RESEARCH ARTICLE

Spatially explicit ecological modeling improves empirical characterization of plant pathogen dispersal

Petteri Karisto[1,2] | Frédéric Suffert[3] | Alexey Mikaberidze[1,4]

[1]Plant Pathology Group, Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland

[2]Plant Health, Natural Resources Institute Finland, Jokioinen, Finland

[3]Université Paris-Saclay, INRAE, AgroParisTech, UMR BIOGER, 78850, Thiverval-Grignon, France

[4]School of Agriculture, Policy and Development, University of Reading, Whiteknights, Reading, UK

Correspondence
Alexey Mikaberidze, School of Agriculture, Policy and Development, University of Reading, Whiteknights, Reading, UK.
Email: a.mikaberidze@reading.ac.uk

Abstract

Dispersal is a key ecological process, but it remains difficult to measure. By recording numbers of dispersed individuals at different distances from the source, one acquires a dispersal gradient. Dispersal gradients contain information on dispersal, but they are influenced by the spatial extent of the source. How can we separate the two contributions to extract knowledge about dispersal? One could use a small, point-like source for which a dispersal gradient represents a dispersal kernel, which quantifies the probability of an individual dispersal event from a source to a destination. However, the validity of this approximation cannot be established before conducting measurements. This represents a key challenge hindering progress in characterization of dispersal. To overcome it, we formulated a theory that incorporates the spatial extent of sources to estimate dispersal kernels from dispersal gradients. Using this theory, we re-analyzed published dispersal gradients for three major plant pathogens. We demonstrated that the three pathogens disperse over substantially shorter distances compared to conventional estimates. This method will allow the researchers to re-analyze a vast number of existing dispersal gradients to improve our knowledge about dispersal. The improved knowledge has potential to advance our understanding of species' range expansions and shifts, and inform management of weeds and diseases in crops.

KEYWORDS
dispersal ecology, dispersal gradient, dispersal kernel, dispersal theory, experimental design, mathematical modeling

1 | INTRODUCTION

Dispersal is an important component of many life histories. Knowledge of dispersal is crucial to our understanding of fundamental population biological processes, e.g., changes in species' geographic ranges (Bonte & Dahirel, 2017; Brooker et al., 2007; Kubisch et al., 2013) as well as size and composition of ecological communities (Leibold et al., 2004; Ron et al., 2018). This knowledge also improves our capacity to conserve endangered species (McConkey et al., 2012; Musciano et al., 2020) and protect crop plants from weeds (Petit et al., 2013) and diseases (Fabre et al., 2012; Mikaberidze et al., 2016). Consequently, empirical characterization of dispersal has been a major research theme for over a hundred years (e.g., Bullock et al., 2017; Fabre et al., 2021; Heald, 1913;

Nathan et al., 2012). However, there is still far fewer datasets describing plant dispersal than plant demography because dispersal remains difficult to measure (Bullock et al., 2017). Here, we identified and resolved one of the key challenges that hinders progress in empirical characterization of dispersal: We incorporated the spatial extent of dispersal sources in the analysis of dispersal measurements.

One approach to measure dispersal is to use spatially localized sources of dispersing individuals and record dispersal gradients produced from them. These gradients do contain relevant information about dispersal, but they are also influenced by the spatial extent of the source (Cousens & Rawlinson, 2001; Ferrandino, 1996; Zadoks & Schein, 1979).

How can we evaluate this influence to extract more general knowledge about dispersal from specific dispersal gradients? This can be achieved using a mathematical description of dispersal with the help of dispersal kernels. A dispersal kernel quantifies the probability of an individual dispersal event from a source point to a destination point. Technically, a dispersal kernel is a probability density function that depends on the location of the destination point ("dispersal location kernel", Nathan et al., 2012).

To characterize dispersal, we need to estimate dispersal kernels based on observed dispersal gradients. An observed dispersal gradient from a point source would correspond to the dispersal kernel. However, sources usually need to have a certain area to yield sufficient number of dispersing propagules to be observed. How can we achieve a sufficiently small source to be considered a point? In an influential book on plant disease epidemics, Zadoks and Schein (1979) formulated a rule of thumb stating that a point source should have "a diameter smaller than 1% of the gradient length". However, this rule of thumb is misleading. The size of the source should be compared with the characteristic distance of dispersal rather than the gradient length. However, we do not know the characteristic dispersal distance in advance of conducting measurements. Therefore, we cannot establish sound criteria for the validity of the point source approximation in advance of conducting measurements.

Due to the lack of clear criteria, "point" sources of various sizes appear in the literature: an adult tree (lichen *Lobaria pulmonaria*, Werth et al., 2006), circles of 1.6m diameter (thistles *Carduus nutans, Carduus acanthoides*, Skarpaas & Shea, 2007), circles of 0.5m diameter (garlic mustard *Alliaria petiolata*, Loebach & Anderson, 2018), 4m² squares (wine raspberry *Rubus phoenicolasius*, Japanese barberry *Berberis thunbergii*, multiflora rose *Rosa multiflora*, and Japanese stiltgrass *Microstegium vimineum*, Emsweller et al., 2018), and even entire agricultural fields (oilseed rape *Brassica napus*, Devaux et al., 2007). These studies reported valuable dispersal gradients, but using these dispersal gradients as proxies for dispersal kernels may be unjustified. Spatially explicit modeling has been suggested (Greene & Calogeropoulos, 2002) to address this problem and was used in some modeling studies (Clark et al., 1999; Shaw et al., 2006), but it was not widely adopted in the literature on experimental dispersal measurements.

In this study, we devised a systematic approach to estimate dispersal kernels from dispersal gradients without using the point source approximation. For this purpose, we combined theory, analysis of empirical data and numerical simulations. We first formulated a theory that incorporates dispersal from a spatially extended source considering each point within the source area as an independent point source (the spatially explicit approach). We highlighted how mathematical properties of widely used kernel functions (exponential, Gaussian and power-law, Nathan et al., 2012) can inform experimental design. Then, we re-analyzed published empirical datasets on three major plant pathogens with contrasting spatial scales of dispersal and conducted comprehensive numerical simulations. In this way, we demonstrated how this approach allows the researchers to estimate dispersal kernels more accurately than using the point source approximation.

2 | DESCRIPTION

2.1 | Theory

The probability of dispersal from a source point $p_s = (x_s, y_s)$ to a destination point $p_d = (x_d, y_d)$ is given by the *dispersal location kernel* (hereafter "dispersal kernel"). It is typically a monotonically decreasing function of the distance between the points.

