Rothamsted Repository Download

A - Papers appearing in refereed journals

Cunniffe, N. J., Stutt, R. O. J. H., Van Den Bosch, F. and Gilligan, C. A. 2012. Time-dependent infectivity and flexible latent and infectious periods in compartmental models of plant disease. *Phytopathology.* 102, pp. 365-380.

The publisher's version can be accessed at:

- https://dx.doi.org/10.1094/PHYTO-12-10-0338

The output can be accessed at: https://repository.rothamsted.ac.uk/item/8qw3y/time-dependent-infectivity-and-flexible-latent-and-infectious-periods-in-compartmental-models-of-plant-disease.

Time-Dependent Infectivity and Flexible Latent and Infectious Periods in Compartmental Models of Plant Disease

N. J. Cunniffe, R. O. J. H. Stutt, F. van den Bosch, and C. A. Gilligan

First, second, and fourth authors: Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK; and third author: Biomathematics and Bioinformatics Division, Rothamsted Research, Harpenden, AL5 2JQ, UK.

ABSTRACT

Cunniffe, N. J., Stutt, R. O. J. H., van den Bosch, F., and Gilligan, C. A. 2012. Time-dependent infectivity and flexible latent and infectious periods in compartmental models of plant disease. Phytopathology 102:365-380.

Compartmental models have become the dominant theoretical paradigm in mechanistic modeling of plant disease and offer well-known advantages in terms of analytic tractability, ease of simulation, and extensibility. However, underlying assumptions of constant rates of infection and of exponentially distributed latent and infectious periods are difficult to justify. Although alternative approaches, including van der Plank's seminal discrete time model and models based on the integrodifferential formulation of Kermack and McKendrick's model, have been suggested for plant disease and relax these unrealistic assumptions, they are challenging to implement and to analyze. Here, we propose an extension to the susceptible, exposed, infected, and removed (SEIR) compartmental model, splitting the latent and infection compartments and thereby allowing time-varying infection rates and more realistic distributions of latent and infectious periods to be represented. Although the model is, in fact, more general, we specifically target plant disease by demonstrating how it can represent both the van der Plank model and the most commonly used variant of the Kermack and McKendrick (K & M) model (in

which the infectivity response is delay Gamma distributed). We show how our reformulation retains the numeric and analytic tractability of SEIR models, and how it can be used to replicate earlier analyses of the van der Plank and K & M models. Our reformulation has the advantage of using elementary mathematical techniques, making implementation easier for the nonspecialist. We show a practical implication of these results for disease control. By taking advantage of the easy extensibility characteristic of compartmental models, we also investigate the effects of including additional biological realism. As an example, we show how the more realistic infection responses we consider interact with host demography and lead to divergent invasion thresholds when compared with the "standard" SEIR model. An ever-increasing number of analyses purportedly extract more biologically realistic invasion thresholds by adding additional biological detail to the SEIR model framework; we contend that our results demonstrate that extending a model that has such a simplistic representation of the infection dynamics may not, in fact, lead to more accurate results. Therefore, we suggest that modelers should carefully consider the underlying assumptions of the simplest compartmental models in their future work.

Additional keywords: infection kernel.

The earliest mechanistic model of plant disease is due to van der Plank (71), and uses a delay differential equation to represent $I(t)$, the density of host tissue first infected at or before time t,

$$\frac{dI(t)}{dt} = R\big(I(t-p) - I(t-p-i)\big)\big(1-I(t)\big). \tag{1}$$

The latent period (p) and the infectious period (i) are constant, and the parameter R is the corrected basic rate of infection. The key threshold is the "progeny-parent ratio", iR: a value >1 indicates that the density of infected tissue will increase. Unfortunately delay differential equations are difficult to analyze (40, 47) and, despite widespread adoption of discrete-time approximations to van der Plank's model in early simulations of plant disease (63,74,78) and a number of often subtle mathematical analyses that followed (28–30,32,37,73), the model is now rarely used in theoretical studies. However, because the model was so influential, it is still of significant historical interest and, indeed, a recent textbook suggested (of van der Plank and equation 1) "he developed this model in 1963, a year that could be seen as the starting of quantitative plant disease epidemiology" (41).

Contemporary work has focused on compartmental SEIR-type (susceptible, exposed, infected, and removed) models (15,16,40,

61), adapting a structure originally proposed for models of human/animal disease (1,2,31,32). The population of hosts is divided according to disease status: susceptible (S), exposed (E), infected (I), and removed (R). Changes in the number of hosts in each compartment follow coupled nonlinear differential equations, which for the basic SEIR model are

$$\begin{aligned}
\frac{dS}{dt} &= -\beta SI, \\
\frac{dE}{dt} &= \beta SI - \gamma E, \\
\frac{dI}{dt} &= \gamma E - \mu I, \\
\frac{dR}{dt} &= \mu I.
\end{aligned} \tag{2}$$

In equation 2, β is the per capita rate of infection of susceptible hosts, γ is the rate at which exposed (i.e., latently infected) hosts become infectious, and μ is the rate at which infected hosts become removed (i.e., become epidemiologically inert by losing infectivity or dying). The corresponding mean latent and infectious periods are $1/\gamma$ and $1/\mu$, respectively, and implicit in the formulation is that these periods are distributed exponentially (i.e., that the rate of leaving a class is independent of the time already spent within it). Although the exponential distributions are selected to ensure that, on average, each host is latently infected or infectious for the correct period, the formulation means

that the latent or infectious period of a particular host may be infinitesimally short which, for plant disease, is unrealistic. The definition of host is flexible and although, by default, a host corresponds to an individual plant, it could potentially represent a plant organ such as an individual leaf or root (6,10), or even large assemblages of plants such as a field or farm (55), depending on the particular scale of interest (15,16). It is also possible for hosts in the model to correspond to individual susceptible sites that could be infected by the pathogen, allowing the results of the model to be directly translated to severity of disease. In the context of plant disease, the S, E, and I compartments are sometimes redubbed healthy (H), latent (L), and infectious (I), respectively (41), although here we exclusively use the S-E-I-R notation and nomenclature that is more typical of the wider literature.

The key threshold is the "basic reproductive number" (sometimes "ratio"), R_0, which is the average number of new infections caused when a single infected host is introduced into a continuously completely susceptible population (14,35) (compare the progeny/parent ratio of van der Plank). For a population of N hosts, the basic reproductive number according to equation 2 is $R_0 = \beta N/\mu$. Compartmental models are currently the dominant mechanistic theoretical paradigm, largely because of the increased flexibility they offer, in particular the ease with which additional biology may be included (15,16). For example, although exponential increase in the density of susceptible tissue was considered by van der Plank himself, more realistic bounded host growth was only analyzed in that framework after a long delay and with some difficulty (37). In contrast, updating the dynamics of the susceptible population in the SEIR model requires only a minor and obvious change to the equation governing the rate of change of susceptible hosts (10). Furthermore, it is the ease of associating parameters and biological processes that distinguishes the mechanistic models we consider in this article from purely descriptive growth functions such as exponential, monomolecular, logistic, and Gompertz. We note that these growth curves are, in fact, more widely used than any of the models we discuss but, because growth curves aim only to summarize rather than to explain, they are often of relatively little interest to most theoreticians.

Following early work that demonstrated that the thresholds $iR > 1$ and $R_0 > 1$ are equivalent (28), the van der Plank and SEIR models were unified for botanical epidemics by Segarra et al. (59), who illustrated how both formulations are special cases of the original Kermack and McKendrick (K & M) epidemic model (36). For sporulating plant pathogens, the K & M model requires a sporulation curve, $\theta(\tau)$, which represents the number of viable spores produced by an infected host τ units of time since it was initially infected. The formulation can model a latent period by setting $\theta(\tau) = 0$ for small τ, can model a finite infectious period by setting $\theta(\tau) = 0$ for large τ, and can readily be generalized to pathogens that do not sporulate (e.g.. soilborne pathogens) by reinterpreting $\theta(\tau)$ as a time-dependent infectivity, although here we will refer to sporulation for concreteness. If the probabilities of a spore being deposited on a particular host and of initiating a new infection conditional upon deposition are ζ and ψ, respectively, then during any epidemic the number of healthy hosts, $S(t)$, evolves according to

$$\frac{dS(t)}{dt} = \zeta \psi S(t) \int_{\tau=0}^{t} \frac{dS(t-\tau)}{dt} \theta(\tau) d\tau, \tag{3}$$

where the integral represents the cumulative production of spores by all infected hosts according to time since infection (because at time t, $-dS(t - \tau)/dt$ is the number of infections produced exactly τ units of time ago and, therefore, that have infectivity $\theta(\tau)$ at time t). Extensions to equation 3 explicitly include an ongoing source of primary inoculum (59) but these elaborations are ignored here for simplicity.

By setting $\theta(\tau)$ appropriately in equation 3, both the van der Plank and SEIR models can be formulated as special cases of the K & M model. However, a number of authors (initially van den Bosch et al., 70) suggest that experimental data measuring sporulation as a function of time is better fitted by the delay Gamma-distributed sporulation curve with parameters k and λ (42,52,53)

$$\theta(\tau) = \begin{cases} 0 & \tau \le p \\ \Omega\left(\dfrac{\lambda^k (\tau - p)^{k-1} e^{-\lambda(\tau - p)}}{\Gamma(k)} \right) & \tau > p \end{cases}. \tag{4}$$

In equation 4, Ω is the total number of spores produced by a single infected host, p is the (fixed) latent period, and λ and k control the time-dependence of the sporulation response. The analysis in the article of Segarra et al. (59) restricts the shape parameter k of the Gamma distribution to be an integer, but this is not too constraining in practice (Appendix 1). However, despite the more realistic representation of sporulation promoted by the K & M model with this infection kernel (the Gamma(k) model), it has subsequently been little used, arguably because further analyses would be challenging and require mathematical or numerical machinery that is unfamiliar to many epidemiologists.

Equations 3 and 4 indicate that $\Omega\zeta\psi N > 1$ is required for the number of infected hosts to increase initially (in a susceptible population of size N). As shown by Segarra et al. (59), this invasion threshold is equivalent to $iR > 1$ and $R_0 > 1$ in the van der Plank and SEIR models, respectively. The final level of infected tissue after a very long time (which we denote as f, the fraction of nonsusceptible hosts at the end of any epidemic, and which we note can take values infinitesimally close to 1) is also conserved across the models, assuming that there are few infected hosts initially. However, the rate of exponential increase of disease in the early part of any epidemic (r) differs depending on the model in question.

In this article, we introduce a suite of SEIR-type models with multiple exposed (i.e., latently infected) and infected (i.e., infectious) compartments (also referred to as classes). Subdividing the exposed and infected classes is increasingly common in models for human and animal disease (34,38,39,50,75) and allows more flexible distributions of latent and infectious periods by relaxing the assumption of exponentially distributed times implicit in equation 2. By carefully selecting the number of compartments and the values of the parameters, we adapt this technique specifically to plant disease, and represent the van der Plank and Gamma(k) models in the compartmental framework of the SEIR model, although we note that the extension we propose is actually more general and additionally is able to represent other responses (Appendix 1). We demonstrate how the multiple compartment reformulation allows us to (i) simplify the analysis of Segarra et al. (59) and characterize the similarities and differences between the SEIR, van der Plank, and Gamma(k) models (in terms of R_0, f, and r) using only elementary mathematical techniques; (ii) simulate the models using standard and widely available numerical routines and packages for solutions of coupled systems of differential equations; and (iii) extend the models very easily; we illustrate this by adding (a) host demography and (b) free-living infectious material.

Additionally, we show how the differences between the SEIR, van der Plank, and Gamma(k) models can be of practical significance. In particular, we note that the nature of the underlying model used to approximate an epidemic affects the inferred relationship between the initial rate of exponential increase of disease (r) (which is often used to estimate early epidemic spread) and the basic reproductive number (R_0) (frequently used as a measure of ability to invade). Moreover, because the final fraction of infected hosts (f) is controlled by R_0, it follows that fitting an inappropriate model formulation to data from the early part of an epidemic leads to an incorrect assessment of the final severity

(Fig. 1). However, data from the early part of the epidemic are all that are available when decisions about control are most pressing. Therefore, any assessment of the level of control required or, indeed, whether control is even necessary, depends upon the model chosen to perform the analysis; we use our theoretical results to illustrate how significant this effect can be.

