

Optimizing dispersal study design by Monte Carlo simulation

OLAV SKARPAAS,* KATRIONA SHEA* and JAMES M. BULLOCK†

*The Pennsylvania State University, Department of Biology and IGDP in Ecology, 208 Mueller Laboratory, University Park, PA 16802, USA; and †NERC Centre for Ecology and Hydrology, CEH Dorset, Winfrith Technology Centre, Dorchester, Dorset DT2 8ZD, UK

Summary

1. The distribution of dispersal distances (the dispersal kernel) is a major determinant of spatial population dynamics, yet little is known about the shape of the dispersal kernel for most species. This is partly because of the relative difficulty of measuring dispersal, exacerbated by a lack of standardized protocols. We suggest that this problem can be addressed by using modelling approaches to aid the design of studies to quantify dispersal.
2. In this study we present such an approach by optimizing seed trap sampling design using stochastic simulations. A number of alternative sampling designs (random placements, grid arrays, transects, sectors and annuli arrangements) for a point source were tested against a common kernel to assess the best methods for estimating the dispersal kernel.
3. For a given source strength and total trap area, transects and sectors of traps usually provided better data for kernel estimation than random placement, grid arrays and annuli. Kernel estimation was improved by increasing the source strength (the number of dispersing propagules) and the trap area, as expected.
4. When the 'true' kernel was unknown, transects were slightly better for identifying the thin-tailed exponential distribution, whereas sectors were better for detecting the fat-tailed half-Cauchy.
5. In the case of anisotropic dispersal (here, dispersal biased in one direction), annuli and grid arrays performed better than transects and sectors when the anisotropy was unknown. However, when the anisotropy was anticipated, and the trap arrangements were adjusted accordingly, transects and sectors were better. This was true regardless of source strength and total trap area.
6. *Synthesis and applications.* This study presents a simulation approach to the design of dispersal experiments. While the general results of our simulations can be used by those designing field studies for plant point sources, the simulation approach itself can be modified for a wide range of organisms, dispersal mechanisms and dispersal measurement approaches. Thus, the approach presented here facilitates improvements of dispersal study designs, which in turn will increase the precision of dispersal kernel estimates and predictions of spatial population dynamics, including modelling of rates of spread or metapopulations. This is invaluable in a range of ecological applications, such as the management of rare or invasive species, predicting species' response to climate change, or planning species reintroductions.

Key-words: dispersal kernel, dispersal measurement, Eulerian approach, long-distance dispersal, sampling design, seed traps

Journal of Applied Ecology (2005) **42**, 731–739

doi: 10.1111/j.1365-2664.2005.01056.x

Introduction

Dispersal is an extremely important biological process. It affects intra- and interspecific interactions and population dynamics (Dieckmann, Law & Metz 2000; Okubo, Hastings & Powell 2001), rates and patterns of spread (Shaw 1995; Shigesada & Kawasaki 1997; Clark, Lewis & Horvath 2001), genetic variation and structure (Hamrick & Nason 1996; Ibrahim, Nichols & Hewitt 1996), regional species persistence (Brachet *et al.* 1999; van Groenendael, Ehrlén & Svensson 2000), community composition and dynamics (Tilman, Lehman & Yin 1997; van Groenendael, Ehrlén & Svensson 2000) and large-scale biogeographical patterns (Cain, Damman & Muir 1998; Kriticos *et al.* 2003). For population forecasting, conservation and management, for example for biological invasions, it is essential to obtain accurate measures of the distribution of dispersal distances (the dispersal kernel). Unfortunately, dispersal data that allow reliable estimates of dispersal kernel parameters are lacking for most species. This is partly because of a lack of guidelines for dispersal measurement and few attempts at optimizing sampling methods for particular questions (Assunção & Jacobi 1996; Stoyan & Wagner 2001). In this study, we used Monte Carlo simulations to investigate the efficiency of different sampling designs for dispersal measurement.

Dispersal study design is a decision process. The optimal design will depend on the dispersal-related question, the species, the mode(s) of dispersal it employs and the context. Three main questions need to be answered at the outset of the design process. What is the objective of the dispersal study? What are the options available for measuring dispersal? What are the constraints on the experimental design?

Of these questions, the first is particularly important. Without a clear objective, optimization is not possible. In many dispersal studies the objective is a proper characterization of the entire dispersal kernel, but different aspects of the kernel (such as the mean, the variance or the tail, i.e. long-distance dispersal) may also be of particular interest. In the simulation approach taken in this study we are able to monitor all of these and optimize for any of them.

The options and constraints in an experimental design are likely to be system specific. Options available for measuring dispersal range from tracking individual propagules to estimates based on offspring counts or genetic markers (Turchin 1998; Bullock, Kenward & Hails 2002; Nathan *et al.* 2003). In this study, we focused on seed shadows of realistic dimensions for well-dispersed herbs (Willson 1993), and we confined ourselves to the Eulerian approach for dispersal measurement, i.e. the direct measurement of numbers/densities of propagules at different distances from the source. However, the approach we present is applicable to a wide range of systems and techniques for measurement and estimation.

We used Monte Carlo simulations to assess the efficiency of various trap designs for dispersal measurements under various assumptions about source strengths, sampling efforts and distributions of dispersal distances and directions. While the results of our simulations can be used more or less directly by those designing new field studies for plant point sources, the general guidelines and the simulation approach are applicable to a wide range of organisms, dispersal mechanisms and methods for dispersal measurement; essentially any directly measurable pattern of dispersal that can be characterized by a probability distribution.

Methods

We present a Monte Carlo simulation approach to optimization of dispersal study design. First, we describe the general procedure and a set of simulations we carried out to illustrate various considerations in designing a trap study for a point source. Second, we describe how models were fitted to the data and compared with the 'true' kernel and to each other. All simulations and analyses were carried out using R (The R Development Core Team 2003).