To estimate a dispersal kernel using a dispersal gradient produced by an area source, we consider the cumulative effect of all point-to-point dispersal events from the source to the destination. This is achieved by taking an integral over the individual points comprising the source to calculate their combined contribution to the dispersed population at a certain destination point (as in Shaw et al., 2006, Equation (4.6)). Similarly, the integral over all points of the destination area gives the total number of individuals that moved there from the source (as in Rimbaud et al., 2018, Equation (16)):

$$N_1(S, D) = \iint_D \iint_S n_0(p_s) \kappa(p_s, p_d) \, dA_S \, dA_D, \tag{1}$$

where $S = \{p_s\}$ is the source area, $D = \{p_d\}$ is the destination area, $n_0(p_s)$ is the density of individuals within S before dispersal, and $\kappa(p_s, p_d)$ is the dispersal kernel (key variables and parameters are listed in Table 1). Equation (1) provides a valid description of the dispersal process when the overall population size is sufficiently large so that stochastic fluctuations in the numbers of dispersed individuals can be neglected. When the populations before dispersal ($n_0(p_s)$) and after dispersal (N_1) are measured, the only unknown in Equation (1) remains the dispersal kernel. Equation (1) offers a way to estimate dispersal kernel parameters that takes into account the spatial extent of both the source and the destination.

A simpler, more common but often inaccurate approach is to fit a function of one spatial coordinate x to dispersal gradient data. For example, the function

$$N_1 = C e^{-x/\alpha} \tag{2}$$

can be fitted to dispersal gradient data to estimate the scale parameter α of the exponential kernel (for example, Saint-Jean et al., 2004). This

TABLE 1 Key variables and parameters.

Symbol	Definition
κ	Dispersal kernel
n_0	Population density in the source area before dispersal [· / m^2]
N_0	Population size in the source area before dispersal
N_1	Population size in the destination area after dispersal
S	Source area
D	Destination area
p_s, p_d	A source point/a destination point
$C_{k,e}/C_{k,g}/C_{k,p}$	Normalization factors of exponential/Gaussian/power-law kernels
$k = 1, 2$	Number of dimensions
I_0	Disease intensity in the source area before dispersal [· / m^2]
I_1	Disease intensity in the destination area after dispersal [· / m^2]
β	Transmission parameter
α	Scale parameter of exponential or Gaussian kernel [m]
γ	Shape parameter of power-law kernel
λ	Scale parameter of power-law kernel [m]
r	Distance from the source [m]
$\bar{r}$	Mean dispersal distance [m]
r_L	L th percentile of the dispersal distance kernel [m]
x_s/y_s	x-/y-coordinate of a source point [m]
x_d/y_d	x-/y-coordinate of a destination point [m]
w_x/w_y	Width of the source area along the x-axis/y-axis [m]
w_d	Width of the destination area along the y-axis [m]
b_s/b_d	Width of the border zone outside the source area/destination area along the y-axis [m]

approach works for any kernel function if both the source and the destination can be considered as points.

However, when the source or the destination is extended in space, the above approach may lead to inaccurate estimates of kernel parameters. In particular, extended sources modify dispersal gradients compared to point sources. Figure 1 illustrates such modifications for exponential, Gaussian and power-law kernels (defined in Box 1). Compare, for example, the gradients produced by the point source (source 1 in Figure 1a) and the area source (source 4 in Figure 1a). Extension of the source leads to a "flattening" of the gradient for the exponential and the power-law kernels, but for the Gaussian kernel, it leads to a "steepening" of the gradient (cf. the dashed green curve with the dashed blue curve in Figure 1b,d). Some studies postulated that gradients produced by spatially extended sources are "flatter" than gradients resulting from more localized sources (Cousens & Rawlinson, 2001; Ferrandino, 1996; Greene & Calogeropoulos, 2002; Zadoks & Schein, 1979). Here, we demonstrated that whether the extension of the source leads to a "flattening" or to a "steepening" of the

gradient depends on the underlying kernel function. Thus, using a dispersal gradient from an extended source as a proxy for a dispersal kernel can lead to either an overestimation or an underestimation of the associated kernel parameters.

Only in special cases, does the shape of the dispersal gradient match to the shape of the dispersal kernel even when the source is extended, whereby the analysis can be simplified. (i) If the source is extended only in the direction of the measured gradient (along the x-axis) and dispersal is governed by the exponential kernel, using Equation (2) will give a correct estimate of α, because exponential kernels are "memoryless" (Box 1). This is visible in Figure 1b where source 1 and source 2 produce the same gradients. However, this does not work for Gaussian and power-law kernels (Figure 1c,d). (ii) If the source is extended along the y-axis, perpendicular to the direction of measured gradient, a similar simplification is possible for the Gaussian kernel (Figure 1c, source 1 and 3). This is due to separability (Box 1) of the kernel, whereby each point source within the line source 3 in Figure 1a produces the same gradient along the x-axis. This holds for any separable kernel, but does not hold for non-separable kernels such as exponential or power-law kernels (Figure 1b,d). Analogous simplifications can be made when considering spatially extended destinations.

Insights presented above inform design and analysis of dispersal experiments. Gaussian and exponential kernels have been used in a number of studies to describe dispersal across a range of taxonomic groups (Table 15.1 in Nathan et al., 2012). When dispersal is governed by a memoryless (exponential) or a separable (e.g., Gaussian) kernel, appropriate line sources should be used to boost the power of the source, while maintaining the validity of the point source approximation to simplify the analysis. However, in most cases dispersal is better described by kernels that are neither memoryless nor separable (Nathan et al., 2012), such as the power-law kernel in Equation (5). In these cases, or when the kernel function is not known before conducting measurements, dispersal gradients should be analyzed using a spatially explicit approach based on Equation (1), as we demonstrate next.

