



HAL
open science

Bayesian inference of a stochastic diffusion process for the dynamic of HIV in closed heterosexual population with simulations and application to Morocco case

Abdellah Abou-Bakre, Zakaria Boufis, Hamid El Maroufy, Tewfik Kernane

► To cite this version:

Abdellah Abou-Bakre, Zakaria Boufis, Hamid El Maroufy, Tewfik Kernane. Bayesian inference of a stochastic diffusion process for the dynamic of HIV in closed heterosexual population with simulations and application to Morocco case. *Applied Mathematical Modelling*, 2022, 106, pp.448-462. 10.1016/j.apm.2022.02.014 . hal-03685912

HAL Id: hal-03685912

<https://hal.science/hal-03685912>

Submitted on 22 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Bayesian inference of a stochastic diffusion process for the dynamic of HIV in closed heterosexual population with simulations and application to Morocco case

Abdellah Abou-Bakre^{a,c*}; Zakaria Boufis^b; Hamid El Maroufy^c and Tewfik Kernane^b

^aUniv Rennes, CNRS, IRMAR [(Institut de Recherche Mathmatique de Rennes)] - UMR 6625, F-35000 Rennes, France;

^bDepartment of Probability and Statistics, Faculty of Mathematics, University of Sciences and Technology USTHB, Algiers, Algeria;

^cMathematics department, Faculty of sciences and technics, University of Sultan Moulay Slimane Beni Mellal, Morocco

ABSTRACT

A model of the HIV dynamic in a heterosexual population of fixed size is considered in this paper. We represent the dynamic by a multidimensional SIR model. We formulate the stochastic diffusion approximation process to describe the dynamic of HIV by using strong approximation theorems for density-dependent Markov chains. Our aim is to estimate the parameters of this model; to reach this goal, we use Bayesian inference with Markov chain Monte Carlo (MCMC) simulations. We prove that the posterior distributions of the parameters are shifted Generalized Inverse Gaussian (GIG). The obtained results are well illustrated by simulations and real application of Morocco's case.

KEYWORDS

Bayesian inference; SIR epidemic model; Diffusion Approximations; Parameter Estimation; Poisson process; HIV dynamic.

1. Introduction

In the last years, there has been an increasing research interest for performing inference methods of the spread of infectious diseases among populations applying mathematical models (deterministic and stochastic). Stochastic inference for epidemic models is generally complicated by the fact that, first, missing data or unobserved several model variables, and second, the high dependence in data that

*Corresponding author: Abdellah Abou-Bakre. Email: abdellah.aboubakre@gmail.com

are often available at discrete time points, while the underlying true process is continuous in time. Historically, deterministic models have received more attention, but our focus, in this paper, is on stochastic models. We believe that both deterministic and stochastic models have an important role to play; which model to use in a specific situation depends on the type of question asked, the type and complexity of the epidemic model, and whether or not there are data from which to infer model parameters (see, [1] and references therein). Therefore, Stochastic models are preferred in the cases of small populations and to describe the beginning of epidemics. In fact, stochastic models usually start with a small number of infectious individuals. This implies that the initial fraction of infectives tend to 0 as the population size $N \rightarrow \infty$. In the deterministic setting we however have to assume a fixed and strictly positive fraction of initially infectives (if we start with a fraction 0 of infectives nothing happens in the deterministic model). This implicitly implies that the deterministic model starts to approximate the stochastic counterpart only when the number of infectives in the stochastic model has grown up to a fraction strictly positive. The earlier part of the stochastic model cannot be approximated by this deterministic model, and as we have seen it might in fact never reach this level (if there is only a minor outbreak).

However, discretely observed models are discrete-time Markov processes, hence, their likelihood depends on the transition probability densities of the diffusion. However, these densities are unavailable in simple closed form (see, [2]). In the case of continuously observed data, parameter estimates can be obtained for complete data (see, [3]), and in the case of incomplete data, estimates have also been developed using some novel methods (see, [4] and references therein). However, for the considered model, the inter-observation times are large; hence, direct estimation is not favored. To overcome this difficulty, we will adopt the Bayesian inference approach (see, [5]).

Recently, the Bayesian inference of the SIR model has been treated by several authors using MCMC methods (see, [6, 7] and references therein). Since the data are only available at discrete times, we impute a finite number of auxiliary latent data points between two consecutive real observations. This method has been treated by several authors (see, [5, 8, 7, 9]).

In the epidemic modeling literature, it is often the case that prior parameters for positive quantities such as rate parameters are assigned the same vague prior distributions. This choice of prior distributions is convenient in terms of Bayesian inference due to conjugacy (see [10]). Furthermore, the flexibility of the Gamma distribution means that it is frequently used in practice as a prior distribution for rate parameters in epidemic models (see e.g. [11, 12, 13]). By the same reasons, in this work, we choose a family of independent gamma distributions as prior densities of the model parameters, which leads to a family of posterior Generalized Inverse Gaussian (GIG) distributions.

This paper is concerned with performing Bayesian statistical inference for parameters of stochastic epidemic model of the HIV dynamic in a heterosexual population. We consider the same model of transmission of HIV in a closed population with fixed size N as in a previous paper by AbouBakre and ElMaroufy [14], where they used a minimum contrast function to estimate directly the parameters of interest. However, here we use the Bayesian approach to estimate these parameters. A multidimensional Susceptible-Infected-Recovered (SIR), as presented in fig. 1, is considered with a five states, in which each individual can find himself at a given time in one of the five mutually exclusive health states named S_F (S_M) susceptible female (male), I_F (I_M) infected female (male) and R the AIDS cases.

The rest of paper is organized as follows. Sections two and three describe the model and its corresponding diffusion approximations. In the fourth section, we develop Bayesian inferences for non-linear diffusion modeling the HIV dynamics. The fifth section illustrates the numerical simulations of the model with real and simulated data. The last section is a concluding one. Some proofs and essential tools and details are left in the appendix.

2. Model description

We consider a closed and mixed heterosexual population of size N , in which the infection can be made only by heterosexual contact. Let, $S_F(t)$, $I_F(t)$, $S_M(t)$, $I_M(t)$ and $Z(t)$ denote the sizes of susceptible females, infected females, susceptible males, infected males, and AIDS cases at time t respectively. The rate λ_F (λ_M) of infection of a susceptible female (male) is $\lambda_F = \beta I_M(t)$ ($\lambda_M = \beta I_F(t)$), λ_F and λ_M are called the forces of infection. We assume that all individuals, including AIDS people, leave the random mixing sexually active population at a rate μ (due to natural death or for reasons other than dying). An individual at stage AIDS dies from the disease at a rate δ . The individuals that leave the system are replaced by an inflow of susceptible, at a proportion $\alpha = \frac{1}{2}$ for females and males ($\alpha(1 - \alpha)$ reflects proportion of females (males) in the population. It is a demographic factor that does not depend on epidemic dynamics. The value $\alpha = 1/2$ is based on the database of Moroccan High commission of planning (HCP)). Thus, the inflow rates for susceptible females and males are $B_F = B_M = \frac{1}{2}(\mu N + \delta Z)$ respectively. The infected individuals develop AIDS at a rate γ .

The situation as described in fig. 1, can be viewed as a multidimensional SIR model. Since, we have considered a closed population that is, $S_F(t) + I_F(t) + S_M(t) + I_M(t) + Z(t)$ is constant and equal to N for all t . The dynamic of epidemic is then, completely determined by the continuous time process $\mathbf{Y}(t) = (S_F(t), I_F(t), S_M(t), I_M(t))$, which is supposed to be a continuous time Markov process with

discrete space $\mathcal{D}_N = \{(a, b, c, d) \in \mathbb{N}^4; a+b+c+d \leq N\}$.

For a small time interval $[t, t + \Delta t]$ only one of the following events occurs; Birth of a susceptible female (male), infection of a susceptible female (male), natural death of a susceptible female (male) and recovery or natural death of an infected female (male). Explicitly the process \mathbf{Y} jumps from state $k = (S_F, I_F, S_M, I_M)$ at time t to state $k + l$ at time $t + \Delta t$ with infinitesimal probability $q_{k,k+l}\Delta t$ according to the following scheme:

$$\begin{array}{ccc}
k & q_{k,k+l}\Delta t & k + l \\
(S_F, I_F, S_M, I_M) & \longrightarrow & (S_F + 1, I_F, S_M, I_M) \\
(S_F, I_F, S_M, I_M) & \longrightarrow & (S_F, I_F, S_M + 1, I_M) \\
(S_F, I_F, S_M, I_M) & \longrightarrow & (S_F - 1, I_F + 1, S_M, I_M) \\
(S_F, I_F, S_M, I_M) & \longrightarrow & (S_F, I_F, S_M - 1, I_M + 1) \\
(S_F, I_F, S_M, I_M) & \longrightarrow & (S_F - 1, I_F, S_M, I_M) \\
(S_F, I_F, S_M, I_M) & \longrightarrow & (S_F, I_F, S_M - 1, I_M) \\
(S_F, I_F, S_M, I_M) & \longrightarrow & (S_F, I_F - 1, S_M, I_M) \\
(S_F, I_F, S_M, I_M) & \longrightarrow & (S_F, I_F, S_M, I_M - 1),
\end{array}$$

where the rates $q_{k-l,k}$, are given by

$$q_{k,k+l} = \left\{ \begin{array}{ll} \frac{1}{2} \cdot N \left(\mu + \delta \frac{Z}{N} \right), & \text{if } l = l^1 = (1, 0, 0, 0); \\ \beta S_F \frac{I_M}{N}, & \text{if } l = l^2 = (-1, 1, 0, 0); \\ \mu S_F, & \text{if } l = l^3 = (-1, 0, 0, 0); \\ (\mu + \gamma) I_F, & \text{if } l = l^4 = (0, -1, 0, 0); \\ \beta S_M \frac{I_F}{N}, & \text{if } l = l^5 = (0, 0, -1, 1); \\ \mu S_M, & \text{if } l = l^6 = (0, 0, -1, 0); \\ (\mu + \delta) I_M, & \text{if } l = l^7 = (0, 0, 0, -1); \\ \frac{1}{2} \cdot N \left(\mu + \delta \frac{Z}{N} \right) & \text{if } l = l^8 = (0, 0, 1, 0). \end{array} \right. \quad (1)$$

3. Diffusion approximation

In this section, we will formulate the diffusion approximation of the process $\mathbf{Y}(t) = (S_F(t), I_F(t), S_M(t), I_M(t))$, describing the evolution of the epidemic. The original model in closed population with transition densities given in (1), which is a random Markov epidemic model, it can be

formulated precisely as a Poisson process driven by stochastic differential equation (see for example [1, 15, 16]). Since we consider a multidimensional SIR with demography in closed population, this gives the following equation for the Markov process \mathbf{Y}

$$Y(t) = Y_0 + \sum_{j=1}^8 h_j P_j \left(\int_0^t \beta_j(s, Y_s) ds \right), \quad (2)$$

where the P_j s are mutually independent standard Poisson processes, and $\beta_j(t, Y_t)$ is the rate of jumps in the direction h_j at time t . Explicitly, the equation of evolution in various components of \mathbf{Y} reads,

$$\begin{aligned} S_F(t) &= S_F(0) + l^1 P_1 \left(\frac{1}{2} \int_0^t N\mu + \delta Z(s) ds \right) - l^2 P_2 \left(N^{-1} \beta \int_0^t S_F(s) I_M ds \right) - l^3 P_3 \left(\mu \int_0^t S_F(s) ds \right) \\ I_F(t) &= I_F(0) + l^2 P_2 \left(N^{-1} \beta \int_0^t S_F(s) I_M ds \right) - l^4 P_4 \left((\mu + \gamma) \int_0^t I_F(s) ds \right) \\ S_M(t) &= S_M(0) - l^5 P_5 \left(N^{-1} \beta \int_0^t I_F(s) S_M(s) ds \right) - l^6 P_6 \left(\mu \int_0^t S_M(s) ds \right) + l^8 P_8 \left(\frac{1}{2} \int_0^t N\mu + \delta Z(s) ds \right) \\ I_M(t) &= I_M(0) + l^5 P_5 \left(N^{-1} \beta \int_0^t I_F(s) S_M(s) ds \right) - l^7 P_7 \left((\mu + \delta) \int_0^t S_M(s) ds \right). \end{aligned} \quad (3)$$

In this system, the various Poisson processes $P_1, P_2, P_3, P_4, P_5, P_6, P_7$ and P_8 are mutually independent and standard with mean rate 1 and l^j for $j = 1, \dots, 8$ are defined in (1).

