



HAL
open science

Numerical convergence of some stochastic compartmental models in epidemiology

Stefana Fandresena Tsilefa, Solym Mawaki Manou-Abi, Angelo Raherinirina

► **To cite this version:**

Stefana Fandresena Tsilefa, Solym Mawaki Manou-Abi, Angelo Raherinirina. Numerical convergence of some stochastic compartmental models in epidemiology. CARI 2022, Oct 2022, Tunis, Tunisia. hal-03715840

HAL Id: hal-03715840

<https://inria.hal.science/hal-03715840v1>

Submitted on 6 Jul 2022

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.

Numerical convergence of some stochastic compartmental models in epidemiology

Stefana Fandresena TSILEFA¹, Solym MANOU-ABI², Angelo RAHERINIRINA¹

¹University of Fianarantsoa, BP 1264 Andrainjato, 301 Fianarantsoa, Madagascar

²CUFR Mayotte & IMAG Montpellier

*E-mail : tsilefastefana@gmail.com

Abstract

We propose a numerical study to analyse the convergence of some stochastic models including some parameter estimation using classical estimation procedure. We illustrate the convergence of a continuous time Markov chain, a Poisson stochastic model, to a deterministic model. Simulations highlight the differences in the speed of convergence of the stochastic models considered. We end with a discussion on some forthcoming theoretical developments.

Keywords

Stochastic epidemic model, Pure jump Markov chain, Poisson process, Itô's integral

I INTRODUCTION

Mathematical modelling of infectious diseases has been used for decades to study the mechanisms of disease spread, to predict the evolution of an epidemic, to evaluate control strategies and to help design public health interventions. Daniel Bernoulli [2] was the first to bring mathematical research to epidemiology and the seminal work of Kermack and Mc Kendrick [6], give a well-established method about epidemic patterns by dividing the population of interest by the abstract notion of compartments, defined by the health status of the pathogen in the system, the demographic or epidemiological characteristics. Recently, various extensions of the epidemics models have been proposed to describe the evolution of this disease in different countries, particularly in Madagascar and it's region, see [11, 13, 14]. In many situation, stochastic compartmental models, [3, 7] are preferred since we have uncertainties on certain model parameters and incomplete data. However, in many cases, the convergence of such stochastic models requires a given framework. In this paper, our objective is to illustrate by simulation, not only the dynamics of the stochastic models that we consider, but also to highlight the convergence depending on the population size. We also deal with the performance of some estimators of the underlined parameters.

II DETERMINISTIC BASELINE MODEL

A classical elementary and simple model that can be derived from the Kermack model is the Markovian SIR model in which the closed population is divided into Susceptible individuals S, Infectious individuals I and Recovered individuals R. The model can be formulated using

ordinary differential equation as follows for all times $t \geq 0$ with constant population size $N = S(t) + I(t) + R(t)$:

$$\begin{cases} \frac{dS(t)}{dt} = & -\beta S(t) \frac{I(t)}{N} \\ \frac{dI(t)}{dt} = & \beta S(t) \frac{I(t)}{N} - \gamma I(t) \\ \frac{dR(t)}{dt} = & \gamma I(t). \end{cases} \quad (1)$$

Several assumptions have been made in the formulation of the above equations. The population of a compartment is differentiable with respect to time and the epidemic process is deterministic. The compartments are homogeneous, which means that an individual in the population contracts the disease with a rate of infection β . A fraction equal to γ representing the average recovery/mortality rate of infectious individuals leaving that compartment per unit time to enter the recovered compartment. An important parameter of the model that determines the magnitude of the epidemic is the basic reproduction number R_0 given by $R_0 = \frac{\beta}{\gamma}$. This is the number of new infections produced by an infected individual in an intact population of susceptible individuals only.

III PURE JUMP MARKOV MODEL AND PARAMETER ESTIMATION

The analogous stochastic model of the above deterministic model is the generalized Continuous Time Markov Chain (CTMC) model constructed by considering the following events : infection I and removal R since only two types of transitions can occur (susceptible person becomes infected or an infected person recover).

