



Promoting the Science of Ecology

---

Multiscale Analysis of Active Seed Dispersal Contributes to Resolving Reid's Paradox

Author(s): James A. Powell and Niklaus E. Zimmermann

Source: *Ecology*, Vol. 85, No. 2 (Feb., 2004), pp. 490-506

Published by: Ecological Society of America

Stable URL: <http://www.jstor.org/stable/3450213>

Accessed: 17/04/2009 10:48

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=esa>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



Ecological Society of America is collaborating with JSTOR to digitize, preserve and extend access to *Ecology*.

<http://www.jstor.org>

## MULTISCALE ANALYSIS OF ACTIVE SEED DISPERSAL CONTRIBUTES TO RESOLVING REID'S PARADOX

JAMES A. POWELL<sup>1,3</sup> AND NIKLAUS E. ZIMMERMANN<sup>2</sup>

<sup>1</sup>Department of Mathematics and Statistics, Utah State University, Logan, Utah 84322-3900 USA  
<sup>2</sup>Department of Landscape Dynamics, WSL—Swiss Federal Research Institute, Zuercherstrasse 111, CH-8903 Birmensdorf (ZH), Switzerland

**Abstract.** “Reid’s paradox” is the mismatch between theoretical estimates of invasion rates for plants and “observed” rates of plant migration, particularly in the Holocene postglacial migration northwards. While Reid couched his paradox in terms of the migration of oaks in Great Britain, observers have documented the same problem in a wide variety of species. In almost all cases, these authors suggest that occasional, long-distance events, probably mitigated by active dispersal factors (ants, birds, rodents) are responsible. Clark and co-workers have shown that order statistics can bridge the gap between theory and predictions, essentially using “fat-tailed” dispersal kernels raised to high powers corresponding to rare, distant dispersal events. However, the spatial caching structure induced by most active dispersers has not been specifically examined.

In this paper we develop a complementary approach to order statistics, based on the theory of homogenization, to describe the long-distance dispersal probabilities of seeds. The homogenization approach includes the spatial structure of active dispersal explicitly, and generates a corrective factor to current estimates of migration rates. This corrective factor is the “caching scale ratio,” the ratio of mean scatter-hoard size to mean separation between cache sites, and is, in principle, directly observable. The methodology is not limited to animal dispersers and can be used to address any set of dispersal agencies operating at differing scales, in one or more spatial dimensions. This approach is tested for three dispersal mutualisms where Reid’s paradox is known to operate: harvester ant/wild ginger, Blue Jay/oaks, and nutcracker/stone pine. Migration rates predicted using the homogenized approach compare favorably with estimates for tree species based on the paleo-ecological record. However, these rates do not seem to explain the Holocene migration of wild ginger.

**Key words:** Asymptotic rates of spread; dispersal kernel; fat-tailed distributions; integrodifference equations; homogenization; invasion; large-scale seed dispersal; small-scale seed dispersal; rapid plant migration; spatial convolution; spatial heterogeneity.

### INTRODUCTION

Dispersal of plant populations is an issue of ongoing prominence in ecology (Portnoy and Willson 1993, Cain et al. 1998, 2000, Clark et al. 1999, Nathan et al. 2000, 2001). How was the current distribution of plants achieved? How do populations respond after disturbances like wind, fire, glaciation, and volcanism? How are gaps filled and successional processes organized in space? What will become of populations if the various predictions of global climate change are realized? How do exotic species invade a landscape? At issue in all these questions are the spatio-temporal mechanisms of and expectations for dispersal of seeds. Much of this debate is related to understanding migration rates and invasion speed, especially with respect to expected global change (Kot et al. 1996, Melillo et al. 1996, Solomon and Kirilenko 1997, Clark 1998, Clark et al.

1998, 2001, Higgins and Richardson 1999, Eriksson 2000, Lewis 2000).

This justifies the quantitative attention that has been paid to dispersal over the last century. Mathematical reasoning going back to Fisher (1937) and Kolmogorov et al. (1937) has predicted asymptotic rates of invasion for reaction–diffusion models of population spread. For one-dimensional spread (rate of progress perpendicular to a front or “wave of invasion”) these models have the form

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + rf(u) \quad (1)$$

with  $f(u)$  being a normalized ( $f(0) = 1$ ), density-dependent fecundity,  $D$  the diffusion constant (with units length squared per time), and  $r$  the intrinsic rate of growth (Malthusian constant). Fisher and Kolmogorov demonstrated that the expected speed of the wave of invasion is  $c = 2\sqrt{Dr}$ . This is also the asymptotic rate of spread in two-dimensional invasions, and figures prominently in the analyses of Skellam (1951), Hengeveld (1989), Andow et al. (1990), Holmes et al. (1994), and Shigesada and Kawasaki (1997).

Manuscript received 29 August 2002; revised 10 March 2003; accepted 24 March 2003; final version received 14 June 2003.  
Corresponding Editor: M. L. Cain.

<sup>3</sup> E-mail: powell@math.usu.edu

The first author to use these results to analyze data on large-scale dispersal observations was Skellam (1951). For dispersal of muskrats (*Ondatra zibethica*) in central Europe the diffusive theory was quite descriptive. However, when applied to oaks (*Quercus* spp.), which had reinvaded England during the Holocene and spread north over several thousand years, the results were quite different. To travel the 960 km required (as estimated by Reid 1899) to reach its current extent, the root mean square distance at which daughter oaks arise from their parent must have been  $>0.93$  km. This, as noted by Reid and later Skellam, is completely beyond the realm of possibility for passive dispersal of acorns, or even vertebrate dispersal by squirrels and rodents, which result in dispersal of some tens of meters. This discrepancy is referred to as "Reid's Paradox" (Clark et al. 1998). Both Reid and Skellam believed that the discrepancy could only be resolved by including the effects of avian dispersers.

As pollen records have amassed in recent decades, more and more species are enrolled in Reid's Paradox. The discrepancy between theory and prediction extends to deciduous and coniferous trees in general (see the recent review by Clark et al. 1998), wind and animal (passive and active) dispersers, as well as herbaceous and understory species (Matlack 1994, Cain et al. 1998). The discrepancy is generally smaller for wind-dispersed species because of larger measured dispersal distances for seeds. Vertebrate-dispersed species have a much larger discrepancy, partially because of their relatively narrow seed patterns. According to the paleoecological record, some of the most rapid actual dispersers have large, nonwinged seeds (e.g., oaks, *Quercus* spp.).

In parallel, but seldom integrated, with the compilation of dispersal discrepancies is the study of animal/plant dispersal mutualisms. Particularly well documented is the mutualism between whitebark pine (*Pinus albicaulis*) and Clark's Nutcracker (*Nucifraga columbiana*), which provided the original inspiration for this manuscript and is co-indicated with establishment of Engelmann spruce (*Picea engelmannii*) in the Rocky Mountain West (Vander Wall 1990, Lanner 1996). Nutcrackers specifically harvest and scatter-hoard the large, wingless seeds of whitebark pine, preferentially dispersing seeds great distances away from the parent tree (Hutchins and Lanner 1982). Seeds are cached in the ground, often in open areas where they can avoid rodent predation, generally in common caching areas (Vander Wall 1990, Lanner 1996). The birds often choose the gradient area on the lee of ridges, where snow will sometimes be swept clear (and thus seeds will be accessible in midwinter), and also where snow will drift. This latter adaptation protects the seeds from hoard-raiding rodents until late spring, when the nutcrackers have young and will need a source of ready food. Not coincidentally, in the arid western United States this is a preferential area for the germination of

nonrecovered pine seeds, which grow and eventually colonize these alpine ridgelines (Perkins 2000). Since nutcrackers habitually cache 1.5–3 times as much food as they need (Vander Wall 1990), there is ample remainder, planted in an ideal environment for germination and establishment. The young whitebark pine stand then becomes a nursery for future species, particularly Engelmann spruce, which are also cached by the nutcracker but require more protection to establish. The nutcracker is thus adapted to offer the pines and spruce dispersal possibilities far superior to random chance and wind.

For its part, *P. albicaulis* has coevolved to favor the nutcracker, developing a cone structure well designed for removal of seeds by these birds (Lanner 1996). The physical structure of the pine, with open branches and cones strongly attached on separate small branches, diverges from the more general pine model of vertical trunk, horizontal branches and easily detached, random cone placement. The pine offers large, rich nuts with nutritional value far in excess of other conifers (Hutchins and Lanner 1982). This allows nutcrackers to maintain year-round residence in the whitebark community. These two species are keystone mutualists, truly (in the words of Lanner 1996) "made for each other."

This story is canonically clear in the case of Clark's Nutcracker/whitebark pine, but is a common theme in many other, less specific, mutualisms. Corvids, intelligent and strong-beaked birds of the family Corvidae (including nutcrackers, jays, magpies, crows, and ravens) often play a prominent role. Eurasian Nutcrackers (*Nucifraga caryocatactes*) have formed mutualisms with a variety of stone pines (genus *Pinus*, subsection *cebrae*) in Europe, Siberia, Korea, and Japan (Stimm and Boeswald 1994, Lanner 1996). Blue Jays (*Cyanocitta cristata*) and European Jays (*Garrulus glandarius*) seem to perform the same function for fagaceous trees (Darley-Hill and Johnson 1981, Johnson and Adkisson 1985, Johnson and Webb 1989, Vander Wall 1990). Chipmunks (*Tamias amoenus* and *T. speciosus*) are much more efficient dispersers of Jeffrey pines (*Pinus jeffreyi*) than wind, preferentially dispersing pine seeds to environments where they germinate and establish more effectively (Vander Wall 1993, 1994). Wild ginger (*Asarum canadense*) and *Sanguinaria canadense*, perennial woodland herbs, are examples of a dispersal mutualism involving harvester ants (Heithaus 1986, Cain et al. 1998). Harvester ant workers carry the seeds to nests, where they are stored temporarily as eliasomes are removed, then dispersed on the gravel disk surrounding the nest (Vander Wall 1990; E. R. Heithaus and M. L. Cain, *personal communications*), 90% of *Sanguinaria* seeds were expelled after four days in one observation site in West Virginia. In all of these cases there are vertebrates also harvesting seeds (mostly rodents), but almost all are pure seed predators, which eat the seeds immediately, rendering them nonviable in the process. In other cases vertebrate hoarders

build one large hoard from which few seedlings emerge. For example, red squirrels (*Tamiasciurus hudsonicus*) stash whitebark cones, then seeds, in large middens. These hoards are often raided by bears, and in general the midden seems to provide a poor habitat for germination; few if any seedlings establish from middens (Hutchins and Lanner 1982).

Determining the probabilities of seeds dispersing a given distance is much more difficult for these mutualisms, since by definition the seeds do not fall passively into seed traps. In the case of terrestrial dispersers (ants, chipmunks) seeds must be marked (painted or irradiated) and then recovered (see, for example, Vander Wall 1993 and 1994, or Heithaus 1986). Birds must be followed visually, and because of the long distances involved seed recovery is much more difficult. Appropriate probability functions are probably leptokurtic, with broad tails to account for this long-distance dispersal (Clark et al. 1998, 1999). In any event, this active dispersal is clearly not diffusive in nature, invalidating predictions from the Fisher/Kolmogorov model.

