

## Research Article

# Spatially Extended SHAR Epidemiological Framework of Infectious Disease Transmission

Damián Knopoff <sup>1,2</sup>, Nicole Cusimano <sup>1</sup>, Nico Stollenwerk <sup>1</sup> and Maíra Aguiar <sup>1,3</sup>

<sup>1</sup>Basque Center for Applied Mathematics, Alameda Mazarredo 14, 48009 Bilbao, Spain

<sup>2</sup>CIEM CONICET, Medina Allende s/n, 5000 Córdoba, Argentina

<sup>3</sup>Ikerbasque, Basque Foundation for Science, Bilbao, Spain

Correspondence should be addressed to Damián Knopoff; [dknopoff@bcamath.org](mailto:dknopoff@bcamath.org)

Received 9 September 2021; Revised 4 November 2021; Accepted 14 December 2021; Published 13 February 2022

Academic Editor: Pritpal Singh

Copyright © 2022 Damián Knopoff et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mathematical models play an important role in epidemiology. The inclusion of a spatial component in epidemiological models is especially important to understand and address many relevant ecological and public health questions, e.g., when wanting to differentiate transmission patterns across geographical regions or when considering spatially heterogeneous intervention measures. However, the introduction of spatial effects can have significant consequences on the observed model dynamics and hence must be carefully analyzed and interpreted. Cellular automata epidemiological models typically rely on simplified computational grids but can provide valuable insight into the spatial dynamics of transmission within a population by suitably accounting for the connections between individuals in the considered community. In this paper, we describe a stochastic cellular automata disease model based on an extension of the traditional Susceptible-Infected-Recovered (SIR) compartmentalization of the population, namely, the Susceptible-Hospitalized-Asymptomatic-Recovered (SHAR) formulation, in which infected individuals either present a severe form of the disease, thus requiring hospitalization, or belong to the so-called mild/asymptomatic class. The critical transmission threshold is derived analytically in the nonspatial SHAR formulation, and this generalizes previously obtained theoretical results for the SIR model. We present simulation results discussing the effect of key model parameters and of spatial correlations on model outputs and propose an algorithm for tracking the evolution of infection clusters within the considered population. Focusing on the role of import and criticality on the overall dynamics, we conclude that the current spatial setting increases the critical transmission threshold in comparison to the nonspatial model.

## 1. Introduction

Infectious diseases have shaped the global population throughout history, and (whether through devastating epidemics or recurrent outbreaks of endemic diseases) they have been—and often still are—responsible for a large number of deaths worldwide. The ongoing COVID-19 pandemic has certainly brought significant attention to the important contribution that mathematical modeling can have in monitoring the spread of infectious pathogens across a population, in better understanding its underlying mechanisms, and in exploring the effects of possible control measures aimed at containing the transmission of the considered disease (and possibly eradicating it from the observed population).

While the foundations of modern epidemiology based on compartmental models were laid in the first few decades of the 20th century, thanks to the work of public health physicians such as Ross, Hamer, McKendrick, and Kermack [1], the last 100 years have seen incredible advances in the field that were made possible also by the advent of computers and modern-day computational resources. The rapid increase in computer power of the last few decades has in fact allowed us to tackle many problems that were previously considered intractable and to simulate numerically complex systems across the entire spectrum of scientific disciplines, testing various working hypotheses and making predictions in many different hypothetical scenarios. This is true also in epidemiological modeling, where modern computers and high-performance simulations

allow the rapid analysis of the detailed models often required in order to provide quantitatively accurate predictions for the design of suitable public health policies.

Probably the most famous compartmental model in epidemiology is the SIR model [2]. In this formulation, individuals in the population belong to one of three mutually exclusive categories: susceptibles (individuals that have never been exposed to the pathogen causing the disease), infected individuals (the ones who have been exposed to the pathogen, becoming ill, and are capable of transmitting it), and recovered (the ones who have gone through the infection, recovering with a certain degree of immunity)—hence the SIR name for the model. The temporal evolution of the number of individuals in each disease-related category is described mathematically via a set of three ordinary differential equations involving three rates describing the probability of susceptible individuals becoming infected, of infected individuals to fight off the disease and recover, and of recovered individuals to become susceptible once again, that is, the so-called infection, recovery, and waning immunity rates, respectively.

This model and its many variations (SI, SIS, SEIR, etc.) give the basic framework for the vast majority of epidemiological models found in the literature, and as such, they have been analyzed in detail and extensively used to describe disease transmission dynamics in many contexts (e.g., [3–5] just to name a few). These classical models can be further refined in order to address disease-specific questions or according to the availability of data, for example, by including multiple strains of the same type of virus (see, e.g., [6–11]) and by subdividing the population into groups according to age or the risk of contracting the disease (e.g., [12–15] and references therein). Adding stochasticity is also an important aspect that might be key in correctly describing observed transmission patterns, especially when population size is small or in general when the number of infectious individuals is low [16, 17].

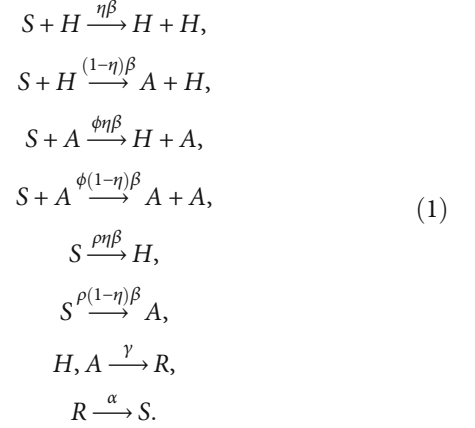
While explicitly including a spatial component in an epidemiological model could be an unnecessary complication to answer certain ecological and public health questions, some degree of spatial resolution is essential when wanting to differentiate transmission patterns across geographical regions [18] or when considering spatially heterogeneous intervention measures [19].

In this work, we consider the SHAR extension of the SIR compartmental model and develop a cellular automata stochastic formulation to explicitly account for spatial effects in the considered framework. We discuss the role of import and study the behavior of the system in various parametric settings, determining the critical transmission threshold analytically in the nonspatial formulation and exploring the effect of spatial correlations on the more general spatiotemporal model. Finally, we introduce the concept of infection clusters and investigate their evolution (in terms of the amount of clusters present in the system as well as their individual size) as a function of different parameter combinations.

## 2. The SHAR Model

The SHAR model is an extension of the well-known SIR framework in which the I class is divided into two groups

labeled  $H$  and  $A$ , respectively:  $H$  stands for individuals developing a severe form of the disease and likely being hospitalized, while  $A$  refers to infected individuals who are asymptomatic or have a mild form of the disease. The model dynamics is described by the following reaction scheme:



In addition to the infection rate  $\beta$ , the recovery rate  $\gamma$ , and the waning immunity rate  $\alpha$  (already present in the general SIR formulation), the SHAR approach includes two extra epidemiological parameters: the severity ratio  $\eta$  that gives the fraction of infected individuals who develop severe symptoms and the scaling factor  $\phi$ , which differentiates the infectivity  $\phi\beta$  of mild/asymptomatic infections with respect to the baseline infectivity  $\beta$  of severe/hospitalized cases [20]. The value of  $\phi$  can be tuned to describe different situations: a value of  $\phi < 1$  indicates that severe cases have larger infectivity than mild cases (which in the case of an infectious respiratory disease, for example, could be linked to enhanced coughing and sneezing), while  $\phi > 1$  would be used to describe the scenario in which asymptomatic individuals and mild cases contribute more to the spread of the infection (e.g., due to their higher mobility and possibility of interaction) than severe cases (that are more likely to be detected and isolated). Here, we also consider the import factor  $\rho$  which refers to the possibility of susceptible individuals contracting the disease via an undetected infection chain which started outside the studied population [21, 22].

