Bayesian inference for spatio-temporal stochastic transmission of plant disease in the presence of roguing: A case study to characterise the dispersal of Flavescence dorée

Hola K Adrakey, Gavin J Gibson, Sandrine Eveillard, Sylvie Malembic-Maher, Frederic Fabre

PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Bayesian inference for spatio-temporal stochastic transmission of plant disease in the presence of roguing: A case study to characterise the dispersal of Flavescence dorée

Hola K. Adrakey [1], Gavin J. Gibson [2], Sandrine Eveillard [3], Sylvie Malembic-Maher [3], Frederic Fabre [1]*

1 UMR SAVE, INRAE, Bordeaux Sciences Agro, Villenave d'Ornon, France, **2** Maxwell Institute for Mathematical Sciences, School of Mathematical and Computer Sciences, Heriot–Watt University, Edinburgh, United Kingdom, **3** UMR BFP, INRAE, Univ. Bordeaux, Villenave d'Ornon, France

* frederic.fabre@inrae.fr

Citation: Adrakey HK, Gibson GJ, Eveillard S, Malembic-Maher S, Fabre F (2023) Bayesian inference for spatio-temporal stochastic transmission of plant disease in the presence of roguing: A case study to characterise the dispersal of Flavescence dorée. PLoS Comput Biol 19(9): e1011399. https://doi.org/10.1371/journal.pcbi.1011399

Editor: Alex Perkins, University of Notre Dame, UNITED STATES

Received: January 27, 2023

Accepted: July 28, 2023

Published: September 1, 2023

Data Availability Statement: The authors confirm that all data underlying the findings are fully available without restriction. The FD dataset, MCMC codes (written in C language) and scripts to generate the figures and tables (written in R language) are available in INRAE Dataverse at https://doi.org/10.57745/YXOEHX.

Funding: This study received financial support from Plan National Deperissement du Vignoble

Abstract

Estimating the distance at which pathogens disperse from one season to the next is crucial for designing efficient control strategies for invasive plant pathogens and a major milestone in the reduction of pesticide use in agriculture. However, we still lack such estimates for many diseases, especially for insect-vectored pathogens, such as Flavescence dorée (FD). FD is a quarantine disease threatening European vineyards. Its management is based on mandatory insecticide treatments and the removal of infected plants identified during annual surveys. This paper introduces a general statistical framework to model the epidemiological dynamics of FD in a mechanistic manner that can take into account missing hosts in surveyed fields (resulting from infected plant removals). We parameterized the model using Markov chain Monte Carlo (MCMC) and data augmentation from surveillance data gathered in Bordeaux vineyards. The data mainly consist of two snapshot maps of the infectious status of all the plants in three adjacent fields during two consecutive years. We demonstrate that heavy-tailed dispersal kernels best fit the spread of FD and that on average, 50% (resp. 80%) of new infection occurs within 10.5 m (resp. 22.2 m) of the source plant. These values are in agreement with estimates of the flying capacity of *Scaphoideus titanus*, the leafhopper vector of FD, reported in the literature using mark–capture techniques. Simulations of simple removal scenarios using the fitted model suggest that cryptic infection hampered FD management. Future efforts should explore whether strategies relying on reactive host removal can improve FD management.

Author summary

The dispersal of pathogen propagules is an important feature of spatial epidemiology that has a major impact on the prevalence and distribution of disease in a population. In

under research contracts RISCA and Co-Act2. H.A. was funded by RISCA and Co-Act2 through a postdoctoral fellowship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

agriculture, properly characterising the dispersal of emerging diseases is of great importance in designing science-based control strategies that allow pesticide use to be reduced. Although field epidemiological surveys can provide informative data, they are by nature rare while resulting from the interactions between disease spread and the undergoing surveillance and control. Here, we take advantage of a general statistical framework to model the epidemiological dynamics of Flavescence dorée (FD), a quarantine disease threatening European vineyards, in a mechanistic manner that can take into account missing hosts in surveyed fields (resulting from infected plant removals). We parameterized the model with a Bayesian approach using mainly two snapshot maps of the infectious status of all plants in three adjacent fields during two consecutive years. We demonstrate that on average, 50% (resp. 80%) of new FD infection occurs within 10.5 m (resp. 22.2 m) of the source plant. Although FD mainly spreads locally from one year to the next, our results also indicate frequent long-distance dispersal events, a feature crucial to consider when designing control strategies.

Introduction

The management of emerging plant pathogens is of foremost importance in the effort to reduce pesticide use in agriculture, particularly since the rate of introduction of alien species is higher than ever due to increasing globalisation [1]. The design of efficient management strategies has much to gain from the recent advances in epidemiological modelling. In particular, spatially explicit stochastic models are able to simulate realistic patterns of epidemic spread and test the effectiveness of a range of management strategies (for a review, see [2] and, more recently, [3–8]). However, for this approach to be feasible, the underlying model parameters must be estimated.

The main issue encountered when fitting spatially explicit mechanistic models is the type and nature of the available data. First, epidemiological surveys most often provide information on epidemic dynamics in the presence of control, which must accordingly be explicitly accounted for in models and fitting procedures [9, 10]. Second, there is a trade-off between the number of fields that can be studied and the number of seasons over which they can be observed. The data collected most often contain only partial spatial and/or temporal information on the location of infected individuals. Moreover, the timing of true infection events is typically not observed. The first symptom expression in an individual plant (the most easily accessible variable during field surveys) can take place a long time after that plant becomes infected and even infectious. The latter situation results from the presence of cryptic infection, during which an infectious host is asymptomatic [2, 11]. All of this context involves the existence of missing data. Data augmentation, a technique first introduced to plant disease epidemiology by [12], enables tractability of model likelihoods and so, in turn, facilitates rigorous Bayesian inference for mechanistic models. Within this framework, unobserved events, typically the timing of infection events, are treated as additional unknown parameters, and the inference is based on Markov chain Monte Carlo (MCMC) methods. These methods have had great success in analysing plant epidemic data over several years [3, 9, 10, 13–18].

Here, Markov chain Monte Carlo (MCMC) was used to fit a spatially explicit stochastic model of Flavescence dorée (FD) spread. FD is a quarantine disease threatening European vineyards. FD is one of the most damaging diseases in European vineyards. It is caused by the FD phytoplasma (taxonomic subgroups 16SrV-C and 16SrV-D), a small bacterium that multiplies in the phloem sap of infected plants. FD phytoplasma is transmitted from grapevine to

grapevine by the leafhopper vector *Scaphoideus titanus* Ball [19]. Typical symptoms are leaf yellowing or reddening, downward rolling, incomplete lignification of canes, abortion of flowers, and grape wilting. FD phytoplasma, which is endemic to European alders, emerged as a disease in South West France in the 1950s, following the accidental introduction of the ampelophagous vector *S. titanus* from North America [20]. Since then, the disease has spread throughout European vineyards [21]. Due to the severe economic consequences of the disease, FD phytoplasma has been classified as a quarantine organism in Europe since 1993. There is currently no means of curing plants of FD phytoplasma. The current regulation of the disease relies mainly on four mandatory measures: (i) the planting of disease-free material, (ii) the application of insecticides to kill the vector, (iii) the establishment of annual vineyard surveys for monitoring plant infection, and (iv) the removal of infected plants.

The epidemiology of FD in vineyards is driven primarily by the spatio-temporal dynamics of its univoltine vector *S. titanus*. The eggs hatch in May, and there are then five nymphal instars before the first adults appear, usually in July. The adults live from one to two months [22, 23]. The fertilised females can lay eggs from around 10 days after their emergence till the end of summer [22]. Considering as starting point a plant infected during year t, the FD epidemic cycle occurs as follows. The vine plant usually becomes infectious from the spring of year $t + 1$ [24–26] as a result, during fall and winter, of its own dormancy stage and of the concomitant diapause of the eggs of the vector. Phytoplasmas are acquired passively in year $t + 1$, through feeding on infected plants, from the first larval stage onwards hatched in May. Winged adults mainly, but also last instar larvae, can transmit the phytoplasma to healthy plants, from late spring to late summer [27, 28]. While some non-specific FD symptoms such as delayed bud break may appear in early spring of year $t + 1$, the set of specific symptoms mostly appear throughout the summer [22, 26]. In the Bordeaux vineyard, the probability of detecting the set of typical FD symptoms is substantially increased after the third week of August [29]. In the vineyard, the epidemiology of FD is also influenced by vine cultivar effects. Although no genetic resistance to FD is known, vine cultivars differ in their susceptibility to (i) phytoplasma transmission by *S. titanus* and (ii) phytoplasma multiplication and diffusion in the plant [30–32]. For example, Merlot and Cabernet Sauvignon, the two main cultivars in the Bordeaux area, display contrasting susceptibilities. Merlot is characterised by a lower proportion of infected shoots and reduced phytoplasma titers in the leaves as compared to the highly susceptible cultivar Cabernet Sauvignon [31]. These features affect FD epidemiology, as phytoplasma acquisition efficiency by *S. titanus* increases with its titers in the plant [27, 33]. Direct cultivar effects on the feeding behaviour of *S. titanus* have also been shown to explain some cultivars' higher susceptibility to FD [34].

A feature of FD epidemiology that is poorly understood is its spatial dispersal kernel. Dispersal kernels represent the statistical distribution of the location of the infected hosts after inoculum dispersal from a focal plant source [35, 36]. Dispersal kernels play a crucial role in the design of control strategies for emerging pathogens that rely on reactive host removal. They consist of removing plants within a particular distance of locations identified as containing the infection. The rationale is to remove (or treat) locations that are likely to be infected but not yet showing symptoms [11, 36–38]. Here, recent georeferenced inspection and eradication data supplied by extensive surveillance at the landscape scale (representing an area of 28 hectares) was used to fit a mechanistic and spatio-temporal model of FD spread from one season to the next. The model builds on the work of [14] regarding the spread of Huanglongbing through citrus groves. It provides insights into the dispersal kernel that underlies FD spread. For coherent integration of information from different observations gathered at plant and field scales, we worked within a Bayesian framework. A key feature of the approach considered here is the use of data augmentation techniques to account for the missing information in the

data related to plants removed before the initial survey date in order to construct the full trajectory of the epidemic. Accounting for this information enables us to use MCMC, a Bayesian computational method to sample from the posterior distribution of the model parameters.

Materials and methods

Data

A plant-by-plant survey of FD symptoms in three fields. The study area is situated in Faleyras (S1(A) Fig), a district near Bordeaux in South West France, where three plots labelled F_1, F_2 and F_3 (S1(B) Fig) were monitored. These plots belong to the same owner. The plot F_1 was planted in 2002, F_2 in 1976 and F_3 in 1998. They are located in an open vineyard landscape at an altitude of 60—70 m with some woods and meadows (S1(A) Fig). The three plots are not separated by any hedgerow. The initial state of the fields before the FD incursion is shown in S1(B) Fig. In that initial state, there were 2259, 677 and 3120 plants distributed in F_1, F_2 and F_3 respectively, with two different cultivars, 2354 Cabernet Sauvignon (CS) and 3702 Merlot (M), the most widely grown cultivars in the region (S1(B) Fig). Of these cultivars, 2259 CS were planted in F_1, 677 M in F_2 and 3025 M and 95 CS in F_3.