2.2 | Experimental design and data analysis

We re-analyzed published empirical data on dispersal gradients using the spatially explicit method that incorporates the spatial extent of the source and compared the outcomes with those based on the conventional point source approximation. We considered three datasets collected in field experiments investigating dispersal of major pathogens of crop plants: (i) the fungus *Zymoseptoria tritici* that causes septoria tritici blotch in wheat (Karisto et al., 2022); (ii) the fungus *Puccinia striiformis* that causes stripe (yellow) rust in wheat (Cowger et al., 2005; Sackett & Mundt, 2005); and (iii) the oomycete *Phytophthora infestans* that causes late blight in potatoes (Gregory, 1968). The three pathosystems represent contrasting mechanisms and spatial scales of dispersal. Asexual spores of *Z. tritici* (pycnidiospores) move pre-dominantly via rain splash, while
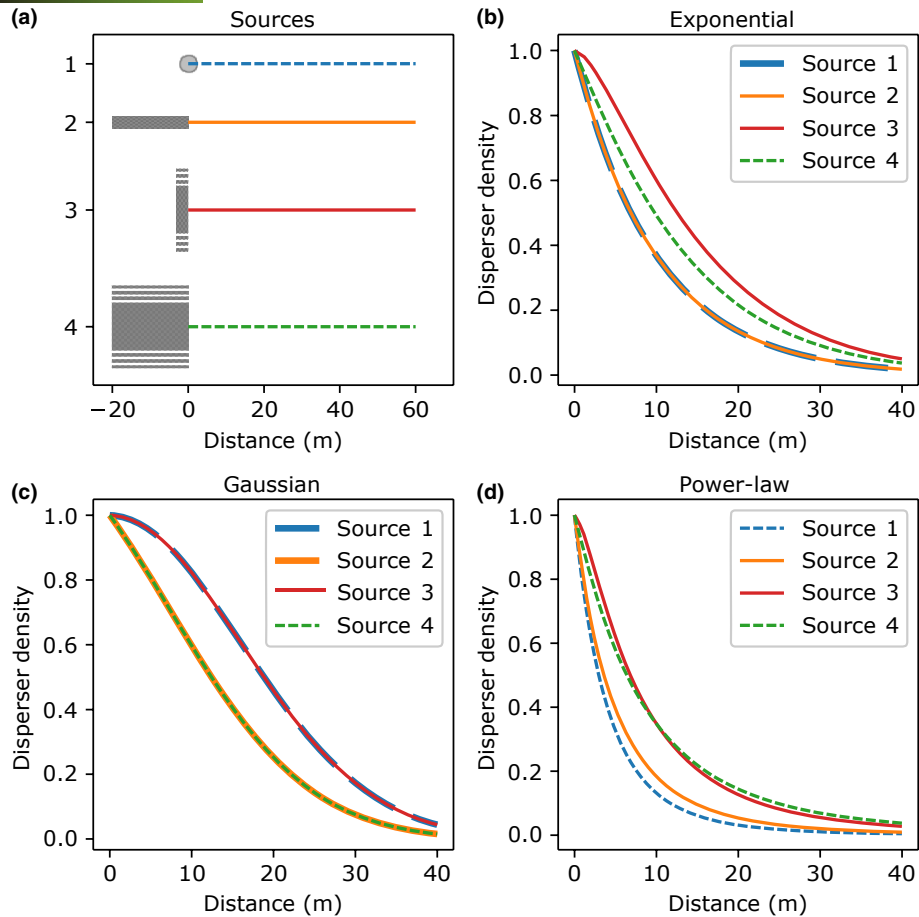
FIGURE 1 Different extensions of the source (panel a) lead to different effects on the dispersal gradients (b–d) depending on the dispersal kernel. Kernel parameters are chosen so that the mean dispersal distance $\bar{r} = 20$ m in all cases. All gradients are normalized to begin at one. (a) Four different sources (gray shapes): (1) a point source; (2) a line source, parallel to the gradient (colored lines) $x_S \in [-20\,\text{m}, 0\,\text{m}]$; (3) a line source perpendicular to the gradient, $y_s \in [-100\,\text{m}, 100\,\text{m}]$; (4) a rectangular area source, $(x_s, y_s) \in [-20\,\text{m}, 0\,\text{m}] \times [-100\,\text{m}, 100\,\text{m}]$. (b) With the exponential kernel ($\alpha = 10$ m), sources 1 and 2 result in identical gradients. (c) With the Gaussian kernel ($\alpha = 16$ m), the gradients are identical between sources 1 and 3 and between sources 2 and 4. (d) With the power-law kernel ($\gamma = 5, \lambda = 20$ m) all gradients are different.

asexual spores of *P. striiformis* (urediniospores) and propagules of *P. infestans* (sporangia) are mainly wind-dispersed. Spatial scales of the experiments varied from 100 cm to 100 m. Design of experimental plots and measurements is shown in Figure 2.

In each experiment, pathogen spores were inoculated across inoculation areas within experimental plots to create area sources of dispersing populations (orange areas in Figure 2). Then, disease gradients (disease intensity versus distance from the source) were recorded outside the inoculation areas across rectangular areas situated at increasing distances from the source (we call these areas "measurement lines"; light brown rectangles in Figure 2). These disease gradients are called primary gradients, because they resulted from a single cycle of pathogen reproduction (based on latent periods and timing of infections). The cycle includes both spore dispersal and infection success, hence, the measured gradients reflected effective dispersal gradients of the pathogen population (analogous to the combination of seed dispersal and establishment, Klein et al., 2013).

In the analysis, we incorporated the spatial extent of the source areas in two dimensions, but considered the measurement lines as thin lines perpendicular to the dispersal direction, since their length along the dispersal direction was short (dark brown lines in Figure 2). For each dataset, we chose an appropriate dispersal kernel function based on the original study, to allow for comparison with the results of the original analysis. Then we derived specific expressions for dispersal gradients, firstly, using the point source approximation (i.e., assuming a point-like source and destinations in the middle of the inoculation area or measurement lines; "1D" in Figure 2) and, secondly, based on the spatially explicit Equation (1). Based on these expressions, we estimated dispersal kernel parameters.

2.2.1 | Septoria tritici blotch

We analyzed a subset of data collected in a larger experiment (Karisto et al., 2022) that characterized dispersal of a specific

---

BOX 1   Dispersal kernels and their special properties.

Exponential kernel is defined as

$$\kappa_e(r) = C_{k,e}e^{-r/\alpha},\tag{3}$$

where α is the scale parameter, $k \in \{1, 2\}$ is the number of dimensions, $r = r(p_s, p_d)$ is the Euclidean distance from the source point $p_s = (x_s, y_s)$ to the destination point $p_d = (x_d, y_d)$ (in one dimension $y_s = y_d = 0$), and $C_{k,e}$ is the normalization factor: $C_{1,e} = 1/(2\alpha)$, $C_{2,e} = 1/(2\pi\alpha^2)$. The mean dispersal distance for $k = 2$ is $\bar{r}_e = 2\alpha$.

Gaussian kernel is defined as

$$\kappa_g(r) = C_{k,g}e^{-r^2/2\alpha^2},\tag{4}$$

where α is the scale parameter, $C_{1,g} = 1/\sqrt{2\pi\alpha^2}$, and $C_{2,g} = 1/(2\pi\alpha^2)$. The mean dispersal distance for $k = 2$ is $\bar{r}_g = \alpha\sqrt{\pi/2}$.