Integrodifference equation (IDE) models for seed dispersal and plant establishment can incorporate non-diffusive, probabilistic processes (Andersen 1991, Kot 1992, Neubert et al. 1995, Lewis 2000, Clark et al. 2001), and may fit more comfortably with the discrete yearly cycle of seed production and dispersal. If  $K(x - \xi)$  is the probability of a seed dispersing a distance  $x - \xi$  from a parent located at  $\xi$  to a target location  $x$ ,  $f$  is the number of seeds produced per adult,  $\sigma$  the fraction of seeds that disperse, and  $A_n(\xi)$  is the density of adults in year  $n$ , then the distribution of seeds produced in year  $n$  is

$$S_n(x) = \sigma f \int_{-\infty}^{\infty} A_n(\xi) K(x - \xi) d\xi \stackrel{\text{def}}{=} \sigma f K * A_n(x). \quad (2)$$

The latter definition in this mathematical statement is the spatial *convolution* of the two functions  $K$  and  $A_n$ . The fraction of seeds that germinate and establish with and without dispersal,  $g_0$  and  $g_1$ , determines how these seeds contribute to the density of adults in the next generation. One IDE rendition of these processes, using a Beverton-Holt nonlinearity (see Cushing et al. [2002: 7]) for a discussion of why this nonlinearity is the discrete analogue of continuous logistic growth) would then read

$$A_{n+1}(x) = (1 - \omega)A_n(x) + fa \frac{(1 - \sigma)g_0 A_n(x) + \sigma q_1 K * A_n(x)}{a + A_n(x)}. \quad (3)$$

The constant  $\omega$  reflects the rate at which adults die off. If the saturation constant  $a$  is chosen so that  $a = \omega A_{\max} / \{f[(1 - \sigma)g_0 + g_1 - \omega]\}$  and spatial dependence is suppressed, the resulting discrete model has two fixed points,  $A = 0$  (unstable) and  $A = A_{\max}$  max (uncon-

ditionally stable). The Beverton-Holt density dependence is intended to reflect the competition for space with adult plants that nondispersed seeds must undergo. This density dependence has the attractive feature that net fecundities are defined and positive for all positive densities of adults, and while it is possible for the population of adults to overshoot the "carrying capacity,"  $A = A_{\max}$ , due to dispersal into a region, the population relaxes monotonically to the asymptotic population at a rate proportional to  $\omega$ . However, so long as the derivative of the nonlinearity is nonzero at  $A = 0$  the functional form of the nonlinearity makes no difference in terms of predictions for migration rates, which depend only on the linear terms in Eq. 3. A discussion of why this is so, as well as methods for predicting speeds for a wave of invasion, are discussed in Kot et al. (1996).

The speeds predicted for IDE applied to Reid's paradox split into two classes. For  $K$  which have exponential or faster decay in their tails, speeds are finite and generally too small to account for the Holocene migration rates (Clark et al. 1998, Lewis 2000). For fat-tailed  $K$ , speeds are much greater, in fact increasing geometrically to infinity as a result of the lack of finite moments for fatter-than-exponential kernels. This can be accounted for by applying order statistics to determine expected distances for the farthest forward dispersers (Clark et al. 2001), yielding finite-speed estimates that agree more closely with observations of rapid migration. Details of the spatial structure of caching, which may be multimodal, are subsumed into the dispersal kernel.

In this paper we suggest a complementary technique based on the spatial behavior of the dispersers, as opposed to the seeds. Spatial heterogeneity is included explicitly (and essentially) at the scale of dispersal from individual plants. At this scale seed motion is assumed to be under the control of active agents, and the spatial dispersion of seeds can be complicated and situation dependent. Small copses can pop up far in advance of the main body of the population, dispersed actively to caching sites. This reflects a multimodality and spatial dependence at the small scale that would render the dispersal kernel approach invalid. At a larger scale, however, the method of homogenization will be used to examine emergent properties of migration. Large-scale probabilities of dispersal, computed deterministically from the small-scale heterogeneity, turn out to depend on distance alone, allowing usage of the dispersal kernel approach. The result is an IDE of the form of Eq. 3, but with spatial scales orders of magnitude larger than the scale of local seed distribution, and with dispersal probabilities conditioned on the spatial behavior of the active agents. Dispersal kernels constructed in this way predict migration rates an order of magnitude larger than "standard" kernels, which do not account directly for active dispersal. This is consistent with the intuition that active dispersers will en-

hance migration rates. Predictions based on this approach are calculated for a selection of mutualist dispersers.

METHODS

*Active dispersal model*

Neubert et al. (1995) describe a model for dispersal and settling of propagules from a point release (described by the Dirac delta function,  $\delta(x)$ ) at the location  $\xi$ :

$$\frac{\partial P}{\partial t} = D \frac{\partial^2 P}{\partial x^2} - \lambda P \quad P(x, t = 0) = \delta(x - \xi) \quad (4)$$

$$\frac{\partial S}{\partial t} = \lambda P \quad S(x, t = 0) = 0. \quad (5)$$

Here  $P$  is the distribution of airborne propagules,  $\xi$  is the location of the original seed source,  $\lambda$  is the rate of settling and  $D$  is the eddy diffusion constant. The distribution of seeds on the ground at any time is  $S$ ; after a season of seed dispersal the distribution of seeds is given by

$$K(x - \xi) = \lim_{t \rightarrow \infty} S(x, t) = \frac{1}{2\alpha} \exp\left[-\frac{|x - \xi|}{\alpha}\right] \quad (6)$$

with  $\alpha = \sqrt{D/\lambda}$ . This is often referred to as the ‘‘Laplace’’ or the ‘‘double exponential’’ distribution. One may interpret the results as the output of a Poisson process in which each particle starting at the origin has equal probability per distance traveled of being precipitated on the ground. The Laplace  $K$  is one of the family of dispersal kernels used in Eq. 3 to predict plant dispersal.

Consider a similar model for actively dispersed seeds. Animal cachers of seeds do not, in general, deposit seeds randomly or with equal probability per distance traveled. Instead they carry seeds to preselected sites rapidly, and then cache in these areas with high probability. Cache areas are chosen for a variety of reasons: ease of defense, proximity to nests and offspring, ease of recovery, avoidance of cache robbers, are a few of the obvious considerations. Nutcrackers and jays seem to choose relatively open areas with distinct visual cues, areas where rodents harvesting the same seeds are unlikely to look. Ants carry seeds to their nest, after which they are dispersed in the nearby vicinity. Chipmunks and squirrels carry nuts to hoard areas or middens in their territories to aid in defense and recovery. This means that for animal dispersers the rate of seed movement (loosely  $D$ ) and ‘‘settling’’ or caching ( $\lambda$ ) are functions of space. In general we would expect these two functions to be anticorrelated; dispersers move seeds most rapidly (high  $D$ ) between cache areas, where seeds are unlikely to be stored (low  $\lambda$ ). Conversely, seeds are most likely to be stashed (high  $\lambda$ ) in smaller cache sites, where little further seed movement occurs (low  $D$ ). These processes occur on

a scale much smaller than forest migration, and to define a model for these small-scale processes we introduce an ‘‘order parameter,’’  $\varepsilon \ll 1$ , which describes the ratio of small and large scales. The small-scale variable,  $y$ , is defined in relation to the large-scale  $x$  by  $y = x/\varepsilon$ . This has the effect of placing the  $x$  variable under a magnifying glass, so that order  $\varepsilon$  changes in  $x$  (i.e., small scale) become order one changes in  $y$ . With respect to this new variable, the simplest possible description of a caching structure would be

$$D(y) = \begin{cases} D_0 & y \in [0, l) \\ D_1 & y \in [l, L) \end{cases} \quad (7)$$

$$\lambda(y) = \begin{cases} \lambda_0 & y \in [0, l) \\ \lambda_1 & y \in [l, L) \end{cases}$$

(see Fig. 1). The functions  $D$  and  $\lambda$  will be assumed periodic in space, with period  $L$ , effectively creating an infinite array of cache areas, a distance  $L$  apart. This will streamline the development, but is not essential. The size of each cache area is given by  $l$ . In keeping with the anticorrelated and active nature of dispersal, we assume  $D_0 \ll D_1$  (rates of movement in cache areas at least an order of magnitude smaller than rates of movement between cache areas) and  $\lambda_0 \gg \lambda_1$ .

Given parameters described as in Eq. 7, the model for active dispersal becomes

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial x} \left[ D(y) \frac{\partial P}{\partial x} \right] - \lambda(y)P \quad P(x, t = 0) = \delta(x - \xi) \quad (8)$$

$$\frac{\partial S}{\partial t} = \lambda(y)P \quad S(x, t = 0) = 0. \quad (9)$$

The output from this model,

$$K(x, \xi, y) = \lim_{t \rightarrow \infty} S(x, y, t)$$

is a seed dispersal function which can be used to examine the consequences of active dispersal. It is not clear at this point that the resulting kernel, which could depend jointly on the location of the initial seed source as well as distance traveled from the seed source, would be suitable to use in an IDE formulation;  $K$  must depend on distance traveled *alone* (that is, only the variable combination  $x - \xi$ ) in order to predict invasion rates analytically. However, as we will discuss in subsequent sections and prove in Appendix A, the dispersal kernel describing large-scale consequences of the small-scale active dispersal described by Eqs. 8 and 9 depends only on distance from the source (measured on the large scale), provided  $D$  and  $\lambda$  depend on the small scale only. Thus, when large-scale dependences can be ignored, small-scale effects can be averaged in such a way as to fit perfectly into an IDE model at the large scale.

To illustrate the behavior of solutions to the joint system of generational adult development and growth,

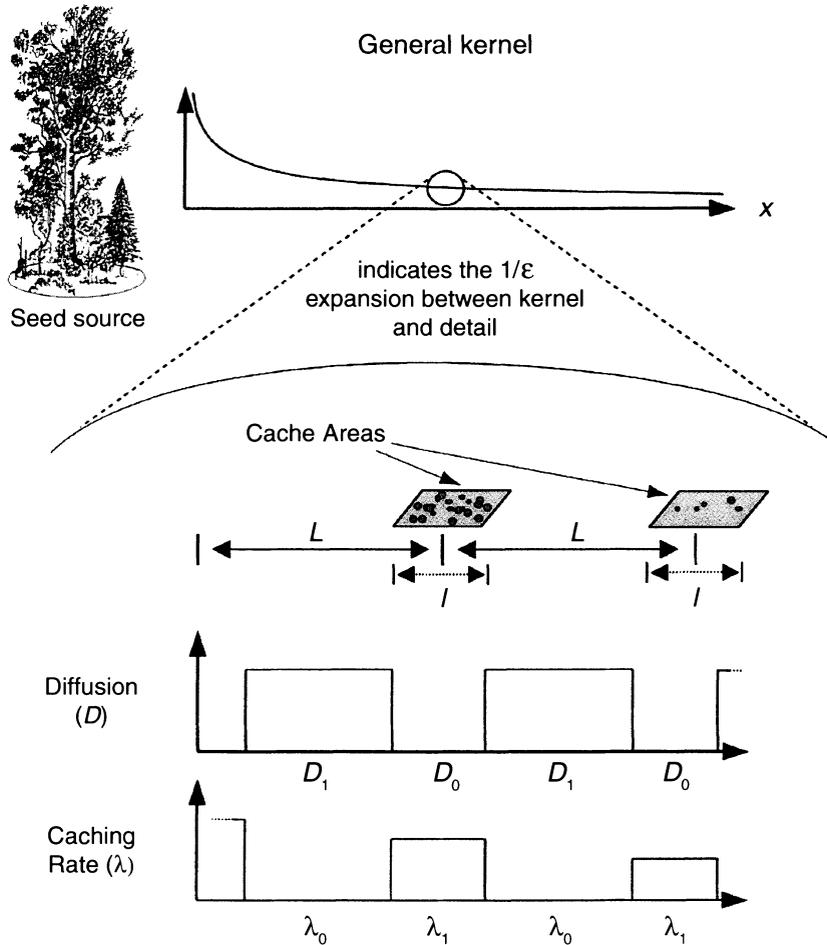


FIG. 1. Spatial structure (along  $x$ ) of the diffusion constant ( $D$ ) and rate of caching ( $\lambda$ ) for active dispersers. Regions of high  $D$  reflect rapid motion between caching sites. Within a caching site the rate of caching is high, since this is where the majority of seed deposition occurs.

(Eq. 3), with seasonal seed dispersal given by Eqs. 8 and 9, operating in a region characterized by active caching parameters (Eq. 7), numerical solutions were computed using parameters for *A. canadense* and harvester ants, discussed in *Dispersal of Asarum canadense by harvester ants*. A total of 280 ant colonies, each 0.4 m wide, were placed in a linear array, 2.15 m apart, over a 600-m interval. A grid size of 0.0366 m was used with 16 385 grid points, so that every caching area was resolved by at least 10 grid cells. An initial distribution of 2/m adults was placed in the interval between  $x = 0$  and  $x = 5$ . We assumed equal germination probabilities for dispersed and nondispersed seed,  $g_0 = g_1 = g$ , and following Cain et al. (1998), intrinsic rate of growth ( $f_g - \omega$ ) was taken to be 0.12  $\text{yr}^{-1}$  both within and between caching areas. Seed movement parameters were chosen to be  $D_0 = 0.89$ ,  $D_1 = 89$ ,  $\lambda_0 = 0.12$ ,  $\lambda_1 = 0.0012$ , and the model (Eqs. 8 and 9) was solved numerically over 250 days of active dispersal to determine the final density of seeds dispersed from adult plants. After this (very) computa-

tionally expensive step, the dispersed seeds underwent density-dependent competition and adults were updated as in Eq. 3, generating one year of dynamics. The entire simulation was run for 100 years, and results are depicted in Fig. 2.