3.1 The model properties

The transition probability from a state (s, i, r) to state $(s - 1, i + 1, r)$ is:

$$P_{(s,i,r) \rightarrow (s-1,i+1,r)} = \frac{\left(\frac{\beta si}{N}\right)}{\left(\frac{\beta si}{N} + \gamma i\right)}$$

and the transition probability from (s, i, r) to $(s, i - 1, r + 1)$ is :

$$P_{(s,i,r) \rightarrow (s,i-1,r+1)} = \frac{\gamma i}{\left(\frac{\beta si}{N} + \gamma i\right)},$$

and

$$P_{(s,i,r) \rightarrow (s,i,r)} = 1 - \left(\frac{\beta si}{N} + \gamma i\right).$$

The system starts in state s_0 at time unit t_0 and enters the last state s_k at t_k , as shown in the following figure (1). There are a total of k transitions in this process. At any time, all observations occur in the time interval (t_0, t) where $t \geq t_k$, and there are no transitions in the interval (t_k, t) . Consider a small time Δt , the infinitesimal generator is defined as follows:

$$P(s + j, i + k, r + l \mid s, i, r) = \begin{cases} \left(\frac{\beta si}{N}\right)\Delta t + o(\Delta t) & (j, k, l) = (-1, 1, 0) \\ (\gamma i)\Delta t + o(\Delta t) & (j, k, l) = (0, -1, 1) \\ \left(1 - \left(\frac{\beta si}{N} + \gamma I\right)\right)\Delta t + o(\Delta t) & (j, k, l) = (0, 0, 0) \end{cases} \quad (2)$$

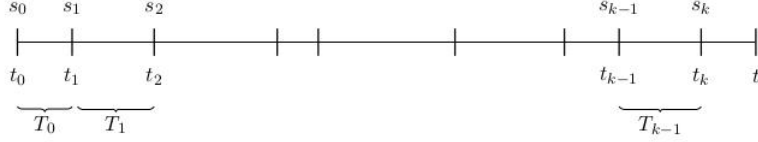


Figure 1: Illustration of a continuous time stochastic process observed between units of time t_0 and units of time t .

The transition process is essentially guided by the first two transitions and for this reason we will consider in the sequel only the first two types of transitions. Set

$$\Theta(T(k-1)) = \left(\frac{\beta S(T(k-1))I(T(k-1))}{N} + \gamma I(T(k-1)) \right).$$

The event times $T(k)$ during which an individual moves from one state to another are modelled as a renewal process with exponentially distributed increments: $\mathbb{P}(T(k) - T(k-1) > t \mid T(j), j \leq k-1) = e^{-\Theta(T(k-1))t}$.

3.2 Parameter estimation by Maximum likelihood

The estimation of parameters in a mathematical model is important since in general many parameters are unknown, see for instance [14] for some epidemiological parameters estimation. In this section, we follow the approach of [5] combined to [10]. Since the transition probabilities depend on the vector parameter $\theta = (\beta, \gamma)$, the likelihood function of the above CTMC model is the product of the transition conditional probabilities :

$$L(\theta) = p(x_0) \prod_{i=1}^n p(x_i/x_{i-1}, \theta).$$

In the case of the above an SIR model, one can split as in [10], the state transition probabilities into the two possible cases of transitions : from susceptible to infectious and infectious to recovered. Let $t_{\beta_1}, t_{\beta_2}, \dots, t_{\beta_n}$ be the set of times when there is a transition from i to $i+1$ and let $t_{\gamma_1}, t_{\gamma_2}, \dots, t_{\gamma_m}$ be the set of times when there is a transition from i to $i-1$. This means that the first transition from i to $i+1$ in the system occurred at time t_{β_1} , the first transition from $i \rightarrow i-1$ in the system occurred at time t_{γ_1} and so on. There are n transitions of $i \rightarrow i+1$ and m transitions of $i \rightarrow i-1$ and we assume that there a total of k transitions in the system from the initial state to the final state. Thus, the likelihood function is obtained as follows

$$L(\beta, \gamma) = \exp \left(- \left(\frac{\beta S(t_k)I(t_k)}{N} + \gamma I(t_k) \right) T_k \right) \prod_{a=t_{\beta_1}}^{t_{\beta_n}} \left[\frac{\beta S(a)I(a)}{N} \exp \left(- \left(\frac{\beta S(a)I(a)}{N} + \gamma I(a) \right) T_a \right) \right] \\ \times \prod_{b=t_{\gamma_1}}^{t_{\gamma_m}} \left[\frac{\beta S(b)I(b)}{N} \exp \left(- \left(\frac{\beta S(b)I(b)}{N} + \gamma I(b) \right) T_b \right) \right],$$

