

REPORT

## Estimating spatial coupling in epidemiological systems: a mechanistic approach

Matt J. Keeling<sup>1</sup> and Pejman Rohani<sup>1,2</sup>

<sup>1</sup>Department of Zoology,  
University of Cambridge,  
Downing Street, Cambridge CB2  
3EJ, U.K. Tel./Fax:  
+44 1223 330110. E-mail:  
matt@zoo.cam.ac.uk

<sup>2</sup>Institute of Ecology, University  
of Georgia, Athens, GA 30602,  
U.S.A.

### Abstract

In recent years, ecologists and epidemiologists have paid increasing attention to the influence of spatial structure in shaping the dynamics and determining the persistence of populations. This is fundamentally affected by the concept of ‘coupling’—the flux of individuals moving between separate populations. In this paper, we contrast how coupling is typically implemented in epidemic models with more detailed approaches. Our aim is to link the popular phenomenological formulations with the results of mechanistic models. By concentrating on the behaviour of simple epidemic systems, we relate explicit movement patterns with observed levels of coupling, validating the standard formulation. The analysis is then extended to include a brief study of how the correlation between stochastic populations is affected by coupling, the underlying deterministic dynamics and the relative population sizes.

### Keywords

Disease models, movement, stochasticity, correlation, dispersal, synchrony, seasonality, population dynamics.

*Ecology Letters* (2002) 5: 20–29

### INTRODUCTION

Spatially structured models and emergent heterogeneity are becoming increasingly important features of ecological and epidemiological modelling (Tilman & Kareiva 1997; Rohani *et al.* 1999; Dieckmann *et al.* 2000; Jansen & Lloyd 2000; Bascompte & Solè 1998). One of the most common forms of capturing the effects of space is the patch or metapopulation model (Levins 1969; Allen 1975; Hassell *et al.* 1991; Rohani *et al.* 1996; Ruxton 1996; Hanski & Gilpin 1997; Keeling 2000). Such models consist of multiple subpopulations with movement (or ‘coupling’) between them, and hence provide an ideal framework for studying the dynamics of diseases within human communities (Grenfell & Harwood 1997).

Generally, increasing the strength of coupling between populations reduces spatial variation and acts to synchronize their dynamics. Sufficiently strong coupling may eventually lead to perfect synchrony or ‘phase-locking’. The extent of this effect is, however, known to depend on the precise geometry of the coupling and the intrinsic patch dynamics (Adler 1993; Allen *et al.* 1993; Ruxton & Rohani 1998; Bjørnstad 2000; Earn *et al.* 2000). Spatial synchrony is a key ecological (and epidemiological) phenomenon since it can profoundly affect the likelihood of metapopulation persistence. In the absence of spatial synchrony we observe rescue

effects, which prevent localized (subpopulation-level) extinctions either from occurring or becoming permanent, thus increasing the probability of global persistence (Brown & Kodric-Brown 1977; Ranta *et al.* 1997; Earn *et al.* 1998, 2000; Hanski 1998; Keeling 2000).

Coupling between distinct populations can take a variety of forms. In most models of ecological population dynamics, coupling arises primarily from migratory dispersal (Hanski 1998). Hence, the important features are those determining individual movement behaviours, which are quantifiable (though the precise estimation of these is often very time-consuming and rarely straightforward). In epidemiological systems, we are primarily concerned with the transfer of infection, which may occur via different routes. For sessile organisms such as plants, coupling is often due to external transport (e.g. by wind or insects) of the pathogen (Jeger 1989; Swinton & Gilligan 1996). A similar mechanism of dispersal largely accounts for the spread of diseases such as foot-and-mouth disease or swine-fever between farms (Keeling *et al.* 2001b). If one is interested in the epidemiology of childhood infections (such as measles or whooping cough), however, two important distinctions must be made.

Firstly, since these infections are directly transmitted, external transportation agents are not involved. Second, and more importantly, coupling no longer refers to the

permanent translocation of individuals or infection from one centre to another. Instead, it is concerned with the relative levels of ‘mixing’ within and between populations, with temporary movement assuming a far more significant role. Unfortunately, good data on relevant human mobility patterns are hard to find. We do, however, have access to excellent data sets on the spatio-temporal incidence of childhood infections (Keeling & Grenfell 1997; Rohani *et al.* 1998, 1999). From these data we can estimate the correlation between epidemics in different populations, and attempt to infer possible movement patterns.

In this paper, we will use the classic SIR (Susceptible–Infectious–Recovered) framework for disease dynamics to investigate the role of movement between human populations. First we consider how models with simple phenomenological coupling can be related to more complex mechanistic models which explicitly describe the movement of individuals between populations. While this relationship takes a simple form when the two populations are of equal size, it is far more complicated given hierarchical populations. In the final section, we attempt to link models with data by relating variation in correlation with coupling strength and how this is affected by population size and the inherent population dynamics.

## THE MODEL

We shall take as our basic framework the standard SIR model (Anderson & May 1991), where individuals are classified according to infection status (susceptible, infectious or recovered):

$$\begin{aligned}\frac{dS}{dt} &= bN - \beta(t)SI/N - dS \\ \frac{dI}{dt} &= \beta(t)SI/N - gI - dI \\ \frac{dR}{dt} &= gI - dR,\end{aligned}\quad (1)$$

where  $N (= S + I + R)$  is the total population size,  $b$  and  $d$  determine the per capita birth and death rates, respectively, and  $g^{-1}$  gives the infectious period. The parameter  $\beta$  represents the disease transmission rate and is often assumed to be constant, giving rise to damped oscillations in the deterministic model. In some of the best studied cases (e.g. childhood microparasitic infections), there is strong seasonal variation in contact rates and  $\beta$  is time-dependent; this may lead to more exotic dynamics (Schwartz & Smith 1983; Glendinning & Perry 1997; Earn *et al.* 2000; Keeling *et al.* 2001a). In most of the analysis that follows, we assume it to be constant, although we comment on numerical findings for seasonally forced models. We also set the birth and death rates to be equal ( $b = d$ ) such that the total population size,  $N$ , remains constant.

