

Modeling of epidemic spreading with white Gaussian noise

GU Jiao^{1*}, GAO ZongMao² & Li Wei^{1,3}

¹Max-Planck-Institute for Mathematics in the Sciences, Leipzig 04103, Germany;

²College of Physical Science and Technology, Central China Normal University, Wuhan 430079, China;

³Complexity Science Center, Institute of Particle Physics, Central China Normal University, Wuhan 430079, China

Received April 6, 2011; accepted May 24, 2011

Motivated by the need to include the different characteristics of individuals and the damping effect in predictions of epidemic spreading, we build a model with variant coefficients and white Gaussian noise based on the traditional SIR model. The analytic and simulation results predicted by the model are presented and discussed. The simulations show that using the variant coefficients results in a higher percentage of susceptible individuals and a lower percentage of removed individuals. When the noise is included in the model, the percentage of infected individuals has a wider peak and more fluctuations than that predicted using the traditional SIR model.

epidemic process, SIR model, white noise, variant coefficients

Citation: Gu J, Gao Z M, Li W. Modeling the epidemic spreading with white Gaussian noise. *Chinese Sci Bull*, 2011, 56: 3683–3688, doi: 10.1007/s11434-011-4753-z

The history of *Homo sapiens* is always closely interrelated with the study of diseases. Whether it be the Black Death in Europe in the fourteenth century or the superbug that recently spread across the globe like wildfire [1,2], scientists have always looked at diseases with great interest [3–10]. In the well-known theoretical SIS model, the population is divided into two disjoint classes, susceptible individuals and infected individuals, for which the percentages at time t are denoted by $s(t)$ and $i(t)$, respectively [11–14]. However, this model is not suitable for describing diseases like flu or malaria. The subsequent Kermack-McKendrick SIR model [15–20] has a third class with percentages denoted by $r(t)$. They are the recovered individuals who are immune to the infection. This is a simple model commonly used for many infectious diseases like measles, mumps and rubella. In these two models, the spread of an infectious disease in a population depends mainly on the character of the disease. The most suitable model can be chosen according to the particular case. However, these models do not work well for some special cases. For example, in the spreading of SARS, although an individual cannot develop lasting immunity,

there is only a very tiny chance of being infected again.

By mean field theory, the course of an epidemic spreading is determined by the contact rates among susceptible, infected and removed individuals, which are assumed to be proportional to the number of encounters among susceptible, infected and removed individuals. Each individual is treated in the same way, even at different times. A more realistic and interesting model of infectious diseases should take into account a change in the environment and the variety of characteristics of individuals [21]. Along with the spreading process, the preventive effect should become stronger because people could find useful methods such as taking pills and avoiding close contacts with infected individuals to prevent the disease from spreading [22,23]. Moreover, how easy or difficult it is for an individual to be infected by others depends on their own characteristic, such as age, nutritional status, sex and so on. Therefore, each individual should be distinguished by different transmission coefficients and recovery coefficients. Furthermore, the part of the model that measures the coefficient of infection during the whole epidemic process should decrease as the process decreases. From this point of view, probabilistic or stochastic models are necessary for dealing with populations in

*Corresponding author (email: jiaogu@mis.mpg.de)

which individuals have different characteristics. Our research is mainly motivated by those observations. In this paper, we study certain classes of equations which can be interpreted as stochastic models of epidemic processes.

1 Our model

Several approaches have been used to find a mathematical description for the spreading of a disease through a population. In many disease models, for simplicity, the population at time t is divided into three disjoint subpopulations [24]:

$s(t)$, the percentage of susceptible individuals;

$i(t)$, the percentage of infected individuals;

$r(t)$, the percentage of removed individuals who are immune to the infection.

