Phytopathology® • 2022 • 112:1686-1697 • https://doi.org/10.1094/PHYTO-10-21-0449-R

Ecology and Epidemiology

e-Xtra*

Field and Landscape Risk Factors Impacting Flavescence Dorée Infection: Insights from Spatial Bayesian Modeling in the Bordeaux Vineyards

Hola Kwame Adrakey,[1] Sylvie Malembic-Maher,[2] Adrien Rusch,[1] Jean-Sauveur Ay,[3] Luke Riley,[4] Lovasoa Ramalanjaona,[1] and Frederic Fabre[1,†]

[1] INRAE, Bordeaux Sciences Agro, Unité Mixte de Recherche SAVE, Villenave d'Ornon F-33882, France
[2] INRAE, Université de Bordeaux, Unité Mixte de Recherche BFP, Villenave d'Ornon F-33882, France
[3] INRAE, Institut Agro, Université Bourgogne Franche-Comté, Unité Mixte de Recherche CESAER, F-21000, Dijon, France
[4] INRAE, Unité de Recherche BioSP, Equipe OPE, Plateforme d'Epidémiosurveillance en Santé Végétale, Avignon, France
Accepted for publication 22 February 2022.

ABSTRACT

Flavescence dorée (FD) is a quarantine disease threatening European vineyards. Its management is based on mandatory insecticide treatments and the uprooting of infected plants identified during annual surveys. Field surveys are currently not optimized because the drivers affecting FD spread in vineyard landscapes remain poorly understood. We collated a georeferenced dataset of FD detection, collected from 34,581 vineyard plots over 5 years in the South West France wine region. Spatial models fitted with integrated nested Laplace approximation were used to identify local and landscape factors affecting FD detection and infection. Our analysis highlights the importance of sampling period on FD detection and of local practices and landscape context on FD infection. At field scale, altitude and cultivar choice were the main factors affecting FD infection. In particular, the odds ratio of FD infection in fields planted with the susceptible Cabernet Sauvignon, Cabernet Franc, or Muscadelle varieties were approximately twice those in fields planted with the less susceptible Merlot. Field infection was also affected by the field's immediate surroundings (within a circle with a radius of 150 to 200 m), corresponding to landscapes of 7 to 12 ha. In particular, the probability of FD infection increased with the proportions of forest and urban land and with the proportion of susceptible cultivars, demonstrating that the cultivar composition impacts FD epidemiology at landscape scale. The satisfactory predictive performance of the model for identifying districts with a prevalence of FD detection >10% of the fields suggests that it could be used to target areas in which future surveys would be most valuable.

Keywords: distribution modeling, INLA, landscape epidemiology, landscape mosaic, varietal landscape, vineyard disease

An understanding of the contribution of environmental variables to the presence and spread of any pathogen is crucial for the design of efficient surveillance and management disease strategies. Species distribution models are used for this purpose in epidemiology (Bebber 2015; Purse and Golding 2015). These correlative models can be used to disentangle the relative effects of multiple environmental variables (e.g., biotic and abiotic factors) on pathogen epidemiology. They also make it possible to provide predictions for all sites in a region of interest for which the factors studied have been mapped. It therefore provides a basis for the implementation of targeted surveillance and improvements in disease control by maximizing the detection of new cases (Parnell et al. 2014). Widely used for the mapping and management of infectious diseases in humans (Kraemer et al. 2016), distribution modeling has also been used to obtain information about the spatial distribution of emerging plant diseases such as sudden oak disease in the United States (Meentemeyer et al. 2008; Václavík et al. 2010), citrus black spot in South Africa (Martínez-Minaya et al. 2018), and *Xylella fastidiosa* in Europe (Cendoya et al. 2020; Godefroid et al. 2019).

On the methodological side, the first step in any species distribution modeling approach is to gather together diverse sources of information about potential risk factors (e.g., host characteristics, cropping practices, climatic data, and land use) and georeferenced records of the presence of the pathogen (Kraemer et al. 2016; Purse and Golding 2015). The next step involves the use of statistical tools to investigate correlations between these variables (Dormann et al. 2012; Elith and Leathwick 2009). This step is not straightforward because the analysis of spatial data is often complicated by a phenomenon known as spatial autocorrelation (Dormann et al. 2007). Spatial autocorrelation occurs when the values of variables sampled at nearby locations are not independent from each other. In such a case, one of the key assumptions of standard statistical analyses, which is that residuals are independent and identically distributed, is violated. The violation of this assumption may bias parameter estimates and can increase type I error rates (falsely rejecting the null hypothesis of no effect) (Dormann et al. 2007). A variety of frequentist and Bayesian statistical approaches have been developed for modeling of spatial data while accounting for spatial autocorrelation (Beguin et al. 2012; Dormann et al. 2007). Among them, integrated nested Laplace approximation (INLA) (Lindgren et al. 2011; Rue et al. 2009) has emerged in the past decade as a highly appealing alternative combining outstanding computational speed and availability in a user-friendly R interface in the R-INLA library (Beguin et al. 2012; Zuur et al. 2017).

Phytoplasmas are plant-pathogenic bacteria, pleiomorphic with no cell walls, that belong to the class of Mollicutes (Namba 2019). They are obligate parasites invading the phloem sieve tube elements of the host plants and colonizing the body of insect vectors. Phytoplasmas are transmitted by phloem-feeding hemipteran insects (leafhoppers, planthoppers, and psyllids) (Weintraub and Beanland 2006) and by vegetative propagation of infected plant material. They are associated with diseases that cause severe economic impacts on many crops worldwide (Namba 2019). Flavescence dorée (FD), one of the most damaging diseases in European vineyards, is caused by the FD phytoplasma (taxonomic subgroups 16SrV-C and 16SrV-D). FD

†Corresponding author: F. Fabre; frederic.fabre@inrae.fr

Funding: This study received financial support from Plan National Dépérissement du Vignoble under research contracts Co-Act, Co-Act2, and RISCA.

*The *e-Xtra* logo stands for "electronic extra" and indicates there are supplementary materials published online.

The author(s) declare no conflict of interest.

phytoplasma is transmitted from grapevine to grapevine by the leafhopper vector *Scaphoideus titanus* (Chuche and Thiéry 2014). Typical symptoms are leaf yellowing or reddening, with downward rolling, incomplete lignification of canes, abortion of flowers, and grape wilting. FD disease emerged in the South West France wine region in the 1950s, following the accidental introduction of *S. titanus* from North America (Caudwell 1957; Papura et al. 2012). However, FD phytoplasma was demonstrated to be European, originating from wild plant reservoirs: the alder tree *Alnus glutinosa* (Fagales, Betulaceae) and the climbing shrub *Clematis vitalba* (Ranunculales, Ranunculaceae) from which the phytoplasma was originally transmitted to cultivated grapevines (Malembic-Maher et al. 2020).

At the field scale, the epidemiology of FD is first driven by the dynamics of its ampelophagous vector *S. titanus*. This species is univoltine. The eggs hatch in April on grapevine, and there are then five nymphal instars before the first adults appear, usually in June and July. The adults live for approximately 1 month. The fertilized females lay eggs in the late summer, from August to September. Phytoplasmas are acquired passively, from the first larval stage onward, through feeding of the phloem sap of infected grapevine plants where the phytoplasma multiply. When infected, the insects carry and transmit the phytoplasma for the rest of their lives (Weintraub and Beanland 2006). The flight activity of the vector depends on vine density and canopy architecture (Lessio and Alma 2004; Lessio et al. 2015). The epidemiology of FD is also driven by the differential propensity of vine cultivars to act as a source of inoculum (Bressan et al. 2005; Galetto et al. 2014). No cultivars are resistant to FD. They rather display a continuum from low (e.g., Merlot) to high (e.g., Cabernet Sauvignon) susceptibility to the presence and multiplication of phytoplasma (Eveillard et al. 2016; Oliveira et al. 2019; Ripamonti et al. 2021). When the disease has become chronically installed in a region, *Vitis* spp. growing in abandoned vineyards or close to cultivated vineyards can constitute an important source of primary infection (Tramontini et al. 2020). Indeed, they provide a reservoir for the phytoplasma and its vector (Lessio et al. 2014; Ripamonti et al. 2020). Although dispersal abilities of the vector is rather small (<30 m), longer-range dispersals have been observed, suggesting that landscape structure may also affect the epidemiology of FD (Lessio et al. 2014).

In recent decades, FD has spread throughout European vineyards (Jeger et al. 2016). As a result of 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 disease is therefore controlled principally by 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 uprooting of infected plants. Two factors complicate the detection of FD in the field. First, FD cocirculates in European vineyards with bois noir, another phytoplasma disease that has similar symptoms but does less economic damage (Quaglino et al. 2013; Tramontini et al. 2020). Molecular tests involving the detection of pathogen DNA in a real-time multiplex PCR assay are therefore required to confirm the presence of FD (Pelletier et al. 2009). Second, typical FD symptoms appear only during the summer of the year following inoculation (Morone et al. 2007; Schvester et al. 1969; Tramontini et al. 2020). Plants newly infected during the spring and early summer of year n therefore constitute a source of inoculum before the symptoms of the disease become visible late in the summer of year $n + 1$, when the monitoring campaigns are performed. When the diagnosis of FD is confirmed, the infected plants must be removed, and insecticide treatments are mandatory within defined perimeters. More than 70% of French vineyards are treated with insecticides against *S. titanus*, with consequences for the environment and human health (Desneux et al. 2007; Tang et al. 2021). Surveillance has been intensified during the past decade under the supervision of legal authorities, leading to a reduction of the number of annual treatments.

In this study, we applied a distribution modeling approach to improve our understanding of FD epidemiology and, more specifically, gain a quantitative view of the field and landscape factors impacting the probabilities of FD detection and infection at field scale. Based on the literature, we first hypothesized that the probability of FD detection improves as fall approaches. We also hypothesized that key local factors would impact the probability of FD infection, with a higher probability of infection associated with older and denser plantation as well as cultivation of more susceptible cultivars. We also hypothesized that the probability of infection would increase with the proportion of susceptible cultivars, the proportion of seminatural habitats, or the proportion of urban areas in the surrounding landscape because these habitats may act as FD reservoirs. In addition, we hypothesized that higher level of fragmentation of these potential reservoirs at the landscape scale would limit vector spillover and therefore be associated with lower probability of infection (Tscharntke and Brandl 2004). To test these hypotheses, we fitted distribution models using INLA to a spatial dataset collected from 34,581 vineyard fields over 5 years of mandatory vineyard surveys. Finally, we investigated whether the model could identify sites with higher probability of infection for targeted surveillance to improve current management strategies.

