

# Epidemics spread in heterogeneous populations

Karol Capała<sup>a</sup> and Bartłomiej Dybiec<sup>b</sup>

Marian Smoluchowski Institute of Physics, and Mark Kac Center for Complex Systems Research, Jagiellonian University, ul. St. Łojasiewicza 11, 30–348 Kraków, Poland

Received 8 December 2016 / Received in final form 15 March 2017

Published online 4 May 2017

© The Author(s) 2017. This article is published with open access at [Springerlink.com](http://Springerlink.com)

**Abstract.** Individuals building populations are subject to variability. This variability affects progress of epidemic outbreaks, because individuals tend to be more or less resistant. Individuals also differ with respect to their recovery rate. Here, properties of the SIR model in inhomogeneous populations are studied. It is shown that a small change in model's parameters, e.g. recovery or infection rate, can substantially change properties of final states which is especially well-visible in distributions of the epidemic size. In addition to the epidemic size and radii distributions, the paper explores first passage time properties of epidemic outbreaks.

## 1 Introduction

The SIR (Susceptible-Infectious-Recovered) model [1–3] is a well-known, classical example of an epidemiological model. Originally, it has been suggested and studied in the continuous limit [4]. The SIR model is considered as a starting point for epidemiological modeling of many diseases. Therefore, various extensions [5] to the model have been suggested and explored. In particular, they include modification of possible states and system topology. These alternations are usually introduced to the discrete version of the SIR model, which is suitable for extensions and modifications.

Observations of epidemics progress and their spatial patterns show that real populations are not homogeneous but intrinsically heterogeneous. Every population consists of individuals which share common features but are not exactly the same. Heterogeneity of a population can be introduced in various ways. The most common approach is to divide populations into smaller groups, which differ with respect to their characteristics/traits [6,7]. Another possibility is to assume that traits, e.g. susceptibility or infectivity, follow continuous distributions [8–13]. Some of such heterogeneous models, by means of nonlinear transformations [9], can be mapped into homogeneous models. Population heterogeneity at the level of infection probability can be studied using percolation theory [14,15] or agent-based modeling [16,17].

Inhomogeneity in epidemiological models not only affects individuals but also connections among them [18,19]. Variable connectivity changes progress of epidemics. Heterogeneity of connections decreases the epidemic

threshold. This fact is especially well-visible in case of epidemics on scale-free networks. In the thermodynamic limit, due to the presence of highly connected hubs, there is no infection threshold on scale-free networks [20,21]. Nevertheless, it can be reintroduced by an appropriate vaccination strategy [22].

Epidemiological processes are inherently stochastic due to the random nature of infection and recovery processes. The contact between susceptible and infectious individuals is necessary for infection but does not guarantee that an infection indeed takes place. In the case of the SIR model in the inhomogeneous population the stochasticity has two sources. The stochasticity of epidemiological systems is extended by the already mentioned variability of individuals. Therefore, a general population consists of individuals which are more or less vulnerable to infections. Similarly, some of individuals are capable of fast recovery, while others recover slowly. The presence of such individuals is a key element for creation of infection paths and islands of isolated, uninfected individuals.

Individuals' heterogeneity affects the average epidemic size and the invasion probability. If only the susceptibility is varied, less homogeneous populations are less likely to be invaded and consequently epidemic size is reduced [10,15,23]. Increasing heterogeneity reduces the chances of large outbreaks due to the possibility of creating impenetrable firewalls [14]. For negatively correlated infectivity and susceptibility the final size of epidemic is larger than in homogeneous case [24]. Finally, using percolation theory, it has been also shown that the increase of the average transmissibility increases the invasion probability. Contrary to the average transmissibility, an increase in the standard deviation of transmissibility reduces the invasion probability and decreases the epidemic

<sup>a</sup> e-mail: [karol.capala@uj.edu.pl](mailto:karol.capala@uj.edu.pl)

<sup>b</sup> e-mail: [bartek@th.if.uj.edu.pl](mailto:bartek@th.if.uj.edu.pl)

size [11,12] showing again that heterogeneity can decrease epidemic size.

Using Monte Carlo methods [25], we study the modification of the Kermack-McKendrick model. We extend the discrete lattice SIR model to take into account simultaneous variability in susceptibility and recovery rates. At the same time the connectivity among individuals is fixed and limited to the von Neuman neighborhood only. The model studied here is similar to ones considered in [11,12] and [14,15]. In contrast to [11,12], within the studied model, the time is discrete, resulting in the replacement of infection and recovery rates by appropriate probabilities per time step. Moreover, within presented studies more quantifiers characterizing epidemics are examined. Here, along with the examination of the epidemic size distribution and the invasion probability we study temporal properties like the epidemic duration and the first time to reach a given radius or size. Contrary to [14,15], variability is also included in the recovery rate.

Exploration of epidemic properties in heterogeneous populations is very important because an epidemic constitutes one of the processes that can significantly affect proper functioning of societies [26,27]. It can result in significant economic losses. Identification of the role of inhomogeneity is crucial for understanding the dynamics of plant [28,29], animal [30] and human epidemics [31].

The manuscript is organized in the following way: Section 2 (Model) presents the lattice SIR model in inhomogeneous populations, Section 3 (Results) discusses obtained results. Finally, the paper ends with Summary and Conclusions (Sect. 4).

## 2 Model

We are studying the lattice SIR (Susceptible-Infectious-Recovered) model [1] using computer simulations. The population of  $N$  individuals is divided into three classes of individuals:  $S$  – susceptible,  $I$  – infected and  $R$  – recovered. In every time step, any susceptible individual, let say  $i$ , can be infected with the probability  $P_I(i)$  due to a contact with an infectious neighbor. After becoming infected, an infectious individual can spontaneously recover with the probability  $P_R(i)$ . Recovered individuals are immune, i.e. they cannot be re-infected. The population is inhomogeneous, therefore every individual is characterized by probabilities per time step to become infected  $P_I(i)$  and to recover  $P_R(i)$ . The (total) probability of infecting a susceptible individual  $i$  during a single iteration (MC step), see below, is

$$P_{S \rightarrow I}(i) = 1 - \prod_{j \in \mathcal{N}(i)} [1 - P_I(i)]^{\Theta(j)}, \quad (1)$$

where  $\mathcal{N}(i)$  represents the nearest neighbors of the  $i$ th individual and  $\Theta(j)$  is the indicator function showing whether a neighbor  $j$  is infectious or not

$$\Theta(j) = \begin{cases} 0 & \text{if } j \text{ is not infectious} \\ 1 & \text{if } j \text{ is infectious.} \end{cases} \quad (2)$$

