PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Studying and approximating spatio–temporal models for epidemic spread and control

J. A. N. Filipe and G. J. Gibson

Phil. Trans. R. Soc. Lond. B 1998 **353**, 2153-2162
doi: 10.1098/rstb.1998.0354

THE ROYAL SOCIETY

Studying and approximating spatio-temporal models for epidemic spread and control

J. A. N. Filipe* **and G. J. Gibson**

Biomathematics & Statistics Scotland, The King's Buildings (JCMB), Edinburgh EH9 3JZ, UK

A class of simple spatio-temporal stochastic models for the spread and control of plant disease is investigated. We consider a lattice-based susceptible–infected model in which the infection of a host occurs through two distinct processes: a background infective challenge representing primary infection from external sources, and a short-range interaction representing the secondary infection of susceptibles by infectives within the population. Recent data-modelling studies have suggested that the above model may describe the spread of aphid-borne virus diseases in orchards. In addition, we extend the model to represent the effects of different control strategies involving replantation (or recovery). The Contact Process is a particular case of this model. The behaviour of the model has been studied using cellular-automata simulations. An alternative approach is to formulate a set of deterministic differential equations that captures the essential dynamics of the stochastic system. Approximate solutions to this set of equations, describing the time evolution over the whole parameter range, have been obtained using the pairwise approximation (PA) as well as the most commonly used mean-field approximation (MF). Comparison with simulation results shows that PA is significantly superior to MF, predicting accurately both transient and long-run, stationary behaviour over relevant parts of the parameter space. The conditions for the validity of the approximations to the present model and extensions thereof are discussed.

Keywords: spatio-temporal stochastic models; primary and secondary infection; epidemic control; Contact Process; cellular-automata simulation; pair approximation

1. INTRODUCTION

In recent years mathematical models have played a major role in contributing to our understanding of spatio-temporal processes in epidemiology, ecology or biological pattern formation. While deterministic models have been most often studied, there is also a long history of stochastic spatio-temporal modelling (e.g. Mollison 1977; Durrett & Levin 1994; Shaw 1994, 1995; Gibson 1997). Since extensive simulation of spatio-temporal stochastic models can be computationally demanding, a major challenge has been to characterize the behaviour of the models analytically or using approximation methods. For example, the pairwise approximation (PA), a member of the general family of cluster approximations, has been frequently used to study stochastic processes in physical systems (e.g. Dickman 1986; Nord & Evans 1985; Ben-Naim & Krapivsky 1994; Filipe & Rodgers 1995), and more recently its application has been extended to biological systems (e.g. Sato *et al.* 1994; Levin & Durrett 1997).

Many existing results on spatio-temporal stochastic models relate to long-term or large-scale behaviour. Theoretical questions studied for epidemic models include thresholds between persistence and extinction (e.g. Buttell *et al.* 1993), or the asymptotic shape and velocity of a spreading cluster (e.g. Cox & Durrett 1988). However, in practical applications, such as assessing strategies for controlling a plant disease epidemic, it may be important to understand system behaviour at smaller spatial and temporal scales and to consider transient as well as asymptotic dynamics. In particular, the development of small-scale spatial structure can significantly affect the progress of a disease. Also, in agricultural systems epidemics may not reach stationarity. Motivated by these considerations, we investigate the use of the simplest cluster approximations (mean-field (MF) and pairwise) to represent both transient and asymptotic behaviour of a stochastic spatio-temporal susceptible–infected model for the spread and control of a disease in a population. The model incorporates both primary infection from sources external to the population and the secondary infection of susceptibles by infected individuals within the population through local interactions. This choice is motivated by the recent study by Gibson (1997) which suggested that experimental observations of the spatio-temporal spread of citrus tristeza virus (CTV) disease, an aphid-borne virus disease of citrus trees, could be explained by a model of this kind. In addition, the model represents the effect of control measures. This paper also generalizes some aspects of work by Levin & Durrett (1997), who used cluster approximations to study asymptotic properties of the Contact Process (Harris 1974), a special case of the model considered here.

Following a description of the model in § 2, we derive the corresponding MF and PAs to it in § 3. In § 4 these

*Author for correspondence (joao@bioss.sari.ac.uk).

Received 29 October 1997 *Accepted* 17 December 1997

approximations are compared with each other and with simulation of the stochastic model. In particular we show that PA is in general significantly superior to MF in predicting both long-run and transient behaviour. Importantly, for many parameter values the PA predicts accurately the asymptotic and transient behaviour of the stochastic simulation model. Discussed in § 5.

2. SPATIO-TEMPORAL STOCHASTIC MODEL FOR THE SPREAD AND CONTROL OF AN EPIDEMIC

We consider a stochastic model for the transmission and control of a disease through a population of individuals located at the vertices of a regular square lattice, with a single individual at each vertex. Each individual, or site, may be in one of two states: susceptible or infected. The state of a site $\boldsymbol{x} = (i,j)$ at time t is given by $S_{\boldsymbol{x}}(t)$, where $S_{\boldsymbol{x}}(t) = 0$ if $\boldsymbol{x}$ is susceptible, and $S_{\boldsymbol{x}}(t) = 1$ if $\boldsymbol{x}$ is infected. We impose fixed boundaries on the lattice, so that almost every individual in the lattice has $z = 4$ nearest neighbours (NNs), except those on the boundary which have $z = 3$ or $z = 2$ NNs.

The dynamics of the model are specified as follows and involve four parameters: $\mathcal{J}_1$, $\mathcal{J}_2$, R, and $\alpha = 0, 1$. A susceptible individual acquires the disease through primary infection (background) at rate $\mathcal{J}_1$ per unit time and through secondary infection by infected neighbours at rate $\mathcal{J}_2\, n_{\boldsymbol{x}}(t)$, where $n_{\boldsymbol{x}}(t)$ denotes the number of NNs of $\boldsymbol{x}$ which are infected at time t. Disease control is effected through a continuous monitoring process such that if $S_{\boldsymbol{x}}(t) = 1$, then the probability that the infection at $\boldsymbol{x}$ is discovered in the time interval $(t, t + \mathrm{d}t)$ is equal to $R\,\mathrm{d}t$. The action taken immediately on discovery of an infection is determined by the value of α. If $\alpha = 0$, then the infected individual at $\boldsymbol{x}$ is replaced by a susceptible individual. If $\alpha = 1$ then, in addition to the replacement of $\boldsymbol{x}$, all infected NNs of $\boldsymbol{x}$ are replaced with susceptible individuals (NN recovery). (It is conceivable that α could vary between 0 and 1, in which case NN recovery would occur with probability α.) Thus infection and recovery, during a small time interval $(t,t + \mathrm{d}t)$, are governed by the conditional probabilities

$$\begin{aligned}\mathrm{Prob}[S_{\boldsymbol{x}}(t+\mathrm{d}t) = 1 | S_{\boldsymbol{x}}(t) = 0] &= \mathrm{Prob}(0 \to 1; \boldsymbol{x}, t)\\ &= [\mathcal{J}_1 + \mathcal{J}_2\, n_x(t)]\mathrm{d}t, \qquad (1)\\ \mathrm{Prob}[S_{\boldsymbol{x}}(t+\mathrm{d}t) = 0 | S_{\boldsymbol{x}}(t) = 1] &= \mathrm{Prob}(1 \to 0; \boldsymbol{x}, t)\\ &= R[1 + \alpha\, n_{\boldsymbol{x}}(t)]\mathrm{d}t. \qquad (2)\end{aligned}$$

This infection process is a limiting case of the more general model considered by Gibson (1997) in which secondary infection could occur through interactions beyond NNs. This replacement strategy is clearly idealized. Our principal aim is to investigate how well a spatially extended stochastic model might be represented by deterministic approximations rather than to provide accurate quantitative predictions of a real system. Nevertheless, the cases $\alpha = 0$ and $\alpha = 1$ are of some practical interest in that they represent control strategies which respectively ignore and take some account of spatial correlation in a pattern of disease. We note that the contact process (Harris 1974) is a special case of the above model, corresponding to $\mathcal{J}_1 = 0$ and $\alpha = 0$.

