# Spatially-explicit modeling improves empirical characterization of dispersal

Petteri Karisto[1†], Frédéric Suffert[2], and Alexey Mikaberidze[1‡]

Article

[1] Plant Pathology Group, Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland.
[2] INRA, Agro Paris Tech, UMR 1290 BIOGER-CPP, Thiverval-Grignon, France.
[†] present address: Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland.
[‡] present address: School of Agriculture, Policy and Development, University of Reading, Whiteknights, Reading, UK.

Corresponding author: P. Karisto
Email address: petteri.karisto@ieu.uzh.ch

Keywords: Dispersal kernel, Dispersal gradient, Point source, Area source, Line source

Running title: Measuring dispersal kernel

Content:

## Abstract

Dispersal is a key ecological process. An individual dispersal event has a source and a destination, both are well localized in space and can be seen as points. A probability to move from a source point to a destination point can be described by a specific probability function, the dispersal kernel. However, when we measure dispersal, the source of dispersing individuals is usually an area, which distorts the shape of the observed dispersal gradient compared to the underlying dispersal kernel. Here, we show with simulations, how different source geometries affect the gradient shape depending on the type of the kernel. We present an explicit mathematical approach for estimating the dispersal kernel from a dispersal gradient data independently of the source dimension. Further, we demonstrate the value of the approach by analysing three experimental dispersal datasets with a conventional method and the proposed method, to show how the estimated dispersal kernels differ between the methods. We use three pre-existing datasets from field experiments measuring dispersal of important plant pathogens. Our results demonstrate how analysis of dispersal data can be improved to achieve more rigorous measures of dispersal. The proposed approach leads to a general measure of dispersal in contrast to results from the conventional method that depend on the design of the dispersal source. This enables a direct comparison between outcomes of different experiments and allows acquiring more knowledge from a large number of previous empirical studies of dispersal.

# Introduction

Individual organisms commonly need to move from one location to another in order to survive and reproduce. Hence, dispersal is an important component of many life histories. Empirical characterization of dispersal has been a major theme in ecological research for a long time (for example Heald, 1913; Bullock et al., 2017). However, Bullock et al. (2017) found much fewer datasets describing plant dispersal than plant demography, indicating that "dispersal is notoriously difficult and resource-consuming to measure".

To measure dispersal, one needs a source of dispersing units and a method to record their displacement. Sources can be natural (e.g. a spawning site of fish) or artificial (a planted patch). To record the displacement, studies on animals often use mark-recapture experiments (Van Houtan et al., 2007; Carrasco et al., 2010), while plant studies commonly use seed traps or genotyping of seedlings around potential parents (Nathan et al., 2000; Goto et al., 2006). Spread of a plant pathogen can be recorded based on visual symptoms and genetic data (Solheim and Hietala, 2017). The appropriate methodology varies depending on the study system.

With an established localized source, a dispersal gradient is expected: normally many individuals will stay close to the source while fewer individuals will travel further, leading to a decreasing trend over distance. This pattern can be described with an appropriate mathematical function, for example a decreasing one-dimensional function (e.g. review of Fitt et al., 1987; Ferrandino, 1996; Werth et al., 2006; Madden et al., 2007). However, the best fitting shape of this function depends among other things, on the dimension of the source (Zadoks and Schein, 1979; Ferrandino, 1996; Cousens and Rawlinson, 2001). Some degree of "flattening" of the gradient is expected with an extended source compared to a point source (Zadoks and Schein, 1979; Ferrandino, 1996). This was noted qualitatively in the previous studies, but that does not answer in what way and how much the source

properties actually affect the observed dispersal gradient.

A more rigorous mathematical description of dispersal is achieved with a dispersal kernel: a probability density function of dispersal to a certain location relative to the source ("dispersal location kernel", Nathan et al., 2012). Hence, dispersal kernel describes a point-to-point dispersal. For characterizing a dispersal kernel experimentally, it would be convenient to have a point source, because in that case the observed dispersal gradient had the same shape as the kernel. Unfortunately, a point source is an impossible requirement in reality. The source have to have certain area to yield any dispersing propagules, and usually the area should be considerably large to secure an observable amount of dispersal events for determining the shape of the gradient.

As having a point source in impossible, the "second best guess" is to use sufficiently small area that can be considered a point. Zadoks and Schein (1979) proposed a rule of thumb, stating that a point source should have "a diameter smaller than 1% of the gradient length; but in many experiments, it is up to 5 or 10%". However, to determine whether the source is small enough so that the dispersal gradient captures the shape of the dispersal kernel, the size of the source should be compared with the characteristic distance of dispersal (i.e., the distance over which the dispersal kernel changes substantially), rather than the gradient length (i.e. extent of measurements). This represents a challenge for the design of dispersal experiments that aim to achieve a point-like source, because whether or not the chosen source size is sufficiently small can be established with certainty only when the measurements are already conducted. As a result of the ambiguity, "point" sources of various sizes are found in literature: an adult tree (Werth et al. (2006); cf. Cousens and Rawlinson (2001) presenting effect of tree canopy morphology on the shape of the gradient), circles of 80 cm (Skarpaas and Shea, 2007) and 25 cm diameter (Loebach and Anderson, 2018), 4 $m^2$ square (Emsweller et al., 2018), route of a single sampling dive (D'Aloia et al., 2015), even an agricultural field as a

whole (Devaux et al., 2006).

This challenge can be resolved using a modeling approach that incorporates the spread from any source considering each point within the source area as an independent point source (Clark et al., 1999; Shaw et al., 2006). This would lead to a better, more mechanistic understanding of the dispersal as recommended by Bullock et al. (2006). While such approach has been suggested (e.g. by Greene and Calogeropoulos, 2002) it is adopted only rarely, as demonstrated by the previous examples of various "point" sources and, for example, Bullock et al. (2017) who excluded line and area sources from their analysis, because those could not be compared to gradients from point sources.

We demonstrate the effect of the spatial extension of the source (i.e. deviations from a point source) on the shape of dispersal gradients considering three different dispersal kernels: exponential, Gaussian, and power-law. The first two were chosen because of their particular mathematical properties and they correspond to constant speed of invasion front (Kot et al., 1996). The latter one is a fat-tailed kernel, which allows for more long-range dispersal resulting in accelerating invasion front, which might be biologically more realistic in many systems (reviewed by Nathan et al., 2012). We present a method for estimating dispersal kernel accurately from disperal gradient data by considering the dimensions of the source explicitly.

Finally, we demonstrate the improvement achieved with this spatially-explicit approach compared to conventional approach, which assumes a point source. For that demonstration, we use three datasets from published dispersal experiments. The datasets consider dispersal of important plant pathogens: *Phytophthora infestans* causing Potato late blight, *Puccinia striiformis* causing Yellow rust of wheat and *Zymoseptoria tritici* causing Septoria tritici blotch of wheat. These pathogens are respectively the 1st, 4th and 3rd most important pests of their hosts, according to a recent global assessment (Savary et al., 2019). Dispersal of important plant pathogens has been studied for

decades due to its practical importance and also convenience of the experimentation, leading to good availability of data in those systems. Our main result is that almost invariably, any deviation from a point source leads to distortion of the dispersal gradient. Hence, dispersal gradients in different experiments are not comparable unless the dispersal kernel is characterized with spatially-explicit consideration of the source.

## Theory

*Dispersal location kernel* (hereafter "dispersal kernel") describes the probability of dispersal from a source point $p_s = (x_s, y_s)$ to a destination point $p_d = (x_d, y_d)$ depending on the distance between the points $r(p_s, p_d) = \sqrt{(x_d - x_s)^2 + (y_d - y_s)^2}$ (cf. *dispersal distance kernel*, Appendix A). To parametrize the dispersal kernel with dispersal data from an area source, we need to consider the cumulative effect of all point-to-point dispersal processes from the source to the destination. This is achieved by taking sum over the individual points comprising the source to calculate their combined contribution to the dispersed population at certain destination point (as in Shaw et al., 2006, Eq. (4.6)). Similarly, the sum over all points of a destination area gives the total population in the area. Population in the destination area after dispersal is then calculated as

$$N_1(S, D) = \iint_D \iint_S N_0(p_s)\kappa(p_s, p_d)\, dA_S\, dA_D \tag{1}$$

where $S = \{p_s\}$ is the source area, $D = \{p_d\}$ is the destination area, $N_0(p_s)$ gives the density of the successfully dispersing individuals within $S$ before dispersal and $\kappa(p_s, p_d)$ is the dispersal kernel. The area integrals sum up the contributions of all source points in $S$ to all destination points in $D$ to get the total observed population in $D$. When the populations before dispersal ($N_0$) and after dispersal ($N_1$) are measured, the only unknown in Eq. (1) is the dispersal kernel. Kernel parameters can be estimated by fitting

this function to the observed data. Eq. (1) gives the general spatially-explicit solution for estimating dispersal kernel parameters while considering sources and destinations as areas.

Fitting a model with the above structure to empirical data can be challenging, although numerical methods have become very efficient recently. Analytical solutions are difficult to achieve with complex formulae and multiple integrations increase the computational demand making the process slower. Therefore, simplifications would be useful for allowing analytical handling of the model and enhancing data analysis. As a simple example, before availability of powerful numerical solvers, fitting a function to data was most practical when the function could be linearized and reduced to one dimension to benefit from simple linear fitting.