The stochastic SHAR epidemic with import is modeled as a time-continuous Markov process to capture population noise. The temporal dynamics for the probability  $p(t, S, H, A)$  of having at time  $t$  an integer number  $S$  of susceptible,  $H$  hospitalized, and  $A$  asymptomatic can be described by the following master equation [23]:

$$\begin{aligned}
 \frac{d}{dt}p(t, S, H, A) &= \eta\phi\frac{\beta}{N}(S+1)Ap(t, S+1, H-1, A) \\
 &\quad + (1-\eta)\phi\frac{\beta}{N}(S+1)(A-1)p(t, S+1, H, A-1) \\
 &\quad + \eta\frac{\beta}{N}(S+1)(H-1)p(t, S+1, H-1, A) \\
 &\quad + (1-\eta)\frac{\beta}{N}(S+1)Hp(t, S+1, H, A-1) \\
 &\quad + \beta\eta\rho(S+1)p(t, S+1, H-1, A)
 \end{aligned}$$

$$\begin{aligned}
& + \beta(1 - \eta)\rho(S + 1)p(t, S + 1, H, A - 1) \\
& + \gamma(A + 1)p(t, S, H, A + 1) \\
& + \gamma(H + 1)p(t, S, H + 1, A) \\
& + \alpha(N - (S - 1) - H - A)p(t, S - 1, H, A) \\
& - \left( \frac{\beta}{N}S(H + \phi A + \rho N) + \gamma(A + H) \right. \\
& \left. + \alpha(N - S - H - A) \right) p(t, S, H, A), \tag{2}
\end{aligned}$$

where the number of recovered individuals follows from the constant population size assumption  $R = N - (S + H + A)$ . From the definition of the mean values

$$\begin{aligned}
\langle S \rangle(t) &= \sum_{S=0}^N \sum_{H=0}^N \sum_{A=0}^N Sp(t, S, H, A), \\
\langle H \rangle(t) &= \sum_{S=0}^N \sum_{H=0}^N \sum_{A=0}^N Hp(t, S, H, A), \tag{3} \\
\langle A \rangle(t) &= \sum_{S=0}^N \sum_{H=0}^N \sum_{A=0}^N Ap(t, S, H, A),
\end{aligned}$$

we can derive dynamic equations for the averages. For example, for the mean susceptible population, we derive

$$\frac{d}{dt} \langle S \rangle(t) = \sum_{S=0}^N \sum_{H=0}^N \sum_{A=0}^N S \frac{d}{dt} p(t, S, H, A), \tag{4}$$

and then insert the master equation (2) on the right-hand side. By using the so-called mean-field approximation [24] (essentially assuming that the covariance between variables is negligible and hence approximating the second order moments as follows:  $\langle SH \rangle \approx \langle S \rangle \langle H \rangle$  and  $\langle SA \rangle \approx \langle S \rangle \langle A \rangle$ ), and by suitably incorporating boundary modifications of Equation (2) for the cases in which  $S, H, A \in \{0, N\}$ , we obtain the mean-field ODE system for the SHAR model

$$\begin{aligned}
\frac{dS}{dt} &= \alpha R - \beta \frac{S}{N} (H + \phi A + \rho N), \\
\frac{dH}{dt} &= \eta \beta \frac{S}{N} (H + \phi A + \rho N) - \gamma H, \\
\frac{dA}{dt} &= (1 - \eta) \beta \frac{S}{N} (H + \phi A + \rho N) - \gamma A, \\
\frac{dR}{dt} &= \gamma(H + A) - \alpha R,
\end{aligned} \tag{5}$$

where we have not included the  $\langle \cdot \rangle$  to simplify notation.

**2.1. The SHAR Model with Import around the Critical Threshold.** Let  $x = (H, A)^T$  (where  $H$  and  $A$  could either be the variables of the deterministic SHAR model or the mean-field approximations of the corresponding stochastic

system) and assuming abundance of susceptibles (i.e.,  $S/N \approx 1$ ), then we can rewrite in vector form the equations for the dynamics of the infected classes of Equation (5) as follows:

$$\frac{dx}{dt} = Mx + v, \tag{6}$$

where the matrix  $M$  and the vector  $v$  are defined, respectively, as

$$\begin{aligned}
M &= \begin{bmatrix} \eta\beta - \gamma & \eta\phi\beta \\ (1 - \eta)\beta & (1 - \eta)\phi\beta - \gamma \end{bmatrix}, \\
v &= \begin{bmatrix} \eta \\ 1 - \eta \end{bmatrix} \beta \rho N.
\end{aligned} \tag{7}$$

The steady-state  $x^* = [H^*, A^*]^T$  of Equation (6) is nothing but  $x^* = -M^{-1}v$ , where the matrix  $M^{-1}$  can be explicitly computed via straightforward calculations as

$$M^{-1} = \frac{1}{\gamma(\gamma - \beta(\eta + \phi(1 - \eta)))} \begin{bmatrix} (1 - \eta)\phi\beta - \gamma & -\eta\phi\beta \\ -(1 - \eta)\beta & \eta\beta - \gamma \end{bmatrix}, \tag{8}$$

leading to

$$\begin{bmatrix} H^* \\ A^* \end{bmatrix} = \frac{\beta \rho N}{\gamma - \beta(\eta + \phi(1 - \eta))} \begin{bmatrix} \eta \\ 1 - \eta \end{bmatrix}. \tag{9}$$

The eigenvalues of the matrix  $M$  are

$$\begin{aligned}
\lambda_1 &= \beta(\eta + \phi(1 - \eta)) - \gamma, \\
\lambda_2 &= -\gamma,
\end{aligned} \tag{10}$$

so that community transmission is controlled (and hence, the steady-state  $x^*$  is an asymptotically stable equilibrium) for values of  $\beta$  below the critical threshold

$$\beta_c = \frac{\gamma}{\eta + \phi(1 - \eta)}. \tag{11}$$

The nonlinear dependence of  $\beta_c$  on  $\eta$  and  $\phi$  for a fixed value of  $\gamma$  is illustrated in Figure 1, highlighting that different parametric combinations could lead to the same value of  $\beta_c$ , but the main underlying modeling assumptions in terms of the scaling parameter  $\phi$  would be preserved (in fact, for  $\eta < 1$ , there are no combinations  $(\eta, \phi)$  with  $\phi > 1$  giving the same value of  $\beta_c$  as any combination of  $(\eta, \phi)$  with  $\phi < 1$ ).

In these settings, letting  $\tilde{I}^* = H^* + \phi A^*$ , we observe that

$$\tilde{I}^* = \frac{\beta}{\beta_c - \beta} \rho N, \tag{12}$$

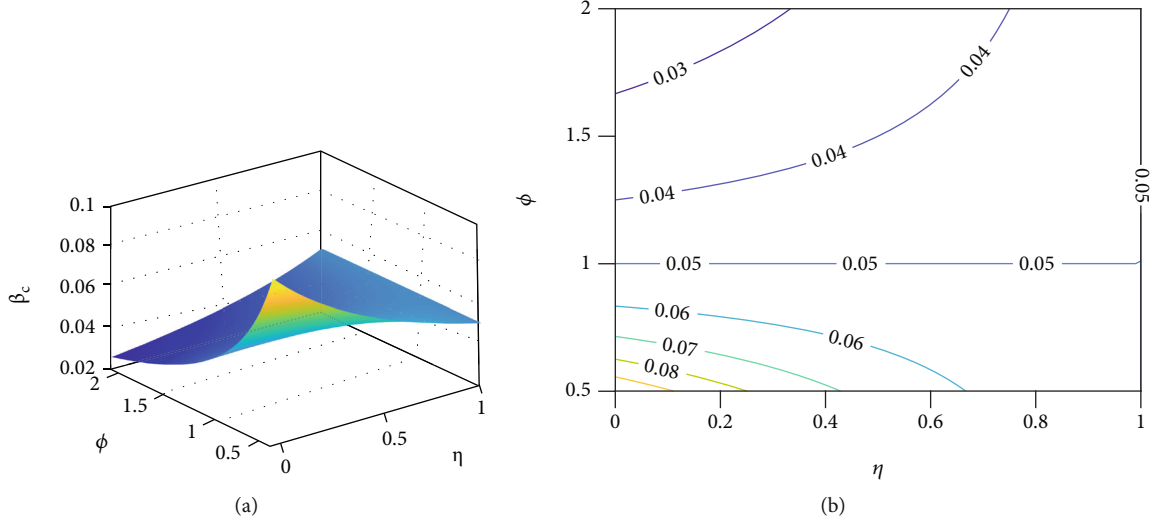


FIGURE 1: Critical threshold surface  $\beta_c$  as a function of  $\eta$  and  $\phi$  for constant  $\gamma = 0.05$ . (a) Surface  $\beta_c$  in three-dimensional space and (b) corresponding contour diagram. For fixed  $\eta < 1$ ,  $\beta_c$  decreases for increasing  $\phi$ , while for fixed  $\phi$ , the behavior of  $\beta_c$  as a function of  $\eta$  depends on whether  $\phi$  is greater than, equal to, or smaller than one.

thus generalizing the critical threshold result obtained for the SIR model with import [25]. Notice that, rather than considering the asymptotic number of infected  $I^* = H^* + A^*$ , in order to generalize previous results on the SIR model, we had to suitably rescale  $A$  by the  $\phi$  parameter. However,  $I^*$  can also be written in compact form as

$$I^* = \frac{\tilde{\beta}}{\beta_c - \beta} \rho N, \quad (13)$$

with  $\tilde{\beta} = \beta(\eta + \phi(1 - \eta))$ , and hence, its qualitative dependence on the import term remains unchanged.