THEORY AND APPROACHES

We extend the SEIR model (equation 2) to include m exposed and n infected compartments (the SEmInR model). Hosts in the ith infected class infect susceptible hosts at per capita rate β_i, and transition rates within and between exposed and infected compartments are scaled by m and n such that the average latent period is fixed at $1/\gamma$ and the average infectious period is fixed at $1/\mu$ (Fig. 2A). The corresponding system of differential equations is

$$\frac{dS}{dt} = -\left(\sum_{j=1}^{n}\beta_j I_j\right)S,$$

$$\frac{dE_1}{dt} = \left(\sum_{j=1}^{n}\beta_j I_j\right)S - m\gamma E_1,$$

$$\frac{dE_j}{dt} = m\gamma E_{j-1} - m\gamma E_j, \quad (2 \leq j \leq m)$$

$$\frac{dI_1}{dt} = m\gamma E_m - n\mu I_1,$$ (5)

$$\frac{dI_j}{dt} = n\mu I_{j-1} - n\mu I_j, \quad (2 \leq j \leq n)$$

$$\frac{dR}{dt} = n\mu I_n.$$

Therefore, an individual host enters the E_1 compartment immediately after being exposed to the pathogen, and moves through the m exposed compartments and then the n infected compartments in strict sequence. While the host remains in one of the exposed compartments, it is incapable of causing infection in other hosts (i.e., it is latently infected). Subdivision of the exposed and infected compartments does not correspond to any biological difference between individuals in the different exposed and infected classes. However, and as will be shown below, using multiple compartments is, instead, a mathematical device to ensure that latent and infectious periods are correctly distributed and which (in combination with the distinct rates of infection in the different infectious compartments, $\beta_1 \ldots \beta_n$) allows much flexibility in the infection rate of infectious hosts as a function of time since infection.

We show how the van der Plank, SEIR, and Gamma(k) models can be represented using equation 5 by choosing appropriate relationships between the values of β_i, m, and n. The basic reproductive number, R_0, is derived by considering a single infected host introduced into a population of N susceptibles (13). By finding a conserved quantity (i.e., a functional expression linking the values of the state variables that remains constant for all time) according to equation 5 and equating its values at the beginning and end of any epidemic, we determine the eventual impact of the pathogen (as f, the fraction of nonsusceptible hosts at the end of the epidemic). Finally, the initial rate of density-independent exponential disease increase, r, is obtained by linearizing equation 5 about the pathogen-free equilibrium with $S = N$ and $E_i = I_i = R = 0$.

We use these results for the general SEmInR system and the relationships between β_i, m, and n for the compartmental analogues of the van der Plank, SEIR, and Gamma(k) models to determine R_0, f, and r for each of these three models and, thereby, replicate the results of Segarra et al. (59). Only r changes between the three models we consider. The relationship between r and R_0 depends not only on which of the van der Plank, SEIR, or Gamma(k) models is in question but also on the duration of the latent or infectious periods. This is investigated numerically. In

particular, we choose illustrative latent and infectious periods ($1/\gamma$ = 5, 10, or 20 days, and $1/\mu$ = 5, 10, 20, 50, or 100 days) and investigate the response of r to changes in R_0 for each model. Also, because the van der Plank and Gamma(k) models are theoretically obtained in the limit of equation 5 with an infinite number of exposed or infected classes, we considered the accuracy of numerical simulation using the SEmInR model with large, but finite, values of m or n. In these (and all) numerical simulations, for the Gamma(k) model, we focus upon the particular case $k = 4$ but the qualitative differences we present are independent of $k > 1$.

We illustrate the extensibility of our framework by considering two elaborations to the SEmInR model. In particular, we add (i) host demography and (ii) free-living infectious material. Host demography is typically included in models targeting pathogens of long-lived hosts (62), pathogens capable of undergoing many generations within a single season (15), or soilborne pathogens for which root production and growth (and so, implicitly, the production of new tissue) is important in bringing susceptible host tissue within striking distance of immobile infective propagules (6,10,60). Models of soilborne pathogens often also include free-living infectious material (2–5,10,17,20–24), representing a reservoir of infective inoculum in the soil that is replenished by infected hosts. This allows the dual pathways of primary and secondary infection characteristic of these pathosystems (1,10,24) to be included in the models, and their effects on the epidemiology to be separately assessed.

In adding host demography (Fig. 2B), we assume that all hosts are subject to removal at per capita rate g, corresponding to either natural death or harvesting, and imposing an average lifetime of $1/g$. By assuming that susceptible hosts are replenished at rate $q(S)$, we allow the pathogen and host to coexist (if $R_0 > 1$). The model equations become

$$\frac{dS}{dt} = q(S) - \left(\sum_{j=1}^{n}\beta_j I_j\right)S - gS,$$

$$\frac{dE_1}{dt} = \left(\sum_{j=1}^{n}\beta_j I_j\right)S - (m\gamma + g)E_1,$$ (6)

$$\frac{dE_j}{dt} = m\gamma E_{j-1} - (m\gamma + g)E_j, \quad (2 \leq j \leq m)$$

$$\frac{dI_1}{dt} = m\gamma E_m - (n\mu + g)I_1,$$

$$\frac{dI_j}{dt} = n\mu I_{j-1} - (n\mu + g)I_j, \quad (2 \leq j \leq n)$$

$$\frac{dR}{dt} = n\mu I_n.$$

The function $q(S)$ in equation 6 is chosen to impose a carrying capacity (i.e., a maximal number of infectible sites, organs, plants, farms, or fields, depending on the spatial scale of the model) in the absence of the pathogen: simple choices of $q(S)$ allow commonly used responses, including monomolecular and logistic, to be recovered (10). We also note, however, that the particular form of $q(S)$ is unimportant and the analysis we present depends only upon susceptible hosts in the absence of the pathogen having a carrying capacity which, for ease of comparison, we assume is N. We calculate R_0 and r for the updated model in equation 6, and demonstrate how host demography has differing effects in each of the SEIR, van der Plank, and Gamma(k) models. We also investigate the level of healthy hosts in the population when the pathogen is endemic, and show how this key metric of disease impact is affected by the value of R_0 and thus, in turn, by the choice of model when host demography is included (for pathogens with latent and infectious periods and rates of infection that are held constant when comparing the different models).

To demonstrate how our model can be extended specifically to target soilborne pathogens (1–6,8,10,16,18,43), we introduce

free-living infectious material (i.e., soilborne inoculum). However, in addition to the multicompartmental representation of exposed or infected hosts, we allow some flexibility in the distribution of the infectious period of free-living infectious tissue by using p subcompartments for inoculum (X_j). The mean infectious period of inoculum is fixed at $1/c$. We assume that infectious hosts in class I_j continuously produce inoculum at particular per capita rate α_j and that inoculum in class X_j causes primary infection at rate β_{Xj}. The equations become (Fig. 2C)

$$\frac{dS}{dt} = -\left(\sum_{j=1}^{n}\beta_j I_j\right)S - \left(\sum_{j=1}^{p}\beta_{Xj}X_j\right)S,$$

$$\frac{dE_1}{dt} = \left(\sum_{j=1}^{n}\beta_j I_j\right)S + \left(\sum_{j=1}^{p}\beta_{Xj}X_j\right)S - m\gamma E_1,$$

$$\frac{dE_j}{dt} = m\gamma E_{j-1} - m\gamma E_j, \quad (2 \le j \le m)$$

$$\frac{dI_1}{dt} = m\gamma E_m - n\mu I_1,$$ (7)

$$\frac{dI_j}{dt} = n\mu I_{j-1} - n\mu I_j, \quad (2 \le j \le n)$$

$$\frac{dX_1}{dt} = \sum_{j=1}^{n}\alpha_j I_j - pcX_1,$$

$$\frac{dX_j}{dt} = pcX_{j-1} - pcX_j, \quad (2 \le j \le p)$$

$$\frac{dR}{dt} = n\mu I_n.$$

By calculating R_0, r, and f for this SEmInXpR model, we demonstrate how our analysis can easily be extended to include this more complex biology.

Finally, we illustrate a practical implication of our results by considering disease control. We consider the percentage change in the values of R_0 and f corresponding to $r = 0.01$ day^{-1} between pairs of models for a range of values of $1/\gamma$ and $1/\mu$ up to 20 and 100 days, respectively. This allows us to demonstrate how misspecification of the distribution of latent and infectious periods or time-varying infectivity can lead to incorrect assessment of the final severity of an epidemic from the data available when it is emerging and therefore can, in turn, lead to poor decisions on the level of control to exert.

RESULTS

For convenience, the principal analytic results are summarized in Table 1.

Preliminary: matching parameters between model frameworks. To compare the models directly, the parameters must be normalized. To match the latent periods across the models, the average latent period in the SEIR model (equation 2), $1/\gamma$, must equal the constant delay before sporulation, p, in the van der Plank (equation 1) and Gamma(k) (equations 3 and 4) models. The average duration of an infection cycle in the SEIR model ($1/\gamma + 1/\mu$) must equal its equivalent in the van der Plank and the Gamma(k) models ($p + i$ and $p + k/\lambda$, respectively), and because $1/\gamma = p$, this requires $1/\mu = i = k/\lambda$. Finally, to ensure that the average infected host causes the same number of infections throughout its infectious period, the rate of infection in the
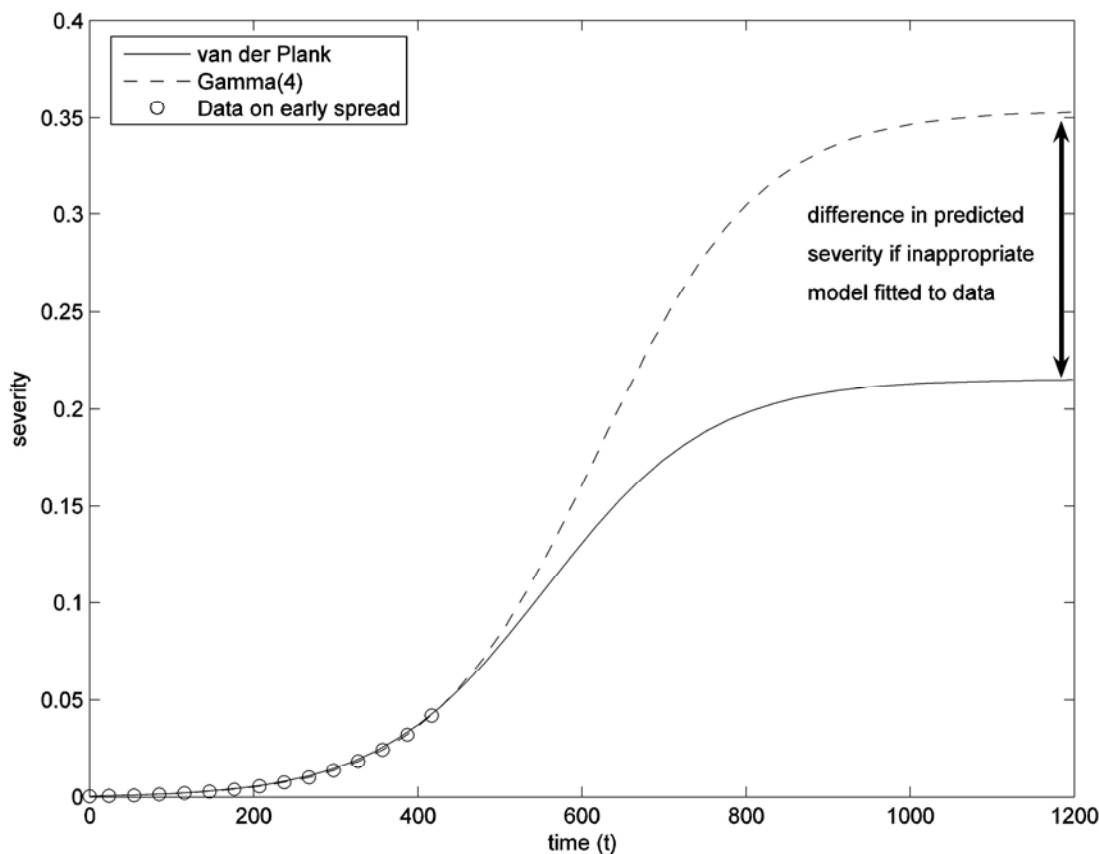

Fig. 1. Final size of an epidemic can vary widely between the models we consider, even if the initial rate of exponential increase r is held constant. This means that, if data from early spread of an epidemic are used to estimate final disease severity (and possibly, in turn, to determine the level of control to exert), the result depends upon the model formulation that is chosen. This is illustrated for a pathogen with latent period $1/\gamma = 2$ days and infectious period $1/\mu = 20$ days, and with infection rate selected to give an initial growth rate $r = 0.01$ according to the van der Plank (solid) and Gamma(4) (dashed) models. Disease progress curves estimated from the early spread of the disease are initially very similar but diverge and give a very different final severity of disease. Note that, for these epidemiological parameters and for this rate of increase, the SEIR model gives a very similar response to that in the Gamma(4) model, and it is omitted.