MATERIALS AND METHODS

FD monitoring and detection. In the vineyards of Bordeaux in South West France, FD surveys are performed by professional organizations known as Groupements de Défense contre les Organismes Nuisibles (GDONs) under the supervision of the French Ministry of Agriculture. GDON des Bordeaux has been in charge of FD monitoring since 2011 in an area of 364,718 ha, of which 83,912 ha (23%) were cultivated with grapevine in 2016 according to the Casier Viticole Informatisé (CVI). CVI is a geographic information system database created by the French directorate general of customs that provides a comprehensive history of each legal piece of land cultivated with grapevine. It gathers information on the spatial boundary, size, year and density of plantation, and cultivars used in each legal field. The region monitored by GDON is subdivided into 347 districts with a mean area of 1,048 ha and 10 and 90% quantiles of 331 ha and 1,849 ha, respectively (Fig. 1A). Although these districts serve as a basis for the practical organization of FD monitoring performed by GDON in the vineyards, the basic observational unit is the field. In the following, we considered the annual survey realized from 2012 to 2016. Each year, the monitoring strategy of GDON was nearly identical. It was conducted by a team of trained inspectors between August and October, the period of the year most favorable for the detection of FD symptoms (Tramontini et al. 2020). The fields selected each year n by GDON can be separated into two categories. The first category consists of the fields already surveyed and detected infected the previous year $n - 1$. When an infected field has been detected, GDON inspectors check the following year that mandatory control measures were applied. These reinspection data were not considered in our study. The second category consists of the fields not surveyed in the previous years. Approximately 10% of the vineyard area was newly inspected each year. These fields were considered in our study. The survey involved the inspection of most of the fields in large parts of newly selected districts each year, most often regardless of prior information on the presence of FD (note that "spontaneous" reporting of FD by the winegrowers is unusual). Overall, a mean of 6,916 fields are newly inspected each year, with numbers ranging from 5,916 in 2013 to 8,712 in 2012. The fields had a mean area of 0.77 ha, with 10 and 90% quantiles of 0.16 ha and 1.6 ha, respectively.

When the inspection teams survey a given area, they inspect a high proportion of fields. We quantify this sampling effort by the proportion of the vineyard sampled by GDON inspectors. For each of the 34,581 fields sampled, the sampling effort was assessed by the ratio between (i) the total vineyard area sampled by GDON inspectors in a radius of 1 km around each field (as assessed from

the area of the polygons drawn by inspectors during their survey; see below) and (ii) the total size of the vineyard in this same radius (as given by the area of legal piece of land cultivated with grapevine provided by the CVI). The mean sampling effort was 0.58, with 10 and 90% quantiles of 0.38 and 0.74, respectively. As previously stated, the basic observational unit was the field (Fig. 1B). For each field surveyed, inspectors walk through the whole field for visual detection of the symptoms and draw its polygon on geographic information system software. If no symptoms are detected, the inspectors do not sample any plants and the field is declared uninfected. If symptoms are detected, symptomatic plants are marked for uprooting, and symptomatic leaves from one to five plants per field are collected and pooled into a single sample. The detection of the FD and bois noir phytoplasmas in samples is performed by accredited laboratories with a real-time PCR triplex molecular test derived from that described by Pelletier et al. (2009). Each field is then classified as infected with FD (if the molecular test is positive for FD phytoplasma, meaning that at least one plant in the sample is infected with FD) or uninfected with FD. Overall, FD was detected in 7.6% of the 34,581 fields surveyed, and bois noir was detected in 4.5%.

Explanatory variables at field scale. In addition to the georeferenced records of FD detection, we collected a set of potentially important local explanatory variables characterizing the production situations of each field (Table 1). First, the altitude was obtained from topography data provided by Institut national de l'information géographique et forestière. Second, the Appellation d'Origine Contrôlée (AOC), a French certification defining the geographic area of production for a particular wine label and common and specific guidelines (e.g., proportions of particular cultivars, growing practices, winemaking practices) was noted to provide information about the socioeconomic context of production. The AOCs were grouped into six levels on the basis of geographic proximity. Third, the CVI was used to determine, for each field, the cultivars grown (seven levels; Fig. 1B), plant density, and the age of the plantation. The protocol for extracting these variables from the CVI is described in the Supplementary Materials. Fourth, the type of viticultural practice (organic or conventional) in each field was provided by the GDON. This information was available only for 2016, but we assumed that growers maintained the same practices over the entire time period considered here.

Explanatory variables at landscape scale. We created annual land-cover maps of the region monitored by GDON des Bordeaux. These maps are raster maps with a 10-m resolution. The maps were created by overlaying the CVI database and the France land cover map for 2017 (Occupation des Sols map; Inglada et al. 2017). We used the Occupation des Sols map for 2017 because (i) the landscape was expected to be constant within the time frame of the study and (ii) the map resolution was more precise in 2017 (pixels of 10 m per side) than previous years (pixels of 20 m per side). For each year from 2012 to 2016, the CVI database was used to map the pixels representing grapevine cultivars. Seven classes depicting the cultivar used in each pixel during the year considered were distinguished (as listed in Table 1), and an independent eighth class indicates whether the pixel belongs to an organic or conventional vineyard farm (data provided by GDON; see previous section). Each pixel of the GDON region not attributed to a vineyard class was classified into one of three separate classes (forest, urban, or other land use) depending on its classification in the Occupation des Sols map (Fig. 1B). The resulting land cover map for 2016 is illustrated in Supplementary Figure S1.

These annual land cover maps were used to characterize the landscape surrounding each field. We calculated composition and configuration metrics in circles of 15 increasing radii (50, 100, 150, 200, 250, 300, 500, 1,000, 1,500, 2,000, 2,500, 3,000, 4,000, 5,000, and 6,000 m) centered on the barycenter of each field i ($i = 1, \ldots,$ 34,581; Table 1). Thereafter, "scale" defined the radii around field barycenters (i.e., the extent of the landscapes). For each field i and

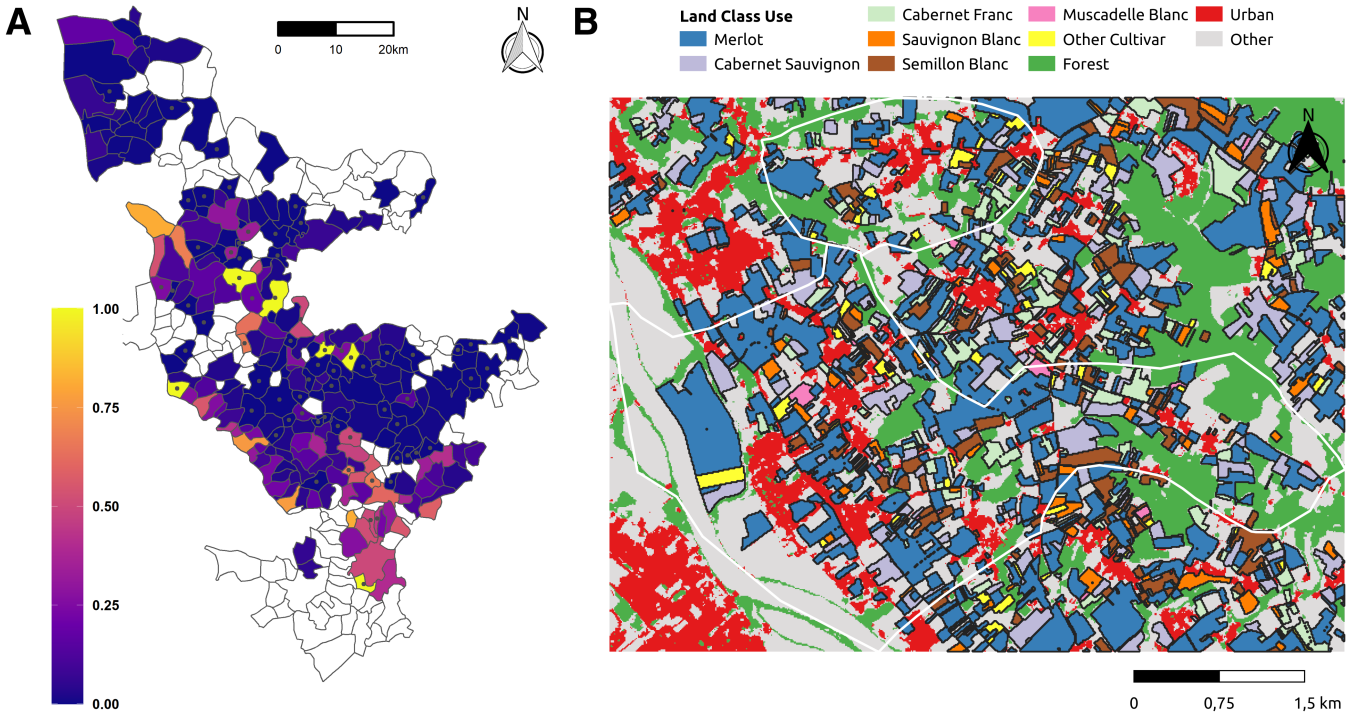

Fig. 1. Field-based flavescence dorée (FD) data at the regional and district scales. **A,** Map of the study region in South West France. The map shows the borders of the 347 districts located within the Groupement de Défense contre les Organismes Nuisibles des Bordeaux area. The color scale indicates the prevalence of FD detection per district in each of the 239 districts in which FD surveys were performed from 2012 to 2016. Unsurveyed districts are shown in white, and districts with fewer than six fields are shown with a dot. **B,** Map at landscape scale (over two entire districts depicted by white border) illustrating the vineyard plots classified according to their cultivar (seven levels) and the other categories of land use (three levels).

scales considered, we calculated the percentage of land occupied by vineyard as L_vineyard (whatever the cultivar grown), forest as L_forest, and urban areas as L_urban. These metrics are assessed as the proportion of pixels attributed to each class within the circle. We also calculated the total number of patches as L_np of these three classes, a simple measure of landscape fragmentation. A patch refers here to a continuous area of the landscape attributed to a single class use. Finally, we calculated the percentages of land (i) cultivated with the cultivar Merlot as L_merlot and (ii) organically managed as L_organic within the total area devoted to grapevine cultivation in the circle considered. These metrics were computed with the R package landscapemetrics (Hesselbarth et al. 2019).

Statistical analysis. *Bayesian inference with INLA.* INLA is a computationally efficient method for fitting complex spatial models within the Bayesian paradigm (Lindgren et al. 2011; Rue et al. 2009). It offers a faster alternative to Markov chain Monte Carlo methods because, in essence, INLA replaces stochastic sampling with deterministic approximation based on a clever use of the Laplace approximation and on numerical integration (Beguin et al. 2012). INLA may be used to fit a large class of latent Gaussian models in a Bayesian framework. Moreover, the method is implemented in the R package R-INLA (https://www.r-inla.org/) (Lindgren and Rue 2015), which allows fitting of models almost as easily as the base R functions for generalized linear models.

Latent Gaussian models include, in particular, several classes of models accounting for spatial autocorrelations. Indeed, combined with the stochastic partial differential equation approach, one can accommodate with INLA all kinds of geographically referenced data, including areal and geostatistical ones, as well as spatial point process data (Lindgren and Rue 2015; Lindgren et al. 2011). Here, we employ INLA to fit spatial logistic regression in a continuous spatial domain. References such as those of Beguin et al. (2012) and Zuur et al. (2017) provide comprehensive text on the subject for applied ecologists and phytopathologists.