Power-law kernel is defined here as

$$\kappa_p(r) = C_{k,p}(\lambda + r)^{-\gamma},\tag{5}$$

where γ is the shape parameter, λ is the scale parameter, $C_{1,p} = (\gamma - 1)\lambda^{\gamma-1}$, $C_{2,p} = (\gamma - 2)(\gamma - 1)\lambda^{\gamma-2}/(2\pi)$. The mean dispersal distance for $k = 2$ is $\bar{r}_p = 2\lambda/(\gamma - 3)$ for $\gamma > 3$ and $\bar{r}_p = \infty$ for $\gamma \leq 3$.

Memorylessness. Exponential kernels are memoryless: when we set any point in the distribution as a starting point, the tail of the distribution will have the same shape as the entire distribution, i.e., the "past" does not affect the "future" probabilities:

$$\frac{f(x_1) - f(x_1 + y)}{f(x_1)} = \frac{f(x_2) - f(x_2 + y)}{f(x_2)},$$

for any starting points x_1, x_2 (see also Ahmad & Alwasel, 1999). Thanks to this property, exponential kernels are unambiguously characterized by the half-distance $\alpha\ln(2)$. For any r-value in Equation (3), moving $\alpha\ln(2)$ units further to $r + \alpha\ln(2)$ will decrease the density by half.

Separability. A function is called separable when it can be expressed as a product of other functions that depend on only one independent variable each: the variables can be separated from each other, e.g., $f(x, y) = f_x(x)f_y(y)$. Separable functions are often considered in connection with separable differential equations (Ahmad & Ambrosetti, 2015). When a dispersal kernel is separable, the shape of the kernel along the x-axis does not depend on the y-coordinate, i.e., dispersal probabilities in x- and y-dimensions are independent random variables.

---

strain of *Z. tritici* (ST99CH_3D7). The plots were 1.125 m wide (y-dimension) and 4 m long (x-dimension). Inoculation areas were 40 cm wide in x-dimension and spanned across entire plot in y-dimension (Figure 2a). Disease intensity in the measurement lines was measured as the density of the fungal fruiting bodies (pycnidia) on leaves using automated digital image analysis (Karisto et al., 2018; Stewart et al., 2016).

Using the point source approximation, we computed the disease intensity after dispersal at a distance $r = x$ from the source with exponential kernel (Equation (3)), $k = 1$ as

$$I_1(x) = \frac{I_0\beta}{2\alpha}e^{-|x|/\alpha},\tag{6}$$

where I_0 is the disease intensity at the source before dispersal and β is the transmission parameter comprising the probability of dispersal and the infection efficiency of fungal spores.

Next, we relaxed the point source approximation and used the spatially explicit approach. We computed the expected disease intensity after dispersal at a destination point by substituting the exponential kernel with $k = 2$ into Equation (1). We specified the integrals in Equation (1) according to the plot design across the inoculation area (source; $w_x = 0.4$ m $\times$ $w_y = 1.125$ m) and along the measurement lines considering them as thin lines in y-dimension (destination; $w_d = w_y - 2b_d = 0.875$ m, where $b_d = 0.125$ m is the width of the border excluded from measurements at each end, along y-dimension). Finally, dividing the total intensity by the length of the

(a) Septoria Tritici Blotch

(b) Wheat Stripe Rust
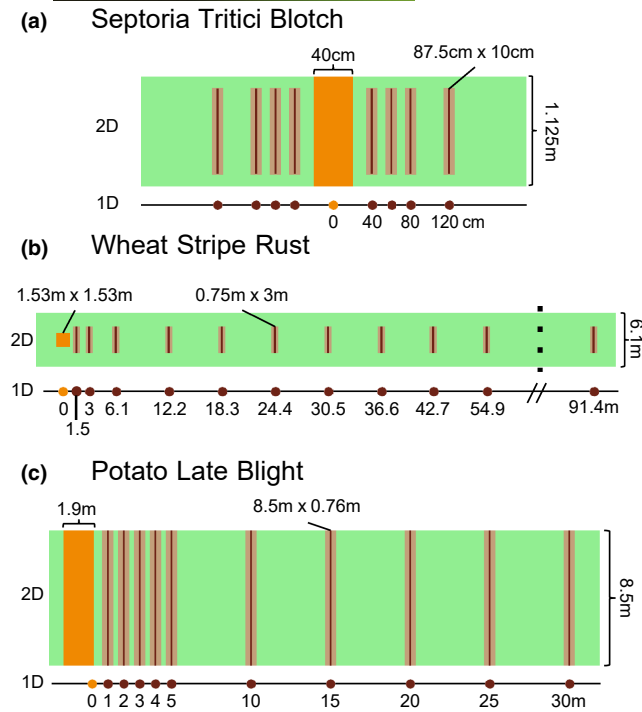
(c) Potato Late Blight

FIGURE 2 Designs of the experimental plots. (a) Septoria tritici blotch experiment (Karisto et al., 2022); (b) yellow rust experiment (Cowger et al., 2005; Sackett & Mundt, 2005); (c) potato late blight experiment (Gregory, 1968). The two-dimensional (2D) view corresponds to the spatially explicit approach; the one-dimensional (1D) view corresponds to the point source approximation. Inoculation areas are shown in orange, measurement lines in light brown and the approximated thin lines in dark brown.

measurement line (w_d) gives the average intensity across the measurement line at a distance x_d as

$$I_1\left(x_d\right) = \frac{I_0 \beta}{w_d 2\pi\alpha^2} \int_{y_d=b_d}^{b_d+w_d} \int_{y_s=0}^{w_y} \int_{x_s=-w_x/2}^{w_x/2} e^{-\sqrt{(x_d-x_s)^2+(y_d-y_s)^2}/\alpha} dx_s\, dy_s\, dy_d, \tag{7}$$

The integrations above incorporate the contribution of each source point $\left(x_s, y_s\right)$ to disease intensity at the destination point (x_d, y_d). In the integrals in Equation (7), we set $x_s = 0$ at the center of the inoculation area and $y_s = 0, y_d = 0$ at the edge of the plot.

We fitted the one-dimensional model Equation (6) and the two-dimensional model Equation (7) to observed dispersal gradients to estimate the scale parameter α.

2.2.2 | Stripe rust

We analyzed a subset of data that corresponds to Hermiston 2002 and Madras 2002 trials (Cowger et al., 2005; Sackett & Mundt, 2005). Asexual spores of *P. striiformis* (urediniospores) were inoculated onto 1.53 m × 1.53 m squares within 6.1 m-wide plots that were at least 100 m long in the downwind direction (Cowger et al., 2005). Disease severity in the measurement lines was measured visually as the percentage of leaf area covered by lesions ("disease severity" is a

specific form of the more general term "disease intensity", Madden et al., 2007).

