Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/140410

Parameter Inference of a Stochastic SIS Model of Transmission of HIV/AIDS With Immigration Effect

*Youness El Ansari**, *Ali El Myr and Lahcen Omari*

Department of Mathematics, Laboratory of Computer Sciences Modeling and systems, Faculty of Sciences Dhar-Mehraz, University Sidi Mohamed Ben Abdellah, B.P. 1796-Atlas, Fez, Morocco

Received: 4 Feb. 2020, Revised: 10 May 2020, Accepted: 22 May 2020
Published online: 1 Jul. 2020

Abstract: In this paper, two techniques of parameter estimation based on the Euler-maximum likelihood are used to estimate some influential parameters of a stochastic *SIS* epidemic model of transmission of HIV/AIDS. The latter is alimented by a constant flow of new members whose fraction is infective. After presenting the two estimation techniques, we adress a complete study of the consistency and convergence of the proposed estimators. Data concerning HIV/AIDS in Morocco are used to simulate the different results and to compare the effectiveness of the used techniques.

Keywords: Euler-Maximum likelihood estimator, Stochastic *SIS* model, Consistency, Unbiased estimator, Asymptotic normality, HIV/AIDS

1 Introduction

The parameter estimation has remained a great non-trivial problem in stochastic differential equations (SDE). Due to its importance in many areas of engineering and sciences, especially in epidemiology, several methods and techniques are described in the pieces of literature to address the parameter estimation issue, see for instance [1–4]. Most papers relative to this subject, particularly in the continuous case, present several methods of parameter estimation that we can only apply in the case of some particular one-dimensional stochastic models [5, 6], or for some particular two-dimensional models with constant population sizes [7].

To study the effect of immigration of new infected individuals, from the outside of a population, to the dynamic of a communicable disease, such as bacterial disease (e.g. meningitis and pneumococcus) or sexually transmitted disease (e.g. HIV/AIDS), F. Brauer and P. van den Driessche [8] have proposed the following deterministic model:

$$\begin{cases} S'_t = (1-p)A - \beta S_t I_t - \mu S_t + \gamma I_t, \\ I'_t = pA + \beta S_t I_t - (\mu+\gamma+\alpha) I_t. \end{cases} \tag{1}$$

In this model, S_t denotes the number of individuals susceptible to have an infection at time t, and I_t denotes the number of infected individuals at time t. The positive parameters in the model are presented by giving the following demographic and epidemiological assumptions:

1. There exists a constant flow A of new individuals into the population in unit time, where a fraction $p, 0 \le p \le 1$, of A is infective.
2. There is a constant per capita natural death rate $\mu > 0$ in each class.
3. There exists a fraction $\gamma \ge 0$ of infected individuals who get recovered and a fraction $\alpha \ge 0$ of infected individuals that die from the infection in unit time.
4. β is the contact rate. Each infected individual produces βN_t contacts sufficient to transmit the infection in unit time, where $N_t = S_t + I_t$.

El Ansari *et al.* [9] have considered the case where the constant per capita natural death rate μ is subject to environmental noise, by considering the following *SIS* stochastic model:

$$\begin{cases} dS_t = ((1-p)A - \beta S_t I_t - \mu S_t + \gamma I_t)\, dt - \sigma S_t dB_t, \\ dI_t = (pA + \beta S_t I_t - (\mu+\gamma+\alpha) I_t)\, dt - \sigma I_t dB_t. \end{cases} \tag{2}$$

The technique of parameter perturbation (see for instance: Zhang *et al.* [10]) is used here to construct this model, the

* Corresponding author e-mail: y.elansari4@gmail.com

parameter μ in the two equations of (1) is replaced by $\mu + \sigma \dot{B}_t$, where B_t is a one-dimensional Brownian motion and σ is the intensity of the perturbation. The choice of the white noise here is appropriate because it is a case of environmental random variability in terrestrial systems (see Steele [11] and Vasseur and Yodzis [12]).
AIDS incubation is the period between initial infection with HIV/AIDS to development of its symptoms. This period varies with the age at which infection occurs and its median is estimated at ten years. However, some studies (see for example, [13]) indicate that the life expectancy of HIV-positive people after starting combination antiretroviral therapy (cART) has improved over time and currently can reach 43.3 years [95% confidence interval (CI) 42.5-44.2 years]. Furthermore, the vast majority of new HIV-positive people is from adults. Due to these reasons and taking into consideration the easy access to the cART, all over the world, we can predict that the rate of infected people who die from the infection tends to decline rapidly with time. Accordingly, to model the spread of HIV/AIDS in a varying population size and in presence of immigration effect, we suppose the case where no death by the infection is registered among infected individuals. That imposes to set a null value to the parameter α in SDE (2). The SDE obtained is:

$$\begin{cases} dS_t = ((1-p)A - \beta S_t I_t - \mu S_t + \gamma I_t)\,dt - \sigma S_t dB_t, \\ dI_t = (pA + \beta S_t I_t - (\mu+\gamma) I_t)\,dt - \sigma I_t dB_t, \end{cases} \tag{3}$$

The equation of the total population size $N(t)$ is obtained by summing the two equations of model (3):

$$dN_t = (A - \mu N_t)dt - \sigma N_t dB_t, \tag{4}$$

with initial condition $N_0 = S_0 + I_0$. One can see that this model describes a special case of the mean reversion process [14], with reversion rate μ, mean reversion level A/μ, volatility σ, and a constant sensitivity of the variance to the level of $N(t)$ that equal 1. It should be noted that there exist many other representations of the mean reversion models, such as the Ornstein-Uhlenbeck model, the *CIR* model proposed by Cox *et al.* [15] and the *CKLS* model (see, Chan *et al.* [16]).