Now, we assume that the rate of infection at time t is proportional to the fraction of the total population that is infectious male of (female) $\frac{I_M(t)}{N}$ ($I_F(t)/N$). Let consider the scaled process $\mathbf{X}(t) = N^{-1} \mathbf{Y}(t) = (x_1(t), x_2(t), y_1(t), y_2(t))$, the equation for the proportions in diverse components leads to the following formulations,

$$\begin{aligned} x_1(t) &= x_1(0) + N^{-1} l^1 P_1 \left(\frac{1}{2} N \int_0^t \mu + \delta z(s) ds \right) - N^{-1} l^2 P_2 \left(N \beta \int_0^t x_1(s) y_2(s) ds \right) \\ &\quad - N^{-1} l^3 P_3 \left(N \mu \int_0^t x_1(s) ds \right) \\ x_2(t) &= x_2(0) + N^{-1} l^2 P_2 \left(N \beta \int_0^t x_1(s) y_2(s) ds \right) - N^{-1} l^4 P_4 \left(N(\mu + \gamma) \int_0^t x_2(s) ds \right) \\ y_1(t) &= y_1(0) - N^{-1} l^5 P_5 \left(N \beta \int_0^t x_2(s) y_1(s) ds \right) - N^{-1} l^6 P_6 \left(N \mu \int_0^t y_1(s) ds \right) \\ &\quad + N^{-1} l^8 P_8 \left(\frac{1}{2} N \int_0^t \mu + \delta z(s) ds \right) \\ y_2(t) &= y_2(0) + N^{-1} l^5 P_5 \left(N \beta \int_0^t x_2(s) y_1(s) ds \right) - N^{-1} l^7 P_7 \left(N(\mu + \delta) \int_0^t y_2(s) ds \right). \end{aligned} \quad (4)$$

To get a diffusion approximation we will need the following well known Theorem (see for example, [1, Prop2.2.5])

Proposition 1. *Let $\{P(t), t \geq 0\}$ be a λ Poisson process. Then,*

$$\frac{P(t)}{t} \longrightarrow \lambda \quad \text{a.s., as } t \longrightarrow \infty \quad (5)$$

Therefore, we have as a consequence of (5), if $\lambda = 1$ then

$$\frac{P(Nt)}{N} \longrightarrow t \quad \text{a.s., as } N \longrightarrow \infty. \quad (6)$$

And also the classical Central limit theorem (see, [1, Lemma2.3.4]) implies, that

$$\lim_{N \rightarrow \infty} \mathbb{P} \left(\frac{P(Nt) - Nt}{\sqrt{N}} \leq x \right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy = \mathbb{P}(B(t) \leq x),$$

where, $B(\cdot)$ is a standard Brownian motion. By this result, we have $\frac{P(Nt) - Nt}{\sqrt{N}} \Longrightarrow B(t)$ with the convergence is in law. Since, $X(t) = (x_1(t), x_2(t), y_1(t), y_2(t))$ is a normalized version of the process Y , hence, we have $\|X(t)\| < 1$. Then, the integrals $\int_0^t \beta_j(s, X_s) ds$ in (4) converge. Therefore,

$$P_j \left(N \int_0^t \beta_j(s, X_s) ds \right) \Longrightarrow \sqrt{N} B_j \left(\int_0^t \beta_j(s, X_s) ds \right) + N \int_0^t \beta_j(s, X_s) ds. \quad (7)$$

Now, using (7) into (4), this yields to the following diffusion approximation,

$$\begin{aligned} x_1(t) &= x_1(0) + l^1 \frac{1}{2} (\mu t + \delta \int_0^t z(s) ds) - l^2 \beta \int_0^t x_1(s) y_2(s) ds - l^3 \mu \int_0^t x_1(s) ds \\ &\quad + l^1 N^{-\frac{1}{2}} B_1 \left(\frac{1}{2} (\mu t + \delta \int_0^t z(s) ds) \right) - l^2 N^{-\frac{1}{2}} B_2 \left(\beta \int_0^t x_1(s) y_2(s) ds \right) - N^{-\frac{1}{2}} l^3 B_3 \left(\mu \int_0^t x_1(s) ds \right) \\ x_2(t) &= x_2(0) - (\mu + \gamma) l^2 \int_0^t x_2(s) ds + \beta l^4 \int_0^t x_1(s) y_2(s) ds + l^2 N^{-\frac{1}{2}} B_2 \left(\beta \int_0^t x_1(s) y_2(s) ds \right) \\ &\quad - l^4 N^{-\frac{1}{2}} B_4 \left((\mu + \gamma) \int_0^t x_2(s) ds \right) \\ y_1(t) &= y_1(0) - \beta l^5 \int_0^t x_2(s) y_1(s) ds + \frac{1}{2} l^8 (\mu t + \delta \int_0^t z(s) ds) - \mu l^6 \int_0^t y_1(s) ds \\ &\quad - l^5 N^{-\frac{1}{2}} B_5 \left(\beta \int_0^t x_2(s) y_1(s) ds \right) - l^6 N^{-\frac{1}{2}} B_6 \left(\mu \int_0^t y_1(s) ds \right) + l^8 N^{-\frac{1}{2}} B_8 \left(\frac{1}{2} (\mu t + \delta \int_0^t z(s) ds) \right) \\ y_2(t) &= y_2(0) + \beta l^5 \int_0^t x_2(s) y_1(s) ds - (\mu + \delta) l^7 \int_0^t y_1(s) ds \\ &\quad + l^5 N^{-\frac{1}{2}} B_5 \left(\beta \int_0^t x_2(s) y_1(s) ds \right) - l^7 N^{-\frac{1}{2}} B_7 \left((\mu + \delta) \int_0^t y_1(s) ds \right) \end{aligned}$$

where, $B_1, B_2, B_3, B_4, B_5, B_6, B_7$ and B_8 are independent standard Brownian motions. However the two processes $B_j \left(\int_0^t \beta_j(s, X_s) ds \right)$ and $\int_0^t \sqrt{\beta_j(s, X_s)} dB_j(s)$ for $j = 1, \dots, 8$, are two centered Gaussian processes which have the same covariance functions. Hence, they have the same distribution. Therefore, using this argument, the diffusion approximation process can be now written as

$$\begin{aligned}
\mathbf{X}(t) = N^{-1}\mathbf{Y}(t) = & X_0 + N^{-\frac{1}{2}}l^1 \int_0^t \sqrt{\frac{1}{2}(\mu + \delta z(s))} dB_1(s) + N^{-\frac{1}{2}}l^2 \int_0^t \sqrt{\beta x_1(s)y_2(s)} dB_2(s) \\
& + N^{-\frac{1}{2}}l^3 \int_0^t \sqrt{\mu x_1(s)} dB_3 + N^{-\frac{1}{2}}l^4 \int_0^t \sqrt{(\mu + \gamma)x_2(s)} dB_4(s) \\
& + N^{-\frac{1}{2}}l^5 \int_0^t \sqrt{\beta x_2(s)y_1(s)} dB_5(s) + N^{-\frac{1}{2}}l^6 \int_0^t \sqrt{\mu y_1(s)} dB_6(s) \\
& + N^{-\frac{1}{2}}l^7 \int_0^t \sqrt{(\mu + \delta)y_2(s)} dB_7(s) + N^{-\frac{1}{2}}l^8 \int_0^t \sqrt{\frac{1}{2}(\mu + \delta z(s))} dB_8(s) \\
& + l^1 \int_0^t \frac{1}{2}(\mu + \delta z(s)) - \mu x_1(s) - \beta y_2(s)x_1(s) ds - l^4 \int_0^t -(\mu + \gamma)y_1(s) + \beta y_2(s)x_1(s) ds \\
& + l^8 \int_0^t \frac{1}{2}(\mu + \delta z(s)) - \mu x_2(s) - \beta y_1(s)x_2(s) ds - l^7 \int_0^t -(\mu + \gamma)y_2(s) + \beta y_1(s)x_2(s) ds.
\end{aligned}$$

The obtained diffusion approximation agrees with that proved by Britton and Pardoux (see, [1, Section 2.4]) for Markovian SEIR.