We used the modified power-law kernel

$$\kappa_{p1}(r) = C_{k,p1}\left(\lambda^2 + r^2\right)^{-\gamma/2}. \tag{8}$$

as defined by Equation (10) of Mikaberidze et al. (2016), because it describes disease gradients of stripe rust better than exponential or Gaussian kernels. Here, r is the distance between the source point and the destination point; $C_{k,p1}$ is the normalization factor, $k = 1, 2$ is the number of dimensions. At $k = 1$, $C_{1,p1} = 2\lambda^{\gamma-1}\Gamma(\gamma/2)/\left[\sqrt{\pi}\Gamma((\gamma-1)/2)\right]$, where $\Gamma(\cdot)$ is the gamma-function; and at $k = 2, C_{2,p1} = (\gamma - 2)/\left(2\pi\lambda^{2-\gamma}\right)$.

The kernel Equation (8) has the same basic properties as the modified power-law kernel in Equation (5) (fat-tailed, power-law). This form was used by Mikaberidze et al. (2016) to analyze the same data and we decided to use the same form to enable an easier comparison. Similarly to Equation (5), γ is the shape parameter, λ is the scale parameter (set to $\lambda = 0.762$ m as in Mikaberidze et al., 2016).

Using Equation (8) and point source approximation ($r = x, k = 1$), we computed the disease severity after dispersal at a distance x from the source

$$I_1(x) = C_{1,p1}I_0\beta\left(\lambda^2 + x^2\right)^{-\gamma/2}. \tag{9}$$

Then, we lifted the point source approximation and we computed the average disease severity in a measurement line at a distance x_d from the middle of the source by substituting the kernel Equation (8) into Equation (1) and specifying the integrals according to the experimental design as

$$I_1(x_d) = \frac{I_0\beta}{w_d}C_{2,p1}\int_{y_d=b_d}^{b_d+w_d}\int_{y_s=b_s}^{b_s+w_x}\int_{x_s=0}^{w_x}\left(\lambda^2 + \left(x_d-x_s\right)^2 + \left(y_d-y_s\right)^2\right)^{-\gamma/2} dx_s\, dy_s\, dy_d, \tag{10}$$

where $w_x = w_y = 1.53$ m is the side length of the square source, $b_s = 2.285$ m is the border area between the inoculation area and the edge of the plot, $w_d = 3$ m is the length of each measurement line (y-dimension), $b_d = 1.525$ m is the width of the border excluded from measurements at each end (Figure 2b).

Following Sackett and Mundt (2005), and Mikaberidze et al. (2016), we performed a natural logarithmic transformation of observed disease gradients to avoid a disproportionate emphasis on the few large values at the beginning of the gradient, and excluded zeros from the log-transformed data. Accordingly, we log-transformed both the one-dimensional model Equation (9) and the two-dimensional model Equation (10). We then fitted both functions to log-transformed disease gradients to estimate the shape parameter γ.

2.2.3 | Potato late blight

We analyzed a subset of data on dispersal of *P. infestans* (Gregory, 1968, Table III, unsprayed experiment). *P. infestans*

zoospores were inoculated across a strip in the middle of each experimental plot (orange area in Figure 2c). Plots were 8.5 m wide and over 60 m long. Rows of potato plants were oriented along the plot length (along the x-axis). Disease severity was measured as numbers of late blight lesions counted visually on all leaves belonging to two adjacent potato stems chosen in every row at several distances from the source: from 1 m up to 30 m.

When describing the experiment, Gregory (1968) did not report the widths of the inoculation areas and the measurement lines (along the x-axis), but only stated that they spanned five plants and two plants along a row, respectively. The paper also did not specify the location of the origin from which the distances between the source and the measurement locations were determined. We inferred these details based on reasonable assumptions. First, we assumed the typical distance between plants along the row to be 0.38 m (15 inches). Based on that, we computed the width of the inoculation area as 1.9 m and the width of a measurement line as 0.76 m. Second, we assumed that the distances were measured from the edge of the inoculation area (Figure 2c), as otherwise the first measurement line would overlap with the inoculation area.

The power-law function

$$\kappa_{p2}(r) = r^{-\gamma} \tag{11}$$

used to describe disease gradients by Gregory (1968) is not a kernel. Nevertheless, we used it in the analysis as if it were a kernel so that the results are comparable with the estimates obtained by Gregory (1968).

Under point source approximation, we used Equation (11) to compute the disease severity after dispersal at a distance $r = x$ from the edge of the source

$$I_1(x) = I_0 \beta x^{-\gamma} \tag{12}$$

Next, we used the spatially explicit approach and computed the disease severity (averaged over the length of a measurement line) at a distance x_d from the source by substituting Equation (11) into Equation (1) and specifying the integrals according to the experimental design

$$I_1(x_d) = \frac{I_0 \beta}{w_d} \int_{y_d=0}^{w_d} \int_{y_s=0}^{w_y} \int_{x_s=-w_x}^{0} \left(\sqrt{(x_d - x_s)^2 + (y_d - y_s)^2} \right)^{-\gamma} dx_s\, dy_s\, dy_d, \tag{13}$$

where $w_x = 1.9$ m and $w_y = 8.5$ m are the dimensions of the inoculation area. Both the inoculation area and the measurement lines spanned the entire width of the plot in y-dimension ($w_y = w_d$). Since the function in Equation (11) cannot be normalized, there is no normalization factor in Equation (12) and Equation (13). For this reason, the prefactor $I_0 \beta$ in Equation (12) and Equation (13) has no biological relevance.

We performed the natural logarithmic transformation of observed disease gradients and excluded zeros from transformed data. Accordingly, we log-transformed both the one-dimensional

Equation (12) and two-dimensional Equation (13). Then, we fitted the two functions to observed disease gradients to estimate the shape parameter γ.

2.2.4 | Data analysis

The fitting was implemented in Python 3.7 using packages numpy (v. 1.17.3, Harris et al., 2020), scipy (v. 1.3.1, Virtanen et al., 2020) and lmfit (v. 1.0.1, Newville et al., 2014). We used least squares optimization via the function 'Model.fit' of the package lmfit (Levenberg–Marquardt or Trust Region Reflective algorithm, function scipy.optimize.least_squares). The same method was used in all analyses of the experimental data and in numerical simulations. Brute force search was used with experimental data to achieve reasonable starting values for the least squares optimization.