The three fields were extensively monitored on October 11, 2018, and on September 23, 2019, following the methodology described in [31]. The three fields were mapped by distinguishing the non-symptomatic, the symptomatic and the missing (due to removal) plants. Over the two studies performed by the same team (this one and [31]), in fields with high FD prevalence and sampling done at the end of summer, 98% of symptomatic Merlot plants ($n = 207$) and 99% of symptomatic Cabernet Sauvignon plants ($n = 202$) were positive for FD. Similarly, 97% of non-symptomatic Merlot plants sampled ($n = 88$) and 98% of non-symptomatic Cabernet Sauvignon plants sampled ($n = 96$) were negative in real-time PCR tests. Genotyping of FD phytoplasma was performed by nested PCR followed by sequencing of the secY-map gene [39] on 4 samples of Merlot (in F_2 and F_3) and 4 samples of Cabernet Sauvignon (in F_1 and F_3) from 2018 and 2019. All samples were infected by the M54 genotype (Map-FD2) which is predominant in outbreaks from Europe and France [20].

For each symptomatic plant, symptom severity was recorded. A severity score from 1 to 4 was assigned depending on the proportion of symptomatic branches on the stock: 1) 0–25%, 2) 26–50%, 3) 51–75% and 4) >75%. The symptomatic plants were marked, and the wine grower was asked to remove them during the winter following the inspection. The growers may have also removed some non-symptomatic plants for other agronomic reasons. The two snapshots of the health status of the fields in October 2018 and September 2019 are shown in Fig 1A and 1B respectively. In 2018, 713 symptomatic plants including 668 CS and 45 M were detected, while 411 were missing (352 CS and 59 M). Field F_1 was entirely removed during winter 2018 in compliance with FD management rules because its prevalence (i.e. the percentage of diseased plant in the field) exceeded 20%. In 2019, 638 new symptomatic plants including 46 CS and 592 M were recorded on F_2 and F_3. During this same year, 50 additional missing plants, arising from the removal of asymptomatic plants, were identified in F_2 and F_3.

Disease prevalence in the extended area of the focal fields. The three fields, F_1, F_2 and F_3, are located in a larger region where FD surveys are performed by a professional organisation known as GDON (Groupements de Défense contre les Organismes Nuisibles) des Bordeaux. The district of Faleyras was surveyed in 2014. The survey is conducted as follows. Inspectors walk around all the fields in the area in search of symptomatic plants. Once symptomatic plants are located in a given field, the inspector collects symptomatic leaves from one to five plants and pools them into a single sample. The detection of the FD phytoplasmas in samples collected is performed by accredited laboratories using a real-time PCR test [40]. Each

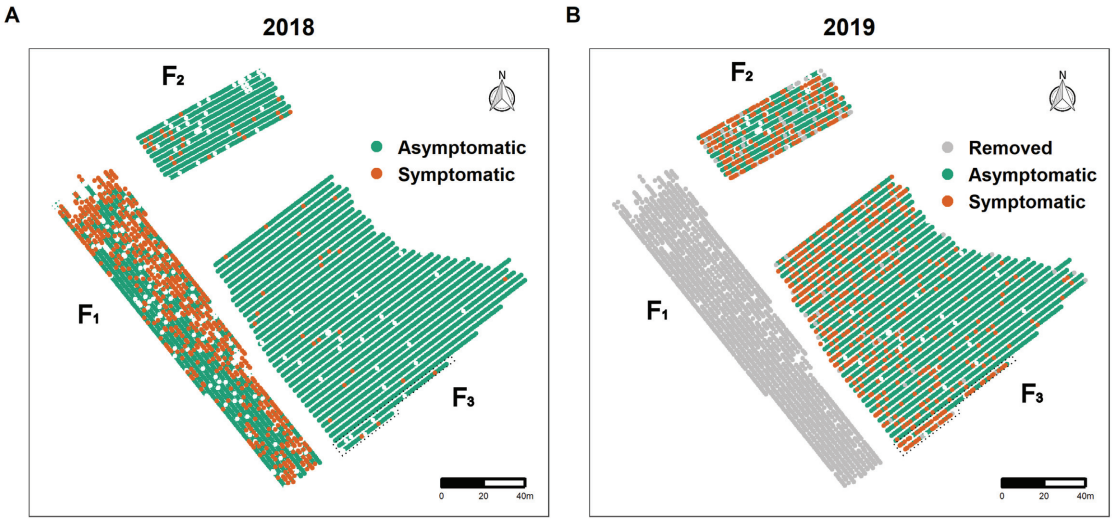

Fig 1. Individual status of the plants in the three vineyard fields F_1, F_2 and F_3 considered. The field F_1 was planted with 2259 Cabernet Sauvignon, the field F_2 with 677 Merlot, and the field F_3 with 3025 Merlot and 95 Cabernet Sauvignon in two rows surrounded by a dotted line. The status of the plants is i) asymptomatic, ii) symptomatic or iii) removed. A: Spatial distribution of 713 FD-symptomatic plants including 668 CS and 45 M detected (red) in 2018 (October 11) as well as the 411 (352 CS and 59 M) holes showing plant removals before 2018. B: Spatial distribution of 638 FD-symptomatic plants including 46 CS and 592 M detected in 2019 (September 23), 2011 plants removed in 2018 (in grey) and the same 411 (352 CS and 59 M) missing plants indicating removal before 2018. Note that real-time PCR tests indicate that more than 98% of symptomatic plants are infected with FD, while more than 97% of asymptomatic plants are FD negative. Maps were drawn by authors using own data set released in the dataverse at https://doi.org/10.57745/YXOEHX.

https://doi.org/10.1371/journal.pcbi.1011399.g001

field is then classified as infected or not. If the field is infected, the inspectors count its total number of symptomatic plants. The field will be monitored in subsequent years. The infection status of the fields monitored by the GDON des Bordeaux in 2014 within 300 m of the three focal fields is shown in Fig 2A. Note that the fields in the neighbouring district of Romagne (the northeastern part of the map) were not inspected. From 2015 to 2017, the survey continued in the fields found to be infected in 2014 and spread to nearby fields after 2017 (Fig 2B, 2C, 2D and 2E). Note that GDON inspectors did not record the precise localisation of symptomatic plants. This information is only available in F_1, F_2 and F_3 for 2018 and 2019 following a dedicated survey performed by INRAE staff.

As shown in Fig 2A, FD was first detected in the area in 2014, most prevalently in field F_1 but also in three other fields. FD has since spread to more fields. The symptoms in the fields F_2 and F_3 were first reported by GDON in 2018 during the second round of inspection of F_2 and F_3 since 2014. Accordingly, their year of initial infection is unknown. In agreement with the set of mandatory measures to control FD, one or two insecticide treatments were recommended from 2015 to 2018 in the area corresponding to the districts of Faleyras and Romagne.

Epidemiological models

A mechanistic compartmental SIR model with a discrete time step of one year was considered to describe the infectious status of each plant from 2014 to 2019 within the two focal fields F_2 and F_3. Three statuses were considered: Susceptible (noted S, *i.e.* healthy plant), Infectious (noted I, *i.e.* infected plant able to transmit infection) and Removed (noted R). The annual discrete time step assumed fits the univoltine cycle of *S. titanus* (*i.e.* one generation per year) and the existence of a dormancy phase (both for the plants and the eggs of the vector). This

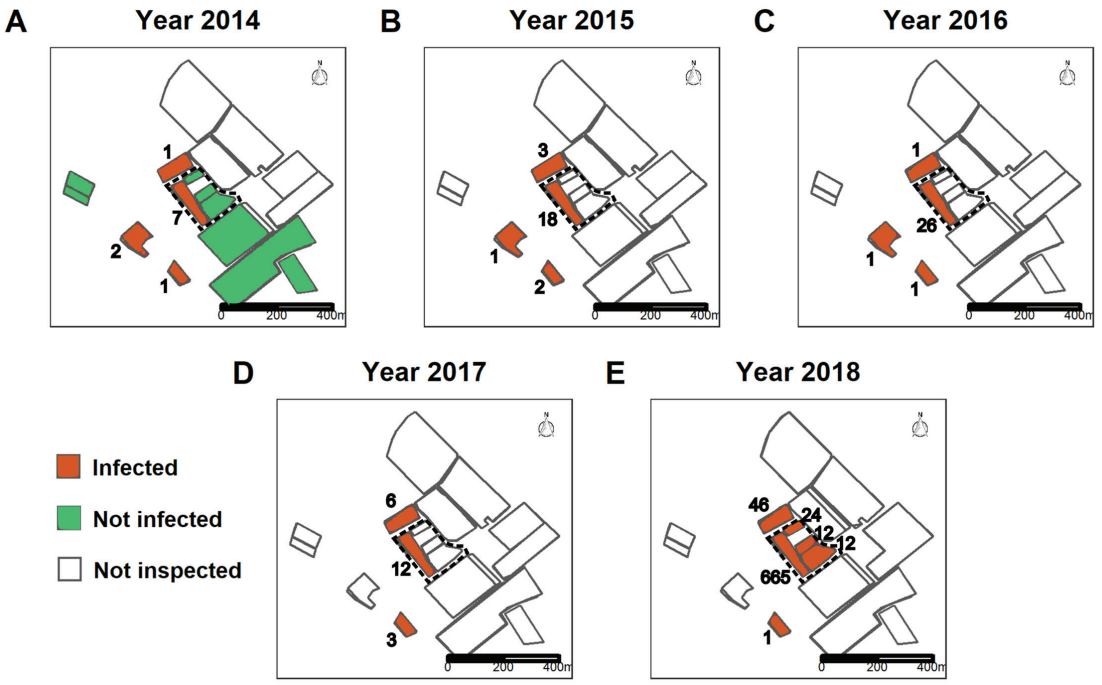

Fig 2. Status of fields within 300 m around the targeted fields from 2014 to 2018. The status of fields is either i) not inspected, ii) not infected by FD (if symptomatic plants were sampled but the PCR test was negative) or iii) infected by FD. For the latter category, the values inside each infected correspond to the number of symptomatic plants recorded and subsequently removed. The area with the three focal fields, F_1, F_2 and F_3, is circled in dotted line. Maps were drawn by authors using own data set released in the dataverse at https://doi.org/10.57745/YXOEHX.

https://doi.org/10.1371/journal.pcbi.1011399.g002

dormancy phase, which results in infected plants in year t becoming infectious in year $t + 1$ (see Introduction), explained why an exposed compartment (host infected, but not yet able to transmit infection) was not considered here. Moreover, all infectious plants in year $t + 1$ were assumed to be removed in year $t + 2$ in compliance with the mandatory measures applied in France to manage FD (see below). Note that, as this transition is deterministic, the model corresponds formally to a SI model.

Modelling the infection pressure. Let first introduce the probability $P_i(t)$ that a susceptible plant i is infected in year t (and therefore becomes infectious in year $t + 1$). This probability is defined as

$$P_i(t) = 1 - exp(-\phi_i(t)). \quad (1)$$

where

$$\phi_i(t) = \epsilon(t) + \beta \sum_{j \in I_{t-1}} c_j K(r_{ij}; \alpha) \mathbb{1}_{j \in F_1} + \beta \sum_{j \in I_{t-1}} c_j K(r_{ij}; \alpha) \mathbb{1}_{j \in F_2 \cup F_3} \quad (2)$$

Eq 2 is the infection pressure on host i at year t. Its first two terms together represent the infection pressure exerted on plants located in fields F_2 and F_3. The first term represents the instantaneous infection pressure exerted by all the fields other than F_2 and F_3 at time t = 2014, 2015, 2016, 2017, and by all the fields other than F_1, F_2 and F_3 in 2018 as an infection snapshot was available for F_1 in 2018. It corresponds to a primary infection rate that varies with time in proportion to the total number of infectious plants in the neighbouring fields (except for the plants in field F_1 in 2018). Specifically, $\epsilon(t) = \epsilon n_{t-1}$, where n_{t-1} is the number of symptomatic

Table 1. Total number of FD-symptomatic plants detected in a radius of 300 m around the focal fields F_2 and F_3. The values of n_t include the number of symptomatic plants in all the neighbouring fields including the field F_1 from 2014 to 2017 but excluding F_1 in 2018. These symptomatic plants are assumed to be infectious in the model.

t	2014	2015	2016	2017	2018
n_t	11	24	29	21	47

https://doi.org/10.1371/journal.pcbi.1011399.t001

plants at year $t-1$ within a radius 300 m measured from the centre of the two focal fields F_2 and F_3 (Fig 2 and Table 1). The second term represents the external force of infection exerted by the infectious plants in field F_1 in 2018 on the plants in the fields F_2 and F_3 in 2019. It incorporates a distance-dependent kernel $K(r_{ij};\alpha)$ shared with the third term, which represents the secondary infections taking place within the fields F_2 and F_3. In the second and third terms, the sum is over infectious individuals j (i.e. individuals infected during year $t-1$). These terms also include a parameter c_j that accounts for the differences in infectivity between cultivars. Differences in infectivity were assumed to be related to the proportion of symptomatic branches on the stock. Accordingly, it was estimated independently using the severity score ranging from 1 to 4 that was described previously. Of the 668 CS plants recorded in 2018, 328 scored 1, 126 scored 2, 87 scored 3 and 127 scored 4, leading to a mean severity score value of 2.02. Of the 637 M plants recorded in 2018 and 2019, 437 scored 1, 147 scored 2, 38 scored 3 and 15 scored 4, leading to a mean severity score value of 1.42. Accordingly, c_j was set to 1.42 ($\approx$ 2.02/1.42) if the cultivar of the source plant is Cabernet Sauvignon and to 1 if the cultivar is Merlot.