Spatial logistic regression. Let y_i be the FD detection status of the field i at location s_i on its year of first inspection $year_i$, i.e., the field i is classified as infected (value 1) or uninfected (value 0). We assumed that y_i follows a Bernoulli random variable and modeled the probability π_i of FD detection in field i using a logit link function. The generic model can be written as follows:

$$\begin{cases} y_i \sim Binomial(1, \pi_i), \quad i = 1, \ldots, 34,581 \\ logit(\pi_i) = \beta X_i + a(year_i) + u(s_i) \\ u(s) \sim GMRF(0, \mathbf{Q}^{-1}(\phi, \sigma_u^2)) \\ a(year_i) \sim Normal(0, \sigma_{year}^2) \end{cases} \quad (1)$$

In this model, X_i is the vector of explanatory variables for field i, and β is the vector of associated coefficients to be estimated. All the explanatory variables X_i listed in Table 1 were treated as fixed effects except the year of first field inspection, which was treated as independent random effects $a(year_i)$. The term $u(s_i)$ is the spatial random effect for field i. It endows the model with a spatial dependence between neighboring fields not explained by the explanatory variables. In R-INLA, a computationally convenient Gaussian Markov random field representation is used to accurately approximate a Gaussian random field with spatial variance and autocorrelation characterized by the Matérn covariance function (Lindgren et al. 2011). This approximation relies on basis functions anchored at a set of discrete points corresponding the nodes of a mesh dividing the study area into a large number of nonoverlapping triangles (Supplementary Fig. S2). Its use allows inference about the latent spatial field over the entire continuous domain of interest, where the spatial interpolation between the Gaussian variables located at the nodes is linear.

The Gaussian Markov random field can be regarded as a multivariate Gaussian distribution (with dimension equal to the number of nodes) with zero mean and sparse precision matrix $\mathbf{Q}$ that depends on two positive hyperparameters, σ_u^2, ϕ, describing the variance and the correlation range (the distance beyond which pairs of observations are approximately not spatially correlated any more), respectively (Krainski et al. 2018; Lindgren et al. 2011). The spatial dependence is encoded in the matrix $\mathbf{Q}$ of size N^*N by using the Matérn covariance function based on the first-order Bessel function K_1. Accordingly, the covariance between $u(s_i)$ and $u(s_j)$ depends on the two unknown hyperparameters and on

TABLE 1. Local and landscape explanatory variables used to characterize the vineyard fields

Variables	Description	Mean (SD)/levels[a]	Source
Local variable			
Area	Field area	0.77 ha (0.77)	GDON
Year	Year of the first field inspection	Five levels	
Season	Summer or autumn inspection	Two levels[b]	
Week	Week number within year	Eleven levels[b]	
Practice	Organic or conventional practices	Two levels[b]	
AOC	Appellation d'origine contrôlée	Six levels[c]	INAO
Age	Age of the plantation at inspection	23 years (14)	CVI
Cultivar	Grape cultivar	Seven levels[d]	
Density	Density of the plantation	4,131 plants/ha (1,008)	
Altitude	Altitude	59.5 m (29)	IGN
Landscape scale			
L_urban	Percentage of urban area	8.2 (10)	CESBIO
L_forest	Percentage of forest area	5.6 (9.8)	
L_vineyard	Percentage of vineyard area	66.7 (20.2)	CVI
L_merlot	Percentage of merlot area in the vineyard	62.2 (32)	
L_organic	Percentage of organic practice area in the vineyard	8 (21.3)	GDON
L_np	Number of vineyard, forest, and urban patches	12.5 (5.5)	

[a] Values for the landscape covariates are given for a scale of 150 m. AOC, Appellation d'Origine Contrôlée; CESBIO, Centre d'Etudes Spatiales de la Biosphère; CVI, Casier Viticole Informatisé; GDON, Groupement de Défense contre les Organismes Nuisibles; IGN, Institut national de l'information géographique et forestière; INAO, Institut national de l'origine et de la qualité; SD, standard deviation.

[b] Season: Summer here corresponds to July/August (28,719 fields), and autumn corresponds to September/October (5,862 fields). Week: 11 levels from week 32 to week 42. Practice: organic (2,745 fields) or conventional (31,836 fields) farming practices.

[c] The 11 controlled designation of origin located in the GDON des Bordeaux area were grouped by spatial proximity in six levels to have ≥700 fields in each level. These levels are AOC1 (Côtes de Bordeaux and Saint-Macaire, n = 1,936), AOC2 (Premières Côtes de Bordeaux, n = 5,032 fields), AOC3 (Loupiac, Sainte-Croix-du-Mont, and Graves supérieures, n = 725 fields), AOC4 (Fronsac and Cannon Fronsac, n = 743 fields), AOC5 (Côtes de Bordeaux Blaye, n = 5,858 fields), and AOC6 (Bordeaux, n = 20,287 fields).

[d] Cultivars include Merlot (21,181 fields), Cabernet Sauvignon (5,266 fields), Cabernet Franc (2,491 fields), Muscadelle (270 fields), Sauvignon (1,985 fields), Semillon (2,587 fields), and other cultivars (801 fields) not previously listed.

the Euclidean distance $d(i, j)$ between the two fields through the following:

$$cov_{Matern}\left[u(s_i), u(s_j)\right] = \sigma_u{}^2\sqrt{8}[d(i,j)/\phi]K_1[\sqrt{8}(d(i,j)/\phi)] \quad (2)$$

We remark that different parametrizations exist for the Matérn covariance, and the preceding one offers the advantage of relatively simple intuitive interpretation. The corresponding correlation function is obtained by replacing $\sigma_u{}^2$ with 1 in cov_{Matern}. In our model (equation 1), we refer to the sparse precision matrix resulting from this covariance model as $Q(\phi,\sigma_u{}^2)$.

Bayesian inference requires to specify prior distributions for the model parameters and hyperparameters (i.e., parameters of prior distributions). These choices are less crucial with large datasets because the effect of prior choice on the posterior estimates is expected to wane as the sample size increases. For the parameters β_j and the hyperparameter σ^2_{year}, we used the default internal vague priors recommended in R-INLA, normal(0, *precision* = 10^{-3}) and loggamma(1,0.00005), respectively. We used penalized complexity priors for the range and the variance of the spatial random effect (Fuglstad et al. 2019; Simpson et al. 2017). Penalized complexity priors can be defined via intuitive probability statements. Because few fields are <0.25 km apart, we assume that the probability that ϕ is <0.25 km is 0.01, leading to a penalized complexity prior (0.25,0.01) for ϕ. Moreover, we assume that the probability that σ_u is >1 is 0.01, leading to a penalized complexity prior (1,0.01) for σ_u.

Model comparison. The baseline spatial model M_0 corresponds to equation 1 with 13 explanatory variables, including seven local variables (area, age, practice, density, altitude, cultivar, and AOC) characterizing the field and six landscape variables (L_organic, L_vineyard, L_np, L_forest, L_merlot, and L_urban) characterizing the landscape within a zone of radius r. No interaction between explanatory variables were considered, and all continuous explanatory variables were standardized. In all, 16 models M_0 corresponding to landscape scales r ranging from 0 to 6,000 m were considered. The case $r = 0$ corresponds to a case in which only local variables are considered. These 16 spatial models were compared with their 16 nonspatial counterparts for which the spatial random effect $u(s_i)$ was removed. We estimated the Watanabe-Akaike information criterion (Gelman et al. 2014; Watanabe 2010), which penalizes model complexity, and the deviance information criterion (Spiegelhalter et al. 2002) to compare all the models fitted.

In addition to the baseline model M_0, we considered a model M_1 to refine the study of the period of field inspection. In model M_1, the variable "season" is eliminated from the set of fixed effects X_i and replaced by a random walk of order 1 indexed by the weeks of inspection. The weeks range from week 32 (second week of August) to week 42 (third week of October). The notation k_i denotes thereafter the week k of inspection of field i (k = 32, … , 42). In model M_1, the logit equation in 1 is replaced by $\text{logit}(\pi_i) = \beta X_i + a(year_i) + u(s_i) + w(k_i)$ with $w_k = w_{k-1} + v_k$ (k = 33, … , 42) and $v_k \sim$ normal(0, $\sigma_k{}^2$). In this model, $\sigma_k{}^2$ is an additional hyperparameter for which we used the default internal INLA prior penalized complexity prior (0.5,0.01).

Preliminary analysis. A set of preliminary analysis was realized with the baseline model M_0 (not shown). First, a test for multicollinearity was performed by fitting the model M_0 without spatial random effect for the 16 landscape scales considered. The variance inflation factors for each variable were less than three (except for "practice" and L_organic at the landscape scale of 50-m radius with values <5), indicating an absence of collinearity between the explanatory variables (Zuur et al. 2009). Second, we compared the model M_0 with a binomial distribution to an alternative model relying on a zero-inflated binomial distribution to account for a possible excess of zeros (Martin et al. 2005), as 92.4% of the fields were FD-negative. We found that the binomial models outperformed the zero-inflated binomial models, suggesting the effectiveness of the surveillance. Finally, we investigated the effect of six meshes on parameter estimations. The meshes differ in the largest triangle edge length (parameter max.edge) and the minimum allowed distance between points (parameter cutoff; Supplementary Fig. S2). Overall, the parameter estimates were consistent from the coarsest to the finest meshes (Supplementary Fig. S3).

Predictive performance of the models. The predictive performance of the baseline model M_0 was estimated at two scales (field, district) by cross-validation. The model was trained on 80% of the data (training set) and tested on the remaining 20% (testing set). We repeated this partitioning for cross-validation 50 times. We first assessed the ability of the model to predict the FD detection status of individual fields with precision-recall (PR) curve. PR curves were preferred over receiver operating characteristic curve because they are more informative when evaluating binary classifiers on imbalanced datasets, which is the case in our study, with 7.6% of FD detection (Saito and Rehmsmeier 2015). For a given threshold (between 0 and 1) above which a field is classified as infected by the model, the recall, also known as sensitivity, is the probability that a field is classified as infected by the model when FD was detected in that field. Similarly, the precision, also known as positive predictive value, is the probability that FD is detected in a field when it is classified as infected by the model. A PR curve is then a plot of the precision (y axis) and the recall (x axis) for all possible thresholds between 0 and 1. Predictive performance was summarized by the area under the PR curve. We also evaluated predictive performance at the district scale. Using the same 80/20% rule applied at the district scale, we evaluated the ability of the model to predict if the proportion (thereafter termed prevalence) of FD detection in a district exceeded a given threshold. Thresholds were varied from 3 to 15%, and, for each threshold used, we calculated precision and recall for each partitioning.