In the traditional SIR model, the mathematical description can be written as

$$\begin{aligned} ds/dt &= -\beta i(t)s(t), \\ di/dt &= \beta i(t)s(t) - \gamma i(t), \\ dr/dt &= \gamma i(t). \end{aligned} \tag{1}$$

In eq. (1), β and γ are constant ratios called the transmission and recovery coefficients, respectively. From the definition of $s(t)$, $i(t)$ and $r(t)$, we have

$$s(t) + r(t) + i(t) = 1. \tag{2}$$

Therefore,

$$\frac{ds}{dt} + \frac{di}{dt} + \frac{dr}{dt} = 0. \tag{3}$$

In this model, it appears that the characteristics for different individuals are lost. In a real system, individuals may have different infection rates so the transmission and recovery coefficients should contain stochastic parts to describe the characteristics for different individuals. Furthermore, the transmission coefficient should contain a factor incarnating the subdued infected effect (enhanced defensive effect). Usually, the factor should be a damping function that is dependent on the evolutionary time t . Hence in our new model, the transmission coefficient β and recovery coefficient γ are as follows:

$$\beta = \beta_0 \cdot s(t) + \zeta_1(t), \tag{4}$$

and

$$\gamma = \gamma_0 + \zeta_2(t), \tag{5}$$

where $s(t)$ represents the damping factor. $s(t)$ is a decreasing function of time because susceptible individuals can change into infected ones, while others cannot normally change into susceptible ones. β and γ are positive constants. $\zeta_1(t)$ and $\zeta_2(t)$ are independent Gaussian white noises which reflect the fluctuations of transmission and recovery, respectively, for individuals. The white noise is caused by different body

constitutions or by other factors. Indeed, different distributions can be chosen to describe the noises here. In the present paper, we have chosen to use the simple Gaussian distribution.

2 Analytic results

First, we analyzed the equations without Gaussian noise. In this case, the parameters can be written as

$$\beta(t) = \beta_0 s(t), \tag{6}$$

and

$$\gamma(t) = \gamma_0. \tag{7}$$

Because the parameters β and γ are positive constants, we denote their ratio β/γ as a positive constant α . The parameters in eqs. (6) and (7) were put into eq. (1) using the conditions in eqs. (2) and (3). Integrating the differential equation for $s(t)$, $i(t)$ and $r(t)$ with initial condition $s(0) > 1$ and $r(0) > 1$ (assuming $i(0) > 0$, a very small initial group of infected individuals), the analytic solutions for $s(t)$ and $i(t)$ were obtained as functions of $r(t)$:

$$s(t) = \frac{1}{\alpha r(t) + 1}, \tag{8}$$

$$i(t) = 1 - r(t) - \frac{1}{\alpha r(t) + 1}.$$

The function $r(t)$, which represents the recovery percentage, monotonically increases because infected individuals have a certain probability of recovery but no individual can switch from the removed class to other classes. Therefore, the function $s(t)$ is a decreasing one. However, $i(t)$ is not monotonous. By calculating the first and second order derivatives of $i(t)$, we can find the maximal value of $i(t)$:

$$i(t)_{\max} = 1 + \frac{1}{\alpha} - \frac{2}{\sqrt{\alpha}}, \tag{9}$$

when

$$\gamma(t) = \frac{1}{\sqrt{\alpha}} - \frac{1}{\alpha}. \tag{10}$$

Because $r(t) \in [0,1]$, the parameter should satisfy the condition $\alpha \geq 1$. If α is not in this range, then $i(t)$ attains its maximal value at the very beginning.