The present paper aims to use two techniques of parameter estimation (which are based on the Euler-maximum likelihood estimator) to estimate the influential parameters μ, σ and A of model (3), to appreciate the credibility of these estimators by studying their asymptotics (consistency and asymptotic convergence), and to apply these results in a case study to extract some information concerning the propagation of HIV/AIDS in Morocco. The theory used to study the asymptotics of the estimators has been derived from the works of (Park and Phillips, [17]) and (Aït-Sahalia and Park, [18, 19]) .
The paper starts out with a reminder of some properties concerning the dynamics of model (3) and some preliminaries concerning the Euler-maximum likelihood estimator and its convergence.
The next section details our main result. In the beginning, we present the first parameter estimation procedure, in which the explicit forms of the estimators of A, μ and σ are presented. Next, in a new subsection, the results of the second parameter estimation method are described. The next subsection is devoted to investigate the consistency and the asymptotic convergence of the previous estimators. In the last subsection and adopting a case study of HIV/AIDS prevalence in Morocco, we simulate the different obtained results in order to extract some information concerning the propagation of HIV/AIDS in Morocco. Then, we compare the effectiveness of the two techniques.
The last section is dedicated to conclusion and some perspectives.

2 Preliminaries

2.1 Asymptotic convergence theorems of model (3)

We present, first, some important theorems concerning the asymptotic dynamics of model (3) (See: El Ansari *et al.*, [9]).

Theorem 1. *Suppose $(S_0, I_0) \in \mathbb{R}^2_+$, then there exists a unique positive solution to SDE (3) for $t \geq 0$. This solution remains in $\mathbb{R}^2_+$ with probability 1.*

Theorem 2. *The solution of model (3) is stochastically ultimate bounded and permanent for any initial value $(S_0, I_0) \in \mathbb{R}^2_+$.*

Theorem 3. *For any positive initial value (S_0, I_0), the variable I_t of model (3) is persistent in the mean a.s., more precisely,*

$$\liminf_{t\to+\infty} \frac{1}{t}\int_0^t I_s ds \geq \frac{pA}{\mu+\gamma} \quad a.s.$$

Theorem 4. *Let (S_t, I_t) be the solution of system (3). Then for every $t > 0$, the distribution of (S_t, I_t) has a density $u(t,x,y)$ and there exists a unique density $u_*(x,y)$ such that:*

$$\lim_{t\to\infty} \iint_{\mathbb{R}^2} |u(t,x,y) - u_*(x,y)|\, dxdy = 0.$$

2.2 Euler-maximum likelihood estimator

We give a description of the Euler-maximum likelihood estimation method. Then, a theorem concerning the consistency and asymptotic convergence of this estimator

is presented.
Consider the time-homogeneous SDE:

$$dX_t = a(X_t, \theta_1)dt + b(X_t, \theta_2)dB_t, \quad (5)$$

with initial condition $X_0 = x_0$. The function a is the drift function and the function b, which is supposed not nil, is the diffusion function. Let $\theta = (\theta_1^T, \theta_2^T)^T$, we denote by $\mathscr{D} = (\underline{x}, \bar{x})$ the domain of the diffusion process X_t. The Euler-Maruyama approximation of this SDE is:

$$X_{i\Delta} - X_{(i-1)\Delta} \simeq a(X_{(i-1)\Delta}, \theta_1)\Delta + b(X_{(i-1)\Delta}, \theta_2) \times (B_{i\Delta} - B_{(i-1)\Delta})$$

with a step Δ, the approximated transition density from x to y is:

$$p(x, y, \theta) = \frac{1}{\sqrt{2\pi\Delta} b(x, \theta_2)} \exp\left[-\frac{(y - x - \Delta a(x, \theta_1))^2}{2\Delta b^2(x, \theta_2)}\right].$$

Giving a sample of time span T and a step Δ and $n = T/\Delta$. Let Θ be the parameter space, a_θ represents the derivative of the function a with respect to the parameter θ, the same for b_θ. The Euler-ML estimator $\hat{\theta}$ is obtained as:

$$\hat{\theta} = \arg\max_{\theta \in \Theta} L(\theta),$$

where $L(\theta)$ is the log-likelihood function defined as:

$$L(\theta) = \sum_{i=1}^{n} \log p(x_{(i-1)\Delta}, x_{i\Delta}, \theta).$$

The following assumptions are given to show the asymptotics of the Euler-ML estimator:

$A.1$: $a(x, \theta_1)$ has its derivatives up to $6th$ order and $b(x, \theta_2)$ has its derivatives up to the $7th$ order, with respect to x on $\mathscr{D}$. $a(x, \theta_1)$ and $b(x, \theta_2)$ and their derivatives with respect to x have their derivatives up to the $6th$ order, with respect to θ on the interior of Θ.

$A.2$: Let f be the functions a or b or one of their derivatives (with respect to x or $\vartheta \in (A, \mu, \sigma)$) or b^{-1}, f is locally bounded on the domain $\mathscr{D}$ and there exists a positive nondecreasing function κ_f such that

$$\frac{1}{\kappa_f(T)} \sup_{t \in [0,T]} |f(X_t)| \to_p 0 \quad \text{and} \quad T^{-p}\kappa_f(T) \to 0$$

as $T \to \infty$ for some $p > 0$. κ_f is called the asymptotic function of f.

$A.3$: There exist positive nondecreasing functions ω_{θ_1} and ω_{θ_2} such that

$$\omega_{\theta_1}^{-2}(T) \int_0^T \frac{a_{\theta_1}^2}{b^2}(X_t)dt \quad \text{and} \quad \omega_{\theta_2}^{-2}(T) \int_0^T \frac{a_{\theta_2}^2}{b^2}(X_t)dt$$

converge in distribution to some almost surely positive definite random variables as $T \to \infty$.

$A.4$: $b^2(x) > 0$, for any $x \in \mathscr{D}$.

$A.5$: Let κ_1 and κ_2 represent any combination of the asymptotic functions in assumption A.2; they satisfy as $T \to \infty$ and $\Delta \to 0$,

$$\Delta T \to 0$$

$$\Delta^{1/4}\kappa_1(T\kappa_2(T)) \to 0$$

$A.6$: Let κ represent one of the asymptotic functions in (A.2), and $\dot{\kappa}$ represent the corresponding asymptotic function of the derivative of f with respect to the parameter. For any $\varepsilon > 0$ we have

$$T^{-\varepsilon}\frac{\dot{\kappa}(T)}{\kappa(T)} \to 0 \quad \text{as} \quad T \to \infty$$

$A.7$: Let $\theta_0 \in \Theta$ and $\omega = \mathrm{Diag}(\omega_{\theta_1}(T), \Delta^{-1/2}\omega_{\theta_2}(T))$, where ω_{θ_1} and ω_{θ_2} are the functions defined in assumption (A.3). Let κ represent one of the asymptotic functions in (A.2). Let Γ be the set of functions of T and Δ such that $v\omega^{-1} \to 0$ as $T \to \infty$. We Define

$$\mathscr{N}_{T,\Delta} = \left\{\theta : \left|v'(\theta - \theta_0)\right| \leq 1, v \in \Gamma\right\}.$$

We have

$$\sup_{\theta \in \mathscr{N}_{T,\Delta}} \left|\frac{\kappa(T, \theta_0)}{\kappa(T, \theta)}\right| \to 1$$

as $T \to \infty$ subject to (A.5).