SIMULATIONS

We carried out a number of simulations to assess the efficiency of various trap designs for different source strengths, sampling efforts and dispersal kernels. In all of the studies the general simulation procedure was as follows.

1. Disperse a number of seeds according to a 'true' dispersal kernel.
2. Sample the dispersed seeds with each of the sampling designs.
3. Fit kernels to each of the sampled data sets.
4. Assess the goodness-of-fit to the data and the true kernel for fitted kernels.
5. Repeat steps 1–4 a number of times to compare replicates.

We used the following general model (Clark *et al.* 1999) to simulate the redistribution of individual propagules from a point source in a two-dimensional landscape:

$$S(Q, r, \theta) = Qf(r)g(\theta) \quad \text{eqn 1}$$

where the seed shadow $S(Q, r, \theta)$ is as a function of the number of propagules dispersed (Q) and distance (r) and angle (θ) from the source. The distance kernel $f(r)$ is the probability distribution of dispersal distances (Fig. 1), and $g(\theta)$ is the probability distribution of dispersal direction θ . Note that equation 1 assumes that distance and direction can be characterized independently, i.e. f is independent of direction (θ) and g is independent of distance (r).

The following sections describe the simulations in greater detail. We started with a baseline scenario

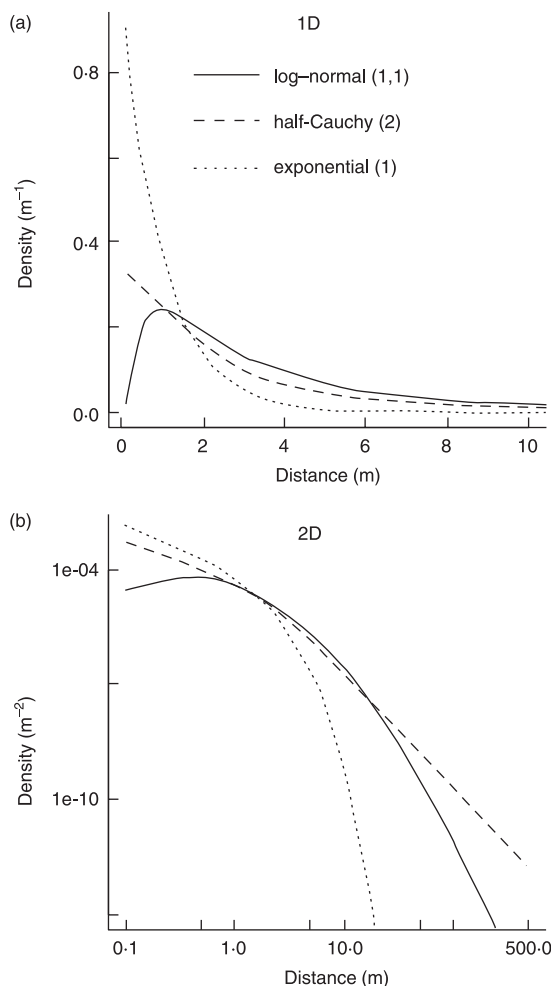


Fig. 1. Distributions of dispersal distances used in this study (equations and shorthand notation described in the Methods). (a) The three distributions plotted on an ordinary scale. (b) The same three distributions divided by $2\pi r$ to give two-dimensional seed densities (assuming isotropic dispersal) and plotted on a log-log scale to accentuate key differences (this plot corresponds to Figs 2b, 5b and 6b).

assuming simple conditions for optimization: large seed sources and sampling efforts (total trap areas), knowledge about the true kernel, and isotropic dispersal. We then relaxed each of these assumptions in scenarios that considered the effects of changing the seed source and sampling effort, model uncertainty and anisotropic dispersal.

BASILINE SCENARIO

In the baseline scenario we simulated seed dispersal of individual seeds from a point source using the log-normal distribution (Fig. 1), and also fitted this function to the data (counts of seeds in each trap; see Model fitting and comparison). The log-normal kernel fits empirical data for several species (Greene & Calogeropoulos 2002) and it has easily interpretable parameters (mean μ and standard deviation σ of $\log r$;

the mean and standard deviation of r are simple functions of μ and σ ; Evans, Hastings & Peacock 2000). Propagule source strength and dispersal kernel parameter values were chosen to produce seed shadows of realistic dimensions for a single plant, or dense patch of plants, of a well-dispersed herb (Willson 1993). The size of the study area (i.e. maximum trap range) was set to 200×200 m on the basis of initial simulations of the true kernel. The baseline scenario involved a total trap area of 100 m^2 (0.25% of the study area) and a source of 100 000 seeds. Based on our own experience with dispersal experiments (e.g. Bullock & Clarke 2000; Skarpaas *et al.* 2004; O. Skarpaas & K. Shea, unpublished data), these are reasonable practical limits to experimental possibilities.

We tested five different trap arrangements in the simulations (Fig. 2, left column): random placement, grid arrays, transects, sectors and annuli. There is an infinite number of possible trap arrangements, but these five represent some of the most commonly used designs and cover the range from unstructured (e.g. random) to structured and directed designs (e.g. sectors). In the random, grid array and transect designs, individual traps were small equal-sized squares (0.25×0.25 m), inspired by the small traps (small sticky boards, trays, pots, buckets, etc.) used in a number of different studies (e.g. Werner 1975; Middleton 1995; Bullock & Clarke 2000). In the sector and annulus designs, the traps were continuous traps of different sizes that covered a fixed proportion of annuli centred on the seed source, such as strips of sticky tape (Skarpaas & Stabbeborg 2003; Skarpaas *et al.* 2004). The annuli differ from the other designs in that a large proportion of the trapping area is concentrated in each annulus. Given the total trap area constraints, the annuli could not be extended equally far away from the source as with the other designs. Because the distance kernel $f(r)$ is non-linear and has two parameters, we needed measurements at a minimum of three distances (i.e. three annuli) to fit the model.