Figure 2(a) shows 100 realizations (obtained by using the Gillespie algorithm [26]) of the stochastic SHAR model in the subcritical regime ( $\beta < \beta_c$ ). Figure 2(b) shows the limiting case of community spreading at criticality  $\beta = \beta_c$ , where we can see a linear growth of infections, while in Figure 2(c), an exponential increase of cases is observed for supercritical community spreading with  $\beta > \beta_c$ .

This transition of the dynamics from subcritical to supercritical regimes is also illustrated in Figure 3, where the mean-field solution is plotted for different values of  $\beta$  both in natural and log-log scales. Note that throughout the manuscript, we use days, labelled by  $d$ , as time units for the proposed simulations. However,  $d$  could be changed as to represent any relevant time unit of interest (which would be typically dictated by the time scale on which real-life data for model validation would be available).

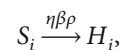
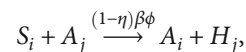
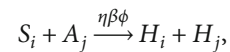
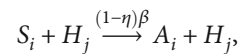
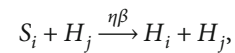
### 3. The Spatial SHAR Model

In this section, we present the spatial version of the stochastic SHAR model. Let us consider a lattice consisting of  $N$  sites, each of which is occupied by an individual and labelled by the indicators  $S_i, H_i, A_i \in \{0, 1\}$ , according to its state. Each site,  $i$ , can either have a susceptible ( $S_i = 1, H_i = A_i =$

0), hospitalized ( $H_i = 1, S_i = A_i = 0$ ), asymptomatic ( $A_i = 1, S_i = H_i = 0$ ), or recovered ( $S_i = H_i = A_i = 0$ ). Notice that there is no need to explicitly introduce the indicator  $R_i$  since in the present settings,  $R_i = 1 - S_i - H_i - A_i$ . The state vector for the system is thus given by  $(S_1, H_1, A_1, \dots, S_N, H_N, A_N)$ . Let  $J$  be the lattice adjacency matrix with entries  $J_{ij} = 1$ , if sites  $i$  and  $j$  are connected, and  $J_{ij} = 0$  otherwise. From the matrix  $J$  and any given state configuration, the following quantities of interest can be straightforwardly computed: the number of neighbors to site  $i$ ,  $Q_i = \sum_{j=1}^N J_{ij}$ , the number of infected neighbors to site  $i$ ,  $\sum_{j=1}^N J_{ij}(H_j + A_j)$ , and the force of infection to site  $i$ ,  $\beta \sum_{j=1}^N J_{ij}(H_j + \phi A_j)$ .

The adjacency matrix  $J$  gives the “neighboring” structure of the lattice. Although different network structures can be considered [27], it is known that the particular choice of lattice structure does not make much difference on the “universality” property of the system [28]. In general, in two-dimensional lattices, interactions are considered with the nearest 4 (Von Neumann neighborhood) or 8 (Moore neighborhood) neighboring lattice sites, as illustrated in Figure 4 for a  $5 \times 5$  lattice. In our present study, we will consider Von Neumann neighborhoods and apply periodic boundary conditions so that  $Q_i = 4$  for all  $i$ .

The reaction scheme for the spatial SHAR model is



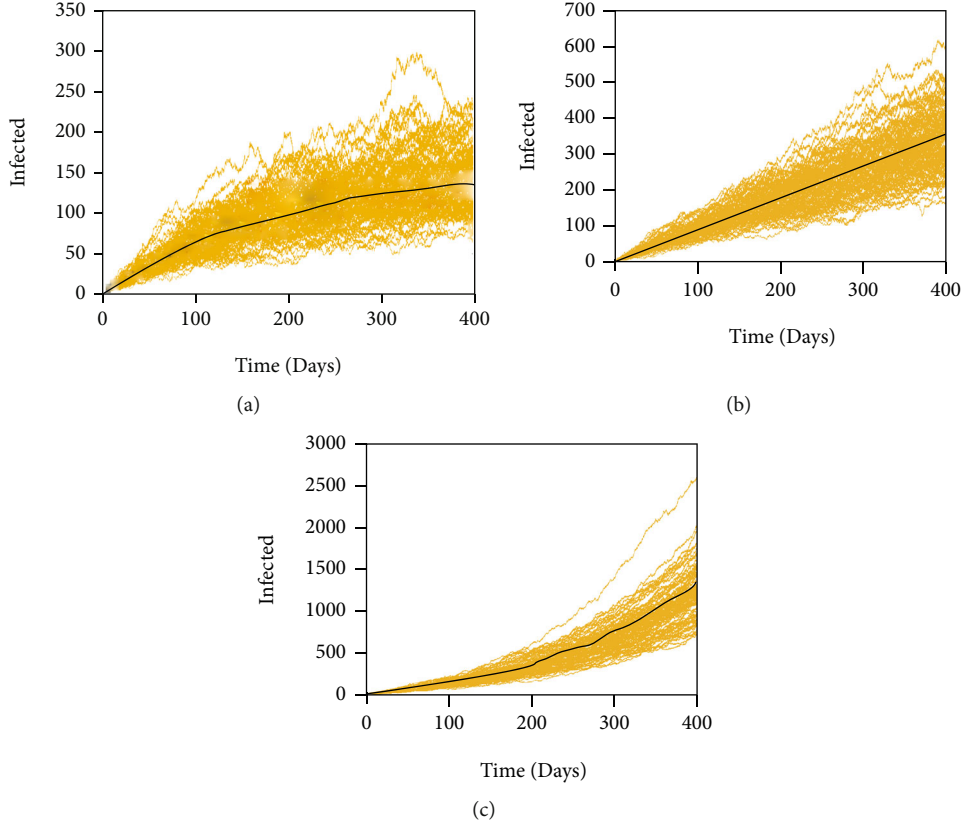


FIGURE 2: Ensemble of stochastic realizations and the analytic mean field solution for infected  $H + A$  in the SHAR model. Here,  $N = 2 \cdot 10^6$  with initially 2 infected (one hospitalized and one asymptomatic),  $\gamma = 0.05 d^{-1}$ ,  $\alpha = 0 d^{-1}$ ,  $\eta = 0.4$ ,  $\phi = 0.5$ ,  $\rho = e^{-12}$ . The infection rate  $\beta$  varies from subcritical to supercritical regimes with (a)  $\beta = 0.9\beta_c$ , (b)  $\beta = \beta_c$ , and (c)  $\beta = 1.1\beta_c$ .

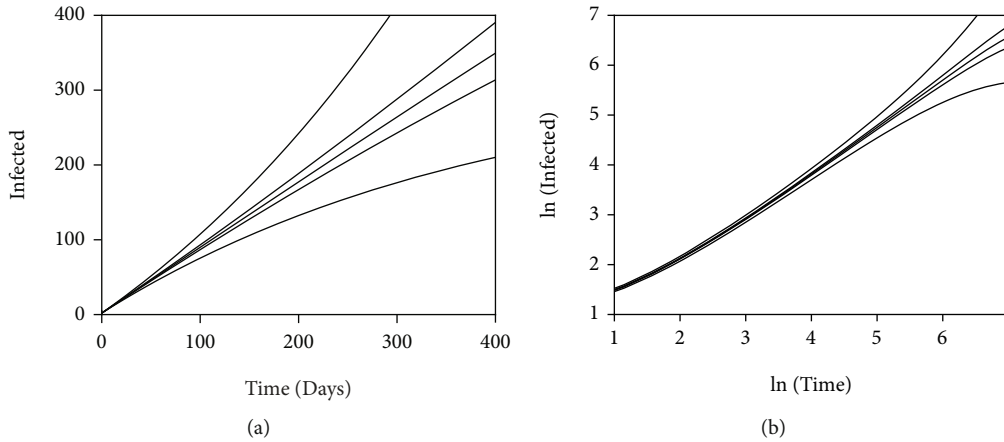
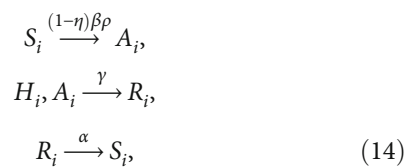


FIGURE 3: Mean-field solution for infected  $H + A$  for different values of  $\beta$  in (a) natural scale and (b) log-log scale. From bottom to top  $\beta = 0.95\beta_c$ ,  $\beta = 0.99\beta_c$ ,  $\beta = \beta_c$ ,  $\beta = 1.01\beta_c$ , and  $\beta = 1.05\beta_c$ .