One common simplification is to fit a one-dimensional model to dispersal gradient data while considering the source and receivers as points. This removes the need for integrations and gives a general model

$$N_1 = C\kappa \tag{2}$$

. For example, a function of the form

$$N_1 = Ce^{-x/\alpha} \tag{3}$$

can be used to estimate the dispersal parameter $\alpha$ in the case of an exponential kernel (for example Saint-Jean et al., 2004). Note that the parameter $C$ does not have a clear biological interpretation here, in contrast to $N_0$ in Eq. (1). If both the source and the destination are points, the above approach provides a correct estimate for the shape parameter $\alpha$, because the function in Eq. (3) is the same as the exponential dispersal kernel [Eq. (4) in Box 1] up to a constant factor. This approach works for any kernel

**Exponential kernel** is defined as

$$\kappa_e(r, \alpha) = C_k e^{-r/\alpha} \tag{4}$$

where $k \in \{1, 2\}$ is the number of dimensions, $r = r(p_s, p_d) > 0$ 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 $C_k$ is normalization factor: $C_1 = 1/(2\alpha)$ and $C_2 = 1/(2\pi\alpha^2)$.

**Gaussian kernel** is defined as

$$\kappa_g(r, \alpha) = C_k e^{-r^2/2\alpha^2}, \tag{5}$$

where $C_1 = 1/\sqrt{2\pi\alpha^2}$ and $C_2 = 1/(2\pi\alpha^2)$.

**Power-law kernel** is defined here as

$$\kappa_p(r, \gamma, \lambda) = C_k(\lambda + r)^{-\gamma}, \tag{6}$$

where $C_1 = (\gamma - 1)\lambda^{\gamma-1}$, $C_2 = (\gamma - 2)(\gamma - 1)\lambda^{\gamma-2}/(2\pi)$. $\lambda$ is a scale parameter defining the finite starting point of the distribution in relation to $r^{-\gamma}$ distribution, which is not defined at $r = 0$.

Box 1

function (e.g., Gaussian or power-law kernels), if both the source and the destination are points.

However, when the dimension of the source and/or destination is more complex, the above approach may lead to wrong estimates. The parameter $\alpha$ estimated with this approach, may depend on the particular experimental design and have no direct relation to the actual kernel shape, as demonstrated in Fig. 1. However, in certain special cases the shape of the dispersal gradient does match to the shape of the dispersal kernel even when the source and the destination are extended. Next, we discuss these special cases in relation to exponential, Gaussian, and power-law kernels (defined in Box 1).

If the source is extended in the direction of the measured gradient, and the underlying kernel is exponential, the Eq. (3) will still give a correct estimate of $\alpha$. This holds, because exponential kernels are memoryless (Box 2). This property allows to sum up all point sources within the source area along the $x$-axis to an equivalent virtual point source at $x = 0$ and in this way simplify the fitting process (see Fig. 1B). Thus, the

**Memoryless kernel.** Exponential kernels have a special feature: they are memoryless. To be memoryless means that setting any point along the gradient as a starting point, the tail of the distribution will have the same shape as entire distribution. This property explains why exponential kernel can be described unambiguously with the half-distance $\alpha \ln(2)$. From any point on an exponential gradient, moving $\alpha \ln(2)$ further along the gradient will decrease the density by half.

**Separable kernel.** Separable functions are those that can be expressed as a product of functions which depend on only one independent variable each, e.g. $f(x, y) = f_x(x)f_y(y)$. The shape of the dispersal gradient in the $x$-direction does not depend on the $y$-coordinate if the kernel is separable.

Most dispersal kernels found in the literature are neither memoryless nor separable (Nathan et al., 2012).

Box 2

extension of the source in the direction of the gradient will only add more power to the source but not change the shape of the gradient outside of the source, leading to a correct estimate of $\alpha$. This is not true for Gaussian and power-law kernels (Fig. 1 C, D).

If the extension of the source is in the other direction, perpendicular to the source, the simplified approach works with Gaussian kernel (Fig. 1C). Gaussian kernel is separable, which means that the shape of the kernel along $x$-dimension does not change when $y_s$ varies (Box 2). Hence, when measuring the dispersal along the $x$-axis, the extension of the source along the $y$-axis only adds to the power of the source but does not modify the shape of the gradient. Thus changing the source from a point to a thin line source perpendicular to the gradient leads to a different estimate of $C$ but the same estimate of $\alpha$. This holds for any separable kernel, but not for non-separable exponential or power-law kernels (Fig. 1B, D).

The situation is analogous when we consider extended destinations. Extended destination here implies that multiple measurements are conducted across the destination area in a uniformly random manner, and subsequently an average is taken over these measurements. When the kernel is exponential, both the source and the destination

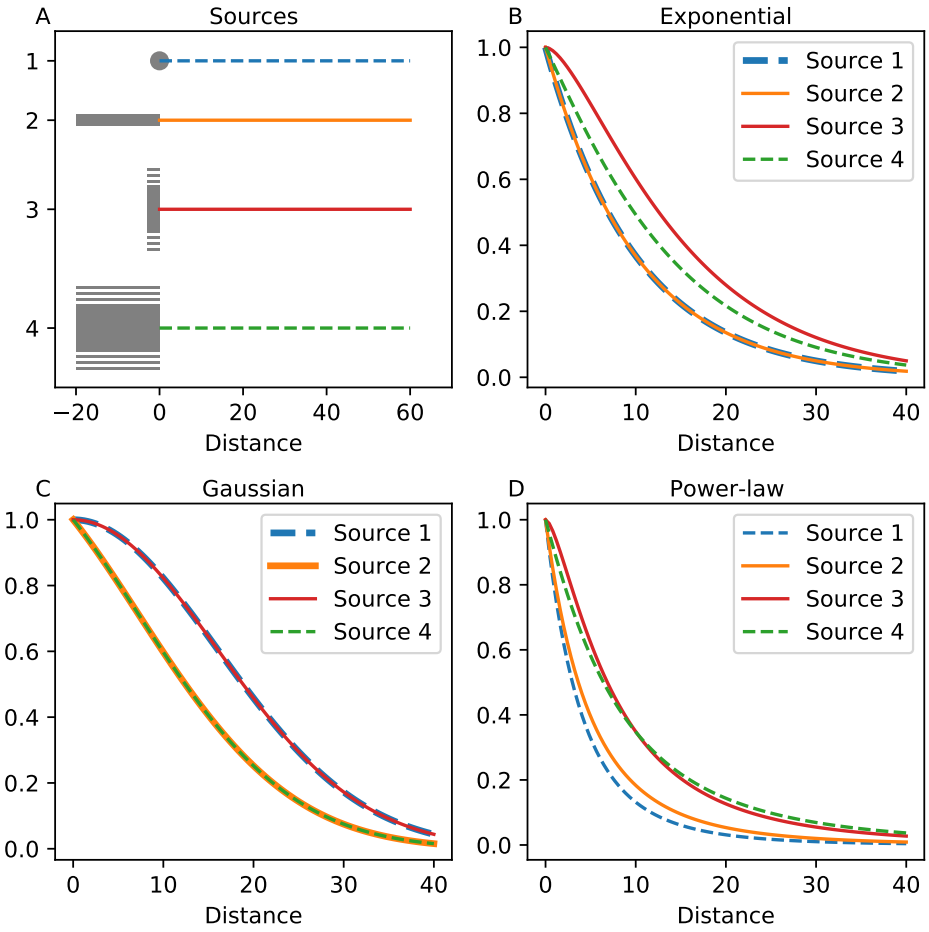


Figure 1: Different extensions of the source (panel A) lead to different effects on the dispersal gradients (B, C, D) depending on the dispersal kernel. Kernel parameters are chosen so that the mean dispersal distance is $\approx 20$ (arbitrary units) in all cases. All gradients are normalized to begin at one. (A) Four different sources: 1) a point source; 2) a line source, parallel to the gradient $x \in [-20, 0]$; 3) a line source perpendicular to the gradient, $y \in [-100, 100]$; 4) a rectangular area source, $(x, y) \in [-20, 0] \times [-100, 100]$. (B) With the exponential kernel ($\alpha = 10$), sources 1 and 2 result in identical gradients. (C) With the Gaussian kernel ($\alpha = 16$) the gradients are identical between sources 1 and 3 and between sources 2 and 4. (D) With the power-law kernel ($\gamma = 5, \lambda = 20$) all gradients are different.

194 can be elongated in the direction of the dispersal gradient. An exponential function in Eq. (3) fitted to dispersal gradient data acquired in this way will have same dispersal 196 parameter $\alpha$ as the dispersal kernel. In the case of separable kernels, both the source and the destination can be elongated perpendicular to the gradient and the gradient will 198 have the same shape as the original kernel. If the spatial configuration of the source or the destination is more complex, for example rectangles, the presented simplifications 200 fail with each of the three kernels.

The simulations in Fig. 1 provide a first glance to the effect of the source shape on 202 the dispersal gradient. Clearly, the gradients' shape can change depending on the source type and the underlying dispersal kernel. Next we demonstrate the effect with simulated 204 and experimental datasets, which we analyse either assuming a point source or using the spatially-explicit approach.

# 206 Simulations

In this section we present simulations to demonstrate the effect of source extensions 208 compared to a point-source. First we show as an example that accounting for the source extensions improves accuracy of the estimated kernel parameters. Then we present 210 analysis of magnitude of the estimate error with different dispersal distances and source sizes. At the end, we take a more detailed look at different parts of the extended source 212 area to see in which direction they will distort the parameter estimate if a point-source is assumed.

## 214 Re-estimation of dispersal parameters from simulated data.

We modelled spread from two different sources with three different two-dimensional dis- 216 persal kernels. The kernels and the parameters are as in Fig. 1: exponential, Gaussian and power-law kernels with mean dispersal distance of 20 cm. The two sources were a

10 cm × 10 cm square and a 20 cm diameter circle. The spread was modelled deterministically to 40 cm long measurement lines perpendicular to the gradient at distances 10, 20, 30, 50, 70, 100 and 150 cm (x-dimension) from the closer edge of the source. The measurement lines were sampled at five locations evenly distributed in the y-dimension. That resulted in five deterministic data points at each distance in x-dimension. Then, we calculated the mean over each measurement line for each replicate and used that dataset for fitting.