The dispersal kernel K is a non-negative function also known as the infection kernel. It characterises the spatial dispersal process with a parameter α. Specifically, it characterises the relationship between infective challenge and the relative positions of infected and susceptible plants. A number of candidate models were proposed for K:

$$\textbf{Exponential}: \quad K(r;\alpha) = \frac{1}{2\pi\alpha^2}\exp\left(-\frac{r}{\alpha}\right) \text{ with } \mu = 2\alpha \tag{3}$$

$$\textbf{Gaussian}: \quad K(r;\alpha) = \frac{1}{\pi\alpha^2}\exp\left(-\frac{r^2}{\alpha^2}\right) \text{ with } \mu = \alpha\sqrt{\pi}/2 \tag{4}$$

$$\textbf{Power law}: \quad K(r;\alpha) = \frac{(b-2)(b-1)}{2\pi\alpha^2}\left(1+\frac{r}{\alpha}\right)^{-b}, \text{ with } \mu = \frac{2\alpha}{b-3}\frac{2\pi}{(b-2)(b-1)} \tag{5}$$

$$\text{for } b = 4, 6, 8$$

where r is the Euclidean distance (measured in m) between a given pair of plants. The mean of the dispersal kernel is given by μ and corresponds to the mean distance to the first infection in a totally susceptible population. The Gaussian and exponential kernels correspond to thin-tailed kernels, while the power law kernel corresponds to thick-tailed kernels. Kernels such as power law result in a rapid and long dispersal ahead of the source [41, 42]. Note that these 1-dimensional kernels are isotropic and can represent different spatial patterns of epidemic dynamics. They also allow these different spatial patterns to be compared directly.

Overall, 20 models were fitted. They represented all combinations of two hypotheses regarding the external source $\epsilon(t)$ (1: constant, by setting to $\epsilon(t) = \epsilon$ in Eq 2, or 2: variable between years) with two hypotheses regarding a cultivar infectivity effect (1: presence, by setting $c_y = 1.42$ for a Cabernet Sauvignon source plant and $c_y = 1$ for a Merlot source plant, or 2:

Table 2. Specification of the 20 models considered. They include 2 × 2 × 5 combinations of hypotheses regarding the primary infection rate (column External source), the presence of a cultivar effect (column Cultivar effect) and the shape of the dispersal kernels (column Kernel). The numbering of equations (columns Eqn) corresponds to the equations detailed in Text A in S1 Text for the infection pressure and in the main text for the kernels.

Label	Infection pressure ($\phi(x;t)$)			Kernel $K(r;\alpha)$	
	External source	Cultivar effect	Eqn	Name	Eqn
Model 0	Constant	No	S1	Gaussian	4
Model 1				Exponential	3
Model 2				Power law, $b = 4$	5
Model 3				Power law, $b = 6$	5
Model 4				Power law, $b = 8$	5
Model 5	Variable	No	S2	Gaussian	4
Model 6				Exponential	3
Model 7				Power law, $b = 4$	5
Model 8				Power law, $b = 6$	5
Model 9				Power law, $b = 8$	5
Model 10	Constant	Yes	S3	Gaussian	4
Model 11				Exponential	3
Model 12				Power law, $b = 4$	5
Model 13				Power law, $b = 6$	5
Model 14				Power law, $b = 8$	5
Model 15	Variable	Yes	S4	Gaussian	4
Model 16				Exponential	3
Model 17				Power law, $b = 4$	5
Model 18				Power law, $b = 6$	5
Model 19				Power law, $b = 8$	5

https://doi.org/10.1371/journal.pcbi.1011399.t002

absence, by setting $c_y = 1$ regardless of the cultivar of the source plant) and five hypotheses regarding the shape dispersal kernel. The 20 models considered are summarised in Table 2, and the corresponding equations for the infection pressure are detailed in Text A in S1 Text. Finally, the parameters to be estimated are summarised in Table 3.

Modelling the initial infection year of the fields F_2 and F_3. The fields F_2 and F_3 were not inspected from 2015 to 2017 (Fig 2). Let t_0 denote the year the inoculum was first transmitted to either of these fields. The unobserved levels of initial infection of the fields F_2 and F_3 (i.e. number and position of infected plants at t_0) are treated as additional model parameters.

Table 3. Definition of model parameters and their mean posterior values along with their 95% credible interval obtained from the MCMC algorithm using Model 18. The credible interval for t_0 is ignored, as there was a negligible posterior weight on years other than 2016.

Process	Symbol	Description	Estimated value	Unit[1]
Dispersal	α	Scale parameter	22.9 [20.1—26.5]	m
Infection	ϵ	Primary infection rate	$10^{-4}[10^{-4}–2.10^{-4}]$	$year^{-1}$
	β	Secondary infection rate	106.1[94.5–118.4]	$m^2.year^{-1}$
	c_y	Cultivar's infectivity[2]	$c_y = 1$ or $c_y = 1.42$	na
Removal	q	Probability of removal for reasons other than FD	0.004 [0.004—0.005]	na
Initial infection date in $F_2 \cup F_3$	t_0	Year of first infection	2016	na

[1] na when not applicable or for dimensionless parameters.

[2] Values estimated independently using data published by [31].

https://doi.org/10.1371/journal.pcbi.1011399.t003

Modelling removal process. The set of specific FD symptoms mostly appear throughout the summer and are best identified between mid-August and mid-October in the Bordeaux vineyard [29]. Symptomatic plants were assumed to be removed during the winter following their identification. This succession during the same year of the periods where a plant is first infectious, then expressed the typical set of FD symptoms, next identified as such and finally removed led us to consider thereafter that the FD survey currently implemented in the Bordeaux area is characterised by the existence of "cryptic" infection from late spring/early summer until optimal detection of symptoms in late summer/early autumn. It corresponds *sensu* [2] to a period of time where infectious hosts go undetected while providing a source of inoculum for spread. Finally, we assumed that any susceptible individual could be removed in a given year for some agronomic reason unrelated to FD infection. The probability q of this event was considered as an additional model parameter.

Model inference

Likelihood. The data **y** record the status of plants at n discrete observation times in $[t_0, t_{max}]$, along with information on removed plants. In addition, let denote the set of removed hosts with $\mathcal{R}$, the corresponding set of reasons hosts are removed with $R = (r_1, r_2, \ldots, r_{|\mathcal{R}|})$ and the corresponding set of removal times with $\underline{t}^{\mathcal{R}} = (t_1^{\mathcal{R}}, t_2^{\mathcal{R}}, \ldots, t_{|\mathcal{R}|}^{\mathcal{R}})$. It is worth noting that $|\mathcal{R}|$ represents the size of the removed hosts in the population. For simplicity, r_j is set to 0 if the host j was removed due to causes other than FD, or to 1 if the host was removed due to FD infection. The set of reasons hosts are removed R comprises those with known reason R_k and those for which the reason for removal is unknown R_u. In other words, $R = R_k \cup R_u$. Finally, let us denote the set of the model parameters with $\theta = (\alpha, \epsilon, \beta, q, t_0)$. Then, taking the reasons hosts are removed and their corresponding times as known, the complete data likelihood function is given as

$$\begin{aligned} L(\theta, R, \underline{t}^{\mathcal{R}}; y) = \quad & \prod_{j\in\mathcal{R}}\left[\left(\prod_{i=1}^{t_j^{\mathcal{R}}-1} \exp\left(-\phi_j(i)\right)(1-q)\right)\left((1-\exp\left(-\phi_j(t_j^{\mathcal{R}})\right))(1-q)\mathbb{1}_{r_j=1} + \right.\right. \\ & \left.\left. \exp\left(-\phi_j(t_j^{\mathcal{R}})\right)q\mathbb{1}_{\{r_j=0\}}\right)\right] \times \\ & \prod_{j\in\bar{\mathcal{R}}}\prod_{i=1}^{t_{max}} \exp\left(-\phi_j(i)\right)(1-q) \end{aligned} \tag{6}$$

The first two lines of Eq 6 represent the contribution to the likelihood arising from the removal events. Two classes of removed hosts are considered: those removed due to the disease ($r_j = 1$) and those removed due to causes other than FD ($r_j = 0$). The term $(1 - q)$ ensures that symptomatic plants are not removed due to any other causes. The third line represents the contribution to the likelihood of plants that escaped the removal with probability $1 - q$ and also escaped infection.

Bayesian inference. The likelihood $L(\theta, R, \underline{t}^{\mathcal{R}}; y)$ is not straightforward to compute, as the locations, dates and natures of infections between 2014 and 2017 were not observed. However, this problem can be overcome using the data-augmented approach by treating the missing plants in 2018 as an additional unknown "parameter" and investigating the joint posterior distribution $P(\alpha, \epsilon, \beta, q, R_u, \underline{t}^{\mathcal{R}_u}|y)$ [3, 13–15, 17, 43]. The algorithm involves proposing accepting changes to the times and nature of removals before 2018 and provides posterior distribution of infection sources as a by-product. This is done using methods which are now standard in computational epidemiology, such as reversible-jump [44, 45] MCMC techniques. Details of algorithms are given in Text B in S1 Text.

Model checking and model comparison

The MCMC algorithm described in Text B in S1 Text resulted in a joint posterior distribution of both the model parameters and the time and nature of missing plants during the first visit to fields F_2 and F_3 in 2018. These outputs can be used both to perform posterior model checking and to select the subset of candidate models that best explains the data. In the context of a spatio-temporal transmission model, natural model selection criteria such as the deviance information criterion (DIC) suffer from important limitations because of their sensitivity and complexity [46]. Following work by [12, 14, 15, 17] in a closed context, model selection was performed by comparing (i) the posterior predictive distribution of the counts of symptomatic and removed plants and (ii) the posterior predictive distribution of the spatial structure of the epidemic to the observed ones.

Count of symptomatic and removed plants in 2018 and 2019. Samples were randomly drawn from the joint posterior distribution of the model parameters to simulate the epidemic process, assuming that the first infection(s) occurred in 2013 on the field F_1. It is worth recalling that the initial infection on $F_2 \cup F_3$ was estimated and therefore also sampled from the joint posterior distribution. The counts of symptomatic and removed plants observed in 2018 and 2019 were compared to those of the simulated epidemics.

Spatial autocorrelation. The spatial structure of the simulated epidemics described above was compared to those of the actual observations. Specifically, Moran's I index and Ripley's L index [47–49] were considered as our spatial summary statistics.