SEIR model, β, must be related to the van der Plank corrected basic rate of infection via $R = \beta N$, and to the sporulation rate in the Gamma(k) model via $\beta/\mu = \Omega\zeta\psi$. In combination, these choices allow the models to be directly compared, and ensure the thresholds $iR > 1$, $R_0 > 1$, and $\Omega\zeta\psi N > 1$ are equivalent. Henceforth, we will work only in terms of the parameters of the equivalent SEIR model (i.e., β, γ, and μ); however, note that all results can be back-translated by reversing the mapping described above.

Representing the models in the SEmInR framework. If $m = n = 1$, there is only one exposed and one infectious class and, therefore, the SEIR model is immediately recovered on making this choice and setting $\beta_1 = \beta$.

In the more general SEmInR model, the overall latent period is the sum of m exponential variates each of mean $1/m\gamma$. By the central limit theorem, for large m, the resulting Gamma-distributed variate is well approximated by a normal distribution with mean $1/\gamma$ and variance $1/\gamma^2 m$. Therefore, in the limit $m \to \infty$, the distribution of times of leaving the final exposed class tends to a δ function (i.e., a sharp infinite spike), and the corresponding latent period becomes fixed at exactly $1/\gamma$ (Fig. 3). Similarly, to represent a fixed infectious period, the number of infective compartments, n, should tend to infinity. Hence, to replicate the van der Plank model in the compartmental framework of the SEmInR model, we require $m,n \to \infty$, with $\beta_i = \beta$ for all i (because, in the van der Plank model, the infection rate remains constant throughout the infectious period).

To replicate the Gamma(k) model, we initially consider how to use the SEmInR model to represent a pathogen with this sporulation response, but with no latent period. The fixed latent period in the Gamma(k) model is subsequently added back in. Consider a single infected host that moves from the final exposed to the first infectious compartment (i.e., from compartment E_m to compartment I_1) at $t = 0$, with no further new infections thereafter. According to equation 5, subsequent dynamics follow

$$\frac{dI_1}{dt} = -n\mu I_1,$$

$$\frac{dI_2}{dt} = n\mu I_1 - n\mu I_2,$$ (8)

$$\vdots$$

$$\frac{dI_n}{dt} = n\mu I_{n-1} - n\mu I_n,$$

with $I_1(0) = 1$, $I_i(0) = 0$ for $i > 1$. Recursively solving these equations

$$I_i(t) = \frac{(n\mu t)^{i-1}}{(i-1)!} e^{-n\mu t},$$ (9)

and, thus, the equivalent sporulation curve in the K & M model (equation 5) is (13)

$$\theta(\tau) = \sum_{i=1}^{n} \beta_i \frac{(n\mu\tau)^{i-1}}{(i-1)!} e^{-n\mu\tau}.$$ (10)

This response is proportional to that in the Gamma(k) model if we set $n = k$ and $\beta_i = 0$ for $i < n = k$, because $\Gamma(k) = (k-1)!$ if k is an integer. However, the rate of infection in the nth compartment, β_n, must be scaled by the number of compartments to obtain the same overall rate of sporulation; therefore, we set $\beta_n = n\beta$. Because the fixed latent period can be added back in by taking an infinite number of exposed compartments, we can represent the Gamma(k) model in the SEmInR framework by taking $m \to \infty$, $n = k$, and $\beta_i = 0$ for $i < n$, $\beta_n = n\beta$. However, we note that the response according to equation 10 is actually more general and, in principle, can represent a wider variety of forms than the Gamma-distributed response (Appendix 1).

Basic reproductive number, R_0. R_0 for the general SEmInR model follows directly from equation 5 by considering a single
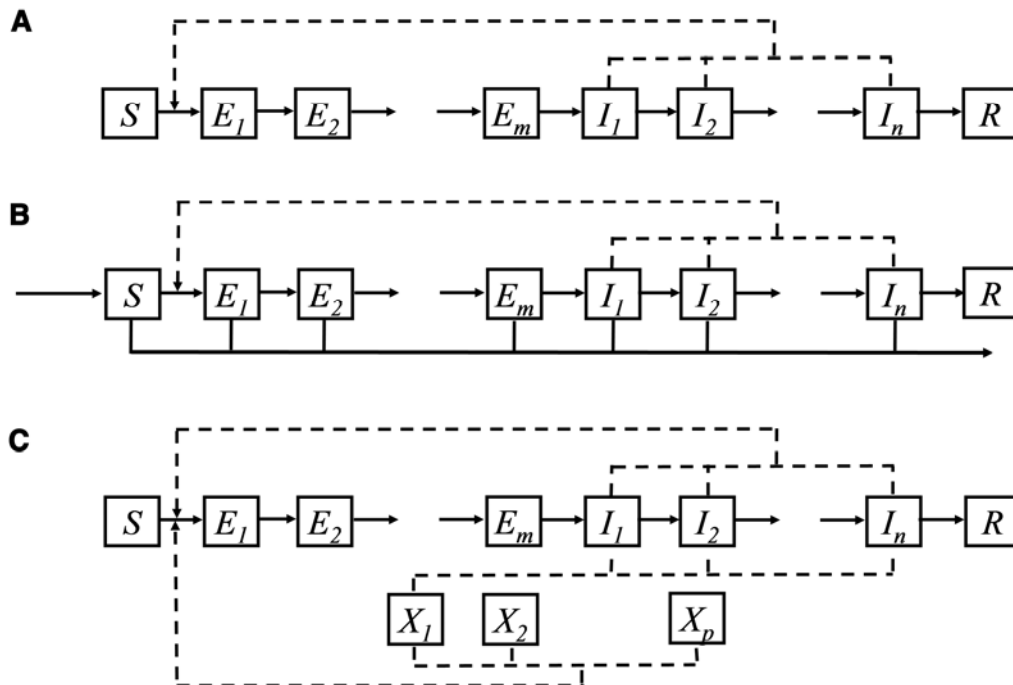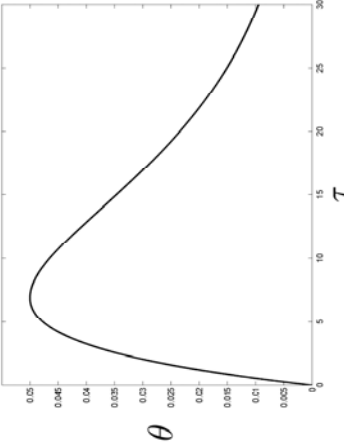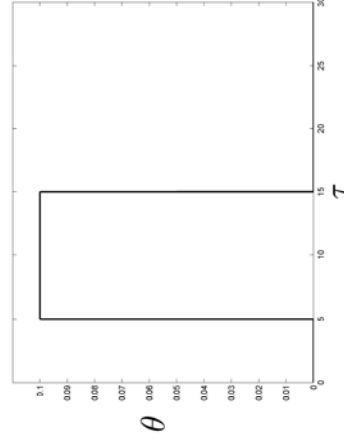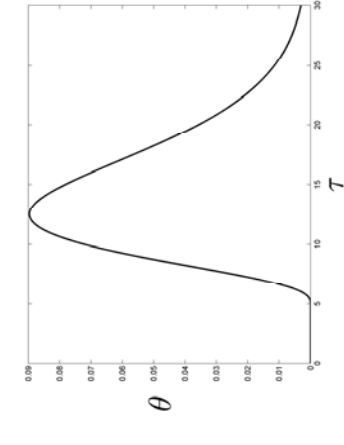

Fig. 2. Model schematics. Solid lines correspond to transitions made by hosts, whereas a dotted line indicates an effect on rate of transition according to the number or density in other compartments. **A,** SEmInR model, with m exposed (i.e., latent) and n infected compartments. The average latent period is $1/\gamma$, the average infectious period is $1/\mu$, and the rate of infection while in the ith infected compartment is β_i. **B,** SEmInR model with host demography. Susceptible hosts are replenished at rate q(S), chosen to impose a carrying capacity N, and there is loss of hosts from all compartments due to natural death at per capita rate g. **C,** SEmInXpR model. Free-living inoculum is produced by hosts in the ith infected compartment at per capita rate α_i, there are p inoculum compartments, with average inoculum infectious period $1/c$, and inoculum in the ith compartment infects each susceptible host at per capita rate of primary infection β_{Xi}.

TABLE 1. A comparison of the models, with a summary of the principal analytic results

Model type	SEIR	van der Plank	Gamma(k)
Sporulation curve[a]			
Underlying assumptions	1) Exponentially distributed latent and infectious periods 2) Constant rate of infection while infective	1) Constant latent and infectious periods 2) Constant rate of infection while infective	1) Constant latent period 2) Variable rate of infection while infective
Sporulation curve in the equivalent K & M model	$\theta(\tau) = \dfrac{\mu\gamma}{\mu-\gamma}\left(e^{-\gamma\tau} - e^{-\mu\tau}\right)$	$\theta(\tau) = \begin{cases} 0 & \text{if } \tau \le p \\ \dfrac{1}{i} & \text{if } p < \tau \le p+i \\ 0 & \text{otherwise} \end{cases}$	$\theta(\tau) = \begin{cases} 0 & \text{if } \tau \le p \\ \lambda^k \dfrac{(\tau-p)^{k-1}\, e^{-\lambda(\tau-p)}}{(k-1)!} & \text{otherwise} \end{cases}$
Representation in the SEmInR model	$m = n = 1$ $\beta_1 = \beta$	$m, n \to \infty$ $\beta_i = \beta$ for all i $\gamma = 1/p,\ \mu = 1/i$	$m \to \infty,\ n = k$ $\beta_n = n\beta,\ \beta_i = 0$ otherwise $\gamma = 1/p,\ \mu = \lambda/k$
Basic reproductive number, R_0	$R_0 = \beta N/\mu$	As SEIR	As SEIR
Fractional epidemic size, f	Given implicitly by root $0 < f < 1$ of $f \approx 1-\exp(-fR_0)$.	As SEIR	As SEIR
Initial growth rate, r	Explicit formula available $r = \dfrac{1}{2}\left(-(\gamma+\mu) + \sqrt{(\gamma+\mu)^2 + 4\mu(R_0-1)}\right)$.	Largest root r of $r = R_0\mu\left(e^{\frac{r}{\gamma}} - e^{-r\left(\frac{1}{\gamma}+\frac{1}{\mu}\right)}\right)$.	Largest root r of $\dfrac{r}{e^{\gamma}} = \dfrac{R_0}{\left(1+\dfrac{r}{n\mu}\right)^n}$.
R_0 when host demography is added to the model	$R_0 = \left(\dfrac{\gamma}{\gamma+g}\right)\left(\dfrac{\beta N}{\mu+g}\right)$	$R_0 = e^{-\frac{g}{\gamma}}\dfrac{\beta N}{g}\left(1 - e^{-\frac{g}{\mu}}\right)$	$R_0 = e^{-\frac{g}{\gamma}}\dfrac{\beta N}{\mu}\left(\dfrac{\mu n}{\mu n + g}\right)^n$

[a] Shown here normalized and with illustrative latent period 5 days, infectious period 10 days and using $k = 4$ in the Gamma(k) model. Note the scales on the y-axis differ between plots.

infected host introduced to a population of N susceptibles. In particular, defining the average rate of infection

$$\bar{\beta} = \frac{1}{n}\sum_i \beta_i, \tag{11}$$

then

$$R_0 = \sum_{i=1}^{n} \frac{\beta_i N}{n\mu} = \frac{\bar{\beta}N}{\mu}, \tag{12}$$

because the average time spent in the ith infected compartment is $1/(n\mu)$ and, while in that compartment, infections are caused in the susceptible population at overall rate $\beta_i N$. This expression for R_0 in the most general SEmInR model reduces to the same value for each of the three models we consider, because the average rate of infection $\bar{\beta}$ is fixed at β by our earlier normalization and thus, in all cases, $R_0 = \beta N/\mu$.

Final size of the epidemic, f. By determining a conserved quantity according to equation 5 and comparing its value at the beginning and end of any epidemic (Appendix 2), we find in the general SEmInR model that, if the number of exposed or infected hosts at the beginning of any epidemic is small in comparison with the population size, then

$$f \approx 1 - \exp(-fR_0), \tag{13}$$

where f is the fraction of hosts eventually infected. This so-called "final size equation" (35) implicitly defines f in terms of $R_0 > 1$, and predicts a very large fraction of hosts eventually infected for $R_0 > 5$. Because R_0 is identical for the models we consider, if the parameters are matched appropriately (compare equation 12), so is f and, therefore, the eventual impact of any epidemic is independent of model structure.