We used the estimates of the kernel parameters for septoria tritici blotch and stripe rust to quantify the characteristic scales of dispersal by computing medians (r_{50}) and 90th percentiles (r_{90}) of dispersal distance kernels (Nathan et al., 2012). We computed the two percentiles numerically by solving the equation $2\pi \int_0^{r_L} r \kappa_i(r) dr = 0.01L$ with respect to r_L at $L = 50, 90$. Here, $\kappa_i(r)$ is the dispersal kernel function, where $i = e, p1$; e stands for the exponential kernel (Equation (3) at $k = 2$) and $p1$ stands for the modified power-law kernel (Equation (8) at $k = 2$).

3 | RESULTS

3.1 | Results of data analysis

In all three cases, the spatially explicit estimation (2D-estimation) resulted in steeper dispersal kernels and shorter dispersal distances compared to the point source approximation (1D-estimation), because the estimated kernel parameters differed between 2D- and 1D-estimation (Table 2, Appendix S1: Figure S1.1). α-estimate for septoria tritici blotch was lower by about 12% in 2D-estimation compared to 1D-estimation; γ-estimate for stripe rust was higher by about 10%; γ-estimate for potato late blight was higher by more than 30%.

For septoria tritici blotch and stripe rust, we further investigated how the differences in kernel parameter estimates affect the characteristic spatial scales of dispersal, quantified by the 50th and 90th percentiles of dispersal distance kernels, r_{50} and r_{90}. For septoria tritici blotch, a moderate reduction in the α-estimate in 2D-estimation compared to 1D-estimation translated into a similarly moderate decrease in r_{50} and r_{90} (Table 2). In contrast, for stripe rust, a modest increase in the γ-estimate in 2D- versus 1D-estimation translated into a dramatic decrease in the spatial scales of dispersal (Table 2). In particular, for Madras dataset, 2D-estimation resulted in a nearly two-fold reduction of r_{50} and a massive, almost seven-fold reduction of r_{90}. Although we could not conduct this analysis with the power-law function defined by Equation (11) since it cannot be normalized,

TABLE 2 Comparison of kernel parameter estimates and associated percentiles of dispersal distance kernels between one- and two-dimensional models (1D and 2D, respectively). The 1D-estimates here correspond to the estimates presented in the earlier publications.

		Septoria tritici blotch	Stripe rust, Madras	Stripe rust, Hermiston	Late blight
1D	Parameter	$\alpha = 0.151$ m	$\gamma = 2.4$	$\gamma = 2.5$	$\gamma = 1.4$
	r_{50}, m	0.25	4.2	3.0	–
	r_{90}, m	0.58	241	76	–
2D	Parameter	$\alpha = 0.135$ m	$\gamma = 2.6$	$\gamma = 2.7$	$\gamma = 1.9$
	r_{50}, m	0.23	2.3	1.9	–
	r_{90}, m	0.52	35.3	20.4	–

Note: The parameter γ appears in the exponent of power-law kernels, but the parameter α enters the denominator of the exponent in the exponential kernel. Hence, the parameter difference has the opposite effect on characteristic dispersal distances in septoria tritici blotch compared to the two other systems.

based on the substantial difference in the γ-estimates between 2D- and 1D-estimation, we expect a comparably strong reduction in estimated characteristic spatial scales of dispersal for potato late blight too. Thus, the three plant pathogens disperse over substantially shorter distances according to more realistic 2D-estimation compared to conventional 1D-estimates.

3.2 | Results of numerical simulations

Are the 2D-estimates acquired above more accurate (i.e., closer to the true values) than 1D-estimates? This is plausible, because the 2D-estimation describes dispersal from spatially extended sources more realistically. However, we cannot answer this question definitively based on the analysis of experimental data alone, because we do not know the true values of dispersal kernel parameters. Here, we addressed this question via numerical simulations. We first simulated the dispersal process according to exponential, Gaussian and power-law kernels with pre-defined parameters. Then, we used both methods to estimate the kernel parameters and compared the two methods in terms of their estimation accuracy across a range of biologically plausible scenarios. Here, we summarize the key outcomes of these simulations, but describe them in more detail in Appendices A, B and C.

We started by conducting idealized simulations (Appendix S2): we assumed that sampling locations were points without spatial extent and that measured values accurately reflected true values. Here, the 2D-estimation provided perfectly accurate estimates, while 1D-estimation exhibited substantial errors. We analyzed how the errors in 1D-estimates depend on the parameters of kernel functions and source sizes. We found that the errors become smaller for organisms with longer mean dispersal distances and when using smaller source sizes.

Next, we wanted to understand the origin of errors in 1D-estimates. We considered different parts of the extended source area separately and analyzed how they 'distort' the parameter estimates when assuming a point source (Appendix S3). We found that different parts of the source generate different errors, depending on the kernel type and the location of the virtual point source (used

in 1D-estimate). For power-law kernels, the errors in 1D-estimation were particularly difficult to minimize even when we decreased the source size. None of the three ways we attempted to improve the accuracy of 1D-estimation led to satisfactory outcomes for power-law kernels (denser sampling, sampling further away from the source, or estimation of both λ- and γ-parameters of the kernel). Hence, the 1D-estimation has serious limitations for parameter estimation from extended sources while the 2D-estimation is accurate.

Finally, we investigated both 1D- and 2D-estimation in more realistic simulations where the design mimics more closely experimental design (Appendix S4). We found that sparser sampling and extension of measurement lines reduced the accuracy of both 1D- and 2D-estimates. However, with sufficiently dense sampling the 2D-estimates recovered their accuracy while the 1D-estimates retained substantial errors in all cases.

Thus, we demonstrated that the effect of extended sources on the accuracy of the point source approximation is complex: the magnitude and even the direction of the error induced by the point source approximation depend in a non-trivial manner on the type of the kernel, the scale of dispersal, the spatial configuration of the source and the properties of sampling. Thus, our outcomes discourage the use of the conventional point source approximation, as it provides accurate estimates only in exceptional cases that are hard to identify. Instead, our outcomes strongly support the use the spatially explicit method for estimating dispersal kernel parameters.

4 | DISCUSSION

We devised a theoretical framework to estimate dispersal kernels from empirical dispersal gradients by incorporating the spatial extent of dispersal sources. We re-analyzed existing dispersal gradients for three major plant pathogens and found that this spatially explicit approach provides considerably different estimates of dispersal kernels compared to the conventional point source approximation. Further, we demonstrated using numerical simulations that the spatially explicit approach yields more accurate estimates across a wide range of biologically plausible scenarios. Combining

these two lines of evidence, we conclude that the three organisms disperse on average over substantially shorter distances compared to estimates from conventional modeling.