RESULTS

Model selection, goodness of fit, and spatial field. The baseline model M_0 was fitted to 16 landscape scales ranging from 0 (only local variables considered) to 6,000 m. As a first step, we compared the 16 spatial models M_0 with their 16 nonspatial counterparts for which the spatial random effect $u(s_i)$ was removed. The Watanabe-Akaike information criterion and deviance information criterion values of the spatial model were much lower at all scales (by ≥400 points), highlighting the importance of the spatial structure underlying FD detection in our data. The spatial model fitted with landscape variables at the 150-m scale was the best of the models fitted for both metrics considered (Fig. 2A). Similar results were obtained with model M_1. The second-best landscape scale was 200 m for both models, parameter estimates being consistent between M_0 and M_1 at these two scales (Supplementary Fig. S5). Thereafter, we then focused on the results obtained with a landscape scale of 150 m, corresponding to a landscape area of 7.06 ha containing six fields on average (90% confidence interval, 2 to 12 fields). Note also that, at this scale, the continuous explanatory variables were weakly correlated (Supplementary Fig. S4).

The goodness of fit of the spatial model for a 150-m landscape scale was highly satisfactory. In particular, the coefficient of determination for the linear regression between the observed and adjusted prevalences of FD detection at the district scale was 0.99. Furthermore, this regression line is very close to the $y = x$ line, suggesting that the model explains most of the variability of the data at district scale (Fig. 2B).

Compared with the nonspatial model, the spatial model strongly reduces spatial autocorrelation, as evidenced by comparing the spatial variograms of the residuals of both models (Fig. 2C). However, a weak spatial autocorrelation remains for distances up to 3.6 km, the estimated range (95% confidence interval, 3.1 to 4.2). This is the distance beyond which the correlation between the FD detection

status of two fields becomes negligible. The shape of the Matérn function is itself instructive (Fig. 2D). In particular, the coefficient of correlation between the detection statuses of fields located ≤1 km apart is >0.5.

The distribution of spatial random effects provided evidence of a strong spatial pattern (Fig. 3A). Positive effects are clustered in the southern and western regions of the GDON des Bordeaux area, whereas negative effects are clustered in the northwestern and the eastern regions. On the logit scale, the spatial random effect ranges from −1.33 (5% quantile) to 1.44 (95% quantile). These values correspond to odds ratios of 0.25 and 4.23, respectively. Moreover, the variation on the standard deviation is attributable to the clustering of FD detection cases over the study area (Fig. 3B).

Intra- and interannual effects on FD detection. Inferences indicate that the probability for the inspectors to detect FD during August is lower (odds ratio, 0.24) than in September and October (Fig. 4A, last panel). The random walk used in model M_1 allows refinement of the estimation of the effect of the weeks of field inspections and highlights a sharp increase from the first week of September (Fig. 5A). Moreover, the probability of FD detection was remarkably consistent for each of the 5 years surveyed except 2014, when it was somewhat significantly lower (Fig. 4A, last panel).

FD risk factors at the field scale. Local variables describe the physical and agronomic characteristics of each field in the vineyard (Table 1). Their effects (posterior mean, 95% credible intervals, and probability of being positive) are displayed in Figure 4A. The

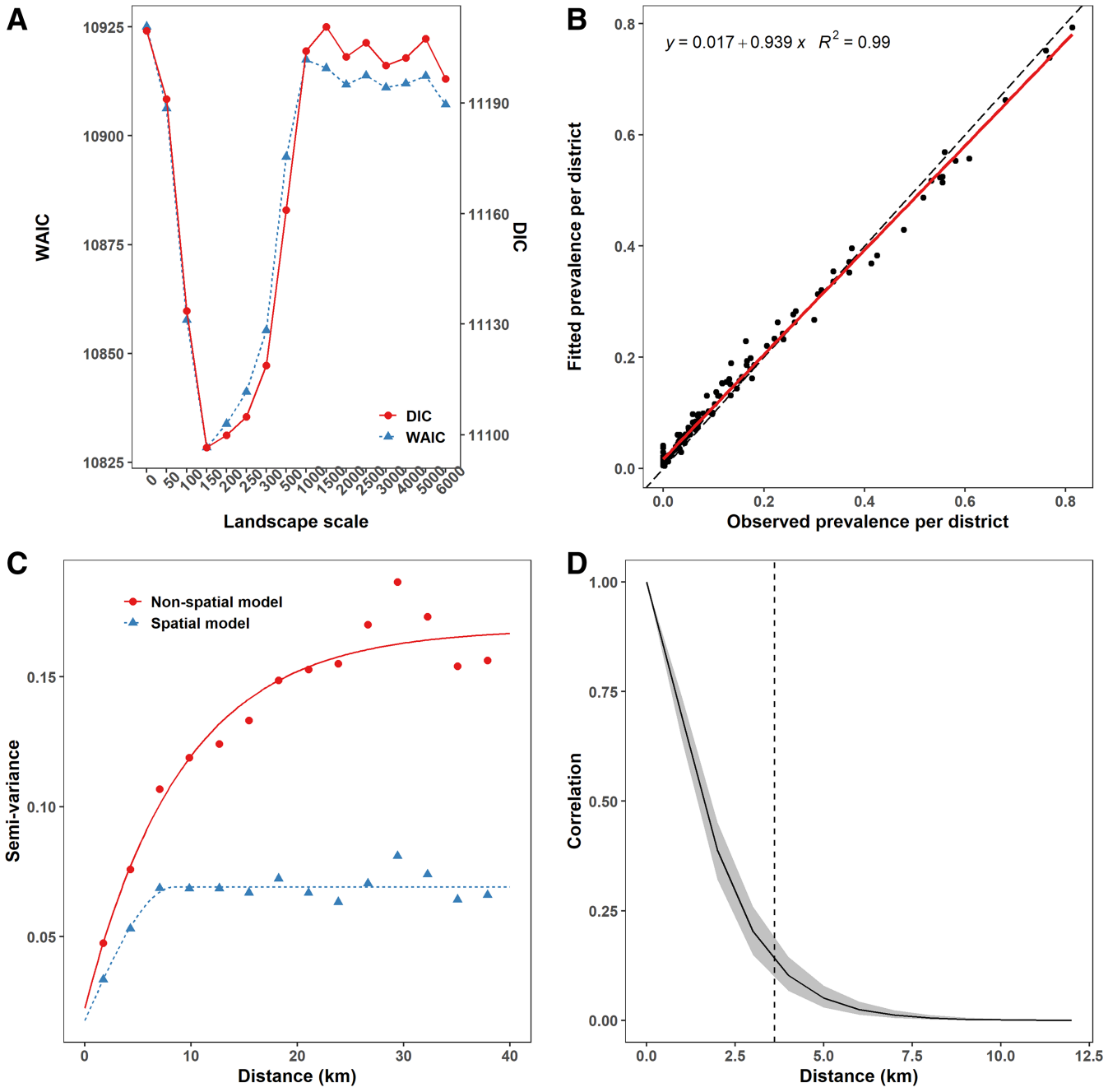

Fig. 2. Comparison and fit of the spatial logistic models M_0. **A,** Watanabe-Akaike information criterion (WAIC) and deviance information criterion (DIC) for the spatial model M_0 in 16 landscapes of increasing scales ranging from 0 m (only local explanatory variables considered) to 6,000 m. **B,** Posterior mean estimate of the adjusted prevalence of FD detection against the observed prevalence aggregated at district scale. The fitted line between the adjusted and observed prevalences is shown in red, and the $y = x$ line is shown in black. **C,** Sample variogram of the Pearson's residuals obtained for the spatial and nonspatial model M_0 at 150 m. **D,** The Matérn correlation function and its 95% credible band obtained with model M_0 at 150 m. The dashed vertical line indicates the estimated range.

corresponding odds ratios are obtained as the exponent of the posterior means. The probability of FD detection increases with the age of the plantation and its area but slightly decreases with planting density. No significant effect of the farming practice was observed. The probability of FD detection decreases significantly with altitude (Fig. 4B). A 20-m increase in altitude is associated with an odds ratio of 0.72.

We analyzed the effect of cultivar choice at field scale using Merlot, the most widespread cultivar in Bordeaux, as the reference (Fig. 4A, second panel from bottom). The odds ratio of FD detection in fields planted with Cabernet Sauvignon, Cabernet Franc, and Muscadelle were, on average, 2.76, 2.29, and 2.08, respectively, compared with fields planted with Merlot. These differences were all highly significantly positive. Conversely, the odds ratio of FD detection in a field planted with Semillon were 0.76. No significant differences were found between Merlot and the cultivars Sauvignon and others (a class containing all the other minority cultivars).

Finally, the variable AOC was used as a proxy for the socioeconomic conditions of vine cultivation, with AOC Bordeaux as the reference. For the six AOC levels considered, the probability of FD detection was significantly higher only in the southwestern part of the GDON des Bordeaux area containing the AOCs Loupiac, Graves supérieures, and Sainte-Croix-du-Mont (Fig. 4A, fourth panel). The odds ratio of FD detection in these AOCs was 2.42 compared with AOC Bordeaux.

FD risk factors at the landscape scale. The effects of the variables describing the composition and configuration of the landscape surrounding a field within a scale of 150 m are displayed in Figure 4A (third panel). The probability of FD detection increases with the proportions of forest and urban land and, more slightly, with the proportion of organic fields. Conversely, the probability of FD detection decreases substantially with the proportion of vineyards. These effects are visualized in Figure 4B. For an increase of 20 percentage points in the proportion of forest, urban areas, or fields with organic practices, the odds ratios of detection are 1.28, 1.34, and 1.13, respectively, whereas a 20 percentage points increase in the proportion of vineyards is associated with an odds ratio of 0.84.

We also tested the effect of the varietal composition of the landscape throughout the proportion of Merlot area in the vineyard. The probability of FD detection decreases substantially with increases in the proportion of fields planted with Merlot (Fig. 5B). An increase of 20 percentage points in the proportion of fields planted with Merlot, for a landscape scale of 150 m, is associated with an odds ratio of 0.93.

Predictive performance of the model. We evaluated the ability of the model to predict the FD detection status of individual fields. The corresponding PR curve indicates a moderate predictive performance (Supplementary Fig. S6A), as summarized by the area under the PR curve of 0.59 (95% credible interval 0.54 to 0.64). The predictive performance was better at district scale. Specifically, we evaluated the ability of the model to predict if the prevalence of FD detection in a district exceeds a given threshold (Supplementary Fig. S6B). Recall decreases with the threshold, whereas precision remains constant. For a threshold of 10%, the posterior mean recall (i.e., sensitivity) is 89%. This is the probability of the model correctly classifying a district as having a prevalence of FD detection >10% of the fields. The posterior mean precision (i.e., positive predictive value) is then 50%. This is the probability that prevalence of FD detection is >10% in a district when the district was classified as such by the model.

