Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica D 215 (2006) 146–158
PHYSICA D
www.elsevier.com/locate/physd

Super-spreaders and the rate of transmission of the SARS virus

Michael Small[a,*], C.K. Tse[a], David M. Walker[b]

[a] Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[b] Biomathematics and Statistics Scotland, Macaulay Institute, Craigbuckler, Aberdeen AB15 8QH, UK

Received 13 June 2005; received in revised form 28 November 2005; accepted 30 January 2006
Available online 10 March 2006
Communicated by H. Levine

Abstract

We describe a stochastic small-world network model of transmission of the SARS virus. Unlike the standard Susceptible-Infected-Removed models of disease transmission, our model exhibits both geographically localised outbreaks and "super-spreaders". Moreover, the combination of localised and long range links allows for more accurate modelling of partial isolation and various public health policies. From this model, we derive an expression for the probability of a widespread outbreak and a condition to ensure that the epidemic is controlled. Moreover, multiple simulations are used to make predictions of the likelihood of various eventual scenarios for fixed initial conditions. The main conclusions of this study are: (i) "super-spreaders" may occur even if the infectiousness of all infected individuals is constant; (ii) consistent with previous reports, extended exposure time beyond 3–5 days (i.e. significant nosocomial transmission) was the key factor in the severity of the SARS outbreak in Hong Kong; and, (iii) the spread of SARS can be effectively controlled by either limiting long range links (imposing a partial quarantine) or enforcing rapid hospitalisation and isolation of symptomatic individuals.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Severe Acute Respiratory Syndrome (SARS); Small-world network; Transmission dynamics; Super-spreader event

1. SARS

The SARS virus first appeared during October 2002 in the Guangdong province of southern China. It then passed over the border to Hong Kong and from there spread to Europe, Africa, Asia, Australia and the Americas [1]. The outbreak in 2003 infected 8422, killing 916 [2]. In this paper we focus on modelling the transmission of SARS within Hong Kong. Beside mainland China, Hong Kong suffered the greatest casualties [2]. In addition, the epidemiological data currently available for Hong Kong is far superior to that of the Chinese mainland.[1]

Two characteristic features were observed during the SARS outbreak in Hong Kong in 2003 (see Fig. 1) [3,4]: so-called super-spread events (SSE), in which a single individual initiates a large number of cases; and persistent transmission within the community. Two widely cited SSEs were observed early in the epidemic and have been the subject of much attention: at the Amoy Gardens housing estate and at the Prince of Wales hospital. Epidemiological studies [5,1] have found that in Hong Kong:

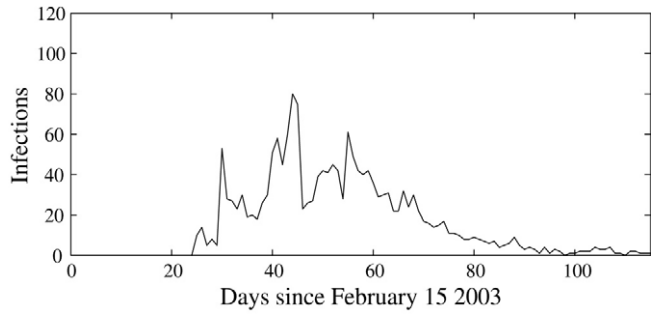
- the fatality rate was approximately 17% (compared to 11% globally);
- the mean incubation period was 6.4 days (range 2–10 days) [6];
- the duration between onset of symptoms and hospitalisation was 3–5 days; and,
- the mean number of individuals infected by each case during the initial phase of the epidemic (excluding SSEs) was 2.7 [4].

Standard deterministic SIR (susceptible-infected-removed) models of the spread of infectious diseases [7] make several important assumptions. An alternative approach [8], particularly popular for the study of sexually transmitted diseases [9–11], is to build an explicit network and model

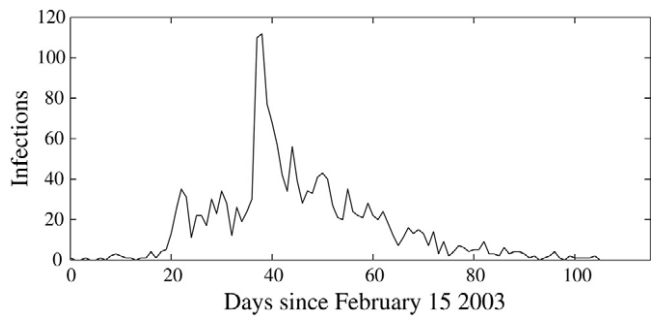
* Corresponding author. Tel.: +852 2766 4744; fax: +852 2362 8439.
E-mail address: ensmall@polyu.edu.hk (M. Small).

[1] During the epidemic, mainland authorities classified information on SARS infections as a state secret. Moreover, bureaucracy caused much of the available information to be concealed. Despite this, subsequent official investigation indicates that the infection rate for the Chinese mainland was significantly over-reported. The reliability of data from China is therefore uncertain.

0167-2789/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2006.01.021



(a) Reported.



(b) Revised.

Fig. 1. Daily SARS infection data for Hong Kong. The top panel is the daily infection rate reported in the Hong Kong media during the outbreak. The lower plot is the (revised) estimated daily number of infections determined after more detailed investigation by the Hong Kong Department of Health. At the end of the epidemic, 105 days after the first recorded case, a total of 1755 confirmed infections had been identified in Hong Kong. Both data sets are freely available.

disease transmission along the links. Under certain network structures, it is then possible to obtain a closed form study of the underlying transmission pattern [10,12]. In particular, [12] studied the transmission of an epidemic among a population whose individuals are connected both locally and globally. From this, they were able to obtain an approximately Poisson distribution for the contact distribution.

In [13], Hufnagel and co-workers also studied the local and non-local transmission of epidemics. The main features modelled by [13] are local SIR infection (with the usual stochastic differential equations) and long distance complex network links to model global aviation routes. They show that transmission dynamics with similar geographical dispersion to the 2003 SARS outbreak can develop in the model. In contrast to [13], we focus on abstract network models within a particular community. A similar global model of SARS transmission could be achieved by introducing an aviation network model on top of the models presented in this paper.

Recently, both small-world (SW) and scale-free (SF) networks have been observed in many areas of natural and physical science, including social relationships [14,15]. The important feature of both SW and SF networks is that they are highly connected: the average path length between random individuals is relatively short. Moreover, for SF networks, the node degree distribution follows a scale-free distribution. Hence, in many areas of natural and physical science, this new model structure has unveiled a rich range of behaviours. In this paper, we apply these methods to the modelling of the spread of SARS in Hong Kong; transmission is only allowed to occur along a limited number of direct links between individuals. By doing this, we will avoid making one of the assumptions underlying standard Susceptible-Infected-Removed (SIR) models: a homogeneous fully connected population. The SIR model assumes that all individuals are susceptible to the disease and all suffer an equal, small, positive probability of contracting the virus. This homogeneous model leads to a continuous and smooth inter-day distribution of infections. Irregularities in this are usually attributed to random variation and non-stationarity in the model parameters.

In proposing an alternative to the standard SIR model, we do not claim that the SIR model has failed. Certainly, the power of any model lies in its simplicity, and its ability to capture the important features of a system. Hence, like the SIR model, the complex network model that we describe here is very simple and is described by a very small number of parameters.

Analysis of the spread of SARS with SIR models shows good correlation between decreasing infection rate r and the introduction of various governmental control measures: quarantine and public awareness campaigns [16]. However, these results are based on a very simplistic model, and localised outbreaks (such as the incident at Amoy Gardens) are not modelled very well. While it is impossible to predict the occurrence of such outbreaks, SW–SF models provide a more realistic physical model of the relationships between individuals in a community and will therefore provide a better picture of the true disease dynamics. Note that, since one cannot actually *predict* when a particular SSE will occur, one cannot accurately model the timing of the peak in the time series of Fig. 1. Moreover, it is not possible to accurately model the initial SSE at the Prince of Wales hospital, which "kick started" the SARS outbreak in Hong Kong. To include the same initial spread of SARS in Hong Kong, via a rather singular SSE, that SSE must be included explicitly in the model.

In other areas where SW–SF models have been applied, a rich range of behaviours has been observed. Unlike standard differential equation based models, SW and/or SF structures model the underlying network of connections between individuals directly. Our model is designed to capture explicitly the small-world features of social interaction. Our model is not scale-free. To generate a scale-free model for disease epidemics, one needs a power-law distribution of infection links; we consider the theoretical implementation of such a model in a separate paper [16].

Finally, we will also provide one possible answer to the question posed by Galvani and May [17]: "Were SARS super-spreaders anomalies, or are super-spreaders characteristic of most infectious diseases?" [17]. We show that, even with uniform rates of infection, super-spreaders will occur, to varying degrees, in a small-world or scale-free [16] network. The implication is that super-spreaders are not (necessarily) a result of variable rate of infection. Nonetheless, we argue that the SARS outbreak in Hong Kong was initiated by a single, rather unfortunate, super-spreader event.