The region was divided into 5×10 square sub-regions, each containing numbers of plants ranging from 1 to 19 (with median value 16), and count of the number of symptomatic sites for the time point used for the computation of Moran's I

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{7}$$

where $y_i = 1$ if plant i is symptomatic before the observation time and $y_i = 0$ otherwise, and $\bar{y}$ is the mean of the observation y_i over n sub-regions. Here, w_{ij} represents a distance-dependent weight between sub-regions i and j. There are many ways to define w_{ij}, and the distance weights between plant i and j was used here. The index was computed by using the ape package [50] available in the statistical software R.

Similarly, the Ripley's L function to characterise clustering/dispersion of point patterns at multiple distances was defined as

$$L(r) = \sqrt{K(r)/\pi} \tag{6}$$

where $K(r)$ is the Ripley K function defined as

$$K(r) = \frac{A}{n(n-1)} \sum_{i,j} \mathbb{1}_{d_{ij} \leq r}.$$

A represents the area of the window, n is the number of data points, and the sum is taken over all ordered pairs of points i and j in the window. d_{ij} is the distance between the two points, and $\mathbb{1}_{d_{ij}}$ is the indicator that equals 1 if the distance is less than or equal to r. This spatial index is computed without requiring any aggregation of the points and therefore gives a more complete description than statistics that utilise aggregation. The index was computed by using the spatstat package available in the statistical software R [51].

Effect of removal efficiency on FD spread

For further insight into how FD spreads, FD epidemics were simulated in a single large square field of 1 km^2 with 167 334 plants. Vines were assumed to be of the same cultivar and planted to mimic the practices in the Bordeaux vineyard, with spacing along rows of 2 m and spacing between rows of 3 m. Disease progress was simulated over 10-year time spans, starting with a single initial infection located at the centre of the field. FD spread only through secondary infections (*i.e.* no primary infection) according to the rate β and to the kernel scale α estimated from the data. Realisations of the epidemic process were generated by sampling sets of parameter (α, β) from the joint posterior distribution of Model 18. Note that, with these assumptions, the maximum distance FD can spread from the initial source is $(1000\sqrt{2})/2 \simeq 707$ m.

Removal due to causes other than FD was ignored ($q = 0$), but the assumption that symptomatic plants will be systematically removed one year after their infection was relaxed by introducing a parameter τ that represents the efficiency of the removal process. More precisely, the parameter τ corresponds to the probability of removing during the autumn/winter of year $t + 1$, that is after the infectiousness period which extends from late spring to late summer, a plant infected during year t. Viewed another way, $1 - \tau$ corresponds to the proportion of infectious plants that escape removal each year, as a result of either flawed identification of symptomatic plants or flawed plant removal. Removed plants are replaced immediately. The algorithm is described in Text C in S1 Text. For each removal efficiency ranging from 0 (no removal) to 100% (full removal) by steps of 20%, 500 realisations of the epidemic process were run. For each run, the observations consisted of the annual snapshots of infected hosts in the 1 km^2 square field considered. Specifically, the distances between the newly infected plants and the initial foci were computed for each year. The median distance, 95% quantile distance and maximum distance between the newly infected plants and the initial source plant were characterized based on the 500 realisations performed. Note that the secondary infection rate β used in the simulation is representative of the average conditions of FD spread in the area and period considered for parameter estimation. In particular, it depends on (and represents) the local population dynamics of *S. titanus* in a context where one to two annual insecticide applications are recommended (but not necessarily done, or done in suboptimal conditions).

Results

Model selection and parameter estimation

The 20 models (Table 2) were fitted to the data using reversible-jump MCMC with data augmentation (Text B in S1 Text and Table A in S2 Text). The MCMC algorithm was iterated for 200 000 iterations, discarding the first 50 000 iterations (the burn-in period) to ensure the convergence to the stationary distribution is reached, with uninformative exponential priors to obtain a joint posterior distribution density for the parameters ϵ, β, α, q and t_0. Trace plots demonstrated that there is no evidence of non-convergence across all models.

We first discuss those parameters whose posterior distributions did not differ greatly with the model fitted. First, the posterior distribution of the starting time of the epidemic revealed that 2016 was the most likely year of introduction of the inoculum into the fields $F_2 \cup F_3$ (Fig A in S2 Text). This introduction probably occurred from field F_1, as infection in this field was first spotted in 2014. Second, the 95% credible interval for the probability q of a plant being removed in any given year due to causes other than FD was consistently estimated to be [0.004, 0.005]. Third, the posterior distribution of ϵ (rate of primary infection from surrounding fields) depended on the assumption made (constant rate or varying with the number of symptomatic plants in the neighbourhood fields, Table 1). Nevertheless, this estimation

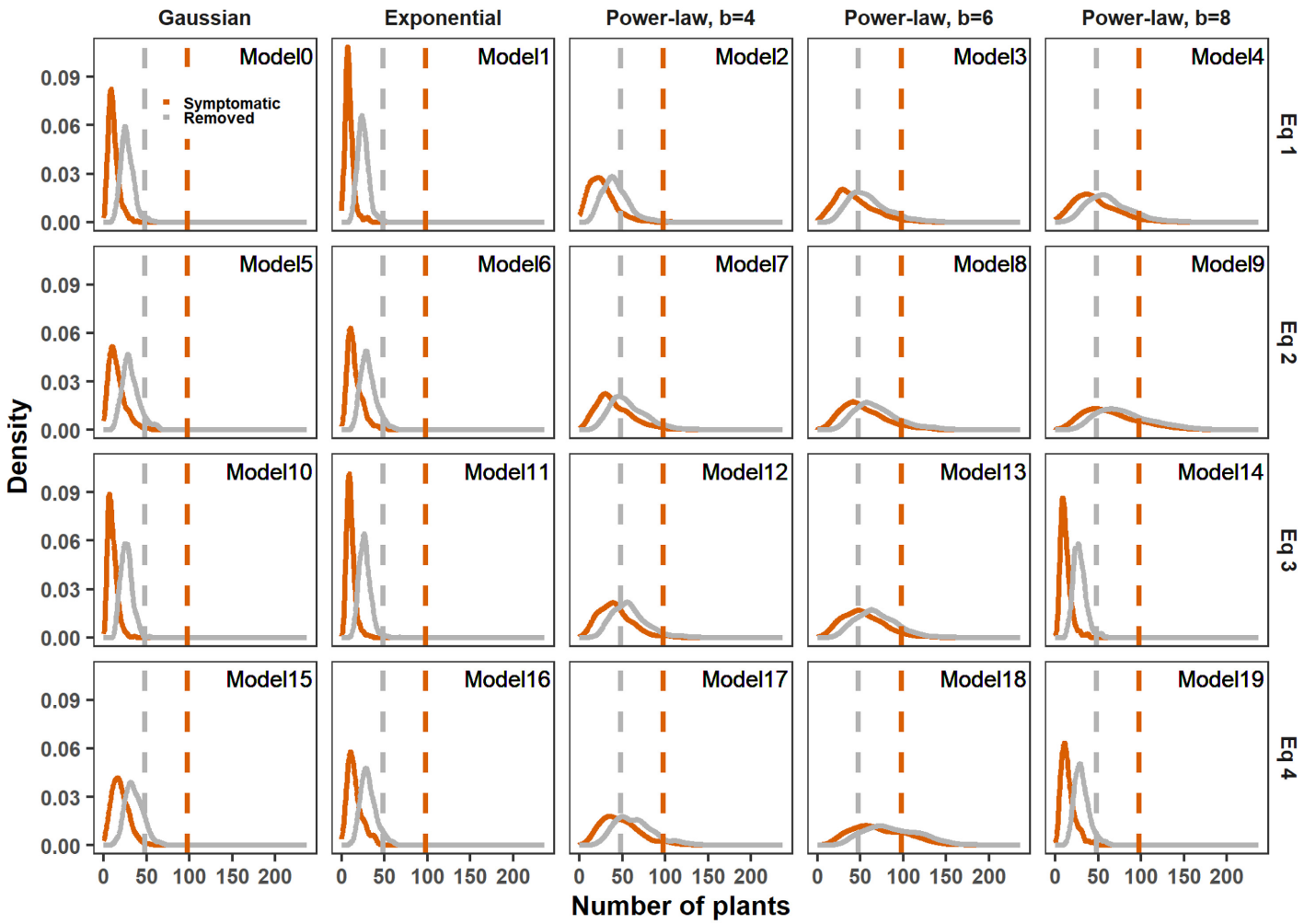

Fig 3. Comparison of the 20 models using the posterior predictive distribution of the counts of observed symptomatic and removed plants in 2018. Each panel corresponds to a model. The dotted line is the actual observed count, and the density of counts is obtained from 1 000 simulations. The colors correspond to the symptomatic (red) and the removed (grey) plants. The 20 models (described in Table 2) differ according to their dispersal kernel (by column) and formulation of infection pressure (by row).

https://doi.org/10.1371/journal.pcbi.1011399.g003

revealed that the number of initial infections arising from primary sources is likely to be 4 for all models.

In contrast, the posterior distribution of other parameters differed significantly with the fitted models. As expected without normalizing the kernel, the infection rate β and the kernel shape α were highly correlated (Table A in S2 Text). Normalizing the kernel by following the approach developed in [52, 53] would have reduced this correlation. The model outputs were used to select the best model(s) by comparing the dynamics and the spatial pattern of simulated epidemics with those of the actual observation. To do so, 1 000 epidemics were simulated for each model using 1 000 samples from their joint posterior distributions.

The first model selection criteria compared the simulated counts of symptomatic and removed plants to observed values in 2018 and 2019 (Fig 3 and S2 Fig). For clarity of these figures and similar ones, the labelling has been chosen so that each line shows the results of different parameterization of the infection pressure and each column the results of different dispersal kernels. The actual counts lay within the distribution of simulated counts only for

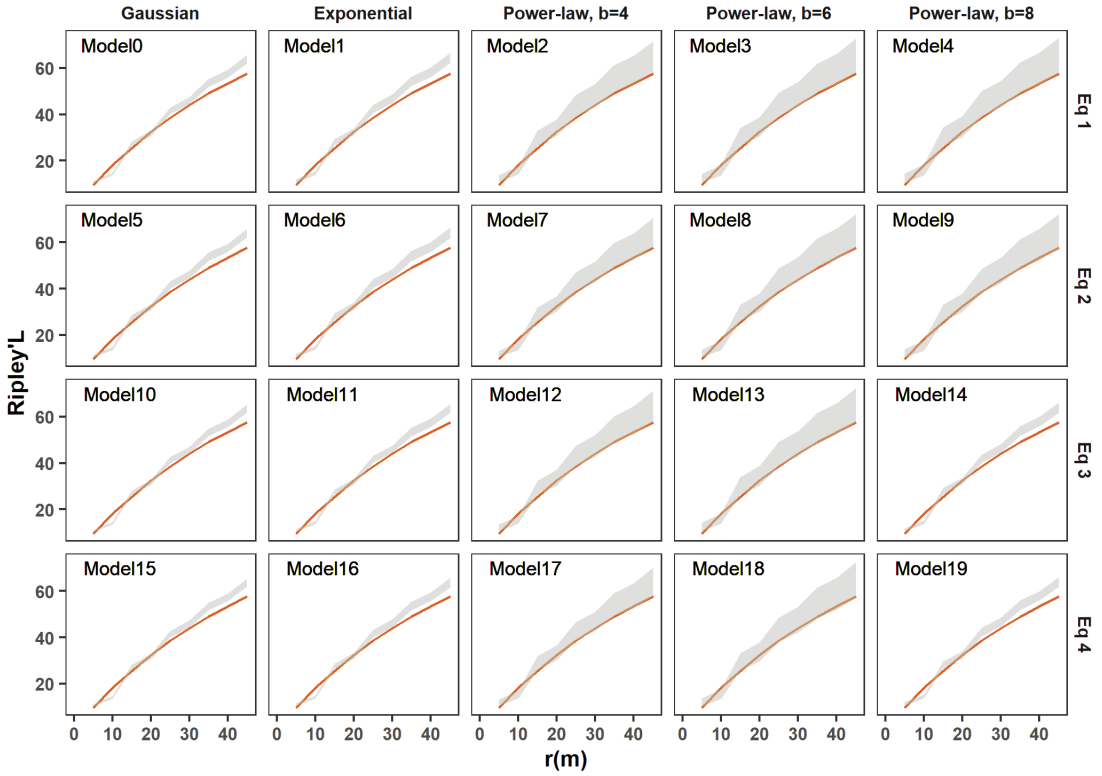

Fig 4. Comparison of the 20 models using the posterior predictive distribution of Ripley's L function. The Ripley's L function determines the clustering/dispersion of point data over a range of distances. In each panel, corresponding to a model, the red line represents the actual measure of Ripley's L function for the observed data, and the grey area indicates the 95% marginal credible region from the simulated epidemic. The 20 models (described in Table 2) differ in their dispersal kernel (column) and formulation of the infection pressure (row).

https://doi.org/10.1371/journal.pcbi.1011399.g004

Model 18. Models relying on Gaussian and exponential thin-tail kernels did not seem to offer an appropriate fit. Observed counts lay in the good range of the posterior count distributions provided by Models 3, 4, 8 and 9, except for the counts of symptomatic plants in 2018.