SOURCE STRENGTH AND SAMPLING EFFORT

We carried out variations on the baseline scenario to test the different trapping designs for variable source strengths and sampling effort. These scenarios involved all combinations of total trap areas of 25, 50 and 100 m^2 and source strengths of 1000, 10 000 and 100 000 seeds. The trap layouts were the same as in the baseline scenario. For the reasons mentioned in the baseline scenario, we always considered annuli designs with at least three annuli, and moved the annuli further in for smaller total trap areas.

MODEL UNCERTAINTY

The functional form of the dispersal kernel is not always known. To address the issue of trap design under model uncertainty, we simulated dispersal using three different

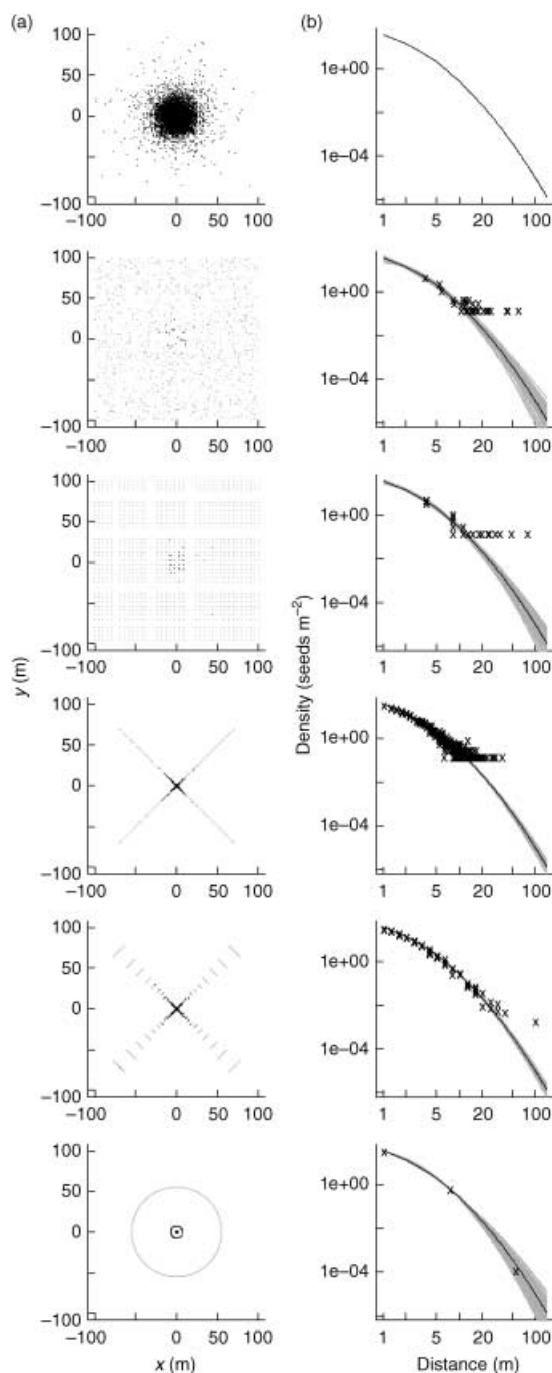


Fig. 2. Top row: 100 000 seeds (a) dispersed in the area defined by x and y in a single realization of the true kernel (log-normal (1,1), b). Five lower rows: sampling designs (grey) and corresponding trapped seeds in a single simulation (black) (a) and estimated dispersal kernels [b; densities of seeds in traps (crosses), estimated kernels in 100 simulations (grey lines) and true kernel (black line)]. The same number of seeds (100 000) was dispersed in all simulations. The sampling designs all have the same total trap area (100 m²). In the random, grid array and transect designs individual traps were small equal-sized squares (0.25 × 0.25 m). In the sector and annulus designs, the traps were continuous traps of different sizes (lines).

kernels (exponential, log-normal and half-Cauchy; Fig. 1) as the 'true' kernel, and fitted each of these kernels to each of the simulated data sets. Equations and shorthand notation for the distributions are:

$$\text{log-normal: } \exp\left(\frac{-[\log(r/\mu)]^2}{2\sigma^2}\right) / r\sigma\sqrt{2\pi} \sim \text{log-normal}(\mu, \sigma)$$

$$\text{half-Cauchy: } 2\beta/\pi(\beta^2 + r^2) \sim \text{half-Cauchy}(\beta)$$

$$\text{exponential: } \lambda \exp(-\lambda r) \sim \text{exponential}(\lambda)$$

These kernels represent a range of shapes used in dispersal studies, from the thin-tailed exponential to the fat-tailed half-Cauchy (Fig. 1; Evans, Hastings & Peacock 2000; Greene & Calogeropoulos 2002; Paradis, Baillie & Sutherland 2002). For these comparisons we used a seed source of 10 000 seeds. Other aspects (trap designs, number of simulations, etc.) were the same as in the baseline scenario. The differences in success rates for identifying the true kernel were tested using a log-linear mixed model with simulations as a random effect (Pinheiro & Bates 2000).