In the next section, we describe our model and study its behaviour analytically. Subsequent sections present some numerical simulations and a summary of our results.

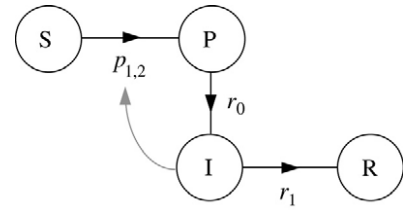
2. The model

In the following subsections, we define our model structure (Section 2.1) and derive some analytical results concerning the likelihood of a widespread outbreak (Section 2.2).

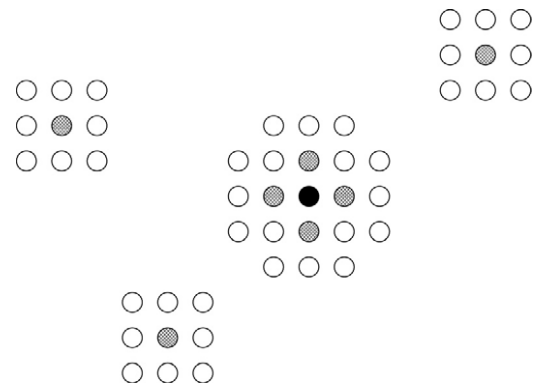
2.1. Model topology

Our aim is to capture accurately the qualitative features of the SARS epidemic with the simplest model (the fewest parameters). We propose four distinct states. Individuals can be susceptible (S), prone (P), infected (I), or removed (R). The transmission path is depicted in Fig. 2. Infected individuals can cause susceptible individuals, to whom they are linked, to become prone with some probability (p_1 or p_2). By *infection* we mean the transition from the susceptible state to the prone state. Infected individuals can cause their immediate neighbours to become infected with probability p_1; long range links cause infection with probability p_2. Prone individuals become infected with probability r_0 and, finally, infected individuals become removed with probability r_1. Just as in the SIR model, we do not distinguish fatalities from recoveries: in either case, the individuals are assumed to have acquired immunity. The model states that we describe here bear a close correspondence to the four states of the SEIR model. The *prone* state is analogous to the *exposed* (E) state in the SEIR model. We choose to use the new term P because the epidemiological state is slightly different. By *prone* we mean *both* the infected but not infectious (incubation) period and the pre-symptomatic period. For SARS, all infections in Hong Kong could be traced to contact with a symptomatic individual.[2] Hence, for the purposes of modelling transmission, we assume that, during the pre-symptomatic period, there is no chance of transmission. Moreover, to calibrate our model with the observed data (which, by definition, are hospital admissions), we further prescribe that the time between the onset of symptoms and hospitalisation is constant (or, equivalently, follows a stationary unimodal distribution).

In our model we explicitly model the geographical structure of the population. We include both "local" and "non-local" links. Because of common transmission of SARS within specific housing estates and districts in Hong Kong, and the (both real and perceived) risk of transmission in workplaces (primarily hospitals and schools) or other public areas, we model these two types of transmission separately. The geographical arrangement of nodes represents the residence of each individual. So, by "local" transmission, we mean only transmission within a family unit (i.e. residents of a single flat), or between adjacent flats.[3] Hence, "non-local" transmission refers to transmission between non-family members due to the mixing of individuals in public spaces. In the context of the SARS outbreak in Hong Kong, this would include transmission within hospitals, schools and public spaces. In our model, we expect SSE to be represented through a single node with a large number of non-local connections.[4]



(a) State transition flow graph.



(b) Small-world network structure.

Fig. 2. The top panel depicts the transmission state diagram: S to P based on the SW structure and the infection probabilities $p_{1,2}$; P to I with probability r_0; and I to R with probability r_1. The lower panel depicts the distinction between "local" (i.e. short range) and "non-local" (long range) network links. The lower panel shows the arrangement of nodes in a small-world network. The black (infected) node may infect its four immediate neighbours with probability p_1 and three other nodes (hashed) with probability p_2.

We fix the population N and assume that there are no other additions to, or deletions from, the population from any other cause for the duration of each simulation. The population of N nodes is arranged in a regular grid, of side length L ($L^2 = N$), and each node is connected directly to n_1 immediate neighbours.[5] An infected individual will infect each of its n_1 neighbours (provided that they are still susceptible) with probability p_1. Furthermore, each node has n_2 non-local (i.e. long distance) links (see Fig. 2). These are links to nodes that are geographically remote from one another; infection occurs along these pathways with probability p_2. For each node i, the number $n_2^{(i)}$ is fixed, and so are the links to its $n_2^{(i)}$ remote neighbours once they are established. The number $n_2^{(i)}$ is chosen to be proportional to a decaying exponential $f_X(x) \propto \mathrm{e}^{-\frac{x}{\mu}}$ with parameter μ proportional to the expected (average) number of

[2] This may be due to the physical mechanism of transmission (via contact or "droplets") or due to the fact that so many transmissions occurred within hospitals. In either case, the implication for our model is the same.

[3] As the vast majority of residents in Hong Kong live in high-density, high-rise buildings, a three-dimensional grid may be more appropriate. However, for the present time, we restrict our interest to the two-dimensional case.

[4] As we will see in what follows, the number of local connections is constant for all nodes.

[5] The grid is assumed to be topologically equivalent to the surface of an annulus in three dimensions, that is, nodes (1, 1) and (L, L) are neighbours.

links to remote nodes, where $n_2^{(i)} = \lfloor x \rfloor$. Hence we have

$$P(n_2^{(i)} = m) = \frac{1}{C} e^{-\frac{m}{\mu}} \tag{1}$$

where

$$C = \sum_{m=0}^{\infty} e^{-\frac{m}{\mu}} = \frac{1}{1 - e^{-\frac{1}{\mu}}}.$$

It is the inclusion of non-local links with a random number of links that can give rise to the network's SW (and, in other cases not considered here, SF) structure. In this paper, we assign an exponentially decaying probability distribution to any number of links, and (for uni-directional links) this is sufficient to generate the necessary SW properties. An SF network requires a power-law distribution of the number of links (which can consequently lead to more nodes with many more links) and we do not examine that case here. The SF distribution of the induced network of actual infections has been considered elsewhere [16].

It is worth considering that, for the model that we present here, the links between nodes are *uni*-directional. That is, infection only spreads in one direction. Clearly, the true network of social interaction consists of *bi*-directional links. But, for the purposes of simulating disease transmission, unidirectional links appear to be a sufficient approximation. The consequence of this is that it becomes easier to generate the small-world (and elsewhere the scale-free) network.

Finally, for each simulation we seed the model with one randomly chosen initial infection. We expect that computational simulations of this network will show that infection spreads locally, just as SARS spread within particular geographical regions of Hong Kong. Moreover, the system can also exhibit non-local infection, as a single individual may infect individuals in distant communities. Occasionally, individuals will infect a large number of others, exactly as was observed at the start of the SARS epidemic in Hong Kong (an SSE). However, unlike the start of the SARS epidemic in Hong Kong, because we seed our population with a single infectious individual, we do not expect to see the initial SSE triggered by that individual (except by chance). Therefore, the initial growth of transmission in our model is exponential rather than an SSE. The only way to overcome this is to include explicitly the "seed" SSE in the model, regardless of network topology.

2.2. Behaviour

The epidemic will eventually be contained if the rate of infection is lower than the rate of removal. Intuitively, provided that $(n_1 p_1 + \mu p_2) \gg r_1$, one would expect the disease to become endemic; conversely, if $(n_1 p_1 + \mu p_2) \ll r_1$, the disease will be contained. In what follows, we study this condition more precisely.

Moreover, with this model we can analytically compute the probability of an outbreak being self-terminating. For a single infectious node, the probability of no further infections on *a given day* is given by

$$\begin{aligned} P_{\text{no1}} &= \sum_{m=0}^{\infty} \left[(1-p_1)^{n_1} (1-p_2)^m P(n_2^{(i)} = m) \right] \\ &= (1-p_1)^{n_1} (1 - e^{-\frac{1}{\mu}}) \sum_{m=0}^{\infty} \left[e^{-\frac{1}{\mu}} (1-p_2) \right]^m \\ &= \frac{(1-p_1)^{n_1} (e^{\frac{1}{\mu}} - 1)}{e^{\frac{1}{\mu}} - 1 + p_2}. \end{aligned} \tag{2}$$

Hence the probability of *no further infections* from this node can be closely approximated by the infinite geometric series using the average P_{no1} computed in Eq. (2):

$$\begin{aligned} P_{\text{none}} &\approx P_{\text{no1}} r_1 + P_{\text{no1}}^2 (1-r_1) r_1 + P_{\text{no1}}^3 (1-r_1)^2 r_1 + \cdots \\ &= \frac{P_{\text{no1}} r_1}{1 - P_{\text{no1}}(1-r_1)} \end{aligned} \tag{3}$$

provided that $|P_{\text{no1}}(1-r_1)| < 1$. Upon substitution of Eq. (2) into (3), we find that

$$P_{\text{none}} = \frac{r_1 (1-p_1)^{n_1}}{1 - (1-r_1)(1-p_1)^{n_1} + p_2 \Big/ \left[e^{\frac{1}{\mu}} - 1 \right]}. \tag{4}$$

Eq. (4) is the probability of no infections from a given individual, and is therefore a weak lower bound on the probability of no general outbreak.