Similar spatially explicit approaches have been used in modeling studies to investigate dispersal in plants (Clark et al., 1999; Shaw et al., 2006) and plant pathogens (Rimbaud et al., 2018). However, such approaches are not adopted in the literature on empirical characterization of dispersal (e.g., not used in Werth et al., 2006; Skarpaas & Shea, 2007; Loebach & Anderson, 2018; Emsweller et al., 2018; Devaux et al., 2007). Also, Bullock et al. (2017) excluded dispersal gradients produced by line and area sources from their analysis, because these gradients could not be compared to dispersal gradients from point sources. However, dispersal kernels estimated using the spatially explicit approach presented here enable such comparisons, because the estimates are independent of specific experimental design. Hence, adoption of this methodology would provide a unifying framework to extract biological knowledge from experimental data on dispersal.

We demonstrated how to use this theory to extract more knowledge from existing dispersal datasets. The improved estimates can potentially enhance our understanding of ecological dynamics. For example, yellow rust pathogen *P. striiformis* has recently expanded its geographic range by adapting to higher temperatures (Milus et al., 2009). Here, we acquired more accurate estimates of *P. striiformis* dispersal kernels using the spatially explicit approach, whereby the characteristic dispersal distance r_{90} is about seven times shorter compared to conventional estimates (Table 1). Thus, our results (combined with knowledge about other relevant biophysical processes) could enable a more accurate prediction of further range expansion of *P. striiformis* populations, which is likely to be slower than expected based on conventional estimates.

Similarly, using this method, a large proportion of other published dispersal gradients (e.g., Devaux et al., 2007; Emsweller et al., 2018; Loebach & Anderson, 2018; Skarpaas & Shea, 2007; Werth et al., 2006) can be re-analyzed to improve our knowledge about spatial scales of dispersal. This could improve our capacity to predict shifts and expansions of species' geographic ranges, and sizes and compositions of plant communities. We demonstrated that depending on several factors (e.g., the functional form of the kernel and the spatial configuration of source/measurement locations), the improved estimates based on spatially explicit approach can result either in shorter or longer dispersal distances compared to conventional estimates. Accordingly, an improved prediction of rate of range expansions or shifts and sizes or compositions of ecological communities can go in either direction, which highlights the importance of acquiring more accurate estimates of dispersal.

We assumed isotropic dispersal in data analysis and simulations. However, anisotropic dispersal is common in nature (Soubeyrand et al., 2007) and the model can be extended to incorporate it. In this extended model, the probability of dispersal from a source point to a destination point will depend not only on the distance between the points, as in our case, but also on the

direction from the source to the destination. Parameters of anisotropic dispersal kernels can then be estimated from measurements of dispersal gradients with the spatially explicit consideration of the source. Empirical data we analyzed characterized populations of passively dispersing plant pathogens. The methodology is applicable to plant and plant pathogen systems that have passive dispersal, but may not be applicable to characterize active dispersal, e.g., of vector-borne plant viruses.

Based on the outcomes of our idealized simulations, it is tempting to propose simple rules of thumb about when the point source approximation provides reasonably accurate estimates of dispersal kernel parameters. This appears to be the case, for example, when the sources are sufficiently small and the spatial scale of dispersal is sufficiently large (Figure S2.2). However, in these idealized simulations, we neglected the spatial extent of measurement areas and limitations in the amount of sampling within these areas. When we considered these features in more realistic simulations, the outcomes revealed a non-trivial influence of several factors (such as the functional form of the kernel, the spatial configuration of the source and the measurement locations as well as sample sizes) on the estimation accuracy. As a result, we are not able to provide simple rules of thumb regarding the validity of the point source approximation.

Instead, based on our results we suggest the following best practices to design future dispersal experiments. First, a proposed experiment should be simulated numerically over a range of plausible parameter values to decide whether the point source approximation is valid or the spatially explicit modeling should be used in the analysis. Second, aspects of experimental design can be optimized by doing further simulations in order to minimize costs while maximizing the estimation accuracy. These aspects include the size of the source, the spatial configuration of measurement areas (such as their sizes, shapes, and measurement distances), and sample sizes. In conclusion, we demonstrated how spatially explicit modeling can improve the analysis of existing dispersal data and optimize design of future dispersal experiments.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

Data and source code are provided in public repositories. Data: https://datadryad.org, TBA; source code: https://doi.org/10.5281/zenodo.6976556.

ORCID

Petteri Karisto https://orcid.org/0000-0003-4807-0190
Frédéric Suffert https://orcid.org/0000-0001-6969-3878
Alexey Mikaberidze https://orcid.org/0000-0003-2278-2288

REFERENCES

Ahmad, I. A., & Alwasel, I. A. (1999). A goodness-of-fit test for exponentiality based on the memoryless property. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 61(3), 681–689.

Ahmad, S., & Ambrosetti, A. (2015). *A textbook on ordinary differential equations*. Springer-Verlag GmbH.

Bonte, D., & Dahirel, M. (2017). Dispersal: A central and independent trait in life history. *Oikos*, 126(4), 472–479.

Brooker, R. W., Travis, J. M., Clark, E. J., & Dytham, C. (2007). Modelling species' range shifts in a changing climate: The impacts of biotic interactions, dispersal distance and the rate of climate change. *Journal of Theoretical Biology*, 245(1), 59–65.

Bullock, J. M., Mallada González, L., Tamme, R., Götzenberger, L., White, S. M., Pärtel, M., & Hooftman, D. A. (2017). A synthesis of empirical plant dispersal kernels. *Journal of Ecology*, 105(1), 6–19.

Clark, J. S., Silman, M., Kern, R., Macklin, E., & HilleRisLambers, J. (1999). Seed dispersal near and far: Patterns across temperate and tropical forests. *Ecology*, 80(5), 1475–1494.

Cousens, R., & Rawlinson, A. (2001). When will plant morphology affect the shape of a seed dispersal "kernel"? *Journal of Theoretical Biology*, 211(3), 229–238.

Cowger, C., Wallace, L. D., & Mundt, C. C. (2005). Velocity of spread of wheat stripe rust epidemics. *Phytopathology*, 95(9), 972–982.

Devaux, C., Lavigne, C., Austerlitz, F., & Klein, E. (2007). Modelling and estimating pollen movement in oilseed rape (Brassica napus) at the landscape scale using genetic markers. *Molecular Ecology*, 16(3), 487–499.

Emsweller, L. N., Gorchov, D. L., Zhang, Q., Driscoll, A. G., & Hughes, M. R. (2018). Seed rain and disturbance impact recruitment of invasive plants in upland forest. *Invasive Plant Science and Management*, 11(2), 69–81.

Fabre, F., Coville, J., & Cunniffe, N. J. (2021). Optimising reactive disease management using spatially explicit models at the landscape scale. In P. Scott, R. Strange, L. Korsten, & M. L. Gullino (Eds.), *Plant diseases and food security in the 21st century* (pp. 47–72). Springer.