DISCUSSION

In this study, we used species distribution models fitted in a Bayesian framework to identify key biotic and abiotic factors driving the detection and epidemiology of flavescence dorée, a major quarantine disease damaging European vineyards. To this end, and based on our ecological knowledge of the disease, we integrated data from different sources (e.g., FD survey data, agricultural management, and environmental explanatory variables) to build a quantitative understanding of FD infection. Our study makes use of a large spatial dataset gathering 34,581 observations of the infection status of vineyard plots. It relies on an accurate monitoring of symptoms realized by trained inspectors completed with molecular detection of FD phytoplasma on symptomatic plants performed by accredited laboratories. As a result of the costs of laboratory testing, no tests are conducted on leaves from nonsymptomatic plants, possibly opening the way to false-negative results attributable to cryptic infection (Parnell et al. 2017). To this respect, our analysis highlights the importance

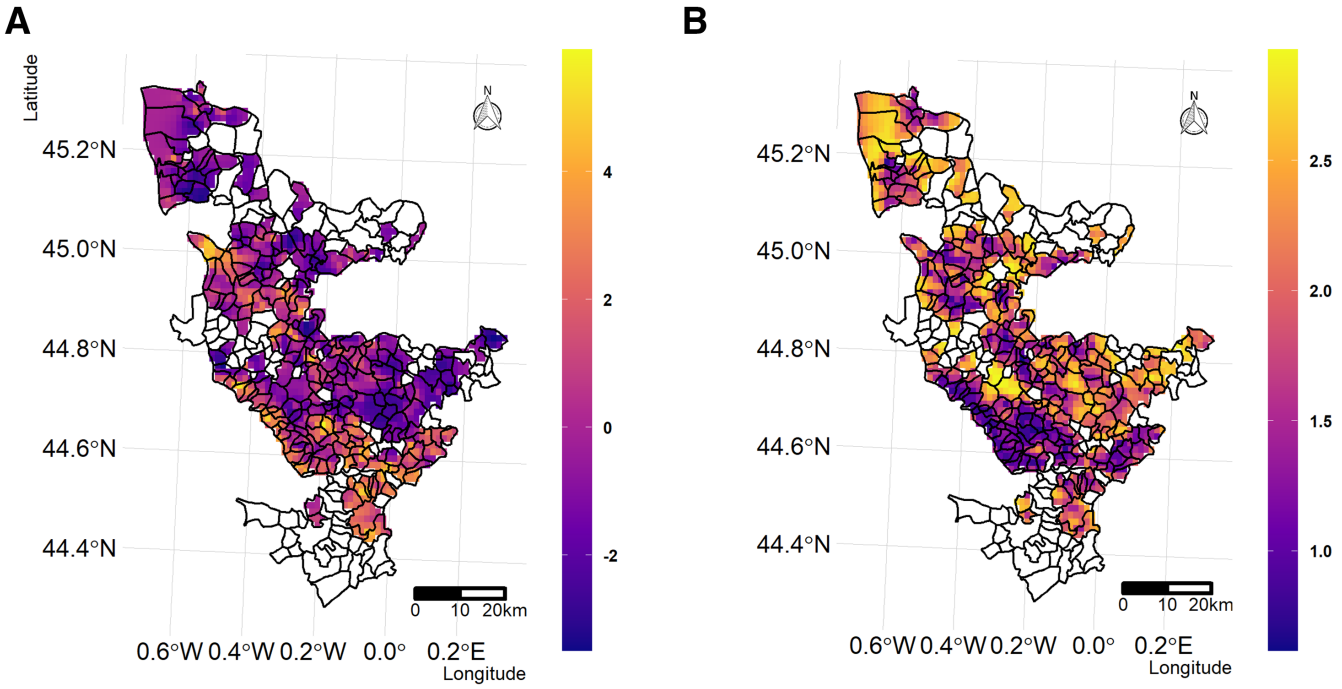

Fig. 3. Spatial random effect estimated in the Groupements de Défense contre les Organismes Nuisibles des Bordeaux area for the baseline model M_0 for a landscape scale of 150 m. No data were available for the districts in white. **A,** Posterior mean of the spatial random effect. **B,** Posterior standard deviation of the spatial random effect.

of sampling period on FD detection (variables "season" in model M_0 and "week" in model M_1). Both models indicate that the probability of FD detection substantially increases after week 36 (first week of September; Figs. 4A and 5A). The visual identification of FD symptoms is then much easier because this period is the ideal time for the expression of certain specific symptoms such as grape shriveling and nonlignification of canes (Tramontini et al. 2020). Accordingly, from September onward, the risk of false-negative results is low. Furthermore, recent studies show that nonsymptomatic plants sampled at the end of the summer in fields with high FD prevalence were actually not infected by the phytoplasma (Eveillard et al. 2016; Ripamonti et al. 2020). In particular, the study of Eveillard et al. (2016) was conducted in the Bordeaux region over hundreds of Cabernet Sauvignon and Merlot plants, the two most widespread cultivars in our dataset (76% of the samples). Moreover, inspectors collected in each field a wide diversity of symptomatic leaves possibly related to phytoplasmas. Among the 4,554 molecular detections performed in as many fields, 47% were positive for FD. Accordingly, based on

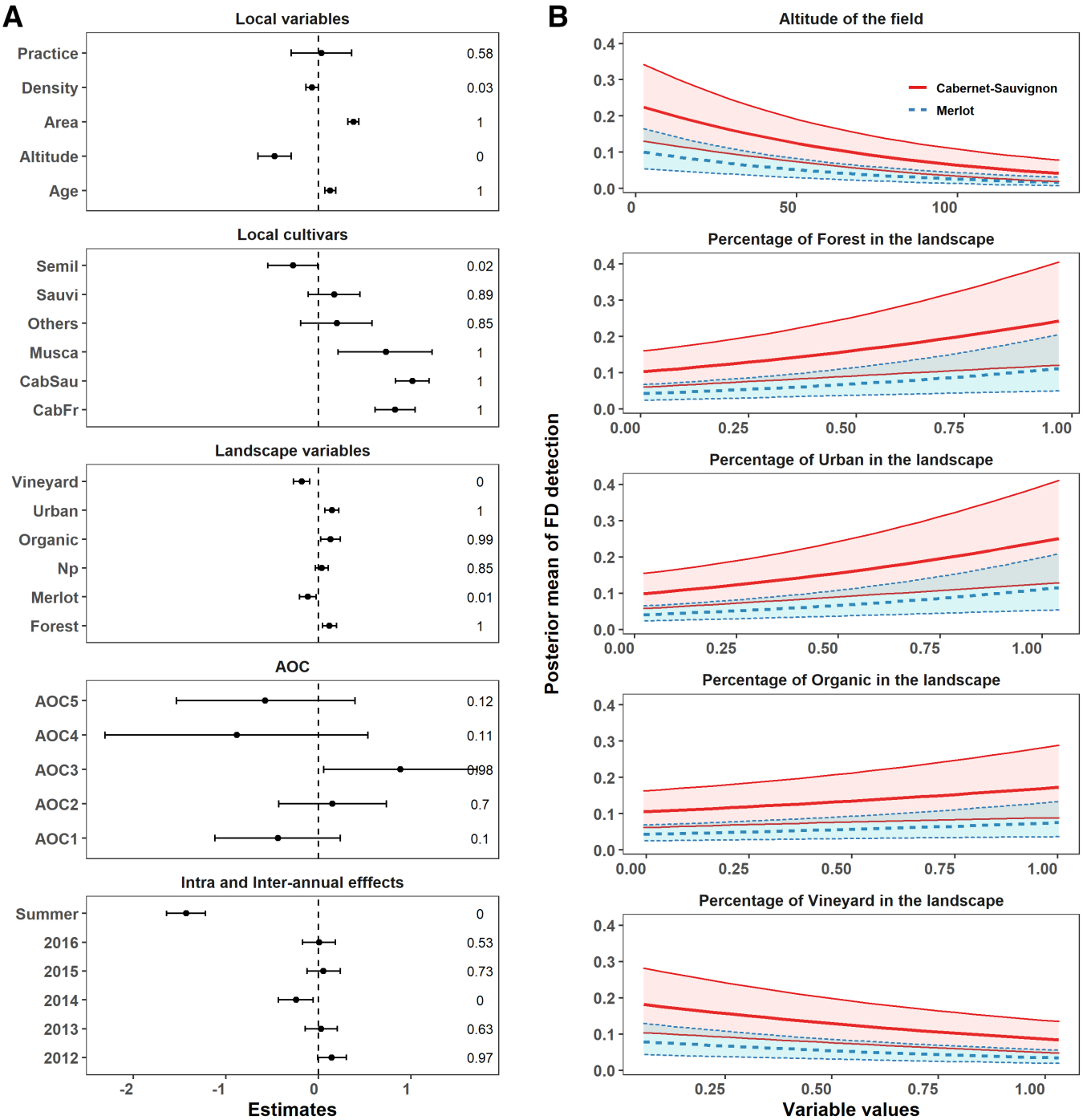

Fig. 4. Estimation of the parameters of the baseline model M_0 for a landscape scale of 150 m. **A,** Fixed and random effects of the local and landscape explanatory variables. For each variable, the posterior mean (dots) and 95% credible intervals (solid lines) are displayed, together with the posterior probability of the effect being positive. The dashed line corresponds to the value 0. **B,** Effects of altitude, proportions of forest, urban area, organic fields, and vineyard in the landscape on the probability of FD detection for the two most widespread cultivars (Merlot and Cabernet Sauvignon) in vineyard plots under conventional practices (mean probability and 95% credible bands). The effects were estimated by fixing all other numerical variables at their mean values, the variable "Appellation d'Origine Contrôlée" to Bordeaux, and the variable "practice" to conventional, and considering a sampling period during the autumn. All variables and their levels are described in Table 1. We dropped the "L_" from the landscape variables to lighten notations.

these premises, we are confident our analyses provide important information on how local farming practices and landscape context affect the probability of FD infection. At the local scale, our results reveal in particular that the probability of detecting infected fields increases with field area, the age of the plantation, and the use of specific cultivars (i.e., Cabernet Sauvignon, Cabernet Franc, and Muscadelle). At the landscape scale, the probability of FD infection increases with the proportion of susceptible cultivars as well as with the proportions of forest, urban land, and organic farming, but decreases with the proportion of vineyard.

At the field scale, cultivar choice, but not other viticultural practices considered, affected the probability of field infection. Cultivar choice was a major determinant of FD infection. No major resistance gene effective against FD is currently available, but substantial differences in cultivar susceptibility have been reported (Eveillard et al. 2016; Oliveira et al. 2019). In the GDON des Bordeaux region, the most frequently grown cultivars are Merlot (in 61% of the fields analyzed), Cabernet Sauvignon (in 15% of the fields), and Cabernet Franc and Semillon (in 7% of the fields each). Previous studies involving controlled inoculations in greenhouses or monitoring in production plots have shown that Cabernet Sauvignon and Cabernet Franc are more susceptible to FD than Merlot, displaying a higher incidence of the disease, more severe symptoms, and higher rates of phytoplasma multiplication (Eveillard et al. 2016). Our results therefore confirmed this ranking in a large set of fields for the first time. They also suggest that Semillon has low susceptibility to FD, whereas Muscadelle is highly susceptible to FD. In addition, our analyses revealed that the altitude substantially affects FD infection: the higher the altitude of the field, the lower its probability of infection (Fig. 4B). This effect fits the empirical observation that, in the region studied, FD is more likely to occur in wetland areas, which are frequently found at lower altitudes, typically close to rivers. The slightly lower goodness of fit obtained by replacing altitude by a wetland index (Merot et al. 2003) led us to favor this variable, which is also easier to access. Identification of the underlying biological factors deserves further study.