The above stochastic integral equation is equivalent to the following non-linear multidimensional Itô stochastic differential equation:

$$d\mathbf{X}_t = \xi(\mathbf{X}_t, \theta)dt + \mathbf{L}(\mathbf{X}_t, \theta)dB(t), \quad X_0 = x_0, \quad (8)$$

where, B is a 8-dimensional standard Brownian motion and ξ is the vector field associated to the non-linear deterministic equation $\dot{\mathbf{X}}(t) = \xi(\mathbf{X}(t))$;

$$\xi(x) = \begin{pmatrix} \frac{1}{2}(\mu + \delta z) - \mu x_1 - \beta y_2 x_1 \\ -(\mu + \gamma)y_1 + \beta y_2 x_1 \\ \frac{1}{2}(\mu + \delta z) - \mu x_2 - \beta y_1 x_2 \\ -(\mu + \gamma)y_2 + \beta y_1 x_2 \end{pmatrix} \quad (9)$$

and L a diffusion coefficient given by

$$\mathbf{L}(x) = \frac{1}{\sqrt{N}} \begin{pmatrix} \sqrt{\frac{1}{2}(\mu + \delta z)} & -\sqrt{\beta y_2 x_1} & -\sqrt{\mu x_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{\beta y_2 x_1} & 0 & -\sqrt{(\mu + \gamma)y_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\sqrt{\beta y_1 x_2} & -\sqrt{\mu x_2} & 0 & \sqrt{\frac{1}{2}(\mu + \delta z)} \\ 0 & 0 & 0 & 0 & \sqrt{\beta y_1 x_2} & 0 & \sqrt{(\mu + \gamma)y_2} & 0 \end{pmatrix}. \quad (10)$$

thus, we have

$$\mathbf{L}\mathbf{L}^t(x) = \mathbf{\Sigma}(x) = \frac{1}{N} \begin{pmatrix} \frac{1}{2}(\mu + \delta z) + \mu x_1 + \beta y_2 x_1 & -\beta y_2 x_1 & 0 & 0 \\ -\beta y_2 x_1 & (\mu + \gamma)y_1 + \beta y_2 x_1 & 0 & 0 \\ 0 & 0 & \frac{1}{2}(\mu + \delta z) + \mu x_2 + \beta y_1 x_2 & -\beta y_1 x_2 \\ 0 & 0 & -\beta y_1 x_2 & (\mu + \gamma)y_2 + \beta y_1 x_2 \end{pmatrix}. \quad (11)$$

Finally, the diffusion process associated to the HIV spread in this case, is given by the non-linear multidimensional Itô stochastic differential equation:

$$d\mathbf{X}_t = \xi(\mathbf{X}_t, \theta)dt + \mathbf{\Sigma}^{\frac{1}{2}}(\mathbf{X}_t, \theta)dW(t), \quad (12)$$

where W is a 4-dimensional standard Brownian motion, ξ is given by (9) and Σ defined by (11). The drift function of eq. (12) is set to be identical with the deterministic behaviour of \mathbf{X} and the diffusion coefficient is chosen such that it appropriately represents the fluctuations of the trajectories around the deterministic course $\dot{\mathbf{x}}(t) = \xi(x(t))$. For the non linear drift coefficient, the identification of the deterministic course with the stochastic expectation in eq. (12) might result in a wrong diffusion coefficient (see, [17]).

For fixed x_0 , let consider, for $t \in \mathbb{R}_+$ and $x \in \mathbb{R}_+^*$ $\mathbb{P}(X(t) = x) = p(t, x)$ be the probability density of finding \mathbf{X} in state $x = (x_1(t), x_2(t), y_1(t), y_2(t))$ at given time $t \geq 0$. So the linear diffusion equation known as the Fokker-Plank equation of the forward Kolmogorov equation is verified by $p(., .)$ is of the form (see [18, 19]).

$$\frac{\partial p(t, x)}{\partial t} = -\frac{\partial}{\partial x} (\xi(X_t)p(t, x)) + \frac{1}{2} \frac{\partial^2}{\partial x^2} (\Sigma(t, x)p(t, x)) \quad (13)$$

with the initial value: $p(0, x) = \delta_{x_0 x}$ with $x_0 = (s_M(0), i_M(0), s_F(0), i_F(0))$ and δ here is the Kronecker delta.

However the goal of this paper is to illustrate an efficient method to estimate $\theta = (\beta, \gamma, \delta)$, because the quantity $R_0 = \frac{\beta}{\delta + \gamma}$ expressed in terms of model parameters is the basic reproduction number

(see, [14]), which is the most important quantity to consider when analyzing any epidemic model for an infectious disease. In particular, R_0 determines whether an epidemic can occur at all; it is used to measure the transmission potential of disease and interpret the threshold behavior of the epidemic (see for example [20]).

4. Bayesian inference

Let consider the inference for the Itô diffusion process of type (12). We assume that the process $\mathbf{X}_t = (x_1(t), x_2(t), y_1(t), y_2(t))$ will be observed at a finite integer of times. The purpose is to inference for the (unknown) parameter vector θ on the basis of partial and discrete observations on $X(t)$. Unfortunately, the exact solution of the Fokker-Plank equation eq. (13) is not available for our model. To overcome the impediment, we use Gaussian approximation of transition density for small times established by [21, §4.7.3] for the process defined by eq. (13).

Thereby, since the diffusion matrix Σ is positive definite so that the inverse of Σ exists and $\det(\Sigma) \neq 0$, then, the expression of transition probability density for small times τ can be written as:

$$p(t + \tau, x | t, x') = (2\sqrt{\pi\tau})^{-4} [\det(\Sigma(t, x))]^{-\frac{1}{2}} \exp \left\{ -\frac{1}{4\tau} \sum_{k,j=1}^4 [\Sigma_{kj}^{-1}(t, x')] [x_k - x'_k - \xi_k(t, x')] [x_j - x'_j - \xi_j(t, x')] \right\}, \quad (14)$$

where the transition density is defined as the conditional probability of a random variable X at time $t + \tau$ has sharp x under the condition that at time t it has the sharp value x' . Here, x_k (x'_k) for $k = 1, \dots, 4$ are the k^{th} component of x (x') and Σ_{kj}^{-1} is the (k, j) component of the inverse of matrix Σ with ξ and Σ are defined in (9) and (11).

In practice, it is necessary to work with a discrete version of (12), given by the Euler approximation (also called Euler-Maruyama approximation), is obtained by setting $X(0) = x_0$ and then successively

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \xi(\mathbf{X}(t), \theta)\Delta t + \Sigma^{\frac{1}{2}}(\mathbf{X}(t), \theta)\Delta B(t), \quad (15)$$

where $\Delta B(t) \sim \mathcal{N}(0, \Delta t I)$ with I is the identity matrix.

In order to perform inference on the parameter θ , we try to approximate the true transition density p_θ of the diffusion process by Gaussian distribution. This is eligible only if inter-observation times of the observed data are small. Since such a requirement is not fulfilled in our case (observation of HIV), we augment the data using the approach developed by Eraker in [5], by imputing intermediate points between each pair of observations. This method is well used to refine the observation density

in Bayesian inference (see for example [8, 7, 9]). The number m of sub-intervals between every two consecutive observations crucially influences the estimation results. A small number of intermediate time points degrades the accuracy of the Euler approximation (15) to the true posterior density and may hence cause a discretization bias. Large numbers of auxiliary time points, on the other hand, are computationally costly. Hence, the number m of subintervals should be chosen both sufficiently large and sustainably small. In general, they will be identified empirically.

From now on, X^{obs} denotes the observed data, and X^{imp} denotes the imputed (missing) data. Furthermore, to inference θ , a MCMC approach is employed to construct a Markov chain $\{\theta^{(i)}, X^{imp}\}_{i=1, \dots, L}$ of length L whose elements are samples **form joint** posterior density $\pi(\theta, X^{imp} | X^{obs})$ of parameter θ and imputed data X^{imp} **conditioning** on X^{obs} (observations). The Markov chain $\{\theta^{(i)}\}_{i=1, \dots, L}$ is regarded as a draw from the marginal density $\pi(\theta | X^{obs})$. In the sequel, $p(\cdot)$ is the prior density and $\pi(\cdot | \cdot)$ denotes the unnormalized posterior density.

4.1. Path update

We aim to approximate the posterior density

$$\pi(\theta | X_{t_1}, \dots, X_{t_M}) \propto \pi(X_{t_1}, \dots, X_{t_M} | \theta) p(\theta).$$

of parameter θ , based on observations $\{X_{t_1}, \dots, X_{t_M}\}$ of the diffusion process $(X_t, t \geq 0)$. Since the diffusion process possesses the Markov property, the likelihood of θ factorises as

$$\pi(X_{t_1}, \dots, X_{t_M} | \theta) = \pi(X_{t_1} | \theta) \prod_{i=2}^M \pi(X_{t_i} | X_{t_{i-1}}, \theta). \quad (16)$$

As motivated above, we impute between each pair of observations $m - 1$ latent data points, thus, the time interval is divided into m sub-intervals which are not necessarily equidistant. To ensure that discretization bias is arbitrary small we put $\Delta t = \frac{t_i - t_{i-1}}{m}$ and $N = mM$ for a chosen positive integer m . Therefore $[0, T]$ is divided into $N + 1$ points $t_0 = 0 < t_1 < \dots < t_m < t_{m+1} < \dots < t_N = T$, then the diffusion process is in state x_{t_k} at time t_k which is only known on times t_j when j is an integer multiple of m and all other points $X_{t_k}, k \neq j$ are treated as missing data.

Let denote by \hat{X} the $4 \times (N + 1)$ matrix obtained by stacking all elements of augmented data (observed and missing), that is

$$\hat{X} = \begin{pmatrix} \hat{x}_{1t_0} & \hat{x}_{1t_1} & \cdots & \hat{x}_{1t_m} & \hat{x}_{1t_{m+1}} & \cdots & \hat{x}_{1t_N} \\ \hat{x}_{2t_0} & \hat{x}_{2t_1} & \cdots & \hat{x}_{2t_m} & \hat{x}_{2t_{m+1}} & \cdots & \hat{x}_{2t_N} \\ \hat{y}_{1t_0} & \hat{y}_{1t_1} & \cdots & \hat{y}_{1t_m} & \hat{y}_{1t_{m+1}} & \cdots & \hat{y}_{1t_N} \\ \hat{y}_{2t_0} & \hat{y}_{2t_1} & \cdots & \hat{y}_{2t_m} & \hat{y}_{2t_{m+1}} & \cdots & \hat{y}_{2t_N} \end{pmatrix}.$$

Let \hat{X}_i denotes the i^{th} column of \hat{X} (if i is a multiple of m , \hat{X}_i is an observed data).

Conditioning on the first observation, the joint posterior density is given by:

$$\pi(\hat{X}, \theta) \propto \prod_{i=1}^N \pi(\hat{X}_i | \theta) p(\theta), \quad (17)$$

where

$$\pi(\hat{X}_i | \theta) = |\Sigma_{i-1}^{-1}|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left[\hat{X}_i - \hat{X}_{i-1} - \xi(\hat{X}_{i-1}, \theta) \Delta t \right]^t (\Sigma_{i-1}^{-1} \Delta t) \left[\hat{X}_i - \hat{X}_{i-1} - \xi(\hat{X}_{i-1}, \theta) \Delta t \right] \right\} \quad (18)$$

The objective is to obtain a sequence of Monte Carlo samples $\{\theta^{(i)}\}_{i=1}^L$. Which is a sample from the marginal posterior density, since $\{\theta^{(i)}\}_{i=1}^L$ is implicitly a sample from the marginal posterior $\pi(\theta | \hat{X}^{obs})$. Because the number of unobservable (missing data and parameter) is large, it is not possible to obtain independent samples of these quantities directly from (17); for this reason we use Gibbs sampler with a block strategy instead of single site to overcome the poor mixing due to high correlation amongst the latent data (see, [22]). The block strategies are based on updating a block of elements, a subset of the imputed data, at each iteration instead of updating the whole path. Applying these methods the convergence results are often achieved.

In the following, we propose that the latent data are updated in blocks of size m . For this, let consider times t_j and t_{j+} , where j is an integer multiple of m and $j^+ = j + m$, the corresponding observation \hat{X}^j and \hat{X}^{j+} which treated as fixed (observed points), the full conditional **density** for the latent path in (t_j, t_{j+}) is

$$\pi(\hat{X}^{j+1}, \dots, \hat{X}^{j^+-1} | \hat{X}^j, \hat{X}^{j^+}, \theta) \propto \prod_{i=j}^{j^+-1} \pi(\hat{X}^{i+1} | \hat{X}^i, \theta). \quad (19)$$

The first step in Gibbs sampler involves to appropriately perform the path update. Direct sampling from the posterior distribution of the latent data given the observed data and parameters is not possible in our case. Hence we use a Metropolis-Hastings (M-H) algorithm (see, [23]) for the general implementation of this step.

