Simulation Modelling Practice and Theory 16 (2008) 1014–1023

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat

Modelling SIR-type epidemics by ODEs, PDEs, difference equations and cellular automata – A comparative study

G. Schneckenreither [a,*], N. Popper [a,b], G. Zauner [a,b], F. Breitenecker [a]

[a] Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstraße 8-10, 1040 Vienna, Austria
[b] 'Die Drahtwarenhandlung' – Simulation Services, Neustiftgasse 57-59, 1070 Vienna, Austria

ARTICLE INFO

Article history:
Received 16 December 2007
Received in revised form 21 April 2008
Accepted 26 May 2008
Available online 6 June 2008

Keywords:
Susceptible-infected-recovered model
Lattice gas cellular automaton
Stochastic cellular automaton
Partial differential equation
Diffusion distribution

ABSTRACT

The Kermack–McKendrick susceptible-infected-recovered (SIR) model describes the dynamics of epidemics in a cumulative way. This contribution compares different approaches for introducing spatial patterns into these dynamics. The applied techniques cover lattice gas cellular automata (LGCA), stochastic cellular automata (SCA) and partial differential equations (PDE). Even though these methods involve distinct types of spatial interaction, it can be shown, that consistent qualitative and quantitative model behaviour can be obtained by means of parameter adaptions and slight technical modifications. These modifications are motivated by stochastic analysis of distributed interaction (PDE, SCA) and diffusion dynamics (LGCA) as well as prevailing physical analogies. The law of large numbers permits to approximate stochastic contacts by distributed interaction. Diffusion of particles can be approximated through empiric adjustment of a Gaussian diffusion distribution.

1. Introduction

The first epidemiological considerations reach back to the 19th century and were conducted by John Snow, a Londoner doctor. In the course of a cholera epidemic in London Snow could detect the source of the outbreak – a contaminated well in Broad Street – by analysing the spatial and social environments of infected people [1]. Nowadays epidemiological considerations also comprise a priori analysis. Governments or health organisations are interested in the number of potentially affected persons, the required amount of vaccine or alternative counter measures.

A very prominent model for simulating spread of diseases is the Kermack–McKendrick SIR model (1927). The basic model assumption consists in the stratification of a closed population into the groups susceptible S, infected I and recovered R. Two parameters α and β describe the probability that in the case of interaction a susceptible individual becomes infected by a contagious (infected) individual, respectively that an infected individual recovers during one time unit (this leads to a geometric distribution of the duration of infection). The number of involved individuals (S, I, R) can be described by the following system of ordinary differential equations (ODE).

$$\begin{aligned} \frac{\partial S(t)}{\partial t} &= -\alpha S(t) I(t), \\ \frac{\partial I(t)}{\partial t} &= \alpha S(t) I(t) - \beta I(t), \\ \frac{\partial R(t)}{\partial t} &= \beta I(t) \end{aligned} \tag{1}$$

* Corresponding author.
E-mail address: gschneck@osiris.tuwien.ac.at (G. Schneckenreither).

1569-190X/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.simpat.2008.05.015

An implicit assumption of this approach is that mixing of the population is perfect, i.e. at any given time all individuals and especially infected and susceptible interact with each other. This becomes evident in the growth term of the group of infected people $\alpha S(t)I(t)$. For a more detailed simulation of epidemic spread, which incorporates spatial spread (limited contacts), a new model or at least some modifications are necessary.

The work presented in this paper emerged from a discussion of ARGESIM Comparison C17 – Temporal and Spatial Evolution of a SIR-type Epidemic [6,7]. Initially the main interest was to examine lattice gas cellular automaton (LGCA) models. Further inquiries gave raise to the idea of using different modelling techniques like stochastic cellular automata (SCA) [13] and partial differential equations (PDE) [3,8]. Because these approaches deliver distinct behaviours, the first question was whether there is an underlying consistency [11,12]. We show that all discussed methods can deliver the qualitative and quantitative behaviour of the SIR model (1). Furthermore with certain adaptions and modifications the spatial dynamics and the qualitative and quantitative behaviours of these approaches correspond to each other. The main difficulty concerning the latter adaptions is that spatial interaction is either generated by diffusion (LGCA) or distributed contacts (SCA and PDE).

Beneath the mentioned approaches stochastic methods like the mean field approach [9,10], agent-based modelling [4] or networks [1] can be applied for simulating epidemic spread but are not subjects of investigation in this paper.

2. LGCA model

For discussion of cellular automaton models, a few introductory words and consistent notations are necessary.

Cellular automaton (CA) [14,15] modelling discretises the observed domain or geographical region on a regular lattice structure. The resulting cells can hold different states, which change only at discrete time steps and for example describe the stage of infection of one or multiple individuals. To calculate the cell states in the next time step an automaton rule is applied on all cells simultaneously. This function is applied locally and depends on the current state of the cell itself and the 'neighbouring' cells. On a square lattice the neighbourhood can be for example the four nearest neighbours (von Neumann neighbourhood) or any stochastically determined set of cells (SCA).

A cellular automaton whose cells contain particles, that interact within the cells, is called lattice gas cellular automaton (LGCA). We denote LGCA with a hexagonal lattice and a capacity of six particles per cell as Frisch–Hasslacher–Pomeau (FHP) automata and LGCA with a square lattice and at most four particles per cell as Hardy–de Pazzis–Pomeau (HPP) automata. Each time step all particles leave their current cell and move in exclusive directions to the neighbouring cells. The moving directions are chosen either randomly or according to collision configurations, which are applied depending on the actual number of particles within the cell and their entering directions. Collision configurations that preserve mass and momentum can be summarised under the term FHP-I rules [14]. For all examinations throughout this paper we will stick to random motion in FHP LGCA.

Boundary conditions of CA determine the behaviour of particles, which hit the border of the domain. They are of minor importance in our observations but we stick to periodic boundary conditions in all approaches, i.e. any particle or interaction process, which leaves the domain at one side, reenters from the other side.

