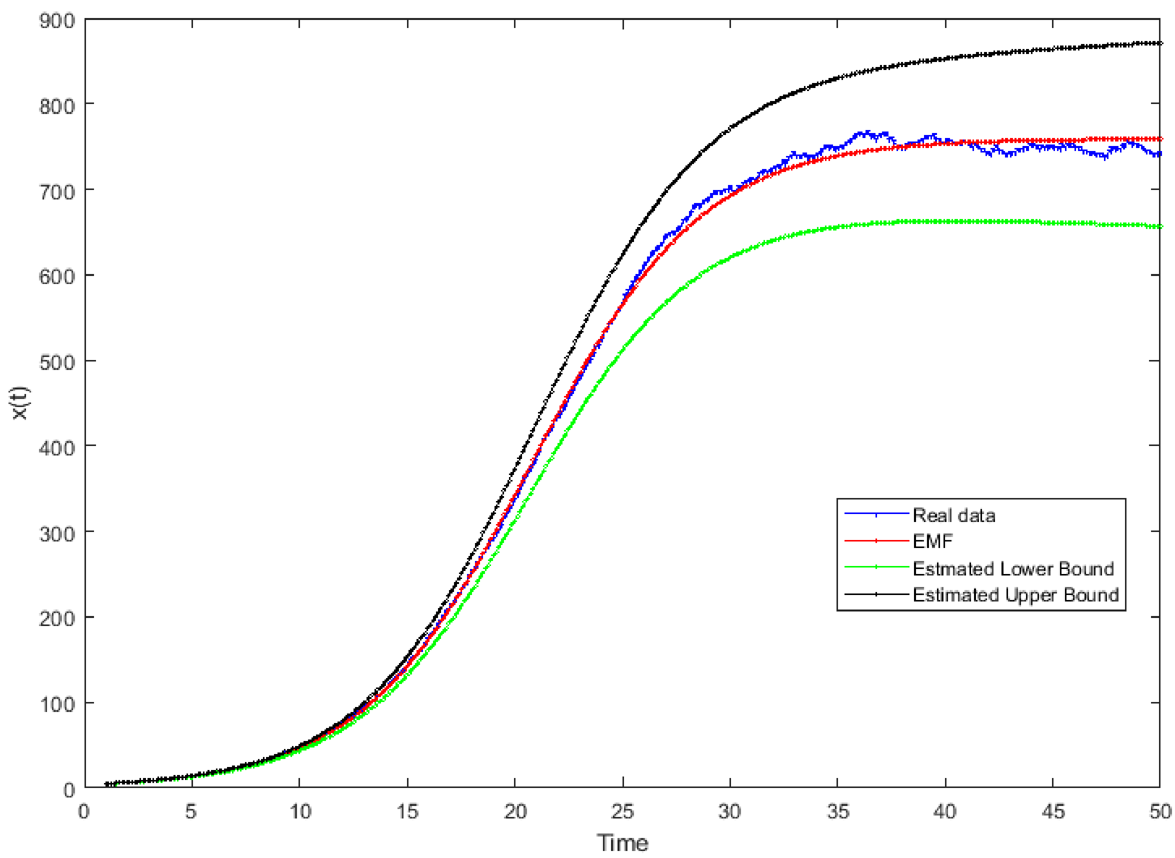


Figure 5.5: Real data versus EMF of the SLDP.