Initial rate of increase of disease, r. The initial rate of exponential disease increase, r, is obtained by linearizing equation 5 about the pathogen-free equilibrium with $S = N$ and $E_i = I_i = R = 0$ (Appendix 2). This corresponds to the density-independent growth at the start of any epidemic. In the general SEmInR model, r is given by the largest root of the equation

$$\left(1 + \frac{r}{m\gamma}\right)^m - \frac{N}{n\mu}\sum_{i=1}^{n}\beta_i\left(\frac{1}{1+\frac{r}{n\mu}}\right)^i = 0 \tag{14}$$

For the SEIR model, this simplifies, and r is given explicitly by

$$r = \frac{1}{2}\left(-(\gamma+\mu) + \sqrt{(\gamma+\mu)^2 + 4\gamma\mu(R_0-1)}\right). \tag{15}$$

For the van der Plank model, r is given implicitly by the largest root of the equation

$$r = R_0\mu\left(e^{-\frac{r}{\gamma}} - e^{-r\left(\frac{1}{\gamma}+\frac{1}{\mu}\right)}\right). \tag{16}$$

For the Gamma(k) model, r is given implicitly by the largest root of

$$e^{\frac{r}{\gamma}} = \frac{R_0}{\left(1+\frac{r}{n\mu}\right)^n}. \tag{17}$$

The response of r to changes in R_0 in the different models for a range of latent ($1/\gamma$) and infectious ($1/\mu$) periods between 5 and
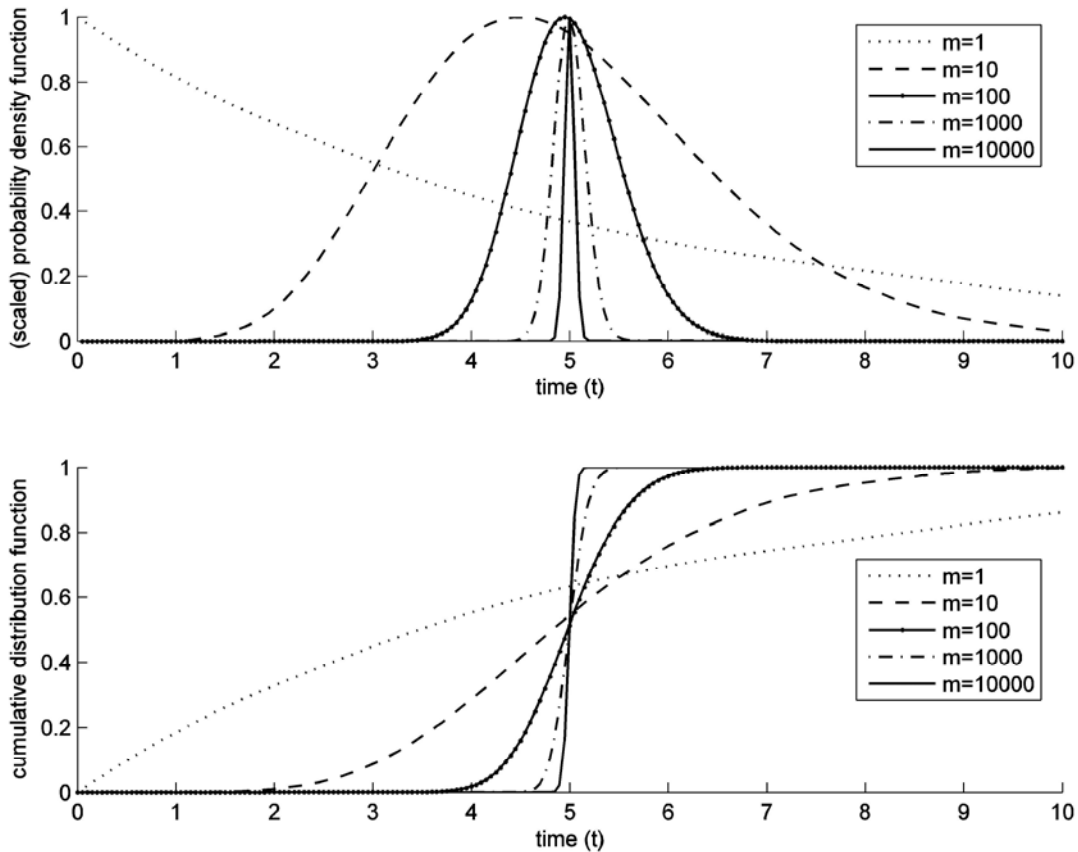

Fig. 3. Probability density and cumulative distribution functions for leaving the exposed classes in the SEmInR model for different values of m (i.e., number of exposed classes) conditioned on entering the first exposed class at $t = 0$. The probability density is renormalized with maximum value 1.0, to allow easy comparison between the responses. The mean latent period is 5 days (i.e., $1/\gamma = 5$ days). As m increases, the probability density function tends to a δ function, and the cumulative density to a step function.

20 days and 5 and 100 days, respectively, is shown in Figure 4. The ordering of the growth rate in the different models depends on the relative lengths of the latent and infectious periods and the value of R_0. In broad terms, and for large values of R_0 (e.g., $R_0 > 10$), the SEIR model has the largest value of r (because there is a non-zero probability of a newly infected host leaving the exposed class and thereby becoming capable of causing new infections very quickly after becoming infected), and the Gamma(k) model has the smallest value of r (because there is a fixed delay before infected individuals can cause new infections and, even when infection is possible, the initial rate is low). The relative value of r in the van der Plank model depends on the relative latent and infectious periods. In particular, as the latent period increases and the delay in infection impacts on the rate of initial growth, the responses of the van der Plank and Gamma(k) models become closer. However, as the infectious period increases, this effect is reversed, and the responses of the van der Plank and SEIR models become closer, because the large rate of infection due to hosts in the I compartments directly after the latent period in the van der Plank model increases its growth rate relative to the Gamma(k) model. We note that, for small values of R_0 (and, therefore, for small growth rates), the value of r in the van der Plank model can, in fact, be larger than that in the SEIR model; this is because the large rate of infection in the van der Plank model immediately after the latent period counteracts any delay in sporulation relative to the SEIR model.

Numerical simulation of the SEmInR model. Numerical solutions of the SEmInR model can be calculated using widely available and standard routines for integrating systems of coupled ordinary differential equations. However, because the van der Plank and Gamma(k) models are obtained in the limit of the SEmInR model with an infinite number of exposed or infected classes, it is possible that relative ease of simulation comes at the cost of reduced accuracy, and we investigate this here.

We consider a pathogen with a long latent period ($1/\gamma = 25$ days) and a short infectious period ($1/\mu = 5$ days). These parameters, although arbitrary, were chosen to maximize differences between the SEIR model and the van der Plank and Gamma(k) models (Fig. 4); the high ratio of latent/infection period means that infection is effectively separated into discrete generations, a response that is difficult to represent in the basic SEIR model. We solve the van der Plank and Gamma(k) models using custom-written numerical routines, and compare the results of these simulations with results of using standard ordinary differential equation solvers to integrate SEmInR compartmental approximations as the number of exposed or infected compartments increases. In particular, we take $\beta N = 1$ (i.e., $R_0 = 5$) and approximate the van der Plank model via a SEmInR model with $m = n = 1, 3, 5, 10, 50$, and 250, and the Gamma(k) model ($n = k = 4$) via a SEmInR model with $m = 1, 3, 5, 10, 50$, and 250.

The equivalent sporulation curve, $\theta(\tau)$ and the proportion of nonsusceptible tissue (i.e., the disease progress curve) are shown in Figure 5 for the different values of m and n, together with the logarithm of the error, E, as the number of classes changes. If the "exact" solution according to the custom simulations of the van der Plank or Gamma(k) model has disease progress curve $f_{exact}(t)$, and the approximate SEmInR model has disease progress curve $f_{approx}(t)$, then we define this error in the SEmInR solution in terms of the difference in area under disease progress curves

$$E = \frac{1}{T_{max}} \int_{t=0}^{T_{max}} \left| f_{approx}(t) - f_{exact}(t) \right| dt \qquad (18)$$

where T_{max} is a normalizing upper bound (here, we take $T_{max} = 220$ days, by which time the disease progress curve has reached its upper asymptote in all cases; however, our qualitative results are independent of this choice).
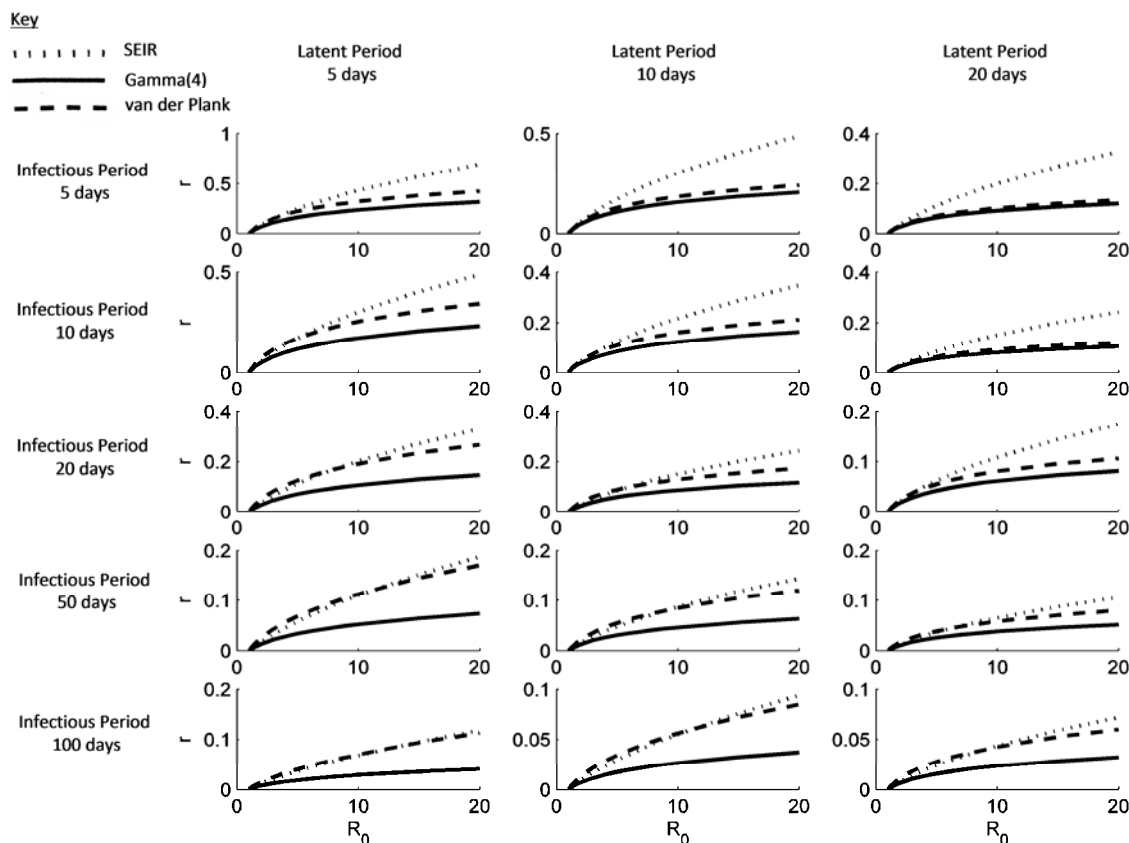

Fig. 4. Response of initial rate of exponential disease increase, r, to basic reproductive number, R_0, for SEIR (dotted line), van der Plank (dashed line), and Gamma(4) (solid line) models. Latent periods ($1/\gamma$) = 5, 10, or 20 days (columns) and infectious periods ($1/\mu$) = 5, 10, 20, 50, or 100 days (rows).

The larger the number of classes, the better the approximation for both the van der Plank and the Gamma(k) models. With $m = n = 250$ (van der Plank) or $m = 250$ (Gamma(k) with $k = 4$), despite minor differences in the equivalent sporulation curve, disease progress curves according to the SEmInR model are almost indistinguishable from those from explicit simulation, and very closely replicate the repeated rise-and-plateau pattern. This is reflected in very small values of E. However, the error continues to decrease with increases in the number of compartments, albeit at a decreasing rate.