Now, let us denote the probability of no further infections occurring, given that there are k infectious nodes by

$$P^{(k)} = P_{\text{none}}^{(k)} = \text{Prob(no further infection} \mid k \text{ infectious nodes)}$$

where, for notational convenience, we will drop the subscript on P_{none}. Treating infections as discrete events (i.e. they occur one at a time), we have that $(1 - P^k)$ is the probability of at least one further infection from k infectious nodes. The probability that the epidemic will terminate is given by

$$\begin{aligned} P_{\text{safe}} &= P + (1-P) \Big[P^2 + (1-P^2) \\ &\quad \times \left[P^3 + (1-P^3) \left[P^4 + (1-P^4) [\ldots] \right] \right] \Big] \\ &= \sum_{m=0}^{\infty} P^{m+1} \prod_{n=1}^{m} (1-P^n) \end{aligned} \tag{5}$$

where $P = P_{\text{none}}$ is given by Eq. (4). By expanding Eq. (5) and comparing to the Pentagonal Number Theorem,[6] we find that Eq. (5) can be rewritten as an infinite sum

$$\begin{aligned} \sum_{m=0}^{\infty} P^{m+1} \prod_{n=1}^{m} (1-P^n) &= P + P^2 - P^5 - P^7 \\ &\quad + P^{12} + P^{15} - P^{22} + \cdots \\ &= \sum_{m=1}^{\infty} (-1)^{m+1} \left[P^{\frac{1}{2}m(3m-1)} + P^{\frac{1}{2}m(3m+1)} \right] \end{aligned} \tag{6}$$

[6] This result was originally proved by Euler in 1775.

where the sequence of indices 0, 1, 2, 5, 7, 12, 15, 22, 26, 35 ... is the generalised pentagonal numbers (described as sequence A001318 in [18]). Eq. (6) may also be re-written in terms of the Dedekind eta function, but for the purposes of this discussion it is unnecessary to do so. Nonetheless, for $0 \le P < 1$ this sequence converges fairly rapidly as the order of the exponent increases.

The exact[7] probability of a general outbreak can alternatively be obtained by using a branching process method. Following [19], we define

$$g_1(x) = (1 - p_1 + p_1 x)^{n_1} \sum_{m=0}^{\infty} \frac{1}{C} \mathrm{e}^{-\frac{m}{\mu}} (1 - p_2 + p_2 x)^m$$
$$= \frac{(1 - p_1 + p_1 x)^{n_1} (1 - \mathrm{e}^{-\frac{1}{\mu}})\mu}{1 - \mathrm{e}^{-\frac{1}{\mu}}(1 - p_2 + p_2 x)}, \tag{7}$$

the probability generating function for the number of secondary cases produced by a single infectious case in a day. Then the probability generating function for the overall number of secondary infections from a single primary case is

$$g(x) = \sum_{j=1}^{\infty} g_1(x)^j r_1 (1 - r_1)^{j-1}$$
$$= \frac{r_1 g_1(x)}{1 - (1 - r_1) g_1(x)}. \tag{8}$$

One can then obtain the probability of no general outbreak (i.e. the probability of the disease not becoming endemic) as the smallest solution $x \in [0, 1]$ of $g(x) = x$.

Unfortunately, Eq. (8) cannot be readily used for further analysis. Similarly, although Eq. (5) can be computed easily, it is not in a form that is immediately amenable for further analysis. However, since $P_{\text{safe}} \ge P_{\text{none}}$, it is clear that $\mu\left[1 - (1 - p_2)\mathrm{e}^{-\frac{1}{\mu}}\right] \gg 1$ will make $P_{\text{safe}} \approx 0$. Hence, either $\mu \gg 1$ or $p_2 \approx 1$ will lead to widespread infection (as expected). Differentiating (5) with respect to $(1 - p_1)^{n_1}$, we can easily verify that P_{safe} is a monotonic function of both p_1 and n_1. One can therefore observe that $P_{\text{safe}} \approx 0$ if $p_1 \approx 1$ or $n_1 \gg 1$.

The most severe limitation on Eq. (5), and also Eqs. (7) and (8), is that we assume that no infected nodes have common neighbours, and that all of the neighbours are susceptible. In reality, the number of potential infections is limited by the fact that some of the potential neighbours are already infected. It is therefore important to estimate the number of neighbours of an infected node that have not been infected. This is equivalent to estimating the ratio of local and non-local infections in an epidemic.[8] One can consider the network of infected individuals as consisting of a number of "clumps": one clump for each non-local infection (i.e. each clump is seeded by a non-local transmission; all other transmissions *within* that clump are local). Provided $p_2 > 0$, this implies that, as the clump gets bigger, the probability of any given infection being a long range infection will increase. Conversely, as the number of clumps increases, the probability of local infection (relative to non-local infection) will increase.[9]

Let us now estimate the expected number of connections from an infected node. Let N_S denote the expected number of susceptible nodes linked to a random node. If this node is the result of a non-local infection, then we suppose that $N_S = n_1 + \mu$; however, if this is the result of a short range infection, then this number should be lower (certainly no more than $\mu + n_1 - 1$). Now,

$$\text{Prob(long range infection | infection)} = \frac{\mu p_2}{n_1 k p_1 + \mu p_2}$$
$$\text{Prob(short range infection | infection)} = \frac{n_1 k p_1}{n_1 k p_1 + \mu p_2}$$

where k is the proportion of local links that support possible infection and $0 < k \le \frac{n_1 - 1}{n_1}$. Hence,

$$N_S = (n_1 + \mu)\frac{\mu p_2}{n_1 k p_1 + \mu p_2} + (n_1 k + \mu)\frac{n_1 k p_1}{n_1 k p_1 + \mu p_2}$$
$$= n_1 \frac{n_1 k^2 p_1 + \mu p_2}{n_1 k p_1 + \mu p_2} + \mu. \tag{10}$$

From the preceding geometric argument, if infection grows in a single clump, then $k \approx \frac{1}{2}$. Moreover, $k < \frac{1}{2}$ only if nodes remain infected when they are on the interior of such "clump" (i.e. when r_1 is very low). We would therefore expect that $\frac{1}{2} \le k \le \frac{n_1 - 1}{n_1}$. Note that k is not a model parameter, but rather it is a term in the model that will both depend on the various model parameters and vary with time.

Finally, we now consider the rate of transmission. Let $P(t)$, $I(t)$ and $R(t)$ be the number of prone, infected and removed individuals at time t (in days). The probabilities r_0 and r_1 can therefore be considered as the rates at which prone nodes become infectious and infectious nodes become removed (respectively). Similarly, $(n_1 p_1 k + \mu p_2) S(t) I(t)$ is the expected number of new infections. Suppose that $S(t) \gg R(t) + I(t) +$

[7] Because of the assumption that infections occur individually and sequentially, the branching process in Eq. (5) is only an approximation to the solution of Eq. (8).

[8] We can achieve this as follows. Suppose that there are no non-local infections (i.e. $p_2 = 0$) and that infections grow in a single (roughly spherical) "clump". Then, if the clump consists of $I(t)$ individuals, then the radius of this clump will be $\sqrt{\frac{I(t)}{\pi}}$ and the number of susceptible individuals is $2\sqrt{I(t)\pi}$. Now, further suppose that all nodes in the clump are infectious (i.e. $r_1 = 0$), then the mean number of links per infected individual is $2\sqrt{\frac{\pi}{I(t)}}$. Even with $r_1 > 0$, as the clump grows, there are, on average, fewer potential infection paths.

[9] Moreover, one can estimate the number of clumps K. Observe that $K \approx \frac{\mu p_2}{n_1 p_1 + \mu p_2} \times$ (number of infections). More precisely,

$$K = \frac{\mu p_2}{(N_s - \mu) p_1 + \mu p_2} \times I$$
$$= \frac{\mu p_2 (n_1 k p_1 + \mu p_2)}{n_1 p_2 (n_1 k^2 p_1 + \mu p_2) + \mu p_2 (n_1 k p_1 + \mu p_2)} \times I \tag{9}$$

where I is the total number of infections.

