Fitting a lattice model with local and global transmission to spread of a plant disease

Alex Best[1*] and Nik J. Cunniffe [2]

[1]School of Mathematical & Physical Sciences, University of Sheffield, Sheffield, S3 7RH

[2]Department of Plant Sciences, University of Cambridge, Cambridge, CB2 3EA

[*]Corresponding author: a.best@shef.ac.uk

Abstract

Understanding, predicting and managing the spread of plant pathogens is crucial given the economic, societal and climatic benefits of plants, including crops and trees. Mathematical models have long been used to investigate disease dynamics in plants. An important component of such models is to account for spatial structure, since plant hosts are immobile and a majority of disease spread will often be localised. Here we apply a lattice-based mathematical modelling approach, a pair approximation, to model disease spread. While this method has previously been used to develop epidemiological theory, it has not been used to predict spread in a specific pathosystem. We fit our lattice-based epidemiological model to experimental data relating to Bahia bark scaling of citrus, an economically-important disease in north-eastern Brazil, and compare its performance to a more commonly used dispersal-kernel modelling approach. We show that the lattice-based model fits the data well, predicting a significant degree of near-neighbour infections, with similar estimated values of epidemiologically-meaningful parameters to the dispersal model. We highlight the pros and cons of the lattice-based approach and discuss how it may be used to predict disease spread and optimise control of plant diseases.

Author Summary

Plant diseases can have significant impacts, including reducing crop yields, limiting the availability of natural spaces, and the knock-on effects on our wellbeing. Mathematical models have long been used to understand how disease spreads through plant populations. Here we apply a form of mathematical model that has not previously been specifically applied to a real disease system that emphasises neighbour-to-neighbour spread of infection. In particular, we use the model to explore the spread of Bahia bark scaling of citrus, for which we have excellent experimental data available of its spatial spread. We show that the model fits the data best when there is significant neighbour-to-neighbour spread with very rare long-range infections. We show that this approach agrees well with a more commonly-used mathematical framework and highlight how it might be used to test disease management strategies.

Introduction

Diseases caused by plant pathogens can have severe consequences. Pathogens affecting crops impact yield as well as food quality and safety, leading to economic and social impacts, and in severe cases food security can even be at risk [1]. When pathogens infect plant species in natural populations, provision of ecosystem services can also be impacted [2]. Many plant pathogens are extremely well-established, causing disease on certain host species in particular locations somewhat predictably. Just two of many possible examples are late blight of potato (which remains the most significant disease of potato worldwide [3], caused by the oomycete *Phytophthora infestans*) and septoria leaf blotch of wheat (one of the most significant diseases of wheat grown in temperate climates [4], caused by the fungus *Zymoseptoria tritici*). While impacts are very difficult to unambiguously quantify, losses of yield due to these and other pathogens of major food crops are generally estimated to be in excess of 10% [5, 6].

In addition to these endemic diseases, many other plant pathogens are emerging [7]. Outbreaks of emerging plant pathogens are increasingly well reported [8, 9], with invasion rates appearing to increase [10, 11]. Drivers are varied and complex, but are generally acknowledged to include changes to farming practices and land use [12, 13], climate change [14, 15], and increased travel and trade [16, 17]. High profile current examples of emerging pathogens of crops include *Xylella fastidiosa* in Europe, affecting many species including olives [18] and almonds [19]; Tropical Race 4 of *Fusarium oxysporum* f.sp. *cubense*, which is spreading in many regions where banana or its wild relatives are grown [20]; *Candidatus* Liberibacter spp. causing huánglóngbìng/citrus greening, particularly in the USA and Brazil [21]; and tomato brown rugose fruit virus, which is spreading to new countries worldwide [22]. Meanwhile, in the United Kingdom alone, trees in forests are threatened by sudden larch death (caused by *Phytophthora ramorum*, the same pathogen which causes sudden oak death elsewhere [23]), chalara ash dieback (caused by *Hymenoscyphus fraxineus*, [24]) and dothistroma (red band) needle blight on conifers (caused by *Dothistroma septosporum*, [25]). Documented and ranked in the UK Plant Health Risk Register are a worryingly large number of pathogens which could potentially invade and threaten UK forests in the future [26].

Mathematical modelling is now routinely used to inform decision making around detection and control of disease outbreaks for pathogens of humans and animals [27]. However, plant disease modelling has its own long and somewhat distinctive history [28]. Models range from purely statistical, which do not attempt to

capture or characterise biological mechanisms underpinning disease risk [29], to process-based compartmental models, which focus closely on interactions between members of host population(s) with different disease statuses [30, 31, 32]. Often, for simplicity, the spatial component of disease spread is omitted, and non-spatial models are used [33]. However, since they are caused by pathogens that are actively spreading, emerging disease epidemics tend to have a particularly strong spatial component, and this requires models to include space [34]. Sometimes space is represented implicitly, focusing on pathogen spread between a small number of patches [35, 36] or constituents of a metapopulation [37]. However, recent models tend to echo trends in ecology more broadly [38] by relying on the concept of a dispersal kernel, which in this context is a characterisation of the probability of inoculum from an infected host challenging a susceptible host elsewhere [34]. Early work often used a continuum description of space, either via reaction-diffusion partial differential equations [39], or integro-differential equations tracking pathogen spread [40]. Probably motivated by applications to particular pathosystems, more recent work has more often tracked disease spread in realistic geographics. Descriptions of the host population range from individual-based models [41, 42, 43] to coupled structured metapopulations, in which the disease status of the hosts in each cell on a rasterised map is tracked via its own compartmental model of disease [44, 45]. Often this type of work has relied on repeated computer simulation of stochastic versions of the models, although analytic techniques are now starting to become available [46, 47].

An additional subset of mathematical models for considering spatial structure in host populations are lattice-based methods [48, 49, 50]. The assumptions of these models are highly appropriate to plant systems, especially managed and agricultural populations where plants are often laid out in regular spacings. Here, individuals in a population are assumed to be fixed in a grid, with each cell occupied by at most one individual. Interactions can be assumed to be mean-field or 'global', but may also be near-neighbour or 'local'. Using a pair approximation, this system can be well described by an extended set of ordinary differential equations [48, 51, 52]. This deterministic model is often paired with fully spatially-explicit stochastic simulations. This approach has been used in a range of studies to explore the epidemiology, ecology and evolution of infectious disease. Early studies highlighted how local transmission leads to smaller and slower epidemics than the mean-field [53]. More recently, studies have explored how we can vary the ratio of local-to-global transmission to understand the dynamics in more detail, both at the epidemiological [54, 55, 56] and evolutionary timescale [57, 58]. However, to date these studies have almost exclusively been used to gain theoretical insights, and

have not been used to directly model particular systems or be fitted to data. The closest example is that of Filipe et al. [59], who fit a purely near-neighbour lattice model to temporal experimental data of *Rhizoctonia solani* infection of radish plants, finding it fit their data well. Our work goes further by allowing the local-to-global parameter itself to be fitted to our dataset, to include limited spatial data in the fitting, and to take a Bayesian fitting procedure.

The aim of this study is to assess the utility of the lattice-based approach in understanding disease spread in real-world systems. In particular, we will focus on previously published data set on Bahia bark scaling of citrus (BBSC), a disease of unknown etiology which has been endemic in north-eastern Brazil since the 1960s [60]. This system is a useful comparison since (a) the experimental study was of citrus trees laid out in a simple grid pattern, matching the assumptions of the model well, and (b) a study fitting an individual-based model representing transmission via a dispersal kernel to the data has already been undertaken [61]. They showed their model could fit the BBSC data well, giving us a good comparison for our parameter estimates.

Materials and methods

Experimental data

We use published experimental data for BBSC to fit our models [60]. In the experiment underlying this work, a 15×16 grid of initially immature citrus trees was established with a 2m x 2m spacing adjacent to a small stand of older trees showing BBSC symptoms. These established founder trees were not in the same spatial arrangement as the experimental grove, with greater distances between each individual, as well as between the established trees and experimental grove. As mentioned below, this initial arrangement of founder trees in the experiment needed to be expressed differently in the model. In particular, we assume (in both models) that there is a single line of infected individuals along the southern edge of the tree stand, since we require equal spacing for the lattice model assumptions to hold. The first infection was detected around 30 months after the experimental grove was planted. Over a period of 5 years, the trees were then inspected 14 times and those showing signs of infection recorded. We therefore know both the total number of trees symptomatically infected at each sample point as well as their spatial location. Snapshots of the experimental data are shown in supplementary Fig. S1.