As the time units are arbitrary there are only three independent parameters in the model, say, $\mathcal{J}_1/R$, $\mathcal{J}_2/R$ and α. Therefore, without loss of generality we set $R = 1$ throughout so that the expected time elapsed between infection and recovery of a given site, in the absence of NN recovery, is one unit of time.

Given $\mathcal{J}_1$, $\mathcal{J}_2$ and α we simulate a realization of the model by specifying the nature, location and times of the infection or replacement events which occur as follows. Suppose at time t we know the state of each site in the lattice. First, the time until the next event, Δt, is simulated from an exponential distribution with mean $1/[Rn_1 + (\mathcal{J}_1 n_0 + \mathcal{J}_2 n_{01})]$, where n_0 and n_1, respectively, represent the total number of susceptible and infected individuals in the lattice and n_{01} denotes the number of NN pairs in the lattice for which one individual is infected and the other susceptible. The current time t is incremented by Δt.

Next the nature of the event occurring at $t + \Delta t$ is determined. With probability $Rn_1/[Rn_1 + (\mathcal{J}_1 n_0 + \mathcal{J}_2 n_{01})]$ this is selected to be the discovery of an infected site and the particular site is chosen randomly from the set of infected sites. The site selected then becomes susceptible and, if $\alpha = 1$, so do any of its NNs which is infected. If the selected event is the infection of a susceptible site, then the site, $\boldsymbol{x}$, is selected from the set of all susceptible sites with probability proportional to the infective challenge experienced by $\boldsymbol{x}$, i.e. $\mathcal{J}_1 + \mathcal{J}_2 n_{\boldsymbol{x}}(t)$, and its status changes to infected.

By repeating this procedure from a set of initial conditions we generate an explicit realization of the stochastic model. For a large lattice the number of events which must be simulated to investigate, for example, the equilibrium behaviour of the model is correspondingly large. Hence, although computer implementation of the model is relatively straightforward to obtain, it may be of limited use in a comprehensive study. In the following section we consider how approximate representations of model behaviour can be obtained.

3. DETERMINISTIC APPROXIMATIONS TO SPATIO-TEMPORAL STOCHASTIC MODELS

To derive a simpler model to represent the behaviour of the stochastic model defined by (1) and (2) we assume that the lattice is large and that boundary effects can be ignored. Furthermore, we assume that the distribution of infected individuals at $t = 0$ arises from a process that is spatially stationary and appears uncorrelated over distances considerably less than the lattice size. Since the dynamic rules governing the system are spatially invariant, isotropic and involve no interactions beyond NNs, we assume that the system remains spatially stationary at large-enough scales of observation, i.e. that it exhibits no long-range correlations at all subsequent times. As a result, ensemble averages over many realizations of the process, denoted by $\langle \ldots \rangle$, should coincide with spatial averages of a quantity for a single realization.

In particular, we consider the densities $P[1](t) \equiv \langle S_{\boldsymbol{x}}(t) \rangle$, and $P[11](t) \equiv \langle S_{\boldsymbol{x}}(t) S_{\boldsymbol{x}+\boldsymbol{u}}(t) \rangle$, which denote, respectively, the probability that $\boldsymbol{x}$ is infected at time t and the probability that both $\boldsymbol{x}$ and its NN at $\boldsymbol{x} + \boldsymbol{u}$ are infected at time t. By the above arguments both these quantities

should be independent of $\boldsymbol{x}$ and the direction of the unit vector $\boldsymbol{u}$.

We now derive kinetic equations for the evolution of these densities with time. The change in $\langle S_{\boldsymbol{x}}\rangle$ during $(t, t+\mathrm{d}t)$ is given by

$$\begin{aligned}\mathrm{d}\langle S_{\boldsymbol{x}}(t)\rangle = &\langle (1-S_{\boldsymbol{x}}(t))\ \mathrm{Prob}(0\to 1, \boldsymbol{x}, t)\rangle \\ &- \langle S_{\boldsymbol{x}}(t)\ \mathrm{Prob}(1\to 0, \boldsymbol{x}, t)\rangle.\end{aligned} \tag{3}$$

Substituting the conditional probabilities defined in equations (1) and (2), and noting that $n_{\boldsymbol{x}}(t) = \sum_{\boldsymbol{u}} S_{\boldsymbol{x}+\boldsymbol{u}}(t)$, where the sum runs over the z NNs of $\boldsymbol{x}$, it follows that

$$\begin{aligned}\frac{\mathrm{d}\langle S_{\boldsymbol{x}}(t)\rangle}{\mathrm{d}t} = &\mathcal{J}_1 - (\mathcal{J}_1 + R)\langle S_{\boldsymbol{x}}\rangle + \mathcal{J}_2 \sum_{\boldsymbol{u}} \langle S_{\boldsymbol{x}+\boldsymbol{u}}\rangle \\ &- (\mathcal{J}_2 + R\alpha)\sum_{\boldsymbol{u}} \langle S_{\boldsymbol{x}} S_{\boldsymbol{x}+\boldsymbol{u}}\rangle.\end{aligned} \tag{4}$$

Using spatial invariance then gives the exact, deterministic equation

$$\frac{\mathrm{d}P[1]}{\mathrm{d}t} = \mathcal{J}_1 - (\mathcal{J}_1 + R - z\mathcal{J}_2)\ P[1] - (\mathcal{J}_2 + R\alpha) z\, P[11] \tag{5}$$

The corresponding equation for $P[11]$ is (see Appendix A)

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}P[11]}{\mathrm{d}t} = &(\mathcal{J}_1 + \mathcal{J}_2)\, P[1] - (\mathcal{J}_1 + \mathcal{J}_2 + R)\, P[11] \\ &+ \mathcal{J}_2 \sum_{\boldsymbol{v}'} \langle S_{\boldsymbol{x}}(1 - S_{\boldsymbol{y}}) S_{\boldsymbol{y}+\boldsymbol{v}'}\rangle - R\alpha \sum_{\boldsymbol{v}'} \langle S_{\boldsymbol{x}} S_{\boldsymbol{y}} S_{\boldsymbol{y}+\boldsymbol{v}'}\rangle,\end{aligned} \tag{6}$$

where $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{u}$ is a NN of $\boldsymbol{x}$, and the $\boldsymbol{v}'$-sum runs over the $z-1$ NNs of $\boldsymbol{y}$ distinct from $\boldsymbol{x}$.

Note that equation (5) involves $P[11]$, while equation (6) involves densities of triplets of adjacent sites. For example, on a square lattice

$$\sum_{\boldsymbol{v}'} \langle S_{\boldsymbol{x}} S_{\boldsymbol{y}} S_{\boldsymbol{y}+\boldsymbol{v}'}\rangle = P[111] + 2\, P\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix},$$

the first term referring to collinear triplets and the second to triplets in a right angle. In fact, equations (5) and (6) are the first in an infinite hierarchy of equations for the densities of consecutively larger clusters, with each equation involving terms relating to higher order clusters so that any subsystem is not closed, a direct consequence of the interactions in the system.