## LINKING COUPLING AND MOVEMENT PATTERNS

In systems where a number of populations are coupled together, the effect of spatial structure is commonly achieved by allowing the ‘force of infection’ (the per capita likelihood of contracting the infection; Anderson & May 1991) in each population to be influenced by the level of infectious individuals in the other population (Lloyd & May 1996; Bolker & Grenfell 1995; Earn *et al.* 1998; Rohani *et al.* 1999; Keeling 2000). For example, for a system of two identical populations ( $N_1 = N_2 = N$ ), the spatially structured set of equations are given by:

$$\begin{aligned}\text{Population 1} \quad & \begin{cases} \frac{dS_1}{dt} = bN - \beta S_1(I_1(1 - \sigma) + I_2\sigma)/N - dS_1 \\ \frac{dI_1}{dt} = \beta S_1(I_1(1 - \sigma) + I_2\sigma)/N - gI_1 - dI_1 \end{cases} \\ \text{Population 2} \quad & \begin{cases} \frac{dS_2}{dt} = bN - \beta S_2(I_2(1 - \sigma) + I_1\sigma)/N - dS_2 \\ \frac{dI_2}{dt} = \beta S_2(I_2(1 - \sigma) + I_1\sigma)/N - gI_2 - dI_2. \end{cases}\end{aligned}\quad (2)$$

The parameter  $\sigma$  defines a ‘coupling strength’ and is related in a phenomenological way to the movement rate between the two populations. Note that  $\sigma$  merely reflects how much ‘mixing’ exists between the two populations, measuring the proportion of infection that occurs between rather than within patches—it does not specify a dispersal rate. The precise formulation of Equation (2) is chosen such that the basic reproductive ratio,  $R_0 = \beta/g$ , remains constant as the strength of the coupling,  $\sigma$ , changes. This is a consistent feature of all the models given in this paper, and is vital if we are to reliably compare models with different couplings.

Such a method of coupling populations is clearly an oversimplification of the true interactions between communities, although it has been widely used in the literature (Grenfell *et al.* 1995; Bolker & Grenfell 1996; Lloyd & May 1996; Swinton *et al.* 1998; Rohani *et al.* 1999; Earn *et al.* 2000). Implicit within this formulation is the idea that infectious individuals from population 1 can infect susceptibles in population 2, but this is achieved without any explicit mechanism for the transfer of infection. The primary aim of this paper, therefore, is to understand how basic movement patterns affect the overall ‘coupling’ between populations.

In order to construct a more mechanistic model, we need to specify the exact mixing patterns between the populations. The precise mechanisms by which infections are transferred to new centres are difficult to establish. Clearly, they involve the interchange of individuals between populations, but without precise data on patterns of human mobility, it is hard to describe what may be ‘realistic’. We shall initially consider the interaction between two

populations of the same size. Individuals may leave their ‘permanent’ population and enter the other (‘temporary’) population, where they will remain for a given period of time. Within each population, we have SIR dynamics.

The nomenclature for this model involves identifying all individuals by two subscripts, such that  $S_{xy}$  refers to susceptibles from population  $x$  that are currently situated in population  $y$ . This allows us to explicitly model the movement of individuals back and forth between the two populations (Fig. 1) and to quantify the effects of this movement on the epidemiological coupling. In this mechanistic model, infection can only occur when both susceptible and infectious individuals are present in the same population. We assume infection operates as pseudo mass-action (de Jong *et al.* 1995; McCallum *et al.* 2001), although if both populations are initially of equal size this assumption has no bearing on the dynamics. The equations have three separate components, births and deaths, infection, and movement between populations; for example, the dynamics of  $S_{xx}$  are given by:

$$\frac{dS_{xx}}{dt} = bN_{xx} - dS_{xx} - \beta S_{xx}(I_{xx} + I_{yx}) / (N_{xx} + N_{yx}) + \tau S_{xy} - \rho S_{xx},$$

where  $\rho$  and  $\tau$  are the rates at which individuals leave and return to their home location, respectively (as shown in Fig. 1). The remaining equations are given in Appendix 1. For convenience, we define and formulate results in terms of the parameter  $\mu = \rho/\tau$ , which measures the ratio of time spent in the temporary to the permanent population.

In this (unforced) system, for most biologically realistic parameters, the population levels rapidly converge to an equilibrium such that the number of individuals in the permanent and temporary location are related by:

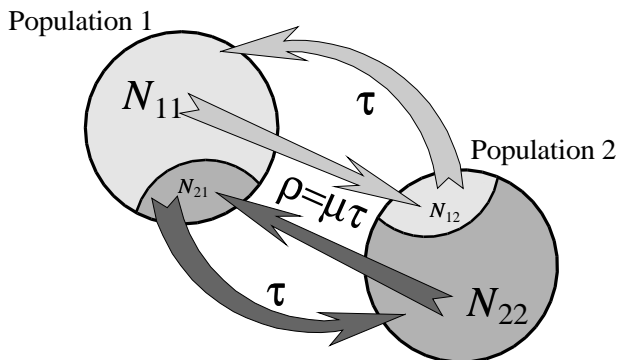
$$\frac{N_{xy}}{N_{xx}} \rightarrow \mu. \tag{3}$$

Assuming this rapid population level convergence has already occurred, our mechanistic model (Appendix 1, equation 6) is an eight dimensional system, which we would like to equate with the much simpler four dimensional phenomenological model (equation 2). We achieve this by considering the eigenvalues of each system, and comparing how rapidly the populations converge in each formulation. In these unforced models, we always observe convergence of the two populations to a stable equilibrium point. By adding coupling between the populations the basic rate of convergence can be increased; and it is this increase that permits us to equate a level of coupling ( $\sigma$ ) from equation (2) with mechanistic movement rates ( $\rho, \tau$ ) from equation (6).