Probabilities (per time step) of becoming infected  $P_I$  and spontaneous recovery  $P_R$  are independent and distributed according to the truncated normal density on  $[0, 1]$

$$f(P, \mu, \sigma) = \begin{cases} 0 & \text{for } x \notin [0, 1] \\ \frac{1}{\sqrt{2\pi\sigma^2}C} \exp\left[-\frac{(P-\mu)^2}{2\sigma^2}\right] & \text{for } x \in [0, 1] \end{cases} \quad (3)$$

for different values of  $\mu_I$ ,  $\sigma_I$  and  $\mu_R$ ,  $\sigma_R$  respectively, where  $C$  is the normalization constant  $C = [\text{erf}(\mu/\sqrt{2}\sigma) - \text{erf}((\mu-1)/\sqrt{2}\sigma)]/2$ . If  $0 \leq \mu_I \leq 1$  and  $0 \leq \mu_R \leq 1$ ,  $\mu_I$  and  $\mu_R$  are modal values of infection and recovery probabilities (per time step) respectively. If  $\mu < 0$  ( $\mu > 1$ ), the modal value is located at 0 (1). Finally,  $\sigma_I$  and  $\sigma_R$  control distribution width, but they do not have a simple interpretation, as appropriate standard deviations are nonlinear functions of  $\sigma_I$  ( $\sigma_R$ ) sensitive to the choice of  $\mu_I$  ( $\mu_R$ ). Moreover, an increase of  $\sigma$  increases (decreases) the mean value of the appropriate probabilities if  $\mu$  is smaller (larger) than 0.5, see below and Table 2. The infection and recovery probabilities per time step  $P_I(i)$  and  $P_R(i)$  have been generated at the system initiation.

A special type of the probability density, see equation (3), originates in the commonness of the normal distributions. The shape of the density follows a general intuition that there are some typical (modal) values of infection and recovery rates. Moreover, within the population there are individuals whose traits take larger or smaller values than typical (most likely) value. Finally, from the simulation point of view, alternations of parameters in equation (3) allow for an easy modification of the most probable (modal) values and transition from full homogeneity ( $\sigma \rightarrow 0$ ) to full heterogeneity ( $\sigma \rightarrow \infty$ ). Thus, the special type of the distribution (3) allows easy control of population properties.

Epidemics spread on the square lattice of size  $101 \times 101$ . Every outbreak is characterized by its size  $s$ , duration  $\mathcal{T}$  and radius  $r$ , which is the Manhattan distance between the first infected (patient zero) individual and the most distant infected agent. Initially, there is only one infected individual which is placed in the center of the lattice. Each agent interacts locally only with its four nearest neighbors in its von Neuman neighborhood. The model is studied by means of Monte Carlo methods [25]. Every MC step consists of: (i) infecting all susceptible individuals (having infected neighbors) and (ii) spontaneous recovery of all infected individuals according to the mechanisms described above. The MC steps are repeated as long as there are infectious individuals in the system. The number of performed steps defines the epidemic duration  $\mathcal{T}$ . Simulations were repeated  $10^4$  times for each set of  $\mu_I$ ,  $\sigma_I$ ,  $\mu_R$  and  $\sigma_R$  parameters. Furthermore, two different boundary conditions were considered: free and periodic. The main scope is to find how the population heterogeneity affects dynamics of epidemics (severity, duration) and properties of final states.

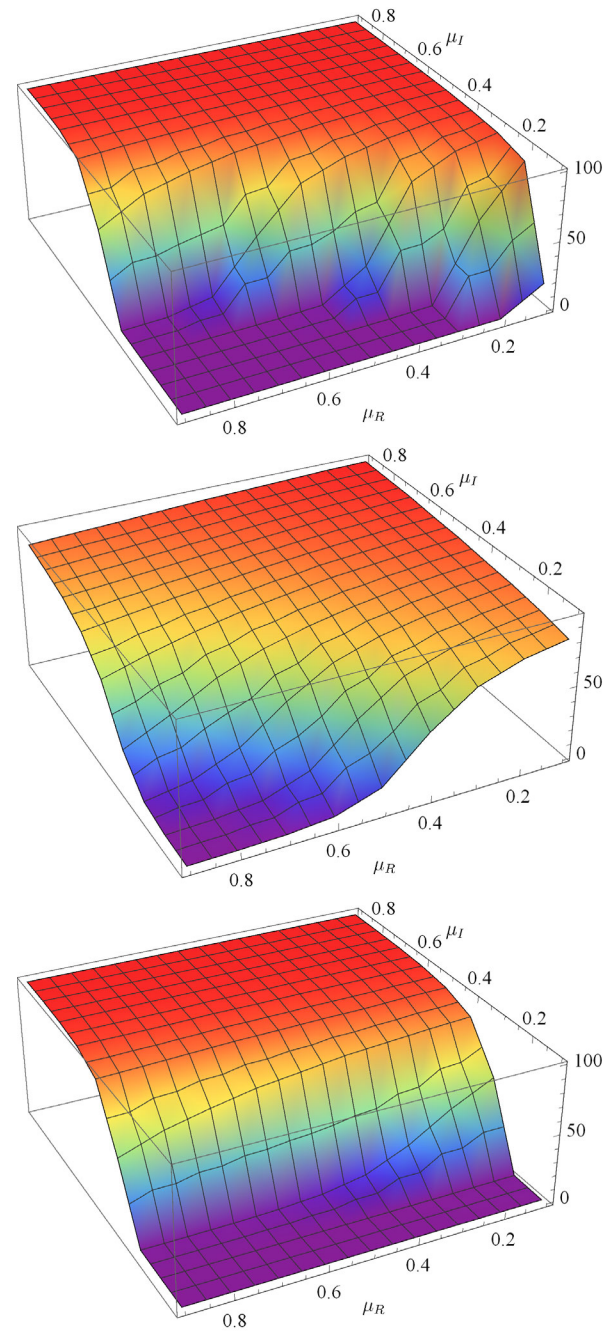
### 3 Results

The epidemic dynamics depends on the four model's parameters  $\mu_I, \sigma_I, \mu_R, \sigma_R$ , which control the shape of distributions from which epidemics' parameters are generated. The mean value of the infection probability per time step  $\langle P_I \rangle$  (recovery probability per time step  $\langle P_R \rangle$ ) grows with the increase of  $\mu_I$  ( $\mu_R$ ). Moreover, for  $\mu_I < 0.5$  ( $\mu_R < 0.5$ ) the mean value of  $\langle P_I \rangle$  ( $\langle P_R \rangle$ ) is an increasing function of  $\sigma_I$  ( $\sigma_R$ ), while for  $\mu_I > 0.5$  ( $\mu_R > 0.5$ ) it is a decreasing function of  $\sigma_I$  ( $\sigma_R$ ). In the limiting case of  $\mu_I = 0.5$  ( $\mu_R = 0.5$ ), the modal value is equal to the mean value  $\langle P_I \rangle$  ( $\langle P_R \rangle$ ). Standard deviations of infection and recovery probabilities are weakly sensitive to the exact choice of  $\mu_I$  and  $\mu_R$ , but they are increasing functions of  $\sigma_I$  and  $\sigma_R$ . Values of these parameters determine the average epidemic size, see Figure 1, and the mean duration, see Figure 2. Figures 1 and 2 present results as a function of modal values of the infection probability per time step  $\mu_I$  and the modal value of the recovery probability per time step  $\mu_R$  for fixed values of  $\sigma_I$  and  $\sigma_R$  for free boundary conditions. One can expect that a small infection probability and a large recovery probability can reduce the number of infected individuals, i.e. the epidemic size. This indeed is visible in Figure 1. Moreover, the small epidemic size is associated with the short epidemic duration, see Figure 2.

