THE ROYAL SOCIETY

The effects of local spatial structure on epidemiological invasions

M. J. Keeling

Zoology Department, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK (matt@zoo.cam.ac.uk)

Predicting the likely success of invasions is vitally important in ecology and especially epidemiology. Whether an organism can successfully invade and persist in the short-term is highly dependent on the spatial correlations that develop in the early stages of invasion. By modelling the correlations between individuals, we are able to understand the role of spatial heterogeneity in invasion dynamics without the need for large-scale computer simulations. Here, a natural methodology is developed for modelling the behaviour of individuals in a fixed network. This formulation is applied to the spread of a disease through a structured network to determine invasion thresholds and some statistical properties of a single epidemic.

Keywords: correlation equations; contact networks; spatial heterogeneity; basic reproductive ratio; final size; invasion

1. INTRODUCTION

Invasion is one of the most fundamental concepts in ecology and epidemiology. It is only with a firm quantitative understanding of this ubiquitous phenomenon that we can accurately determine vaccination thresholds and evolutionary selection as well as more common forms of invasion (Kornberg & Williamson 1987). Invading organisms are initial highly aggregated, with only limited spatial spread, and therefore suffer from far more intraspecific competition than non-spatial models would predict. As infectious diseases provide the best documented and most accurately modelled problems in ecology, this paper shall concentrate on analytical results for the invasion of an infection into a spatially distributed host population.

Spatial models, from meta-population models to partial differential equations (PDEs), have become increasingly popular in both ecology and epidemiology. It has become obvious that in many situations spatial patterns and correlations play a vital role; this is especially true for invasions. Using the correlations between individuals to capture the essential spatial characteristics is not new to ecology (Hassell & May 1974), although only recently have these correlations been treated as dynamic variables (Dickman 1986; Matsuda 1987; Matsuda *et al*. 1992; Sato *et al*. 1994; Levin & Durrett 1996; Keeling & Rand 1999).

Correlation models, and in particular pair-wise models, have been primarily used to describe the behaviour of simple spatial models (such as probabilistic cellular automata) in terms of a set of ordinary differential equations (ODEs). However, these correlation models can be used in their own right (Dietz & Hadeler 1988; Altmann 1995; Keeling *et al*. 1997) and can provide a more general framework and neighbourhood structure than is feasible in traditional spatial models. Correlation models are of most use when the interactions between individuals (or sites) can be considered as occurring on a network—this is the case for communicable diseases.

When considering the spread of an epidemic, it is the contact structure between individuals that determines the progress of the disease through the population (Barbour & Mollison 1990). One of the simplest models which captures the fundamental features of infection dynamics is the SIR (susceptible–infectious–recovered) model (Anderson & May 1992; Mollison 1995; Grenfell & Dobson 1995).

The first section examines the structure of a network and the behaviour of various quantities when the network is specified by a graph or contact matrix. In § 3, a method for closing the system of ODEs at the level of pairs is formulated, (i.e. we express the number of triples in terms of the number of pairs). Section 4 combines the theoretical arguments to derive the equations for an SIR disease spreading across a network. Sections 5 and 6 examine two fundamental properties of the SIR model the basic reproductive ratio R_0 and the final size of the epidemic. The final section considers how these two properties are changed by vaccination, and predicts vaccination thresholds.

2. GENERAL THEORY

One of the simplest assumptions for disease transmission is that the contact structure forms a network of links between individuals (or nodes), with all links being of equal strength. Such a network is often referred to as a graph. We can describe a network involving N individuals by a matrix $G \in \{0,1\}^{N^2}$,

$$G_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connected} \\ 0 & \text{otherwise} \end{cases}$$

Received 8 December 1998 *Accepted* 25 January 1999

As all links are bidirectional and self-contact is not allowed, this places two constraints upon the matrix: $G = G^T$ and $G_{ii} = 0$. From this matrix, we can calculate the number of connected pairs and triples in the graph,

$$\begin{aligned}\text{number of pairs} &= \|G\| = n\mathcal{N},\\ \text{number of triples} &= \|G^2\| - \text{trace}(G^2).\end{aligned}$$

Here, $\|G\| = \sum_{i,j} G_{ij}$ is the sum of all the elements in the matrix and n is therefore the average number of neighbours per node. The number of triples is calculated as the number of nodes which are joined by two connections, given that the nodes are distinct.

It should be noted that there is ambiguity in the precise form of a triple; for three connected nodes, it is possible to form triangular loops as well as linear arrangements (Keeling *et al*. 1997; Morris 1997; Rand 1999; Van Baalen 1999). These loops are very important in the spread of a disease, as we will show later. Let ϕ be defined as the ratio of triangles to triples, this is a simple measure of how interconnected the local neighbourhoods are. When ϕ is large, the members of a connected pair will be connected to many common nodes, your neighbour's neighbours are also your neighbours (figure 1*a*), whereas when ϕ is small there are few common nodes, and long-range connections dominate (figure 1*b*). As triangles are three linked nodes with the same start and end point, ϕ can be expressed in terms of the graph

$$\phi = \frac{\text{number of triangles}}{\text{number of triples}} = \frac{\text{trace}(G^3)}{\|G^2\| - \text{trace}(G^2)}. \qquad (1)$$

For most communicable diseases, we are likely to find that each individual experiences only a small proportion of the population ($n \ll \mathcal{N}$) and yet ϕ is generally large as there tends to be complete interaction within small social groups (cf. Watts & Strogatz 1998). We believe that in general, much of the underlying structure of the network can be characterized in terms of the average number of neighbours (n) and the interconnectedness (ϕ). Throughout this paper, we will assume that both n and ϕ are fixed for all sites, i.e. the network is homogeneous.