where the processes involving contagion between individuals within the considered population may only occur

when sites  $i$  and  $j$  are connected. The master equation for the stochastic spatial SHAR model with import is given by

$$\begin{aligned}
 \frac{d}{dt} p(t, S_1, H_1, A_1, \dots, S_i, H_i, A_i, \dots, S_N, H_N, A_N) \\
 = \sum_{i=1}^N \left[ \beta\eta \left( \rho + \sum_{j=1}^N J_{ij}(H_j + \phi A_j) \right) H_i p(t, S_1, H_1, A_1, \dots, 1 \right. \\
 \left. - S_i, 1 - H_i, A_i, \dots, S_N, H_N, A_N) + \beta(1 - \eta) \right]
 \end{aligned}$$

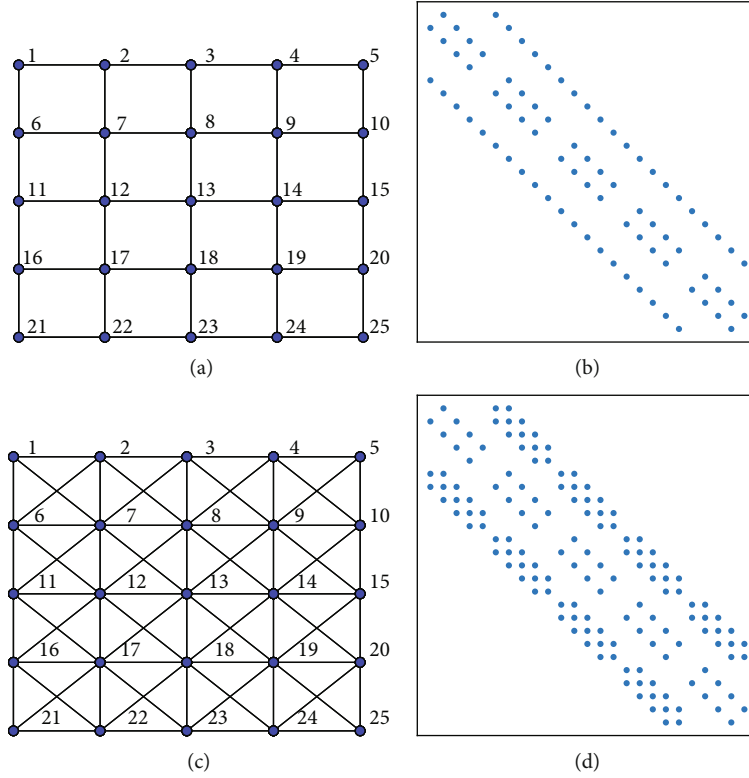


FIGURE 4: Schematic representation of a two-dimensional  $5 \times 5$  lattice. (a) and (b) (resp., (c) and (d)) show the edges connecting neighboring sites according to the Von Neumann (resp., Moore) neighborhood and the sparsity pattern of the corresponding adjacency matrix for the considered lattice.

$$\begin{aligned}
& \cdot \left( \rho + \sum_{j=1}^N J_{ij}(H_j + \phi A_j) \right) A_i p(t, S_1, H_1, A_1, \dots, 1 \\
& - S_i, H_i, 1A_i, \dots, S_N, H_N, A_N) \\
& + \gamma(1 - S_i - H_i - A_i) p(t, S_1, H_1, A_1, \dots, S_i, 1 - H_i, A_i, \dots, S_N, H_N, A_N) \\
& + \gamma(1 - S_i - H_i - A_i) p(t, S_1, H_1, A_1, \dots, S_i, H_i, 1 - A_i, \dots, S_N, H_N, A_N) \\
& + \alpha S_i p(t, S_1, H_1, A_1, \dots, 1 - S_i, H_i, A_i, \dots, S_N, H_N, A_N) \Big] \\
& - \sum_{i=1}^N \left[ \beta \left( \rho + \sum_{j=1}^N J_{ij}(H_j + \phi A_j) \right) S_i + \gamma(H_i + A_i) + \alpha(1 - S_i - H_i - A_i) \right] \\
& \times p(t, S_1, H_1, A_1, \dots, S_i, H_i, A_i, \dots, S_N, H_N, A_N). \quad (15)
\end{aligned}$$

**3.1. Critical Threshold.** In Section 2, the critical threshold for the stochastic SHAR model was obtained analytically from the mean-field approximation, see Equation (11). In the spatial case, the problem is much more challenging and since we do not have an analytical expression for this threshold, we must rely on numerical methods, which are often computationally expensive.

To get a first insight into the spatial dynamics, let us consider the one-dimensional SHAR process in which each lattice site is connected to its nearest neighbors. Using periodic boundary conditions, we can essentially think about a

ring lattice (a network in which each node is connected to its 2 nearest neighbors, i.e.,  $Q_i = 2$  for all  $i$ ). Figure 5 shows three different realizations of the SHAR process on a lattice of size  $N = 100$ . Each pixel row of the given plots shows the spatial configuration of the considered one-dimensional system at a given time point, while time increases from top to bottom. In all cases, we start from an entirely susceptible population (green) with one hospitalized individual in the center of the lattice and run the simulation up to the final time 1000 d. Hospitalized individuals are shown in red, and asymptomatic cases are yellow, while recovered are blue. The set of parameter values for these simulations is  $\gamma = 0.05 \text{ d}^{-1}$ ,  $\alpha = 0 \text{ d}^{-1}$ ,  $\eta = 0.4$ ,  $\phi = 0.5$ , and  $\rho = e^{-12}$ , while  $\beta = 0.1 \text{ d}^{-1}$  in Figure 5(a),  $\beta = 1 \text{ d}^{-1}$  in Figure 5(b), and  $\beta = 2 \text{ d}^{-1}$  in Figure 5(c). We can see that for the smallest value of  $\beta$  considered here, only a few infections occur (only asymptomatic for the particular realization shown in Figure 5(a)) before extinction, but for increasing values of  $\beta$ , larger outbreaks take place and import also starts to play a role (see the small outbreak generated by import towards the end of the simulation shown in Figure 5(c)).

For the two-dimensional case, let us consider a population of  $N = 62500$  individuals distributed in a  $250 \times 250$  square lattice, once again adopting a Von Neumann neighborhood with periodic boundary conditions. The population is assumed to be entirely susceptible at the beginning, except for one initial infected individual located in the center of the