The simplified fitting was performed using one-dimensional kernels and assuming a point source in the middle of the source area. Spatially explicit fitting was performed with two-dimensional kernels with integration over the source area, and over the measurement lines as if they were sampled densely. In all cases of two different sources and three different kernels, the spatially explicit method resulted in a better estimate of the kernel parameter, although not always perfect due to the difference in the "collection" and analysis of the data in the measurement lines (five sampling points vs. continuous, dense sampling) (Table 1).

Table 1: Re-estimated parameter values from simulated data. The spatially-explicit approach ("2D") always gives better estimates.

| Source | Kernel | Parameter | 1D | 2D | True |
|---|---|---|---|---|---|
| Square | Exponential | $\alpha$ | 11.34 | 10.26 | 10.0 |
| | Gaussian | $\alpha$ | 16.25 | 16.00 | 16.0 |
| | Power-law | $\gamma$ | 4.27 | 4.83 | 5.0 |
| Circle | Exponential | $\alpha$ | 11.11 | 10.21 | 10.0 |
| | Gaussian | $\alpha$ | 16.77 | 16.00 | 16.0 |
| | Power-law | $\gamma$ | 4.54 | 4.85 | 5.0 |

Note: “2D” and “1D” stand for two- and one-dimensional models, respectively. For power-law, fixed $\lambda = 20$ in all cases to only estimate $\gamma$.

## Effect of dispersal distance and source size on the error

Next we investigated the magnitude of the error induced by the source extension as a function of dispersal distance ("gradient length") and as a function of the source size. We simulated dispersal from a square source onto a thin line using a two-dimensional dispersal kernel and explicit source area, and then we re-estimated the shape parameter of the kernel assuming that the source was a point. The virtual point-source was assumed to be located either in the middle of the original square source (or at the edge of the source, Appendix B). We expected the estimate error to be in the direction of longer estimated dispersal when assuming a point source, because dispersal kernels are usually steep in the beginning and flat in the tail. Extensions of the would source lead to more contributions from the tail part of the kernel, because most points in the extended area source are further away from the gradient than the assumed point source. Additionally, perpendicular offset of the source points from the measured gradient line can lead to flatter gradient as the gradient resulting from that offset-point is not sampled along the steepest gradient directly away from the point but in an angle. Hence, the observed gradient would be a result of combined contributions of flatter gradients from many points in within the area source. When an extended source is considered as a point, the flattening effect of the source extension would be incorrectly accounted for in the "flatter" estimates of the kernel. However, while this reasoning seems to hold for exponential kernel it does not necessarily hold for Gaussian and power-law kernels.

For the analysis of the effect dispersal distance, a square source of size $1\,\mathrm{m} \times 1\,\mathrm{m}$ was used and the mean dispersal distance was varied from 1 m to 50 m (1 m resolution). For studying effect of source size, the mean dispersal distance was fixed to 20 m and the square source size was varied from $1\,\mathrm{m} \times 1\,\mathrm{m}$ to $30\,\mathrm{m} \times 30\,\mathrm{m}$ (1 m resolution). The mean dispersal distance is $2\alpha$ for exponential kernel, $\alpha\sqrt{\pi/2}$ for Gaussian kernel and $2\lambda/(\gamma - 3)$ for power-law kernel.

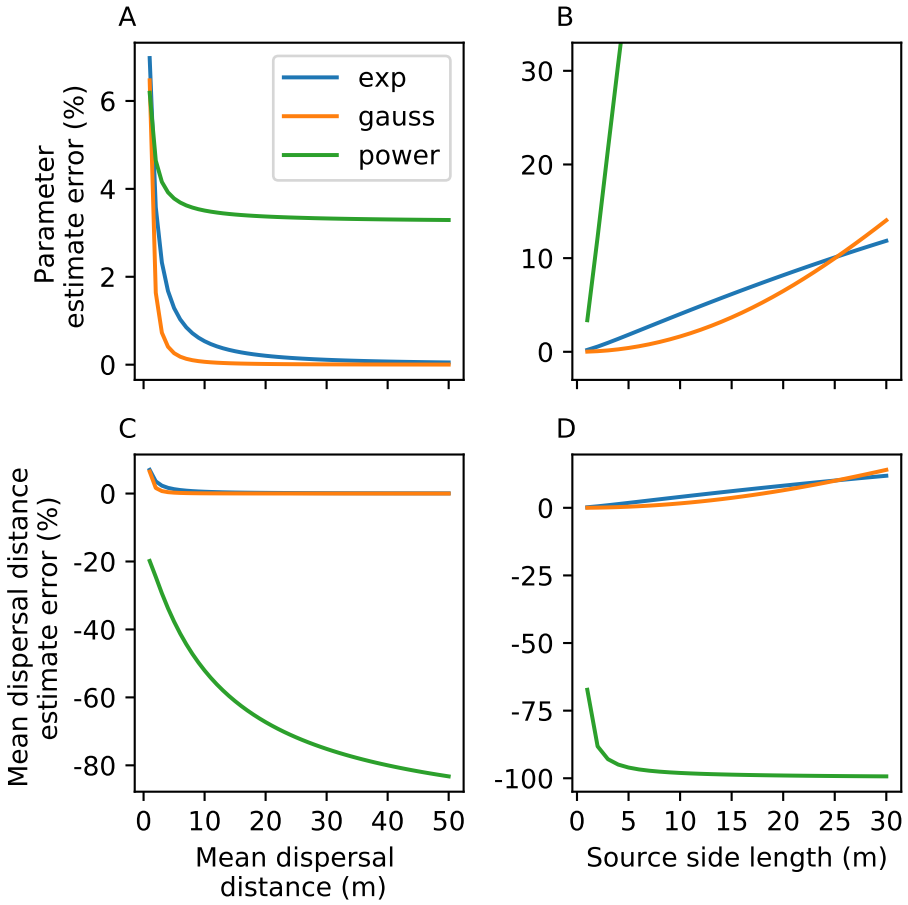


Figure 2: Effect of real mean dispersal distance (left column) and source size (right column) on the accuracy of the point source estimate when assuming a point in the middle of the square source. Errors in the parameter estimate (upper row) and in the estimated mean dispersal (lower row) as a function of real mean dispersal distance and source size. Parameter estimates improve when dispersal distance is increase or the source size is decreased.

In both simulations, the gradient was sampled starting from 10 cm away from the edge of the source with 10 cm spacing until reaching a point where the measured value (number of dispersed individuals) was less than 1% of that of the first measurement point. The dispersal parameter was then re-estimated from the simulated data using a one-dimensional dispersal kernel assuming a point-source either in the middle of the source. Parameter $\lambda$ of power-law kernel was fixed to 0.5 in simulation and in fitting to only estimate the shape parameter $\gamma$ (in Appendix B we show that re-estimating also $\lambda$ improves the fit but does not improve accuracy of the estimated dispersal distance). Errors in the estimate of shape parameter and in the corresponding estimated mean dispersal distance are presented as functions of real mean dispersal distance (Fig. 2A, C) and as functions of the source size for the latter (Fig. 2B, D).

As seen in Figure 2, the errors in the parameter estimates decrease with increasing dispersal distance when the source extension becomes relatively smaller, as expected. However, the decreased error in parameter estimate does not translate into more precise estimate of dispersal distance in the case of power-law kernel, instead the error increases. That seemingly counter-intuitive relationship follows from the formula of mean dispersal distance of the power-law kernel ($2\lambda/(\gamma-3)$). With longer dispersal, the real value of $\gamma$ gets closer to the critical value of three, and hence the mean dispersal distance becomes more sensitive to errors in the parameter estimate. Analogously to varying dispersal distance, when increasing the source size, the errors in the parameter estimates and mean dispersal distance increase with all three kernels. The power-law kernel seems much more sensitive, reaching 230% overestimation of the shape parameter with 30 m-sized source, while for exponential and Gaussian kernels the error is only around 10%.

As a summary, the dispersal distance was overestimated with the point-source approximation when using exponential and Gaussian kernels, but underestimated with the power-law kernel. However, if assuming a point source at the edge of the source, the

situation changes drastically. In Appendix B we show how that assumption leads to underestimation of the dispersal distance with Gaussian kernel but overestimation with power-law kernel. As a conclusion from these simulations, we note that the errors are not consistent but their magnitude and even their sign depend on the type of the kernel and on the location of the imaginary point-source. We also suggest, that different source shapes may complicate the situation even further. Hence, we cannot give any useful rule of thumb for when an area source could be considered as a point. Instead, each individual situation with its particular (experimental) design and dispersal distance should be resolved individually by simulations like we the ones presented here.

## Which parts of the source cause the distortion

To understand the observed patterns better, we studied next how each part of an extended source contributes to the distortion of the estimated dispersal. For this purpose, we simulated dispersal from a 1 m × 1 m source with 2 m mean dispersal distance ($\lambda = 0.5$ for power-law), recording the gradient again from 10 cm outside of the source with 10 cm spacing to six meters further, similar to the previous simulations. However, we did not record to cumulative dispersal gradient from entire source but instead we created a separate dispersal gradient from each individual point over the source area (2 cm-by-2 cm grid). From each of these gradients originating from different points, we re-estimated the dispersal distance while assuming a point-source in the middle of the source and we recorded whether the estimated mean dispersal distance was overestimated or underestimated. The results are presented in Figure 3.

The exponential kernel follows a simple pattern where points along the extension of the measured gradient give precise estimate, whereas other points result in flatter gradient and consequently overestimated dispersal distance. This follows from the memorylessness of the kernel. The Gaussian kernel has similarly simple pattern, where a line

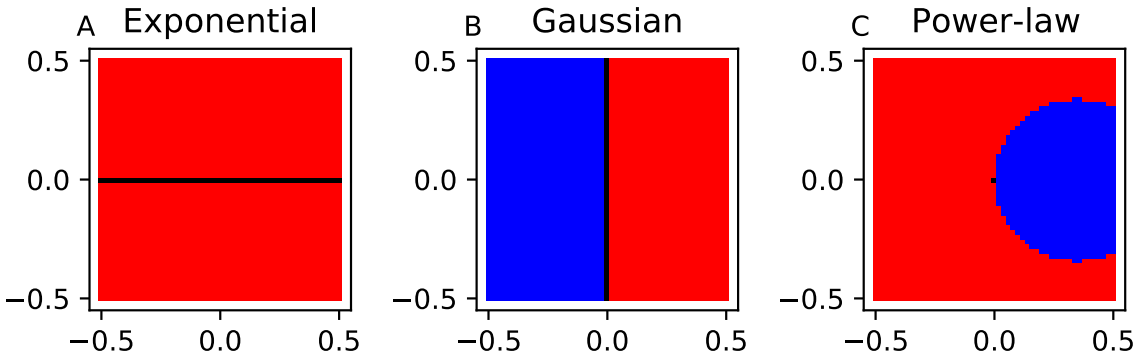


Figure 3: Contribution of individual points within an extended source to the estimated mean dispersal distance. Dispersal distance re-estimated assuming a point-source in the middle of the actual source. For each point in the source (2 cm × 2 cm grid) the color shows, whether that point contributes to underestimation (blue) or overestimation (red) of the dispersal distance. Black: precise estimate. Gradient measured to right from the source.

perpendicular to the gradient through the center of the source (imaginary source point) gives precise estimation due to separability, the points towards the gradient result in overestimation and points away from gradient result in underestimation of the dispersal distance. The power-law kernel creates more complex pattern where points that are close to the beginning of the gradient in a relatively circular area cause underestimation of the dispersal distance but the rest of the source create overestimation. Note that the points that are closer to the measured gradient (right side in the figure) impose higher dispersal pressure onto the gradient and thus have relatively stronger contribution on the measured gradient than points that are further away from the gradient. Finally, if the point source was assumed at the edge of the source (as in Appendix B), the source area to the right of the imaginary point-source would be excluded, which explains why the estimate error changes sign for Gaussian and power-law kernels when changing the location of the imaginary point-source.

# Empirical datasets

In this section we present experimental datasets and re-analyse them using both the conventional method that assumes point source and the spatially explicit method accounting for source shape. We show the extent of estimate error in real-life examples and also demonstrate the use of spatially explicit method to achieve design-independent dispersal estimates that allow comparison between different experiments.

The first dataset comes from a field experiment with Septoria tritici blotch of wheat (Karisto et al., 2021). The second is a field experiment with wheat stripe rust conducted by Sackett and Mundt (2005) (data described in Cowger et al., 2005). The third dataset comes from a field experiment with potato late blight by Gregory et al. (1968) (data extracted from Gregory, 1968, Table 3). In each experiment, the disease was inoculated in an inoculation area within experimental plots to create a source of dispersing population. Then, primary disease gradient was recorded outside the inoculation area, to measure the dispersal of the pathogen assuming uniform infection success. In the septoria experiment, rain-splash dispersal is expected, while the two other pathogens are mainly airborne. The spatial scales of the experiments vary from 100 cm to 100 m (Fig. 4). For each of the three datasets, we choose the appropriate dispersal kernel function and derive the expressions for the expected dispersal gradients first assuming the point-like source and second applying spatially-explicit consideration of the source area.

## Septoria tritici blotch

Central areas of experimental plots were inoculated with spores of the fungal pathogen *Z. tritici* and the subsequent spread of the pathogen was measured. Plot design and measurement distances are shown in Fig. 4A. The dataset is a subset from a larger experiment (pathogen strain "3D7" out of thee treatments) described by Karisto et al.

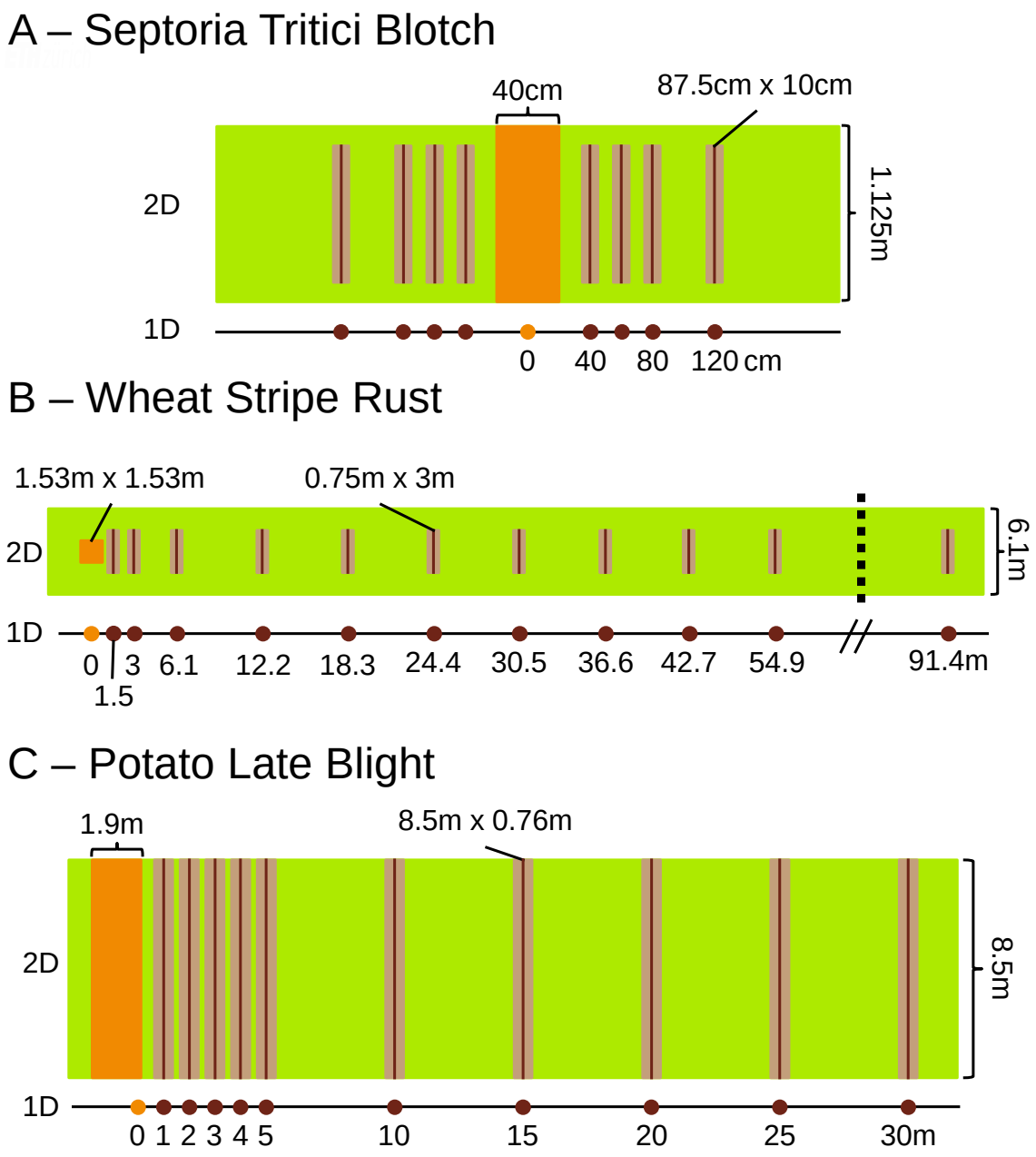


Figure 4: Designs of the experimental plots. The two-dimensional (2D) view was used for the spatially-explicit approach, while the one-dimensional (1D) simplification was used in the conventional approach assuming one-dimensional dispersal, point source, and point destinations. Inoculation areas in orange, measurement areas in light brown and measurement lines (used in the fitting) in dark brown. (A) Septoria experiment.(B) Yellow rust experiment. (C) Late blight experiment. Note that in the Late blight experiment, the virtual point source was located in the near edge of the inoculation area. Also, the experimental plots were symmetric around the inoculation area, but only one half is shown.

(2021). For the data-analysis we first assume that the dispersing individuals originate from an infinitesimally small point in the middle of the inoculated area, at $x = 0$, and that the dispersal gradient is measured at four points located at increasing distances from the source (Fig. 4A, 1D). Hence, we model the dispersal as a one-dimensional process. Applying the exponential kernel (Eq. (4)) in the one-dimensional model (Eq. (2)) the expected disease intensity after dispersal $I_{t_1}$ at a distance $x$ from the source is

$$I_{t_1}(x) = I_0 \beta \frac{e^{-|x|/\alpha}}{2\alpha} \tag{7}$$

where $I_0$ is the disease intensity at the source and $\beta$ is the transmission parameter comprising the probabilities of dispersal and successful infection of fungal spores. We use the exponential kernel as it is expected to fit well in the case of splash dispersal of spores (Fitt et al., 1987).

Next, we relax the assumption that the source is a point and use a spatially-explicit model (1) to account for the spatial extension of the source. In this case, the expected disease intensity at a destination point $(x_d, y_d)$ is given by

$$I_{t_1}(x_d, y_d) = I_0 \beta \int_{y_s=0}^{w_y} \int_{x_s=-w_x/2}^{w_x/2} \frac{e^{-\sqrt{(x_d-x_s)^2+(y_d-y_s)^2}/\alpha}}{2\pi\alpha^2} dx_s\, dy_s. \tag{8}$$

where the integration is conducted across the area of the source, ($w_x = 125\,\text{cm}$ x $w_y = 40\,\text{cm}$), and in this way the contribution of each source point $(x_s, y_s)$ to the disease intensity at the destination point is taken into account.

We adjust the Eq. (8) to incorporate two specific features of our experimental design: (i) the values measured at different distances from the source were in fact averages over multiple values acquired from measurements within small rectangular areas (shown as light brown rectangles in Fig. 4A) and (ii) these areas did not span the whole width of the plot but excluded its borders. With this in mind, the average disease intensity at a

distance $x$ from the source now reads

$$I_{t_1}(x) = \frac{I_0\beta}{w_y - 2b} \int_{y_d=b}^{w_y-b} \int_{y_s=0}^{w_y} \int_{x_s=-w_x/2}^{w_x/2} \frac{e^{-\sqrt{(x-x_s)^2+(y_d-y_s)^2}/\alpha}}{2\pi\alpha^2} dx_s\, dy_s\, dy_d, \qquad (9)$$

where there is an additional integration over the vertical extension of the destination lines $y_d$ and the outcome is divided by the length of the destination line ($w_y - 2b$, where $b = 12.5$ cm is the width of the border excluded from the measurement) to give the mean disease intensity over the line. Here, we consider the destinations as infinitesimally thin lines, neglecting their 10 cm extension in x-dimension, as indicated by the dark brown lines within the receiver areas in Fig. 4A. We fit the one-dimensional Eq. (7) and the two-dimensional Eq. (9) to the empirical data on dispersal gradients of *Z. tritici* to estimate and compare the shape parameters $\alpha$.

## Stripe rust

Stripe rust (caused by the fungus *P. striiformis*) was inoculated in 1.53 m x 1.53 m squares in 6.1 m-wide plots that were at least 100 m-long in the down-wind direction by Cowger et al. (2005). Dispersal was measured visually as the percentage of leaf area covered by lesions. Measurement distances were 1.5 m, 3 m, 6.1 m, 12.2 m and then every 6.1 m, from the center of the inoculated area. Disease severity was assessed over 75 cm-wide and 3 m-long strips across the middle of the plot (Fig. 4B).

We estimate the kernel parameters using the modified power-law kernel as in Mikaberidze et al. (2016), which follows from the power-law kernel (Box 1, eq. (6)) taking square of lambda and distance and then square-root of their sum. In one dimension, the disease severity at a distance $x$ from the source is

$$I_{t_1}(x) = C I_0 \beta (\lambda^2 + x^2)^{-\gamma/2} \qquad (10)$$

where $\lambda$ is fixed to 0.762 m, and C is normalization factor of the modified kernel

$$C = \frac{2\lambda^{\gamma-1}\Gamma(\frac{\gamma}{2})}{\sqrt{\pi}\Gamma(\frac{\gamma-1}{2})}$$

where $\Gamma(\cdot)$ is the gamma-function.

Following the approach of Sackett and Mundt (2005) and Mikaberidze et al. (2016) we performed the fitting on logarithmic scale to avoid over-emphasis on the few large values in the beginning of the gradient. Zeros were excluded from the log-transformed data. Hence, we fitted function

$$\log(I_{t_1}(x)) = \log(CI_0\beta(\lambda^2 + x^2)^{-\gamma/2}) = \log(CI_0\beta) - \frac{\gamma}{2}\log(\lambda^2 + x^2) \tag{11}$$

to the dataset to estimate the shape parameter $\gamma$.

In the two-dimensional approach with the explicit consideration of the spatial extension of the source, we used similarly modified power-law kernel applied in eq. (1), resulting in the following dispersal gradient

$$I_{t_1}(x_d) = \frac{I_0\beta}{w_m}C\int_{y_d=b_m}^{b_m+w_m}\int_{y_s=b_s}^{b_s+w_s}\int_{x_s=0}^{w_s}(\lambda^2 + (x_d - x_s)^2 + (y_d - y_s)^2)^{-\gamma/2}dx_s\,dy_s\,dy_d. \tag{12}$$

Here, $x_s = 0$ is set at the middle of the inoculated area in the direction of the gradient and $y = 0$ is set at the edge of the plot, $w_s = 1.53$ m is the linear extension of the square inoculated area, $b_s = 2.285$ m is the gap between the inoculated area and the edge of the plot, $w_m = 3$ m is the extension of the measurement line along the $y$-axis and $b_m = 1.525$ m is the width of the border zone excluded from the measurement line at each end (see Fig. 4B). $C = (\gamma - 2)/(2\pi\lambda^{2-\gamma})$. The fitting was performed on log-transformed data and taking a logarithm of both sides of Eq. (12). We fit the one-dimensional Eq. (11) and the two-dimensional Eq. (12) to the empirical data on dispersal gradients of *P. striiformis* to estimate and compare the shape parameters $\gamma$.

### Potato late blight

Potato late blight (caused by the oomycete *P. infestans*) was inoculated in a five plant-wide strip across the middle of the experimental plots (Gregory, 1968). Plots were 8.5 m wide and over 30 m long in each the two directions from the inoculated area. Two details of the experimental design were not reported in the original paper (Gregory, 1968) and for this reason we inferred them based on reasonable assumptions. First, we assumed the typical distance between plants along the row to be 0.38 m, which resulted in the extension of the inoculated area in the $x$-direction (along the rows) of 1.9 m. The disease was observed in two plants (spanning 0.76 m along the row) in every row at different distances from the source (see Fig. 4C). Second, we assumed the reported distances to be measured from the edge of the inoculation area that is closer to the considered gradient (Fig. 4C), as measuring them from the middle of the inoculation area would cause the measurement location closest to the source to overlap with the inoculated area.

The power-law function was fitted to the observed disease gradient. Without modifications, the power-law function ($r^{-\gamma}$) is not defined at zero and is not a probability density function. Hence, it is formally not a kernel, but can be used to describe dispersal gradients nevertheless. In one-dimensional model, the disease gradient follows the function

$$I_{t_1}(x) = I_0 \beta x^{-\gamma} \tag{13}$$

The gradient data was log-transformed, zeros were excluded, and the fitting was performed on logarithmic scale using the log-transformed function

$$\log(I_{t_1}(x)) = \log(I_0\beta) - \gamma \log(x) \tag{14}$$

In the spatially-explicit approach that takes into account the spatial extension of the source, we describe the dispersal gradient based on the mean intensity over the

measurement lines at a distance $x_d$ from the source applying the power-law function in eq. (1) resulting in a function

$$I_{t_1}(x_d) = \frac{I_0\beta}{w} \int_{y_d=0}^{w} \int_{y_s=0}^{w_y} \int_{x_s=0}^{w_x} \left(\sqrt{(x_d - x_s)^2 + (y_d - y_s)^2}\right)^{-\gamma} dx_s\, dy_s\, dy_d, \qquad (15)$$

where $w_x = 1.9\,\mathrm{m}$ and $w_y = 8.5\,\mathrm{m}$ are the extensions of the inoculated area and both the inoculated area and the measurement line span across the width of the plot, from 0 to $w_y = 8.5\,\mathrm{m}$. Here, the distances are measured from $x = 0$ at the closer edge of the inoculated area. The fitting was performed on log-transformed data and taking a logarithm of the fitting functions. Note that the unmodified power-law kernel cannot be normalized, hence there is no normalization factor C in the functions (13) and (15). Due to this, the variable $\beta$ has no biological relevance in these equations unlike in the equations for the spread of septoria tritici blotch and stripe rust, which emphasizes the benefits of using proper dispersal kernels. We fit the one-dimensional Eq. (14) and the two-dimensional Eq. (15) to empirical data on dispersal gradients of *P. infestans* to estimate and compare the shape parameters $\gamma$.

## Results of the experiments

Figure 5 shows the dispersal gradients for three case studies, the fits obtained using the one-dimensional method and the two-dimensional method. The estimated kernel shapes are also shown. The differences in the parameter estimates between the approach assuming point-like source and the spatially-explicit approach were around 10% for Septoria tritici blotch and stripe rust, but exceeded 30% for potato late blight (Table 2). The goodness of fits, measured as a sum of squared residuals, is not much different between the two methods, but slightly favouring the one-dimensional fit. Hence, goodness of fit does not offer a straight-forward guidance for choosing the right method, as we know that the spatially-explicit method is more realistic and describes the dispersal process

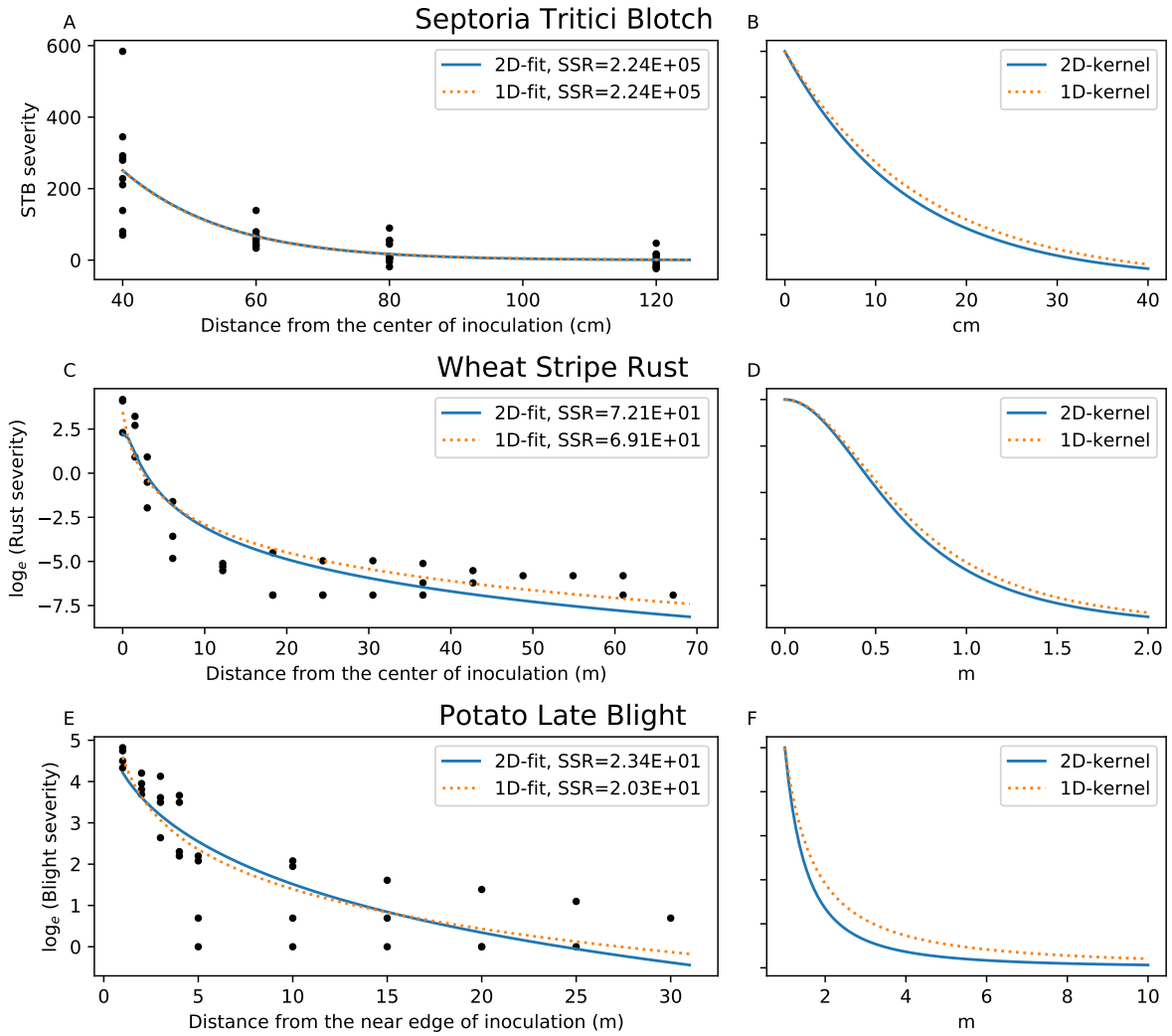


Figure 5: Results of the fittings for estimating the dispersal kernels. The first column, the datasets and the best fitting disease gradients following the spatially-explicit consideration of the source area ("2D-fit", blue line) or the conventional approach assuming a point source ("1D-fit", orange dotted line). The sum of squared residuals (SSR) shown in the legend. The second column, the resulting kernel shapes after parametrization following the two different methods, blue line for spatially explicit estimation and orange dotted line for point-source-assuming approach. With spatially explicit approach, the estimated kernel is steeper in all cases, implying shorter dispersal distance than expected if assuming a point source.

more correctly mechanistically, albeit being a simplification. Note also that in Appendix B we show that re-estimating parameter $\lambda$ of power-law kernel improved the quality of fit when using the point-source approximation, but it simultaneously led to worse estimate of the shape parameter and dispersal distance.

Table 2: Comparison of parameter estimates between one- and two-dimensional models (1D and 2D, respectively).

| Pathosystem | 1D-result | 2D-result | difference (%) |
|---|---|---|---|
| Septoria | $\alpha = 15.09\,\mathrm{cm}$ | $\alpha = 13.51\,\mathrm{cm}$ | 11.7 |
| Rust - Madras | $\gamma = 2.41$ | $\gamma = 2.64$ | 9.5 |
| Rust - Hermiston | $\gamma = 2.51$ | $\gamma = 2.73$ | 8.8 |
| Blight | $\gamma = 1.39$ | $\gamma = 1.89$ | 36.0 |

Note: Parameter $\gamma$ appears in the exponent in power-law kernels, but $\alpha$ is in the denominator of the exponent in the exponential kernel. Hence the direction of the parameter difference is opposite for Septoria compared to others.

In all three examples, an explicit account of the extension of the source revealed that the dispersal kernels were steeper than those estimated assuming a point-like source. This means that in all three organisms the characteristic dispersal distances are somewhat shorter than previously thought. This demonstrates the "flattening" effect of source extensions , which was noted qualitatively for example by Gregory (1968), but in contrast to previous studies we provide a quantitative insight. Note also that extended source can sometimes lead to underestimation of the dispersal ("steepening") as shown by our simulations.

## Discussion

We propose an approach for estimating dispersal kernel parameters, where the spatial extension of the source is explicitly incorporated in the model. This provides a solution for correcting inaccurate estimates caused by unjustified simplifying assumptions (See

Fig. 1 and Table 2). This approach also provides a quantitative answer to the question "By how much and in which way does the source dimension affect the observed dispersal gradient?", instead of more qualitative statements regarding the "flattening" of the gradient with a larger source (Zadoks and Schein, 1979; Ferrandino, 1996; Cousens and Rawlinson, 2001).

Using our method, we are able to relax the requirement of having a point source in a dispersal experiment. This helps experimental design by increasing the output from the source and consequently the availability of samples to record, which may be a limiting factor in many systems. Additionally, with our approach, results acquired from different experimental designs (e.g. those cited in Fitt et al., 1987) can be re-analyzed to estimate dispersal kernels in each case, as we demonstrated. Importantly, the presented method provides estimates of kernel shape that are independent of the source shape. Hence, results from different experiments can be compared directly, when the estimates are achieved with the spatially-explicit method, whereas different source dimensions would invalidate the comparison if a point source was assumed. Moreover, our approach allows to estimate actual kernel parameters in a much wider range of empirical studies than it was recognized previously, including all studies with spatially extended sources. From theoretical perspective, using the spatially explicit consideration of the source geometry, results in a more correct estimate of the kernel shape than assuming that the source is a point. In this way, "we can move from descriptions of pattern to a grasp of process" (Bullock et al., 2006).

We show, with simulations (Fig. 1) that different source geometries may lead to similar gradients when the kernel is either memoryless or separable. However, most kernels are neither memoryless nor separable, and thus distortions of the gradient shape are expected with varying sources. In any case, such simulations can be used when planning an experiment to guide the experimental design and to test predicted outcomes.

Simulated outcomes of an experiment can also help to determine when the source can be considered a point, to simplify the analysis, and what kind of errors this simplification may introduce. Furthermore, the presented two-dimensional models of dispersal are naturally simplifications of the three-dimensional process of dispersal (e.g. Vidal et al., 2018). If the third dimension is of great importance in the system, as perhaps in aquatic environments or with tree canopies (Cousens and Rawlinson, 2001), modeling the source and dispersal processes in three dimensions may be necessary for accuracy.

Besides the simulations, we used the spatially-explicit accounting of source area to analyze three experimental datasets with variable spatial scales demonstrating the effect of source extensions on the estimated dispersal kernels. In all experiments, the assumption of point-source led to notable difference in parameter estimates compared to the spatially- explicit method, suggesting that the size of the distorting effect of the point-source approximation is of relevance in actual field experiments. All experiments were conducted using artificial experimental design with passively dispersing organisms, which provide conducive conditions for experiments, but similar analysis approach can be used in observational studies in nature and with actively dispersing organisms whenever the source area can be characterized and the dispersal process can be described with dispersal kernels.

We assumed isotropic dispersal in the analysis of the data. However, the presented method can be readily modified to take into account the anisotropy of the kernel. In the modified model, the probability of dispersal from a source point to a destination point should depend not only on the distance between the points, as in our case, but also on the direction from the source to the destination. In this way, parameters for anisotropic dispersal can be estimated from measurements of dispersal gradients with the spatially-explicit consideration of the source.

Our simulations showed that most pronounced differences between the dispersal gra-

dients originating from different source geometries appear close to the source, while at larger distances from the source these differences disappear (Fig. 1B, C, D). The effect is seen in each of the three very different types of kernels, suggesting that it is a universal feature. Therefore, even when the size of the source is much smaller than the gradient length, it could be that the size of the source is still comparable to the characteristic dispersal distance (i.e., the distance over which the dispersal kernel changes substantially). In this case, measurements close to the source will be substantially distorted due to the finite area of the source. As an example, the difference in the kernel parameter estimates was clear even in the case of rust experiments, where the source was obviously small compared to the measured gradient (1.53m vs. 100m). Furthermore, the error in parameter estimate of power-law kernel was significant even with 50 m dispersal distance (Fig. 2A). Therefore, simple rules of thumb stating that to be considered as a point, the size of the source should be smaller than 1 % of the length of the gradient (Zadoks and Schein, 1979; McCartney et al., 2006), can be quite misleading, and result in inaccurate estimates of dispersal parameters. This emphasizes the importance of the spatially-explicit modeling of the source as we have done it here, considering that it is often the case that most measurements are conducted close to the source even when the overall gradient is long (Werth et al., 2006; Skarpaas and Shea, 2007; Loebach and Anderson, 2018).

Experiments that measure dispersal are difficult and laborious (Bullock et al., 2017). Spatial dimension of the source, location of sampling areas, amount of sampling at different locations and other components of the experimental design may have a large effect on the precision and generalizability of the results. We support Skarpaas et al. (2005) calling for optimization of dispersal study designs by simulations, to make the most out of the effort.

# Acknowledgements

PK and AM gratefully acknowledge financial support from the Swiss National Science Foundation through the Ambizione grant PZ00P3_161453. Genetic Diversity Centre at ETH Zurich helped with the genetic analyses. PK and AM would like to thank Andreas Hund and Hansueli Zellweger for maintenance of the field experiment. FS would like to thank Christophe Montagnier, Sandrine Gélisse and Nicolas Lecutier for maintenance of the similar field experiment prepared at the facilities of INRA Bioger in Thiverval-Grignon, France.

# Appendix A: Dispersal distance kernel

**Dispersal distance kernel.** Dispersal kernels are not only used to describe distributions of locations of dispersed individuals, but also to summarize dispersal distances, such as mean distance travelled. Dispersal distance kernel is a one-dimensional function describing the probability of individuals to end up at a certain distance from the source. It can be derived from a two-dimensional dispersal location kernel by integrating it around the source, essentially by multiplying it with $2\pi r$ (Nathan et al., 2012). The shape of the dispersal location kernel can differ substantially from the shape of the dispersal distance kernel (e.g. Cousens and Rawlinson, 2001; Nathan et al., 2012). The dispersal distance kernel corresponding to exponential dispersal location kernel is given by

$$\kappa_{e,dist}(r) = 2\pi r \frac{e^{-r/\alpha}}{(2\pi\alpha^2)} = \frac{re^{-r/\alpha}}{\alpha^2} = \frac{r^{k-1}e^{-r/\alpha}}{\alpha^k\Gamma(k)}, \quad k = 2, \tag{A1}$$

where $\Gamma(2) = 1$ is the gamma function. Equation (A1) gives the one-dimensional gamma-distribution with the shape parameter $k = 2$ and the scale parameter $\alpha$.

It is important to keep in mind that means and medians of dispersal location kernels do not generally correspond to means and medians of population dispersal distances (i.e. means and medians of dispersal distance kernels). The example of the exponential kernel is of particular importance, as this kernel is often described with the half-distance. Considering the one-dimensional exponential kernel, the dispersal parameter $\alpha$ gives the mean and $\alpha \ln(2)$ (half-distance) gives the median of the distribution. However, in the case of the two-dimensional exponential location kernel the mean dispersal distance is not $\alpha$ but $2\alpha$ (i.e. mean of the gamma-distribution in Eq. (A1)). Furthermore, median, or any percentile, of the dispersal distance distribution can be determined by solving the equation

$$\int_0^{x_L} \kappa_{e,dist}(r)dr = 0.01L \tag{A2}$$

where $x_L$ is the $L^{\text{th}}$ percentile of dispersal distance ($L \in [0, 100]$). After integration, the Eq. (A2) reads

$$e^{-x_L/\alpha}\left(1 + \frac{x_L}{\alpha}\right) = 1 - 0.01L. \tag{A3}$$

We solve Eq. (A3) numerically to obtain the median dispersal distance $x_{50} \approx 1.7\alpha \gg 0.69\alpha \approx \ln(2)\alpha$. Considering the limits of population dispersal, Golan and Pringle (2017) defined 99th percentile of the dispersal distance distribution as a limit for the long-distance dispersal in fungi. At $L = 99$, we find $x_{99} \approx 6.6\alpha$. These numbers have applied relevance, for example in conservation biology or in precision agriculture when a treatment is targeted to a certain fraction of a dispersing population. Mean or median dispersal distances or other characteristic numbers should be determined using the two-dimensional location kernel and the corresponding distance kernel.

The limit of long-distance dispersal ($x_{99}$, Eq. (A3)) corresponding to observed values of $\alpha$ is 90 cm for treatment D and 142 cm for treatment B. In an agricultural field, a visible disease focus (Zadoks and van den Bosch, 1994) and significant host damage (Shaw and Royle, 1993) would occur close to the source due to higher density, while the edge of the population is likely to incur less damage because of lower pathogen density. This hidden pathogen population in the tails of the distribution should be taken into account when attempting spatially targeted treatments, for example in precision agriculture involving focal fungicide spraying.

# Appendix B: Further analysis on the errors caused by the point-source approximation

## Point source at the edge of the source area

In the main text, we presented estimate errors when assuming a point source in the middle of the source. Here, we conducted a similar analysis while assuming the point-source at the edge of the source closest to the measured gradient. Analogously to the centered point approach, the errors in the parameter estimates decrease with longer dispersal and increase with larger source (Fig B1A, B). With the point-source at the edge, exponential kernel and power-law kernel overestimate but Gaussian kernel underestimates the dispersal distance (Fig B1C, D). Hence, the sign of the error is opposite to that with centered point source, for Gaussian and power-law. The distorted parameter estimate for power-law kernel actually implies infinite mean dispersal distance (real $\gamma > 3$, estimated $\gamma < 3$). (Note that for exponential and Gaussian kernels, the error in the parameter and in the dispersal distance have the same sign, but the sign is opposite for power-law kernel.)

The change of direction in the estimate error for Gaussian and power-law kernels can be explained by cumulative distorting effects of the source points within the area source (as in Fig. 3. When placing the imaginary source points at the edge of the source, in the beginning of the measured gradient, the distorting effect of entire source is simply leading to underestimation of the dispersal distance with Gaussian kernel and to overestimation with power-law kernel (Fig. B2).

## "Remaining error" in the estimate of power-law kernel parameter

In Figure 2A, the error of the parameter estimate approached zero quite fast for exponential and Gaussian kernel, but remained at around 3% for power-law kernel even with

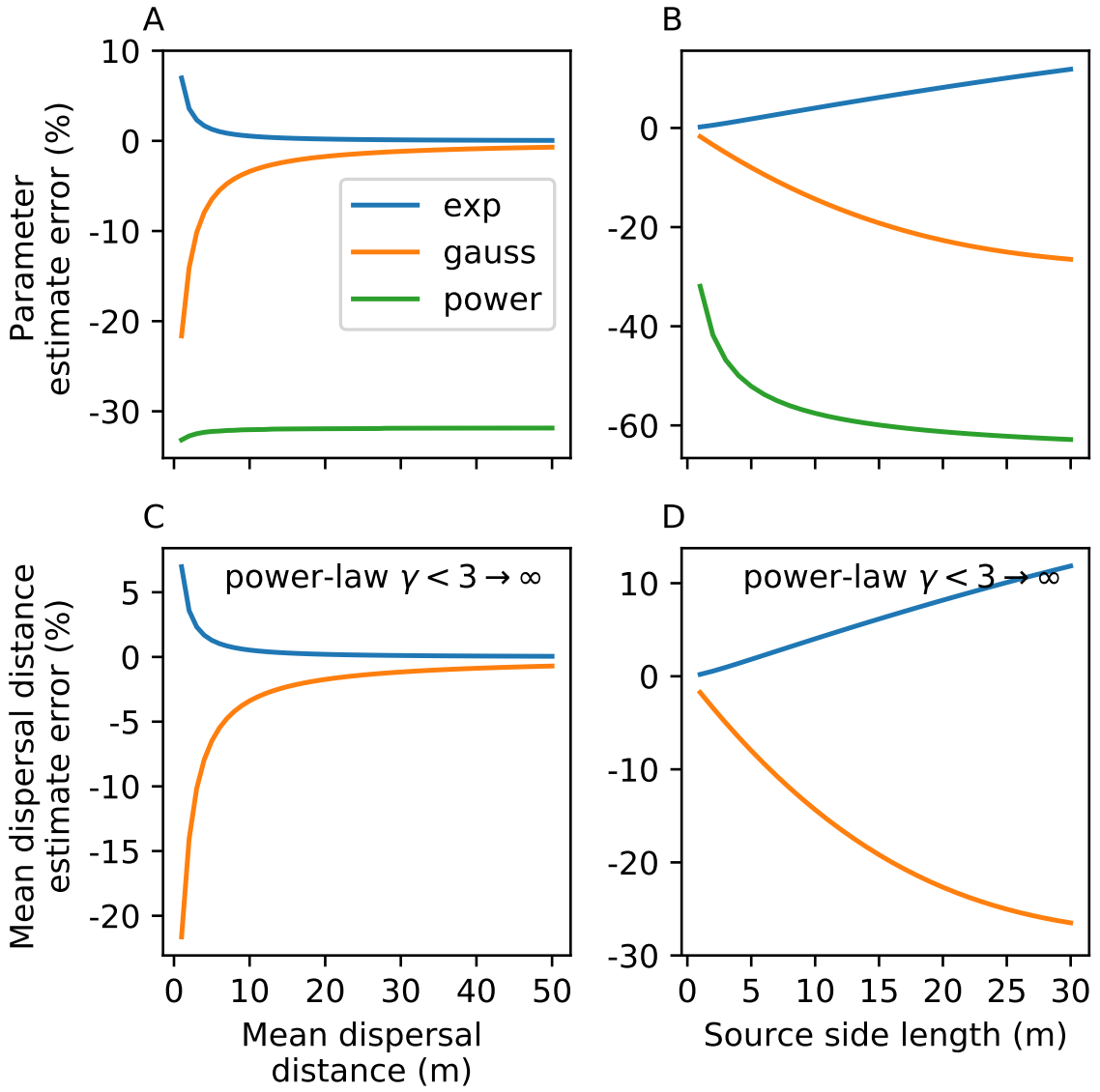


Figure B1: Errors in parameter estimate (upper row) and in estimated mean dispersal (lower row) as a function of source size (side length) when using the point-source estimate assuming the point in the middle of the source (left column) or on the edge (right column).

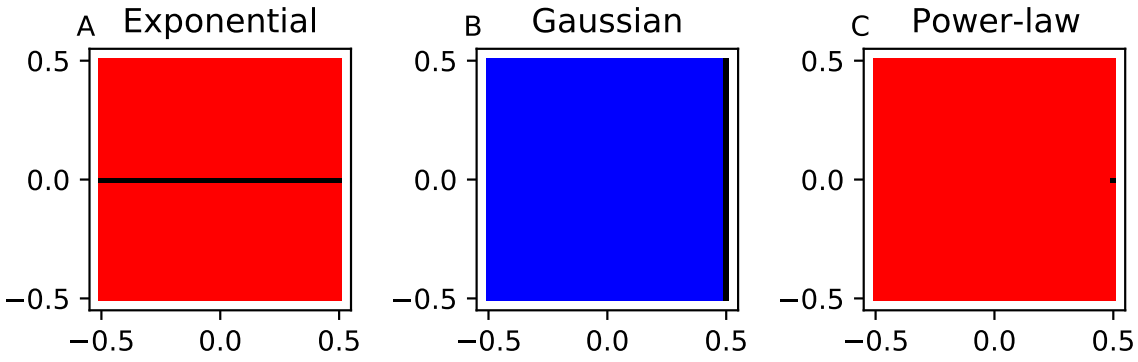


Figure B2: Contribution of individual points within an extended source to the estimated mean dispersal distance. Dispersal distance re-estimated assuming a point-source at the edge of the actual source. For each point in the source (2 cm × 2 cm grid) the color shows, whether that point contributes to underestimation (blue) or overestimation (red) of the dispersal distance. Black: precise estimate. Gradient measured to right from the source.

the longest dispersal distance. To investigate whether we could decrease the remaining error, we tried implementing various changes in the simulations or fitting. First we increased the sampling density of the gradient, then sampled the gradient further into the tail, and finally re-estimated the scale parameter $\lambda$ alongside with the shape parameter $\gamma$ (originally $\lambda$ was kept constant). In these tests we used fixed dispersal distance of 20 m, 1 m-by-1 m source, as for Figure 2. With the regular 10 cm sampling interval and sampling until reaching 1% of the intensity of the first measurement point, the parameter estimate was $\gamma_{est} = 3.153$ exhibiting a 3.37% overestimation compared to the real value ($\gamma_{real} = 3.05$). The estimated mean dispersal distance was 6.54 m.

### Sample more densely

We increased the sampling of the simulated gradient to ten-fold density, starting the measurements from 1 cm outside of the source and sampling every 1 cm instead of 10 cm. With this ten-fold sampling, the error decreased slightly to 2.25% ($\gamma_{est} = 3.119$) and estimated mean dispersal distance was 8.4 m.

### Sample further into the tail

We decreased the threshold for stopping the sampling so that we included all points that had more than 0.1% of the intensity of the first sampling point (original threshold: 1%). This led to $\gamma_{est} = 3.152$ (error: 3.36%) corresponding to 6.56 m estimated mean dispersal distance, representing only negligible improvement.

### Re-estimate $\lambda$

In the final modification, we fitted parameter $\lambda$ alongside with $\gamma$. This improved quality of the one-dimensional fit, leading to chi-squared value of 0.699 compared to 1.278 of the original fitting. In contrast, the error in the shape parameter increased ($\gamma_{est} = 3.337$ (error: 9.33%)) and so did the error in the mean dispersal distance (estimated mean dispersal distance = 3.50 m, $\lambda_{est} = 0.585$). Hence, while fitting two parameters instead of one improved the fitting, the results were more erroneous, suggesting that the quality of the fit does not imply quality of the parameter estimates. However, this is what would be done with real experimental data, as we cannot know the real value of the scale parameter $\lambda$ (or any other parameter). Thus, the one-dimensional fitting with point source approximation has serious limitations for estimating the properties of dispersal kernels.

### Conclusion

To summarize, the modifications of the sampling and fitting procedure led to minor improvements either in the parameter estimates or the quality of the fit. The largest improvement was achieved by ten-fold sampling, which decreased the error from 3.37% to 2.25%. However, this improvement in the parameter estimate seems not satisfying considering the increased effort to achieve ten-fold sampling density. When sampling further into the tail, the improvement was negligible, while the practical issues of de-

tecting dispersing population at 0.1% of the highest density would likely be considerable. Finally, fitting both of the parameters simultaneously (as in reality) improved the quality of the fit, but also made the estimates of dispersal worse. Hence, there seems to be quite hard limit to the precision of the point-source approximation when using power-law kernel. Luckily, the spatially-explicit method provided precise estimates in all examples of simulated dispersal experiments.

# Literature Cited

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

Bullock, J. M., K. Shea, and O. Skarpaas. 2006. Measuring plant dispersal: an introduction to field methods and experimental design. Plant Ecology **186**:217–234.

Carrasco, L., T. Harwood, S. Toepfer, A. MacLeod, N. Levay, J. Kiss, R. Baker, J. Mumford, and J. Knight. 2010. Dispersal kernels of the invasive alien western corn rootworm and the effectiveness of buffer zones in eradication programmes in Europe. Annals of Applied Biology **156**:63–77.

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

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

Cowger, C., L. D. Wallace, and C. C. Mundt. 2005. Velocity of spread of wheat stripe rust epidemics. Phytopathology **95**:972–982.

D'Aloia, C. C., S. M. Bogdanowicz, R. K. Francis, J. E. Majoris, R. G. Harrison, and P. M. Buston. 2015. Patterns, causes, and consequences of marine larval dispersal. Proceedings of the National Academy of Sciences **112**:13940–13945.

Devaux, C., C. Lavigne, F. Austerlitz, and E. K. Klein. 2006. Modelling and estimating pollen movement in oilseed rape (Brassica napus) at the landscape scale using genetic markers. Molecular Ecology **16**:487–499.

Emsweller, L. N., D. L. Gorchov, Q. Zhang, A. G. Driscoll, and M. R. Hughes. 2018. Seed Rain and Disturbance Impact Recruitment of Invasive Plants in Upland Forest. Invasive Plant Science and Management **11**:69–81.

Ferrandino, F. 1996. Length scale of disease spread: fact or artifact of experimental geometry. Phytopathology **86**:806–811.

Fitt, B. D. L., P. H. Gregory, A. D. Todd, H. A. McCartney, and O. C. Macdonald. 1987. Spore dispersal and plant disease gradients; a comparison between two empirical models. Journal of Phytopathology **118**:227–242.

Golan, J. J., and A. Pringle. 2017. Long-distance dispersal of fungi. Microbiology spectrum **5**:FUNK–0047–2016.

Goto, S., K. Shimatani, H. Yoshimaru, and Y. Takahashi. 2006. Fat-tailed gene flow in the dioecious canopy tree species Fraxinus mandshurica var. japonica revealed by microsatellites. Molecular Ecology **15**:2985–2996.

Greene, D. F., and C. Calogeropoulos, 2002. Measuring and modelling seed dispersal of terrestrial plants. Pages 3–23 *in* Dispersal ecology: the 42nd symposium of the British Ecological Society.

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

Gregory, P. H., D. R. Henden, and D. H. Lapwood, 1968. Dispersal gradients of potato blight. Page 128 *in* J. M. Hirst, editor. Rothamsted Report for 1967. Rothamsted Research.

Heald, F. 1913. The Address of the President for 1912: The Dissemination of Fungi Causing Disease. Transactions of the American Microscopical Society **32**:5–29.

Karisto, P., F. Suffert, and A. Mikaberidze. 2021. Measuring splash-dispersal of a major wheat pathogen in the field. bioRxiv page 2021.03.23.436423.

Kot, M., M. A. Lewis, and P. van den Driessche. 1996. Dispersal data and the spread of invading organisms. Ecology **77**:2027–2042.

Loebach, C. A., and R. C. Anderson. 2018. Measuring short distance dispersal of Alliaria petiolata and determining potential long distance dispersal mechanisms. PeerJ **6**:e4477.

Madden, L. V., G. Hughes, and F. van den Bosch. 2007. The study of plant disease epidemics. American Phytopathological Society (APS Press).

McCartney, H., B. D. Fitt, and J. S. West, 2006. Dispersal of foliar plant pathogens: mechanisms, gradients and spatial patterns. Pages 159–192 *in* B. M. Cooke, D. G. Jones, and B. Kaye, editors. The epidemiology of plant diseases. Springer.

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

Nathan, R., E. Klein, J. J. Robledo-Arnuncio, and E. Revilla, 2012. Dispersal kernels: review. Chapter 15, pages 187–210 *in* J. Clobert, M. Baguette, T. G. Benton, and J. M. Bullock, editors. Dispersal ecology and evolution. Oxford University Press Oxford.

Nathan, R., U. N. Safriel, I. Noy-Meir, and G. Schiller. 2000. Spatiotemporal variation in seed dispersal and recruitment near and far from Pinus halepensis trees. Ecology **81**:2156–2169.

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

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

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

Shaw, M., T. Harwood, M. Wilkinson, and L. Elliott. 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**:1705–1713.

Shaw, M., and D. Royle. 1993. Factors determining the severity of epidemics of Mycosphaerella graminicola (Septoria tritici) on winter wheat in the UK. Plant Pathology **42**:882–899.

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

Skarpaas, O., K. Shea, and J. M. Bullock. 2005. Optimizing dispersal study design by Monte Carlo simulation. Journal of Applied Ecology **42**:731–739.

Solheim, H., and A. M. Hietala. 2017. Spread of ash dieback in Norway. Baltic Forestry **23**:144–149.

760 Van Houtan, K. S., S. L. Pimm, J. M. Halley, R. O. Bierregaard Jr, and T. E. Lovejoy. 2007. Dispersal of Amazonian birds in continuous and fragmented forest. Ecology 762 letters **10**:219–229.

Vidal, T., C. Gigot, C. de Vallavieille-Pope, L. Huber, and S. Saint-Jean. 2018. Con- 764 trasting plant height can improve the control of rain-borne diseases in wheat cultivar mixture: modelling splash dispersal in 3-D canopies. Annals of botany **121**:1299–1308.

766 Werth, S., H. H. Wagner, F. Gugerli, R. Holderegger, D. Csencsics, J. M. Kalwij, and C. Scheidegger. 2006. Quantifying dispersal and establishment limitation in a popu- 768 lation of an epiphytic lichen. Ecology **87**:2037–2046.

Zadoks, J., and F. van den Bosch. 1994. On the spread of plant disease: a theory on 770 foci. Annual review of phytopathology **32**:503–521.

Zadoks, J. C., and R. D. Schein. 1979. Epidemiology and plant disease management. 772 Oxford University Press Inc.