In order to determine the asymptotic convergence of the Euler-ML estimator, we give the following theorem (see, [20, Theorem 1., pp. 17]). Hereafter, $A \approx B$ means $A - B$ is of smaller order than B.

Theorem 5. *If the assumptions A.1 to A.7 hold, the asymptotic first order terms of Euler-ML estimator are given, as follows*:

$$\hat{\theta}_1 - \theta_1 \approx \left(\int_0^T \frac{a_{\theta_1} a_{\theta_1}^T}{b^2}(X_t)dt\right)^{-1} \int_0^T \frac{a_{\theta_1}}{b}(X_t)dB_t$$

$$\hat{\theta}_2 - \theta_2 \approx \sqrt{\frac{\Delta}{2}} \left(\int_0^T \frac{b_{\theta_2} b_{\theta_2}^T}{b^2}(X_t)dt\right)^{-1} \int_0^T \frac{b_{\theta_2}}{b}(X_t)dV_t$$

as $T \to \infty$ and $\Delta \to 0$ under (A.5). V_t is a standard Brownian motion independent of B_t.

3 Main results

The first method detailed in the following is a direct application of the Euler-maximum likelihood estimation.

3.1 First method: Direct Euler-maximum likelihood estimator

The numerical approximation of system (4) by Euler-Maruyama scheme is:

$$N(t_i) = N(t_{i-1}) + (A - \mu N(t_{i-1}))\,\Delta_i + N(t_{i-1})\xi_{t_i},\ 1 \le i \le K.$$

where $(t_i)_{1\le i\le K}$ is a subdivision of $[0,T]$, $N_i = N(t_i), \Delta_i = t_i - t_{i-1}$ and $\xi_{t_i} \sim N(0,\sigma^2\Delta_i)$. We suppose that the steps are all constant and $\Delta_i = \Delta,\ 1 \le i \le K$, and we define the new variable:

$$Y_i = \frac{N_i - N_{i-1} - (A - \mu N_{i-1})\Delta}{N_{i-1}} = \xi_i,\ \ 1 \le i \le K.$$

Note that the random variables $Y_i, i = 1,2,\ldots,K$, are independent and identically distributed (iid), and that all follow the law $N(0,\sigma^2\Delta)$. Let $\theta = (A,\mu,\sigma)^T$, the density function of Y_i can be written as:

$$f(Y_i,\theta) = \frac{1}{\sigma\sqrt{2\pi\Delta}} \times \exp\left[\frac{-1}{2\sigma^2\Delta}\left(\frac{N_i - N_{i-1} - (A - \mu N_{i-1}))\,\Delta}{N_{i-1}}\right)^2\right],$$

where $1 \le i \le K$. Thus,

$$\log L(\theta) = -\frac{K}{2}\log(2\pi\sigma^2\Delta) - \frac{1}{2\sigma^2\Delta}\sum_{i=1}^{K}\left(\frac{N_i - N_{i-1} - (A - \mu N_{i-1})\,\Delta}{N_{i-1}}\right)^2.$$

The corresponding partial derivatives with respect to A, μ and σ^2 are:

$$\frac{\partial \log L(\theta)}{\partial \mu} = -\frac{1}{\sigma^2}\sum_{i=1}^{K}\left(\frac{N_i - N_{i-1} - (A - \mu N_{i-1})\,\Delta}{N_{i-1}}\right)$$
$$\frac{\partial \log L(\theta)}{\partial A} = \frac{1}{\sigma^2}\sum_{i=1}^{K}\left(\frac{N_i - N_{i-1} - (A - \mu N_{i-1})\,\Delta}{N_{i-1}^2}\right)$$
$$\frac{\partial \log L(\theta)}{\partial \sigma^2} = -\frac{K}{2\sigma^2} + \frac{1}{2\Delta(\sigma^2)^2} \times \sum_{i=1}^{K}\left(\frac{N_i - N_{i-1} - (A - \mu N_{i-1})\,\Delta}{N_{i-1}}\right)^2$$

The estimator $\hat{\theta}$ of θ verifies $\hat{\theta} = \arg\max_{\theta}(\log(L))$, which is equivalent to $\frac{\partial \log L(\theta,\xi_i)}{\partial \mu} = \frac{\partial \log L(\theta,\xi_i)}{\partial A} = \frac{\partial \log L(\theta,\xi_i)}{\partial \sigma^2} = 0, \frac{\partial^2 \log L(\theta,\xi_i)}{\partial \mu^2} \le 0, \frac{\partial^2 \log L(\theta,\xi_i)}{\partial A^2} \le 0$ and $\frac{\partial^2 \log L(\theta,\xi_i)}{\partial (\sigma^2)^2} \le 0.$

Resolving this last problem, we found the estimators $\hat{\mu}$, $\hat{A}$ and $\hat{\sigma^2}$ of μ, A and σ^2:

$$\hat{\mu} = \frac{-Y(Y-Z) + W(K-X)}{\Delta\,(WK - Y^2)} \tag{6}$$
$$\hat{A} = \frac{KZ - YX}{\Delta\,(WK - Y^2)} \tag{7}$$
$$\hat{\sigma^2} = \frac{1}{\Delta K}\sum_{i=1}^{K}\left(\frac{N_i - N_{i-1} - (\hat{A} - \hat{\mu}N_{i-1})\,\Delta}{N_{i-1}}\right)^2 \tag{8}$$

where

$$X = \sum_{i=1}^{K}\frac{N_i}{N_{i-1}},\quad Y = \sum_{i=1}^{K}\frac{1}{N_{i-1}},\quad Z = \sum_{i=1}^{K}\frac{N_i}{N_{i-1}^2}$$
$$\text{and}\ \ W = \sum_{i=1}^{K}\frac{1}{N_{i-1}^2} \tag{9}$$