At each iteration, as a first step in path update, we choose a block of size m that is, a time interval $[t_j, t_{j+}]$ in whose the path interior is to be updated. Having decided about the block update strategy, we denote by X^* a simulated path, which is a candidate to update \hat{X} . We simulate each column of the block $X_{(t_j, t_{j+})}^* = \{X_{t_j}^*, \dots, X_{t_{j+}}^*\}$ with proposal density

$$q(X_{k+1}^* | X_k^*, \theta) = \mathcal{N}(X_k^* + \mu(X_k^*, \theta) \Delta t_k, \Sigma(X_k^*, \theta) \Delta t_k), \quad (20)$$

for $k = t_{j+1}, \dots, t_{j+}-2$, with $X_{t_j}^* = \hat{X}_{t_j}$ and $X_{t_{j+}}^* = \hat{X}_{t_{j+}}$.

We accept the new candidate of the path $X_{(t_j, t_{j+})}^*$ with probability

$$\alpha(X_{(t_j, t_{j+})}^*, \hat{X}_{(t_j, t_{j+})}) = 1 \wedge \frac{\pi(X_{t_{j+}}^* | X_{t_{j+}-1}^*, \theta)}{\pi(\hat{X}_{t_{j+}} | \hat{X}_{t_{j+}-1}, \theta)}. \quad (21)$$

4.2. Posterior distribution of model parameters

The last step in Gibbs sampler is the update of parameter $\theta = (\beta, \delta, \gamma)$ conditioning on its current state and the augmented data. Due to the form of **the likelihood** function derived in eq. (16), a family of independent gamma distributions is seen as a natural set of conjugate priors in the context of epidemic model, where the parameters are positive (see, [24]). This choice of prior distributions is convenient in terms of Bayesian inference due to conjugacy (see, [10]). Furthermore, the flexibility of the gamma distribution means that it is frequently used in practice as a prior distribution for rate parameters in epidemic models (see, [12]). When this type of priors is considered, the application of Bayes formula leads to a shifted Generalized Inverse Gaussian GIG as posterior distribution. In the following proposition, we prove this claim, for which the proof is left to Appendix appendix A.

Proposition 2. *If β , δ and γ follow independent gamma distributions: $\pi(\beta) \propto \Gamma(m_\beta, \lambda_\beta)$, $\pi(\delta) \propto$*

$\Gamma(m_\delta, \lambda_\delta)$ and $\pi(\gamma) \propto \Gamma(m_\gamma, \lambda_\gamma)$, then

$$\pi(\beta - C_\beta | \hat{X}, \delta, \gamma) \propto \beta^{m_\beta - N - 1} \exp \left\{ -\frac{1}{2} \left[\frac{M_\beta}{\beta} + M'_\beta \beta \right] \right\}; \quad (22)$$

$$\pi(\delta - C_\delta | \hat{X}, \beta, \gamma) \propto \delta^{m_\delta - N - 1} \exp \left\{ -\frac{1}{2} \left[\frac{M_\delta}{\delta} + M'_\delta \delta \right] \right\}; \quad (23)$$

$$\pi(\gamma - C_\gamma | \hat{X}, \delta, \beta) \propto \gamma^{m_\gamma - N - 1} \exp \left\{ -\frac{1}{2} \left[\frac{M_\gamma}{\gamma} + M'_\gamma \gamma \right] \right\}, \quad (24)$$

where the constants C_k, M_k and M'_k for $k \in \{\beta, \delta, \gamma\}$ are calculated from data and they are given in Appendix A.

5. Implementation and simulation

To implement the described estimation procedures given here, we adopt the following algorithm, which summarises our strategy simulation.

5.1. Algorithm

- i- Initialize all unknown: $\theta = (\beta, \delta, \gamma)$ and the path \hat{X}_i using linear interpolation between observed values of X_i .
- ii- At iteration h
 - a- Choose j multiple of m , set $j^+ = j + m$, that is a time interval $[t_j, t_{j^+}]$ in whose the path interior is to be updated.
 - b- Set $X_{t_j}^* = \hat{X}_{t_j}$ and $X_{t_{j^+}}^* = \hat{X}_{t_{j^+}}$. Simulate each column of the block $X_{(t_j, t_{j^+})}^* = \{X_{t_j}^*, \dots, X_{t_{j^+}}^*\}$ with proposal density given in (20)
 - c- Accept $X_{(t_j, t_{j^+})}^*$ with probability (21)
- iii- Update β^h according to (22)
- iv- Update δ^h according to (23)
- v- Update γ^h according to (24)
- vi- Increase the value of h and return to step ii-

5.2. Simulation

The MCMC scheme is applied to the diffusion stochastic epidemic model with known parameters. The observations $\{X_k, k = 1, \dots, n\}$ comes from the simulation of the original Markov chain $(X_t, t \geq 0)$,

which defined by its transitions rates given by (1). The exact Gillespie algorithm established by [25], is adapted to our situation with true values of the model parameters β , δ and γ as given in the Table 1 (see Appendix B for details). To illustrate our methodology, the MCMC sampler is run for 20000 iterations of our algorithm, with $m = 1$, $m = 2$ and $m = 5$ (number of missing data). Note that the algorithm is coded in Matlab and executed on a laptop i5. Figure 2 gives the plots of the MCMC chains. Looking at Table 1 and Figure 3 which summarizes respectively the posterior mean with variance and the simulated posterior densities. We see that the estimates are close to the true values as the number of latent data increases; therefore, for greater values of m there is a less improvement, but it leads to a heavy computational cost. The histograms in Figure 3 reveal that the convergence of the algorithm towards a Gaussian distribution as limit. In real application (see Table 2), it should be better to choose m as large as possible, to keep the balance between the computational cost and goodness of fit. A suitable value of m should be considered, as suggested by Eraker [5] by starting with small m and to subsequently increase this number after convergence of the Markov chain has been achieved. This procedure is pursued until further increases of m have negligible impact on the estimation results.

5.3. Real application: Morocco's case

We consider the case of Morocco. The database considered here, is the constructed data of HIV/AIDS in dynamic Morocco, which contains susceptible and infected males and females and AIDS cases, given in our recent work (see [14]). This database is obtained by combining statistics and data from Moroccan High Commission of Planing HCP [26], Ministry of Public Health with the support of UNAIDS [27, 28], the database of Wold bank [29] and the SPECTRUM software [30].

The results presented in Table 2, give an estimation parameters. The estimated basic reproductive number \hat{R}_0 is bigger than one, which means that the number of secondary cases which an infected person would produce in a completely susceptible Moroccan population is more than two, in other words, the epidemic will keep growing in Morocco.

Figure 4 and Figure 5 give the exact proportions of observed number the infected males and females in the active sexually Moroccan population from 1985 to 2014, and the projection of these subpopulation from 2015 up to 2020 using the estimated value of parameters given in Table 2. As illustrated in Figure 4 and Figure 5, for example, the exact proportion of the observed number of infected females in Morocco during the years 2015 and 2016 are 4.757×10^{-4} and 4.833×10^{-4} respectively, well the predicted values corresponding to these proportions are respectively 4.703×10^{-4}

and 4.842×10^{-4} . Since the predicted results are coherent with the real results in the case of Morocco database, hence, the model and the approach presented in this paper can be useful and adapted to study and clarify the dynamic transmissions of HIV or similar epidemic in a heterosexual population.

6. Conclusion

This paper is concerned with Bayesian estimation of HIV/AIDS dynamic parameters in a closed heterosexual population. A stochastic diffusion approximation is adopted. We are essentially concerned with the Bayesian analysis of non-linear, discretely observed stochastic diffusion. The MCMC methods has been used to infer the estimation of the parameters. We have proposed Gamma distribution as prior, which is an adequate choice for the rate parameters in epidemic models (see [10]) as mentioned in the introduction. Accordingly, the results presented for both simulation and real data are efficient.

The inference, established here, can be likely applicable to some other compartmental models. This work represents an alternative method to overcome discretely observed process considered in the work by AbouBakre and ElMaroufy [14]. Further research will be devoted to study the dynamic of HIV in heterosexual population with random environment and general mechanism as in [2].

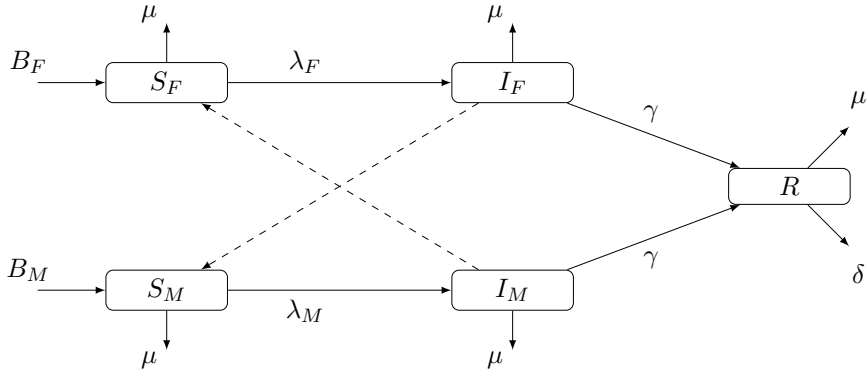


Figure 1. The model scheme. The dashed arrows indicate infection made via sexual contact only.

$m = 1$								
	true value of β	$\hat{\beta}$	true value of δ	$\hat{\delta}$	true value of γ	$\hat{\gamma}$	R_0	\hat{R}_0
mean	0.5	0.5537	0.5	0.4996	0.5	0.4976	0.5	0.5552
variance		0.1290		0.0846		0.0188		
mean	1	1.0006	0.5	0.4753	0.5	0.4983	1	1.0277
variance		0.2541		0.1150		0.0051		
mean	2	1.8343	0.5	0.5199	0.5	0.4934	2	1.8102
variance		0.3385		0.1054		0.0046		
$m = 2$								
	true value of β	$\hat{\beta}$	true value of δ	$\hat{\delta}$	true value of γ	$\hat{\gamma}$	R_0	\hat{R}_0
mean	0.5	0.5040	0.5	0.5252	0.5	0.5070	0.5	0.4883
variance		0.0903		0.0885		0.0188		
mean	1	1.0378	0.5	0.5620	0.5	0.4997	1	0.9774
variance		0.2311		0.0833		0.0042		
mean	2	2.1406	0.5	0.5177	0.5	0.5101	2	2.0827
variance		0.2999		0.0598		0.0009		
$m = 5$								
	true value of β	$\hat{\beta}$	true value of δ	$\hat{\delta}$	true value of γ	$\hat{\gamma}$	R_0	\hat{R}_0
mean	0.5	0.5008	0.5	0.5154	0.5	0.5004	0.5	0.4930
variance		0.0282		0.0896		0.0568		
mean	1	0.9990	0.5	0.5295	0.5	0.5016	1	0.9689
variance		0.2450		0.1273		0.0051		
mean	2	2.0392	0.5	0.5001	0.5	0.5047	2	2.0293
variance		0.3517		0.1001		0.0047		

Table 1. Posterior mean and variance for β , δ , γ and R_0 for $m = 1$, $m = 2$ and $m = 5$.

m		$\hat{\beta}$	$\hat{\delta}$	$\hat{\gamma}$	\hat{R}_0
1	mean	0,9619	0,2186	0,2288	2,1502
	variance	0,0790	0,0002	0,0048	
2	mean	0,9781	0,2145	0,2187	2,2578
	variance	0,0355	0,0003	0,0032	
5	mean	0,9795	0,2145	0,2166	2,2722
	variance	0,0352	0,0003	0,0031	
10	mean	0,9798	0,2145	0,2154	2,2791
	variance	0,0355	0,0003	0,0032	

Table 2. Posteriors mean and variance for $\hat{\beta}$, $\hat{\delta}$, $\hat{\gamma}$ and \hat{R}_0 in cases of $m = 1$, $m = 2$, $m = 5$ and $m = 10$, for data of Morocco, with 20000 iterations.