The simulation shows a wave of invasion with interesting behavior. Since seeds end up preferentially dispersed to the vicinity of ant colonies, the effective reproduction rate is much higher in caching areas than between caching areas, creating the comb-like profile for adult plants. The asymptotic density between colonies (where very few seeds are deposited) is the location of the lower boundary of the wave of invasion above. Within caching areas the adult density has a fixed point of 3.4 adults/m, which is the value approached asymptotically behind the wave of invasion. Intermediate between these two values is  $A_{\text{max}} = 2$ , corresponding to the average dispersal. Directly behind the wave of invasion, dispersal serves as an averaging process, creating an overall carrying capacity commensurate with averaged dispersal (as predicted by the

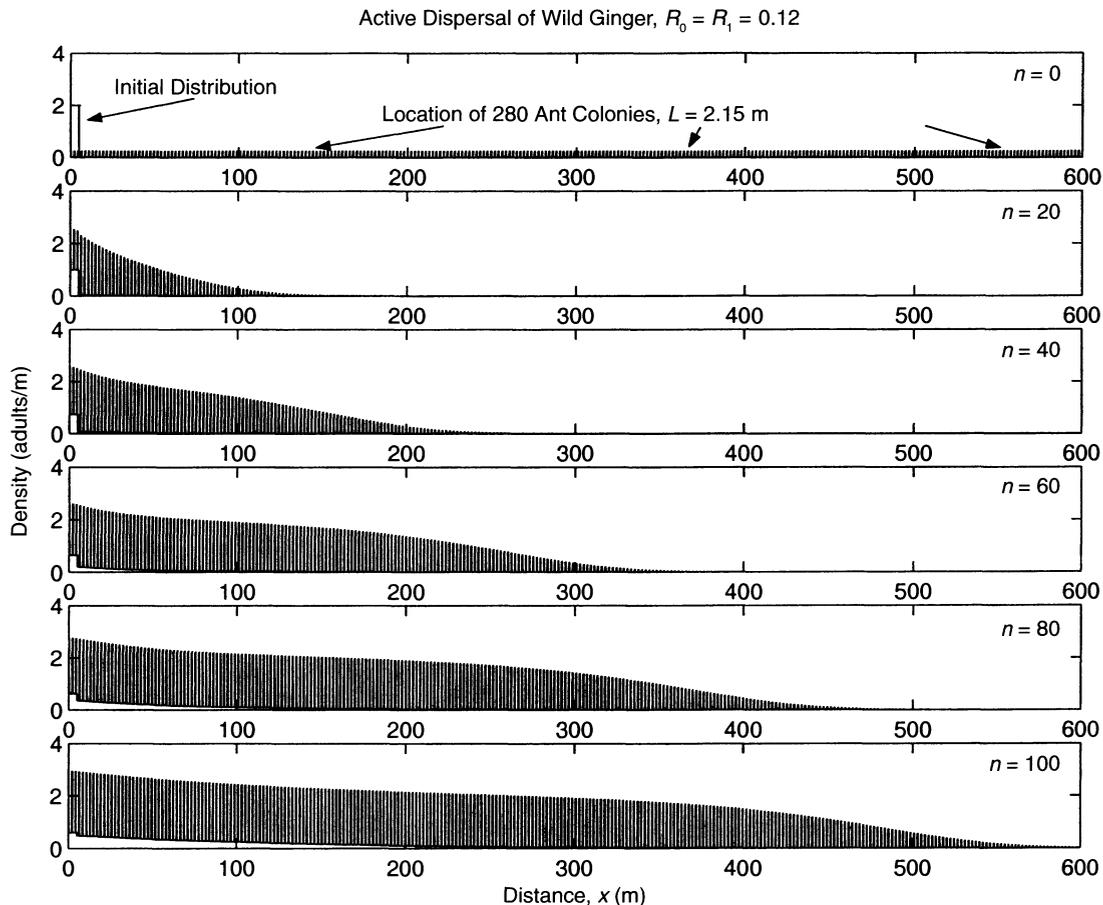


FIG. 2. Simulation of small-scale, active dispersal of seeds coupled with an integro-difference model for growth of adults. Ant colonies are of 0.4 m extent, separated by a distance of 2.15 m, and the initial location of adult plants is for  $x$  values between 0 and 5 m. Time slices ( $n$ ) are plotted every 20 generations and indicate the density of adult plants in that generation. Since seeds end up preferentially dispersed to the vicinity of ant colonies, the effective reproduction rate is much higher in caching areas than between caching areas, creating the comb-like profile for adult plants. The asymptotic density between colonies (where very few seeds are deposited) is set to 0.5, which is the location of the lower boundary of the wave of invasion above. Within caching areas the adult density has a fixed point of 3.4, which is the initial value approached behind the wave of invasion. The asymptotic adult density consistent with averaged dispersal is 2, which is the value approached immediately behind the wave of invasion. Subsequently, the population values relax to their local asymptotic values, defining the upper and lower values of the structure far behind the wave of invasion.

IDE model). Farther behind the initial colonization, dispersal matters less and the observed densities relax to the local carrying capacities, creating the stable, fine-scaled structure well behind the front. Predicting the speed of the wave of invasion will require the application of homogenization techniques, described in the next section.

#### *Multi-scale analysis and the method of homogenization*

Seeds may be transported great distances by active dispersal. Moreover, unlike the constant “rain of seeds” at a rate  $\lambda$  described by Eqs. 4 and 5, seeds in Eqs. 8 and 9 are not deposited rapidly (or at all) over much of the domain. Thus more seeds are available to disperse farther. It is important to allow for these long-distance effects and for the large scale of observation

at the outset. Accordingly the short-distance variable,  $y$ , and the observation scale,  $x$ , are related by  $x = \varepsilon y$ . The dimensionless parameter  $\varepsilon$  is taken to be small, describing the separation of scales between local, individual dispersal events and the much larger scale of observed movements over epochs. As an example, consider dispersal of *A. canadense* seeds by harvester ants, as described by Cain et al. (1998). The spatial structure of ant nests and dispersal varies on the order of meters; as discussed below in the case studies, we will take  $L = 2.15$  m and  $l = 0.4$  m (based on measurements from E. R. Heithaus, *personal communication*), and thus we would think of  $y$  in meters. On the other hand, dispersal more on the scale of hectares is required to describe the Holocene migration of *A. canadense*, so we would think of  $x$  in hundreds of meters. In this case, then,  $\varepsilon = 1/100 = 0.01$ .

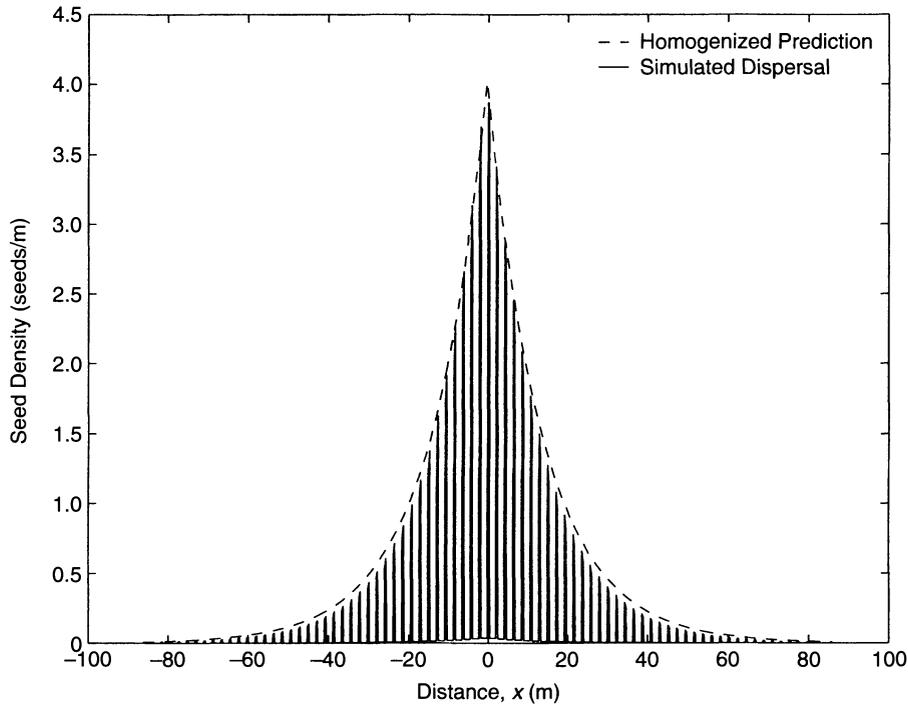


FIG. 3. Comparison of simulated (solid vertical lines) and predicted (dashed line) seed densities according to simulated diffusion and caching of seeds in a spatial grid of caching areas, as in Eq. 7. Parameters are selected to model dispersal of *A. canadense* by ants, with  $D_0 = 0.89$ ,  $D_1 = 89$ ,  $\lambda_0 = 0.12$ ,  $\lambda_1 = 0.0012$ ,  $L = 2.25$  m, and  $l = 0.4$  m. A source of 100 seeds is initially placed at  $x = 0$ , and dispersal proceeds as described by the spatially varying PDE model Eqs. 8 and 9 simulated for 250 days, resulting in the spiky distribution of seeds in cache areas. The homogenized prediction is the Laplace kernel, Eq. 19. In addition to the total seed number, the homogenized dispersal kernel is multiplied by a factor of  $\gamma^{-1}$  to facilitate comparison of distribution peaks. Because of the discontinuous parameters, errors are of the order  $\Delta x \approx 0.1$ , meaning that the agreement between the two curves is as good as can be resolved given the numerical accuracy of the simulation.

Since small-scale and large-scale variability are at some level independent, inclusion of both scales results in a change of variables:

$$P = P(x, y = \varepsilon^{-1}x, t)$$

$$\frac{\partial}{\partial x} \rightarrow \frac{\partial}{\partial x} + \frac{1}{\varepsilon} \frac{\partial}{\partial y}. \quad (10)$$

(See Holmes [1995] for a full discussion of multiscale analysis.) This gives a version of the classic multiscale diffusion equation:

$$\frac{\partial P}{\partial t} = \left( \frac{\partial}{\partial x} + \frac{1}{\varepsilon} \frac{\partial}{\partial y} \right) \left[ D(y) \left( \frac{\partial}{\partial x} + \frac{1}{\varepsilon} \frac{\partial}{\partial y} \right) P \right] - \lambda(y)P. \quad (11)$$

The idea of the homogenization approach is that the small-scale variability is such a large effect (of size  $1/\varepsilon^2$  in Eq. 11) that it must equilibrate very rapidly. In dispersal terms, active dispersers will rapidly move seeds to caching areas. In fact, at leading order the effect of caching areas being separated by zones of rapid dispersal (large  $D$  in Fig. 1) is to segment the double exponential distribution (Eq. 6) and then spread the segments apart to occupy only the caching areas, exactly as though the regions between caching areas were so easy to pass through that they present

no barrier to seed movement. An illustration of results of numerical solution to Eqs. 8 and 9, with parameters given in a spatial array according to Eq. 7, is shown in Fig. 3.