where $S(a) = N - I(a) - R(a)$ and $S(b) = N - I(b) - R(b)$

The maximum likelihood estimator $\hat{\beta}, \hat{\gamma}$ is obtained as follows:

$$\frac{\partial \log L(\beta, \gamma)}{\partial \beta} = 0 \text{ and } \frac{\partial \log L(\beta, \gamma)}{\partial \gamma} = 0$$

so that

$$\hat{\beta} = \frac{n}{\sum_{i=0}^k \left[\frac{S(t_i)I(t_i)T_i}{N} \right]} \text{ and } \hat{\gamma} = \frac{m}{\sum_{i=0}^k [I(t_i)T_i]} \quad (3)$$

where

$$n = \sum_{a=t_{\beta_1}}^{t_{\beta_n}} \text{ and } m = \sum_{a=t_{\gamma_1}}^{t_{\gamma_m}} .$$

3.3 Probability of an outbreak

In general for the above birth and death model, the process approaches state 0 or approaches infinity. The only absorbing state is 0. The probability of absorption for the state 0 depends on the birth probability $\lambda_i := \beta \frac{si}{N}$, death probability $\mu_i := \gamma i$ for a state (s, i) . If we take $X(0) = x_0 > 0$, it is well-known that :

$$\lim_{t \rightarrow \infty} P(X(t) = 0) = \begin{cases} 1 & \text{if } \lambda \leq \mu \\ \left(\frac{\mu}{\lambda}\right)^{x_0} & \text{if } \lambda \geq \mu \end{cases}$$

Assume that the initial number of infected individuals i_0 is small and the initial population size N is large. Applying this identity, the considered probability of absorption for the state 0 is

$$P(I(t) = 0) \simeq \begin{cases} 1 & \text{if } R_0 \leq 1 \\ \left(\frac{1}{R_0}\right)^{i_0} & \text{if } R_0 > 1 \end{cases}$$

Therefore, one obtain the following well-know approximation probability of a major outbreak

$$\text{Probability of an outbreak} \simeq \begin{cases} 0 & \text{if } R_0 \leq 1 \\ 1 - \left(\frac{1}{R_0}\right)^{i_0} & \text{if } R_0 > 1. \end{cases}$$

One observe a close agreement between the numerical values and the estimated probability of an epidemic.

3.4 Numerical estimation and convergence

We deal here with numerical simulation of the above parameters in the CTMC model. We simulate $k = n + m = 2000$ observations on a 1000 trajectory per population size and at each iteration, we calculate the maximum likelihood estimators (3) of the parameters β, γ and also R_0 . We make a comparison between their averages (blue line) and the true values (red line).

By the law of large numbers, the stochasticity disappears. Thus, the model fails to correctly reproduce the stochastic dynamics, the model converges to the deterministic model in the limit (at infinity). Indeed, we observe a decrease in the relative stochasticity of the trajectories in large population. In the model, the infection disappears much earlier than in the corresponding deterministic models. The parameter R_0 is obtained in a similar way as in the deterministic model except for the probability of an epidemic. As shown in the above figure, there is a probability where R_0 is less than 1 i.e. there is no epidemic.

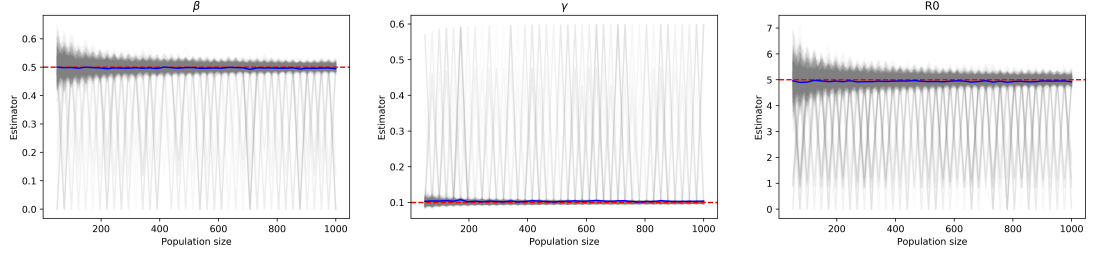


Figure 2: Estimation of the parameters $\beta = 0.50$, $\gamma = 0.10$ and $R_0 = 5$ for the CTMC model according to time-varying population size N . The red line represents the parameters and the blue line represents the mean of the estimate per population size

