Kalpa Publications in Computing

Volume XXX, 2022, Pages 132–135

Proceedings of V XoveTIC Conference. XoveTIC 2022

Chemical Master Equation model to predict the COVID-19 epidemic spread

Manuel Pájaro[1], Antonio A. Alonso[2], and Irene Otero-Muras[3]

[1] Department of Mathematics and CITIC research center, University of A Coruña, Campus Elviña s/n, A Coruña, 15071, Spain. manuel.pajaro@udc.es

[2] BioProcess Engineering Group, IIM-CSIC. Spanish National Research Council, Eduardo Cabello 6, 36208, Vigo, Spain

[3] Institute for Integrative Systems Biology I^{2}SysBio (UV, CSIC) Spanish National Research Council, 46980, València, Spain. irene.otero.muras@csic.es

Abstract

The COVID-19 pandemics has shown relevant stochasticity in the spread of viral infection. Here, we formulate the dynamics of the epidemics spread as a stochastic SIR model, handling the compartmental model (of the SIR type) as a set of chemical reactions, and applying the Chemical Master Equation that describes the dynamics of the equivalent stochastic chemical reaction system. In this way, the solution (evolution of the probability distribution over time) can be obtained via the classical Stochastic Simulation Algorithm. The proposed methodology has been used to predict the COVID-19 evolution in small and medium size municipalities of Galicia.

1 Introduction

The spread of viral infectious diseases was typically modelled using the classical deterministic SIR type compartmental models [3]. In these models the trajectories are determined by the initial conditions and the inherent stochasticity of the process is not taken into account. Therefore, several stochastic models were proposed to address the prediction of COVID-19 evolution. In this context, we reinterpret the transitions between each possible state of a SIR model (Susceptible, Infected, Recovered) as chemical reactions. This system is modelled using a Chemical Master Equation (CME) which incorporates the inherent stochasticity of the COVID-19 infection process [4, 5]. We have tested this technique for small and medium municipalities in Galicia using data of the Public Health System (SERGAS) and data of viral load in Wastewater Treatment Plants (WWTP).

The CME is a widespread model used to represent stochastic biochemical systems, whose solution is unavailable for most cases. We use a classical Monte Carlo method, the Stochastic Simulation Algorithm (SSA) by Gillespie [2], which in each iteration produces an exact realization of the Markov Process, to approximate the solution of the CME. In this way, obtain not only the probability density functions of the infected persons, but also a number of trajectories of the infected persons evolution. Some of these trajectories are found to coincide exactly with the data of daily infected persons provided SERGAS.

Moreover, a more complex model is proposed, the SIRO model, aiming to incorporate the data of viral load detected and quantified by RT-qPCR in Wastewater Treatment Plants (WWTP). The population is split in Susceptible, Total Infected (estimated from the viral load in WWTP), Observed (detected by SERGAS) and Recovered. We construct the associated CME and solve it via the SSA, which allows forecasting the evolution of the number of infected persons starting from the viral load detected in sewage the previous week [4, 5].

2 Mathematical modeling

According with SIR models for infectious diseases [3], we assume that the N individuals of the total population can be split into the three categories S I and R previously described. These three categories can be associated to the corresponding three model state variables, which are related through the following set of reactions [4, 5]:

$$S + I \xrightarrow{\beta \frac{S}{N} I} 2I; \qquad I \xrightarrow{\gamma I} R, \tag{1}$$

with β being the infection rate, γ the recovered rate, while S, I and R represent the numbers of susceptible, infected and recovered persons, respectively.

The CME is a set of ordinary differential equations of the probability density function of being in each possible state of the system. For a general set of r reactions and m species the CME can be written as:

$$\frac{\mathrm{d}\mathcal{P}(\boldsymbol{n},t)}{\mathrm{d}t} = \sum_{j=1}^{r}\Big(t_j(\boldsymbol{n}+\boldsymbol{\nu}_j)\mathcal{P}(\boldsymbol{n}+\boldsymbol{\nu}_j,t) - t_j(\boldsymbol{n})\mathcal{P}(\boldsymbol{n},t)\Big), \tag{2}$$

where $\mathcal{P} : \mathbb{N}^m \times \mathbb{R}^+ \to [0\ 1]$ is the probability that there are n_i molecules of each specie i for $i = 1, \ldots, m$ at time t, t_j is the propensity function and ν_j is the state change vector (see for example [1]). The associated CME for the SIR model, set of reactions in (1), reads:

$$\begin{aligned}\frac{\mathrm{d}\mathcal{P}(\mathbf{n},t)}{\mathrm{d}t} \quad = \quad & \beta\frac{S+1}{N} I\mathcal{P}(S+1, I-1, R, t) - \beta\frac{S}{N} I\mathcal{P}(\mathbf{n},t) \\ & +\alpha(I+1)\mathcal{P}(S, I+1, R-1, t) - \alpha I\mathcal{P}(\mathbf{n},t),\end{aligned} \tag{3}$$

where $\boldsymbol{n} = (S, I, R)$ is the vector of states of the system, $[0\ N]$ is a interval in $\mathbb{N}$ and $\mathcal{P} : [0\ N]^3 \times \mathbb{R}^+ \to [0\ 1]$ is the probability density function of the state $\boldsymbol{n}$ for any time $t \geq 0$. The solution of the CME, equation (3), is obtained by using several realizations of the SSA [2].