To begin to consider the dynamics of individuals, it is necessary to define a set of functions which inform us about the state of each node. Let A_i be equal to one if the individual at node i is of type A or zero otherwise. This allows us to define rigorously the number of single, pairs and triples of each type,

$$\begin{aligned}\text{singles of type A} = [A] &= \sum_i A_i,\\ \text{pairs of type A–B} = [AB] &= \sum_{i,j} A_i B_j G_{ij},\\ \text{triples of type A–B–C} &= [ABC] = \sum_{i,j,k} A_i B_j C_k G_{ij} G_{jk}.\end{aligned} \qquad (2)$$

This method of counting means that pairs are counted once in each direction so that $[AB] = [BA]$ and that $[AA]$ is even. From equations (2), we can recover the natural rules for summing singles, pairs and triples,

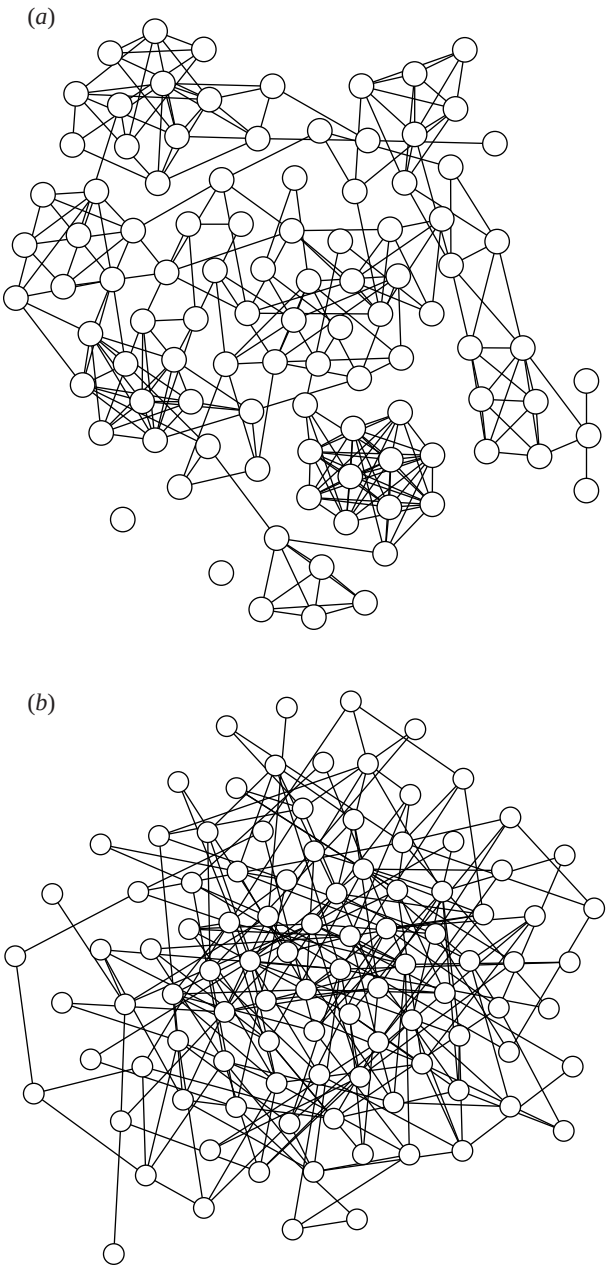

Figure 1. Examples of a network of 100 nodes, with an average of five connections per node ($n = 5$). In (*a*) $\phi = 0.7$ and triangles are common, whereas in (*b*) $\phi = 0.2$ and there is less obvious structure. These graphs were obtained by placing nodes randomly in two dimensions and weighting the probability of a connection between nodes by the distance.

$$\sum_A [A] = \sum_A \sum_i A_i = \sum_i \left(\sum_A A_i \right) = \mathcal{N}, \qquad (3)$$

$$\sum_B [AB] = \sum_B \sum_{i \neq j} A_i B_j G_{ij} = \sum_{i,j} A_i G_{ij} = n[A], \qquad (4)$$

$$\sum_C [ABC] = \sum_C \sum_{j, i \neq k} A_i B_j C_k G_{ij} G_{jk} = \frac{n(n-1)}{\mathcal{N}} [A][B]. \qquad (5)$$

Note that, in general, there is no corresponding formula for $\sum_B [ABC]$.

As well as the number of pairs, it is often useful to consider the multiplicative correlation between connected nodes of various types. We shall define $\mathcal{C}_{AB}$ to be the correlation between nodes of type A and B,

$$\mathcal{C}_{AB} = \mathcal{N}^2 \frac{\sum_{i\neq j} A_i B_j G_{ij}}{\sum_{i\neq j} A_i B_j \sum_{i\neq j} G_{ij}} = \frac{\mathcal{N}}{n}\frac{[AB]}{[A][B]}. \quad (6)$$

From this, it can be seen that $\mathcal{C}_{AB} \in [0,\mathcal{N}]$. When $\mathcal{C}_{AB} = 1$, then A and B are uncorrelated so their placement with respect to each other is random.

3. FORMULATING A MODEL BY CLOSING THE SYSTEM

Let f be any function of the states of nodes in the network. If f can be assumed to be continuous in the limit of large populations, then its behaviour can be captured by the differential equation (Morris 1997; Rand 1999)

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \sum_{\text{events}} \text{rate of event} \times \text{change in } f \text{ due to event} \quad (7)$$

Normally, f is considered to be either the number of nodes of a given type ($[A]$) or the number of pairs of a given type ($[AB]$), allowing the formulation of equations for the behaviour of the system.

In the vast majority of ecological and epidemiological systems, any change in the behaviour or state of an individual will be dependent on the state of its neighbours. For example, a susceptible individual surrounded by infectious neighbours is likely to become infected. Let $Q_i(B|A)$ be the number of B neighbours surrounding node i, given that node i is state A. Q_i is considered to be comprised two parts: the expected value $\overline{Q}$, together with some associated error σ which is not necessarily small.

$$Q_i(B|A) = \overline{Q}(B|A) + \sigma_i(B|A) = \frac{[AB]}{[A]} + \sigma_i(B|A).$$