Our study reveals that field infection by FD is also affected by the landscape context within the 150 to 200 m surrounding focal fields. The best fit was obtained with the landscape scale of 150 m, followed by 200 m. These landscape extents correspond to areas ranging from 7 ha (150-m scale) to 12.5 ha (200-m scale) and contain an average of 6 to 10 fields in the study area (the mean field area is 0.76 ha). This short range may result from the low dispersal capacity of the vector of the FD phytoplasma. Indeed, 80% of *S. titanus* adults disperse within 30 m of their source, although long-range dispersal over distances of as great as 330 m has occasionally been observed (Lessio et al. 2014). Several effects of landscape composition were identified (Fig. 4B). Increases in the proportions of two generic land-use classes, forest and urban land, increased the probability of FD infection. The interpretation of such effects is not straightforward because generic land-use classes provide only an imperfect description of the presence of host habitats involved in epidemiological dynamics (Vanwambeke et al. 2019). In our case study, it is tempting to consider the positive effect of the proportion of forest as favoring wild reservoirs of FD, such as wild alders and *Clematis* sp. However, recent studies have demonstrated that the transfer of the phytoplasma from these plants to grapevine is rare, so their impact as chronic reservoirs is likely to be minimal (Filippin et al. 2009; Malembic-Maher et al. 2020). The hypothesis involving vines, whether abandoned or cultivated, is more likely. Indeed, abandoned *Vitis* sp. plants in forest margins and outskirts constitute an important source of FD inoculum and also of insect vectors (Lessio et al. 2014; Ripamonti et al. 2020). In this respect, it could be interesting to refine landscape description to investigate the role of the size of forest patches because smaller patches offer more surface on their outskirts compared with larger ones. Similarly, inventories and cartography of *Vitis* sp. in two districts of the Bordeaux area have shown that cultivated vines in individual gardens or uncultivated *Vitis* sp. in wasteland are frequent in urbanized areas neighboring

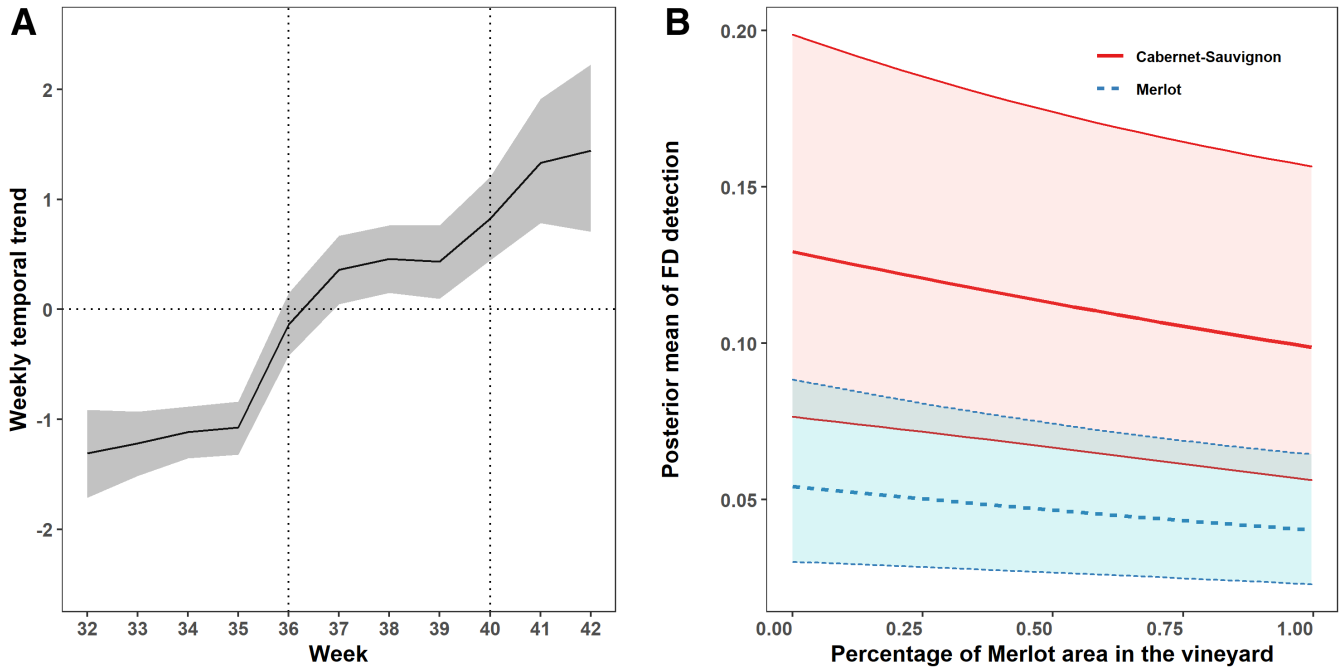

Fig. 5. Effects of the week of field inspection and of the cultivar composition of the landscape on flavescence dorée (FD) detection for a landscape scale of 150 m. **A,** Effects of the week of field inspected for FD detection as estimated with model M_1 at 150 m. In model M_1, the fixed effect variable "season" is replaced by a random walk of order 1 indexed by weeks. The mean standardized temporal effect (from week 32, the second week of August, to week 42, the third week of October) is displayed along with its 95% credible band. **B,** Effect of the proportion of Merlot in the landscape on the probability of FD detection, using the model M_0 for the two most widespread cultivars (Merlot and Cabernet Sauvignon) in vineyard plots under conventional practices (mean probability and 95% credible bands). The effects were estimated by fixing all other numerical variables at their mean values, the variable "Appellation d'Origine Contrôlée" to Bordeaux, and the variable "practice" to conventional, and considering a sampling period during the autumn.

vineyards (*unpublished data*). They could also constitute an important source of FD inoculum and insect vectors, especially as residents and gardeners have little awareness of the problem of FD. Alternatively, it is also likely that fewer insecticide treatments are performed by farmers in vineyards close to private houses or public facilities and also by amateur gardeners, which might also explain this result.

In addition, we identify that the probability of FD infection decreases with the proportion of vineyards in the landscape and also with planting density. This effect, which appears counterintuitive at first glance, may come from several processes. It may reflect a direct dilution effect of the available inoculum or of the vector population within larger cultivated vineyards. In line with this hypothesis, Delaune et al. (2021) have recently shown that univoltine pest species such as *S. titanus* are negatively correlated with the host crop area in the landscape during the ongoing growing season. This effect may also reflect that larger cultivated vineyards can benefit from a more homogeneous insecticide protection (Meehan et al. 2011). Alternatively, it might be an indirect effect mediated by the economic value of vineyards. Indeed, areas with highly valued vineyards are generally more specialized, with more financial incentives for farmers to control the infection (Ay and Gozlan 2020).

Interestingly, effects of cultivars detected at the field scale were also found at the landscape scale, suggesting upscaling of these effects that may be mediated by dispersal abilities of *S. titanus* populations. We show that the probability of FD infection decreases with increasing proportions of Merlot (Fig. 5B). Several studies have shown that the landscape composition in plant species (i.e., between-species diversity) affects plant disease epidemiology (Meentemeyer et al. 2012; Plantegenest et al. 2007). However, studies on the effect of varietal (i.e., within-species diversity) composition at the landscape scale are much rarer and mainly based on theoretical analyses (Rimbaud et al. 2021). Few empirical studies have suggested such an effect, and they mainly considered regional or national scales (Finckh and Wolfe 2017; Papaïx et al. 2011; Priestley and Bayles 1980). By contrast, our study suggests an effect of the varietal composition of the landscape at an intermediate spatial scale, typically over an area of 7 to 12 ha, containing approximately 6 to 10 fields. Finally, although no effect of farming practices (organic versus conventional) was detected at the field scale, our results suggest that the probability of FD infection may slightly increase with the proportion of organic fields in the landscape (Fig. 4B). Note first that this variable indicates if organic practices are used at a farm scale rather than for each individuals field (fields being attached to their farm in our study). It could possibly explain this mismatch between the results at local and landscape scales. The positive effect at landscape scale could suggest a lower efficacy of vector control in organic systems. The natural products used in organic vineyards to control *S. titanus* are as effective as synthetic insecticides in controlling nymphs but less effective in controlling adults, the most dispersive instar (Chuche and Thiéry 2014; Gusberti et al. 2008; Tacoli et al. 2017). Combined with the stricter conditions governing field applications, it might account for this effect. However, this hypothesis needs further investigation because we did not have the information about the number of mandatory treatments actually realized. Moreover, other practices used in conventional and organic viticulture (e.g., soil tillage, cover crops, or canopy management) that were not considered here might have affected the spread of FD and the dynamics of its vector (Muneret et al. 2018).

The local and landscape factors identified can be used to improve vineyard surveys, thereby decreasing the economic and environmental impacts of managing this disease. The use of the model as a predictive tool is potentially interesting for the targeting of new districts in which future surveys are most likely to be valuable, as suggested by the ability of the model to discriminate districts with prevalence of FD detection >10% of the fields (Supplementary Fig. S6B). Out-of-sample predicted probabilities of FD presence could be used in a benefit-cost framework to determine the optimal size of future surveys or to improve the current control policy (Ay and Gozlan 2020). The grubbing costs associated with FD presence would be derived from the farm's gross margins or vineyard prices and compared with the monitoring costs from professional organizations such as GDON. Moreover, the Bayesian method used (INLA) allows uncertainty to be quantified and visualized in risk map outputs. This is important for fostering dialogue with stakeholders and policymakers. However, and importantly, use of the model as a predictive tool must first be restricted to make interpolation to unsampled sites within the GDON des Bordeaux area and not as a tool that is able to make predictions in remote vineyards. Moreover, given the importance of the spatial effects (as discussed below), the model has also to be fitted again annually by taking into account the new FD survey to update the spatial field to the current state of the epidemic in the area. A short-term operational output of the result to improve FD monitoring concerns the effect of the sampling period. It suggests that the harvest period in Bordeaux (September) is a more favorable period for maximizing the efficacy of vineyard surveys (Figs. 4A and 5A). Note that part of the size of this effect could also be explained by the GDON internal organization of field surveys because the last weeks of the survey, mostly after week 38 (i.e., >2.8% of the sample), are realized by more experienced teams of inspectors sometimes targeting districts closer to previously identified infected fields. This result strongly suggests that equipping grapevine harvesting machines with cameras and automatic symptom-recognition tools could greatly improve FD detection, especially because harvesting machines travel through a large part of the vineyard every year. Recently, the use of convolutional neural networks has shown promise for detecting FD symptoms from in-field images (Boulent et al. 2020). Finally, our results may have important consequences for the management of vineyard landscapes, echoing the land sparing/land sharing debate (Phalan et al. 2011). Our study suggests that promoting landscapes with large amounts of vineyards and lower proportions of forests limit FD infections, while maintaining seminatural habitats in such landscapes is known to be of major importance for biodiversity conservation and mitigation of climate change effects (Barbaro et al. 2021; Rusch et al. 2022). Policymakers and land-use planners may therefore include this if they are to design multifunctional vineyard landscapes.