DISPERSAL GEOMETRY

In the baseline scenario, dispersal was assumed to be isotropic, i.e. the same in all directions. Anisotropic dispersal, i.e. dispersal patterns that are not the same in all directions, can result from systematic processes such as prevailing winds or orientated movement of dispersal vectors. In contrast to the isotropic baseline scenario, where the true distribution of dispersal directions $g(\theta)$ was uniform, we considered two anisotropic scenarios where dispersal was concentrated in one direction (with g determined according to a normal distribution with fixed standard deviation of 0.5 radians). In one of these cases the anisotropy was assumed to be unknown to the experimenters. In this case, the traps were laid out as in the baseline scenario. In the other anisotropic case, the anisotropy (i.e. the distribution function g and the mean and variance of the dispersal direction θ in radians) was assumed to be known. In this case the sampling designs were adjusted by shifting all the traps to the half of the study area in the main dispersal direction, while maintaining the basic sampling design (random, regular array, etc.). Real-world situations are likely to be somewhere between the ignorant case of unknown anisotropy and the omniscient case of fully known anisotropy.

MODEL FITTING AND COMPARISON

The dispersal models were fitted to data using numerical maximum likelihood estimation assuming a Poisson error distribution for seed counts (Clark *et al.* 1999). The Poisson likelihood of the model given the data and parameters is:

$$L = \prod_{i=1}^n \frac{\lambda_i^{c_i} \exp(-\lambda_i)}{c_i!} \quad \text{eqn 2}$$

where c_i is the observed count of individually dispersed seeds reaching seed trap i (based on the portion of the

seed shadow, equation 1, falling within trap i), n is the number of traps and λ_i is the expected seed count in trap i using the fitted probability distribution for distance (f) and the direction distribution (g). In the isotropic case, direction was uniform. In the case of unknown anisotropy, the true direction distribution was normal, but since the anisotropy was assumed to be unknown, the uniform distribution was used in the calculation of expected seed counts. In the known anisotropic case, the known true distribution of direction was used to calculate expected seed counts. The negative log-likelihood, $-\ln L$, was numerically minimized (using the simulated annealing option in the R-function 'optim'; The R Development Core Team 2003). The parameter values at the minimum of this function are the maximum likelihood parameter estimates.

To evaluate goodness-of-fit to the simulated data we used the Akaike Information Criterion (AIC; Burnham & Anderson 2002). The AIC for a given model is simply $-2 \ln L + 2K$, where K is the number of fitted parameters. By this commonly used criterion for model selection, the model with the lowest AIC is considered the best model. To compare the fitted models to the true models we used the Kullback–Leibler Information (Burnham & Anderson 2002):

$$I = \int_{r=0}^{\infty} f(r) \ln \frac{f(r)}{h(r)} dr \quad \text{eqn 3}$$

I measures the distance between the two models f and h , the true and the fitted models, respectively; the lower I , the better the fit to the true model.

Results

In the baseline (log-normal) scenario, transects and sectors performed significantly better than the other sampling designs, and annuli performed better than random and grid array arrangements (Fig. 2). The means of the parameter estimates across simulations were close to the true values for all of the sampling designs. However, the precision of the estimates varied greatly among designs; the confidence intervals were much narrower for transects and sectors than for the other designs (Fig. 3). This was especially true for long-distance dispersal (LDD).

The goodness-of-fit (I) was generally better for higher source strengths and sampling efforts, with transect and sector usually performing better than random and grid array designs (Fig. 4). The annuli design deviated from this pattern: for this design the overall fit was better for lower sampling efforts (Fig. 4). This was because of the constraint of a minimum of three traps. For lower total trap areas, the two outermost annuli were shifted inwards and therefore trapped more seeds. However, the fit improved with increasing source strength for this design as well. On average, all designs tended to converge to the true parameter values at large source strengths but the variability in the estimates was usually higher for the random and grid array than for the other designs.

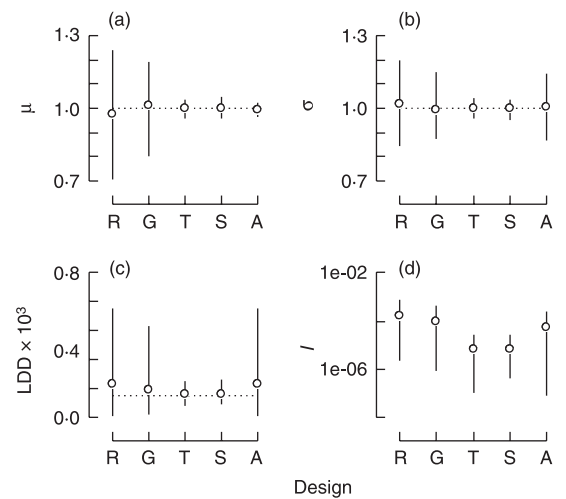


Fig. 3. Estimated parameter values (means and 95% confidence limits) for models fitted to data from 100 simulations of 100 000 seeds each dispersed according to the true kernel [$f \sim \text{log-normal}(1,1)$, $g \sim \text{uniform}(0,2\pi)$] using different sampling designs with the same total trap area (approximately 100 m²; Fig. 2). Sampling designs are random (R), grid array (G), transect (T), sector (S) and annuli (A). The panels show the mean (μ ; a) and the standard deviation (σ ; b) of the log-normal distance distribution, the frequency of long-distance dispersal (LDD; c) and goodness-of-fit (I ; d). True values are shown by the dotted line where appropriate. The frequency of LDD is defined here as the fraction of seeds dispersing > 100 m. The Kullback–Leibler information (I ; Burnham & Anderson 2002) indicates the distance between the true and the fitted model (equation 3).