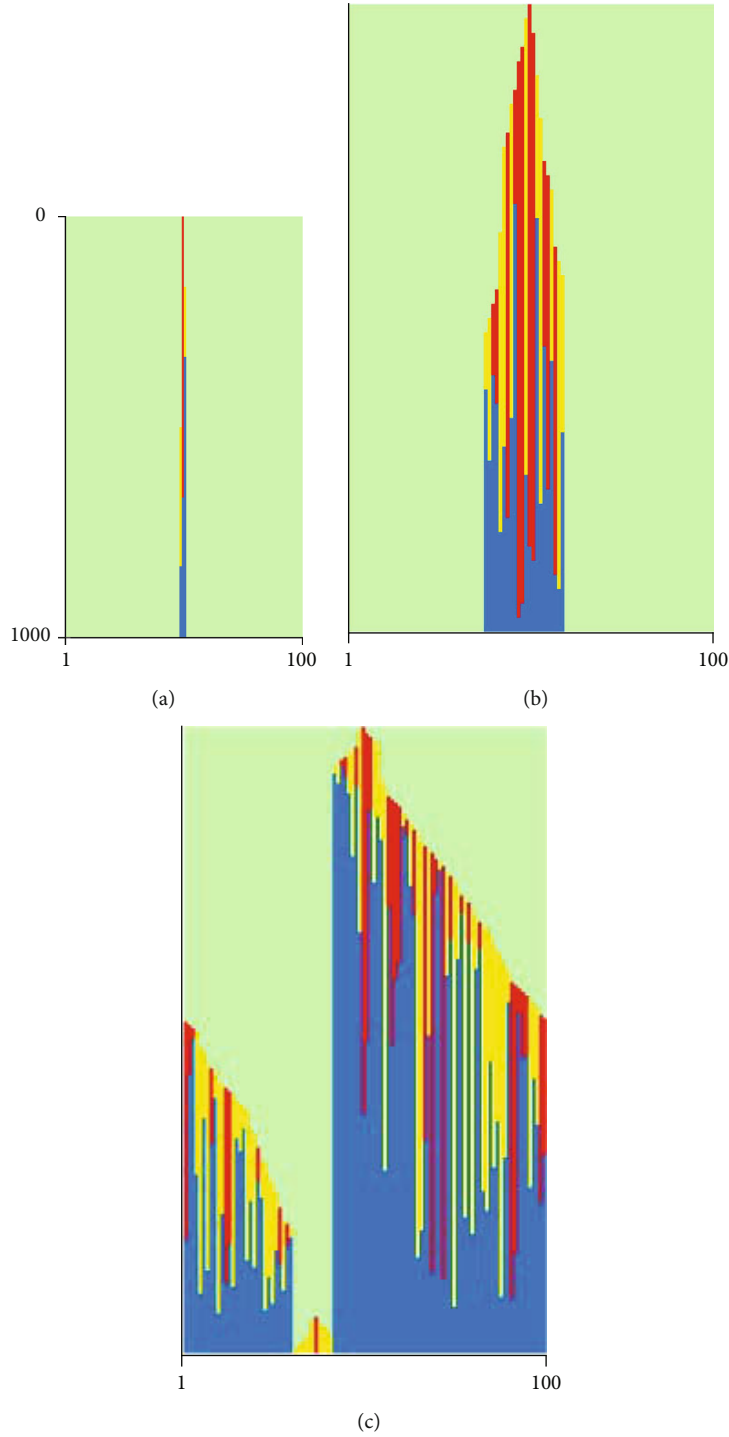


FIGURE 5: Three stochastic realizations of the one-dimensional SHAR model in a population of size  $N = 100$  divided into susceptible (green), hospitalized (red), asymptomatic (yellow), and recovered (blue). The horizontal axis gives the one-dimensional space, while time grows from top to bottom and the scale is indicated on the left figure. Parameters are  $\gamma = 0.05 \text{ d}^{-1}$ ,  $\alpha = 0 \text{ d}^{-1}$ ,  $\eta = 0.4$ ,  $\phi = 0.5$ , and  $\rho = e^{-12}$ , while (a)  $\beta = 0.1 \text{ d}^{-1}$ , (b)  $\beta = 1 \text{ d}^{-1}$ , and (c)  $\beta = 2 \text{ d}^{-1}$ .

lattice. Parameter values are  $\gamma = 0.05 \text{ d}^{-1}$ ,  $\alpha = 0 \text{ d}^{-1}$ ,  $\eta = 0.4$ ,  $\phi = 0.5$ , and  $\rho = e^{-12}$ , as in the simulations for the nonspatial model of Section 2. Figures 6 and 7 show four snapshots in time of a single realization of the spatial SHAR system for two different values of the infection rate, namely,  $\beta = 0.071$

$\text{d}^{-1}$  and  $\beta = 0.1 \text{ d}^{-1}$ , respectively. In both cases, a first outbreak is generated by the infected individual initially considered in the lattice center. Then, the import term produces more isolated outbreaks of different sizes. While some of these outbreaks involve only a few individuals and quickly

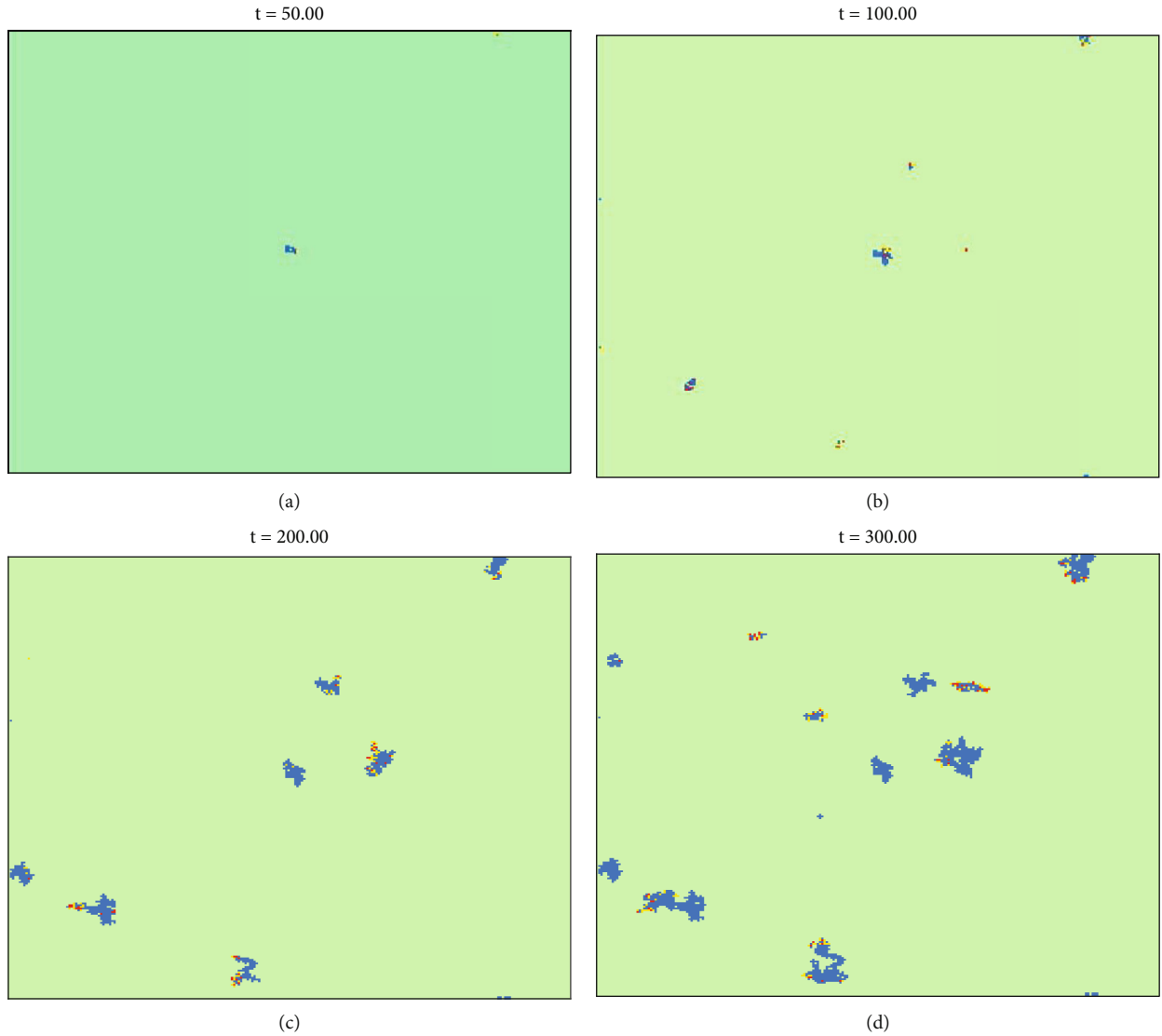


FIGURE 6: Spatial configuration of the two-dimensional SHAR system with import at four different time points. The population of  $N = 250^2$  is divided into susceptible (green), hospitalized (red), asymptomatic (yellow), and recovered (blue). Parameter values are  $\gamma = 0.05 \text{ d}^{-1}$ ,  $\beta = 0.071 \text{ d}^{-1}$ ,  $\alpha = 0 \text{ d}^{-1}$ ,  $\eta = 0.4$ ,  $\phi = 0.5$ , and  $\rho = e^{-12}$ . One initial infected is present in the center of the lattice, and a Von Neumann neighborhood is considered.

die out, with all individuals recovering, others may involve many more active infections (severe and/or mild and asymptomatic), spreading widely and eventually collapsing with neighboring clusters, leading to much larger outbreaks. An exponential growth of cases is to be expected when the community transmission is supercritical (see discussion below).

In order to decide whether there is critical percolation threshold, numerous stochastic realizations are run for different infection rates  $\beta$ . Figure 8 shows the mean infected population of 100 stochastic realizations for 10 different values of  $\beta$  in log-log scale. As for the nonspatial case, we observe that the mean number of infected individuals for low values of  $\beta$  eventually levels off, approaching a constant infected population size, while for the largest values of  $\beta$ , we see a very fast increase in the mean number of infected (an exponential growth is expected for values of  $\beta$  above critical

threshold). However, we notice that for  $\beta = 0.0714 \text{ d}^{-1}$  (i.e., when  $\beta$  coincides with the  $\beta_c$  of the nonspatial model in the considered parametric setting), the mean number of infected no longer grows linearly in time but rather levels off, indicating a subcritical behavior for this particular parameter value. Hence, we can infer that spatial correlations for the considered model contribute to maintaining disease transmission under control for larger values of the infectivity  $\beta$ . Nevertheless, giving a more precise estimate of the critical threshold for this spatial model would require considering much larger lattice sizes (in order for the “abundance of susceptible” assumption to remain valid for longer times) and longer simulations (with the corresponding heavier computational burden), thus allowing to disambiguate between values of  $\beta$  corresponding to a super- or subcritical regime. Obtaining accurate estimates of the