The second model selection criterion compared the spatial structures of the simulated epidemics to those of the actual observation using the indices of Moran's I (S3 Fig) and Ripley's L (Fig 4). While the actual value of Moran's I index lay in a good range of the posterior distribution for all models, models with thin-tail kernels were completely ruled out when considering Ripley's L index. Power law kernel models fitted reasonably well for this criterion, especially for the exponent parameters $b = 4$ and $b = 6$.

To summarise, four of the 20 models considered (Model 3, Model 4, Model 8 and Model 18) displayed the best fit for both observed counts and observed spatial structures. Of these, Model 18 performed slightly better. Interestingly, similar to Model 18, Models 3 and 8 rely on a power law kernel with $b = 6$. They differed by the hypothesis underlying the infection pressure, with Model 8 utilising a variable primary infection rate (but no cultivar effect) and Model 18 assuming both a variable primary infection rate and a cultivar effect, Cabernet Sauvignon being a better source of infection than the cultivar Merlot. In the following, we will focus on the results obtained from these four models, particularly Model 18.

Dispersal kernels and FD transmission distance

A key epidemiological feature obtained in this study was the dispersal kernel of FD. It represents the rate of disease transmission between a pair of hosts at a given distance. The estimated spatial kernels are shown in Fig 5 for the 4 selected models. The posterior distribution of the kernel means μ showed little variation. This was evidenced by the 95% credible intervals lying within [9, 18] m (Fig 5A). In particular, the posterior mean of μ estimated from the best model (Model 18) was 13 m. It was comparable to the 15 m obtained from Model 3 and Model 8 (that only differ by the formulation of the primary infection rate). Model 4 resulted in a slightly lower estimate of 10 m compared to the three others. Besides the posterior distribution of the kernel mean, the probability of transmission was consistently observed to vanish beyond 20 m for the 4 models considered in a uniformly distributed host landscape (Fig 5B).

Running simulations without removal in the 1 km^2 square field considered enable to locate infected hosts after a single cropping season and thus estimate the posterior predictive distribution of the transmission distance, i.e. the mean distance of hosts infected from a given source in a totally susceptible population (Fig 5C). The posterior median (resp. mean) of the

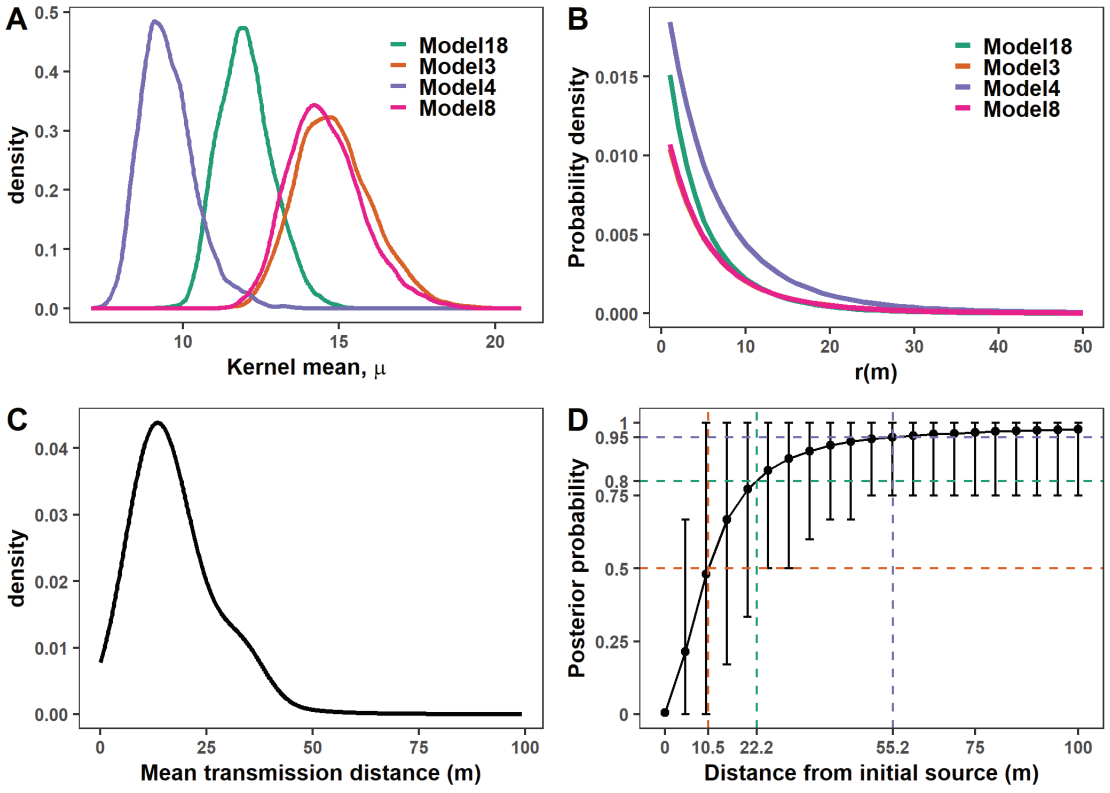

Fig 5. Posterior distribution of the spatial transmission under the four best models. The set of best models is Models 3, 4, 8 and 18, the last being the best one. A: Posterior distribution of the kernel mean, the mean distance to the first infection in a totally susceptible population. B: Probability density of the distance between an infectious plant and the first secondary infection that it causes assuming a uniformly distributed population of susceptible hosts when the scale parameter α is set to its posterior mean. C: Posterior predictive distribution of the mean transmission distance obtained by simulating epidemic realisations using 500 samples from the posterior distribution of parameters estimated from Model 18. D: Corresponding posterior predictive distribution of the cumulative distribution function (CDF) of the distance between the initial source plant and new infections in a single epidemic season (i.e. one year). Error bars represent 95% credible intervals of the CDF. The dotted lines show that on average, 50%, 80% and 95% of the new infections occur within 10.5, 22.2 and 55.2 m of the source, respectively.

https://doi.org/10.1371/journal.pcbi.1011399.g005

transmission distance was 15.01 m (resp. 16.9 m), with the 95% credible interval estimated to be [5.67, 33.59] m. The corresponding posterior predictive distribution of the cumulative distribution function of the distance of the newly infected plants after a single annual epidemic cycle is shown in Fig 5D. On average, 50%, 80% and 95% of the new infection occurred within 10.5, 22.2 and 55.2 m of the source, respectively.

Effect of removal efficiency on FD spread

The simulations also allowed us to estimate how removal efficiency affects the distribution of the distance of the newly infected plants over 10 annual epidemic cycles (Fig 6A, 6B and 6C). The median distance remained in the same range regardless of the level of removal efficiency during the first 5 years (Fig 6A). Then, they substantially increased for low removal efficiencies ($\tau \leq 20\%$) while remaining in the same range for higher removal efficiencies ($\tau \geq 60\%$). Ten seasons after epidemic onset, infected plants were at a median distance of 273 m from the source without removal. This distance was reduced to 212 m for low removal efficiencies ($\tau = 20\%$) while remaining close to 140 m for removal efficiencies $\geq 60\%$. The higher level of removal efficiencies was particularly important for reducing the upper quantiles of the distance between newly infected plants of the initial source (Fig 6B and 6C). The maximum distance without removal is 604 m, meaning that FD infection almost reached the border of the square field of 1 km^2 considered, as the maximum distance FD can spread from the initial source is 707 m. With removal, the maximum distance decreased almost linearly to 443 m with a removal efficiency of 80%, while remaining at a close value with full removal. Importantly, a full removal strategy slowed down but did not stop FD spread. This is mainly due to the presence of a delay between plant infectiousness and optimal symptom detection, followed by removal. For example, ten seasons after epidemic onset, 166 m separated the maximum distance of localisation of newly infected plants between the no removal and full removal strategies. Besides these effects on the distributions of the distance of the new infections, increasing removal efficiencies strongly reduces the number of newly infected plants in each cropping season (Fig 6D), and thus, epidemic size. Ten years after the epidemic onset, this reduction amounted to nearly 85% between the no removal and the full removal strategy.

Discussion

While the estimation of dispersal parameters of plant pathogens using spatio-temporal stochastic mechanistic models has been an active research area for more than 20 years [12], information is still lacking for many diseases, especially for insect-vectored pathogens [36]. In this study, data incorporating two consecutive annual snapshots indicating the symptomatic status of thousands of host plants in three close vineyard plots were used to fit a spatially explicit model representing the epidemiology of FD, a major quarantine disease damaging European vineyards. We assumed that symptom expressions reveal the infected status of plants to FD, based on quantitative real-time PCR tests indicating that more than 98% of symptomatic (resp. more than 97% of asymptomatic) plants noted by the team in charge of the survey were positive (resp. negative) for the phytoplasma. The corresponding high sensitivity and specificity of FD detection are strongly supported for the sampling period in autumn. Indeed, the visual identification of FD symptoms is much easier in autumn [29], as this period is the ideal time for the expression of certain specific symptoms, such as grape shrivelling and the non-lignification of canes [26]. Based on this premise, model inference relies on Markov chain Monte Carlo methods with data augmentation and allows extraction of new insights into the spread of FD. In particular, as far as we know, this study provides the first characterisation of the dispersal kernel of FD. The inference was challenging in our case study because the observed data