In addition to the effects of the fixed factors, a large proportion of the biological process of interest was captured by the spatial effects included in our models. In our case study with >35,000 observations, the INLA method proved effective for implementing complex Bayesian hierarchical models, taking spatial autocorrelation into account. The spatial component, as revealed by the Matérn function (Fig. 2D), had a strong effect, suggesting that some biotic (e.g., specific cultural practices) and abiotic (e.g., climatic variables) factors are missing or misspecified (e.g., nonlinear terms are required) (Elith and Leathwick 2009). However, the strong spatial pattern observed may also be the result of several epidemics developing simultaneously at a local level over small scales, consistent with the short dispersal distances of the vector, whereas the distribution model is based primarily on a static vision of the disease risk factor in an environment at equilibrium (Elith and Leathwick 2009). This assumption does not hold for invading pathogens (Purse and Golding 2015). In recent years, the idea of coupling niche modeling with spatially explicit models of disease dispersal to provide better information about potential disease spread and ultimately improve epidemiological surveillance strategies has been explored by several authors (Adrakey et al. 2017; Cunniffe et al. 2016; Hyatt-Twynam et al. 2017; Martinetti and Soubeyrand 2019; Parnell et al. 2017; Purse and Golding 2015; Rimbaud et al. 2018). This approach merits implementation in this context to improve prediction accuracy at the field scale. This implementation could make use of existing models of FD epidemics at the field scale (Lessio et al. 2015). The insights gained from our correlative study could be incorporated into mechanistic models of epidemic spread (Hartemink et al. 2011; Meentemyer et al. 2011). Moreover, it should also be recognized

that correlative models, which can easily take into account large numbers of predictors, sometimes outperform mechanistic models and can have a major impact on policy decisions (Leach and Scoones 2013). From this point of view, it may also be useful to explore the rapidly developing research fields of machine learning and data mining. These methods have been successfully applied to the surveillance of *X. fastidiosa* (Martinetti and Soubeyrand 2019). This possibility is particularly attractive given that machine-learning methods have recently been adapted to take spatial autocorrelation into account (Georganos et al. 2021).

ACKNOWLEDGMENTS

We thank the people who sampled or provided data: Sophie Bentejac, Morgane Le Goff, and Charlotte Labit (GDON des Bordeaux); Dominique Vergnes (FREDON Aquitaine); Gontran Casella (Syndicat des Bordeaux et Bordeaux Supérieur); and Soungalo Dembele (Master student involved in early work on the geographic information system database). We also thanks Sophie Bentejac, Charlotte Labit, and Thomas Opitz (INRAE) for their critical and careful reading of the manuscript. We thank the team at INRAE in charge of the Migale server used to perform calculations. We also thanks Ruairi Donnelly for reviewing an early draft.

LITERATURE CITED

Adrakey, H., Streftaris, G., Cunniffe, N., Gottwald, T., Gilligan, C., and Gibson, G. 2017. Evidence-based controls for epidemics using spatio-temporal stochastic models in a Bayesian framework. J. R. Soc. Interface 14:20170386.

Ay, J.-S., and Gozlan, E. 2020. Disease dispersion as a spatial interaction: The case of flavescence dorée. Nat. Resour. Model. 33:e12265.

Barbaro, L., Assandri, G., Brambilla, M., Castagneyrol, B., Froidevaux, J., Giffard, B., Pithon, J., Puig-Montserrat, X., Torre, I., Calatayud, F., Gaüzère, P., Guenser, J., Macià-Valverde, F. X., Mary, S., Raison, L., Sirami, C., and Rusch, A. 2021. Organic management and landscape heterogeneity combine to sustain multifunctional bird communities in European vineyards. J. Appl. Ecol. 58:1261-1271.

Bebber, D. P. 2015. Range-expanding pests and pathogens in a warming world. Annu. Rev. Phytopathol. 53:335-356.

Beguin, J., Martino, S., Rue, H., and Cumming, S. G. 2012. Hierarchical analysis of spatially autocorrelated ecological data using integrated nested Laplace approximation. Methods Ecol. Evol. 3:921-929.

Boulent, J., St-Charles, P.-L., Foucher, S., and Théau, J. 2020. Automatic detection of flavescence dorée symptoms across white grapevine varieties using deep learning. Front. Artif. Intel. 3:564878.

Bressan, A., Spiazzi, S., Girolami, V., and Boudon-Padieu, E. 2005. Acquisition efficiency of flavescence dorée phytoplasma by *Scaphoideus titanus* Ball from infected tolerant or susceptible grapevine cultivars or experimental host plants. Vitis 44:143-146.

Caudwell, A. 1957. Deux années d'études sur la flavescence dorée nouvelle maladie grave de la vigne. Ann. Amelior. Plant. 4:359-393.

Cendoya, M., Martinez-Minaya, J., Dalmau, V., Ferrer, A., Saponari, M., Conesa, D., Lopez-Quilez, A., and Vicent, A. 2020. Spatial Bayesian modeling applied to the surveys of *Xylella fastidiosa* in Alicante (Spain) and Apulia (Italy). Front. Plant Sci. 11:1204.

Chuche, J., and Thiéry, D. 2014. Biology and ecology of the flavescence dorée vector *Scaphoideus titanus*: A review. Agron. Sustain. Dev. 34:381-403.

Cunniffe, N. J., Cobb, R. C., Meentemeyer, R. K., Rizzo, D. M., and Gilligan, C. A. 2016. Modeling when, where, and how to manage a forest epidemic, motivated by sudden oak death in California. Proc. Natl. Acad. Sci. U.S.A. 113:5640-5645.

Delaune, T., Ouattara, M. S., Ballot, R., Sausse, C., Felix, I., Maupas, F., Chen, M., Morison, M., Makowski, D., and Barbu, C. 2021. Landscape drivers of pests and pathogens abundance in arable crops. Ecography 44: 1429-1442.

Desneux, N., Decourtye, A., and Delpuech, J.-M. 2007. The sublethal effects of pesticides on beneficial arthropods. Annu. Rev. Entomol. 52:81-106.

Dormann, C. F., McPherson, J. M., Araùjo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Kissling, W. D., Kühn, I., Ohlemüller, R., Peres-Neto, P., Reineking, B., Schröder, B., Schurr, F. M., and Wilson, R. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. Ecography 30:609-628.

Dormann, C. F., Schymanski, S. J., Cabral, J., Chuine, I., Graham, C., Hartig, F., Kearney, M., Morin, X., Römermann, C., Schröder, B., and Singer, A. 2012. Correlation and process in species distribution models: Bridging a dichotomy. J. Biogeogr. 39:2119-2131.

Elith, J., and Leathwick, J. R. 2009. Species distribution models: Ecological explanation and prediction across space and time. Annu. Rev. Ecol. Evol. Syst. 40:677-697.

Eveillard, S., Jollard, C., Labroussaa, F., Khalil, D., Perrin, M., Desqué, D., Salar, P., Razan, F., Hévin, C., Bordenave, L., Foissac, X., Masson, J. E., and Malembic-Maher, S. 2016. Contrasting susceptibilities to flavescence dorée in *Vitis vinifera*, rootstocks and wild *Vitis* species. Front. Plant Sci. 7:1762.

Filippin, L., Jovic, J., Cvrkovic, T., Forte, V., Clair, D., Tosevski, I., Boudon-Padieu, E., Borgo, M., and Angelini, E. 2009. Molecular characteristics of phytoplasmas associated with flavescence dorée in clematis and grapevine and preliminary results on the role of *Dictyophara europaea* as a vector. Plant Pathol. 58:826-837.

Finckh, M. R., and Wolfe, M. S. 2017. Biodiversity enhancement. Pages 153-174 in: Plant Diseases and Their Management in Organic Agriculture. M. R. Finckh, A. H. C. van Bruggen, and L. Tamm, eds. American Phytopathological Society, St. Paul, MN.

Fuglstad, G. A., Simpson, D., Lindgren, F., and Rue, H. 2019. Constructing priors that penalize the complexity of Gaussian random fields. J. Am. Stat. Assoc. 114:445-452.

Galetto, L., Miliordos, D., Roggia, C., Rashidi, M., Sacco, D., Marzachi, C., and Bosco, D. 2014. 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. 87:671-679.

Gelman, A., Hwang, J., and Vehtari, A. 2014. Understanding predictive information criteria for Bayesian models. Stat. Comput. 24:997-1016.

Georganos, S., Grippa, T., Gadiaga, A. N., Linard, C., Lennert, M., Vanhuysse, S., Mboga, N., Wolff, E., and Kalogirou, S. 2021. Geographical random forests: A spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling. Geocarto Int. 36:121-136.

Godefroid, M., Cruaud, A., Streito, J.-C., Rasplus, J.-Y., and Rossi, J.-P. 2019. *Xylella fastidiosa*: Climate suitability of European continent. Sci. Rep. 9:8844.

Gusberti, M., Jermini, M., Wyss, E., and Lindner, C. 2008. Efficacité d'insecticides contre *Scaphoideus titanus* en vignobles biologiques and effets secondaires. Rev. Suisse Vitic. Arboric. Hortic. 40:173-177.

Hartemink, N., Vanwambeke, S. O., Heesterbeek, H., Rogers, D., Morley, D., Pesson, B., Davies, C., Mahamdallie, S., and Ready, P. 2011. Integrated mapping of establishment risk for emerging vector-borne infections: A case study of canine leishmaniasis in southwest France. PLoS One 6:e20817.

Hesselbarth, M. H., Sciaini, M., With, K. A., Wiegand, K., and Nowosad, J. 2019. Landscapemetrics: An open-source r tool to calculate landscape metrics. Ecography 42:1648-1657.

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

Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., and Rodes, I. 2017. Operational high resolution land cover map production at the country scale using satellite image time series. Remote Sens. 9-95.

Jeger, M., Bragard, C., Caffier, D., Candresse, T., Chatzivassiliou, E., Dehnen-Schmutz, K., Gilioli, G., Miret, J. A. J., et al. 2016. Risk to plant health of flavescence dorée for the EU territory. EFSA J. 14:e04603.