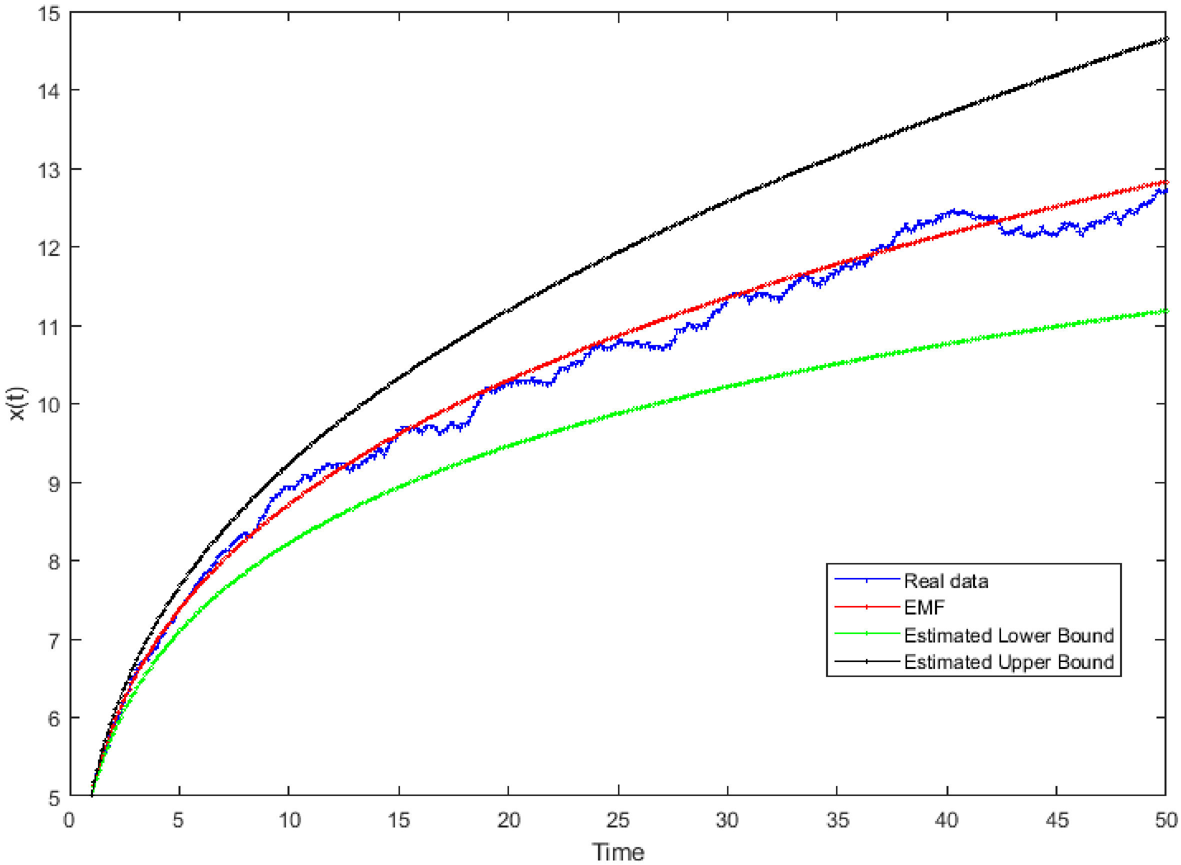


Figure 5.6: Real data versus EMF of the SLLDP.

Table 5.5: Estimation of the parameters of the models.

Model	$\hat{a}$	$\hat{b}$	$\hat{\sigma}$
SLDP	0.005146	0.775972	0.010301
SLLDP	0.004988	0.784343	0.009851

Table 5.6: Goodness of fit of the models.

Model	MAE	RMSE	MAPE
SLDP	5.7666	8.0990	1.427%
SLLDP	0.1438	0.1750	1.335%

5.4 Conclusions

In this chapter, first, we introduce a new diffusion process related to a reformulation of the logistic curve. Its distribution and main characteristics were analyzed, and its mean function as well as its conditional mean function was found to be logistic growth curve. The problem of maximum likelihood estimation for the parameters of the process was also considered. Since a complex system of equation appeared, which cannot be solved via classical numerical procedures, we suggest the use of metaheuristic optimization algorithms such as simulated annealing. One of the fundamental problems for the application of these methods is the space of solutions, since in this case it is continuous and unbounded, which could lead to unnecessary calculation and long algorithm running times. To this end, some strategies are suggested for bounding the space of solutions. Simulations were performed in order to test the validity of the bounding method for the space of solutions, showing that it is indeed very useful. The suggested bounding procedure was used in the application of the SA algorithms to estimating the parameters of the process. We illustrate this study, an application of the proposed model to an example for the growth of a microorganism culture illustrates the predictive possibilities of the new process.

Secondly, we introduce a new diffusion process related to the log-logistic curve. Its distribution and main characteristics were analyzed, and its mean function as well as its conditional mean function was found to be log-logistic growth curve. The problem of maximum likelihood estimation for the parameters of the process was also considered.Since a complex system of equation appeared, which cannot be solved via classical numerical procedures,we suggest the use of metaheuristic optimization algorithms such as simulated annealing. One of the fundamental problems for the application of these methods is the space of solutions, since in this case it is continuous and unbounded, which could lead to unnecessary calculation and long algorithm running times. To this end, some strategies are suggested for bounding the space of solutions. Simulations were performed in order to test the validity of the bounding method for the space of solutions, showing that it is indeed very useful. The suggested bounding procedure was used in the application of the SA algorithms to estimating the parameters of the process. Finally, an application to real data of the proposed methodology showed its usefulness in practice then, the capability of the SLLDP for forecasting and predicting is shown. The behavior of the SLLDP diffusion process is finally compared with that of the SLDP diffusion process.