Extending the models to include host demography. When demography is added, the invasion threshold R_0 depends on the variant of the model in question (Table 1). In equation 6, a latently infected host has probability $m\gamma/(m\gamma + g)$ of passing from compartment E_{i-1} to E_i without dying and, thus, with m exposed compartments, the probability of an exposed host successfully reaching the first infective compartment is

$$\left(\frac{m\gamma}{m\gamma + g}\right)^m. \tag{19}$$

Conditioned on reaching any particular infected compartment, the probability of reaching the next is given by $n\mu/(n\mu + g)$ and, thus, the probability of surviving to reach the ith infected compartment conditioned upon reaching I_1 is

$$\left(\frac{n\mu}{n\mu + g}\right)^{i-1}. \tag{20}$$

Because an infected host remains in the ith infected compartment for an average of $1/(n\mu + g)$ units of time and, while in that compartment, causes infections at overall rate $\beta_i N$, the general expression for R_0 according to equation 6 becomes

$$R_0 = \left(\frac{m\gamma}{m\gamma + g}\right)^m \sum_{i=1}^{n} \frac{\beta_i N}{n\mu + g}\left(\frac{n\mu}{n\mu + g}\right)^{i-1}. \tag{21}$$

For the SEIR model

$$R_0 = \left(\frac{\gamma}{\gamma + g}\right)\left(\frac{\beta N}{\mu + g}\right). \tag{22}$$

For the van der Plank model

$$R_0 = e^{-\frac{g}{\gamma}}\frac{\beta N}{g}\left(1 - e^{-\frac{g}{\mu}}\right). \tag{23}$$

For the Gamma(k) model

$$R_0 = e^{-\frac{g}{\gamma}}\frac{\beta N}{\mu}\left(\frac{\mu n}{\mu n + g}\right)^{n}. \tag{24}$$

We note that now the value of R_0 depends upon the model in question.

The value of R_0 when $\beta N = 1$ is shown in Figure 6 for a range of $1/\gamma$ and $1/\mu$ between 5 and 20 days and 5 and 100 days, respectively, with illustrative average lifetimes ($1/g = 10$, 50, or 100 days). Unsurprisingly, the effects of host demography on R_0 are more striking when the average lifetime is short. In the SEIR and van der Plank models, the value of R_0 increases with the infectious period. However the Gamma(k) model exhibits a non-monotone response to $1/\mu$; in particular, as the infectious period increases, the value of R_0 first increases, then decreases. This is because extending the infectious period increases the number of spores released (and, thus, acts to reduce R_0) but, as $1/\mu$ increases, a larger fraction of potential sporulation occurs at later times, which are never reached in practice (because hosts are dying). Finally, whereas, in the underlying models, the latent period does not affect the value of R_0, because all infected hosts survive to produce the same number of new infections, this is no longer true if the host dies. Death of hosts puts a premium on becoming infectious immediately, and this is reflected in larger values of R_0 for pathogens with shorter latent periods, in all variants of the models.
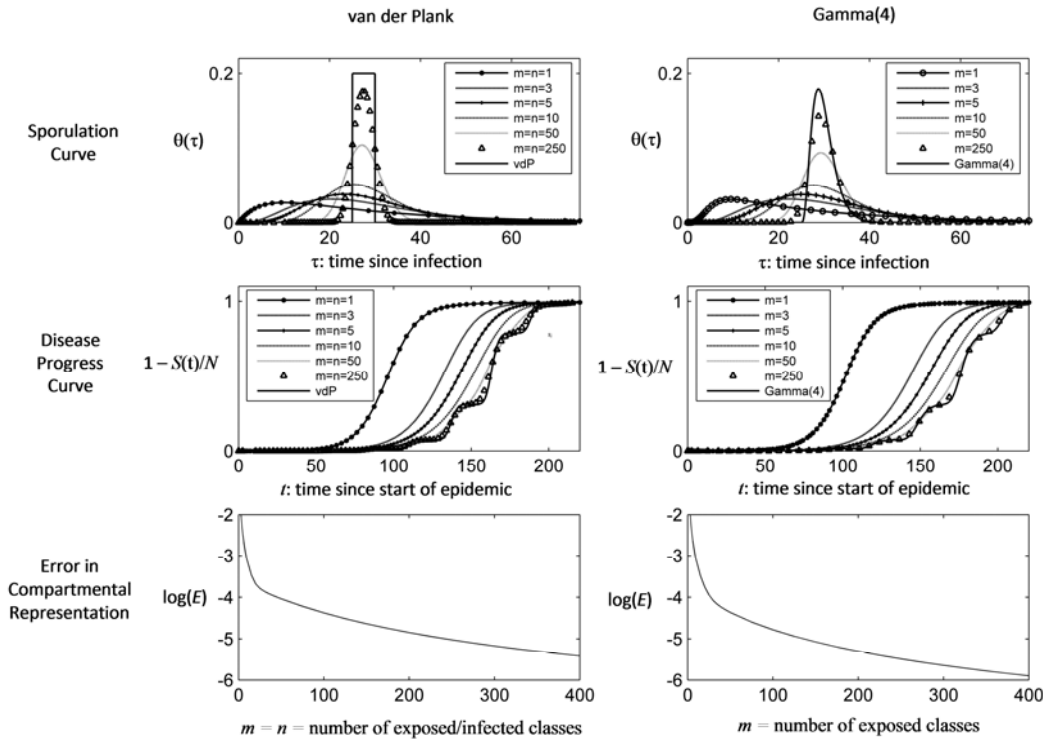

Fig. 5. Accuracy of SEmInR approximations to the van der Plank and Gamma(4) models as the number of compartments increases. The sporulation curve, $\theta(\tau)$, and the proportion of nonsusceptible tissue (i.e., the disease progress curve) are shown for the van der Plank and Gamma(n) models. These are compared with the SEmInR approximations to the van der Plank model with $m = n = 1$, 3, 5, 10, 50, and 250, and to the Gamma(4) model with $m = 1$, 3, 5, 10, 50, and 250. The latent ($1/\gamma = 25$ days) and infectious ($1/\mu = 5$ days) periods are selected to give a response that is difficult to represent in the underlying SEIR model. Also shown is the accuracy of the approximation in terms of E, the normalized difference in area under disease progress curves (equation 17).

In general (and, therefore, for each of the models we consider) the number of susceptible hosts ($\overline{S}$) at the endemic equilibrium according to equation 6 is given by

$$\overline{S} = \frac{N}{R_0},\qquad(25)$$

at least if $R_0 > 1$. This shows how differences in the invasion thresholds identified in equations 22 to 24 also have a crucial effect on the long-term endemic behavior in host populations with demography. The initial exponential growth rate follows from an analysis similar to that used before (details available on request) and is given by $r - g$, where r is the growth rate in the particular variant of the underlying model without demography. This is unsurprising; the initial exponential increase of the epidemic is retarded by the continuous death of infected hosts, which occurs at per capita rate g.

Extending the models to include free-living infectious material. To calculate R_0 for the SEmInXpR model (equation 7), we again consider a single infectious host introduced to an otherwise susceptible population. We consider this rather than the introduction of a small quantity of inoculum because, although the initial introduction of inoculum would result in a number of new infections, it is actually the *dynamics* of these resulting infections that determine whether or not the disease successfully invades (10,24). For convenience recall the mean rate of secondary infection,

$$\overline{\beta} = \frac{1}{n}\sum_{j=1}^{n}\beta_j\qquad(26)$$

and introduce mean rates of production of inoculum,

$$\overline{\alpha} = \frac{1}{n}\sum_{j=1}^{n}\alpha_j\qquad(27)$$

and of primary infection,

$$\overline{\beta}_X = \frac{1}{p}\sum_{j=1}^{p}\beta_{Xj}\qquad(28)$$

As before, the initial infected host causes an average of $\overline{\beta}N/\mu$ secondary infections (equation 12). However, it also produces an average of $\overline{\alpha}/\mu$ units of inoculum over its lifetime, each unit of which produces an average of $\overline{\beta}_X N/c$ new primary infections. Because R_0 is the total number of new infected hosts corresponding to the initial single introduction, it is the sum of these two contributions

$$R_0 = \frac{\overline{\beta}N}{\mu} + \frac{\overline{\alpha}}{\mu}\cdot\frac{\overline{\beta}_X N}{c} = \frac{N}{\mu}\left(\overline{\beta} + \frac{\overline{\alpha}\,\overline{\beta}_X}{c}\right)\qquad(29)$$

Extending our earlier analysis, the final size equation is unchanged (i.e., $f \approx 1 - \exp(-fR_0)$) whereas, again, the initial rate of increase of infected tissue differs between model variants and here, too, for different numbers of inoculum compartments (Appendix 3 provides details of the analysis).

Estimating f from r in the SEmInR model. As demonstrated above, the relation between the initial growth rate, r, and the value of R_0 (and thus, in turn f, the eventual extent to which the pathogen will affect the host population) depends on the model. This means that, if we use an experimentally determined r together with prior knowledge on the life history of the pathogen to estimate the values of R_0 and, therefore, f, our results depend on the model we use (Fig. 1). We illustrate this by considering the estimates of R_0 and f corresponding to $r = 0.01$ day^{-1} in all three models, for a range of values of $1/\gamma$ and $1/\mu$ up to 20 and 100 days, respectively. This is then translated into percent changes in R_0 and f between pairs of models (Fig. 7).

The choice of model can greatly affect the estimated R_0 and f. The results broadly follow the patterns in r identified previously (Fig. 4) and depend strongly on the latent and infectious periods. Differences of up to 60% in the estimated final severity are possible and are promoted by long latent and short infectious periods; in this case, the differences between the SEIR and the van der Plank/Gamma(k) models are largest (Fig. 5). We note that the latent period does not affect the van der Plank versus Gamma(k) model comparison because, in both models, there is no possibility of sporulation during the fixed latent period.
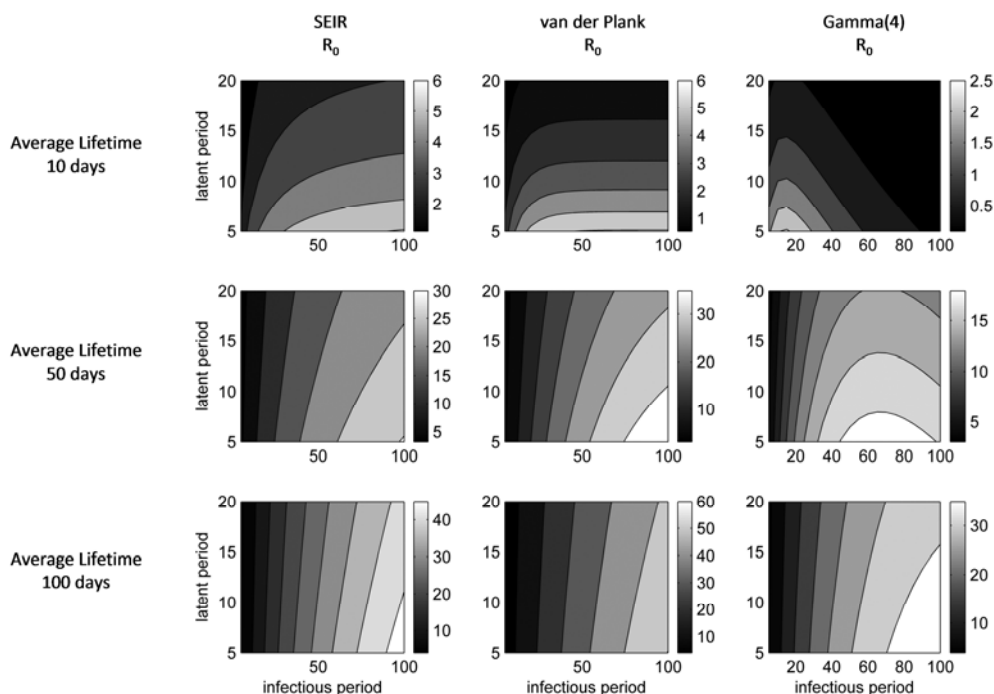

Fig. 6. Effect of host demography on the basic reproductive number, R_0, in the SEIR (column one), van der Plank (column two), and Gamma(4) (column three) models. The latent period ($1/\gamma$) ranges up to 20 days and the infectious period ($1/\mu$) up to 100 days. Rows correspond to different average host lifetimes ($1/g$); 10 days, row one; 50 days, row two; 100 days, row three.

DISCUSSION

We have introduced the SEmInR model, an extension of the SEIR model with subdivided exposed (i.e., latent) and infectious compartments. This extension allows the assumptions of exponentially distributed infectious and latent periods implicit in the SEIR model to be relaxed, and more complex time-dependent infectivity to be represented. Although splitting the exposed or infectious classes has become somewhat commonplace in models of human or animal disease (34,38,39,50,75) as a practically useful example of the so-called "linear chain trick" to convert certain delay differential equations to systems of ordinary differential equations (46), to our knowledge, this is the first use of multiple exposed or infected compartments in the mathematical analysis of a model of plant disease. We do note that the early simulation literature introduced delays using effectively the same multiple compartment formalization (via so-called "boxcar trains") (12,19, 74) but also note that no mathematical analysis was performed in these simulation studies. Additionally, whereas modelers of human or animal disease have concentrated solely on changing the number of classes to increase flexibility in latent or infectious periods, in our model, we allow infection rates to vary by infected compartment, allowing us to capture the complex time-dependence in infection rate that has been measured experimentally for plant pathogens (52,67,68).