The epidemiological model

Our model follows a Susceptible - Exposed - Infected compartmental framework, with no births, deaths or recovery. This is in line with the previous modelling study of BBSC [61] and with the nature of the experimental system [60]. Let P_S, P_E, P_I and P_F be the probabilities that, if we choose a random site on the lattice, it is occupied by, respectively, a susceptible, exposed, infected or 'founder'. The founder trees, P_F are those trees initially inoculated at the start of the experiment, one row outside of the main experimental tree stand. Since there is no recovery, P_F is therefore a constant. Further, as there are no empty sites on the lattice, $P_S + P_E + P_I + P_F = 1$ always. As such, we need only track two of these variables with differential equations. Also let the term $q_{I/S}$ be the conditional probability that if we choose a susceptible individual, it has a neighbour that is infected, and similarly for $q_{F/S}$. We can describe the dynamics of the 'singlet' densities by the following equations,

$$\frac{dP_S}{dt} = -\beta\left(Lq_{I/S} + (1-L)P_I\right)P_S - \beta_F\left(Lq_{F/S} + (1-L)P_F\right)P_S, \tag{1}$$

$$\frac{dP_E}{dt} = \beta\left(Lq_{I/S} + (1-L)P_I\right)P_S + \beta_F\left(Lq_{F/S} + (1-L)P_F\right)P_S - \rho P_E. \tag{2}$$

The parameter β gives the rate of infection from infected trees, and β_F that from founder trees. We allow the infection rates from the two types of tree to differ for a number of reasons: the founder trees were much larger, they had been infected for some time at the start of the experiment, and the spatial distribution of founder trees in the model is slightly different to that of the experiment (described in detail above). The parameter ρ is the rate of progression from being exposed to being infectious (with $1/\rho$ giving the latent period). The parameter $L \in [0, 1]$ controls the ratio of local-to-global infection. We stress here that 'local' infections are purely near-neighbour, and 'global' infections are random/mean-field, such that $L = 0.5$ means that half of an individual's infectious contacts are from neighbouring sites and half from anywhere else on the lattice. If $L = 1$ the equations (1)-(2) are simply the standard mean-field SEI model. Note we assume this is the same for all trees - we felt justified in this assumption given initial exploratory work where we included a separate parameter, L_F, for founder trees identified that the posterior distributions of the parameters L and L_F were almost identical.

This system is not closed, since, for example, $q_{I/S} = P_{SI}/P_S$, where P_{SI} is the probability that if we choose a neighbouring pair of sites, one is susceptible and the other infected. We therefore need to write down

equations describing the dynamics of these 'pair' densities. We can find these by considering the possible transitions between pairs of sites. For example, consider the creation and loss of an $[S, I]$-pair:

- $[S, E] \to [S, I]$ (created) E member of an $[S, E]$-pair becomes symptomatic at rate ρ
- $[S, I] \to [E, I]$ (lost) S member of the pair becomes exposed after infection from
  - its known I neighbour (i.e. from the $[S, I]$-pair) at rate $\beta L/4$
  - another I neighbour (since each individual has 4 neighbours) at rate $3\beta L q_{I/SI}/4$
  - a random I contact at rate $\beta(1-L)P_I$
  - a F neighbour at rate $3\beta_F L q_{F/SI}/4$
  - a random F contact at rate $\beta_F(1-L)P_F$.

Considering all such possible transitions we reach the following set of equations,

$$\frac{dP_{SS}}{dt} = -2\left[\beta\left(\frac{3}{4}Lq_{I/SS} + (1-L)P_I\right)P_{SS} + \beta_F\left(\frac{3}{4}Lq_{F/SS} + (1-L)P_F\right)P_{SS}\right], \tag{3}$$

$$\begin{aligned}\frac{dP_{SE}}{dt} =& -\beta\left(\frac{3}{4}Lq_{I/SE} + (1-L)P_I\right)P_{SE} - \beta_F\left(\frac{3}{4}Lq_{F/SE} + (1-L)P_F\right)P_{SE} - \rho P_{SE} \\ &+ \beta\left(\frac{3}{4}Lq_{I/SS} + (1-L)P_I\right)P_{SS} + \beta_F\left(\frac{3}{4}Lq_{F/SS} + (1-L)P_F\right)P_{SS},\end{aligned} \tag{4}$$

$$\begin{aligned}\frac{dP_{SI}}{dt} =& -\beta\left(L\left[\frac{1}{4} + \frac{3}{4}q_{I/SI}\right] + (1-L)P_I\right)P_{SI} - \beta_F\left(\frac{3}{4}Lq_{F/SI} + (1-L)P_F\right)P_{SI} \\ &+ \rho P_{SE},\end{aligned} \tag{5}$$

$$\begin{aligned}\frac{dP_{EE}}{dt} =& -2\rho P_{EE} + 2\beta\left(\frac{3}{4}Lq_{I/SE} + (1-L)P_I\right)P_{SE} \\ &+ 2\beta_F\left(\frac{3}{4}Lq_{F/SE} + (1-L)P_F\right)P_{SE},\end{aligned} \tag{6}$$

$$\begin{aligned}\frac{dP_{EI}}{dt} =& -\rho P_{EI} + \beta\left(L\left[\frac{1}{4} + \frac{3}{4}q_{I/SI}\right] + (1-L)P_I\right)P_{SI} + \beta_F\left(\frac{3}{4}Lq_{F/SI} + (1-L)P_F\right)P_{SI} \\ &+ \rho P_{EE},\end{aligned} \tag{7}$$

$$\frac{dP_{II}}{dt} = 2\rho P_{EI}, \tag{8}$$

$$\frac{dP_{SF}}{dt} = -\beta\left(L\frac{3}{4}q_{I/SF} + (1-L)P_I\right)P_{SF} - \beta_F\left(L_F\left[\frac{1}{4} + \frac{3}{4}q_{F/SF}\right] + (1-L_F)P_F\right)P_{SF}, \tag{9}$$

with $P_{SS} - P_{II} - P_{EE} - P_{FF} + 2(P_{SE} + P_{SI} + P_{EI} + P_{EF} + P_{IF}) = 1$. This system is also not closed, since we have conditional probabilities such as $q_{I/SS} = P_{SSI}/P_{SS}$, which would require writing down the dynamics

of 'triplet' densities. At this point we can apply a pair approximation, letting $q_{I/SS} = q_{I/S}$, which closes the system [48]. While this model comprises a large set of equations, it only includes four parameters: the infection parameters from infected (β) and founder (β_F) trees, the emergence of infectivity rate, ρ, and the parameter controlling the ratios of local-to-global transmission, L.

We also produce stochastic simulations of the system using a direct-method ('Gillespie') algorithm [62]. This uses the singlet equations from the lattice model (equations (1)-(2)) to define the event rates, but unlike the deterministic model it is fully spatially explicit, with dispersal either near-neighbour (with probability L) or random (with probability $1 - L$). While not used in the model fitting, results from this model allow us to check that the output of fully spatially-explicit simulations match with the predictions of the deterministic but approximate model. These simulations are also required for looking at detailed spatial questions, such as the level of infection on certain rows of the grid.

For comparison, we also produce a dispersal model. This has the same underlying epidemiological assumptions, but now rather than the simple local/global delineation, the way in which the probability of infection decays with increasing distance is parameterised using a dispersal kernel. In keeping with the previous study [61], we assume the dispersal kernel is exponential. We calculate the distance, d_{ij} between two trees i and j (treated as the central point of the grid cells for direct comparison to the lattice-based pairs model), and then take the dispersal kernel between a susceptible individual S_i and infected individual I_j to be,

$$\phi_{ij} = \Lambda \frac{\exp(-d_{ij}/\alpha)}{2\pi\alpha^2},$$

where α is the scale parameter (and the corresponding mean dispersal distance in a homogeneous landscape is 2α [34]). In our model one 'unit' of distance is the space between two trees, which in the experiment was 2m. Therefore the 'real-world'-adjusted mean dispersal distance would be 4α. The parameter Λ accounts for the typical (crown) area of trees, and makes the dispersal kernel a dimensionless quantity. Therefore the transmission rate experienced by a susceptible individual, S_i, is, $\beta \sum_j \phi_{ij} I_j$. As described in [61], we arbitrarily set $\Lambda = 1\text{m}^2$, noting that any value could have been used since any choice can be scaled into the fitted infection rate parameters, β and β_F.

Note that this system cannot be analysed deterministically. As such we produce an additional direct-method stochastic algorithm for this system. The parameters for this model are again infection rate parameters, β and β_F, progression rate ρ and now dispersal scale parameter, α.

As in the previous study that focussed on a dispersal model for this data [61], we also include an initial time delay in both models to account for the time it takes for the (non-founder) plants within the grove to reach epidemiological maturity, and so the delay before the first infections can occur, δ. Since the first infections were detected 30 months after the beginning of the experiment, we assume $0 \leq \delta \leq 30$ in our model fitting (additional investigations where we allowed $\delta > 30$ found extremely low numbers of accepted fits beyond this threshold).

Model fitting

We use an Approximate Bayesian Computation (ABC) approach to fit the models to the data. In a Bayesian framework, the parameter values are treated as unknown random variables. Our aim is to calculate the likelihood that certain parameter values (θ) would generate the data (D). In particular, we can use prior information we might have about the distribution of the parameter values, $\pi(\theta)$, in combination with the likelihood of the data, $\pi(D|\theta)$, to arrive at posterior distributions for the parameter values,

$$\pi(\theta|D) = \frac{\pi(D|\theta)\pi(\theta)}{\int_\theta \pi(D|\theta)\pi(\theta)d\theta}$$

A considerable difficulty is computing the likelihood, $\pi(D|\theta)$. The ABC approach negates this issue by simulating the model and considering how closely model outputs (for example the number of infected plants) match the recorded data points. The key to this approach is, rather than looking for a single 'best fit', we keep all parameter combinations where the fit is 'good enough', and use the proportion of simulations for a given parameterisation that meet this constraint as a proxy for the likelihood function. ABC approaches are increasingly applied to a range of biological models including epidemic models [63, 64, 65, 66, 67].