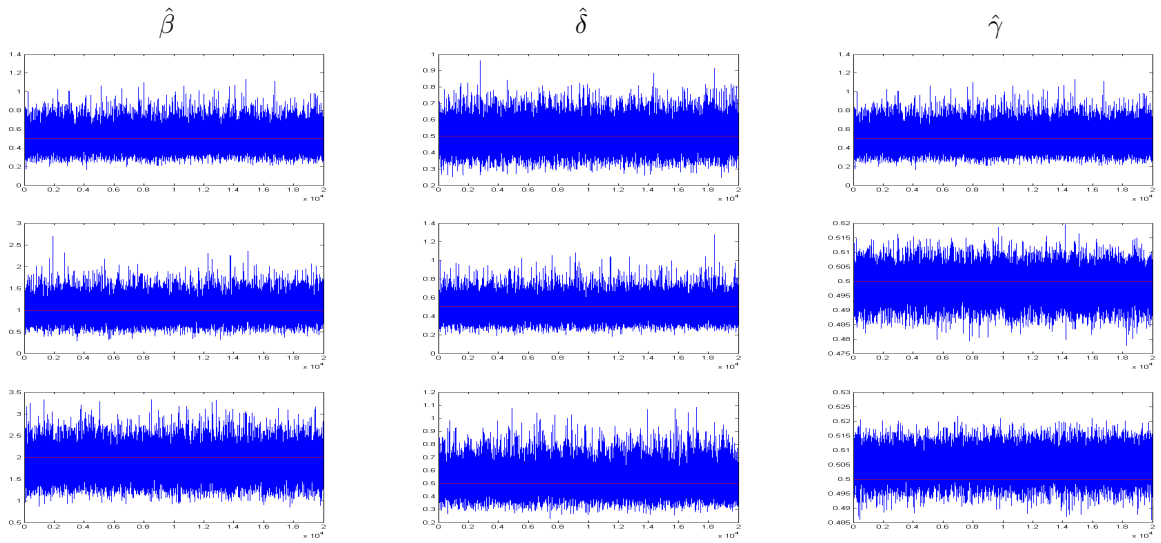


Figure 2. Trace plots of the MCMC chains for the parameters after 20000 iterations. The red lines show the true values of the parameters $\beta = \{0.5, 1, 2\}$, which correspond to $R_0 < 1, R_0 = 1, R_0 > 1$ and $\delta = \gamma = 0.5$.

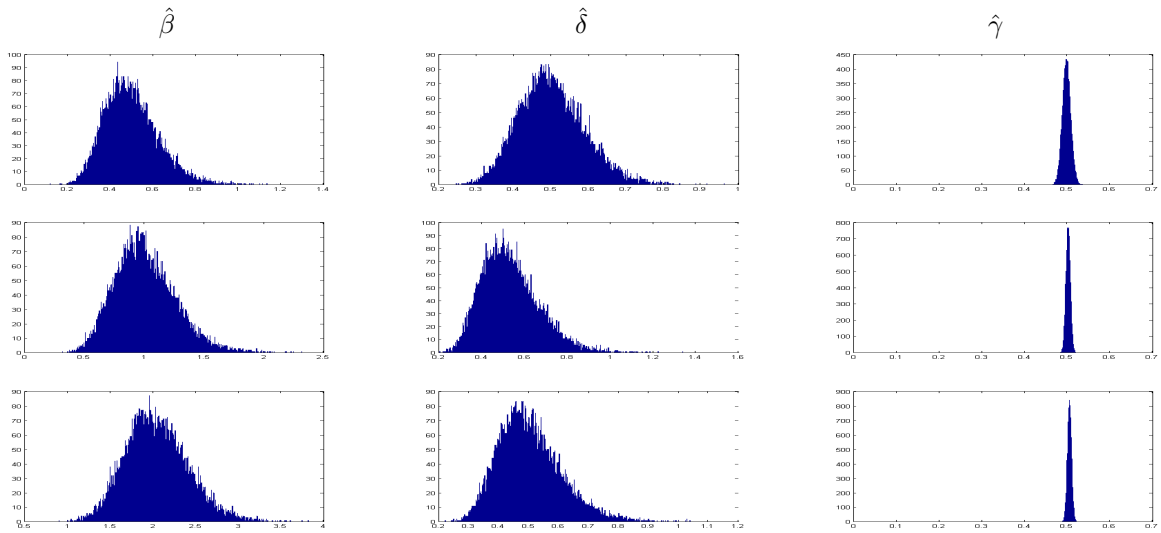


Figure 3. Posterior densities for the parameters of the model (after 20000 iterations), in three cases: $R_0 < 1$ (first row), $R_0 = 1$ (second row) and $R_0 > 1$ (third row).

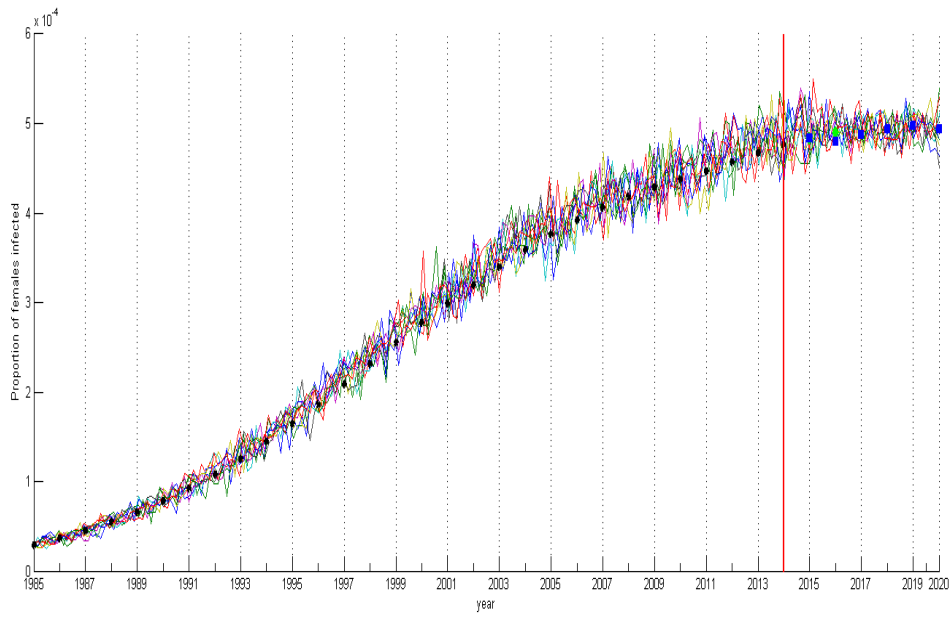


Figure 4. Plots of 10 trajectories of females infected proportion from 1985 to 2020. The black stars design the true observed data of females from 1985 up to 2014. The green diamonds design real data for 2015 and 2016. Well the blue squares gives the mean predicted value of the proportion of infected females. (For interpretation of the references to colour of these illustrations, the reader is referred to the electronic version of this paper.)

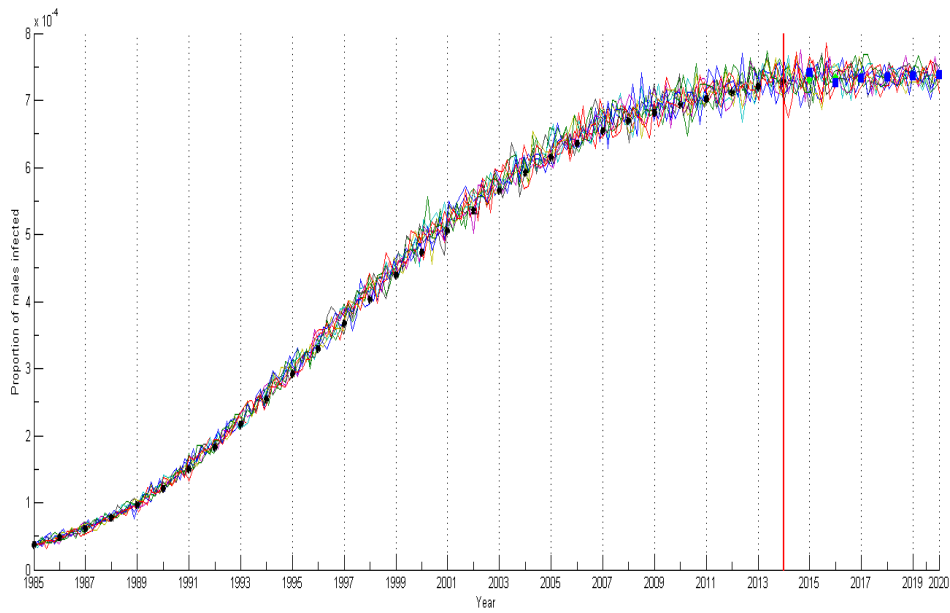


Figure 5. Plots of 10 trajectories of males infected proportion from 1985 to 2020. The black stars design the true observed data of males from 1985 up to 2014. The green diamonds design real data for 2015 and 2016. Well the blue squares gives the mean predicted value of the proportion of infected males. (For interpretation of the references to colour of these illustrations, the reader is referred to the electronic version of this paper.)

Acknowledgment

The authors would like to thank the editorial staff and the referees for the comments that helped in improving this work. Also, the first and third authors would like to thank the supporter of this work: the Germano-Moroccan Program for scientific research PMARS 3-060.