We demonstrate how the SEmInR model can specifically target plant pathogens by recasting the van der Plank and Gamma(k) (i.e., K & M model with Gamma-distributed infectivity) models in the compartmental framework. As well as demonstrating the wide applicability of our multiple compartment reformulation, representing the van der Plank and Gamma(k) models in this fashion allows the relatively simple analysis typical of compartmental models. Compartmental models have "entered the mainstream" of plant pathology in the past twenty years or so (61) and the requisite mathematical techniques to analyze compartmental models are familiar from a number of articles (9,10,16,24–26,31,42,54, 55,69,77). Therefore, by unifying the van der Plank and Gamma(k)

models, together with the SEIR model itself, in a single and well-known theoretical framework, we simplify the more mathematically demanding analysis of Segarra et al. (59).

For each of the three models in question, we determined the initial rate of exponential increase of disease (r), the basic reproductive number (R_0), and the final fraction of infected hosts (f). As first identified by Segarra et al. (59), only r changes among the models we consider. However, we illustrate a possible practical significance of this result by demonstrating that it can lead to widely divergent assessments of the likely severity of an emerging epidemic from data on its initial spread (Fig. 1). Data from the early part of the epidemic are all that are available when decisions on control are most pressing; therefore, any assessment of the level of control required, or whether it is even necessary, could depend crucially on the seemingly esoteric phenomena addressed by our models (see also Roberts and Heesterbeek [57], who give a detailed treatment in the context of human or animal disease, for a variety of simple infection kernels).

Although the above provides an interesting illustration of the potential significance of our results, it is, alas, rather unlikely that a particular farmer or grower would use our techniques to determine how heavily to control an individual field, farm, or grove. Doubtless, the same could be said of almost any mathematically motivated method although, for crop plants, there is the additional difficulty that any epidemic may not reach the final size predicted by our models before being interrupted by harvesting or by periods of unsuitable environmental conditions. However, we contend that, at larger spatial (i.e., over entire regions or countries) and temporal (i.e., for perennials) scales, our methodology is both practical and relevant, in particular for use by policy makers. A timely and pressing example is sudden oak death in the United Kingdom, caused by *Phytophthora ramorum* (7); significant efforts are already being made to predict the likely progress of this outbreak and, thus, to determine an appropriate level of control to exert. These decisions become particularly pressing in the light of the ongoing and devastating sudden oak death outbreak in California (45,56), and intense scrutiny of the
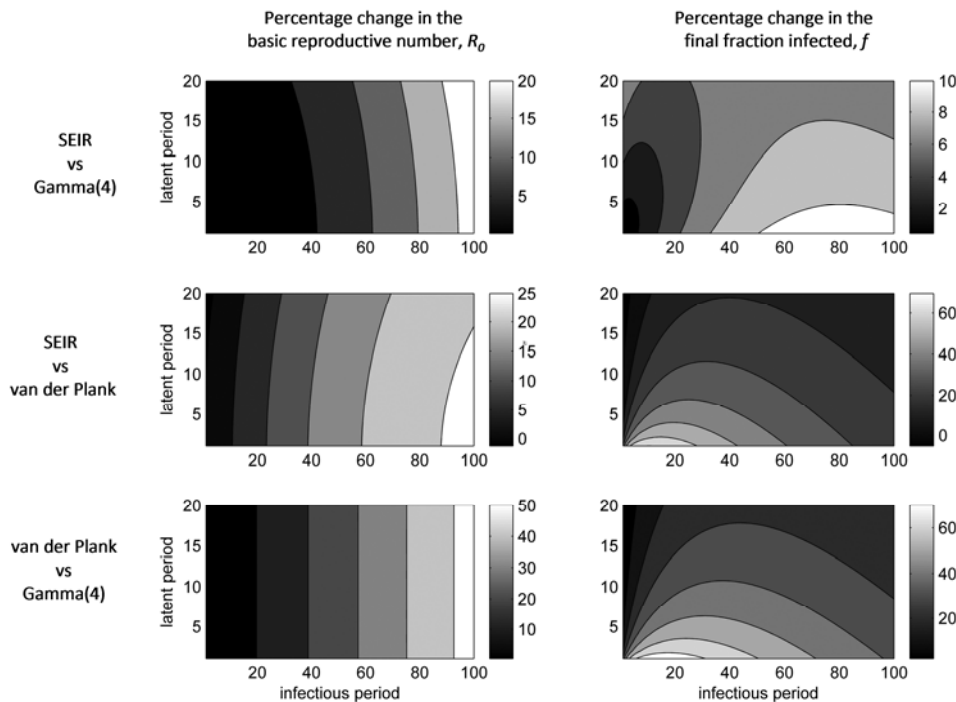

Fig. 7. Effects of model structure on estimates of eventual severity from initial rates of exponential increase. Percent difference in the basic reproductive number, R_0 (column one), and the final fraction of infected hosts, f (column two), are shown between pairs of models. The initial rate of disease increase is $r = 0.01$ day^{-1}. The latent period ($1/\gamma$) ranges up to 20 days and the infectious period ($1/\mu$) up to 100 days. SEIR versus Gamma(4) is shown in row one, SEIR versus van der Plank in row two, and van der Plank versus Gamma(4) in row three.

potential mismanagement of the Dutch elm disease outbreak in the late 1960s and 1970s (64). Our method of informing control by predicting eventual epidemic size from data on early spread could also be used by policy makers to justifiably select those regions upon which to focus efforts when resources for control are limited by budgetary constraints (48,49). In this context, a fairly obvious large-scale strategy would be to attempt to preferentially control those regions for which a mathematical model predicts the most devastating eventual impact, using data on initial spread within individual regions to predict localized long-term behavior. However, initial results from a caricature of a different but related problem via a coupled metapopulation model and optimal control theory indicate that, when controlling with limited resources, in that case the obvious strategy is not always optimal (58). We will return to this using the framework we have outlined here in our future work.

In the context of these and other examples, the focus shifts to how to correctly and justifiably select a model from data on early spread. Despite the SEmInR model being able to represent an infectivity response closer to that observed in experiment (52, 67,68) and therefore, a priori, being expected to both provide the better fit to population level epidemic data and lead to more reliable predictions of eventual epidemic size, the additional flexibility of the most general SEmInR model comes at the cost of many more parameters. This indicates that model selection criteria may work against it being selected as the best-fitting model and, thus, any prediction based upon it may be difficult to justify in practice. We do not address this potential practical difficulty here, although we do note that the (perhaps most biologically realistic) Gamma(k) variant of the SEmInR model in fact requires only one additional parameter when compared with the simple SEIR model; therefore, this may not be a serious problem when selecting one from that pair of models.

In matching the parameters between the SEIR and van der Plank models, we followed well-established practice by choosing the "obvious" normalization in which the average latent and infectious periods are equal for the two models (28,33,41,51,59). However this immediately implies that the mean age at infection, i.e., the quantity

$$\int_{\tau=0}^{\infty} \tau \hat{\theta}(\tau) d\tau \qquad (30)$$

where $\hat{\theta}(\tau)$ is the infection kernel $\theta(\tau)$ normalized to be a probability distribution) differs. Mathematically elegant work (46) focuses on deriving approximations to the epidemic growth rate r as a function of $\ln(R_0)$ and the moments of the infection kernel, and indicates that the mean age at infection is an important driver of the response (an effect which is most pronounced for small $\ln(R_0)$). We will investigate the effects of the standard choice of normalization in terms of these approximations in our future work.

Representing the van der Plank and Gamma(k) models in the compartmental framework allows numerical solution using standard software tools (40) and, in particular, the coupled system of ordinary differential equations specifying the SEmInR model can be directly inputted and solved using widely available routines available in all mathematical software packages. The model can also be simulated using graphical "systems dynamics" approaches, such as Stella, Similie, Berkeley Madonna, and Vensim, although handling fixed delays may require some effort in these packages. Another option is to use the well-known and relatively simple Fortran simulator translator scripting language (72), which is based on the Continuous Systems Modeling Package (CSMP); a numerical recipe for building boxcar models with multiple compartments in CSMP is given by Goudriaan and van Roermund (19). Taken together, this relative ease of implementation means that our reformulation of the model makes the more realistic sporulation responses more accessible to modelers, and avoids any need for custom-written, complex numerical routines to handle delay differential (van der Plank) or integro-differential (K & M) equations, surely reducing the chance of errors in implementation.

A major advantage of compartmental models is the increased flexibility they offer and the ease with which additional biology can be included (40,61). We have illustrated this by extending our underlying SEmInR model to include free-living infectious material and host demography. Adding inoculum to the basic SEmInR model (moving to the SEmInXpR model) (Fig. 2C) increases the value of R_0 by adding a new component corresponding to primary infection, and allows the complex balance between primary and secondary infection most characteristic of soilborne pathogens (1–6,8,10,16,18,43) to be represented. The updated R_0 does not, in fact, differ between the analogues of the SEIR, van der Plank, and Gamma(k) models. However, the initial rate of increase of infected tissue again differs between model variants and for different numbers of inoculum compartments, reiterating that initial disease progress is highly sensitive to assumptions concerning infection rates and distributions of latent and infectious periods.

We note that our model of inoculum dynamics itself subdivides the X compartment and, thus, promotes increased flexibility in the survivorship of free-living infectious material. More realistic sigmoidal survivorship curves (70) are natural within our framework, and will be the subject of future work (particularly over multiple seasons, a situation which is both more complex and relatively little studied) (20,21,23,43). In the context of multiseason dynamics, differences in the initial growth rate r revealed by our analysis will become particularly important. Often (e.g., for crops), epidemics do not, in fact, run to completion but are instead interrupted by harvesting; the initial rate of increase is then an important determinant of the level of disease at the end of the growing season and therefore, in turn, of the amount of inoculum (and, thus, primary infection) at the start of the next (21,24,43).

Adding host demography (Fig. 2B) has arguably an even more pronounced effect. Depending on the variant of the model in question, different values of the key threshold R_0 are obtained, even if all pathogen parameters are matched, because, when demography is included, the endemic level at which the pathogen persists in the host population depends on R_0, failure to match the estimation of r with an appropriate model for the epidemic under consideration leads to incorrect conclusions about the long-term prevalence of disease and of healthy hosts. Although host demography is a relatively simple example, a large and ever-increasing number of articles extend the SEIR model to encompass more biology (1–6,10,11,16,17,20–27,31,42,44,48,49,53,54,65,66,69,76, 77) and, therefore, purportedly extract more biologically realistic thresholds. However, as we have shown here, even small changes to the biology introduced into epidemic models can interact with more realistic infection responses and lead to divergent conclusions. Therefore, the "standard" assumptions exemplified by the simplest SEIR model could have significant consequences whatever additional biology is included, and we urge modelers to consider carefully whether these assumptions are justified in their future work. We contend that this is the most important message of our article.

APPENDIX 1:
Fitting the SEmInR Model to Sporulation Data

We fitted infection kernels to a representative sporulation data set (*Puccinia lagenphorae* on *Senecio vulgaris*) (41). We examine whether (i) the restriction imposed by the SEmInR model that the shape parameter for Gamma infection kernels must take integer values markedly constrains the fit and (ii) the extra flexibility available in the "full" SEmInR model leads to an improved fit

compared with the Gamma infection kernel with integer shape parameter.

Fixing the shape parameter of the Gamma distribution to be an integer. For a Gamma-distributed infection kernel to be represented in the SEmInR framework, its shape parameter must be an integer. We used nonlinear least squares to fit the Gamma distribution in the (nonnormalized) form

$$I(t) = \frac{\Omega\lambda(\lambda t)^{k-1} e^{-\lambda t}}{\Gamma(k)}. \qquad (31)$$

Our estimates of the parameters were $\Omega = 11.52$, $\lambda = 0.30$, and $k = 4.20$ (Fig. 8A). The data were normalized such that the maximum daily spore production was 1.0; this fixes the value of the scale factor Ω and is independent of the values of λ and k; $t = 0$ corresponded to the very first sign of spore production, allowing us to ignore the latent period in fitting. The fit according to equation 31 was compared with that according to the submodel

$$I(t) = \frac{\Omega\lambda(\lambda t)^{3} e^{-\lambda t}}{3!}, \qquad (32)$$

that is, equation 31 with fixed shape parameter $k = 4$ ($\Omega = 11.67$, $\lambda = 0.28$). Because the models are nested, the fits can be compared using a F test. The extra flexibility corresponding to non-integer values of the shape parameter k does not lead to any significant improvement ($F = 0.121$, $P = 0.73$). In fact, visual inspection reveals no clearly visible difference between the fits (Fig. 8A). For this data set, the restriction of k to integer values imposed by our analysis is unimportant.