One of the main features of individual-based models with a small neighbourhood, is that individuals often experience large fluctuations in their environment, so σ may be large compared to $\overline{Q}$ (Keeling & Rand 1999). By assuming a distribution for the Q_i terms and, hence, a distribution for the errors σ_i, one can produce expressions for the expected number of triples and higher-order connections in terms of pairs (Rand 1999). For a fixed number of neighbours per site, the most likely form for the errors is multinomial, in which case

$$[ABC] = \sum_{j,B_j=1} Q_j(A|B)Q_j(C|B), \quad (8)$$

$$= \sum_{j,B_j=1} \left(\frac{[AB]}{[B]} + \sigma_j(A|B)\right)\left(\frac{[BC]}{[B]} + \sigma_j(C|B)\right),$$

$$= \frac{[AB][BC]}{[B]} + \sum_{j,B_j=1} \sigma_j(A|B)\sigma_j(C|B),$$

$$= \left(\frac{n-1}{n}\right)\frac{[AB][BC]}{[B]}. \quad (9)$$

A similar technique can be applied if other nonlinear arrangements of the Q_i terms arise or if other distributions for σ are plausible. In the SIR example given below, only linear combinations of pairs and triples will appear, and therefore equation (9) will be all that is needed. In the remainder of the paper, for convenience of notation, we shall set $\zeta = (n-1)/n$.

In the above calculation of the number of triples, no consideration has been given to the structure of the network; in particular the number of triangular and higher order loops has been ignored. If A–B–C form a triangle, we must also allow for the fact that the correlation between A and C is also important. The most natural way to include this correlation is

$$[ABC] \approx \zeta\frac{[AB][BC]}{[B]}\mathcal{C}_{AC} = \frac{\zeta\mathcal{N}}{n}\frac{[AB][BC][AC]}{[A][B][C]}.$$

Hence, bringing in the full network structure by including the ratio between triangles and triples,

$$[ABC] \approx \zeta\frac{[AB][BC]}{[B]}\left((1-\phi) + \phi\frac{\mathcal{N}}{n}\frac{[AC]}{[A][C]}\right). \quad (10)$$

It can be hoped that in choosing ϕ for any particular applications, it may capture the effects due to all loops of three nodes and higher. It should be expected that the effect from four node loops is far less than that from three node loops etc., and so the ϕ which describes the system best should be slightly larger than the exact ratio of triangles to triples.

4. THE SIR MODEL

With these tools, we are now in a position to consider the spread of a disease through a network of nodes; in particular, a network with n neighbours per site and with an interconnectedness of ϕ. Throughout this work, the mean field limit ($n \to \mathcal{N} \to \infty$, $\phi \to 1$) will be calculated as a check and a comparison.

The most basic epidemic model with a recovered status is the simple epidemic, described by the SIR model without the demographic processes of birth and death (Kermack & McKendrick 1927; Anderson & May 1992; Mollison 1995). Such a model exhibits a single epidemic. Each individual can be in one of three states: susceptible to the disease, infectious when they can spread the disease to susceptibles, and recovered when they have been infectious but can no longer spread or catch the disease (cf. the rapid spread of influenza). This framework forms the basis of almost all epidemiological models (Grenfell *et al*. 1992; Grenfell & Dobson 1995; Keeling & Grenfell 1997). Denoting the number of susceptible, infectious and recovered individuals by S, I and R respectively, the mean field equations are

$$\begin{aligned}\dot{S} &= -\beta\frac{S}{\mathcal{N}}I,\\ \dot{I} &= \beta\frac{S}{\mathcal{N}}I - gI,\\ \dot{R} &= gI.\end{aligned} \tag{11}$$

Here, β is the contact parameter, and $1/g$ is the infectious period. This model has been analysed and applied to numerous situations (Anderson & May 1992; Mollison 1995 and references therein), the main results are as follows.

(i) An epidemic can only occur if $R_0 = \beta/g > 1$.
(ii) S is monotonically decreasing, R is monotonically increasing and I is unimodal.
(iii) The epidemic eventually dies out with some proportion of susceptibles, S_∞, remaining

$$S_\infty = \exp\left((S_\infty - 1)R_0\right).$$

For the correlation model, the equations describe the behaviour of A–B pairs instead of the behaviour of individuals. As we have moved from a global to a more individual approach, we will define τ the transmission rate across a connection to be β/n; hence, the potential for spreading infection is equal between the two approaches.

There exist nine distinct types of pairs; however, due to symmetries ($[AB] = [BA]$) and the fact that the sum over all pairs remains constant, only five differential equations are necessary.

$$\begin{aligned}[\dot{SS}] &= -2\tau[SSI],\\ [\dot{SI}] &= \tau([SSI] - [ISI] - [SI]) - g[SI],\\ [\dot{SR}] &= -\tau[RSI] + g[SI],\\ [\dot{II}] &= 2\tau([ISI] + [SI]) - 2g[II],\\ [\dot{IR}] &= \tau[RSI] + g([II] - [IR]),\end{aligned} \tag{12}$$

Using equation (10), the system can be closed at the level of pairs, assuming a multinomial distribution of neighbours. Figure 2 shows a comparison between the results of the correlation equations (12) and the average of astochastic simulation modelling the spread of a disease across a network. There is good quantitative agreement between the equations and the full model. It is the characteristic shape of the S–I correlation (figure 2*b*) with its rapid initial decline that leads to the major difference between the correlation and mean field equations.

When $\phi = 0$, the ODEs can be uncoupled in a natural way by expressing the system in terms of $\overline{Q}(S|S) = [SS]/[S]$, $\overline{Q}(I|S) = [SI]/[S]$ and the number of new cases $C = \tau[SI]$,

$$\begin{aligned}\dot{C} &= C\left(\tau\frac{n-1}{n}\overline{Q}(S|S) - \frac{n-1}{n}\tau\overline{Q}(I|S) - \tau - g\right),\\ \overline{Q}(\dot{S}|S)) &= -\tau\frac{n-2}{n}\overline{Q}(S|S)\overline{Q}(I|S),\\ \overline{Q}(\dot{I}|S) &= \tau\frac{n-1}{n}\overline{Q}(S|S)\overline{Q}(I|S) - (\tau+g)\overline{Q}(I|S) + \frac{\tau}{n}\overline{Q}(I|S)^2.\end{aligned} \tag{13}$$

Therefore, when there are no triangular connections, the pairwise ODEs can be reduced to three dimensions, only one more than the mean-field model. The extra dimension accounts for the correlation between susceptibles and infectious individuals.