Intuitively, one could predict the effect of this spatial grid of dispersal/caching zones by comparing residence times. Following Okubo (1980), the residence time of individual dispersers can be thought of as inversely proportional to the diffusion constant. In a small patch of extent  $\Delta y$  (in the  $x$  direction, and unit extent in the perpendicular direction), residence time is proportional to  $\Delta y/D(y)$ . The average residence time in a region located at the large-scale coordinate  $x$  is then

$$\frac{1}{2z} \int_{-z}^z \frac{dy}{D(y)}. \quad (12)$$

The average is taken over a length  $2z$  which is long enough to sample the small-scale variation, but small compared to changes in  $x$ . Similarly, the average per time probability of seed caching in the same area is

$$\frac{1}{2z} \int_{-z}^z \lambda(y) dy. \quad (13)$$

Mathematically formalizing these notions and remov-

ing explicit dependence on any particular averaging length, we define average residence time ( $\widehat{D}^{-1}$ ) and average caching probability ( $\widehat{\lambda}$ ) as follows:

$$\widehat{D}^{-1} = \lim_{z \rightarrow \infty} \frac{1}{2z} \int_{-z}^z \frac{dy}{D(y)}$$

$$\widehat{\lambda} = \lim_{z \rightarrow \infty} \frac{1}{2z} \int_{-z}^z \lambda(y) dy. \quad (14)$$

The probability *per area* of seeds being cached should be the product of the residence time in the area and the per time probability of caching, and thus we can write the probability *per distance* as the square root:

$$\text{per distance probability of caching} = \sqrt{\widehat{D}^{-1} \widehat{\lambda}}. \quad (15)$$

At the large scale, we might expect that the output of this averaged procedure would be a Poisson process, with seeds distributed proportionally to  $\exp(-\sqrt{\widehat{D}^{-1} \widehat{\lambda}} |x - \xi|)$ .

As it turns out, the results of this intuitive homogenization can be derived rigorously, as discussed in Appendix A. Eq. 11 has been considered extensively as a model for transport of contaminants in porous media or diffusion of heat in layered materials, and the result that net diffusion relates to the harmonic mean of the diffusion constant was derived by Shigesada et al. (1987) for the diffusive dispersal and reproduction of a population in a very heterogeneous environment. A leading-order (in  $\epsilon$ ) solution to Eq. 11 is given by  $P = P_0(x, t) + O(\epsilon)$ , with  $P_0$  satisfying the homogenized equation

$$\frac{\partial}{\partial t} P_0 = \frac{\partial}{\partial x} \frac{1}{\widehat{D}^{-1}} \frac{\partial}{\partial x} P_0 - \widehat{\lambda} P_0 \quad (16)$$

and cached seeds satisfying

$$\frac{\partial}{\partial t} S = \widehat{\lambda} P_0. \quad (17)$$

It is important to note that the averaging procedure has removed dependence of the parameters on the small-scale variable,  $y$  (although they could still include terms reflecting the *trend* in  $y$  at the larger scale, as described in Appendix A). The coefficients could still depend on the large scale,  $x$ , but that would invalidate use of the integrodifference approach and dispersal kernel to resolve the effects of homogenized seed dispersal. For the particular case of the parameter model (Eq. 7), the homogenized constants become

$$\widehat{D}^{-1} = \frac{(L - l)}{LD_0} + \frac{l}{LD_1}$$

$$\widehat{\lambda} = \frac{(L - l)}{L} \lambda_0 + \frac{l}{L} \lambda_1. \quad (18)$$

These parameters are now spatially constant, and consequently the large-scale dispersal kernel corresponding to solution of (Eq. 16) is as follows:

$$K_{\text{hom}} = \lim_{t \rightarrow \infty} S_0 = \frac{1}{2} \sqrt{\widehat{\lambda} \widehat{D}^{-1}} \exp[-\sqrt{\widehat{\lambda} \widehat{D}^{-1}} |x - \xi|]. \quad (19)$$

This kernel is compared with simulated solutions to the full set of Eqs. 8 and 9 in Fig. 3.

### Homogenized model for active dispersal

The homogenized version of Eq. 3 becomes

$$A_{n+1} = (1 - \omega)A_n$$

$$+ \left[ \frac{(1 - \sigma)fg_0 A_n}{R_0 + (1 - \sigma)\omega} + \frac{\sigma fg_1 K_{\text{hom}} \times A_n}{R_1 + \sigma\omega} \right] \frac{a}{a + A_n} \quad (20)$$

where the convolution is taken over the long space scale. The net growth rates for dispersed and undispersed seeds,  $R_0$  and  $R_1$ , are defined to include appropriate relative weights of the adult death rate. The parameter  $a$  relates to the asymptotic density ( $A = A_{\text{max}}$ , the ‘‘carrying capacity’’) via

$$a = \frac{\omega A_{\text{max}}}{f[(1 - \sigma)R_0 + \sigma R_1]}. \quad (21)$$

Returning to the homogenized coefficients in Eq. 14, recall that movement between cache sites is large compared with movement inside ( $D_0 \ll D_1$ ), while the rate of caching between cache areas is negligible ( $\lambda_1 \ll \lambda_0$ ). We may therefore neglect  $\lambda_1$  and  $1/D_1$  in Eq. 18. If we define  $\gamma = l/L$  to be the spatial scale ratio of caching, then Eq. 18 is approximated by  $\widehat{D} \approx \gamma/D_0$  and  $\widehat{\lambda} \approx \gamma\lambda_0$ . The large-scale dispersal kernel, Eq. 19, becomes Eq. 6 with  $\alpha = \gamma^{-1} \sqrt{D_0/\lambda_0}$ .

### Asymptotic rates of spread

In the previous section we saw that the mean dispersal distance is increased by a factor inversely proportional to the caching scale ratio. It should not then be too surprising that the rate of invasion also scales with  $\gamma^{-1}$ . For a model of the form of Eq. 20, with a dispersal kernel (Eq. 6), the asymptotic distance traveled per iteration,  $c$ , satisfies a system of equations (following Kot et al. 1996) involving the spatial decay rate of the wave of invasion ( $s$ ) and the moment-generating function,  $M(s; \alpha)$ , associated with  $K$ . The parameter  $\alpha$  is the mean distance seeds are placed from the parent plant. (In the case of Eq. 6 the mean distance is  $\alpha = \sqrt{D\lambda^{-1}}$ .) In general these equations can not be solved in closed form, which makes comparison of invasion rates under parametric changes difficult, as well as rendering it problematic to determine parametric sensitivities of invasion rates. However, as shown in Appendix B, if  $R_0 = \rho R_1$  (rate of germination in the vicinity of the parent is  $\rho$  times greater than away from the parent) and  $R_1 \ll 1$  (small intrinsic rate of growth per dispersal season),  $c$  may be approximated as follows:

$$\begin{aligned}
c \approx & 4\sqrt{2}\alpha R_1 \rho^{3/2} \frac{\sqrt{3 + 2\rho - \sqrt{9 + 8\rho}}}{(9 + 8\rho - 3\sqrt{9 + 8\rho})^2} \\
& + \langle \{16\alpha R_1^2 \rho^{5/2} \sqrt{6 + 4\rho - 2\sqrt{9 + 8\rho}} \\
& \quad \times [27 + 33\rho + 8\rho^2 - (9 + 7\rho)\sqrt{9 + 8\rho}] \rangle \\
& \div [(-3 + \sqrt{9 + 8\rho})^4 (-9 - 8\rho + 3\sqrt{9 + 8\rho})].
\end{aligned}
\tag{22}$$

It can be shown that this is an underprediction of the actual speed, and that the size of the error scales with  $\alpha R_1^3 (1 + \rho^3)$ . Small intrinsic growth rates are reasonable for the three case studies we consider: dispersal of *A. canadense* by harvester ants, of oak trees by jays, and of whitebark pine by Clark's Nutcracker, which allows us to use this approximate expression for  $c$  as a surrogate for the actual solution. Most important to note here is that predicted migration rates relate directly to the mean dispersal distance,  $\alpha$ , at all orders in the asymptotic representation. The mean dispersal distance for homogenized dispersal scales directly, in turn, with the "standard," in-patch dispersal distance ( $\alpha_{\text{std}}$ ) and inversely with the caching scale ratio,  $\gamma = l/L$ . Comparison between standard and homogenized estimates for migration speeds can be calculated by the simple expedient of multiplying by  $\gamma^{-1}$ . Thus, not only is  $\gamma$  an easy parameter to understand, observe and measure, but also the net effect of the homogenization is to directly re-scale standard estimates of migration rates by a factor of  $\gamma^{-1}$ .

## RESULTS

### *Dispersal of Asarum canadense by harvester ants*

*Asarum canadense* is a woodland herb which grows in the understory of deciduous forests. Plants reproduce vegetatively or by seeds, which are dispersed by harvester ants (Heithaus 1986, Cain et al. 1998). Numerous authors have measured average carry distance for seeds from artificial depots, reporting means in the vicinity of 1 m (0.8 m by Heithaus, 1.54 m by Cain et al.). This suggests a spacing between ant nests on the order of 2 m. Heithaus mapped two  $10 \times 10$  m plots in West Virginia, finding 21 and 25 nests per plot, distributed randomly (E. R. Heithaus, *personal communication*), which gives a mean spacing between nests of  $\sim 2.15$  m. To connect with the spatial diffusion/deposition model (Eq. 7), we therefore take  $L = 2.15$  m. This falls within extremes of an alternate measure, twice the mean carry distance from artificial depots (giving 1.6 m and 3.08 m in the cases of Heithaus and Cain, respectively).

Most seeds are subsequently expelled from nests, more or less randomly at a distance of  $<0.5$  m (E. R. Heithaus, *personal communication*). Heithaus (1986) reported a mean clump size of 1.51 seeds among 150 scattered seeds, where a clump was defined as  $>1$  seed within 2 cm. Working backwards, one may estimate

that those seeds were scattered radially to a distance of 0.2 m. We therefore take  $l$  to be 0.4 m.

Cain et al. (1998) estimate an in-patch diffusivity  $D_0 = 0.89$  m<sup>2</sup>/yr, as well as an intrinsic growth rate of  $r = 0.12$  yr<sup>-1</sup>. This gives a "standard" length scale for dispersal,  $\alpha_{\text{std}} = \sqrt{D_0/r} = 2.72$  m. Heithaus (1986) reports that  $<1\%$  of seeds scattered by ants survive predation, but that other processes (germination, establishment) are not particularly affected by ant dispersal. However, Cain et al. report that all seeds they put out were found and dispersed. We therefore take  $R_0 = r = 0.12$  and  $R_1 = R_0 = 0.12$ . This generates a standard estimate for the rate of progress  $C_{\text{std}} = 0.4199$  m/yr.  $\alpha_{\text{std}} = 1.14$  m/yr. This is larger than the 0.65 m/yr estimated by Cain et al. using the purely diffusive methods of Skellam (1951), Okubo (1980) and Andow et al. (1990). The difference is due to two factors: the perennial form of the IDE model (Eq. 3) and the relatively broad profile of the exponential dispersal kernel.

However, over 16 000 years this new, "improved"  $C_{\text{std}}$  still does not resolve the paradox of rapid Holocene dispersal. Estimated travel distance for 16 000 years of dispersal is  $\mathcal{D} = 16 C_{\text{std}} = 18.27$  km. By contrast, Cain et al. (1998) estimate that *A. canadense* must have traveled a minimum of 450 km, and more likely 700–1900 km. The effect of homogenization is to increase the standard estimates by a factor of  $\sim \gamma^{-1}$ , where  $\gamma = l/L$  is the caching scale ratio 0.186 for ant/*A. canadense* complex). Thus, estimates for the yearly progress become  $c_{\text{hom}} = (1/\gamma)c_{\text{std}} = 6.14$  m/yr, and total distance traveled during the Holocene increases  $\mathcal{D}_{\text{hom}} = (1/\gamma)\mathcal{D}_{\text{std}} = 98$  km.

It should be noted that this estimate is optimistic in the sense that it ignores the age structure of the population. While a fully developed population with stable age structure will always have seed producing individuals, at the leading edge of an invasion there will not be a stable age structure and the wave of invasion cannot be expected to travel as fast as the prediction above. A recent analysis of migration rates for age-structured populations by Neubert and Caswell (2000) indicates that the actual speed of advance may lag the predicted speed (using the growth rate of the stable age distribution) by as much as a factor of two. On the other hand, the speeds are much greater than those predicted by using only the raw seed-to-seed-producer generation time. In either case, it is clear that the effect of active dispersal, incorporated through the technique of homogenization, greatly accelerates migration, although not enough, in the case of the parameters we have assumed, to fully account for the Holocene migration.

To test the predicted values (using homogenization) and the asymptotic theory, as well as to determine what sort of parameter values would be required to generate sufficiently rapid rates of migration, we performed simulations of invasions for a variety of growth parameters. Growth parameters were fixed at either  $R_0$  (holding

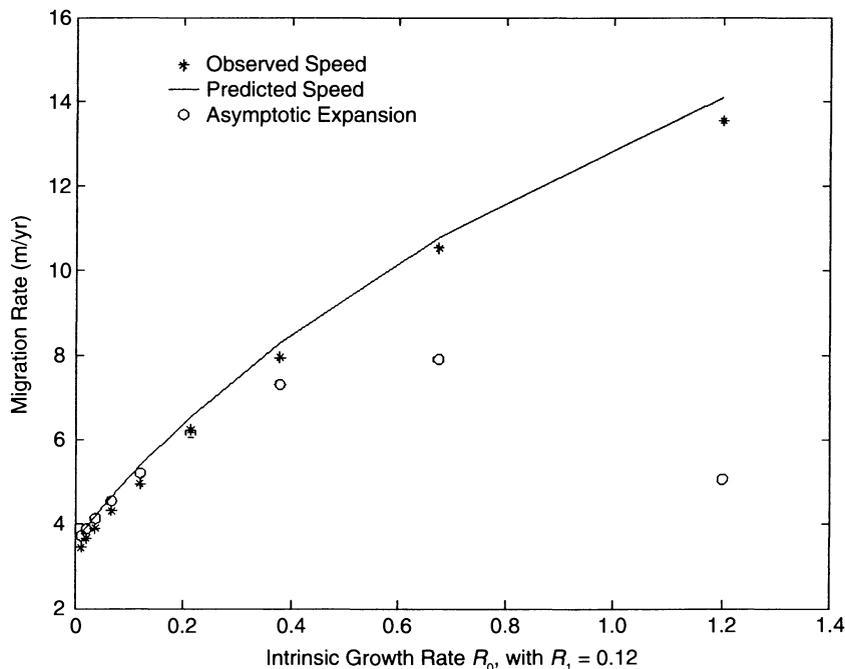


FIG. 4. Comparison of migration rates observed in simulations (stars, \*) with predicted speeds calculated via numerical evaluation of the minimum speed criteria, Eqs. B.8 and B.9 of the appendix (solid line), and speeds predicted by the asymptotic expansion (Eq. 22) (open circles). The intrinsic growth rate of dispersed seeds was held fixed at  $R_1 = 0.12$ , and growth rates of non-dispersed seeds were varied in a logarithmic range from one-tenth to ten times as large as  $R_1$ . The simulations are for dispersal (by ants) and growth of *A. canadense* (see *Results: Dispersal of Asarum canadense by harvester ants*), with spatially explicit locations for colonies and caching behavior and dispersal evaluated by direct numerical integration of (Eqs. 8 and 9) for each season of dispersal.

growth rate of nondispersed seeds constant) or  $R_1 = 0.12$  (holding growth rate of dispersed seeds constant). The parameter  $\rho$  was then varied in an exponential range varying from 0.1 to 10 to test either the sensitivity of speed results to variations in growth rates of dispersed seeds, (see Fig. 4) or to test sensitivity to growth rates of seeds left in the vicinity of parents, (see Fig. 5). Simulations were performed with 16 385 grid points at a spacing of 0.0366 m, so that every caching area was resolved by at least 10 grid cells. An initial distribution of one adult per meter was placed in the interval between  $x = 0$  and  $x = 5$ , and then allowed to evolve according to the system (Eqs. 3, 8, 9). Each year dispersal was evaluated by numerically solving the spatially explicit diffusion equations (Eqs. 8 and 9) over 250 days, followed Beverton-Holt competition among seedlings and mortality among adults (Eq. 3). The “farthest forward” position of adult density exceeding  $A_n = 0.5$  was recorded in each of 50 simulated generations, and a line was fit to the position in the last 10 generations to determine the asymptotic rate of advance. These results are compared with numerical solutions for the speed of the wave of advance (solving Eqs. B.8 and B.9 of Appendix B numerically) as well as the asymptotic expansion given by Eq. 22. Results are shown in Figs. 4 and 5.

Predicted and observed speeds compared well for all parameter ranges, indicating that the homogenized theory agrees with the much more complicated numerical computations. Moreover, the asymptotic theory, which will be used in the absence of numerical simulation for examples below, agreed quite well with both predicted and observed speeds in regimes for which both  $R_0$  and  $R_1$  are small, as should be expected. Perhaps more important, the increase of intrinsic growth rate required to produce results minimally consistent with Cain et al.’s 1998 prediction for required Holocene migration can be estimated. More than a fourfold increase in speeds would be required to raise the estimate of distance traveled to the minimum of 450 km required. This, in turn, would require  $R_1$  to be  $>1.2$ , or  $\rho$  to be 0.1 or less. Biologically, this would mean that dispersed seeds would need to be at least 10 times more likely to germinate than non-dispersed seeds—which is probably not the case for *A. canadense*.

#### Two less detailed examples

*Dispersal of oaks by jays.*—The spread of the genus *Quercus* during the Holocene is the essence of Reid’s paradox; over a period of  $\sim 5000$  years oaks spread from refuges in the Gulf Coast states of the United States to their current range. This feat required an av-

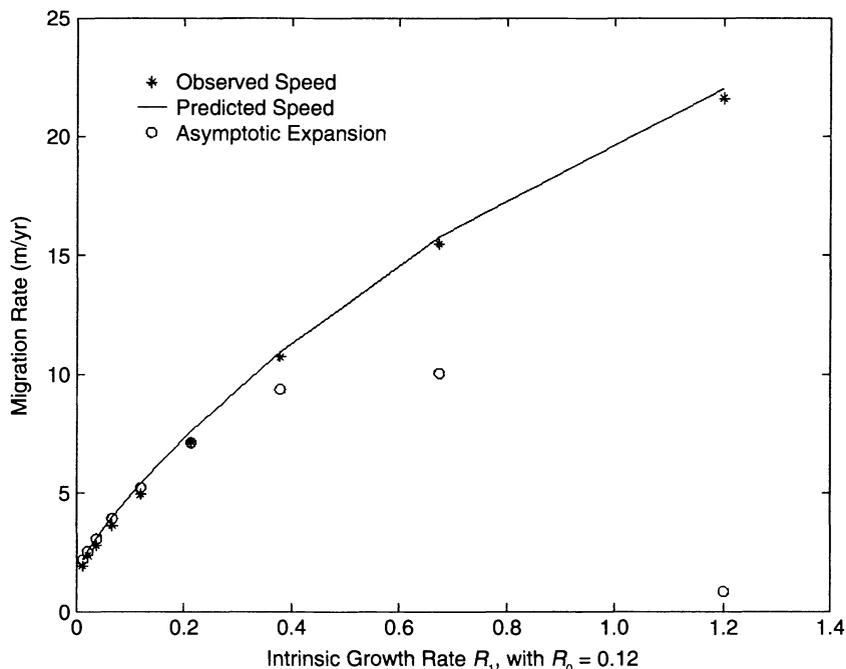


FIG. 5. Comparison of migration rates observed in simulations (stars, \*) with predicted speeds calculated via numerical evaluation of the minimum speed criteria, Eqs. B.8 and B.9 of the Appendix B (solid line), and speeds predicted by the asymptotic expansion (Eq. 22) (open circles). The simulations are for dispersal (by ants) and growth of *A. canadense* (see Results: Dispersal of *Asarum canadense* by harvester ants) with spatially explicit locations for colonies and caching behavior and dispersal evaluated by direct numerical integration of Eqs. 8 and 9 for each season of dispersal. Intrinsic growth rate of dispersed seeds was held fixed at  $R_0 = 0.12$ , and growth rates of nondispersed seeds were varied in a logarithmic range from 0.1 to 10 times as large as  $R_0$ .

erage speed of 350 m/yr (Davis 1981). In the case of England, oaks expanded 1000 km north over a period of <10 000 years (Clark et al. 1998). During this period, migration rates varied between 150 and 500 m/yr, with fastest speeds at early stages of the invasion, and with a considerable slowing (to 75–150 m/yr) toward the end of the migration process (Huntley and Birks 1983).

Skellam's (1951) mathematical treatment of Reid's (1899) estimates for the theoretical spread of oaks is based on an intrinsic growth rate ( $\gamma^2$  in Skellam 1951) derived from the estimate that  $9 \times 10^6$  oaks are produced over 300 years of seed production, including a 1% survival rate of acorns. Hence, the intrinsic growth rate,  $r$ , of oaks satisfies

$$(e^r)^{300} = 9 \times 10^6$$

or

$$r_{\text{Skell}} = \frac{1}{300} \ln(9 \times 10^6) \approx 0.0534. \quad (23)$$

In the absence of any other estimate for the intrinsic growth rate of oaks, we will use this parameter in the following discussion.

Estimates for transport and predation fractions of acorns taken by jays vary, and little is known of their eventual fate. Darley-Hill and Johnson (1981) estimate

that one-third of seeds are eaten, while Johnson and Webb (1989) estimate only 5–6% of seeds produced were actually cached. On the other hand, Vander Wall (1990: 189) states that 54% of acorns are transported and 20% eaten immediately by jays. One must add to this picture acorn masting. In most years many acorns are produced on a particular tree, and it is conceivable that harvesting/caching percentages vary widely compared to a nonmasting tree. All authors agree, however, that jays preferentially select high-quality seed (by tapping it), leaving low-quality seed near the parent. Observers (Johnson and Adkisson 1985, Vander Wall 1990) note that seeds are cached in areas more generally favorable to germination and establishment than near the parent. These sites include meadows and fields, where competitors for acorns are less likely to be looking for the seeds.

Vander Wall, Johnson, and Adkisson all mention that cache sites are scattered "within a few meters of one another," which suggests  $l = 10$  m. Johnson and Webb (1989) mention that pairs of jays inhabit forest patches of size 0.8–24 ha. Taking 25 ha to be the size of an area that strictly includes one jay territory and abuts another (thus it may include meadows around a 0.8-ha copse of trees, for example), the average spacing between the center of territories would be  $L = [25 \times (100 \text{ m})^2]^{1/2} = 500$  m. This seems reasonably consistent with

observations that jays fly a mean of 1 km on their caching flights (Darley-Hill and Johnson 1981, Johnson and Webb 1989).

We still lack an average within-patch mean spatial distance of dispersal,  $\alpha_{\text{std}}$ . Fortunately, a recent statistical analysis by Clark et al. (1999) gives  $\alpha_{\text{std}} = 13.4$  m. The relative intrinsic growth rates of dispersed and nondispersed seeds are basically unknown. Certainly jays must eat many of the seeds, after which the remainder take their chances, but in more favorable environments. Based on the various considerations above, we take  $R_1 = r_{\text{skel}}$  and then estimate that this is only 0.1 of the net fecundity (that is, <10% of seeds are cached and *not* recovered by jays). Thus  $\rho = 10$  in Eq. 22. Making the assumption that  $\alpha_{\text{std}} = \sqrt{D_0/\lambda_0}$ , we get the standard estimate for rate of spread  $c_{\text{std}} = 10.34R_1\alpha_{\text{std}} = 7.40$  m/yr. Including now the scaling factor from homogenization,  $\gamma = l/L = 1/50$ , gives  $c_{\text{hom}} = (1/\gamma)c_{\text{std}} = 370$  m/yr, which is certainly in the ballpark with Davis' (1981) estimate of 350 m/yr. In the case of Great Britain, if recolonization by *Quercus* occurred over 5000 years (as in North America), then the total distance traveled was 1850 km, which is more than sufficient to resolve Reid's Paradox.

*Clark's nutcracker and whitebark pine.*—As with oaks, pines apparently migrated very rapidly during the Holocene. Palynological evidence can only distinguish between the two pollen types haploxylon (mostly five-needle pines) and diploxylon (mostly two- and three-needle pines). Pines had their glacial retreats close to the ice shields (Huntley and Birks 1983, Delcourt and Delcourt 1987). The diploxylon pollen type is reported to have migrated extremely fast, as much as 1500 m/yr in the early Holocene, with migration rates slowing down considerably later in the Holocene (Huntley and Birks 1983). For the haploxylon pollen type, migration is more difficult to trace. In Europe, this subgenus (mostly *Pinus cembra*) seems to have survived the Ice Age even in the Alps (Huntley and Birks 1983). It has spread from there along the Alps with little additional terrain gained. Similar information can be derived from the North American Pollen database (NAPD). Data from NAPD show that pine has been present in the late glacial time at 45° N along the Rocky Mountains, with many sites from New Mexico/Arizona reporting the presence of *Pinus* spp. at 15 000 yr BP. Lacking better taxonomic resolution, we assume that *Pinus albicaulis* had its refugia somewhere between 35° N and 45° N. Today its range extends to 55° N, where the first pine pollens are reported as early as 9000 BP. This makes for a migration rate of 10–20° latitude within 6000–15 000 years, equating to a migration rate of 74–370 m/yr, with 180 m/yr being a conservative but realistic estimate.

Mimicking Skellam's reasoning for oak, we take the average seeding life of *P. albicaulis* as 400 years (taking into account a conservative ~100 years to grow

from seedling to fruiting adult, following Lanner [1996]). This estimate is reinforced by comparison with Eurasian stone pines (*Pinus cembra*). Lang (1994) in his review cites the maximum age for *P. cembra* as 1200 years and the minimum age of reproduction as 40 years in open stands. Schmidt (1918) reports the minimum age of reproduction more generally as between 50 and 100 years for *P. cembra*. Assuming fewer seeds per parent than oak (by a factor of three) and a survival probability of 0.001% per seed (relative to the 1% probability in Skellam's oak estimate), each adult whitebark may produce 3000 seedlings over its life. Thus,

$$(e^r)^{400} = 3 \times 10^3 \Rightarrow r = \frac{1}{400} \ln(3000) \approx 0.02. \quad (24)$$

This agrees with a recent study of succession in whitebark communities (Perkins 2000). Perkins estimates a total of 899 trees/ha on five undisturbed whitebark stands in 1930 and 3160 trees/ha on those stands 68 years later, in 1998. This gives

$$(e^r)^{68} = \frac{3160}{899} \Rightarrow r = \frac{1}{68} \ln(3.515) \approx 0.018. \quad (25)$$

In addition, Perkins' estimates were for exposed, high-altitude ridgelines in south-central Idaho, where one might expect population growth to be suppressed. Based on these considerations we will take the intrinsic growth rate of whitebark pine to be 0.02/yr<sup>-1</sup>.

According to Lanner (1980, 1982), Hutchins and Lanner (1982), and Vander Wall (1990), predation on nondispersed whitebark seeds is extreme, verging on 100%, and vertebrates other than nutcrackers seem to function as pure seed predators. We therefore take  $R_0 = 0$  and  $R_1 = r = 0.02$ ; that is, all reproduction is from dispersed seeds. This gives a prediction for rates of invasion conditioned on the dispersal scale,  $\alpha$ :  $c = 2.56 \times R_1 \times \alpha = 0.0512 \alpha$ . It remains to determine dispersal scales and the caching ratio,  $\gamma$ .

In the western United States, the interdependence of Clark's Nutcracker and whitebark pine has been well documented by Lanner and co-workers. Hutchins and Lanner (1982), Vander Wall (1990) and Lanner (1996) all attest to the high frequency of 5–15 km dispersal flights by nutcrackers, often to open areas and ridgelines. Caches are spread in lees and exposures of ridgelines where differing areas of seed cache are exposed through the year as snow accretes, melts, and is blown free. This simultaneously protects caches from other seed predators and allows for year-round access. Nutcrackers use communal scatter-hoarding areas, reported by Lanner (1980, 1996) to be 100 m<sup>2</sup> in area, and by VanderWall (1990) to be composed of individual caches within meters of one another. A like fraction of seeds are cached near the parent tree. We therefore take  $l = \alpha_{\text{std}} = 10$  m as a reasonable cross-sectional length for a communal storage area. The length scale between

caches we take to be  $L = 5$  km, closer to the shortest (1 km) than the longest (22 km) reported nonlocal caching flights (reported by Hutchins and Lanner 1982 and Vander Wall and Balda 1977, respectively).

These parameters give a "standard" estimate for whitebark pine migration rates of  $c_{\text{std}} = 0.0512 \times (10) = 0.512$  m/yr. In this case the caching scale ratio is  $\gamma = l/L = 1/500$ , and multiplying by a factor of  $\gamma^{-1}$  gives the homogenized estimate of migration rates,  $c_{\text{hom}} = 500 \times c_{\text{std}} = 256$  m/yr. This estimate compares favorably with a conservative assumption of estimated migration rates derived from NAPD data.

#### DISCUSSION

It seems that a spatially explicit model for active dispersal of seeds, coupled with a minimal model for foraging/caching behavior, may be sufficient to account for hitherto anomalous migration rates during the Holocene, without using the order-statistics approach. Even without active dispersal using exponentially tailed dispersal kernels, the perennial plant model (Eq. 3) generally predicts more rapid spread than a purely diffusive (Gaussian kernel) dispersal. When a spatial caching scale parameter

$$\gamma = \frac{\text{extent of cache area}}{\text{mean separation between cache areas}} \quad (26)$$

is included via the method of homogenization, estimates of invasion rate increase again by a factor of  $\gamma^{-1}$ . These two factors are capable of accounting for the two to three order of magnitude discrepancy between dispersal predictions and observations, often referred to as Reid's Paradox.

This approach is only theoretical, with "proof of principle" supplied by three examples gleaned from the literature. Predictions are quite favorable, but can vary by an order of magnitude depending on the parameters. These parameters may even have changed during the Holocene, partly due to the evolution of animal behavior and the ecology of plants, and partly due to the fact that the competitive balance of plants and animals has changed considerably during this period of time. In addition, age structure will generally decrease these predictions, and of course one may question the realism of the PDE-based mechanistic dispersal model for seeds and active dispersers. Confirmation of this theory must wait on careful assessment of parameter values discussed in this manuscript, particularly the caching scale ratio and the two intrinsic growth rates (from dispersed and nondispersed seeds). Given these parameters and a "standard" estimate of RMS distance of seeds from parents, the proposed theory delivers unequivocal migration rate estimates, which can be tested in the arena of observation. As suggested by Cain et al. (1998), the sorts of measurements required to get at these parameters would be very useful for bridging the gap between theory and observation.

It should be noted that the method of homogenization achieves a natural integration across scales, from local (plant–daughter) to medium (active dispersal) to large (landscape migration). The homogenized results presented here are appropriate for comparison with the epochal migration of forest trees across a continent, and predict a pattern of uniform spread at the landscape scale. At the intermediate scale the process is quite different. Small stands of trees spring up from seed caches at some distance from parents, while a more gentle, local dispersal process fills in the gaps. In the meantime, however, the forest has jumped ahead again, to caches farther afield. The mechanism of advance is explicitly abrupt and patchy at this intermediate scale, but at a larger scale quite predictable and uniform. Best of all, the machinery of homogenization mathematically connects these processes. One need not propose different models at different scales, and then struggle to connect them; the correct large-scale model emerges naturally through homogenization applied at smaller scales.

Homogenization techniques are not restricted to one spatial dimension, though the results presented here are unidimensional. One may calculate homogenized coefficients for multidimensionally structured landscapes and dispersal behaviors. The calculations are more intensive, but explicit. Also, the method is not restricted to dispersal caused by hoarding birds and small rodents as dispersal agents. The method is generally applicable to integrating two dispersal processes, one rapid and one slow. In this generalized form,  $L$  can be thought of as the mean scale over which the fast process works and  $l$  as the mean scale over which the slow process works. Our goal is to adapt these techniques to construct a landscape-scale reforestation model, with realistic seed dispersal kernels and a dynamic tree population submodel. This model will allow us to test hypotheses regarding the relative importance of individual life history characteristics related to seed dispersal on tree migration rates.

Such simulations, combined with extended investigations of past and present migration rates of trees, will also help to resolve the controversy regarding whether trees *are* actually limited in their migration by dispersal or by climate, i.e., whether dispersal by plants is an expression of dynamic equilibrium with climate (Webb 1986, Prentice et al. 1991), or instead shows signs of lags in the readjustment of vegetation composition due to difference in migration rates (Bennett 1983, Davis 1989, Lotter 2000). This problem is not easy to resolve due both to uncertainties in dating paleoecological records and to the coarse spatial resolution of the paleo samples. Recent work suggests that some tree species migrated very fast to form advance colonies far from the moving front (Kullmann 1996, 1998), indicating dynamic equilibrium, but that local disequilibrium with climate also occurred due to mi-

grational lags when long-range dispersal was absent (Lotter 2000).

Modern analyses of paleo-ecological data have shown that in some regions tree species have migrated at rates of up to 4 km/yr in eastern North America, regardless of their dispersal mechanism (King and Herstrom 1997). To put this in perspective, some climatic zones are expected to shift polewards at average rates of 4–6 km/yr as the global climate warms (Solomon et al. 1984), which is an order of magnitude greater than the average past migration rate of trees (Solomon and Kirilenko 1997), and even exceeds King and Herstrom's (1997) maximum estimate. Such observations suggest that seed dispersal may well limit future migration, particularly in highly fragmented landscapes where successful colonization will be limited to relatively few suitable regeneration sites (Melillo et al. 1996, Iverson and Prasad 1998).

#### ACKNOWLEDGMENTS

We would like to thank Heike Lischke, Jesse Logan, and Jim Haefner for fruitful discussions, as well as Michael Cain and E. Raymond Heithaus for willingness to discuss and share their data. Comments of Michael Cain, Jim Clark, Mark Lewis, Michael Neubert, and an anonymous reviewer helped us to improve the manuscript considerably. Portions of this research were funded by the Swiss National Science Foundation (grant number 31-55985.98) and the U.S. National Science Foundation (grants DMS-0077663 and EEP-9903947), and by the Novartis Foundation.