$P(t)\,\forall t$. Then

$$\begin{aligned} R(t+1) &= R(t) + r_1 I(t) \\ I(t+1) &= I(t) - r_1 I(t) + r_0 P(t) \\ P(t+1) &= P(t) - r_0 P(t) + n_k I(t) \end{aligned} \tag{11}$$

where $n_k = (n_1 p_1 k + \mu p_2)$ is the expected number of links for each infectious node. We are now modelling the inter-day process assuming discrete day-to-day dynamics. The reason for this approximation is that the available time series data (which will be the basis of our comparison) is similarly course-grained. Assuming that the population is seeded with a single infectious individual, the solution of Eq. (11) is given by

$$\begin{aligned} \begin{bmatrix} R(t) \\ I(t) \\ P(t) \end{bmatrix} &= \begin{bmatrix} 1 & r_1 & 0 \\ 0 & (1-r_1) & r_0 \\ 0 & n_k & (1-r_0) \end{bmatrix}^t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\ &= P D^t P^{-1} [0\ 1\ 0]^{\mathrm{T}} \end{aligned} \tag{12}$$

where $P = [v_1\ v_2\ v_3]$ is the matrix of eigenvectors and

$$D^t = \begin{bmatrix} \lambda_1^t & 0 & 0 \\ 0 & \lambda_2^t & 0 \\ 0 & 0 & \lambda_3^t \end{bmatrix}$$

is formed from the corresponding eigenvalues, given by

$$\lambda_1 = 1$$
$$\lambda_{2,3} = 1 - \frac{r_0 + r_1}{2} \pm \sqrt{\frac{1}{4}(r_0 - r_1)^2 + n_k r_0}.$$

It then follows that the system has a marginally stable focus (i.e. the epidemic will terminate) if $|\lambda_{2,3}| < 1$, i.e.

$$n_k < r_1 \tag{13}$$
$$n_k r_0 < (2 - r_0)(2 - r_1). \tag{14}$$

The second condition (14) is only violated if $n_k > 1$, which would also violate condition (13). Therefore, the epidemic is controllable provided that $n_k = n_1 p_1 k + \mu p_2 < r_1$. The right hand side of this inequality is the rate of infection and the left hand side is the rate of removal, as expected. In fact, this result is exactly analogous to the equivalent result for the continuous SIR model [7]. Moreover,

$$\max_{i=1,2,3} |\lambda_i| = 1 - \frac{r_0 + r_1}{2} - \sqrt{\frac{1}{4}(r_0 - r_1)^2 + n_k r_0}. \tag{15}$$

Computationally, we can see that, as r_0 or n_k increases, then the rate of growth of the epidemic also increases. Conversely, as r_1 increases, the rate of growth decreases. This is as one would expect, as increasing r_1 will decrease the number of infectious individuals while increasing either r_0 or n_k increases this quantity.

3. Computation

In the following subsections, we confirm the preceding relationships and numerically explore the behaviour of our models under a variety of conditions.

3.1. Model parameters

As stated, our model has seven explicit parameters: L, n_1, μ, p_1, p_2, r_1, and r_2. For the population of Hong Kong, we set $L = 2700$ ($N = L^2 = 7{,}290{,}000$). We arbitrarily choose $n_1 = 4$ and set $E(n_2) = \mu = 7$ [16]. One further parameter is the time interval between successive steps in the discrete time simulations from our model. We choose the natural scale of one day. However, it is not obvious that this is the best choice. Given that the data is discretised to this interval, it is perhaps the best choice in this situation. Nonetheless, to confirm that this choice is appropriate, we have repeated our analysis with both longer and shorter time steps. The results are equivalent.

The average incubation period between infection and an individual becoming symptomatic is 6.4 days [6]. The number of days in the prone state can therefore be modelled as the result of a series of independent Bernoulli trials with a mean $\frac{1}{r_0}$ and so it follows a geometric distribution $f_X(x) = (1-p)^{x-1}p$.[10] For a general disease model, it would possibly be more appropriate to have a separate state for the pre-symptomatic but infectious period. Although contact tracing of all SARS patients in Hong Kong has demonstrated that this is not a significant period for SARS, we have repeated the simulations with this state. If the hiatus in this state is relatively short, the results are not altered significantly.

In a similar spirit, the time before hospitalisation is 3–5 days, and we model this as a series of independent Bernoulli trials. In an effort to establish the degree to which hospital transmission in our model matches what was observed, we first assume that hospitalisation is equivalent to isolation of infectious individuals. That is, we suppose that infectious individuals are only infectious for 3–5 days. The current weight of evidence suggests that hospital transmission was a crucial factor for the SARS outbreak in Hong Kong during 2003. We will show that this is also the case for our model.

Hence, we suppose for now that the average amount of time prior to isolation is 4 days (we will consider larger values later). In our model, the number of days prior to hospitalisation is also made to follow a geometric distribution with mean $\frac{1}{r_1}$.

Hence, the only free parameters are μ, p_1 and p_2. Without active control, we also know that the average number of new infections per case (excluding SSEs) is 2.7 [4]. In this state, each infectious individual will infect, on average, $n_1 p_1 + E(n_2^{(i)}) p_2$ new individuals every day. Since $E(n_2^{(i)}) = \mu$ and we suppose that the time before hospitalisation is d days, we have

$$n_1 p_1 + \mu p_2 \approx \frac{2.7}{d}. \tag{16}$$

We set $d = 4$ and therefore $p_1 \approx \frac{1}{n_1}(0.675 - \mu p_2) = 0.135 - \frac{\mu}{n_1} p_2$. In subsequent simulations, we also consider $d = 3, 4, 6$ and also a larger number of new infections than 2.7.

[10] We do it this way simply because it is convenient to do so. For simulation purposes, this sequence of Bernoulli trials is both easy to implement and (more importantly) requires very modest computational storage resources.

In each case, the results are equivalent to those presented here. However, we maintain the values described above in the following discussion. Any deviation of results between these and other values will be highlighted in the text.

Our initial model parameters are therefore [16]:

$$L = 2700,$$
$$r_0 = \frac{1}{7.4} \approx 0.14,$$
$$r_1 = \frac{1}{4} = 0.25,$$
$$n_1 = 4,$$
$$\mu = 7, \text{ and}$$
$$p_1 = 0.135 - \frac{7}{4} p_2.$$

Note that, because we have the possibility of P to I transition after zero days, $r_0 = \frac{1}{7.4}$ rather than $\frac{1}{6.4}$. This does not have a significant effect on our results; it is merely a computational convenience. Now, from Eq. (15) we can deduce that the rate of growth of infection is approximately

$$1 - 0.5 \times \left((0.135 + r_1) - \sqrt{(0.135 - r_1)^2 + 0.675 \times (3p_1 + 7p_2)} \right)$$

which, for $r_1 = 0.25$, yields either a growth rate significantly less than exhibited in the data or rates of infection significantly greater. Even for reasonable variation of d and the average number of secondary infections, we obtain similar results. Hence, we conclude that the assumption of no nosocomial transmission is inconsistent with the observed data. Increasing the average infectious time to 6 days gives a substantially higher rate of infection: consistent with the observed data.

3.2. Feasible region

Subject to Eq. (5), we explore which parameter values give a significant probability of the epidemic becoming endemic. From Eq. (13), we have that $P_{\text{safe}} = 1$ if $n_k < r_1$. Consistent with the discussion of the preceding section, we start with the assumption that $k = \frac{n_1 - 1}{n_1}$ and set $n_1 = 4$ and $\mu = 7$. Fig. 3 is a plot of the probability of complete infection for various parameter values estimated from the data and Eq. (5). We see that there is close agreement between the theoretical and experimental results. Moreover, we note that a smaller value of k is appropriate for scenarios with a relatively high proportion of local spreading (i.e. p_1 larger and $p_2 \approx 0$). This is consistent with the case of spreading within a single clump, and we approach the situation of $k = \frac{1}{2}$. However, in these simulations the best choice is $k > \frac{1}{2}$ in all cases. Only for extremely small values of r_1 and p_2 would we expect smaller values of k. Typically, $k = \frac{n_1 - 1}{n_1} = \frac{3}{4}$ seems to be a good choice.

Furthermore, in Fig. 3 it is evident that the greatest probability of an epidemic becoming endemic is when there is both local and non-local infection. In the model that we have constructed here, local infection spreads approximately geometrically,[11] while non-local infection spreads exponentially. Yet, it is some combination of both that provided the greatest possibility of an outbreak spreading without control. Clearly, exponential growth will lead (all else being equal) to more rapid growth of an epidemic than geometric growth. But, by combining some geometric growth, the epidemic becomes even more dangerous. It seems that the additional local transmissions allow each non-local transmission the possibility of seeding a new infection cluster (rather than just a single point), and each new cluster is more difficult to eradicate than a single point infection.