Conclusions & perspectives

This thesis was devoted principally to studying a new stochastic diffusion process in which the mean function is proportional to specific growth curves. These models are useful for survival populations, reliability studies and life-testing experiments, energy consumption, biological, chemotherapy, radiotherapy, and economic phenomena.

Our first contribution, presented in the second chapter, was to elaborate on the new stochastic diffusion process related to the Schumacher growth curve. In this contribution, the main features of the process were analyzed, including the transition probability density function and conditional and non-conditional mean functions. In addition, the parameters of the process were estimated by maximum likelihood using discrete sampling. Then, we will apply this process with their associated statistical methodology to simulated data based on a discretization of the exact solution of the stochastic differential equation of the process.

In second contribution, we have introduced the stochastic process based on the Lundqvist-Korf growth curve, where we started by obtaining the probabilistic characteristics of this process as: the explicit expression of the process, its trends, and its distribution by transforming the diffusion process in a Wiener process as shown in the Ricciardi theorem [83]. Then, we have developed the statistical inference of this model using the maximum likelihood methodology. The simulated annealing algorithm is proposed to solve the problem of maximum likelihood estimation of the parameters. We will also join this study with simulated data. Then, we introduce a new stochastic diffusion process defined by the order power of the Lundqvist-Korf diffusion process. We obtain all its probabilistic characteristics. Finally, we apply the proposed models and statistical results to simulated data.

In our third contribution, we have introduced the new diffusion process related to the modified Lundqvist-Korf growth curve. Subsequently, we developed the theoretical and the practical aspects of the modified Lundqvist-Korf diffusion process as a particular case of the stochastic lognormal diffusion process. Then, we applied the simulated annealing algorithm to solve inference problems. Finally, an application to simulated data of the proposed model showed its usefulness in practice and demonstrated that the strategy used for bounding the parametric space behaves well. Then, we applied the process to study the total emission of CO_2 in Morocco. By fitting the SMLKDP to the real data from the period corresponding from 1987 to 2018, we obtained a good description of the series and good short-medium term forecasts for 2019 to 2020.

In our last contribution, first, we introduce a new diffusion process related to the reformulation of the logistic curve. We analyzed its distribution and main characteristics. Then, we found its mean function, as well as its conditional mean function, to be a logistic growth curve. Also, we considered the problem of maximum likelihood estimation for the parameters of the process. Since a complex system of equations appeared, which we cannot solve via classical numerical procedures, we suggest using metaheuristic optimization algorithms such as simulated annealing. One of the fundamental problems for applying this method is the space of solutions. Since in this case, it is continuous and unbounded, which could lead to unnecessary calculation and constant algorithm running times. To this end, we suggested some strategies for bounding the space of solutions. We performed simulations to test the validity of the bounding method for the space of solutions, showing that it is indeed useful. Finally, we present an application of the proposed model to an example of the growth of a microorganism culture. The results illustrate the predictive possibilities of the new process. Secondly, we present a new stochastic diffusion process related to the log-logistic curve. We analyze its distribution and main characteristics. Then, we found its mean function, as well as its conditional mean function. It corresponds to the log-logistic growth curve. Moreover, we considered the problem of maximum likelihood estimation for the parameters of the process. Since a complex system of equations appeared, which we cannot solve via classical numerical procedures, we suggest using metaheuristic optimization algorithms such as simulated annealing. One of the fundamental problems for applying this method is the space of solutions. Since in this case, it is continuous and unbounded, which could lead to unnecessary calculation and constant algorithm running times. We suggested to this end, some strategies are for bounding the space of solutions. We performed simulations to test the validity of the bounding method for the space of solutions, showing that it is indeed useful. We used the suggested bounding procedure in aaplying the SA algorithms to estimate the parameters of the process. Finally, an application to an example for the growth of a microorganism culture of the proposed methodology showed its usefulness in practice. Also, it showed the capability of the SLLDP for forecasting and predicting. Finally, we compared the behavior of the SLLDP diffusion process with that of the well-known stochastic logistic diffusion process obtained by the parameterizing logistic curve.

Bibliography

[1] R. Capocelli, L. Ricciardi, A diffusion model for population growth in random environment, Theoretical Population Biology 5 (1) (1974) 28–41.

[2] A. N. Giovanis, C. H. Skaidas, A stochastic logistic innovation diffusion model studying the electricity consumption in greece and the united states, Technological Forecasting and Social Change 61 (1999) 235–246.

