

# Why Trees Migrate So Fast: Confronting Theory with Dispersal Biology and the Paleorecord

James S. Clark

Department of Botany, Duke University, Durham, North Carolina 27708

Submitted August 25, 1997; Accepted January 22, 1998

**ABSTRACT:** Reid's paradox describes the fact that classical models cannot account for the rapid ( $10^2$ – $10^3$  m yr<sup>-1</sup>) spread of trees at the end of the Pleistocene. I use field estimates of seed dispersal with an integrodifference equation and simulation models of population growth to show that dispersal data are compatible with rapid spread. Dispersal estimates lay to rest the possibility that rapid spread occurred by diffusion. The integrodifference model predicts that, if the seed shadow has a long "fat" tail, then rapid spread is possible, despite short average dispersal distances. It further predicts that velocity is more sensitive to life history than is classical diffusion. Application of such models is frustrated because the tail of the seed shadow cannot be fitted to data. However, the data can be used to test a "long-distance" hypothesis against alternative ("local") models of dispersal using Akaike's Information Criterion and likelihood ratio tests. Tests show that data are consistent with >10% of seed dispersed as a long ( $10^2$  m) fat-tailed kernel. Models based on such kernels predict spread as rapid as that inferred from the pollen record. If fat-tailed dispersal explains these rapid rates, then it is surprising not to see large differences in velocities among taxa with contrasting life histories. The inference of rapid spread, together with lack of obvious life-history effects, suggests velocities may have not reached their potentials, being stalled by rates of climate change, geography, or both.

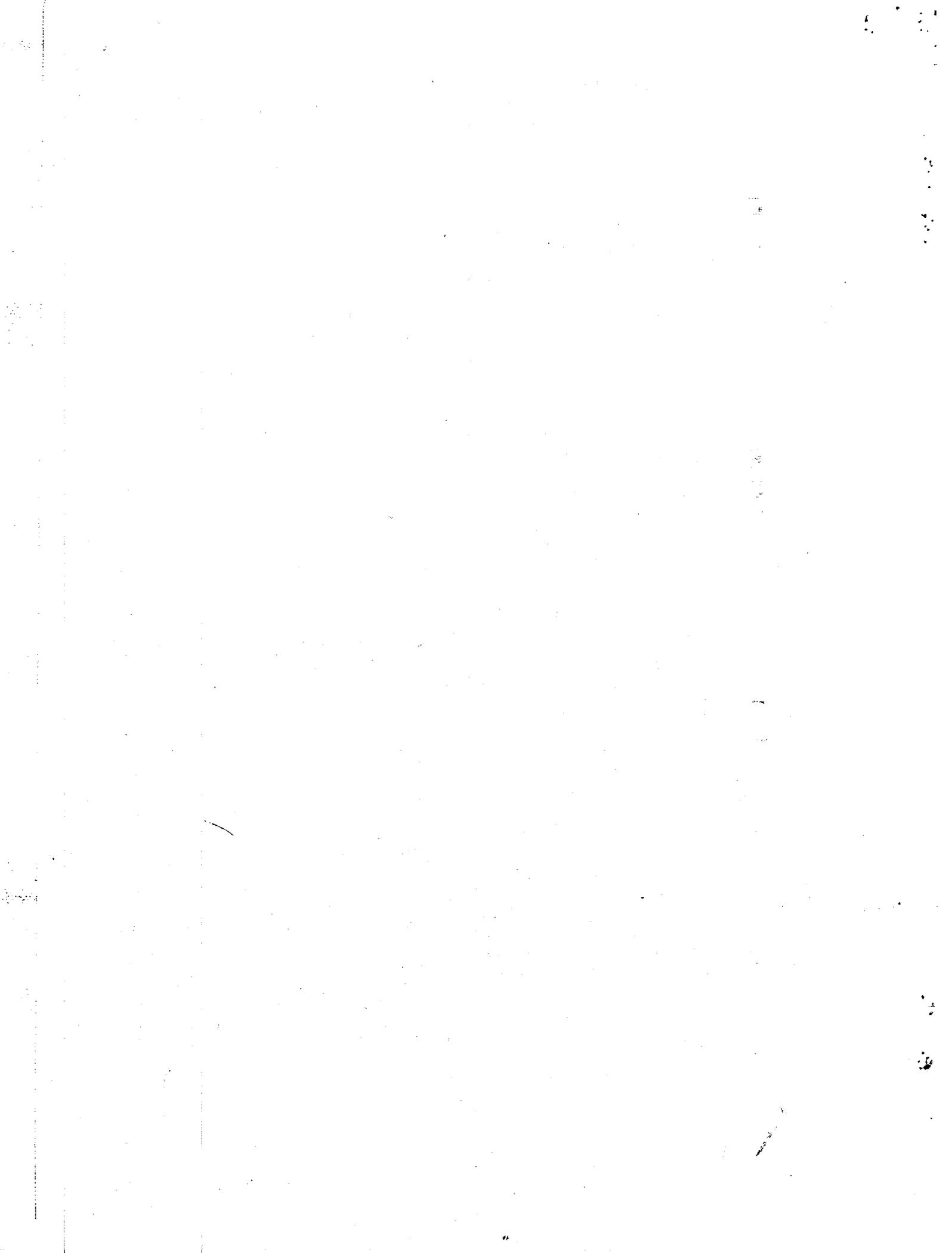
**Keywords:** climate change, diffusion, dispersal, Holocene, migration, pollen data.

About the time a Bernese patent clerk thought to describe random movements by a diffusion equation (Einstein 1905), the underlying principle of diffusion was on the mind of a Victorian botanist pondering the invasion of the British Isles by plants (Reid 1899). This early deliberation on the spread of tree populations at the end of the Ice Age, and the crossing of the English Channel in

particular, was the start of a long-standing puzzle in ecology over the disparity between the life histories of trees and their past rates of spread. This century-old problem was termed *Reid's paradox* (Clark et al. 1998a), in honor of Clement Reid's calculations of seemingly impossible seed dispersal distances needed to spread oaks into Great Britain at the end of the Pleistocene. Although diffusion was the assumption from the start, the first formal application to this paradox came a half century later and with disappointing results (Skellam 1951). The effort helped focus the conflict between life history and paleoevidence (fig. 1) and emphasized that such rates required dispersal distances impossibly great in the absence of some external aid.

Tree migration rates following the last Ice Age are truly remarkable. Since Reid's early calculations, paleoecologists have used <sup>14</sup>C-dated pollen sequences from temperate lakes (Davis 1976; Huntley and Birks 1983; Delcourt and Delcourt 1987; Birks 1989; MacDonald 1993) to estimate rates of spread exceeding  $10^2$  m yr<sup>-1</sup> for many tree genera. The fossil record is virtually the sole evidence for these high rates (but see Fastie 1995); without it, ecologists might arguably believe such rates impossible. For instance, tropical biogeography is widely viewed to reflect Pleistocene refugia (reviewed by Bush 1994) because populations are thought incapable of extended, rapid migration. Limited migration is a notion still tenable in the Tropics, where hypotheses are poorly constrained by paleodata. The possibility that limited dispersal could threaten plant populations with future global change has renewed interest in explaining rapid migrations of the past (Clark 1993; Pacala and Hurtt 1993; Mellilo et al. 1996; Pitelka et al. 1997).

Just as enigmatic as rapid spread is an apparent unimportance of life history. Simple diffusion predicts that reproduction and dispersal determine migration rate. Yet genera as different in these traits as birch and beech obtained high rates of spread. If dispersal and reproduction provide for the rapid rates in the paleorecord, we expect that taxa displaying the most rapid spread to be those that mature early, possess high reproductive capacity,



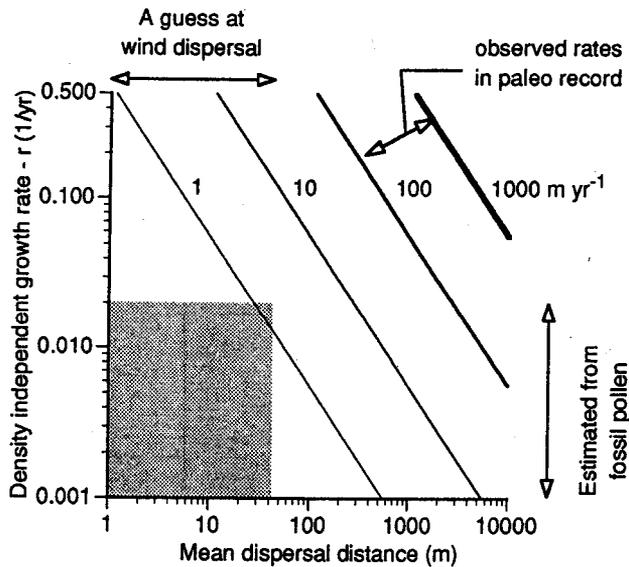


Figure 1: Rates of spread (eq. [6]) predicted by a diffusion model for mean dispersal distances  $\mu'$  with a Gaussian ( $c = 2$ ) dispersal kernel and density-independent growth rates  $\sqrt{\ln R_0'/T}$ . The shaded region indicates parameter space conforming to estimated dispersal distances from field data (Clark et al. 1998b) and growth rates typically reported from rates of increase in fossil pollen data (Clark 1993).

and disperse seed long distances. Either such relationships are masked in the paleorecord or taxa do not differ sufficiently in life history to have much effect.

Reid's paradox persists because traditional theory has been unable to account for rates of early Holocene spread and because seed dispersal has been too poorly understood to motivate much new model application. Since Skellam's demonstration that a simple model had little to offer beyond showing that tree migration is not diffusion, there have been few efforts to fit migration models to paleodata (Dexter et al. 1987). Migration theory has moved beyond simple diffusion, but models typically require information that dispersal biologists and paleoecologists lack. Models of population spread appropriately concern idealized life histories that often consist of no more than a dispersal parameter and an "intrinsic rate of increase" or net reproductive rate. There has been little guidance on how to collapse complex life histories into such parameters and, then, how to estimate them. Analytical models are likewise of limited application if dispersal is poorly characterized (Skellam 1951; Mollison 1991; Higgins et al. 1996). Until recently, seed production and dispersal estimates have been highly indirect; many estimates come from spatial patterns of seedlings (rather than of seed) relative to adults, which confounds seed dispersal with density-dependent germination and survival (Janzen 1970; Hubbell 1980). Identifying parents

of seedlings under closed canopies where seed shadows overlap is difficult. Seed distributions measured in open fields may be unrepresentative of closed canopies. Long-distance dispersal is especially difficult to estimate (Willson 1993; Greene and Johnson 1995).

Advances in migration models and in dispersal biology permit reanalysis of Reid's paradox. Paleoecologists (Davis 1987; Prentice 1992; Clark 1993) and theorists (Mollison 1972, 1991; Shigesada et al. 1995; Kot et al. 1996) alike have begun to embrace a concept of dispersal that extends beyond simple diffusion to include rare, long-distance dispersal. Recent models of spread that accommodate a variety of dispersal patterns (Kot et al. 1996; Lewis 1997) together with new methods to quantify dispersal (Ribbens et al. 1994; Clark et al. 1998b) provide means for parameter estimation and a framework for analysis.

Resolving the paradox means finding a minimal model that, when parameterized with life-history data, is consistent with paleoevidence. Here I use a tractable model of spread parameterized with dispersal data to account for Holocene tree spread. I preface this analysis with an outline of the model, life-history details needed for its application to trees, and predictions relevant for post-Glacial spread. The subsequent description of dispersal modeling and data is the basis for the first of two components in this analysis. The first component tests the "long-distance dispersal hypothesis," which entails finding a range of dispersal patterns that are compatible with field data. The second component uses these dispersal patterns with life-history data in analytical and simulation models to test whether the data can produce the velocities of spread inferred from paleorecords.

#### Velocities of Spread and the Necessary Parameters

Recent models capable of predicting rapid spread differ from classical diffusion (Skellam 1951) only in the shape of the dispersal kernel. "Local dispersal," assumed by diffusion, yields traveling wave solutions (Kolmogorov et al. 1937). Diekmann (1978, 1979; see also Thieme 1979; van den Bosch et al. 1990; Mollison 1991) describes a general approach that begins with the reproductive schedule and seed shadow as a joint "reproduction and dispersal" density. The population model involves a spatial renewal equation that describes the fecundity schedule and dispersal. Assuming a traveling wave solution exists, we arrive at a moment (Mollison 1977, 1991) or (upon taking logs) cumulant (van den Bosch et al. 1990) generating function for the dispersal kernel. For "exponentially bounded" kernels (i.e., lacking a fat tail) the moment integrals converge, permitting asymptotic solutions for wave velocity (app. A). For kernels fatter than exponen-

tial, the moment integrals do not converge, and the velocity of spread increases indefinitely. Kot et al. (1996) use moment-generating functions to solve integrodifference equations for exponentially bounded kernels but also provide an approximation when dispersal is fatter than exponential. Here I summarize model assumptions, describe how its parameters relate to life history, demonstrate how the shape of the kernel might affect velocity, and derive the influence of life history on that velocity.

#### Model Assumptions

Assume dispersal and population spread occur along a one-dimensional transect. An integrodifference equation model expresses density  $N_t$  at time  $t + T$  as a function  $R(N_t(y))$  of density at time  $t$  and dispersal

$$N_{t+T}(x) = \int_{-\infty}^{\infty} f(x-y)R(N_t(y))dy.$$

The dispersal kernel  $f(x-y)$  is a probability density function, and  $T$  is generation time. The production of seed at location  $y$  is  $R(N_t(y))$ , and the fraction of seed produced on the interval  $(y, y+dy)$  at time  $t$  and arriving at  $x$ ,  $T$  yr later, is approximately  $f(x-y)dy$ . The dispersal kernel  $f(x-y)$  does not depend on location of propagule source but only on source distance  $|x-y|$ . The growth function  $R(N_t(y))$  is complex for tree populations. Fortunately, much of this complexity can be ignored for purposes of estimating population spread, because nonlinear reproduction, growth, and survival in a competitive environment have minimal impact at the population frontier. If reproductive success is higher in the unoccupied habitat than in the competitive population interior (van den Bosch et al.'s [1990] "linear conjecture"), then we can estimate rates of spread from the linearized equation

$$N_{t+T}(x) = R'_0 \int_{-\infty}^{\infty} f(x-y)N_t(y)dy, \quad (1)$$

where  $R'_0$  is the net reproductive rate at low density,  $R'_0(0)$ . The unknowable effects of interspecific interactions with resident populations in the Holocene will reduce rates of spread below those we derive from idealized equation (1).

#### Model Parameters from Tree Life History

Equation (1) describes spread of an idealized population with discrete reproduction summarized by three elements: net reproductive rate  $R'_0$ , generation time  $T$ , and the dispersal kernel  $f(x)$ . These elements predict diffusion, and they provide basis for analysis of more complex types of spread. Net reproductive rate  $R'_0$  and generation

time  $T$  summarize complex life-history schedules, and they depend on the rate of population growth. For comparison with simulations that follow, I estimate generation time weighted by density-independent population growth rate  $r$  from Lotka's equation

$$1 = \int_{t_1}^{t_2} e^{-ra} m(a)l(a) da, \quad (2)$$

where  $m(a)$  and  $l(a)$  are fecundity and survivorship schedules, respectively,  $t_1$  is maturation age, and  $t_2$  is longevity. This estimate of  $r$  permits calculation of generation time  $T$ :

$$T = \int_{t_1}^{t_2} a e^{-ra} m(a)l(a) da \quad (3)$$

(Leslie 1966). Net reproductive rate is

$$R_0 = \int_{t_1}^{t_2} m(a)l(a) da. \quad (4)$$

All parameters ( $r$ ,  $R'_0$ , and  $T$ ) apply to the density-independent case (see below).

I use a two-parameter kernel that has some useful attributes: it includes published dispersal kernels as special cases, it can be exponentially bounded or fat tailed and so yields both diffusive and accelerating spread, and it fits seed dispersal data (Clark et al. 1998b):

$$f(x) = \frac{c}{2\alpha\Gamma(1/c)} \exp\left(-\left|\frac{x}{\alpha}\right|^c\right), \quad (5)$$

where  $\Gamma(\cdot)$  is the gamma function, and  $\alpha$  and  $c$  are distance and shape parameters, respectively (fig. 2). Although fitted to two-dimensional seed rain data (see next section), I use the one-dimensional kernel (5) to emphasize spread in a single direction and to place my results within the context of many classical models. The symmetry of the distribution means that odd moments (mean, skewness, etc.) equal 0, and kernel shape is summarized by even moments, the first two being the mean-squared distance traveled by a seed (second moment)—

$$\mu_2 = \frac{\alpha^2\Gamma(3/c)}{\Gamma(1/c)}$$

—and kurtosis, or "fatness," of the kernel, which depends only on parameter  $c$ —

$$\frac{\mu_4}{\mu_2^2} = \frac{\Gamma(5/c)\Gamma(1/c)}{\Gamma^2(3/c)}$$

The conventional "diffusion coefficient"  $D$  is proportional to the mean square displacement (second moment),  $2D = \mu_2$ , when  $c = 2$ , provided we scale time in discrete units. Although mean dispersal is 0 (seeds traveling in opposite directions balance), it is useful to relate

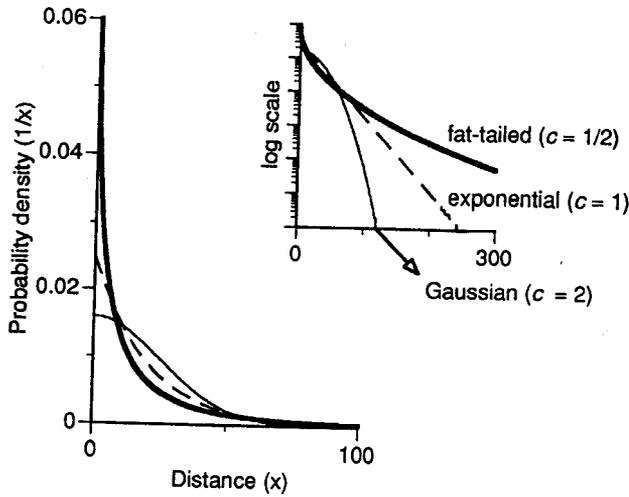


Figure 2: Example dispersal kernels from equation (5) with three different values of the shape parameter  $c$  shown on linear and log (inset) scales (bold solid line = fat tailed; dashed line = exponential; thin solid line = Gaussian). “Exponentially bounded” kernels ( $c \geq 1$ ) decay more rapidly than the exponential ( $c = 1$ ). “Fat-tailed” kernels ( $c < 1$ ) decay more slowly than exponential.

these quantities to the average distance a seed travels from the source, the quantity dispersal biologists unconcerned with direction typically estimate as

$$\mu'_1 = \frac{\alpha \Gamma(2/c)}{\Gamma(1/c)}$$

Velocities

The velocity of expansion that a population achieves depends on life history and on kernel shape (app. A). An exponentially bounded kernel tends to the traveling wave; for  $c = 2$ ,

$$\lim_{t \rightarrow \infty} C(t) \rightarrow \frac{\alpha}{T} \sqrt{\ln R'_0} \tag{6}$$

This result obtains as the limit of equation (A4), derived from integrodifference equation (1). The discrete time velocity in equation (6) describes population spread that is paced by delayed maturation and thus is slower than purely continuous reproduction, which yields a velocity of

$$C_c = \alpha \sqrt{\frac{\ln R'_0}{T}} \tag{7}$$

(Kolmogorov et al. 1937). The “stepwise” spread described by (6) results in a coherent population front moving at asymptotically constant “average” velocity (fig.

3A, B), which depends on  $R'_0$  and  $T$ , and on the distance parameter of the dispersal kernel,  $\alpha$ .

“Fat-tailed” kernels ( $c < 1$ ) are not exponentially bounded (are fatter than exponential; fig. 2) and do not approach constant rates of spread. Instead, spread accelerates as rare, long-distance dispersal produces outlying populations. The “population front” is noisy and hard to characterize as outliers establish and coalesce (Mollison 1977; Lewis 1997).

Mixed kernels describe dispersal when distinct processes control movement under different circumstances (Goldwasser et al. 1994; Metz and van den Bosch 1995; Shigesada et al. 1995; Lewis 1997). I use a mixed kernel to describe long-distance dispersal due to rare events such as storms or by animals. Appendix A derives the velocity of spread for a mixed kernel that includes a fraction  $p$  that is dispersed nearby according to an exponentially bounded component  $f_s(x)$  and fraction  $(1 - p)$  dispersed long distances according to a fat-tailed kernel  $f_l(x)$ ,

$$f(x) = pf_s(x) + (1 - p)f_l(x) \tag{8}$$

For the case of  $c = 1/2$ , the rate of spread increases linearly (A6), and the front flattens over time (fig. 3B).

How Much Should Life History Matter?

The shape of the kernel affects not only the rate of spread but also the importance of life history. Appendix B demonstrates these effects using sensitivity coefficients  $S_j$ , the proportionate effect on velocity  $C(t)$  produced by a proportionate change in parameter  $j$ . For diffusion ( $c = 2$ ), fecundity is important only if net reproductive rate is extremely low (B1). Dispersal and generation time are more important than fecundity, having  $S_\alpha = 1$  and  $S_T = -1$ , respectively.

Fat-tailed dispersal amplifies the importance of life history. Fecundity and generation time have stronger effects on velocity for fat-tailed kernels than for diffusion, the limits of equations (B2b) and (B2c) tending to  $\lim_{t \rightarrow \infty} S_\beta \rightarrow 2/\ln R'_0$  and  $\lim_{t \rightarrow \infty} S_T \rightarrow -2$ , respectively (fig. 3D). As in the case of diffusion, velocity scales with dispersal distance, the asymptote of equation (B2a) being  $\lim_{t \rightarrow \infty} S_\alpha \rightarrow 1$ . It is straightforward to demonstrate that net reproductive rate and generation time become increasingly important with increasing fatness of the kernel.

Thus, the generalization of dispersal to include fat-tailed kernels provides one means of obtaining spread more rapid than diffusion. Discrete reproduction predicts slower spread than continuous reproduction. Fat-tailed kernels exaggerate the effects of life history on rates of spread. Assessing whether fat-tailed kernels predict the rates observed in the paleorecord requires estimates of

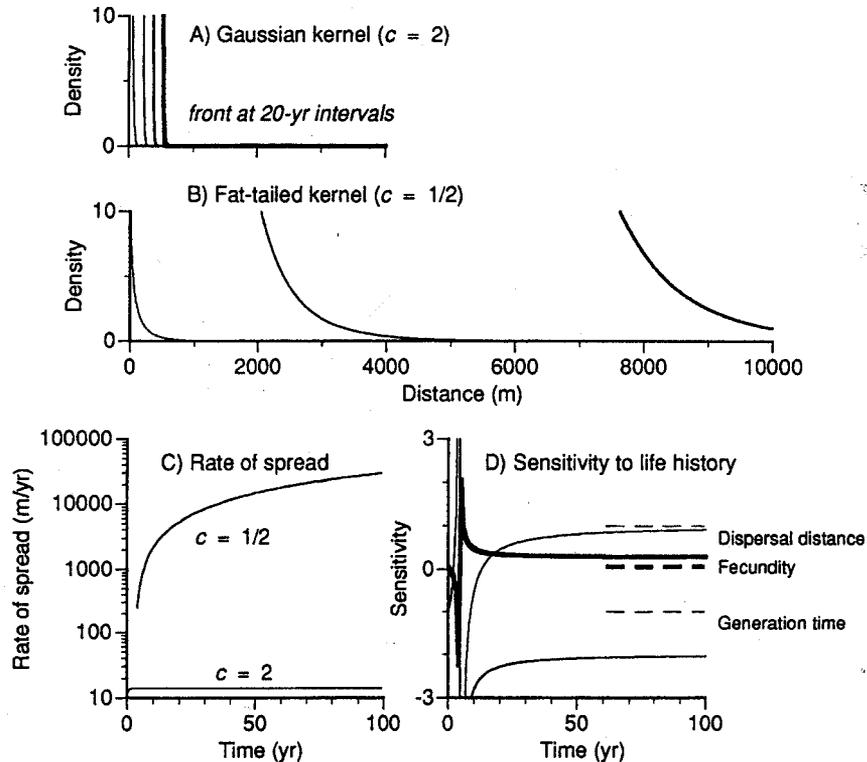


Figure 3: Population fronts at 20-yr intervals for (A) Gaussian and (B) fat-tailed kernel. C, Rates of spread and (D) sensitivities to life-history parameters for a fat-tailed kernel ( $c_1 = 1/2$ ,  $p = 0.05$ ,  $\alpha_1 = 100$  m) compared with Gaussian dispersal for *Acer rubrum*.  $T = 5.8$  yr,  $R_0' = 1325$ ,  $\alpha_s = 30.8$  m. The sensitivities in (D) show the proportionate effect on spread velocity of a proportionate change in life-history parameters. Solid lines are for the fat-tailed kernel in (B), and they are contrasted with the dashed lines for a diffusion model (A).

net reproductive rate, generation time, and the dispersal kernel.

#### What Do the Dispersal Data Say?

Recent estimates of seed dispersal in forests are incompatible with post-Glacial migration rates. Although Skellam (1951) and others suspected this to be the case, there existed no estimates of seed dispersal under closed canopies that could be used to test models of population spread (e.g., eq. [1]). Most seed shadows have been estimated from spatial patterns of seedlings, often in open fields, clearcuts, parking lots, or hedgerows. More recently, dispersal has been estimated from actual seed rain in closed stands (Clark et al. 1998b). The outline of the estimation procedure that follows is the basis for my analysis of life-history effects on migration rate and a test of the long-distance dispersal hypothesis.

Seed shadows were estimated from 5 yr of seed rain collected in 100 traps located in five 60-m  $\times$  60-m mapped stands in the southern Appalachians (stand de-

scriptions, methods, and data are detailed in Clark et al. 1998b). The summed seed shadow (SSS) model permits estimation of the parameters describing fecundity ( $\beta$ , seed production per centimeter-squared basal area), dispersal distance ( $\alpha$ , meters), and clumping ( $\theta$ , dimensionless). Values of  $\theta \ll 1$  indicate highly clumped seed, whereas large values tend to a Poisson process. The likelihood function is a negative binomial with "mean" seed arrival  $\hat{s}(\mathbf{x}, \mathbf{b}|\beta, \alpha)$ , depending on a matrix of distances between trees and seed traps  $\mathbf{x}$  and a vector tree basal areas  $\mathbf{b}$ ,

$$L(\mathbf{S}|\beta, \alpha, \theta) = \prod_{j=1}^{m \text{ seed traps}} \frac{\Gamma(s_j + \theta) \hat{s}(\mathbf{x}, \mathbf{b}|\beta, \alpha)^{s_j} \theta^\theta}{\Gamma(s_j + 1) \Gamma(\theta) [\hat{s}(\mathbf{x}, \mathbf{b}|\beta, \alpha) + \theta]^{s_j + \theta}}, \quad (9)$$

where  $\mathbf{S}$  is the ensemble of seed traps and  $s_j$  is the density of seed observed in the  $j$ th seed trap. The mean seed shadow is the summed contributions of individual trees  $i$  weighted by their distances  $x_{ij}$  and basal areas  $b_i$ ,

$$\hat{s}_i(\mathbf{x}, \mathbf{b}|\beta, \alpha) = \sum_{i=1}^{n \text{ trees}} \beta b_i f(x_{ij}|\alpha), \quad (10)$$

with dispersal kernel given by equation (5) with  $c = 2$ . The fitted seed shadow is represented by values of  $\beta$ ,  $\alpha$ , and  $\theta$  that maximize the likelihood of the observed seed rain densities  $S$ . The fitted two-dimensional kernel in (10) describes seed rain in any given direction that differs by a constant from that of the one-dimensional kernel (5). The proper normalization constant is critical for parameterization of the kernel from field data (Clark et al. 1998b). Because the analysis below includes a density-independent germination  $g$  (see "Parameterizing Life-History Schedules"), the scaling difference between one- and two-dimensional kernels is absorbed by  $g$ .

The model was fitted to data within individual stands and tested against the null model of nonlocal dispersal, that is, seed rain is independent of locations of parent trees (Clark et al. 1998b). The null model was rejected for 11 of 14 tree taxa, the three exceptions being taxa with few trees, rare seed, or both. Comparisons among stands demonstrated that dispersal for a species is consistent among stands for all but the three best-dispersed taxa because some distant trees (outside mapped plots) contributed seed to our traps (Clark et al. 1998b).

Dispersal distances and fecundities estimated for these species (table 1) lay to rest the hypothesis that diffusion (eq. [6]) might explain rapid post-Glacial migrations.

The growth rates needed to produce rates of spread  $>10^2$  m yr<sup>-1</sup> are impossibly great (fig. 1). These mean dispersal estimates confirm the need to examine seed dispersed long distances.

Understanding the contribution of long-distance dispersal to rapid population spread requires the three elements of the integrodifference equation model (1), including life-history schedules (summarized by  $R'_0$  and  $T$ ) and seed dispersal  $f(x)$ . Despite providing confident predictions at local ( $<10^2$  m) scales, the SSS model does not estimate "long-distance" dispersal. Seeds arriving from distant points are rare; they constitute noise in the data that cannot be resolved with a parametric model. Although we cannot fit the tail, we can test hypotheses about the tail using seed rain data from mapped plots. The next section describes an inverse approach to assess the degree to which long-distance dispersal may contribute to seed rain.

### Methods

My analysis consists of three steps. First, I test the "long-distance dispersal hypothesis," wherein I establish the extent to which dispersal data from modern forests admit a long-distance tail. Second, I construct a simulation model of tree population spread based on fitted parameters to assess characteristics of the tail needed to match the rates of spread inferred from the paleorecord. Finally, I invert

Table 1: Parameters used in simulations and for calculation of life-history summary parameters  $R'_0$  and  $T$  in the integrodifference model

Taxon	Summed seed shadow model*			Life-history parameters	
	Fecundity $\beta$ (cm <sup>-2</sup> yr <sup>-1</sup> )	Dispersal $\alpha$ (m)	Clumping $\theta$ (dimensionless)	Maturation age $t_1$ (yr)	Longevity $t_2$ (yr)
<i>Acer rubrum</i>	73.1 ± 7.31	30.8 ± 3.80	10.1 ± 4.36	8	80
<i>Betula lenta</i>	1,418 ± 125	34.2 ± 3.08	6.40 ± 1.94	40	150
<i>Carya glabra</i>	1.01 ± .276	10.1 ± 4.10	.490 ± 1.61	30	200
<i>Cornus florida</i>	2.53 ± 1.50	7.84 ± 4.45	.394 ± 6.31	10	125
<i>Fraxinus americana</i>	3.40 ± .788	19.3 ± 4.33	.632 ± .716	37	260
<i>Liriodendron tulipifera</i>	133 ± 13.9	30.7 ± 2.30	26.5 ± 1.37	20	200
<i>Nyssa sylvatica</i>	7.33 ± 1.63	6.68 ± 5.87	2.28 ± 2.10	30	200†
<i>Pinus rigida</i>	2.16 ± .466	15.1 ± 2.92	2.01 ± 2.69	12	100
<i>Quercus rubra</i>	5.56 ± 3.07	12.9 ± 1.80	1.22 ± 1.60	50	200
<i>Robinia pseudoacacia</i>	2.19 ± 3.46	3.52 ± 2.10	.133 ± .143	6	200†
<i>Tilia americana</i>	8.95 ± 4.26	17.34 ± 4.83	.680 ± 2.66	30	140
<i>Tsuga canadensis</i>	9.14 ± 2.38	22.8 ± 8.78	∞‡	15	500

\* Parameter estimates ± 1 SE. Confidence intervals, parameter correlations, and hypothesis tests are included in Clark et al. (1998b).

† Approximate estimates were used where unavailable (Burns and Honkala 1990).

‡ *Tsuga* seed rain was overdispersed (more regular than Poisson), and thus the clumping parameter estimate tends to ∞ (Clark et al. 1998b).

the problem and ask whether dispersal kernels consistent with paleorecords are compatible with field data.

The notion of "compatibility" between data and model bears explanation. I pose two opposing hypotheses regarding the nature of dispersal. Hypothesis  $H_0$  maintains that dispersal is local, being described by a single kernel with low dispersal parameter and kurtosis. Hypothesis  $H_1$  is the competing view that some dispersal moves long distances, as described by a large dispersal parameter and kurtosis. I confront dispersal data with kernels possessing long-distance tails that vary in three respects, each of which might affect rates of spread: the "fraction" of the dispersal kernel in the tail, the "length" of the tail (summarized by a distance parameter), and the shape (variance and kurtosis) of the tail. The test of the long-distance dispersal hypothesis sidesteps the problem of "fitting" a tail. Instead I ask how much and what sort of long-distance dispersal the data will admit. The approach consists of appending tails to kernels fitted to seed rain data (eq. [8]) and using the data to arbitrate between the competing hypotheses  $H_0$  and  $H_1$ . Data incompatible with long-distance tails are unlikely relative to the competing model of local dispersal. Tests are used to generate probability surfaces for combinations of tail fraction and length for "fat" ( $c = 1/2$ ) and "thin" tails ( $c = 2$ ).

#### *How Much Tail Will the Data Allow?*

I use a model of long-distance dispersal that admits as a null model the kernel for local dispersal. The likelihood of the long-distance model is identical to equation (9), with the exception of two additional parameters  $\alpha_1$ , a dispersal parameter for the tail, and  $1 - p$ , the fraction of the kernel allocated to the tail. The mixed kernel from equation (8),

$$f(x|\alpha, \alpha_1, p) = [pf_1(x_j|\alpha, c = 2) + (1 - p)f_1(x_j|\alpha_1, c_1 = 1/2 \text{ or } 2)], \quad (11)$$

collapses to the local model (5) when  $p = 1$ . Decreasing values of  $p$  represent increasingly larger fractions devoted to the tail. The shape parameter for the tail was assumed to be either fat ( $c_1 = 1/2$ ) or Gaussian ( $c_1 = 2$ ).

Standard data modeling would involve finding the parameters for equation (11) that maximize the likelihood (9). Using a likelihood ratio test, a comparison of this likelihood with that for the nested local model ( $p = 1$ ) would determine whether the long-distance model provided a significant improvement over the local model. Unfortunately, model (11) is "ill-conditioned" due to parameter redundancy; equivalent likelihoods are achieved with alternative parameter combinations (Seber and Wild

1989). We thus require a creative approach to model comparison. I first treat the models as nonnested with the view that each represents the data under different assumptions concerning  $p$  ( $p = 1$  for the local model, and  $p < 1$  for the regional model), but neither model actually fits this parameter. The lowest value of Akaike's Information Criterion,  $AIC = -2 \ln L + 2 \times$  (number of parameters) (Akaike 1992), represents the best fit. I determined the difference in AICs for competing models:

$$\ln L_0(S|\hat{\alpha}, \hat{\theta}, p = 1) \text{ vs. } \ln L_1(S|\hat{\alpha}, \alpha_1, \hat{\theta}, p < 1),$$

where each has two fitted parameters but different assumptions about  $p$ . AIC identifies the model most compatible with the data, but it does not have an associated probability.

Second, I assumed that the data contained a long-distance component ( $p < 1$ ), and thus treated  $p$  as an added parameter. Although we cannot fit  $p$ , we can calculate the likelihood of the data under different assumptions about  $p$ . The local model is viewed as being nested within the long-distance model when  $p = 1$ . Likelihood ratio tests compare the long-distance model  $H_1$  ( $p < 1$ ) with the local model  $H_0$  ( $p = 1$ ). The two models have different fitted parameters that determine degrees of freedom. The shape parameters  $c$  and  $c_1$  are fixed in both models. I completed fits for two shapes of long-distance tails (two values of  $c_1$ ). The fraction in the tail ( $1 - p$ ) and the "length" of the tail  $\alpha_1$  are dimensions I explore explicitly; they are not fitted parameters, but are fixed for values of  $p$  ranging from 1 (i.e., local dispersal only) to 0.9 (10% of the kernel in the tail) and for  $\alpha_1$  ranging from 25 to 1,000 m. Because  $c$  and  $\alpha_1$  are fixed (i.e., held constant for any given fit), they do not contribute degrees of freedom. Fecundity parameter  $\beta$  is fixed because the maximum likelihood (ML) estimate for the local model (eqq. [9]–[10]) agreed with estimates calculated independently (Clark et al. 1998b). For each combination of fixed parameters, I found ML estimates of  $\alpha$  and  $\theta$  for  $L_1$  that include a long-distance tail. I determined the probability of this likelihood relative to the local model  $L_0$  using the deviance

$$D = -2[\ln L_0(S|\hat{\alpha}, \hat{\theta}) - \ln L_1(S|\hat{\alpha}, \hat{\theta}, p)], \quad (12)$$

which is distributed as  $\chi^2$  with 1 degree of freedom. A probability of .05, for example, indicates we might reject the long-distance dispersal hypothesis  $H_1$  with 95% confidence.

The degree to which the data admit a long-distance tail depends not only on the fraction of seed arriving in traps that derive from long-distance dispersal but also on other sources of variance. In order to identify factors affecting the fits, I determined percentiles for local kernels

both for the estimation error (confidence intervals on  $\hat{s}(x_{ij}, b_i | \beta, \alpha)$ ; hereafter,  $\hat{s}$ ) and for the degree of clumping about  $\hat{s}$  (eq. [9]). The estimation error about  $\hat{s}$  was propagated from 500 bootstrapped estimates of  $\beta$  and  $\alpha$  for each data set (each species). Percentiles for the negative binomial at distance  $x$  are defined by mean parameter  $\hat{s}$  and clumping parameter  $\theta$ .

The above method uses the model of local dispersal as the basis for evaluating long-distance dispersal. If there is low power in this baseline fit, then the long-distance hypothesis can describe the data as well as the local model simply because we lack power to discriminate. The probability of identifying that the true  $\alpha$  lies outside the confidence interval (the probability of not making a Type II error) describes whether "local" dispersal might in fact be greater than the ML estimate of  $\alpha$  seems to indicate. The confidence with which we could discriminate large distance parameters was assessed with bootstrapped power curves for  $\alpha$ . Probabilities of  $\alpha$  for each bootstrap sample were found using the likelihood profile based on deviance:

$$D = -2[\ln L_k(S|\hat{\beta}, \alpha_k, \hat{\theta}) - \ln L(S|\hat{\beta}, \hat{\alpha}, \hat{\theta})] \quad (13)$$

(Efron and Tibshirani 1993). Parameter  $L_k$  is the likelihood of the data under that hypothesis that the dispersal parameter assumes the value  $\alpha_k$  and all other parameters are ML estimates. Likelihood ratio tests with 1 degree of freedom determine a probability for each  $\alpha_k$ . The fraction of samples in which the probability of each  $\alpha_k$  falls below 0.025 defines the power curve, that is, the chances of rejecting the hypothesis when it is indeed false.

*Parameterizing Life-History Schedules*

The analytical (eqq. [1]–[4] and app. A) and simulation models use parameters estimated from the same stands used to estimate seed dispersal. The fecundity schedule is

$$m(a) = \begin{cases} g\beta b(a) & t_1 \leq a \leq t_2 \\ 0 & \text{otherwise} \end{cases}, \quad (14)$$

where fecundity parameter  $\beta$  maximizes the likelihood of the data given the model (9), and  $g$  is the germination fraction. Maturation ages  $t_1$  and longevities  $t_2$  were taken from the literature (table 1). The "growth schedule"  $b(a)$  was estimated from diameter increments measured on the five long-term census plots (P. Wyckoff and J. S. Clark, unpublished manuscript). Although these growth schedules include the effects of density, those effects have minimal impact on population spread (see "Discussion"). The "germination fraction"  $g$  subsumes the many factors affecting seedling success that have weak correla-

tion with density. These fractions from our stands range from 0 to 0.8 seedlings per seed. Rather than fit this parameter, I chose a low value ( $g = 0.001$ ) in order to make rapid migration difficult. The survivor function  $l(a)$  (eqq. [2]–[4]) assumed minimal mortality once established, in keeping with density-independent growth. The simulation model, however, does include density-dependent thinning (next section).

*Simulating Population Spread*

The stand simulation model is modified from that in a previous article (Clark and Ji 1995) to incorporate age structure and fitted kernels. I assume a transect of contiguous 10 m  $\times$  10 m patches. The population is initialized with four individuals of reproductive age at one end of the transect. Individuals grow in basal area and thin, depending on crowding level:

$$G \propto \sum_{a=1}^{t_2} n_a b_a, \quad (15)$$

where  $n_a$  and  $b_a$  are density and basal area of age class  $a$ . New individuals become established in the understory so long as crowding  $G < 0.99$ . Thinning is driven by growth

$$\frac{1}{N} \frac{dN}{da} \propto -G \frac{1}{B} \frac{dB}{da}$$

(Clark 1992), where  $N$  and  $B$  are density and basal area summed over age. Seed dispersal is a negative binomial random variable (eqq. [9]–[10]). The mean number of seeds dispersed to a patch is the integral of the dispersal kernel,

$$F(x_1, x_2) = \int_{x_1}^{x_2} f(x) dx \quad (16)$$

$$= \frac{1}{2\Gamma(1/c)} [\gamma(1/c, X_2) - \gamma(1/c, X_1)],$$

where  $x_1$  and  $x_2$  are distances to the near and far sides of the 10-m wide patches, the limits  $X_1$  and  $X_2$  are  $X_i \equiv (X_i/\alpha)^c$ , and incomplete gamma functions are

$$\gamma(a, u) = \int_0^u z^{a-1} e^{-z} dz.$$

The mean  $F(x_1, x_2)$  and fitted clumping parameter are the basis for negative binomial seed arrivals. The expected seeds landing on a patch  $j$  is the summed contribution from all trees along the transect,

$$E[s_j] = g\beta \sum_{i=1}^{\# \text{ patches}} \sum_{k=1}^{\# \text{ trees}} b_{ik} F(x_{ij}, x_{ij}).$$

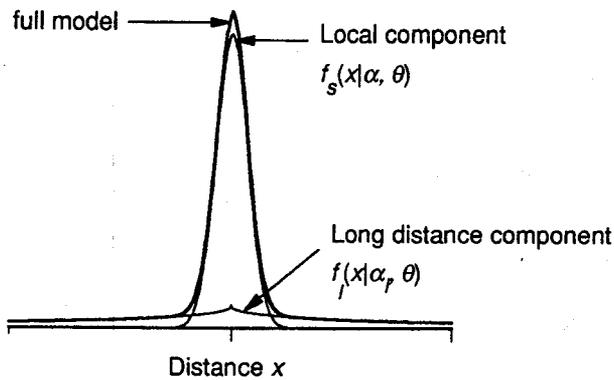


Figure 4: The mixed kernel used to test the long-distance dispersal hypothesis and to simulate a fraction of seed dispersed long distances (eq. [11]). The local component  $f_s(x)$  is Gaussian ( $c = 2$ ) and is fitted to seed rain data in modern stands. The long-distance component  $f_l(x)$  is fat-tailed ( $c = 1/2$ ) and prescribes the amount and length of the tail.

The dispersal kernel in equation (16) includes a fraction  $p$  parameterized with seed rain data (eq. [9]) and a fraction  $(1 - p)$  allocated to a "tail" represented by  $f_l(x)$  in equation (11) (fig. 4).

Several variables were calculated from simulations to permit comparison with analytical models. Intrinsic rate of increase, generation time, and net reproductive rate were calculated from simulated populations using equations (3), (4), and (5), respectively. Velocities were calculated both for the edge of the continuous population and for the most distant colonist and compared with rates predicted by discrete (eq. [6]) and continuous (eq. [7]) reproduction.

## Results

### *The Data Accommodate a Tail*

Seed rain data were compatible with dispersal kernels having a long, fat tail. AIC comparisons (fig. 5) showed that the long-distance model fit the data better than did a local model for low values of  $(1 - p)$  for most animal-dispersed types (*Carya*, *Cornus*, *Nyssa*, and *Quercus*) and one wind-dispersed type (*Tilia*). For all but three wind-dispersed types (*Acer*, *Liriodendron*, and *Betula*) the AIC differences  $D$  were small ( $\ll 1$  in fig. 5) compared with the AIC values themselves, which ranged from 24 (*Robinia*) to 257 (*Betula*). Only *Acer*, *Liriodendron*, and *Betula* reached  $D > 1$  for  $p = 0.1$ , and these were the types with the highest AICs (139 for *Acer*, 244 for *Liriodendron*, and 257 for *Betula*). Likelihood ratio tests supported interpretations of AIC comparisons. Seed shadows containing up to 10% in a tail  $(1 - p)$  and dispersal distances to  $\alpha_1 = 1,000$  m did not approach sufficiently low

probabilities for any taxa to conclude that the appended tail makes the fit substantially worse than the fitted local model. Up to 10% in long tails were all well above probabilities of 0.05 (figs. 5, 6). Long-distance dispersal is compatible with data almost independent of the dispersal parameter, with shapes fitting equally well for dispersal parameters of 200–1,000 m. Near vertical contours in figures 5 and 6 indicate minimal effect of tail length ( $\alpha_1$ ) when the fraction in the tail  $(1 - p)$  is less than 0.1.

Thus, the fraction in the tail has greater influence on the compatibility between model and data than does the shape or dispersal parameter. Surprisingly, the decline in probability with increasing tail fraction is greatest for wind-dispersed taxa. These taxa have the largest dispersal distances and, thus, were expected to be most compatible with long-distance dispersal. With 10% in the tail, probabilities for all animal-dispersed taxa remain above 0.70, whereas the longest wind-dispersed taxa *Acer*, *Liriodendron*, and *Betula* are all near 0.30 (figs. 5, 6).

### *Purely Local Dispersal Is Insufficient to Explain Seed Rain*

Does this compatibility with a long tail mean that the data might include seed arrivals from long distances or that the local model is simply a poor fit? Three factors other than presence of a "tail in the data" might cause us to accept the hypothesis of long-distance dispersal for the "wrong" reasons: large estimation error (a mediocre fit) is compatible with nearly everything, high clumping (broad spread about the mean) permits unrealistic latitude in model form and parameter estimates, or low power (e.g., inadequate sample size) does not allow us to reject the hypothesis that dispersal is actually only local. Error propagation, fitted negative binomial distributions, and power curves (bootstrapped probabilities of rejecting the null model when it is false) help identify how sources of variability affect likelihood ratio tests.

Narrow estimation errors on seed shadows help allay concerns that data might admit long tails simply because fits are poor (fig. 7). If data poorly fit the local dispersal model (eq. [9]), then estimation errors in figure 7 would be broad, and any composite kernel might fit as well as the local model; compatibility with a long tail (i.e., figs. 5, 6) might reflect, in part, low confidence in the model fit. With the exceptions of several taxa having low seed recovery (*Amelanchier*, *Cornus*, *Robinia*, and *Tilia*), 90 percentiles for estimation errors were narrow relative to the same percentiles for the negative binomial sampling density.

Negative binomial percentiles for animal-dispersed species are large relative to wind-dispersed taxa (fig. 7) due to a high degree of clumping (fig. 8). Wind-dispersed

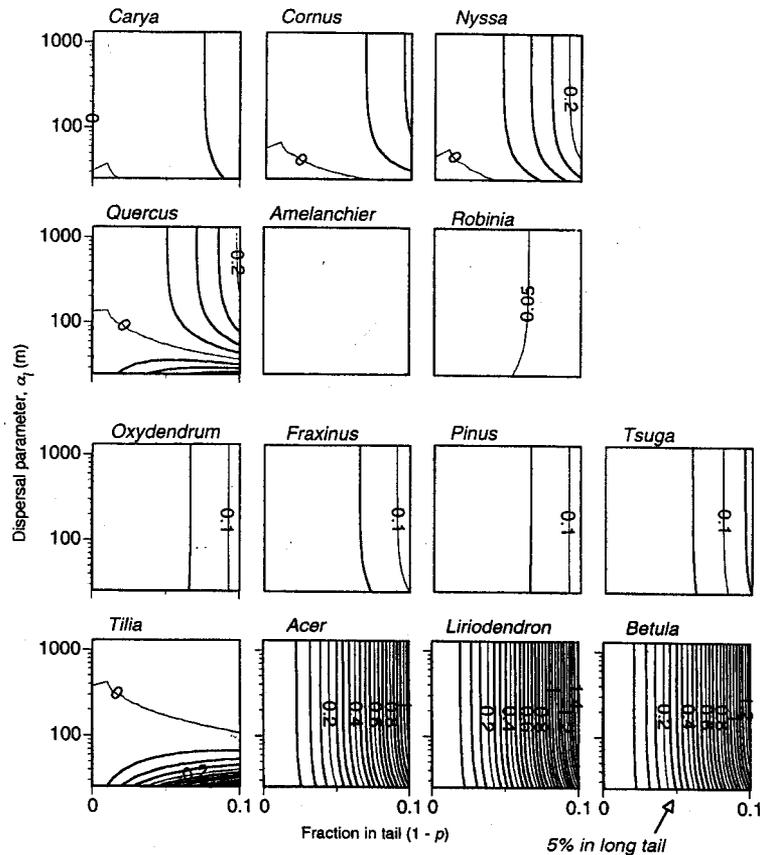


Figure 5: The difference between AIC values for the two-component kernel and the local model. The long-distance component  $f_i(x)$  has a fat ( $c = 1/2$ ) tail. Positive values indicate a better fit with the local model and vice versa. AIC differences near 0 indicate that data are compatible with the prescribed fraction ( $1 - p$ ) and length ( $\alpha_1$ ) of the long-distance component. Increasing AIC with the increasing fraction and length of the tail (upper right-hand corner of each square) describe the increasing incompatibility of the data with the model as the model assumes too much long-distance dispersal.

seed tends to a Poisson distribution. Seed shadows with high clumping yield 90 percentiles that extend to zero seeds even close to the seed source for all animal-dispersed species except *Nyssa* (fig. 7). Thus, despite relatively high average seed rain near parent trees, there are still many square meters of ground receiving little or no seed. High clumping could explain the compatibility with distant dispersal for some taxa.

Compatibility between data and long-distance tails does not result from low power for most species. If we failed to resolve dispersal parameters in the local model, then compatibility of data with a model including a long tail could occur due to insufficient sample size. In addition to low estimation error on seed shadows (fig. 7), power curves indicate our sample sizes are sufficiently large such that, if present, we could confidently fit large dispersal parameters for all taxa except *Fraxinus*, *Cornus*, *Carya*, and *Amelanchier* (fig. 9). Low power for *Amelanchier* and *Cornus* is consistent with wide estimation error in seed shadows (fig. 7). Tight confidence intervals on

*Fraxinus* and *Carya* (fig. 7) belie low power to reject the possibility that the dispersal parameter is large. Despite moderately wide estimation error on seed shadows for *Tilia* (fig. 7), there is high power to reject the possibility of a large dispersal parameter (fig. 9).

Taken together, the data are compatible with >10% of seed dispersed long distances. Compatibility between data and model depends on the tail fraction, but it is insensitive to the tail shape and distance parameter. In the case of some rare types, this compatibility may reflect poor fits (*Amelanchier*, *Cornus*, *Robinia*, and *Tilia*), low power to discriminate long dispersal (*Fraxinus*, *Cornus*, *Carya*, and *Amelanchier*), or both. But many of the best-dispersed taxa also have tight seed shadows, high power, and still admit a long-distance tail in fractions >10%.

#### Diffusion Rates from Tree Life Histories

Simulated populations parameterized with field data migrated at rates described by the diffusion model for dis-

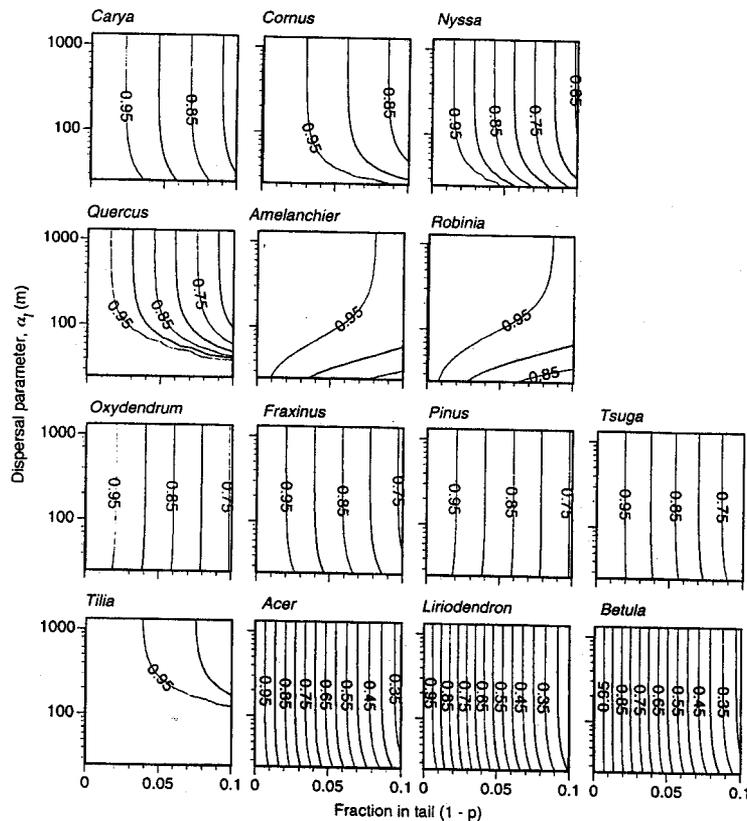


Figure 6: Probability surfaces for likelihood ratio tests between local and long-distance dispersal (eq. [12]). Axes as in figure 5.

crete reproduction (eq. [6]) and slower than predicted by the model for continuous reproduction (eq. [7]; fig. 10A). Although individuals reproduce continuously once established, delayed maturation makes population spread a stepwise process better described by a model that assumes discrete reproduction (eq. [1]).

Asymptotic rates of spread using just the fitted (local) seed shadow provide insights into migration potentials in the absence of long-distance dispersal. Overall, migration is slow. Among species, migration rates ranged over two orders of magnitude, but all were  $< 20 \text{ m yr}^{-1}$ . As predicted by diffusion (eq. [6]), simulated rates of spread increased with decreasing generation time (fig. 10B), increasing fecundity (fig. 10C), and increasing dispersal distance (fig. 10D). Correlations among species in their life-history parameters, however, made fecundity appear more important for migration than it actually is. For  $R_0$  values typical of these data,  $S_\beta$  is low across the several orders of magnitude of  $\beta$  values (fig. 10E). This effect,  $C \propto \beta^2$ , is plotted together with simulated rates in figure 10C to demonstrate that migration rates continue to rise with increasing fecundity beyond what is predicted by diffusion. This apparent sensitivity results from a correlation among species in fecundity and dispersal ( $r =$

0.376), which has a strong (i.e., strictly proportional) influence on velocity  $C$ . Sensitivity experiments with the simulation model confirmed this relatively small role for fecundity compared to dispersal and generation time. The relationships between generation time and  $C$  (fig. 10B) and dispersal and  $C$  (fig. 10D) are consistent with sensitivities predicted by diffusion.

Although spread was slower than observed in the paleorecord, the growth rates were faster than estimates from fossil pollen data (Clark 1993). Rates of pollen increase are mostly below  $0.02 \text{ yr}^{-1}$ , whereas simulated rates were  $0.03\text{--}0.5 \text{ yr}^{-1}$  (fig. 11). The pollen estimates predict rates of spread  $< 1 \text{ m yr}^{-1}$ ; these simulations suggest rates higher by an order of magnitude ( $1\text{--}10 \text{ m yr}^{-1}$ ), but still slower than past tree migrations ( $100\text{--}1,000 \text{ m yr}^{-1}$ ).

#### *The Tails Admitted by Real Data Allow for Rapid Spread*

A small fraction of seed allocated to a long-distance tail permit spread as rapid as observed in the paleorecord, provided the tail has the right shape. The example for *Acer rubrum* (fig. 12) is typical. The traveling wave ob-

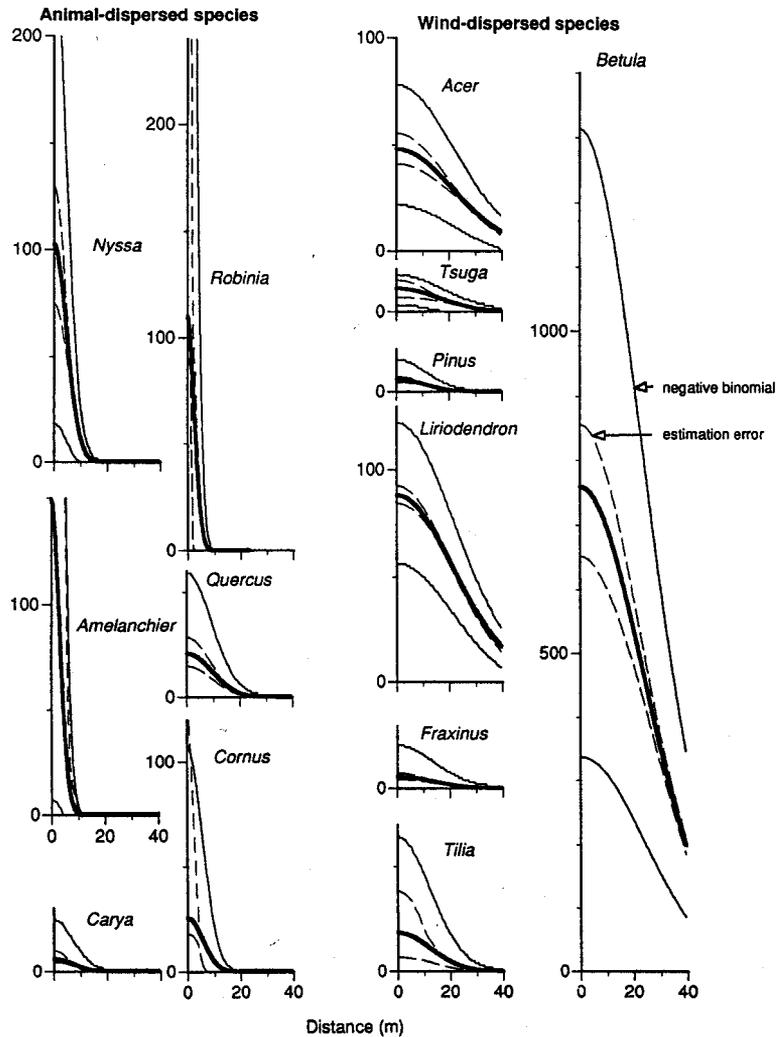


Figure 7: Comparison of maximum likelihood kernels (*thick solid lines*), 95% confidence intervals for estimation error (*dashed lines*), and the negative binomial distributions of seed arrivals (*thin solid lines*). In most cases, estimation errors are substantially tighter than negative binomial distributions. Estimation errors are bootstrapped from and negative binomials are fitted to equation (9).

tained in the absence of long-distance dispersal is a coherent front that spreads at constant rate (fig. 12A) as predicted by simple diffusion (eq. [6]). Incorporating a small fraction in a tail of  $c_1 = 1/2$  that is compatible with dispersal data (fig. 6) alters the pattern to one of incoherent spread at accelerating rate (fig. 12C). The shape of the front increasingly flattens as the outlying satellite colonies coalesce and increasingly pull away from the main front.

Although data accommodated long tails regardless of shape ( $c = 1/2$  or 2), only the fat-tailed distribution produced rates approaching those obtained in the early Holocene. The Gaussian long-distance tail (fig. 12B) produced spread more rapid than the local dispersal model, but there was still an asymptotic rate of spread as pre-

dicted by simple diffusion. A comparison of rates of spread over time (fig. 12D) illustrates how the long-distance exponentially bounded tail ( $c = 2$ ) differs from the local model only to the extent that it increases the velocity. The linear increase in velocity for the thin tail is predicted by equation (A5).

Because fecundity differences among taxa are large (table 1), and fat-tailed dispersal exaggerates fecundity effects on spread (eq. [B2b]), fat-tailed dispersal resulted in large differences among species. For example, *Betula*, which produces an order of magnitude more seed than any other species in this community, rapidly outpaced other life histories, almost immediately occupying the entire 10 km transect as soon as the initial trees achieved reproductive maturity (fig. 13A). *Acer rubrum* achieved

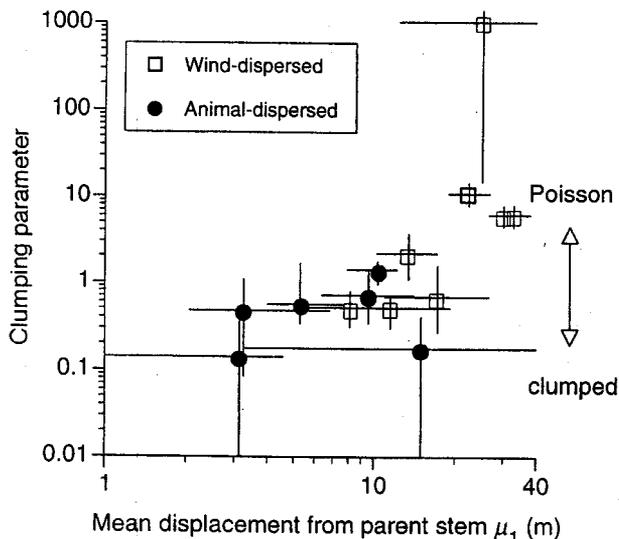


Figure 8: Comparison of 95% confidence intervals for mean dispersal distances and degree of clumping. Wind-dispersed species have longer dispersal distances and less clumped seed. Wind-dispersed species = empty square. Animal-dispersed species = solid circle.

rapid spread, in part due to short generation time (fig. 13B). In both of these cases, the long-distance tail tended to pull the continuous population front forward, accelerating the rate beyond that predicted by simple diffusion (fig. 13A, B). *Quercus* and *Nyssa* have lower fecundities, and the tail did not tend to pull the front forward (fig. 13C, D). Rather, the tail tended to be supported by seed dispersal from the population interior. These large contrasts in rate differ substantially with those obtained by the same taxa with simple diffusion (fig. 10).

The rates obtained in simulation tended to be slower than predicted by integrodifference equation (A5). The slowest moving taxa (e.g., *Nyssa* and *Quercus*) had simulated velocities closest to those predicted by the analytical model, whereas simulated rates for the faster moving *Betula* and *Quercus* were far slower than predicted by the analytical model (fig. 14).

### Discussion

Reid's paradox represents the failure of classical models of dispersal to describe the reality of rapid spread. Finding a model consistent both with modern dispersal and with past migrations resolves the paradox by providing at least one hypothesis for rapid spread. Lacking has been a model that incorporates these assumptions together with parameters estimated from data. Two challenges to re-

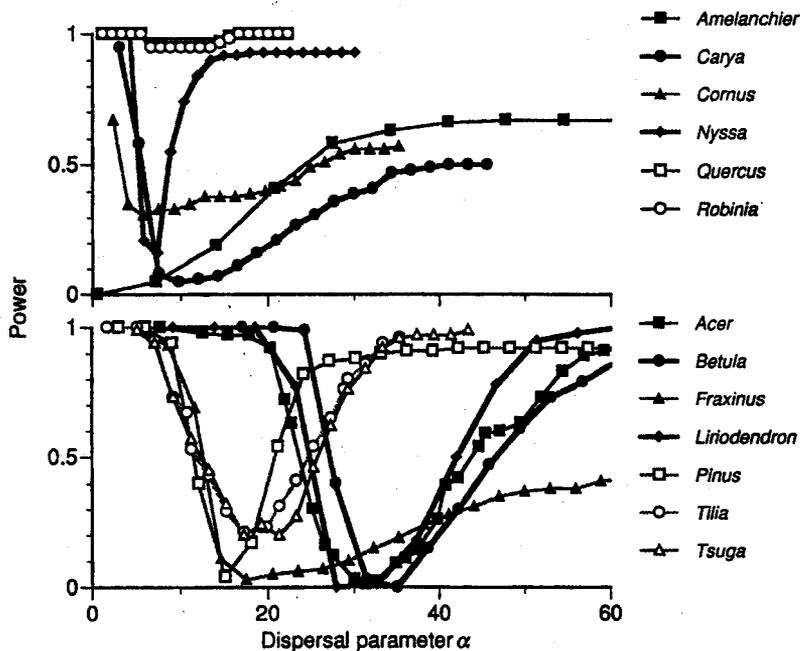


Figure 9: Bootstrapped power curves (eq. [13]) for dispersal parameter estimates. Power curves show the probability of mistakenly estimating local dispersal (low values of  $\alpha$ ) if dispersal was, in fact, long distance. The fact that power curves for *Amelanchier*, *Cornus*, *Carya*, and *Fraxinus* remain well below 1 out to distances of 60 m does indicate a low degree of confidence that dispersal is, in fact, local.

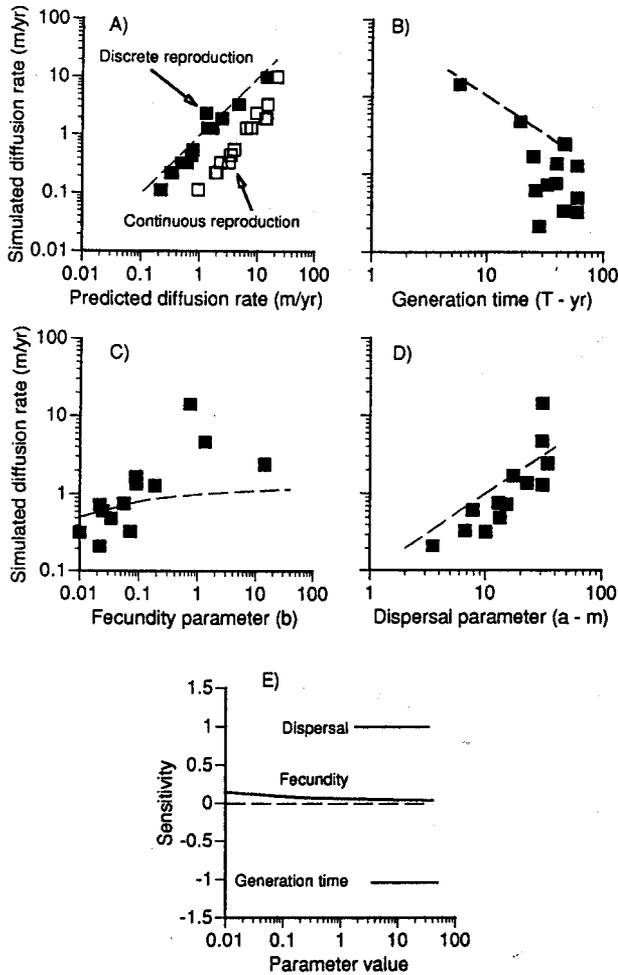


Figure 10: Comparison of simulated diffusion rates (A–D) against the predictions of equations (6) and (7) and sensitivities to life-history parameters (E). A, The predicted rates for discrete reproduction (eq. [6]) slightly exceed simulated rates, whereas predictions from continuous reproduction (7) far exceed simulated rates. Simulated rates are plotted against generation time (calculated from simulation) (B), fecundity (C), and dispersal (D) parameters (ML estimates from eq. [9] and used in simulation). E, Sensitivities of velocities (app. B) are small for fecundity relative to dispersal and generation time.

solving the conflict between past migration and modern dispersal included an inability to estimate dispersal kernels and implementation of a model that incorporates dispersal and life history that can be measured in forests.

This analysis is not motivated by observations of accelerating spread in paleorecords. The coarse resolution of paleorecords do not permit the spatial and temporal resolution needed to resolve such patterns. Rather, the rates averaged over centuries to millennia are sufficiently high to motivate search for mechanisms of rapid spread. Rather than claim to match observed patterns of accel-

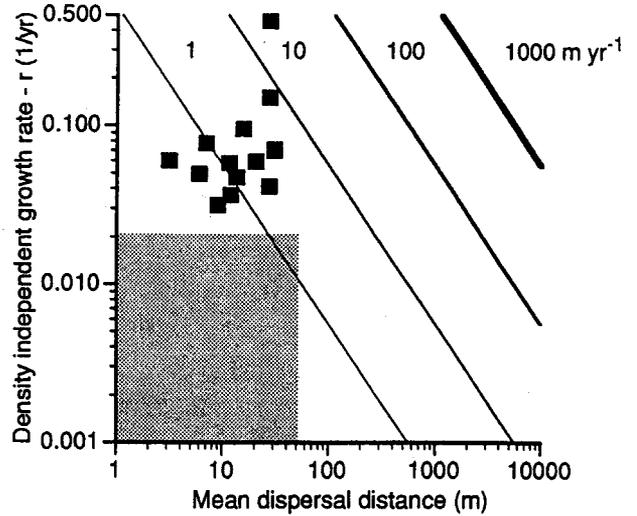


Figure 11: Simulations predict growth rates higher than estimates from fossil pollen data (shaded parameter space) but still far too low to match the high rates of spread in the paleorecord ( $10^2$ – $10^3$  m yr<sup>-1</sup>).

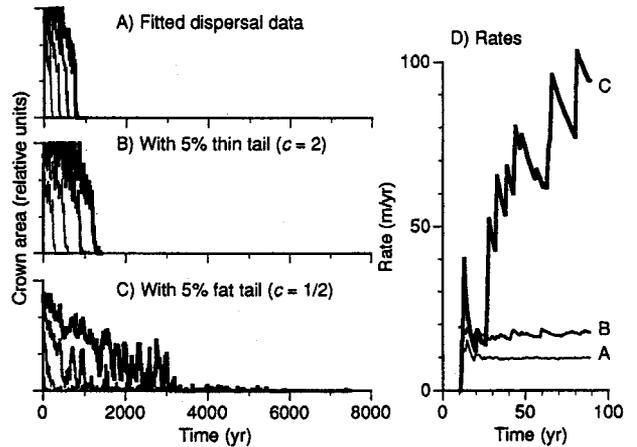


Figure 12: Example population fronts and velocities for *Acer rubrum* contrasting (A) purely local dispersal (eq. [5]) with parameters from table 1, (B) a mixed kernel (eq. [11]) with a long ( $\alpha_1 = 100$  m) but Gaussian ( $c_1 = 2$ ) tail, and (C) a mixed kernel with fat tail ( $c_1 = 1/2$ ) of the same length as in (B). Parameter values in B and C are compatible with seed rain data for *Acer rubrum* (P values of .65 for likelihood ratios in fig. 5). The exponentially bounded tail has a higher velocity than purely local dispersal, but it does not accelerate (D). “Crown area” is proportional to basal area given by G in equation (15).

ation in the paleorecord, I demonstrate here explanations for the high rates.

#### Rapid Spread Is Compatible with Dispersal Data

Lack of dispersal estimates has been the “immediate outstanding problem” (Mollison 1977, p. 311) for under-

## Population fronts at 80 yr

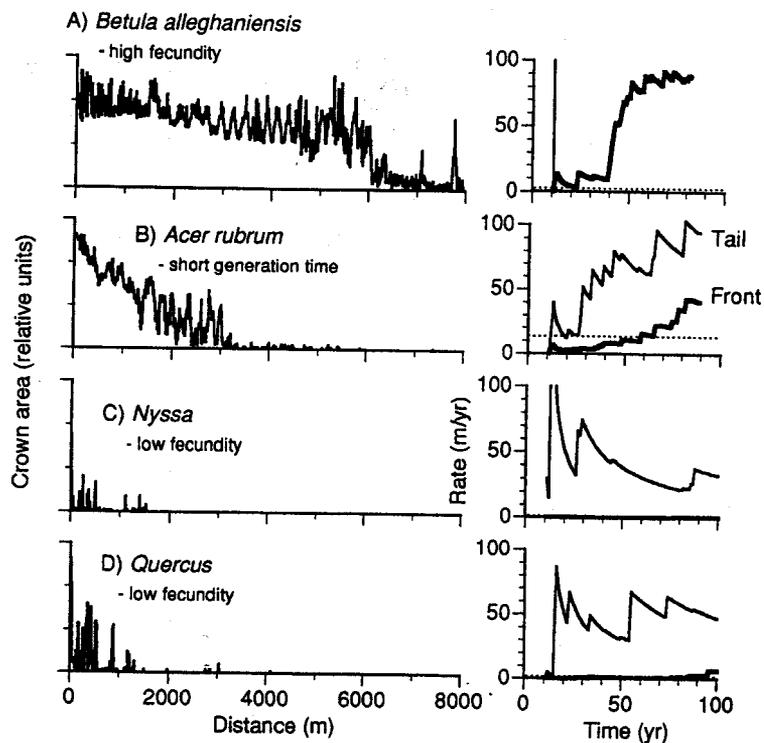


Figure 13: Effect of a 5% fat tail with dispersal parameter  $\alpha_1 = 100$  m on rates of spread of two wind-dispersed (A, B) and two animal-dispersed (C, D) species. At left are population fronts at 80 yr after the initiation of spread. At right are velocities for the spread of the farthest established individual (*thin lines*) and for the edge of the "continuous population," to the left of which all patches are occupied.

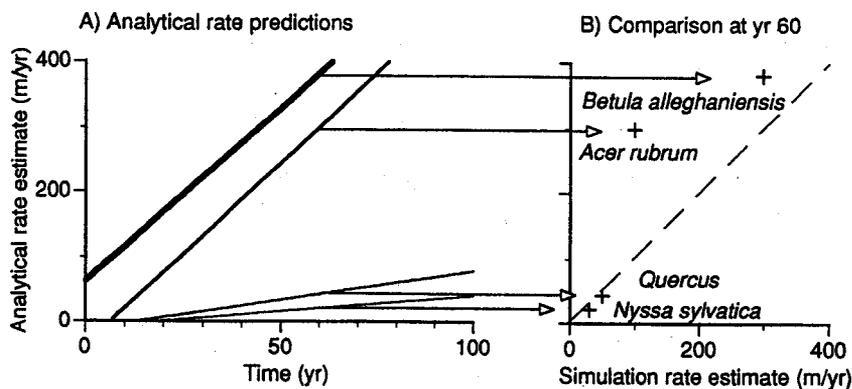


Figure 14: A, Analytical predictions (eq. [A5]) of spread for species in figure 13 having 5% of the kernel in a fat tail; B, comparison with simulated rates. *Nyssa sylvatica* = thinnest line. *Quercus* = thin line. *Acer rubrum* = thick line. *Betula alleghaniensis* = thickest line.

standing population spread. In the absence of data, it has been impossible to convincingly reject diffusion and to postulate the actual dispersal kernels for seed. Fitted local kernels (Clark et al. 1998b) demonstrate the failure of a diffusion model, and compatibility with rare, long-distance dispersal (figs. 5, 6) provides a hypothesis in its place.

Using parameter estimates from field data, I obtained migration rates surprisingly similar to (within the level of uncertainty of) those estimated by fossil pollen data (100–1,000 m yr<sup>-1</sup>; Davis 1976; Huntley and Birks 1983; Delcourt and Delcourt 1987; Birks 1989). Those rates depended on relaxing the assumption of local dispersal only to a degree that was consistent with data (figs. 13, 14). The dispersal estimates are sufficiently powerful to resolve local dispersal (fig. 9), yet they admit 2%–10% long-distance dispersal (figs. 5, 6). Although animal-dispersed taxa have shorter average dispersal distances, they were more compatible with the long-distance dispersal hypothesis than were wind-dispersed types. Although our seed traps do not quantify many types of secondary dispersal by wind and animals, they do include seed dispersed by frugivorous birds (Clark et al. 1998b). Higher clumping (fig. 8), larger estimation error (fig. 7), and lower power (fig. 9) for animal-dispersed taxa might contribute to the differences observed in tests of the long-distance dispersal hypothesis (figs. 5, 6).