#### LITERATURE CITED

- Andersen, M. 1991. Properties of some density-dependent integro-difference equation population models. *Mathematical Biosciences* **104**:135–157.
- Andow, D. A., P. M. Kareiva, S. A. Levin, and A. Okubo. 1990. Spread of invading organisms. *Landscape Ecology* **4**:177–188.
- Bennett, K. D. 1983. Postglacial population expansion of forest trees in Norfolk, UK. *Nature* **303**:164–167.
- Cain, M. L., H. Damman, and A. Muir. 1998. Seed dispersal and the Holocene migration of woodland herbs. *Ecological Monographs* **68**:325–347.
- Cain, M. L., B. G. Milligan, and A. E. Strand. 2000. Long-distance seed dispersal in plant populations. *American Journal of Botany* **87**:1217–1227.
- Clark, J. S. 1998. Why trees migrate so fast: confronting theory with dispersal biology and the paleorecord. *American Naturalist* **152**:204–224.
- Clark, J. S., C. Fastie, G. Hurrst, S. T. Jackson, C. Johnson, G. A. King, M. Lewis, J. Lynch, S. Pacala, C. Prentice, E. W. Schupp, T. Webb III, and P. Wyckoff. 1998. Reid's paradox of rapid plant migration. *BioScience* **48**:13–24.
- Clark, J. S., M. A. Lewis, and L. Horvath. 2001. Invasion by extremes: population spread with variation in dispersal and reproduction. *American Naturalist* **157**:537–554.
- Clark, J. S., M. Silman, R. Kern, E. Macklin, and J. HilleRisLambers. 1999. Seed dispersal near and far: patterns across temperate and tropical forests. *Ecology* **80**:1475–1494.
- Cushing, J. M., R. F. Costantino, B. Dennis, R. A. Desharnais, and S. M. Henson. 2002. *Chaos in ecology: experimental nonlinear dynamics*. Academic Press, Boston, Massachusetts, USA.
- Darley-Hill, S., and W. C. Johnson. 1981. Acorn dispersal by the Blue Jay (*Cyanocitta cristata*). *Oecologia* **50**:231–232.
- Davis, M. B. 1981. Quaternary history and the stability of forest communities. Pages 132–153 in D. C. West, H. H. Shugart, and D. B. Botkin, editors. *Forest succession: concepts and application*. Springer Verlag, New York, New York, USA.
- Davis, M. B. 1989. Lags in vegetation response to greenhouse warming. *Climatic Change* **15**:75–82.
- Delcourt, P. A., and H. R. Delcourt. 1987. *Long-term forest dynamics of the temperate zone*. Springer, New York, New York, USA.
- Eriksson, O. 2000. Seed dispersal and colonization ability of plants—assessment and implications for conservation. *Folia Geobotanica* **35**:115–123.
- Fisher, R. 1937. The wave of advance of an advantageous gene. *Annual Eugenics* **7**:355–369.
- Heithaus, E. R. 1986. Seed dispersal mutualism and the population density of *Asarum canadense*, an ant-dispersed plant. Pages 199–210 in A. Estrada and T. H. Fleming, editors. *Frugivores and seed dispersal*. Dr. W. Junk, Dordrecht, The Netherlands.
- Hengeveld, R. 1989. *Dynamics of biological invasions*. Chapman and Hall, London, UK.
- Higgins, S. I., and D. M. Richardson. 1999. Predicting plant migration rates in a changing world: the role of long-distance dispersal. *American Naturalist* **153**:464–475.
- Holmes, E. E., M. A. Lewis, J. E. Banks, and R. R. Veit. 1994. Partial differential equations in ecology: spatial interactions and population dynamics. *Ecology* **75**:17–29.
- Holmes, M. H. 1995. *Introduction to perturbation methods*. Springer-Verlag, New York, New York, USA.
- 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, UK.
- Hutchins, H. E., and R. M. Lanner. 1982. The central role of Clark's nutcracker in the dispersal and establishment of whitebark pine. *Oecologia* **55**:192–201.
- Iverson, L. R., and A. M. Prasad. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecological Monographs* **68**:465–485.
- 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 (*Cyanocitta cristata*) in the postglacial dispersal of fagaceous trees in eastern North America. *Journal of Biogeography* **16**:561–571.
- King, G. A., and A. A. Herstrom. 1997. Holocene tree migration rates objectively determined from fossil pollen data. Pages 91–101 in B. Huntley, W. Cramer, A. V. Morgan, H. C. Prentice, and J. R. M. Allen, editors. *Past and future rapid environmental changes: the spatial and evolutionary responses of terrestrial biota*. NATO ASI Series I: Global Environmental Change. Springer-Verlag, Berlin, Germany.
- Kolmogorov, A., I. Petrovsky, and N. Piscounov. 1937. Etude de l'équation de la diffusion avec croissance de la quantité de la matière et son application à un problème biologique. *Bulletin of University of Moscow, Series A* **1**:1–25.
- Kot, M. 1992. Discrete-time travelling waves: ecological examples. *Journal of Mathematical Biology* **30**:413–436.
- Kot, M., M. A. Lewis, and P. van den Driessche. 1996. Dispersal data and the spread of invading organisms. *Ecology* **77**:2027–2042.
- Kullman, L. 1996. Norway spruce present in the Scandes Mountains, Sweden at 8,000 BP: new light on Holocene tree spread. *Global Ecology and Biogeography Letters* **5**:94–101.
- Kullman, L. 1998. Paleoeological, biogeographical and paleoclimatological implications of early Holocene immigration of *Larix sibirica* Ledeb. into the Scandes mountains,

- Sweden. *Global Ecology and Biogeography Letters* **7**:181–188.
- Lang, G. 1994. *Quartäre Vegetationsgeschichte Europas*. Fischer, Jena, Germany.
- Lanner, R. M. 1980. Avian seed dispersal as a factor in the ecology and evolution of limber and whitebark pines. Pages 15–48 in B. P. Dancik and K. O. Higginbotham, editors. Sixth North American Forest Biology Workshop Proceedings. University of Alberta, Edmonton, Alberta, Canada.
- Lanner, R. M. 1982. Adaptations of whitebark pine *Pinus albicaulis* for seed dispersal by Clark's nutcracker *Nucifraga columbiana*. *Canadian Journal of Forest Research* **12**:391–402.
- Lanner, R. M. 1996. *Made for each other*. Oxford University Press, New York, New York, USA.
- Lewis, M. A. 2000. Spread rate for a nonlinear stochastic invasion. *Journal of Mathematical Biology* **41**:430–454.
- Lotter, A. F., H. J. B. Birks, U. Eicher, W. Hofmann, J. Schwander, and L. Wick. 2000. Younger Dryas and Allerød summer temperatures at Gerzensee (Switzerland) inferred from fossil pollen and cladoceran assemblages. *Palaeogeography, Palaeoclimatology, Palaeoecology* **159**:349–361.
- Matlack, G. R. 1994. Plant species migration in a mixed-history forest landscape in eastern North America. *Ecology* **75**:1491–1502.
- Melillo, J. M., I. C. Prentice, G. D. Farquhar, E. D. Schulze, and O. E. Sala. 1996. Terrestrial biotic responses to environmental change and feedbacks to climate. Pages 445–481 in J. T. Houghton, L. G. Meira Filho, B. A. Callander, N. Harris, A. Kattenburg, and K. Maskell, editors. *Climate change 1995: the science of climate change*. Cambridge University Press, Cambridge, UK.
- Nathan, R., U. N. Safriel, and I. Noy Meir. 2001. Field validation and sensitivity analysis of a mechanistic model for tree seed dispersal by wind. *Ecology* **82**:374–388.
- Nathan, R., U. N. Safriel, I. Noy Meir, and G. Schiller. 2000. Spatiotemporal variation in seed dispersal and recruitment near and far from *Pinus halepensis* trees. *Ecology* **81**:2156–2169.
- Neubert, M. G., and H. Caswell. 2000. Demography and dispersal: calculation and sensitivity analysis of invasion speed for structured populations. *Ecology* **81**:1613–1628.
- Neubert, M. G., M. Kot, and M. A. Lewis. 1995. Dispersal and pattern formation in a discrete-time predator-prey model. *Theoretical Population Biology* **48**:7–43.
- Okubo, A. 1980. *Diffusion and ecological problems: mathematical models*. Springer-Verlag, New York, New York, USA.
- Perkins, D. L. 2000. An assessment of successional status and factors affecting recruitment and mortality of treeline whitebark pine populations. Dissertation. Department of Forest Resources, Utah State University, Logan, Utah, USA.
- Portnoy, S., and M. F. Willson. 1993. Seed dispersal curves: behavior of the tail of the distribution. *Evolutionary Ecology* **7**:25–44.
- Prentice, I. C., P. J. Bartlein, and T. W. 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 British flora*. Dulau, London, UK.
- Schmidt, T. 1918. Die Verbreitung von Samen und Blütenstaub durch die Luftbewegungen. *Oesterreichische Botanische Zeitschrift* **67**:313–402.
- Shigesada, N., and K. Kawasaki. 1997. *Biological invasions: theory and practice*. Oxford University Press, Oxford, UK.
- Shigesada, N., K. Kawasaki, and E. Teramoto. 1987. Traveling periodic waves in heterogeneous environments. *Theoretical Population Biology* **36**:311–338.
- Skellam, J. G. 1951. Random dispersal in theoretical populations. *Biometrika* **38**:196–218.
- Solomon, A. M., and A. P. Kirilenko. 1997. Climate change and terrestrial biomass: what if trees do not migrate? *Global Ecology and Biogeography Letters* **6**:139–148.
- Solomon, A. M., M. L. Tharp, D. C. West, G. E. Taylor, J. W. Webb, and J. L. Trimble. 1984. Response of unmanaged forests to carbon dioxide-induced climate change: available information, initial tests and data requirements. Technical Report TR-009, United States Department of Energy, Washington, D.C., USA.
- Stimm, B., and K. Boeswald. 1994. Jays and nutcrackers: the ecology and silvicultural importance of seed dispersal by birds. *Forstwissenschaftliches Centralblatt* **113**(3–4):204–223.
- Vander Wall, S. B. 1990. *Food hoarding in animals*. University of Chicago Press, Chicago, Illinois, USA.
- Vander Wall, S. B. 1993. Cache site selection by chipmunks (*Tamias* spp.) and its influence on the effectiveness of seed dispersal in Jeffrey pine (*Pinus jeffreyi*). *Oecologia* **96**:246–252.
- Vander Wall, S. B. 1994. Removal of wind-dispersed pine seeds by ground-foraging vertebrates. *Oikos* **69**(1):125–132.
- Vander Wall, S. B., and R. B. Balda. 1977. Coadaptations of the Clark's nutcracker and the piñon pine for efficient seed harvest and dispersal. *Ecological Monographs* **47**:89–111.
- Webb, T. L. 1986. Is vegetation in equilibrium with climate? How to interpret late-Quaternary pollen data. *Vegetatio* **67**:75–91.

## APPENDIX A

### THE METHOD OF HOMOGENIZATION

A general diffusive model for the dispersal of seeds by active dispersers is given by

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial x} \left[ D(x) \frac{\partial P}{\partial x} \right] - \lambda(x)P, \quad P(x, t = 0) = \delta(x - \xi), \quad (\text{A.1})$$

$$\frac{\partial S}{\partial t} = \lambda(x)P, \quad S(x, t = 0) = 0. \quad (\text{A.2})$$

Here  $P$  is the density of seeds in movement from source (at location  $\xi$ ) to caching site and  $S$  is the density of seeds cached by active dispersers. The independent variables  $x$  and  $t$  are space and time variables on the scale of kilometers and years. We wish to take into account the effect of active dispersers on seed movement, and assume that the dispersers cache seeds

in a structure resolved on a much finer scale than  $x$ . Therefore, we may view the parameters in (Eqs. A.1, A.2) as functions of *two* spatial variables,  $x$  and  $y = x/\varepsilon$ , where  $\varepsilon$  is the ratio of the migration scale and the caching scale, and is assumed to be small. Thus, in the case of an ant-dispersed plant, one may assume that the small-scale distribution of seeds varies on the order of meters, while the plant population is distributed over hectares. In this case,  $\varepsilon = 0.01 = 1 \text{ m}/100 \text{ m}$ . In the main text we also assume that parameters depend only on the small scale in order to apply integrodifference equation techniques, but that is not a requirement of the homogenization theory.

The parameter in this model reflecting movement among cache sites,  $D$ , and the rate of caching seeds,  $\lambda$ , are functions of both the large ( $x$ ) and small ( $y$ ) scales. As an example, ant

colonies in a riparian lowland may be relatively close together, spaced at a mean distance of 2 m. Over kilometers, moving away from the riparian zone and into plains or foothills, the density of undergrowth and amount of forage decrease, and correspondingly ant mounds are spaced farther apart. A functional model for this situation might be periodic with period  $L$  on the small scale, but the period itself varies with the large scale, as in

$$D(y) = D_0 + (D_1 - D_0) \left[ \cos^2 \left( \pi \frac{y}{L_0 + L_1 x} \right) \right]^\theta \quad (\text{A.3})$$

$$\lambda(y) = \lambda_0 + (\lambda_1 - \lambda_0) \left[ \cos^2 \left( \pi \frac{y}{L_0 + L_1 x} \right) \right]^\theta. \quad (\text{A.4})$$

Both parameters oscillate between their within-cache ( $\lambda_0, D_0$ ) and between-cache ( $\lambda_1, D_1$ ) values; the parameter  $\theta$  controls the width of the caching zone. The larger the value of  $\theta$ , the more tightly focussed the cache area.

Now we may begin the homogenization procedure. The goal of the homogenization is to find approximate solutions to Eqs. A.1 and A.2, asymptotic in  $\varepsilon$ . That is, we wish to find solutions whose error can be measured by  $\varepsilon$  and whose structure can be written as a power series in  $\varepsilon$ . Since  $\varepsilon$  is thought of as varying (in the limit as it approaches zero), the variables  $x$  and  $y$  can be thought of as independent. Most of the discussion from this point on is derived from Chapter 5 of Holmes (1995). Treating  $x$  and  $y$  as independent induces a change of derivatives:

$$\frac{\partial}{\partial x} \rightarrow \frac{\partial}{\partial x} + \frac{1}{\varepsilon} \frac{\partial}{\partial y}. \quad (\text{A.5})$$

One must also account for multiple time scales; a fast time,  $\tau = \varepsilon^{-2}t$ , captures rates at which caching occurs (very fast) compared with time scales on which migration occurs. Treating  $\tau$  and  $t$  independently induces another change of variables:

$$\frac{\partial}{\partial t} \rightarrow \frac{\partial}{\partial t} + \frac{1}{\varepsilon^2} \frac{\partial}{\partial \tau}. \quad (\text{A.6})$$

Introducing these changes of derivative into Eq. A.1 and multiplying through by  $\varepsilon^2$  gives a diffusive equation with small parameters:

$$(\partial_\tau + \varepsilon^2 \partial_t)P = (\partial_y + \varepsilon \partial_x)[D(x, y)(\partial_y + \varepsilon \partial_x)P] - \varepsilon^2 \lambda(x, y)P. \quad (\text{A.7})$$

Above we have used the compact notation  $\partial_x = \partial/\partial x$  for differentiation. We can now proceed with a regular perturbation approach to Eq. A.7, but must remember that  $P$  is a function of *four* independent variables,  $x, y, t$ , and  $\tau$ .