Furthermore, the value of $i(\infty)$ is equal to 0. In a real system, infected individuals originate from susceptible ones. However, the number of susceptible individuals is limited and all infected individuals will become removed ones with $s(t)$ being zero or nonzero. This means that, when $t \rightarrow \infty$, $i(t) = 0$ and $s(\infty) = 1 - r(\infty)$. Putting this into eq. (8), we get

$$s(\infty) = \frac{1}{\alpha r(\infty) + 1} = 1 - r(\infty), \tag{11}$$

which also means that

$$r(\infty) = 1 - \frac{1}{\alpha}. \tag{12}$$

The function $r(t)$ monotonically increases, so the maximal value of $r(t)$ always occurs at the final moment. The maximal value is less than 1 and approaches 1 when the ratio β/γ is large enough. This result has a biological interpretation. If the recovery parameter γ is larger than the spreading parameter β , then infected individuals can become recovery ones more easily than the susceptible individuals become infected ones. When all the infected individuals become removed ones, the remaining susceptible ones have no chance of becoming infected ones. Thus, the number of susceptible individuals is not equal to 0, and the percentage of removed individuals will obviously be less than 1. On the other hand, if β is larger than the recovery parameter γ , after almost all the susceptible individuals become infected ones, all the infected individuals will slowly become recovery ones as the spreading process proceeds. In this case, the percentage of removed individuals is close to 1. This result varies when noise is included in the models.

We can analyze the effect that the introduction of noise parameters would have on the epidemic spreading. Because the percentage of susceptible individuals is monotonically decreasing, the value of $\beta(t)s(t)$ will tend to zero when $s(t)$ is close to zero. From eqs. (4) and (5), we can see that the noise parameters will become the main factors under this condition. Therefore, the $\beta(t)s(t)$ value can be ignored when $\beta(t)s(t)$ is close to zero. In this case, the coefficients are

$$\beta = \zeta_1(t) \tag{13}$$

and

$$\gamma = \gamma_0 + \zeta_2(t). \tag{14}$$

If the fluctuations of the noises are large enough, even when $s(t)$ is small, the change from susceptible individuals to infected ones can occur even with a tiny number of infected seeds. Indeed, we used the noise to describe the difference among individuals, and this satisfied simply Gaussian distribution. Because of the presence of noise, the spreading process contains more turbulence.

3 Numerical analysis

Some simulations which were performed are briefly described below.

Figure 1 shows a comparison between our model, the traditional SIR model and the real data for SARS in Beijing in 2003. The population of Beijing in 2003 was considered to be 1.5×10^7 . From Figure 1, we can see that the traditional SIR model fits the peak value well, but not the peak width. The shorter predicted epidemic process may be because this

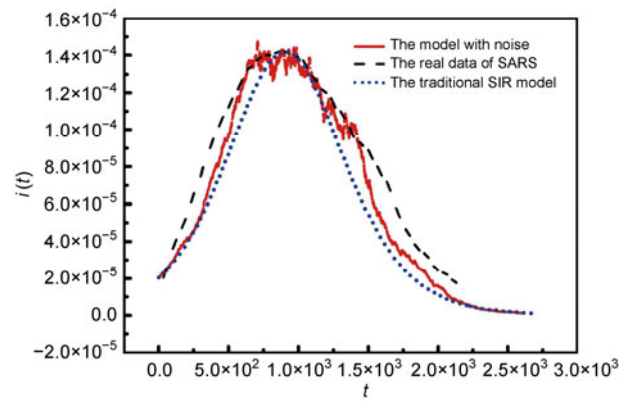


Figure 1 (Color online) A comparison between the real data of SARS in Beijing in 2003 and the numerical simulations using the traditional SIR model and our model. One day in the real data is approximated by 33.3 steps in the simulations. The parameters in the traditional SIR model were $\beta = 0.21223$, $\gamma = 0.21$; in our model they were $\beta=0.153373$, $\gamma = 0.15$. The mean and the mean square values of the variable $\zeta_1(t)$ were 0 and 0.01. The mean and the mean square values of the variable $\zeta_2(t)$ were 0 and 0.005. The same initial condition were used in both models: $s(0) = 0.9999762$, $i(0) = 2.04 \times 10^{-5}$ and $r(0) = 3.4 \times 10^{-6}$.

model ignores the different characteristics of individuals. Because of the differences between infected individuals, as expected, the epidemic process is longer and contains more fluctuations than the traditional SIR. Considering the differences among the individuals, including the preventive methods used, several susceptible persons still become infected because of their close contact with infected ones. Doctors or individuals with weak immunities would be in this category. This will prolong the whole epidemic process, especially the epidemic peak. Therefore, the model that includes the noise can, under certain conditions, fit the real data better.