Although fits to the data are insensitive to the tail shape, the capacity for rapid spread relies on it. I found little effect of tail shape on compatibility with data when the fraction in the tail ( $1 - p$ ) was  $< 0.10$ . So the data cannot resolve the shape of the tail, but they do clarify that long-distance dispersal could account for  $> 5\%$  of seed rain for all taxa studied here. It is interesting to note that we find from 1% to 5% seed for many taxa in stands where trees are absent, which is compatible with the results in figure 5. This small amount is enough to insure rapid spread, provided the tail is "fat" (figs. 12, 13). The Gaussian tail is effectively truncated in comparison to the fat-tailed kernel (fig. 2). The small difference in probabilities at large distances makes for extreme differences in velocity (fig. 12D).

#### Why Does Life History Not Matter?

The tree populations analyzed here vary over five orders of magnitude in fecundity ( $10^{-1}$ – $10^4$  seeds cm<sup>-2</sup> basal area yr<sup>-1</sup>) and over one order of magnitude in mean dispersal distance (2–40 m; Clark et al. 1998b) and generation time (5–50 yr; table 2). On the one hand, results presented here emphasize the importance of fat-tailed dispersal for the Holocene spread of trees. Analytical models and simulations predict slow ( $< 20$  m yr<sup>-1</sup>) diffusion in the absence of long-distance dispersal (figs. 11,

Table 2: Life-history parameters calculated from simulation

Taxon	Generation time T (yr)	Net reproductive rate $R'_0$
<i>Acer rubrum</i>	5.76	1,325
<i>Betula lenta</i>	45.7	50,800
<i>Carya glabra</i>	59.4	36.6
<i>Cornus florida</i>	26.0	64.0
<i>Fraxinus americana</i>	59.2	116
<i>Liriodendron</i>	19.1	4,624
<i>Nyssa sylvatica</i>	45.7	193
<i>Pinus rigida</i>	32.9	11
<i>Quercus rubra</i>	39.0	189
<i>Robinia</i>	28.1	19
<i>Tilia americana</i>	24.8	10
<i>Tsuga</i>	39.6	293

Note: The germination fraction  $g$  used in simulations is 0.001.

12). On the other hand, long-distance dispersal is expected to amplify the role of fecundity and generation time (fig. 3D), a prediction borne out in simulation (e.g., fig. 13). There are some differences in Holocene rates of spread that may relate to life history, but most taxa achieve high velocities. Among the taxa for which rates of spread exceed  $10^2$  m yr<sup>-1</sup> are taxa having the full range of life-history traits (reviews of Davis 1976; Huntley and Birks 1983; Ritchie and MacDonald 1986; Delcourt and Delcourt 1987; MacDonald 1993); high rates are not restricted to the most fecund, earliest-maturing, and best-dispersed taxa.

The contradiction represented by the need for fat-tailed dispersal to explain the high rates and the absence of differences that might reflect life history suggests Holocene spread may have failed to achieve potential velocities. Although not all paleoestimates are the same, estimates for many taxa exceed  $10^2$  m yr<sup>-1</sup> and few exceed  $10^3$  m yr<sup>-1</sup>. The accelerating spread predicted by fat-tailed dispersal cannot proceed indefinitely. The fact that the highest velocities estimated for many taxa come at the onset of spread (e.g., Delcourt and Delcourt 1987; Birks 1989) suggests that climate or geographic constraints set in rapidly and restricted velocities before life-history differences could become evident. The lack of life-history effects supports the notion that migrations were constrained by factors other than migration potential. Potential factors have been reviewed by Webb (1986), Bennett (1988), and Prentice et al. (1991).

#### Too Much Emphasis on Low Probabilities?

Unlike the fat-tailed kernels analyzed here, travel of seeds is not potentially infinite. The flat tail that makes approx-

imation (A2) possible means that probabilities approach 0 too slowly to realistically represent dispersal at impossibly large distances. Because dispersal is not infinite, it is important to ask if mechanisms exist that can transport seed long distances and whether results obtained here depend on the nature of the kernel beyond distances that realistically describe dispersal. Observation suggests that the part of the tail that matters extends in the range of 1–10 km. This range is consistent with estimates for frugivorous mammals and birds that cache fagaceous nuts (Storm and Montgomery 1975; Vanderwall and Balda 1977; Johnson and Adkisson 1985; Clark et al. 1998a). Storms likewise appear capable of moving some seed such distances (Snow et al. 1995). Reid (1899), Webb (1987), Birks (1989), Johnson and Webb (1989), and Woods and Davis (1989) make plausible arguments for dispersal over such distances. Greene and Johnson (1995) suggest causes for fat-tailed dispersal in wind-dispersed species.

Leaps beyond the range of several kilometers are not essential to achieve the high velocities simulated here. The rates reported here occur along a transect <10 km long (e.g., figs. 12, 13). To insure that rates predicted by simulation did not depend on kernel shape at impossibly long distances, I completed simulations with kernels truncated at 5,000 m. These simulations displayed the same accelerating rates observed with "infinite" tails, indicating that observable processes (long-distance dispersal by wind, birds, and mammals) could produce such results.

#### *Limitations of Simplified Life History*

Simulations and the integrodifference equation model disagree on velocities, slightly, in the case of diffusion (fig. 10A), and substantially, in the case of fat-tailed dispersal (fig. 14). Simplifying assumptions regarding life history and stochasticity represent two reasons for potential differences. In my application of an integrodifference equation model, I summarize life-history schedules by  $R'_0$  and  $T$ . Mollison (1991) points out that, if reproduction were strictly concentrated at age  $T$  and distance  $\mu'_1$ , net reproductive rate would have no effect on velocity whatsoever, provided only that it is large enough to permit spread (i.e.,  $R'_0 \geq 2$ , one new recruit to move, on average, one pace to the left and one to the right). In this case the population simply steps forward at rate  $\mu'_1/T$ . Net reproductive rate appears in the calculation with the higher moments of the reproduction and dispersal kernel, the modest contribution in approximation (6) representing Gaussian dispersal. Variance in the life-history schedule increases velocity, at least through its effect on early reproduction (van den Bosch et al. 1990). Higher

moments in the dispersal kernel also increase the importance of net reproductive rate. Positive correlation between dispersal and reproduction can reduce rate of spread below the prediction that treats the marginal distributions of dispersal and reproduction as independent. This effect of positive correlation arises as a negative term contributed by covariance (the product cumulant  $\kappa_{21}$ ) in van den Bosch et al.'s (1990) equation (6.4) and its effects are discussed by Mollison (1991). My simulations do not contain this correlation because fecundity, but not dispersal, change with tree size. Collapsing reproductive schedules down to two parameters,  $R'_0$  and  $T$ , could explain slight discrepancies for diffusion (fig. 10A) and larger ones for fat-tailed dispersal (fig. 10B).

Stochasticity may explain why simulated population spread tended to be lower than predicted by the integrodifference equation model. Mollison (1972) reports that certain types of nonlinear stochastic models in one dimension can have finite velocities for kernels more leptokurtic than exponential provided only that the variance is finite. While the deterministic model assumes growth of noninteger (infinitely small) founding populations, the stochastic model takes the predicted seed rain as the expectation of a clumped distribution with probabilistic establishment. The uncertainty of establishment in the stochastic model probably accounts for lower rates in simulation (fig. 14). If so, the approximation (A2) may increasingly overestimate velocities with increasing kernel kurtosis. Although integrodifference models provide valuable insights into qualitative behavior of migrating populations, it is possible that they might overestimate rates of spread for long-lived plants with delayed maturation.

#### *Implications for Potential Velocities*

The prospect that plants can migrate in response to climate change at rates well in excess of those predicted from average dispersal distances holds promise for populations in the face of future global change. Fossil pollen demonstrate a migration potential for trees still unrecognized in the Tropics, with its comparatively meager paleorecord. Relatively few rare events can be enough to support this spread. Contemporary invasions face a host of conditions unlike those prevailing in the late- and postglacial times, however, when temperate forest expansion was so dynamic (Clark et al. 1998a; Pitelka et al. 1997). Results of this analysis are important, not as a prediction of rates to be expected in the future, but rather to shift the focus to processes that govern these long-distance dispersal events as the ones that ultimately control the velocity of spread.

Acknowledgments

For helpful discussion and/or reviews of the manuscript, I thank I. Hanski, G. Hurtt, M. Lewis, J. Lynch, E. Macklin, J. McLachlan, S. Pacala, P. Wyckoff, and two anonymous reviewers. This research was supported by National Science Foundation grants DEB-9419677 and DEB-9632854.

APPENDIX A

Rate of Spread for Fat and Mixed Kernels

Estimating the rate of population expansion proceeds in one of two ways, depending on kernel shape. I outline these two approaches and then use a hybrid to contrast velocities for different shapes of the kernel (5), including the mixed kernel used in simulation. The traveling wave solution for an integrodifference equation (Kot et al. 1996) makes use of the moment-generating function and provides an asymptotic velocity that obtains with the passage of time. Moment integrals converge provided the kernel is exponentially bounded ( $c \geq 1$  in eq. [5]).

Kot et al. (1996) provide a new approximation that extends (and is limited) to the case of fat-tailed kernels. For purposes of estimating spread of tree populations, I first scale time in generations, such that generation  $k = t/T$ . Taking Fourier transforms in equation (1) gives

$$\hat{N}_{k+1}(\omega) = \hat{N}_k(\omega) R_0^k \hat{f}(\omega).$$

For initial density  $N_0$  concentrated at  $x = 0$ , this linear model becomes

$$\hat{N}_k(\omega) = N_0 R_0^k (\hat{f}(\omega))^k. \tag{A1}$$

Kot et al. (1996) sidestepped the problem of inverting  $(\hat{f}(\omega))^k$  by showing that, as distance  $x$  becomes large (or  $\omega$  tends to 0) for fat-tailed kernels that possess moments, this inverse tends to  $f(x)$ , regardless of  $k$ . In the limit as density approaches some low level of detection, say one individual, that density at large  $x$  is approximately

$$N_k(x) \approx N_0 R_0^k f(x). \tag{A2}$$

Some specific results applicable to tree migration and a mixed kernel can be derived by extension. The kernel analyzed here is a two-part mixture that includes a large fraction  $p$  of seed dispersed nearby  $f_1(x)$  and a small fraction  $(1 - p)$  that is dispersed a long distance, described by  $f_2(x)$  (eq. [8]). I first consider the migration rate that results from purely local dispersal and then provide an approximation for their combined effect. The "local" component fitted to seed rain data is exponentially bounded ( $c = 2$  in eq. [5]). Because  $f(x)$  is a density function,  $(\hat{f}(\omega))^k$  is the characteristic function of a den-

sity after  $k$  iterations of the process (i.e., the characteristic function of the sum is the product of their individual characteristic functions). For the case of seed shadows fitted to field data ( $c = 2$ ), that transform is

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} f(x) dx = \exp\left(-\left(\frac{\alpha\omega}{2}\right)^2\right),$$

and the transform needed for (A1) is

$$(\hat{f}(\omega))^k = \exp\left(-k\left(\frac{\alpha\omega}{2}\right)^2\right).$$

Upon inversion we obtain

$$\begin{aligned} f_k(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega x} (\hat{f}(\omega))^k d\omega \\ &= \frac{1}{\alpha \sqrt{\pi k}} \exp\left(-\frac{x^2}{\alpha^2 k}\right), \end{aligned} \tag{A3}$$

which satisfies Einstein's (1905) diffusion equation.