3.2 Second method: Indirect Euler-maximum likelihood estimator

In this subsection, the Euler-maximum likelihood estimation is used differently to establish explicit forms for the estimators of the parameters A,μ and σ of model (3).
Let $r(t) = \mathbb{E}(N(t)), t \ge 0$. From (4), we have:

$$N(t) = N(0) + \int_0^t (A - \mu N(s))\,ds - \sigma\int_0^t N(s)dB_s.$$

Taking the expectation and using the Fubini theorem, we get:

$$\mathbb{E}(N(t)) = \mathbb{E}(N(0)) + \int_0^t (A - \mu\mathbb{E}(N(s)))\,ds - \sigma\mathbb{E}\left(\int_0^t N(s)dB_s\right).$$

According to (Mao, [21]), we have $\mathbb{E}\left(\int_0^t N(s)dB(s)\right) = 0$. Thus, we obtain the following ordinary differential equation:

$$\dot{r}(t) = A - \mu r(t), \tag{10}$$

which has the following solution:

$$r(t) = (r(0) - \frac{A}{\mu})e^{-\mu t} + \frac{A}{\mu}.$$

Substituting the value $\frac{A}{\mu}$ by $r(t)+\frac{\dot{r}(t)}{\mu}$ in (4), we get:

$$dN(t)=\mu\left(r(t)+\frac{\dot{r}(t)}{\mu}-N(t)\right)dt-\sigma N(t)dB_t. \quad (11)$$

With a subdivision $(t_i)_{0\le i\le K}$ of $[0,T]$ (with, $t_0=0$), we can get a discrete observation $(N_i)_{i=0...K}$ of $N(t)$, where $N_i=N(t_i)$. Let $r_i=r(t_i)$ and $\dot{r}_i=\dot{r}(t_i)$, $i=0,1,...,K$. We suppose that $\Delta_i=t_i-t_{i-1},\ \ i=1,...,K$, are all constants and equal Δ. Considering the initial value N_0 and applying the Euler-Maruyama scheme to equation (11), we get:

$$N_i=N_{i-1}+(\mu r_{i-1}+\dot{r}_{i-1}-\mu N_{i-1})\Delta+N_{i-1}\xi_i, \quad i=1,...,K$$

where $\xi_i=-\sigma(B_{t_i}-B_{t_{i-1}})\sim N(0,\sigma^2\Delta)$. We have:

$$\xi_i=\frac{N_i-N_{i-1}-(\mu(r_{i-1}-N_{i-1})+\dot{r}_{i-1})\Delta}{N_{i-1}}, \quad i=1,\ldots,K$$

and ξ_i, $i=1,\ldots,K$, are iid random variables. Thus, the normal density function for each ξ_i is given by:

$$f(\xi_i,\theta)=\frac{1}{\sigma\sqrt{2\pi\Delta}}\times\exp\left[\frac{-1}{2\sigma^2\Delta}\left(\frac{N_i-N_{i-1}-(\mu(r_{i-1}-N_{i-1})+\dot{r}_{i-1})\Delta}{N_{i-1}}\right)^2\right],$$

where $\theta=\begin{bmatrix}A\\ \mu\\ \sigma\end{bmatrix}$.

Since the variables $\xi_i, i=1,2,...,K$, are independents, the joint density function is:

$$f(\xi_1,\xi_2,...,\xi_K)=\prod_{i=1}^{K}f(\xi_i,\theta).$$

Thus, the likelihood function is given as:

$$L(\theta,\xi_i)=\left(\frac{1}{2\pi\sigma^2\Delta}\right)^{\frac{K}{2}}\times\exp\left[\frac{-1}{2\sigma^2\Delta}\sum_{i=1}^{K}\left(\frac{N_i-N_{i-1}-(\mu(r_{i-1}-N_{i-1})+\dot{r}_{i-1})\Delta}{N_{i-1}}\right)^2\right],$$

consequently,

$$\log L(\theta,\xi_i)=\frac{-K}{2}\log(2\pi\sigma^2\Delta)+\left[\frac{-1}{2\sigma^2\Delta}\sum_{i=1}^{K}\left(\frac{N_i-N_{i-1}-(\mu(r_{i-1}-N_{i-1})+\dot{r}_{i-1})\Delta}{N_{i-1}}\right)^2\right].$$

The MLEs $\hat{\mu}$ and $\hat{\sigma}$ for the parameters μ and σ verify:

$$\hat{\mu}=\arg\max_{\mu}(\log L(\theta,\xi_i)) \text{ and } \hat{\sigma}=\arg\max_{\sigma}(\log L(\theta,\xi_i))$$

i.e. $\frac{\partial\log L(\theta,\xi_i)}{\partial\mu}=\frac{\partial\log L(\theta,\xi_i)}{\partial\sigma}=0, \frac{\partial^2\log L(\theta,\xi_i)}{\partial\mu^2}\le 0$ and $\frac{\partial^2\log L(\theta,\xi_i)}{\partial\sigma^2}\le 0$. Resolving this problem, we found the estimators $\hat{\mu}$ and $\hat{\sigma}$ of μ and σ:

$$\hat{\mu}=\frac{\sum_{i=1}^{K}\left((N_i-N_{i-1}-\dot{r}_{i-1}\Delta)(r_{i-1}-N_{i-1})/N_{i-1}^2\right)}{\sum_{i=1}^{K}[(r_{i-1}-N_{i-1})/N_{i-1}]^2\Delta} \quad (12)$$

and

$$\hat{\sigma}=\sqrt{\frac{1}{K\Delta}\sum_{i=1}^{K}\left(\frac{N_i-N_{i-1}-[\hat{\mu}(r_{i-1}-N_{i-1})+\dot{r}_{i-1}]\Delta}{N_{i-1}}\right)^2} \quad (13)$$

From equation (10), we have

$$A=\dot{r}(t)+\mu r(t)$$

Thus, an estimator $\hat{A}$ of A is given by:

$$\hat{A}=\frac{1}{K}\sum_{i=1}^{K}\dot{r}_i+\frac{\hat{\mu}}{K}\sum_{i=1}^{K}r_i=\bar{\dot{r}}+\hat{\mu}\bar{r}. \quad (14)$$

It remains to estimate $(r_i)_{i=1...K}$ and $(\dot{r}_i)_{i=1...K}$ from the observation $N=(N_0,N_1,...N_K)$. For this purpose, we refer to (Tifenbach, [22]). The expected value $r(t)$ can be approximated by taking a convolution of the sample path $N=(N_0,N_1,...N_k)$ of the process $N(t)$:

$$r_i=\sum_{j=-M}^{M}c_jN_{i-j}, \quad i=0,...,K \quad (15)$$

where $c_j(j=-M,\ldots,M)$ are the weights of the convolution. Several convolutions exist: The moving average denoted by c^1 remains the most common, in which all the weights are constants,

$$c_j=\frac{1}{2M+1}, \quad j=-M,\ldots,M$$

And, the convolution c^2, where the weights are given by,

$$c_j=\frac{(2M-|j|)}{M(3M+1)}, \quad j=-M,\ldots,M.$$

The simulation in Figure 1 with different values of parameters and initial conditions shows the path of the process $N(t)$ compared with its expected value that uses the moving average convolution c^1, and the one that uses the convolution c^2.
A simple view of these simulations shows that, for all $t \geq 0$ these expected values are close to each others, but we can observe that the expected value with the convolution c^2 may approximate $r(t) = \mathbb{E}(N(t))$ more precisely than the expected value with the moving average c^1.
To approximate the derivative $\dot{r}(t)$ of $r(t)$, we can use the three points rule which is a numerical derivative technique based on Taylor's formula:

$$\dot{r}_i = \frac{2r_{i+1} - 3r_i + r_{i-1}}{\Delta}, \quad \text{for } i = 1, 2, \ldots, K-1. \tag{16}$$

For $i = 0$ or $i = K$, the derivate is given by:

$$\dot{r}_0 = \frac{r_1 - r_0}{\Delta}, \text{ and } \dot{r}_K = \frac{r_K - r_{K-1}}{\Delta} \tag{17}$$

3.3 Convergence of estimators

The following theorem deals with the asymptotic convergence of the estimators $\hat{\mu}, \hat{\sigma}$ and $\hat{A}$ of the two techniques.

Theorem 6. *If the condition* (24) *holds, the estimators* $\hat{\mu}, \hat{\sigma}$ *and* $\hat{A}$, *are consistent as* $T \to \infty$ *and* $\Delta \to 0$. *Furthermore, we have*:

$$\sqrt{T}(\hat{\mu} - \mu) \to_d N(0, \sigma^2) \quad as\ T \to \infty \tag{18}$$

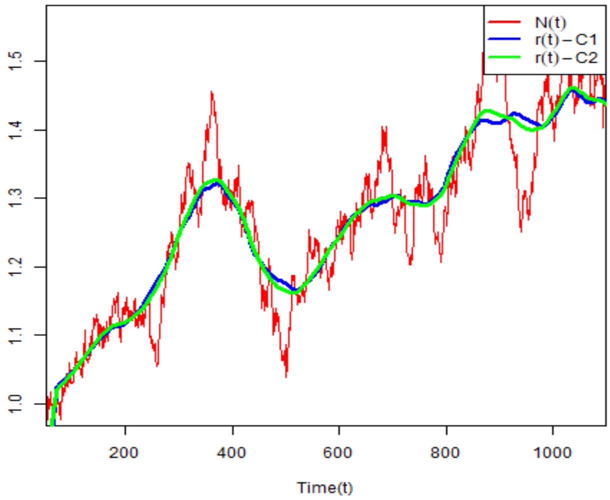
$$\sqrt{T/\Delta}\,(\hat{\sigma} - \sigma) \to_d N(0, \frac{\sigma^2}{2}) \quad as\ T \to \infty \tag{19}$$

$$\sqrt{T}(\hat{A} - A) \to_d N(0, (\bar{r}\sigma)^2) \quad as\ T \to \infty \tag{20}$$

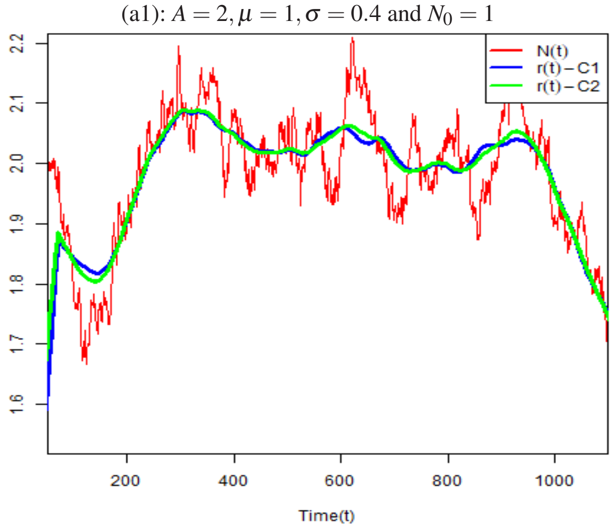
Proof. To ensure the asymtotics of the above Euler-ML estimators $\hat{\theta}$ for each technique, we check first whether the assumptions 1 to 7 are fulfilled:
The drift function and the diffusion function of SDE (4) are $a(x) = A - \mu x$ and $b(x) = \sigma x$, respectively. The first assumption (A.1) is clearly checked. The assumption (A.4) is satisfied in $\mathscr{D} = (0, \infty)$ since all the parameters A, μ and σ are positive and $N(t) > 0$ for all $t \geq 0$.
Let $X_t = N_t^{-1}$, and applying Itô formula, we have,