Here we describe the effect of the damping factor. It can increase the saturation value of susceptible individuals. To show this effect, the numerical simulations were performed using the traditional SIR model and our model with and without noise and the results are shown in Figure 2. By including the damping factor under the same conditions, the saturation value of $s(t)$ predicted by the traditional model is larger than the saturation value from our model. The saturation value of $r(t)$ predicted by the traditional model is smaller than that from our model. This is because, under the control of the damping factor, the probability of individuals becoming infected would decrease. Additionally, the effect of including noise in our model is clearly seen in Figure 2(b) because, in that model, $i(t)$ shows distinct fluctuation.

In a series of simulations, we explored the effect of noise in our model. Figure 3 shows that the larger the mean squares, the greater the predicted fluctuations on the epidemic process. When the mean squares of $\zeta_1(t)$ and $\zeta_2(t)$ are 1 and 0.2, two peaks are observed in the process. Clearly, the inclusion of noise in the model leads to predictions of a prolonged epidemic process with more fluctuations, especially

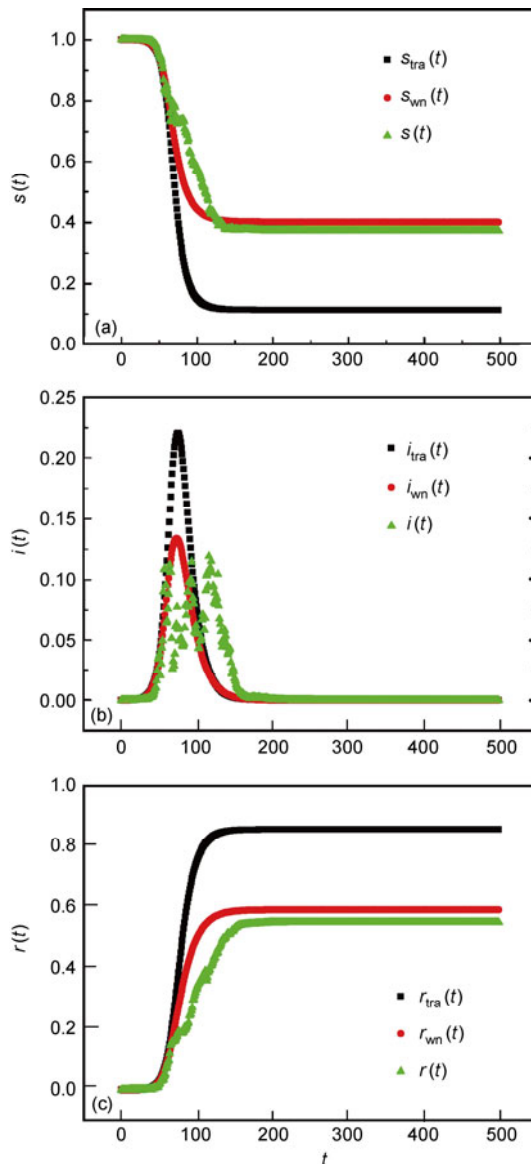


Figure 2 (Color online) Comparisons between the traditional model (with subscript “tra”) and our model with (with no subscript) and without noise (with subscript “wn”). The initial conditions used were $s(0) = 0.99997$, $i(0) = 0.00003$ and $r(0) = 0$. The parameters in the traditional model were $\beta = 0.25$, $\gamma = 0.1$; in our model without noise they were $\beta = 0.25$, $\gamma = 0.1$; in our model with noise they were $\beta = 0.25$, $\gamma = 0.1$. The mean and the mean square values of the noise $\zeta_1(t)$ were 0 and 0.25, and the mean and the mean square values of the noise $\zeta_2(t)$ were 0 and 0.1. (a)–(c) show the effect of the damping factor on the percentages of susceptible, infected and removed individuals, respectively, in the different models.

during the peak time.