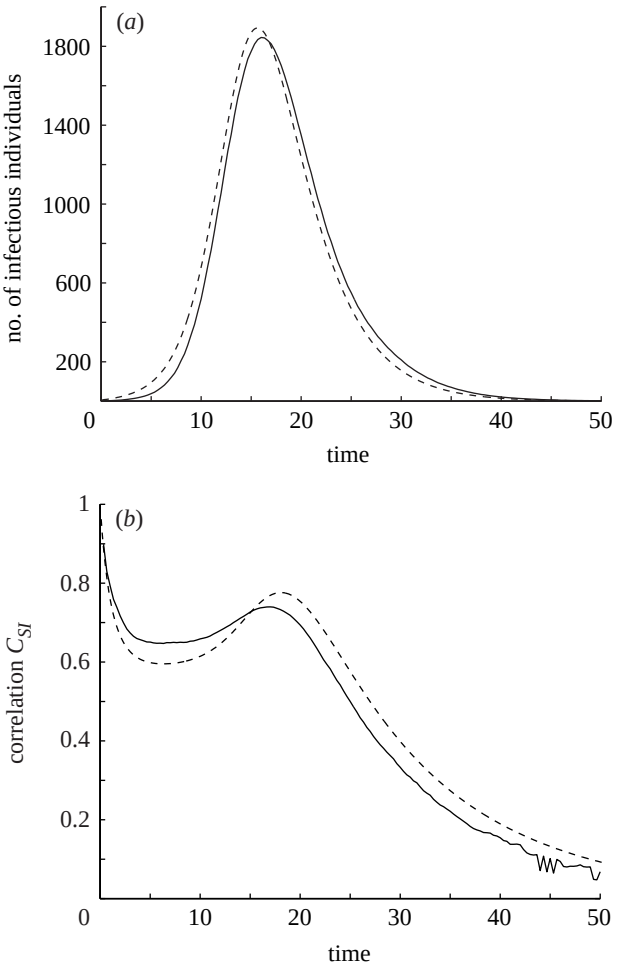

Figure 2. Results from the correlation equations (12) (solid line) and from the average of 100 stochastic simulations of an epidemic on a network (dashed line). The network contained $\mathcal{N} = 6000$ nodes, with each individual having $n = 6$ neighbours and an interconnectedness $\phi = 0.2$. (*a*) shows the number of infectious individuals over time, (*b*) shows the correlation between infectious and susceptibles $\mathcal{C}_{SI}$. The discrepancies between the two results may in part be due to the imprecise synchronization of epidemics across the many realizations (because of the stochastic nature of the system), leading to a smoothing-out of the simulation results.

5. THE BASIC REPRODUCTIVE RATIO, R_0

The basic reproductive ratio, R_0, is the most fundamental quantity in epidemiology (Diekmann *et al*. 1990; Anderson & May 1992; de Jong *et al*. 1994). R_0 is defined as the average number of secondary cases produced by an infectious individual in a totally susceptible population. It informs us whether a disease can ever invade a population, and is useful in the calculation of many other quantities. For the simple epidemic,

$$R_0 = \frac{\beta}{g}.$$

Let us consider the initial phase of an infection invading a total susceptible population.

$$\dot{[I]} = \tau[SI] - g[I],$$
$$= \left(\beta\frac{[S]}{\mathcal{N}}\mathcal{C}_{SI} - g\right)[I],$$

Because $[S]$ is assumed to be equal to $\mathcal{N}$ initially, this gives $R_0 = \mathcal{C}_{SI}\beta/g$. As a single infectious individual is placed in a sea of susceptibles, initially we find that susceptibles and infectious nodes must be uncorrelated ($\mathcal{C}_{SI} = 1$). It would therefore appear that R_0 is the same for the mean-field and the network model. However, this approach is flawed; consider the early behaviour of $\mathcal{C}_{SI}$

$$\dot{\mathcal{C}_{SI}} = \frac{\mathcal{N}}{n}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{[SI]}{[S][I]}\right) \rightarrow$$
$$-\tau\left(\mathcal{C}_{SI} + \mathcal{C}_{SI}^2 - n\zeta(\mathcal{C}_{SI} - \mathcal{C}_{SI}^2)(1-\phi) + n\zeta\mathcal{C}_{SI}^2\phi\frac{[I]\mathcal{C}_{II}}{\mathcal{N}}\right) \tag{14}$$

as $[S]/\mathcal{N} \rightarrow 1$, $C_{SS} \rightarrow 1$ and $[I]/\mathcal{N} \rightarrow 0$. The initial growth in the proportion of infectious nodes is small, however equation (14) shows that the correlation between S and I decays at order one. This means that in the region of the network that has been invaded, there is rapid development of the spatial structure as captured by the local correlations. In the early development of the epidemic $\mathcal{C}_{SI}$ converges to a quasi-equilibrium $\mathcal{C}_{SI}^*$. It is therefore advisable to measure R_0 once the local spatial pattern has formed and $\mathcal{C}_{SI}$ has equilibrated. As $\mathcal{C}_{SI}^*$ will be less than one, the value of R_0 will be similarly reduced.

From equation (14), it is clear that, in general, the quasi-equilibrium value also depends on the value of $\mathcal{C}_{II}[I]/\mathcal{N}$. The correlations between infectious nodes grow fast enough that $[I]\mathcal{C}_{II}/\mathcal{N}$ (which is interpreted as the probability that a neighbour of an infectious individual is also infectious) is of order one, even when the density of infectious individuals is small.