The top panel of Figure 1 demonstrates results for a homogeneous population ( $\sigma_I = \sigma_R = 0$  resulting in  $P_I = \mu_I$  and  $P_R = \mu_R$ ), while the middle and bottom panels for inhomogeneous ( $\sigma_I > 0$  and  $\sigma_R > 0$ ) cases. The inhomogeneity of infection and recovery rates changes the domain in which epidemic outbreaks are observed. Heterogeneity seriously affects the duration of epidemics, which is significantly longer than for homogeneous populations, compare the top, middle and bottom panels of Figure 2. The differences in epidemic progress are due to variability of individuals. The variability of the infection rate determines which individuals can be easily infected and which individuals are resistant. Resistant individuals are capable of building firewalls [14], while irrisistant individuals determine paths along which epidemics can easily spread. Figure 1 indicates that increasing  $\sigma_I$ , in line with earlier investigations [10,23], might decrease the epidemic size, see the top, bottom and middle panels of Figure 1.

Subsequently, Figures 3–6 display epidemic characteristics (the distribution of epidemic size  $s$  and radius  $r$ , the average first time  $\langle T(s) \rangle$  to infect a chosen size  $s$ , the average time  $\langle T(r) \rangle$  to infect a given radius  $r$  for the first time, and the distribution of epidemic duration  $\mathcal{T}$ ) for periodic (full symbols) and free (empty symbols) boundary conditions. Different figures correspond to various model parameters. These figures demonstrate in detail how heterogeneity changes epidemic dynamics. Consequently, they are used to explain the observed phenomena.

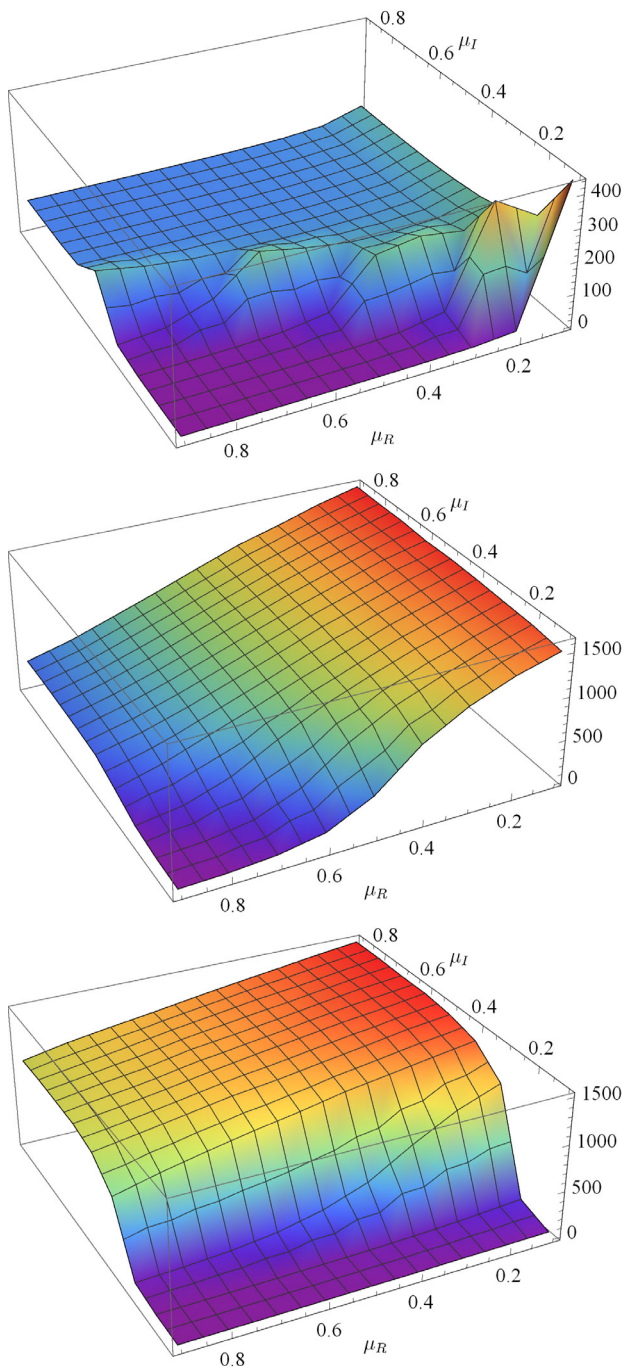
First of all, heterogeneity changes the mean value of probabilities  $\langle P_I \rangle$  and  $\langle P_R \rangle$  except the case when  $\mu_I = 0.5$  and  $\mu_R = 0.5$ . It causes significant differences in dynamics of epidemics between sets of parameters with very similar values of  $\sigma$ . For example, it can be seen for  $\mu_I = 0.2$ ,  $\mu_R = 0.1$ ,  $\sigma_I = 0.1$  with  $\sigma_R = 0.2$  (Fig. 3) and  $\sigma_R = 0.3$



**Fig. 1.** The average epidemic size  $\langle s \rangle$  (as a percentage of the system size) as a function of modal values of infection  $\mu_I$  and recovery  $\mu_R$  probabilities per time step for  $\sigma_I = 0, \sigma_R = 0$  (top panel),  $\sigma_I = 0.5, \sigma_R = 0.3$  (middle panel) and  $\sigma_I = 0.1, \sigma_R = 0.5$  (bottom panel).

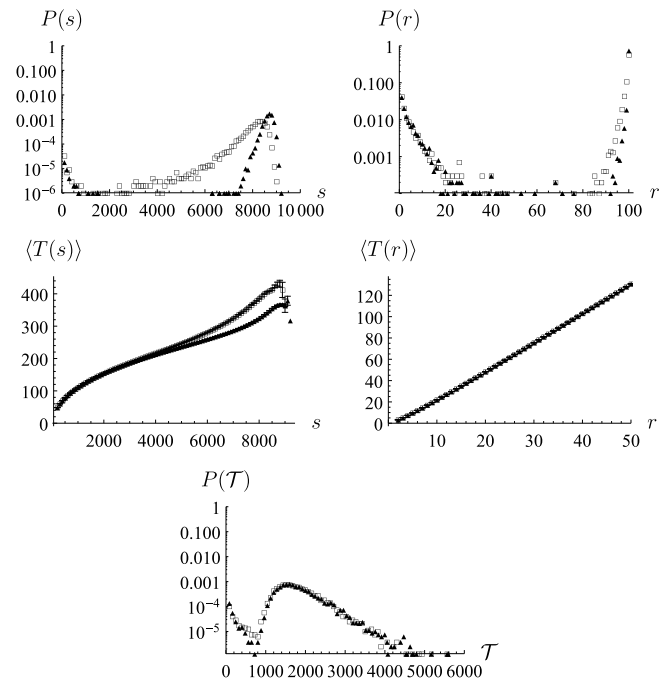
(Fig. 4). For the lower value of  $\sigma_R$ , if a disease does not end on the first individual, it spreads through almost 90% of a population. In this case an epidemic either reaches a radius smaller than 4 or equal to the maximum possible radius ( $r = 100$ ), which is the Manhattan distance from the initially infected individual to the agent located in the grid corner. For  $\sigma_R = 0.3$ , see Figure 4, the description is drastically changed. The probability that an epidemic