2.1. LGCA model and parameters

In order to apply the SIR LGCA model the population must be discretised on a two dimensional LGCA-lattice structure depending on the desired initial conditions. Each individual is represented by a particle and can hold one of the discrete states susceptible, infected or recovered. The state transition parameters α_{LG} and β are defined analogous to those in (1) but interaction is restricted to particles, which are located in the same cell.

The current probability of infection for a susceptible individual Ψ_c depends on the number of infected individuals or particles in the same cell I_c and can be written as

$$\Psi_c = 1 - (1 - \alpha_{LG})^{I_c}. \tag{2}$$

It is clear that the probability of infection Ψ_c is different for all cells.

The expected or average probability of infection Ψ (we assume a uniform distribution of the individuals) in the LGCA is

$$\Psi = 1 - (1 - \alpha_{LG})^{\frac{I}{N}}, \tag{3}$$

where I is the overall number of infected individuals and N is the number of cells. A Taylor series expansion [5,7] shows that

$$(1 - \alpha_{LG})^{\frac{I}{N}} = 1 - \frac{\alpha_{LG} I}{N} + \frac{\alpha_{LG}^2 I(I-N)}{2N^2} + \dots. \tag{4}$$

Accordingly for small α_{LG} the probability of infection can be approximated by $\Psi \approx \alpha_{LG} \frac{I(t)}{N}$, what forces us to set $\alpha_{LG} = \alpha N$ if we want a behaviour similar to the ODE model.

2.2. Consistency – LGCA model and SIR model

The following considerations proof that the time-discretisation of the ODEs is concerning speed of spread an upper bound and concerning spatial inhomogeneities a lower bound for FHP simulations.

In order to accelerate epidemic spread in the LGCA it is necessary to dissolve the local character of motion by repositioning the individuals randomly every time step [7]. Increasing the infection rate in order to accelerate the epidemic has no effect after a certain value has been reached, because the local character of contagion (restricted to the individuals within the same cell) limits the speed of spread. To dissolve the local character of motion, the lattice must be rearranged very often so that the average distribution of individuals, which then is uniform, can be applied to the lattice.

With the previously obtained (average) infection rate $\Psi = \alpha I(t)$, which is valid for all cells, and under the assumption of a uniformly distributed population the expected cumulative growth of the number of infected individuals in the LGCA approach is $S(t) \cdot \Psi = S(t) \cdot \alpha I(t)$. Consequently it is now possible to describe the evolution of the epidemic on the lattice by the difference equations system

$$\begin{aligned} S(t+1) &= S(t) - S(t) \cdot \alpha I(t), \\ I(t+1) &= I(t) + S(t) \cdot \alpha I(t) - I(t) \cdot \beta, \\ R(t+1) &= R(t) + I(t) \cdot \beta, \end{aligned} \tag{5}$$

which actually corresponds to the time-discretisation of the classical ODE model.

Accordingly by dissolving the local character and simultaneously reducing the step size, the LGCA converges towards the ODE model. Therefore, the LGCA model is a (discretised) extension of the classical model by a spatial character of disease propagation.

3. SCA model

By stochastic cellular automata (SCA) we refer to ordinary cellular automata (without considering motion of particles) with a stochastically determined neighbourhood. The application of SCA for modelling epidemic spread can be found in [13] for example.

To define a 'sociological' neighbourhood in the SCA model, that provides a gradation for the occurrence of interaction between individuals, a decaying likelihood of interaction between cells depending on the distance between them can be used. This approach delivers a radial-symmetric distribution of the contacts for each cell, what principally can be described by an arbitrary probability distribution or a similar function, which we denote likelihood functions.

It is not difficult to show that this approach extends the classical model by a spatial component. The ODEs are again an upper bound concerning the speed of spread and a lower bound for spatial inhomogeneities. If the SCA establishes contact between each two cells (dissolution of the local character and increase of speed), a probability of infection for every individual of $\Psi = \alpha_{SCA} I$ is obtained and consequently the growth of the number of infected individuals is $\alpha_{SCA} IS$, which corresponds to the discretised ODE system if $\alpha_{SCA} = \alpha$.

Finding a connection to the LGCA approach on the contrary is not straight forward. It is possible to find some rules of thumb concerning the weight factors for the SCA, but these rules deliver rather imprecise results and only apply to specific conditions. The difficulty lies in finding a tool that relates motion of particles in the LGCA with distributed contacts in SCA (Section 6.1).

4. PDE model

A classical spatial approach is by partial differential equations. The investigated PDEs can be found in works of Jones and Sleeman [8]. They are based on distributed contacts like the SCA approach and emerge from a Taylor series simplification of an integral equation system. The derivation of the second order simplification can be found in [3] for example, which also refers to [8].

4.1. Basic PDE model

Spatial inhomogeneities are introduced by distributing the individuals on a domain and observing the densities of each group $S(t,x,y)$, $I(t,x,y)$ and $R(t,x,y)$, respectively their change in time. For every location (x,y) on the domain $S(t,x,y) + I(t,x,y) + R(t,x,y)$ is always $\leqslant 1$.

Further a rate or likelihood of infection for every location is necessary. The basis for this is an interaction coefficient $\tilde{\alpha}(\vec{x},\vec{v})$ for each two points $\vec{x} = (x,y)$ and $\vec{v} = (v,w)$ on the domain, which provides the impact that these cells have on each other. As already mentioned in the section on SCA, a prototype for this 'likelihood' of interaction is a function that depends on the distance between the two points $\|\vec{v} - \vec{x}\|$. We only observe interaction functions that satisfy $\|\tilde{\alpha}\|_{L^1(\mathbb{R}^2)} = \alpha_{PDE} \leqslant 1$ and thus have similar features like probability distributions (compare [8]).