[3] P. Román-Román, F. Torres-Ruiz, Modelling logistic growth by a new diffusion process: Application to biological systems, BioSystems 110 (1) (2012) 9–21.

[4] R. Gutiérrez, R. Gutiérrez-Sánchez, A. Nafidi, The trend of the total stock of the private car-petrol in spain: Stochastic modelling using a new gamma diffusion process, Applied Energy 86 (2009) 18–24.

[5] P. Román-Román, F. Torres-Ruiz, A stochastic model related to the richards-type growth curve. estimation by means of simulated annealing and variable neighborhood search, Applied Mathematics and Computation 266 (2015) 579–598.

[6] L. S. Istoni da, P. Román-Román, F. Torres-Ruiz, The hubbert diffusion process: Estimation via simulated annealing and variable neighborhood search procedures-application to forecasting peak oil production, Appl Stochastic Models Bus Ind. 34 (3) (2018) 376–394.

[7] S. Karlin, H. W. Taylor, A second course in stochastic processes, Academic Press, New York, 1981.

[8] P. Kloeden, E. Platen, The numerical solution of stochastic differential equations, Springer, Berlin, Germany, 1992.

[9] R. Gutiérrez, R. Gutiérrez-Sánchez, A. Nafidi, E. Ramos, Three-parameter stochastic lognormal diffusion model: Statistical computation and simulating annealing–application to real case, Journal of Statistical Computation and Simulation 79 (1) (2009) 25–38.

[10] B. M. Al Eideh, A. S. A. Al Refai, W. M. Sbeiti, Modelling the cpi using a lognormal diffusion process and implications on forecasting inflation, IMAA Journal of Management Mathematics 15 (1) (2004) 39–51.

[11] R. Gutiérrez, P. Román, D. Romero, F. Torres, Forcasting for the univariate lognormal diffusion process with exogenous factors, Cybernetics and Systems 34 (8) (2003) 709–724.

[12] R. Gutiérrez, P. Romàn, F. Torres, Inference on some parametric functions in the univariate lognormal diffusion process with exogenous factors, Test 10 (2) (2001) 375–373.

[13] J. C. Cox, J. E. Ingersoll, J. S. A. Ross, A theory of the term structure of interest rates, Econometrica 53 (2) (1985) 385–408.

[14] R. Gutiérrez, A. Nafidi, R. Gutiérrez-Sànchez, Inference in the stochastic gompertz diffusion model with continuous sampling, Monografias del Seminario Garcia de Galdeano 31 (10) (2004) 247–253.

[15] R. Gutiérrez, R. Gutiérrez-Sánchez, A. Nafidi, Electricity consumption in morocco: Stochastic gompertz diffusion analysis with exogenous factors, Applied Energy 83 (10) (2006) 1139–1151.

[16] R. Gutiérrez, R. Gutiérrez-Sánchez, A. Nafidi, Modelling and forecasting vehicle stocks using the trends of stochastic gompertz diffusion models: The case of spain, Applied Stochastic Models in Business and Industry 25 (3) (2009) 385–405.

[17] R. Gutiérrez, A. Nafidi, R. Gutiérrez-Sánchez, Forecasting total natural-gas consumption in spain by using the stochastic gompertz innovation diffusion model, Applied Energy 80 (2005) 115–124.

[18] R. Gutiérrez-Jáimez, P. Román, D. Romero, J. Juan Serrano, F. Torres, A new gompertz-type diffusion process with application to random growth, Mathematical Biosciences 208 (2007) 147–165.

[19] R. Gutiérrez, R. Gutiérrez-Sánchez, A. Nafidi, Trend analysis and computational statistical estimation in a stochastic rayleigh model: Simulation and application, Mathematics and Computers in Simulation 77 (6) (2008) 209–217.

[20] R. Gutiérrez, R. Gutiérrez-Sánchez, A. Nafidi, The stochastic rayleigh diffusion model: Statistical inference and computational aspects. applications to modelling of real cases, Applied Mathematics and Computation 175 (1) (2006) 628–644.

[21] V. Giorno, A. Nobile, L. Ricciardi, L. Sacerdote, Some remarks on the rayleigh process, Journal of Applied Probability 23 (2) (1986) 398–408.

[22] B. M. Bibby, M. Sørensen, A hyperbolic diffusion model for stock prices, Finance and Stochastics 1 (1) (1996) 25–41.