**Fig. 2.** The average epidemic duration  $\langle T \rangle$  as a function of modal values of infection  $\mu_I$  and recovery  $\mu_R$  probabilities per time step for  $\sigma_I = 0, \sigma_R = 0$  (top panel),  $\sigma_I = 0.5, \sigma_R = 0.3$  (middle panel) and  $\sigma_I = 0.1, \sigma_R = 0.5$  (bottom panel).

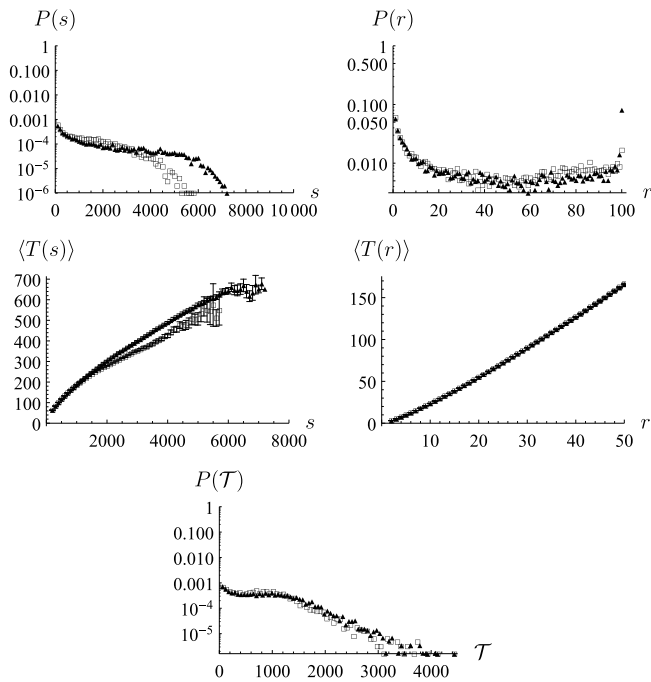
reaches a given size decreases when the epidemic size is growing. Most epidemics still end on the lowest or the largest radius, but it is more likely to observe the final radius somewhere between these two limiting values. Moreover, epidemics spread faster, as measured by  $\langle T(s) \rangle$  and  $\langle T(r) \rangle$ , when heterogeneity in  $P_R$  is lower, i.e. when  $\sigma_R$  is reduced and consequently the variance of  $P_R$  is smaller. The epidemic duration displays a different pattern because



**Fig. 3.** Distributions of the epidemic size  $s$  (top left panel) and radius  $r$  (top right panel). The middle panel shows the average time  $\langle T \rangle$  needed to infect a domain of the size  $s$  (middle left panel) or the radius  $r$  (middle right panel) for the first time. The bottom panel shows the distribution of epidemic duration. Full symbols correspond to periodic while empty symbols to free boundary conditions. Simulation parameters  $\mu_I = 0.2$ ,  $\mu_R = 0.1$ ,  $\sigma_I = 0.1$  and  $\sigma_R = 0.2$ .

it is not only determined by the infection but also by the recovery, see below.

Different effects are observed for  $\mu_I = 0.1$ ,  $\mu_R = 0.5$ ,  $\sigma_I = 0.5$  with  $\sigma_R = 0.2$  (Fig. 5) and  $\sigma_R = 0.5$  (Fig. 6). For these parameters the mean value of the recovery probability is equal to  $\mu_R$  ( $\langle P_R \rangle = \mu_R$ ). At the same time  $\langle P_I \rangle$  is different than  $\mu_I$  but it is the same in both figures. The increase of  $\sigma_R$  increases the probability that epidemics reach at least one of the lattice corners (maximum radius). Even a greater  $\sigma_R$  increases the chances that epidemics spread beyond the initially infected individuals and reach large final sizes. For  $\sigma_R \geq 0.3$ , more outbreaks reach a modal value of epidemic size and the modal value increases, see Figure 6. The rate of epidemic spread is also slightly increased, which is manifested by the decreasing time needed to reach the same size or radius. This behavior can be explained by the heterogeneity of a population, which makes individuals with extreme recovery rates appear more often in the population. If one of the initially infected individuals has a high recovery probability (per time step)  $P_R$ , an epidemic ends early reaching a small size. In the opposite case, every individual with a low recovery probability per time step has more opportunities to infect neighbors, thus chances of spreading a disease through neighborhood are higher, even if neighbors are characterized by a low infection rate. This dependence can be found for every set of parameters when the mean value

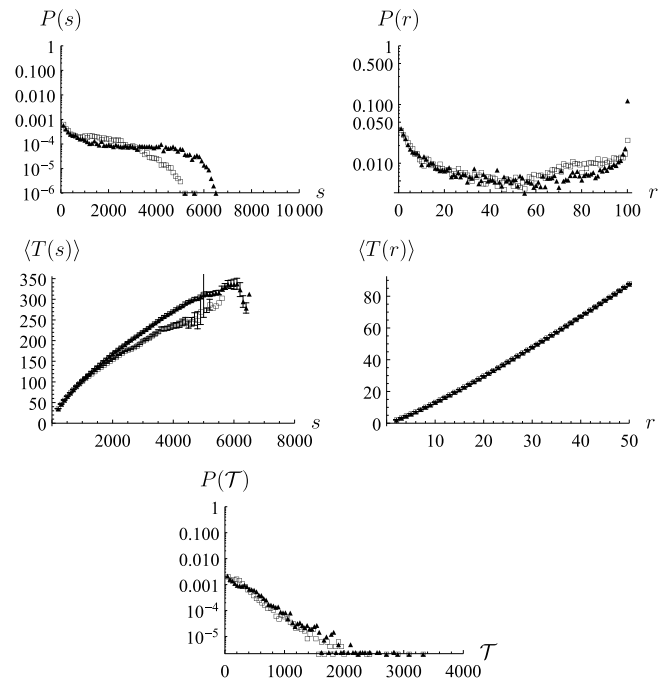


**Fig. 4.** The same as in Figure 3 for  $\sigma_R = 0.3$ .

of the recovery probability per time step is equal to the modal value  $\langle P_R \rangle = \mu_R = 0.5$  and the mean value of the infection probability per time step is constant. Nevertheless, for many sets of parameters for which this effect appears, it can be hard to recognize due to influence of  $P_I$ .