Cluster approximations (Bethe 1935; Kikuchi 1951) can be used to approximate a system of equations such as (5) and (6) with a closed system. In this approach the full spatial correlations between individuals inside a basic cluster are incorporated but simplifying assumptions are made regarding correlations between individuals inside and outside a cluster. These simplifications allow densities of clusters larger than the basic cluster to be expressed in terms of densities of smaller clusters, thereby closing the system of equations. As long as the system's correlation length (Yeomans 1992; Stauffer & Aharony 1992) remains finite it is in principle possible to improve the accuracy of the approximation by increasing the size of the basic cluster (Burley 1972; Nord & Evans 1985). However, as the cluster size increases, the approach becomes increasingly convoluted and there is no *a priori* indication of the degree of the resulting improvement. Nevertheless, low-order cluster approximations are relatively straightforward to implement and have proved to yield surprisingly good descriptions, at least in the absence of large-scale, critical phenomena. (For a review of cluster approximations applied to systems in thermodynamic equilibrium see, for example, Burley (1972).)

Before describing the construction of the two simplest cluster-approximations, MF and PA, for the system of equations (5) and (6), we analyse the equations in a soluble limit. We consider the low-infection regime where a few infectives are scattered in a sea of susceptibles, i.e. $P[1] \ll 1$ and $P[11] \ll P[1]$. In this case all but the first terms on the right-hand side of equations (5) and (6) are small and can be neglected. Not surprisingly the background infection process dominates over all others, as we would expect in the early stages of pathogen invasion of a healthy field. Starting with zero density of infectives we find, to leading order,

$$P[1](t) = \mathcal{J}_1\, t + \ldots, \quad (t \ll 1), \tag{7}$$

$$P[11](t) = \mathcal{J}_1\, (\mathcal{J}_1 + \mathcal{J}_2)\, t^2 + \ldots, \quad (t \ll 1), \tag{8}$$

which implies that

$$\lim_{t\to 0} \frac{P[11]}{P[1]^2} = [1 + \mathcal{J}_2/\mathcal{J}_1]. \tag{9}$$

This result indicates that even in this non-structured regime there are correlations between NN infectives as a consequence of the small, but non-vanishing effect of NN transmission. The validity of equations (7) and (8) requires that the time t since the beginning of the epidemic is small. For fixed parameters $\mathcal{J}_1$ and $\mathcal{J}_2$, the necessary condition $P[11] \ll P[1]$ implies the upper bound $t \ll 1/(\mathcal{J}_1 + \mathcal{J}_2)$.

(a) *Mean-field approach*

Due to its simplicity, mean-field (MF), the lowest order, has been the most widely used cluster approximation in spatial problems. Within the MF approximation the basic cluster is a single site, so we assume that the states of any two sites in the lattice are independent of each other. This implies that

$$P[11] = P[1]^2. \tag{10}$$

The resulting lack of consistency with equation (6) is resolved by simply ignoring the equation.

Using equation (10), equation (5) can be recast as

$$\frac{\mathrm{d}P[1]}{\mathrm{d}t} = a_0 - a_1\, P[1] - a_2\, P[1]^2, \tag{11}$$

with

$$a_0 = \mathcal{J}_1, \quad a_1 = \mathcal{J}_1 + R - z\mathcal{J}_2, \quad a_2 = (\mathcal{J}_2 + R\alpha)z. \tag{12}$$

This is a particular case of Riccati's equation (see, for example, Stephenson & Radmore 1990), with full transient solution

$$P[1](t) = \frac{1}{2a_2}[\gamma \tanh(\gamma\, t/2 + C) - a_1], \tag{13}$$

with

$$\gamma = \sqrt{a_1^2 + 4a_0a_2} = \sqrt{(\mathcal{J}_1 + R + z\mathcal{J}_2)^2 + 4\mathcal{J}_1 R\alpha z}, \quad (14)$$

$$C = \tanh^{-1}\left(\frac{2a_0\ P_0 + a_1}{\gamma}\right), \quad (15)$$

and $P_0 = P[1](0)$ the initial density of infectives. The asymptotic equilibrium level of infection

$$P[1]_\infty = \frac{\sqrt{a_1^2 + 4a_0a_2} - a_1}{2a_2} \quad (16)$$

is obtained on taking the limit $t \to \infty$ in equation (13). It is easily verified that $P[1]_\infty \leqslant 1$ for all parameter values.

The MF solution reveals that the effect of NN recovery ($\alpha = 1$) is quantitative: it reduces the fraction of infected individuals at any time. However, the effect of background infection ($\mathcal{J}_1 > 0$) is more dramatic: it prevents the system from undergoing a phase transition from a parameter region where the epidemic persists to another where it dies out. In the limiting case where $a_0 = \mathcal{J}_1 = 0$, the solution to equation (11) is

$$P[1](t) = \frac{P_0}{\exp(a_1 t) + \mathrm{sgn}(a_1)[\exp(a_1 t) - 1]\ P_0/P[1]_\infty}, \quad (17)$$

with the equilibrium limit

$$P[1]_\infty = \begin{cases} 0, & R \geqslant z\mathcal{J}_2 \\ |a_1|/a_2 = (\mathcal{J}_2 - R/z)/(\mathcal{J}_2 + R\alpha), & R < z\mathcal{J}_2 \end{cases}. \quad (18)$$

While the MF prediction for the critical threshold value of $(\mathcal{J}_2/R)$ is independent of α,

$$(\mathcal{J}_2/R)_c = 1/z, \quad (19)$$

i.e. it is the same as for the MF approximation to the Contact Process ($\alpha = 0$), the asymptotic infection level above the threshold is reduced as a result of NN recovery.

Derivations of threshold conditions for disease persistence in non-spatial deterministic models representing control measures have been carried out by several authors (e.g. Jeger & van den Bosch 1994*a*,*b*; van den Bosch & Roos 1996).

(b) *Pairwise approach*

Previous examples of pairwise approximation (PA) of epidemic models are provided by Sato *et al.* (1994) and Levin & Durrett (1997). Here the basic cluster is a pair of NN sites. The key assumption in the derivation of the approximation is that for a given site, conditional on its state, the states of any two of its NNs are independent. For example, if x, y and z denote the states of three adjacent sites this assumption allows us to write

$$P[xyz] = P[xy]P[z|xy] \simeq P[xy]P[z|y] = P[xy]P[yz]/P[y]. \quad (20)$$

We can therefore make the following substitutions for triplet densities in equation (6):

$$\sum_{v'} \langle S_x(1 - S_y)S_{y+v}\rangle = (z-1)\frac{P[10]^2}{P[0]} \quad (21)$$

Table 1. *Critical thresholds*

$(\mathcal{J}_2/R)_c$	$\alpha = 0$	$\alpha = 0.5$	$\alpha = 1$
MF	0.25	0.25	0.25
pairwise	0.333 ...	0.533 ...	0.757 ...
simulation	0.412 ...	0.837 ...	1.285 ...

$$\sum_{v'} \langle S_x S_y S_{y+v}\rangle = (z-1)\frac{P[11]^2}{P[1]}. \quad (22)$$