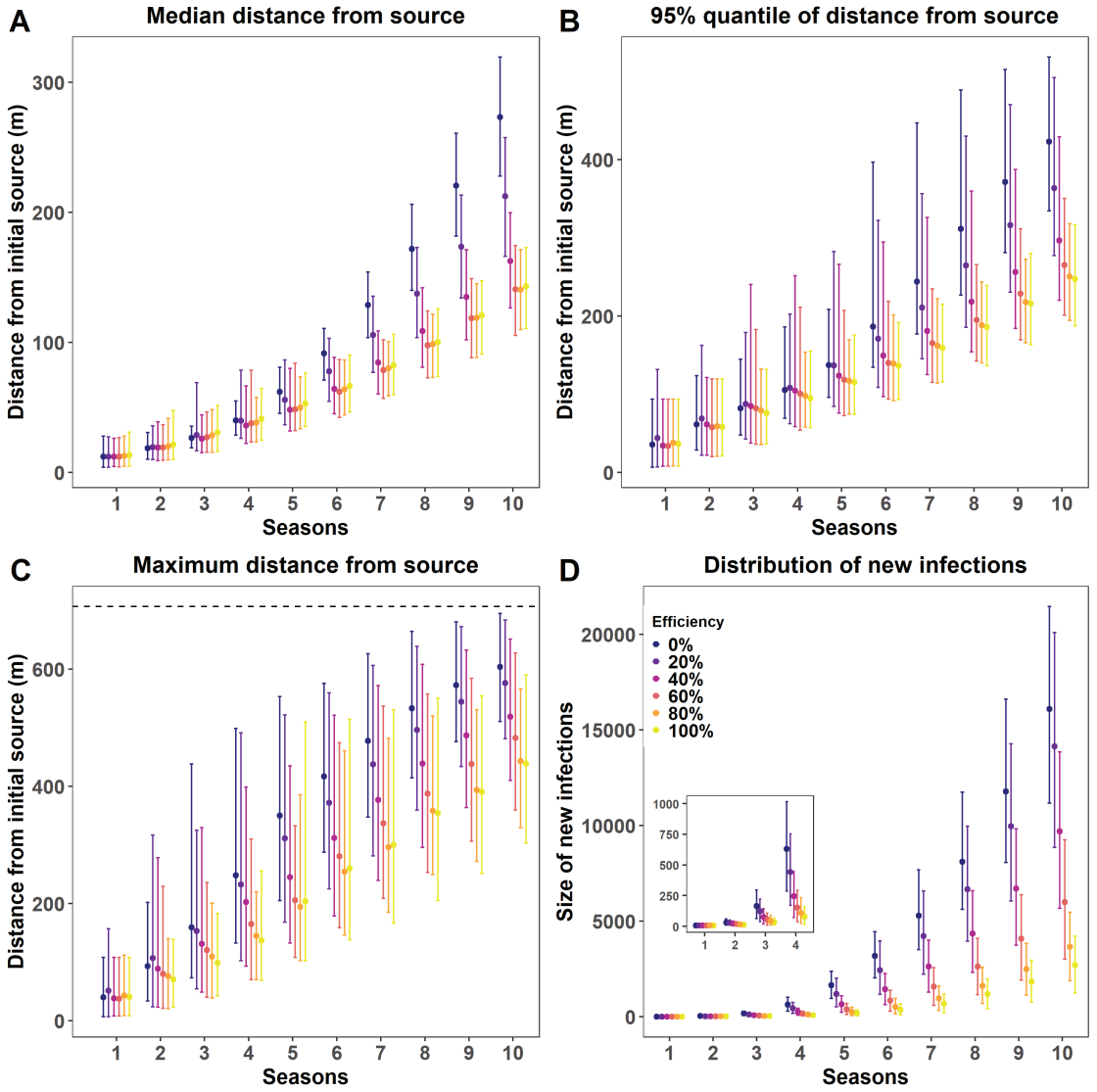

Fig 6. Posterior predictive distribution of the distance between newly infected plants and the initial FD source and posterior predictive distribution of numbers of new infections as a function of the number of seasons after epidemic onset and removal efficiency. Distance and new infections are obtained by simulating epidemics in a square field of 1 km^2 with 167334 vine plants. Epidemics, initiated by a single infected plant at the centre of the field, spread as a result of secondary infections and of the removal efficiency of FD-infectious plants. A: Posterior distribution of the median distance during 10 cropping seasons as a function of removal efficiency (6 levels, from no to full removal). B: Same as A for the 95% quantile of the distance. C: Same as A for the maximum distance. The dashed line indicates the maximum distance FD can spread in the field considered. D: Posterior distributions of the mean number of newly infected plants occurring each year during 10 cropping seasons as a function of 6 levels of removal efficiency.

https://doi.org/10.1371/journal.pcbi.1011399.g006

mainly consist of only two snapshots taken in two consecutive years, a situation similar to the data analysed by [12]. Such situations can be frequent for quarantine disease. Many detailed studies of plant epidemics have benefitted from more frequent spatio-temporal observations (e.g. 4 in [14], 6 in [16] and $\geq$ 15 in [10, 17, 18]). The inference framework, which reconstructs the infection dynamics in the area of interest, notably relies on the presence of missing plants (i.e. vine plants removed) before the year of observation of the first snapshot (2018) in the two

target fields. The data augmentation methodology used here makes it feasible to infer the causes of these observed removals by considering the probability q that a vine plant is removed due to agronomic causes other than FD. Including this parameter allows the spread of FD during the first years after its initial incursion in the area to be slowed in a manner reflecting the situation in the field. Our results strongly suggest that the most likely year of the first infection of fields F_2 and F_3 is 2016 (Fig A in S2 Text).

Dispersal kernels are the key to understanding the spatial transmission of pathogens. They are firstly characterised by their scale parameter. The mean of the posterior distribution of the kernel mean μ obtained from the best model is 13 m (Fig 5A), and the values estimated with the 4 best models range from 10 to 15 m (Fig 5A). In addition, the predictions-based epidemic simulations indicate that the mean (resp. the median) of the posterior distribution of the mean transmission distance is 16.9 m (resp. 15.02 m) (Fig 5C). It also indicates that 50% (resp. 80%) of the new infections are likely to be found within 10.5 m (resp. 22.2 m) of the infectious source plant (Fig 5D). Accordingly, FD mainly spreads locally from one year to the next. Interestingly, these estimations are in broad agreement with the space-time point pattern analysis of an FD epidemic (surveyed from 2011 to 2015) in a field located in Italy [54]. These authors found that the dominant feature was an aggregated cluster of symptomatic plants in moving windows of 8 × 8 plants (equivalent to a square of 11 m/side) to 12 × 12 plants (equivalent to a square of 17 m/side). They also estimated that about 50% of newly symptomatic plants (on year t) were found within 3 m of plants already symptomatic (on year $t - 1$) while discussing how this figure was somewhat blurred by their statistical approaches, which could not necessarily disentangle primary from secondary infections. Our estimations are also in agreement with observations on the fly activity of *S. titanus*, the vector of FD [55]. These authors have shown using mark-recapture techniques that 80% of *S. titanus* adults disperse over distances of less than 30 m along the whole season, a value in line with the aggregated spatial distribution of this monophagous leafhopper previously observed [56–58]. Some authors suggest that the crepuscular flight activity of *S. titanus* adults makes them likely to rely on an active wandering movement rather than on passive wind-borne transport [59]. More broadly, few dispersal kernels are known among insect-vectored pathogens. Huanglongbing, a major disease of citrus worldwide, is mainly transmitted by the Asian citrus psyllid, *Diaphorina citri*. The estimation of its dispersal kernel also indicates a short mean transmission distance of 5 m for 5-year-old trees and 10 m for 18-year-old trees [14]. By contrast, the mean transmission distance of plum pox virus by *Aphis gosypii* amounts to 92.8 m [10], a higher figure possibly linked to the better dispersal capacity of winged aphids that rely on wind. More broadly, it should be noted that the mean dispersal distance of fungal plant pathogens propagated by airborne spores is often much higher (ranging from 213 to 2560 m) in the four pathosystems listed by [36].

Dispersal kernels can be further defined by their shape, which in particular informs the "fatness" of their tails. Tail fatness can be used to categorise kernels in a binary fashion [41]. When at a relatively large distance the shape of the tail decreases less slowly than exponential distribution, or equally slowly, a kernel is termed "short-tailed" or "thin-tailed" [60]. In contrast, if the probability of dispersal decreases more slowly than an exponential distribution at long distances from the source, kernels are termed "long-tailed" or "fat-tailed". Long-distance dispersal events are more frequent than with an exponential kernel that shares the same mean dispersal distance. Here, the annual spread of FD is best described by a fat-tailed power law with exponent parameter $b = 6$, whereas the "thin-tailed" kernels considered (Gaussian and exponential kernels) were clearly discarded by model selection. This result is in line with the maximum distances occasionally observed for *S. titanus* by [14] (330 m) and [58] (600 m). The small-scale dispersion ability of *S. titanus* does not exclude the larger-scale passive dissemination of infectious *S. titanus* over longer distances by the wind, as is the case for many insects [61]. A

consequence of fat tails is the patchiness of the epidemics that can arise. Rather than having a wave that propagates from an initial focus, you get patterns that appear to have multiple foci [12, 62] and in some occasions accelerating epidemic waves [63]. These features are crucial to consider when designing control strategies for emerging plant disease [36].

Given the importance of properly characterising the speed of propagation of the front of the spread of an emerging disease, it would be interesting to consider other fat-tailed dispersal kernels, such as exponential power kernels. Indeed, fat-tailed kernels can be further distinguished depending on whether they are "regularly varying" (e.g. power-law kernels) or "rapidly varying" (e.g. exponential power kernels) [60]. Mathematically, this implies that power law kernels decrease even more slowly than any exponential power function. Biologically, this means that fat-tailed exponential power kernels display rarer long-distance dispersal events than power-law kernels. This affects the dynamics of an epidemic. Indeed, e.g. with weak Allee effects, subtle interactions between tail fatness (rapidly versus regularly varying) and per capita growth rate of the epidemic near zero determine whether the spread is accelerating [64, 65]. However, our observations are mainly collected in three fields within a bounding box of 232 m by 202 m (i.e. over 4.7 *ha*). It is well known that data confined to relatively small spatial scales can blur the precise estimates of the form of dispersal at large distances, and in particular the shape of the kernel's tail [66]. The selected dispersal at long distances depends on both the kernel considered and the distances over which the dataset is collected [67].

The inference framework used considered a non-spatialized primary infection rate term, $\epsilon(t)$, modelling the introduction of FD in the two targeted fields from a set of remote fields. Recent work by [14, 17] also incorporates seasonality into the primary or external rate of infection in modelling plant epidemics, but assumes the rate to be homogeneous. Here, we utilise the data on the disease dynamic in adjacent fields by building the primary infection rate as a function of the density of symptomatic plants in those adjacent fields. This could be further refined by considering the primary infection rate as a spatially decaying source of inoculum as, in our case study, the spatial polygons of the remote fields are known (Fig 2). Thus, a further improvement of the inference framework could consist in using the same dispersal kernel for handling both the data collected at plant scale (in the fields F_1, F_2 and F_3) and the data aggregated at field scale (number of symptomatic plants recorded in remote fields). Regarding this, the method developed in [68] is of interest, as it allows the quantification of the flow of particles over a heterogeneous area by integrating a pointwise dispersal function over source and target polygons. A simpler approach could also be used, e.g. through connectivity matrices between fields. However, the latter approach neglects field geometry and can be expected to bias connectivity estimates, such as when field shapes and sizes are disparate. Furthermore, given the evidence that the FD epidemic spreads more rapidly along than across rows ([54]), anisotropic kernels, in which propagules can disperse differently depending on the direction (e.g. [69]), are worth considering. A different aspect, in light of [10], that could be interesting to take into account in the inference framework is the sensitivity and specificity of FD detection (as detailed in the materials and methods section), given that FD detection in the focal fields was assumed to be errorless. Finally, no records of vector population densities were available for the focal fields. Undoubtedly, incorporating such data, typically throughout a secondary infection rate depending annually on sticky traps data, would have improved model fit.