$$dX_t = \left(-AX_t^2 + (\mu + \sigma^2)X_t\right) dt + \sigma X_t dB_t,$$

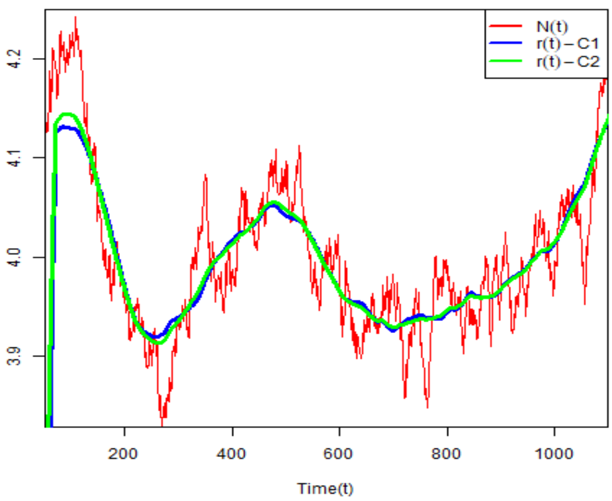
So $a^{-1}(x) = -Ax^2 + (\mu + \sigma^2)x$.
Let f be the function a or one of its derivatives, with respect to x or $\vartheta \in (A, \mu, \sigma)$, or b or one of its derivatives, with respect to x or $\vartheta \in (A, \mu, \sigma)$, or a^{-1}, J. Minsoo [20] clarified that checking assumption (A.2) for the terms of



(a1): $A = 2, \mu = 1, \sigma = 0.4$ and $N_0 = 1$



(a2): $A = 2.5, \mu = 1.5, \sigma = 0.3$ and $N_0 = 2$



(a3): $A = 3, \mu = 0.7, \sigma = 0.1$ and $N_0 = 4$.

Fig. 1: The paths of $N(t)$ and $r(t)$, by two convolutions c^1 and c^2, and for random sets of parameters and initial conditions. In all simulations $K = 1200$.

f is sufficient to conclude that assumption (A.2) is verified for f. From [9], we have

$$\lim_{t\to+\infty}\frac{N_t}{t}=0,\text{ a.s.}$$

Thus, we can easily see that

$$\lim_{T\to+\infty}\frac{\sup_{t\in[0,T]}|N_t|}{T^2}=0,\text{ a.s}$$

which gives

$$\sup_{t\in[0,T]}|N_t|=o_p(T^2),\text{ a.s.} \tag{21}$$

Thus, we can consider $\kappa_f(x)=x^2$.
In addition, Rudnicki [23] showed that $\lim_{t\to+\infty}X_t=0$ a.s, then $\lim_{t\to+\infty}\frac{X_t}{t}=0$ a.s, so

$$\lim_{T\to+\infty}\frac{\sup_{t\in[0,T]}|X_t|}{T^2}=0,\text{ a.s.}$$

and

$$\lim_{T\to+\infty}\frac{\sup_{t\in[0,T]}|X_t^2|}{T^4}=0,\text{ a.s.}$$

Consequently,

$$\sup_{t\in[0,T]}|X_t|=o_p(T^2),\text{ a.s.} \tag{22}$$

Thus, we can take $\kappa_f(x)=x^2$. And

$$\sup_{t\in[0,T]}|X_t^2|=o_p(T^4),\text{ a.s} \tag{23}$$

Hence, $\kappa_f(x)=x^4$.
For $p>0$, the p-order derivatives (with respect to x or $\vartheta\in(A,\mu,\sigma)$) of a and b are equal to 0. Thus, from (21), (22) and (23), we deduce that assumption (A.2) is verified for a and all its derivatives, b and all its derivatives and for a^{-1} in $\mathscr{D}=(0,\infty)$.
The assumption (A.3) can be clearly verified by taking $\omega_A=\omega_\mu=\omega_\sigma=T$.
Let κ_1 and κ_2 be a combination of the asymptotics function of (A.2). It is sufficient to verify assumption (A.5) with the biggest order of k_f, which is 4. Thus, (A.5) is verified if

$$\Delta^{1/4}T^{4\times5}=\Delta^{1/4}T^{20}\to0,\text{ as }T\to\infty\text{ and }\Delta\to0. \tag{24}$$

Practicaly, we can always make sure that this condition is verified.
For assumption (A.6), we remark that for all asymptotic function κ of (A.2), and for all parameter $\vartheta\in(A,\mu,\sigma)$ the value $\frac{\kappa_\vartheta(T)}{\kappa(T)}$ tends to 0 or 1 as $T\to\infty$. Thus,

$$T^{-\varepsilon}\frac{\kappa_\vartheta(T)}{\kappa(T)}\to0\quad\text{for any }\varepsilon>0,$$

as $T\to\infty$. Where κ_ϑ is the derivative of κ with respect to ϑ.
It remains to check the assumption (A.7). It is obvious to see that for all $\theta,\theta_0\in\mathbb{R}_+$, $\kappa(T,\theta)$ and $\kappa(T,\theta_0)$ have the same order independently of the set $\mathscr{N}_{T,\Delta}=\{\theta:|v'(\theta-\theta_0)|\le1\}$ with v satisfying $v\omega^{-1}\to0$ as $T\to\infty$. Thus,

$$\sup_{\theta\in\mathscr{N}_{T,\Delta}}\left|\frac{\kappa(T,\theta_0)}{\kappa(T,\theta)}\right|\to1\quad\text{as }T\to\infty.$$

All the assumptions 1 to 7 are checked for model (4), an application of Theorem 5 gives:

$$\hat{\mu}-\mu\approx\frac{\int_0^T\frac{1}{\sigma}dB_t}{\int_0^T\frac{1}{\sigma^2}dt}\approx\frac{\frac{1}{\sigma}B_T}{\frac{1}{\sigma^2}T}\approx\sigma\frac{B_T}{T},\text{ as }T\to\infty \tag{25}$$

We know that B_T is normally distributed with mean zero and variance T, so

$$\sqrt{T}(\hat{\mu}-\mu)\to_d N(0,\sigma^2),\text{ as }T\to\infty. \tag{26}$$

We also have

$$\hat{\sigma}-\sigma\approx\sqrt{\frac{\Delta}{2}}\frac{\int_0^T\frac{N_t}{\sigma N_t}dV_t}{\int_0^T\frac{N_t^2}{\sigma^2N_t^2}dt}\approx\sqrt{\frac{\Delta}{2}}\frac{\sigma V_T}{T},\text{ as }T\to\infty. \tag{27}$$