The situation when the mean value of the infection probability per time step is equal to the modal value  $\langle P_I \rangle = \mu_I = 0.5$  and the mean value of the recovery probability per time step is constant can be also easily analyzed. In this case the modal value of epidemic size is decreasing with the increasing heterogeneity of the infection rate. Moreover, the rate of epidemics' spread is lower.

The top panels of Figures 3 and 4 show spatial properties of final states. The top left panels demonstrate epidemic size distributions  $P(s)$ , while the top right panels epidemic radii distributions  $P(r)$  i.e. the Manhattan distance between the initially infected individual and the most distant infected agent. The maximal possible distance is obtained when the outbreak reaches corners of the grid, i.e.  $r = 100$  for the lattice  $101 \times 101$ . Depending on parameters' values, epidemic outbreaks can be characterized by a typical size and radius. In the epidemic size distribution and epidemic radii distribution there are peaks corresponding to small sizes and small radii. These peaks are due to epidemics which end at a very small size due to the immediate recovery. Moreover, if the recovery rate is large enough and the infection rate is small enough, epidemics die out rapidly giving rise to a fast decay of the epidemic size distribution. Nevertheless, there are situations when despite a significant probability that epidemics end immediately there are many outbreaks (realizations) infecting the majority of the population, see Figure 3. Surprisingly, this behavior can be significantly changed by a small change in  $\sigma_R$ , see Figure 4. In Figure 4, all param-

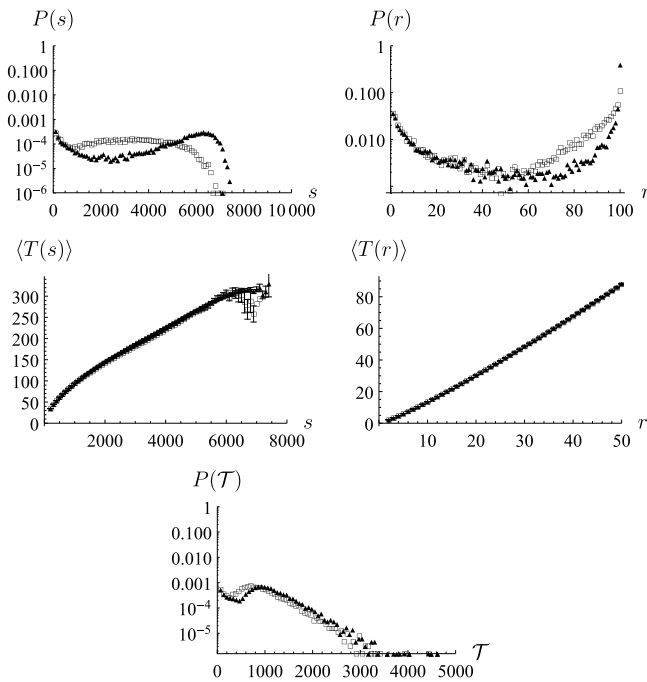


**Fig. 5.** Distributions of the epidemic size  $s$  (top left panel) and radius  $r$  (top right panel). The middle panel shows the average time  $\langle T \rangle$  needed to infect a domain of the size  $s$  (middle left panel) or the radius  $r$  (middle right panel) for the first time. The bottom panel shows the distribution of epidemic duration. Full symbols correspond to periodic while empty symbols to free boundary conditions. Simulation parameters  $\mu_I = 0.1$ ,  $\mu_R = 0.5$ ,  $\sigma_I = 0.5$  and  $\sigma_R = 0.2$ .

eters except  $\sigma_R$ , are the same as in Figure 3, which is set to  $\sigma_R = 0.3$ . Therefore, it is slightly larger than in Figure 3 where  $\sigma_R = 0.2$ . The increase in  $\sigma_R$  results in the increase of  $\langle P_R \rangle$  which determines the recovery rate. Consequently, the recovery rate is large enough to completely diminish the pronounced peak visible in Figure 3 around  $s \approx 8000$ . This also implies that epidemics reach mainly small radii. In Figures 5 and 6, the average infection and recovery probabilities per time step are larger than in Figures 3 and 4, although, on average, epidemics are less severe because larger chances of infection are compensated by a faster recovery. Only in Figure 6 a local maximum of  $P(s)$  is well-visible especially for periodic boundary conditions. Otherwise,  $P(s)$  is a decreasing function of the epidemic size  $s$ .

In general, the distance distribution is not only a function of the model's parameters but also of a system size. In order to verify the sensitivity of obtained results to the system size some of simulations have been performed for the two times larger system. When epidemics infect large fraction of individuals, like in Figures 3 and 6, there are only minor quantitative differences. Otherwise, in situations when epidemic outbreaks are limited, like in Figures 4 and 5, despite quantitative differences qualitative behavior is the same.

The middle panels of Figures 3–6 demonstrate temporal properties of epidemic spread. The middle left panels present the average first time  $\langle T(s) \rangle$  to infect a given



**Fig. 6.** The same as in Figure 5 for  $\sigma_R = 0.5$ .

size  $s$ , while the middle right panels the average first time,  $\langle T(r) \rangle$ , to reach a given radius  $r$ . The mean first time  $\langle T(r) \rangle$  needed to infect a cluster of the radius  $r$  scales almost linearly with  $r$ . This can be intuitively explained by the fact that in every time step the radius can increase by 1 at most giving rise to the linear growth of  $r$ . In practical situations the slope is larger than 1 because the (average) infection probability is smaller than 1. Moreover, the recovery process further slows down the growth rate of the infected region. An increase in the average recovery probability per time step  $\langle P_R \rangle$  leads to a slower growth of the  $\langle T(r) \rangle$ , compare Figures 3 and 4. Analogously, the rate of the epidemic spread is larger in Figures 5 and 6 than in Figures 3 and 4 due to the increase in the average infection probability per time step  $\langle P_I \rangle$  which facilitates the epidemic spread, see the middle panels of Figures 3–6 and Table 1.

The growth of the mean first time to reach a given size  $s$  is limited by topological properties of the system. The type of the neighborhood restricts the maximal possible growth rate of the epidemic size which is further reduced by the recovery process and finite (smaller than 1) infection probability. Initially,  $\langle T(s) \rangle$  grows like  $s^{1/2}$ , because the epidemic size grows with the distance  $r$  to the initially infected individual as  $r^2$  and the radius grows (approximately) linearly in time, i.e.  $r \propto t$ . Here, analogously, like for  $\langle T(r) \rangle$ , the increase in the mean infection probability per time step  $\langle P_I \rangle$  increases the growth rate, see for example Figures 4 and 5. The increase in  $\langle P_R \rangle$  reduces the growth rate, see Figures 3 and 4.