Fig. 4 depicts level curves for the probability of fixed levels of infection. That is, we compute the parameter values $p_{1,2} \in [0, 1]$ and $r_1 \in [0.1, 0.5]$ required to achieve specific values of P_{safe}. We see that only for $p_{1,2} \ll 0.2$ is the outbreak likely to be controlled. Moreover, the variation with r_1 is not critical i.e. the range of behaviour for specific $p_{1,2}$ is not great.

Finally, in Fig. 5 we plot the probability of the epidemic being controlled for various values of μ with $p_{1,2} \in [0, 1]$, $r_0 = 0.1$ and $r_1 = 0.25$. Consistent with Fig. 4, results for different values of $r_{0,1}$ did not change significantly. Moreover, from Fig. 5 we see that only for $\mu > 2$ does the infection probability p_2 have a significant effect. That is, to limit long range infection, one should aim to reduce the average number of long range links below 2. We find that, for all values of $p_{1,2}$ and $\mu > 2$, there was remarkably little variation in the value of P_{safe}. However, the rate of infection (Eq. (15)) did change significantly.

3.3. Unconstrained growth

Subject to the choice of parameters in the previous sections, we now simulate the expected dynamics and compare this to the theoretical bounds of Section 2.2.

According to theory, restricting the infectious period to five or fewer days does not yield a growth rate large enough to be consistent with the observed data. We test that assertion numerically and find that the observed data (over 1000 individuals infected after 50 days) is inconsistent with the simulation for $r_1 = 0.25$. However, by lowering r_1 to 0.165 (and therefore increasing the mean infectious period to six days, we achieve results more consistent with the observed data. We see that only with $r_1 \geq 0.165$ do we obtain results for which the true data is not statistically atypical. Moreover, this result is robust to moderate changes of the other relevant parameters, as illustrated in Fig. 6. We conclude that the obtained results for which the true data is not atypical; we require $r_1 \geq 0.165$ (i.e. a mean exposure time of six or more days). Moreover, we see from Fig. 6 that widespread infection is associated with a large number of clusters. The results of the next section corroborate this.

3.4. Clustering

To examine the role of clustering more closely, Fig. 7 is a snapshot of a single simulation for parameter values $r_1 = 0.25$ and $p_2 = 0.006$. This simulation shows SSE resulting from clustering and highly connected nodes. Moreover, the gradual spread of the disease within a single cluster is evident; one can

[11] Proportional to the size of the clump, to be precise.

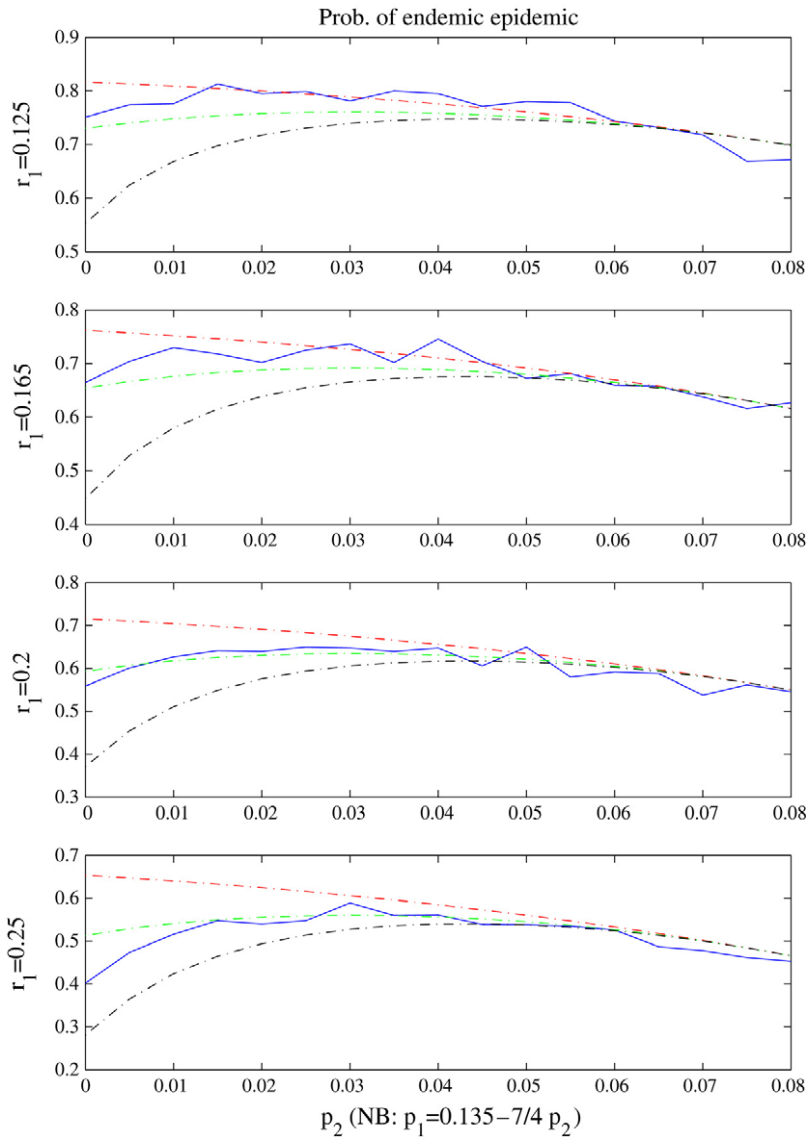


Fig. 3. Probability of complete infection. The probability of an endemic epidemic for various values of $p_{1,2}$ and r_1 with $r_0 = 0.1$, $n_1 = 4$ and $\mu = 7$. The probability is estimated from 1000 simulations of the model and according to Eq. (5) with $k = 1$ (upper dashed line), $k = \frac{3}{4}$ (middle dashed) and $k = \frac{1}{2}$. The convex shape of each of these plots is extremely interesting. This indicates that neither purely local nor purely non-local infection results in the greatest probability of overall infection, rather an endemic outbreak is most likely when both local and non-local infection is possible.

see that the number of clusters is far less than the number of infections. Time series of the infection total appear qualitatively similar to the Hong Kong SARS data. Specifically, burstiness typical of SSEs is evident. However, the quantitative behaviour of this model is remarkably different to the dynamics observed in the SARS outbreak. The *daily* number of reported infections in Fig. 1 far exceeds the total number of active infected and prone individuals in Fig. 7.

3.5. Simulations

Finally, we provide simulations of the Hong Kong epidemic and from multiple simulations estimate the likelihood of various outcomes based on the model. We initiate the model with a single infected individual and a relatively low removal rate r_1. Fig. 8 depicts our results.

We can see from Fig. 8 that many of the features of the true data are reproduced well in the simulations. However, two important aspects of the simulations are not sufficiently similar to the data. Firstly, the initial spreading of the disease is exponential, rather than the single SSE observed in the real data. Secondly, the magnitude of the SSEs in the simulations is somewhat smaller than the largest SSEs in the data.

This second aspect can be overcome by simply altering the distribution of non-local links. In [16] we describe how a power-law distribution of links can lead to many more extreme events. Conversely, the initial SSE in the data cannot be modelled well by our simulations, except by chance. Therefore, to achieve similar initial events, we would expect that we would have to execute many simulations (and choose only those that suit our purpose), or simply build the SSE into the model. Neither of these approaches are desirable. We prefer to focus on

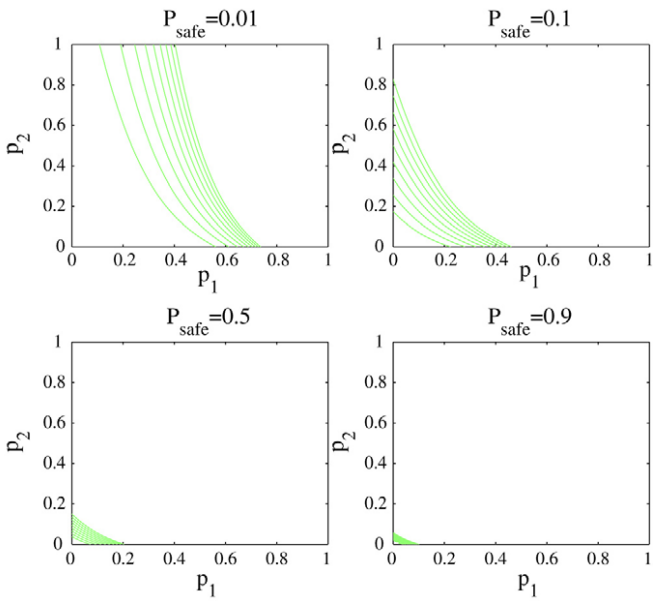


Fig. 4. Level curves for probability of complete infection. The probability of an endemic epidemic for various values of $p_{1,2}$ and r_1 with $r_0 = 0.1$, $n_1 = 4$ and $\mu = 7$. The probability is estimated according to Eq. (5) with $k = \frac{3}{4}$. The four panels correspond to level curves for the probability of the disease being self-contained of 0.01, 0.1, 0.5 and 0.9, respectively. The curves and the contours for $r = 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45$, and 0.5. The smooth dependence of P_{safe} on each of the parameters is evident.