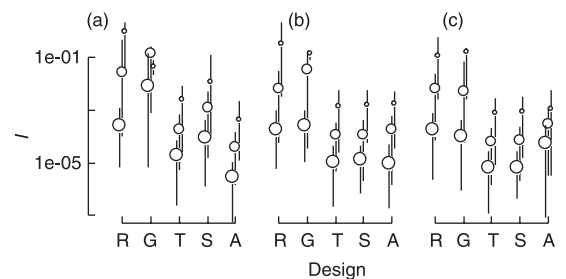


Fig. 4. Goodness-of-fit (I) with 95% confidence intervals for log-normal kernels fitted to data from different sampling designs [random (R), grid array (G), transect (T), sector (S) and annuli (A)] with increasing source strength ($Q = 1000$ seeds, 10 000 seeds, 100 000 seeds, indicated by increasing bubble size) and trap area (a, 25 m²; b, 50 m²; c, 100 m²). I expresses the difference between the true and the fitted model; a lower value means better fit (see the Methods).

This was because the random and grid array designs did not always provide sufficient data for parameter estimation for small source strengths, and both tended to overestimate the mean, μ , but underestimate the standard deviation, σ . This led to broader ranges on the estimates of LDD and worse overall fit.

When the underlying kernel was unknown, all of the trapping designs most often selected the right kernel, but the success rates were higher for the log-normal than for exponential and half-Cauchy (Table 1). None of the trapping designs was better at distinguishing between unknown kernels. Thus, as expected, the ability

Table 1. Kernel selection using different trapping designs: for each true kernel (Fig. 1), the table gives the percentage of simulations in which the fitted kernels were selected (using AIC; see the Methods). Success rates (i.e. correct selection) are marked in bold. The differences in success rates are statistically significant among true kernels (log-linear mixed model, $P < 0.01$) but not among sampling designs ($P > 0.1$)

True kernel	Exponential (E)			Log-normal (L)			Half-Cauchy (C)		
	E	L	C	E	L	C	E	L	C
Random	70	30	0	3	90	7	13	37	50
Grid array	70	30	0	5	91	4	14	36	50
Transect	76	24	0	3	88	9	14	38	48
Sector	72	26	2	3	89	8	3	39	58
Annulus	70	30	0	0	87	13	18	32	50

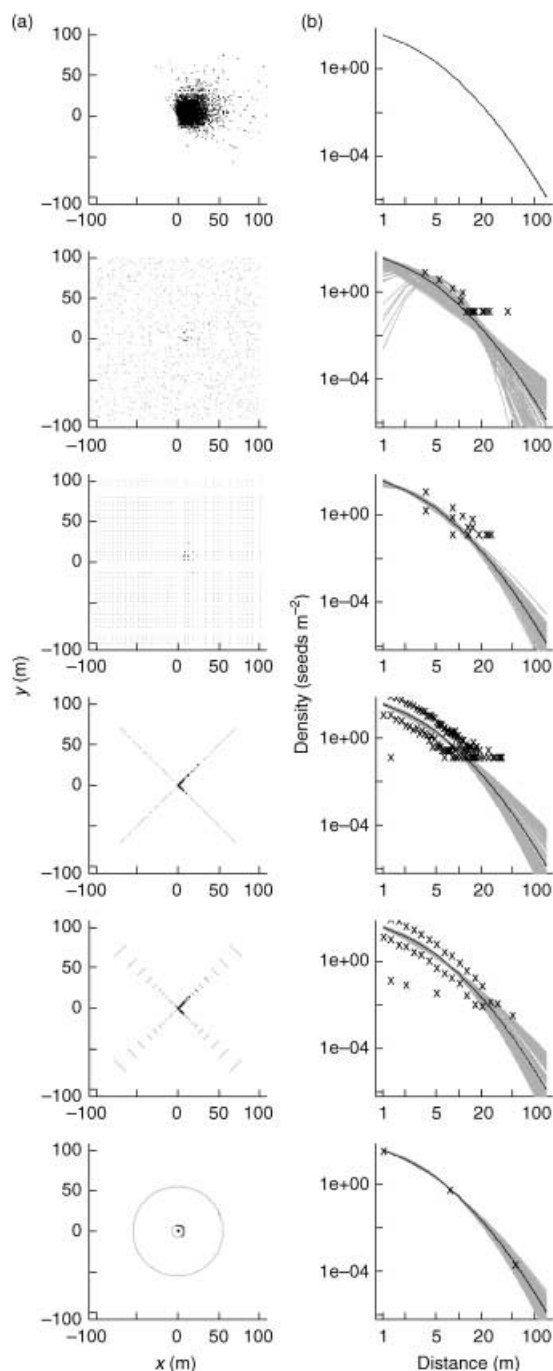


Fig. 5. As Fig. 2 but for unknown anisotropic dispersal (see the Methods).

to detect fat-tailed kernels was related to sampling effort at the tail. When no seeds were caught at the tail, the half-Cauchy was often mistaken for the log-normal (Table 1). This was because these kernels are similar at shorter distances (Fig. 1) and, while the half-Cauchy only has one free parameter, the log-normal has two and it is therefore more flexible to vary in shape.

In the case of unknown anisotropy (Fig. 5) the annuli ($I = 0.575 \times 10^{-2}$) and grid arrays ($I = 2.911 \times 10^{-2}$) performed better on average than transects and sectors ($I = 3.829 \times 10^{-2}$ and 3.759×10^{-2} , respectively). This was not unexpected given that the transect and sector designs sample only a small proportion of the circumference whereas the annuli samples the entire circumference and the grid array effectively covers more of the circumference than the transect and sector designs. However, when the anisotropy was known, and the trap arrangements were adjusted accordingly, transects and sectors were better (Fig. 6; $I = 0.018 \times 10^{-2}$ and 0.017×10^{-2} for transects and sectors, respectively, and 0.080×10^{-2} and 0.713×10^{-2} for annuli and grid arrays). This was true regardless of source strength and total trap area. Observed trap densities were higher than the true and fitted distance kernels (Fig. 6) because the fitting procedure took into consideration the known prevailing direction.