[23] G. D'Onofrio, M. Tamborrino, P. Lansky, The jacobi diffusion process as a neuronal model, Chaos 28 (4) (2018) 103–119.

[24] J. L. Forman, M. Sorensen, The pearson diffusions: A class of statistically tractable diffusion processes, Scandinnavian Journal of Statistics 35 (2005) 438–465.

[25] P. Román-Román, R. Desirée, F. Torres-Ruiz, A diffusion process to model generalized von bertalanffy growth patterns: Fitting to real data, Journal of Theoretical Biology 263 (1) (2010) 59–69.

[26] A. Nafidi, G. Moutabir, R. Gutiérrez-Sánchez, Stochastic brennan–schwartz diffusion process: Statistical computation and application, Mathematics 7 (11) (2019) 1062.

[27] A. Nafidi, M. Bahij, R. Gutiérrez-Sánchez, B. Achchab, Two-parameter stochastic weibull diffusion model: Statistical inference and application to real modeling example, Mathematics 8 (2) (2020) 160.

[28] A. Nafidi, A. El Azri, A stochastic diffusion process based on the lundqvist-korf growth: Computational aspects and simulation, Mathematics and Computers in Simulation 182 (2021) 25–38.

[29] L. M. Ricciardi, Diffusion processes and related topics in biology, Vol. 14, Springer Science & Business Media, 2013.

[30] L. Ferrante, S. Bompadre, L. Leone, M. Montanari, A stochastic formulation of the gompertzian growth model for in vitro bactericidal kinetics: Parameter estimation and extinction probability, Biometrical Journal 47 (3) (2005) 309–318.

[31] L. Ferrante, S. Bompadre, L. Possati, L. Leone, Parameter estimation in a gompertzian stochastic model for tumor growth, Biometrics 56 (4) (2000) 1076–1081.

[32] R. Gutiérrez, R. Gutiérrez-Sànchez, A. Nafidi, E. Ramos, A diffusion model with cubic drift: Statistical and computational aspects and application to modelling of the global CO_2 emission in spain, Environmetrics 18 (2007) 55–69.

[33] R. Gutiérrez, R. Gutiérrez-Sànchez, A. Nafidi, A. Pascual, Detection, modelling and estimation of non-linear trends by using a non-homogeneous vasicek stochastic diffusion. application to CO_2 emissions in morocco, Stoch Environ Res Risk Assess 26 (2012) 533–543.

[34] R. Gutiérrez, R. Gutiérrez-Sànchez, A. Nafidi, A bivariate stochastic gompertz diffusion model: Statistical aspects and application to the joint modeling of the gross domestic product and CO_2 emissions in spain, Environmetrics 19 (2008) 643–658.

[35] M. Abbas, S. Dalia, C. Fausto, L. Nanthakumar, K. Masoumeh, Carbon dioxide (CO_2) emissions and economic growth: A systematic review of two decades of research from 1995 to 2017, Science of The Total Environment 649 (2019) 31–49.

[36] C. Magazzino, G. Cerulli, The determinants of CO_2 emissions in mena countries: A responsiveness scores approach, International Journal of Sustainable Development and World Ecology 26 (6) (2019) 522–534.

[37] S. Solaymani, CO_2 emissions patterns in 7 top carbon emitter economies: The case of transport sector, Energy 168 (2019) 989–1001.

[38] R. Capocelli, L. Ricciardi, Growth with regulationin random environment, Kybernetik 15 (3) (1974) 147–157.

[39] J. Heydari, C. Lawless, D. A. Lydall, D. J. Wilkinson, Fast bayesian parameter estimation for stochastic logistic growth models, BioSystems 122 (2014) 55–72.

[40] S. Tang, E. A. Heron, Bayesian inference for a stochastic logistic model with switching points, Ecological modelling 219 (2008) 153–169.

[41] R. B. Prakasa, Statistical inference for diffusion type processes, Arnold, London, 1999.

[42] Y. Ait-Sahalia, Maximum-likelihood estimation of discretely sampled diffusion: A closed-form approximation approach, Econometrica 113 (2002) 223–262.

[43] A. V. Egorov, H. Li, Y. Xu, Maximum likelihood estimation of time-inhomogeneous diffusions, Journal of Econometrics 114 (1) (2003) 107–139.

[44] A. Dadgar, K. Shafie, M. Emadi, Evidential inference for diffusion-type processes, Journal of Statistical Computation and Simulation 86 (1) (2016) 183–194.

[45] E. B. Wilson, J. Worcester, The normal logarithmic transform, Review of Economics and Statistics 27 (1945) 17–22.