We take a relatively straightforward rejection-ABC approach, using the experimental data described above. For our fitting function, we take the sum of squared error between the data (expressed as proportions) at each of the recorded time-points and the model prediction for the number of infections (P_I) and the infected-infected pair densities (P_{II}) at these time-points. Since the experimental data only records the number of infections in the experimental tree stand, we re-calculate P_I and P_{II} to be the relative proportion of *non-founder* trees. To be accepted, the errors of both P_I and P_{II} must be within some threshold. Clearly, the posteriors will vary depending on the chosen acceptance threshold. We present results in the main text for a single chosen threshold, but show results for other values in the supplementary information and comment

on our choice of default acceptance threshold. We use uniform prior distributions for all parameters, with $L \in [0, 1]$ (dimensionless), $\beta, \beta_F \in [0, 5]$ months^{-1}, $1/\rho \in [0, 20]$ months (note we are using the latent period rather than the direct parameter ρ here as it has clearer biological significance), and $\delta \in [0, 40]$ months. In the dispersal model we take $\alpha \in [0, 10]$ (α has units of distance, with 2α the mean dispersal distance, but in the model one 'unit' of distance equates to 2 metres in the experiment).

For both the lattice model and dispersal model we run 5 million simulations on a High Performance Computing cluster. All simulations are conducted in Python. The lattice model is run using solve_ivp() with the Radau solver from the SciPy package [68]. Python code to run the models is available at `https://github.com/abestshef/plant_lattice_dispersal` .

Results

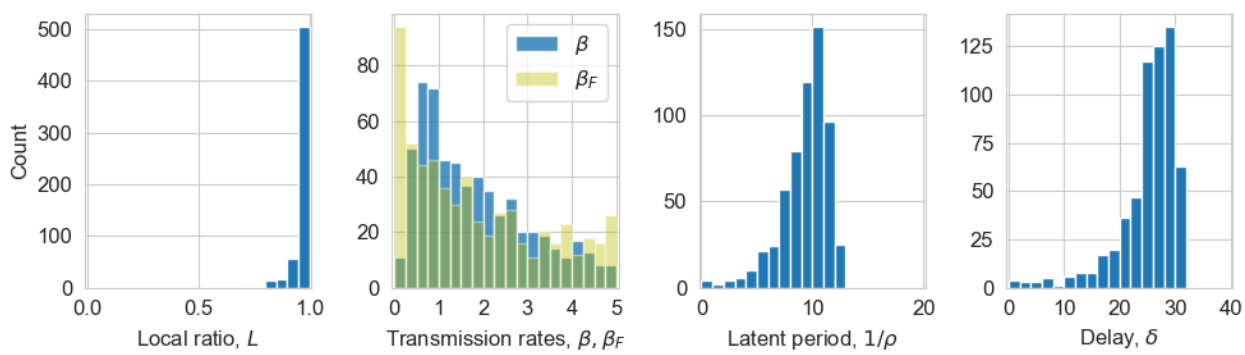

Figure 1: **Posterior distributions for the lattice model.** Histograms of the number of simulation accepted by the ABC procedure for local-to-global ratio, L, transmission parameters for regular, β, and founder β_F infections, latent period $1/\rho$ and delay, δ.

Lattice model

We begin by examining the posterior distributions for the lattice model. We choose an error acceptance threshold of 0.025 for the results presented in the main text, for which 598 (0.012% of the 5 million runs) parameter sets were accepted. In the supplementary information we show posterior distributions for a number of different thresholds for comparison. We find that our chosen threshold has enough accepted fits to form

Parameter	Mode	90% Smallest Credible Interval
L	0.99	[0.93, 1.00]
β (months^{-1})	0.82	[0.18, 3.70]
β_F (months^{-1})	0.48	[0.01, 4.26]
$1/\rho$ (months)	10.53	[6.58, 12.40]
δ (months)	27.57	[18.66, 31.87]

Table 1: Mode parameter values for the lattice model (using a Gaussian kernel density estimator) and the 90% Smallest Credible Interval/Highest Density Interval from the posterior distributions with an error threshold of 0.025.

'smooth' posteriors, and the distributions remain consistent for small changes to the threshold.

The posterior distributions are plotted in Fig. 1 and the mode for each parameter (after smoothing from a Gaussian kernel density estimator) and its 90% Smallest Credible Interval (sometimes known as Highest Density Interval) are shown in Table 1. The local-to-global ratio parameter, L, shows a significant skew, with a mode of 0.99, and no values of $L < 0.70$ accepted, indicating that predominantly localised spread is likely. For the two transmission parameters - β_F for infections from founders and β for all secondary infections - the distributions are notably different, with a much broader distribution for the founder infections. The distribution for the latent period is reasonably symmetric around a mode of 10.53 months, while δ is left-skewed with a mode of 27.57 months. We show correlations between the parameters in supplementary Fig. 2, which highlight that β and $1/\rho$ have a strong non-linear correlation, with a much broader spread of transmission rates for high latent periods. We also see that low L values require low β values.

We conduct a posterior predictive check of the dynamics in Fig. 2 by randomly choosing 100 sets of parameter values from our joint posterior distributions and plotting the dynamics against the data. We see remarkably good fits visually from the model to both data sets (single and pair densities).

We also used the posterior distributions from the deterministic model to run fully spatially-explicit stochastic simulations of the model to check that the results are similar to the approximate deterministic model. Fig. 3a reveals that while the visual fit is still good, the stochastic simulations do show more variation than the deterministic model; this is to be expected due to the inherent stochastic nature of these simulations, but may potentially also be due to the more detailed spatial organisation included in the simu-

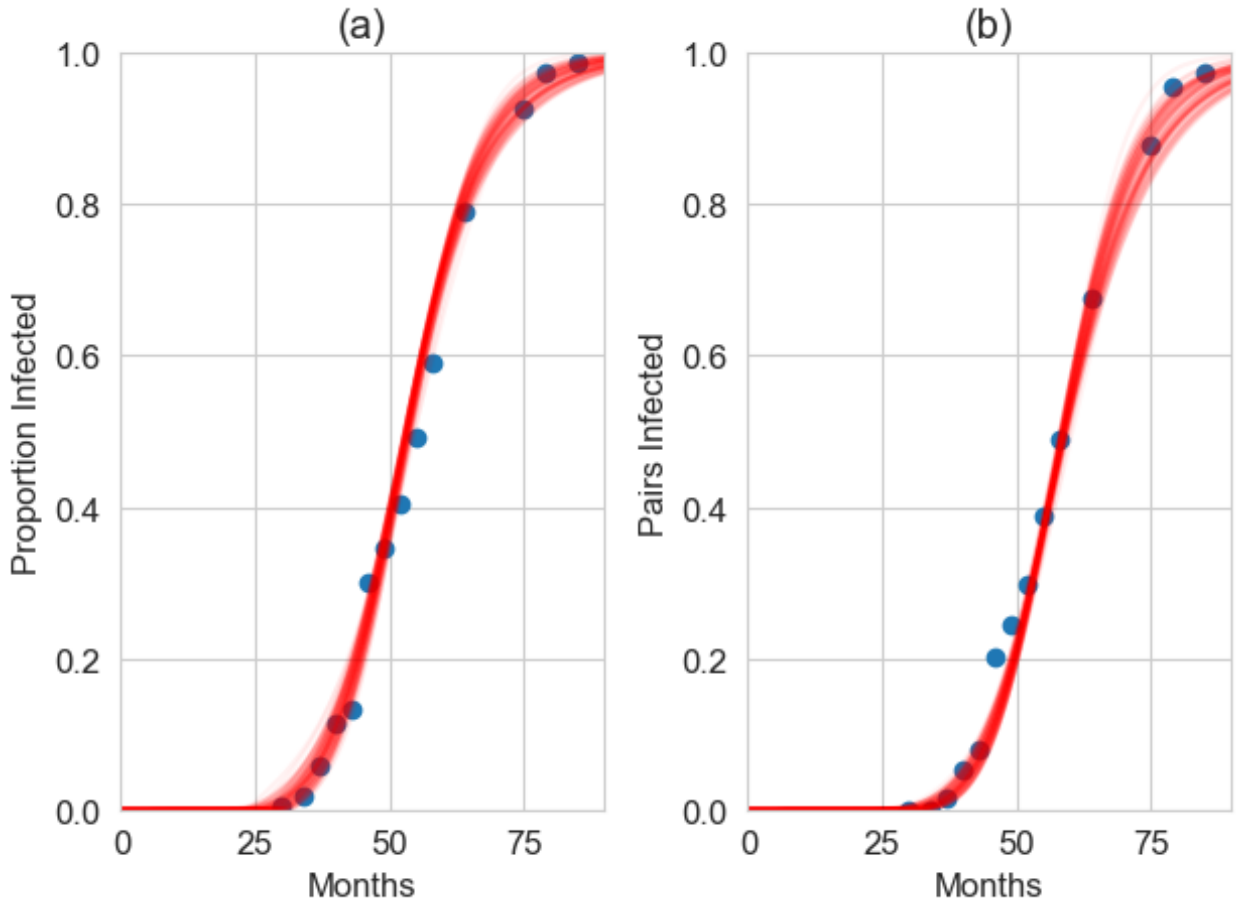

Figure 2: **Posterior predictive check of the lattice model.** 100 simulation runs of the lattice model (equations (1)-(9)) using parameters randomly drawn from the joint posterior distributions in figure 1 shown in red. The blue dots mark the data points from the BBSC experiments. Panel (a) shows the total proportion infected, panel (b) the proportion of Infected-Infected neighbouring pairs.

lations. We can also use these simulations to examine the spatial spread of the infection by considering how quickly it spreads to each row (Fig. 3b), as was similarly done in the previous modelling study of these experimental data [61]. We see that across the first 6-8 rows there is a trend of increasing time for the infection to reach the row that is roughly in line with the data. Past this point, the model predicts a fairly steady arrival time, due to the effect from the lattice model of infection being either near-neighbour or random; the rows at a distance from the source are almost equally likely to be infected through random infection.