Discussion

Our simulation approach to the design of dispersal studies is useful for comparing alternative dispersal study designs (see also Stoyan & Wagner 2001) and complements analytical optimization methods for simpler design problems (such as the allocation of a sampling effort at different distances in an isotropic setting, given a fixed total sampling effort; Assunção & Jacobi 1996). The simulation approach can be used for a wide range of organisms, dispersal mechanisms, dispersal kernels and different dispersal measurement techniques. As we demonstrate, some designs are better than others, and choosing the right design for a given problem will increase the precision of dispersal kernel estimates. Such precision is not just of academic interest. Many applied problems require accurate estimates of dispersal, especially of the tail of the kernel: the responses of species to

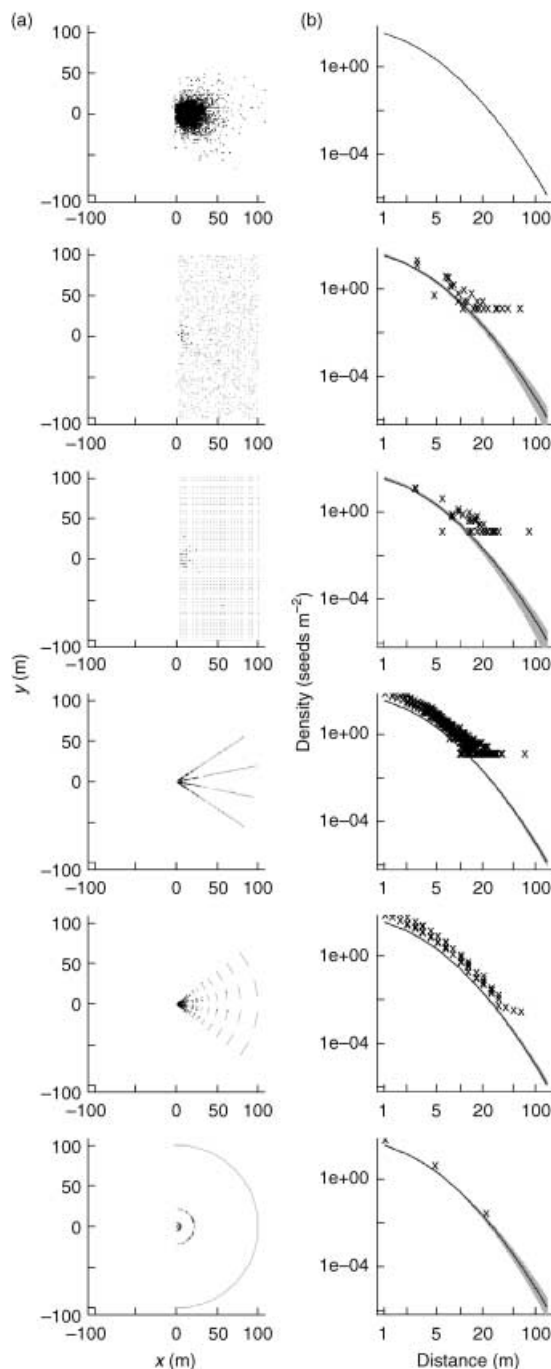


Fig. 6. As Fig. 2 but for known anisotropic dispersal (see the Methods).

habitat fragmentation (Soons & Heil 2002), the ability of species to track climate change (Watkinson & Gill 2002), range expansion of non-natives (Higgins, Richardson & Cowling 2001), recolonization by endangered species (Coulson *et al.* 2001).

To illustrate the approach with a realistic example, we have focused on a typical case in plant dispersal studies: the layout of seed traps to measure dispersal from a point source. Aspects of this design problem were addressed by Stoyan & Wagner (2001) who used simulations to evaluate alternative transect trapping arrangements assuming a log-normal dispersal kernel,

a fixed point source strength, fixed sampling effort and isotropic dispersal. In this study we explored the efficiency of several realistic two-dimensional trapping designs for different kernels, source strengths, sampling efforts and dispersal geometries.

We found that the arrangement of traps had a large effect on the quality of the data with respect to the reconstruction of the true dispersal kernel. For any given source strength and total trap area, transects and sectors were the best of the five trap arrangements tested in this study, only matched by annuli in the case of unknown anisotropic dispersal. For fat-tailed kernels (half-Cauchy), the sector design was best, albeit mediocre, at identifying the true kernel.

The confidence intervals illustrate the precision of parameter estimates based on data from different sampling designs. With a large seed source and total trap area, all of the sampling designs may give good results if the parameter estimates are averaged across a number of replicate studies. However, with a poor design the chances are great that in any one study the parameter estimates will be far from the 'true' values. Because of the time and labour costs, dispersal studies are rarely replicated. Therefore, choosing an appropriate sampling design is crucial.

As pointed out above, dispersal study design is a decision process. The optimal design will depend on the species, the dispersal mode(s) and the context. At the outset of the study design one must identify the objective of the dispersal study, the options available for measuring dispersal and the constraints on the experimental design (economic, physical and other). The simulation approach presented in this study can be used to help the design of experiments for dispersal kernel estimation, given the objectives, options and constraints. Our results suggest five general guidelines for dispersal measurement. When applying these guidelines, the objective, options and constraints must be kept in mind.

1. *Undertake pilot studies.* Pilot studies of dispersal (or data from similar species) can be used to improve layout in several ways. They help identify the appropriate scale of the study, and may reveal patterns that are useful to know in order to focus the sampling effort (e.g. anisotropy). Experimental results and simulations can be used to optimize study design in an iterative adaptive learning process. The first step in such a process is the identification of a suitable parametric or non-parametric dispersal kernel from pilot data. In our baseline example, we assumed a log-normal kernel. The second step would be to simulate n dispersal events from this kernel and compare the different sampling designs, as we have done. The third step would be an empirical study based on the best design in the simulations. The data from this experiment could then be used to refine the kernel for a new simulation to improve further the sampling design.