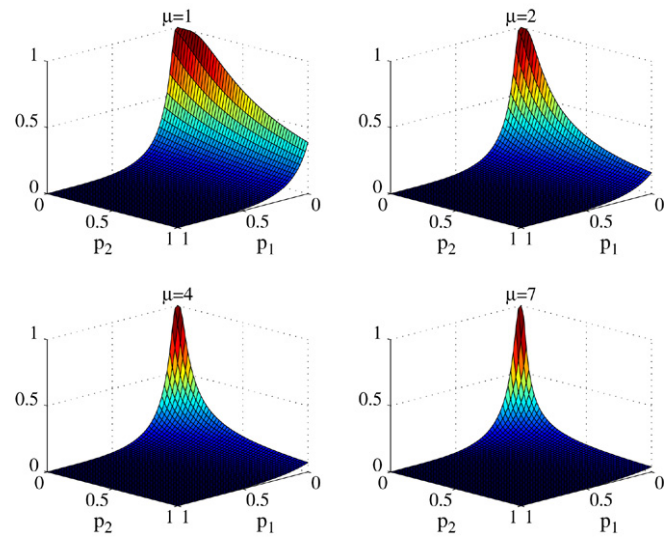


Fig. 5. Probability of control of infection. The probability of the epidemic terminating without intervention for various values of $p_{1,2}$ and μ with $r_0 = 0.1$, $r_1 = 0.25$ and $n_1 = 4$. The probability is estimated according to Eq. (5) with $k = \frac{3}{4}$. The four panels correspond to $\mu = 1, 2, 4$ and 7, respectively. Reducing μ greatly increases the probability of being able to control the epidemic, but increasing μ beyond moderate values (i.e. $\mu \geq 4$) does not change the behaviour substantially.

the possibility of SSEs occurring randomly, without explicitly adding them in the model. Hence, we do see SSEs, but not necessarily immediately after the start of the simulation (as we suppose occurred in the true data).

We now wish to demonstrated how the SW model described here improves on stochastic SIR-type models [7]. One very simple way to achieve this is to compare the distribution of statistic values measured from simulations of either type of model to the true data. In Fig. 9 we do this, and we find that the small-world model exhibits statistical properties much closer to the observed data. In Fig. 9 we compare the results for the SIR-type model (with stochastic inputs) to the SW model described in this paper. By generating multiple simulations of both models and estimating simple statistical measures from both sets of simulations, we see that the model dynamics of the SW model are much closer to the true data.

Assuming that this model and parameters are accurate, or at least appropriate, we compute the likely behaviour for various epidemics. We generated 1000 realisations of the model depicted in Fig. 8 and computed the total number of casualties.

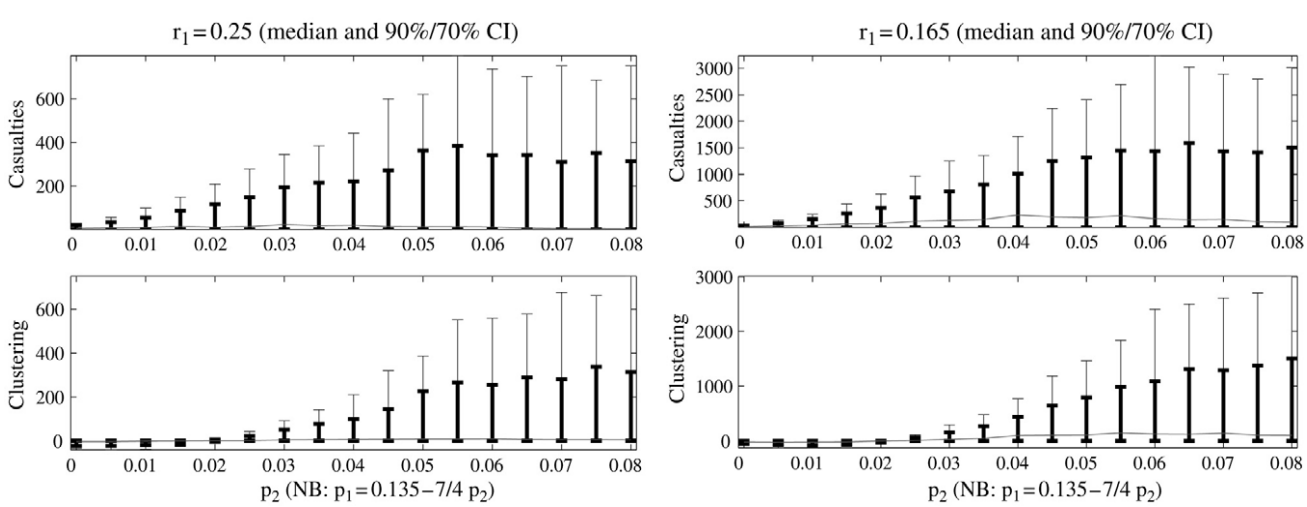


Fig. 6. Unconstrained growth of the infectious population. The upper panels show the number of individuals infected after 50 days; the lower plots show the number of distinct clusters detected after the same time. The parameter $p_1 = 0.135 - \frac{7}{4}p_2$, $r_0 = 0.1$, $n_1 = 4$ and $\mu = 7$. The left hand plots are for $r_1 = 0.25$ (i.e. no nosocomial transmission) and the right panels are for $r_1 = 0.165$ (a mean infection period of 6 days). The results are median, 70% and 90% confidence intervals from 1000 simulations.

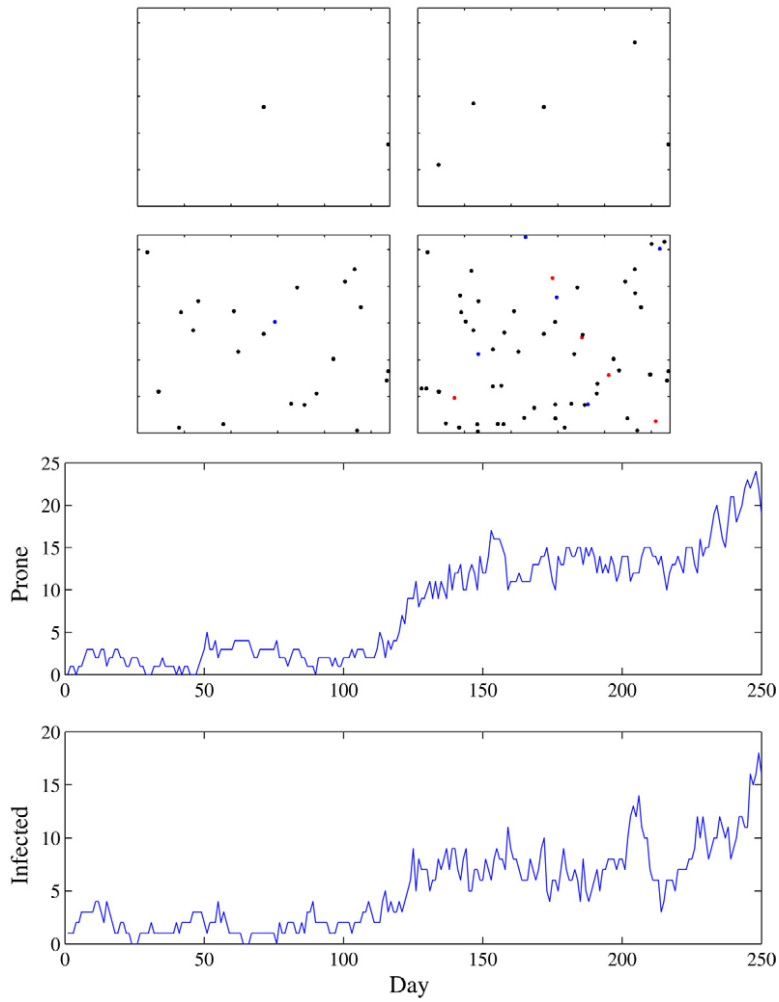


Fig. 7. Outbreak of clusters in a single simulation. The top four panels show the location of infected individuals after 50, 100, 150 and 250 days. Removed, infected and prone individuals are illustrated as solid dots (R are black, I are red, and P are blue). Parameter values used are $r_1 = 0.25$, $p_1 = 0.093$ and $p_2 = 0.06$, and the results for these parameter values are typical. The bottom two panels show the number of prone and infected individuals over the same period as time series. One can clearly see the irregular bursts of activities in the time series corresponding to the explosion in clusters. The total number of infections (R + I + P) is 22, 40, 117, and 309 after 50, 100, 150, and 200 days, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 10 is a plot of the probability distribution of the number of fatalities for these simulations, and Fig. 11 is the probability distribution for the daily number of infections. We found that the probability of infecting fewer than 20 people was approximately 0.18, while the probability of infecting more than 1000 was 0.27. One can see that, with respect to these gross statistics, the true situation for Hong Kong (1755 casualties) is quite typical.[12] It is interesting to note, however, that there is a large variation in the number of casualties. In all cases, the parameter values of Fig. 8 provided effective control of SARS transmission after approximately 150 days.