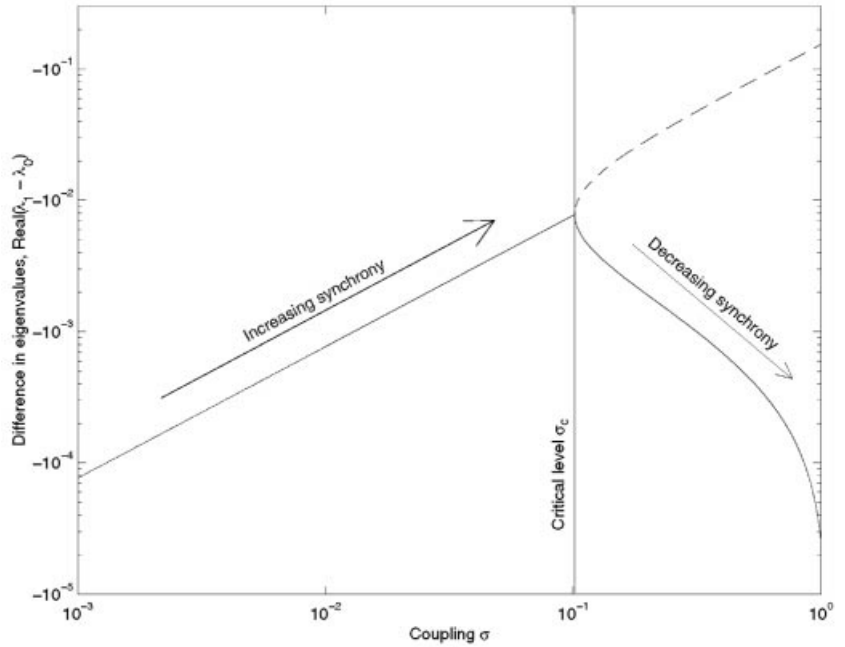
For all reasonable choice of parameters, the eigenvalues of the simple uncoupled and unforced model (equation 1) are complex conjugates,  $\lambda_0^+$  and  $\lambda_0^-$ , with negative real parts. This corresponds to damped oscillations towards the fixed point (Anderson & May 1991). We expect the behaviour dictated by this pair of eigenvalues to dominate the behaviour of both the phenomenological and mechanistic models. Since the coupled system (equation 2) is four dimensional, it has two pairs of eigenvalues ( $\lambda_1^+, \lambda_1^-$ ) and ( $\lambda_2^+, \lambda_2^-$ ). One pair is always equal to the eigenvalues of the uncoupled system (without loss of generality, we assume that  $\lambda_2^\pm = \lambda_0^\pm$ ) and provides information on the convergence to the fixed point. The second set of eigenvalues ( $\lambda_1^\pm$ ) determine the speed with which trajectories in the two populations approach each other.

In Fig. 2(a), we plot the value of  $Real(\lambda_1^\pm - \lambda_0^\pm)$ , which measures the extra rate of convergence of the two populations towards each other in addition to the normal convergence to the fixed point. The figure demonstrates that below some critical level ( $\sigma_c \approx 0.1$ ) greater coupling increases the synchronizing effect, and this is reflected by a more rapid convergence of trajectories in different populations towards each other. Above the critical coupling level, the eigenvalues  $\lambda_1^\pm$  become real and so two values are observed in Fig. 2(a). In this region, increasing the coupling  $\sigma$  generally leads to slower convergence, as the larger eigenvalue dominates. Therefore, at the critical level of coupling, convergence and hence synchrony of the two populations is maximized.

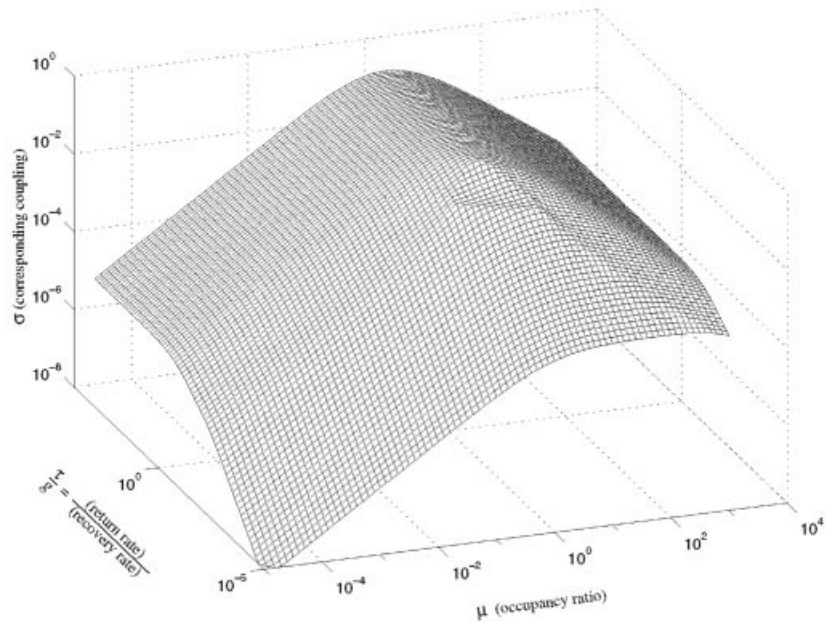
The mechanistic model is described by a set of eight differential equations and therefore has four pairs of eigenvalues, which we denote by,  $\Lambda_n^\pm$ , ( $n = 1, \dots, 4$ ). Two of these pairs (say,  $\Lambda_3^\pm$  and  $\Lambda_4^\pm$ ) are large and negative; these correspond to the rapid distribution of individuals between the permanent and temporary populations, and can be ignored (for our purposes). Another pair of eigenvalues, say  $\Lambda_2^\pm$ , are equal to the eigenvalues of the uncoupled system



**Figure 1** Schematic representation of the movement of infected individuals between two populations. Individuals in their permanent population ( $N_{11}$  and  $N_{22}$ ) visit the other, temporary population at rate  $\rho = \mu\tau$ , those individuals in the temporary population ( $N_{12}$  and  $N_{21}$ ) return at rate  $\tau$ .



**Figure 2** (a) Comparison between the eigenvalues of the 4-dimensional coupled system (equation 2),  $\lambda_1^\pm$ , and the eigenvalue of the 2-dimensional SIR model (equation 1),  $\lambda_0^\pm$ . We note that  $Real(\lambda_1^\pm - \lambda_0^\pm)$  is always negative indicating faster convergence in the coupled model. For levels of coupling,  $\sigma$ , above 0.1 the eigenvalues  $\lambda_1^\pm$  are real and hence two curves exist on the graph. (b) The relationship between occupancy ratio,  $\mu$ , the relative return rate,  $\frac{\tau}{g}$ , and the phenomenological coupling,  $\sigma$ . Note that the coupling reaches a maximum when  $\mu = 1$  such that equal time is spent in both populations. Asymptotic behaviour is reached whenever the movement rate  $\tau$  is large compared to the recovery rate  $g$ . ( $b = d = 5.5 \times 10^{-5}$  days $^{-1}$ ,  $g^{-1} = 13$  days,  $R_0 = \frac{\beta}{g} = 17$ ).



(equation 1). This again leaves just one pair of eigenvalues remaining ( $\Lambda_1^\pm$ ) to determine the effects of coupling on the disease dynamics. Comparing  $\Lambda_1^\pm(\mu, \tau)$  with  $\Lambda_1^\pm(\sigma)$ , allows us to determine a relationship between the parameters of the two models.