As a guide, we also produce a series of snapshots of the grid for one randomly chosen stochastic realisation

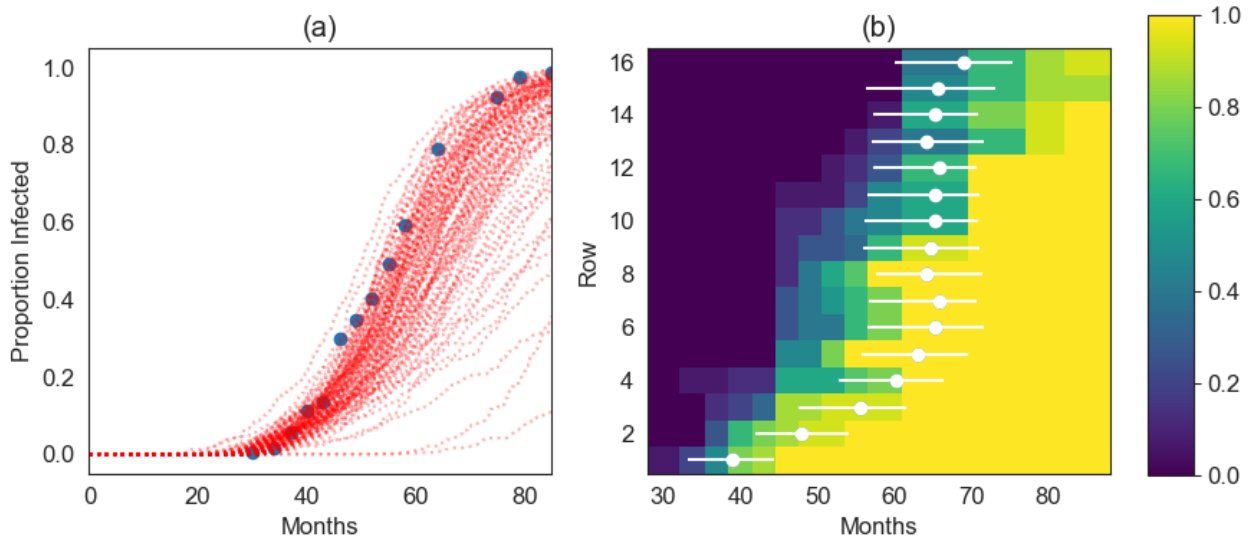

Figure 3: **Stochastic realisations of the lattice-fitted model using the joint posterior distributions.** (a) Time-courses from 100 stochastic simulation runs alongside the data (blue circles). (b) The month when infection arrives on each row. The colormap shows the proportion of trees infected from the data at each time-point. The white dots are the the average of the 100 simulations for the median (8th) tree on the row being infected and the lines for the 4th and 11th trees.

of the lattice model using parameters from the joint posterior distribution in Fig. 4, with all parameters within the 90% smallest credible intervals reported in Table 1. While the expansion is primarily localised, we see occasional longer-range infections which seed new disease foci.

Dispersal model

We now examine the dispersal model to which we compare our lattice model. We again show results in the main text for a threshold of 0.025 for consistency, and show the posteriors for other thresholds in the supplementary information. We note that we do see slightly higher rates of acceptance for this model (3849 accepted sets, which is 0.08% of the 5 million runs), but caution is needed when comparing the acceptance rates with those of the lattice model given the differing natures of the parameters L and α.

The posterior distributions are shown in Fig. 5 and the modes and 90% smallest credibile intervals in Table 2. Similarly to the lattice model, there is clear skew towards more localised infection, with mode $\alpha = 0.74$ (giving a 'real-world' mean dispersal distance of 2.96m), and only 26 parameter sets were below

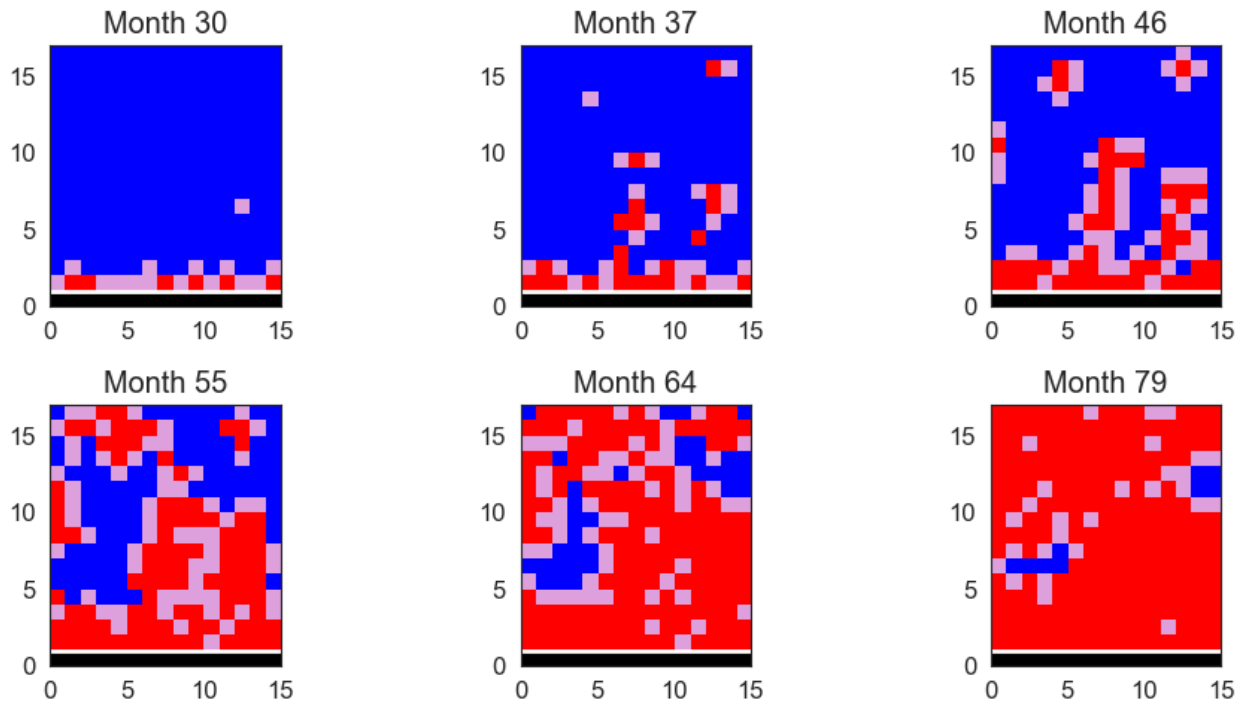

Figure 4: **Snapshots of the grid showing the spatial spread of the disease in one stochastic realisation of the lattice-fitted model.** Blue denotes a cell with a susceptible plant, purple a cell with an exposed plant and red a cell with an infected plant, with the row of founder trees in black. In this simulation run, $L = 0.98$, $\beta = 1.00$ months^{-1}, $\beta_F = 1.61$ months^{-1}, $1/\rho = 11.06$ months and $\delta = 22.60$ months.

the acceptance threshold where $\alpha > 5$. The distributions for the transmission parameters are similar to the lattice model, with a much broader distribution for β_F than for β. Again, similarly to the lattice model, we see a symmetric distribution for ρ with a mode of 8.19 months, and a slightly left-skewed distribution for δ (mode 30.00 months). We again show correlations between the parameters in supplementary Fig. 3, which show similar correlations between β, $1/\rho$ and α, although with much more noise than for the lattice model.

Conducting the posterior predictive checks we see more variation than in the deterministic lattice model (Fig. 6a). As might be expected, they are visually more similar to the stochastic simulations of the lattice model (Fig. 3a). This is again due to the fact that even using the posterior distributions, the stochastic nature of the simulations means that we can have very different time-courses for the same parameter sets.

We also compare the times for the infection to spread to each row (Fig. 6b). We see a clear trend

Parameter	Mode	90% Smallest Credible Interval
α	0.74	[0.35, 2.02]
β (months^{-1})	0.74	[0.25, 4.29]
β_F (months^{-1})	2.93	[0.39, 4.78]
$1/\rho$ (months)	8.19	[2.82, 12.27]
δ (months)	30.00	[25.24, 36.51]

Table 2: Mode parameter values for the dispersal model (using a Gaussian kernel density estimator) and the 90% Smallest Credible Interval/Highest Density Interval from the posterior distributions with an error threshold of 0.025.