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{[I]\mathcal{C}_{II}}{\mathcal{N}} = \frac{1}{n}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{[II]}{[I]}\right) \Rightarrow \left(\frac{I\mathcal{C}_{II}}{\mathcal{N}}\right) \rightarrow \frac{2\tau\mathcal{C}_{SI}}{g + \beta\mathcal{C}_{SI} - 2\zeta\beta\mathcal{C}_{SI}^2\phi} \tag{15}$$

Therefore, from equations (14) and (15), it is found that $\mathcal{C}_{SI}^*$ satisfies

$$n\zeta(1-\mathcal{C}_{SI}^*)(1-\phi) - \frac{2\zeta\beta\phi\mathcal{C}_{SI}^{*\,2}}{g + \beta\mathcal{C}_{SI}^* - 2\zeta\beta\mathcal{C}_{SI}^{*\,2}\phi} - \mathcal{C}_{SI}^* = 1 \tag{16}$$

The values of $\mathcal{C}_{SI}$ and hence R_0 can be seen to depend only of the values of n, ϕ and β/g. Figure 3 shows the behaviour of $\mathcal{C}_{SI}$ (and hence R_0) and the associated values of $[I]\mathcal{C}_{II}/\mathcal{N}$ for a range of n and ϕ.

As R_0 is proportional to the S–I correlation, it is clear that the network models have a lower basic reproductive ratio than their mean-field counterparts. Therefore, there will exist conditions when the mean-field equations will predict that an epidemic will take off, but the network model will show that this is not the case. We can examine this result analytically in the limiting case when there are no triangular loops, $\phi = 0$.

$$n\zeta(1-\mathcal{C}_{SI}^*) - \mathcal{C}_{SI}^* = 1 \Rightarrow \mathcal{C}_{SI}^* = 1 - \frac{2}{n}$$
$$\Rightarrow R_0 = \left(1 - \frac{2}{n}\right)\frac{\beta}{g}$$

Similarly, by considering the limiting case when β is large, we find that the disease always dies out if

$$\phi > \frac{n-2}{n-1}.$$

This limiting value of ϕ corresponds to a highly interconnected system, where all but one of the neighbours is within a completely interconnected group.

It should be clear that whether a disease can invade a totally susceptible population (whether $R_0 > 1$) is dependent upon the network structure as well as the individual disease parameters. We find that the potential to invade is reduced by having few neighbours or a highly interconnected network structure, with the effects of ϕ being reduced as the number of neighbours increases.

6. THE FINAL SIZE OF A SINGLE EPIDEMIC

One of the main characteristics of an SIR epidemic in a population without births is the final size of the epidemic. This is the total number of individuals that become infected during the course of the epidemic and is equal to the final number of recovered individuals $R_\infty = 1 - S_\infty$. The final size of simple epidemics has been much studied in deterministic and stochastic models (Kermack & McKendrick 1927; Ball & Nasell 1994; Islam *et al*. 1996). It should be noted that when the birth rate is low compared to the epidemic time (e.g. influenza), this formulation still gives an accurate prediction of the number infected.

Given a fixed number of neighbours n per site and $\phi = 0$, analytical results have been developed (Diekmann *et al*. 1998) to show how, for a network of connected nodes, R_∞ is smaller than predicted by mean-field assumptions. For the network model

$$R_\infty = 1 - \left(1 - \frac{\tau}{g+\tau} + \theta\frac{\tau}{g+\tau}\right)^n, \tag{17}$$

where

$$\theta = \left(1 - \frac{\tau}{g+\tau} + \theta\frac{\tau}{g+\tau}\right)^{n-1}.$$

When $n \rightarrow \infty$, and keeping $n\tau = \beta$, this returns to the standard final size result of Kermack & McKendrick (1927) calculated as the long-term limit of the SIR equations. Unfortunately, the long-term limit of equations (12) cannot be solved as easily. However, as shown in Appendix A, for the case where $\phi = 0$, using the reduced form (13) the final size can be obtained and agrees with the value calculated by Diekmann *et al*. (1998). There is no obvious means of extracting the final size equations when $\phi \neq 0$, instead numerical results from integration of the network equations will be used to demonstrate the behaviour (figure 4).

The correlation model predicts a lower final size than the mean-field equations, and as ϕ is increased the final