From these simulations, we therefore conclude that, with effective control measures in place, the likelihood of a significant outbreak is low.

[12] From 1000 simulations, 106 exhibited a larger number of casualties.

4. Conclusion

We summarise the main results of the preceding sections as follows.

- The probability of the disease spread being controlled, P_{safe}, is approximately

$$\sum_{m=0}^{\infty} P^{m+1} \prod_{n=1}^{m} (1 - P^m)$$

where

$$P = \frac{r_1(1-p_1)^{n_1}}{1 - (1-r_1)(1-p_1)^{n_1} + p_2 \Big/ \left[\mathrm{e}^{\frac{1}{\mu}} - 1\right]}$$

(Section 2.2). An equivalent value can also be obtained from the smallest zero of Eq. (8), subject to (7), in [0, 1].

- The epidemic is under control if $n_k = n_1 k p_1 + \mu p_2 < r_1$ (Section 2.2).

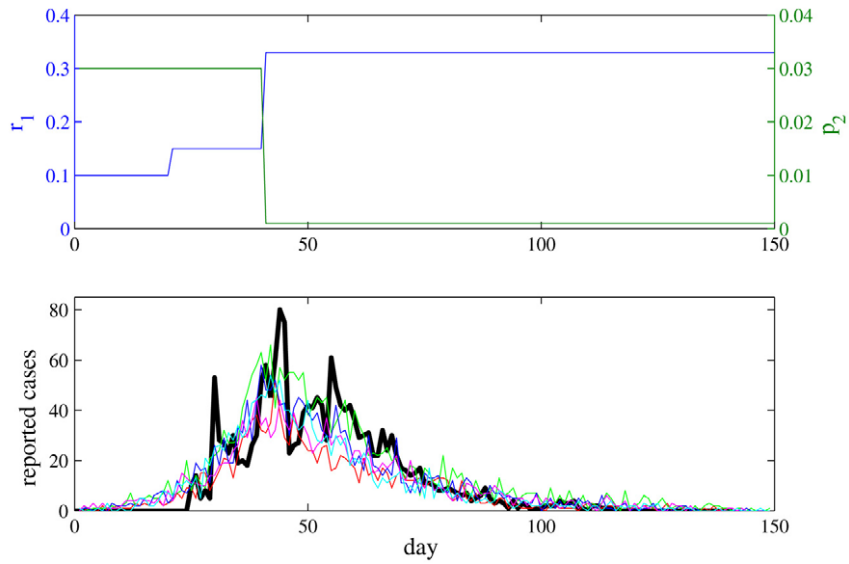
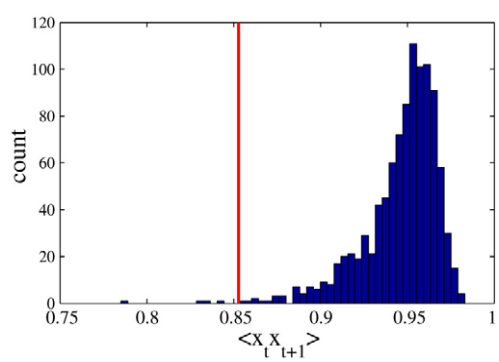
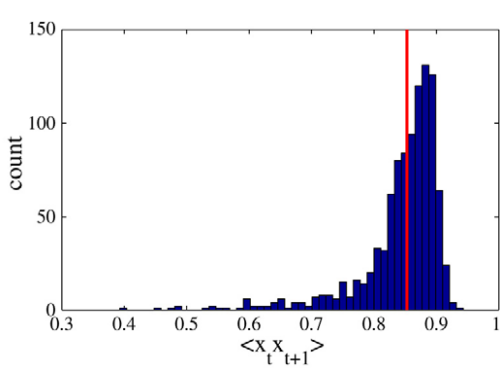


Fig. 8. Model simulations. The top panel shows the change in parameters r_1 and p_2 with time (all other parameters are constant: $p_1 = 0.08$, $n_1 = 4$ and $\mu = 7$). The bottom plot shows five model simulations and the true SARS data for Hong Kong. The five model simulations were selected to ensure that a "full" outbreak occurred (a total number of infections greater than 1000). The true data is plotted as a heavy solid line.



(a) x_t.



(b) x_t.

Fig. 9. Probability distribution on the lag one autocorrelation for 1000 simulations of the SW model, and the standard SIR model with stochastic perturbations. The heavy vertical line is the equivalent quantity for the true data. We see that the simulations from the SW model (b) are indistinguishable from the data; the SIR model (panel (a)) is not.

- For the endemic case, the rate of growth of infection is

$$1 - \frac{r_0 + r_1}{2} + \sqrt{\frac{1}{4}(r_0 - r_1)^2 + n_k r_0}.$$

- This model exhibits behaviour consistent with both SSEs and persistent localised transmission (Section 3.4). More extreme SSEs can be observed simply by fattening the tail of the distribution of links. By assigning a power-law distribution to the number of nonlocal links that a node has, one can readily obtain SSE involving many hundreds of secondary infections from a single source.
- An SSE does not imply highly infectious individuals, only highly connected ones (Section 3.4).
- Theoretical results and model simulations are unlike the true data, unless exposure time is significantly greater than an average of three days (Sections 2.2 and 3.3). If the exposure time is three days or less, the rate of growth of the epidemic is significantly lower than that observed in the true data.
- Nosocomial transmission was therefore a key factor in the acuteness of the SARS epidemic in Hong Kong in 2003. Effective control of hospital transmissions would have prevented a serious outbreak (Section 3.3). With respect to nosocomial transmission, our model therefore confirms what has been observed independently, and suggested by many authors.
- Simulations of our model, with minimal parameter variation, produced dynamics indistinguishable from the true data. These same parameter values exhibit a wide variety of behaviours (Section 3.5). Hence, any effort to actually obtain maximum likelihood parameter estimates from the observed data is futile. Despite this, our calculations show that the true data is certainly typical of our models. Moreover, our models exhibit a marked long-tailedness both in infection times and epidemic lifetimes.
- The likelihood of infecting fewer than 20 people was approximately 0.18, and the likelihood of infecting more than 1000 was 0.27 (Section 3.5).

Apart from deriving analytic expression for the spreading and control of an epidemic, the main results of this study, when

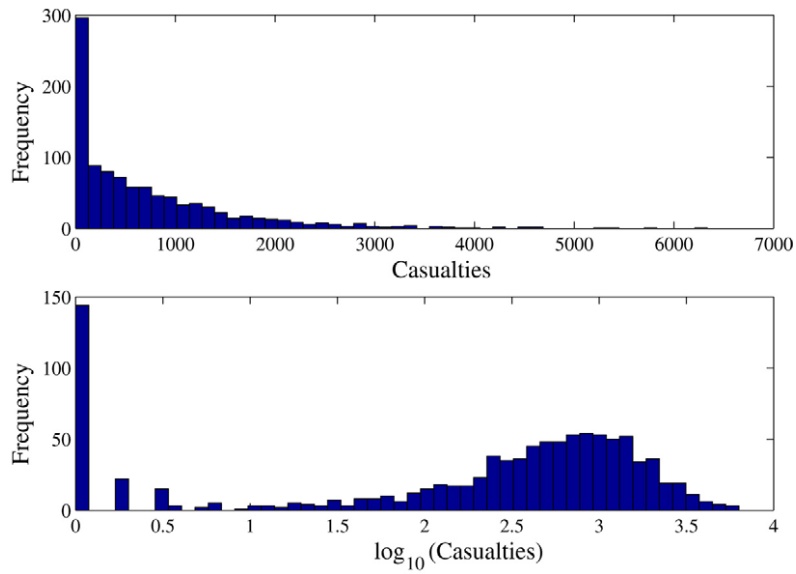


Fig. 10. Probability distribution on the total number of infections. The probability distribution on the total number of infections for 1000 simulations of the model in Fig. 8 is shown on a linear (top plot) and a logarithmic scale.