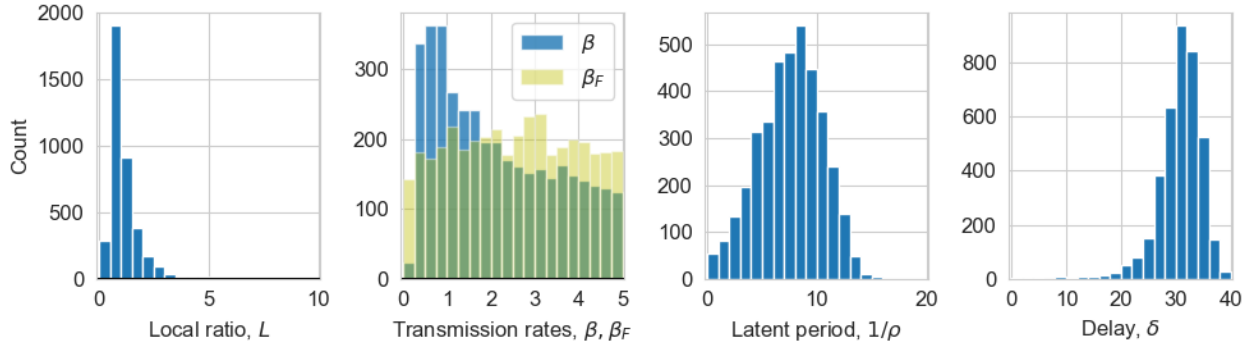

Figure 5: **Posterior distributions for parameters from the dispersal model.** Histograms of the number of simulation runs kept by the ABC procedure for dispersal (α), regular (β) and founder (β_F) transmission, latent period $1/\rho$ and delay, δ.

of increasing times to reach each row, and the dispersal model visually matches the data better than the equivalent plots for the lattice model (Fig. 3b). For the snapshots of the grid from one stochastic realisation of the dispersal model (Fig. 7), with parameters chosen from within the 90% credible intervals in Table 2 we see a steady, though fragmented, spread of the infection through the rows.

Comparison of run times

An important point to note when considering the comparison of the lattice and dispersal models is their run times. Since the lattice model is deterministic, it runs considerably faster than the stochastic dispersal model. As an example, the runtime for a batch of 100,000 simulations of the lattice model on an HPC node

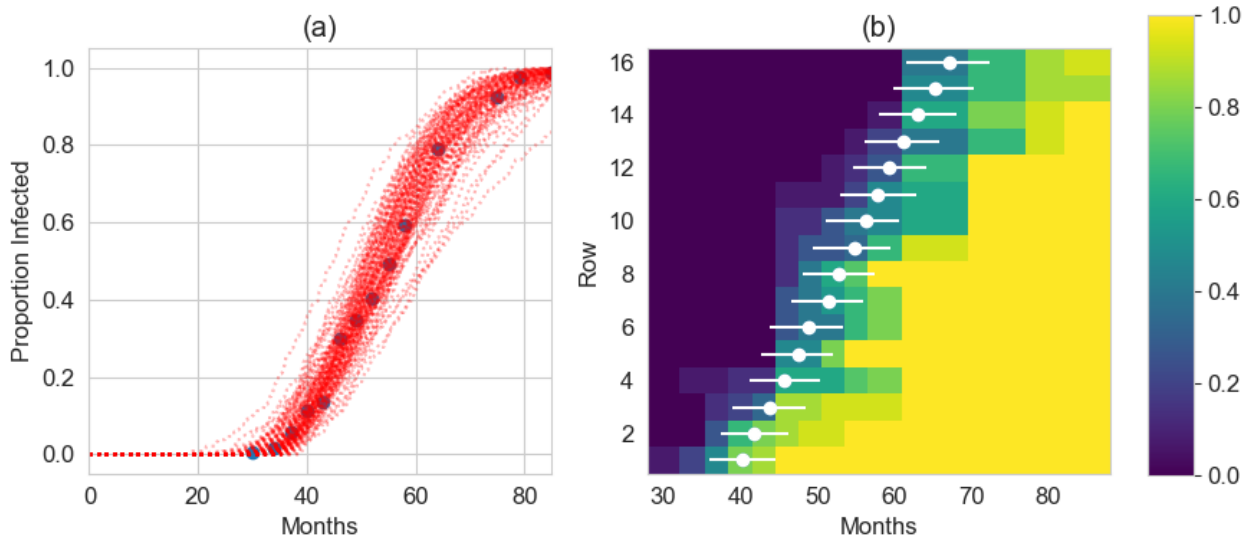

Figure 6: **Stochastic implementations of the dispersal model using the joint posterior distributions .** Left: time-courses from 100 stochastic simulation runs using parameter values selected from the joint posterior distributions alongside the data (blue circles). Right: the month when infection arrives on each row. The colormap shows the proportion of trees infected from the data at each time-point. The white dots are the the average of the 100 simulations for the median (8th) tree on the row being infected and the lines for the 4th and 11th trees.

was around 50 minutes, but for the dispersal model was over 60 hours, more than 70 times slower. While the particular run times are not important, and while efficiency savings might be made (for example by coding in Fortran or C rather than Python), the important conclusion is that based on this result we would certainly expect the lattice model to be at least an order of magnitude faster than the stochastic dispersal model. We note that this runtime for the dispersal model will only increase with larger populations, but should stay near constant for the lattice model since, as a deterministic ordinary differential equation model, there is no dependence on population size.

Discussion

We have demonstrated that a lattice-based, deterministic epidemiological model with a mix of local and global infection can provide a good fit to experimental data concerning the spread of a plant pathogen. In

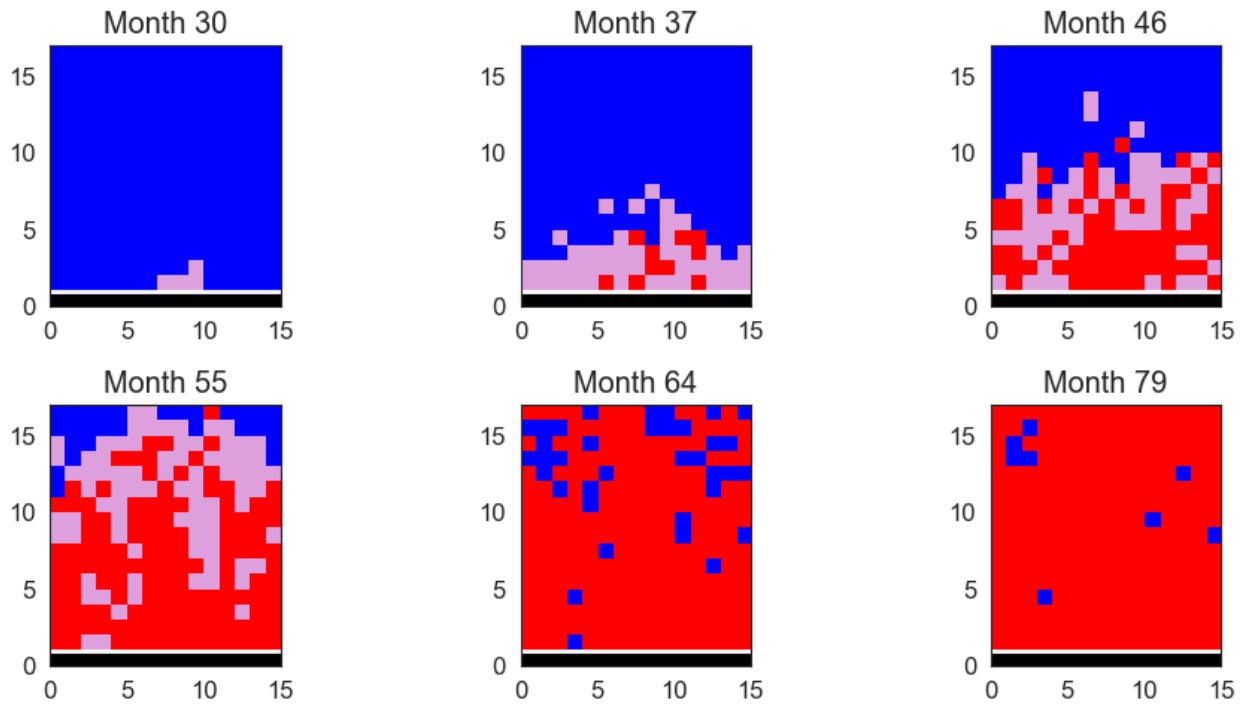

Figure 7: **Snapshots of the grid showing the spatial spread of the disease in one stochastic realisation of the dispersal model.** Blue denotes a cell with a susceptible plant, pink a cell with an exposed plant and red a cell with an infected plant, with the row of founder trees in black. In this simulation run, $\alpha = 0.85$, $\beta = 3.68$,$\beta_F = 2.76$, $1/\rho = 8.65$ and $\delta = 29.70$.

particular, an experimental study tracking spread of Bahia bark scaling of citrus (BBSC) through a grid-like stand of fruit trees [60] is well approximated by a model with predominantly local transmission, but including occasional global (i.e., longer distance) dispersal. In broad terms, and although the wave-like spread of the disease front in space was not fully captured by the lattice model, its performance was similar to that of a dispersal kernel model (a fully spatially explicit individual based model) as used previously both for this pathosystem [61] and a number of others (e.g. citrus canker [69, 42, 70, 41], citrus greening/huanglongbing [71, 72, 43], citrus tristeza virus [73], sudden oak death [74, 75], cassava brown streak [76] and *Xylella fastidiosa* [77]). For BBSC, both the lattice-based and dispersal kernel models predict highly localised spread with modal parameter values of 99% local interactions / 5.6m dispersal distance, a latent/asymptomatic period of 8-11 months and a maturation delay of 27-30 months. Our 90% smallest credible intervals reliably overlay the reported 95% confidence intervals of the previous model based on the dispersal kernel formulation

fitted to these data by Cunniffe et al. [61] .