References

- [1] Tom Britton and Etienne Pardoux. *Stochastic Epidemic Models with Inference*. Mathematical Biosciences Subseries. Springer International Publishing, 2019. ISBN: 978-3-030-30899-5. DOI: [10.1007/978-3-030-30900-8](https://doi.org/10.1007/978-3-030-30900-8).
- [2] Hamid El Maroufy. “Note on the threshold theorem of a heterogeneous SIR epidemic”. In: *Stat. Probab. Lett.* 83.1 (2013), pp. 292–296. ISSN: 01677152. DOI: [10.1016/j.spl.2012.09.013](https://doi.org/10.1016/j.spl.2012.09.013).
- [3] Niels G. Becker. “On a general stochastic epidemic model”. In: *Theoretical Population Biology* 11.1 (Feb. 1, 1977), pp. 23–36. ISSN: 0040-5809. DOI: [10.1016/0040-5809\(77\)90004-1](https://doi.org/10.1016/0040-5809(77)90004-1).
- [4] Hakan Andersson and Tom Britton. *Stochastic Epidemic Models and Their Statistical Analysis*. Lecture Notes in Statistics. New York: Springer-Verlag, 2000. ISBN: 978-0-387-95050-1.
- [5] Bjørn Eraker. “MCMC Analysis of Diffusion Models With Application to Finance”. In: *Journal of Business & Economic Statistics* 19.2 (Apr. 1, 2001), pp. 177–191. ISSN: 0735-0015. DOI: [10.1198/073500101316970403](https://doi.org/10.1198/073500101316970403).
- [6] Tom Britton and Philip D. O’Neill. “Bayesian Inference for Stochastic Epidemics in Populations with Random Social Structure”. In: *Scandinavian Journal of Statistics* 29.3 (Sept. 2002), pp. 375–390. ISSN: 0303-6898, 1467-9469. DOI: [10.1111/1467-9469.00296](https://doi.org/10.1111/1467-9469.00296).
- [7] Abdelaziz Qaffou, Hamid El Maroufy, and Tewfik Kernane. “Bayesian parameter inference for stochastic SIR epidemic model with hyperbolic diffusion”. In: *Communications in Statistics - Simulation and Computation* 46.9 (Oct. 21, 2017), pp. 6907–6922. ISSN: 0361-0918. DOI: [10.1080/03610918.2016.1217013](https://doi.org/10.1080/03610918.2016.1217013).
- [8] Hamid El Maroufy et al. “Bayesian inference for nonlinear stochastic SIR epidemic model”. In: *Journal of Statistical Computation and Simulation* 86.11 (July 23, 2016), pp. 2229–2240. ISSN: 0094-9655, 1563-5163. DOI: [10.1080/00949655.2015.1107561](https://doi.org/10.1080/00949655.2015.1107561).

- [9] Chouaib Beldjoudi, Tewfik Kernane, and Hamid El Maroufy. “Bayesian inference for a susceptible-exposed-infected-recovered epidemic model with data augmentation”. In: *Mathematical Population Studies* 27.4 (Oct. 1, 2020). Publisher: Routledge .eprint: <https://doi.org/10.1080/08898480.2019.1656491>, pp. 232–258. ISSN: 0889-8480. DOI: [10.1080/08898480.2019.1656491](https://doi.org/10.1080/08898480.2019.1656491).
- [10] P. D. O’Neill and G. O. Roberts. “Bayesian inference for partially observed stochastic epidemics”. In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 162.1 (), pp. 121–129. ISSN: 1467-985X. DOI: [10.1111/1467-985X.00125](https://doi.org/10.1111/1467-985X.00125).
- [11] Kari Auranen et al. “Transmission of Pneumococcal Carriage in Families: A Latent Markov Process Model for Binary Longitudinal Data”. In: *Journal of the American Statistical Association* 95.452 (Dec. 1, 2000). Publisher: Taylor & Francis .eprint: <https://www.tandfonline.com/doi/pdf/10.1080/01621459.2000.10474301>, pp. 1044–1053. ISSN: 0162-1459. DOI: [10.1080/01621459.2000.10474301](https://doi.org/10.1080/01621459.2000.10474301).
- [12] S. Cauchemez et al. “A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data”. In: *Statistics in Medicine* 23.22 (), pp. 3469–3487. ISSN: 1097-0258. DOI: [10.1002/sim.1912](https://doi.org/10.1002/sim.1912).
- [13] George Streftaris and Gavin J Gibson. “Bayesian inference for stochastic epidemics in closed populations”. In: *Statistical Modelling* 4.1 (Apr. 1, 2004). Publisher: SAGE Publications India, pp. 63–75. ISSN: 1471-082X. DOI: [10.1191/1471082X04st065oa](https://doi.org/10.1191/1471082X04st065oa).
- [14] Abdellah AbouBakre and Hamid ElMaroufy. “Parameter inference for HIV stochastic diffusion model in closed heterosexual population”. In: *Mathematical Methods in the Applied Sciences* 41.18 (2018). .eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mma.4940>, pp. 9081–9091. ISSN: 1099-1476. DOI: <https://doi.org/10.1002/mma.4940>.
- [15] Thomas G. Kurtz. “Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes”. In: *Journal of Applied Probability* 7.1 (1970). Publisher: Applied Probability Trust, pp. 49–58. ISSN: 0021-9002. DOI: [10.2307/3212147](https://doi.org/10.2307/3212147).
- [16] Thomas G. Kurtz. “Strong approximation theorems for density dependent Markov chains”. In: *Stochastic Processes and their Applications* 6.3 (Feb. 1, 1978), pp. 223–240. ISSN: 0304-4149. DOI: [10.1016/0304-4149\(78\)90020-0](https://doi.org/10.1016/0304-4149(78)90020-0).
- [17] Christiane Fuchs. *Inference for Diffusion Processes: With Applications in Life Sciences*. Berlin Heidelberg: Springer-Verlag, 2013. ISBN: 978-3-642-25968-5.

- [18] Sabir Umarov, Marjorie Hahn, and Kei Kobayashi. *Beyond the Triangle: Brownian Motion, Ito Calculus, and FokkerPlanck Equation Fractional Generalizations*. WORLD SCIENTIFIC, Apr. 2018. ISBN: 978-981-323-091-0 978-981-323-092-7. DOI: [10.1142/10734](https://doi.org/10.1142/10734).
- [19] Johan Grasman, Onno A, and van Herwaarden. *Asymptotic Methods for the Fokker-Planck Equation and the Exit Problem in Applications*. Springer Series in Synergetics. Berlin Heidelberg: Springer-Verlag, 1999. ISBN: 978-3-540-64435-4. DOI: [10.1007/978-3-662-03857-4](https://doi.org/10.1007/978-3-662-03857-4).
- [20] Fred Brauer, Pauline van den Driessche, and J. Wu. *Mathematical Epidemiology*. Mathematical Biosciences Subseries. Berlin Heidelberg: Springer-Verlag, 2008. ISBN: 978-3-540-78910-9. DOI: [10.1007/978-3-540-78911-6](https://doi.org/10.1007/978-3-540-78911-6).
- [21] Hannes Risken and Till Frank. *The Fokker-Planck Equation: Methods of Solution and Applications*. 2nd ed. Springer Series in Synergetics. Berlin Heidelberg: Springer-Verlag, 1996. ISBN: 978-3-540-61530-9. DOI: [10.1007/978-3-642-61544-3](https://doi.org/10.1007/978-3-642-61544-3).
- [22] Ola Elerian, Siddhartha Chib, and Neil Shephard. “Likelihood Inference for Discretely Observed Nonlinear Diffusions”. In: *Econometrica* 69.4 (July 2001), pp. 959–993. ISSN: 0012-9682, 1468-0262. DOI: [10.1111/1468-0262.00226](https://doi.org/10.1111/1468-0262.00226).
- [23] Christian Robert and George Casella. *Monte Carlo Statistical Methods*. 2nd ed. Springer Texts in Statistics. New York: Springer-Verlag, 2004. ISBN: 978-0-387-21239-5.
- [24] Nikolaos Demiris and Philip D. O’Neill. “Bayesian inference for stochastic multitype epidemics in structured populations via random graphs”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67.5 (), pp. 731–745. ISSN: 1467-9868. DOI: [10.1111/j.1467-9868.2005.00524.x](https://doi.org/10.1111/j.1467-9868.2005.00524.x).
- [25] Michael Gibson and Jehoshua Bruck. “An Efficient Algorithm for Generating Trajectories of Stochastic Gene Regulation Reaction”. In: *The journal of physical chemistry A* 104.9 (1998), pp. 1876–1889.
- [26] HCP. *Population du Maroc selon l’age et le sexe*. Aug. 25, 2014. URL: <http://www.data.gov.ma/data/fr/dataset/population-du-maroc-selon-l-age-et-le-sexe/resource/ba4d6e16-5901-42c5-9aff-1e308ca266c8> (visited on 08/12/2018).
- [27] unaids. *Mise en oeuvre de la dclaration politique sur le VIH/sida, Rapport National 2014*. 2014.
- [28] unaids. *Mise en oeuvre de la dclaration politique sur le VIH/sida. Rapport National 2015*. 2015.
- [29] banquemondiale. *Maroc Data*. Aug. 12, 2018. URL: <https://donnees.banquemondiale.org/pays/maroc> (visited on 08/12/2018).
- [30] Avenir Health. *software-spectrum*. Aug. 12, 2018. URL: <http://www.avenirhealth.org/software-spectrum.php> (visited on 08/12/2018).

Appendix A. Posterior densities of parameters

The joint posterior density, conditioning on the first observation, is given by:

$$\pi(\hat{X}, \theta) \propto \prod_{i=1}^N \pi(\hat{X}_i | \theta) p(\theta), \quad (\text{A1})$$

where

$$\pi(\hat{X}_i | \theta) = |\Sigma_{i-1}^{-1}|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left[\hat{X}_i - \hat{X}_{i-1} - \xi(\hat{X}_{i-1}, \theta) \Delta t \right]^t (\Sigma_{i-1}^{-1} \Delta t) \left[\hat{X}_i - \hat{X}_{i-1} - \xi(\hat{X}_{i-1}, \theta) \Delta t \right] \right\}. \quad (\text{A2})$$

Let denote $\mathbf{T}_i = \hat{X}_{i+1} - \hat{X}_i - \xi(\hat{X}_i, \theta) \Delta t$, then,

$$\begin{aligned} \mathbf{T}_i &= (\mathbf{T}_{1i}, \mathbf{T}_{2i}, \mathbf{T}_{3i}, \mathbf{T}_{4i})^t \\ &= \begin{pmatrix} \beta x_{1i} y_{2i} + \frac{x_{1_{i+1}} - x_{1i}}{\Delta t} + \mu x_{1i} - \alpha(\mu + \delta z_i) \\ -\beta x_{1i} y_{2i} + \frac{x_{2_{i+1}} - x_{2i}}{\Delta t} + (\mu + \gamma) y_{1i} \\ \beta x_{2i} y_{1i} + \frac{y_{1_{i+1}} - y_{1i}}{\Delta t} + \mu x_{2i} - (1 - \alpha)(\mu + \delta z_i) \\ -\beta x_{2i} y_{1i} + \frac{y_{2_{i+1}} - y_{2i}}{\Delta t} + (\mu + \gamma) y_{2i} \end{pmatrix} \Delta t \\ &= \beta \begin{pmatrix} x_{1i} y_{2i} \\ -x_{1i} y_{2i} \\ x_{2i} y_{1i} \\ -x_{2i} y_{1i} \end{pmatrix} \Delta t + \begin{pmatrix} A(1)_i \\ A(2)_i \\ A(3)_i \\ A(4)_i \end{pmatrix} \Delta t. \end{aligned}$$