We discussed the effect of the ratio β/γ when describing the analytical results from eqs. (11) and (12) above. The effect is similar to the effect predicted by the numerical simulations with noise. In Figure 4, when the ratio is large, for example $\alpha = 10$, susceptible individuals are infected more quickly than infected ones recover. Even with a small number of susceptible individuals, infected ones change into removed ones slowly. This results in smaller saturated

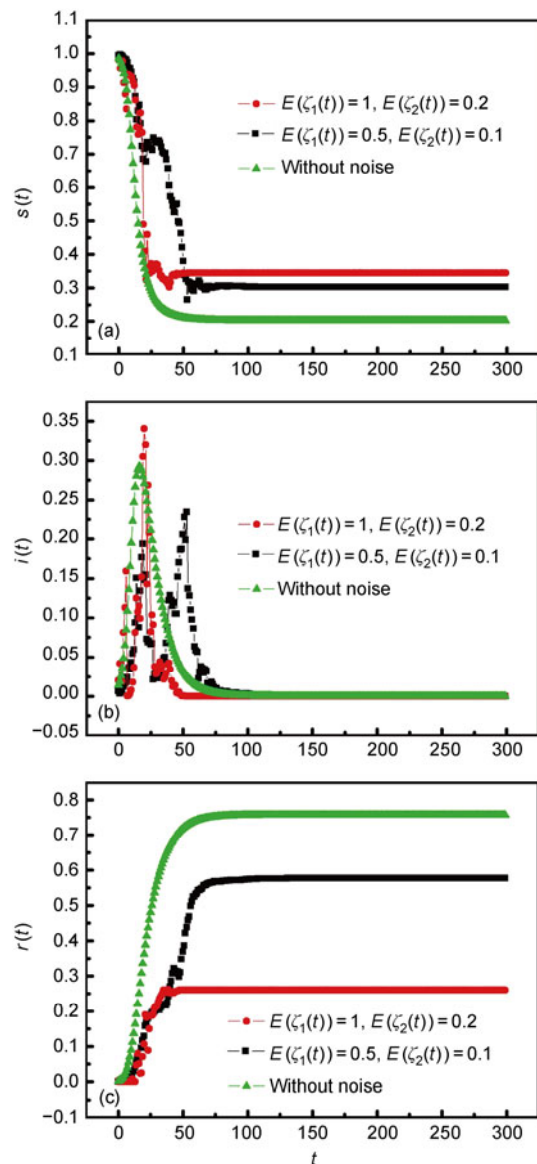


Figure 3 (Color online) The variations of $s(t)$ (a), $i(t)$ (b) and $r(t)$ (c) with the time step t using different noise parameters in our model under the conditions: $s(0) = 0.99$, $i(0) = 0.01$, $r(0) = 0$, $\beta = 0.5$ and $\gamma = 0.1$.

values for $s(t)$ and $r(t)$. Conversely, susceptible individuals cannot become infected ones quickly when there are a small number of infected individuals because there are not enough infected seeds. An extreme example is AIDS spreading. The recovery coefficient is near to zero (because the infected person usually dies after several or many years), so the ratio is almost infinity. If no cure is found for the infected people or no preventive method to prevent its spread to susceptible people is available, then this disease will spread across the globe even over a long time period.