Despite its simplicity, the spatially explicit SIR model developed and fitted here captures the main features of FD epidemiology in vineyard fields: (i) a single annual infection cycle resulting from the univoltine life cycles of *S. titanus*, (ii) a dormancy phase in the life cycle of the plants and the vector that results in infected plants in year t becoming infectious in spring of year $t + 1$ and (iii) the existence of a delay between plant infectiousness from spring of year $t + 1$ and plant removal the next autumn/winter. This time gap results from practical

considerations for implementing FD management in the field: the beginning of the survey in August as the set of specific symptoms mostly appear throughout the summer, molecular verification of diagnosis, and finally plant removal. The simulations consider a basic scenario, all other factors impacting FD epidemiology being equal: (i) FD is spreading in an isolated and large field (100 ha) after its introduction in a single infected plant, (ii) plants infected during the year t are removed with an efficiency τ infected plants during autumn/winter of year $t + 1$ and (iii) removed plants are replanted immediately. Such removal strategy mimics the current uprooting practice in France, with the exception of the rule that states that fields must be entirely removed as soon as their FD prevalences exceed 20%. This rule would be important for modelling FD control strategies in agricultural landscapes, which is out of the scope of this study. Overall, in the context of a FD survey as currently carried out in the area, simulations suggest that even a fully efficient uprooting strategy ($\tau = 1$) is unable to avoid FD spread (Fig 6D). It illustrates, as already emphasised in several pathosystems, that the control of emerging disease is the most difficult when it is hampered by invisible cryptic infection [11, 16]. In our case study, "cryptic" infections are the consequence of taking into account practical constraints in defining the modalities of FD management in the field. It imposes a time gap between the infectious period and the period of optimal detection of symptoms (and an even larger gap until effective plant removal). The major interest of mandatory insecticide treatments directed against nymphs and adults of the vector is precisely to reduce the risk of FD phytoplasma acquisition/transmission during this time gap.

In a context of a necessary reduction of insecticide treatments, previous results should lead us to reconsider the component of the global FD control strategy which is identifying and removing infected plants. Currently, only symptomatic plants are removed in France. In the presence of "cryptic" infection, strategies relying on reactive host removal are of particular interest. They consist in removing plants, regardless of their infectious status (i.e. both asymptomatic and symptomatic plants), within a particular distance of infection foci. The rationale is to remove (and/or treat) locations that are likely to be infected without yet showing symptoms [11, 37, 38]. In the case of FD, the plants should be removed during the winter of year t, before their infectiousness. Note that, in line with these reasons, insecticide treatments against *S. titanus* are already advised in a buffer of 500 m around infected plots. Another strategy could consist in cutting the vine twigs as soon as the very first symptoms appear on them, all along the summer, as suggested by [33] and as practiced in some vineyards in Italy. It must be followed by a complete removal in winter as imposed by quarantine measures. The model developed in this work is a first step in redesigning FD control strategies. Future research should take advantage of the (S)usceptible-(E)xposed-(C)ryptic-(D)etectable-(I)nfected-(R)emoved framework proposed by [11] to optimise control strategies for invasive plant disease. Research should also benefit from one of the few attempts to model the long-term epidemiology of FD [70]. Using a mean-field model approximation (i.e. ignoring the transmission dispersal kernel), these authors explore the role of hotbeds (defined as vine-growing areas such as abandoned vineyards and woods containing wild grapevines) and of insecticide spray on FD dynamics. In keeping with these and our approaches, future works should consider the full spectrum of existing control measures (e.g. healthier planting material, removal of hotbeds, improving vector control with insecticides but also vibrational disturbances and "push-pull" strategies, see [22] for a review) when we only consider here the uprooting of infected plants. Finally, following [7], this approach should account for economic criteria that balance benefits (generated by the cultivation of healthy/asymptomatic plants) against costs (particularly due to FD control: surveillance, plant removal, replanting and insecticide treatments).

Supporting information

S1 Text. A: Equations used to model the infection pressure. B: Description of the MCMC algorithm. C: Simulation algorithm.
(PDF)

S2 Text. Parameters' estimates for all models.
(PDF)

S1 Fig. Map of the three vineyard fields F_1, F_2 and F_3 in their landscape.
(PDF)

S2 Fig. Comparison of the 20 models using the counts of observed symptomatic and removed plants in 2019.
(PDF)

S3 Fig. Comparison of the 20 models using the spatial correlation index of Moran.
(PDF)

Acknowledgments

The authors wish to thank the people who sampled or provided data: Sophie Bentejac, Morgane Le Goff and Charlotte Labit (GDON des Bordeaux); Dominique Vergnes (FREDON Aquitaine); and Gontran Casella (Syndicat des Bordeaux et Bordeaux Supérieur). We also thank Sophie Bentejac and Charlotte Labit for their critical and careful reading of the manuscript.

Author Contributions

Conceptualization: Hola K. Adrakey, Gavin J. Gibson, Frederic Fabre.

Formal analysis: Hola K. Adrakey, Gavin J. Gibson.

Funding acquisition: Sylvie Malembic-Maher, Frederic Fabre.

Investigation: Hola K. Adrakey, Sandrine Eveillard, Sylvie Malembic-Maher, Frederic Fabre.

Visualization: Hola K. Adrakey, Frederic Fabre.

Writing – original draft: Hola K. Adrakey, Frederic Fabre.

Writing – review & editing: Hola K. Adrakey, Gavin J. Gibson, Sandrine Eveillard, Sylvie Malembic-Maher, Frederic Fabre.

References

1. Seebens H, Blackburn TM, Dyer EE, Genovesi P, Hulme PE, Jeschke JM, et al. No saturation in the accumulation of alien species worldwide. Nat Commun. 2017; 8(1):14435. https://doi.org/10.1038/ncomms14435
2. Parnell S, van den Bosch F, Gottwald T, Gilligan CA. Surveillance to inform control of emerging plant diseases: an epidemiological perspective. Annu Rev Phytopathol. 2017; 55(1):591–610. https://doi.org/10.1146/annurev-phyto-080516-035334 PMID: 28637378
3. Adrakey H, Streftaris G, Cuniffe N, Gottwald T, Gilligan CA, Gibson GJ. Evidence-based controls for epidemics using spatio-temporal stochastic models in a Bayesian framework. J R Soc Interface. 2017; 14(136):20170386. https://doi.org/10.1098/rsif.2017.0386 PMID: 29187634
4. Mastin AJ, Gottwald TR, van den Bosch F, Cunniffe NJ, Parnell S. Optimising risk-based surveillance for early detection of invasive plant pathogens. PLoS Biology. 2020; 18(10):e3000863. https://doi.org/10.1371/journal.pbio.3000863 PMID: 33044954

5. Bussell EH, Cunniffe NJ. Applying optimal control theory to a spatial simulation model of sudden oak death: ongoing surveillance protects tanoak while conserving biodiversity. J R Soc Interface. 2020; 17 (165):20190671. https://doi.org/10.1098/rsif.2019.0671 PMID: 32228402

6. White SM, Bullock JM, Hooftman DAP, Chapman DS. Modelling the spread and control of Xylella fastidiosa in the early stages of invasion in Apulia, Italy. Biol Invasions. 2017; 19(6):1825–1837. https://doi.org/10.1007/s10530-017-1393-5 PMID: 32025190

7. Rimbaud L, Dallot S, Bruchou C, Thoyer S, Jacquot E, Soubeyrand S, et al. Improving management strategies of plant diseases using sequential sensitivity analyses. Phytopathology. 2019; 109(7):1184–1197. https://doi.org/10.1094/PHYTO-06-18-0196-R PMID: 30844325

8. Rimbaud L, Bruchou C, Dallot S, Pleydell DRJ, Jacquot E, Soubeyrand S, et al. Using sensitivity analysis to identify key factors for the propagation of a plant epidemic. R Soc Open Sci. 2018; 5(1):171435. https://doi.org/10.1098/rsos.171435 PMID: 29410846

9. Cook AR, Gibson GJ, Gottwald TR, Gilligan CA. Constructing the effect of alternative intervention strategies on historic epidemics. J R Soc Interface. 2008; 5:1203–1213. https://doi.org/10.1098/rsif.2008.0030 PMID: 18302995

10. Pleydell DRJ, Soubeyrand S, Dallot S, Labonne G, Chadœuf J, Jacquot E, et al. Estimation of the dispersal distances of an aphid-borne virus in a patchy landscape. PLoS Comput Biol. 2018; 14(4): e1006085. https://doi.org/10.1371/journal.pcbi.1006085 PMID: 29708968

11. Cunniffe N, Stutt R, DeSimone RE, Gottwald T, Gilligan C. Optimising and communicating options for the control of invasive plant disease when there is epidemiological uncertainty. New Phytol. 2015; 11 (4):e1004211. https://doi.org/10.1371/journal.pcbi.1004211 PMID: 25874622

12. Gibson GJ, Austin EJ. Fitting and testing spatio-temporal stochastic models with application in plant epidemiology. Plant Pathology. 1996; 45(2):172–184. https://doi.org/10.1046/j.1365-3059.1996.d01-116.x

13. Gibson GJ, Renshaw E. Estimating parameters in stochastic compartemental model using Markov chain methods. IMA J Math Appl Med Biol. 1998; 15:19–40. https://doi.org/10.1093/imammb/15.1.19

14. Parry M, Gibson GJ, Parnell S, Gottwald TR, Irey MS, Gast TC, et al. Bayesian inference for an emerging arboreal epidemic in the presence of control. PNAS. 2014; 111(17):6258–6262. https://doi.org/10.1073/pnas.1310997111 PMID: 24711393

15. Lau M, Gibson G, Adrakey H, McClelland A, Riley S, Zelner J, et al. A mechanistic spatio-temporal framework for modelling individual-to-individual transmission—With an application to the 2014-2015 west Africa Ebola outbreak. PLoS Comput Biol. 2017; 13(10):e1005798. https://doi.org/10.1371/journal.pcbi.1005798 PMID: 29084216

16. Filipe JAN, Cobb RC, Meentemeyer RK, Lee CA, Valachovic YS, Cook AR, et al. Landscape epidemiology and control of pathogens with cryptic and long-distance dispersal: Sudden oak death in northern Californian forests. PLoS Comput Biol. 2012; 8(1):e1002328. https://doi.org/10.1371/journal.pcbi.1002328 PMID: 22241973

17. Neri F, Cook F, Gibson G, Gottwald T, Gilligan C. Bayesian analysis for inference of an emerging epidemic: citrus canker in urban landscapes. PLoS Comput Biol. 2014; 10(4):e1003587. https://doi.org/10.1371/journal.pcbi.1003587 PMID: 24762851

18. Cunniffe N, Laranjeira F, Neri F, DeSimone R, Gilligan C. Cost-effective control of plant disease when epidemiological knowledge is incomplete: modelling Bahia bark scaling of citrus. PLoS Comput Biol. 2014; 10(8):e1003753. https://doi.org/10.1371/journal.pcbi.1003753 PMID: 25102099

19. Schvester D, Carle P, Moutous G. Transmission de la Flavescence dorée de la vigne par Scaphoideus littoralis Ball. Ann Epiphyt. 1963; 14:175–198.

20. Malembic-Maher S, Desqué D, Khalil D, Salar P, Bergey B, Danet JL, et al. When a palearctic bacterium meets a nearctic insect vector: genetic and ecological insights into the emergence of the grapevine Flavescence dorée epidemics in Europe. PLoS Pathogens. 2020; 16(3):e1007967. https://doi.org/10.1371/journal.ppat.1007967 PMID: 32210479

21. Jeger M, Bragard C, Caffier D, Candresse T, Chatzivassiliou E, Dehnen-Schmutz K, et al. Risk to plant health of Flavescence dorée for the EU territory. EFSA Journal. 2016; 14(12):e04603.

22. Chuche J, Thiéry D. Biology and ecology of the Flavescence dorée vector Scaphoideus titanus: a review. Agron Sustainable Dev. 2014; 34(2):381–403. https://doi.org/10.1007/s13593-014-0208-7

23. Bocca FM, Picciau L, Alma A. New insights on Scaphoideus titanus biology and their implications for integrated pest management. Entomol Gen. 2020; 40(4):337–349. https://doi.org/10.1127/entomologia/2020/0977

24. Schvester D, Carle P, Moutous G. Nouvelles données sur la transmission de la Flavescence dorée de la vigne par Scaphoideus littoralis Ball. Ann Zool Ecol Anim. 1969; 1(4):445–465.