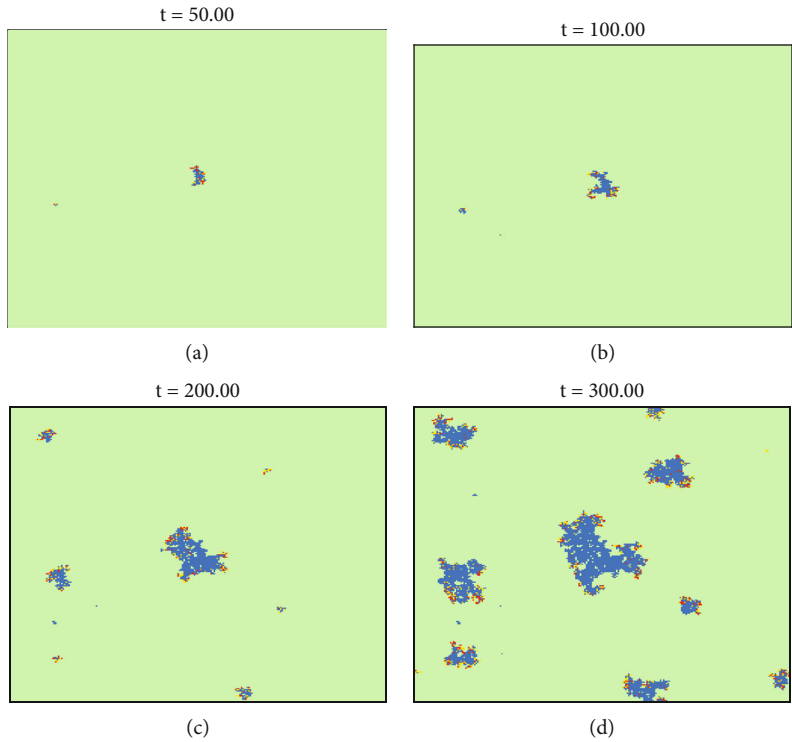


FIGURE 7: Spatial configuration of the two-dimensional SHAR system with import at four different time points. The population of  $N = 250^2$  is divided into susceptible (green), hospitalized (red), asymptomatic (yellow), and recovered (blue). Parameter values are  $\gamma = 0.05 \text{ d}^{-1}$ ,  $\beta = 0.1 \text{ d}^{-1}$ ,  $\alpha = 0 \text{ d}^{-1}$ ,  $\eta = 0.4$ ,  $\phi = 0.5$ , and  $\rho = e^{-12}$ . One initial infected is present in the center of the lattice, and a Von Neumann neighborhood is considered.

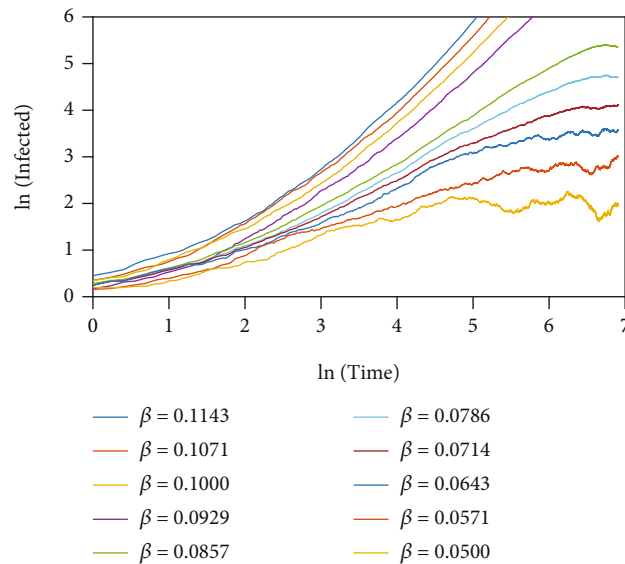


FIGURE 8: Numerically computed mean of the infected  $H + A$  for different values of  $\beta$  in log-log scale, for 100 realizations of the spatial SHAR model. Parameter values are  $\gamma = 0.05 \text{ d}^{-1}$ ,  $\alpha = 0 \text{ d}^{-1}$ ,  $\eta = 0.4$ ,  $\phi = 0.5$ , and  $\rho = e^{-12}$ .

critical exponents for the proposed spatial model is out of the scope of the presented study but will surely be a topic of our future investigation.

**3.2. The Effect of Import and Cluster Formation.** In this subsection, we focus on two key concepts of epidemiological

dynamics, namely, import and cluster formation. For the sake of simplicity, let us assume that the considered infectious disease is spread via human-to-human contact (i.e., ignoring vector-borne diseases and disease transmission across species). When no infections are present within a considered population, imports are responsible for

```

If a new infection occurred at  $t_n$  in location  $j$  then update cluster info as follows:
  Create  $n$ -th column of cluster configuration matrix  $C$ 
  Copy all entries of  $C(:, n-1)$  in  $C(:, n)$ 
  Update information in  $C(j, n)$  as follows:
  If no neighbours of  $j$  are infected use new cluster label
     $C(j, n) = \max \{C(:, n-1)\} + 1$ 
  Else
    Define vector of indicators of neighboring sites
    Remove zeros and possible repetitions
    If only one label  $k$  remains then assign the same label to location  $j$ 
       $C(j, n) = k$ 
    Else
      Compute number of individuals in each neighboring cluster
      Assign to location  $j$  the label of the largest cluster
    End if
    Update total number of clusters
    Compute number of individuals per cluster
  End if
End if

```

ALGORITHM 1: Pseudocode for the infection cluster computing algorithm.

triggering disease outbreaks. Imports typically take the form of external visitors who remain for a short period of time within the considered population (so that their presence does not change the population size) either introducing or reintroducing the pathogen, or of individuals belonging to the studied population who travel outside and return infected. A disease outbreak may become an epidemic but may also frequently die out due to the stochastic nature of the process. However, import of the pathogen from outside the population can prevent permanent extinction and generate new and sometimes unexpectedly large outbreaks.

In the case of infectious diseases, the development of useful testing and tracing strategies is fundamental. If we know that a particular individual has been infected and know her connections (in this context, her closest neighbors), then by looking at their state, we can effectively monitor the evolution of the outbreak and plan suitable containment measures accordingly. Fast and reliable testing strategies are clearly essential in obtaining a clear picture of the disease dynamics in the population of interest, and tracking individuals who enter from outside is particularly important, due to their previously mentioned key role in potentially triggering new outbreaks.

*3.2.1. An Infection Cluster Computing Algorithm.* Here, we describe an algorithm to monitor the evolution of an epidemic by tracking the number of infected individuals generated by each isolated outbreak. We assume that the population under study is often and regularly tested as to quickly identify new infections (either due to import or direct contact with an infected neighbor). For this purpose, we define the concept of an infection cluster, that is, a group of connected individuals who have been infected by the same index case and are either still infected or have recovered but have not lost their immunity yet.

Our proposed algorithm is coupled to the stochastic SHAR model solver described above and stores the cluster information in a cluster configuration matrix,  $C$ . Rows of the matrix  $C$  correspond to different lattice sites, and a new column is added to  $C$  for every cluster configuration change, i.e., whenever the number of clusters present in the system or any of the cluster sizes changes.  $C$  is initialized as a column vector of zeros having the size of the spatial lattice used for the simulation, and the label “1” is assigned to the row corresponding to the only infected individual present in the center of the lattice. A new isolated cluster is created when there are no infected neighbors to the newly infected site. Whenever this happens, our algorithm assigns to the new cluster a new label (an integer corresponding to the total number of clusters present in the system at that particular time point) and all future infections in the same cluster are then assigned the same label. When two clusters collide, we choose for the new infection the label of the larger cluster involved in the collision, but other choices (e.g., random selection) are possible. Moreover, in all the presented simulations, we assume  $\alpha = 0 \text{ d}^{-1}$  so that the number of clusters and their individual size never decreases. Clearly, as soon as  $\alpha > 0$ , this would be no longer the case and care should be taken in ensuring the proper labeling of new clusters. Notice that in order to monitor the number of clusters and their sizes, it is not necessary to store the state of the entire lattice for each time step. Instead, a new column of  $C$  is only created when a new infection occurs. The pseudocode is given in Algorithm 1.