These lead directly to the PA given by the set of equations (equation (5) is unchanged)

$$\frac{\mathrm{d}P[1]}{\mathrm{d}t} = \mathcal{J}_1 - (\mathcal{J}_1 + R - z\mathcal{J}_2)\,P[1] - (\mathcal{J}_2 + R\alpha)zP[11] \quad (23)$$

$$\frac{1}{2}\frac{\mathrm{d}P[11]}{\mathrm{d}t} = (\mathcal{J}_1 + \mathcal{J}_2)P[1] - (\mathcal{J}_1 + \mathcal{J}_2 + R)P[11] + \mathcal{J}_2(z-1)\frac{(P[1]-P[11])^2}{1-P[1]} - R\alpha(z-1)\frac{P[11]^2}{P[1]}. \quad (24)$$

It is a straightforward task to solve these nonlinear equations numerically for given initial conditions. The equilibrium level of infection can be derived from equations (23) and (24), by setting their right-hand sides to zero, recasting equation (23) as

$$P[11]_\infty = \frac{\mathcal{J}_1 - (\mathcal{J}_1 + R - \mathcal{J}_2 z)P[1]_\infty}{(\mathcal{J}_2 + R\alpha)z}, \quad (25)$$

and substituting into equation (24) to give a cubic equation in $P[1]_\infty$, which can then be obtained as the real solution satisfying $0 \leqslant P[1] \leqslant 1$. (We have checked that there is only one such solution.) Although this solution can be obtained analytically, the resulting expression is convoluted and not given here.

Like the MF approximation, in the absence of background infection ($\mathcal{J}_1 = 0$) the PA displays critical behaviour (Yeomans 1992; Stauffer & Aharony 1992). In this case the resulting cubic can be solved more easily to give

$$P[1]_\infty = \frac{(z-1)\alpha(z-\lambda)^2 + z(1+\alpha\lambda)[z(1-\alpha) - (1+\lambda)]}{(1+\alpha\lambda)[z^2(z-2)\alpha + z(z-1) - \lambda]}, \quad \text{for} \quad \mathcal{J}_2/R > (\mathcal{J}_2/R)_c, \quad (26)$$

where $\lambda = R/\mathcal{J}_2$, and $P[1]_\infty = 0$ otherwise, with (see table 1)

$$(\mathcal{J}_2/R)_c = \begin{cases} 1/(z-1) & \text{for } \alpha = 0 \\ 2\alpha/\{(1+z\alpha)\sqrt{z[z(1+\alpha)^2 - 4\alpha]} \\ \quad -z[z\alpha(1+\alpha) + 1 - \alpha]\} & \text{for } \alpha > 0. \end{cases} \quad (27)$$

The threshold when $\alpha = 0$ is the same as for percolation on a tree (Stauffer & Aharony 1992).

4. COMPARISON OF APPROXIMATIONS WITH STOCHASTIC MODEL

We now investigate the ability of the MF and pairwise approximations to capture the behaviour of the full

stochastic spatio-temporal model over a range of parameter values. First we consider the asymptotic value of $P[1]$ as predicted by simulation and by the MF and PA from equations (11), (23) and (24). Later in the section, the ability of the approximations to capture transient behaviour is investigated.

The simulations were carried out on a two-dimensional system with a square boundary (fixed boundary conditions, FBC). Lattice sizes ranging from $N = 50^2$ to 300^2 were used to probe any systematic dependence of the results on the system size. As expected, the fluctuations in the transient curves as well as differences between process realizations were found to decrease with n. The size of the system considered depended on the quantities being monitored and on the parameter values input; for example, in the vicinity of a critical point (Yeomans 1992; Stauffer & Aharony 1992) we used larger systems to lower the probability of extinction.

We also compared the results with simulations on a torus (periodic boundary conditions, PBC). So far as asymptotic disease levels are concerned, edge effects were found to be significant for $N = 100^2$, but were negligible for $N = 200^2$. There was, however, a slight delay in the disease build-up with FBC that appeared to persist even in the larger systems considered by us. For example, we compared random realizations of the system with different boundary conditions and with $N = 200^2$ and $\mathcal{J}_1 = 0$, in the cases (i) $\mathcal{J}_2 = 1$, $r = 0$ ($P[1]_\infty = 0.72$, as in figure 1*a*), and (ii) $\mathcal{J}_2 = 2$, $r = 1$ ($P[1]_\infty = 0.38$, as in figure 1*d*), and observed a tendency during the transient period for the disease level in the FBC system to be lower than that in the PBC system. The maximum relative difference between disease levels for the FBC and PBC systems during the transient period was typically around 5% for parameter sets (i) and (ii).

The following parameter values were used in the simulations: $R = 1$, $\mathcal{J}_1 = 0$, 0.001, 0.01, 0.1; $\mathcal{J}_2 = 0.1, \ldots, 5$, $\alpha = 0, 1$.

(a) *Stationary behaviour*

To measure the equilibrium level of infection, we increased the system size as necessary to observe a convincing stationary behaviour. The system was then simulated in equilibrium for a large number of time steps (up to 150×300^2) and the resulting series of values of $P[1]$ was averaged over the total series and over blocks of five samples. A check on stationarity was provided by the requirement of small variations between successive block averages.

Figure 1 shows the equilibrium level of infection $P[1]_\infty$ as a function of $\mathcal{J}_2$, for different values of $\mathcal{J}_1$ and α, as predicted by simulation and the MF and pairwise approximations. Both approximations overestimate the simulated equilibrium values. Since both approximations neglect, to some extent, spatial correlation in the distribution of disease, they are expected to overestimate the degree of mixing of susceptibles and infectives and hence the effectiveness of epidemic spread. The PA, which models explicitly the number of 'reaction pairs' 1–0, is always in better quantitative agreement with simulation and provides accurate predictions over a wider parameter range than MF.

Some general trends can be discerned from these results. When $\alpha = 0$, the PA provides a good prediction of the simulation results over much of the parameter space. It is generally less effective when $\alpha = 1$, but accurate predictions of simulations are nevertheless obtained for some parameter values. However, MF approximations are particularly poor when $\alpha = 1$. For a given value of the asymptotic infection level in the stochastic model the disparity between approximations and simulation becomes greater as (i) $\mathcal{J}_2$ increases and $\mathcal{J}_1$ decreases; and (ii) α is increased from 0 to 1. Thus, in general, the agreement diminishes as the strength of the local interactions in either the infection or recovery processes increases. Note the drop in the disease level (e.g. about 50% for $\mathcal{J}_2 = 2$) when the control strategy corresponding to $\alpha = 1$ is applied.

When $\mathcal{J}_1 = 0$ there is a critical threshold value of $\mathcal{J}_2$ separating regions of epidemic persistence and extinction (figures 1*a* and *b*) and the performance of the approximations is particularly poor near this point. This is to be expected because long-range spatial correlations (as in figure 4) develop in the system at the transition (Stauffer & Aharony 1992), while the pairwise approach assumes exponential decay of pair correlations. Table 1 shows the thresholds values $(\mathcal{J}_2/R)_c$ from theory and simulation for $\alpha = 0, 1$ (and $\alpha = 1/2$, for comparison). The considerable mismatch between threshold values is not unexpected, as argued, but once again the pairwise approach yields a considerable improvement over the MF prediction, which does not even depend on α. The simulation threshold for the Contact Process ($\alpha = 0$) agrees well with known values reported in the literature (see, for example, Buttell *et al.* 1993; the authors use the notation $\beta_c = z(\mathcal{J}_2/R)_c$). The corresponding pairwise prediction also agrees with recent results (Levin & Durrett 1997).