IV EQUIVALENT SIR STOCHASTIC DIFFERENTIAL EQUATION MODEL

There is a need to investigate other approaches that include stochastic noise that does not vanish for large populations. The idea is to form a model of stochastic differential equation of the SIR model from a discrete time markov chain (DTMC) using the procedure developed by [3, 4]. This method considers $S(t)$, $I(t)$, $R(t)$ as discrete random variables in $[0, N]$ and approximately normally distributed. Set $\Delta X(t) = (\Delta S, \Delta I)^T$. We have in a small time step Δt ,

$$E(\Delta X(t)) = \begin{pmatrix} -\frac{\beta}{N}SI \\ \frac{\beta}{N}SI - \gamma I \end{pmatrix} \Delta t.$$

and the covariance matrix

$$V(\Delta X(t)) = \begin{pmatrix} \frac{\beta}{N}SI & -\frac{\beta}{N}SI \\ -\frac{\beta}{N}SI & \frac{\beta}{N}SI + \gamma I \end{pmatrix} \Delta t.$$

The random vector $X(t + \Delta t)$ can be approximated as follows:

$$X(t + \Delta t) = X(t) + \Delta X(t) \simeq X(t) + E(\Delta X(t)) + \sqrt{V(\Delta X(t))} \quad (4)$$

Since the covariance matrix is symmetric and positive definite, it has a unique square root $B\sqrt{\Delta t} = \sqrt{V}$ where $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$. For sufficiently smooth coefficients, the solution $X(t)$ of (4) converges to the solution of the following system of

$$\begin{cases} dS(t) = -\beta S(t)\frac{I(t)}{N}dt + B_{11}dW_1 + B_{12}dW_2 \\ dI(t) = \beta S(t)\frac{I(t)}{N}dt - \gamma I(t)dt + B_{21}dW_1 + B_{22}dW_2 \end{cases} \quad (5)$$

with $R(t) = N - S(t) - I(t)$; W_1, W_2 are independent copies of a Brownian motion[1]. The mean of the above SDE is the deterministic model in a small time, but the noise builds up a variance that allows for stochastic dynamics even in large population as shown in the following figure (3) for 100 trajectories.

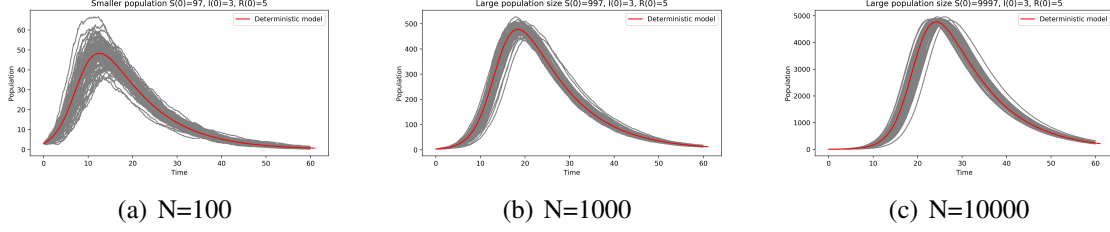


Figure 3: Simulated trajectories of the equivalent SIR stochastic differential equation model (infected compartment) with different population sizes and $I(0) = 3$.

This model allows to consider the stochastic aspect of the dynamics in a large population. This stochasticity is decreasing with the increase of the population size.