4 Conclusion

In this paper, we report the development and application of

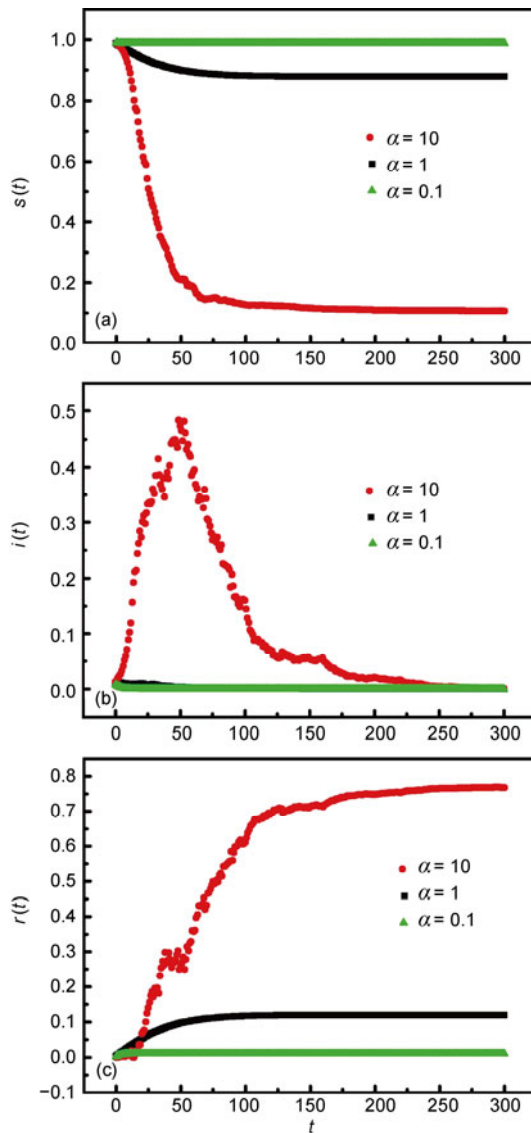


Figure 4 (Color online) The variations of $s(t)$ (a), $i(t)$ (b) and $r(t)$ (c) with the time step t using different ratios of β/γ in our model. The initial conditions were $s(0) = 0.99$, $i(0) = 0.01$ and $r(0) = 0$. The mean and the mean square values of the variable $\zeta_1(t)$ and $\zeta_2(t)$ were 0 and 0.05. $\alpha = 10$, means $\beta = 0.25$ and $\gamma = 0.025$; $\alpha = 1$ means $\beta = 0.25$ and $\gamma = 0.25$; and $\alpha = 0.1$ means $\beta = 0.025$ and $\gamma = 0.25$.

a modified SIR model with varied coefficients and white Gaussian noise. In the traditional SIR model, it assumed that the ensemble transmission effect can be described by a constant. However, this ignores the use of preventive and control methods along the epidemic process. Therefore, we included a damping factor, which is normally a decreasing function, to control the epidemic. Because of the presence of this factor, the saturation value of $s(t)$ was larger than in the traditional SIR model and the saturation value of $r(t)$ was smaller than that in the traditional SIR model.

Additionally, considering the different characteristics of individuals, the transmission coefficient and recovery coefficient were supposed to contain independent Gaussian

noises representing the differences among individuals in a population. Because the damping factor weakens the infected effect, when the noise is included and the time is large enough, the susceptible individuals can still be infected. Thus, by including the noise the model can predict a prolonged epidemic process producing the peak with some fluctuations.

The analytic results also showed, that without the noise, the percentage of infected individuals $i(t)$ can attain its maximal value under given conditions. The percentage of removed individuals $r(\infty)$ is close to 1 when the ratio β/γ is large enough. When the noise and the modified coefficients were included, our model fitted the real data of SARS in Beijing in 2003 well, because it predicted the wider peak for $i(t)$. In summary, we propose the use of variant coefficients and white Gaussian noise to describe the epidemic process. Including these parameters in the models might give us some new understandings about the epidemic process. Exploring this idea for other models and networks might produce useful results and is worthy of further study.

We would like to thank the Max-Planck-Institute for Mathematics in the Sciences, Leipzig, Germany for the support. We also thank Shipping Liu for useful discussion and helpful suggestions. This work was supported in part by the National Natural Science Foundation of China (10647125, 10635020, 10975057, 10975062) and the Programme of Introducing Talents of Discipline to Universities (B08033).