With the assumptions from above and the substitution $\vec{z} := \vec{v} - \vec{x}$, the probability of infection Ψ at a certain point $\vec{x}$ is

$$\Psi = \int \tilde{\alpha}(\vec{x},\vec{v}) I(\vec{v})\,\mathrm{d}\vec{v} = \int \tilde{\alpha}(\|\vec{v}-\vec{x}\|) I(\vec{v})\,\mathrm{d}\vec{v} = \int \tilde{\alpha}(\|\vec{z}\|) I(\vec{x}+\vec{z})\,\mathrm{d}\vec{z}. \tag{6}$$

In [3] the zeroth- and second-order Taylor series expansion of $I(\vec{x}+\vec{z})$ is observed in order to simplify the resulting integral equations ($\vec{x} = (x,y)$ and $\vec{z} = (u,v)$).

$$I(\vec{x}+\vec{z}) = I + uI_x + vI_y + \frac{1}{2}(u^2 I_{xx} + 2uvI_{xy} + v^2 I_{yy}) + \ldots \tag{7}$$

After simplification of the integral equation with the zeroth-order expansion the probability of infection $\Psi_0 = \alpha_{\mathrm{PDE}} I$ is the same as in the original ODE model if $\alpha_{\mathrm{PDE}} = \alpha$. Using the second-order expansion delivers

$$\Psi_2 = I\int \tilde{\alpha}\,\mathrm{d}z + \frac{1}{2} I_{xx}\int u^2\tilde{\alpha}\,\mathrm{d}z + \frac{1}{2} I_{yy}\int v^2\tilde{\alpha}\,\mathrm{d}z \tag{8}$$

for the probability of infection. Because of the radial symmetry of the likelihood function, integration of the terms $u\tilde{\alpha}$, $v\tilde{\alpha}$ and $uv\tilde{\alpha}$ delivers 0 so that they can be neglected. This is also the reason for $\Psi_1 = \Psi_0$.

Furthermore we only observe domains with finite range and assume that $r = \sqrt{u^2+v^2} \leqslant \rho$ or $\tilde{\alpha}(r) = 0$ for $r > \rho$. This allows to define a parameter γ with a finite number

$$\gamma = \frac{1}{2}\int u^2\tilde{\alpha}\,\mathrm{d}z = \frac{\pi}{2}\int_0^{\rho} r^3\tilde{\alpha}(r)\,\mathrm{d}r. \tag{9}$$

Using this substitution the probability of infection can be written as

$$\Psi_2 = I\alpha_{\mathrm{PDE}} + I_{xx}\gamma + I_{yy}\gamma = \alpha_{\mathrm{PDE}} I + \gamma\Delta I. \tag{10}$$

For a second order simplification of the probability of infection at a point $\vec{x}$, a (new) partial differential equations system, which describes the densities of the three groups, can be set up

$$\begin{aligned}
\frac{\partial S(t,\vec{x})}{\partial t} &= -S(t,\vec{x})(\alpha_{\mathrm{PDE}} I(t,\vec{x}) + \gamma\Delta I(t,\vec{x})),\\
\frac{\partial I(t,\vec{x})}{\partial t} &= S(t,\vec{x})(\alpha_{\mathrm{PDE}} I(t,\vec{x}) + \gamma\Delta I(t,\vec{x})) - \beta I(t,\vec{x}),\\
\frac{\partial R(t,\vec{x})}{\partial t} &= \beta I(t,\vec{x}).
\end{aligned} \tag{11}$$

With the Hölder inequality and because $\alpha_{\mathrm{PDE}} = \|\tilde{\alpha}\|_{L^1(\mathbb{R}^2)} = 2\pi\|r\tilde{\alpha}\|_{L^1[0,\rho]}$ an upper bound for γ can be calculated.

$$\gamma = \frac{\pi}{2}\cdot\|r^3\tilde{\alpha}(r)\|_{L^1[0,\rho]} \leqslant \frac{\pi}{2}\cdot\|r^2\|_{L^\infty[0,\rho]}\cdot\|r\tilde{\alpha}(r)\|_{L^1[0,\rho]} = \frac{1}{4}\rho^2\alpha_{\mathrm{PDE}}. \tag{12}$$

4.2. Refined PDE model

Generally the probability of infection can be approximated with the Taylor series expansion of arbitrary order n, what delivers a differential equations system of the form

$$\begin{aligned}
\frac{\partial S(t,\vec{x})}{\partial t} &= -S(t,\vec{x})\Psi_n(t,\vec{x}),\\
\frac{\partial I(t,\vec{x})}{\partial t} &= S(t,\vec{x})\Psi_n(t,\vec{x}) - \beta I(t,\vec{x}),\\
\frac{\partial R(t,\vec{x})}{\partial t} &= \beta I(t,\vec{x}),
\end{aligned} \tag{13}$$

where $\Psi_n(t,\vec{x})$ is the probability of infection at a certain point $\vec{x}$ at time t, which has been approximated using the nth order Taylor series expansion I_n of the density of infected individuals I.

In order to analyse the behaviour of Ψ_n when $n \to \infty$, some transformations are required ($\vec{x} = (x,y)$ and $\vec{z} = (u,v)$).

$$\Psi_n(t,\vec{x}) = \int \tilde{\alpha}(\|\vec{z}\|) I_n(\vec{x}+\vec{z})\,\mathrm{d}\vec{z} = \int \tilde{\alpha}(\|\vec{z}\|) \sum_{\substack{i+j\leqslant n\\ i,j\in 2\mathbb{N}}} \frac{\partial^i}{\partial x^i}\frac{\partial^j}{\partial y^j}\frac{I(x,y)}{i!j!} u^i v^j\,\mathrm{d}\vec{z} = \sum_{\substack{k=0\\ i+j=k}}^{\left\lfloor\frac{n}{2}\right\rfloor} \frac{\partial^{2i}}{\partial x^{2i}}\frac{\partial^{2j}}{\partial y^{2j}}\frac{I(x,y)}{(2i)!(2j)!}\int \tilde{\alpha}(\|\vec{z}\|) u^{2i}v^{2j}\,\mathrm{d}\vec{z} \tag{14}$$

In the second line of (14) a summation representation of I_n was used; simultaneously terms for which $i\cdot j$ is an odd number ($\int u^i v^j\tilde{\alpha} = 0$) were neglected. The third line results after rearrangement of the summation expression.

We propose that

$$\Psi_n(t,\vec{x}) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \gamma_{2k} \Delta^k I(x,y), \tag{15}$$

where

$$\gamma_{2k} = \frac{1}{(2k)!} \int \tilde{\alpha}(\|\vec{z}\|) u^{2k} \, \mathrm{d}\vec{z} = \frac{2\pi}{(2k)!} \prod_{l=1}^{k} \frac{2l-1}{2l} \int_0^{\rho} \tilde{\alpha}(r) r^{2k+1} \, \mathrm{d}r. \tag{16}$$

The second equal sign in (16) follows from the Wallis product and the periodicities of the cosine function. Furthermore we can develop a relation between γ_{2k} and $\gamma_{2(k-1)}$, respectively $\gamma_0 = \alpha_{\mathrm{PDE}}$ $(k > 0)$.

$$\gamma_{2k} \leqslant \frac{1}{4k^2} \rho^2 \gamma_{2(k-1)} \quad \text{and} \quad \gamma_{2k} \leqslant \frac{1}{4^k (k!)^2} \rho^{2k} \gamma_0. \tag{17}$$

Our proposition (15) is true because

$$\Delta^k I(x,y) = \sum_{j=0}^{k} \binom{k}{j} \frac{\partial^{2(k-j)}}{\partial x^{2(k-j)}} \frac{\partial^{2j}}{\partial y^{2j}} I(x,y) \tag{18}$$

and (γ_{2k} is involved!)

$$\frac{1}{(2(k-j))!(2j)!} \int_0^{2\pi} \cos^{2k-2j} \varphi \cdot \sin^{2j} \varphi \, \mathrm{d}\varphi = \frac{\binom{k}{j}}{(2k)!} \int_0^{2\pi} \cos^{2k} \varphi \, \mathrm{d}\varphi, \tag{19}$$

which can be applied after rearranging the result of (14) to a double series.

Because the Taylor series expansion converges locally towards the function value (actually depending on the smoothness of the function), the final result is

$$\lim_{n\to\infty} \Psi_n(t,\vec{x}) = \lim_{n\to\infty} \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \gamma_{2k} \Delta^k I(x,y) = \int \tilde{\alpha}(\|\vec{z}\|) I(\vec{x}+\vec{z}) \; \mathrm{d}\vec{z}. \tag{20}$$

4.3. Implementation

For implementation of a solution method for the second order simplified PDE model on a square grid the calculation of the Laplacian of the density of infected individuals at every discrete location is required.

$$\Delta I(x,y) \approx \frac{1}{\varepsilon^2} (I(x+\varepsilon) + I(x-\varepsilon) + I(y+\varepsilon) + I(y-\varepsilon) - 4I). \tag{21}$$

If we assume a lattice corresponding to the SCA approach, the minimal value for ε is 1. Accordingly this solution procedure (finite differences) for the PDE uses a von Neumann neighbourhood of size one to calculate the growth of the density of infected individuals in the cell.

$\cdots$			
γ_4	$\cdots$		
γ_2 $-4\gamma_4$	$2\gamma_4$	$\cdots$	
γ_0 $-4\gamma_2$ $+4\gamma_4$	γ_2 $-4\gamma_4$	γ_4	$\cdots$

Fig. 1. Weight factors in the neighbourhood for the fourth-order simplified PDE implementation (upper-right quadrant).

For the fourth order approximated PDE model we additionally need to know $\Delta\Delta I(x,y)$, which can be calculated from $\{\Delta I(u,v),\ |x-u|+|y-v|\leqslant 2\}$. This leads to a von Neumann neighbourhood of size two for Ψ_4. In this case the weight factors for the cells in the neighbourhood are determined by the parameters γ_{2k} (Fig. 1). This also explains why these parameters can not be totally arbitrary but must satisfy the condition (17). Otherwise problems arise at the latest in the implementation, especially if all densities should be $\leqslant 1$.

Because of (20) we know that the 'approximations' Ψ_n converge for $n\to\infty$ towards the integral representation (6) of the model. Accordingly we can approximate the impact of infected contacts Ψ directly by

$$\Psi(x,y)\approx\sum_{(u,v)}\tilde{\alpha}(\|(u,v)-(x,y)\|)\cdot I(u,v). \tag{22}$$

Although the resulting approach does not directly involve derivatives or differential equations we refer to it as PDE approach because it is the basis for PDE systems that emerge from Taylor series simplifications.

5. Relation SCA–PDE model

The just mentioned solution methods for the PDE approach provide the densities of the three groups on the lattice (domain). With the help of a uniformly distributed random number for each cell, that is constant over time, a discrete state can be obtained by locating the random number in the segmentation of $[0,1]=[0,S]\cup(S,S+I]\bigcup(S+I,S+I+R]\cup(S+I+R,1]$. In this case Ψ describes the growth of the probability of the cell being infected.

The behaviour of this lattice representation during evaluation can be compared to a SCA simulation. In the following we provide the theoretical link between these two approaches.

With the results from Section 4.2 and a factor α_{PDE} for the likelihood function $\tilde{\alpha}$, which corresponds to a probability density function, the probability of a distinct cell in the PDE model being infected Φ_{PDE} (in the next time step) can be written as

$$\Phi_{\mathrm{PDE}}=I_c+S_c\cdot\Psi_{\mathrm{PDE}}=I_c+S_c\cdot\alpha_{\mathrm{PDE}}\cdot\sum\tilde{\alpha}(u,v)I(u,v), \tag{23}$$

minus the recoveries, which are neglected here for reasons of simplification but reintroduced later on without further mentioning. I_c and S_c are the densities of infected, respectively susceptible individuals in the cell. Ψ_{PDE} is the growth of the density of infected individuals or the growth of the probability of the cell being infected.

In the SCA approach the probability of a cell being infected (in the next time step) Φ_{SCA} is

$$\Phi_{\mathrm{SCA}}=I_c+S_c\cdot\Psi_{\mathrm{SCA}}=I_c+S_c\cdot(1-(1-\alpha_{\mathrm{SCA}})^{I_{\Omega_\lambda}}). \tag{24}$$

The number of contacts is λ and the number of infected contact cells is $I_{\Omega_\lambda}.\ S_c$, respectively I_c is one if the cell is susceptible, respectively infected otherwise it is zero.

We now assume equal conditions in both models, which allows to confine on comparing the growth of the probability of a cell being infected. Therefore S_c and I_c are supposed to take values between 0 and 1 also in the SCA approach. The number of contacts λ in the SCA is 1 and the density of infected individuals in this contact cell, which is chosen according to a probability distribution $\tilde{\alpha}$, is I_1.

$$\begin{aligned}&\Psi_{\mathrm{SCA}}\approx\Psi_{\mathrm{PDE}}\\&1-(1-\alpha_{\mathrm{SCA}})^{I_1}\approx\alpha_{\mathrm{PDE}}\cdot\sum\tilde{\alpha}(u,v)I(u,v)\\&\alpha_{\mathrm{SCA}}\cdot I_1\approx\alpha_{\mathrm{PDE}}\cdot\int\tilde{\alpha}(u,v)I(u,v).\end{aligned} \tag{25}$$

Because the location of the SCA contact cell is distributed with the probability distribution $\tilde{\alpha}$, the law of large numbers delivers an approximation of the density of infected individuals in this cell.

$$I_1\to\int\tilde{\alpha}(u,v)I(u,v). \tag{26}$$

Accordingly it is necessary that $\alpha_{\mathrm{SCA}}\approx\alpha_{\mathrm{PDE}}$. And with the same idea we obtain for $\lambda>1$ that $\alpha_{\mathrm{SCA}}\approx\frac{\alpha_{\mathrm{PDE}}}{\lambda}$.

6. Relation LGCA–PDE and SCA model

The aim of this section is to adapt the behaviour of the PDE and SCA approach to the LGCA. As a basic connection we use that random walk of particles can at least in a scaling limit be regarded as a diffusion process (i.e. Brownian motion). One way to describe diffusion processes is by Gaussian semigroups [2] of the form

$$T(t)f(\vec{x})=\frac{1}{4\pi t}\int\exp\left(\frac{-\|\vec{z}\|^2}{4t}\right)f(\vec{x}+\vec{z})\,\mathrm{d}\vec{z} \tag{27}$$

where the argument $f(\vec{x})$ (in our specific case $I(\vec{x})$) of the evolution operator $T(t)$ describes the density of a substance or of a large number of particles. If only one particle is observed the kernel (probability density function of a normal distribution) can describe the probability of the particle having a specific distance r from its point of origin after t time units.

6.1. LGCA specific diffusion

In order to be sure that the mathematical diffusion process corresponds to diffusion of particles in random motion LGCA we made an analysis of random walk particles on a hexagonal grid. We measured the deviations $r_k(t)$ and square deviations $s_k(t)$ of a test set of $n = 10,000$ particles from their initial location after $t = 0, 1, \ldots$ time units.

The mean square deviations correspond to $\bar{s}(t) \approx t$. In accordance to the semigroup approach for diffusion we assume that the positions of the particles are distributed with a two dimensional symmetric normal distribution $N_2(0, \sigma^2)$. This leads to the radial density function

$$\tilde{\alpha}(r) = f(r) = \frac{r}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}}. \tag{28}$$

The expectation values $\mathbb{E}(r)$ and $\mathbb{E}(r^2)$ are $\sqrt{\frac{\pi\sigma^2}{2}}$, respectively $2\sigma^2$. By setting $\mathbb{E}(r^2) = t \approx \bar{s}(t)$ the following results can be obtained:

$$\sigma^2 = \frac{\mathbb{E}(r^2)}{2} = \frac{t}{2} \quad \text{and} \quad \mathbb{E}(r) = \sqrt{\frac{\pi\sigma^2}{2}} = \sqrt{\frac{\pi t}{4}}. \tag{29}$$

The second equation in (29) is a result of the first one and complies to the results from the test sample $\bar{r}(t) \approx \mathbb{E}(r)$ (Fig. 2a). Additionally for given t the empirical distribution function $\widehat{F}(r)$ of the test sample corresponds to the cumulative distribution function $F(r)$ of $f(r)$.

Further we also measured the deviations and square deviations of moving particles in SCA. Therefore a particle, that moves or jumps to its contact cell according to the probability distribution $\tilde{\alpha}$ with parameter $t = 1$ at each time step, was observed. The mean deviations corresponded to those from the LGCA, but the mean square deviations showed a different behaviour (Fig. 2b). The SCA delivers much better correspondence to the LGCA when no jumps are performed, but the particle location after $\tilde{t}$ time units is simulated (as one single jump starting from the initial location) according to the probability distribution with parameter $t = \tilde{t}$ (Fig. 2c).

Therefore in further SCA simulations the contacts of an infected cell are established according to this distribution with parameter $t = h$ where h is the number of time units since infection of this individual cell.

6.2. Parameter adaption

Until now only the diffusion process was analysed. For corresponding model behaviour further parameter adaptions concerning the probability of infection are necessary.

We want to relate the PDE and LGCA behaviour and therefore identify positions on the PDE domain (cells in the SCA) with cells in the LGCA. Accordingly not the number $I_{c,\text{abs}}$ but the density I_c of for example infected individuals within LGCA cells is of interest.

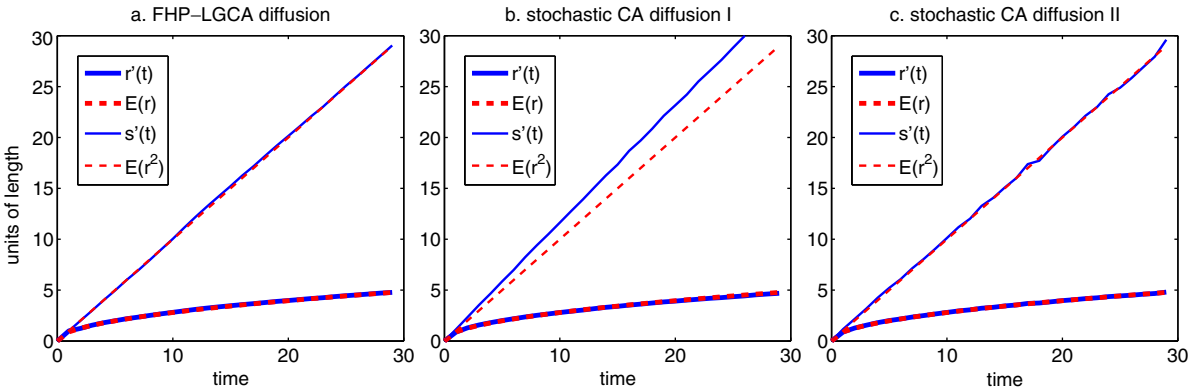

Fig. 2. These three plots show the mean deviation $r'(t)$ and mean square deviation $s'(t)$ of a sample of 10,000 diffusing particles over time t in a hexagonal LGCA (a), a SCA (b) and a modified SCA (c). The grid consists of 100×100 cells and the time of observation is 30 time units. The modified SCA does not directly simulate moving or jumping particles but randomly selects a final location by using an increased variance ($\sigma^2 = \frac{t}{2}$) in the diffusion distribution. Additionally the expectation values of the diffusion distribution $E(r)$ and $E(r^2)$ are shown.

The probability of a cell being infected in the FHP LGCA is

$$\Phi_{\mathrm{FHP}} = \frac{1}{6} \cdot (I_{c,\mathrm{abs}} + (1 - (1 - \alpha_{\mathrm{FHP}})^{I_{c,\mathrm{abs}}}) \cdot S_{c,\mathrm{abs}}) = I_c + S_c \cdot (1 - (1 - \alpha_{\mathrm{FHP}})^{6I_c}). \tag{30}$$

The amount of infected particles in the LGCA cell depends on the motion behaviour, which in the case of one discrete time step (compare Section 2) can be described by $\tilde{\alpha}$ (28) with variance $\sigma^2 = 1$ $(t = 2)$ for example. Therefore it is possible to approximate the density of infected individuals in the current cell I_c by $\int \tilde{\alpha}(\vec{x}) I(\vec{x})$. This allows to relate the growth of the probability of infection in both model approaches and to extract a parameter relation for the probabilities of infection if the densities of infected and susceptible individuals I_c and S_c are the same in the observed LGCA and PDE cell.

$$\begin{aligned} 1 - (1 - \alpha_{\mathrm{FHP}})^{6I_c} &= \Psi_{\mathrm{FHP}} \approx \Psi_{\mathrm{PDE}} = \alpha_{\mathrm{PDE}} \cdot \sum \tilde{\alpha}(u,v) I(u,v) \\ 6 \cdot \alpha_{\mathrm{FHP}} \cdot I_c &\approx \alpha_{\mathrm{PDE}} \cdot \int \tilde{\alpha} I \\ \alpha_{\mathrm{FHP}} &\approx \frac{\alpha_{\mathrm{PDE}}}{6} \end{aligned} \tag{31}$$

7. Testing and results

For testing of the parameter relations found in the previous sections simulations of susceptible-infected-recovered-susceptible (SIRS) model extensions implemented in MATLAB with the LGCA, SCA and PDE approach are observed (Figs. 4–7). The key output value is the number of currently infected individuals but also a visualisation of the density of infected individuals on the domain permits conclusions on the model behaviours (compare Fig. 3). The input parameters involve the size of the domain (grid size), the number of performed time steps, the disease stage transition parameters (α, β, etc.), the number

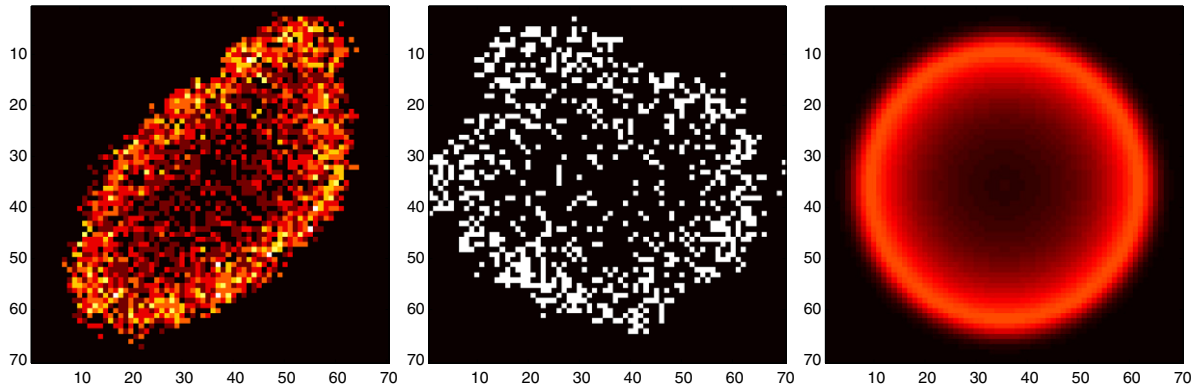

Fig. 3. A lattice representation of the FHP LGCA approach (left), the SCA approach (center) and the PDE approach (right). Light pixels mark high densities of infected particles and dark pixels mark low densities. This figure does not show specific results but is intended to demonstrate general lattice representations.

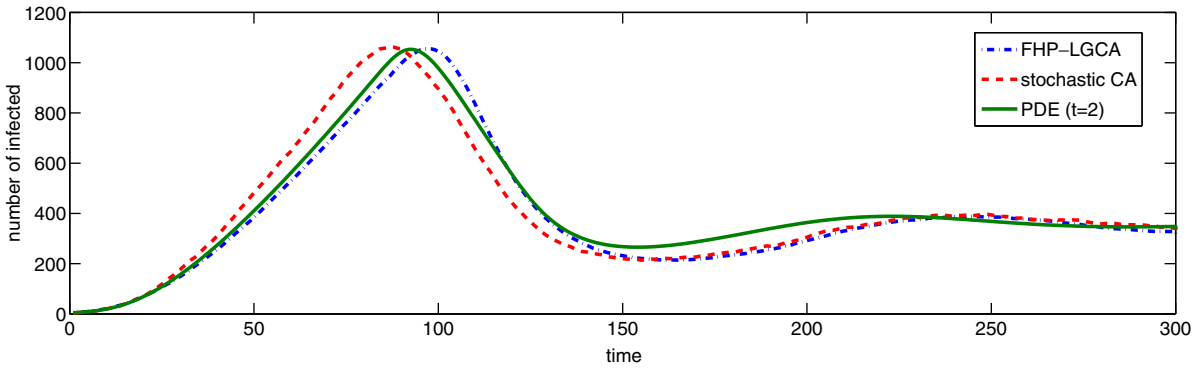

Fig. 4. Number of infected individuals over 300 time units. The infection rate is 0.4 and the initial population density is 1. The density function parameter for the PDE approach is 2. Additional parameters are given in Section 7.

of runs used for generating an averaged result (Monte Carlo method) and the parameter t of the diffusion distribution that was used in the PDE approach.

In the test simulations (Figs. 4–7) the results of the LGCA and SCA approach were averaged over 20 simulation runs. The grid size was 70×70 cells, except for the simulation in Fig. 6, where it was 30×30 cells. The rates for recovery and loss of immunity were 0.08, respectively 0.008 except for the simulation in Fig. 7, where both were 0.

An optimal model identification would persist, if the differences are minimal for arbitrary parameter settings. In our case there exist two (heuristic) parameter regions, which deliver good, respectively bad correspondence in the model behaviours.

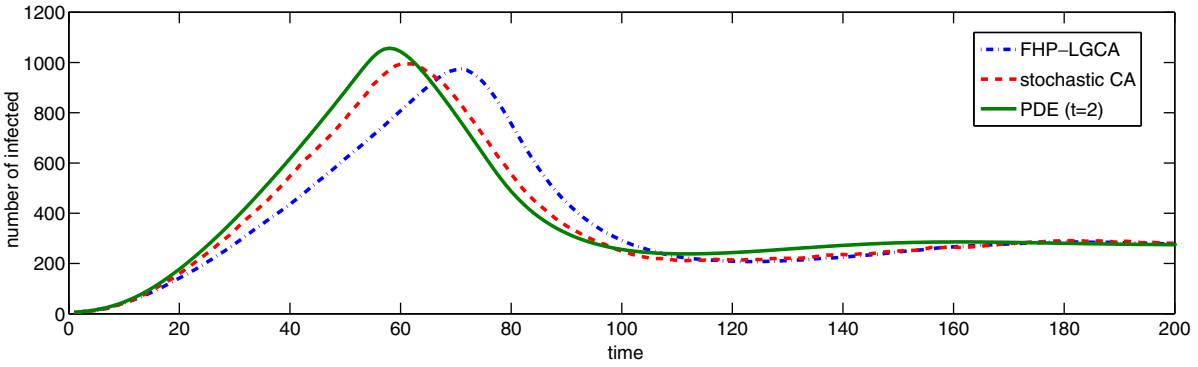

Fig. 5. Number of infected individuals over 200 time units. The infection rate is 1 and the initial population density is 0.7. The density function parameter for the PDE approach is 2. Additional parameters are given in Section 7.

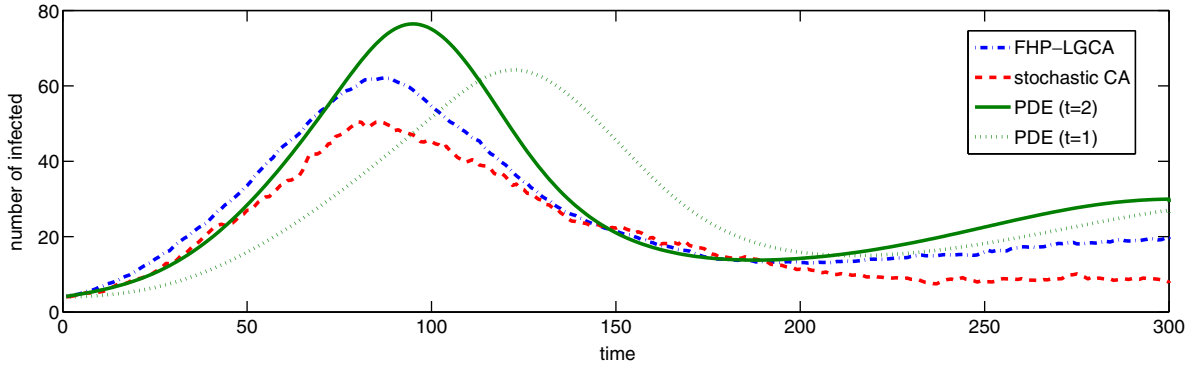

Fig. 6. Number of infected individuals over 300 time units. The infection rate is 0.4 and the initial population density is 0.5. The density function parameter for the PDE approach is 1 and 2. Additional parameters are given in Section 7.