V MODELING EPIDEMIC DYNAMICS WITH POISSON PROCESSES

5.1 The model

Consider again a close population of N individuals. Assume that, while infectious, an individual has infectious contacts according to a Poisson process. Each contact is with an individual chosen uniformly at random from the rest of the population, and if the contacted individual is susceptible he/she becomes infected, otherwise the infectious contact has no effect, see [7]. More precisely,

- Infection of a susceptible event happen at rate :

$$\beta \frac{S(t)I(t)}{N}$$

- Recovered of an infective event happen at rate

$$\gamma I(t)$$

Hence we have the following equations with P_{inf} and P_{rec} two standard independent Poisson processes :

$$\begin{cases} S(t) = S(0) - P_{inf} \left(\beta \int_0^t \frac{S(s)I(s)}{N} ds \right) \\ I(t) = I(0) + P_{inf} \left(\beta \int_0^t \frac{S(s)I(s)}{N} ds \right) - P_{rec} \left(\gamma \int_0^t I(s) ds \right) \\ R(t) = R(0) + P_{rec} \left(\gamma \int_0^t I(s) ds \right) \end{cases} \quad (6)$$

5.2 Study of convergence

Let $A_N(t) = \frac{S(t)I(t)}{N}$. By taking the expectation in each equation of model (6), we have for time horizon $T > 0$:

$$\beta = \frac{S(0) - \mathbb{E}S(T)}{\mathbb{E} \int_0^T A_N(s) ds}, \quad \gamma = \frac{R(0) - \mathbb{E}R(T)}{\mathbb{E} \int_0^T I(s) ds},$$

and for $R(0) = 0$ one can derive an estimation of β and γ using classical estimation procedure, for instance the least squares method or method of moments during a given time horizon T . Then :

$$\hat{\beta} = \sum_{k=1}^T \sum_{j=1}^T \frac{S(0)-S(k)}{T \times A_N(j)}, \quad \hat{\gamma} = \sum_{k=1}^T \sum_{j=1}^T \frac{R(k)}{T \times I(j)} \quad (7)$$

We simulate 1000 trajectory model with a time horizon $T=50$ per population size and at each iteration, we calculate the estimators (7).

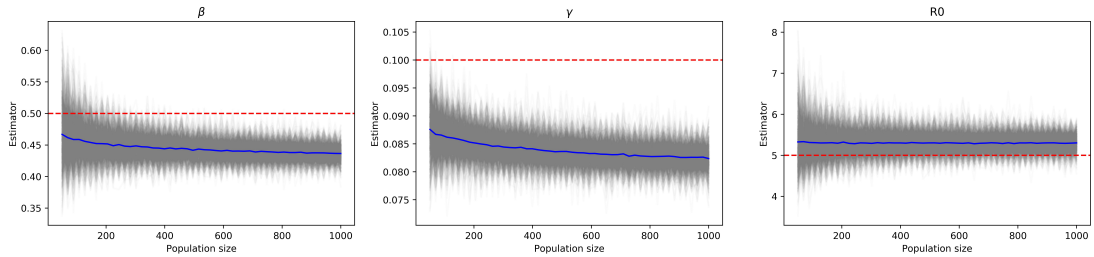


Figure 4: Numerical illustration of the estimators with parameters $\beta = 0.50$, $\gamma = 0.10$ and $R_0 = 5$ for the above Poisson processes type SIR model according to the population size with 1000 trajectories. The red line represents the initial parameter and the blue line represents the mean of the estimate

The estimators seem not to be efficient. We see then that the model does not lose the stochastic dynamic property in large population.

VI CONCLUSION AND FURTHER STUDIES

We develop a numerical approach to study the convergences of three stochastic compartmental models. Considering as reference, the basic deterministic model, we illustrate through simulations the behaviour in large populations of these three extensions: the continuous-time Markov chain model, the equivalent stochastic differential equation model and the model with Poisson process.

The Markov extension model allows the evolution of the epidemic to be modelled in a natural way and some results already established in probability theory allow for further studies. In our case, simulation shows a relatively fast convergence to the deterministic model in a growing population.

The stochastic equivalent SIR model with differential equation allows to keep a stochastic pace for a large-scale population. Converging to the deterministic case, the model slightly loses its stochasticity when the size of the population increases.

The stochastic extension with the Poisson process is interesting when we are interested in a large population. It preserves stochastic dynamics even if the size of the population increases. In this case, our studies do not lead to good quality estimators. A more in-depth approach must be carried out to have effective estimators. This will be down in a next work.