25. Morone C, Boveri M, Giosuè S, Gotta P, Rossi V, Scapin I, et al. Epidemiology of Flavescence dorée in vineyards in northwestern Italy. Phytopathology. 2007; 97(11):1422–1427. https://doi.org/10.1094/PHYTO-97-11-1422 PMID: 18943511

26. Tramontini S, Delbianco A, Vos S. Pest survey card on Flavescence dorée phytoplasma and its vector Scaphoideus titanus. EFSA Supporting Publications. 2020; 17(8):1909E.

27. Bressan A, Spiazzi S, Girolami V, Boudon-Padieu E. Acquisition efficiency of Flavescence dorée phytoplasma by Scaphoideus titanus Ball from infected, tolerant or susceptible grapevine cultivars or experimental host plants. Vitis—J Grape Res. 2005; 44(3):143–146.

28. Boudon-Padieu E. Flavescence dorée of the grapevine: knowledge and new developments in epidemiology, etiology and diagnosis. ATTI Giornate Fitopatologiche. 2002; 1:15–34.

29. Adrakey K, Malembic-Maher S, Rusch A, Ay J, Riley L, Ramalanjaona L, et al. Field and landscape risk factors impacting Flavescence dorée infection: insights from spatial Bayesian modelling in the Bordeaux vineyards. Phytopathology. 2022; 112(8):1686–1697. https://doi.org/10.1094/PHYTO-10-21-0449-R PMID: 35230150

30. Roggia C, Caciagli P, Galetto L, Pacifico D, Veratti F, Bosco D, et al. Flavescence dorée phytoplasma titre in field-infected Barbera and Nebbiolo Grapevines. Plant Pathol. 2014; 63(1):31–41. https://doi.org/10.1111/ppa.12068

31. Eveillard S, Jollard C, Labroussaa F, Khalil D, Perrin M, Desqué D, et al. Contrasting susceptibilities to Flavescence dorée in Vitis vinifera, rootstocks and wild vitis species. Front Plant Sci. 2016; 7:1762. https://doi.org/10.3389/fpls.2016.01762 PMID: 27965681

32. Ripamonti M, Pegoraro M, Rossi M, Bodino N, Beal D, Panero L, et al. Prevalence of Flavescence dorée phytoplasma-infected Scaphoideus titanus in different vineyard agroecosystems of northwestern Italy. Insects. 2020; 11(5):301. https://doi.org/10.3390/insects11050301 PMID: 32413986

33. Galetto L, Miliordos D, Roggia C, Rashidi M, Sacco D, Marzachì C, et al. Acquisition capability of the grapevine Flavescence dorée by the leafhopper vector Scaphoideus titanus Ball correlates with phytoplasma titre in the source plant. J Pest Sci. 2014; 87(4):671–679. https://doi.org/10.1007/s10340-014-0593-3

34. Ripamonti M, Maron F, Cornara D, Marzachì C, Fereres A, Bosco D. Leafhopper feeding behaviour on three grapevine cultivars with different susceptibilities to Flavescence dorée. J Insect Physiol. 2022; 137:104366. https://doi.org/10.1016/j.jinsphys.2022.104366 PMID: 35122779

35. Nathan R, Klein E, Robledo-Arnuncio JJ, Revilla E. Dispersal kernels: review. Oxford University Press; 2012.

36. Fabre F, Coville J, Cunniffe NJ. Optimising reactive disease management using spatially explicit models at the landscape scale. In: Scott P, Strange R, Korsten L, Gullino ML, editors. Plant Diseases and Food Security in the 21st Century. Plant Pathology in the 21st Century. Cham: Springer International Publishing; 2021. p. 47–72.

37. Thompson R, Cobb R, Gilligan C, Cunniffe N. Management of invading pathogens should be informed by epidemiology rather than administrative boundaries. Ecol Modell. 2016; 324:28–32. https://doi.org/10.1016/j.ecolmodel.2015.12.014 PMID: 27019546

38. Parnell S, Gottwald T, Gilligan C, Cunniffe N, van den Bosch F. The effect of landscape pattern on the optimal eradication zone of an invading epidemic. Phytopathology. 2010; 100:638–644. https://doi.org/10.1094/PHYTO-100-7-0638 PMID: 20528181

39. Arnaud G, Malembic-Maher S, Salar P, Bonnet P, Maixner M, Marcone C, et al. Multilocus sequence typing confirms the close genetic interrelatedness of three distinct Flavescence dorée phytoplasma strain clusters and group 16SrV phytoplasmas infecting grapevine and alder in Europe. Appl Environ Microbiol. 2007; 73(12):4001–4010. https://doi.org/10.1128/AEM.02323-06 PMID: 17468266

40. Pelletier C, Salar P, Gillet J, Cloquemin G, Very P, Foissac X, et al. Triplex real-time PCR assay for sensitive and simultaneous detection of grapevine phytoplasmas of the 16SrV and 16SrXII-A groups with an endogenous analytical control. Vitis—J Grape Res. 2009; 48(2):87–87.

41. Mollison D. Spatial contact models for ecological and epidemic spread. J R Stat Soc. 1977; 39:283–326.

42. Clark JS, Silman M, Kern R, Macklin E, Hillerislambers J. Seed dispersal near and far: patterns across temperate and tropical forests. Ecology. 1999; 80:1475–1494. https://doi.org/10.1890/0012-9658(1999)080%5B1475:SDNAFP%5D2.0.CO;2

43. MacCalman L, McKendrick IJ, Denwood M, Gibson G, Catterall S, Innocent G, et al. MAPRA: modelling animal pathogens: review and adaptation. EFSA Supporting Publications. 2016; 13(12):1112E.

44. Green PJ. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika. 1995; 82:711–732. https://doi.org/10.1093/biomet/82.4.711

45. Gibson GJ. Markov chain Monte Carlo methods for fitting spatio-temporal sotchastic models in plant epidemiology. J R Stat Soc Ser C. 1997; 46:215–233. https://doi.org/10.1111/1467-9876.00061

46. Lau MSY, Marion G, Streftaris G, Gibson GJ. New model diagnostics for spatio-temporal systems in epidemiology and ecology. J R Soc Interface. 2014; 11(93):20131093. https://doi.org/10.1098/rsif.2013.1093 PMID: 24522782

47. Getis A. Spatial interaction and spatial autocorrelation: a cross-product approach. Environ Plann A: Econ Space. 1977; 23(9):1269–1277. https://doi.org/10.1068/a231269

48. Ripley BD. The second-order analysis of stationary point processes. J App Prob. 1976; 13(02):255–266. https://doi.org/10.2307/3212829

49. Ripley B. Statistical inference for spatial processes. Cambridge university press; 1991.

50. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 2019; 35:526–528. https://doi.org/10.1093/bioinformatics/bty633 PMID: 30016406

51. Baddeley A, Turner Rea. Spatstat: an R package for analyzing spatial point patterns. J Stat software. 2005; 12(6):1–42. https://doi.org/10.18637/jss.v012.i06

52. Lindström T, Håkansson N, Wennergren U. The shape of the spatial kernel and its implications for biological invasions in patchy environments. Proc R Soc Ser B Biol Sci. 2011; 278(1711):1564–1571. https://doi.org/10.1098/rspb.2010.1902 PMID: 21047854

53. Catterall S, Cook AR, Marion G, Butler A, Hulme PE. Accounting for uncertainty in colonisation times: a novel approach to modelling the spatio-temporal dynamics of alien invasions using distribution data. Ecography. 2012; 35(10):901–911. https://doi.org/10.1111/j.1600-0587.2011.07190.x

54. Maggi F, Bosco D, Galetto L, Palmano S, Marzachì C. Space-time point pattern analysis of Flavescence dorée epidemic in a grapevine field: disease progression and recovery. Front Plant Sci. 2017; 7:1987. https://doi.org/10.3389/fpls.2016.01987 PMID: 28111581

55. Lessio F, Tota F, Alma A. Tracking the dispersion of Scaphoideus titanus Ball (Hemiptera: Cicadellidae) from wild to cultivated grapevine: use of a novel mark-capture technique. Bull Entomol Res. 2014; 104 (4):432–443. https://doi.org/10.1017/S0007485314000030 PMID: 24725361

56. Bosco D, Alma A, Arzone A. Studies on population dynamics and spatial distribution of leafhoppers in vineyards (Homoptera: Cicadellidae). Ann Appl Biol. 1997; 130(1):1–11. https://doi.org/10.1111/j.1744-7348.1997.tb05778.x

57. Lessio F, Alma A. Dispersal patterns and chromatic response of Scaphoideus titanus Ball (Homoptera Cicadellidae), vector of the phytoplasma agent of grapevine Flavescence dorée. Agric For Entomol. 2004; 6(2):121–128. https://doi.org/10.1111/j.1461-9563.2004.00212.x

58. Riolo P, Minuz RL, Landi L, Nardi S, Ricci E, Righi M, et al. Population dynamics and dispersal of Scaphoideus titanus from recently recorded infested areas in central-eastern Italy. Bull Insect. 2014; 67 (1):99–107.

59. Lessio F, Alma A. Seasonal and daily movement of Scaphoideus titanus Ball (Homoptera: Cicadellidae). Environ Entomol. 2004; 33(6):1689–1694. https://doi.org/10.1603/0046-225X-33.6.1689

60. Klein EK, Lavigne C, Gouyon PH. Mixing of propagules from discrete sources at long distance: comparing a dispersal tail to an exponential. BMC Ecology. 2006; 6(1):3. https://doi.org/10.1186/1472-6785-6-3 PMID: 16504013

61. Compton SG. Sailing with the wind: dispersal by small flying insects. Dispersal ecology: The 42nd Symposium of the British Ecological Society held at the University of Reading, UK on 2-5 April 2001. 2002; p. 113–133.

62. Shaw MW. Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance. Proc R Soc Ser B Biol Sci. 1995; 259(1356):243–248. https://doi.org/10.1098/rspb.1995.0036

63. Kot M, Lewis MA, van den Driessche P. Dispersal data and the spread of invading organisms. Ecology. 1996; 77(7):2027–2042. https://doi.org/10.2307/2265698

64. Alfaro M, Coville J. Propagation phenomena in monostable integro-differential equations: acceleration or not? J Differ Equations. 2017; 263(9):5727–5758.

65. Bouin E, Coville J, Legendre G. Sharp exponent of acceleration in general nonlocal equations with a weak Allee effect. 2021. arXiv:2105.09911

66. Ferrandino F. Length scale of disease spread: Fact or artifact of experimental geometry. Phytopathology. 1996; 86:685–691.

67. Kuparinen A, Snäll T, Vänskä S, O'Hara RB. The role of model selection in describing stochastic ecological processes. Oikos. 2007; 116(6):966–974. https://doi.org/10.1111/j.0030-1299.2007.15563.x

68. Bouvier A, Kiêu K, Adamczyk K, Monod H. Computation of the integrated flow of particles between polygons. Environ Modell Software. 2009; 24(7):843–849. https://doi.org/10.1016/j.envsoft.2008.11.006

69. Soubeyrand S, Laine AL, Hanski I, Penttinen A. Spatiotemporal structure of host-pathogen interactions in a metapopulation. Am Nat. 2009; 174:308–320. https://doi.org/10.1086/603624 PMID: 19627233

70. Lessio F, Portaluri A, Paparella F, Alma A. A mathematical model of Flavescence dorée epidemiology. Ecol Modell. 2015; 312:41–53. https://doi.org/10.1016/j.ecolmodel.2015.05.014