Fabre, F., Rousseau, E., Mailleret, L., & Moury, B. (2012). Durable strategies to deploy plant resistance in agricultural landscapes. *New Phytologist*, 193, 1064–1075.

Ferrandino, F. (1996). Length scale of disease spread: Fact or artifact of experimental geometry. *Phytopathology*, 86, 806.

Greene, D. F., & Calogeropoulos, C. (2002). Measuring and modelling seed dispersal of terrestrial plants. In J. Bullock, R. Kenward, & R. Hails (Eds.), *Dispersal ecology: The 42nd Symposium of the British Ecological Society* (pp. 3–23). Blackwell.

Gregory, P. (1968). Interpreting plant disease dispersal gradients. *Annual Review of Phytopathology*, 6, 189–212.

Harris, C. R., Millman, K. J., Walt, S. J. v. d., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. v., Brett, M., Haldane, A., Río, J. F. d., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362.

Heald, F. (1913). The address of the president for 1912: The dissemination of fungi causing disease. *Transactions of the American Microscopical Society*, 32(1), 5–29.

Karisto, P., Hund, A., Yu, K., Anderegg, J., Walter, A., Mascher, F., McDonald, B. A., & Mikaberidze, A. (2018). Ranking quantitative resistance to Septoria tritici blotch in elite wheat cultivars using automated image analysis. *Phytopathology*, 108, 568–581.

Karisto, P., Suffert, F., & Mikaberidze, A. (2022). Measuring splash dispersal of a major wheat pathogen in the field. *Phytofrontiers™*, 2(1), 30–40.

Klein, E. K., Bontemps, A., & Oddou-Muratorio, S. (2013). Seed dispersal kernels estimated from genotypes of established seedlings: Does density-dependent mortality matter? *Methods in Ecology and Evolution*, 4(11), 1059–1069.

Kubisch, A., Degen, T., Hovestadt, T., & Poethke, H. J. (2013). Predicting range shifts under global change: The balance between local adaptation and dispersal. *Ecography*, 36(8), 873–882.

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Holt, R. D., Shurin, J. B., Law, R., Tilman, D., Loreau, M., & Gonzalez, A. (2004). The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters*, 7(7), 601–613.

Loebach, C. A., & Anderson, R. C. (2018). Measuring short distance dispersal of alliaria petiolata and determining potential long distance dispersal mechanisms. *PeerJ*, 6, e4477.

Madden, L. V., Hughes, G., & Bosch, F. v. d. (2007). *The study of plant disease epidemics*. American Phytopathological Society (APS Press).

McConkey, K. R., Prasad, S., Corlett, R. T., Campos-Arceiz, A., Brodie, J. F., Rogers, H., & Santamaria, L. (2012). Seed dispersal in changing landscapes. *Biological Conservation*, 146(1), 1–13.

Mikaberidze, A., Mundt, C. C., & Bonhoeffer, S. (2016). Invasiveness of plant pathogens depends on the spatial scale of host distribution. *Ecological Applications*, 26, 1238–1248.

Milus, E., Kristensen, K., & Hovmoller, M. (2009). Evidence for increased aggressiveness in a recent widespread strain of Puccinia striiformis f. sp. tritici causing stripe rust of wheat. *Phytopathology*, 99(1), 89–94.

Musciano, M. D., Cecco, V. D., Bartolucci, F., Conti, F., Frattaroli, A. R., & Martino, L. D. (2020). Dispersal ability of threatened species affects future distributions. *Plant Ecology*, 221(4), 265–281.

Nathan, R., Klein, E., Robledo-Arnuncio, J. J., & Revilla, E. (2012). Dispersal kernels: Review. In J. Clobert, M. Baguette, T. G. Benton, & J. M. Bullock (Eds.), *Dispersal ecology and evolution* (Vol. 15, p. 187). Oxford University Press.

Newville, M., Stensitzki, T., Allen, D. B., & Ingargiola, A. (2014). Lmfit: Non-linear leastsquare minimization and curve-fitting for python. *Zenodo*, 598352.

Petit, S., Alignier, A., Colbach, N., Joannon, A., Le C÷ur, D., & Thenail, C. (2013). Weed dispersal by farming at various spatial scales. A review. *Agronomy for Sustainable Development*, 33(1), 205–217.

Rimbaud, L., Papaïx, J., Rey, J.-F., Barrett, L. G., & Thrall, P. H. (2018). Assessing the durability and efficiency of landscape-based strategies to deploy plant resistance to pathogens. *PLoS Computational Biology*, 14(4), e1006067.

Ron, R., Fragman-Sapir, O., & Kadmon, R. (2018). Dispersal increases ecological selection by increasing effective community size. *Proceedings of the National Academy of Sciences*, 115(44), 11280–11285.

Sackett, K. E., & Mundt, C. C. (2005). Primary disease gradients of wheat stripe rust in large field plots. *Phytopathology*, 95(9), 983–991.

Saint-Jean, S., Chelle, M., & Huber, L. (2004). Modelling water transfer by rain-splash in a 3d canopy using Monte Carlo integration. *Agricultural and Forest Meteorology*, 121(3), 183–196.

Shaw, M., Harwood, T., Wilkinson, M., & Elliott, L. (2006). Assembling spatially explicit landscape models of pollen and spore dispersal by wind for risk assessment. *Proceedings of the Royal Society B: Biological Sciences*, 273(1594), 1705–1713.

Skarpaas, O., & Shea, K. (2007). Dispersal patterns, dispersal mechanisms, and invasion wave speeds for invasive thistles. *The American Naturalist*, 170(3), 421–430.

Soubeyrand, S., Enjalbert, J., Sanchez, A., & Sache, I. (2007). Anisotropy, in density and in distance, of the dispersal of yellow rust of wheat: Experiments in large field plots and estimation. *Phytopathology*, 97(10), 1315–1324.

Stewart, E. L., Hagerty, C. H., Mikaberidze, A., Mundt, C., Zhong, Z., & McDonald, B. A. (2016). An improved method for measuring quantitative resistance to the wheat pathogen Zymoseptoria tritici using high throughput automated image analysis. *Phytopathology*, 106, 782–788.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... Vázquez-Baeza, Y. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, 17(3), 261–272.

Werth, S., Wagner, H. H., Gugerli, F., Holderegger, R., Csencsics, D., Kalwij, J. M., & Scheidegger, C. (2006). Quantifying dispersal and establishment limitation in a population of an epiphytic lichen. *Ecology, 87*(8), 2037–2046.

Zadoks, J. C., & Schein, R. D. (1979). *Epidemiology and plant disease management*. Oxford University Press Inc.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.