It might well be questioned how realistic the underlying assumption of the lattice model is, i.e., that infection is *either* local *or* global. Of course, in many real-world systems, the more granular dispersal kernel is likely to be a more realistic interpretation of disease spread. However, we might propose situations where this assumption makes some sense, at least as an initial approximation. For example, spores of *Dothistroma septosporum*, a commercially important disease of pine and fir trees, have two morphs - asexually reproduced conidia that are primarily splash-dispersed over short distances and sexually reproduced ascospores which are wind dispersed over much greater distances [78]; *Hymenoscyphus fraxineus*, the fungal pathogen causing ash dieback, also disperses in this fashion [24]. Similar separations of dispersal scale caused by different types of propagules are a feature of the epidemiology of a number of other diseases caused by ascomycete pathogens of crops, including septoria on wheat [4], powdery mildew on a number of host species [79], and blast diseases of wheat [80] and rice [81]. Although it would also involve a separation of timescales, a somewhat similar mechanism could partially underpin soil-borne pathogen spread patterns, with spread often very restricted within a growing season, but with movement over long distances by cultivation at harvest time [82]. We might also consider vector-borne diseases where there is dimorphism in the vector between winged and non-winged forms. Aphids, for example, have morphs who can only crawl between neighbouring plants and those who can fly long distances [83], likely leading to a mix of more local and more global dispersal. This would be particularly apparent for diseases which are vectored by a mix of colonising aphids and transient aphids [84], a distinction based on whether the aphid population is resident on plants in a particular field or instead only briefly pass through probing a small number of plants essentially at random before flying off elsewhere. Beyond infectious diseases, our model framework would also be relevant for insect pests such as oak processionary moth, where caterpillars will tend to only spread short distances but moths much farther [85].

Perhaps the strongest reason for using the lattice model over the dispersal model is a practical one. Since the lattice-based approach yields a deterministic model, the Monte Carlo fitting procedure, with many thousands of model runs needed to form meaningful posterior distributions, is orders of magnitude quicker compared to a stochastic implementation. While more advanced computational techniques would improve the speed - utilising GPUs for example - this time difference between the approaches will always remain to some extent. This difference in computational speed will only grow as we look at larger populations of plants. The

speed of the deterministic model will remain almost unchanged at higher population sizes, but the stochastic implementations will become significantly slower and may start needing their own approximations, such as tau-leaping, to become practical to investigate [86]. We would thus propose the lattice model as, at the least, an initial choice for investigating disease spread in larger plant populations. Indeed, given the similarity in posterior distributions between the two approaches found here, we might find it useful to take a sequential approach, fitting the lattice model first to find parameter distributions that could then be taken forward to be used as priors in a dispersal model.

While the speed of the model fitting will be considerably faster for the determinstic lattice model in a larger population, the assumption of mean-field transmission will likely become less reasonable. A possible approach here would be to extend the framework of Ellner [50] for multiple interaction scales in a pair approximation model. In the original study, this was applied such that two different ecological processes - competition and infection, say - could be considered to act over different distances. However, we could equally use this approach to define two or more transmission scales, in some sense giving a discretised dispersal kernel. Here we used a simple von Neumann neighbourhood of four near-neighbours on a square lattice, but alternative neighbourhood definitions could also be used to alter the scale of local infection.

The dispersal kernel model used here for comparison with our lattice-based model was inherited from the original study of Cunniffe et al. [61], and used an exponential dispersal kernel. The probability of transmission decays exponentially with distance, a so-called thin-tailed kernel. Fat-tailed kernels - which decay more slowly than exponentially in their tails - are often supported by empirical plant disease data (see Fabre et al. [34] for a review). Such kernels are well known to have important epidemiological consequences, most notably permitting disease wavefronts that accelerate over time [**? ? ?**]. The choice of a thin-tailed kernel in [61] was motivated by apparently highly localised transmission and a relatively constant wave speed in the experimental data, and strong support from exploratory analyses comparing alternative dispersal kernel formulations. In this context, it is therefore striking that we were able to recover a good fit to the experimental data using the local–global model, since the equivalent dispersal structure is in fact closer to a fat-tailed kernel, since the global component of dispersal corresponds to an equal probability of dispersal between any pair of trees, irrespective of distance. However, perhaps as a consequence of the limited spatial extent of the experimental plot, this did not result in substantial systematic lack of fit.

We have used this initial study as a 'proof of concept' of using lattice-based models to study disease

spread in plant populations. Only one previous study had used a lattice-based model to fit to experimental data [59], which assumed entirely near-neighbour infection, and also showed good agreement between the model and data. There are a number of developments that might be made in future work. We may well wish to explore more efficient and/or detailed statistical methods for fitting the models. The ABC rejection approach used here is relatively simple to implement but can have its drawbacks [87]. As already alluded to, the fitting procedure could be improved with a sequential approach - undertaking an initial run to determine the broad shapes of the distributions, and then subsequent runs to explore these in more detail. Recent research has also explored using machine learning to improve the ABC algorithm [88]. It would also be possible to include more detail on the spatial pattern of disease in the pseudo-likelihood used in the ABC procedure, for example by going beyond simple statistics on numbers/pairs of infected/susceptible hosts, and instead including information on the full distribution of distances between infected plants on successive surveys [66] , or the Earth Mover's Distance [89] between the simulation runs and the spatial data. Indeed, the Markov Chain Monte Carlo inference algorithm presented in the original study of [61] used information on the location of each host, as well as its symptom status over time, in order to calculate the full likelihood of the data given model parameters. However, doing so relied on data augmentation, an approach dating back to the pioneering work of [73] (in fact first used in the context of epidemiology for citrus tristeza virus, another citrus disease). Briefly, the approach allows unobserved and/or censored determinants of the model's likelihood function - here the unobservable time of first infection of each host, and the censored time of symptom emergence - to be treated as additional unknown parameters to be estimated within ranges consistent with the constraints from the data, thereby allowing likelihood-based inference to be performed. While MCMC with data augmentation has subsequently been used successfully for a number of plant disease systems (see, e.g., [90, 69, 71, 70]), it tends to be most useful for systems consisting of a relatively small number of individuals, since the number of additional parameters to estimate scales roughly linearly with the system size. ABC approaches are therefore becoming increasingly common for fitting models at larger spatial scales [30].

More practically, an important development will be to use the models to explore prevention and mitigation strategies. For example, roguing, removing (and potentially replacing) symptomatic plants, is commonly used as a control strategy [91]. The effects of this have been explored by a number of mathematical models [92, 93, 94, 95, 96, 61, 97, 72, 98, 45] and could easily be included in our lattice model. Moreover, it would be

relatively straightforward to include a strategy of removing plants that neighbour symptomatic cases in our model set-up. It is clear from the snapshots that, given the incubation period of 8-11 months in both models, there are often a large number of asymptomatic trees present at any time, and only removing symptomatic trees is unlikely to fully control the epidemic [99, 100, 44, 75, 42, 70, 34, 77]. Exploring these management strategies is an important tool offered by mathematical models, and can lead to evidence-based strategies for disease control in real populations.

Acknowledgments

Thanks to Chico Laranjeira for making the BBSC data available via their 2006 publication [60].

References

[1] R. N. Strange and P. R. Scott. Plant disease: a threat to global food security. *Annual Review of Phytopathology*, 43:83–116, 2005.

[2] I. L. Boyd, P. H. Freer-Smith, C. A. Gilligan, and H. C. J. Godfray. The consequence of tree pests and diseases for ecosystem services. *Science*, 342:1235773, 2013.

[3] S. Dong and S. Zhou. Potato late blight caused by *Phytophthora infestans*: From molecular interactions to integrated management strategies. *Journal of Integrative Agriculture*, 21:3456–3466, 2022.

[4] H. Fones and S. Gurr. The impact of *Septoria tritici* blotch disease on wheat: An EU perspective. *Fungal Genetics and Biology*, 79:3–7, 2015.

[5] E.-C. Oerke. Crop losses to pests. *The Journal of Agricultural Science*, 144:31–43, 2006.

[6] S. Savary, L. Willocquet, S.J. Pethybridge, P. Esker, N. McRoberts, and A. Nelson. The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution*, 3:430–439, 2019.

[7] J. Ristaino, K. Anderson, D. P. Bebber, K. A. Brauman, N. J. Cunniffe, N. V. Fedoroff, C. Finegold, K. A. Garrett, C. A. Gilligan, C. Jones, M. D. Martin, G. K. MacDonald, P. Neenan, A. Records, D. G. Schmale, L. Tateosian, and Q. Wei. The persistent threat of emerging plant disease pandemics to global food security. *Proceedings of the National Academy of Sciences*, 118:e2022239118, 2021.

[8] M. J. Jeger, H. Fielder, T. Beale, A. M. Szyniszewska, S. Parnell, and N. J. Cunniffe. What can be learned by a synoptic review of plant disease epidemics and outbreaks published in 2021? *Phytopathology*, 113:1141–1158, 2023.

[9] H. Fielder, T. Beale, M. J. Jeger, G. Oliver, S. Parnell, A. M. Szyniszewska, P. Taylor, and N. J. Cunniffe. A synoptic review of plant disease epidemics and outbreaks published in 2022. *Phytopathology*, 2024.

[10] M. C. Rosace, M. Cendoya, G. Mattion, A. Vicent, A. Battisti, G. Cavaletto, L. Marini, and V. Rossi. A spatio-temporal database of plant pests' first introductions across the EU and potential entry pathways. *Scientific Data*, 10:731, 2023.

[11] D. P. Bebber, T. Holmes, and S. J. Gurr. The global spread of crop pests and pathogens. *Global Ecology and Biogeography*, 23:1398–1407, 2014.

[12] P. K. Anderson, A. A. Cunningham, N. G. Patel, F. J. Morales, P. R. Epstein, and P. Daszak. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology and Evolution*, 19:535–544, 2004.