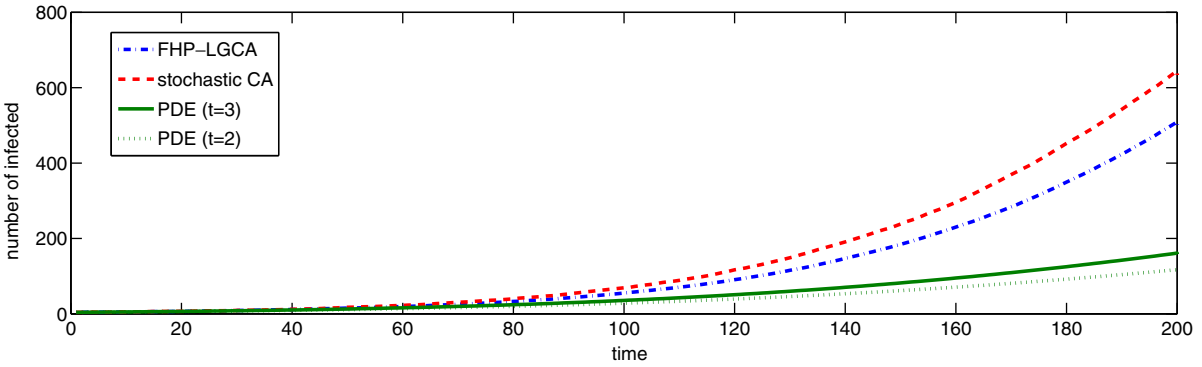

Fig. 7. Number of infected individuals over 200 time units. The infection rate is 0.05 and the initial population density is 0.7. The density function parameter for the PDE approach is 2 and 3. Additional parameters are given in Section 7.

Notably a high population density and high infection rates as in Figs. 4 and 5 produce greater correspondence in the evolution of the number of infected individuals than lower values as in Figs. 6 and 7. In the former cases also the PDE approach with the heuristic parameter choice $t = 2$ for the diffusion distribution $\tilde{\alpha}$ delivers similar behaviour. A hypothetical reason for the differences in the latter cases is that lower parameter values decrease speed of spread and therefore approximation errors provoked by imprecise assumptions on the diffusion and interaction behaviour are amplified because a larger number of iterative steps must be performed. Meanwhile the qualitative behaviour remains roughly the same.

In Figs. 4 and 6 a small reverberation of the peek of the number of infected individuals is visible. This phenomenon corresponds to reoccurring epidemic waves, which can be observed in real life and are of great interest in epidemic forecasts [4].

8. Conclusions and discussion

The investigated spatial extensions of the Kermack–McKendrick SIR model show similar qualitative and quantitative behaviour in space and time. Furthermore the cumulative results (sum of currently infected individuals) of all approaches coincide with the typical qualitative behaviour of the SIR ODE model.

For a certain parameter range (low density and infection rate) greater differences in the quantitative behaviours occur. Accordingly additional investigations especially concerning the cell interaction behaviour in the PDE and LGCA approach could be done. Further convergence analysis of stochastic processes or regarding physical features like for example Graham's law of diffusion could deliver additional information regarding the configuration of a diffusion distribution or the choice of the infection rates.

Regarding the construction of specific spatial epidemiological models including hybrid approaches, the results of these investigations could help to find appropriate parameters and techniques or to extract the advantages of several techniques at the same time.

References

[1] C.L. Barrett, S.G. Eubank, J.P. Smith, Pocken in Portland – simuliert, Spektrum der Wissenschaft, Dossier Seuchen II 3 (2006) 18.
[2] M. Blümlinger, M. Kaltenbäck, H. Langer, M. Langer, Funktionalanalysis 2, Lecture Notes, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2003/2004.
[3] J. Callahan, The spread of an infection through a region, Lecture Notes, Department of Mathematics and Statistics, Smith College Northampton, 1996.
[4] Š. Emrich, Comparison of Mathematical Models and Development of a Hybrid Approach for the Simulation and Forecast of Influenza Epidemics within Heterogeneous Populations, Diploma Thesis, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2007.
[5] H. Fukś, A. Lawniczak, Individual-based lattice model for spatial spread of epidemics, Discrete Dynamics in Nature and Society, 6 (3) (Hindawi, 2001), pp. 191–200.
[6] H. Hötzendorfer, F. Breitenecker, A Directly Programmed Implementation of ARGESIM Comparison C17 SIR-type Epidemic using Matlab, Simulation News Europe, vol. 41/42 (ARGESIM, 2004), p. 45.
[7] H. Hötzendorfer, N. Popper, F. Breitenecker, Temporal and Spatial Evolution of a SIR-type Epidemic – ARGESIM Comparison C17 – Definition, Simulation News Europe, vol. 41/42 (ARGESIM, 2004), pp. 42–45.
[8] D.S. Jones, B.D. Sleeman, Differential Equations and Mathematical Biology, George Allen & Unwin, London, 1983.
[9] J. Joo, J.L. Lebowitz, Pair Approximation of the stochastic susceptible-infected-recovered-susceptible epidemic model on the hypercubic lattice, DIMACS Technical Report 2004/13, Department of Mathematics and Physics, Rutgers University, 2004.
[10] S.P. Lalley, Critical scaling of stochastic epidemic models, IMS Lecture Notes-Monograph Series, vol. 55 (Institute of Mathematical Statistics, 2007), pp. 167–178.
[11] G. Schneckenreither, F. Breitenecker, CA models for SIR type epidemics, Simulation News Europe, vol. 16/3 (ARGESIM, 2006), pp. 27–36.
[12] G. Schneckenreither, N. Popper, G. Zauner, Modelling SIR-type epidemics by ODEs, PDEs, difference equations and cellular automata – a comparative study, in: B. Zupančič, R. Karba, S. Blažič, (Eds.), Proceedings of the 6th EUROSIM Congress on Modelling and Simulation, vol. 2 (Ljubljana, 2007).
[13] S. Venkatachalam, A.R. Mikler, Towards computational epidemiology: using stochastic cellular automata in modeling spread of diseases, in: Proceedings of the 4th Annual International Conference on Statistics, Mathematics and Related Fields (Honolulu, 2005).
[14] D.A. Wolf-Gladrow, Lattice-Gas Cellular Automata and Lattice Boltzmann Models, Springer, Berlin, 2000.
[15] S. Wolfram, A New Kind of Science, (B&T, 2002).