(b) *Transient behaviour*

To investigate the transient behaviour of the model and the ability of the approximations to capture it, we consider first the evolution of the process on a lattice of 100×100 sites, with parameters $R = 1$, $\mathcal{J}_1 = 0.01$ and $\mathcal{J}_2 = 0.25$, and with $\alpha = 0$ (figure 2*a*,*b*) and $\alpha = 1$ (figure 2*c*,*d*). This scenario represents a situation where the strength of the NN interaction is quite large relative to the background process and leads to a stationary level of disease under 5% for both values of α. The results of a study where related models were fitted to observations of virus spread in orchards (Gibson 1997) suggest that for some real systems the strength of the local interactions may be considerably lower. Nevertheless, since the difficulty in capturing the dynamics increases with the strength of local interactions, this scenario is considered in order to better test the approximation technique.

Two different initial conditions were considered. In figure 2*a*,*c* the simulations were initiated with a small number of randomly located infections, in 0.1% of the hosts. This corresponds to a situation where an epidemic is controlled from the early stages of pathogen invasion from external sources. In figure 2*b*,*d* the system is initiated with a correlated pattern of disease with 20% of sites infected. This initial state was generated by iterating the stochastic model with recovery parameter R set to zero until 20% infection was attained, at which point the control was 'switched on'. This scenario corresponds to a situation where the decision to initiate disease control is

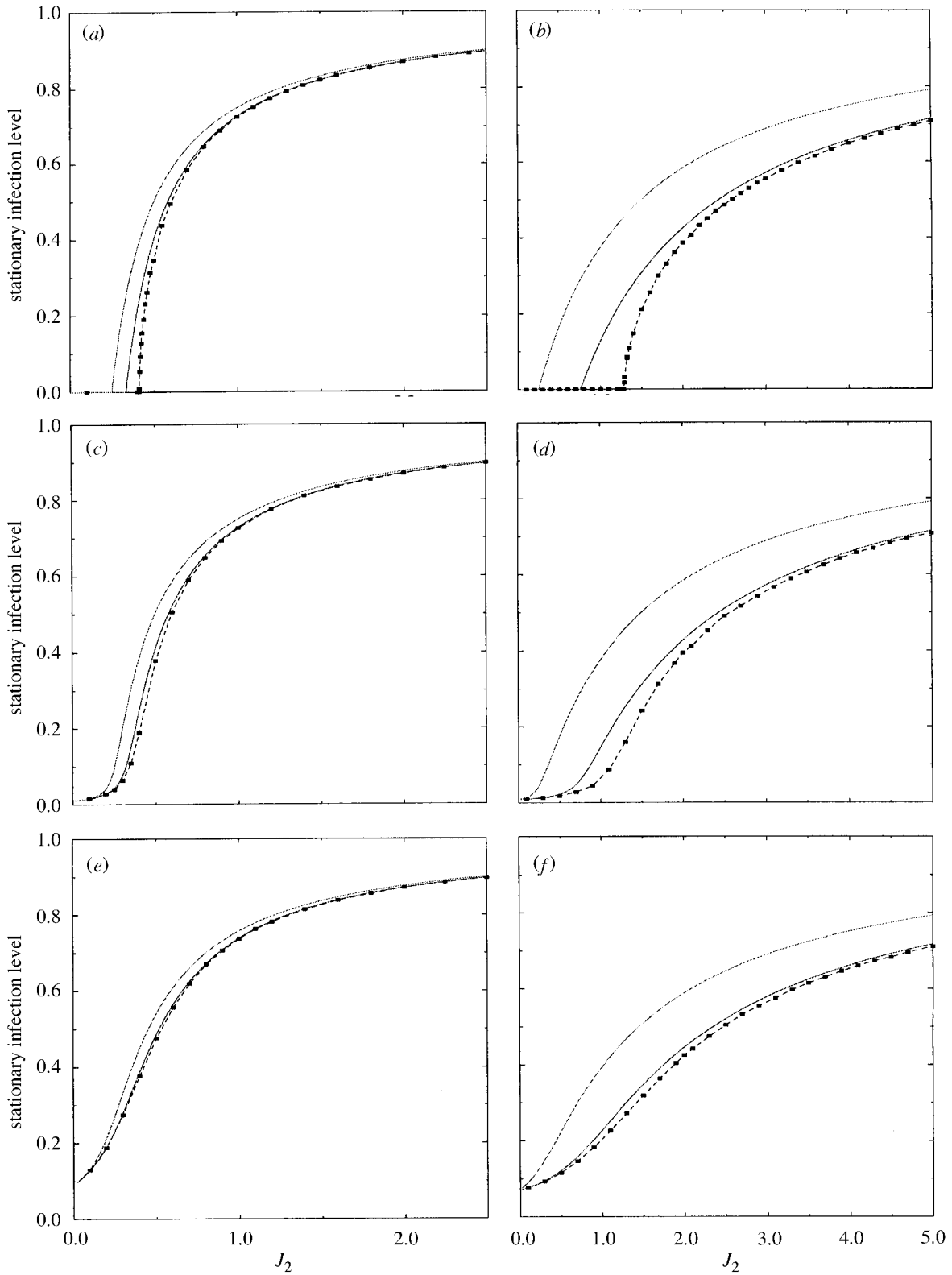

Figure 1. Phase diagram for $P[1]_\infty$, with $\alpha = 0$: (*a*) $\mathcal{J}_1 = 0$; (*c*) $\mathcal{J}_1 = 0.01$; (*e*) $\mathcal{J}_1 = 0.1$; and with $\alpha = 1$: (*b*) $\mathcal{J}_1 = 0$; (*d*) $\mathcal{J}_1 = 0.01$; (*f*) $\mathcal{J}_1 = 0.1$. Simulation (bottom curves), PA (intermediate curves) and MF (top curves).

based on disease levels becoming unacceptably high. The evolution towards the initial non-random pattern is included to illustrate the epidemic growth in the absence of recovery and the ability of the approximations to capture this aspect of model behaviour. Figure 2*a*,*b* ($\alpha = 0$) and 2*c*,*d* ($\alpha = 1$), display different ways by which the epidemic level $P[1]$ attains the same stationary controlled level, about 4% and 1%, respectively (cf. figure 1*c*,*f*). Note the much faster drop in disease level obtained using the NN-replacement strategy, $\alpha = 1$ (figure 2*d*). Comparison is made with MF and PA: the pairwise estimate of $P[1]$ follows the simulation very closely. The evolution of $P[11]$ is also shown (lower curves) with the PA accurately capturing its evolution for both α-values.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