Figure 2(b) shows the level of simple phenomenological coupling  $\sigma$  corresponding to the mechanistic parameters  $\tau, g$

and  $\mu$ . We note from the graph that the coupling is maximized when the returning rate ( $\tau$ ) is large (such that there is rapid mixing between the populations) and when individuals spend equal times in each population ( $\mu = 1$ ). When  $\tau$  is small, individuals spend a long time away from their permanent population, it is therefore quite likely that if they catch the disease they pass into the recovered class before they return

home. Small values of  $\tau$  mean that individuals cannot carry the disease between the two populations and hence there is a reduction in the amount of coupling.

It is interesting to note that in general as the infectious period changes, so does the relationship between  $\mu$ ,  $\tau$  and  $\sigma$ . Thus, although the movement patterns are determined by the population and are independent of the disease, the epidemiological differences in the infectious period mean that the precise value of  $\sigma$  will be disease dependent. Analysis shows that it is the length of the infectious period relative to the average time spent away from home ( $\tau/g$ ) that determines the level of coupling (assuming that the infectious period is much shorter than the life expectancy). We notice that when the time spent in the temporary population is smaller than the infectious period  $\tau^{-1} < g^{-1}$ , the eigenvalues are very close to the large  $\tau$  limit (see below). Intuitively, in most applied settings we might expect short stays in comparison to the infectious period so the precise value of  $\tau$  and  $g$  are likely to be unimportant. (For infections like measles and whooping cough, we are dealing with infectious periods of 5 and 15 days, respectively, and an intuitive assumption would be that ‘mixing’ occurs on the scale of a day). Therefore, in such cases it may be appropriate to study the large  $\tau$  limit, where precise disease parameters do not affect the level of coupling.

When the returning rate ( $\tau$ ) is very high, it is possible to obtain a simple algebraic relationship between the strength of coupling ( $\sigma$ ) and distribution of individuals ( $\mu$ ). This is due to the rapid convergence of the distribution of individuals between the two populations such that,

$$N_{xy} \rightarrow \mu N_{xx} \quad S_{xy} \rightarrow \mu S_{xx} \quad I_{xy} \rightarrow \mu I_{xx}.$$

Summing the differential equations for the number of susceptible individuals whose permanent home is population  $x$  gives:

$$\begin{aligned} \frac{dS_x}{dt} &= \frac{d}{dt}(S_{xx} + S_{xy}) \\ &= bN - \beta S_x \left( \left[ \frac{1 + \mu^2}{(1 + \mu)^2} \right] I_x + \left[ \frac{2\mu}{(1 + \mu)^2} \right] I_y \right) / N - dS_x \end{aligned}$$

with a similar equation for  $I_x$ . These are clearly identical to the phenomenological equation (2) with the strength of the coupling as given by

$$\sigma = \frac{2\mu}{(1 + \mu)^2}. \tag{4}$$

Hence in the majority of applied settings when  $\tau$  is large compared to  $g$ , the full mechanistic model rapidly converges to the much simpler coupled model, and there is a simple relationship between the parameters of the two models.

A similar theory relating phenomenological coupling to movement parameters can be developed for the more realistic situation where the two coupled populations are of

different sizes (Appendix 2). The analysis shows that given identical intrinsic movement rates in the two populations, coupling levels are identical except for a scaling factor determined by the relative population sizes.

### CORRELATIONS FOR COUPLED STOCHASTIC MODELS

Although a detailed social survey may allow us to estimate the parameters  $\mu$  and  $\tau$  and therefore find a corresponding value of  $\sigma$ , for many diseases we are simply faced with the number of reported cases in each community. In this section, using a stochastic or Monte Carlo version of equation (2) (Renshaw 1991), we compare the correlation between cases in two model populations, with the level of coupling  $\sigma$ . With this stochastic model, all events (birth, death, infection and recovery) occur at random, but with the same underlying rate as predicted by the differential equations. In addition, to prevent permanent stochastic extinctions of the disease, a small immigration rate ( $\varepsilon$ ) of infectious individuals was added to the model.

For the time-series of infectious cases from the two populations,  $I_1(t)$  and  $I_2(t)$  the correlation is defined as,

$$C = \frac{\langle I_1 I_2 \rangle - \langle I_1 \rangle \langle I_2 \rangle}{\sqrt{(\langle I_1^2 \rangle - \langle I_1 \rangle^2)(\langle I_2^2 \rangle - \langle I_2 \rangle^2)}}$$

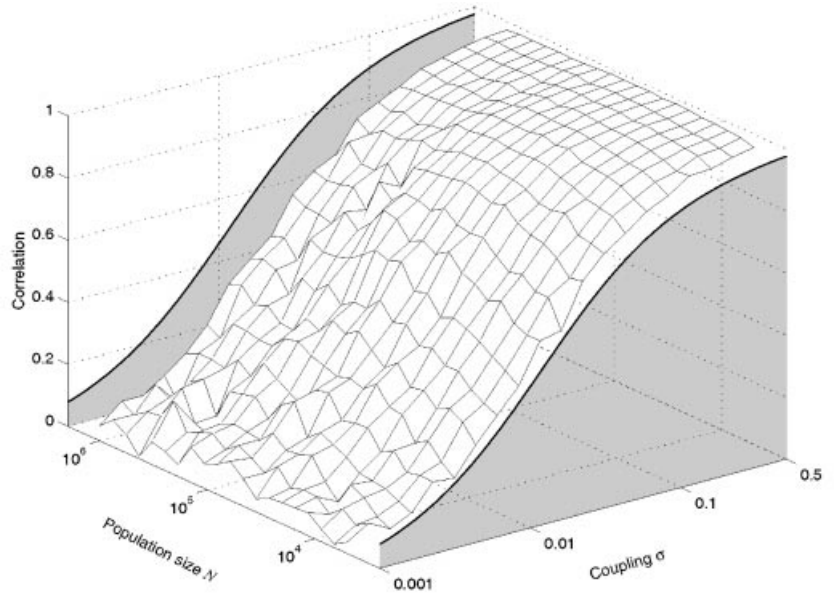
Here  $\langle \cdot \rangle$  corresponds to the long-term average. Figure 3(a) shows the correlation (as calculated from a 5000-year sample) for two populations of equal size  $N$ , with coupling  $\sigma$ .