Assuming that the correct solution may be expanded in a power series in  $\varepsilon$  we write

$$P = p_0(x, y, t, \tau) + \varepsilon p_1(x, y, t, \tau) + \varepsilon^2 p_2(x, y, t, \tau) + O(\varepsilon^3). \quad (\text{A.8})$$

Introducing this perturbation *Ansatz* into Eq. A.7, reorganizing, and using the subscript convention for derivatives (that is,  $\partial_x p_0 = p_{0x}$ ), gives

$$\begin{aligned} p_{0\tau} + \varepsilon p_{1\tau} + \varepsilon^2 (p_{2\tau} + p_{0t}) &= \partial_y [D(x, y) p_{0y}] \\ &+ \varepsilon \{ \partial_x [D(x, y) p_{0y}] + \partial_y [D(x, y) (p_{1y} + p_{0x})] \} \\ &+ \varepsilon^2 \{ \partial_y [D(x, y) (p_{1x} + p_{2y})] \\ &+ \partial_x [D(x, y) (p_{1y} + p_{0x})] \} + O(\varepsilon^3). \end{aligned} \quad (\text{A.9})$$

Eq. A.9 can be integrated sequentially through the expedient of equating terms at comparable orders of  $\varepsilon$  and solving the resulting, simplified equations in respective order.

At leading order ( $\varepsilon^0$ ) the governing equation relates only the most rapid scales,

$$\partial_\tau p_0 = \partial_y [D(x, y) p_{0y}] \quad (\text{A.10})$$

which has the general solution

$$p_0 = c_1(x, t) + c_0(x, t) \frac{1}{2} \int_{-y}^y \frac{dy'}{D(x, y')} + \tilde{p}_0(x, y, \tau). \quad (\text{A.11})$$

Here the first two terms on the right hand side are the equilibrium solutions (in fast variables) and  $\tilde{p}_0$  comprises the transients, which in the case of periodic small scales must decay exponentially to zero. The equilibrium solution can be derived by direct integration (assuming time derivatives are zero). The first integration gives

$$c_0(x, t) = D(x, y) \partial_y p_0 \quad (\text{A.12})$$

where  $c_0$  is a constant function with respect to  $y$  and  $\tau$ . Dividing by  $D$  and integrating once again gives the equilibrium solution:

$$p_0 = c_1(x, t) + c_0(x, t) \frac{1}{2} \int_{-y}^y \frac{dy'}{D(x, y')}. \quad (\text{A.13})$$

Since any solution to a linear PDE can be written as the sum of steady state and transient components, we arrive immediately at Eq. A.11.

In fact, given that  $D_0$  is the minimum diffusion constant and  $L(x)$  the local period, one may show that

$$|\tilde{p}_0(x, y, \tau)| \leq C(x) e^{-4D_0(\pi^2/L^2)\tau}. \quad (\text{A.14})$$

Thus, we may take the equilibrium solution for  $p_0$  given in Eq. A.13 as the full solution and make only exponentially small errors. Finally, we have to apply boundary conditions as  $y \rightarrow \pm\infty$ ; namely, we ask for bounded solutions. Since  $1/D_1 \leq 1/D(x, y) \leq 1/D_0$ , we know that the integral in the right-hand side of Eq. A.13 grows linearly in  $y$ , as a function bounded below by  $y/D_1$  and above by  $y/D_0$ . This clearly violates boundedness for large  $y$  (such terms are called *secular* in the parlance of asymptotic theory), and the only consistent solution is therefore to set  $c_0(x, t) = 0$  and totally remove the secularity. Thus, we have learned that the steady-state solution to Eq. A.10 (and therefore the leading order solution to Eq. A.9) is a function *only* of the slow variables,  $x$  and  $t$ .

At the next order,  $O(\varepsilon^1)$ , the governing equation is

$$\partial_\tau p_1 = \partial_x [D(x, y) p_{0y}] + \partial_y [D(x, y) (p_{1y} + p_{0x})]. \quad (\text{A.15})$$

Integrating directly once and isolating  $p_1$  gives

$$\partial_\tau p_1 = b_0(x, t) \frac{1}{D(x, y)} - p_{0x}. \quad (\text{A.16})$$

Integrating to find the steady-state solution and then adding the transient solution gives a general solution of the form

$$\begin{aligned} p_1 &= b_1(x, t) + b_0(x, t) \frac{1}{2} \int_{-y}^y \frac{dy'}{D(x, y')} - y p_{0x} \\ &+ \tilde{p}_1(x, y, \tau). \end{aligned} \quad (\text{A.17})$$

Again, the transient terms vanish exponentially fast in time, and we can therefore say that

$$p_1 = b_1(x, t) + \underbrace{b_0(x, t) \frac{1}{2} \int_{-y}^y \frac{dy'}{D(x, y')} - y p_{0x}}_{\text{Secular Terms}}. \quad (\text{A.18})$$

Here the secular terms are identified, but in this case one is not constrained to choose  $b_0 = 0$ . The secularity is removed asymptotically in  $y$  if we choose the following:

$$p_{0x} = b_0(x, t) \lim_{y \rightarrow \infty} \underbrace{\frac{1}{2} \int_{-y}^y \frac{dy'}{D(x, y')}}_{\widehat{D}^{-1}(x)} \quad (\text{A.19})$$

Here is where the harmonic average of the diffusion constant crops up, which has such large ramifications in the main text.

Finally, at order  $\varepsilon^2$  we see the last of the solvability conditions that specifies  $p_0$ , the leading order solution we seek. The governing equation at this order is

$$p_{0\tau} + p_{2\tau} - \lambda(x, y)p_0 = \partial_y [D(x, y)(p_{1x} + p_{2y})] + \partial_x [D(x, y)(p_{1y} + p_{0x})]. \quad (\text{A.20})$$

We begin by seeking the steady state (in  $\tau$ ) solution for  $p_2$ . Integrating once (in  $y$ ) gives

$$D(x, y)(p_{1x} + p_{2y}) = a_0(x, t) + yp_0 - p_0 \frac{1}{2} \int_{-y}^y \lambda(x, y') dy' - y\partial_x b_0 \quad (\text{A.21})$$

where we have used Eq. A.16 to simplify the second term on the right-hand side of Eq. A.20. Isolating  $p_2$  and a subsequent integration gives

$$p_2 = \underbrace{a_1(x, t) - a_0(x, t) \frac{1}{2} \int_{-y}^y \frac{dy'}{D(x, y')} - yp_{1x}}_{\text{Secular, linear in } y} + \underbrace{\frac{1}{2} y^2 (\partial_y p_0 - \partial_x b_0) + p_0 \int \left( \int^{y'} \lambda(x, y') dy' \right) dy''}_{\text{Secular, quadratic in } y} \quad (\text{A.22})$$

There are two sets of secular terms in this expression, one of which grows linearly in  $y$  (which would give rise to solvability conditions for  $p_1$ , if we were calculating the first correction to  $p_0$ ) and the other of which grows quadratically in  $y$ .

To continue the perturbation analysis and solve for  $p_2$  it is necessary that both sets of secular terms be removed independently, and therefore it is necessary that the quadratic terms vanish. To simplify the analysis, we note that if  $\lambda(x, y)$  is locally periodic in  $y$  and bounded by  $\lambda_0$  and  $\lambda_1$ , then we can be certain that the integral of  $\lambda(x, y)$  grows linearly in  $y$  and we may write

$$\hat{\lambda}(x) = \lim_{y \rightarrow \infty} \frac{1}{2y} \int_{-y}^y \lambda(x, y') dy'. \quad (\text{A.23})$$

Then, in the limit of large  $y$ , all terms in the quadratically secular group have a common factor of  $\frac{1}{2}y^2$ , and to remove these secular terms it is necessary that

$$\partial_\tau p_0 - \partial_x b_0 + \hat{\lambda}(x)p_0 = 0. \quad (\text{A.24})$$

Using Eq. A.19 gives

$$\partial_\tau p_0 = \partial_x \left[ \frac{1}{\widehat{D}^{-1}(x)} \partial_x p_0 \right] - \hat{\lambda}(x)p_0. \quad (\text{A.25})$$

This is the homogenized version of Eq. A.1, correct at leading order in  $\varepsilon$ .

The upshot of the homogenization procedure is that, given we are working at a scale large enough to integrate over small-scale dispersal structure, the multiscale dispersal system Eqs. A.1 and A.2 can be replaced with the following equations:

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial x} \left[ \frac{1}{\widehat{D}^{-1}(x)} \frac{\partial}{\partial x} P \right] - \hat{\lambda}(x)P \quad P(x, t = 0) = \delta(x - \xi) \quad (\text{A.26})$$

$$\frac{\partial S}{\partial t} = \hat{\lambda}(x)P, \quad S(x, t = 0) = 0. \quad (\text{A.27})$$

These homogenized equations generate correct leading-order solutions to the multiscale dispersal equations.

The homogenized parameters  $\widehat{D}^{-1}$  and  $\hat{\lambda}(x)$  can be calculated, in principle, by evaluating the necessary indefinite integrals and taking the limit of the ratio with  $y$ , as indicated in the definitions above. In the case of periodic (in the small scale) functions, it is sufficient to take the average over a single period, as is shown in Holmes (1995). For example, using the example of Eqs. A.3 and A.4 and choosing  $\theta = 1$ ,

$$\hat{\lambda}(x) = \lim_{y \rightarrow \infty} \frac{1}{2y} \int_{-y}^y \left[ \lambda_0 + (\lambda_1 - \lambda_0) \cos^2 \left( \frac{\pi y'}{L_0 + L_1 x} \right) \right] dy' \quad (\text{A.28})$$

$$= \lim_{y \rightarrow \infty} \frac{1}{2y} \left[ (2\lambda_0 + \lambda_1 - \lambda_0)y + (\lambda_1 - \lambda_0) \times \frac{(L_0 + L_1 x)}{\pi} \sin \left( \frac{2\pi y}{L_0 + L_1 x} \right) \right] \quad (\text{A.29})$$

$$= \frac{1}{2}(\lambda_0 + \lambda_1). \quad (\text{A.30})$$

Turning to the homogenized diffusion coefficient,

$$\widehat{D}^{-1}(x) = \lim_{y \rightarrow \infty} \frac{1}{2y} \int_{-y}^y \frac{dy'}{D_0 + (D_1 - D_0) \cos^2 \left( \frac{\pi y'}{L_0 + L_1 x} \right)} \quad (\text{A.31})$$

$$= \lim_{y \rightarrow \infty} \frac{1}{y} \frac{(L_0 + L_1 x)}{\pi \sqrt{D_0 D_1}} \tan^{-1} \left[ \frac{\sqrt{D_0}}{\sqrt{D_1}} \tan \left( \frac{\pi y}{L_0 + L_1 x} \right) \right] \quad (\text{A.32})$$

$$= \frac{1}{\sqrt{D_0 D_1}}. \quad (\text{A.33})$$

If  $\lambda_0, \lambda_1, D_0, D_1$  are now taken to be functions of the long scale,  $x$ , then direct substitution shows that the homogenized parameters may be functions of the long space variable, averaged over the short space variable. In the case of solutions for the dispersal kernel,  $K$ , to be used in calculations in the main text, we assume that dispersal parameters vary only on the small scale to avoid dependence of the homogenized parameters on the large scale.

**APPENDIX B**

A set of equations showing the asymptotic speed of invasion is available in ESA's Electronic Data Archive: *Ecological Archives* E085-010-A1.