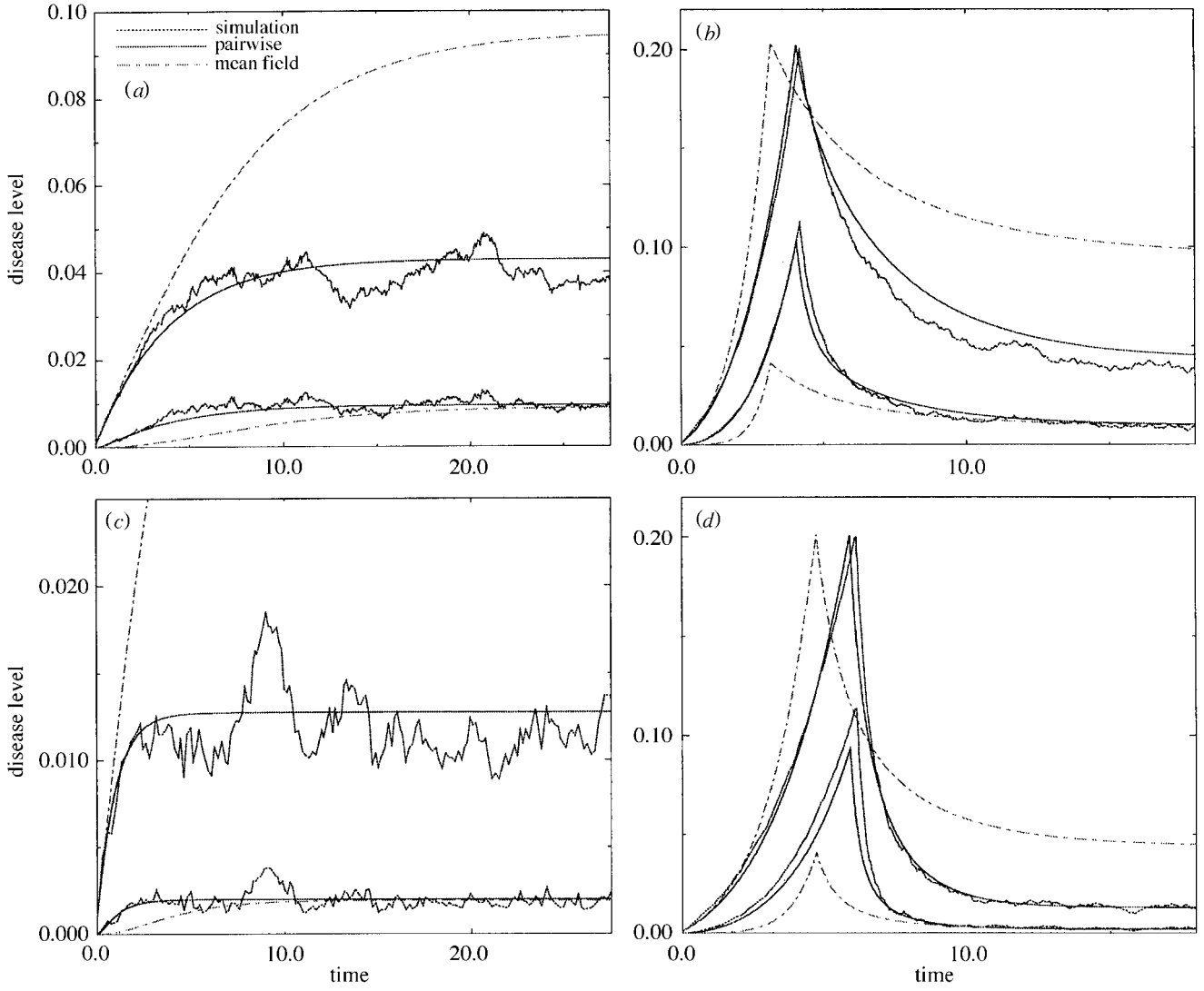

Figure 2. System's evolution: $P[1]$ (upper curves) and $P[11]$ (lower curves) with $\mathcal{J}_1 = 0.01$, $\mathcal{J}_2 = 0.25$ and random initial condition with $P_0 = 0.001$. (*a*) and (*b*) $\alpha = 0$, in (*b*) the infection level grows steadily up to $P[1] = 0.2$ in the absence of recovery ($R = 0$), after which R is set to 1 and the level stabilizes; (*c*) and (*d*) $\alpha = 1$, the same conditions as in (*a*) and (*b*).

A good agreement between the transient behaviour of the stochastic model and the PA is found over a considerable part of the parameter space. The approach has its limitations, though, so the agreement is not universal. For example, we compared figure 2*a*,*b* with the case where $\alpha = 1$ and a similar stationary level of infection is reached ($\mathcal{J}_2 = 0.8$). As NN interactions are over three times stronger the performance of the PA is poorer (figure 1*d*). Naturally, we expect infectives to be more clustered in the latter case. This is illustrated by the sequences of system snapshots in figure 3, the first of which is along the plot in figure 2*b*. A measure of patchiness is given by the correlation between NNs, $(P[11]-P[1]^2)/(P[1](1-P[1]))$, and plottings of this were consistent with the previous statement. In more extreme situations, i.e. close to the critical point, some assumptions of the pairwise approach breakdown completely. Figure 4 shows snapshots of the system in the stationary regime, with $(\mathcal{J}_1,\alpha,P[1]_\infty) = (0.001,1,1\%)$, $(0,0,4\%)$ and $(0,1,1\%)$. Although the level of disease is similar to figure 3, the patterns show structure characterized by length scales of the order or larger than the system size. Some of them ($\mathcal{J}_1 = 0$) give clear signs of self-similarity, typical of power-law decaying correlations.

5. DISCUSSION

For the class of models considered, the results show that the PA captures well the main features of both transient and asymptotic behaviour of the full stochastic spatio-temporal implementation. This is true for a large part of the parameter space, and in particular in parameter regions of practical interest. The MF approach, on the other hand, is seen to offer a generally poorer description, which illustrates the importance of adopting spatially explicit models for representing plant populations.

While we have interpreted the model within a specific epidemiological context, many other scenarios are of course possible. In particular, the processes of recovery have been thought of as part of a control strategy, but in other situations these might occur naturally and α may vary between 0 and 1. Furthermore, although the model studied is simple, it allows for local interactions of