### Theoretical prediction

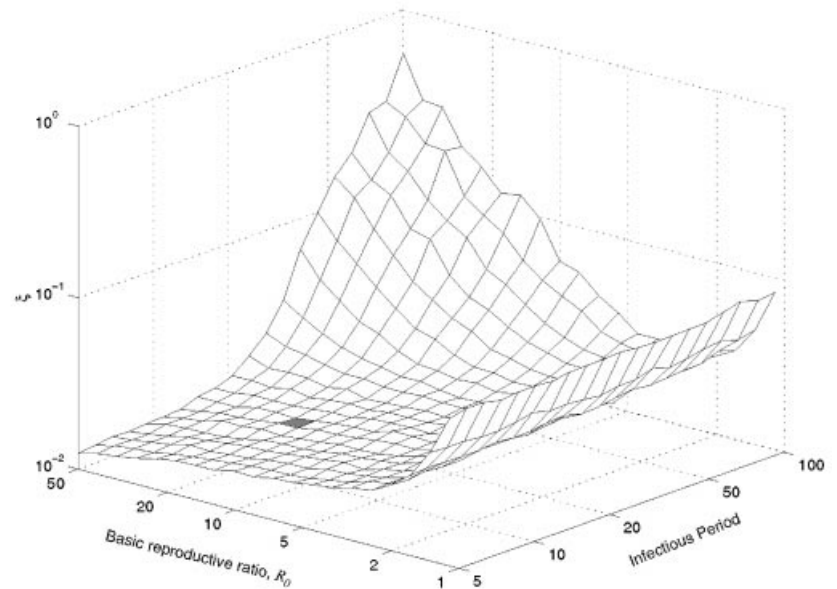
By formulating an equation for the covariance between the number of infectious individuals in the two populations, we can begin to understand the behaviour of the correlation. This is achieved using moment closure approximations (Renshaw 1991; Keeling 2000) in which means, variances and covariances are modelled but third order cumulants are ignored. As detailed in Appendix 3, we obtain the following approximation for the correlation,

$$C = \frac{\sigma}{\xi + \sigma}. \tag{5}$$

where  $\xi$  is a parameter which can be estimated from numerical simulations. In Fig. 3(a), we find that  $\xi \approx 0.0119$ , and is independent of population size—the shaded curves show the theoretically predicted correlation using this value. Thus, for a given disease and set of demographic parameters, there is a unique curve relating the amount of correlation observed with the strength of coupling. Hence, this result allows us to infer individual level parameters from global behaviour. Given the disease dynamics of two simple



**Figure 3** The correlation between number of infectious individuals in two populations using a stochastic (Monte Carlo) version of the coupled model (equation 2). Part (a) clearly shows that the correlation is largely unaffected by population size, and that the simulation results are a close fit to the theoretical predictions (thick lines). ( $b = d = 5.5 \times 10^{-5}$  days $^{-1}$ ,  $g^{-1} = 13$  days,  $R_0 = \frac{\beta}{\varepsilon} = 17$ ,  $\xi = 0.0119$ , following the work of Bartlett (1956) the import rate was chosen as  $\varepsilon = 5.5 \times 10^{-5} \sqrt{N}$  days $^{-1}$ ). Part (b) gives the value of  $\xi$  which provides the best fit to the correlation, in terms of minimizing the mean squared error. The shaded square corresponds to the epidemiological parameters used throughout the rest of the paper. ( $b = d = 5.5 \times 10^{-5}$  days $^{-1}$ ,  $\sigma = 10^{-3}$  to 0.5,  $N = 10^5$ ).



coupled populations, it is possible to calculate the phenomenological level of coupling and therefore the individual movement rates between the populations.

Figure 3(b) shows the value of  $\xi$  which provides the best fit to stochastic simulation results for a range of disease parameters. Again we find that equation (5) provides a good description of the correlation. The parameter,  $\xi$ , is large whenever both  $R_0$  and the infectious period ( $g^{-1}$ ) are large; in this region of parameter space the correlation  $C$  increases slowly and almost linearly with the coupling. The increase in  $\xi$  for small values of  $R_0$  may be attributed to frequent extinctions within the model populations. In this regime the theoretical prediction is less accurate, but still gives a good

qualitative description of the relationship between correlation and coupling.

### Spatial complexity and seasonality

So far we have only considered the unforced system (which possesses a globally stable fixed point) and two isolated populations. In this section we shall consider the difficulties that arise in real applications when there is often strong seasonality and as well as coupling between numerous locations.

To capture the cyclic behaviour observed for many diseases, seasonality in the contact parameter  $\beta$  has to be

included (Fine & Clarkson 1982; Schwartz & Smith 1983). Here we use the formulation estimated for measles dynamics, where  $\beta$  takes a high value during term-times and a lower value during school holidays (Fine & Clarkson 1982; Finkenstädt & Grenfell 2000). This produces biennial dynamics in the deterministic model for measles parameters in developed countries (Anderson & May 1991; Keeling *et al.* 2001a).

When seasonality is included, we need greater care in comparing the dynamics of the two populations. What we wish to measure is the correlation in the deviation from the average cyclic pattern. With measles parameters, for example, it becomes necessary to define a separate correlation  $C_t$  for every point  $t$  ( $0 \leq t \leq 2$  years) on the biennial attractor. The correlation between two cyclic populations can then be taken as the average over the entire cycle of the separate correlations. With this modified definition of  $C$ , we find that the theoretical prediction (equation 5) remains a reliable approximation. We note that if the cyclic dynamics were ignored, and we simply calculated the correlation in the standard manner, then synchronization (or phase-locking) of the epidemics in the two communities would produce much higher correlations than expected when the populations are large ( $N > 10^5$ ).

This methodology must be used with caution. For some diseases, such as whooping cough (Rohani *et al.* 1999) or rubella (Anderson & May 1991), seasonal changes drive irregular epidemics. Other diseases (for example influenza) are dominated by localized extinctions and subsequent recolonization, as well as other complications relating to antigenic variation. For such diseases, it is difficult to define an average cyclic pattern, and our theoretical approximation breaks-down whenever deterministic epidemics become out of phase.

Most applied situations also require us to consider multiple populations with different degrees of coupling between them. A complete analysis of all the vast number of possible scenarios is not feasible. Here, we note in brief the results from two simple spatial configurations—arranging  $n$  identical populations in a line with either local or global coupling. For both these situations the same basic theoretical relationship between coupling and correlation still holds. With global coupling (every population coupled to every other population with strength  $\sigma$ ) the value of  $\xi$  shows a very slight decrease with the number of populations. For local coupling the situation is more complex,  $\xi$  increases as we look at the correlation between ever more distant populations.