Now inverting in the manner of Kot et al. (1996), we have after  $k$  generations the linear distance occupied by densities  $N \geq 1$ ,

$$x_k = \alpha \left[ k^2 \ln R_0 + k \ln\left(\frac{N_0}{\alpha \sqrt{\pi k}}\right) \right]^{1/2}.$$

Calculating  $x_{k+1}$  in the same manner, taking their difference, and dividing by elapsed time (a generation) gives the rate

$$\begin{aligned} C(t) &\equiv \frac{x_{k+1} - x_k}{T} \\ &= \frac{\alpha}{T} \left\{ \left[ (k+1)^2 \ln R_0 + (k+1) \ln\left(\frac{N_0}{\alpha \sqrt{\pi(k+1)}}\right) \right]^{1/2} \right. \\ &\quad \left. - \left[ k^2 \ln R_0 + k \ln\left(\frac{N_0}{\alpha \sqrt{\pi k}}\right) \right]^{1/2} \right\}, \end{aligned} \tag{A4}$$

which is diffusion, with asymptotic velocity

$$\lim_{t \rightarrow \infty} C(t) \rightarrow \frac{\alpha}{T} \sqrt{\ln R_0}.$$

My analysis of tree migration assumes a long-distance component of the mixed kernel having a fat tail ( $c = 1/2$ ). If migration is dominated by the tail of the kernel  $f_2(x)$ , densities at large  $x$  tend to

$$N_k(x) \approx N_0(0) R_0^k (1 - p) f_2(x).$$

Although I do not fully justify exclusive focus on the tail, the asymptotic rate of expansion implied by (A5) below demonstrates that the tail of the kernel rapidly outpaces

the diffusive spread that can be supported by (A4). We can further appeal to Kot et al.'s (1995) demonstration that the shape of the density profile,  $N_k(x)$  at large  $x$  is proportional to the kernel itself (their app. A for kernels lacking moment-generating functions). Solving for  $x$  and setting the critical density  $N_k(x) = 1$  yields the length along this transect that is occupied by the population at densities exceeding one individual in generation  $k$ :

$$x_k = \alpha \left[ k \ln R'_0 + \ln \left( \frac{(1-p)N_0}{4\alpha} \right) \right]^2.$$

Repeating the process for generation  $k + 1$  permits calculation of the rate of spread

$$C_t(k) \equiv \frac{x_{k+1} - x_k}{T} \approx \frac{\alpha_t \ln R'_0}{T} \times \left[ (2k+1) \ln R'_0 + 2 \ln \left( \frac{(1-p)N_0}{4\alpha_t} \right) \right]. \quad (\text{A5})$$

This rate increases linearly with number of generations  $k$  (i.e., time =  $kT$ ) since the population was initiated at  $x = 0$  and density  $N_0$ . I have not pursued here the possibility of thresholds that might depend on  $R'_0$ ,  $T$ , and/or  $p$  needed for spread to occur at all.

## APPENDIX B

### The Importance of Life History

This appendix uses a sensitivity analysis to demonstrate how fat-tailed dispersal amplifies life-history differences among species. For a diffusing population, the rate of spread (eq. [6]) is weakly affected by fecundity. Let  $S_\beta$  be the proportionate effect on the asymptotic rate of spread  $C$  for a fractional change in fecundity  $\beta$ . This sensitivity tends to

$$S_\beta \equiv \frac{\partial C / \partial \beta}{C / \beta} = \frac{1}{2 \ln R'_0}. \quad (\text{B1})$$

Fat-tailed kernels amplify the effects of fecundity and generation time. For the above case of  $c = 1/2$  sensitivity coefficients are

$$S_\alpha \equiv \frac{\partial C / \partial \alpha}{C / \alpha} = \frac{(k + 1/2) \ln R'_0 + 2Z_0 - 1}{(k + 1/2) \ln R'_0 + Z_0}; \quad (\text{B2a})$$

$$S_\beta \equiv \frac{\partial C / \partial \beta}{C / \beta} = \frac{2}{\ln R'_0} \left[ \frac{(2k+1) \ln R'_0 + Z_0}{(2k+1) \ln R'_0 + 2Z_0} \right]; \quad (\text{B2b})$$

and

$$S_T \equiv \frac{\partial C / \partial T}{C / T} = -\frac{(2k + 1/2) \ln R'_0 + Z_0}{(k + 1/2) \ln R'_0 + Z_0}, \quad (\text{B2c})$$

where constant  $Z_0 \equiv \ln((1-p)N_0/4\alpha_1)$  depends on initial conditions, the fraction in the tail, and the length of the tail. It can be readily demonstrated that the fatter the tail the greater the effect of net reproductive rate and generation time.

### Literature Cited

- Akaike, H. 1992. Information theory and an extension of the maximum likelihood principle. Pages 610–624 in S. Kotz and N. Johnson, eds. *Breakthroughs in statistics*. Springer, New York.
- Bennett, K. D. 1988. Post-glacial vegetation history: ecological considerations. Pages 699–725 in B. Huntley and T. Webb III, eds. *Vegetation history*. Kluwer, Dordrecht.
- Birks, H. J. B. 1989. Holocene isochrone maps and patterns of tree-spreading in the British Isles. *Journal of Biogeography* 16:503–540.
- Burns, R. M., and B. H. Honkala. 1990. *Silvics of North America*. Agricultural Handbook 654. U.S. Department of Agriculture, Washington, D.C.
- Bush, M. B. 1994. Amazon speciation: a necessarily complex model. *Journal of Biogeography* 21:5–17.
- Clark, J. S. 1992. Density-independent mortality, density compensation, gap formation, and self-thinning in plant populations. *Theoretical Population Biology* 42:172–198.
- . 1993. Paleocological perspectives on modeling broad-scale responses to global change. Pages 315–332 in P. Kareiva, J. Kingsolver, and R. Huey, eds. *Biotic interactions and global change*. Sinauer, Sunderland, Mass.
- Clark, J. S., and Y. Ji. 1995. Fecundity and dispersal in plant populations: implications for structure and diversity. *American Naturalist* 146:72–111.
- Clark, J. S., C. Fastie, G. Hurtt, S. T. Jackson, C. Johnson, G. King, M. Lewis, et al. 1998a. Reid's paradox of rapid plant migration. *BioScience* 48:13–24.
- Clark, J. S., E. Macklin, and L. Wood. 1998b. Stages and spatial scales of recruitment limitation in southern Appalachian forests. *Ecological Monographs* 68:213–235.
- Davis, M. B. 1976. Pleistocene biogeography of temperate deciduous forests. *Geoscience and Man* 13:13–26.
- . 1987. Invasion of forest communities during the Holocene: beech and hemlock in the Great Lakes region. Pages 373–393 in A. J. Gray, M. J. Crawley, and P. J. Edwards, eds. *Colonization, succession, and stability*. Blackwell Scientific, Oxford.
- Delcourt, P. A., and H. R. Delcourt. 1987. Long-term forest dynamics of the temperate zone. Springer, New York.

- Dexter, F., H. T. Banks, and T. Webb III. 1987. Modeling Holocene changes in the location and abundance of beech populations in eastern North America. *Review of Palaeobotany and Palynology* 50:273–292.
- Diekmann, O. 1978. Thresholds and traveling waves for the geographical spread of infection. *Journal of Mathematical Biology* 6:109–130.
- . 1979. Run for your life: a note on the asymptotic speed of propagation of an epidemic. *Journal of Differential Equations* 33:58–73.
- Efron, B., and R. J. Tibshirani. 1993. *An introduction to the bootstrap*. Chapman & Hall, New York.
- Einstein, A. 1905. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik* 17:549–560.
- Fastie, C. L. 1995. Causes and ecosystem consequences of multiple pathways of primary succession at Glacier Bay, Alaska. *Ecology* 76:1899–1916.
- Goldwasser, L., J. Cook, and E. D. Silverman. 1994. The effects of variability on metapopulation dynamics and rates of invasion. *Ecology* 75:40–47.
- Greene, D. F., and E. A. Johnson. 1995. Long-distance dispersal of tree seeds. *Canadian Journal of Botany* 73:1036–1045.
- Higgins, S. I., D. M. Richardson, and R. M. Cowling. 1996. Modeling invasive plant spread: the role of plant-environment interactions and model structure. *Ecology* 77:2043–2054.
- Hubbell, S. P. 1980. Seed predation and the coexistence of tree species in tropical forests. *Oikos* 35:214–229.
- Huntley, B., and H. J. B. Birks. 1983. *An atlas of past and present pollen maps for Europe 0–13,000 years ago*. Cambridge University Press, Cambridge.
- Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. *American Naturalist* 104:501–528.
- Johnson, W. C., and C. S. Adkisson. 1985. Dispersal of beech nuts by blue jays in fragmented landscapes. *American Midland Naturalist* 113:319–324.
- Johnson, W. C., and T. Webb III. 1989. The role of blue jays in the postglacial dispersal of fagaceous trees in eastern North America. *Journal of Biogeography* 16:561–571.
- Kolmogorov, A., I. Petrovsky, and N. Piscounoff. 1937. Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique. *Bulletin de l'Université d'Etat à Moscou Série Internationale Section A* 1:1–25.
- Kot, M., M. A. Lewis, and P. van den Driessche. 1996. Dispersal data and the spread of invading organisms. *Ecology* 77:2027–2042.
- Leslie, P. H. 1966. The intrinsic rate of increase and the overlap of successive generations in a population guillemots (*Uria aalge* Pont.). *Journal of Animal Ecology* 35:291–301.
- Lewis, M. A. 1997. Variability, patchiness and jump dispersal in the spread of an invading population. Pages 46–49 in D. Tilman and P. Kareiva, eds. *Spatial ecology*. Princeton University Press, Princeton, N.J.
- MacDonald, G. M. 1993. Fossil pollen analysis and the reconstruction of plant invasions. *Advances in Ecological Research* 24:67–110.
- Melillo, J., I. C. Prentice, E.-D. Schulze, G. Farquhar, and O. Sala. 1996. Terrestrial biotic responses to environmental change and feedbacks to climate. Pages 445–482 in J. T. Houghton, L. G. Meira Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. Maskell, eds. *Climate change 1995: the science of climate change*. Cambridge University Press for the Intergovernmental Panel on Climate Change, Cambridge.
- Metz, H., and F. van den Bosch. 1995. Velocities of epidemic spread. Pages 187–210 in D. Mollison, ed. *Epidemic models: their structure and relation to data*. Cambridge University Press, Cambridge.
- Mollison, D. 1972. The rate of spatial propagation of simple epidemics. *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability* 3:579–614.
- . 1977. Spatial contact models for ecological and epidemic spread. *Journal of the Royal Statistical Society Series B* 39:283–326.
- . 1991. Dependence of epidemic and population velocities on basic parameters. *Mathematical Biosciences* 107:255–287.
- Pacala, S. W., and G. C. Hurtt. 1993. Terrestrial vegetation and climate change: integrating models and experiments. Pages 57–74 in P. Kareiva, J. G. Kingsolver, and R. B. Huey, eds. *Biotic interactions and global change*. Sinauer, Sunderland, Mass.
- Pitelka, L. F., J. Ash, S. Berry, R. H. W. Bradshaw, L. Brubaker, J. S. Clark, M. B. Davis, et al. 1997. Plant migration and climate change. *American Scientist* 85:464–473.
- Prentice, I. C. 1992. Climate change and long-term vegetation dynamics. Pages 293–339 in D. C. Glenn-Lewin, R. A. Peet, and T. Veblen, eds. *Plant succession: theory and prediction*. Chapman & Hall, London.
- Prentice, I. C., P. J. Bartlein, and T. Webb III. 1991. Vegetation and climate change in eastern North America since the last glacial maximum. *Ecology* 72:2038–2056.
- Reid, C. 1899. *The origin of the British flora*. Dulau, London.
- Ribbens, E., J. A. Silander, and S. W. Pacala. 1994. Seedling recruitment in forests: calibrating models to predict patterns of tree seedling dispersion. *Ecology* 75:1794–1806.

- Ritchie, J. C., and G. M. MacDonald. 1986. The patterns of post-glacial spread of white spruce. *Journal of Biogeography* 13:527-540.
- Seber, G. A. F., and C. J. Wild. 1989. *Nonlinear regression*. Wiley, New York.
- Shigesada, N., K. Kawasaki, and Y. Takeda. 1995. Modeling stratified diffusion in biological invasions. *American Naturalist* 146:229-251.
- Skellam, J. G. 1951. Random dispersal in theoretical populations. *Biometrika* 38:196-218.
- Snow, J. T., A. L. Wyatt, A. K. McCarthy, and E. K. Bishop. 1995. Fallout of debris from tornadic thunderstorms: a historical perspective and two examples from VORTEX. *Bulletin of the American Meteorological Society* 76:1777-1790.
- Storm, G. L., and G. G. Montgomery. 1975. Dispersal and social contact among red foxes: results from telemetry and computer simulation. Pages 237-246 in M. W. Fox, ed. *The wild Canids: their systematics, behavioral ecology, and evolution*. Van Nostrand Reinhold, New York.
- Thieme, H. R. 1979. Density-dependent regulation of spatially distributed populations and their asymptotic speed of spread. *Journal of Mathematical Biology* 8:173-187.
- van den Bosch, F., J. A. J. Metz, and O. Diekmann. 1990. The velocity of population expansion. *Journal of Mathematical Biology* 28:529-565.
- Vanderwall, S. B., and R. B. Balda. 1977. Coadaptations of the Clark's nutcracker and the pinon pine for efficient seed harvest and dispersal. *Ecological Monographs* 47:89-111.
- Webb, S. L. 1987. Beech range extension and vegetation history: pollen stratigraphy of two Wisconsin lakes. *Ecology* 68:1993-2005.
- Webb, T., III. 1986. Is vegetation in equilibrium with climate? how to interpret late-Quaternary pollen data. *Vegetatio* 67:75-91.
- Willson, M. F. 1993. Dispersal mode, seed shadows, and colonization patterns. *Vegetatio* 107/108:261-280.
- Woods, K. D., and M. B. Davis. 1989. Paleoecology of range limits: beech in the upper peninsula of Michigan. *Ecology* 70:681-696.

Associate Editor: Ilkka Hanski