3 Results and Discussion

As an example of the performance of the stochastic SIR model presented, we show in Figures 1 and 2 the predictions obtained in Ares for a week period from dates 2020/12/31 to 2021/01/06. The recovered rate $\alpha = 1/14$ and the infection rate $\beta = 0.14$ have been previously calibrated from data. The solution of the CME (3), has been approximated using 10^4 SSA realizations. Comparing the simulations obtained with the data of new daily infected persons reported by SERGAS we obtain several exact realizations, as plotted in Figure 1 (b). In Figure 1 (a), we compared the mean of all realizations and their standard deviation (blue lines) with the real data of cumulative infected persons provided by SERGAS (black squares). In Figure 2 we represent the CME solution, probability density distribution of the 14 days cumulative

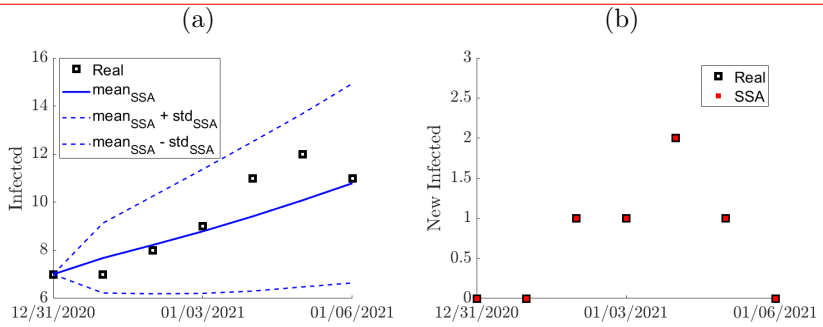

Figure 1: Cumulative infected persons for 14 days and new infected persons are depicted in plots (a) and (b), respectively. The CME solution is approximated by 10^4 SSA realizations. Mean and standard deviation (dashed lines) are depicted in (a), and one exact realization is shown in (b), red squares. Black empty squares are the real data provided by SERGAS.

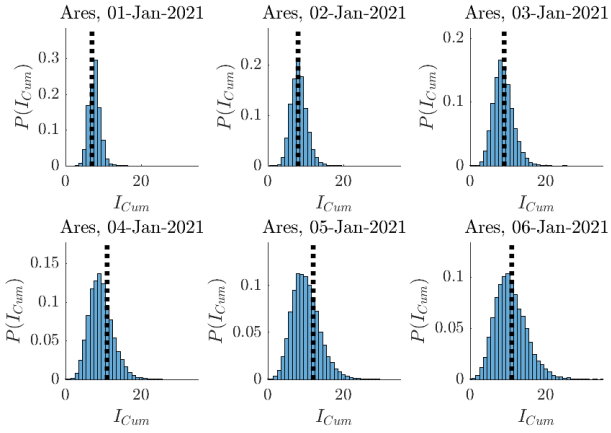

Figure 2: The histograms represent the CME solution approximated by 10^4 SSA realizations. The vertical dashed line shows the cumulative infected persons provided by the SERGAS.

infected persons, together with the real data provided by SERGAS, for each day of the week considered. Another dates, localities and stochastic models, as the SIRO to incorporate data from viral load in sewage, can be simulated using the available scripts under GPLv3 license at https://github.com/manuelpajaro/stochasticSIR_O [5].

Acknowledgments

MP acknowledges grant FJC2019-041397-I funded by MCIN/AEI/10.13039/501100011033.

References

[1] C. Gardiner. *Stochastic Methods. A Handbook for the Natural and Social Sciences*. Springer, Berlin, fourth edition, 2009.

[2] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of couple chemical reactions. *J. Comput. Phys.*, 22:403–434, 1976.

[3] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proc. R. Soc.*, 115(772):700–721, 1927.

[4] B. Novoa, R. Rios-Castro, I. Otero-Muras, S. Gouveia, and et al. Wastewater and marine bioindicators surveillance to anticipate COVID-19 prevalence and to explore SARS-CoV-2 diversity by next generation sequencing: one-year study. *Sci. Total Environ.*, 833:155140, 2022.

[5] M. Pájaro, N. Fajar, A.A. Alonso, and I. Otero-Muras. Stochastic SIR model predicts the evolution of COVID-19 epidemics from public health and wastewater data in small and medium-sized municipalities: a one year study. *Chaos Solitons Fractals*, page 112671, 2022.