Extra flexibility in the SEmInR model. Although, in the main text, we concentrated on the Gamma distribution with integer shape parameter, this corresponded to setting all but the last of the infection rates (β_i) in our SEmInR model to zero. By allowing the other β_i to be non-zero, the infectivity kernel in the SEmInR model may be generalized. We examined whether such an extension was justified for this data set using forward stepwise regression. In particular, we compared the Gamma-distributed infection kernel with fixed $k = 4$ (i.e., equation 32) with the three variants of the following form

$$I(t) = \sum_{i=1}^{4} \frac{\Omega_i \lambda(\lambda t)^{i-1} e^{-\lambda t}}{(i-1)!}, \qquad (33)$$

in which Ω_4 and one of Ω_1, Ω_2, or Ω_3 were non-zero. This is the infection kernel corresponding to taking β_4 and one of β_1, β_2, or β_3 to be non-zero in the SEmInR model (equation 10). None of the fits according to equation 33 were visually distinguishable from that according to equation 32 (Fig. 8B), and for the most significant of the fits ($\Omega_2 \neq 0$), $F = 0.378$, and $P = 0.54$. The extra complexity of generalizing the infection kernel is not warranted, at least for these data.

APPENDIX 2:
Mathematical Details of the Analysis of the SEmInR Model

Final size of the epidemic, f. According to equation 5, if we define

$$\phi(t) = S + \sum_{j=1}^{m} E_j + \frac{1}{n\bar\beta}\sum_{j=1}^{n}\sum_{i=j}^{n}\beta_i I_j - \frac{N}{R_0}\log(S), \qquad (34)$$

then $d\phi/dt = 0$, and so $\phi(t)$ is constant. At the end of the epidemic, because $E_j = I_j = 0$ for all j,

$$\phi(\infty) = S(\infty) - \frac{N}{R_0}\log(S(\infty)) \qquad (35)$$

where $S(\infty)$ is the final number of susceptible hosts. Assuming that the number of exposed or infected hosts at the beginning of the epidemic is small in comparison with the population size, $S(0) \approx N$ and, thus,

$$\phi(0) \approx N - \frac{N}{R_0}\log(N). \qquad (36)$$

Because ϕ is constant, $\phi(0) = \phi(\infty)$,

$$S(\infty) - \frac{N}{R_0}\log(S(\infty)) \approx N - \frac{N}{R_0}\log(N). \qquad (37)$$

However, defining f as the fraction of hosts infected during the course of any epidemic, $S(\infty) = (1-f)N$ and, thus,

$$f \approx 1 - \exp(-fR_0). \qquad (38)$$

Initial rate of increase of disease, r. The linearization of equation 5 about the pathogen-free equilibrium, $S = N$, $E_i = I_i = R = 0$ is $d\mathbf{x}/dt = J\mathbf{x}$, where $\mathbf{x} = (S, E_1 \ldots E_m, I_1 \ldots I_n, R)^{\mathrm{T}}$, and where J is the $(m + n + 2)$ by $(m + n + 2)$ Jacobian matrix
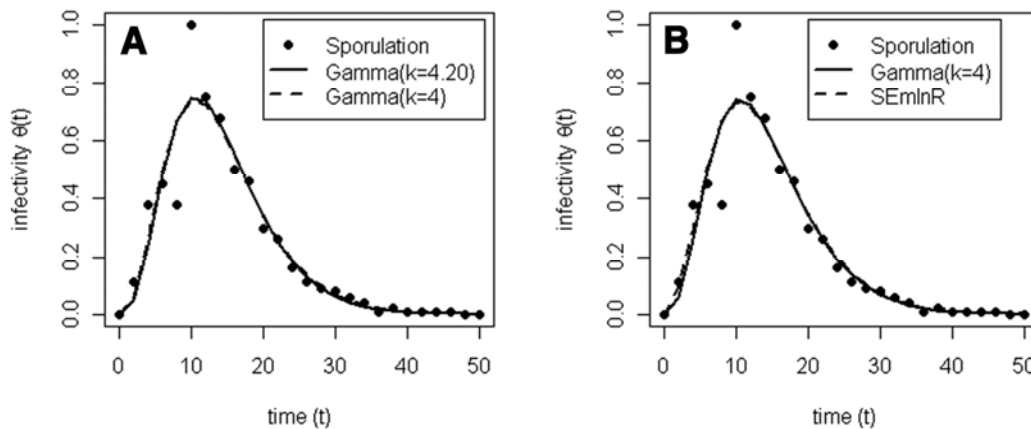

Fig. 8. Generalizing the infection kernel in the SEmInR model. **A,** Fit to real sporulation data and the (almost indistinguishable) consequence of the restriction of the shape parameter to an integer that is required to represent the Gamma(k) response in the SEmInR model. **B,** Extra flexibility available in the "full" SEmInR model with more than one infection rate non-zero does not lead to an improved fit when compared with the Gamma-distributed integer shape parameter response we considered in the main text of the article. Statistical assessment of these results is given in Appendix 1.

$$J = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\beta_1 N & \cdots & -\beta_n N & 0 \\ 0 & -m\gamma & & \vdots & \beta_1 N & \cdots & \beta_n N & 0 \\ 0 & m\gamma & \ddots & 0 & & & & \\ & 0 & \ddots & -m\gamma & 0 & & & \\ & & & m\gamma & -n\mu & & & \\ & & & 0 & n\mu & \ddots & 0 & \\ & & & & 0 & \ddots & -n\mu & 0 \\ & & & & & & n\mu & 0 \end{pmatrix}. \tag{39}$$

In the vicinity of the equilibrium (i.e., at the beginning of any epidemic, before the number/density of susceptibles becomes limiting), the approximate solution follows

$$\mathbf{x} = \sum_i \mathbf{v}_i e^{-r_i t}, \tag{40}$$

where r_i is the ith Eigenvalue of the matrix J, and the vectors $\mathbf{v}_i$ depend upon the initial condition. The time dependence is dominated by the largest positive Eigenvalue which, therefore, controls the initial exponential growth of disease. In this case, the Eigenvalues of J are given implicitly by the roots, r, of its characteristic equation

$$r^2 \left(\left(1 + \frac{r}{m\gamma}\right)^m - \frac{N}{n\mu} \sum_{i=1}^n \beta_i \left(\frac{1}{1 + \frac{r}{n\mu}}\right)^i \right) = 0, \tag{41}$$

and, thus, the rate of exponential increase in the initial density-independent growth is given by the largest root of equation 41.

In particular, because for the SEIR model $m = n = 1$ and $\beta_1 = \beta$, the initial rate of increase is given by the positive root of the (quadratic) equation

$$\left(1 + \frac{r}{\gamma}\right) - \frac{\beta N}{\mu} \left(\frac{1}{1 + \frac{r}{\mu}}\right) = 0, \tag{42}$$

which is (using $R_0 = \beta N/\mu$)

$$r = \frac{1}{2}\left(-(\gamma + \mu) + \sqrt{(\gamma + \mu)^2 + 4\gamma\mu(R_0 - 1)}\right). \tag{43}$$

For the van der Plank model, because $\beta_i = \beta$ for all i

$$\left(1 + \frac{r}{m\gamma}\right)^m - \frac{R_0}{n} \sum_{i=1}^n \left(\frac{1}{1 + \frac{r}{n\mu}}\right)^i = 0. \tag{44}$$

Summing the geometric progression

$$\left(1 + \frac{r}{m\gamma}\right)^m - \frac{R_0\mu}{r}\left(1 - \left(\frac{1}{1 + \frac{r}{n\mu}}\right)^n\right) = 0. \tag{45}$$

However,

$$\lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n = e^x \tag{46}$$

and, thus, in the limit $m, n \to \infty$

$$r = R_0\mu\left(e^{-\frac{r}{\gamma}} - e^{-r\left(\frac{1}{\gamma} + \frac{1}{\mu}\right)}\right). \tag{47}$$

The initial growth rate is fixed by the largest root of equation 47.

For the Gamma(k) model, only $\beta_n \neq 0$ and the relevant part of equation 41 becomes

$$\left(1 + \frac{r}{m\gamma}\right)^m - \frac{R_0}{\left(1 + \frac{r}{n\mu}\right)^n} = 0. \tag{48}$$

Because $m \to \infty$, the initial rate of disease increase is then given implicitly by the largest root of

$$e^{\frac{r}{\gamma}} = \frac{R_0}{\left(1 + \frac{r}{n\mu}\right)^n}. \tag{49}$$

APPENDIX 3:
Mathematical Details
of the Analysis of the SEmInXpR Model

To obtain a final size equation in the SEmInXpR model, the conserved quantity ϕ (equation 34) can be augmented,

$$\phi_X = -\log(S) + \frac{\bar{\beta}}{\mu}\left(S + \sum_{j=1}^m E_j + \frac{1}{n\bar{\beta}}\sum_{j=1}^n \sum_{k=j}^n \beta_k I_j\right)$$
$$+ \frac{\bar{\beta}_X}{c}\left(\frac{\bar{\alpha}}{\mu}\left(S + \sum_{j=1}^m E_j + \frac{1}{n\bar{\alpha}}\sum_{j=1}^n \sum_{k=j}^n \alpha_k I_j\right) + \frac{1}{p\bar{\beta}_X}\sum_{j=1}^p \sum_{k=j}^p \beta_{Xk} X_j\right) \tag{50}$$

As before, all terms involving E_j, I_j, or X_j drop in the limit of long time, allowing us to fix $\phi_X(\infty)$ and, thus, to show that $f \approx 1 - \exp(-fR_0)$, where R_0 now involves components corresponding to primary and secondary infection and is given by equation 29.

To find the initial rate of disease increase, r, we follow our earlier analysis. In particular, the $(m + n + p + 2)$ by $(m + n + p + 2)$ Jacobian matrix for linearizing about the disease-free equilibrium is given by

$$J = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\beta_1 N & \cdots & -\beta_n N & -\beta_{X1}N & \cdots & -\beta_{Xp}N & 0 \\ 0 & -m\gamma & & & +\beta_1 N & \cdots & +\beta_n N & +\beta_{X1}N & \cdots & +\beta_{Xp}N & 0 \\ 0 & m\gamma & \ddots & 0 & & & & & & & \\ & 0 & \ddots & -m\gamma & 0 & & & & & & \\ & & & m\gamma & -n\mu & & & & & & \\ & & & 0 & n\mu & \ddots & 0 & & & & \\ & & & & 0 & \ddots & -n\mu & & & 0 & \\ & & & & \alpha_1 & \cdots & \alpha_n & -pc & & & \\ & & & & 0 & \cdots & 0 & pc & \ddots & & \\ & & & & & & & 0 & \ddots & -pc & \\ 0 & & & & 0 & n\mu & 0 & & & & 0 \end{pmatrix}. \tag{51}$$

Manipulation of the corresponding characteristic equation indicates that the initial growth rate is given by the largest root, r, of the equation

$$\left(1 + \frac{r}{m\gamma}\right)^m - \frac{N}{n\mu}\sum_{k=1}^n\left[(\beta_k + Q\alpha_k)\left(1 + \frac{r}{n\mu}\right)^{-k}\right] = 0 \tag{52}$$

where

$$Q = \frac{1}{pc}\sum_{j=1}^p \beta_{Xj}\left(1 + \frac{r}{pc}\right)^{-j}. \tag{53}$$