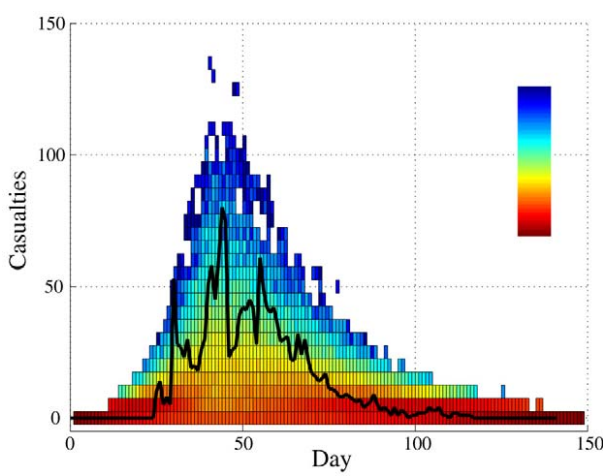


Fig. 11. Probability distribution infection dynamics. The probability distribution of the daily number of infections for 1000 simulations of the model in Fig. 8 are shown on a logarithmic scale. Blue represents low probability, while red represents high probability of a particular infection tally for any number of days after onset.(For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

applied to the SARS epidemic in Hong Kong in 2003, is that our data is consistent with the observation that the epidemic that occurred was largely preventable. The primary factor for the severity of the outbreak was poor infection control in the hospital setting and the delayed introduction of quarantine and community isolation practices. This is consistent with various discussions presented elsewhere. However, we note that nosocomial transmission was considered only by supposing (admittedly erroneously) that the time between becoming symptomatic (and infectious) and hospitalisation is constant. Nosocomial transmission is simulated by changing the expected duration between becoming symptomatic (I) and becoming removed (R). Although this is perhaps unconventional from an epidemiological viewpoint, it is probably the best that we can do, as the only data that we have to work from is the date of hospital admission. This conclusion is in contrast with work done by Pastor-Satorras and colleagues [20,21] with scale-free SIS type [7] disease models. In that case, they found that the disease would almost always persist and that random immunisation was ineffective [21].

The model presented here emphasised the small-world transmission dynamics created by providing a small number of highly connected nodes. We model the inter-connection between nodes with one-directional links (rather than the bi-directional links we would expect for social contacts) because this makes the network much easier to construct, without sacrificing any realism. This model could easily be extended to exhibit scale-free characteristics, by simply altering the distribution of non-local links (to follow a power-law distribution). Doing so would produce simulations with a larger number of highly connected individuals, and therefore we could simulate the largest SSEs in the data. The results presented in this paper show small (less than 100 individuals) SSEs; extension to the power-law distribution would provide SSEs with a larger number of secondary infections.

We should note that it is very difficult to reliably and accurately fit even a moderate number of parameters to a stochastic model from such limited data. Especially when the model, for the same parameter values, can exhibit a wide variety of behaviour. To overcome this, we (a) make our model so simple that the number of parameters is very few (certainly comparable in number to ordinary SIR), and (b) only "fit" the model with typical parameter values and test that, for these parameter values, the observed data is typical. Certainly, we cannot exclude the (quite likely) probability that other parameter values also produce behaviour of which the observed data is typical. Therefore, in this paper we only draw conclusions based on our observations of parameters that produce behaviours that are either typical or (more importantly)

atypical of the observed data. Although we have observed that our model exhibits more realistic dynamics than SIR models, it is only a model. And, like the SIR model, our model is a compromise between complexity and realism. We intend to address this problem in the future by applying Monte Carlo Markov Chain models [22] to this sparse data [23].

Even our choice of a discrete time model is open to debate. The choice, with a time interval chosen as one day, was motivated by the data and what was more intuitive to us. We have conducted calculations with both longer and shorter inter-interval time scales and find that the results are invariant under a suitable change in time scale (i.e. we get equivalent results by using a shorter time step).

Moreover, recent epidemiological case studies have found that, overall, in Hong Kong 8% of individuals sharing a flat with a SARS patient contracted SARS [24]. Given an average infectious period of four days (prior to hospitalisation), this implies that the daily probability of transmission is 0.02. This is approximately consistent with the values of p_1 used in this study, and therefore provides further support of our results.

Conversely, analysis of the SSE at the Prince of Wales Hospital found that, for a group of medical students, the probability of direct contact with the index patient leading to SARS infection was $\frac{10}{27}$ [5]. As this study dealt with a single meeting, this implies a daily infection probability of approximately 0.371: significantly higher than p_2 in our simulations. Hence, although we have concluded that SSEs may occur without individuals *necessarily* being highly infectious, epidemiological evidence suggests that, in some cases, this may still be the case.

Finally, we note that, although the methods presented here are applied only to the SARS outbreak in Hong Kong in 2003, these methods are not limited to this situation. Apart from repeating, or modifying, this analysis for different infectious diseases (such as HIV AIDS), we imagine that these techniques could be useful in the theoretical study of quarantine and isolation practices, as well as disease transmission among isolated communities. Quarantine can be effectively modelled as a limitation of long range transmission. One can easily model change in quarantine by varying the parameter μ (or alternatively, but not equivalently, p_2). This should provide a new, simpler approach to the study of quarantine to supplement compartmental models such as those described in [25].

Acknowledgments

This work is supported by funding from the Health, Welfare and Food Bureau of the government of the Hong Kong Special Administrative Region under the Research Fund for Control of Infectious Diseases (RFCID). DW thanks the Scottish Executive Environment and Rural Affairs Department (SEERAD) for support.

References

[1] C.-Y. Moira, R.-H. Xu, SARS: epidemiology, Respirology 8 (2003) S9–S14.
[2] World Health Organisation, Consensus document on the epidemiology of SARS, Technical Report, World Health Organisation, 17 October 2003. http://www.who.int/csr/sars/en/WHOconsensus.pdf.
[3] Severe Acute Respiratory Syndrome Expert Committee, Report of the severe acute respiratory syndrome expert committee. Technical Report, Hong Kong Department of Health, 2 October 2003. http://www.sars-expertcom.gov.hk/english/reports/report.html.
[4] S. Riley et al., Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions, Science 300 (2003) 1961–1966.
[5] T.-W. Wong et al., Cluster of SARS among medical students exposed to single patient, Hong Kong, Emerg. Infect. Dis. 10 (2004).
[6] C.A. Donnelly et al., Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong, Lancet 361 (7) (2003) 1761–1766. http://image.thelancet.com/extras/03art4453web.pdf.
[7] J.D. Murray, Mathematical Biology, 2nd edition, in: Biomathematics Texts, vol. 19, Springer, 1993.
[8] J. Koopman, Modeling infection transmission, Annu. Rev. Public Health 25 (2004) 303–326.
[9] N.M. Ferguson, G.P. Garnet, More realistic models of sexually transmitted disease transmission dynamics: Sexual partnership networks, pair models and moment closure, Sex. Transm. Dis. 27 (2000) 600–609.
[10] C.T. Bauch, A versatile ODE approximation to a network model for the spread of sexually transmitted diseases, Math. Biol. 45 (2002) 375–395.
[11] M.S. Handcock, J.H. Jones, Likelihood-based inference for stochastic models of sexual network formation, Theor. Popul. Biol. 65 (2004) 413–422.
[12] F. Ball, P. Neal, Poisson approximation for epidemics with two levels of mixing, Ann. Probab. 32 (2004) 1168–1200.
[13] L. Hufnagel, D. Brockmann, T. Geisel, Forecast and control of epidemics in a globalized world, Proc. Nat. Acad. Sci. USA 101 (2004) 15124–15129.
[14] D.J. Watts, Six Degrees: The Science of a Connected Age, W.W. Norton & Company, 2003.
[15] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440–442.
[16] M. Small, C.K. Tse, Small world and scale free model for transmission of SARS, Int. J. Bifurcation Chaos 15 (2005) 1745–1755.
[17] A.P. Galvani, R.M. May, Dimensions of superspreading, Nature 438 (2005) 293–295.
[18] N.J.A. Sloane, On-line encyclopedia of integer sequences, Online, 2005. http://www.research.att.com/~njas/sequences/Seis.html.
[19] O. Diekmann, J.A.P. Heesterbeek, Mathematical Epidemiology of Infectious Diseases, John Wiley & Son, Chichester, 2000.
[20] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Phys. Rev. Lett. 86 (2001) 3200–3203.
[21] R. Pastor-Satorras, A. Vespignani, Epidemic dynamics and endemic states in complex networks, Phys. Rev. E 63 (2001) 066117.
[22] P.D. O'Neill, G.O. Roberts, Bayesian inference for partially observed stochastic epidemics, J. Roy. Statist. Soc. A 162 (Part 1) (1999) 121–129.
[23] D.M. Walker, F.J. Pérez-Barbería, G. Marion, Stochastic modelling of ecological processes using hybrid Gibbs samplers, Ecol. Model. (2005) (submitted for publication).
[24] J.T.F. Lau et al., Probable secondary infections in households of SARS patients in Hong Kong, Emerg. Infect. Dis. 10 (2004).
[25] L. Sattenspeil, D.A. Herring, Simulating the effect of quarantine on the spread of the 1918–19 flu in central Canada, Bull. Math. Biol. 65 (2003).