[13] C.-S. Ma, B.-X. Wang, X.-J. Wang, Q.-C. Lin, W. Zhang, X.-F. Yang, J. van Baaren, D. P. Bebber, S. D. Eigenbrode, M. P. Zalucki, J. Zeng, and G. Ma. Crop pest responses to global changes in climate and land management. *Nature Reviews Earth and Environment*, 6:264–283, 2025.

[14] D. P. Bebber. Range-expanding pests and pathogens in a warming world. *Annual Review of Phytopathology*, 53:335–356, 2015.

[15] B. K. Singh, M. Delgado-Baquerizo, E. Egidi, E. Guirado, J. E. Leach, H. Liu, and P. Trivedi. Climate change impacts on plant pathogens, food security and paths forward. *Nature Reviews Microbiology*, 21: 640–656, 2023.

[16] C. M. Brasier. The biosecurity threat to the UK and global environment from international trade in plants. *Plant Pathology*, 57:792–808, 2008.

[17] B. A. Sikes, J. L. Bufford, P. E. Hulme, J. A. Cooper, P. R. Johnston, and R. P. Duncan. Import

volumes and biosecurity interventions shape the arrival rate of fungal pathogens. *PLOS Biology*, 16: e2006025, 2018.

[18] M. Morelli, J. M. García-Madero, A. Jos, P. Saldarelli, C. Dongiovanni, M. Kovacova, M. Saponari, A. Baños Arjona, E. Hackl, S. Webb, and S. Compant. *Xylella fastidiosa* in olive: A review of control attempts and current management. *Microorganisms*, 9:1771, 2021.

[19] E. Marco-Noales, S. Barbé, A. Monterde, I. Navarro-Herrero, A. Ferrer, V. Dalmau, C. M. Aure, M. L. Domingo-Calap, B. B. Landa, and M. Roselló. Evidence that *Xylella fastidiosa* is the causal agent of almond leaf scorch disease in Alicante, mainland Spain (Iberian Peninsula). *Plant Disease*, 105: 3349–3352, 2021.

[20] R. C. Ploetz. Fusarium wilt of banana. *Phytopathology*, 105:1512–1521, 2015.

[21] T. R. Gottwald. Current epidemiological understanding of citrus huanglongbing. *Annual Review of Phytopathology*, 48:119–139, 2010.

[22] N. M. Salem, A. Jewehan, M. A. Aranda, and A. Fox. Tomato brown rugose fruit virus pandemic. *Annual Review of Phytopathology*, 61:137–164, 2023.

[23] C. Brasier and J. Webber. Plant pathology: sudden larch death. *Nature*, 466:824–825, 2010.

[24] M. Coombes, J. Webber, and L. Boddy. Current understanding and future prospects for ash dieback disease with a focus on Britain. *Forestry*, 97:678–691, 2024.

[25] M. S. Mullett, A. V. Brown, S. Fraser, R. Baden, and K. V. Tubby. Insights into the pathways of spread and potential origins of *Dothistroma septosporum* in Britain. *Fungal Ecology*, 26:85–98, 2017.

[26] N. M. Salem, A. Jewehan, M. A. Aranda, and A. Fox. Implementation of the GB Plant Health and Biosecurity Strategy 2014–2019 with foresight on a new strategy for 2020. *Outlook on Agriculture*, 49: 5–12, 2020.

[27] M. Keeling and P. Rohani. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2008.

[28] L. V. Madden, G. Hughes, and F. van den Bosch. *The Study of Plant Disease Epidemics*. American Phytopathological Society Press, 2007.

[29] D. A. Shah, E. D. De Wolf, P. A. Paul, and L. V. Madden. Functional data analysis of weather variables linked to Fusarium head blight epidemics in the United States. *Phytopathology*, 109:96–110, 2019.

[30] N. J. Cunniffe and C. A. Gilligan. Using mathematical models to predict epidemics and to optimise disease detection and management. In J. B. Ristaino and A. R. Records, editors, *Emerging Plant Diseases and Global Food Security*, chapter 12. American Phytopathological Society Press, 2020.

[31] N. J. Cunniffe, B. Koskella, C. J. E. Metcalf, S. Parnell, T. R. Gottwald, and C. A. Gilligan. Thirteen challenges in modelling plant disease. *Epidemics*, 10:6–10, 2015.

[32] L. E. Wadkin, J. Branson, A. Hoppit, N. G. Parker, A. Golightly, and A. W. Baggaley. Inference for epidemic models with time-varying infection rates: Tracking the dynamics of oak processionary moth in the uk. *Ecology and Evolution*, 12(5):e8871, 2022.

[33] L. V. Madden. Botanical epidemiology: some key advances and its continuing role in disease management. *European Journal of Plant Pathology*, 115:3–23, 2006.

[34] F. Fabre, J. Coville, and N. J. Cunniffe. Optimising reactive disease management using spatially explicit models at the landscape scale. In P. Scott, R. Strange, L. Korsten, and L. Gullino, editors, *Plant Diseases and Food Security in the 21st Century*, chapter 4. Springer, 2021.

[35] E. H. Bussell and N. J. Cunniffe. Optimal strategies to protect a sub-population at risk due to an established epidemic. *Journal of the Royal Society Interface*, 19:20210718, 2022.

[36] L. E. Wadkin, A. Golightly, J. Branson, A. Hoppit, N. G. Parker, and A. W. Baggaley. Quantifying invasive pest dynamics through inference of a two-node epidemic network model. *Diversity*, 15(4), 2023.

[37] Y. D. DeWoody, Z. Feng, and R. K. Swihart. Merging spatial and temporal structure within a metapopulation model. *The American Naturalist*, 166:42–55, 2005.

[38] R. Nathan, E. Klein, J. J. Robeldo-Arnuncio, and E. Revilla. Dispersal kernels: review. In J. Clobert, M. Baguette, T. G. Benton, and J. M. Bullock, editors, *Dispersal Ecology and Evolution*, chapter 15, pages 186–210. Oxford University Press, 2012.

[39] H. Scherm. On the velocity of epidemic waves in model plant disease epidemics. *Ecological Modeling*, 87:217–222, 1996.

[40] F. van den Bosch, J. A. J. Metz, and J. C. Zadoks. Pandemics of focal plant disease, a model. *Phytopathology*, 89:495–505, 1999.

[41] N. J. Cunniffe, R. O. J. H. Stutt, R. E. DeSimone, T. R. Gottwald, and C. A. Gilligan. Optimising and communicating options for the control of invasive plant disease when there is epidemiological uncertainty. *PLOS Computational Biology*, 11:e1004211, 2015.

[42] S. R. Hyatt-Twynam, S. Parnell, R. J. O. H. Stutt, T. R. Gottwald, C. A. Gilligan, and N. J. Cunniffe. Risk-based management of invading plant disease. *New Phytologist*, 214:1317–1329, 2017.

[43] J.M.J. Koh, N.J. Cunniffe, and S. Parnell. Assessing delimiting strategies to identify the infested zones of quarantine plant pests and diseases. *Scientific Reports*, 15:5610, 2025.

[44] N. J. Cunniffe, R. C. Cobb, R. K. Meentemeyer, D. M. Rizzo, and C. A. Gilligan. Modeling when, where, and how to manage a forest epidemic, motivated by sudden oak death in California. *Proceedings of the National Academy of Sciences*, 113:5640–5645, 2016.

[45] J. Ellis, E. Lázaro, B. Duarte, T. Magalhães, A. Duarte, J. Benhadi-Marín, J. Alberto Pereira, A. Vicent, S. Parnell, and N. J. Cunniffe. Developing epidemiological preparedness for a plant disease invasion: Modelling citrus huánglóngbìng in the european union. *Plants, People, Planet*, Early view, 2025.

[46] F. van den Bosch, J. Helps, and N. J. Cunniffe. The basic-reproduction number of infectious diseases in spatially structured host populations. *OIKOS*, page e10616, 2024.

[47] Y. F. Suprunenko, S. J. Cornell, and C. A. Gilligan. Predicting the effect of landscape structure on epidemic invasion using an analytical estimate for infection rate. *Royal Society Open Science*, 12: 240763, 2025.

[48] H. Matsuda, N. Ogita, A. Sasaki, and K. Sato. Statistical mechanics of population: the lattice Lotka-Volterra model. *Progress of Theoretical Physics*, 88(6):1035–1044, 1992.

[49] D. A. Rand, M. Keeling, and H. B. Wilson. Invasion, stability and evolution to criticality in spatially extended, artificial host-pathogen ecologies. *Proceedings of the Royal Society of London B: Biological Sciences*, 259:55–63, 1995.

[50] S. P. Ellner. Pair approximation for lattice models with multiple interactions scales. *Journal of Theoretical Biology*, 210:435–447, 2001.

[51] K. Sato, H. Matsuda, and A. Sasaki. Pathogen invasion and host extinction in lattice structured populations. *Journal of Mathematical Biology*, 32:251–268, 1994.

[52] JAN Filipe and GJ Gibson. Studying and approximating spatio–temporal models for epidemic spread and control. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 353(1378):2153–2162, 1998.

[53] M. J. Keeling. The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London B: Biological Sciences*, 266:859–867, 1999.

[54] S. D. Webb, M. J. Keeling, and M. Boots. Host-parasite interactions between the local and the mean-field: how and when does spatial population structure matter? *Journal of Theoretical Biology*, 249: 140–152, 2007.

[55] A. Maltz and G. Fabricius. Sir model with local and global infective contacts: A deterministic approach and applications. *Theoretical Population Biology*, 112:70–79, 2016.