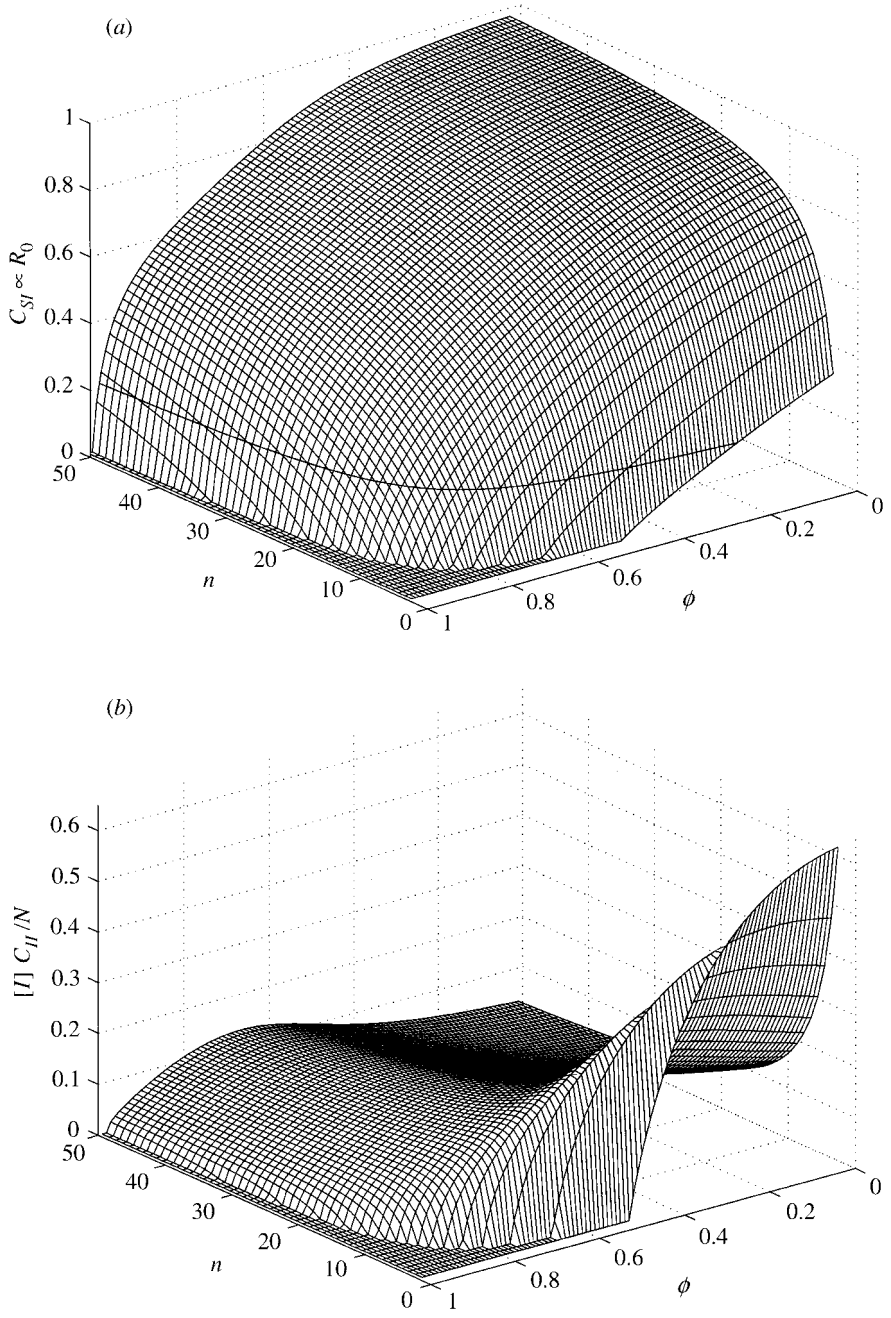

Figure 3. From equations (14) and (15), we can calculate the correlation between susceptibles and infectious nodes in the initial stages of invasion (*a*) and the probability that a neighbour of an infectious node is also infectious (*b*). For both these models, we have taken $g = 0.1$ and $\beta = 0.5$. The contour on graph *a* represents the line where $R_0 = 1$, above this line epidemics can start; for the comparative mean-field model, $R_0 = 5$.

size is further reduced. In all cases, it is true that $R_\infty \to 1$ as $\beta \to \infty$. Figure 4 also illustrates the effect ϕ has on the invasion threshold ($R_0 = 1$) which can be seen to occur at increasing transmissibilities for increasing ϕ. Therefore, not only does the proportion of triangular contacts (ϕ) reduce the initial spread of an epidemic, it also limits the final proportion of the population that the epidemic reaches.

7. VACCINATION

One of the main aims of epidemiology is to understand the role of vaccination and hence predict the level of vaccination necessary to eradicate a disease. If we assume that vaccination confers lifelong immunity to the disease, then we can amalgamate the vaccination and recovered individuals into a single non-susceptible class. For the simple SIR model without births or deaths, the effects of vaccination can be captured by starting the population with a fraction $V = 1 - S$ in the non-susceptible class, so equations (12) and (13) still hold.

When starting with only a proportion S of the population being susceptible, the invasion is characterized by the effective reproductive ratio R. For the mean-field models, there is a simple relationship between R and R_0,

$$R = SR_0 = \frac{\beta S}{g}.$$

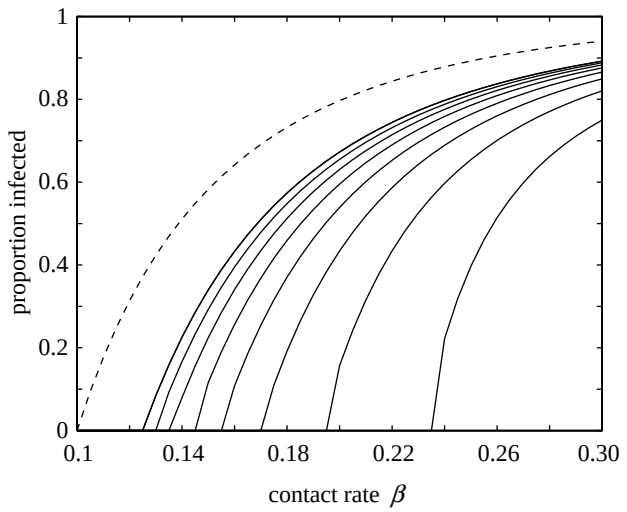

Figure 4. Comparison of theoretical and numerical results for $n = 10$ and $g = 0.1$. The dashed line is the theoretical proportion infected in the mean-field model. The black line is the theoretical result (and numerical solution of the ODEs (12)) when $\phi = 0$. The grey lines are for increasing ϕ (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8) and show how on a more interconnected graph we should expect smaller epidemics. The points where the curves meet the horizontal axis correspond to the invasion threshold when $R_0 = 1$. Again, $g = 0.1$.

Hence, using this approximation, the disease can only invade if the proportion susceptible is greater than some threshold value $S_T = g/\beta$. When considering the invasion on a network, we must take into account the initial spatial distribution of the susceptibles as well as their overall density. In general, we should expect susceptibles to be aggregated so that $\mathcal{C}_{SS} > 1$. As shown in § 5, local spatial structure develops over a short time-scale; the effect this has on R can be captured by examining the correlation $\mathcal{C}_{SI}$.

$$\mathcal{C}_{SI} \to \frac{(n-1)[S]\mathcal{C}_{SS}(1-\phi) - \mathcal{N}}{n[S] + (n-1)[I]\mathcal{C}_{II}\phi - (n-1)[S]\mathcal{C}_{SS}\phi},$$
$$[I]\mathcal{C}_{II} \to \frac{2\tau\mathcal{N}[S]\mathcal{C}_{SI}}{g\mathcal{N} + n\tau[S]\mathcal{C}_{SI} - 2(n-1))\tau[S]\mathcal{C}_{SI}^2\phi}.$$

When, $\phi = 0$, so that there is no spatial structure present in the network, we find

$$\mathcal{C}_{SI} \to \frac{(n-1)[S]\mathcal{C}_{SS} - [\mathcal{N}]}{n[S]},$$
$$\Rightarrow R = \frac{\tau}{g}\left[(n-1)\frac{[S]}{\mathcal{N}}\mathcal{C}_{SS} - 1\right]. \qquad (18)$$