## CONCLUSIONS

In human epidemiology, the interaction between discrete settlements can have a profound dynamic effect and important public health consequences. Most notably,

re-infection from an external source (rescue effects) plays a significant role in the global persistence of many diseases (Grenfell *et al.* 1995; Grenfell & Harwood 1997). It is therefore important to have reliable models for the interaction between communities. We have established a clear equivalence between the standard phenomenological models of coupling, and a more mechanistic approach based on individual movement patterns. While this research has been primarily driven by patterns of human mobility, our approach should be applicable to any territorial organism, which spends short periods away from its permanent population.

The relationship between the phenomenological coupling and the mechanistic movement rates can be calculated numerically for any combination of disease and demographic parameters. It is interesting to note that the level of coupling is sensitive to the length of the infectious period, as well as on the movement rates. Despite identical underlying movement behaviours, the coupling between communities is disease dependent. This is because when the infectious period is short and the length of stay in the temporary population long, it is unlikely that an individual can acquire the disease in one population and return home before recovering. However, for most human infections the time spent out of the home population is short relative to the incubation period of the disease. In this case, there exists an exact analytical relationship between the phenomenological and the mechanistic models,

$$\sigma = \frac{2\mu}{(1 + \mu)^2}.$$

For most human populations, it is difficult to assess the level of movement between different communities. Even when records of travel do exist, they are rarely stratified sufficiently for us to identify the potential mixing between susceptible and infectious individuals. For childhood diseases, the distribution of disease is age dependent, so information on the movement rates of recovered adults is irrelevant; similarly for sexually transmitted diseases, the risk associated with each individual is highly variable and not easily ascertained. Often, however, we have extremely detailed reports of the number of cases in each community. Stochastic simulations are used to generate surrogate data for two coupled populations. Results from these simulations agree with simple analytical predictions that in many situations the correlation,  $C$ , is related to the coupling,  $\sigma$ , by

$$C \approx \frac{\sigma}{\xi + \sigma}$$

where  $\xi$  is a function of the particular disease parameters, but does not depend upon population size.

This work has mainly focused on the most tractable scenario, the interaction between two communities of equal size where the underlying disease dynamics possess a fixed

point attractor. However, some consideration has also been given to the three main complications: different sized populations, cyclic dynamics and multiple interacting populations.

- When the populations are of different sizes, we again find that a phenomenological model is equivalent to the mechanistic form, although four distinct coupling parameters become necessary. If we again assume that visits to another population are of much shorter duration than the infectious period then a simplified analytical approximation with just two coupling parameters is achieved. Correlations between populations of vastly different sizes do not agree well with the theoretical formula, but it still provides a good approximation when the population sizes are within an order of magnitude of each other.
- If the dynamics have a regular cyclic pattern, then we must subtract this pattern from the time-series before calculating the correlation; hence we are looking at the correlation between the deviations away from the attractor. The correlation then obeys the same formula as before, with a similar  $\xi$  value as a non-cyclic (non-seasonally forced) model. When diseases are strongly influenced by stochastic forces and epidemics are irregular and out of phase, it is difficult to assess the effects of any small degree of coupling.
- When many populations interact, the array of possible mixing patterns is vast. Concentrating only on local and global interactions, with equal coupling strengths, we firmly believe that the same relationship between correlation and coupling still exists, although the value of  $\xi$  may depend on the number of populations and the nature of the coupling.

Throughout we have assumed an SIR-type formulation, so that individuals are infectious as soon as they are infected. However, for many diseases the SEIR model (Anderson & May 1991) is more appropriate; this introduces an exposed class when individuals are have caught the pathogen, but are incubating the disease and are not yet infectious. All of our results transfer in the obvious manner to this SEIR model, with very little change in the behaviour or quantitative predictions.

We have found clear analytical relationships between the individual movement patterns and phenomenological coupling, and then between coupling and the observed correlations. These relationships fairly robust, and therefore provide us with an important tool for analysing disease dynamics in a spatial context.

## ACKNOWLEDGEMENTS

We thank the four anonymous referees for their comments on this paper. This research was supported by the Royal Society (MJK and PR).

## REFERENCES

- Adler, F.R. (1993). Migration Alone can Produce Persistence of Host-Parasitoid Models. *Am. Naturalist*, 141, 642–650.
- Allen, J.C. (1975). Mathematical models of species interactions in time and space. *Am. Naturalist*, 109, 319–342.
- Allen, J.C., Schaffer, W.M. & Rosko, D. (1993). Chaos reduces species extinction by amplifying local population noise. *Nature*, 364, 229–232.
- Anderson, R.M. & May, R.M. (1991). *Infectious Diseases of Humans*. Oxford University Press.
- Bartlett, M.S. (1956). Deterministic and stochastic models for recurrent epidemics. *Proceedings of the of the Third Berkeley Symposium on Mathematics, Statistics and Probability*, 4, 81–108.
- Bascompte, J. & Solè, R.V. (1998). Spatiotemporal patterns in nature. *Trends Ecol. Evol.*, 5, 173–174.
- Bjørnstad, O.N. (2000). Cycles and synchrony: two historical ‘experiments’ and one experience. *J. Anim. Ecol.*, 69, 869–873.
- Bolker, B.M. & Grenfell, B.T. (1995). Space, persistence and dynamics of measles epidemics. *Phil. Trans. Royal Soc. Lond. B*, 348, 309–320.
- Bolker, B.M. & Grenfell, B.T. (1996). Impact of vaccination on the spatial correlation and persistence of measles dynamics. *Proc. Natl. Acad. Sci USA*, 93, 12648–12653.
- Brown, J.H. & Kodric-Brown, A. (1977). Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology*, 58, 445–449.
- Dieckmann, U., Law, R. & Metz, J. (2000). *The Geometry of Ecological Interactions*. Cambridge University Press.
- Earn, D.J.D., Rohani, P., Bolker, B.M. & Grenfell, B.T. (2000). A simple model for complex dynamical transitions in epidemics. *Science*, 287, 667–670.
- Earn, D.J.D., Rohani, P. & Grenfell, B.T. (1998). Persistence, chaos and synchrony in ecology and epidemiology. *Proc. Royal Soc. Lond. B*, 265, 7–10.
- Fine, P.E.M. & Clarkson, J.A. (1982). Measles in England and Wales I. An analysis of factors Underlying Seasonal Patterns. *Internat. J. Epidemiol.*, 11, 5–14.
- Finkenstädt, B. & Grenfell, B. (2000). Time series modelling of childhood diseases: a dynamical systems approach. *J. Royal Statistical Soc. C*, 49, 187–205.
- Glendinning, P. & Perry, L.P. (1997). Melnikov analysis of chaos in a simple epidemiological model. *Math. Bioscience*, 35, 359–374.
- Grenfell, B.T., Bolker, B.M. & Kleczkowski, A. (1995). Seasonality and extinction in chaotic metapopulations. *Proc. Royal Soc. Lond. B*, 259, 97–103.
- Grenfell, B. & Harwood, J. (1997). (Meta) population dynamics of infectious diseases. *Trends Ecol. Evol.*, 12, 395–399.
- Hanski, I. (1998). Metapopulation dynamics. *Nature*, 396, 41–49.
- Hanski, I. (1999). *Metapopulation Ecology*. Oxford University Press, Oxford.
- Hanski, I. & Gilpin, M.E., eds. (1997). *Metapopulation Biology: Ecology, Genetics and Evolution*. Academic Press.
- Hassell, M.P., Comins, H. & May, R.M. (1991). Spatial structure and chaos in insect population dynamics. *Nature*, 353, 255–258.
- Jansen, V.A.A. & Lloyd, A.L. (2000). Local stability analysis of spatially homogeneous solutions of multi-patch systems. *J. Mathemat. Biol.*, 41, 232–252.
- Jeger, M.J. (1989). *Spatial Components of Plant Disease Epidemics*. Prentice Hall.