V_T is normally distributed with mean zero and variance T. Thus,

$$\sqrt{\frac{T}{\Delta}}(\hat{\sigma}-\sigma)\to_d N(0,\frac{\sigma^2}{2}),\text{ as }T\to\infty. \tag{28}$$

From (14) we have

$$\hat{A}=\bar{\dot{r}}+\hat{\mu}\bar{r},$$

so,

$$\hat{A}-A=\bar{\dot{r}}+\hat{\mu}\bar{r}-\dot{r}(t)-\mu r(t).$$

$r(t)$ is continuous, according to the intermediate value theorem, there exists $\tau\in(0,T)$ such that $r(\tau)=\bar{r}$. We can easily check that $\dot{r}(\tau)=\bar{\dot{r}}$ thus,

$$\hat{A} - A = \bar{r} - \dot{r}(\tau) + \bar{r}(\hat{\mu} - \mu) = \bar{r}(\hat{\mu} - \mu)$$

Hence,

$$\hat{A} - A = \bar{r}(\hat{\mu} - \mu) \approx \sigma\bar{r}\frac{B_T}{T}, \quad (29)$$

together with (25), we obtain:

$$\sqrt{T}(\hat{A} - A) \rightarrow_d N(0, (\bar{r}\sigma)^2), \text{ as } T \rightarrow \infty \quad (30)$$

the equations (25), (27) and (29), together with $\lim_{t\to\infty} \frac{B_t}{t} = 0$ (see: strong law of large numbers (Mao, [21]) affirm the consistency of the estimators $\hat{\mu}, \hat{\sigma}$ and $\hat{A}$ as $T \rightarrow \infty$.

Remark. We have from (26), (28) and (30), $\mathbb{E}(\hat{\theta}) \rightarrow \theta$ as $T \rightarrow \infty$. The estimators $\hat{\mu}, \hat{\sigma}$ and $\hat{A}$ are asymptotically unbiased.

3.4 Case study: HIV/AIDS in Morocco

Table 1 presents the dynamics of HIV/AIDS in Morocco between the years 1986 and 2014. This database is obtained by aggregation of statistics and data from Moroccan High Commission of Planning (HCP (2018), [25]) and Ministry of Public Health with the support of UNAIDS (Ministère de la santé du Maroc (2014), [26]).
Based on this data and on the good reason that the estimators are consistent and unbiaised for a large T, we can simulate the two techniques above (see, Subsection 3.1 and Subsection 3.2) to estimate the parameters μ, σ and A. We summarize the results in Table 2. These results reveal that for the two techniques, the values of the estimators of the three parameters are close , but with a primary preference to the second technique (for each parameter, the standard deviations (SDs) of the estimators of the second technique, either for the convolution c^1 or for c^2, are closer to zero than the SDs of the estimators of the first technique).
For all the estimators, the estimated natural death rate $\hat{\mu}$ is very close to the real rate $\mu = 0.0145$ calculated by the World Bank data (Maroc Data) [27]. Moreover, the estimators of the annual flows of new individuals added to the population are almost the same. A fraction $\hat{p}$ of this flow is infective. A simple view of the numbers of infected individuals compared with susceptible individuals can show that this fraction is very small (*i.e.* $\hat{p} << 1$), however; we can not suppose that it is inconsiderable. Consequently, the immigration of new members infected by HIV/AIDS to Morocco slightly affect the evolution of HIV/AIDS in the country. Moreover, the intensity of perturbation of the parameter μ does not exceed the value 0.127 in the two estimators. This relatively small value explains how much the randomness may affect the natural death rate in Morocco. Stability of the disease is guaranteed by Theorem 4.

Table 1: Evolution of AIDS susceptible people and AIDS infected people in sexually active population in the period from 1986 to 2014 in Morocco.

Year(t)	$S(t)$	$I(t)$	$N(t)$
1986	10741840	145	10741985
1987	11031650	326	11031976
1988	11327176	551	11327727
1989	11629750	831	11630581
1990	11940106	1176	11941282
1991	12256252	1597	12257849
1992	12580404	2121	12582525
1993	12914674	2762	12917436
1994	13259076	3528	13262604
1995	13612534	4416	13616950
1996	13975394	5497	13980891
1997	14342946	6773	14349719
1998	14700834	8272	14709106
1999	15030004	10048	15040052
2000	15299530	12068	15311598
2001	15534066	14208	15548274
2002	15743746	16368	15760114
2003	15942988	18506	15961494
2004	16139624	20576	16160200
2005	16334864	22524	16357388
2006	16528246	24414	16552660
2007	16715348	26234	16741582
2008	16894294	28008	16922302
2009	17067916	29700	17097616
2010	17242432	31365	17273797
2011	17410406	32867	17443273
2012	17569800	34095	17603895
2013	17717800	35292	17753092
2014	17853318	36455	17889773

Table 2: Simulation of estimators for the first and second techniques using Morocco's HIV database in Table 1
.

Methods	*Indirect E-ML estimator*		*Direct E-ML estimator*
	c^1	c^2	
$\hat{\mu}$	0.01514	0.01561	0.0162
SD	0.0243	0.0179	0.0243
$\hat{\sigma}$	0.1156912	0.123263	0.12626
SD	0.0231	0.0121	0.0254
$\hat{A}$	190387.94	192341.94	195643.94
SD	0.112	0.091	0.095

4 Conclusion and perspectives

We have applied two techniques to estimate some influential parameters of a stochastic model of transmission of HIV/AIDS that supposes the arrival of infective individuals from the outside. These estimators are very credible because of their consistency and asymptotic normality as shown Theorem 6.
The results are simulated in a case study of HIV/AIDS spread in Morocco. The two estimators of each parameter are close, but a comparison that depends on the standard deviation (SD) gave a prior preference to the estimators of the indirect Euler-maximum likelihood estimators. In addition, we concluded some pieces of information concerning the prevalence of HIV/AIDS in Morocco. However, the estimation of the other parameters (study in progress) is required to give more clarity about the spread of HIV/AIDS in Morocco in the presence of infected immigrants. Also, to answer some questions as: Does the disease have a tendency to grow or to perish? Taking into consideration the effect of immigration of infected individuals, to what extent does the estimated contact rate $\hat{\beta}$ between susceptible and infected individuals affect the evolution of HIV/AIDS in Morocco?.