LITERATURE CITED

1. Bailey, D. J., and Gilligan, C. A. 1999. Dynamics of primary and secondary infection in take-all epidemics. Phytopathology 89:84-91.
2. Bailey, D. J., and Gilligan, C. A. 2004. Modeling and analysis of disease-induced host growth in the epidemiology of take-all. Phytopathology 94:535-540.
3. Bailey, D. J., Kleczkowski, A., and Gilligan, C. A. 2006. An epidemiological analysis of the role of disease-induced root growth in the differential response of two cultivars of winter wheat to infection by *Gaeumannomyces graminis* var. *tritici*. Phytopathology 96:510-516.
4. Bailey, D. J., Kleczkowski, A., and Gilligan, C. A. 2004. Epidemiological dynamics and the efficiency of biological control of soil-borne disease in consecutive crops. New Phytol. 161:561-575.
5. Bailey, D. J., Paveley, N., Pillinger, C., Foulkes, J., Spink, J., and Gilligan, C. A. 2005. Epidemiology and chemical control of take-all on seminal and adventitious roots of wheat. Phytopathology 95:62-68.
6. Bailey, D. J., Paveley, N., Spink, J., Lucas, P., and Gilligan, C. A. 2009. Epidemiological analysis of take-all decline in winter wheat. Phytopathology 99:861-868.
7. Brasier, C. W. J. 2010. Plant pathology: Sudden larch death. Nature 466:824-825.
8. Brassett, P. R., and Gilligan, C. A. 1988. A model for primary and secondary infection in botanical epidemics. Z. Pflanzenkrankh. Pflanzenschutz 95:352-360.
9. Chan, M.-S., and Jeger, M. J. 1994. An analytical model of plant virus disease dynamics with roguing and replanting. J. Appl. Ecol. 21:413-427.
10. Cunniffe, N. J., and Gilligan, C. A. 2010. Invasion, persistence and control in models of soil-borne plant pathogens: The effect of host demography. J. R. Soc. Interface 7:439-451.
11. Cunniffe, N. J., and Gilligan, C. A. 2011. A theoretical framework for biological control of soil-borne plant pathogens: Identifying efficient strategies. J. Theor. Biol. 278:32-43.
12. de Wit, C. T., and Goudriaan, J. 1978. Simulation of Ecological Processes. Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands.
13. Diekmann, O., and Heesterbeek, J. A. P. 2000. Mathematical Epidemiology of Infectious Diseases. Model Building, Analysis and Interpretation. Wiley, New York.
14. Diekmann, O., Heesterbeek, J. A. P., and Metz, J. A. J. 1990. On the definition and the computation of the basic reproductive ratio R_0 in models for infectious diseases in heterogeneous populations. J. Math. Biol. 28:365-382.
15. Gilligan, C. A. 2002. An epidemiological framework for disease management. Pages 1-64 in: Advances in Botanical Research. J. A. Callow, ed. Academic Press, San Diego, CA.
16. Gilligan, C. A. 2008. Sustainable agriculture and plant disease: An epidemiological perspective. Philos. Trans. R. Soc. London B 363:741-759.
17. Gilligan, C. A., Gubbins, S., and Simons, S. A. 1997. Analysis and fitting of an SIR model with host response to infection load for a plant disease. Philos. Trans. R. Soc. London B 352:353-364.
18. Gilligan, C. A., and Kleczkowski, A. 1997. Population dynamics of botanical epidemics involving primary and secondary infection. Proc. R. Soc. London Ser. B 352:591-608.
19. Goudriaan, J., and van Roermund, H. J. W. 1993. Modelling of ageing, development, delays and dispersion. Pages 89-126 in: On Systems Analysis and Simulation of Ecological Processes, with Examples in CSMP and Fortran. P. A. Leffelaar, ed. Kluwer Academic Publishers, Dordrecht, The Netherlands.
20. Gubbins, S., and Gilligan, C. A. 1997. Biological control in a disturbed environment. Philos. Trans. R. Soc. London B 352:1935-1949.
21. Gubbins, S., and Gilligan, C. A. 1997. Persistence of host-parasite interactions in a disturbed environment. J. Theor. Biol. 188:241-258.
22. Gubbins, S., and Gilligan, C. A. 1996. Population dynamics of a parasite and hyperparasite in a closed system: Model analysis and parameter estimation. Proc. R. Soc. London B263:1071-1078.
23. Gubbins, S., and Gilligan, C. A. 1997. A test of heterogeneous mixing as a mechanism for ecological persistence in disturbed environments. Proc. R. Soc. London B264:227-232.
24. Gubbins, S., Gilligan, C. A., and Kleczkowski, A. 2000. Thresholds for invasion in plant-parasite systems. Theor. Pop. Biol. 57:219-234.
25. Hall, R. J., Gubbins, S., and Gilligan, C. A. 2007. Evaluating the performance of chemical control in the presence of resistant pathogens. Bull. Math. Biol. 69:525-537.
26. Hall, R. J., Gubbins, S., and Gilligan, C. A. 2004. Invasion of drug and pesticide resistance is determined by a trade-off between biocide efficacy and relative fitness. Bull. Math. Biol. 66:835-840.
27. Hamelin, F., Castel, M., Poggi, S., Andrivon, D., and Mailleret, L. 2011. Seasonality and the evolutionary divergence of plant parasites. Ecology 92(12):2159-2166.
28. Jeger, M. J. 1986. Asymptotic behaviour and threshold criteria in model plant disease epidemics. Plant Pathol. 35:355-361.
29. Jeger, M. J. 1984. Relation between rate parameters and latent and infectious periods during a plant disease epidemic. Phytopathology 74:1148-1152.
30. Jeger, M. J. 1982. The relation between total, infectious, and postinfectious diseased plant tissue. Phytopathology 72:1185-1189.
31. Jeger, M. J., Holt, J., van den Bosch, F., and Madden, L. V. 2004. Epidemiology of insect-transmitted plant viruses: Modelling disease dynamics and control interventions. Physiol. Entomol. 29:291-304.
32. Jeger, M. J., and van den Bosch, F. 1994. Threshold criteria for model plant disease epidemics. I. Asymptotic results. Phytopathology 84:24-27.
33. Jeger, M. J., and van den Bosch, F. 1994. Threshold criteria for model plant disease epidemics. II. Persistence and endemicity. Phytopathology 84:28-30.
34. Keeling, M. J., and Grenfell, B. T. 1997. Disease extinction and community size: Modeling the persistence of measles. Science 275:65-66.
35. Keeling, M. J., and Rohani, P. 2007. Modeling Infectious Diseases in Humans and Animals. Princeton University Press.
36. Kermack, W. O., and McKendrick, A. G. 1927. A contribution to the mathematical theory of epidemics. Proc. R. Soc. London A 115:700-721.
37. Kushalappa, A. C., and Ludwig, A. 1982. Calculation of apparent infection rate in plant diseases: Development of a method to correct for host growth. Phytopathology 72:1373-1377.
38. Lloyd, A. L. 2001. Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. Proc. R. Soc. London B 268:985-993.
39. Lloyd, A. L. 2001. Realistic distributions of infectious periods in epidemic models: Changing patterns of persistence and dynamics. Theor. Pop. Biol. 60:59-71.
40. Madden, L. V. 2006. Botanical epidemiology: Some key advances and its continuing role in disease management. Eur. J. Plant Pathol. 115:3-23.
41. Madden, L. V., Hughes, G., and van den Bosch, F. 2007. The Study of Plant Disease Epidemics. American Phytopathological Society, St. Paul, MN.
42. Madden, L. V., Jeger, M. J., and van den Bosch, F. A. 2000. Theoretical assessment of the effects of vector-virus transmission mechanism on plant virus disease epidemics. Phytopathology 90:576-594.
43. Madden, L. V., and van den Bosch, F. A. 2002. population-dynamic approach to assess the threat of plant pathogens as biological weapons against annual crops. BioScience 52:65-74.
44. Mailleret, L., Castel, M., Montarry, J., and Hamelin, F. M. From elaborate to compact seasonal plant epidemic models and back: Is competitive exclusion in the details? Theor. Ecol. (In press.)
45. Meentemeyer, R., Cunniffe, N. J., Cook, A. J., Filipe, J. A. N., Hunter, R. D., Rizzo, D., and Gilligan, C. A. 2011. Epidemiological modeling of invasion in heterogeneous landscapes: Spread of sudden oak death in California (1990–2030). Ecosphere 2:art17.
46. Metz, J. A. J., and Diekmann, O. 1986. The Dynamics of Physiologically Structured Populations. Springer, Berlin.
47. Murray, J. D. 2002. Mathematical Biology: I. An Introduction. Springer, Berlin.
48. Ndeffo-Mbah, M. L., and Gilligan, C. A. 2010. Balancing detection and eradication for control of epidemics: Sudden oak death in mixed-species stands. PLOS One 5:e12317.
49. Ndeffo-Mbah, M. L., and Gilligan, C. A. 2010. Optimization of control strategies for epidemics in heterogeneous populations with symmetric and asymmetric transmission. J. Theor. Biol. 262:757-763
50. Nguyen, H. T. H., and Rohani, P. 2008. Noise, nonlinearity and seasonality: The epidemics of whooping cough revisited. J. R. Soc. Interface 5:403-413.
51. Onstad, D. W., and Kornkven, E. A. 1992. Persistence and endemicity of pathogens in plant populations over time and space. Phytopathology 82:561-566.
52. Papastamati, K., and Bosch, F. 2007. The sensitivity of the epidemic growth rate to weather variables, with an application to yellow rust on wheat. Phytopathology 97:202-210.
53. Park, A. W., Gubbins, S., and Gilligan, C. A. 2001. Invasion and persistence of plant parasites in a spatially structured host population. OIKOS

94:162-174.

54. Parnell, S., Gilligan, C. A., and van den Bosch, F. 2005. Small-scale fungicide spray heterogeneity and the coexistence of resistant and sensitive pathogen strains. Phytopathology 95:632-639.

55. Parnell, S., van den Bosch, F., and Gilligan, C. A. 2006. Large-scale fungicide spray heterogeneity and the regional spread of resistant pathogen strains. Phytopathology 96:549-555.

56. Rizzo, D., Garbelotto, M., and Hansen, E. 2005. *Phytophthora ramorum*: Integrative research and management of an emerging pathogen in California and Oregon forests. Annu. Rev. Phytopathol. 43:309-335.

57. Roberts, M. G., and Heesterbeek, J. A. P. 2007. Model-consistent estimation of the basic reproduction number from the incidence of an emerging infection. J. Math. Biol. 55:803-816.

58. Rowthorn, R. E., Laxminarayan, R., and Gilligan, C. A. 2009. Optimal control of epidemics in metapopulations. J. R. Soc. Interface 6:1135-1144.

59. Segarra, J., Jeger, M. J., and van den Bosch, F. 2001. Epidemic dynamics and patterns of plant disease. Phytopathology 91:1001-1010.

60. Shaw, M. W. 2006. Pathogen population dynamics. Pages 193-214 in: The Epidemiology of Plant Diseases. B. M. Cooke, G. D. Jones, and B. Kaye, eds. Springer, The Netherlands.

61. Sherm, H., Ngugi, H. K., and Ojiambo, P. S. 1996. Trends in theoretical plant epidemiology. Eur. J. Plant Pathol. 115:61-73, 2006.

62. Swinton, J., and Gilligan, C. A. Dutch elm disease and the future of the elm in the U.K.: A quantitative analysis. Philos. Trans. R. Soc. London B 351:605-615.

63. Teng, P. S. 1985. A comparison of simulation approaches to epidemic modeling. Annu. Rev. Phytopathol. 23:351-379.

64. Tomlinson, I., and Potter, C. 2010. 'Too little, too late'? Science, policy and Dutch elm disease in the UK. J. Hist. Geogr. 36:121-131.

65. Truscott, J. E., Webb, C. R., and Gilligan, C. A. 1997. Asymptotic analysis of an epidemic model with primary and secondary infection. Bull. Math. Biol. 59:1101-1123.

66. van den Berg, F., Gilligan, C. A., Bailey, D. J., and van den Bosch, F. 2010. Periodicity in host availability does not account for evolutionary branching as observed in many plant pathogens: An application to *Gaeumannomyces graminis* var. *tritici*. Phytopathology 100:1169-1175.

67. van den Berg, F., and van den Bosch, F. 2007. The elasticity of the epidemic growth rate to observed weather patterns with an application to yellow rust. Phytopathology 97:1512-1518.

68. van den Bosch, F., Frinking, H. D., Metz, J. A. J., and Zadoks, J. C. 1988. Focus expansion in plant disease. III: Two experimental examples. Phytopathology 78:919-925.

69. van den Bosch, F., Jeger, M. J., and Gilligan, C. A. 2007. Disease control and its selection for damaging plant virus strains in vegetatively propagated staple food crops; A theoretical assessment. Proc. R. Soc. London B 274:11-18.

70. van den Bosch, F., Zadoks, J. C., and Metz, J. A. J. 1988. Focus expansion in plant disease. II: Realistic parameter sparse models. Phytopathology 78:59-64.

71. van der Plank, J. E. 1963. Plant Diseases: Epidemics and Control. Academic Press, London.

72. van Kraalingen, D. W. G., Rappoldt, C., and van Laar, H. H. 2003. The Fortran simulation translator, a simulation language. Eur. J. Agron. 18:359-361.

73. Waggoner, P. E. 1977. Contributions of mathematical models to epidemiology. Ann. N. Y. Acad. Sci. 287:191-206.

74. Waggoner, P. E. 1974. Simulation of epidemics. Pages 137-160 in: Epidemics of Plant Diseases: Mathematical Analysis and Modelling. J. Kranz, ed. Springer-Verlag, Berlin.

75. Wearing, H. J., Rohani, P., and Keeling, M. J. 2005. Appropriate models for the management of infectious diseases. PLOS Med. 2(7):3174. doi:10.1371/journal.pmed.0020174.

76. Webb, C. R., Gilligan, C. A., and Asher, M. J. C. 1999. A model for the temporal buildup of *Polymyxa betae*. Phytopathology 89:30-38.

77. Webb, C. R., Gilligan, C. A., and Asher, M. J. C. 2000. Modelling the effect of temperature on the dynamics of *Polymyxa betae*, the vector of Rhizomania disease. Plant Pathol. 49:600-607.

78. Zadoks, J. C. 1971. Systems analysis and the dynamics of epidemics. Phytopathology 61:600-610.