- de Jong, M.C.M. *et al.* (1995). How does transmission of infection depend on population size? In: *Epidemic Models: Their Structure and Relation to Data*, ed. Mollison, D. Cambridge University Press, pp. 84–94.
- Keeling, M.J. (2000). Metapopulation moments: coupling, stochasticity and persistence. *J. Anim. Ecol.*, 69, 725–736.
- Keeling, M.J. & Grenfell, B.T. (1997). Disease extinction and community size: modeling persistence measles. *Science*, 275, 65–67.
- Keeling, M.J., Rohani, P. & Grenfell, B.T. (2001a). Seasonally-forced disease dynamics explored as switching between attractors. *Physica D*, 148, 317–335.
- Keeling, M.J., Woolhouse, M.E.J., Shaw, D.J., Matthews, J., Chase-Topping, M., Haydon, D.T., Cornell, S.J., Kappey, J., Wilesmith J. and Grenfell, B.T. (2001b) Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science*, 294, 813–817.
- Levins, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull. Entomol. Soc. America*, 15, 237–240.
- Lloyd, A.L. & May, R.M. (1996). Spatial Heterogeneity in Epidemic Models. *J. Theoret. Biol.*, 179, 1–11.
- McCallum, H. *et al.* (2001). How should pathogen transmission be modelled? *Trends Ecol. Evol.*, 16, 295–300.
- Ranta, E., Kaitala, V. & Lundberg, P. (1997). The spatial dimension in population fluctuations. *Science*, 278, 1621–1623.
- Renshaw, E. (1991). *Modelling Biological Populations in Space and Time*. University Press, Cambridge.
- Rohani, P., Earn, D.J., Finkenstadt, B. & Grenfell, B.T. (1998). Population dynamic interference among childhood diseases. *Proc. Royal Soc. Lond. B*, 265, 2033–2041.
- Rohani, P., Earn, D.J.D. & Grenfell, B.T. (1999). Opposite patterns of synchrony in sympatric disease metapopulations. *Science*, 286, 968–971.
- Rohani, P., May, R.M. & Hassell, M.P. (1996). Metapopulations and equilibrium stability—the effects of spatial structure. *J. Theoret. Biol.*, 181, 97–109.
- Ruxton, G.D. (1996). Dispersal and chaos in spatially structured models—an individual-level approach. *J. Anim. Ecol.*, 65, 161–169.
- Ruxton, G.D. & Rohani, P. (1998). Fitness-dependent dispersal in metapopulations and its consequences for persistence and synchrony. *J. Anim. Ecol.*, 67, 530–539.
- Schwartz, I.B. & Smith, H.L. (1983). Infinite subharmonic bifurcation in an SEIR epidemic model. *J. Mathemat. Biol.*, 18, 233–253.
- Swinton, J. & Gilligan, C.A. (1996). Dutch elm disease and the future of the elm in the UK: a quantitative analysis. *Phil. Trans. Royal Soc. Lond. B*, 351, 605–615.
- Swinton, J., Harwood, J., Grenfell, B.T. & Gilligan, C.A. (1998). Persistence thresholds for phocine distemper virus infection in harbour seal *Phoca vitulina* metapopulations. *J. Anim. Ecol.*, 67, 54–68.
- Tilman, D. & Kareiva, P. (1997). *Spatial Ecology*. Princeton University Press.

Editor, M. Hochberg

Manuscript received 28 June 2001

First decision made 8 August 2001

Manuscript accepted 20 August 2001

## APPENDIX I

### Mechanistic equations for populations of equal sizes

$$\begin{aligned} \frac{dS_{xx}}{dt} &= bN_{xx} - \beta S_{xx}(I_{xx} + I_{yx}) / (N_{xx} + N_{yx}) \\ &\quad - dS_{xx} + \tau S_{xy} - \rho S_{xx} \\ \frac{dS_{xy}}{dt} &= bN_{xy} - \beta S_{xy}(I_{xy} + I_{yy}) / (N_{xy} + N_{yy}) \\ &\quad - dS_{xy} + \rho S_{xx} - \tau S_{xy} \\ \frac{dI_{xx}}{dt} &= \beta S_{xx}(I_{xx} + I_{yx}) / (N_{xx} + N_{yx}) - gI_{xx} \\ &\quad - dI_{xx} + \tau I_{xy} - \rho I_{xx} \\ \frac{dI_{xy}}{dt} &= \beta S_{xy}(I_{xy} + I_{yy}) / (N_{xy} + N_{yy}) - gI_{xy} \\ &\quad - dI_{xy} + \rho I_{xx} - \tau I_{xy} \\ \frac{dN_{xx}}{dt} &= bN_{xx} - dN_{xx} + \tau N_{xy} - \rho N_{xx} \\ \frac{dN_{xy}}{dt} &= bN_{xy} - dN_{xy} + \rho N_{xx} - \tau N_{xy} \end{aligned} \quad (6)$$

where  $x = 1$  or  $2$  and  $y \neq x$ . This formulation corresponds to individuals leaving their permanent population at rate  $\rho$  and returning at rate  $\tau$  (Fig. 1). Note that individuals inherit their parents' permanent location, irrespective of where they are born. Susceptible individuals can only catch the infection from infectious individuals in the same location.