Finally, the bottom panels of Figures 3–6 show distribution of the epidemic duration  $P(T)$ . The exact shape of  $P(T)$ s depends on the model parameters, which also determine the size and the rate of epidemic spread and

**Table 1.** Average values of parameters and characteristic quantifiers for Figures 3–6: the average infection probability per time step  $\langle P_I \rangle$ , the average recovery probability per time step  $\langle P_R \rangle$ , the average transmissibility  $\langle \Psi \rangle$ , the average epidemic size  $\langle s \rangle$ , the average epidemic duration  $\langle T \rangle$  and the (linear) slope  $\alpha$  in the average time to reach a given radius  $r$ , i.e.  $\langle T(r) \rangle = \alpha r + \beta$ .

	Figure 3	Figure 4	Figure 5	Figure 6
$\mu_I$	0.2	0.2	0.1	0.1
$\sigma_I$	0.1	0.1	0.5	0.5
$\mu_R$	0.1	0.1	0.5	0.5
$\sigma_R$	0.2	0.3	0.2	0.5
$\langle P_I \rangle$	$0.21 \pm 0.09$	$0.21 \pm 0.09$	$0.38 \pm 0.25$	$0.38 \pm 0.25$
$\langle P_R \rangle$	$0.20 \pm 0.13$	$0.27 \pm 0.19$	$0.5 \pm 0.19$	$0.5 \pm 0.26$
$\langle \Psi \rangle$	$0.58 \pm 0.22$	$0.52 \pm 0.23$	$0.52 \pm 0.27$	$0.54 \pm 0.28$
$\langle s \rangle$	$6152 \pm 34$	$808 \pm 11$	$935 \pm 11$	$2139 \pm 19$
$\langle T \rangle$	$1405.8 \pm 8.8$	$561.3 \pm 6.1$	$196.0 \pm 2.4$	$638.0 \pm 5.7$
$\alpha$	2.71	3.57	1.88	1.85

in turn the epidemic duration  $\mathcal{T}$ . The general shape of  $P(T)$  curves follow predictions of [32]. If epidemics do not finish immediately at small sizes, they are likely to last long. The duration of epidemics is determined by two main factors. The first factor is the moment when the last infection takes place, which is the epidemic first time to reach its final size. The last infection is followed by the recovery of infected individuals only. After reaching the maximal size (maximal value of  $R + I$ ), the remaining duration time depends on the recovery of infected individuals only. For non-homogeneous populations, the rate of the recovery process is determined by individuals with extremely low recovery probabilities per time step, which significantly contribute to the epidemic duration because the average time that an individual needs to recover is  $\propto 1/P_R$ . Such individuals always exist in inhomogeneous populations. Moreover, the growth in the distribution of  $P_R$  width increases the chances of observing individuals with the very low recovery rates. This relation is visible in Figures 3–6 and Table 1, namely in situations when an outbreak is large. Then its duration is likely to be extended.

The probabilities (per time step) of becoming infected  $P_I$  and recovery  $P_R$  have been adjusted at the beginning of simulations. Consequently, the same set of parameters has been used for all realizations. Nevertheless, we have checked whether the model properties change when  $P_I$  and  $P_R$  vary among realizations. We have not observed any differences between these two ways of assigning probabilities (results not shown).

In order to verify the role of boundary conditions, simulations with free and periodic boundary conditions have been performed. Boundary conditions have an observable influence when epidemics reach edges of the lattice. In this case, the epidemic size is characterized by a larger modal value for periodic boundary conditions than for free boundary conditions. Also for the periodic boundary conditions the modal value is reached by more outbreaks but fewer epidemics end at an intermediate size between 0 and the mode. This difference is related to the number

of trajectories connecting individuals on the edge of the lattice with any other individual. Otherwise, quantitative differences between results for free and periodic boundary conditions are not significant, compare filled and empty points in Figures 3–6.

The condition when the epidemic outbreak is observed can be deduced from transmissibility [11,12]. The transmissibility  $\Psi(P_I, P_R)$  is a probability that an epidemic is transmitted from an infected individual to a susceptible one, i.e. it is a probability of transmission along the link connecting these two agents. It plays a similar role to the bond or site opening probability in the percolation theory [33,34]. The transmissibility takes into account the infection and the recovery processes, as well as inhomogeneity of the population. If  $P_I$  is the infection probability (per time step) and  $P_R$  is the recovery probability (per time step), the transmissibility  $\Psi(P_I, P_R)$  is

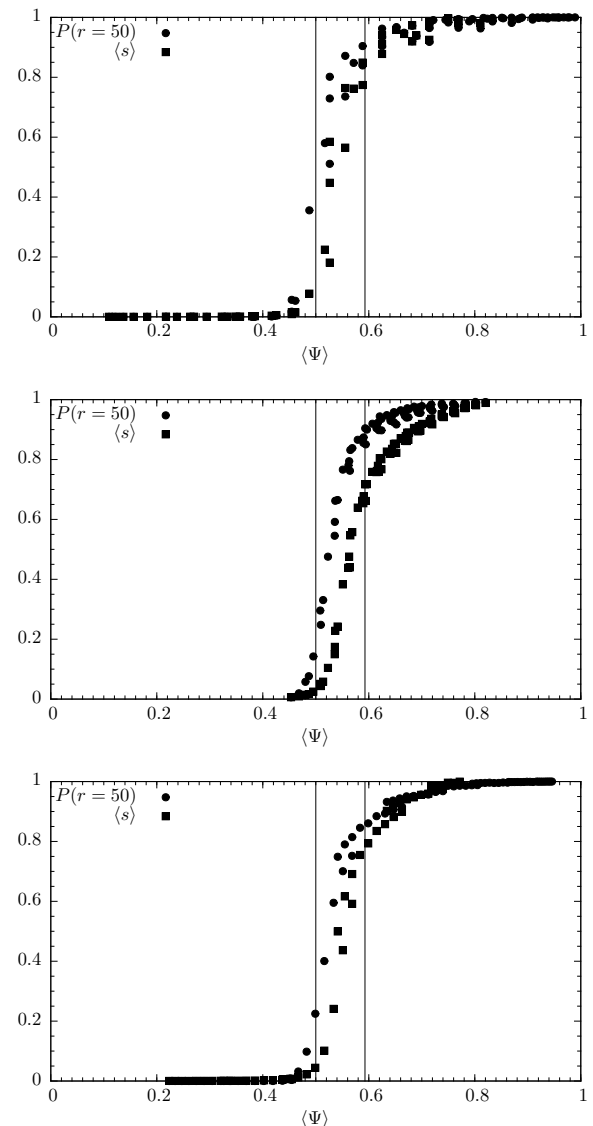
$$\begin{aligned} \Psi(P_I, P_R) &= \sum_{k=1}^{\infty} (1 - P_R)^{k-1} P_R \times [1 - (1 - P_I)^k] \\ &= \frac{P_I(P_R - 1)}{(1 - P_R)(P_I P_R - P_I - P_R)}. \end{aligned} \quad (4)$$

In equation (4),  $(1 - P_R)^{k-1} P_R$  is a probability that an infected individual stays infectious for  $k$  iterations and  $1 - (1 - P_I)^k$  is the probability of infecting a susceptible individual during these  $k$  consecutive iterations. An inhomogeneous system can be characterized by the average transmissibility

$$\langle \Psi \rangle = \int_0^1 \int_0^1 \Psi(P_I, P_R) f(P_I, \mu_I, \sigma_I) \times f(P_R, \mu_R, \sigma_R) dP_I dP_R, \quad (5)$$

where  $f(x, \mu, \sigma)$  is the  $P_I$  and  $P_R$  distribution given by equation (3). The average transmissibility is calculated numerically. Figure 7 presents a percolation-like analysis corresponding to the results depicted in Figures 1 and 2. The figure shows two characteristics used in order to detect the percolation: the average epidemic size  $\langle s \rangle$  and the probability that epidemics reaches a given radius, i.e.  $r = 50$ . Due to the fact that epidemics start from a single infected site, the average epidemic size is equivalent to the probability that a given site belongs to the largest (dominating) cluster, which is used as a standard percolation quantifier [34]. Figure 7 suggests that epidemic start to be observed for the average transmissibility located somewhere between the bond (0.5) and the site percolation (0.592) thresholds, which is in accordance with earlier theoretical and experimental considerations [11,12]. Mean values of the model parameters along with selected properties of final states are included in Table 1.

Complementary to the previous analysis based on infection and recovery rates, the average transmissibility can be used to explain phenomena depicted in Figures 3–6. The average transmissibility  $\langle \Psi \rangle$  in the subsequent figures is equal to 0.58, 0.52, 0.52, 0.54, see Table 1. The change in  $\sigma_R$  from  $\sigma_R = 0.2$  to  $\sigma_R = 0.3$ , see Figures 3 and 4,



**Fig. 7.** The probability that a given site belongs to the largest cluster ( $\blacksquare$ ) and the probability that an epidemic leaves a circle with  $r = 50$  ( $\bullet$ ) for  $\sigma_I = 0, \sigma_R = 0$  (top panel),  $\sigma_I = 0.5, \sigma_R = 0.3$  (middle panel) and  $\sigma_I = 0.1, \sigma_R = 0.5$  (bottom panel) as a function of the average transmissibility  $\langle \Psi \rangle$ . Straight solid lines represent the bond (0.5) and the site (0.592) percolation thresholds. Free boundary conditions are assumed.

decreases the  $\langle \Psi \rangle$  from 0.58 (a value close to the site percolation) to 0.52 (a value close to the bond percolation) explaining changes in epidemic severity as measured by the epidemic size distribution  $P(s)$ . In Figures 5 and 6, the increase in  $\sigma_R$  resulted in the increase of the average transmissibility. This in turn is responsible for higher chances of observing outbreaks characterized by larger radii and sizes.

The studied model displays non-trivial sensitivity to its parameters. Sometimes general conclusions can be difficult to draw because changes in heterogeneity usually also affect the mean values of parameters. Therefore, as indicated



**Table 2.** The dependence of the average transmissibility  $\langle\Psi\rangle$  on  $\sigma_I$  and  $\sigma_R$  as a function of  $\mu_I$  and  $\mu_R$ .

	$0 < \mu_I < 0.5$	$0.5 \leq \mu_I < 1$
$0 < \mu_R < 0.5$	$\sigma_I \uparrow: \langle P_I \rangle \uparrow, \langle \Psi \rangle \uparrow$	$\sigma_I \uparrow: \langle P_I \rangle \downarrow, \langle \Psi \rangle \downarrow$
	$\sigma_R \uparrow: \langle P_R \rangle \downarrow, \langle \Psi \rangle \uparrow$	$\sigma_R \uparrow: \langle P_R \rangle \downarrow, \langle \Psi \rangle \uparrow$
$0.5 \leq \mu_R < 1$	$\sigma_I \uparrow: \langle P_I \rangle \uparrow, \langle \Psi \rangle \uparrow$	$\sigma_I \uparrow: \langle P_I \rangle \downarrow, \langle \Psi \rangle \downarrow$
	$\sigma_R \uparrow: \langle P_R \rangle \uparrow, \langle \Psi \rangle \downarrow$	$\sigma_R \uparrow: \langle P_R \rangle \uparrow, \langle \Psi \rangle \downarrow$

in the manuscript, it can be difficult to distinguish whether an observable effect can be attributed to changes in average values of parameters or to their standard deviations.

Changes in the average transmissibility responsible for differences among Figures 3–6 are due to changes in the average values of the infection  $\langle P_I \rangle$  and recovery  $\langle P_R \rangle$  probabilities per time step. These changes can be introduced by modifications of parameters of the probability density (3). Table 2 indicates how the average transmissibility  $\langle\Psi\rangle$  changes with changes in  $\sigma_I$  and  $\sigma_R$  depending on the modal values  $\mu_I$  and  $\mu_R$ . The dependence of the average transmissibility on  $\mu_I$  and  $\mu_R$  is simpler: it is an increasing function of  $\mu_I$  and a decreasing function of  $\mu_R$ .

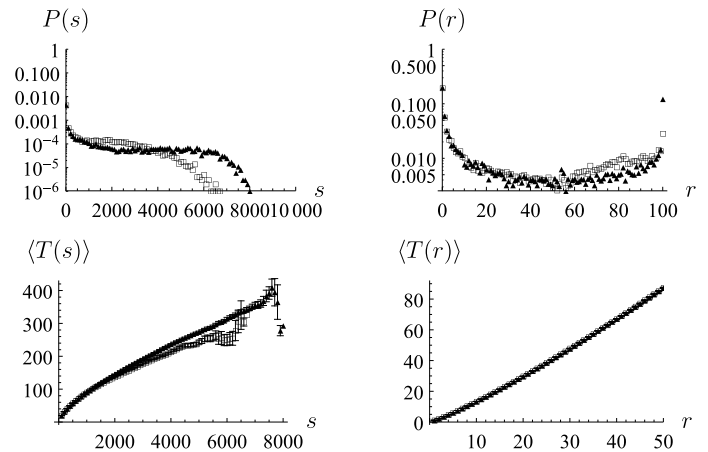
Within the model it is assumed that the probability per time step of infecting a susceptible individual is assigned to this particular susceptible individual. Therefore, we have considered the modified version of the model in which probabilities (per time step) to infect a susceptible agent are assigned to its infectious neighbors. In such a case the probability of infecting a susceptible individual  $i$  during a single iteration is not given by equation (1) but by the following formula

$$P_{S \rightarrow I}(i) = 1 - \prod_{j \in \mathcal{N}(i)} [1 - P_I(j)]^{\Theta(j)}, \quad (6)$$

where the interpretation of the parameters is the same as in equation (1). As it can be seen in Figure 8, the modified model leads to qualitatively the same results as the basic model, see equation (1) and Figure 5. On the one hand, the epidemic radius distribution  $P(r)$  and the mean first passage time to reach a given radius  $\langle T(r) \rangle$  are practically the same. On the other hand, the size distribution  $P(s)$  and the first passage time  $T(s)$  to infect the domain of the size  $s$  are enlarged due to the increase in the epidemic size. This can be explained by the fact that the modified model results in a larger average infection probability per time step  $\langle P_{S \rightarrow I} \rangle$ . If a susceptible individual  $i$  has  $n$  ( $n > 1$ ) infected neighbors,  $\langle P_{S \rightarrow I} \rangle = 1 - (1 - \langle P_I \rangle)^n$  for the modified model is larger than for the basic model  $\langle P_{S \rightarrow I} \rangle = 1 - \langle (1 - P_I)^n \rangle$ . At the same time the average transmissibilities, see equation (5), stay the same.