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

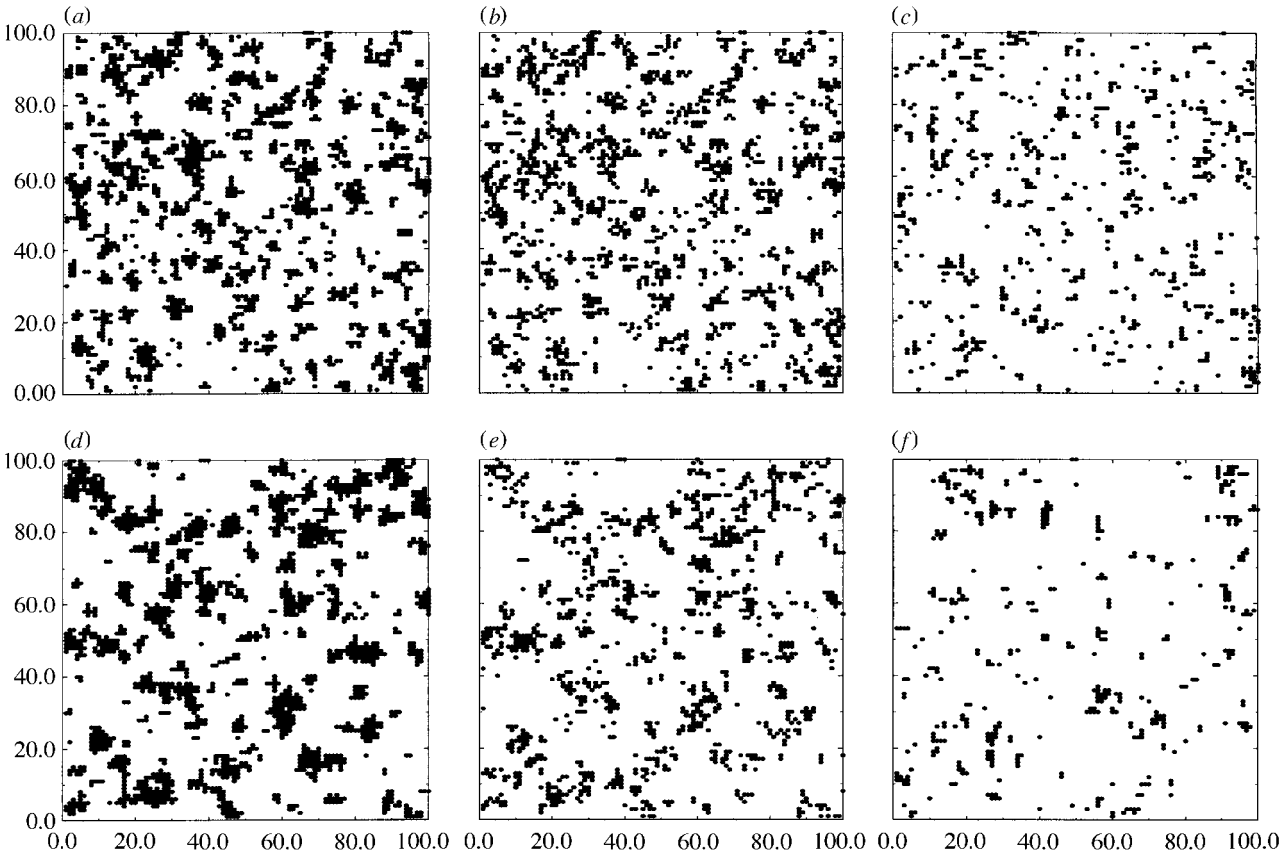

Figure 3. Snapshots of the system at three different time instants: along the evolution plot in figure 2*b* (*a*) $t = 4.37$; (*b*) $t = 4.80$; (*c*) $t = 8.10$, and with $\alpha = 1$ but a comparable level of infection ($\mathcal{J}_2 = 0.8$, $\mathcal{J}_1 = 0.01$) (*d*) $t = 2.17$; (*e*) $t = 2.48$; (*f*) $t = 5.65$). Dots represent infected hosts.

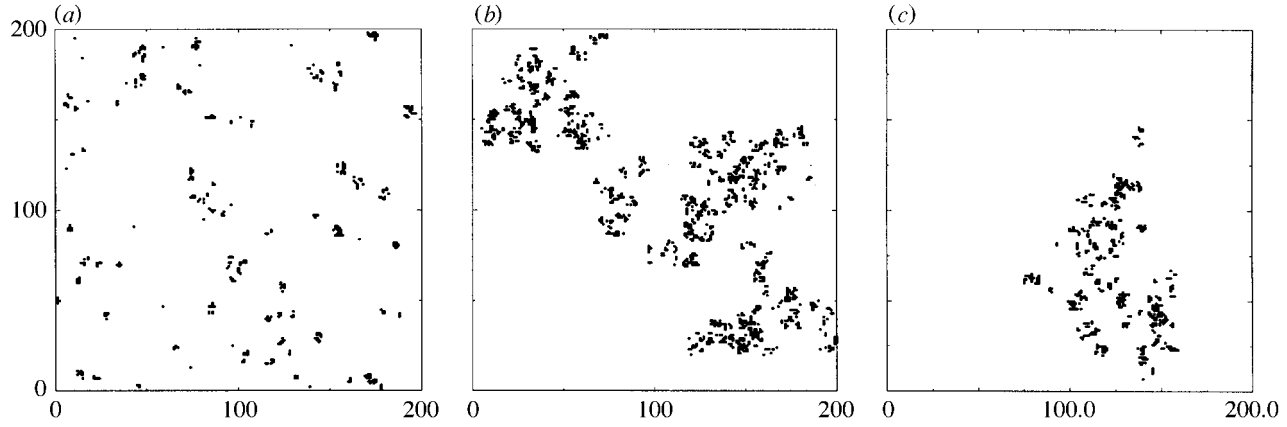

Figure 4. A configuration of the system in the stationary regime corresponding to different values of $(\mathcal{J}_1, \mathcal{J}_2, \alpha)$: (*a*) (0.001,1.01,1), $P[1]_\infty \simeq 1\%$, $t = 111.2$; (*b*) (0,0.417,0), $P[1]_\infty \simeq 4\%$, $t = 727.5$; (*c*) (0,1.287,1), $P[1]_\infty \simeq 1\%$, $t = 340.202$. Dots represent infected hosts.

arbitrary strength in both the disease spread and control processes. It is therefore sufficiently complex for the results to suggest a wide potential of cluster approximations for studying more complex processes in epidemiology and ecology. In fact, it is our belief that the addition of further local states and other features to the model would not, in essence, change the present conclusions as long as the interactions remain short-ranged.

In particular, our expectation of how model behaviour and approximation performance would vary by changing the range of interactions is as follows. The addition of finite-range interactions beyond NNs (e.g. decaying exponentially with distance) would increase the number of 'reaction pairs' at any time and reduce the heterogeneity of mixing between susceptibles and infectives. This would bring the system closer to the conditions assumed in the MF approach, so the MF prediction would become more accurate though still overestimating the level of infection. The PA prediction would still lie between these two bounds. While the location of the critical threshold will depend on the details of the interactions, critical behaviour should not be affected. In the

extreme case of interactions which are not exponentially bounded important modifications should occur in spatial structure (see, for example, Minogue (1989), Shaw (1995) for models without recovery), as well as in critical behaviour. Disease patterns might then display long-range correlations and, in the transient regime, resemble those observed near the critical threshold (figure 4).

There are several practical benefits arising from the use of cluster approximations and the efficient description of system behaviour offered by them. Even though the resulting system of differential equations may require numerical solution, the computational demands for this are orders of magnitude less than for simulation of the stochastic model. This difference may not be of great practical significance so far as investigating model behaviour for a single parameter set is concerned. However, in cases where stochastic models are fitted to experimental data (e.g. Gibson 1997) large uncertainty in parameter estimates is typically found. Predictions of system behaviour using the fitted model must take account of this uncertainty by considering a space of parameters. Exhaustive simulation of the full stochastic model over this space may not be feasible and the availability of deterministic approximations may be valuable. Moreover, there is a well-established body of mathematical methods, such as stability analysis, which can be applied to the deterministic cluster approximations to characterize their behaviour over parameter space. These methods can yield *a priori* insight into the general model behaviour and provide guidance on an appropriate range of parameter sets on which to base a full stochastic study. These advantages will be even more apparent when studying stochastic models of greater complexity than considered here.

We are currently investigating modifications to the PA method that may lead to more accurate predictions. One way to progress, is to note that the assumption made to close the system of equations (5) and (6) is not unique and that other assumptions may lead to different approximate solutions. For example, for $\alpha = 1$ (the case of worst performance, figure 1), it is possible to build an approach that underestimates the size of the infected population, thus providing us with both upper and lower bounds to the exact solution (Filipe 1998). Another possibility is based on redefining the basic cluster (e.g. Sato *et al*. 1994). Note that all terms in the sums (21) and (22) were assumed to be equal, while in general we expect the following inequalities

$$P[101] < P\begin{bmatrix}\mathbf{1} & 0\\ & \mathbf{1}\end{bmatrix}, \quad P[111] < P\begin{bmatrix}\mathbf{1} & 1\\ & \mathbf{1}\end{bmatrix}, \tag{28}$$

to hold between densities of triplets of NN sites. In fact, in the configurations on the right-hand side the extreme sites (bold faced) are closer to each other and should therefore be more strongly correlated. One attempt to refine the method may be to consider explicit equations for the densities of NN and next-nearest-neighbour pairs.