Data Availability

The data has been obtained by aggregation of statistics and data from Moroccan High Commission of Planing (HCP (2018), [25]) and Ministry of Public Health with the support of UNAIDS (Ministére de la santé du Maroc (2014), [26]). The simulated data used to support the findings of this study are provided by R software and are included within the attached document.

Acknowledgement

The authors are grateful to the anonymous referee for their careful checking of the details and the constructive comments that improved this paper.

References

[1] R. Sandhu, M. Khalil, A. Sarkar and D. Poirel, Bayesian model selection for nonlinear aeroelastic systems using wind-tunnel data, Computer Methods in Applied Mechanics and Engineering, 282, 161-183 (2014).
[2] T. G. Ritto and L. C. S. Nunes, Bayesian model selection of hyperelastic models for simple and pure shear at large deformations, Computers & Structures, 156, 101-109 (2015).
[3] K. M. Yenkie and U. Diwekar, The "No Sampling Parameter Estimation (NSPE)" algorithm for stochastic differential equations, Chemical Engineering Research and Design, 129, 376-383 (2018).
[4] E. Lindström, Estimating parameters in diffusion processes using an approximate maximum likelihood approach, Annals of Operations Research, 151, 269-288 (2006).
[5] M. C. Mariani, M. A. M. Bhuiyan and O. K. Tweneboah, Estimation of stochastic volatility by using Ornstein-Uhlenbeck type models, Physica A: Statistical Mechanics and Its Applications, 491, 167-176 (2018).
[6] G. B. Durham and A. R. Gallant, Numerical Techniques for Maximum Likelihood Estimation of Continuous-Time Diffusion Processes, Journal of Business & Economic Statistics, 20(3), 297-338 (2002).
[7] J. Pan, A. Gray, D. Greenhalgh and X. Mao, Parameter estimation for the stochastic SIS epidemic model, Statistical Inference for Stochastic Processes, 17(1), 75-98 (2014).
[8] F. Brauer and P. van den Driessche, Models for transmission of disease with immigration of infectives, Mathematical Biosciences, 171(2), 143-154 (2001).
[9] Y. El Ansari, A. El Myr, L. Omari and A. Lahrouz, The Effect of Immigration on The Persistence and Ergodicity of a Stochastic SIS Model for Transmission of Disease, Applied Mathematics & Information Sciences, 12(6), 1127-1138 (2018).
[10] C. Zhang, Y. Zhao, Y. Wu and S. Deng, A Stochastic Dynamic Model of Computer Viruses, Discrete Dynamics in Nature and Society, 2012, 1-16 (2012).
[11] J. H. Steele, A comparison of terrestrial and marine ecological systems, Nature, 313, 355-358 (1985).
[12] D. A. Vasseur and P. Yodzis, The color of environmental noise, Ecology, 85(4), 1146-1152 (2004).
[13] S. Teeraananchai, S. Kerr, J. Amin, K. Ruxrungtham and M. Law, Life expectancy of HIV-positive people after starting combination antiretroviral therapy: a meta-analysis, HIV Medicine, 18(4), 256-266 (2016).
[14] J. Schwill, Energy Risk, Valuing and Managing Energy Derivatives, Zeitschrift Für Energiewirtschaft, 32(2), 139-139 (2008).
[15] J. C. Cox, J. E. Ingersoll and S. A. Ross, A Theory of the Term Structure of Interest Rates, Econometrica, 53(2), 385 (1985).
[16] K. C. Chan, G. A. Karolyi, F. A. Longstaff and A. B. Sanders, An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, The Journal of Finance, 47(3), 1209-1227 (1992).
[17] J. Y. Park and P. C. B. Phillips, Nonlinear Regressions with Integrated Time Series, Econometrica, 69(1), 117-161 (2001).
[18] Y. Aït-Sahalia and J. Y. Park, Nonparametric Estimation of Difiusion Models, working paper Texas A&M University, (2008).
[19] Y. Aït-Sahalia and J. Y. Park, Speciflcation Testing for Nonstationary Difiusions, working paper Texas A&M University, (2008b).
[20] J. Minsoo. Asymptotics for the maximum likelihood estimators of diffusion models, Doctoral dissertation, Texas A&M University, (2008).
[21] X. Mao, Stochastic Differential Equations and Applications, Elsevier Science, (2008).
[22] B. Tifenbach, Numerical Methods for Modeling Energy Spot Prices, University of Calgary, (2000).
[23] R. Rudnicki, Long-time behaviour of a stochastic prey-predator model, Stochastic Processes and Their Applications, 108(1), 93-107 (2003).

[24] D. B. H. Cline , M. Jeong and J. Y. Park, Limit Theorems for the Integral of Difiusion Processes, working paper Texas A&M University, (2008).
[25] High Commission for Planning of Morocco, Recensement général de la population et de l'habitat 2014, Population du Maroc selon l'âge et le sexe, (2004).
[26] Ministry of Health of Morocco, Mise en oeuvre de la déclaration politique sur le VIH/SIDA, Rapport National, (2014).
[27] The World Bank, Incidence of HIV in Morocco, (2018).

Youness El Ansari received his Engineering Diploma in Statistics and BI (Business Intelligence) from the faculty of Sciences and Techniques, Abdelmalek Essaadi University of Tangier. He obtained in 2020 his PhD diploma in Stochastic modelisation and statistics at the faculty of sciences, Sidi Mohamed Ben Abdellah Univeristy, laboratory of Computer Sciences, Modeling and Systems, departement of Mathematics, Fez, Morocco.

Ali El Myr obtained his Master diploma degree in mathematical sciences at Sidi Mohammed Ben Abdellah University in Morocco. He obtained in 2020 his PhD diploma in Stochastic modelisation and statistics at the faculty of sciences, Sidi Mohamed Ben Abdeklah Univeristy, laboratory of Computer Sciences, Modeling and Systems, departement of Mathematics, Fez, Morocco.

Lahcen Omari Is a professor of higher education of mathematics at the faculty of sciences, Sidi Mohamed Ben Abdellah University in Morocco. He obtained his Phd diploma in epidemiology and stochastic processes. His main areas of research are: stochastic dynamic systems and stochastic epidemic systems.