## 4 Summary and conclusions

The SIR model [2] is an archetypal epidemiological model, which is almost a hundred years old. Despite its simplicity



**Fig. 8.** The same as in Figure 5 for the modified model, see equation (6). Full symbols correspond to periodic while empty symbols to free boundary conditions. Simulation parameters  $\mu_I = 0.1$ ,  $\mu_R = 0.5$ ,  $\sigma_I = 0.5$  and  $\sigma_R = 0.2$ .

the Kermack-McKendrick model can be used to describe complex and complicated situations even in realistic realms. Within the current manuscript, excessive computer simulations have been used to investigate how the heterogeneity of a population affects temporal and spatial properties of epidemics on lattices.

Variable infection and recovery rates result in non-trivial system properties. On the one hand, individuals which can be easily infected create paths along which epidemics can propagate. On the other hand, resistant individuals make impenetrable firewalls responsible for the creation of islands (patches) of uninfected individuals. Therefore, the system is sensitive to exact values of the mean infection probabilities per time step and (mean) transmissibilities. A minimal change in one of the system parameters can significantly modify mean values of parameters. Consequently, as it was demonstrated in Section 3, epidemic spread can be facilitated or reduced.

The distribution of epidemic radii has two maxima corresponding to epidemics which die out fast (small  $r$ ) and those that propagate to peripheral parts of the system (large  $r$ ). The mean first time that an epidemic needs to reach a given radius grows almost linearly with  $r$ . The epidemic size distribution always has a maximum corresponding to small outbreaks, but it can also have another local maximum corresponding to large epidemics.

Within the model epidemics end spontaneously, because there are no mechanisms of disease eradication. Further extensions of the model could introduce such mechanisms in order to provide answers to questions such as what the optimal control strategy is and how eradication strategies are affected by heterogeneity of individuals.

Computer simulations have been performed at the Academic Computer Center Cyfronet, Akademia Górniczo-Hutnicza (Kraków, Poland). Suggestions from Adam Kleczkowski are greatly acknowledged.



## Author contribution statement

Both authors participated in developing the model, drafting and revising the manuscript as well in analyzing and interpreting the data. K.C. carried out simulations and prepared the graphics.

**Open Access** This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## References

1. R.M. Anderson, R.M. May, in *Infectious diseases of humans: Dynamics and control* (Oxford University Press, Oxford, 1991)
2. H.W. Hethcote, *SIAM Rev.* **42**, 599 (2000)
3. T. Britton, *Math. Biosci.* **225**, 24 (2010)
4. W.O. Kermack, A.G. McKendrick, *Proc. R. Soc. A* **115**, 700 (1927)
5. B. Dybiec, A. Kleczkowski, C.A. Gilligan, *Phys. Rev. E* **70**, 066145 (2004)
6. J.M. Hyman, J. Li, *Math. Biosci. Eng.* **3**, 89 (2006)
7. F. Ball, *Adv. Appl. Probab.* **17**, 1 (1985)
8. V.M. Veliov, *J. Math. Biol.* **51**, 123 (2005)
9. A.S. Novozhilov, *Math. Biosci.* **215**, 177 (2008)
10. G. Katriel, *J. Math. Biol.* **65**, 237 (2012)
11. F.M. Neri, A. Bates, W.S. Füchtbauer, F.J. Pérez-Reche, S.N. Taraskin, W. Otten, D.J. Bailey, C.A. Gilligan, *PLoS Comput. Biol.* **7**, e1002174 (2011)
12. F.M. Neri, F.J. Pérez-Reche, S.N. Taraskin, C.A. Gilligan, *J. R. Soc. Interface* **8**, 201 (2010)
13. R. Hickson, M. Roberts, *J. Theor. Biol.* **350**, 70 (2014)
14. L.M. Sander, C.P. Warren, I.M. Sokolov, C. Simon, J. Koopman, *Math. Biosci.* **180**, 293 (2002)
15. L. Sander, C. Warren, I. Sokolov, *Physica A* **325**, 1 (2003)
16. S. Wolfram, *Rev. Mod. Phys.* **55**, 601 (1983)
17. S. Wolfram, *A new kind of science* (Wolfram Media, Champaign, IL, 2002)
18. J.C. Miller, *Phys. Rev. E* **76**, 010101(R) (2007)
19. Y. Moreno, R. Pastor-Satorras, A. Vespignani, *Eur. Phys. J. B* **26**, 521 (2002)
20. R. Pastor-Satorras, A. Vespignani, *Phys. Rev. Lett.* **86**, 3200 (2001)
21. M. Boguñá, R. Pastor-Satorras, A. Vespignani, *Phys. Rev. Lett.* **90**, 028701 (2003)
22. Z. Dezsó, A.L. Barabási, *Phys. Rev. E* **65**, 055103 (2002)
23. R.D. Boylan, *Math. Biosci.* **105**, 133 (1991)
24. V. Andreasen, *Bull. Math. Biol.* **73**, 2305 (2011)
25. M.E.J. Newman, G.T. Barkema, *Monte Carlo methods in statistical physics* (Oxford University Press, Oxford, 1999)
26. T.C. Reluga, *PLoS Comput. Biol.* **6**, e1000793 (2010)
27. S. Maharaj, A. Kleczkowski, *BMC Public Health* **12**, 679 (2012)
28. J. Swinton, C.A. Gilligan, *Proc. R. Soc. London Ser. B* **351**, 605 (1996)
29. A.J. Stacey, J.E. Truscott, M.J.C. Asher, C.A. Gilligan, *Phytopathology* **94**, 209 (2004)
30. M.J. Keeling, M.E.J. Woolhouse, D.J. Shaw, L. Matthews, M. Chase-Topping, D.T. Haydon, S.J. Cornell, J. Kappey, J. Wilesmith, B.T. Grenfell, *Science* **294**, 813 (2001)
31. C. Dye, N. Gay, *Science* **300**, 1884 (2003)
32. A.D. Barbour, *Biometrika* **62**, 477 (1975)
33. P. Grassberger, *Math. Biosci.* **63**, 157 (1983)
34. A. Aharony, D. Stauffer, *Introduction to percolation theory* (Taylor & Francis, London, 2003)