This research was supported by funding from the Scottish Office Agriculture, Environment and Fisheries Department. We thank Chris Glasbey, Elizabeth Austin and Milena Maule for valuable comments and suggestions on the manuscript.

APPENDIX A

In this appendix we derive equation (6). To obtain an expression for $\mathrm{d}\langle S_{\boldsymbol{x}} S_{\boldsymbol{y}}\rangle$, with $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{u}$, we note that the probability that a susceptible at $\boldsymbol{y}$ becomes diseased, given that one of its NNs is diseased, during $(t, t + \mathrm{d}t)$ is

$$\begin{aligned}\mathrm{Prob}[S_{\boldsymbol{x}}(t+\mathrm{d}t)\cdot S_{\boldsymbol{y}}(t+\mathrm{d}t) = 1 \mid S_{\boldsymbol{x}}(t) = 1, S_{\boldsymbol{y}}(t) = 0] \\ = \mathrm{Prob}(10 \to 11; \boldsymbol{x}, \boldsymbol{y}, t) \\ = [\mathcal{J}_1 + \mathcal{J}_2\,(1 + n'_{\boldsymbol{y}}(t))]\mathrm{d}t, \end{aligned} \tag{29}$$

and the probability that at least one of the hosts in a given pair of NN infectives is replaced by a susceptible during $(t, t + \mathrm{d}t)$ is

$$\begin{aligned}\mathrm{Prob}[S_{\boldsymbol{x}}(t+\mathrm{d}t)\cdot S_{\boldsymbol{y}}(t+\mathrm{d}t) = 0 \mid S_{\boldsymbol{x}}(t)\cdot S_{\boldsymbol{y}}(t) = 1] \\ = \mathrm{Prob}(11 \text{ changes}; \boldsymbol{x}, \boldsymbol{y}, t) \\ = R\,[2 + \alpha\, n'_{\boldsymbol{x}}(t) + \alpha\, n'_{\boldsymbol{y}}(t)]\mathrm{d}t, \end{aligned} \tag{30}$$

where $n'_{\boldsymbol{x}} = \sum_{\boldsymbol{v}' \neq \boldsymbol{u}} S_{\boldsymbol{x}+\boldsymbol{v}'}$ excludes $\boldsymbol{y}$, and conversely, the summation $n'_{\boldsymbol{y}} = \sum_{\boldsymbol{v}' \neq -\boldsymbol{u}} S_{\boldsymbol{x}+\boldsymbol{u}+\boldsymbol{v}'}$ excludes $\boldsymbol{x}$. Thus, the expression for the variation of $S_{\boldsymbol{x}} S_{\boldsymbol{y}}$, during $(t, t + \mathrm{d}t)$ reads

$$\begin{aligned}\mathrm{d}\langle S_{\boldsymbol{x}}(t)\, S_{\boldsymbol{y}}(t)\rangle = \langle S_{\boldsymbol{x}}(t)\,[1 - S_{\boldsymbol{y}}(t)]\mathrm{Prob}(10 \to 11; \boldsymbol{x}, \boldsymbol{y}, t)\rangle \\ + \langle [1 - S_{\boldsymbol{x}}(t)]\, S_{\boldsymbol{y}}(t)\mathrm{Prob}(01 \to 11; \boldsymbol{x}, \boldsymbol{y}, t)\rangle \\ - \langle S_{\boldsymbol{x}}(t)\, S_{\boldsymbol{y}}(t)\mathrm{Prob}(11 \text{ changes}; \boldsymbol{x}, \boldsymbol{y}, t)\rangle. \end{aligned} \tag{31}$$

Substituting the probability expressions into equation (31), and using the assumptions of spatial invariance, one finally obtains equation (6).

REFERENCES

Ben-Naim, E. & Krapivsky, P. L. 1994 Cluster approximation for the contact process. *J. Phys.* A **27**, L481–L487.

Bethe, H. A. 1935 Statistical theory of superlattices. *Proc. R. Soc. Lond.* A **150**, 552–575.

Burley, D. M. 1972 Closed form approximations for lattice systems. In *Phase transitions and critical phenomena*, vol. 2 (ed. C. Domb & M. S. Green), pp. 329–374. London: Academic Press.

Buttell, L., Cox, J. T. & Durrett, R. 1993 Estimating the critical values of stochastic growth models. *J. Appl. Prob.* **30**, 455–461.

Cox, J. T. & Durrett, R. 1988 Limit theorems for the spread of epidemics and forest fires. *Stoch. Proc. Appl.* **30**, 171–191.

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

Durrett, R. & Levin, S. A. 1994 Stochastic spatial models—a user's guide to ecological applications. *Phil. Trans. R. Soc. Lond.* B **343**, 329–350.

Filipe, J. A. N. 1998 Hybrid closure-approximation to epidemic models. *Physica* A. (In the press.)

Filipe, J. A. N. & Rodgers, G. J. 1995 Theoretical and numerical studies of chemisorption on a line with precursor layer diffusion. *Phys. Rev.* E **52**, 6044–6054.

Gibson, G. J. 1997 Investigating mechanisms of spatio-temporal epidemic spread using stochastic models. *Phytopathology* **87**, 139.

Harris, T. E. 1974 Contact interactions on a lattice. *Ann. Prob.* **2**, 969.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Jeger, M. J. & van den Bosch, F. 1994*a* Threshold criteria for model-plant disease epidemics. 1. Asymptotic results. *Phytopathology* **84**, 24–27.

Jeger, M. J. & van den Bosch, F. 1994*b* Threshold criteria for model-plant disease epidemics. 2. Persistence and endemicity. *Phytopathology* **84**, 28–30.

Kikuchi, R. 1951 Theory of cooperative phenomena. *Phys. Rev.* **81**, 988–1003.

Levin, S. A. & Durrett, R. 1997 From individuals to epidemics. *Phil. Trans. R. Soc. Lond.* B **351**, 1615–1621.

Minogue, K. P. 1989 Diffusion and spatial probability models for disease spread. In *Spatial components of plant disease epidemics* (ed. M. J. Jeger), pp. 127–143. Englewood Cliffs, NJ: Prentice Hall.

Mollison, D. 1977 Spatial contact models for ecological and epidemic spread. *J. R. Statist. Soc.* B **39**, 283–326.

Nord, R. S. & Evans, J. W. 1985 Irreversible immobile random adsorption of dimmers and trimers on 2D lattices. *J. Chem. Phys.* **82**, 2795–2810.

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

Shaw, M. 1994 Modelling stochastic processes in plant population. *A. Rev. Phytopathol.* **32**, 523–544.

Shaw, M. 1995 Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance. *Proc. R. Soc. Lond.* B **259**, 243–248.

Stauffer, D. & Aharony, A. 1992 *Introduction to percolation theory.* London: Taylor & Francis.

Stephenson, G. & Radmore, P. M. 1990 *Advanced mathematical methods for engineering and science students*. Cambridge University Press.

van den Bosch, F. & de Roos, A. M. 1996 The dynamics of infectious diseases in orchards with roguing and replanting as control strategy. *J. Math. Biol.* **35**, 129–157.

Yeomans, J. M. 1992 *Statistical mechanics of phase transitions*. Oxford: Clarendon Press.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES