Applications of percolation theory to fungal spread with synergy

Jonathan J. Ludlam (Churchill College, Cambridge, CB3 0DS, UK), Gavin J. Gibson (Maxwell Institute for Mathematical Sciences School of Mathematical and Computer Sciences, Heriot-Watt University, Riccarton Edinburgh, EH14 4AS, UK), Wilfred Otten (University of Abertay Dundee, SIMBIOS, Kydd Building, Dundee, DD1 1HG, UK), and Christopher A. Gilligan[1] (Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge, CB2 3EA, UK)

Abstract

There is increasing interest in the use of the percolation paradigm to analyze and predict the progress of disease spreading in spatially-structured populations of animals and plants. The wider utility of the approach has been limited, however, by several restrictive assumptions, foremost of which is a strict requirement for simple nearest-neighbour transmission, in which the disease history of an individual is influenced only by that of its neighbours. In a recent paper the percolation paradigm has been generalised to incorporate synergistic interactions in host infectivity and susceptibility and the impact of these interactions on the invasive dynamics of an epidemic has been demonstrated. In the current paper we elicit evidence that such synergistic interactions may underlie transmission dynamics in real-world systems by first formulating a model for the spread of a ubiquitous parasitic and saprotrophic fungus through replicated populations of nutrient sites and subsequently fitting and testing the model using data from experimental microcosms. Using Bayesian computational methods for model fitting, we demonstrate that synergistic interactions are necessary to explain the dynamics observed in the replicate experiments. The broader implications of this work in identifying disease control strategies that deflect epidemics from invasive to non-invasive regimes are discussed.

Keywords: Percolation, synergistic interactions, saprotrophic invasion, microcosm experiments, Bayesian inference, Markov chain methods.

---

[1]**Corresponding author** Telephone - +44 1223 333904, email - cag1@cus.cam.ac.uk.

1 INTRODUCTION

The statistical analysis of observations of spatial disease spread using contact distribution models and percolation methods has been recognised as an extremely useful approach for characterizing transmission mechanisms since it was first applied to virus diseases of citrus [1]. Standard epidemiological models such as SIR cellular automata can be conveniently considered as percolation processes [2, 3], with concepts such as the percolation threshold being analogous to the threshold between invasive and non-invasive behaviour [3]. This enables the properties of large-scale phenomena, in this case an epidemic, to be related to the interactions among individuals in the population, or vice-versa. Work by Gibson et al. [4] has advanced this to the stage of obtaining estimates of the percolation parameters using Markov chain Monte Carlo (MCMC) methods on experimental data, comprising successive spatial snap-shots of infected and susceptible individuals in plant populations. The use of the percolation paradigm for epidemiological spread has been advocated for an increasingly wide range of host-pathogen systems. Hence, previous work has adapted the paradigm not only to the spread of pathogenic fungi through populations of live hosts in soil [5] but also to the spread of saprotrophic fungi through fragments of organic matter in soil [6]. Within animal populations, the percolation paradigm has recently been used to analyse the spread of plague in gerbil populations in Kazakstan [7]. There are, however, certain assumptions that may restrict the application of the standard percolation paradigm to epidemiology. Foremost amongst these is the assumption that the challenge presented by an infected to a susceptible site does not depend upon the state of any other neighbours of these sites. This would suggest that multiple challenges to a susceptible site from one, two or more neighbouring infected sites are mutually independent. Other assumptions include the absence of a latent period, whereby infected sites are immediately infectious, and restrictions to nearest-neighbour transmission of infection. Recently the need for the first assumption has been removed thanks to a generalisation of the percolation paradigm [8] that includes synergistic interactions. Here, we focus on methods of inference for this more general model and on applying it directly in the analysis of experimental data to provide the first demonstration of the potential presence of synergy in a real system.

We use a tractable experimental model system and a particular type of synergy that relates to fungal infection in soil-borne plant pathogens [9]. For many soil-borne fungi, infection occurs by mycelial growth from infected to susceptible sites. Translocation of nutrients along mycelium to hyphal tips is known to occur for many fungi [10, 11, 12] and hence the connectivity of infected sites may be assumed to influence the probability of infection of newly encountered susceptible sites. In such a model, the infection process is therefore assumed to depend upon the number of other infectious individuals connected to the infectious site (Fig. 1a). We introduce a method for estimation of the parameters defining the infectious properties of individuals in the presence of synergy by analyzing the large-scale dynamics of an epidemic. Our experimental methods also apply to saprotrophic dynamics involving soil-borne fungi that spread by mycelial extension through populations of organic fragments and which are also known to exhibit translocation and synergy [13]. We have shown previously that the epidemiological paradigm can be used to analyze the spatial and temporal dynamics of saprotrophic colonization by treating uncolonized, colonized and nutrient-exhausted sites as the equivalents of susceptible, infected and removed sites in

epidemiological models [14, 6].

We first introduce a Bayesian computational method to estimate parameters for a percolation model with synergistic enhancement of infectivity, hereafter referred to, for brevity, as synergy and test the sensitivity of the inferences using simulated data. Thereafter we apply the methods to data from a simple experimental system that mimics both the pathogenic and saprotrophic systems. Successive snapshots for the spread of the fungus *Rhizoctonia solani* Kühn over a lattice of nutrient sites, are used to estimate the parameters for a percolation model with synergy and to infer mechanisms underlying the synergy. *Rhizoctonia solani* is a ubiquitous soil-borne fungus with both pathogenic and saprotrophic activity that is representative of a broad range of fungi. It causes damping-off disease of seedlings and root or stem-base rots in a very broad range of host species that include economically important agricultural crops as well as species in natural and semi-natural habitats [15]. It is therefore an economically and ecologically-important fungus in its influence on plant and crop health and in nutrient cycling in soils. Specifically, we investigate the following biological hypotheses:

1. do the transmission rates depend upon the local state of the infection neighborhood?

2. how does the *number* of infectious neighbours connected to a site affect the transmission rate?

3. does the relationship change with the nutrient state of the sites?

The consequences of the results for the interpretation and prediction of fungal and pathogen invasion are briefly discussed together with the robustness of the results to the assumptions that underpin the model and analyses and the general role of the methodology in epidemiological inference.

2 SYSTEM AND MODEL

The spread of *R. solani* (R5, AG2-1, IMI385769) was monitored through a lattice of nutrient sites as a series of spatio-temporal snapshots (*cf* Fig. 1b). The experimental system is the same as that detailed in [6] and described in detail in the methods section. In essence, the system consists of a circular Petri plate with 177 nutrient sites comprising potato dextrose agar available for fungal colonization arranged in a square lattice. The central site is inoculated with *R. solani*, and the progress of colonization during successive 24h periods is mapped for a maximum of 30 days. There were two treatments: one with high nutrient density (HND) at vertices to encourage sharing of nutrients and enhance synergy in colonization, and one with lower nutrient density (LND) which is expected to suppress the synergy. Note that differences in nutrient density at nutrient sites is expected to influence both the standard rate of transmission as well as an additional synergy parameter. There were 20 replicates of the HND treatment, and 10 for the LND.

Since synergy depends upon the connectivity between infected sites, the model for the system depends upon the state of the sites (vertices) and the links (or bonds) between sites. The model is based upon a standard SIR formulation. Each site is in one of three states: Susceptible, Infectious or Removed. On becoming infectious, a site remains so for a fixed period τ after which it enters the removed state. During the infectious period a bond to each neighbour is created as the first event in a Poisson process with

instantaneous rate λ. Should this first event fall outside of the infectious period, the bond is not created. Creating a bond between an infectious and a susceptible individual causes the infection of the susceptible. Bonds can also be created between two infectious sites, enabling sharing of nutrients, possibly changing the rate of bond creation of each. A site becomes removed a time τ after it has become infected.

We consider two models for synergistic enhancement. In both cases, the transmission rate λ for an infectious site varies with the number n of infected neighbours to which it is connected. For example, Fig. 1a shows an infectious site that is surrounded by two infectious and two susceptible sites and creates a bond with each susceptible site independently during its infectious period according to a Poisson process with rate λ. Since the central infectious site has a connection to only one of its infected neighbours, $n = 1$. In the first synergy formulation, λ has the simplest linear dependence whereby $\lambda = \alpha + (n-1)\beta$, and α is the initial rate of infection, and β is a parameter quantifying the amount of synergy. The appearance of $n - 1$ in the rate as opposed to n is motivated as follows: initially each infectious site (apart from the first) has one connected neighbour - the source of the infection. The above definition of λ therefore ensures that the *initial* rate of infection is always α, making comparison with the standard percolation model (which corresponds to $\beta = 0$) more convenient. If λ is a constant, this model precisely corresponds with bond percolation [3].

The second model is a generalization of the first that allows for any specified dependence on n, and allows us to assess whether the linear model is adequate to describe the synergy. In this case, the rate for each distinct number of infectious neighbours is estimated independently leading to four parameters, $\lambda(n)$ for $n \in (0, 1, 2, 3)$. This allows for unknown effects, such as, for example, saturation of the rate. Both models exemplify the concept of *d-synergy* as described in [8]. The first corresponds to the linear situation considered therein [8] while the second provides a more flexible framework for representing synergistic interactions. The results of this paper suggest that the generality of the latter model may be required to explain real-world synergistic interactions. The focus on d-synergy is felt to be appropriate as this form of synergy captures the known ability of fungi to translocate nutrients over considerable distances. A site with multiple connections to other infectious sites may therefore be expected have access to a greater supply of nutrients through this mechanism and consequently exhibit greater infectivity.

3   EXPERIMENTS AND MODEL FITTING

The fungal plant pathogen *Rhizoctonia solani* Kühn (R5, AG2-1, IMI385769) was introduced onto a Petri plate (140mm diameter) containing a population of potato dextrose agar (PDA) sites. Each site was a 3mm diameter dot of either 10% (w/v) PDA for the LND treatment, or 40% (w/v) PDA for the HND treatment. The sites were arranged in a square lattice with an inter-site distance of 8mm, extending as far as the circular boundary of the plate, giving a total of 177 sites. The central site was inoculated with a single hyphal strand, 1 mm in length, removed from the edge of a 4-day-old colony of *R. solani* growing on water agar. To avoid desiccation of the agar, moist filter paper was placed into the lid of each plate. The plates were sealed and incubated in the dark at 23^oC, and assessed daily for 28 d using a dissecting microscope (magnification $\times 40$) and the number and location of colonised sites recorded. The LND treatment was replicated 10 times, the HND treatment 20 times. The results of the 10% PDA

treatment were first published and analysed previously [6], whereas the 40% treatment has not previously been published. Note that the path of infestation and presence or absence of hyphal links between sites was not recorded.

To fit the mathematical model to the experimental data, a Bayesian analysis using Markov chain Monte Carlo (MCMC) methods has been performed. These are a set of algorithms used in this work to determine a distribution of belief in the values of parameters that underlie the data - in this case, α, β and τ, or $\lambda(0..3)$ and τ - by simulating the distribution using a Markov chain whose state space is the parameter space. A general MCMC calculation requires the use of any prior knowledge of the parameters, encoded as a probability distribution $p(\theta)$ where θ is the set of parameters defined above, and a likelihood function $L(\mathbf{x}|\theta)$, where $\mathbf{x}$ is the set of observed data. This likelihood is, in essence, the sampling density of obtaining the data given values of the parameters. In the Bayesian paradigm, these are combined to produce the posterior density:

$$\pi(\theta|\mathbf{x}) \propto L(\mathbf{x}|\theta)p(\theta). \tag{1}$$

The prior distribution represents the beliefs regarding the parameters before the experiment is performed. In these experiments, the spacing and nutrient density of the agar dots were chosen to ensure the system was close to the critical point separating invasive from non-invasive behaviour. For the linear model, this corresponds to the value of α being close to the critical point of a pure percolation process, i.e. $\alpha_c = \ln 2/\tau$. Similarly for the unconstrained model, this corresponds with $\lambda(1)$ being close to α_c. The priors on these parameters (α and $\lambda(1)$) can therefore be fairly informative, whereas the other parameters (β and $\lambda(0, 2, 3)$) have vaguer priors. In all cases, independent uniform prior distributions have been used to represent prior belief regarding parameters. The specific hyperparameters used to specify each uniform distribution are tabulated in table 1.

The likelihood $L(\mathbf{x}, \theta)$ in this case is the joint probability density of creation times of the bonds $\mathbf{x} = \{t_i\}$ considered as a function of the parameters of the model $\theta = (\alpha, \beta, \tau)$. By exploiting the independence of the Poisson processes underlying bond creation, the likelihood can be expressed as a product of the individual contributions l_i from each bond, both created and absent:

$$L(\mathbf{x}, \theta) = \prod_i l_i(t_i, \theta), \tag{2}$$

where

$$l_i(t_i, \theta) = \lambda_i(t_i, \theta) \exp\left(-\int_0^{t_i} \lambda_i(t, \theta)dt\right), \tag{3}$$

for the created bonds, where $\lambda(t, \theta)$ is the rate of bond creation as a function of time. For absent bonds, l_i is simply the probability that the bond was not created, and is equal to

$$l_i = \exp\left(-\int_0^\infty \lambda_i(t, \theta)dt\right). \tag{4}$$

The rate $\lambda_i(t, \theta)$ is calculated from the history of the environment of bond i over time. The bond connects two sites A and B, and its rate of creation is a function of the creation times of the neighbouring bonds, as well as the removal times of the neighbouring sites of A and B. As soon A or B becomes

5

infectious, $\lambda_i(t, \theta)$ takes a value of $\alpha + (n-1)\beta$ or $\lambda(1)$. For this initial event, n will always be one, since it takes one and only one infectious site connecting to infect A, hence $\lambda_i(t, \theta)$ is initially α (or $\lambda(1)$). As the environment changes, $\lambda_i(t, \theta)$ changes too, increasing as bonds are added to A and B, and decreasing as neighbours of A and B become removed. If B becomes infectious, the rate $\lambda_i(t, \theta)$ becomes the average of the rate due to A and the rate due to B. Once A or B becomes removed, $\lambda_i(t, \theta)$ drops to zero, since bonds cannot be made with removed sites. Biologically, the removal of a site corresponds to the exhaustion of the nutrients at that site, and therefore no sharing can happen.

In order to obtain the above likelihood, not only are the exact infection times for sites required, but also knowledge of the bonds created for nutrient sharing and their creation times. However, this information is not contained explicitly in the experimental data analysed. Observations were taken at daily intervals, where the state of each site was recorded. The exact path of infection (and hence the states of the bonds) were not recorded. Thus the data give the infection times of each site rounded up to the next day. These observations are denoted $\mathbf{y}$.

Since the infection process is determined entirely by the process of bond creation, we view creation times of bonds as the modelled variables rather than the infection times of the sites. The experimental observations then specify constraints on the bond creation times. A bond's existence implies that A and B are both infectious (as the bond's creation immediately infects any non-infectious sites). The minimum time the bond can have been created is therefore $\max[\min(t_A), \min(t_B)]$ where t_X is the range of possible infection times of site X. Similarly, a bond cannot be created linking a removed cell, therefore the maximum time the bond can have been created is $\min[\max(t_A), \max(t_B)] + \tau$ where τ is the infectious period. Any set of exact bond creation times $\mathbf{x} = \{t_i\}$ that satisfy the observations $\mathbf{y}$ gives a particular value of the likelihood. An average should be taken over the set $\mathbf{X}$ of such $\mathbf{x}$ to obtain the marginal likelihood as a function of the model parameters θ:

$$L(\theta|\mathbf{y}) = \int_{\mathbf{x} \in \mathbf{X}} L(\theta|\mathbf{x}) d\mathbf{x}. \tag{5}$$

The integration is not tractable, but can be approximated via computational Bayesian methods. Firstly, the unobserved bond creation times are treated as additional unknown parameters, and MCMC methods are then used to investigate the posterior distribution of this 'extended' parameter vector.

Following standard procedure, a joint posterior density on $\mathbf{\Theta} \times X$ is defined by

$$\pi(\theta, \mathbf{x}|\mathbf{y}) \propto p(\theta)L(\theta|\mathbf{x}) \tag{6}$$

where $p(\theta)$ is the prior distribution and $L(\theta|\mathbf{x})$ is evaluated using the method described above. The integration is performed by constructing a Markov chain on $\mathbf{\Theta} \times \mathbf{X}$. Steps are taken by either proposing new times for the individual bonds, proposing that bonds change their state from 'created' to 'absent' or vice-versa, or updating the parameters θ. These proposed steps are then accepted or rejected according to standard Metropolis-Hastings updating. For each component of θ, α, β and τ (or $\lambda(0..3)$ and τ), a

new value is proposed, giving the new parameter vector θ^p. This is then accepted with probability

$$P_{acc} = min\left(1, \frac{p(\theta^p)\pi(\theta^p, \mathbf{x}|\mathbf{y})}{p(\theta)\pi(\theta, \mathbf{x}|\mathbf{y})}\right). \tag{7}$$

The proposal of changes for the individual bonds is complicated by the fact that it is not known *a priori* which bonds are created and which are not. Thus the updating of an individual bond can either modify its creation time, or change the state ('on' to 'off' or vice-versa). This necessitates the use of Reversible Jump MCMC [19], which modifies the standard Metropolis-Hastings acceptance probability to become

$$p_{acc} = min\left(1, \frac{\pi(\theta, \mathbf{x}'|\mathbf{y})j_{m^r}(\mathbf{x}')g'_{m^r}(u')}{\pi(\theta, \mathbf{x}|\mathbf{y})j_m(\mathbf{x})g_m(u)}\left|\frac{\partial(x', u')}{\partial(x, u)}\right|\right) \tag{8}$$

where $j_m(\mathbf{x})$ is the probability of proposing move m from $\mathbf{x}$, for example the probability of proposing to change the time at which a bond was created, or of proposing to remove the bond entirely. The term $j_{m^r}(\mathbf{x}')$ denotes the probability of proposing the reverse move m^r from state $\mathbf{x}'$. Here $g_m(u)$ is the probability density at the particular value of the random quantity u required to make the jump. When moving up in dimension, this is the probability density of the new variable, in this case, the probability density of choosing the new bond creation time which is selected uniformly from its permissable range. The corresponding term is unity when moving down in dimensionality. The Jacobian $\frac{\partial(x', u')}{\partial x, u}$ is in this case also unity, since the dimension-changing moves are simply projections.

From a state $\mathbf{x}$, the new proposed state is constructed by selecting one of four possibilities, depending on the current state of the bond. If the bond is present in state $\mathbf{x}$, it can either be removed, or its time altered. If the bond is currently not present, it can either be inserted with an allocated creation time, or left as it is. The two alternatives for each state are chosen randomly with equal probability, so that j_m, the probability of choosing a particular jump, is identical for all m, and cancels from the above acceptance probability. Therefore when the dimensionality remains the same, the acceptance probability reduces to standard Metropolis-Hastings. When moving up or down in dimensionality, a factor of $1/\Delta$ is introduced, where Δ is the width of the permissable range of the bond-creation time. In order to analyse all the experimental replicates jointly, the likelihood function is simply replaced with the joint likelihood function, $L^J(\theta|\mathbf{x}) = \prod_j L(\theta|\mathbf{x}_j)$, where $\mathbf{x}_j$ represents the state vector of the jth replicate. This new likelihood function is used when updating components of the model parameter vector θ.

This updating scheme specifies a Markov chain whose stationary distribution is the target distribution $\pi(\theta, \mathbf{x}|\mathbf{y})$, and by simulating this chain the marginal posterior distributions of the parameters $\theta = (\alpha, \beta, \tau)$ or $(\lambda(0..3), \tau)$ can be obtained. These distributions represents the belief regarding parameter values given the experimental data.

Applying this technique to real data presents additional challenges in that infections between next-nearest neighbours occurred in the experimental results. These should not be permitted by the model outlined so far. There are several techniques for dealing with this problem, e.g. allowing for long-range infection processes [20] or by introducing a small rate of primary infection [4]. However, in this case, we have chosen to model these rare events as spontaneous infection of sites, with the exact infection time an additional parameter over which to integrate the likelihood, constrained only by the observation and not

by the creation time of an appropriate bond.

4 RESULTS

The statistical methodology was applied to fit the two models of synergy to the data from the microcosm experiments. Figure 1b illustrates spatio-temporal maps for the high (HND) and low nutrient density (LND) experiments at 7, 12 and 20d after initiation. It is evident that the HND sites are being colonized at a much higher rate than the LND sites: there are also far fewer gaps in the resulting pattern for colonization in the HND treatment. This is significant in that the presence of gaps provides important information with which to infer the lifetime τ of the fungus on a site. Inspection of the time-course for individual replicates (Fig. 2) clearly shows that the higher rate of spread occurs in the HND treatment. The data from the LND treatment also show that as well as having a lower average rate of spread, there is much more variability in the final size of the epidemic, e.g. in two replicates, no site other than the first became colonized.

The principal results of the MCMC analysis are shown in Figs 3-4 and offer some immediate insights into the questions outlined in the introduction. In the first instance there is clear evidence that for both treatments LND and HND a synergistic effect is present with the marginal posterior density for β being concentrated far from zero in both cases. This can be seen by inspection of the MCMC samples from the posterior density of (α, β) for the respective treatments (Fig. 3). The joint posterior density from the MCMC estimation for the transmission parameters, α and β, shows that nutrient density affects not only the initial rate of infection, α, but also the the nature and strength of synergy (Fig. 4a,b) when it is characterised by fitting the linear synergy model. The HND treatment exhibits both a higher initial rate of infection as well as a greater synergistic effect (Fig. 4b,d).

There is clear evidence that the nature of the synergistic interaction varies with the treatment. The linear and unconstrained models are compared in Figs 4c,d, along with the 95% HPDI calculated from the posterior densities of α and β. The models characterize the rates in a very similar manner for the low density treatment (Fig. 4c), with good overlap for all values of n, indicating that the linear model gives a parsimonius description of the synergy effect. However, for the high density experiment (Fig. 4d), there is a strong deviation (concave upwards) from the linear model in the rates fitted using the unconstrained model. Note that the linear model attempts to capture this feature, but is bounded by the requirement that $\lambda(n) > 0$, and is, perhaps most influenced by the rates $\lambda(1)$ and $\lambda(2)$ for which there are most data. Predictions made using the linear model in the HND should be expected to underestimate significantly the infectivity of sites with 3 infected neighbours.

Examination of the time-series plots for the parameters α, β and τ as the Markov chain progresses (Fig. 4a,b) verifies that the Markov chains mix well and that the estimated posterior densities are likely to be accurate. However we note that the trace-plot for τ covers smaller values and exhibits lower-frequency fluctuations in the high-density treatment indicating less efficient mixing of the chain in this case. This phenomenon may be anticipated given that the acceptable range for the lifetime (τ) is constrained by the current configuration of bonds and creation times. Should the proposed lifetime invalidate the presence or creation time of any bond, the proposal would immediately be rejected. This is more likely to happen

under the high density treatment for which there were more replicates and more bonds per replicate. An accurate estimate of the posterior density of τ is nevertheless possible within the length of simulation runs described here (Fig 4a,b). We infer that the smaller value for τ, and hence the shorter 'infectious' period associated with the high nutrient source, reflects differences in modes of growth, whereby the fungus with ready access to nutrients grows faster and more vigorously, and hence uses the nutrients more quickly.

5 DISCUSSION

The percolation paradigm is widely invoked as a plausible mechanism for the spread of epidemics [16]. The impact of percolation models on decision making is generally achieved through studies that first estimate key parameters in order to identify whether the system is in the invasive or non-invasion regions of parameter space. Control measures can then be designed that affect the parameters so as to move the system from the invasive to the non-invasive region. Until recently, the utility of this approach was hampered by several of the simplifying assumptions in the standard percolation paradigm. One such assumption that is likely to be violated in many biological systems is the restriction to nearest-neighbour interactions in which interactions between a given pair of neighbours are not influenced by the states of other neighbours. Failure to account for these interactions could give highly misleading inferences about the probability of invasion and the effectiveness of potential control strategies. Our methods exploit the flexible paradigm of percolation with synergy, as introduced in [8], to provide a statistical scheme to avoid this risk by explicitly detecting and characterizing synergy in the transmission process. In particular, we formulate Bayesian computational techniques that can distinguish synergistic interactions at the individual scale that underlie the statistical properties of large-scale patterns of disease spread. The methods use modest amounts of experimental data and replication.

We have used the methods to answer the three questions listed in Section 1 concerning the effects of synergy on transmission rates. In particular, under both low-nutrient (LND) and high-nutrient (HND) treatments the inferences on parameters suggest that the transmission dynamics are not captured by a standard percolation model; rather, there is overwhelming evidence that favors the presence of synergistic enhancement of infectiousness. We conclude that the transmission rates depend upon the local state of the infection neighborhood, since the posterior distribution for β is far from zero. Moreover, while the analysis shows that a linear dependence of infectiousness on the number of infected neighbours is supported by the data for the LND treatment, the parameter estimates for $(\lambda(0), \lambda(1), \lambda(2), \lambda(3))$ when the unconstrained model is fitted to the HND treatment show that the dependence is highly nonlinear. This confirms the hypothesized effect of nutrient status of sites on the characteristics of local transmission. It is plausible that this nonlinearity could arise through a synergistic effect due to connections with infectious sites beyond the nearest neighbours in the HND case. Any site with connections to 3 infectious neighbours may tend to be located in a region of high infection density and may consequently have synergistic connections with other infectious sites. The influence of these sites may be being captured in the nearest-neighbour model with general synergy by $\lambda(3)$ taking an inflated value, leading to the nonlinearity. A further plausible mechanism may be a tendency of the fungus to explore preferentially directions that avoid areas already colonised. These mechanisms are speculative at this stage.

9

Attention here is focussed on a particular system using models that make a number of strong assumptions. We remark that the methods can be readily applied, with appropriate modification, in more general settings. For example, the 'memoryless' assumption that underlies the formation of bonds can be dispensed with, as can the assumption of a constant infectious period (cf [17, 18]) without compromising the feasibility of the parameter estimation techniques. It is also possible to investigate synergistic effects for models that include transmission beyond nearest neighbours. We are confident that there exist numerous opportunities for widespread application of the approach.

References

[1] Frisch, H.L. and Hammersley, J.M. (1963) Percolation processes and related topics *J. SIAM* **11**, 894–918.

[2] Cardy, J. L. and Grassberger, P. (1985) Epidemic models and percolation, *Journal of Physics A-Mathematical and General* **18**, L267–L271.

[3] Grassberger, P. (1983) On the Critical Behaviour of the General Epidemic Process and Dynamical Percolation, *Math. Biosci.* **63**, 157–172.

[4] Gibson, G. J. Otten, W. Filipe, J. A. N. Cook, A. Marion, G. and Gilligan, C. A. (2006) Bayesian estimation for percolation models of disease spread in plant populations, *Statistics and Computing* **16**, 391-402.

[5] Bailey, D. J. Otten, W. and Gilligan, C. A. (2000) Percolation, heterogeneity and the saprotrophic invasion of soil by the fungal plant pathogen Rhizoctonia solani, *New Phytologist* **146**, 535–544.

[6] Otten, W. Bailey, D. J. and Gilligan, C. A. (2004) Empirical evidence of spatial thresholds to control invasion of fungal parasites and saproptrophs, *New Phytologist* **163**, 125–132.

[7] Davis, S. Trapman, P. Leirs, H. Begon, M. and Heesterbeek, J. A. P. (2008) The abundance threshold for plague as a critical percolation phenomenon, *Nature* **454**, 634–637.

[8] Perez-Reche, F. J., Ludlam, J. J., Taraskin, S. N., and Gilligan, C. A. (2011) Synergy in spreading processes: from exploitative to explorative foraging strategies, *Phys. Rev. Lett.*, **106**, 218701-4.

[9] Lockwood, J. L. (1988) Evolution of concepts associated with soilborne plant pathogens, *Annual Review of Phytopathology* **26**, 732–746.

[10] Olsson, S. (1995) Mycelial density profiles of fungi on heterogeneous media and their interpretation in terms of nutrient reallocation patterns, *Mycological Research* **99**, 143–153.

[11] Wells, J. M. Boddy, L. and Donnelly, D. P. (1998) Wood Decay and Phosphorus Translocation by the Cord-Forming Basidiomycete Phanerochaete velutina: The Significance of Local Nutrient Supply, *New Phytologist* **138**, 607–617.

[12] Davidson, F. A. (1998) Modelling the qualitative response of fungal mycelia to heterogeneous environments, *Journal of Theoretical Biology* **195**, 281–292.

[13] Zakaria, A. J. and Boddy L. (2002) Mycelial foraging by Resinicium bicolor: interactive effects of resource quantity, quality and soil composition, *FEMS Microbiology Ecology* **40**, 135–142.

[14] Gilligan, C. A. (2002) An epidemiological framework for disease management, *Advances in Botanical Research* **38**, 1–64.

[15] Sneh, B., Jabaji-Hare, S., Neate, S. and Dijst, G. eds. (1996) *Rhizoctonia species, taxonomy, molecular biology, ecology, pathology and control.* (Kluwer, Dordrecht).

[16] Cox, J. T. and Durrett, R. (1988) Limit-Theorems for the Spread of Epidemics and Forest Fires, *Stochastic Processes And Their Applications* **30**, 171–191.

[17] O'Neill, P. D. and Becker, N. G. (2001) Inference for an epidemic when susceptibility varies, *Biostatistics* **2**, 99-108.

[18] Streftaris, G. and Gibson, G. J. (2004) Bayesian analysis of experimental epidemics of foot-and-mouth disease, *Proceedings of the Royal Society Series B* **271**, 1111 - 1117.

[19] Green, P. J. (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika* **82**, 711–732.

[20] Gibson, G. J. (1997) Markov Chain Monte Carlo methods for fitting and testing spatio-temporal stochastic models in plant epidemiology, *Applied Statistics* **46**, 215–233.

Figure legends

Figure 1 (a) An example of the calculation of the rate of infection (bond creation). The site in the center is surrounded by two infectious neighbours and two susceptible neighbours. There is a connection with only one of its infectious neighbours, so the instantaneous rate of infection $\lambda = \alpha + (n-1)\beta$ is determined by n, the number of connected infectious neighbours, taking the value 1, giving $\lambda = \alpha + \beta$ in this case. (b) Snapshots from three times of representative replicates from the HND and LND experiments. Susceptible sites are white, and infected sites are dark grey. The LND replicate is clearly spreading more slowly and in a more filamentary manner.

Figure 2 Time-course data for the individual replicates for each of the two treatments, LND (10 replicates) and HND (20 replicates), showing the number of infected sites $I(t)$ over time. The LND treatment exhibits much greater variability than the HND treatment in the final numbers of affected sites, suggesting the LND system is closer to the percolation threshold.

Figure 3 Scatter plots of samples from the joint posterior densities of infection parameters (α, β) for the two treatments. The HND treatment has both a higher initial rate of infection and synergy parameter than the LND treatment. For each treatment the sample is generated from a run of 10^6 iterations of the MCMC scheme described in Section 3, after discarding to first 10^5 iterations as a 'burn-in' period. Parameters α, β and τ were initialised to the mean of the corresponding prior distribution from Table 1. Bond creation times were initialised by setting each to the earliest time consistent with the observations.

Figure 4 (a,b) Infection parameters α, β and τ plotted against iteration number for the LND (a) treatment and HND (b) treatment. There is no systematic deviation with iteration number, and hence convergence has been achieved. Note however that fluctuations are clearly visible in the HND treatment. (c,d) Comparison of the linear and general models for the LND (c) and HND (d) treatments. Plotted are 95% credible intervals (estimated from an MCMC sample of 9×10^5 iterations after discarding an initial burn-in 10^5 as a burn-in) for transmission rate (λ) against the number of connected neighbours (n). There is a high degree of overlap in the estimates from the two models in (c), but a significant difference in (d), particularly for high and low values of n. Note also that the estimate with the least uncertainty in both treatments is $\lambda(1)$.

Short title: Models for fungal spread with synergy