Figure 9 shows three different aspects of the monitored cluster information for two realizations of the spatial stochastic SHAR model with import obtained, respectively, with  $\beta = 0.071 \text{ d}^{-1}$  (left column) and  $\beta = 0.1 \text{ d}^{-1}$  (right column). Figures 9(a) and 9(d) give the cluster configurations at the end of the simulation, namely, at  $t = 300 \text{ d}$ , for both realizations, and in both plots, a different colour is assigned

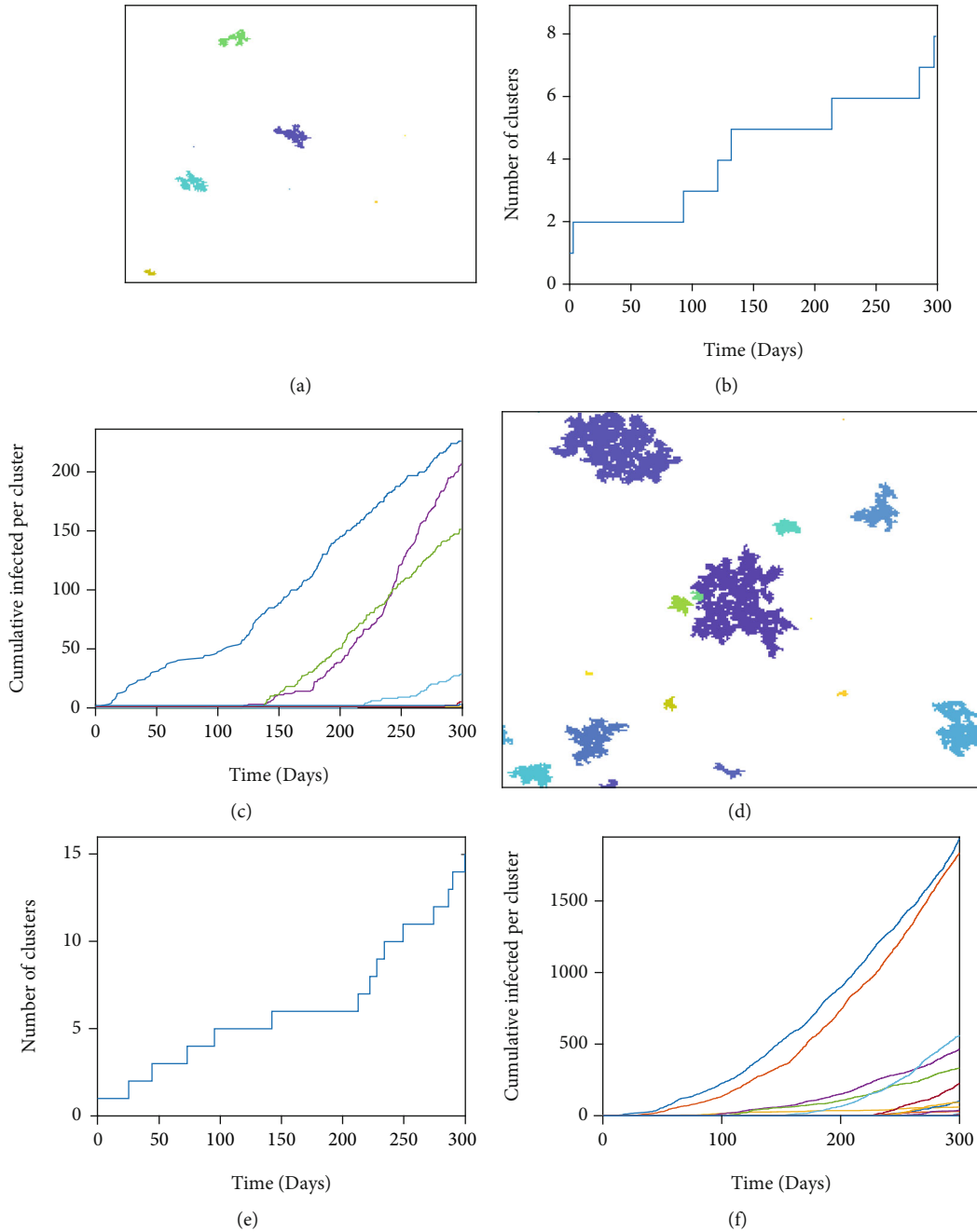


FIGURE 9: Tracing strategy and cluster formation. (a–c)  $\beta = 0.071 d^{-1}$ ; (d–f)  $\beta = 0.1 d^{-1}$ . (a, d) cluster configuration at time  $t = 300 d$ , (b, e) number of clusters, and (c, f) cumulative infected population; each curve represents a cluster.

to each cluster in order to aid their visualization. Figures 9(b) and 9(e) track the temporal evolution of the number of clusters in each simulation, while Figures 9(c) and 9(f) display changes in each individual cluster size over time.

By comparing these two individual realizations, we see that when  $\beta$  is larger, many more clusters form in the time interval considered and their individual size is generally much bigger than for the other value of the considered parameter. In order to confirm this trend, we perform a simple statistical analysis over several realizations of the spatial SHAR model. In particular, we monitor infection cluster

information (i.e., cluster size and number of clusters) of 500 realizations for both values of  $\beta$ , plotting in Figures 10 and 11 the corresponding histograms. The obtained results confirm our expectations. Specifically, Figure 10 shows that clusters of much larger size form for the larger value of  $\beta$  (which is to be expected due to the larger infectivity and hence faster transmission of the pathogen), while in Figure 11, we observe that, in average, a slightly larger number of clusters (i.e., new isolated outbreaks) forms for the larger value of  $\beta$  due to its inflating effect on the net import rate of the system.

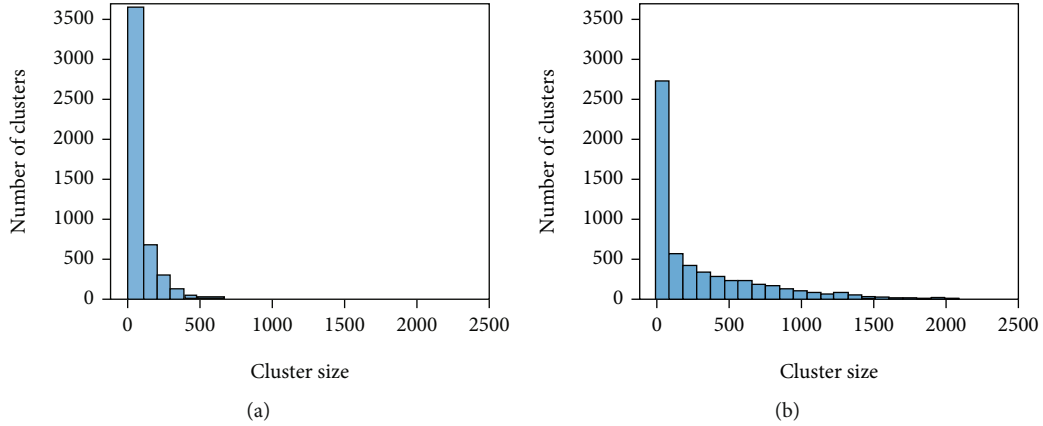


FIGURE 10: Distribution of the cluster sizes at final simulation time  $t = 300$  d for 500 realizations of the spatial SHAR model with (a)  $\beta = 0.071 \text{ d}^{-1}$  and (b)  $\beta = 0.1 \text{ d}^{-1}$ .

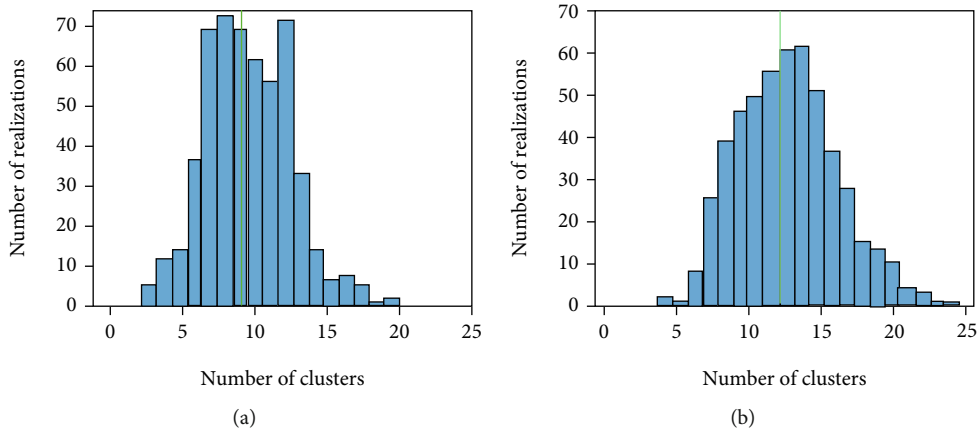


FIGURE 11: Distribution of the number of clusters (at final simulation time  $t = 300$  d) in 500 realizations of the spatial SHAR model with (a)  $\beta = 0.071 \text{ d}^{-1}$  and (b)  $\beta = 0.1 \text{ d}^{-1}$ . The green vertical line indicates the average number of clusters and is equal to 9.5 in (a) and 12.1 in (b).