[46] A. C. Cohen, Estimating parameters of logarithmic-normal distributions by maximum likelihood, Journal of American Statistical Association 46 (254) (1951) 206–212.

[47] J. A. Lambert, Estimation of parameters in the three-parameter lognormal distribution, Australian Journal of Statistics 6 (2) (1964) 29–32.

[48] H. L. Harter, A. H. Moore, The local-maximum-likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples, Journal of American Statistical Association 61 (315) (1966) 842–851.

[49] F. Calitz, Maximum likelihood estimation of the parameters of the three-parameter lognormal distribution-a reconsideration, Australian Journal of Statistics 15 (3) (1973) 185–190.

[50] K. C. Chan, G. A. Karolyi, F. A. Longstaff, A. B. Sanders, An empirical comparison of alternative models of the short-term interest rate, The Journal of Finance 3 (1992) 1209–1227.

[51] M. Arapis, J. Gao, Empirical comparison in short-terme interest rate models using nonparametric methods, Journal of Financial Econometrics 4 (2) (2006) 310–345.

[52] G. J. Jiang, J. L. Knight, A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model, Econometric Theory 13 (5) (1997) 615–645.

[53] O. ELerian, S. Chib, N. Shephad, Likelihood inference for discretely observed non-linear diffusions, Econometrica 69 (2006) 959–993.

[54] A. Nafidi, M. Bahij, B. Achchab, R. Gutiérrez-Sanchez, The stochastic weibull diffusion process: Computational aspects and simulation, Applied Mathematics and Computation 348 (2019) 575–587.

[55] P. Román-Román, D. Romero, M. A. Rubio, F. Torres-Ruiz, Estimating the parameters of a gompertz-type diffusion process by means of simulated annealing, Applied Mathematics and Computation 218 (2012) 5121–5131.

[56] J. Fernando Vera, A. Jose Díaz-García, A global simulated annealing heuristic for the three-parameter lognormal maximum likelihood estimation, Computational Statistics and Data Analysis 52 (12) (2008) 5055–5065.

[57] A. R. Pedersen, A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations, Scandinavian Journal of Statistics 22 (1) (1995) 55–71.

[58] S. Kirkpatrick, C. d. Gelatt, M. P. Vecch, Optimization by simulated annealing, Science 220 (1983) 671–680.

[59] V. Cerny, Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm, Journal of Optimization Theory and Applications 45 (1) (1985) 41–51.

[60] A. Nafidi, A. El Azri, R. Gutiérrez-Sánchez, The stochastic modified lundqvist-korf diffusion process: Statistical and computational aspects and application to modeling of the CO_2 emission in morocco, Stochastic Environmental Research and Risk Assessment 36 (4) (2022) 1163–1176.

[61] A. Nafidi, A. El Azri, R. Gutiérrez-Sánchez, A stochastic schumacher diffusion process: Probability characteristics computation and statistical analysis, Preprint submitted to Methodology and Computing in Applied Probability (Under Revision) 00 (2022) 000–000. doi:10.21203/rs.3.rs-1805443/v1.

[62] F. X. Schumacher, A new growth curve and its application to timber yield studies, J. Forestry Res. 37 (1939) 819–820.

[63] M. Schulin-Zeuthen, E. Kebreab, J. Dijkstra, S. Lopez, A. Bannink, H. Darmani Kuhi, J. France, A comparison of the schumacher with other functions for describing growth in pigs, Animal Feed Science and Technology 143 (1–4) (2008) 314–327.

[64] F. R. Silva, V. S. M. Silva, S. O. Miranda, F. H. Gava, R. Drescher, Aplicação do modelo de schumacher para expressar a capacidade produtiva em um plantio jovem de tectona grandis [application of the schumacher model to express the productive capacity in a young plantation of tectona grandis], Advances in Forestry Science 2 (4) (2015) 79–84.

[65] S. M. L. Marques da, B. D. Henrique Breda, G. J. Marinaldo, L. H. Garcia, Adjustment of the schumacher and hall model and application of artificial neural networks to estimate volume of eucalypt trees, Revista Árvore 33 (6) (2009) 1133–1139.

[66] L. Fang, S. Liu, Z. Huang, Uncertain johnson–schumacher growth model with imprecise observations and k-fold cross-validation test, Soft Comput. 24 (2020) 2715–2720.

[67] A. Katsamaki, C. Skiadas, Analytic solution and estimation of parameters on a stochastic exponential model for technology diffusion process, Applied Stochastic Models and Data Analysis 11 (1) (1995) 59–75.

[68] P. W. Zehna, Invariance of maximum likelihood estimators, Annals of Mathematical Statistics 37 (3) (1966) 744–744.

[69] C. D. Lewis, Industrial and Business Forecasting Methods: A practical guide to exponential smoothing and curve fitting, Butterworth-Heinemann, 1982.

[70] A. El Azri, A. Nafidi, A γ-power stochastic lundqvist-korf diffusion process: Computational aspects and simulation, Moroccan Journal of Pure and Applied Analysis 8 (3) (2022) 364–374.

[71] R. Gutiérrez, P. Román, F. Torres, Inference and first-passage time for the lognormal diffusion process with exogenous factors: Application to modelling in economics, Appl. Stoch. Models Bus. Ind. 15 (4) (1999) 325–332.

[72] V. Korf, Příspevěk k matematické formulaci vzrůstového zákona lesních porostů [contribution to mathematical definition of the law of stand volume growth], Lesnická práce 18 (1939) 339–379.

[73] B. Zeide, Analysis of growth equations, Forest Science 39 (1993) 594–616.

[74] R. Zamovican, Fonction de la croissance de korf, The Foretry Chronicle 55 (5) (1979) 194–197.

[75] T. Sghaier, M. Palahi, S. Garchi, J. A. Bonet, Y. Ammari, M. Pique, Modeling dominant height growth in planted pinus pinea stands in northwest of tunisia, International Journal of Forestry Research 2012 (2012) 1–12.

[76] A. Amaro, D. Reed, M. Tomé, I. Themido, Modelling dominant height growth: Eucalyptus plantations in portugal, Forest Science 44 (1) (1998) 37–46.

[77] M. Sánchez-González, M. Tomé, G. Montero, Modelling height and diameter growth of dominant cork oak trees in spain, Annals of Forest Science 62 (7) (2005) 633–643.

[78] F. Bolzan Martins, C. P. Boechat Soares, G. Fernandes da Silva, Individual tree growth models for eucalyptus in northern brazil, Scientia Agricola 71 (3) (2014) 212–225.

[79] H. Krisnawati, Y. Wang, K. Peter Ades, Generalized height-diameter model for acacia mangium willd. plantations in south sumatra, Journal of Forestry Research 7 (1) (2010) 1–19.

[80] R. Gutiérrez-Jáimez, P. Román, D. Romero, J. Juan Serrano, F. Torres, Analysis of a growth model inspired by gompertz and korf laws, and an analogous birth-death process, Mathematical Biosciences 282 (2016) 121–134.

[81] M. Duflo, Random iterative models (Applications of Mathematics), Vol. 34, Springer-Verlag, Berlin, 1997.

[82] L. Arnold, Stochastic differential equation, John Wiley and Sons, New York, 1973.

[83] L. M. Ricciardi, On the transformation of diffusion processes into the wiener process, Journal of Mathematical Analysis and Applications 54 (1976) 185–199.

[84] A. Nafidi, R. Gutiérrez, E. Gutiérrez-Sánchez, Rand Ramos-Abalos, S. El Hachimi, Modelling and predicting electricity consumption in spain using the stochastic gamma diffusion process with exogenous factors, Energy 113 (2016) 309–318.

[85] A. D. Crescenzo, P. Paraggio, Logistic growth described by birth-death and diffussion processes, Mathematics 7 (2019) 489.

[86] K. Amit, A. Balram, K. S. Tapan, S. S. Shrikanta, K. Sudarshan, Source identification and health risk assessment of atmospheric $PM_{2.5}$-bound polycyclic aromatic hydrocarbons in jamshedpur, india, Sustainable Cities and Society 52 (101801) (2020) 1–12.

[87] K. Amit, K. S. Tapan, S. S. Shrikanata, A. Balram, Characteristics, toxicity, source identification and seasonal variation of atmospheric polycyclic aromatic hydrocarbons over east india, Environmental Science and Pollution Research 27 (1) (2020) 678–690.

[88] A. Balram, K. Amit, K. S. Lokesh, Characterization and health risk assessment of particulate bound polycyclic aromatic hydrocarbons (pahs) in indoor and outdoor atmosphere of central east india, Environmental Science and Pollution Research 28 (2021) 56269–56280.

[89] A. Balram, S. S. Shrikanta, K. Amit, K. S. Tapan, K. Sudarshan, Health risk assessment, composition, and distribution of polycyclic aromatic hydrocarbons (pahs) in drinking water of southern jharkhand, east india, Archives of Environmental Contamination and Toxicology 80 (1) (2021) 120–133.

[90] A. Balram, K. Sudarshan, K. S. Tapan, K. Amit, Emission reduction of black carbon and polycyclic aromatic hydrocarbons during COVID-19 pandemic lockdown, Air Quality, Atmosphere & Health 14 (2021) 1081–1095.

[91] A. Balram, S. S. Shrikanta, Health risk assessment and characterization of polycyclic aromatic hydrocarbon from hydrosphere, Journal of Hazardous, Toxic, and Radioactive Waste 25 (05020008) (2021) 1–11.

[92] K. P. Susanta, A. Balram, Extractive separation of rare earth elements and their determination by inductively coupled plasma optical emission spectrometry in geological samples, Journal of Analytical Atomic Spectrometry 35 (7) (2020) 1395–1404.

[93] S. Wang, J. Zeng, X. Liu, Examining the multiple impacts of technological progress on CO_2 emissions in china: A panel quantile regression approach, Renewable and Sustainable Energy Reviews 103 (2019) 140–150.

[94] B. Lin, B. Xu, Factors affecting CO_2 emissions in china's agriculture sector: A quantile regression, Renewable and Sustainable Energy Reviews 94 (2018) 15–27.

[95] S. M. Hosseini, A. Saifoddin, R. Shirmohammadi, A. Aslani, Forecasting of CO_2 emissions in iran based on time series and regression analysis, Energy Reports 5 (2019) 619–631.

[96] T. Nguyen, Q. Le, Impact of globalization on CO_2 emissions in vietnam: An autoregressive distributed lag approach, Decision Science Letters 9 (2) (2020) 257–270.

[97] A. Nafidi, A. El Azri, A stochastic log-logistic diffusion process: Statistical computational aspects and application to modelling real example, Preprint submitted to Stochastic Models (Under Revision) 00 (2022) 000–000.

[98] A. Nafidi, A. El Azri, Stochastic logistic-type diffusion process with application to random growth: Estimating the parameters by means of simulated annealing, Preprint submitted to Stochastics: An International Journal of Probability and Stochastic Processes.

[99] P. F. Verhulst, Notice sur la loi que la population suit dans son accroissement [notice on the law that the population follows in its growth], Correspondance Mathématique et Physique 10 (1838) 113–121.

[100] P. F. Verhulst, Recherches mathématiques sur la loi d'accroissement de la population [mathematical research on the law of population increase], Nouveaux mémoires de l'Academie Royale des Science et Belles-Lettres de Bruxelles 18 (1845) 1–41.

[101] M. E. Solomon, Population dynamics, 2nd ed. Arnold, London, 1976.

[102] R. M. May, Stability in randomly fluctuating versus deterministic environments, The American Naturalist 107 (1973) 621–650.

[103] S. Y. Tang, L. S. Chen, The effect of seasonal harvesting on stage-structured population models, Journal of Mathematical Biology 48 (2004) 357–374.

[104] C. Xu, M. S. Boyce, D. J. Daley, Harvesting in seasonal environments, Journal of Mathematical Biology 50 (2005) 663–682.

[105] S. Bordy, Bioenergetics and growth, Reinhold, New York, 1945.

[106] D. A. Charles-Edwards, A model of leaf growth, Annals of Botany 44 (3) (1979) 523–535.

[107] G. Albano, V. Giorno, A stochastic model in tumor growth, J. Theor. Biol. 242 (2) (2006) 329–336.

[108] N. K. Jerne, E. C. Wood, The validity and meaning of the results of biological assays, Biometrics 5 (1949) 273–299.

[109] J. Berkson, Why i prefer logits to probits, Biometrics 7 (1951) 327.

[110] J. C. Streibig, P. Kudsk, Herbicide bioassays, CRC Press, Boca Raton, FL., 1993.

[111] S. S. Seefeldt, J. E. Jensen, E. P. Fuerst, Log-logistic analysis ofherbicide dose-response relationships, Weed Technology 9 (2) (1995) 218–227.

[112] P. Román-Román, D. Romero, M. A. Rubio, F. Torres-Ruiz, Estimating the parameters of a gompertz-type diffusion process by means of simulated annealing, Applied Mathematics and Computation 218 (2012) 5121–5131.

[113] R. Gutiérrez, N. Rico, P. Román, F. Torres, Approximate and generalized confidence bands for some parametric functions of the lognormal diffusion process with exogenous factors, Scientiae Mathematicae Japonicae 64 (2) (2006) 313–330.