Hence, it is both the proportion of susceptibles, $[S]/\mathcal{N}$, and the correlation between susceptibles, $\mathcal{C}_{SS}$ which controls the initial spread of the disease. The disease can only invade if

$$\frac{[S]}{\mathcal{N}} > \frac{1}{\mathcal{C}_{SS}}\frac{g+\tau}{\tau(n-1)},$$
$$\Rightarrow \frac{[S]}{\mathcal{N}} > \frac{1}{(n-1)\mathcal{C}_{SS}}\left[\frac{1}{R_0}(n-2)+1\right] > \frac{S_T}{\mathcal{C}_{SS}}.$$

In general, irrespective of the contact rate β, a disease cannot invade if

$$\frac{[S]}{\mathcal{N}} < \frac{1}{(n-1)(1-\phi)\mathcal{C}_{SS}},$$

which gives us a vaccination threshold that takes into account the network structure as well as the aggregation of the susceptibles. The aggregation of susceptible individuals is a monotonic function of the aggregation of vaccinated individuals. We therefore find that the vaccination threshold necessary for disease eradication increases with the number of neighbours and the aggregation of vaccination, but decreases with increasing ϕ.

As detailed in § 6, it is also possible to formulate equations for the final size of an epidemic when $\phi = 0$. Using the same approach, starting with a proportion S_0 of susceptibles with correlation $\mathcal{C}_{SS}$, we find

$$S_\infty = S_0\left(1 - \frac{\tau}{\tau+g}\mathcal{C}_{SS}S_0 + \frac{\tau}{\tau+g}\mathcal{C}_{SS}S_0^{1/n}S_\infty^{1-1/n}\right)^n. \qquad (19)$$
$$\text{Final proportion infected} = S_0 - S_\infty.$$

From which we find that both larger initial densities and larger initial aggregations will lead to a greater proportion of the population being infected.

8. DISCUSSION

Using the simple epidemiological SIR model, we have shown that the spatial structure which develops during the early stages of invasion can determine its success or failure. The limited spread of invading organisms means that they suffer far greater intraspecific competition than the homogeneously mixed, mean-field equations would suggest. This greater intraspecific competition leads to reduced success at invasion. This paper has demonstrated, primarily by numerical integration, that intraspecific competition, and therefore departure from the mean-field model, is greatest when the neighbourhood is small and there are many local connections (ϕ is large).

We have defined the reproductive ratio (R_0) as being the average number of individuals produced by an invading organism once the local spatial structure has reached equilibrium. Thus, this is a measure of whether an organism can enter the population and persist in the short-term; the standard mean-field result informs us whether the organism can simply enter the population. Therefore, for some parameter values, an organism may be able to enter a population, but cannot survive in the environment that rapidly develops around it. However, such situations are likely to be indistinguishable from chance events. It should be realized that invasion is a highly stochastic phenomenon and many invasions will fail simply by 'bad luck'. We believe that, in general, by the time the local spatial structure has developed, there should be sufficiently many cases that stochastic extinctions can be ignored. Thus, invasion is best characterized by a short stochastic entry phase followed by a more deterministic growth phase governed by the value of R_0 calculated above.

For communicable diseases, assuming the infection is spread via a network of contacts is the natural way to model the effects of spatial correlations. We have shown that departure from the standard mean-field results is greatest when n is small and ϕ is large. This is often the case for sexually transmitted diseases where most of the

partners come from a small social group. For other human diseases, such as childhood epidemics, there are many more contacts, but they are still highly interconnected. Therefore we should expect that for many common communicable diseases the fact that a network consists of a limited number of connected pairs will play an important role in the dynamics. If we allow the network to be dynamic, such that connections can break and reform (cf. Dietz & Hadeler 1988), then the behaviour becomes more like the mean-field model, with the correlations decaying to one as the movement of individuals becomes large.

For these correlation models to prove useful in practical epidemiological problems requires greater parameterization at the level of an individual, therefore what is needed is a method of extracting local parameters from global results. However, this method forms a valuable link between mathematically simplistic mean-field equations and computationally intensive individual-based models, and provides insights into the role of individuals and spatial correlations in ecological invasions.

This research was supported by the Wellcome Trust and the Royal Society. I wish to thank David Rand for his many helpful comments and insights as well as Minus van Baalen and Odo Diekmann.

APPENDIX A

(a) *Long-term limiting behaviour of the correlation model*

Equation (13) cannot be solved for the final size of the epidemic so easily as the standard SIR model. It is necessary to consider the behaviour of some new combinations of parameters; let us define

$$\zeta=\frac{n-1}{n},\quad P=\frac{[SI]}{[S]^{\zeta}},\quad R=\frac{[SS]}{[S]^{\zeta}},\quad \mathcal{T}=\frac{[SS]}{\exp(n[S]^{1/n})S^{2\zeta}}.$$

Now, from equation (12)

$$\begin{aligned}\dot{P}&=\tau\zeta\frac{[SS]}{[S]}P-(\tau+g)P,\\ \dot{R}&=-2\tau\zeta\frac{[SS]}{[S]}P+[SS]\frac{\zeta}{[S]^{\zeta+1}}\tau[SI],\\ &=-\tau\zeta\frac{[SS]}{[S]}P,\\ \dot{\mathcal{T}}&=\tau\mathcal{T}P,\end{aligned}\qquad\text{(A1)}$$

and from equation (13),

$$\overline{Q}(\dot{S}|S)=\frac{\mathrm{d}}{\mathrm{d}t}\frac{[SS]}{[S]}=-\tau\frac{n-2}{n}\overline{Q}(S|S)\overline{Q}(I|S),$$

and

$$\frac{\mathrm{d}[S]}{\mathrm{d}t}=-\tau[S]\overline{Q}(I|S)\Rightarrow[SS]=n[S]^{2\zeta}.$$

In particular, $[SS]_{\infty}=n[S]_{\infty}^{2\zeta}$.

Integrating the three equations (A1) from $t=0$ to ∞,

$$P(\infty)-P(0)=0=\tau\zeta\int_0^{\infty}\frac{[SS]}{[S]}P\mathrm{d}t-(\tau+g)\int_0^{\infty}P\mathrm{d}t,\quad\text{(A2)}$$

$$R(\infty)-R(0)=\frac{[SS]_{\infty}}{[S]_{\infty}^{\zeta}}-n=-\tau\zeta\int_0^{\infty}\frac{[SS]}{[S]}P\mathrm{d}t,\qquad\text{(A3)}$$

$$\ln(\mathcal{T}(\infty))-\ln(\mathcal{T}(0))=n-n[S]_{\infty}^{1/n}=\tau\int_0^{\infty}P\mathrm{d}t.\qquad\text{(A4)}$$

Therefore, from equation (A2) and substituting the expressions from equations (A3) and (A4),

$$-\frac{(\tau+g)}{\tau}(n-n[S]_{\infty}^{1/n})=\frac{[SS]_{\infty}}{[S]_{\infty}^{\zeta}}-n.$$

So setting $S_{\infty}=[S]_{\infty}/\mathcal{N}$, the equation for the final size of an epidemic can be recovered.

$$\begin{aligned}-(\tau+g)(1-S_{\infty}^{1/n})&=\tau(S_{\infty}^{\zeta}-1),\\ S_{\infty}^{1/n}&=1+\frac{\tau}{\tau+g}(S_{\infty}^{\zeta}-1),\\ S_{\infty}&=\left(1-\frac{\tau}{\tau+g}+S_{\infty}^{\zeta}\frac{\tau}{\tau+g}\right)^{n}.\end{aligned}$$

This is the same solution as equation (17) because $\theta=S_{\infty}^{\zeta}$. It is doubtful whether a similar technique could be applied to the situation when $\phi\neq0$.

REFERENCES

Altmann, M. 1995 Susceptible–infectious–recovered epidemic models with dynamic partnerships *J. Math. Biol.* **33**, 661–675.

Anderson, R. M. & May, R. M. 1992 *Infectious diseases of humans.* Oxford University Press.

Ball, F. & Nasell, I. 1994 The shape of the size distribution of an epidemic in a finite population. *Math. Biosci.* **123**, 167–181.

Barbour, A. & Mollison, D. 1990 Epidemics and random graphs. In *Stochastic processes in epidemic theory* (ed. J. P. Gabriel, C. Lefevre & P. Picard), pp. 86–89. New York: Springer.

de Jong, M. C. M., Diekmann, O. & Heesterbeek, J. A. P. 1994 The computation of R_0 for discrete-time epidemic models with dynamic heterogeneity. *Math. Biosci.* **119**, 97–114.

Dickman, R. 1986 Kinetic phase transitions in a surface reaction model: mean field theory. *Phys. Rev.* A **34**, 4246–4250.

Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. 1990 On the definition and the computation of the basic reproduction ratio R_0, in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382.

Diekmann, O., de Jong, M. C. M. & Metz, J. A. J. 1998 A deterministic epidemic model taking account of repeated contacts between the same individuals. *J. Appl. Prob.* **35**, 448–462.

Dietz, K. & Hadeler, K. P. 1988 Epidemiological models for sexually transmitted diseases *J. Math. Biol.* **26**, 1–25.

Grenfell, B. T. & Dobson, A. P. 1995 *Ecology of infectious diseases in natural populations*. Cambridge University Press.

Grenfell, B. T., Lonergan, M. E. & Harwood, J. 1992 Quantitative investigations of the epidemiology of phocine distemper virus (PDV) in European common seal populations. *Sci. Total Environ.* **115**, 15–29.

Hassell, M. P. & May, R. M. 1974 Aggregation in predators and insect parasites and its effects on stability. *J. Anim. Ecol.* **43**, 567–594.

Islam, M. N., Oshaughnessy, C. D. & Smith, B. 1996 A random graph model for the final-size distribution of household infections. *Statist. Med.* **15**, 837–843.

Keeling, M. J. & Grenfell, B. T. 1997 Disease extinction and community size: modeling the persistence of measles. *Science* **275**, 65–67.

Keeling, M. J. & Rand, D. A. 1999 Spatial correlations and local fluctuations in host–parasite systems. In *From finite to infinite dimensional dynamical systems* (ed. P. A. Glendinning). Amsterdam: Kluwer.

Keeling, M. J., Rand, D. A. & Morris, A. J. 1997 Correlation models for childhood epidemics. *Proc. R. Soc. Lond.* B **264**, 1149–1156.

Kermack, W. O. & Mc.Kendrick, A. G. 1927 Contributions to the mathematical theory of epidemics. 1. *Proc. R. Soc. Edinb.* A **115**, 700–721 (reprinted in *Bull. Math. Biol.* **53**, 33–55).

Kornberg, H. & Williamson, M. H. (eds) 1987 *Quantitative aspects of the ecology of biological invasions*. London: The Royal Society.

Levin, S. A. & Durrett, R. 1996 From individuals to epidemics. *Proc. R. Soc. Lond.* B **351**, 1615–1621.

Matsuda, H. 1987 Conditions for the evolution of altruism. In *Animal societies: theories and facts* (ed. Y. Ito, J. P. Brown & J. Kikkawa), pp. 67–80. Tokyo: Japanese Scientific Society Press.

Matsuda, H., Ogita, N., Sasaki, A. & Sato, K. 1992 Statistical mechanics of population: the lattice Lotka–Volterra model. *Prog. Theor. Phys.* **88**, 1035–1049.

Mollison, D. 1995 *Epidemic models: their structure and relation to data*. Cambridge University Press.

Morris, A. J. 1997 Representing spatial interactions in simple ecological models. PhD thesis, University of Warwick, UK.

Rand, D. A. 1999 Correlation equations for spatial ecologies. In *Advanced ecological theory* (ed. J. McGlade), pp. 99–143. London: Blackwell Scientific Publishing.

Sato, K., Matsuda, H. & Sasaki, A. 1994 Pathogen invasion and host extinction in lattice structured populations. *J. Math. Biol.* **32**, 251–268.

Van Baalen, M. 1999 Pair dynamics and configuration approximations. In *The geometry of ecological interactions: simplifying spatial complexity* (ed. U. Dieckmann, & J. Metz). Cambridge University Press. (In the press.)

Watts, D. J. & Strogatz, S. H. 1998 Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442.

As this paper exceeds the maximum length normally permitted, the author has agreed to contribute to production costs.