[56] L. Wren and A. Best. How local interactions impact the dynamics of an epidemic. *Bulletin of Mathematical Biology*, 83:1–14, 2021.

[57] M. Boots and A. Sasaki. 'Small worlds' and the evolution of virulence: infection occurs locally and at a distance. *Proceedings of the Royal Society of London B: Biological Sciences*, 266:1933–1938, 1999.

[58] A. Best, S. D. Webb, A. White, and M. Boots. Host resistance and coevolution in spatially structured populations. *Proceedings of the Royal Society of London B: Biological Sciences*, 278:2216–2222, 2011.

[59] JAN Filipe, W Otten, GJ Gibson, and Gilligan CA. Inferring the dynamics of a spatial epidemic from time-series data. *Bulletin of Mathematical Biology*, 66:373–391, 2004.

[60] F Laranjeira, C Barbosa, H Santos Filho, T Gonçalves, and O Nickel. Progress, spread and natural transmission of bahia bark scaling of citrus in brazil. *Annals of Applied Biology*, 148:187–195, 2006.

[61] N. Cunniffe, F. Laranjeira, F. Neri, E. DeSimone, and C. Gilligan. Cost-effective control of plant disease when epidemiological knowledge is incomplete: Modelling bahia bark scaling of citrus. *PLoS Computational Biology*, 10:e1003753, 2014.

[62] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22:403–434, 1977.

[63] T. McKinley, A. Cook, and R. Deardon. Inference in epidemicmodels without likelihoods. *The International Journal of Biostatistics*, 5:24, 2009.

[64] T. Kypraios, P. Neal, and D. Prangle. A tutorial introduction to Bayesian inference for stochastic epidemic models using Approximate Bayesian Computation. *Mathematical Biosciences*, 287:42–53, 2017.

[65] M. A. Irvine and T. D. Hollingsworth. Kernel-density estimation and approximate bayesian computation for flexible epidemiological model fitting in python. *Epidemics*, 25:80–88, 2018.

[66] A. Minter and R. Retkute. Approximate Bayesian Computation for infectious disease modelling. *Epidemics*, 29:100368, 2019.

[67] A. Cunha Jr, D.A.W. Barton, and T.G. Ritto. Uncertainty quantification in mechanistic epidemic models via cross-entropy approximate Bayesian computation. *Nonlinear dynamics*, 111:9649–9679, 2023.

[68] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

[69] F. M. Neri, A. R. Cook, G. J. Gibson, T. R. Gottwald, and C. A. Gilligan. Bayesian analysis for

inference of an emerging epidemic: citrus canker in urban landscapes. *PLOS Computational Biology*, 10:e1003587, 2014.

[70] H. K. Adrakey, G. Streftaris, N. J. Cunniffe, T. R. Gottwald, C. A. Gilligan, and G. J. Gibson. Evidence-based controls for epidemics using spatio-temporal stochastic models in a bayesian framework. *Journal of The Royal Society Interface*, 14:20170386, 2017.

[71] M. Parry, G. J. Gibson, S. Parnell, T. R. Gottwald, M. S. Irey, T. C. Gast, and C. A. Gilligan. Bayesian inference for an emerging arboreal epidemic in the presence of control. *Proceedings of the National Academy of Sciences*, 111:6258–6262, 2014.

[72] A. P. Craig, N. J. Cunniffe, M. Parry, F. F. Laranjeira, and C. A. Gilligan. Grower and regulator conflict in management of the citrus disease huanglongbing in Brazil: a modelling study. *Journal of Applied Ecology*, 55:1956–1965, 2018.

[73] G. J. Gibson and E. J. Austin. Fitting and testing spatio-temporal stochastic models with application in plant epidemiology. *Plant Pathology*, 45:172–184, 1996.

[74] I. Demon, N. J. Cunniffe, B. P. Marchant, C. A. Gilligan, and F. van den Bosch. Spatial sampling to detect an invasive pathogen outside of an eradication zone. *Phytopathology*, 101:725–731, 2011.

[75] R. N. Thompson, R. C. Cobb, C. A. Gilligan, and N. J. Cunniffe. Management of invading pathogens should be informed by epidemiology rather than administrative boundaries. *Ecological Modelling*, 324: 28–32, 2016.

[76] R. E. Murray-Watson, F. M. Hamelin, and N. J. Cunniffe. How growers make decisions impacts plant disease control. *PLOS Computational Biology*, 18:e1010309, 2022.

[77] M. Cendoya, E. Lazaro, A. Navarro-Quiles, A. Lopez-Quiles, D. Conesa, and A. Vicent. Performance of outbreak management plans for emerging plant diseases: The case of almond leaf scorch caused by *Xylella fastidiosa* in mainland Spain. *Phytopathology*, 114:1566–1576, 2024.

[78] A. Dale, K. Lewis, and B. Murray. Sexual reproduction and gene flow in the pine pathogen dothistroma septosporum in british columbia. *Phytopathology*, 101:68–76, 2011.

[79] D. A. Glawe. The powdery mildews: A review of the world's most familiar (yet poorly known) plant pathogens. *Annual Review of Phytopathology*, 46:27–51, 2008.

[80] P. C. Ceresini, V. L. Castrogudin, F. A. Rodrigues, J. A. Rios, C. E. Aucique-Perez, S. I. Moreira, E. Alves, D. Croll, and J. L. N. Maciel. Wheat blast: Past, present, and future. *Annual Review of Phytopathology*, 56:427–456, 2018.

[81] N. J. Talbot. On the trail of a cereal killer: Exploring the biology of *Magnaporthe grisea*. *Annual Review of Microbiology*, 57:177–202, 2003.

[82] J. E. Truscott and C. A. Gilligan. The effect of cultivation on the size, shape, and persistence of disease patches in fields. *Proceedings of the National Academy of Sciences*, 98:7128–7133, 2001.

[83] H. van Emden and R. Harrington. *Aphids as crop pests*. CABI, 2007.

[84] M. Zaffaroni, L. Rimbaud, L. Mailleret, N. J. Cunniffe, and D. Bevacqua. Modelling interference between vectors of non-persistently transmitted plant viruses to identify effective control strategies. *PLOS Computational Biology*, 17:e1009727, 2021.

[85] I. Tomlinson, C. Potter, and H. Bayliss. Managing tree pests and diseases in urban settings: The case of oak processionary moth in london, 2006–2012. *Urban Forestry and Urban Greening*, 14:286–292, 2015.

[86] Y. Cao, D. T. Gillespie, and L. R Petzold. Efficient step size selection for the tau-leaping simulation method. *The Journal of Chemical Physics*, 124:044109, 2006.

[87] C. P. Robert. Bayesian computational tools. *Annu. Rev. Stat. Appl.*, 1:153–177, 2014.

[88] T. Sanchez, J. Cury, G. Charpiat, and F. Jay. Deep learning for population size history inference: Design, comparison and combination with Approximate Bayesian Computation. *Molecular Ecology Resources*, 21(8):2645–2660, 2021.

[89] Y. Rubner, C. Tomasi, and L.J. Guibas. The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 40:99–121, 2000.

[90] R. K. Meentemeyer, N. J. Cunniffe, A. R. Cook, J. A. N. Filipe, R. D. Hunter, D. M. Rizzo, and C. A. Gilligan. Epidemiological modeling of invasion in heterogeneous landscapes: spread of sudden oak death in California (1990–2030). *Ecosphere*, 2:1–24, 2011.

[91] M. Dasgupta. *Principles of plant pathology.* Allied Publishers, 1999.

[92] M.-S. Chan and M. J. Jeger. An analytical model of plant virus disease dynamics with roguing and replanting. *Journal of Applied Ecology*, 31:413–427, 1994.

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

[94] M. J. Jeger, J. Holt, F. van den Bosch, and L. V. Madden. Epidemiology of insect-transmitted plant viruses: modelling disease dynamics and control interventions. *Physiological Entomology*, 29:291–304, 1994.

[95] K. Jacobsen, J. Stupiansky, and S. S. Pilyugin. Mathematical modeling of citrus groves infected by huanglongbing. *Mathematical Biosciences & Engineering*, 10(3):705–728, 2013.

[96] M. S. Sisterson and D. C. Stenger. Roguing with replacement in perennial crops: conditions for successful disease management. *Phytopathology*, 103:117–128, 2013.

[97] F. M. Hilker, L. J. S. Allen, V. A. Bokil, C. J. Briggs, Z. Feng, K. A. Garrett, L. J. Gross, F. M. Hamelin, M. J. Jeger, C. A. Manore, A. G. Power, M. G. Redinbaugh, M. A. Rua, and N. J. Cunniffe. Modeling virus coinfection to inform management of maize lethal necrosis in Kenya. *Phytopathology*, 107:1095–1108, 2017.

[98] R. E. Murray-Watson and N. J. Cunniffe. How the epidemiology of disease-resistant and disease-tolerant varieties affects grower behaviour. *Journal of the Royal Society Interface*, 19:20220517, 2022.

[99] S. Parnell, T. Gottwald, F. van den Bosch, and C. Gilligan. Optimal strategies for the eradication of asiatic citrus canker in heterogeneous host landscapes. *Phytopathology*, 99:1370–1376, 2009.

[100] S. Parnell, T. R. Gottwald, C. A. Gilligan, N. J. Cunniffe, and F. van den Bosch. The effect of landscape pattern on the optimal eradication zone of an invading epidemic. *Phytopathology*, 100:638–644, 2010.