- George C. Interdisciplinary Public Health Reasoning and Epidemic Modeling: The Case of Black Death. Berlin: Springer, 2005
- James K. Superbugs spread to the community. *New Scientist*, 2007, 195: 37–39
- Weng P X, Zhao X Q. Spreading speed and traveling waves for a multi-type SIS epidemic model. *J Diff Equat*, 2006, 229: 270–296
- Han X P. Disease spreading with epidemic alert on small-world networks. *Phys Lett A*, 2007, 365: 1–5
- Liu T, Li X, Liu X P. Integration of small world networks with multi-agent systems for simulating epidemic spatiotemporal transmission. *Chinese Sci Bull*, 2010, 55: 1285–1293
- Cao Z D, Zeng D J, Wang Q Y. An epidemiological analysis of the Beijing 2008 Hand-Foot-Mouth epidemic. *Chinese Sci Bull*, 2010, 55: 1142–1149
- He Y G, Ding G H, Bian C, et al. The mutation network for the hemagglutinin gene from the novel influenza A (H1N1) virus. *Chinese Sci Bull*, 2009, 54: 2168–2170
- Chen H L. H5N1 avian influenza in China. *Sci China Ser C-Life Sci*, 2009, 52: 419–427
- Gu H, Chen H F, Wei D Q, et al. Molecular dynamics simulations exploring drug resistance in HIV-1 proteases. *Chinese Sci Bull*, 2010, 55: 2677–2683
- Cao C X, Xu M, Chang C Y, et al. Risk analysis for the highly pathogenic avian influenza in Mainland China using meta-modeling. *Chinese Sci Bull*, 2010, 55: 4168–4178
- Richard J K, Claude L. On the extinction of the S-I-S stochastic logistic epidemic. *J Appl Probab*, 1989, 27: 685–694
- Martin-Löf A. Diffusion approximation near the threshold in the stochastic SIS epidemic. In: *Luminy Conference on Stochastic Modeling for Infectious Diseases*, 1991
- Näsell I. The threshold of the stochastic SIS epidemic. In: *Conference on Stochastic Modeling for Infectious Diseases*, 1991
- Linda R. Asymptotic results for a multi-type contact birth-death process and related SIS epidemic. *Math Biosci*, 2007, 208: 552–570

- 15 Kermack W O, McKendrick A G. A contribution to the mathematical theory of epidemics. *Proc R Soc A*, 1927, 115: 700–721
- 16 Rvachev A L, Longini Jr I M. A mathematical model for the global spread of influenza. *Math Biosci*, 1985, 75: 3–22
- 17 Hethcote H W. Qualitative analyses of communicable disease models. *Math Biosci*, 1976, 28: 335–356
- 18 Kermack W O, McKendrick A G. Contributions to the mathematical theory of epidemics—I. 1927. *Bull Math Biol*, 1991, 53: 33–55
- 19 Kermack W O, McKendrick A G. Contributions to the mathematical theory of epidemics—II. The problem of endemicity. *Bull Math Biol*, 1991, 53: 57–87
- 20 Kermack W O, McKendrick A G. Contributions to the mathematical theory of epidemics—III. Further studies of the problem of endemicity. *Bull Math Biol*, 1991, 53: 89–118
- 21 Pei Y, Liu S, Li C, et al. The dynamics of an impulsive delay SI model with variable coefficients. *Appl Math Model*, 2009, 33: 2766–2776
- 22 Li W, Gu J, Cai X. Message spreading and forget—Remember mechanism on a scale-free network. *Chin Phys Lett*, 2008, 25: 2303–2306
- 23 Gu J, Li W, Cai X. The effect of the forget-remember mechanism on spreading. *Eur Phys J B*, 2008, 62: 247–255
- 24 Cooke K L, Yorke J A. Some equations modelling growth processes and gonorrhoea epidemics. *Math Biosci*, 1973, 16: 75–101

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.