2. *Maximize the source strength.* Increasing the seed source improved kernel estimation for all of the dispersal geometries, sampling designs and total trap areas

considered in this study. In most cases, it is probably easier to increase source strength than trap area (see below). Increasing the source strength will in many cases involve placing several individual sources (e.g. plants) close together. However, there are a few physical constraints that must be taken into account; there is a limit to what may be considered a point source, and wind and current patterns, and hence the dispersal kernel, may be different around a dense patch of source individuals than around a single individual.

3. Maximize sampling effort. Increasing the total trap area also improved kernel estimation in all of our simulations. Increasing the trap area is often more costly than increasing the source strength. We suggest maximizing the source strength (within the constraints) and then adjusting the sampling effort to the set source strength. For a given source strength, there may be diminishing returns in increasing the trap area.

4. Sample along lines outwards from the source. The transects and sectors were the best designs in most of the cases considered in this study because they concentrate the sampling effort along the direction of dispersal outwards from the source. This gives a number of data points to determine the shape of the dispersal kernel. In some cases, hybrid sampling strategies may be useful. For instance, in the case of likely but still somewhat uncertain main dispersal directions (e.g. prevailing winds or currents), one could use a combination of annuli near the source to detect anisotropy, and transects or sectors extended in the likely main dispersal directions to detect long-distance dispersal.

5. Sample heavily at large distances to quantify long-distance dispersal. Long-distance dispersal is least easy to quantify but strongly determines the kernel fit. To measure dispersal where seed densities are low (at large distances), the trap area needs to be large, as demonstrated by the success of the sector and transect approaches. Few studies have maintained or increased the effort at large distances (but see Bullock & Clarke 2000; Skarpaas *et al.* 2004). This is costly, however, and must be balanced against sampling at high densities (short distances) to get sufficient data for kernel estimation, as illustrated by the results for the annuli design (see also Assunção & Jacobi 1996; Stoyan & Wagner 2001). Again, pilot studies are likely to be helpful in striking the appropriate balance.

These guidelines for sampling design are based on results from a wide range of conditions but still relatively simple scenarios. In some cases it may be necessary to consider more complicated situations. Two cases in point are direction-dependent and stratified or mixed dispersal kernels (Shigesada & Kawasaki 1997; Higgins, Nathan & Cain 2003). The simulation approach can still be used to address such problems, but may need different model formulations and/or additional information. Direction-dependent kernels will require a different specification of the seed shadow (equation 1) in which the distance and direction kernels (f and g) are joined. Stratified kernels may be problematic if the researcher

does not have any idea that the true kernel is stratified, or of the scale of different strata. However, once some information about the strata is available (e.g. through pilot studies or from basic knowledge of the dispersal mechanisms of the species), the approach we have described is applicable and useful. The simulation method is also applicable to other systems and sampling methods, such as collections of point sources (briefly explored by Stoyan & Wagner 2001 for small collections of trees), mark–release–recapture (Purse, Hopkins & Day 2003) and tracking of individuals (J. M. Bullock, K. Shea & O. Skarpaas, unpublished data). It requires only minor changes to the algorithms to represent appropriately the positions and strengths of propagule sources and sampling technique (e.g. recording the distances moved by all released propagules rather than the numbers of seeds in a sample determined by trap locations and sizes).

In conclusion, the simulation approach presented here is a useful framework for the design of dispersal experiments. While the results of our simulations can be used by those designing field studies of plant point sources without resorting to modelling, the general guidelines and the simulation approach are applicable to a wide range of organisms, dispersal mechanisms, dispersal models and methods for dispersal measurement. Regardless of the system, this approach can easily be incorporated in an active adaptive learning process to further improve dispersal study design. Improved dispersal designs and increased precision of dispersal kernel estimates will improve our ability to address a wide range of applied problems involving dispersal and spatial population dynamics.

Acknowledgements

We would like to thank the NCEAS working group on dispersal and demography for inspiration, members of the Mortensen and Shea laboratories at PSU for thoughtful comments and stimulating discussions, Matt Ferrari and Ottar Bjørnstad for tips on optimization in R and statistical advice, and Andy South, Eckart Winkler and several anonymous referees for helpful comments. This research was funded by the National Science Foundation (grant no. DEB-0315860 to K. Shea) and the Norwegian Research Council (grant no. 161484/V10 to O. Skarpaas).

References

- Assunção, R. & Jacobi, C.M. (1996) Optimal sampling design for studies of gene flow from a point source using marker genes or marked individuals. *Evolution*, **50**, 918–923.
- Brachet, S., Olivieri, I., Godelle, B., Klein, E., Frascaria-Lacoste, N. & Gouyon, P.-H. (1999) Dispersal and metapopulation viability in a heterogeneous landscape. *Journal of Theoretical Biology*, **198**, 479–495.
- Bullock, J.M. & Clarke, R.T. (2000) Long distance seed dispersal by wind: measuring and modelling the tail of the curve. *Oecologia*, **124**, 506–521.
- Bullock, J.M., Kenward, R.E. & Hails, R.S. (2002) *Dispersal Ecology*. Blackwell Science, Oxford, UK.

- Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Inference: A Practical Information-Theoretic Approach*, 2 edn. Springer, New York, NY.
- Cain, M.L., Damman, H. & Muir, A. (1998) Seed dispersal and the holocene migration of woodland herbs. *Ecological Monographs*, **68**, 325–347.
- Clark, J.S., Lewis, M. & Horvath, L. (2001) Invasion by extremes: population spread with variation in dispersal and reproduction. *American Naturalist*, **157**, 537–554.
- Clark, J.S., Silman, M., Kern, R., Macklin, E. & HilleRis-Lambers, J. (1999) Seed dispersal near and far: patterns across temperate and tropical forests. *Ecology*, **80**, 1475–1494.
- Coulson, S.J., Bullock, J.M., Stevenson, M.J. & Pywell, R.F. (2001) Colonization of grassland by sown species: dispersal versus microsite limitation in responses to management. *Journal of Applied Ecology*, **38**, 204–216.
- Dieckmann, U., Law, R. & Metz, J.A.J. (2000) *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*. Cambridge University Press, Cambridge, UK.
- Evans, M., Hastings, N. & Peacock, B. (2000) *Statistical Distributions*. Wiley, New York, NY.
- Greene, D.F. & Calogeropoulos, C. (2002) Measuring and modelling seed dispersal of terrestrial plants. *Dispersal Ecology* (eds J.M. Bullock, R.E. Kenward & R.S. Hails), pp. 3–23. Blackwell Science, Oxford, UK.
- van Groenendael, J.M., Ehrlén, J. & Svensson, B.M. (2000) Dispersal and persistence: population processes and community dynamics. *Folia Geobotanica*, **35**, 107–114.
- Hamrick, J.L. & Nason, J.D. (1996) Consequences of dispersal in plants. *Population Dynamics in Time and Space* (eds O.E. Rhodes, R.K. Chesser & M.H. Smith), pp. 203–234. University of Chicago Press, Chicago, IL.
- Higgins, S.I., Nathan, R. & Cain, M.L. (2003) Are long-distance dispersal events in plants usually caused by nonstandard means of dispersal? *Ecology*, **84**, 1945–1956.
- Higgins, S.I., Richardson, D.M. & Cowling, R.M. (2001) Validation of a spatial simulation model of a spreading alien plant population. *Journal of Applied Ecology*, **38**, 571–584.
- Ibrahim, K.M., Nichols, R.A. & Hewitt, G.M. (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, **77**, 282–291.
- Kriticos, D.J., Sutherst, R.W., Brown, J.R., Adkins, S.W. & Maywald, G.F. (2003) Climate change and the potential distribution of an invasive alien plant: *Acacia nilotica* ssp. *indica* in Australia. *Journal of Applied Ecology*, **40**, 111–124.
- Middleton, B.A. (1995) Sampling devices for the measurement of seed rain and hydrochory in rivers. *Bulletin of the Torrey Botanical Club*, **122**, 152–155.
- Nathan, R., Perry, G., Cronin, J.T., Strand, A.E. & Cain, M.L. (2003) Methods for estimating long-distance dispersal. *Oikos*, **103**, 261–273.
- Okubo, A., Hastings, A. & Powell, T. (2001) Population dynamics in temporal and spatial domains. *Diffusion and Ecological Problems: Modern Perspectives* (eds A. Okubo & S.A. Levin), pp. 298–373. Springer, New York, NY.
- Paradis, E., Baillie, S.R. & Sutherland, W.J. (2002) Modelling large-scale dispersal distances. *Ecological Modelling*, **151**, 279–292.
- Pinheiro, J.C. & Bates, D.M. (2000) *Mixed-Effects Models in S and S-Plus*. Springer, New York, NY.
- Purse, B.V., Hopkins, G.W. & Day, K.J. (2003) Dispersal characteristics and management of a rare damselfly. *Journal of Applied Ecology*, **40**, 716–728.
- The R Development Core Team (2003) *R 1.6.2: A Language and Environment*. The Foundation for statistical Computing, Vienna, Austria. <http://www.r-project.org>.
- Shaw, M.W. (1995) Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, **259**, 243–248.
- Shigesada, N. & Kawasaki, K. (1997) *Biological Invasions: Theory and Practice*. Oxford University Press, Oxford, UK.
- Skarpaas, O. & Stabbe, O.E. (2003) A simple seed trap and two small dispersal experiments with colt's foot *Tussilago farfara*. *Blyttia*, **61**, 218–223.
- Skarpaas, O., Stabbe, O.E., Rønning, I. & Svenningsen, T.O. (2004) How far can a hawk's beard fly? Measuring and modeling the dispersal of *Crepis praemorsa*. *Journal of Ecology*, **92**, 747–757.
- Soons, M.B. & Heil, G.W. (2002) Reduced colonization capacity in fragmented populations of wind-dispersed grassland forbs. *Journal of Ecology*, **90**, 1033–1043.
- Stoyan, D. & Wagner, S. (2001) Estimating the fruit dispersion of anemochorous forest trees. *Ecological Modelling*, **145**, 35–47.
- Tilman, D., Lehman, C.L. & Yin, C. (1997) Habitat destruction, dispersal, and deterministic extinction in competitive communities. *American Naturalist*, **149**, 407–435.
- Turchin, P. (1998) *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants*. Sinauer Associates, Sunderland, MA.
- Watkinson, A.R. & Gill, J.A. (2002) Climate change and dispersal. *Dispersal Ecology* (eds J.M. Bullock, R.E. Kenward & R.S. Hails), pp. 410–429. Blackwell Science, Oxford, UK.
- Werner, P.A. (1975) A seed trap for determining patterns of seed deposition in terrestrial plants. *Canadian Journal of Botany*, **53**, 810–813.
- Willson, M.F. (1993) Dispersal mode, seed shadows and colonization patterns. *Vegetatio*, **107/108**, 261–280.

Received 9 November 2004; final copy received 7 April 2005
Editor: Rob Freckleton