

# Simulation of the spread of infectious diseases in a geographical environment

ZHONG ShaoBo<sup>1†</sup>, HUANG QuanYi<sup>1</sup> & SONG DunJiang<sup>2</sup>

<sup>1</sup>Department of Engineering Physics, Tsinghua University, Beijing 100084, China;

<sup>2</sup>Institute of Policy and Management, Chinese Academy of Sciences, Beijing 100190, China

**The study of mathematical models for the spread of infectious diseases is an important issue in epidemiology. Given the fact that most existing models cannot comprehensively depict heterogeneities (e.g., the population heterogeneity and the distribution heterogeneity) and complex contagion patterns (which are mostly caused by the human interaction induced by modern transportation) in the real world, a theoretical model of the spread of infectious diseases is proposed. It employs geo-entity based cellular automata to simulate the spread of infectious diseases in a geographical environment. In the model, physical geographical regions are defined as cells. The population within each cell is divided into three classes: Susceptible, Infective, and Recovered, which are further divided into some subclasses by states of individuals. The transition rules, which determine the changes of proportions of those subclasses and reciprocal transformation formulas among them, are provided. Through defining suitable spatial weighting functions, the model is applied to simulate the spread of the infectious diseases with not only local contagion but also global contagion. With some cases of simulation, it has been shown that the results are reasonably consistent with the spread of infectious diseases in the real world. The model is supposed to model dynamics of infectious diseases on complex networks, which is nearly impossible to be achieved with differential equations because of the complexity of the problem. The cases of simulation also demonstrate that efforts of all kinds of interventions can be visualized and explored, and then the model is able to provide decision-making support for prevention and control of infectious diseases.**

cellular automata, infectious disease, modeling, geographic information systems, spatial weighting function

Infectious diseases are caused by pathogenic microorganisms, such as bacteria, viruses, parasites or fungi, and can spread, directly or indirectly, from one person (or animal) to another. A serious infectious disease might cause mass deaths, social panics, or even a national crisis. Some examples of infectious diseases are the Black Death during the middle fourteenth century, the Spanish Flu pandemic in 1918, the Severe Acute Respiratory Syndrome (SARS) in 2002, and more recently, the Highly Pathogenic Avian Influenza (HPAI)<sup>[1]</sup>. In order to prevent and control infectious diseases, the study of mathematical models for the spread of infectious diseases is an important issue in epidemiology. Through scientifically modeling the spread of an infectious dis-

ease, some key factors that disclose the mechanism of the spread of the disease can be found out and the effect of some interventions (e.g., therapy, isolation, immunization) on the spread of the disease can be evaluated. Modeling infectious diseases using differential equations has been a popular approach. In differential equation models, the population is generally divided into three classes: Susceptible, Infective, and Recovered. The Susceptible refers to those healthy individuals who are

Received September 18, 2008; accepted January 21, 2009; published online February 26, 2009

doi: 10.1007/s11430-009-0044-9

<sup>†</sup>Corresponding author (email: zhongshaobo@tsinghua.edu.cn)

Supported by Postdoctoral Foundation of China (Grant No. 20070410552) and Youth Fund of Institute of Policy and Management (IPM), the Chinese Academy of Sciences (Grant No. O700481Q01)

susceptible to a certain infectious disease and are likely to be infected with the disease. The Infective refers to the individuals under infection. The Recovered refers to those who are recovered from infection with the disease and are contemporarily or permanently immune to the disease. To characterize the varying infection force of an infected individual, the infection period of a disease is further divided into three phases: incubation, when the individual is already infected and infectious but presents no symptoms; proper infection, when the individual is infectious and shows the symptoms; and latency, when the infected individual is not infectious anymore but still has the symptoms<sup>[2]</sup>. Two basic differential equation models were extensively studied. One is SIR model for modeling the infectious diseases that confer immunity to the infected individuals after recovery (e.g., measles, chickenpox, mumps, HIV, and poliomyelitis). The other is SIS model for modeling the infectious diseases that do not confer immunity to the infected individuals after recovery (e.g. meningitis, plague, venereal diseases, and malaria). Some other differential equation models are more or less adapted from SIR or SIS, such as SIRS, SEIRS and so on<sup>[3]</sup>. These models all assume that the population under study is homogeneous and the interaction among individuals is equiprobable. These assumptions are unrealistic, which are also the intrinsic drawbacks of these differential equation models, because the local characters are neglected in these models<sup>[4]</sup>. They do not include variable susceptibility of individuals, and cannot handle the complicated boundary and initial conditions<sup>[5]</sup>. Furthermore, these models are all adimensional, i.e., they can depict the changes of the number of all classes of individuals in time, but not in space.

Some types of methods are developed to deal with the heterogeneous characteristics. Spatial infectious disease models and Cellular Automata (CA) models are two outstanding types of them. In essence, spatial infectious disease models<sup>[6]</sup> are composed of a system of differential equations and understandably have the common shortcomings as differential equation models. However, CA models are a new kind of computational techniques for simulating the real world, introduced first by Von Neumann<sup>[7]</sup>. At first, they were mainly used for simulation of physics systems. In recent years, many researchers from all kinds of fields have had an increasing interest in them thanks to the development of the computer techniques. Some authors have applied CA to model infectious diseases<sup>[1,2,5,8,9]</sup>. They basically construct the CA based on differential equation models. By that way,

some epidemiological features of an infectious disease can be expressed more directly. Moreover, CA simulation can present visual results that enable us to ascertain some key features of the spread of an infectious disease through experiment and observation. In most existing studies, CA models use cellular space that is generated by equably dividing the research area into regular lattices. And in those studies, the classic neighbor patterns, such as Moore and von Neumann neighbors, are applied and the transition rules of the CA are derived from the relevant differential equation models. These, however, have some shortcomings. First of all, the classic neighbor patterns are not suitable for some special diseases while they are suitable for characterizing the spread of some infectious diseases. SARS is one of the typical examples. The spread of SARS shows the small world characteristic<sup>[10]</sup> which is completely different from the behavior expressed by the classic neighbor patterns. On the other hand, the real background space where diseases spread comprises irregular geographical regions (which generally have different sizes and shapes), so it is also not suitable to impersonate physical geographical regions using regular lattices.

As mentioned above, the classic CA model is too simple for modeling the dynamics in a geographical environment. It has some limitations with respect to the cellular shape, the neighbor pattern, and the transition rule, which constrains its ability to simulate the real world. To overcome these drawbacks, some researchers have studied the theories and applications of geographical CA<sup>[11,12]</sup>. In this paper, we employ a geo-entity based CA model<sup>[13,14]</sup>, which defines physical geographical regions as cells. In addition to the consideration of the epidemiological characteristics of infectious diseases (e.g., effective contact rate, recovery rate, and the varying infection force), we pay more attention to the influence of contagion patterns on the spread of infectious diseases. Two typical kinds of contagion patterns are explored. One is the local contagion good for portraying local diffusion and the other is the global contagion, which is potentially likely to express the contact behavior of infectious diseases on complex networks<sup>[15-20]</sup>. The cases of simulation for the two situations are also presented and analyzed. It has been shown that the results are reasonably consistent with the spread of infectious diseases in the real world. The model is supposed to model dynamics of infectious diseases in complex networks, which is nearly impossible to be achieved with differential equations because of the complexity of

the problem. The additional cases of simulation also demonstrate that efforts of all kinds of interventions can be visualized and examined, and then the model is able to provide decision-making support for prevention and control of infectious diseases.

## 1 SIRS model

The main features of an infectious disease and the environment where it is spreading are as follows.

(i)  $S(t)$ ,  $I(t)$  and  $R(t)$  denote the proportions of Susceptible, Infective and Recovered. The population within the study area is constant. There are no death and birth during the epidemic, i.e.,  $S(t)+I(t)+R(t)=1$ .

(ii) The effective contact rate is constant  $\lambda$ , i.e., the average number of Susceptible infected by each infected individual in a given time (e.g., one day) is constantly proportioned to the number of the current Susceptible.

(iii) The cured rate is constant  $\nu$ , i.e., the number of Infective cured in a given time is constantly proportioned to the number of the current Infective. The cured Infective are contemporarily immune to the disease.

(iv) The recovery rate is constant  $\gamma$ , i.e., the number of Recovered losing immunity in a given time is constantly proportioned to the number of the current Recovered. After losing immunity, Recovered are susceptible to the disease again.

According to the above assumptions, The SIRS model can be formulated using the following differential equations:

$$\begin{cases} \frac{dI}{dt} = \lambda SI - \nu I, \\ \frac{dS}{dt} = -\lambda SI + \gamma R, \\ \frac{dR}{dt} = \nu I - \gamma R, \\ S + I + R = 1. \end{cases} \quad (1)$$

Eq. (1) can be reduced to the following ones:

$$\begin{cases} \frac{dS}{dt} = -\lambda SI + \gamma(1 - S - I), \\ \frac{dI}{dt} = \lambda SI - \nu I. \end{cases} \quad (2)$$

It is difficult to calculate the analytical solution of eq. (1). A phase plane method is generally employed to analyze some characteristics of the equations<sup>[21]</sup>.

## 2 Classic CA model

Bidimensional CA are discrete dynamical systems

formed by a finite number of  $row \times col$  identical objects called cells, which are arranged uniformly in a two-dimensional cellular space. Each cell is endowed with a state (from a finite state set  $Q$ ) that changes at every step of time according to a local transition rule. The state of a cell at the step of time  $t$  depends on the states of the cell and its neighbors. Generally, CA can be defined by a 4-uplet  $(C, Q, V, f)$ , where  $C = \{(i, j), 1 \leq i \leq row, 1 \leq j \leq col\}$  is the cellular space;  $Q$  is the finite state set whose elements are the all possible states of the cells;  $V = \{(\alpha_k, \beta_k), 1 \leq k \leq n\} \subset Z \times Z$ , is the finite set of indices defining the neighborhood of each cell, such that the neighborhood of the cell  $(i, j)$  is  $V_{ij} = \{(i+\alpha_1, j+\beta_1), (i+\alpha_2, j+\beta_2), \dots, (i+\alpha_n, j+\beta_n)\}$ ; finally, the function  $f$  is the local transition function:

$$s_{ij}^t = f(s_{i+\alpha_1, j+\beta_1}^{t-1}, \dots, s_{i+\alpha_n, j+\beta_n}^{t-1}) \in Q, \quad (3)$$

where  $s_{ij}^t$  stands for the state of the cell  $(i, j)$  at the step of time  $t$ .

## 3 Geo-entity based CA model

We consider a geographical environment composed of  $k$  geographical regions. These regions can be any type of geo-entity, such as administrative region, community, building. Label these geo-entities with 1, 2, 3, ...,  $k$  in arbitrary order.

First, we define the following notations:  $N_i$ , the total amount of the population in region  $i$ .

$A_i$ , the area of region  $i$ ;  $S_i^t, I_i^t, R_i^t$ , the proportions of Susceptible, Infective and Recovered in region  $i$  at the step of time  $t$  respectively, and certainly  $S_i^t + I_i^t + R_i^t = 1$ .

Now, we consider the geographical area under study as the cellular space, and look on a geo-entity (e.g., jurisdiction, community, residence) in the area as a cell. Use a set  $M = \{m_i, i=1, 2, \dots, k\}$  to denote the cellular space. Each cell  $m_i$  comprises three classes of individuals, i.e., Susceptible, Infective and Recovered. Every cell is further divided into subclasses by the states of individuals. The state of every subclass will dynamically change with time. Here we assume, for a certain disease, its phases' average times are fixed.  $T_i, T_p, T_l$ , and  $T_r$  denote the average times of the incubation, proper infection, latency, and immunity respectively (unit: day). Thus, we define the state set which each subclass may go through as  $P = \{0, 1, 2, \dots, T_i + T_p + T_l + T_r\}$ . The subclass with state 0 consists of the individuals susceptible to the disease. The subclass with state 1 consists of the

individuals on the first day of incubation, state  $T_i+1$  on the first day of proper infection, ..., and state  $T_i+T_p+T_l+1$  on the first day of immune. When the state of a subclass is  $T_i+T_p+T_l+T_r$ , which means that the individuals with the state are on the last day of the immune phase, the state of the subclass should be changed to 0 at the next step of time. Every subclass will go through each state in the state set orderly except that they are cured during the infection period. If the subclass is cured during the infection period, it will directly enter into the first day of the immune phase, i.e. its state should be changed to  $T_i+T_p+T_l+1$ .

Figure 1 illustrates the process each subclass may go through.

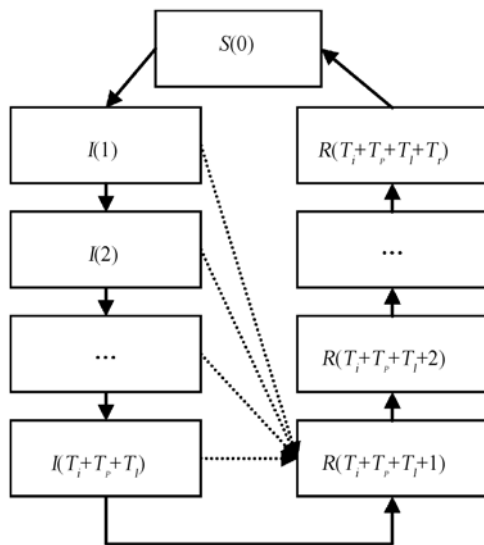


Figure 1 The process every subclass may go through.

In order to describe the transition rules of our CA, we further define the following notations:

$S_i^t$ , the proportion of Susceptible with state 0 in region  $i$  at the step of time  $t$ ;  $I_i^t(q)$ , the proportion of Infective with state  $q$  in region  $i$  at the step of time  $t$ ,  $q \in \{1, 2, \dots, T_i+T_p+T_l\}$ ;  $R_i^t(q)$ , the proportion of Recovered with state  $q$  in region  $i$  at the step of time  $t$ ,  $q \in \{T_i+T_p+T_l+1, T_i+T_p+T_l+2, \dots, T_i+T_p+T_l+T_r\}$ .

Now, let us examine the state set of the CA (Please note that the state set of a CA is different from the state set of a subclass described above). We define the state set of a cell according to proportions of all subclasses in it. Set  $DS_i^t$ ,  $DI_i^t$  and  $DR_i^t$  suitable discretizations of the fractions of Susceptible, Infective and Recovered of cell  $i$  at the step of time  $t$ , respectively. To get elements of the finite state set  $Q$ , we will consider the state set:

$$Q = \overbrace{K \times K \times \dots \times K}^{T_i+T_p+T_l+T_r+1 \text{ times}}, \quad (4)$$

where  $K = \{0.00, 0.01, 0.02, 0.03, \dots, 0.99, 1.00\}$ , which is formed by 101 elements. Consequently, the discretization used is

$$\begin{cases} DS_i^t = \frac{[100S_i^t]}{100}, \\ DI_i^t(q) = \frac{[100I_i^t(q)]}{100}, \\ q \in [1, 2, \dots, T_i + T_p + T_l], \\ DR_i^t(q) = \frac{[100R_i^t(q)]}{100}, \\ q \in [T_i + T_p + T_l + 1, T_i + T_p + T_l + 2, \dots, T_i + T_p + T_l + T_r], \end{cases} \quad (5)$$

where  $[x]$  is the nearest integer to  $x$ .

Then, the state of the CA used in the model is  $s_i^t = (DS_i^t, DI_i^t(1), DI_i^t(2), \dots, DI_i^t(T_i+T_p+T_l), DR_i^t(T_i+T_p+T_l+1), DR_i^t(T_i+T_p+T_l+2), \dots, DR_i^t(T_i+T_p+T_l+T_r)) \in Q$ .

The subclasses of all cells of the cellular space interact with each another. The interaction of every cell involves two portions: Internal and External. Internal refers to reciprocity of the subclasses inside the same cell, whereas External refers to reciprocity of subclasses from different cells.

Based on the above definitions and concepts, we formulate the following transition rules:

$$\begin{cases} R_i^{t+1}(q) = R_i^t(q-1), & q \in \{T_i + T_p + T_l + 2, T_i + T_p + T_l + 3, \dots, T_i + T_p + T_l + T_r\}, \quad (6a) \\ S_i^{t+1} = S_i^t - \sum_{q \in \{1, 2, \dots, T_i + T_p + T_l\}} \frac{\lambda(q)\rho_i^t S_i^t I_i^t(q)}{\rho} + R_i^t(T_i + T_p + T_l + T_r) - \sum_{\substack{j \in \{1, 2, \dots, k\} \\ q \in \{1, 2, \dots, T_i + T_p + T_l\}}} \frac{c_{ij}\lambda(q)\rho_i^t S_i^t I_j^t(q)}{\rho}, \quad (6b) \\ I_i^{t+1}(1) = \sum_{q \in \{1, 2, \dots, T_i + T_p + T_l\}} \frac{\lambda(q)\rho_i^t S_i^t I_i^t(q)}{\rho} + \sum_{\substack{j \in \{1, 2, \dots, k\} \\ q \in \{1, 2, \dots, T_i + T_p + T_l\}}} \frac{c_{ij}\lambda(q)\rho_i^t S_i^t I_j^t(q)}{\rho}, \quad (6c) \\ I_i^{t+1}(q) = (1 - \gamma(q-1))I_i^t(q-1), & q \in \{2, 3, \dots, T_i + T_p + T_l\}, \quad (6d) \\ R_i^{t+1}(T_i + T_p + T_l + 1) = I_i^t(T_i + T_p + T_l) + \sum_{q \in \{1, 2, \dots, T_i + T_p + T_l - 1\}} \gamma(q)I_i^t(q). \quad (6e) \end{cases}$$



Eq. (6a) depicts that the state of Recovered who do not reach the end of the immune phase increases one.

Eq. (6b) reflects that Susceptible will lose some individuals due to the internal interaction and the external interaction (the first summation and the second summation) while they gain some individuals since some Recovered reach the end of the immune phase. The first summation depicts Susceptible infected by Infective inside the same cell at the step of time  $t$ . The second summation depicts Susceptible infected by Infective from the neighbor cells at the step of time  $t$ .  $\lambda(q)$ ,  $q \in \{1, \dots, T_i+T_p+T_r\}$  denote the effective contact rates under a benchmark density of individual distribution  $\rho$  (e.g. 1000 person/km<sup>2</sup>). We call them Benchmark Effective Contact Rate (BECR). The BECR may be different for each state, which is a function of state  $q$ . In eq. (6b),  $\rho_i^t = S_i^t N_i / A_i$ . Here, we assume the effective contact rate between Susceptible and Infective is positively proportioned to the density of Susceptible distribution.  $c_{ij}$  is the correlation factor, which ranges from 0 to 1. The greater the distance between two cells is, the smaller the  $c_{ij}$  is.  $R_i^t(T_r)$  means Recovered who become susceptible to the disease again because their immune phases end at the step of time  $t$ .

Eq. (6c) expresses that the new Infective are equal to the loss of Susceptible.

Eq. (6d) means that some Infective will be cured, and the cured rate  $\gamma(q)$  may be different for each state, which is a function of state  $q$ .

Eq. (6e) expresses that some Infective at step of time  $t$  will enter into the immune phase. These individuals come from two parts: one is Infective reaching the end of the infection period, and the other is Infective cured.

#### 4 Computation for correlation factors

In the above model, the correlation factor  $c_{ij}$  reflects the force of the external interaction between region  $i$  and region  $j$ . Quantifying the variable reasonably is crucial for effective simulation of an infectious disease. In this paper, we will quantify  $c_{ij}$  based on the spatial weighting function.

The Tobler's first law of geography points out that everything is related to everything else, but near things are more related than distant things<sup>[22]</sup>. The law illustrates the universal correlation. In classic geography, the spatial weighting function is used for quantifying the spatial autocorrelation. A spatial weighting function is a

set of rules that assign values or "weights" to every pair of locations in a study area<sup>[23]</sup>. In essence, the aim of defining a spatial weighting function is to represent nearness and distance as consistently as possible for a set of irregular regions or irregularly spaced points.

The simplest weighting function for regional data is the binary weight. In that case, if two regions have shared boundaries, then the weight equals to 1, or else the weight equals to 0. This function assigns the same values to pairs of regions with very short boundaries and pairs with very long boundaries. An improved approach is to define the weight of pairs of regions considering their shared boundary length. Cliff and Ord<sup>[24]</sup> proposed a weighting function as follows

$$w_{ij} = z_{ij} / l_i, \quad (7)$$

where  $z_{ij}$  is the length of the boundaries of region  $i$  that are shared with region  $j$ , and  $l_i$  is the perimeter of region  $i$ . This approach gives greater weight to regions that share long boundaries. Whereas the weights calculated by this means are unsymmetrical, i.e.,  $w_{ij} \neq w_{ji}$ , so we modify eq. (7) as follows:

$$w_{ij} = w_{ji} = (z_{ij} / l_i + z_{ji} / l_j) / 2. \quad (8)$$

Any set of regions can also be represented by points located at their centers so that weighting functions defined for point data can also be applied to regional data. Tobler listed a series of weighting functions based on distance relation for point data<sup>[25]</sup>. There are also some other definitions which are not limited to distance relation, such as travel and individual contact, journey-to-work, journey-to-school, and journey-to-hospital<sup>[23]</sup>.

In order to model an infectious disease, first we select a suitable spatial weighting function given the epidemiological characteristics of the infectious disease. Then standardize the resulting weights. The standardization process is to transform the weights to meet the following conditions: the correlation factors approach zero for small weight but do not become negative and they approach one for large weight but are not greater than one, i.e., the weights and the correlation factors meet the following mapping relationship

$$w_{ij} \xrightarrow{f} c_{ij} \in (0, 1). \quad (9)$$

The following is an alternative of the transformation functions:

$$c_{ij} = e^{-\frac{1}{w_{ij}}}. \quad (10)$$

## 5 Simulation analysis

If an infectious disease is locally contagious, in which case it would spread mainly around nearby regions, the weighting functions based on distance relation may be appropriate alternatives. However, while an infectious disease is not limited to local contagion, the weighting functions based on global contagion patterns, such as travel and individual contact, journey-to-work, journey-to-school, and journey-to-hospital<sup>[23]</sup>, would be preferred.

In the cases of simulation, we demonstrate the state of each cell by means of pie charts. The sectors filled with colors white, grey, black represent the proportions of Susceptible, Infective and Recovered respectively.  $T_i, T_p, T_l, T_r$  are assigned with the following artificially chosen parameters: 2, 2, 2, 7 days, and  $\lambda(q)=0.2, \gamma(q)=0.1, q \in \{1, 2, \dots, T_i+T_p+T_l\}$  (except explicitly claimed). The steps of time of the cases of simulation are all 100 days.

Figure 2 illustrates the sketch map of the geographical environment of simulation. The environment comprises tens of geographical regions and several transportation lines. There are entries/exits distributed unevenly along these lines through which people can get up and off vehicles.

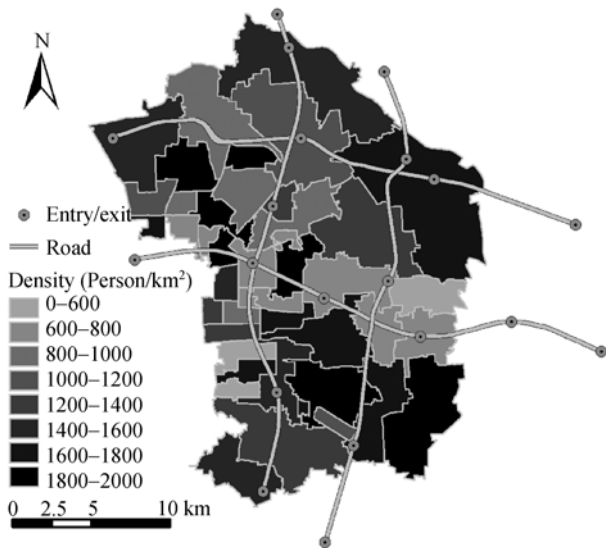


Figure 2 A sketch map of the geographical environment of simulations.

In the following cases of simulation, we will examine the spatial dynamics of infectious diseases with the local contagion and the global contagion, and the effect of two typical interventions (isolation and quarantine) on the

spread of infectious diseases.

The first case we present is the situation with local contagion. Here the spatial weighting function defined by eq. (8) is employed. Since the calculated weights are within (0, 1), we define  $c_{ij}=w_{ij}$ .

Four configurations considering only the local contagion are shown in Figure 3: those at the steps of time  $t=0, 10, 50, 100$  days.

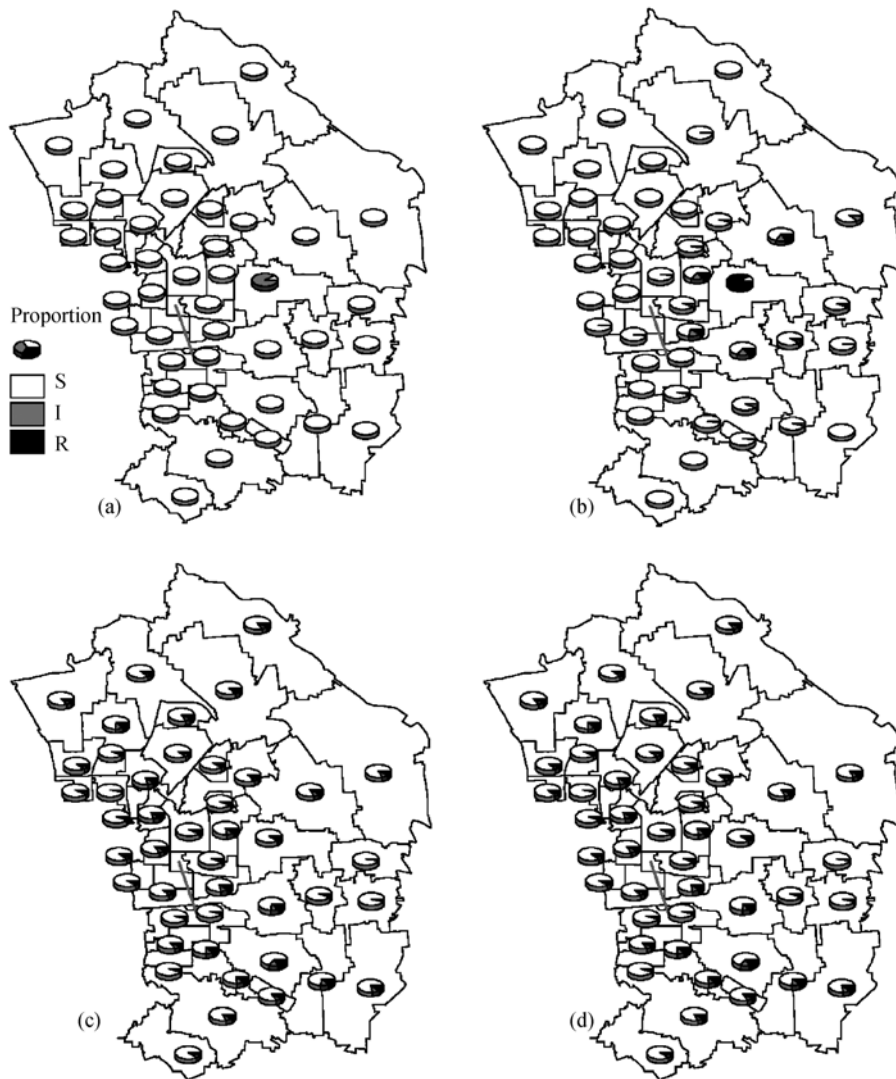
The curves that depict the numbers of Susceptible, Infective and Recovered change according to the steps of time are shown in Figure 4. As seen, the spread will arrive at a global stationary state, i.e., the numbers of the individuals of every subclass will not change any more after a long period. Furthermore, by comparing Figure 3(c) with Figure 3(d), we found that the local stationary state is also reached (the proportions of the individuals of each subclass are nearly same in the two configurations). Through running the simulation program with different values of  $\lambda(q)$  and initial configurations, we also found that a threshold  $\lambda_0$  exists. When  $\lambda(q)>\lambda_0$ , the prevalence can hold and arrive at a stationary state finally. However, when  $\lambda(q)<\lambda_0$ , the prevalence cannot persist. The final state (whether stationary or not) has nothing to do with the initial configuration. In this case of simulation, the spread presents a typical character of local diffusion.

The second case of simulation considers not only the local contagion but also the global contagion. The weights are calculated according to the proximity  $\delta_{ij}$  (0, not adjacent; 1, adjacent) and the number of road links  $n_{ij}$  between two regions (eq. (11)):

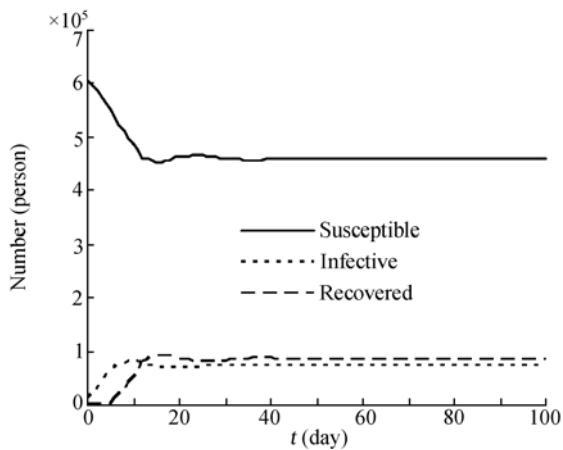
$$c_{ij} = \begin{cases} 0, & \delta_{ij} = 0 \text{ and } n_{ij} = 0, \\ 0.3, & \delta_{ij} = 0 \text{ and } n_{ij} > 0, \\ (z_{ij}/l_i + z_{ji}/l_j)/2, & \delta_{ij} = 1 \text{ and } n_{ij} = 0, \\ \max(0.3, (z_{ij}/l_i + z_{ji}/l_j)/2), & \delta_{ij} = 1 \text{ and } n_{ij} > 0. \end{cases} \quad (11)$$

Four configurations considering the additional global contagion caused by transportation lines are shown in Figure 5.

Figures 6 and 7 show the change curves of the numbers with different BECR under the local and the global contagions respectively. Comparing the two figures, we found that a threshold also exists but the threshold is slightly lower when considering the global contagion. With same BECR, the spread with the global contagion



**Figure 3** The simulation results considering only local contagion. The steps of time  $t=0$  days (a), 10 days (b), 50 days (c), 100 days (d).



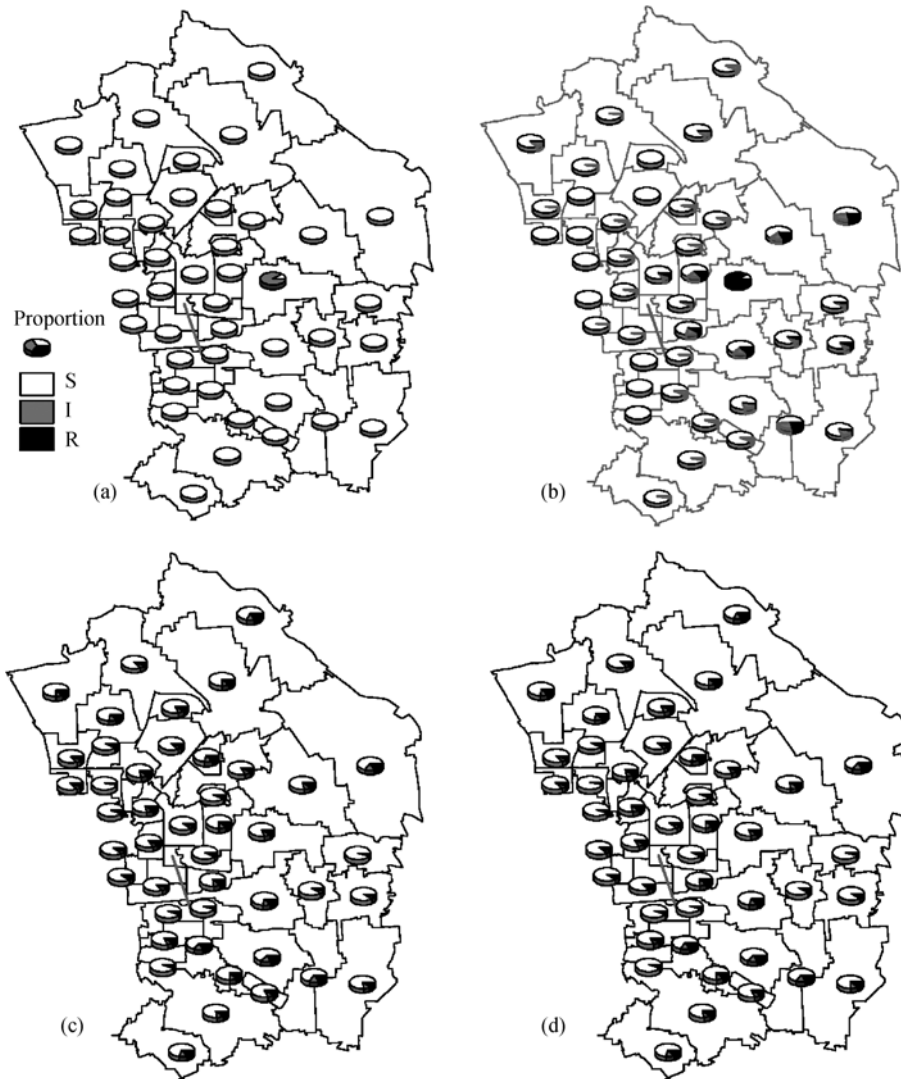
**Figure 4** The change curves of the number for each subclass individuals.

is faster and the number of Infective under the stationary state is more. In fact, with more and more complex road

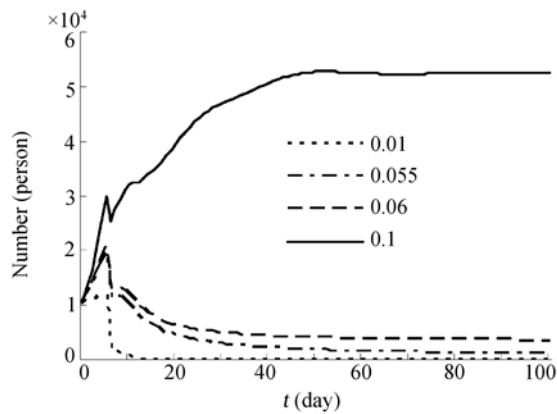
links, correspondingly the threshold is getting progressively lower. With the global contagion, the spread shows aggression along transportation lines.

In the third case of simulation, we consider the isolation of individuals after they are infected. The operation of isolation is equivalent to lowering the density of Infective distribution in regions. At every step of time, the individuals meeting the conditions of isolation are removed from Infective in a region and the density of Infective distribution is calculated dynamically. The curves with different isolation times (those on 1th, 3th, 5th days after infected) are drawn in Figure 8. As seen, the earlier isolation is carried out, the better control effort can be obtained.

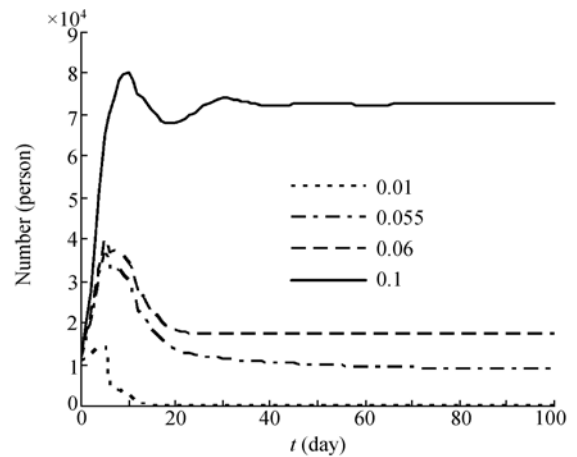
In the fourth case, we will simulate the effect of quarantine which prevents Infective from moving among



**Figure 5** The simulation results considering global contagion. The steps of time  $t=0$  day (a), 10 days (b), 50 days (c), 100 days (d).



**Figure 6** The change curves of the numbers of Infective with different BECR considering only local contagion.

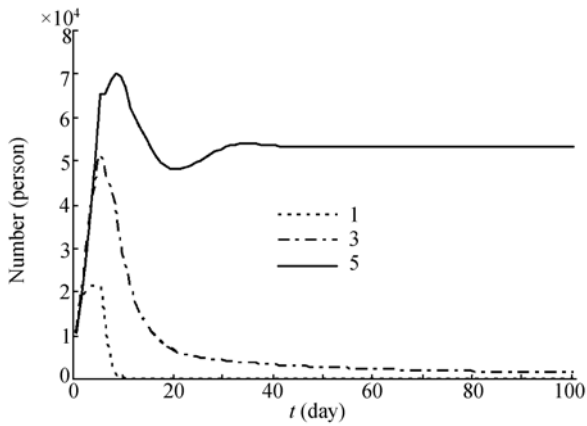


**Figure 7** The change curves of the numbers of Infective with different BECR considering global contagion.

regions through exit/entry. Given the isolation strength is  $\omega \in [0, 1]$ , the quarantine effort is equal to changing

the weighting function in eq. (11) into the following form:

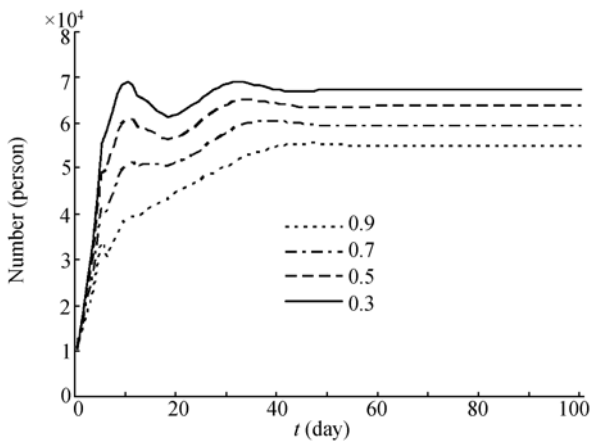




**Figure 8** The curves with different isolation times (those on the 1th, 3th, 5th day after infected).

$$c_{ij} = \begin{cases} 0, & \delta_{ij} = 0 \text{ and } n_{ij} = 0, \\ 0.3(1 - \omega), & \delta_{ij} = 0 \text{ and } n_{ij} > 0, \\ (z_{ij}/l_i + z_{ji}/l_j)/2, & \delta_{ij} = 1 \text{ and } n_{ij} = 0, \\ \max(0.3(1 - \omega), (z_{ij}/l_i + z_{ji}/l_j)/2), & \delta_{ij} = 1 \text{ and } n_{ij} > 0. \end{cases} \quad (12)$$

Figure 9 illustrates the curves of the number of Infective with different  $\omega$ . The curves show that when the quarantine strength  $\omega$  is getting higher and higher, the number of Infective under stationarity is getting lower and lower. Nevertheless, as  $\omega$  increases, the decrease of the number of final infected individuals is not remarkable. The reason is that the weights assigned for region pairs without road links between them are comparatively high. And yet the quarantine for this situation is not considered in this case of simulation. Imaginably, given this kind of quarantine or equivalent (e.g., refrain the interaction between two adjacent regions), the number under stationarity will decrease greatly.



**Figure 9** The curves of the number of Infective with different  $\omega$  (those with 0.3, 0.5, 0.7 and 0.9).

## 6 Case study

In the following study, we use the epidemic data of SARS in Beijing of China to validate the proposed model in the above sections.

SARS is a new infectious disease first reported in November 2002 in the Guangdong province of China (www.who.int/sarsarchive/2003-03-12/en). In Beijing, the first SARS case was confirmed on March 5, 2003. From then on, SARS spread rapidly over the city since people have hardly recognized SARS and some interventions for prevention and control of the disease have not been carried out in the early phase of the epidemic. Here, we will compare the execution results of the proposed model with the real observed disease data. In order to reduce the bias caused by interventions, the disease data from March 5 to April 5 are selected (these data are supposed to be affected hardly due to little interventions such as isolation, quarantine and so on) and they are compared with the simulating results from the model.

The map in Figure 10 demonstrates the political areas of Beijing, which are looked upon as cells. And we calculate the distribution of Beijing of 2003 by means of YUE-SMPD<sup>[26]</sup>.

Given the fact that the whole city is connected with roads in length and breadth, eq. (13) is used to calculate the connection distance between two areas  $i$  and  $j$ :

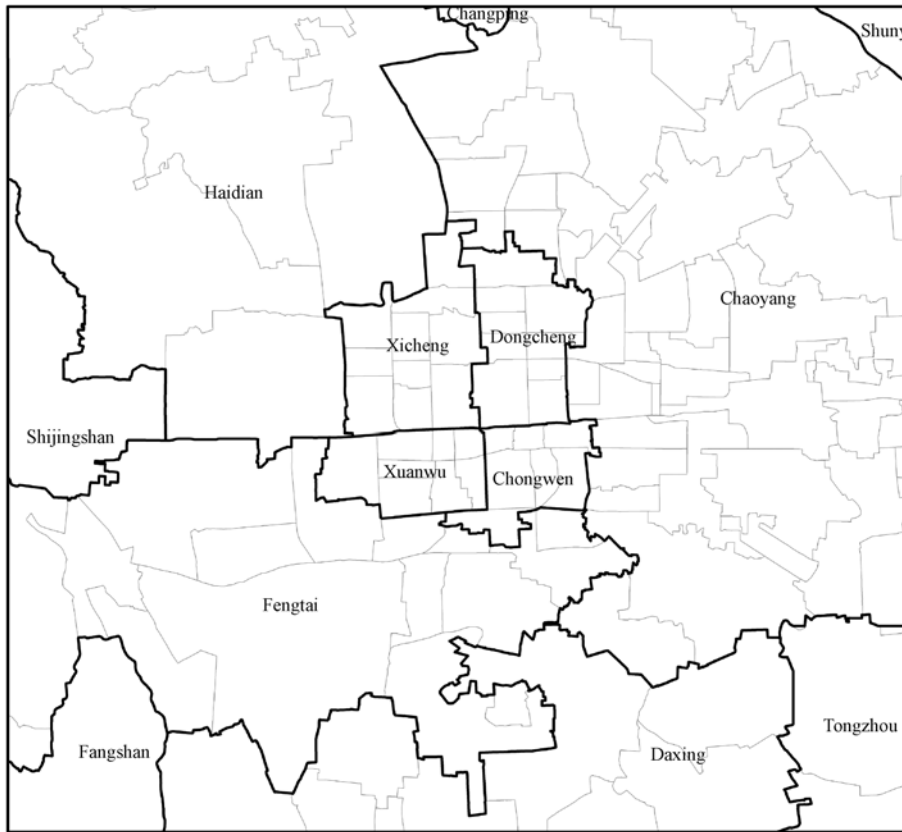
$$d_{ij} = \frac{1}{|x_i - x_j| + |y_i - y_j|}, \quad (13)$$

where  $x_i, y_i; x_j, y_j$  are the coordinates of the centroids of areas  $i, j$  respectively. Thus the connection factor between areas  $i$  and  $j$  can be quantified according to eq. (14):

$$c_{ij} = \frac{1}{1 + \sqrt{d_{ij}}}, \quad (14)$$

where the unit for  $d_{ij}$  is km.

Based on the existing research<sup>[27-32]</sup>,  $T_i, T_p, T_b, T_r$  are assigned with the following parameters: 5, 5, 7, 9999 days (this means the recovered are permanently immune to SARS, which is based on the fact that nobody is re-infected with SARS when recovered from the disease).  $\gamma(q), q \in \{1, 2, \dots, T_i + T_p + T_b\}$  is fixed to 0 (in the early phase of SARS, little interventions are carried out). The simulation time is 30 days (from March 5 to April 5). We change the model parameters  $\lambda$  to run the model and the results of model simulation are drawn in Figure



**Figure 10** The political area map of Beijing.

11 when  $\lambda(q)$ ,  $q \in \{1, 2, \dots, T_i + T_p + T_i\}$  is assigned with 0.0025.

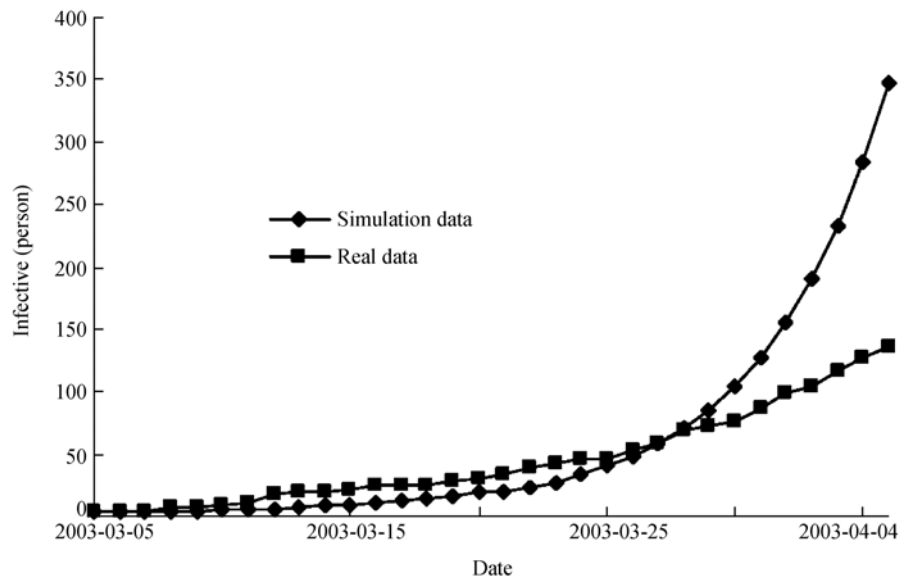
Figure 11 shows that the simulated data have good consistence with the real disease data before March 22. However, from March 22 on, the curve for the simulated data has remarkable rise while that for the real data are still slight increase. The reason for this phenomenon is likely that the interventions have been augmented with increasing respect with the disease. Furthermore, the real road network is quite complex so that it is very difficult to precisely calculate the connection factors without detailed road data. But imaginably, given perfect road data and well defined road network topology, the more practical connection factors can be calculated and the model will achieve higher accuracy.

## 7 Conclusions

Since the SIR model was first published, it has being widely used in modeling transmission of contagious diseases. It suffers from the assumption of mass action, or the homogenous interaction between populations. A lot of efforts are required to relax the assumption. The

paper proposes to embed CA model under SIR framework; a new effort has been made to revise SIR model and the study on the simulation and real-world cases shows that work. In this paper, a theoretical model of infectious diseases is introduced. It incorporates two main heterogeneities: population heterogeneity and distribution heterogeneity which relax the assumption of homogeneity. It employs geo-entity based CA considering physical geographical regions as cells. The model incorporates two categories of parameters (the BECR and the cured rate) that may express man's interventions, through which the evaluation of the control effect (e.g., therapy, isolation, immunization) is executable.

In the proposed model, the parameter correlation factor  $c_{ij}$  reflects the force of the external interaction between region  $i$  and region  $j$ . Quantifying the variable reasonably is crucial for effective simulation of an infectious disease. A spatial weighting function is a set of rules that assign values or "weights" to every pair of locations in a study area and it is appropriate for quantifying the correlation factors. Based on spatial weighting functions, some new neighbor patterns can be applied in the model, which are different from those used in the



**Figure 11** The curves of simulated data and real data (the data are all the individuals with Infective states).

classic CA and supposed to more reasonably depict the spatial interaction of infectious diseases. Consequently, the model is applicable for not only the diseases with local contagion but also ones with global contagion (e.g., SARS, Foot and mouth disease and Avian Influenza). The model has some advantages over the classic CA. It can couple cells with geographical phenomena closely. Thus, the incorporation of region related parameters, such as various attributes for geographical regions, spatial weighting functions and so on, becomes more direct and convenient.

According to the model, a simulation program is developed in GIS. For two typical contagion patterns: the local contagion and the global contagion, some results of simulation are obtained. And the results show that the model is adaptable for modeling the complicated phenomena of the spread of diseases in real world, which is nearly impossible to be achieved with differential equations because of the complexity of the problem. Fur-

thermore, the model couples visualization with complexity and is supposed to effectively simulate the dynamics of infectious diseases on complex networks which is unreachable with either the classic CA or analytical complex networks. In this paper, the cases of simulation considering isolation and quarantine are also presented, which demonstrate several sensible uses of the model in control and prevention of infectious diseases. The case study of Beijing SARS also shows that the simulation results of the model are in good agreement with the real disease data, which in a sense explains the applicability of the proposed model. Through visualizing and examining efforts of all kinds of interventions, the model can provide decision-making support for prevention and control of infectious diseases.

*We thank Prof. Weng W. G. for his critical reading of the manuscript and Li B. T. for his careful copy-editing of it. The anonymous reviewers are thanked for their constructive reviews.*

- White S H, del Rey A M, Sánchez G R. Modeling epidemics using cellular automata. *Appl Math Comput*, 2007, 186: 193–202
- Fuentes M A, Kuperman M N. Cellular automata and epidemiological models with spatial dependence. *Physica A*, 1999, 267: 471–486
- Anderson R M, May R M. *Infectious Diseases of Humans: Dynamics and Control*. Oxford: Oxford University Press, 1991
- Ahmed E, Agiza H N. On modeling epidemics including latency, incubation and variable susceptibility. *Physica A*, 1998, 253: 347–352
- Sirakoulis G C, Karafyllidis I, Thanailakis A. A cellular automaton model for the effects of population movement and vaccination on epidemic propagation. *Ecol Model*, 2000, 133: 209–223
- Wang J F. *Spatial Analysis*. Beijing: Science Press, 2006
- Von Neumann J. *Theory of Self-Reproducing Automata*. Urbana: University of Illinois Press, 1966
- Liu Q X, Jin Z. Cellular automata modelling of SEIRS. *Chin Phys*, 2005, 14: 1370–1377
- Mikler A R, Venkatachalam S, Abbas K. Modeling infectious diseases using global stochastic cellular automata. *J Biol Syst*, 2005, 13: 421–439
- Huang C Y, Sun C T, Hsieh J L, et al. Simulating SARS: Small-world

- epidemiological modeling and public health policy assessments. *J Artif Soc Soc Simul*, 2004, 7(4), <http://jasss.soc.surrey.ac.uk/7/4/2.html>
- 11 Zhou C H, Sun Z L, Xie Y C. *Geographical Cellular Automata*. Beijing: Science Press, 1999
  - 12 Liu X P, Li X, Anthony G O E, et al. Discovery of transition rules for geographical cellular automata by using ant colony optimization. *Sci China Ser-D Earth Sci*, 2007, 50(10): 1578—1588
  - 13 Flache A, Hegselmann R. Do irregular grids make a difference? Relaxing the spatial regularity assumption in cellular models of social dynamics. *JASSS*, 2001, 4(4), <http://jasss.soc.surrey.ac.uk/4/4/6.html>
  - 14 Moreno N, Ménard A, Marceau D J. VecGCA: An vector-based geographic cellular automata model allowing geometric transformations of objects. *Environ Plann B*, 2008, 35(4): 647—665
  - 15 Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. *Phys Rev Lett*, 2001, 86: 3200
  - 16 Wang X F. Complex networks: Topology, dynamics and synchronization. *Int J Bifurcat Chaos*, 2002, 12: 885—916
  - 17 Barthélemy M, Barrat A, Pastor-Satorras R, et al. Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. *Phys Rev Lett*, 2004, 92(1): 178701
  - 18 Shirley M D F, Rushton S P. The impacts of network topology on disease spread. *Ecol Complex*, 2005, 2: 287—299
  - 19 Draief M. Epidemic processes on complex networks: The effect of topology on the spread of epidemics. *Physica A*, 2006, 363: 120—131
  - 20 Silva S L, Ferreira J A, Martins M L. Epidemic spreading in a scale-free network of regular lattices. *Physica A*, 2007, 377: 689—697
  - 21 Allman E S, Rhodes J A. *Mathematical Models in Biology: An Introduction*. Cambridge: Cambridge University Press, 2004
  - 22 Tobler W R. A computer movie simulating urban growth in the Detroit region. *Econ Geol*, 1970, 46: 234—240
  - 23 Odland J. *Spatial Autocorrelation*. California: Sage Publications, 1988
  - 24 Cliff A D, Ord J K. *Spatial Processes: Models and Applications*. London: Pion, 1981
  - 25 Tobler W R. *Linear Operators Applied to Areal Data*. London: John Wiley, 1975
  - 26 Yue T X, Wang Y A, Zhang Q, et al. YUE-SMPD scenarios of Beijing population distribution (in Chinese). *Geo-information Sci*, 2008, 10(4): 479—488
  - 27 Meng B, Wang J F. Understanding the spatial diffusion process of SARS in Beijing. *Public Health*, 2005, 119: 1080—1087
  - 28 Wang J F, McMichael A J, Meng B, et al. Spatial dynamics of an epidemic of severe acute respiratory syndrome in an urban area. *Bull World Health Organ*, 2006, 84: 965—968
  - 29 Bombardt J N. Congruent epidemic models for unstructured and structured populations: Analytical reconstruction of a 2003 SARS outbreak. *Math Biosci*, 2006, 203: 171—203
  - 30 Bowen Jr J T, Laroe C. Airline networks and the international diffusion of severe acute respiratory syndrome (SARS). *Geogr J*, 2006, 72: 130—144
  - 31 Ruan S, Wang W, Levin S A. The effect of global travel on the spread of SARS. *Math Biosci Eng*, 2006, 3: 205—218
  - 32 Wang J F, Christakos G, Han W G, et al. Data-driven exploration of “spatial pattern-time process-driving forces” associations of SARS epidemic in Beijing, China. *J Public Health*, 2008, 30(3): 234—244