## 4. Conclusions

In this work, we have presented the stochastic SHAR model with import and its mean-field approximation. We have computed the critical threshold of the system under the assumption of abundance of susceptibles and the asymptotic number of infected for controlled community transmission as a function of the model parameters, thus generalizing previously obtained results for the SIR model with import. Then, we introduced a cellular automata extension of the SHAR framework accounting for spatial variation within the system and we presented the master equation for the proposed model. Modifications to the critical threshold behavior in the presence of spatial correlations have been discussed in Section 3.1, showing that, in the parametric setting considered, these correlations pushed the system critical threshold towards larger values, hence contributing to maintaining community transmission controlled for a larger range of  $\beta$  values. Finally, in Section 3.2, we have focused on the role of import in triggering and maintaining an epidemic and proposed a way to track the number of infected individuals generated by each index case by looking at cluster

formation and their evolution in time. We see the proposed algorithm as a way to mimic the contact tracing procedure that might be implemented by the health system of a given region in order to monitor disease transmission. To conclude the manuscript, we would like to mention some of the limitations of the presented work and comment on possible directions of future work. The proposed analysis is aimed at providing an overview of the potential offered by the SHAR framework in a spatially extended context, focusing in particular on tools such as the infection cluster computing algorithm to study relevant features of the resulting spatiotemporal dynamics. We explored the effect of varying the infectivity parameter  $\beta$  on the model output but kept all other parameters fixed at some given (and typically arbitrary) value. A much more in-depth exploration of the parameter space remains to be performed. Clearly, according to the considerations made in Section 2.1, we expect variations in  $\gamma$ ,  $\eta$ ,  $\phi$  to affect the critical threshold of the system, and increasing values of  $\rho$  to produce more and more isolated outbreaks, following our considerations of Section 3.2. However, studying the combined effect of multiple parametric changes (including different population sizes) will

surely be a natural extension of this work. We remark also that throughout this manuscript, we always assumed  $\alpha = 0$   $d^{-1}$  and hence the case  $\alpha > 0$  (which will likely produce much richer dynamical scenarios) should also be investigated. Moreover, this particular work did not focus on describing the dynamics of any specific disease and therefore specific parameter values (or ranges) were not validated against any real-life data. Nevertheless, the SHAR framework already captures many important features of infectious disease transmission, and as such, we expect it to be a very versatile tool for further disease-specific investigation (e.g., in the context of COVID-19, seasonal influenza, etc.).

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare no potential conflict of interests.

## Authors' Contributions

M.A. and N.S. devised the mathematical framework and supervised the development of this study. D.K. and N.C. designed the computational framework and conceived and performed the numerical simulations. All authors contributed to the analysis of the results and to the writing of the manuscript.

## Acknowledgments

This research is supported by the Basque Government through the “Mathematical Modeling Applied to Health” Project, BERC 2018-2021 program and by the Spanish Ministry of Sciences, Innovation and Universities: BCAM Severo Ochoa Accreditation SEV-2017-0718. N.C. acknowledges also the ELKARTEK KK-2020/00008 funding scheme.

## References

- [1] F. Brauer, C. Castillo-Chavez, and Z. Feng, *Mathematical Models in Epidemiology*, Springer Science+Business Media, New York, 2019.
- [2] K. Dietz, “Epidemics and rumours: a survey,” *J. Roy. Statist. Soc. Ser. A*, vol. 130, no. 4, pp. 505–528, 1967.
- [3] C. Gilligan, S. Gubbins, and S. Simons, “Analysis and fitting of an SIR model with host response to infection load for a plant disease,” *Philosophical Transactions of the Royal Society of London B*, vol. 352, no. 1351, pp. 353–364, 1997.
- [4] M. Keeling, P. Rohani, and B. Grenfell, “Seasonally forced disease dynamics explored as switching between attractors,” *Physica D*, vol. 148, no. 3-4, pp. 317–335, 2001.
- [5] L. J. S. Allen, V. L. Brown, C. B. Jonsson et al., “Mathematical modeling of viral zoonoses in wildlife,” *Natural Resource Modeling*, vol. 25, no. 1, pp. 5–51, 2012.
- [6] A. Kucharski, V. Andreasen, and J. Gog, “Capturing the dynamics of pathogens with many strains,” *Journal of Mathematical Biology*, vol. 72, no. 1, pp. 1–24, 2015.
- [7] M. Aguiar, B. Kooi, and N. Stollenwerk, “Epidemiology of dengue fever: a model with temporary cross-immunity and possible secondary infection shows bifurcations and chaotic behaviour in wide parameter regions,” *Mathematical Modelling of Natural Phenomena*, vol. 3, no. 4, pp. 48–70, 2008.
- [8] M. Aguiar, B. Kooi, and N. Stollenwerk, “Multi-strain deterministic chaos in dengue epidemiology, a challenge for computational mathematics,” *AIP Conference Proceedings*, vol. 1168, pp. 1555–1558, 2009.
- [9] M. Aguiar, S. Ballesteros, and N. Stollenwerk, “Two strain dengue model with temporary cross immunity and seasonality,” *AIP Conference Proceedings*, vol. 1281, pp. 732–735, 2010.
- [10] M. Aguiar, S. Ballesteros, B. Kooi, and N. Stollenwerk, “The role of seasonality and import in a minimalistic multi-strain dengue model capturing differences between primary and secondary infections: complex dynamics and its implications for data analysis,” *Journal of Theoretical Biology*, vol. 289, no. 1, pp. 181–196, 2011.
- [11] B. Kooi, M. Aguiar, and N. Stollenwerk, “Analysis of an asymmetric two-strain dengue model,” *Mathematical Biosciences*, vol. 248, no. 1, pp. 128–139, 2014.
- [12] L. Xue-Zhi, J. Yang, and M. Martcheva, *Age Structured Epidemic Modeling*, Springer, 2020.
- [13] M. Aguiar, N. Stollenwerk, and S. Halstead, “The impact of the newly licensed dengue vaccine in endemic countries,” *PLoS Neglected Tropical Diseases*, vol. 10, no. 12, article e0005179, 2016.
- [14] M. Aguiar and N. Stollenwerk, “Dengvaxia efficacy dependency on serostatus: a closer look at more recent data,” *Clinical Infectious Diseases*, vol. 66, no. 4, pp. 641–642, 2018.
- [15] M. Aguiar and N. Stollenwerk, “Dengvaxia: age as surrogate for serostatus,” *The Lancet Infectious Diseases*, vol. 18, no. 3, p. 245, 2018.
- [16] M. Aguiar, B. Kooi, J. Martins, and N. Stollenwerk, “Scaling of stochasticity in dengue hemorrhagic fever epidemics,” *MMNP*, vol. 7, no. 3, pp. 1–11, 2012.
- [17] M. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals*, Princeton University Press, Princeton, 2008.
- [18] N. Stollenwerk, T. Götz, L. Mateus et al., “Modelling spatial connectivity in epidemiological systems, dengue fever in Thailand on networks from radiation models,” *AIP Conference Proceedings*, vol. 1738, article 390011, 2016.
- [19] S. Riley, K. Eames, V. Isham, D. Mollison, and P. Trapman, “Five challenges for spatial epidemic models,” *Epidemics*, vol. 10, pp. 68–71, 2015.
- [20] M. Aguiar and N. Stollenwerk, “SHAR and effective SIR models: from dengue fever toy models to a COVID-19 fully parametrized SHARUCD framework,” *Commun. Biomath. Sci.*, vol. 3, pp. 60–89, 2020.
- [21] L. Mateus, P. Ghaffari, U. Skwara et al., “Semiclassical approximations of stochastic epidemiological processes towards parameter estimation using as prime example the SIS system with import,” *Ecological Complexity*, vol. 27, pp. 63–73, 2016.
- [22] L. Mateus, D. Masoero, F. Rocha et al., “Epidemiological models in semiclassical approximation: an analytically solvable model as a test case,” *Math. Method. Appl. Sci.*, vol. 39, no. 16, pp. 4914–4922, 2016.
- [23] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, Elsevier, 1992.

- [24] N. Stollenwerk and V. Jansen, *Population Biology and Criticality: From Critical Birth-Death Processes to Self-Organized Criticality in Mutation Pathogen Systems*, World Scientific, 2011.
- [25] M. Aguiar, J. B. van-Dierdonck, J. Mar et al., “Critical fluctuations in epidemic models explain COVID-19 post-lockdown dynamics,” *Scientific Reports*, vol. 11, no. 1, pp. 1–12, 2021.
- [26] D. T. Gillespie, “A general method for numerically simulating the stochastic time evolution of coupled chemical reactions,” *Journal of Computational Physics*, vol. 22, no. 4, pp. 403–434, 1976.
- [27] M. Aguiar, G. Dosi, D. Knopoff, and M. E. Virgillito, *A multi-scale network-based model of contagion dynamics: heterogeneity, spatial distancing and vaccination*, Mathematical Models and Methods in Applied Sciences, 2021.
- [28] D. Stauffer, “Scaling theory of percolation clusters,” *Physics Reports*, vol. 54, no. 1, pp. 1–74, 1979.