## APPENDIX 2

### Populations of different sizes

When the two populations are of different sizes then we necessarily have to consider that the parameter  $\mu$  may be dependent upon  $N_x$  and  $N_y$ . However, it would seem sensible to assume that the time spent in the other community ( $\tau^{-1}$ ) is independent of the population sizes. To approximate the value of  $\mu$ , let us assume that individuals (in either population) randomly visit other individuals (irrespective of location) at a rate  $\gamma\tau$ . The value of  $\rho$  and therefore  $\mu$  can be calculated from the rate of visiting an individual in the other population.  $\rho_x$  is the rate at which someone in  $x$  visits someone in  $y$ , and is calculated as the visiting rate times the probability that the visit is to someone in  $y$ :

$$\rho_x = \gamma\tau \frac{N_y}{N_x + N_y} \Rightarrow \mu_x = \gamma \frac{N_y}{N_x + N_y},$$

hence the value of  $\mu$  is maximized when the other (temporary) population is very large in comparison to the permanent population. We note that the total number of

individuals currently in a population  $P_x = N_{xx} + N_{yx}$  is no longer equal to the number of individuals that belong to that population; but again we expect rapid convergence of the distribution of individuals between patches such that,

$$P_x \rightarrow N_x \frac{1}{1 + \mu_x} + N_y \frac{\mu_y}{1 + \mu_y}$$

We can now reformulate the full model,

$$\begin{aligned} \frac{dS_{xx}}{dt} &= N_{xx}b - \beta S_{xx}(I_{xx} + I_{yx})/P_x - dS_{xx} + \tau S_{xy} - \mu_x \tau S_{xx} \\ \frac{dS_{yy}}{dt} &= N_{yy}b - \beta S_{yy}(I_{xy} + I_{yy})/P_y - dS_{yy} + \mu_x \tau S_{xx} - \tau S_{xy} \\ \frac{dI_{xx}}{dt} &= \beta S_{xx}(I_{xx} + I_{yx})/P_x - gI_{xx} - dI_{xx} + \tau I_{xy} - \mu_x \tau I_{xx} \\ \frac{dI_{xy}}{dt} &= \beta S_{xy}(I_{xy} + I_{yy})/P_y - gI_{xy} - dI_{xy} + \mu_x \tau I_{xx} - \tau I_{xy} \end{aligned} \quad (7)$$

Such a formulation still retains the same value of  $R_0$  (irrespective of movement rates) and still allows us to reduce the system to four dimensions, although the coupling terms are more complex.

$$\begin{aligned} \frac{dS_x}{dt} &= bN - \beta S_x(I_x(1 - \sigma_{xx}) + I_y \sigma_{xy})/N_x - dS_x \\ \frac{dI_x}{dt} &= \beta S_x(I_x(1 - \sigma_{xx}) + I_y \sigma_{xy})/N_x - gI_x - dI_x. \end{aligned} \quad (8)$$

In the limit where  $\tau$  is large, the four coupling terms  $\sigma_{xx}$ ,  $\sigma_{xy}$ ,  $\sigma_{yx}$  and  $\sigma_{yy}$  are found to converge to two asymptotic values,

$$\begin{aligned} \sigma_{xx} \rightarrow \sigma_{yx} \rightarrow \sigma_x &= \frac{N_y}{(1 + \mu_x)(1 + \mu_y)} \left[ \frac{\mu_y}{P_x} + \frac{\mu_x}{P_y} \right] \\ \sigma_{yy} \rightarrow \sigma_{xy} \rightarrow \sigma_y &= \frac{N_x}{(1 + \mu_x)(1 + \mu_y)} \left[ \frac{\mu_y}{P_x} + \frac{\mu_x}{P_y} \right]. \end{aligned}$$

### APPENDIX 3

#### Theoretical calculation of the correlation

Throughout this section we shall assume that the two populations are of equal sizes, as this greatly simplifies the

mathematics and notation. Let  $C_{XY}$  be the covariance between  $X$  and  $Y$  in the same population,

$$C_{XY} = C_{YX} = \langle X_1 Y_1 \rangle - \langle X_1 \rangle \langle Y_1 \rangle = \langle X_2 Y_2 \rangle - \langle X_2 \rangle \langle Y_2 \rangle$$

and  $\bar{C}_{XY}$  be the covariance when  $X$  and  $Y$  are in different populations,

$$\bar{C}_{XY} = \langle X_1 Y_2 \rangle - \langle X_1 \rangle \langle Y_2 \rangle = \langle X_2 Y_1 \rangle - \langle X_2 \rangle \langle Y_1 \rangle.$$

Where  $\langle \cdot \rangle$  refers to the long-term time average of a given quantity. We can now formulate an equation for the covariance,  $\bar{C}_{II}$  if we make a moment closure approximation and ignore the third order cumulants (Renshaw 1991; Keeling 2000).

$$\begin{aligned} \frac{d}{dt} \bar{C}_{II} &= 2\langle \beta(1 - \sigma)I_1 S_1 I_2 / N + \beta \sigma I_2 S_1 I_2 / N - gI_1 I_2 - dI_1 I_2 \rangle \\ &\quad - 2I_2 \langle \beta(1 - \sigma)I_1 S_1 / N + \beta \sigma I_2 S_1 / N - gI_1 - dI_1 \rangle \\ &= 2\beta(1 - \sigma)S \bar{C}_{II} / N + 2\beta \sigma S C_{II} / N - 2g \bar{C}_{II} \\ &\quad - 2d \bar{C}_{II} + 2\beta I \bar{C}_{SI} / N. \end{aligned}$$

If we make the simplifying assumption that  $I \bar{C}_{SI} / N$  is small, the differential equation has the fixed point,

$$\bar{C}_{II}^* = \frac{\beta \sigma S^* C_{II}^*}{gN + dN - \beta(1 - \sigma)S^*}.$$

From the equation for the number of infectious individuals, we find that

$$S^* = \frac{gN + dN}{\beta} - \frac{C_{SI}^*}{I}$$

and therefore obtain the following equation for the correlation,

$$C = \frac{\bar{C}_{II}^*}{C_{II}^*} = \frac{\sigma}{\frac{\beta C_{SI}^*}{gI^* N + dI^* N - \beta C_{SI}^*} + \sigma}. \quad (9)$$

For simplicity we set,

$$\xi = \frac{\beta C_{SI}^*}{gI^* N + dI^* N - \beta C_{SI}^*} \quad (10)$$

and calculate the value of  $\xi$  from numerical simulations.