We calculate the product inside the exponential operator, then

$$\begin{aligned} \mathbf{T}_k(\theta)^t \Sigma_{i-1}^{-1} \mathbf{T}_k(\theta) &= (\mathbf{T}_{1i})^2 \Sigma_{11_{i-1}}^{-1} + 2\mathbf{T}_{1i} \mathbf{T}_{2i} \Sigma_{12_{i-1}}^{-1} + (\mathbf{T}_{2i})^2 \Sigma_{22_{i-1}}^{-1} \\ &\quad + (\mathbf{T}_{3i})^2 \Sigma_{33_{i-1}}^{-1} + 2\mathbf{T}_{3i} \mathbf{T}_{4i} \Sigma_{34_{i-1}}^{-1} + (\mathbf{T}_{4i})^2 \Sigma_{44_{i-1}}^{-1}, \end{aligned}$$

where

$$\Sigma_i^{-1} = \begin{pmatrix} \frac{(\mu + \gamma) y_{1i} + \beta y_{2i} x_{1i}}{\Delta_{1i}^2} & \frac{\beta y_{2i} x_{1i}}{\Delta_{1i}^2} & 0 & 0 \\ \frac{\beta y_{2i} x_{1i}}{\Delta_{1i}^2} & \frac{(1 - \alpha)(\mu + \delta z_i) + \mu x_{2i} + \beta y_{1i} x_{2i}}{\Delta_{1i}^2} & 0 & 0 \\ 0 & 0 & \frac{(\mu + \gamma) y_{2i} + \beta y_{2i} x_{1i}}{\Delta_{2i}^2} & \frac{\beta y_{2i} x_{2i}}{\Delta_{2i}^2} \\ 0 & 0 & \frac{\beta y_{2i} x_{2i}}{\Delta_{2i}^2} & \frac{\alpha(\mu + \delta z_i) + \mu x_{2i} + \beta y_{2i} x_{2i}}{\Delta_{2i}^2} \end{pmatrix},$$

where $\Delta_{1_i} = \beta C_{1_i} + C_{2_i}$ with $C_{1_i} = x_1 y_2 ((\mu + \gamma) y_1 + (\alpha(\mu + \delta z) + \mu x_1))$, $C_{2_i} = (\mu + \gamma) y_1 (\alpha(\mu + \delta z) + \mu x_1)$, and $\Delta_{2_i} = \beta C_{1'_i} + C_{2'_i}$ with $C_{1'_i} = x_2 y_1 ((\mu + \gamma) y_2 + (\alpha(\mu + \delta z) + \mu x_2))$; $C_{2'_i} = (\mu + \gamma) y_2 (\alpha(\mu + \delta z) + \mu x_2)$. Therefore, the expression becomes

$$\begin{aligned} \mathbf{T}_k(\theta)^t \Sigma_{i-1}^{-1} \mathbf{T}_k(\theta) = & (\mathbf{T}_{1_i})^2 \left(\frac{(\mu + \gamma) y_{1_i} + \beta y_{2_i} x_{1_i}}{\Delta_{1_k}^2} \right) + 2\mathbf{T}_{1_i} \mathbf{T}_{2_i} \left(\frac{\beta y_{2_i} x_{1_i}}{\Delta_{1_k}^2} \right) + (\mathbf{T}_{2_i})^2 \left(\frac{(1 - \alpha)(\mu + \delta z_i) + \mu x_{2_i} + \beta y_{1_i} x_{2_i}}{\Delta_{1_k}^2} \right) \\ & (\mathbf{T}_{3_i})^2 \left(\frac{(\mu + \gamma) y_{2_i} + \beta y_{2_i} x_{1_i}}{\Delta_{1_k}^2} \right) + 2\mathbf{T}_{3_i} \mathbf{T}_{4_i} \left(\frac{\beta y_{2_i} x_{2_i}}{\Delta_{1_k}^2} \right) + (\mathbf{T}_{4_i})^2 \left(\frac{\alpha(\mu + \delta z_i) + \mu x_{2_i} + \beta y_{2_i} x_{2_i}}{\Delta_{1_k}^2} \right) \end{aligned} \quad (\text{A3})$$

Then, after calculus, (A3) becomes

$$\mathbf{T}_k(\theta)^t \Sigma_{i-1}^{-1} \mathbf{T}_k(\theta) = \frac{1}{\beta C_{1_i} + C_{2_i}} (K_{0_i} + K_{1_i} \beta + K_{2_i} \beta^2) + \frac{1}{\beta C_{1'_i} + C_{2'_i}} (K_{0'_i} + K_{1'_i} \beta + K_{2'_i} \beta^2) \quad (\text{A4})$$

with

$$\begin{aligned} K_{0_i} &= (\mu + \gamma) y_{1_i} A(1)_i^2 \Delta t + (\alpha(\mu + \delta z_i) + \mu x_{1_i}) A(2)_i^2 \Delta t, \\ K_{1_i} &= 2(\mu + \gamma) y_{1_i} x_{1_i} y_{2_i} A(1)_i - 2x_{1_i} y_{2_i} (\alpha(\mu + \delta z_i) + \mu x_{1_i}) A(2)_i + (A(1)_i + A(2)_i)^2 \Delta t, \\ K_{2_i} &= ((\mu + \gamma) y_{1_i} (x_{1_i} y_{2_i})^2 + (\alpha(\mu + \delta z_i) + \mu x_{1_i} (x_{1_i} y_{2_i})^2)) \Delta t, \end{aligned}$$

where the constants $K_{0'_i}$, $K_{1'_i}$ and $K_{2'_i}$ are obtained from the above constants by exchanging x_{1_i}, y_{1_i} by x_{2_i}, y_{2_i} and α by $1 - \alpha$, and exchanging $A(1)_i, A(2)_i$ respectively by $A(3)_i$ and $A(4)_i$ respectively.

If $C_{1_i} = 0$ ($C_{1'_i} = 0$) that correspond to the case where we have no female (male) infected, in this situation the study is reduced to a simple SIR model see for example [7].

Let the following assumption hold true from now on: “the numbers of infected people female and male are not equal to zero” which means that $y_{1_i} \neq 0$ and $y_{2_i} \neq 0$ for all i .

In this case, if $\mathbf{p}(\beta) \propto \beta^{m_\beta - 1} \exp(-\lambda_\beta \beta)$, then we have

$$\begin{aligned} \pi(\beta | \hat{Y}_i, \delta, \gamma) \propto & \beta^{m_\beta - 1} \exp(-\lambda_\beta \beta) \prod_{i=1}^N (C_{1_i} C'_{i1})^{-\frac{1}{2}} (\beta + C_\beta)^{-1} \\ & \times \exp \left\{ -\frac{1}{2} \left[\frac{1}{\beta + C_\beta} \left(\frac{K_{0_i}}{C_{1_i}} + \frac{K_{1_i}}{C_{1_i}} \beta + \frac{K_{2_i}}{C_{1_i}} \beta^2 \right) + \frac{1}{\beta + C_\beta} \left(\frac{K_{0'_i}}{C'_{i1}} + \frac{K_{1'_i}}{C'_{i1}} \beta + \frac{K_{2'_i}}{C'_{i1}} \beta^2 \right) \right] \right\}, \end{aligned}$$

where $C_\beta = \min_{1 \leq i \leq N} \left\{ \frac{C_{2_i}}{C_{1_i}}, \frac{C_{2'_i}}{C_{1'_i}} \right\}$. Let use the same notation for the new constants $K_{ij} = \frac{K_{ij}}{C_{1_i}}$ and

$K'_{ij} = \frac{K'_{ij}}{C'_{1i}}$, then we have the following result

$$\pi(\beta|\hat{Y}_i, \delta, \gamma) \propto h(\beta)^{m_\beta - N - 1} \exp \left\{ -\frac{1}{2} \left[\frac{M_\beta}{h(\beta)} + M'_\beta h(\beta) \right] \right\}, \quad (\text{A5})$$

where $h(\beta) = \beta + C_\beta$, $M_\beta = \sum_{i=0}^N (K_{0i} + K_{2i}C_\beta^2 - K_{1i}C_\beta) + \sum_{i=0}^N (K'_{0i} + K'_{2i}C_\beta^2 - K'_{1i}C_\beta)$ and $M'_\beta = 2\lambda_\beta + \sum_{i=1}^N K_{2i} + K'_{2i}$.

Since C_β is a positive constant which not depends on β , however, using a transfer theorem, (A5) becomes

$$\pi(\beta - C_\beta|\hat{Y}_i, \delta, \gamma) \propto \beta^{m_\beta - N - 1} \exp \left\{ -\frac{1}{2} \left[\frac{M_\beta}{\beta} + M'_\beta \beta \right] \right\}. \quad (\text{A6})$$

Hence, the posterior distribution of $\beta - C$ is proportional to a Generalized inverse Gaussian distribution (GIG).

Using the similar technique of decomposition on (A4), we construct the posterior distribution of the parameter δ and we have:

$$\mathbf{T}_i(\theta) = \delta \begin{pmatrix} -\alpha z_i \\ 0 \\ -(1-\alpha)z_i \\ 0 \end{pmatrix} \Delta t + \begin{pmatrix} \frac{x_{1_{i+1}} - x_{1_i}}{\Delta t} + \mu x_{1_i} - \alpha \mu \\ \frac{y_{1_{i+1}} - y_{1_i}}{\Delta t} + (\mu + \gamma) y_{1_i} \\ \frac{x_{2_{i+1}} - x_{2_i}}{\Delta t} + \mu x_{2_i} - \mu(1-\alpha) \\ \frac{y_{2_{i+1}} - y_{2_i}}{\Delta t} + (\mu + \gamma) y_{2_i} \end{pmatrix} \Delta t = \delta \begin{pmatrix} -\alpha z_i \\ 0 \\ -(1-\alpha)z_i \\ 0 \end{pmatrix} \Delta t + \begin{pmatrix} F(1)_i \\ F(2)_i \\ F(3)_i \\ F(4)_i \end{pmatrix} \Delta t,$$

if $\mathbf{p}(\delta) \propto \delta^{m_\delta - 1} \exp(-\lambda\delta)$, then, by a similar calculus, we have

$$\pi(\delta - C_\delta|\hat{Y}_i, \gamma, \beta) \propto \delta^{m_\delta - N - 1} \exp \left\{ -\frac{1}{2} \left[\frac{M_\delta}{\delta} + M'_\delta \delta \right] \right\},$$

where $M_\delta = \sum_{i=0}^N (E_{0i} + E_{2i}C_\delta^2 - E_{1i}C_\delta) + \sum_{i=0}^N (E'_{0i} + E'_{2i}C_\delta^2 - E'_{1i}C_\delta)$,

$$C_\delta = \min_{1 \leq i \leq N} \left\{ \frac{C_{2i}}{C_{1i}}, \frac{C'_{2i}}{C'_{1i}} \right\} \text{ and } M'_\delta = 2\lambda_\delta + \sum_{i=1}^N E_{2i} + E'_{2i} \text{ with:}$$

$$\begin{aligned}
E_i &= C_{1i}\delta + C_{2i} \\
&= \alpha z_i ((\mu + \gamma)y_{1i} + \beta x_{1i}y_{2i})\delta + (\mu + \gamma)y_{1i}(\alpha\mu + \mu x_{1i} + \beta x_{1i}y_{2i}) + \beta x_{1i}y_{2i}(\alpha\mu + \mu x_{1i}), \\
E_{0i} &= ((\mu + \gamma)y_{1i} + \beta x_{1i}y_{2i})F(1)_i^2 + 2\beta x_{1i}y_{2i}F(1)_iF(2)_i + (\alpha\mu + \mu x_{1i} + \beta x_{1i}y_{2i})F(2)_i^2, \\
E_{1i} &= \alpha z_i (F(2)_i^2 - 2(F(1)_i((\mu + \gamma)y_{1i} + \beta x_{1i}y_{2i}) + F(2)_i\beta x_{1i}y_{2i})), \\
E_{2i} &= \alpha^2 z_i^2 ((\mu + \gamma)y_{1i} + \beta x_{1i}y_{2i}),
\end{aligned}$$

where the constants with prime symbol ($'$) are obtained from the above constants by exchanging x_{1i}, y_{1i} by x_{2i}, y_{2i} and α by $1 - \alpha$ respectively, and exchanging $F(1)_i, F(2)_i$ respectively by $F(3)_i$ and $F(4)_i$.