Kraemer, M. U. G., Hay, S. I., Pigott, D. M., Smith, D. L., Wint, G. R. W., and Golding, N. 2016. Progress and challenges in infectious disease cartography. Trends Parasitol. 32:19-29.

Krainski, E., Gomez-Rubio, V., Bakka, H., Lenzi, A., Castro-Camilo, D., Simpson, D., Lindgren, F., and Rue, H. 2018. Advanced spatial modeling with stochastic partial differential equations using R and INLA. Chapman and Hall/CRC, London.

Leach, M., and Scoones, I. 2013. The social and political lives of zoonotic disease models: Narratives, science and policy. Soc. Sci. Med. 88:10-17.

Lessio, F., and Alma, A. 2004. Dispersal patterns and chromatic response of *Scaphoideus titanus* Ball (Homoptera Cicadellidae), vector of the phytoplasma agent of grapevine flavescence dorée. Agric. For. Entomol. 6:121-128.

Lessio, F., Portaluri, A., Paparella, F., and Alma, A. 2015. A mathematical model of flavescence dorée epidemiology. Ecol. Modell. 312:41-53.

Lessio, F., Tota, F., and Alma, A. 2014. Tracking the dispersion of *Scaphoideus titanus* Ball (Hemiptera: Cicadellidae) from wild to cultivated grapevine: Use of a novel mark-capture technique. Bull. Entomol. Res. 104: 432-443.

Lindgren, F., and Rue, H. 2015. Bayesian spatial modelling with R-INLA. J. Stat. Softw. 63:1-25.

Lindgren, F., Rue, H., and Lindstrom, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B. 73:423-498.

Malembic-Maher, S., Desqué, D., Khalil, D., Salar, P., Bergey, B., Danet, J.-L., Duret, S., et al. 2020. 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 Pathog. 16: e1007967.

Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., Tyre, A. J., and Possingham, H. P. 2005. Zero tolerance ecology: Improving ecological inference by modelling the source of zero observations. Ecol. Lett. 8:1235-1246.

Martinetti, D., and Soubeyrand, S. 2019. Identifying lookouts for epidemio-surveillance: Application to the emergence of *Xylella fastidiosa* in France. Phytopathology 109:265-276.

Martínez-Minaya, J., Conesa, D., Lopez-Quilez, A., and Vicent, A. 2018. Spatial and climatic factors associated with the geographical distribution of citrus black spot disease in South Africa. A Bayesian latent Gaussian model approach. Eur. J. Plant Pathol. 151:991-1007.

Meehan, T., Werling, B., Landis, D., and Gratton, C. 2011. Agricultural landscape simplification and insecticide use in the midwestern United States. Proc. Natl. Acad. Sci. U.S.A. 108:11500-11505.

Meentemeyer, R. K., Anacker, B. L., Mark, W., and Rizzo, D. M. 2008. Early detection of emerging forest disease using dispersal estimation and ecological niche modeling. Ecol. Appl. 18:377-390.

Meentemeyer, R. K., Haas, S. E., and Vàclavik, T. 2012. Landscape epidemiology of emerging infectious diseases in natural and human-altered ecosystems. Annu. Rev. Phytopathol. 50:379-402.

Meentemyer, R. K., Cunniffe, N. J., Cook, A. R., Filipe, J. A. N., Hunter, R. D., Rizzo, D. M., and Gilligan, C. A. 2011. Epidemiological modelling of invasion in heterogeneous landscapes: Spread of sudden oak death in California (1990-2030). Ecosphere 2:1-24.

Merot, P., Squividant, H., Aurousseau, P., Hefting, M., Burt, T., Maitre, V., Kruk, M., Butturini, A., Thenail, C., and Viaud, V. 2003. Testing a climato-topographic index for predicting wetlands distribution along a European climate gradient. Ecol. Modell. 163:51-71.

Morone, C., Boveri, M., Giosue, S., Gotta, P., Rossi, V., Scapin, I., and Marzachi, C. 2007. Epidemiology of flavescence dorée in vineyards in northwestern Italy. Phytopathology 97:1422-1427.

Muneret, L., Mitchell, M., Seufert, V., Aviron, S., Djoudi, E. A., Pétillon, J., Plantegenest, M., Thiéry, D., and Rusch, A. 2018. Evidence that organic farming promotes pest control. Nat. Sustain. 1:361-368.

Namba, S. 2019. Molecular and biological properties of phytoplasmas. Proc. Jpn. Acad., Ser. B, Phys. Biol. Sci. 95:401-418.

Oliveira, M. J. R., Roriz, M., Vasconcelos, M. W., Bertaccini, A., and Carvalho, S. M. P. 2019. Conventional and novel approaches for managing "flavescence dorée" in grapevine: Knowledge gaps and future prospects. Plant Pathol. 68:3-17.

Papaïx, J., Goyeau, H., Du Cheyron, P., Monod, H., and Lannou, C. 2011. Influence of cultivated landscape composition on variety resistance: An assessment based on wheat leaf rust epidemics. New Phytol. 191:1095-1107.

Papura, D., Burban, C., van Helden, M., Giresse, X., Nusillard, B., Guillemaud, T., and Kerdelhué, C. 2012. Microsatellite and mitochondrial data provide evidence for a single major introduction for the Neartic leafhopper *Scaphoideus titanus* in Europe. PLoS One 7:e36882.

Parnell, S., Gottwald, T. R., Riley, T., and van den Bosch, F. 2014. A generic risk-based surveying method for invading plant pathogens. Ecol. Appl. 24:779-790.

Parnell, S., van den Bosch, F., Gottwald, T. R., and Gilligan, C. A. 2017. Surveillance to inform control of emerging plant diseases: An epidemiological perspective. Annu. Rev. Phytopathol. 55:591-610.

Pelletier, C., Salar, P., Gillet, J., Cloquemin, G., Very, P., Foissac, X., and Malembic-Maher, S. 2009. 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 48:87-95.

Phalan, B., Onial, M., Balmford, A., and Green, R. E. 2011. Reconciling food production and biodiversity conservation: Land sharing and land sparing compared. Science 333:1289-1291.

Plantegenest, M., Le May, C., and Fabre, F. 2007. Landscape epidemiology of plant diseases. J. R. Soc. Interface 4:963-972.

Priestley, R. H., and Bayles, R. A. 1980. Varietal diversification as a means of reducing the spread of cereal diseases in the United Kingdom. J. Natl. Inst. Agric. Bot. 15:205-214.

Purse, B. V., and Golding, N. 2015. Tracking the distribution and impacts of diseases with biological records and distribution modelling. Biol. J. Linn. Soc. 115:664-677.

Quaglino, F., Zhao, Y., Casati, P., Bulgari, D., Bianco, P. A., Wei, W., and Davis, R. E. 2013. '*Candidatus* Phytoplasma solani', a novel taxon associated with stolbur- and bois noir-related diseases of plants. Int. J. Syst. Evol. Microbiol. 63:2879-2894.

Rimbaud, L., Bruchou, C., Dallot, S., Pleydell, D. R. J., Jacquot, E., Soubeyrand, S., and Thébaud, G. 2018. Using sensitivity analysis to identify key factors for the propagation of a plant epidemic. R. Soc. Open Sci. 5:171435.

Rimbaud, L., Fabre, F., Papaïx, J., Moury, B., Lannou, C., Barrett, L. G., and Thrall, P. H. 2021. Models of plant resistance deployment. Annu. Rev. Phytopathol. 59:125-152.

Ripamonti, M., Pegoraro, M., Morabito, C., Gribaudo, I., Schubert, A., Bosco, D., and Marzachi, C. 2021. Susceptibility to flavescence dorée of different *Vitis vinifera* genotypes from north-western Italy. Plant Pathol. 70:511-520.

Ripamonti, M., Pegoraro, M., Rossi, M., Bodino, N., Beal, D., Panero, L., Marzachì, C., and Bosco, D. 2020. Prevalence of flavescence dorée phytoplasma-infected *Scaphoideus titanus* in different vineyard agroecosystems of northwestern Italy. Insects 11:301.

Rue, H., Martino, S., and Chopin, N. 2009. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations. J. R. Stat. Soc. Ser. B. 71:319-392.

Rusch, A., Beaumelle, L., Giffard, B., and Alonso Ugaglia, A. 2022. Harnessing biodiversity and ecosystem services to safeguard multifunctional vineyard landscapes in a global change context. Pages 305-335 in: Advances in Ecological Research. D. A. Bohan, A. J. Dumbrell, and A. J. Vanbergen, eds. Academic Press, San Diego.

Saito, T., and Rehmsmeier, M. 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One 10:e0118432.

Schvester, D., Carle, P., and Moutous, G. 1969. Nouvelles données sur la transmission de la flavescence dorée de la vigne par *Scaphoideus littoralis* Ball. Ann. Zool. Ecologie Anim. 1:445-465.

Simpson, D., Rue, H., Riebler, A., Martins, T. G., and Sørbye, S. H. 2017. Penalising model component complexity: A principled, practical approach to constructing priors. Stat. Sci. 32:1-28.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and der Linde, A. V. 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc. 64:583-639.

Tacoli, F., Mori, N., Pozzebon, A., Cargnus, E., Da Via, S., Zandigiacomo, P., Duso, C., and Pavan, F. 2017. Control of *Scaphoideus titanus* with natural products in organic vineyards. Insects 8:129.

Tang, F. H. M., Lenzen, M., McBratney, A., and Maggi, F. 2021. Risk of pesticide pollution at the global scale. Nat. Geosci. 14:206-210.

Tramontini, S., Delbianco, A., and Vos, S. 2020. Pest survey card on flavescence dorée phytoplasma and its vector *Scaphoideus titanus*. EFSA Supporting Publ. 17:1909E.

Tscharntke, T., and Brandl, R. 2004. Plant-insect interactions in fragmented landscapes. Annu. Rev. Entomol. 49:405-430.

Václavík, T., Kanaskie, A., Hansen, E. M., Ohmann, J. L., and Meentemeyer, R. K. 2010. Predicting potential and actual distribution of sudden oak death in Oregon: Prioritizing landscape contexts for early detection and eradication of disease outbreaks. For. Ecol. Manage. 260:1026-1035.

Vanwambeke, S. O., Linard, C., and Gilbert, M. 2019. Emerging challenges of infectious diseases as a feature of land systems. Curr. Opin. Environ. Sustain. 38:31-36.

Watanabe, S. 2010. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. J. Mach. Learn. Res. 11:3571-3594.

Weintraub, P. G., and Beanland, L. 2006. Insect vectors of phytoplasmas. Annu. Rev. Entomol. 51:91-111.

Zuur, A. F., Ieno, E. N., and Saveliev, A. A. 2017. Beginner's Guide to Spatial, Temporal and Spatial-Temporal Ecological Data Analysis with r-INLA. Highland Statistics, Newburgh, U.K.

Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., and Smith, G. M. 2009. Mixed Effects Models and Extensions in Ecology with R. Springer Science, New York.