In terms of perspectives, we are currently interested in theoretical developments of more general non-Markovian SIR models as in [9]. By exploiting real data, we estimate the model parameters with the Expectation-Maximization algorithm [8]. The use of a Bayesian approach can also be a solution to take into account the opinions of epidemiological specialists in the estimation procedure.

REFERENCES

- [1] P. E. Kloeden, E. Platen, and H. Schurz. “Stochastic Differential Equations”. en. In: *Numerical Solution of SDE Through Computer Experiments*. Edited by P. E. Kloeden, E. Platen, and H. Schurz. Universitext. Berlin, Heidelberg: Springer, 1994. ISBN: 978-3-642-57913-4.
- [2] K. Dietz and J. A. P. Heesterbeek. “Daniel Bernoulli’s epidemiological model revisited”. en. In: *Mathematical Biosciences* 180.1 (Nov. 2002), pages 1–21. ISSN: 0025-5564.
- [3] E. J. Allen, L. J. S. Allen, A. Arciniega, and P. E. Greenwood. “Construction of Equivalent Stochastic Differential Equation Models”. In: *Stochastic Analysis and Applications* 26.2 (Mar. 2008). Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/07362990701857129>, pages 274–297. ISSN: 0736-2994.
- [4] L. J. S. Allen. “An Introduction to Stochastic Epidemic Models”. en. In: *Mathematical Epidemiology*. Edited by F. Brauer, P. van den Driessche, and J. Wu. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer, 2008, pages 81–130. ISBN: 978-3-540-78911-6.
- [5] T. Perkins. *Maximum likelihood trajectories for continuous-time Markov chains*. Journal Abbreviation: Advances in Neural Information Processing Systems 22 - Proceedings of the 2009 Conference Pages: 1445 Publication Title: Advances in Neural Information Processing Systems 22 - Proceedings of the 2009 Conference. Jan. 2009.
- [6] N. Bacaër. “McKendrick and Kermack on epidemic modelling (1926–1927)”. en. In: *A Short History of Mathematical Population Dynamics*. Edited by N. Bacaër. London: Springer, 2011, pages 89–96. ISBN: 978-0-85729-115-8.
- [7] T. Britton and E. Pardoux. “Stochastic epidemics in a homogeneous community”. In: *arXiv:1808.05350 [math]* 2255 (2019). arXiv: 1808.05350.
- [8] D. Castillo-Barnes, F. J. Martínez-Murcia, J. Ramírez, J. M. Górriz, and D. Salas-Gonzalez. “Expectation–Maximization algorithm for finite mixture of -stable distributions”. In: *Neurocomputing* 413 (Nov. 6, 2020), pages 210–216. ISSN: 0925-2312.
- [9] R. Forien, G. Pang, and É. Pardoux. “Estimating the state of the Covid-19 epidemic in France using a non-Markovian model”. In: *medRxiv* (2020).
- [10] O. Ige. “Markov Chain epidemic models and parameter estimation”. PhD thesis. Marshall University, 2020.
- [11] S. M. Manou-Abi and J. Balicchi. “Analysis of the COVID-19 epidemic in french overseas department Mayotte based on a modified deterministic and stochastic SEIR model”. In: *MedRxiv* (2020).
- [12] A. Raherinirina, T. S. Fandresena, A. R. Hajalalaina, H. Rabetafika, R. A. Rakotoarivelo, and F. Rafamatanantsoa. “Probabilistic Modelling of COVID-19 Dynamic in the Context of Madagascar”. en. In: *Open Journal of Modelling and Simulation* 9.3 (May 2021). Number: 3 Publisher: Scientific Research Publishing, pages 211–230.
- [13] A. Raherinirina, F. Rafamatanantsoa, T. S. Fandresena, and R. A. Rakotoarivelo. “Mathematical Model and Non-Pharmaceutical Control of the Coronavirus 2019 Disease in Madagascar”. en. In: *Open Journal of Modelling and Simulation* 9.3 (May 2021). Number: 3 Publisher: Scientific Research Publishing, pages 259–274.
- [14] S. M. MANOU-ABI, Y. Slaoui, and J. Balicchi. “Estimation of some epidemiological parameters with the Covid-19 data of Mayotte”. In: *Frontiers in Applied Mathematics and Statistics* (), page 67.