For the parameter γ , we use the same techniques, then, the component inside exponential operator in function of the parameter γ can be written as:

$$\mathbf{T}_i(\theta) = \gamma \begin{pmatrix} 0 \\ y_{1i} \\ 0 \\ y_{2i} \end{pmatrix} \Delta t + \begin{pmatrix} \frac{x_{1_{i+1}} - x_{1i}}{\Delta t} + \mu x_{1i} - \alpha(\mu + \delta z) \\ \frac{y_{1_{i+1}} - y_{1i}}{\Delta t} + \mu y_{1i} \\ \frac{x_{2_{i+1}} - x_{2i}}{\Delta t} + \mu x_{2i} - (1 - \alpha)(\mu + \delta z) \\ \frac{y_{2_{i+1}} - y_{2i}}{\Delta t} + \mu y_{2i} \end{pmatrix} \Delta t = \gamma \begin{pmatrix} 0 \\ y_{1i} \\ 0 \\ y_{2i} \end{pmatrix} \Delta t + \begin{pmatrix} B(1)_i \\ B(2)_i \\ B(3)_i \\ B(4)_i \end{pmatrix} \Delta t.$$

So, if $\mathbf{p}(\gamma) \propto \gamma^{m_\gamma - 1} \exp(-\lambda_\gamma \gamma)$ then, using the same technique, we have after tedious calculus:

$$\pi(\gamma - C_\gamma | \hat{Y}_i, \delta, \beta) \propto \gamma^{m_\gamma - N - 1} \exp \left\{ -\frac{1}{2} \left[\frac{M_\gamma}{\gamma} + M'_\gamma \gamma \right] \right\},$$

$$\text{where } M_\gamma = \sum_{i=0}^N (D_{0i} + D_{2i}C_\gamma^2 - D_{1i}C_\gamma) + \sum_{i=0}^N (D'_{0i} + D'_{2i}C_\gamma^2 - D'_{1i}C_\gamma), \quad C_\gamma = \min_{1 \leq i \leq N} \left\{ \frac{C_{2i}}{C_{1i}}, \frac{C'_{2i}}{C'_{1i}} \right\}$$

and $M'_\gamma = 2\lambda_\gamma + \sum_{i=1}^N D_{2i} + D'_{2i}$ with:

$D_i = C_{1i}\gamma + C_{2i} = \mu y_{1i}(\alpha(\mu + \delta z) + \mu x_{1i} + \beta x_{1i}y_{2i})\gamma + (\mu y_{1i} + \beta x_{1i}y_{2i})(\alpha(\mu + \delta z) + \mu x_{1i}) + \beta x_{1i}y_{2i}$, $D_{0i} = (\mu y_{1i} + \beta x_{1i}y_{2i})B(1)_i^2 + \beta x_{1i}y_{2i}(2B(1)_i^2 + B(1)_iB(2)_i)$, $D_{1i} = 2y_{1i}(B(1)_i + B(2)_i) + B(1)_i^2$, and $D_{2i} = y_{1i}(\mu y_{1i} + \beta x_{1i}y_{2i})$. The remaining constants D'_{0i} , D'_{1i} , D'_{2i} and D'_i are obtained from the above constants by exchanging x_{1i}, y_{1i} by x_{2i}, y_{2i} and α by $1 - \alpha$ respectively, and exchanging $B(1)_i, B(2)_i$ respectively by $B(3)_i$ and $B(4)_i$.

Appendix B. Gillespie's algorithm

Gillespie's algorithm as given in [25], is an exact simulation algorithm for continuous-time pure Markov jump processes. In our case the implementation of the algorithm is based on two main steps. The first is the draw of the waiting time τ until the occurrence of the next event, which is exponentially distributed $\exp(\lambda)$ with rate $\lambda = \sum_{k=1}^{11} \lambda_k$. We have the rates $\lambda_1 = \alpha(\mu N + \delta(Z + 1))$, $\lambda_2 = (1 - \alpha)(\mu N + \delta(Z + 1))$ for birth of new susceptible, $\lambda_3 = \beta \frac{I_M S_F}{S_M + I_M}$, $\lambda_4 = \beta \frac{I_F S_M}{S_F + I_F}$ infection of a susceptible, $\lambda_5 = \mu S_F$, $\lambda_6 = \mu S_F$ for death of a susceptible, $\lambda_7 = \mu I_F$, $\lambda_8 = \mu I_F$ for death of an infected, $\lambda_9 = \gamma I_M$, $\lambda_{10} = \gamma I_F$ recovery of an infected and $\lambda_{11} = \mu Z$ death of a recovered. In the second step, events are randomly selected according to the probabilities $p_k = \frac{\lambda_k}{\lambda}$. The algorithm is implemented, until stopping criteria is fulfilled, as follows:

- i- At $t = 0$, set initial population numbers $(S_f(0), I_f(0), S_m(0), I_m(0)) = (s_f, i_f, s_m, i_m)$.
- ii- Choose τ from the exponential distribution $\exp(\lambda)$.
- iii- Choose $k = 1, \dots, 11$ according to the probabilities $p_k = \frac{\lambda_k}{\lambda}$.
- iv- Change the number of individuals to reflect the event k drawn in step **iii**-.

Update time $t \leftarrow t + \tau$ and go to step **ii**-.

Appendix C. Acceptance probability

The probability acceptance $\alpha(X_{(t_j, t_{j+})}^*, \hat{X}_{(t_j, t_{j+})})$ to accept a candidate path $X_{(t_j, t_{j+})}^*$ in favor of $\hat{X}_{(t_j, t_{j+})}$ (j is a multiple of m) is given by:

$$\alpha(X_{(t_j, t_{j+})}^*, \hat{X}_{(t_j, t_{j+})}) = 1 \wedge \frac{\pi \left(X_{(t_j, t_{j+})}^*, \hat{X}_{-(t_j, t_{j+})} | \hat{X}^{obs}, \theta \right) q \left(\hat{X}_{(t_j, t_{j+})} | X_{(t_j, t_{j+})}^*, \hat{X}_{-(t_j, t_{j+})}, \hat{X}^{obs}, \theta \right)}{\pi \left(\hat{X}_{(t_j, t_{j+})}, \hat{X}_{-(t_j, t_{j+})} | \hat{X}^{obs}, \theta \right) q \left(X_{(t_j, t_{j+})}^* | \hat{X}_{(t_j, t_{j+})}, \hat{X}_{-(t_j, t_{j+})}, \hat{X}^{obs}, \theta \right)}, \quad (C1)$$

where $\hat{X}_{-(t_j, t_{j+})} = \hat{X} \setminus \hat{X}_{(t_j, t_{j+})}$. Due to the Markov property, we obtain

$$\begin{aligned} \frac{\pi \left(X_{(t_j, t_{j+})}^*, \hat{X}_{-(t_j, t_{j+})} | \hat{X}^{obs}, \theta \right)}{\pi \left(\hat{X}_{(t_j, t_{j+})}, \hat{X}_{-(t_j, t_{j+})} | \hat{X}^{obs}, \theta \right)} &= \prod_{k=t_j+1}^{t_{j+}-1} \frac{\pi \left(X_{k+1}^* | X_k^*, \theta \right)}{\pi \left(\hat{X}_{k+1} | \hat{X}_k, \theta \right)} \\ &= \prod_{k=t_j+1}^{t_{j+}-1} \frac{p_\theta \left(\Delta_{t_k}, X_{k+1}^* | X_k^* \right)}{p_\theta \left(\Delta_{t_k}, \hat{X}_{k+1} | \hat{X}_k \right)}. \end{aligned}$$

The time step Δt_k in p_θ are now supposed to be small enough such that an approximation by Euler scheme is adopted, then p_θ may be replaced by

$$\pi\left(\hat{X}_{k+1}|\hat{X}_i, \theta\right) = \mathcal{N}\left(\hat{X}_{k+1}|\hat{X}_i + \mu(\hat{X}_i, \theta)\Delta t_k, \Sigma(\hat{X}_i, \theta)\Delta t_k\right).$$

If the density of \hat{X}_{k+1} is normal, $\hat{X}_{k+1} \sim \mathcal{N}\left(\hat{X}_{k+1}|\hat{X}_i + \mu(\hat{X}_i, \theta)\Delta t_k, \Sigma(\hat{X}_i, \theta)\Delta t_k\right)$ for $k = j, \dots, j^+ - 2$ with $\hat{X}_j = X_j$, then the proposal density for $X_{(t_j, t_{j^+})}^*$ becomes

$$q\left(X_{(t_j, t_{j^+})}^*|\hat{X}_{t_j}, \hat{X}_{t_{j^+}}, \theta\right) = \prod_{k=t_j+1}^{t_{j^+}-2} q\left(X_{k+1}^*|X_k^*, \theta\right) = \prod_{k=t_j+1}^{t_{j^+}-2} \pi\left(X_{k+1}^*|X_k^*, \theta\right)$$

The acceptance probability for the proposal $X_{(t_j, t_{j^+})}^*$ (C1) is reduced to the following expression

$$\begin{aligned} \alpha\left(X_{(t_j, t_{j^+})}^*, \hat{X}_{(t_j, t_{j^+})}\right) &= 1 \wedge \left(\prod_{k=t_j+1}^{t_{j^+}-1} \frac{\pi\left(X_{k+1}^*|X_k^*, \theta\right)}{\pi\left(\hat{X}_{k+1}|\hat{X}_k, \theta\right)} \right) \left(\prod_{k=t_j+1}^{t_{j^+}-2} \frac{\pi\left(\hat{X}_{k+1}|\hat{X}_k, \theta\right)}{\pi\left(X_{k+1}^*|X_k^*, \theta\right)} \right) \\ &= 1 \wedge \frac{\pi\left(X_{t_{j^+}}^*|X_{t_{j^+}-1}^*, \theta\right)}{\pi\left(\hat{X}_{t_{j^+}}|\hat{X}_{t_{j^+}-1}, \theta\right)}. \end{aligned}$$